Modeling of hot-electron effects in Si MOS devices

Citation for published version (APA):
Elias, P. J. H. (1993). *Modeling of hot-electron effects in Si MOS devices*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR402034

DOI:
10.6100/IR402034

Document status and date:
Published: 01/01/1993

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 04. Oct. 2023

MODELING OF HOT-ELECTRON EFFECTS IN SI MOS DEVICES

PETER ELIAS

MODELING OF HOT-ELECTRON EFFECTS IN SI MOS DEVICES

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof. dr. J.H. van Lint, voor een commissie aangewezen door het College van Dekanen in het openbaar te verdedigen op maandag 20 september 1993 om 16.00 uur

door

Petrus Johannes Henricus Elias

geboren te Tilburg

Dit proefschrift is goedgekeurd door de promotoren:
prof.dr.F.M.Klaassen
en
prof.dr.K.De Meyer
en co-promotor: dr.ir.Th.G.van de Roer

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Elias, Petrus Johannes Henricus

Modeling of hot-electron effects in Si MOS devices /
Petrus Johannes Henricus Elias. - [S.l. : s.n.]. - Fig.,
tab.
Thesis Eindhoven. - With ref. - With summary in Dutch.
ISBN 90-9006316-1 bound
NUGI 832
Subject headings: metal oxide semiconductor devices; hot
electrons.

Druk: WIBRO dissertatiedrukkerij, Helmond

O...

Uit: De boeken der kleine zielen
Louis Couperus

This work was supported by the Dutch Innovative Research Program (IOP-IC Technology).

LIST OF CONTENTS

SUMMARY

The Metal-Oxide-Silicon Field Effect Transistor (MOSFET) has proven to be the work horse of modern IC technology. In the development of the MOSFET over the last decade, the most striking feature is the reduction of the dimensions of the active region into the deep-submicron range. However, this decrease in device size leads to an increase in complexity of the device physics. First of all, the free charges gain much energy due to the corresponding high and steep peak in the electric field and the so called hot-carrier effects will play an essential role in the device behavior. Secondly, quantization effects become increasingly important.

The main goal of this thesis is to obtain a better insight in the hot-electron effects in submicron Si n-channel MOSFETs, by making use of simulations. In these simulations the electron temperature distribution and avalanche generation distribution in the MOSFETs are calculated. In addition the device currents are calculated, in particular the substrate current, which is a measure of the overall avalanche generation. The simulation results obtained are compared to experimental data obtained from other research institutes (in particular Philips Research Laboratories). In the end, this research may lead to further improvement of the devices in the future.

Much effort has been spent in solving numerically the Boltzmann transport equation, including high-energy effects. To solve the transport equation, basically two different methods are available. One is the Ensemble Monte Carlo method, in which the trajectories of a large number of simulation carriers are followed through the device. The great advantage of the Monte Carlo method is that complicated band structures and scatter mechanisms, and all sorts of quantum effects can be included. Furthermore, no a priori knowledge of the energy distribution function is needed. But the Monte Carlo method has its price: it is extremely expensive in computing time and storage space. The method also introduces much numerical noise.

The second method, the hydrodynamic method, is an extension of the well-established drift-diffusion method and takes into account the first three moments of the Boltzmann transport equation. This method is faster and less noisy than the Monte Carlo method, and in this thesis the hydrodynamic method has been preferred as a device simulator. However, its main disadvantage is that implicitly a shifted, asymmetric Maxwell-Boltzmann distribution function is assumed. Therefore in theory, its range of applicability is less than that of the Monte Carlo method. Nevertheless, our simulations show that the

hydrodynamic method can be used in the deep-submicron region, which fulfills our requirements. However, it has to be noted that the hydrodynamic method yet does give rise to serious convergence problems.

Although the Ensemble Monte Carlo method in this thesis is not used for the actual simulation of MOSFETs, in order to obtain the transport parameters as a function of electron energy, a Monte Carlo simulator has been developed for bulk material. The results show a nearly linear relation between average electron energy and applied field. Furthermore a constant energy relaxation time of 0.4 picosecond has been found. This value is nearly independent of the doping concentration.

In addition a Monte Carlo one-dimensional device simulator has been developed for studying the electron transport in diode structures. Results obtained with this simulator have been compared with results obtained by the hydrodynamic method. Simulations show that the heat-flow in the hydrodynamic model is better described by the adiabatic model than by the more conventional Wiedemann-Franz model.

Next a quasi two-dimensional hydrodynamic MOSFET simulator has been developed. In this simulator the transport parameters have been defined as a function of local electron energy. Using the adiabatic heat-flow model in combination with the avalanche parameters at the Si-SiO_2 interface, as defined by Slotboom, good agreement between the calculated and the experimental values of the substrate current has been found.

Furthermore the full two-dimensional hydrodynamic simulator TRENDY, developed at the University of Twente, has been improved and tailored to our problem. However TRENDY yields substrate currents that are much lower than the experimental data. This is due to the fact that TRENDY uses the Wiedemann-Franz heat flow model. Probably better results can be obtained if a good discretization scheme for the adiabatic model has been found.

Finally further studies have shown that the use of the drift-diffusion method, in combination with a post-routine to calculate the electron temperature, yields very useful first-order approximations of the hot-electron effects.

SAMENVATTING

De Metaal-Oxide-Silicium Veld-Effekt-Transistor (MOSFET) kan met recht het werkpaard van de moderne IC-technologie genoemd worden. Meest opvallend in de ontwikkeling van de MOSFET gedurende de afgelopen tien jaar, is de verkorting van het aktieve gebied tot diep in het submikron-gebied. Echter, in deze kort-kanaal MOSFETs is de fysika aanmerkelijk meer complex. Op de eerste plaats winnen de aktieve ladingsdragers veel energie als gevolg van de corresponderende hoge en steile piek in het elektrisch veld en gaan de zogenaamde hete-ladingsdragerseffekten een belangrijke rol spelen in het transistorgedrag. Ten tweede worden de quantumeffekten belangrijker.

Het doel van dit proefschrift is het verkrijgen van meer inzicht in de hete-elektroneneffekten in submikron silicium n-kanaals MOSFETs door middel van simulaties. In deze simulaties worden o.a. de verdeling van de elektronentemperatuur en de generatie van extra ladingsdragers als gevolg van ladingsvermenigvuldiging in de MOSFETs berekend. Ook de stromen aan de contacten worden berekend, in het bijzonder de substraatstroom, die een maat is voor de totale vermenigvuldigingsgeneratie. De verkregen simulatieresultaten worden vergeleken met experimentele data, die verkregen zijn van andere onderzoeksinstituten (in het bijzonder van het Philips Natuurkundig Laboratorium). Uiteindelijk kan dit onderzoek leiden tot verdere verbetering van de MOS transistoren in de toekomst.

Veel werk is besteed aan het numeriek oplossen van de Boltzmann transportvergelijking, waarin de hoge-elektroneneffekten meegenomen worden. In principe zijn er twee verschillende methoden beschikbaar om de transportvergelijking op te lossen. Eén is de Ensemble Monte Carlo methode, waarin een groot aantal simulatie-elektronen gevolgd worden op hun weg door de transistor. Het grote voordeel van de Monte Carlo methode is dat gekompliceerde bandenstrukturen, verstrooiingsmechanismen en ook allerlei quantumeffekten meegenomen kunnen worden. Maar dit voordeel heeft ook zijn prijs: de Monte Carlo methode vergt zeer veel rekentijd en geheugenruimte. Bovendien is de methode behept met ruis.

De tweede methode, de hydrodynamische, is een uitbreiding van de veel gebruikte drift-diffusie methode, waarbij nu de eerste drie momenten van de Boltzmann transportvergelijking als uitgangspunt genomen worden. De hydrodynamische methode is sneller en produceert minder ruis dan de Monte Carlo methode. In dit proefschrift is dan ook de voorkeur gegeven aan het hydrodynamisch model. Het belangrijkste nadeel is echter dat dit model impliciet een verschoven, asymmetrische Maxwell-Boltzmann verdelingsfunktie verondersteIt. Daardoor is het toepassingsgebied in theorie kleiner dan dat van de Monte Carlo methode. Onze simulaties tonen aan dat dit in de praktijk nogal meevalt, en dat de hydrodynamische methode tot diep in het submikrongebied bruikbaar is. Helaas is echter ook gebleken, dat deze methode nog ernstige convergentieproblemen heeft.

Hoewel de Ensemble Monte Carlo methode in dit proefschrift niet gebruikt wordt voor de simulatie van MOSFETs, is, om de transportparameters als funktie van de elektronenenergie te kunnen beschrijven, wel een Monte Carlo simulator ontwikkeld voor bulkmateriaal. De resultaten laten een vrijwel lineaire relatie zien tussen de gemiddelde elektronenenergie en het aangelegde veld; bovendien is een konstante energie-relaxatietijd van ongeveer 0.4 picosekonde gevonden. De gevonden waarden zijn praktisch onafhankelijk van de aanwezige verontreinigingskoncentratie.

Bovendien is een Monte Carlo één-dimensionale devicesimulator ontwikkeld voor de studie van het elektronentransport in diodestrukturen. Resultaten verkregen met deze simulator zijn vergeleken met resultaten verkregen met een hydrodynamische simulator. Hierbij blijkt dat de warmtestroming in het hydrodynamische model beter beschreven wordt door een adiabatisch model dan door het vaker gebruikte Wiedemann-Franz model.

Vervolgens is een quasi twee-dimensionale hydrodynamische MOSFET simulator ontwikkeld. In deze simulator zijn de transportparameters als funktie van de lokale elektronenenergie beschreven. Gebruikmakend van het adiabatische warmte-stromingsmodel, in combinatie met de vermengvuldigingsparameters aan het Si-SiO_2 grensvlak, zoals gedefinieerd door Slotboom, zijn de berekende substraatstromen in goede overeenstemming met de experimentele data.

Daarnaast is de twee-dimensionale simulator TRENDY, ontwikkeld aan de Universiteit van Twente, verbeterd en aangepast aan ons specifieke probleem. De substraatstromen berekend met TRENDY zijn veel lager dan de experimentele waarden. Dit komt doordat in TRENDY het Wiedemann-Franz model gebruikt is. Betere resultaten worden waarschijnlijk verkregen als een goed discretisatieschema voor het adiabatische model gevonden is.

Tenslotte heeft verder onderzoek aangetoond dat resultaten, verkregen met behulp van een drift-diffusie oplossing, in combinatie met een post-processor om de elektronentemperatuur te berekenen, een zeer bruikbare eerste-orde benadering oplevert van de hete-elektroneneffekten.

LIST OF SYMBOLS

a_0	Lattice parameter
$D_n;D_p$	Electron and hole diffusion constants
$\vec{e}_i$	Unit vectors in i-direction
$\vec{\mathbf{E}};E;E_{eff}$	Electric field; $E=\lvert\vec{\mathbf{E}}\rvert$; Effective electric field
$GR;GR_{AV}$	Generation/recombination rate; Avalanche generation rate
$\mathcal{G}$	Overlap factor
$\hbar$	Planck constant divided by 2π
$\mathbf{I}$	Identity matrix
$I_{DS};I_{AV}$	Drain-Source current; Avalanche generation current
$\vec{\mathbf{J}}_n;\vec{\mathbf{J}}_p;J_n;J_p$	Electron and hole current densities; $J_n=\lvert\vec{\mathbf{J}}_n\rvert$ and $J_p=\lvert\vec{\mathbf{J}}_p\rvert$
$\vec{\mathbf{k}};k$	Electron wave vector; $k=\lvert\vec{\mathbf{k}}\rvert$
k_B	Boltzmann constant
L	Device length
L_D	Debye length
m_0	Free electron rest mass
$m_n^*;m_p^*$	Conductivity electron and hole effective masses
M	Number of simulation electrons
$n;n_I$	Electron density; Intrinsic charge density
$N_A;N_D;N_I$	Acceptor, donor and total ionized impurity concentrations
p	Hole density
$\mathbf{P}_n;P_n$	Gas pressure tensor; Hydrostatic gas pressure
q	Unit charge
$\vec{\mathbf{Q}}_n;Q_n$	Electron heat flow; $Q_n=\lvert\vec{\mathbf{Q}}_n\rvert$
$\mathscr{r}$	Random number uniform distributed in $[0,1)$
$\vec{\mathbf{S}}_n;S_n$	Electron energy flow; $S_n=\lvert\vec{\mathbf{S}}_n\rvert$
$t;\mathrm{T}$	Time; Total simulation time
$T_n;T_0$	Electron and lattice temperature
$\mathbf{T}_i$	Herring-Vogt transformation matrix of i-th X-valley
$\mathbf{T}_L$	Herring-Vogt transformation matrix of L-valley
u	Electron kinetic energy
$u_G;u_{XL}$	Energy gap of Si; Energy gap between X- and L-valleys
$\mathbf{U}$	Transformation matrix between different coordination systems
U_{ex}	Externally applied voltage
$\vec{\mathbf{v}}_n;\vec{\mathbf{v}}_p;v_n;v_p$	Electron and hole mean velocities; $v_n=\lvert\vec{\mathbf{v}}_n\rvert$ and $v_p=\lvert\vec{\mathbf{v}}_p\rvert$
v_s	Velocity of sound in the crystal
v_{sat}	Saturation velocity
$V_{DS};V_{GS}$	Drain-Source and Gate-Source voltage
$\vec{\mathbf{w}};w$	Electron velocity; $w=\lvert\vec{\mathbf{w}}\rvert$
W_0	Thermal energy of carriers in equilibrium
W_n	Electron mean energy

$\alpha;\alpha_X;\alpha_L$	Non-parabolicity parameter; X-valley and L-valley respectively
$\alpha_n;\alpha_p$	Electron and hole impact ionization rates
α_S	Surface scattering ratio
$\alpha_2;\alpha_3$	Ionized impurity clustering parameters
β_s	Inverse screening length
δ	Dirac function
Γ	Maximal total scatter rate, including self-scattering
$\epsilon;\epsilon_0$	Dielectric constant; Dielectric constant in vacuum
ϵ_r	Relative dielectric constant
ϑ	Polar scatter angle
θ_{op}	Equivalent optical phonon temperature
$\vec{\kappa}$	Transformed electron wave vector
$\vec{\kappa}_n$	Electron thermal conductivity
λ_{ee}	Electron-electron scatter rate
λ_k	Scatter rate of k-th scatter process
λ_{wn}	Electron energy relaxation length
Λ_k	Total scatter rate
$\mu_n;\mu_p$	Electron and hole electrical mobility
Ξ_{AX}	Deformation potential of acoustic scattering in X-valley
Ξ_{AL}	Deformation potential of acoustic scattering in L-valley
Ξ_{ij}	Optical coupling constant
ρ	Space-charge density
ρ_S	Crystal density
σ_I	Cross section
$\sigma_n;\sigma_p$	Electron and hole electrical conductivities
τ_c	Average collision time
τ_H	Heating time
$\tau_{pn};\tau_{pp}$	Electron and hole momentum relaxation times
τ_{wn}	Electron energy relaxation time
φ	Azimuthal scatter angle
Ψ	Electric potential
ω_{pn}	Plasma frequency

ACKNOWLEDGEMENT

I would like to thank all following persons:

- My promotors prof.dr.F.M. Klaassen and prof.dr.K. De Meyer; the co-promotor dr.ir.Th.G. van de Roer; furthermore prof.dr.J. Zimmermann and prof.dr.M. Kleefstra for critically reviewing this work.
- All colleagues at the Semiconductor Device Group of the Eindhoven University of Technology for their assistance, in particular my room mates during the years: Karel Nederveen, Theo Smedes, Paul Tegelaar, Omar Abu-Zeid, Jan Otten, and Peter Konijn.
- At Philips Research Laboratories, for stimulating discussions and supplying me with experimental data: Jan Slotboom, Jos in de Kleef, Anko Heringa, Piere Woerlee, Maarten van Dort, Rudolph Velghe, Wil Schilders, Henk Boots and Frans Widdershoven.
- Philip Wolbert at Twente University of Technology for his assistance in using TRENDY.
- Herman Rohenkohl and several other system managers of the group Automatic IC Design for their aid in using the computer systems.
- Wim Schoenmakers for supplying me with some simulation results obtained at IMEC, Leuven.
- Students that contributed to this work: Grego Peeters, Ernst van Dijke, and Pieter van Boxtel.
- My family and friends for much needed moral support, in particular Helène Meurs, Annette van Esch and Bettina Garcia.
- Radboud Schmitz for his advice when my spirits were very low.
- My daughter Martha, just for being there.

Chapter 1

MOTIVATION AND INTRODUCTION

1.1 Motivation

The Metal-Oxide-Silicon Field Effect Transistor (MOSFET) has proven to be the work horse of modern IC technology and is used as a switching device, amplifier and load. A general description of the working of the MOS transistor is given in section 1.2. In the development of the MOS transistor over the last decade, the most striking feature is the reduction of the dimensions (down scaling) of the active regions into the submicron range (i.e. less than 1 micrometer). The major driving forces behind down scaling are speed improvement, reduction of power consumption and improvement of packing density. At the moment devices are produced with channel lengths less than 100 nanometers. However, this decrease in device size leads to an increase in complexity of the device physics. First of all, the free charges moving from source to drain gain much energy due to the corresponding high peak in the electric field and the so called hot-carrier effects will play an essential role in device behavior. Secondly, quantization effects in the inversion layer become increasingly important. Both quantization and hot-electron effects are described in detail in section 1.3.

The main goal of this thesis is to obtain a better insight in the hot-electron effects in submicron Si n-channel MOSFETs, by making use of simulations. In these simulations the electron temperature distribution and avalanche generation distribution in the MOSFETs are calculated. In addition the device currents are calculated, in particular the substrate current, which is a measure of the overall avalanche generation. The simulation results obtained are compared to experimental data obtained from other research institutes (in particular Philips Research Laboratories). In the end this may lead to further improvement of the devices in future.

The use of numerical simulation is common practice in the world of electronic device engineering nowadays. However different goals may be served. The technologist is mainly interested in process simulations; the designer of integrated electronic circuits uses circuit analysis programs like SPICE and other programs that contain analytical models of the terminal behavior of the devices. For our purpose a detailed numerical physical simulator is needed that gives insight in what exactly is going on inside the device. Traditionally the well-known drift-diffusion model has been used for this kind of simulation. Although in theory the drift-diffusion model is no longer valid in case of small devices, in practice its usefulness can be extended to regions far beyond the range of theoretical applicability. So for a long time more complicated models were hardly needed. However, as devices do get smaller and smaller, more advanced models have to be used. Therefore in our work most effort has been spent in solving numerically the Boltzmann transport equations, including high-energy effects.

To solve the Boltzmann transport equation basically two different methods are available. One is the Ensemble Monte Carlo method, in which the trajectories of a large number of simulation carriers are followed through the device. The great advantage of the Monte Carlo method is that complicated band structures and scatter mechanisms can be included. If the Monte Carlo method is combined with Schrödinger's equation also all sorts of quantum effects can be included. Furthermore, no a priori knowledge of the energy distribution function is needed. But the Monte Carlo method has its price: it is extremely expensive in computing time and storage space. The method also introduces much numerical noise.

The second method, the hydrodynamic method, is an extension of the well established drift-diffusion method and takes into account the first three moments of the Boltzmann transport equation. The derivation of the basic hydrodynamic equations is presented in section 1.4. These equations are solved iteratively. This method is faster and less noisy than the Monte Carlo method. However, its main disadvantage is that it operates only on average quantities and that a priori a shifted, asymmetric Maxwell-Boltzmann distribution is assumed. Furthermore it cannot take quantization and some hot-carrier effects (such as ballistic carriers, see section 1.3.2) into account. Therefore in theory its range of applicability is less than that of the Monte Carlo method.

In this thesis the hydrodynamic method has been preferred over the Monte Carlo method for the simulation of submicron MOSFETs. The Monte Carlo method has been used for other purposes. In order to obtain the transport parameters needed by the hydrodynamic method as a function of electron energy a Monte Carlo simulator has been developed for bulk material. In addition for studying the electron transport in diode structures a Monte Carlo device simulator in one dimension has been developed. This simulator has been used as a numerical reference to check the validity and range of applicability of the hydrodynamic method. In this thesis the development and use of the various simulators mentioned, combined with simulation results, is described in detail. Section 1.5 gives an overview.

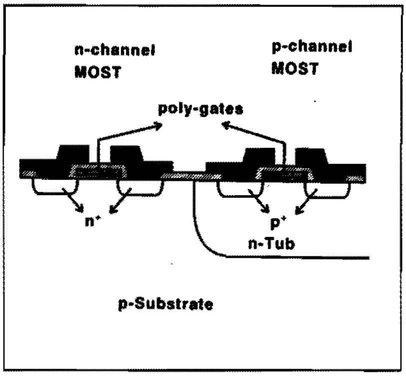

Fig 1.1 Schematic cross section of a n-MOST and a p-MOST in a CMOS process.

1.2 Basic principles of the MOS transistor

MOS transistors come in two versions, the n-channel MOST, in which the current is generated mainly by electrons moving from source to drain, and the p-channel type, in which holes are moving between the contacts. In modern technology both are made simultaneously in the so called CMOS-process (see figure 1.1). We will limit ourselves here to a description of the n-channel MOST (n-MOST). The working of the p-MOST is dual to the n-MOST. In principle the n-MOST is a very simple and straightforward device. In the substrate a p-well is created by implantation of acceptors. In this p-well two n^+-contact areas are created by heavy donor implantation. On top of the device a thin oxide is grown, on which a gate contact is deposited. Now voltage is applied to the contacts. If the gate-source voltage is below a certain threshold voltage, all p-n junctions are in reverse and very little current is flowing. The device is said to operate in the sub-threshold region; as a switching device the MOST is considered to be in the off-state. In modern technology there is an extra p-implant directly underneath the gate to increase the threshold voltage and so decrease the leakage currents in the sub-threshold region. If the gate-source voltage is above the threshold voltage, a large concentration of electrons is drawn towards the Si-SiO_2 interface directly underneath the gate. The electron channel connecting source and drain is called the inversion channel. Now electrons can move easily from source to drain and the device is in the on-state. At increasing drain-source voltages the drain-gate voltage finally drops below the threshold voltage and the current is pinched off at the drain side, allowing no further increase in the current. The device is said to be in saturation.

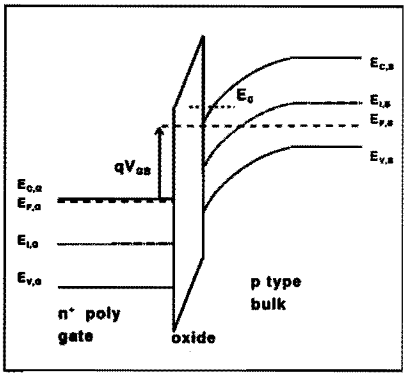

Fig 1.2 Band structure of a n^+ poly gate on a p-type substrate.

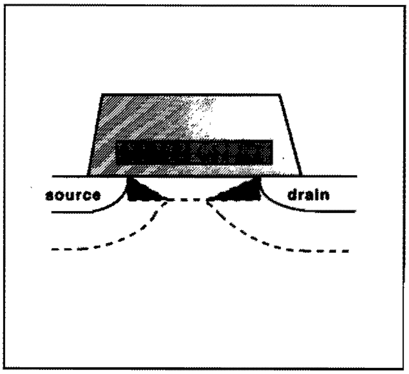

Fig 1.3 Shape of the depletion region in a short-channel MOSFET at zero drain bias; the charge shared by the gate and the junctions is indicated by the shaded areas.

1.3 Quantization and hot-electron effects in submicron MOSFETs

1.3.1 Quantization effects

To understand the physical background of the quantization effects in MOS transistors, it is necessary to look in detail at the Si-SiO_2 interface. In figure 1.2 a band diagram of this interface is given. Due to the difference in the work function between gate contact and Si, the difference between the Fermi level E_F in the bulk and the intrinsic Fermi level E_i, and due to the applied gate voltage V_{GB}, the Si band shows a strong band bending near the interface. This way a well-defined potential well is created. It is a well known fact that in such a potential well the electrons cannot be considered to be free in all three dimensions but the motion in the direction perpendicular to the Si-SiO_2 interface is quantized. The potential well, also called quantum well, has roughly a triangular form. An important parameter defining the exact shape of this well is the doping concentration at the interface: the higher the doping, the steeper and narrower the well.

It should be noted that the quantum well is a fundamental property of all MOS transistors, independent of the channel length. In large-device processes the doping concentration at the interface is fairly low and the quantum well is wide, which means that many discrete states close to one another are present, and quantization effects are not noticed in the device behavior. However, in short channel devices the situation is different. In modern MOS technology it is good practice to increase the p-implant directly underneath the gate for decreasing channel lengths. This needs some explication. Junction depletion layers are surrounding the source and drain contact areas, and there also is a depletion layer underneath the gate. As can be seen in figure 1.3, part of the depletion charge is shared by the source and drain junction. In long channel devices this part is negligible compared to the total depletion charge, but this is not the case in short channel transistors. Consequently less charge is controlled by the gate contact and it becomes easier to turn on the device. In other words the threshold voltage is lower in short channel devices. A high p-implant is used to counteract this effect. A high p-dope decreases the width of the depletion layer at the source and drain junctions, and consequently the charge shared is lowered. This way the threshold voltage is kept constant over a larger range of device lengths. This high p-implant also increases the voltage at which punch-through occurs.

However, when the doping at the Si-SiO_2 interface is increased, the potential well becomes narrower, and quantization effects become more prominent. In a narrow well E_0, the lowest energy level available does not coincide with the bottom of the conduction band (figure 1.2), and the work function difference is increased, which in its turn increases the threshold voltage. The quantization mentioned above decreases the number of states in the quantum well and this makes it even less easy to turn on the device, and increases the threshold voltage even further. A quantitative analysis is given by Van Dort and Woerlee [1991]. Their method is used in the MOST models presented in this thesis.

1.3.2 Hot-electron effects

By applying an electric field the electrons gain momentum and energy. If the field is strong enough the energy gain is substantial and the electron temperature, which is a measure for the amount of energy, becomes significantly higher than the lattice temperature. These high-energy electrons are called hot electrons, and they play an important role in submicron devices. The effect of hot electrons on the device behavior could be substantially reduced if the applied voltages were down-scaled according to the scaling rules, but for practical reasons this is not done. The most important hot-electron effects are given here:

1. Oxide injection: some electrons travelling from source to drain can gain so much energy that they force their way into the oxide layer underneath the gate. Some tunnel through and cause a gate current to flow, but most get trapped in the oxide, creating a negative oxide charge. This degenerates the device performance very rapidly.
2. Weak avalanche: the energy of a substantial number of electrons can become higher than the forbidden gap and these electrons are able to create new electron-hole pairs. The newly created electrons add to the source-drain current, while most of the holes move into the substrate and give rise to a substrate current. Some holes are injected in the oxide, in analogy to the electron injection described before.

3. Velocity overshoot: electrons at the drain side can reach a velocity that is higher than the saturation velocity. To understand this phenomenon qualitatively one has to consider the individual carriers to be in constant interaction with the phonons in the lattice. This interaction is energy dependant: an increase of electron energy also gives an increase of interaction. At high electric fields the gain of energy in equilibrium on the average is equal to the loss of energy to the lattice, and the electron velocity always equals the saturation velocity. If however the electric field is suddenly increased, the electrons gain momentum faster than it can be relaxed to the lattice, and velocity overshoot occurs. Such a situation can occur in a submicron MOS device, where the field in the channel rises to a high peak at the drain side.
4. Ballistic carriers: if the source-drain distance is smaller than the mean free path of the electrons, a considerable number of electrons reaches the drain without having had any interaction with the lattice or impurities. This gives rise to a so-called ballistic peak in the energy distribution function.
5. Energy diffusion current: since the electrons do not have a uniformly distributed average energy along the channel, an extra component is added to the total device current.

To make practical devices in the submicron range possible, it is necessary to reduce the peak in the electric field and the hot-electron effects. This is done by implanting a low doped extension of the drain (LDD). For practical reasons the source also is extended. Often there is an extra anti-punch-through (ATP) p-implant to avoid breakdown between the contacts. A three-dimensional plot of the doping profile in a modern submicron device is given in figure 1.4.

1.4 Derivation of the hydrodynamic equations

Models based on the Boltzmann transport equations have been in use since the sixties, the most well known iterative schemes having been developed by Gummel [1964] and Scharfetter and Gummel [1969]. The early models were based on the Poisson equation and the zero- and first-order moments of the Boltzmann transport equation, the so called drift-diffusion approach. The hydrodynamic method is an extension of the drift-diffusion method and also includes the second-order moment. In this section the full set of the hydrodynamic equations for both electrons and holes is given. However the derivation of these equations only is given for electrons. For holes a similar description holds.

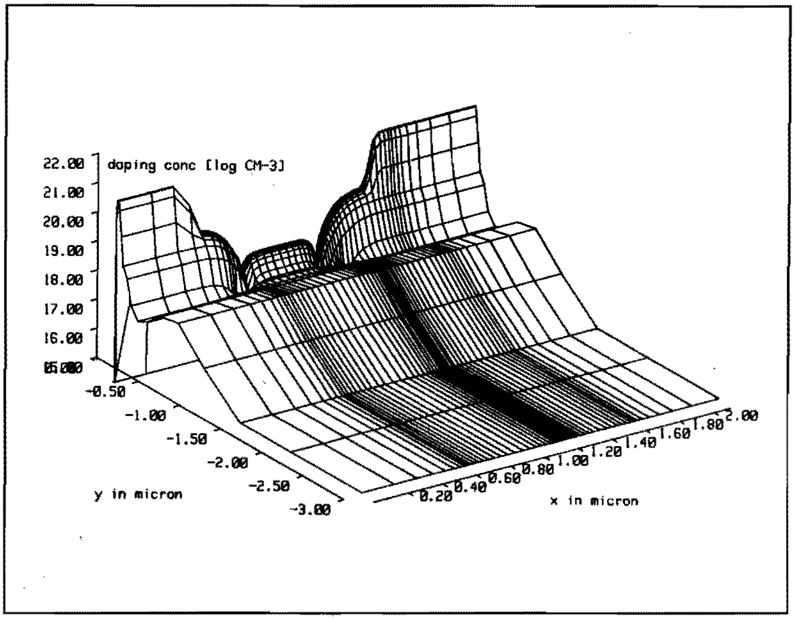

Fig 1.4 Doping profile of a submicron LDD MOSFET

Important is the Poisson equation that is given by

$$\nabla^2 \psi = -\frac{q}{\epsilon}\left(N_D^+ - N_A^- + p - n\right) \tag{1.4.1}$$

The electric field $\vec{E}$ is defined by

$$\vec{E} = -\nabla \psi \tag{1.4.2}$$

Assuming all particles to have zero volume, and assuming all particle interactions to be instantaneous, it is possible to define for each kind of particle a distribution function $f_c(\vec{k},\vec{r},t)$ in the six-dimensional phase-space. If both electrons and holes are present generation and recombination of particles may occur. An extra term $f_{GR}(\vec{k},\vec{r},t)$ is needed to correct for this process. The derivative of the distribution function with respect to time along a particle trajectory $\vec{k},\vec{r}$ equals zero as a result of the need to conserve the number of states, the so called Liouville theorem

$$\frac{d}{dt}\left(f_c(\vec{k},\vec{r},t) - f_{GR}(\vec{k},\vec{r},t)\right) = 0 \tag{1.4.3}$$

Expansion of equation (1.4.3). leads to the implicit form of the Boltzmann transport equation, that reads

$$\frac{\partial}{\partial t} f_c + \nabla_k f_c \cdot \frac{d\vec{k}}{dt} + \nabla_r f_c \cdot \frac{d\vec{r}}{dt} = \frac{d}{dt} f_{GR} \tag{1.4.4}$$

Next the group velocity $\vec{w}$ is defined by

$$\vec{w} = \frac{d\vec{r}}{dt} \tag{1.4.5}$$

The derivative of momentum with respect to time is equal to the sum of all forces acting on the particles. This is written as the sum of all external forces $\vec{F}_E$ and all internal forces $\vec{F}_I$ due to collisions, as long as the external forces can be considered to be constant over a distance comparable with the dimensions of the carrier, which is true in all our problems. Thus

$$\hbar \frac{d\vec{k}}{dt} = \vec{F}_E + \vec{F}_I \tag{1.4.6}$$

Combining (1.4.4) and (1.4.6) yields

$$\frac{\partial}{\partial t} f_c + \frac{\vec{F}_E}{\hbar} \cdot \nabla_k f_c + \vec{w} \cdot \nabla_r f_c = \frac{d}{dt} f_{GR} - \frac{\vec{F}_I}{\hbar} \cdot \nabla_k f_c \tag{1.4.7}$$

In both the hydrodynamic method and the Monte Carlo method all changes in f_c by scatter processes are assumed to be instantaneous and the right-hand-side of equation (1.4.7) is written explicitly as a collision term

$$\frac{\partial}{\partial t} f_c + \frac{\vec{F}_E}{\hbar} \cdot \nabla_k f_c + \vec{w} \cdot \nabla_r f_c = \frac{d}{dt} f_{GR} + \left(\frac{\partial}{\partial t} f_c \right)_{coll} \tag{1.4.8}$$

Equation (1.4.8) is called the explicit form of the Boltzmann transport equation. This form acts as a starting point for the Monte Carlo method, that is described in detail in chapter 2.

In the hydrodynamic method the first three moments of (1.4.8) are needed. The zero-order moment is obtained by direct integration of (1.4.8) over the momentum space V_k and gives the carrier conservation equation. The first-order moment is obtained by multiplying (1.4.8) with the group velocity $\vec{w}$ before integration and this gives the momentum conservation equation. The energy conservation equation is obtained by multiplying (1.4.8) with $\frac{1}{2} m^* (\vec{w} \cdot \vec{w})$ before integrating over V_k (second-order moment). Before the integrations are actually carried out and the results presented some definitions and further assumptions have to be discussed.

The electron density $n(\vec{r},t)$ (function of position in space $\vec{r}$ and time t) is defined by integrating the distribution function for electrons $f_n(\vec{k},\vec{r},t)$ over the momentum space V_k

$$n(\vec{r},t) = \int_{V_k} f_n(\vec{k},\vec{r},t)\; d\vec{k} \tag{1.4.9}$$

Here $d\vec{k}$ stands symbolically for $dk_x dk_y dk_z$. The generation-recombination rate $GR(\vec{r},t)$, the average electron velocity $\vec{v}_n(\vec{r},t)$, electron energy $W_n(\vec{r},t)$ and heat flow vector $\vec{Q}_n(\vec{r},t)$ are defined in a similar way

$$GR(\vec{r},t) = \frac{d}{dt} \int_{V_k} f_{GR}(\vec{k},\vec{r},t)\; d\vec{k} \tag{1.4.10}$$

$$\vec{v}_n(\vec{r},t) = \frac{1}{n} \int_{V_k} f_n(\vec{k},\vec{r},t)\; \vec{w}(\vec{k})\cdot d\vec{k} \tag{1.4.11}$$

$$W_n(\vec{r},t) = \frac{1}{n} \int_{V_k} \tfrac{1}{2}\, m^* |\vec{w}(\vec{k})|^2\; f_n(\vec{k},\vec{r},t)\; d\vec{k} \tag{1.4.12}$$

$$\vec{Q}_n(\vec{r},t) = \frac{1}{n} \int_{V_k} \tfrac{1}{2}\, m^* |\vec{w}(\vec{k})-\vec{v}_n|^2\; f_n(\vec{k},\vec{r},t)\; (\vec{w}(\vec{k})-\vec{v}_n)\cdot d\vec{k} \tag{1.4.13}$$

The electron current density $\vec{J}_n$ reads

$$\vec{J}_n = -\, q\, n\, \vec{v}_n \tag{1.4.14}$$

For holes a similar set of definitions is given with $\vec{J}_p = q\, p\, \vec{v}_p$.

In addition a few more assumptions are needed in the hydrodynamic method:
1. All particle velocities are non-relativistic.
2. External forces are only due to the electric field (in case of electrons $\vec{F}_E = -\, q\, \vec{E}$).
3. No degeneration effects occur.
4. In the second-order moment terms appear that represent the gas pressure tensor. If the distribution function is considered isotropic with respect to $\vec{v}$ this term can be simplified to the hydrostatic pressure, and then the concept of electron (or hole) temperature can be introduced in the equations. These temperatures are indicated by T_n and T_p.

In the first- and second-order moments the collision term represents the loss in momentum and energy from the electron system to the external world. In semiconductors this in general is the crystal lattice. In the hydrodynamic method this loss is represented by a relaxation time. In general the momentum relaxation time τ_p differs from the energy relaxation time τ_w. Because each function f_c can be given as the sum of an asymmetrical part f_{ca} and a symmetrical part f_{cs}, the collision term can be defined by

$$\left(\frac{\partial}{\partial t} f_c \right)_{coll} = - \frac{f_{ca}}{\tau_p} - \frac{f_{cs} - f_{c0}}{\tau_w} \qquad (1.4.15)$$

Here f_{c0} represents the equilibrium distribution. In equation (1.4.15) f_{c0} is assumed to be symmetrical, which is the case in our problems. The integration over the momentum space, using the definitions and assumptions given for both holes and electrons is a tedious but rather straight forward matter that is dealt with in most standard text books on the subject of plasma and particle dynamics and is omitted here. The final results, however, are given for both electrons and holes.

- The carrier conservation equations

$$\frac{\partial n}{\partial t} - \frac{1}{q} \nabla_r \cdot \vec{J}_n = GR \qquad (1.4.16)$$

$$\frac{\partial p}{\partial t} + \frac{1}{q} \nabla_r \cdot \vec{J}_p = GR \qquad (1.4.17)$$

- The momentum conservation equations

$$\vec{J}_n + n\, \tau_{pn} \frac{\partial}{\partial t} \left(\frac{\vec{J}_n}{n} \right) - \frac{\tau_{pn}}{q} \left(\vec{J}_n \cdot \nabla_r \right) \frac{\vec{J}_n}{n} = q\, \mu_n n\, \vec{E} + q\, \tau_{pn}\, \nabla_r \left(\frac{n\, k_B T_n}{m_n^*} \right) \qquad (1.4.18)$$

$$\vec{J}_p + p\, \tau_{pp} \frac{\partial}{\partial t} \left(\frac{\vec{J}_p}{p} \right) + \frac{\tau_{pp}}{q} \left(\vec{J}_p \cdot \nabla_r \right) \frac{\vec{J}_p}{p} = q\, \mu_p p\, \vec{E} - q\, \tau_{pp}\, \nabla_r \left(\frac{p\, k_B T_p}{m_p^*} \right) \qquad (1.4.19)$$

Here the mobility is given by

$$\mu_n = \frac{q\, \tau_{pn}}{m_n^*} \quad \text{and} \quad \mu_p = \frac{q\, \tau_{pp}}{m_p^*} \qquad (1.4.20)$$

- The energy conservation equations

$$\frac{\partial\, (n W_n)}{\partial t} + \nabla_r \cdot \vec{S}_n = \vec{E} \cdot \vec{J}_n - n \frac{W_n - W_0}{\tau_{wn}} + GR \cdot W_n \qquad (1.4.21)$$

$$\frac{\partial\, (p W_p)}{\partial t} + \nabla_r \cdot \vec{S}_p = \vec{E} \cdot \vec{J}_p - p \frac{W_p - W_0}{\tau_{wp}} + GR \cdot W_p \qquad (1.4.22)$$

Here $W_n = \frac{3}{2} k_B T_n + \frac{1}{2} m_n^* |\vec{v}_n|^2 \qquad W_p = \frac{3}{2} k_B T_p + \frac{1}{2} m_p^* |\vec{v}_p|^2$ (1.4.23)

$$\vec{S}_n = - \vec{Q}_n - (W_n + k_B T_n) \frac{\vec{J}_n}{q} \qquad \vec{S}_p = - \vec{Q}_p + (W_p + k_B T_p) \frac{\vec{J}_p}{q} \qquad (1.4.24)$$

$\vec{S}$ is called the energy flow vector. It should be noted that implicitly the assumption has been made that the momentum-energy relation is parabolic.

In traditional device modeling some extra simplifications are made. First of all no energy effects are taken into account and the carrier temperature is always considered to be equal to the lattice temperature. As a result no energy conservation equation is needed and in addition to the Poisson and carrier conservation equations, only simplified momentum equations are used. In general these are even further simplified by neglecting the time derivative and the convection term $(\vec{J}\cdot\nabla)\,\vec{v}$. As a result the so-called drift-diffusion equations are obtained

$$\vec{J}_n = q\,\mu_n\,n\,\vec{E} + q\,D_n\,\nabla_r\,n \tag{1.4.25}$$

$$\vec{J}_p = q\,\mu_p\,p\,\vec{E} - q\,D_p\,\nabla_r\,p \tag{1.4.26}$$

At relatively low electric fields (large devices) the drift-diffusion model has proven to work very well (which also justifies the simplifications that have been made) and is still an important tool in device modeling.

In the derivation thus far it has been assumed everywhere that the transport parameters m^* (conductivity effective mass), τ_p and τ_w (momentum and energy relaxation times) are all constants. However in practice this is not true. Depending on the model used these parameters are a function of electric field, position in the device, local electron energy or a combination of these. Including all these factors complicates the Boltzmann transport equation enormously and the derivation of manageable moments is impossible. Therefore the assumption is made that the transport parameters locally can be considered constant, and the field or energy dependency is introduced after the derivation of the equations (1.4.15) to (1.4.23). Although formally not correct, this has been common practice and has worked very well in device simulations, where the gradient in the parameters in fact is small compared to the gradient in the charge concentration. In the drift-diffusion model the parameters are taken to be a function of the local field. The hydrodynamic approach offers the possibility to use transport parameters as a function of local carrier energy, which is more in accordance with physical reality, and this has been used in this thesis. Also the hydrodynamic method makes it possible to include the effects of non-parabolic energy bands in an approximate way by allowing the average effective mass to be a function of the average energy.

A final note on the Generation-Recombination GR term has to be made. This term consists of three different mechanisms:

1. The Shockley-Read-Hall recombination

$$GR_{SRH} = \frac{n_I^2 - n\,p}{\tau_p(n_I + n) + \tau_n(n_I + p)} \tag{1.4.27}$$

In the above the intrinsic charge concentration is represented by n_I. τ_n and τ_p represent the electron and hole life time.

2. The Auger recombination

$$GR_{AU} = (n_I^2 - n\,p)\ (\ C_n^{AU}\,n + C_p^{AU}\,p\ ) \tag{1.4.28}$$

Here C_n^{AU} and C_p^{AU} are the Auger constants.

3. The avalanche generation

$$GR_{AV} = \frac{1}{q}\ (\ \alpha_n|\vec{J}_n| + \alpha_p|\vec{J}_p|\ ) \tag{1.4.29}$$

Here α_n and α_p are the impact ionization rates.

1.5 Outline of this thesis

This thesis will focus on the study of a hydrodynamic device model in which electron transport is dominant.

In chapter 2 special attention is given to the determination of the transport parameters as a function of local electron energy and doping concentration. An Ensemble Monte Carlo simulation program has been developed for this goal. In this chapter the Monte Carlo method is described in detail. The main questions we try to answer in this chapter are:

- Does a multi-valley model of the Si conduction band suffice for our purpose.
- What model for the ionized impurity scattering and the impact ionization scattering can best be used in the Monte Carlo program.
- How do the transport parameters have to be defined to get a consistent set of parameters.
- How can the transport parameters be described as a function of local doping concentration and local electron energy.

- What effect has an interface with an applied field perpendicular to this interface on the transport parameters.

In chapter 3 the Monte Carlo method is extended to a one-dimensional drift device simulator. The program is used as a reference for the hydrodynamic device simulator that is described in the next chapter. Special attention has been given to:

- The modeling of the ohmic contacts.
- To avoid that virtually all simulation electrons pile up in the highly doped contact regions, the simulation electrons in the contact regions represent a larger number of real electrons than in the other regions.
- The study of velocity and energy distributions in one-dimensional drift devices.

In chapter 4 a time-dependent hydrodynamic one-dimensional drift-device simulator is described. To study the validity of the hydrodynamic method, simulation results of this model are compared with simulation results of the Monte Carlo simulator. Major topics that are studied:

- How well compare the results obtained by the Monte Carlo method with the results obtained by the hydrodynamic method.
- How important are the inclusion of the velocity convection term $(\vec{J}_n \cdot \nabla_r)$ $(\vec{J}_n/n)$ and the energy convection term $\frac{1}{2}\, m^* |\vec{v}|^2$.
- What definition of the heat-flow term $\vec{Q}_n$ is the best to use in device simulators.

In chapter 5 the one-dimensional hydrodynamic device simulator is extended to a quasi two-dimensional device simulator suited for the simulation of n-channel MOSFETs. Special attention is given to:

- The calculation of the thickness of the inversion channel and the effective gate-voltage correction due to quantization effects.
- The simulation of MOS devices in the deep-submicron range. Simulation results are compared to experimental data from the literature and provided by Philips Research Laboratories, which gives an indication of the usefulness of the model proposed. Furthermore, the results will be used to study the range of applicability of the drift-diffusion model compared to the hydrodynamic model.
- An evaluation of the weak-avalanche parameters.

In chapter 6 a study is presented of the hot-electron models in the commercially available two-dimensional MOS device simulator MINIMOS and the two-dimensional hydrodynamic simulator TRENDY (Van Schie [1990] and Wolbert [1991]). Attention is given to:

- How well are the hot-electron effects calculated by the various simulators.
- Does a drift-diffusion simulator combined with a post-processing routine for calculating the hot-electron effects, offer a fast, robust, and above all, useful alternative to the hydrodynamic method.
- How well does the analytical method to calculate the electron temperature developed by Slotboom [1991] compare to more detailed numerical methods.

Finally in chapter 7 our final conclusions and advice for further research are given.

Chapter 2

THE TRANSPORT PARAMETERS

2.1 Introduction

To solve the hydrodynamic equations describing electron transport a number of transport parameters are needed as a function of local electron energy. The parameters of interest are the mobility μ_n, the momentum relaxation time τ_{pn} and the energy relaxation time τ_{wn}, the conductivity effective mass m_n^*, the diffusion coëfficient D_n and the impact ionization rate α_n. All parameters are calculated by analyzing the steady-state results of a large number of simulations of electron transport under different applied electric fields. The simulation method used is the Monte Carlo method.
The name Monte Carlo is used for any method of solution where random numbers are involved. Since the early seventies this method has also been applied to electron transport in semiconductor devices. The Monte Carlo method is used to solve the Boltzmann transport equation given by equation (1.4.8) without any a priori assumptions of the distribution function. The method in principle can be applied to any material. The pioneers have been Kurosawa [1966] and Fawcett, Boardman and Swain [1970] who applied the method to bulk GaAs. Initially only one simulation electron was taken into account, and its behavior was followed during a long period of time. Due to the increase in computing and storage capability of available computers, the Monte Carlo method later was applied to a large number of electrons, called the Ensemble Monte Carlo method. An excellent review of the Monte Carlo method in semiconductors was given by Jacoboni and Reggiani [1983]. In section 2.2 of this chapter the basic principles of the Ensemble Monte Carlo method are discussed. Section 2.3 discusses how the Monte Carlo method has to be applied to Si. In particular the band structure of Si and the scatter mechanisms in this material have to be considered in detail.

In section 2.4 methods to extract the transport parameters from the Monte Carlo simulation results in a consistent way are presented. The resulting transport parameters in the bulk are given in section 2.5. In addition the effects of the Si-SiO_2 interface on the above parameters is studied. In section 2.6 the conclusions are summarized. Finally section 2.7 can be considered an appendix to this chapter.

2.2 Basic principles of the Ensemble Monte Carlo method

The basic idea behind the Ensemble Monte Carlo method is to follow on a time step basis the trajectories of an ensemble of simulation carriers. The number of carriers in the program typically is several thousands. The Monte Carlo approach of carrier transport is based on the assumption that the motions of the fixed ions in the lattice around their equilibrium position and the motions of the free carriers are decoupled. Now the carriers during their flight do not interact continuously with the lattice, but only at discrete moments in time. Furthermore it is assumed that the interaction has effect only on the carriers , and that there is negligible effect on the lattice. In this way the history of each carrier is a series of free flights during which the carrier is accelerated due to the forces applied to the carriers, followed by a scatter event.

A basic element in the Monte Carlo method is the use of random numbers in the program to make decisions (hence the name Monte Carlo). In this program pseudo random numbers r with an uniform distribution on [0,1) were produced by the mixed congruential scheme (Hammersley and Handscomb [1964]). It has been proven that this scheme for all practical purposes performs satisfactorily (Graybeal and Pooch [1980]).

In this thesis only electrons are taken into account in the Monte Carlo method and no holes. During free flight the external forces $\vec{F}_E$ operate on the free electron, and it is assumed that $\vec{F}_E$ is only due to the electric field $\vec{E}$. Then the time derivative of the wave vector of a given carrier is given by

$$\frac{d\,\vec{k}}{d\,t} = \frac{1}{\hbar}\vec{F}_E = -\frac{q}{\hbar}\vec{E} \tag{2.2.1}$$

Each free flight ends by an abrupt change in $\vec{k}$-space from one state to another when the carrier is subject to a scatter process. All scatter processes are considered as being instantaneous and then for each scatter process k a rate of transition $S_k(\vec{k}_i,\vec{k}_f)$ from initial state $\vec{k}_i$ to final state $\vec{k}_f$ is defined (see section 2.3). Now the probability that an electron in state $\vec{k}_i$ scatters to any other state is given by the scatter rate $\lambda_k(\vec{k}_i)$, being the integration of S_k over the momentum space V_k

$$\lambda_k(\vec{k}_i) = \frac{V}{(2\pi)^3}\int_{V_k} S_k(\vec{k}_i,\vec{k}_f)\ d\vec{k}_f \tag{2.2.2}$$

The factor $V/(2\pi)^3$ accounts for the density of states in $\vec{k}$-space. Since the spin is not altered by scattering it is not necessary to include an extra factor of two. The above formula is only valid for non-degenerate materials. Due to the high density of states in Si, degeneracy is not relevant in most devices and is not considered in this work.

Once the scatter process is chosen, the new wave vector $\vec{k}_f$ can be calculated. In this way the wave vector $\vec{k}$ is known at each moment of time, and if the band structure of the material concerned is known, knowledge of the wave vector of a given carrier leads to knowledge of the kinetic energy u of that carrier, which leads to the carrier group velocity $\vec{w}$ via the equation

$$\vec{w} = \frac{1}{\hbar} \left(\frac{\partial}{\partial k_x}, \frac{\partial}{\partial k_y}, \frac{\partial}{\partial k_z} \right) u \qquad (2.2.3)$$

By integration of $\vec{w}$ the real space position $\vec{x}$ of each carrier is obtained. If $\vec{k}$, u, $\vec{w}$ and $\vec{x}$ of all carriers are known the drift velocity $\vec{v}_n$, the average carrier energy W_n, the carrier temperature T_n, the conductivity effective mass m_n^*, the momentum and energy relaxation times τ_{pn} and τ_{wn}, and the diffusion coefficient D_n can be calculated.

The strength of the Monte Carlo method lies in its ability to deal with complex band structures and various complex scatter mechanisms. Definite knowledge about scatter processes is combined with random numbers which decide if a scatter process occurs, and if so, random numbers also decide exactly which scatter process occurs and how strong its effect is on the carrier.

Disadvantages are that statistical errors tend to accumulate when long periods of time are simulated. In addition extensive computing and storage power is needed. Especially at weak fields the thermal noise generated is very strong compared to the drift velocity and very large numbers of simulation carriers or very long simulation times are needed to find representative averages.

2.2.1 Free-flight calculations

Each free flight of an electron starts at the given initial state of that electron. The force field $\vec{E}$ applied to the electron changes the kinetic energy, momentum and (in an elliptic, non-parabolic valley) also the effective mass of the electron continuously during that free flight. Now if the total scatter rate is given by $\Lambda(\vec{k}) = \sum \lambda_k(\vec{k})$, then the probability that the electron will suffer its next collision during dt around time t since the last collision is given by (Jacoboni and Reggiani [1983])

$$\mathcal{P}(t)dt = \Lambda(\vec{k}(t)) \exp \left(-\int_0^t \Lambda(\vec{k}(t'))\, dt' \right) dt \qquad (2.2.4)$$

As $\vec{k}$ changes constantly due to the external forces, $\Lambda(\vec{k})$ is not an analytical function of $\vec{k}$, and (2.2.4) can only be evaluated at high computational costs. To simplify the calculations a new fictitious self-scatter process was introduced by Rees [1968, 1969]. If Γ is the maximum value of $\Lambda(\vec{k})$ in the region of $\vec{k}$-space of interest, the self-scatter rate is chosen in such a way that the total scatter probability, including this self-scattering, is constant and equal to Γ. If the carrier is subject to such a self-scattering, its state $\vec{k}$ after the collision is taken to be equal to its state $\vec{k}$ before the collision, so that in practice the electron path continues unperturbed as if no scattering at all had occurred. The inclusion of the self-scattering does not change the rate of occurrence of real scatter processes, as was proven by Fawcett et al. [1970]. Introducing this self-scatter process, equation (2.2.4) is simplified to

$$\mathcal{P}(t) = \Gamma \exp(-\Gamma t) \tag{2.2.5}$$

and a random number r is used to generate the free-flight duration t_r

$$t_r = -\frac{1}{\Gamma} \ln (r) \tag{2.2.6}$$

The extra computing time needed in taking care of the self-scattering is more than compensated for by the simplification of the calculation of the free-flight duration.

In general the scatter rate $\Lambda(\vec{k})$ in state $\vec{k}$, is an increasing function of kinetic energy u and in the self-scatter method is taken $\Gamma = \Lambda(u_m)$, where u_m is the maximum kinetic energy with a negligible probability of being reached by the carrier during simulation. However this gives a very large value of Γ which makes the simulation very inefficient.

For a more efficient estimate of Γ the maximal kinetic energy obtained during a particular free flight has to be known a priori. This problem can be solved using iterative techniques. One method using an iterative approximation has been given by Hockney and Eastwood [1981], which can be used if $\Lambda(u)$ is monotonic in u.

The method proposed by Hockney starts by estimating the value of Γ from the initial value of the energy u_0; then with equation (2.2.6) the time of free flight t_r and the final energy u_1 at the end of the free flight are calculated; now if $\Gamma < \Lambda(u_1)$, the process is restarted with the new $\Gamma = \Lambda(u_1)$; this procedure is repeated until $\Gamma \geq \Lambda(u_1)$, using every time the same random number to calculate t_r. In practice the number of iterations is on the average only two or three. A problem of this approach is that Γ is now an implicit function of the random number r, which is formally forbidden (Jacoboni and Reggiani, [1983]). In a Monte Carlo simulation all carriers are followed during a series of time intervals. During each interval Δt all carriers scatter at least once (self-scattering at the end of the time

interval). After each time interval the ensemble of carriers is evaluated (e.g. calculation of the distribution function or the drift velocity). Our simulations have shown that if the Hockney algorithm is used, the results appear to become dependent on the value of Δt, which certainly cannot be explained from a physical point of view.
To avoid this problem a different approach is used. Again a first estimate of Γ is obtained from the initial value of the kinetic energy. A second estimate is obtained from the energy that the electron would get if it were allowed to proceed its free flight up to the end of the time interval. As final value of Γ the maximum of the two estimates is taken. This approach does not show any dependency of simulation results on the time interval Δt. If Δt is below the average free-flight duration, both methods give the same results. Although the Hockney algorithm is more efficient, and correct as long as the time interval is small enough, the second algorithm has been preferred and used in our program.

For most scatter processes the scatter rates are a function of kinetic energy, which makes it easy to tabulate them. In addition, the maximum sum of scatter rates in each energy interval u to $u + \delta u$ ($u \in [0,u_{max}]$) is calculated and tabulated. In this way it is not necessary to calculate Γ at each electron free flight and considerable computing time is saved. However not all scatter processes can be tabulated. In the ionized impurity scattering model and the electron-electron scattering model there is also a dependence on the electron temperature, which is an average value over all electrons. For this reason those two scatter processes have not been tabulated, but are calculated each time over and again.

2.2.2 Selection of the scatter process

Once it is decided how long the free flight will be, the wave-vector $\vec{k}_i$ and the kinetic energy u_i just before the scattering occurs are calculated. To decide to which scatter process the electron is subject, the maximum total scatter rate Γ is used. Now given a random number r, the product $r\Gamma$ is compared with successive sums of the scatter rates $\lambda_k(u_i)$ and a scatter mechanism is selected. A table of scatter rates is used, which saves much computing effort. Only when the electron energy exceeds u_{max} of the table, all scatter rates have to be calculated. Of course u_{max} is chosen large enough to ensure this occurring only very occasionally. If none of the scatter mechanisms has been selected, a self-scattering occurs.

After the scatter process is known, the final energy u_f is calculated, and from u_f also $|\vec{k}_f|$. The normalized cumulative distribution $P_\vartheta(\vartheta)$ for the angle ϑ (figure 2.1) which determines the spatial orientation of $\vec{k}_f$ with respect to

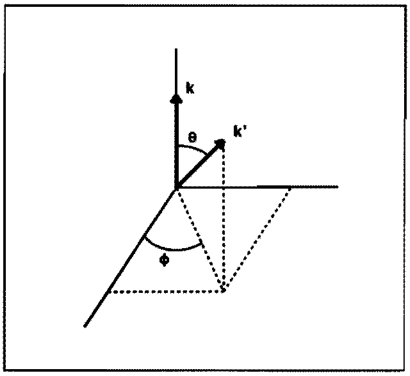

Fig 2.1 The scatter angles θ and φ that determine the spatial orientation of $\vec{k}'$ with respect to $\vec{k}$.

$\vec{k}_i$ is given by

$$P_\theta(\theta) = \frac{V}{(2\pi)^3 \lambda_k(\vec{k}_i)} \int_0^\infty \int_0^\theta \int_0^{2\pi} S_k(\vec{k}_i,\vec{k}_f)\, k_f^2 \sin\theta_f \, dk_f d\theta_f d\phi_f \qquad (2.2.7)$$

With the help of a random number r_θ in [0,1), the new value θ is calculated

$$\theta = P_\theta^{-1}(r_\theta) \qquad (2.2.8)$$

The distribution over the angle φ is always uniform and therefore, given a random number r_φ in [0,1)

$$\varphi = 2\,\pi\, r_\varphi \qquad (2.2.9)$$

2.3 Band structure and scatter processes

2.3.1 Band structure

In a Monte Carlo program complicated band structures can be incorporated. Most commonly in semiconductors the band is considered as a set of valleys all modeled by the effective mass approximation. In Si this multi-valley model already has proven to yield adequate results if the applied electric field is not too high (E < 100 kV/cm), e.g. Canali et al. [1975]. However, even at high fields the multi-valley model can be used if high energy valleys and impact ionization scattering are included. These two mechanisms strongly reduce the electron energy at high fields and only very few electrons ever gain an energy higher than 2 eV. Then the effective mass approximation is still valid. Furthermore the multi-valley model is easy to implement in a Monte Carlo program and yields results fast.

In our simulations the multi-valley model has been preferred to a more realistic band model as described by Tang and Hess [1983] and Fischetti and

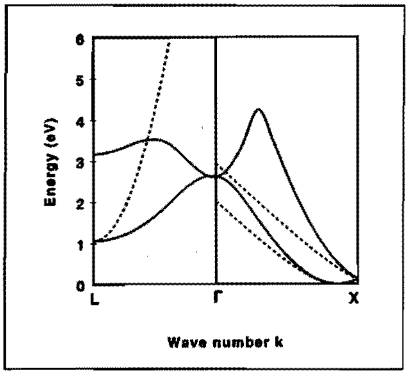

Fig 2.2 First two conduction bands of Si. The solid lines represent empirical pseudo-potential calculations (Tang and Hess [1983]); the dashed lines represent the multi-valley model.

Laux [1988], who use the empirical pseudo-potential method to describe the conduction bands. Their band structure has been plotted in figure 2.2 (solid lines). The multi-valley model is given in the same figure (dashed lines). More comments on this choice of band model is given in section 2.3.3.

Basically the effective mass approximation gives a parabolic dispersion relation in each valley but a non-parabolicity factor α is introduced according to Kane [1957]

$$u(1 + \alpha u) = \frac{\hbar^2}{2} \left(\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y} + \frac{k_z^2}{m_z} \right) \tag{2.3.1}$$

In first approximation the conduction band of Si consists of six X-valleys distributed along the principal axes. This valley can clearly be observed in figure 2.2. For all practical purposes two valleys lying on the same axis are indistinguishable, so the number of X-valleys used in the program is reduced to three, each having a multiplicity of two. The X-valleys are strongly elliptic: $m_{Xl} = 0.9163\ m_0$, and $m_{Xt} = 0.1905\ m_0$. Like Jacoboni and Reggiani [1983] a non-parabolicity factor $\alpha_X = 0.5$ has been chosen.

Since our interest is in high energy effects it is necessary also to include in the model the four valleys that lie along the (1,1,1) directions, the so called L-valleys. Data on the L-valley were obtained from the data by Tang and Hess [1983]. They give a 1.06 eV gap between the bottom of the L- and X-minima. The data from Hess show the L-valley to be strongly anisotropic with $m_{Ll} = 1.53\ m_0$ and $m_{Lt} = 0.16\ m_0$. Although the L-valley can be described accurately, there is little quantitative knowledge on the scatter rates into and out of the L-valleys. So a detailed description of the L-valleys serves no purpose and to keep our program manageable the L-valleys have been considered spherical with an effective mass $m_L = 0.3\ m_0$, which is the density of state effective mass of the L-valleys.

$a_0 = 5.43$ Å	$\epsilon_r = 11.7$		
$\rho_s = 2329$ kg/m^3	$p_{ii} = 0.01$		
$v_s = 9040$ m/s			
X-valley (Jacoboni et al. [1983]		*L-valley (Tang and Hess [1983])*	
$u_{GAP} = 1.107$ eV		$u_{XL} = 1.06$ eV	
$m_{Xl} = 0.9163\ m_0$		$m_L = 0.3\ m_0$	
$m_{Xt} = 0.1905\ m_0$			
$\alpha_X = 0.5$ eV^{-1}		$\alpha_L = 0.0$ eV^{-1}	
$\Xi_{AX} = 9.0$ eV		$\Xi_{AL} = 9.0$ eV	
$\theta_{f1} = 220$ K	$\Xi_{f1} = 3.0\ 10^{9}$ eV/m	$\theta_{XL1} = 672$ K	$\Xi_{XL1} = 3.0\ 10^{10}$ eV/m
$\theta_{f2} = 550$ K	$\Xi_{f2} = 2.0\ 10^{10}$ eV/m	$\theta_{XL2} = 634$ K	$\Xi_{XL2} = 3.0\ 10^{10}$ eV/m
$\theta_{f3} = 685$ K	$\Xi_{f3} = 2.0\ 10^{10}$ eV/m	$\theta_{XL3} = 480$ K	$\Xi_{XL3} = 3.0\ 10^{10}$ eV/m
$\theta_{g1} = 140$ K	$\Xi_{g1} = 5.0\ 10^{9}$ eV/m	$\theta_{XL4} = 197$ K	$\Xi_{XL4} = 3.0\ 10^{10}$ eV/m
$\theta_{g2} = 215$ K	$\Xi_{g2} = 8.0\ 10^{9}$ eV/m		
$\theta_{g3} = 720$ K	$\Xi_{g3} = 11.0\ 10^{10}$ eV/m		

Table 2.1: Scatter, band and material parameters of Si.

Now the L-valleys are indistinguishable for all practical purposes and in our model the number of L-valleys is reduced to one with a multiplicity equal to four. Furthermore, the L-valley included in this model is considered parabolic, so $\alpha_L = 0$. Table 2.1 gives an overview of all band parameters. Also included are the scatter parameters that are discussed in more detail in the next section.

The X-valleys in Si are strongly anisotropic and have to be dealt with by transforming $\vec{k}$-space to $\vec{\kappa}$-space by the transformation proposed by Herring and Vogt [1956]. Each valley has its own transformation matrix T_i, defined by

$$T_i = m_0^{1/2} \begin{pmatrix} m_{ix}^{-1/2} & 0 & 0 \\ 0 & m_{iy}^{-1/2} & 0 \\ 0 & 0 & m_{iz}^{-1/2} \end{pmatrix} \tag{2.3.2}$$

Defining in the *i*-th X-valley $\vec{\kappa} = T_i \vec{k}$ the dispersion relation (2.3.1) reads

$$u(\ 1 + \alpha_X u\ ) = \frac{\hbar^2}{2m_0}\,\kappa^2 \tag{2.3.3}$$

This means that each of the X-valleys is spheric in the $\vec{\kappa}$-u relation with an isotropic effective mass m_0. The Herring-Vogt transformation has to be extended to equations (2.1.1) and (2.1.2), which give in $\vec{\kappa}$-space

$$\frac{d\vec{\kappa}}{dt} = -\frac{q}{\hbar} T_i \vec{E} \tag{2.3.4}$$

$$\vec{w} = \frac{\hbar}{m_0(1 + 2\alpha_X u)} \vec{T}_i \kappa \tag{2.3.5}$$

Since the transformation matrix T_i is valley dependent an important consequence of the Herring-Vogt transformation is that the effective field $\vec{E}_i = T_i \vec{E}$ differs in each valley. And since The L-valleys are considered isotropic the transformation matrix is defined (I is the identity matrix) by

$$T_L = (m_0/m_L)^{1/2}\, I \tag{2.3.6}$$

In general we are interested in the properties of the system in the direction of the applied electric field. This direction is not necessarily parallel to one of the principal axes of the lattice. Now two independent coordinate systems are chosen: the external coordinate system (x',y',z') with unit vectors $\vec{e}_{x'}$, $\vec{e}_{y'}$ and $\vec{e}_{z'}$, and the internal system (x,y,z) with unit vectors $\vec{e}_x$, $\vec{e}_y$ and $\vec{e}_z$. The internal system has its axes in the direction of the principal axes of the lattice, while the external system is defined such that the externally applied electric field is aligned to the x'-axis, so $\vec{E}' = E\,\vec{e}_{x'}$. Now an orthonormal transformation matrix U, where $\vec{x}' = U\,\vec{x}$, is needed. Two orthonormal vectors $\vec{d}_x$ and $\vec{d}_y$ are chosen to define the transformations $\vec{e}_{x'} = d_{xx}\vec{e}_x + d_{xy}\vec{e}_y + d_{xz}\vec{e}_z$ and $\vec{e}_{y'} = d_{yx}\vec{e}_x + d_{yy}\vec{e}_y + d_{yz}\vec{e}_z$. Now define $\vec{d}_z = \vec{d}_x \times \vec{d}_y$, the matrix U can be given by

$$U = \begin{pmatrix} d_{xx} & d_{xy} & d_{xz} \\ d_{yx} & d_{yy} & d_{yz} \\ d_{zx} & d_{zy} & d_{zz} \end{pmatrix} \tag{2.3.7}$$

Basically we are interested in the external variables, while most calculations deal with the internal variables. Using $\vec{k}' = U\,\vec{k}$, $\vec{w}' = U\,\vec{w}$ we find the final set of equations

$$u(1 + \alpha_X u) = \frac{\hbar^2}{m_0}\kappa^2 \tag{2.3.8}$$

$$\frac{d\vec{\kappa}}{dt} = -\frac{q}{\hbar}\vec{E}_i \qquad \text{with } \vec{E}_i = T_i U^T \vec{E}' \tag{2.3.9}$$

$$\vec{k}' = U_{1i}\vec{\kappa} \qquad \text{with } U_{1i} = U\,T_i^{-1} \tag{2.3.10}$$

$$\vec{w}' = \frac{1}{1+2\alpha_X u}\frac{\hbar}{m_0} U_{2i}\vec{\kappa} \qquad \text{with } U_{2i} = U\,T_i \tag{2.3.11}$$

2.3.2 Phonon scatter processes

Two different types of phonon scatter processes are relevant in Si: acoustic phonon scattering and non-polar optic phonon scattering. These processes are described here. Because the valleys in Si are considered elliptic, a Herring-Vogt transformation is used to transform $\vec{k}$-space to $\vec{\kappa}$-space (see section 2.3.1). Now the transition probability S_k and the scatter rate λ_k for each scatter process k also have to be defined in $\vec{\kappa}$-space in analogy with equation (2.2.2).

2.3.2.1 Acoustic phonon scattering

In this type of scatter process little energy is involved and the process is always an intra-valley process, i.e. the electron does not change valley. In general this process can be considered to be elastic. However this assumption breaks down at very low lattice temperatures (few degrees K) or at very high fields. The first situation is of no concern to us, but the second is. However, at high fields the optic scatter processes become dominant and the exact value of the acoustic scatter process has little importance. Therefore the elastic model can be used even at high fields (Jacoboni and Reggiani [1983]). The transition probability in $\vec{\kappa}$-space from the initial wave vector $\vec{\kappa}_i$ (with energy u_i) to the final wave vector $\vec{\kappa}_f$ (with energy u_f) was given by Rode [1978] and reads

$$S(\vec{\kappa}_i,\vec{\kappa}_f) = \frac{2\pi}{\hbar}\;\frac{2\pi\,\Xi_A^2\,k_B T_0}{2V\,\rho_s v_s^{\,2}}\;\mathcal{G}(\vec{\kappa}_i,\vec{\kappa}_f)\;\delta(u_i - u_f) \tag{2.3.12}$$

Here T_0 represents the lattice temperature, Ξ_A the deformation potential of acoustic scattering, ρ_s the crystal density and v_s the velocity of sound in Si and V the crystal volume. In case of a parabolic band structure, the overlap factor $\mathcal{G}(\vec{\kappa}_i,\vec{\kappa}_f)$ equals one. But in a Kane type band structure $\mathcal{G}$ reads

$$\mathcal{G}(\vec{\kappa}_i,\vec{\kappa}_f) = \left\{ \sqrt{\frac{(1+\alpha u_i)(1+\alpha u_f)}{(1+2\alpha u_i)(1+2\alpha u_f)}} + \sqrt{\frac{\alpha^2 u_i u_f}{(1+2\alpha u_i)(1+2\alpha u_f)}}\;\cos\vartheta \right\}^2 \tag{2.3.13}$$

Here ϑ is the angle between the initial and final state $\vec{\kappa}_i$ and $\vec{\kappa}_f$, and in the X-valley the acoustic scatter rate λ_A as a function of energy u reads

$$\lambda_A(u) = 2\;\frac{(m_{Xd} m_0)^{3/2}\,(\Xi_A q)^2\,k_B T_0}{\rho_s v_s^2\,\pi\,\hbar^4\,\sqrt{2}}\;[u(1+\alpha u)]^{1/2}\;\frac{(1+\alpha u)^2+(\alpha u)^2/3}{1+2\alpha u} \tag{2.3.14}$$

In which the density of state effective mass in the X-valley m_{Xd} is given by

$$m_{Xd} = (m_{Xl} m_{Xt}^2)^{1/3} \tag{2.3.15}$$

Using the random number r, the scatter angle θ can be calculated

$$\theta = \arccos\left[\frac{[\,(1+2\alpha u)^3\ (1-r) + r\,]^{1/3} - 1}{\alpha u} - 1 \right] \tag{2.3.16}$$

The values of the different material parameters ρ_s ,v_s and Ξ_A are given in table 2.1.

2.3.2.2 Non-polar optic phonon scattering

Two modes of scattering are present, either by absorption or by emission of a phonon, and scatter processes take place in the L-valley as well as in the X-valley. Now all scatter processes are inter-valley, i.e. electrons move from one valley to another. The overlap factor is given by

$$\mathcal{G}(\vec{\kappa}_i,\vec{\kappa}_f) = \frac{(1+\alpha_i u_i)(1+\alpha_f u_f)}{(1+2\alpha_i u_i)(1+2\alpha_f u_f)} \tag{2.3.17}$$

Because the scattering is inter-valley and each valley does not necessarily have the same non-parabolicity factor α, a distinction is made between the initial non-parabolicity factor α_i and the final α_f. Then the scatter rate is (Fawcett et al. [1970]) given by

$$\lambda_{op}(u) = \frac{Z_v \Xi_{op}^2\ (m_0 m_d)^{3/2} N_{op}}{\pi \rho_S \sqrt{2}\ \hbar^2 k_B \theta_{op}}\ [u_f(1+\alpha_f u_f)]^{1/2}\ \frac{(1+\alpha_i u_i)(1+\alpha_f u_f)}{1+2\alpha_i u_i} \tag{2.3.18}$$

Here Ξ_{op} represents the optical coupling constant, θ_{op} the equivalent optical phonon temperature and Z_v is the number of the valleys attainable by the electron. If the scattering occurs by emission of a photon the energy after scattering is given by $u_f = u_i - k_B\theta_{op} - \Delta u$, and in case of absorption by $u_f = u_i + k_B\theta_{op} - \Delta u$. Δu represents the energy difference between initial and final valley. The phonon number N_{op} is given by (2.3.19) in case of emission and by (2.3.20) in case of absorption

$$N_{op} = \frac{\exp\left(\frac{k_B\theta_{op}}{k_B T_0}\right)}{\exp\left(\frac{k_B\theta_{op}}{k_B T_0}\right) - 1} \tag{2.3.19}$$

$$N_{op} = \frac{1}{\exp\left(\frac{k_B\theta_{op}}{k_B T_0}\right) - 1} \tag{2.3.20}$$

The overlap factor $\mathcal{G}$ does not depend on the angle θ between $\vec{\kappa}_i$ and $\vec{\kappa}_f$ and neither is the transition rate, so the scattering is uniform and

$$\theta = \arccos (2r - 1) \tag{2.3.21}$$

Three kinds of inter-valley scattering have been taken into account, g-scattering, f-scattering and X-L-scattering, all of which are caused either by emission or by absorption.

1. g-scattering: in this scatter process an electron from one X-valley scatters to the other X-valley that is situated on the same axis. The number of attainable valleys $Z_v = 1$ and the energy shift $\Delta u = 0$. There are three different g-scatter processes involved, g_1, g_2 and g_3 each with its own phonon energy and coupling constant.
2. f-scattering: this is the process complementary to g-scattering, in which an electron in an X-valley scatters to another X-valley, not situated on the same axis. Now $Z_v = 4$ and $\Delta u = 0$. All four valleys have equal probability to be attained by the scattering electron. There are three different f-scattering processes.
3. X-L scattering: this process deals with the scattering between the X-valleys and the L-valleys. In case an electron scatters from an L-valley to an X-valley then $Z_v = 6$ and $\Delta u = -u_{XL}$; u_{XL} is the energy gap between the bottom of the L-valley and the X-valley. If an electron scatters from X- to L-valley $Z_v = 4$ and $\Delta u = u_{XL}$. All valleys have equal probability to be attained by the scattering electron. There are four different X-L scattering processes (Tang and Hess [1983]).

Probably there is also some inter-valley scattering among the L-valleys. However, since no data is available it is not considered here. The values of θ_{op} and Ξ_{op} in case of f- or g-scattering were obtained from Canali et al. [1975], and in case of X-L scattering they were obtained from Tang and Hess [1983]. The scatter rates have been plotted in figure 2.3.

2.3.3 Discussion of the band structure and the phonon scatter parameters

In the previous two sections the model of the conduction band and the phonon scatter processes have been described and values have been assigned to the various parameters. In figure 2.4 the drift velocity in intrinsic Si calculated by our Monte Carlo program is compared to the experimental results obtained by Canali et al. [1971]. Good agreement is obtained for $E < 100$ kV/cm. However, a discrepancy occurs for higher fields. This probably is due to the simplifications made in the model of the L-valley and the X-L scatter mechanism. In order to obtain a better fit, the effective mass and non-parabolicity factor of the L-valley, the X-L scatter parameters and the

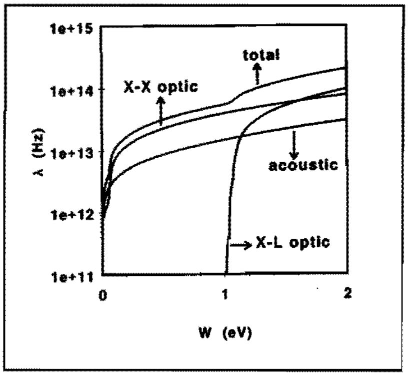

Fig 2.3 Various scatter rates in intrinsic Si.

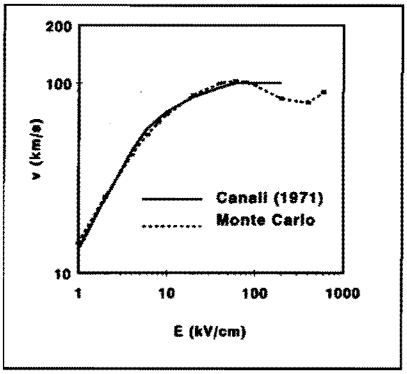

Fig 2.4 Electron velocity as a function of the applied electric field; experimental (solid line) and simulated data.

impact ionization parameter (see section 2.3.4.1) have been modified. However, no improvement was obtained.

A second comment has to be made on the overlap factor $\mathcal{G}$. Some authors always set $\mathcal{G}$ to one (e.g. Zimmermann [1980]). However, simulations show that $\mathcal{G}$ has a considerable effect on the electron energy at high fields: the energy is increased by some 20%. Therefore in all our simulations we use the $\mathcal{G}$ described by equations (2.3.13) and (2.3.17).

2.3.4 Non-phonon scatter processes

Electrons not only interact with the phonons, but are also subject to several other scatter processes. The most important ones are dealt with in this section: impurity scattering, impact ionization scattering, electron-electron scattering and interface scattering.

2.3.4.1 Impurity scattering

Impurities in the crystal cause extra scatter processes. In the first place they disturb the lattice, causing neutral impurity scatter processes. In general these have a small effect on the carriers and are difficult to describe. They have not been taken into account here. More important is the scatter process due to the charge of the impurities, the ionized impurity scattering.

Several methods to model this scatter process have been proposed:

1. The model proposed by Conwell and Weisskopf [1950] uses the Born approximation and is essentially a two-body model. The Coulomb field is cut off at a radius equal to one half the mean distance between the scattering centers.

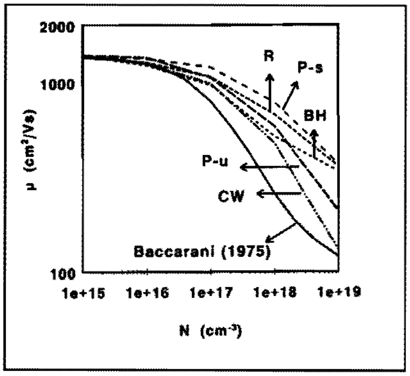

Fig 2.5 Low-field mobility as a function of the ionized impurity concentration for various models of the scatter mechanism: Conwell-Weisskopf (CW), Brooks-Herring (BH), Ridley (R), and screened and unscreened Perri model (P-s and P-u). The solid line represents experimental data.

2. The Conwell-Weisskopf model was improved by Brooks and Herring (Brooks [1955]). They introduced screening into the model, thereby allowing the scattering potential to fall with distance more rapidly. A complete overview of the Conwell-Weisskopf and Brooks-Herring models was given by Chattopadhyay and Queisser [1981].
3. If the average distance between the scattering centers is smaller than the screening length, more than one center can play a role and the Brooks-Herring approach is in error, as it simply adds the effects of all centers. A refinement of the model was given by Ridley [1977] by introducing the restriction that only one scattering center is acting on the electron. The model contains both the Conwell-Weisskopf and the Brooks-Herring models as limiting cases and was first used in a Monte Carlo simulator by Van de Roer and Widdershoven [1986]. An extensive description is given by Nederveen [1989]. The equations needed are given in the appendix 2.7.1.
4. A fourth model was proposed by Perri [1986]. In this model the electron is scattered only by the nearest scattering center. All other centers are assumed to be screened completely. This model comes in two versions: one in which the scattering angle is calculated assuming the scattering center to be unscreened, the second in which a screening of the scattering center is assumed to be present.

All above models have one serious draw-back in common: they do not fit experimental data. All of them give too high values for the low-field mobility as is shown in figure 2.5. In this figure the solid line represents experimental data, the dashed lines simulation data using different ionized impurity models. In all simulations full ionization of the impurities was assumed, which is in fact not correct. However, accounting for this would make matters even worse. In addition to the possibility that the models are just not good enough, some other explanations can be given:

1. In the models the distortion of the lattice by the doping atoms has not been taken into account. However this effect is expected to become appreciable only at very high doping concentrations ($N_I > 10^{19}$ cm^{-3}) and the models fail already at much lower concentrations.
2. Electron-electron interaction has not yet been taken into account. Especially at high doping levels (which also means high electron concentration in the calculations done) this could be an important scattering mechanism. This aspect is dealt with in section 2.3.4.3. However simulations show only a minor effect of electron-electron scattering on the low-field mobility.
3. Another contribution to the discrepancy could be that in practice the ions are not distributed homogeneously in the crystal; some clustering of ions is to be expected (Widdershoven [1988]). An approximate model of the possible effects of clustering is described further on: see (2.3.28).

In our simulations the Perri model that assumes one unscreened scattering center has been used because the results obtained by this model fit experimental data better than most other models. Furthermore, if the unscreened Perri model is used in a Monte Carlo device simulator, this model has the important advantage that in the scatter process only the nearest impurity is included. In the other models also distant charges have effect on the scattering. But the effects of distant charges already have been taken into account by the electric potential distribution in the device, so there is some double-counting. Furthermore, the unscreened Perri model is easy to implement and easily can be extended with the clustering ion concept.

The unscreened Perri model is described here in more detail:
If N_I is the ionized impurity concentration, the average distance between scatter centers equals $N_I^{-1/3}$. Now assume that only one impurity acts on an electron moving through the lattice with velocity $|\vec{w}|$ at the moment of each interaction, then the averaged geometrical cross section is defined (see figure 2.6a) by

$$\sigma_s = \pi \left(\frac{N_I^{-1/3}}{2} \right)^2 \tag{2.3.22}$$

Now in an infinitesimal time step dt the movement of the electron is assumed to be in a straight line, and an elementary volume dV is defined (figure 2.6b)

$$dV = \sigma_s \, |\vec{w}| \, dt \tag{2.3.23}$$

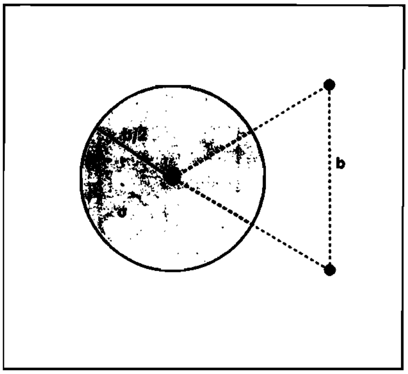

Fig 2.6a The average geometrical cross section is defined as that section that contains exactly one impurity.

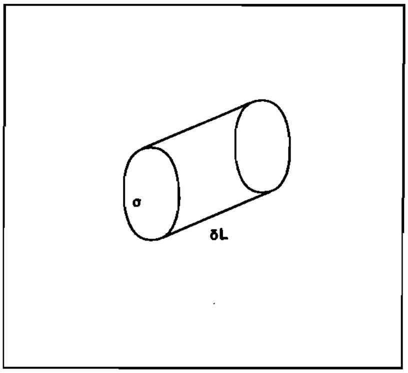

Fig 2.6b Elementary volume dV.

Now d*n*, the number of ions present in the volume d*V* is defined by $dn = N_I dV$. The scatter rate λ_P, which is equal to the number of ions d*n*/d*t* encountered by the moving electron per time unit, is given by

$$\lambda_P(u) = N_I \; \sigma_s \; |\vec{w}| = \frac{\pi}{4} \; |\vec{w}| \; N_I^{1/3} \tag{2.3.24}$$

The scattering angle between the electron wave vector before and after the scatter process is calculated by taking the Lagrangian of the electron-impurity system assuming a pure Coulomb potential. In a parabolic band this gives (McQuarrie [1976])

$$\theta = \arccos \left(\frac{64 \; \pi^2 \epsilon^2 \sigma_I^2 u^2 - Z_I^2 q^4}{64 \; \pi^2 \epsilon^2 \sigma_I^2 u^2 + Z_I^2 q^4} \right) \tag{2.3.25}$$

Z_I is the multiplicity of the impurity. The impact parameter σ_I is calculated from the geometrical cross section by using a random number r

$$\sigma_I^2 = r \; \pi \; N_I^{-2/3} \tag{2.3.26}$$

If the band is non-parabolic according to equation (2.3.1) then the energy *u* has to be substituted by the modified energy u_1 (Perri [1986]) given by

$$u_1 = \frac{u \; (1 + \alpha u)}{(1 + 2\alpha u)^2} \tag{2.3.27}$$

The unscreened Perri model easily can be extended to include clustering of ions. Our clustering model has the basic assumption that ions possibly form clusters of two or three ions, and that these clusters are distributed homogeneously through the lattice. If the fraction of the ions that is in a

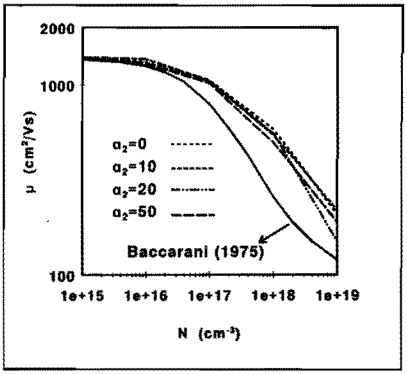

Fig 2.7 Low-field mobility as a function of the ionized impurity concentration; the solid line represents experimental data, the dashed lines simulated data using the unscreened Perri model; the parameter a_2 gives the percentage of the ions that form clusters of two.

cluster of two equals a_2 and the part that is in a cluster of three equals a_3 then the effective number of impurities is reduced

$$N_{I,eff} = [\ 1 - \tfrac{1}{2} a_2 - \tfrac{2}{3} a_3\ ]\ N_I \tag{2.3.28}$$

The effective scatter rate and the impact parameter σ_I now may be calculated by substituting $N_{I,eff}$ instead of N_I in (2.3.24) and (2.3.26). If a scatter process occurs, a random number decides whether a one, a two or a three ion cluster is involved, i.e. Z_I equals 1, 2 or 3. In figure 2.7 simulation results for different a_2 ratios is given (in this plot a_3 equals zero). Figure 2.7 shows that in general clustering does decrease the mobility a little, but the discrepancy between simulations and experiment can not be fully explained by the mechanism of clustering.

2.3.4.2 Impact ionization scattering

Electrons with sufficiently high energies are involved in impact ionization processes. This process can be treated as an additional scatter mechanism in Monte Carlo simulation. The impact ionization scatter rate is calculated by the empirical formula given by Keldysh [1965])

$$\lambda_{ii}(u) = p_{ii} \lambda_{ph}(u_{th}) \left(\frac{u - u_{th}}{u_{th}} \right)^2 \tag{2.3.29}$$

u_{th} is the threshold energy for impact ionization which is equal to the energy gap. $\lambda_{ph}(u_{th})$ is the total phonon scattering rate at the threshold energy. As electrons situated in the X-valleys are subject to different scatter mechanisms as those in the L-valleys, λ_{ph} has a different value in both cases. p_{ii} is a constant; following Tang and Hess [1983], the value $p_{ii} = 0.01$ has been used. An electron subject to an impact ionization process loses an amount of energy equal to u_{th}. Furthermore it is assumed that the scatter process is intra-valley. In figure 2.8 simulation results are

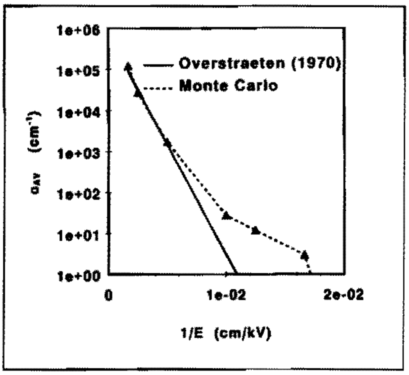

Fig 2.8 Impact ionization rate as a function of the inversed electric field in intrinsic Si; experimental data are represented by the solid line; simulated data by the dashed line.

compared to the experimental results obtained by Van Overstraeten and the Man [1970]. At high fields a good agreement has been found. At low fields the agreement is much less, but in that range the practical importance of the avalanche generation is not very great.

2.3.4.3 Electron-electron scattering

Not only do carriers interact with the lattice and impurities, but also with other carriers. A model was proposed by Lugli and Ferry [1983]. The assumptions made are that the electrons interact elastically only via a short-range screened Coulomb potential, and that the valleys are parabolic and spherical. In Si these last assumptions are invalid. The Herring-Vogt transform provides spherical valleys in $\vec{\kappa}$-space, but a serious problem remains: the electron-electron scattering essentially absorbs neither energy nor momentum, but the latter condition is not fulfilled in $\vec{\kappa}$-space. The model used here is a modification of the model proposed by Brunetti et al. [1985] who introduced an extra self-scatter rate for electron-electron scattering. Then the total electron-electron scatter rate is given by

$$\lambda_{ee}(u) = \frac{n\, m_{Xd} q^4}{8\pi\, \epsilon_r^2 \epsilon_0^2 \hbar^3 \beta_s^3} \tag{2.3.30}$$

Here β_s is the inverse Debye length associated with screening

$$\beta_s = \left(\frac{n\, q^2}{\epsilon_r \epsilon_0 k_B T_n} \right)^{1/2} \tag{2.3.31}$$

These equations are only valid if both electrons are situated in an X-valley and have both the same effective mass m_{Xd}, but as the effective mass in the L-valleys m_L has nearly the same value as m_{Xd}, the L-valleys may be included. If a simulation electron attempts an inter-carrier collision, a counterpart electron is selected at random from the electron distribution. Define $\vec{g}_i$ is the relative initial wave vector between the simulation electron and the counterpart $\vec{g}_i = \vec{\kappa}_{0i} - \vec{\kappa}_{ci}$. In the process a self-scattering is included, and the electron-electron collision is accepted only if a random number r results to be less than the acceptance condition C_a

$$C_a = \frac{2\,\beta_s|\vec{g}_i|}{|\vec{g}_i|^2 + \beta_s^2} \tag{2.3.32}$$

Now the angle θ between the relative wave vectors $\vec{g}_i$ and $\vec{g}_f$, before and after the scattering is given by

$$\theta = \arccos\left[\,1 - \frac{2r}{1 + \left(\frac{|\vec{g}_i|}{\beta_s}\right)^2 (1-r)}\,\right] \tag{2.3.33}$$

where r is a random number, uniformly distributed over $[0,1)$. Now

$$\vec{\kappa}_{0f} = \frac{1}{2}\,(\,\vec{\kappa}_{0i} + \vec{\kappa}_{ci} + \vec{g}_f\,) \tag{2.3.34}$$

$$\vec{\kappa}_{0f} = \frac{1}{2}\,(\,\vec{\kappa}_{0i} + \vec{\kappa}_{ci} - \vec{g}_f\,) \tag{2.3.35}$$

Using this algorithm, the state of the simulation electron is changed at the time of the collision; the state of the counterpart is changed at the time in which its simulation was suspended. In this procedure the distribution of the counterpart carriers is not exactly synchronous with the simulation time of the simulation electron, but this difference is not relevant if the time step duration Δt is smaller than the time between two inter-particle collisions.

Considering these equations some conclusions can be drawn:

1. At very low electron densities the scatter rate λ_{ee} becomes very high. This is due to the self-scattering, as can be seen by the acceptance condition (2.3.32) that becomes very low at low electron densities. Furthermore, the scatter angle nearly equals zero at low electron concentrations. All this results in a negligible scatter effect at low electron concentration, but it also means that the method proposed is very inefficient in this case.
2. At very high electron densities both the scatter rate λ_{ee} and the acceptance condition approach zero. This means that at high electron densities there is hardly any electron-electron scattering at all. This is a consequence of the screening effect in the model, that becomes very strong and prohibits the electron-electron interactions.

3. The influence of the electron temperature on the electron-electron scatter process is opposite to the influence of the electron concentration.

In practice the time step duration has to be smaller than the time between two inter-carrier interactions. At $n = 10^{21}\ m^{-3}$ and $m_X = 0.3\ m_0$ the average scatter time is $\lambda_{ee}^{-1} \approx 8 \cdot 10^{-16}$ s (self-scattering included). From simulations it is found that the acceptance condition factor typically has a value of 0.2 which leads to a time step of $\Delta t = 5 \cdot 10^{-15}$ s.
Simulations at room temperature in the low-field region show that electron-electron scatter processes have only a minor effect on the transport parameters e.g. the low-field mobility. In the high energy region the simulations by Brunetti et al. [1985] show a minor decrease of the energy relaxation time due to electron-electron scattering and this was confirmed by our simulations.

2.3.4.4 Interface scattering

This scatter process is of no relevance if the pure bulk material is considered. However in studying an interface (e.g. the Si-SiO_2 interface) this scatter process becomes relevant. Here the interface scattering is considered to be elastic, and scattering occurs in two forms:

1. Specular reflection, in which the component of the wave vector perpendicular to the interface changes its sign by reflection.
2. Diffuse reflection, in which the angle of the reflected wave vector is distributed uniformly over the interval $[0,\pi]$ away from the interface.

The selection between these two forms at the moment of scattering is made by random numbers. The ratio between the two different mechanisms cannot be decided on theoretical grounds. A best fit to experimental data in case of the Si-SiO_2 interface was found by Laux and Fischetti [1988] if both scatter mechanisms were given equal probability.

2.4 Definition of the transport parameters

The Ensemble Monte Carlo method can be used to calculate the transport parameters of interest, such as the mobility μ_n, the momentum relaxation time τ_{pn} and the energy relaxation time τ_{wn}, the conductivity effective mass m_n^*, the diffusion coëfficient D_n and the electron temperature T_n. In this chapter it is described how these transport parameters can be calculated in a Monte Carlo program.

2.4.1 Average electron velocity and energy

At the end of each time step Δt the kinetic energy u and velocity $\vec{w}$ of each of the M simulation electrons are known. Then the average electron energy $W_n = \langle u \rangle$ and the average electron velocity $\vec{v}_n = \langle \vec{w} \rangle$ can be calculated at time $t_N = N \cdot \Delta t$

$$W_n(t_N) = \frac{1}{M} \sum_{i=1}^{M} u_i(t_N) \tag{2.4.1}$$

$$\vec{v}_n(t_N) = \frac{1}{M} \sum_{i=1}^{M} \vec{w}_i(t_N) \tag{2.4.2}$$

In the steady state W_n and $\vec{v}_n$ are time independent. The number of simulation electrons M determines the statistical fluctuations in W_n and $\vec{v}_n$. For our purpose $M = 5000$ gives sufficiently accurate results.

A more sophisticated way of averaging is to integrate over each free flight, and then taking the sum over all free flights and over all electrons. Defining S the number of different free flights during Δt, and taking the duration of the j-th free flight $\Delta t_j = t_j - t_{j-1}$, yields

$$\vec{v}_n(t_N) = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{\Delta t} \sum_{j=1}^{S} \int_{t_{j-1}}^{t_j} \vec{w}_i(t')\, dt' \tag{2.4.3}$$

It must be emphasized however that now the average is taken over a period of time instead of the average at a definite point in time. In the steady state the two averaging techniques given by (2.4.2) and (2.4.3) represent essentially the same quantity (ergodic theorem), however when studying transient problems, this does not hold anymore.

In general the integral in (2.4.3) is approximated by a simple average over the free flight

$$\vec{v}_n(t_N) = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{\Delta t} \sum_{j=1}^{S} \frac{1}{2} \left[\vec{w}_i(t_{j-1}) + \vec{w}_i(t_j) \right] \Delta t_j \tag{2.4.4}$$

As long as the variable $\vec{w}_i(t)$ is linear in t or as long as the change of the variable during one free flight is small enough, this approximation is allowed. A similar expression is given for W_n. In principle, compared to the first method, the second method gives a reduction in the statistical uncertainty. In practice this is hardly noticed, and in case of short time steps Δt the advantage of the second method disappears completely.

2.4.2 Mobility

In the steady state in the infinite homogeneous bulk material, μ_n can be easily defined. If the electric field E applied is pointing in the x-direction then

$$\mu_n = \frac{v_{nx}}{E_x} \tag{2.4.5}$$

Using this definition μ_n is given only in the direction parallel to E, and therefore is a function of lattice orientation.

2.4.3 Relaxation times

To calculate the relaxation times τ_{pn} and τ_{wn}, two different methods are used. The momentum relaxation time τ_{pn} essentially is a tensor. However in practice τ_{pn} is taken by definition in the direction of E, and is dealt with as a scalar. The first method follows strictly the definitions

$$\left(\frac{\partial w_{//}}{\partial t}\right)_c = - \frac{v_{n//}}{\tau_{pn}} \qquad \Rightarrow \tau_{pn} = - \frac{v_{n//}}{\left(\frac{\partial w_{//}}{\partial t}\right)_c} \tag{2.4.6}$$

$$\left(\frac{\partial u}{\partial t}\right)_c = - \frac{W_n - W_0}{\tau_{wn}} \qquad \Rightarrow \tau_{wn} = - \frac{W_n - W_0}{\left(\frac{\partial u}{\partial t}\right)_c} \tag{2.4.7}$$

The drift velocity $v_{n//}$ in (2.4.6) is defined by (2.4.4). To calculate τ_{pn} also an expression for the collision term has to be found. This is done by the summation of all changes in the electron velocity in the direction of the field at all scattering moments during the interval Δt

$$\left(\frac{\partial w_{//}}{\partial t}\right)_c = \frac{1}{M\, \Delta t} \sum_{i=1}^{M} \sum_{j=1}^{S} \left(w_{i//}(t_j) - w_{i//}(t_{j-1})\right) \tag{2.4.8}$$

In a similar way (2.4.7) is dealt with.

A second method uses the transport equations in the steady state in homogeneous material, which gives

$$\tau_{pn} = - \frac{\hbar\, k_{nx}}{q\, E_x} \tag{2.4.9}$$

$$\tau_{wn} = - \frac{q\, E_x\, v_{nx}}{W_n - W_0} \tag{2.4.10}$$

In Si the valleys are elliptic and each valley has its own effective wave vector $\vec{\kappa}_i$ and its own effective field $\vec{E}_i$ (see section 2.3.1), which gives a new definition for τ_{pn}

$$\tau_{pn} = -\frac{\hbar}{q}\sum_{i=1}^{N}\frac{\langle\kappa_{xi}\rangle}{E_{ix}} \tag{2.4.11}$$

Here N is the total number of valleys. Using the definitions (2.4.10) and (2.4.11) τ_{pn} and τ_{wn} is given only in the direction parallel to the applied field. therefore the relaxation times are also a function of lattice orientation.

Calculations using either (2.4.6) and (2.4.7) or (2.4.10) and (2.4.11) show that as long as the strength of the applied electric field is high enough (E > 1 kV/cm), both methods give the same results; however, the statistical uncertainty in the first method is larger.

2.4.4 Conductivity effective mass

Essentially, the conductivity effective mass m_n^* is a tensor, the elements of which are defined by

$$\frac{1}{m_{nij}^*} = \frac{1}{\hbar^2}\left\langle\frac{\partial^2 u}{\partial k_i \partial k_j}\right\rangle \tag{2.4.12}$$

Three different practical definitions are considered

1. Because the conduction band is modeled as non-parabolic and non-spherical (2.3.1), the diagonal tensor elements have values

   $$\frac{1}{m_{nii}^*} = \left\langle -\frac{2\alpha\hbar^2 k_i^2}{m_i^2(1+2\alpha u)^3} + \frac{1}{m_i(1+2\alpha u)}\right\rangle \tag{2.4.13}$$

   and the non-diagonal elements are given by

   $$\frac{1}{m_{nij}^*} = \left\langle -\frac{2\alpha\hbar^2 k_i k_j}{m_i m_j(1+2\alpha u)^3}\right\rangle \tag{2.4.14}$$

   If the field is parallel to one of the principal lattice axes, equation (2.4.13) can be used to calculate m_n^* in the direction of the electric field, with the appropriate choice of index i. However if the field is not parallel to one of the principal lattice axes, then a coordinate transformation with transformation matrix U (2.3.9) is needed.

2. Zimmermann [1980] has proposed a more simplified approximation of m_n^*

$$m_n^* = m_0^* \left(1 + 6\alpha <u> \right) \tag{2.4.15}$$

Here m_0^* is defined by

$$\frac{1}{m_0^*} = \frac{1}{m_0} \frac{3}{\frac{1}{m_l} + \frac{2}{m_t}} \tag{2.4.16}$$

3. m_n^* also can be calculated directly from the momentum relation, yielding

$$m_n^* = \frac{\hbar\, k_{nx}}{v_{nx}} \tag{2.4.17}$$

Simulations show that (2.4.17) and (2.4.13) give very similar results, with values of 0.3 m_0 at low fields (E < 10 kV/cm) and a value of 0.5 m_0 at E = 100 kV/cm. The simplified expression (2.4.14) gives much higher values for m_n^* at high fields, and should not be used if E > 50 kV/cm.

2.4.5 Electron temperature

It is not a priori clear which definition of the electron temperature T_n is the best. Here, four different definitions are compared.

1. Assuming a shifted Maxwell-Boltzmann distribution, T_n is defined by

$$T_n = \frac{2}{3k_B} \mathbf{W}_n \quad \text{and} \quad \mathbf{W}_n = W_n - \frac{1}{2} m_n^* |\vec{v}_n|^2 \tag{2.4.18}$$

2. A more complicated definition has been given by Williams [1982] for GaAs to include non-parabolicity, also assuming a Maxwell-Boltzmann distribution

$$T_n = \frac{5\alpha \mathbf{W}_n - 3 + [\,(5\alpha \mathbf{W}_n - 3)^2 + 120\alpha \mathbf{W}_n\,]^{1/2}}{30\alpha k_B} \tag{2.4.19}$$

In the limit $\alpha \to 0$ equations (2.4.18) and (2.4.19) are identical.

3. A third definition of T_n makes use of the definition of the electron gas pressure tensor $\mathbf{P}_n$. This method is very well suited for Monte Carlo simulations. In a parabolic band the diagonal elements of $\mathbf{P}_n$ are given by

$$P_{nii} = n\, m_n^* (<w_i^2> - v_{ni}^2) \tag{2.4.20}$$

Then the hydrostatic pressure P_n is defined by

$$P_n = n\, k_B T_n = \frac{1}{3} \sum_i P_{nii} \tag{2.4.21}$$

All this yields he following definition for T_n

$$T_n = \frac{m_n^*}{3k_B} \sum_i (\langle w_i^2 \rangle - v_{ni}^2) \tag{2.4.22}$$

4. Definition (2.4.22) can be modified by defining the hydrostatic pressure only in the direction of the field. This yields

$$T_n = \frac{m_n^*}{k_B} (\langle w_{//}^2 \rangle - v_{n//}^2) \tag{2.4.23}$$

All four different definitions have been compared, yielding the following conclusions:

1. The kinetic term in (2.4.18) can be neglected.
2. The results obtained by (2.4.23) hardly can be distinguished from the results using (2.4.22).
3. The values of T_n calculated by (2.4.19) are some 10% lower than the values obtained by (2.4.22), which, in its turn, are 10% lower than the values obtained by (2.4.18).

2.4.6 Diffusion coefficient

A definition of the diffusion coefficient parallel to the electric field $D_{n//}$ is given by the generalized Einstein relation

$$D_{n//} = \frac{k_B T_n}{q} \mu_n \tag{2.4.24}$$

This method needs a priori an adequate definition of T_n, which is a problem (see section 2.4.5).

The Monte Carlo method offers an alternative way to calculate the diffusion coefficient by looking at the position of all electrons. At the start of the simulation all electrons are situated in the origin of space. The way they spread out during the simulation gives a measure for the diffusion coefficient D_n. In appendix 2.7.2 it is shown how the position of each electron can be calculated. Using the second Fick equation and the time derivative of the second central moment of electron density an expression for $D_{n//}$ is found (Jacoboni and Reggiani [1983])

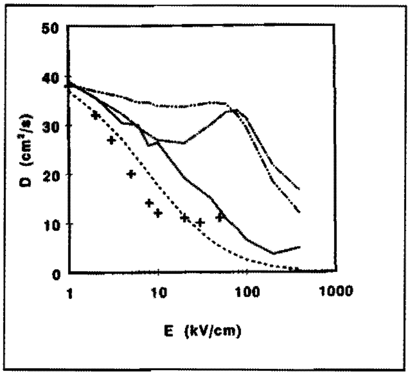

Fig 2.9 Diffusion coefficient as a function of the parallel field; both experimental data (crosses) and simulated data have been plotted various definitions:
eq. (2.4.25) yields the solid line;
eq. (2.4.24) with $T_n=T_0$ yields the dashed line;
eq. (2.4.24) in combination with (2.4.18) yields the dash-dot line;
eq. (2.4.24) in combination with (2.4.22) yields the dash-dot-dot line.

$$D_{n//} = \frac{1}{2}\frac{\partial}{\partial t} < (x - <x>)^2 > \tag{2.4.25}$$

In a Monte Carlo simulator, time is discretized to time steps Δt, and so is equation (2.4.25). In addition this method gives us the opportunity of calculating the diffusion coefficient perpendicular to the applied field in relation to that field

$$D_{n\perp} = \frac{1}{2}\frac{\partial}{\partial t} < (x - <x>)\ (y - <y>) > \tag{2.4.26}$$

A drawback of the method described is that statistical errors are accumulating during simulation time, so the statistical error in D_n increases as a function of time. Especially as the steady-state conditions are studied, which require a large amount of simulation time, this can be a serious problem.

In figure 2.9 various results have been plotted. Experimental data, given by Canali et al. [1975], are represented by the crosses. The solid line represents results obtained by equation (2.4.25); the other lines represent the result obtained if equation (2.4.24) is used. If the electron temperature is assumed to be equal to the lattice temperature the dashed line is obtained. If T_n is defined either by (2.4.18) or (2.4.22), the dash-dot line and the dash-dot-dot line are obtained respectively.

Surprisingly, at relatively low fields (E < 20 kV/cm), the best fit between experimental and theoretical data is obtained if the electron temperature equals the lattice temperature, and the worst fit if the Einstein relation is used. No physical explanation for this has as yet been found, but the same also has been observed in GaAs. Several attempts have been made to try to explain the phenomenon (e.g. by taking into account that also T_n is a function of electron position which modifies the Fick equation) but all failed. At high fields (E > 50 kV/cm) no experimental data were available, but it is clear that the Einstein relation (dash-dot and dash-dot-dot lines in figure 2.9) also will fail in this region.

Although the problem is not solved, it probably offers no difficulties for the rest of our work: in later studies the transport parameters are mainly used to model submicron MOS transistors, in which electron transport by drift is much more important than transport by diffusion. So the exact value of the diffusion coefficient probably is of little interest.

2.4.7 Impact ionization rate

The impact ionization rate for electrons α_n can be calculated directly from the Monte Carlo simulations. First the impact ionization scatter rate $\lambda_I(E)$ is calculated by counting the number of impact ionization scatter events that occur during a certain period of time while the system is in the steady state, and dividing the result by the total number of simulation electrons. Due to the high noise rate the period of time mentioned has to be taken rather long (several picoseconds). Now $\alpha_n(E)$ can be calculated by

$$\alpha_n(E) = \lambda_I(E)/\ v_d(E) \tag{2.4.27}$$

The results have been plotted in figure 2.8 (dashed line) combined with experimental data and as stated before, there is good agreement between the two sets of data, especially at high fields. At low fields the agreement is less but this should not worry us too much as at low fields the impact ionization is of little practical relevance.

2.5 Simulation results and discussion

All transport parameters have been obtained by analyzing the steady-state results of a large number of simulations of bulk Si. As explained in section 2.3.4.1 the unscreened Perri method has been preferred to model the ionized impurity scattering. Electron-electron scattering consumes enormous amounts of computing time with little effect and therefore has been neglected. Impact ionization is included. Furthermore the lattice temperature T_0 is assumed to be 300 K, and all X and L valleys are included. In each simulation the electric field is switched on from zero to a finite value, assuming that all electrons are initially in thermal equilibrium with the crystal lattice (the initialization are described in appendix 2.7.3). After several picoseconds the system reaches the steady state. Repeating this procedure while the electric field E is varied over a wide range of values (typically from 1 to 600 kV/cm), for various doping concentrations N_D (from zero to 10^{19} cm^{-3}), the transport parameters are found as a function of the applied field and

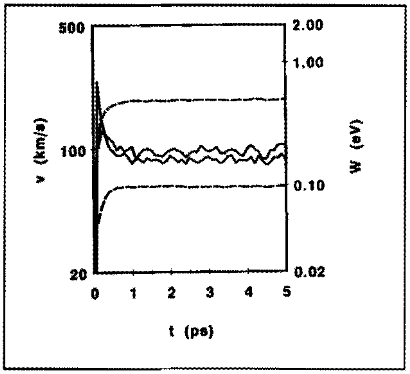

Fig 2.10 Average electron velocity (solid lines) and average electron energy (dashed lines) in intrinsic Si as a function of time after the applied field has been switched on; the electric field is the parameter (20 and 200 kV/cm respectively).

doping concentration. As is shown, the electron energy W_n increases monotonically with applied field, and only a simple transformation gives the electron transport parameters as a function of W_n instead of E. In section 2.5.1 simulation results of the bulk material are presented. The applied electric field is always directed in the (1,0,0) direction. The mobility is calculated using (2.4.5), the momentum relaxation time using (2.4.9), the energy relaxation time using (2.4.10), and the average effective mass using definition (2.4.13). In section 2.5.2 the transport parameters are described using analytical expressions based on the Monte Carlo data. In section 2.5.3 interface properties are dealt with. Part of the electrons arriving at the interface are scattered diffusively, part are reflected as described in section 2.3.4.4. The driving field is applied in a direction parallel to the interface. However also a field perpendicular to the interface is applied. Simulation results are given and discussed. Section 2.6 summarizes our conclusions.

2.5.1 Simulation results in the bulk

In figure 2.10 the averaged electron velocity (solid lines) and energy (dashed lines) in intrinsic Si are presented as a function of time at two different applied electric fields (20 kV/cm and 200 kV/cm respectively). Immediately after the field is switched on at $t=0$ the electrons gain energy and momentum. the gain in momentum is greater than the momentum relaxation to the lattice.

As a result of this the velocity reaches values higher than the saturation velocity. This phenomenon is called velocity overshoot. This velocity overshoot should not be confused with the phenomenon that in some materials (e.g. GaAs) the velocity at a certain critical field reaches a maximum value higher than the saturation velocity at much higher fields. The latter effect is from an entirely different nature and is caused by electrons being transferred to another valley with a higher effective mass. This leads to a

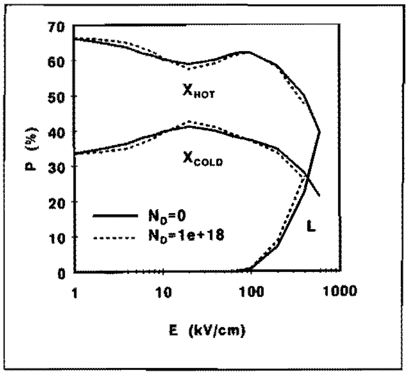

Fig 2.11 Calculated valley occupancy (X-valleys, both "hot" and "cold", and L-valleys) as a function of the applied field; the solid line represents the values in intrinsic material, the dashed line in n-type doped Si.

decreasing mobility and consequently a decrease in velocity is found. For a more detailed description of this phenomenon see e.g. Aspnes [1976].

After 2 ps equilibrium is reached for all electric fields applied. In figure 2.11 the valley occupancy in the steady state has been plotted as a function of the applied electric field. The field is directed in the (1,0,0) direction, and because of different orientations of the X-valleys, four of the X-valleys have a smaller effective mass than the other two, and electrons are able to gain more energy in those valleys. For this reason they are called the "hot" X-valleys. The other are called "cold" X-valleys. At low fields all X-valleys hold the same number of electrons. At increasing fields, due to the fact that in the hot valleys the electrons are more subject to scattering, relatively more electrons are situated in the cold valleys. At even higher fields the scattering satturates in all valleys and there is a tendency to an even distribution over all valleys. This even distribution is not reached because at that point (E = 100 kV/cm) many electrons gain enough energy to make the transition to the L-valley. In figure 2.11 the solid line indicates the situation for intrinsic material, the dashed line for extrinsic material with $N_D = 10^{18}$ cm^{-3}. As can be seen, the doping concentration is hardly relevant for the valley occupancy.

In figures 2.12a and 2.12b the velocity and energy distribution functions have been plotted when an electric field of 100 kV/cm is applied. The broken lines represent the result of our simulations. The continuous lines represent the Maxwell-Boltzmann distribution assuming a parabolic band. In the latter case the velocity distribution in the direction of the field is described by

$$F(v_{//}) = \left(\frac{m^*}{2\pi k_B T_n} \right)^{1/2} \exp \left(- \frac{m^*}{2k_B T_n} (v_{//} - v_0)^2 \right) \qquad (2.5.1)$$

Neglecting the kinetic term, the energy distribution is described by

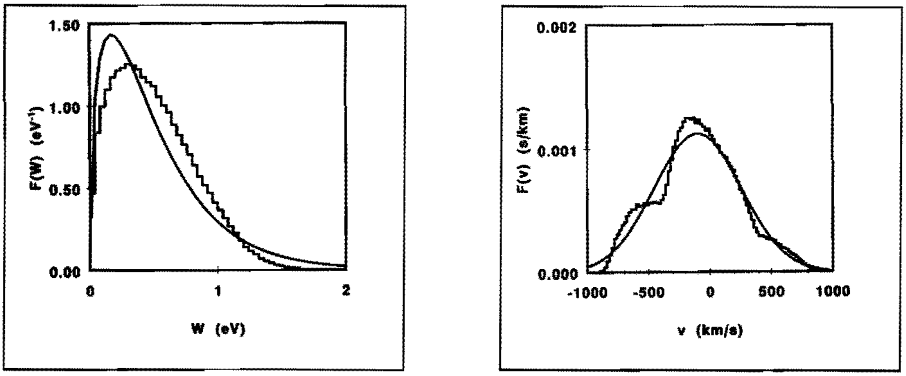

Fig 2.12 Energy (a) and velocity (b) distribution in intrinsic Si at E=100 kV/cm; the smooth curve has been calculated assuming a Maxwell-Boltzmann distribution; the other curve results directly from the Monte Carlo simulations.

$$F(u) = 2\pi \left(\pi k_B T_n \right)^{-3/2} u^{1/2} \exp\left(- \frac{u}{k_B T_n} \right) \qquad (2.5.2)$$

The parameters T_n , v_0 and m_n^* are obtained from the simulation $k_B T_n = \frac{2}{3} \langle u \rangle$, $v_0 = \langle v_{//} \rangle$ and $m_n^* = 0.47\ m_0$.
It is clearly seen that at this high field the distribution definitely is not Maxwellian. The velocity distribution function is very asymmetric, and the high velocity tails are truncated: the velocity hardly ever exceeds 10^8 cm/s. Also the high energy tail is lower than the Maxwell-Boltzmann distribution function predicts. Other simulations show that the Maxwell-Boltzmann distribution function holds more or less up to 20 kV/cm.

In figure 2.13a the electron energy in the steady state has been plotted as a function of the applied electric field. The doping concentration is used as parameter. The solid line represents intrinsic material, the dashed line extrinsic, with $N_D = 10^{18}$ cm^{-3}. The triangles in the energy plot represent results for intrinsic Si obtained by Fischetti [1991]. In Fischetti's work a more realistic band model has been used. Nevertheless, their values are much higher, probably due to underestimating the impact ionization rate, that has a value nearly a factor 10 lower than used in this work.

From these results the following conclusions can be drawn:

1. The energy W_n increases monotonically with the field E. For this reason it is possible to define the inverse E-W_n relation (see section 2.5.2.2).
2. For fields $E < 100$ kV/cm our results are very similar to the results obtained by the Fischetti group. A discrepancy arises for higher fields, where our results tend to be lower.

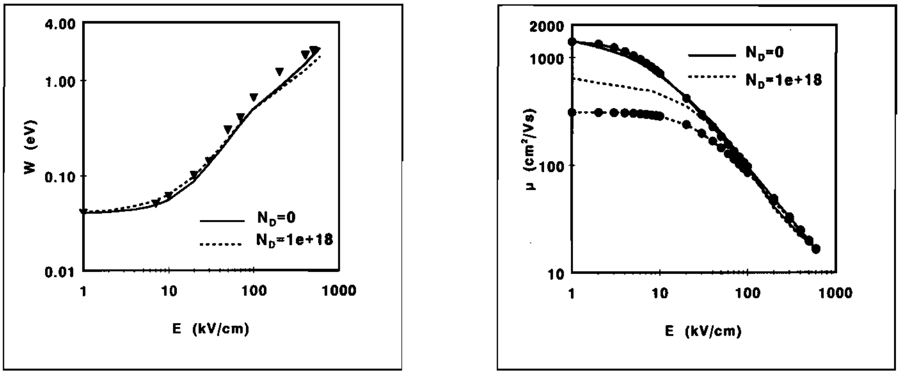

Fig 2.13 Calculated average electron energy (a) and electron mobility (b) in both intrinsic (solid line) and n-type material (dashed line) as a function of applied electric field; the triangles in (a) represent data obtained by Fischetti [1991]; in (b) the lines with dots represent data obtained by Selberherr [1989].

3. The average electron energy has little dependency on the doping concentration, which is not very surprising as the model used assumes elastic interaction between ionized impurities and electrons.

In figure 2.13b the electron mobility in the steady state has been plotted as a function of E. Again N_D is used as a parameter. The curves without dots represent our Monte Carlo results. The results obtained from the mobility model presented by Selberherr [1989], if no surface scattering is included, are marked by large dots. The latter model has a good fit to experimental data.

Now the following conclusions can be drawn from figure 2.13b:

1. In intrinsic material the Monte Carlo results and the Selberherr model are indistinguishable for all electric fields considered.
2. In extrinsic material the Monte Carlo simulations give too high values for the mobility at low fields (E < 50 kV/cm). This is a result of an inadequate ionized impurity scattering model (see 2.3.4.1).
3. At high fields (E > 50 kV/cm) the doping concentration has negligible effect on the mobility.

Other relevant transport parameters are τ_{wn} and m_n^*. The conductivity effective mass m_n^* is given in figure 2.14a and appears to increase monotonically as a function of energy, which is not surprising, given a band structure with non-parabolic valleys. The solid line gives the value for intrinsic material, the dashed line for material doped with $N_D = 10^{18}$ cm^{-3}. The effect of doping on m_n^* is negligible.

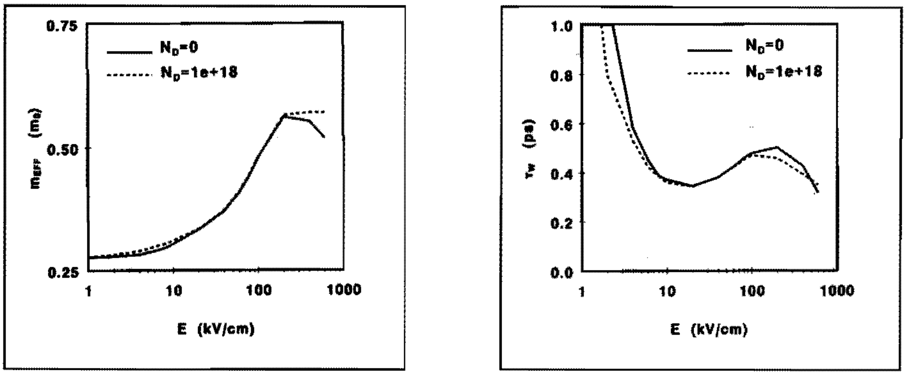

Fig 2.14 Calculated electron conductivity effective mass (a) and energy relaxation time (b) in both intrinsic and n-type Si as a function of the applied electric field.

The energy relaxation time has been plotted in figure 2.14b. At low fields ($E < 10$ kV/cm) the energy gained by the electrons is negligible (see figure 2.13a) and therefore the value of τ_{wn} in this region is not of any relevance. In figure 2.14b it can be seen that τ_{wn} has a nearly constant value in the high-field region ($E > 10$ kV/cm), and ranges from 350 to 500 femtoseconds. The solid line gives the values found for intrinsic Si, the simple dashed line for doped material ($N_D = 10^{18}$ cm^{-3}). It can be seen that the effect of the doping concentration on τ_{wn} is little.

The results are in good agreement with results given by Jacoboni and Reggiani [1983]. However, the results published by Fischetti [1991], using a different, more realistic conduction band model, give a τ_{wn} that has a value of 500 fs, in the energy range up to 0.1 eV, and a nearly constant value of 900 fs if $W > 0.5$ eV. This latter value seems rather high. Furthermore a decrease of the τ_{wn} is expected due to an increase in the impact ionization scatter rate λ_{ii}. The Fischetti group probably has underestimated λ_{ii}.

2.5.2 Analytical form of the transport parameters

Using the Monte Carlo simulation results, the transport parameters are described in an analytical form that will be used in the hydrodynamic equations. In sections 2.5.2.1 to 2.5.2.4 the various relevant transport parameters (μ_n, m^*_n, τ_{wn} and α_n) are given as a function of the applied electric field E.

To obtain these parameters as a function of the local electron energy, E has to be substituted by an effective electric field E_{eff}. E_{eff} is a function of the local electron energy and is dealt with in section 2.5.2.5.

2.5.2.1 Mobility

As shown in the previous section the Selberherr mobility model for the bulk gives good agreement with the Monte Carlo data for intrinsic material. Furthermore it has been explained in section 2.3.4.1 that the ionized impurity scatter model in the Monte Carlo method does not lead to results that fit with experimental data. For this reason the Selberherr mobility model has been used in our hydrodynamic model that is described in the next chapter. Thus the mobility is defined by

$$\mu_n^{LI} = \mu_{n,min} + \frac{\mu_n^L - \mu_{n,min}}{1 + \left(\frac{N_I}{N_{ref}} \right)^{\alpha_n}} \tag{2.5.3}$$

$$\mu_n = \frac{2\ \mu_n^{LI}}{1 + \left(1 + \left(\frac{2\ \mu_n^{LI}\ E}{v} \right)^2 \right)^{1/2}} \tag{2.5.4}$$

In section 2.5.4 N_I is the impurity concentration. The values of the different parameters at 300 K are given in table 2.2.

μ_n^L	*1430*	*cm²/Vs*
$\mu_{n,min}$	*80*	*cm²/Vs*
N_{ref}	*1.12 10^{17}*	*cm^{-3}*
α_n	*0.72*	
v_{sat}	*10^7*	*cm/s*

Table 2.2 Mobility parameters at 300 K (Selberherr [1989]).

2.5.2.2 Conductivity effective mass

From figure 2.14a it can be seen that m_n^* also is a function of the applied field. However, in the rest of this work the value of m_n^* only appears in the kinetic term of the energy, and this term hardly is of any relevance in Si. Therefore the field dependency of m_n^* can safely be neglected and the constant low field value $m_n^* = 0.256\ m_0$ can be used.

2.5.2.3 Energy relaxation time

A first order approximation for τ_{wn} gives the constant value $\tau_{wn} = 0.4$ ps. A very good fit with the data given in figure 2.14b (if $E > 10$ kV/cm) has been found by fitting two cosine functions on a logarithmic scale. The expression used is

If $E < E_0$ then $\tau_{wn} = \tau_{w0}$ (2.5.5)

If $E_0 \leq E \leq E_1$ then

$$\tau_{wn} = \frac{1}{2} (\tau_{w0} + \tau_{w1} + (\tau_{w1} - \tau_{w0}) \sin \left[\frac{\pi}{2} \frac{2 \ln(E) - \ln(E_0) - \ln(E_1)}{\ln(E_1) - \ln(E_0)} \right]$$

If $E_1 \leq E \leq E_2$ then

$$\tau_{wn} = \tau_{w2} + (\tau_{w2} - \tau_{w1}) \sin \left[\frac{\pi}{2} \frac{\ln(E) - \ln(E_2)}{\ln(E_2) - \ln(E_1)} \right]$$

The values of the different parameters at 300 K are given in table 2.3.

τ_{w0}	0.35 ps	E_0	20 kV/cm
τ_{w1}	0.50 ps	E_1	200 kV/cm
τ_{w2}	0.35 ps	E_2	600 kV/cm

Table 2.3 Energy relaxation time parameters at 300 K, extracted from Monte Carlo simulation data.

2.5.2.4 Impact ionization rate

The impact ionization as a function of applied electric field can be described by the expression given by Chynoweth [1958]

$$\alpha_n = \alpha_n^0 \exp [-E_{crit}/E] \tag{2.5.6}$$

The values of α_n^0 and E_{crit} for Si were obtained by Van Overstraeten and De Man [1970] from experimental data $\alpha_n^0 = 7.03 \cdot 10^5$ [cm^{-1}] and $E_{crit} = 1.23 \cdot 10^6$ [V/cm]. As has been shown in figure 2.8 the Monte Carlo results give good agreement with these experimental data, especially at high fields. For this reason expression (2.5.6) is considered to give an adequate description of the impact ionization rate in the bulk material.

2.5.2.5 Effective electric field

In sections 2.5.2.1 to 2.5.2.4 all transport parameters have been defined as a function of the electric field. However we are interested in the parameters as a function of the energy. In figure 2.13a the energy W_n is given as a function of E. We are interested in the inverse relation: given W_n, what is the so-called effective electric field E_{eff}. Because W_n increases monotonically as a function of E it is possible to define this inverse relation. The procedure is as follows:

First consider the energy conservation equation in the steady state

$$E \cdot J_n = n \frac{W_n - W_0}{\tau_{wn}} \tag{2.5.7}$$

Now two different versions are considered:

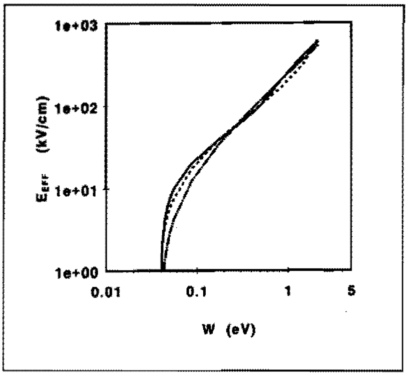

Fig 2.15 The effective electric field as a function of electron energy; the solid line results directly from the Monte Carlo results; the dashed line is calculated by eq. (2.5.10) combined with (2.5.5); the dash-dot is obtained by eq. (2.5.8) with constant energy relaxation time.

1. Because in the region of interest v_n equals, as a first approximation, the saturation velocity v_{sat}, J_n can be defined by $J_n = - q\, n\, v_{sat}$. This yields

$$E_{eff} = \frac{W_n - W_0}{q\, \tau_{wn} v_{sat}} \tag{2.5.8}$$

2. A more complete version uses the definition $v_n = - \mu_n E$. This yields

$$q\, \mu_n E^2 = \frac{W_n - W_0}{\tau_{wn}} \tag{2.5.9}$$

Combining (2.5.9) with the Selberherr mobility given by (2.5.4) and solving E yields

$$E_{eff} = \left(\frac{(W_n - W_0)^2}{q^2 \tau_{wn}^2 v_{sat}^2} + \frac{W_n - W_0}{q\, \mu_n^{LI}\, \tau_{wn}} \right)^{1/2} \tag{2.5.10}$$

In figure 2.15 the exact electric field (solid line) as a function of electron energy is compared to the calculated effective fields using (2.5.10), combined with (2.5.5). Good agreement is found. Also the simple expression (2.5.8), assuming a constant $\tau_{wn} = 0.4$ ps yields a good agreement (dash-dot line) at high energies.

2.5.3 Interface transport parameters

The transport parameters will be used in a hydrodynamic device simulator, mainly to simulate MOS devices, where electron transport is concentrated at the Si-SiO$_2$ interface. In this paragraph the effects of the interface are discussed. No quantization effects at the interface have been taken into account, as has been done by e.g. Basu [1978] and Chu-Hao [1985].

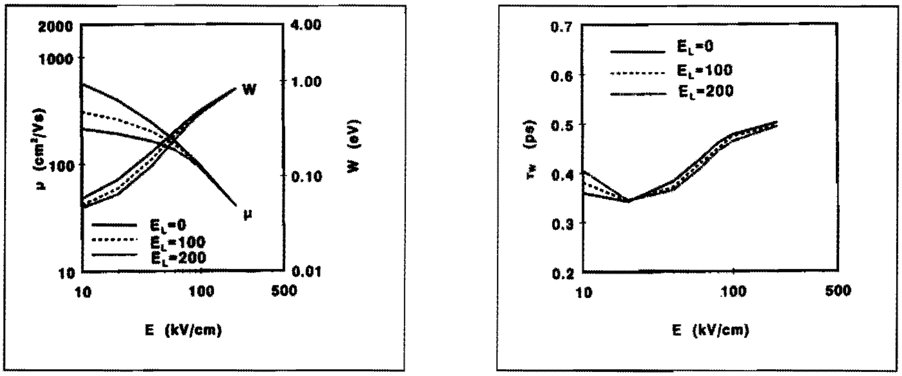

Fig 2.16 Surface electron mobility and energy (a) and energy relaxation time (b) in n-type Si ($Nd=3.3\ 10^{17}\ cm^{-3}$) as a function of the applied parallel electric field; the perpendicular electric field is the parameter (values in kV/cm).

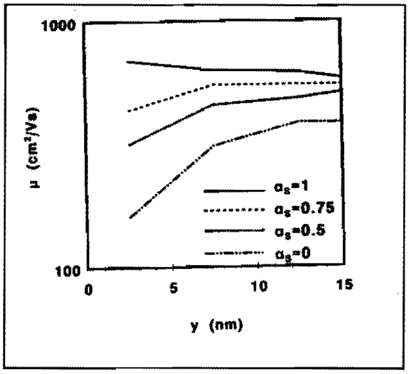

Fig 2.17 Low-field electron mobility in intrinsic Si as a function of the distance from the interface; the interface scatter ratio α_s is the parameter.

The applied field has two components, $E_{//}$ parallel to the interface, and $E_{\perp}$ perpendicular to the interface. Both can be varied independently. In addition, the ratio α_s between specular reflection and diffuse reflection at the interface has to be defined.

In figure 2.16 the surface mobility, the average energy, and the energy relaxation time have been plotted as a function of $E_{//}$ for doped material ($N_D = 3.3 \cdot 10^{17}\ cm^{-3}$). The parameter is $E_{\perp}$. In these simulations the ratio α_s equals 0.5, being the value found by Laux and Fischetti [1988].

From the simulation results the following conclusions can be drawn:

1. At high parallel fields ($E_{//} > E_{\perp}$) the interface scattering has negligible effects.
2. The surface mobility decreases substantially with increasing $E_{\perp}$.
3. Also the average electron energy decreases with increasing $E_{\perp}$; however the effect is small in the region of interest.
4. The energy relaxation time is nearly independent of $E_{//}$ and $E_{\perp}$.

Finally the value of α_s has been varied. Simulation data for intrinsic Si have been plotted in figure 2.17 ($E_{//}$= 10 kV/cm and $E_{\perp}$= 100 kV/cm). If α_s= 1, the interface scattering is purely specular reflexion, and if α_s= 0 all interface scattering is diffuse. The perpendicular field forces the electrons towards the interface, and the number of simulation electrons decreases fast with increasing distance from the interface. Therefore the error in the mobility increases fast with increasing distance.

The following conclusions can be drawn:

1. A conducting channel of some 10 nm has been found.
2. If α_s= 0 the interface moblity equals the bulk mobility (i.e. 680 cm^2/Vs at $E_{//}$= 10 kV/cm).
3. For increasing α_s the interface mobility decreases strongly.
4. Furthermore it has been found that a value α_s= 0.75 gives the best agreement with the Selberherr interface mobility model.

2.6 Summary and conclusions

The conclusions that are relevant to the rest of our work are summarized:

1. The electron energy is a monotonic function of the applied electric field and therefore the inverse relation can be defined.
2. Because no adequate results of the ionized impurity scattering from first principles is available, the measured bulk mobility is better described by the model given by Selberherr [1989]. To obtain the energy dependence of the mobility, in the latter model the electric field can be replaced by the effective electric field given by (2.5.8) or (2.5.10).
3. The calculated impact ionization rate is well described by the model given by Van Overstraeten and De Man [1970]. Energy dependence of the impact ionization rate is obtained by using the effective electric field concept.
4. The energy relaxation time is almost independent of the applied field. A value of 400 fs seems appropriate both for bulk and interface material. A small, nearly negligable, dependence on doping concentration has been found in our simulations. Our results are in agreement with Jacoboni and Reggiani [1983]. The value τ_{wn} = 900 fs, found by Fischetti [1991] is probably too high, due to an underestimation of the impact ionization scatter rate.

2.7 Appendices

2.7.1 The Ridley model of the ionized impurity scattering

The Ridley approach is a classical approach to the problem of ionized impurity scattering [1977]. From that theory Van de Roer and Widdershoven [1986] deduced definitions of the scatter rate and scatter angle that can be used in a Monte Carlo simulator. They found that, if λ_{BH} is the scatter rate calculated using the Brooks-Herring model, the scatter rate in the Ridley model can be defined by

$$\lambda_R(\vec{\kappa}) = \frac{|\vec{v}(\vec{\kappa})|}{a} \left[1 - \exp\left(- \frac{a\lambda_{BH}(\kappa)}{|\vec{v}(\vec{\kappa})|}\right) \right] \quad \text{with } a = N_I^{-1/3} \tag{2.7.1}$$

With λ_{BH} defined by

$$\lambda_{BH}(\vec{\kappa}) = \frac{Z^2 N_I \sqrt{2}\ (k_B T_n)^2 (m_0 m_d)^{3/2}}{\pi \hbar^4 n^2} \ \frac{(1+2\alpha u)\ [u(1+\alpha u)]^{1/2}}{1 + 8u(1+\alpha u)\ \dfrac{m_0 m_d}{\hbar^2 \beta_s^2}} \tag{2.7.2}$$

Here β_s is the inverse Debye length associated with screening as given by equation (2.3.31).

Now the distribution of the angle θ between initial velocity vector and final velocity vector in $\vec{\kappa}$-space is non-isotropic, and is a function of a random number r. Van de Roer and Widdershoven found that in the Ridley model this angle $\theta_R(r) = \theta_{BH}(r_R)$, with r_R given by

$$r_R = - \frac{|\vec{v}(\vec{\kappa})|}{a\lambda_{BH}(\vec{\kappa})} \log \left[1 - (1-r)\ \frac{a\lambda_R(\vec{\kappa})}{|\vec{v}(\vec{\kappa})|} \right] \tag{2.7.3}$$

This yields

$$\theta_R(r) = \theta_{BH}(r_R) = \arccos \left[1 - \frac{2(1-r_R)}{1 + 8r_R u(1+\alpha u)\ \dfrac{m_0 m_d}{\hbar^2 \beta_s^2}} \right] \tag{2.7.4}$$

2.7.2 Calculation of the electron position

Knowledge of the position of each of the simulation electrons is needed in Monte Carlo bulk simulations in calculating the diffusion coefficients, and it is needed in Monte Carlo interface and device simulations. The electron position does not change at the moment of scattering; it only changes during the free flight. If the position of the electron at the beginning of the free flight is $\vec{x}_0$ and the time of flight is t_r, then

$$\vec{x} = \vec{x}_0 + \int_0^{t_r} \vec{v}(t)\, dt \tag{2.7.5}$$

With equation (2.3.5) this yields

$$\vec{x} = \vec{x}_0 + \frac{\hbar}{m_0} T_i \vec{s} \tag{2.7.6}$$

Here $\vec{s}$ is path length, defined by

$$\vec{s} = \int_0^{t_r} \frac{\vec{\kappa}\, dt}{1+2\alpha u} \tag{2.7.7}$$

From equation (2.3.3) it can be deduced that

$$1+2\alpha u = [\ 1+2\alpha \frac{\hbar^2}{m_0} \kappa^2\ ]^{1/2} \tag{2.7.8}$$

And equation (2.2.1), with $\vec{\kappa}_f$ the transformed wave vector at the start of the free flight, yields

$$\vec{\kappa}(t) = \vec{\kappa}_f - \frac{q}{m_0} t \vec{E}_i \tag{2.7.9}$$

Then with $\kappa = |\vec{\kappa}|$, $\kappa_f = |\vec{\kappa}_f|$ and $E_i = |\vec{E}_i|$ the combination of equations (2.7.8) and (2.7.9) yields

$$1+2\alpha u = [\ 1 + 2\alpha\kappa_f^2 + 2\alpha t^2 E_i^2 - 4\alpha t\ (\vec{\kappa}_f\cdot\vec{E}_i)\ ]^{1/2} \tag{2.7.10}$$

Next substituting (2.7.9) and (2.7.10) in (2.7.7) and using two standard integrals (Gradshteyn, [1980]), with $R=a+bt+ct^2$

$$\int\frac{dt}{R^{1/2}} = \frac{1}{\sqrt{c}} \log\ [2\cdot(cR)^{1/2}+2ct+b] \tag{2.7.11}$$

$$\int\frac{t\, dt}{R^{1/2}} = \frac{R^{1/2}}{c} - \frac{b}{2c} \int \frac{dt}{R^{1/2}} \tag{2.7.12}$$

Taking $R = 1+2\alpha u$ it can be deduced in a straight forward way that

$$\vec{s} = -\vec{E}_i \frac{2\alpha}{c} (u_i-u_f) + \\ (\vec{\kappa}_f + \vec{E}_i \frac{b}{2c})\frac{1}{\sqrt{c}} \log \left(\frac{2\sqrt{c}(1+2\alpha u_i) + 2ct_r + b}{2\sqrt{c}(1+2\alpha u_f) + b} \right) \tag{2.7.13}$$

Here u_f is the energy at the start of the free flight and u_i at the end; $b = -4\alpha\ (\vec{\kappa}_f \cdot \vec{E}_i)$ and $c = 2\alpha\ E_i^2$. If $\vec{s}$ is known, the position $\vec{x}$ can be calculated with (2.7.6).

Note: in case $\alpha=0$ the log-function in (2.7.13) is not defined. However the limit $\alpha \to 0$ does exist and yields

$$\lim_{\alpha \to 0} \vec{s} = \vec{\kappa}_f t_r - \frac{1}{2} \vec{E}_i t_r^2 = \frac{1}{2}\ t_r\ (\vec{\kappa}_f + \vec{\kappa}_i) \tag{2.7.14}$$

$\vec{\kappa}_i$ is the transformed wave vector at t_r.

2.7.3 Initialization of the electron energy and wave vector

Before a Monte Carlo run is started, all variables have to be initialized. We are mainly interested in parameters as a function of the applied electric field at steady-state conditions in an homogeneous bulk material. In principle the initial position of the electrons is not of interest. Only in calculating the diffusion, by studying the way electrons diverge, is the position of the carriers used. For this reason at $t=0$ all electrons are situated at the origin. We assume thermal equilibrium at $t=0$ at a lattice temperature T_0, and assume all electrons to be in the lowest valleys. In case of Si this means that all electrons are uniformly distributed over all six X-valleys. The electron temperature is assumed to be equal to the lattice temperature, and the initial drift velocity is assumed to be zero.

More of a problem is to calculate the distribution of electrons over $\vec{k}$-space. One way to do this is to look at the $\vec{k}$ distribution which is chosen to be Maxwellian. Another way, and that is the way it is done in this program, is to look at the energy distribution. If the valleys in the material are spheric and parabolic, the energy distribution function is given by

$$P(u) = \frac{2\pi\ u^{1/2}}{(\pi k_B T_0)^{3/2}} \exp\left(- \frac{u}{k_B T_0} \right) \tag{2.7.15}$$

This equation can be rewritten as a chi-square distribution of order n=3

$$P(u) = \frac{2}{k_B T_0} \mathcal{X}_3(\chi) \qquad \text{with} \quad \chi = \left(\frac{2u}{k_B T_0}\right)^{1/2} \tag{2.7.16}$$

$$\text{And} \quad \mathcal{X}_n(\chi) = \frac{\chi^{n-2}}{2^{n/2}\ \Gamma(\frac{n}{2})} \exp\left(- \frac{\chi^2}{2} \right) \tag{2.7.17}$$

Now χ^2 is found using random numbers. In this case (n=3) χ^2 can be found by

$$\chi^2 = -2\ [\ \log(r_1) + \log(r_2)\cos\ (2\pi r_3)\ ] \tag{2.7.18}$$

Here r_1, r_2 and r_3 are random numbers distributed uniformly on [0,1). If χ^2 is known, u can be calculated, and if u is known so is $\vec{k}$. This procedure of generating random numbers, calculating χ^2, u and $\vec{k}$ is repeated for all simulation electrons. The direction of $\vec{k}$ in a spheric valley is chosen to be isotropic, which means that the azimuthal angle φ equals $2\pi\ r_4$, and the tangential angle θ equals arccos $(1 - 2\ r_5)$, where r_4 and r_5 are random numbers distributed uniformly on [0,1)

The problem appears to be more complicated when the valleys are elliptic instead of spheric, as is the case in Si. However by applying a suitable transformation, it is possible to define a transformed wave vector $\vec{\kappa}$, and in $\vec{\kappa}$-space the valleys are again spherical and the same procedure can be applied. More serious is the problem when the valleys are non-parabolic with a non-parabolicity factor α, and the energy distribution function becomes much more complicated. Using an appropriate expansion of the energy distribution function leads to the following first order approximation

$$\mathrm{P}(u) = \frac{2}{k_B T_0}\ (C_3\mathcal{X}_3(\chi)+C_5\mathcal{X}_5(\chi)) \tag{2.7.19}$$

$$C_3 = \frac{1}{1 + \frac{15}{4}\ \alpha k_B T_0} \tag{2.7.20}$$

$$C_5 = \frac{\frac{15}{4}\ \alpha k_B T_0}{1 + \frac{15}{4}\ \alpha k_B T_0} \tag{2.7.21}$$

Now two different χ^2 have to be calculated, one χ_3^2 belonging to $\mathcal{X}_3$, which can be calculated using equation (2.7.18), and one χ_5^2 belonging to $\mathcal{X}_5$, which can be calculated by

$$\chi_5^2 = -2\ [\ \log(r_1 r_2) + \log(r_3)\cos\ (2\pi r_4)\ ] \tag{2.7.22}$$

Here r_1, r_2, r_3 and r_4 are random numbers distributed uniformly over [0,1). This results in

$$u = \frac{k_B T_0}{2}\ (\ C_3\chi_3^2 + C_5\chi_5^2\ ) \tag{2.7.23}$$

If $\vec{k}$ is known, the electron velocity can be calculated too, and all variables have been initialized.

Chapter 3

MONTE CARLO DEVICE SIMULATION IN ONE DIMENSION

3.1 Introduction

The Ensemble Monte Carlo method not only can be used for simulating bulk material or interface properties, but also for simulating the characteristics of complete devices. A description of a full two-dimensional Monte Carlo simulator of a GaAs HEMT was given by Nederveen [1989], a description of a two-dimensional MOST simulator by Laux and Fischetti [1988, 1991]. The major problem of these simulators is that the simulation electrons have to be distributed over a two-dimensional space. Because of limited computing capacity the number of electrons has to be limited, and in large parts of the device the simulated electron density is very low. This makes the results very noisy. This noise has no physical background, but is purely due to the Monte Carlo method.
In our study we are not so much interested in developing a Monte Carlo device simulator, but more in developing a hydrodynamic device simulator. The Monte Carlo method is a tool in this work and is used as a reference to simulation results obtained by the hydrodynamic method for determining the accuracy and the limits of applicability of the hydrodynamic method. For this purpose a one-dimensional Monte Carlo device simulator suffices. In this chapter this simulator is described. Simulation results are used to answer the question how to model the device contacts best. More simulation results are presented in the next chapter, after the introduction of the hydrodynamic drift device simulator.

3.2 The one-dimensional Monte Carlo device simulator

In the Monte Carlo device simulator only electron transport is dealt with. The device consist of three layers: two n-doped contact layers and one p- or n-type drift layer in between. Ohmic contacts are defined at the left hand and right hand side of the device, and the potential at the contacts is applied externally. In analogy with MOST structures the device contacts are called source and drain.

The most important differences between bulk material simulations and device simulations are the following:

1. The applied field is not distributed uniformly over the device and depends on the device structure and the charge density. Therefore the Poisson

equation has to be introduced in the program.

2. Parameters which can be considered as uniform in a bulk simulator (e.g. impurity concentration, electron concentration, electron temperature) are functions of the position in the device.
3. The range of electron densities varies by several orders of magnitude. If each simulation electron had to represent the same electric charge, then the majority of the simulation electrons would accumulate in the contact layers, which are the least interesting parts of the device. To avoid this, the charge represented by each simulation electron is multiplied by a weight factor, depending on its position in the device.
4. A model of the contacts has to be incorporated in the program, describing how electrons are absorbed in a contact and how they are injected from it.

These topics are dealt with in this chapter more extensively. The global set up of the program is as follows:

1. Definition of device geometry and initialization conditions.
2. Calculating the charge distribution in the device, solving the Poisson equation and calculating the electric field distribution.
3. Use of the Monte Carlo method during a certain time step Δt to obtain the new position, velocity and energy distribution of the simulation electrons in the device.

The last two steps are repeated as long as is considered necessary. Stability considerations needed to decide on the spatial discretization parameter Δx and on the temporal discretization parameter Δt are discussed in section 3.3.

3.2.1 Device geometry

A uniform one-dimensional grid is used to discretize the drift device. The number of intervals is N_x, each of size Δx, and the total device length L equals $N_x \Delta x$. The number of grid points equals N_x+1, being labelled 0 to N_x. The i-th interval is defined between grid points i-1 and i. The doping concentration is defined at the grid points. Between two grid points the doping concentration is considered constant having the averaged value of the concentration at the grid points. The charge density, the potential, the average velocity and the average energy are defined at the grid points; the electric field and the electron temperature are considered to be constants between two grid points. The device is divided in three different layers: the source region (from grid point 0 to L_s), the drain region (from L_d to L), and the drift region in between. The electric charge represented by each simulation electron differs in each region. If each simulation electron situated in the drift region represents M_r real electrons, then each

simulation electron in the source region represents $q_s \cdot M_r$ real electrons, and in the drain region $q_d \cdot M_r$. q_s and q_d are the charge weights of respectively source and drain. If an electron moves from one region to another its weight changes as is described in more detail in section 3.2.5.
It is not a priori known how the electron density distribution in the device looks like, but it seems a logical choice to choose the weight factor q_s to be equal to the ratio of the source region dope and the drift region dope, and similar for q_d. Then the simulation electrons are distributed homogeneously over the device. However simulations show that we have to take care that the electron charge weight in the contact regions does not become so large that the movement of one electron has a severe effect on the potential distribution. This would make the program very unstable. In general it is better to restrict q_s and q_d to not too large values (up to a factor of ten).

3.2.2 Initializations

Before the actual simulation is started, the electrons are considered to be in thermal equilibrium and the energy and wave-vector distributions at the onset of the simulation are calculated as described in appendix 2.7.3.

The initial position of the simulation electrons needs some attention. The initial total number of simulation electrons is M (M can be up to 20000 electrons). All donors and acceptors are assumed to be fully ionized. Now the number of simulation electrons in the i-th interval M_i is proportional to the doping concentration divided by the factor $q_i M_r$. M_i is given by

$$M_i = O_s \Delta x \, \frac{\frac{1}{2}\,(N_{Di-1} + N_{Di})}{q_i M_r} \qquad (3.2.1)$$

Here O_s represents the contact area. In the drift region the weight factor q_i equals one, in the source and drain regions q_i equals q_s and q_d, respectively. Now M_r can be calculated

$$M_r = \frac{O_s \Delta x}{M} \sum_{i=1}^{N_x} \frac{\frac{1}{2}\,(N_{Di-1} + N_{Di})}{q_i} \qquad (3.2.2)$$

If the number of simulation electrons situated in each interval is known, the probability that an electron is situated in the i-th interval is given by

$$P_i = \frac{M_i}{M} = \frac{\frac{1}{2}\,(N_{Di-1} + N_{Di})\, O_s \Delta x}{q_i M_r M} \qquad (3.2.3)$$

The initial spatial position of each simulation electron in each interval is calculated by assuming an uniform distribution over that interval.

3.2.3 Charge distribution

To calculate the electron density from a given electron distribution the cloud in cell (CIC) method is used, in which the charge of each simulation electron is fairly distributed over both nearest grid points: if the distance of the simulation electron to one of these grid points is $\alpha\ \Delta x$, this point is assigned a fraction $1-\alpha$ of the charge the electron is carrying, and the other one a fraction α of that charge. In calculating the average electron velocity and average electron energy at each grid point the same method is used.

A different method is used in calculating the electron temperature T_{ni} in each interval by averaging the energy of all electrons in that interval, using the simplified definition

$$T_{ni} = \frac{2}{3k_B} \langle u_i \rangle \tag{3.2.4}$$

In the program the impurity concentration is defined at each grid point and the charge density is given by

$$\rho_i = q\ (N_{Di} - n_i) \tag{3.2.5}$$

3.2.4 The Poisson equation

If the charge density is known, the Poisson equation (3.2.6) is used to calculate the potential and the field everywhere in the device

$$\nabla^2 \Psi = - \frac{\rho}{\epsilon} \tag{3.2.6}$$

A three point differential scheme is used, and the tridiagonal matrix solver DGTSL from the LINPACK-library is used to calculate the potential at each grid point, given the external applied potentials at the final grid points. Since the potential is defined at the grid points, the electric field can be calculated in a straightforward way from the potential distribution

$$E_i = (\Psi_{i-1} - \Psi_i) \frac{1}{\Delta x} \tag{3.2.7}$$

Because of the various valleys and the fact that the X-valleys in Si are strongly elliptical, each electron has its own effective field, as described in chapter 2.3. Using Herring-Vogt transformations, the effective electric field of each electron can be calculated. If by scattering the electron jumps from one valley to another, this effective field has to be readjusted.

3.2.5 Electron transport

The situation in a device is much more complicated than in bulk material. As the electron is moving through the device it changes its position and by doing so the force working on the electron, the material properties (and so the scatter rates) change. To be able to deal with these problems it is necessary to limit the maximum time of flight in such a way that the electron never moves more than one interval Δx during each flight. Because it is also necessary to reevaluate the potential distribution at regular time intervals, the introduction of a time step at the end of which all free flights are ended with a self-scatter process, is a natural one. The actual size of this time step is discussed in section 3.3.

Assuming that the position of an electron indeed does not change more than the size of one interval Δx, and assuming all parameters to be sufficiently smooth, it is reasonable to state that during each flight the electric field, the impurity- and electron concentrations and the electron temperature remain constant, and that these parameters depend only on the initial position of the electron. All this makes the Monte Carlo process in a device very similar to the process in bulk material, and the same conduction band model and scatter mechanisms are involved. To calculate the position of each electron the equations in appendix 2.7.2 are used.

A problem occurs at the interface between drift region and contact regions owing to the difference in weight factor between the regions, as described in section 3.2.1. In the case that the electron moves from the drift region (charge weight equal to one) to the source region (weight equal to q_s) it is stated that the probability that the electron actually moves into the source region equals $1/q_s$ (the decision being made by the random number generator). If this situation occurs, the electron changes its charge weight to q_s. All other times the electron is completely annihilated. A similar procedure is followed if an electron moves from the drift to the drain region. If however an electron moves from the source to the drift region, the problem is more complex. The electron having a weight q_s (an integer value) is split up in q_s new electrons, all having a weight equal to one. All these electrons have identical electrical properties and positions, and are distributed randomly among the already existing simulation electrons. Naturally all this coming and going of electrons demands a very accurate administration. If an electron moves from the drain to the drift region, the problem is similar.

3.2.6 Avalanche generation

In the Monte Carlo program extra simulation electrons actually can be generated by weak avalanche processes. If a hot electron is subject to impact ionization scattering, it loses part of its energy, as described in section 2.3.4.2. In addition a new simulation electron is generated at the same position in the device with the same charge weight as the scattering electron, and with zero energy and zero wave vector. Because the Monte Carlo program can only deal with electrons, no holes are generated.

3.2.7 Modeling of the contacts

If the Si device is highly doped at the metal contact but not yet degenerated, and if externally a voltage U_{ex} is applied to the contact, then the internal voltage at the contact is given by

$$\Psi_0 = U_{ex} - \frac{k_B T_0}{q} \ln \left(\frac{\sqrt{N_C N_V}}{N_D} \right) \tag{3.2.8}$$

where N_C and N_V are the number of states in conduction and valence band, respectively.

To describe the contacts two different models are used.

1. The regenerating contact model: the exact moment an electron reaches a contact, it is absorbed at that contact and a new electron is generated instantaneously at the other contact in a hemi-Maxwellian distribution (velocity directed into the device). An iterative process is used to calculate the exact moment a simulation electron reaches the contact. A possible difference in charge weight between source and drain region is taken into account. If the avalanche generation is negligible, the device as a whole always is necessarily charge neutral.
2. The neutral contact model: this model uses the assumption of charge neutrality at the metal contacts. In the Monte Carlo program this implies that each time step the charge density at the contacts is calculated, and simulation electrons are injected, or removed at the contact to get as close as possible to charge neutrality. The injected electron is placed at a random position in the first interval bordering the contact and drawn from a Maxwellian distribution. A possible difference in charge weight between source and drain is taken into account. In this model the device as a whole is not necessarily charge neutral.

3.2.8 Calculation of the device currents

In order to calculate the primary device currents, three different methods are used. They all can be implemented easily in the program. By implementing different methods that all calculate the device currents an extra check on the validity of the calculations is obtained.

1. The number of electrons that is absorbed by the drain, minus the number of electrons absorbed by the source during a certain interval of time Δt_c

equals the total current through the device. Of course the charge weight of each electron has to be taken into account. In a similar way the avalanche current is calculated by simply counting the number of simulation electrons generated by the impact ionization scattering. Because of the counting of single electrons the noise calculated is increasing at decreasing device current and decreasing time interval Δt_c. Especially if the transient behavior of a device is studied, this can be a problem. In the steady-state studies Δt_c can in principle chosen to be arbitrarily long, just to reduce the noise.

2. The primary device current density j_x at each position x can be defined by $j_x = - q\, n_x v_{Dx}$. Now if M_i is the number of simulation electrons situated in the i-th interval between x and Δx, and if $q_i M_r$ is the total number of electrons represented by each simulation electron in the i-th interval (see section 3.2.2), then by definition the electron density n_x and electron drift velocity v_{Dx} at position x read

   $$n_x = \frac{q_i M_r M_i}{O_s \Delta x} \quad \text{and} \quad v_{Dx} = \frac{1}{M_i} \sum_{j=1}^{M_i} w_{ji} \tag{3.2.9}$$

   Here w_{ji} represents the velocity of the j-th electron at the i-th interval. Then the primary device current i_x at each position x reads

   $$i_x = j_x \cdot O_s = - q\, \frac{q_i M_r}{\Delta x} \sum_{j=1}^{M_i} w_{ji} \tag{3.2.10}$$

   Equation (3.2.10) is also valid at the contacts. The numerical noise is high if M_i is low. A noise reduction is obtained if the current is calculated regularly and averaged during a long total simulation time. Again this is only applicable in the steady-state situation.

3. A noise reduction also is obtained if the current is averaged over the whole length of the device. Then the total device current I_D can be defined by

   $$I_D = \frac{1}{L} \int i_x \, dx \cong \frac{1}{L} \sum_{i=1}^{N_x} i_x \cdot \Delta x \tag{3.2.11}$$

   Here N_x represents the number of spatial intervals in the device. Combining (3.2.10) and (3.2.11) yields

   $$I_D = - \frac{q}{L} \cdot M_r \cdot \sum_{j=1}^{M} q_j w_j \tag{3.2.12}$$

q_j is the charge weight of the j-th simulation electron and w_j the velocity of the j-th electron. Because equation (3.2.12) gives the average current over the device, this is not very accurate if the avalanche current (which is generated locally somewhere in the device) is of the same order of magnitude as the primary device current.

3.3 Stability considerations

Because of discretization in time and space, the stochastic character of the Monte Carlo method and computational errors the simulation results are not exact. If the errors are too large the simulation results have no practical significance at all. To ensure that the errors are small, simulation and discretization parameters have to satisfy a number of conditions. Since in Monte Carlo simulations electrons can be considered as a collisionless electron plasma with scatter processes superimposed, these conditions can be derived from plasma theory. An extensive treatment of the stability considerations has been given by Hockney and Eastwood [1981].

1. To achieve integration stability, the following condition must be fulfilled

   $$\omega_{pn} \cdot \Delta t < 2 \tag{3.2.13}$$

   Here ω_{pn} is the plasma frequency defined by

   $$\omega_{pn} = \left[\frac{n\, q^2}{\epsilon_r \epsilon_0 m_n^*} \right]^{1/2} \tag{3.2.14}$$

   This conditions guarantees that errors in the electron position caused by integration over the electron velocity remain small enough for the simulation to stay stable.

2. To be sure that electrons do not pass field fluctuations unnoticed, Δx and Δt are chosen in such a way that during a single free flight never more than one grid point is passed. This means

   $$w_{max} \cdot \Delta t < \Delta x \tag{3.2.15}$$

   Here w_{max} is the maximal electron velocity.

3. Considering an uniform collisionless and field-free plasma, any moving charge passes this plasma undisturbed. However if a Monte Carlo model is

used, there are some stochastic errors in the uniform charge distribution. Therefore small electric fields will arise, and a charge cannot pass undisturbed anymore. A measure of this disturbance is given by the collision time τ_c, the average time it takes for a passing charge to be deflected 90° from its original course. As long as the collision frequency $\lambda_c = \tau_c^{-1}$ is much smaller than the minimal scatter frequency λ_{min}, this effect can be neglected. Using the expression of Hockney for λ_c

$$\lambda_c = \tau_c^{-1} = \frac{L}{M\ (L_D + \Delta x)} \frac{\omega_{pn}}{2\pi} << \lambda_{min} \tag{3.2.16}$$

Here L_D is the Debye length at equilibrium, defined by

$$L_D = \left[\frac{\epsilon_r \epsilon_0 k_B T_0}{q^2 n} \right]^{1/2} \tag{3.2.17}$$

4. Because of errors in the calculation of the electron momentum, the total energy of the system is increasing continuously. So the total simulation time should not be too long. A heating time τ_H is defined as the time in which the average increase in electron energy is $\frac{1}{2} k_B T_0$. As long as the total simulation time T is smaller than τ_H, the effect can be neglected. Using Hockney's expression for τ_H we find

$$\tau_H = \left(\frac{L_D}{\Delta x} \right)^2 K_H \tau_c \quad > T \tag{3.2.18}$$

The heating constant K_H depends on the algorithm used and has a value of 1 in our program.

3.4 Simulation results

The program described has been used to simulate various modifications of the drift device. These devices also have been simulated using the hydrodynamic device simulator described in the next chapter. Simulation results are compared. In this way it is possible to evaluate the validity of the use of the hydrodynamic simulator. The results of this comparison are discussed in detail in section 4.3 after the introduction of the hydrodynamic simulator.
In this section we describe several devices and check whether the stability considerations have been met. Simulations are performed concentrating on the question which choice of contact model (neutral contact or regenerating contact model, see section 2.3.7) is the best.

The simulated devices have a total device length of 1000, 800 and 700 nm, and a corresponding drift region of 400, 200 and 100 nm respectively. In the drift regions the doping rate is 10^{16} cm^{-3}, in the contact regions 10^{18} cm^{-3}. The source is grounded and a positive voltage is applied to the drain.

Taking into account that $m_n^* \geq 0.256\ m_0$ (see figure 2.14a) and that $10^{16} \leq n \leq 10^{18}$ cm^{-3} everywhere in each device, a maximal plasma frequency $\omega_{p,max} \cong 30$ rad/ps is found. This fixes the upper limit for the time step to Δt: $\Delta t < 0.06$ ps. Simulations in bulk show that in general $w_{max} < 10^6$ m/s (figure 2.12b), so a choice of $\Delta t \cong 0.01$ ps and $\Delta x \cong 10$ nm satisfies both conditions (3.2.13) and (3.2.15). The Debye length varies per region, but has a minimum value $L_{D,min} \cong 4$ nm. In our simulations the number of simulation electrons M equals 7500 or more, and, given a maximal device length of $L = 1000$ nm, the collision frequency λ_c always is less than $5 \cdot 10^{10}$ Hz. As in Si the scatter frequency is about 10^{12} Hz (figure 2.3), also condition (3.2.16) is fulfilled. Finally with the parameters defined above, it is found that $\tau_H \geq 2.5$ ps. However it has to be noted that this is really the worst case situation for these devices. In general the number of electrons M exceeds the initial number, m_n^* is larger than $0.256\ m_0$, and the electron density is less than 10^{18} cm^{-3} in most of the device. So the value of τ_H is sufficiently large to allow a total simulation time of 10 ps that is needed for these devices that reach a steady state after 2 or 3 ps.

Simulations using the parameters given above ($\Delta t = 0.01$ ps, $\Delta x = 10$ nm, $T = 10$ ps and $M = 7500$) have been performed for the three devices having two different contact models. The applied potential has been varied from 0 to 5 Volts. Although the parameters have been chosen in such a way to guarantee numerical stability, strong instabilities may occur if the regenerating contact model is used. This is caused by electron accumulation at the source contact due to the fact that all electrons entering the drain contact with a high velocity are generated at the source contact, having only a low average velocity. This accumulation can become so strong that condition (3.2.13) is violated. The problem can be solved by a further decrease of the time step. A time step $\Delta t = 0.001$ is found to be adequate, but causes a substantial rise in simulation time. In the neutral contact model these instabilities do not occur as electron accumulation at the contacts is actively suppressed by the model.

In figure 3.1a the characteristics of the three devices have been plotted for the neutral contact model (solid lines). In addition the characteristics of the 100 nm device obtained by the regenerating contact model (dashed line) is given. In the plot error bars have been included. The total simulation time is 10 ps. The regenerating contact model gives slightly smaller primary

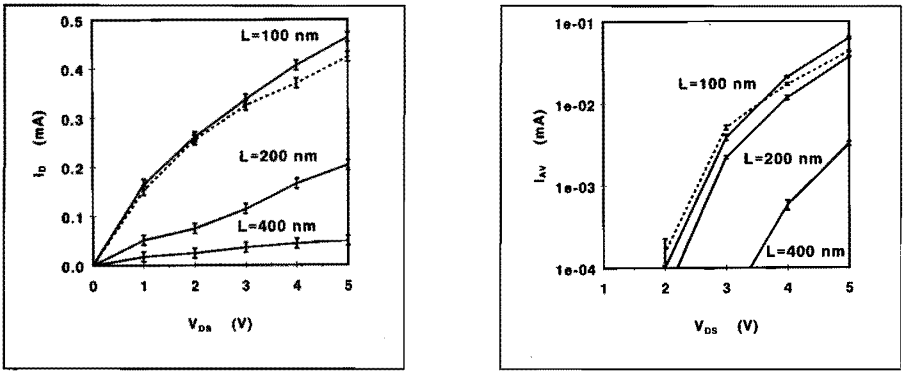

Fig 3.1 Primary device current (a) and avalanche current (b) for various drift devices as a function of applied voltage; results obtained using both the neutral contact model (solid line) and the regenerating contact model (dashed line) have been plotted.

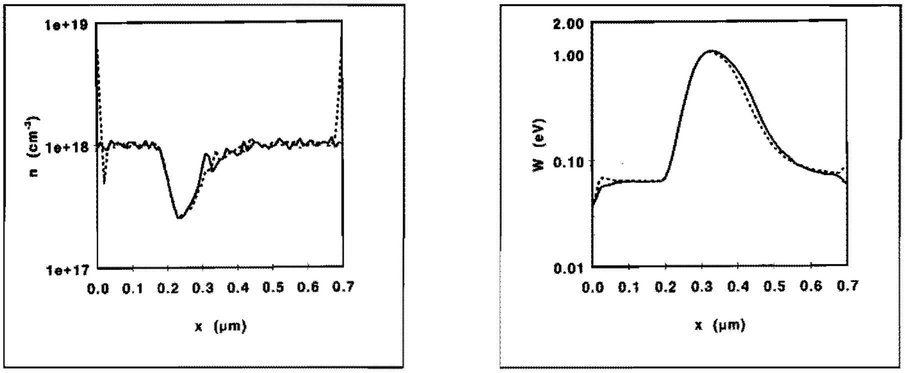

Fig 3.2 Electron density (a) and energy distribution (b) in a 100 nm drift device; V_{ds}= 4 V; the solid line represents data obtained using the neutral contact model, the dashed line using the regenerating contact model.

currents, but for most part of the curve this difference lies within the calculated error margin due to the stochastic character of the Monte Carlo method.

In figure 3.1b the calculated avalanche currents have been plotted. Again the difference between the different contact models is small. In figure 3.2 the electron energy and electron distribution of the 100 nm device are given for both contact models. The strong charge accumulation obtained by the regenerating model is clearly seen. In the energy plot the difference is only marginal.

In all further simulations the neutral contact model has been preferred due to superior stability properties. Furthermore this model seems more in accordance with the physical reality of contact behavior.

Chapter 4

TIME-DEPENDENT HYDRODYNAMIC DEVICE SIMULATION IN ONE DIMENSION

4.1 Introduction

In chapter 1 the first three moments of the Boltzmann transport equation are given by equations (1.4.16) to (1.4.24) for both holes and electrons respectively. In this chapter it will be described how these are solved for one-dimensional devices. If all three moments are taken into account the model is called hydrodynamic; if only the first two moments are considered and all energy effects are neglected, the model is called a drift-diffusion model. The drift-diffusion method is widely used and well known in the field of semiconductor device modeling. In section 4.2 it is described how the full set of equations is discretized and implemented in a one-dimensional hydrodynamic simulator.

Several drift devices have been simulated. The results are presented in section 4.3. Transient as well as steady-state results are compared to results obtained by the Monte Carlo method described in the previous chapter. This comparison will give insight into the limits of applicability of the hydrodynamic model. The simulator will also be used to look at some terms in the hydrodynamic equations in more detail, in particular the velocity convection $(\vec{J}_n \cdot \nabla)\, \vec{v}_n$, the energy convection $\frac{1}{2}\, m^*_n |\vec{v}_n|^2$, and the heat flow $\vec{Q}_n$. An overview of the conclusions is given in section 4.4.

4.2 The one-dimensional hydrodynamic device simulator

A one-dimensional device simulator is developed using the first three moments of the Boltzmann Transport Equation. Only electron transport is taken into consideration. The program allows a number of optional simplifications.

1. The transport parameters are defined as a function of electric field or as a function of energy.
2. The inclusion of the energy conservation equation is optional (in this way, the classical drift-diffusion method is considered as a special case of the more general hydrodynamic method).
3. The inclusion of the velocity convection $(\vec{J}_n \cdot \nabla)\, \vec{v}_n$ is optional.
4. The inclusion of the energy convection $\frac{1}{2}\, m^*_n |\vec{v}_n|^2$ is optional.
5. Several models of the heat flow $\vec{Q}_n$ are present.
6. The inclusion of generation-recombination is optional.

The simulator has been structured as follows:

1. The device to be analyzed has to be defined.
2. The equilibrium situation is calculated (zero applied voltage).
3. Potentials are applied at the contacts and the Poisson equation is solved to calculate the electric field in the device.
4. The transport parameters μ_n, m_n^*, τ_{pn} and τ_{wn} everywhere in the device are calculated as a function of the local electron energy (or, optional, as a function of the local electric field).
5. The current density and electron velocity are calculated by solving the momentum conservation equation (1.4.18).
6. The electron energy and temperature are calculated by solving the energy conservation equation (1.4.21).
7. The electron density is calculated using the carrier conservation equation (1.4.16).

Steps 3 to 7 are repeated until sufficient accuracy is obtained.

4.2.1 Device definition

A non-uniform one-dimensional grid is used to discretize the device. The number of intervals is N_x, the size of the i-th interval equals Δx_i. The grid points are labelled 0 to N_x. The i-th interval is defined between grid points i-1 and i. The charge density, the potential, the doping concentration and the average electron energy are defined at the grid points; the electric field, the current density and the electron velocity are defined in the intervals. Only electron transport is considered, so everywhere in the device $n >> p$. Ohmic contacts are defined at both ends of the device. In analogy with MOS devices, the contacts are called source and drain. Band gap narrowing is included by defining the intrinsic carrier concentration n_I as a function of the local doping concentration

$$n_I = n_{I0} \exp \left(V_G^0 \left(F + (F^2 + \tfrac{1}{2})^{1/2} \right) \right) \tag{4.2.1}$$

Here F is defined by

$$F = \log \left(\frac{N_A + N_D}{N_{norm}} \right) \tag{4.2.2}$$

The values of the parameters are (Slotboom and De Graaff [1976]): $n_{I0} = 1.48 \cdot 10^{10}$ cm^{-3}, $V_G^0 = 9$ mV, $N_{norm} = 10^{17}$ cm^{-3}.

4.2.2 Equilibrium solution

If no external voltages are applied to source and drain, no currents will be flowing in the device and the electron temperature will be equal to the lattice temperature. The electron and hole density are given everywhere in the device by the Boltzmann factors

$$n = n_I \exp \left[\frac{q \Psi}{k_B T_0} \right] \tag{4.2.3}$$

$$p = n_I \exp \left[- \frac{q \Psi}{k_B T_0} \right] \tag{4.2.4}$$

Here Ψ is the potential difference between the mid-gap and the Fermi-level. Combining these equations with the Poisson equation (1.4.1) yields a non-linear differential equation in Ψ. A three point discretization scheme is used; the Newton-Raphson method combined with the tridiagonal matrix solver DGTSL is used to calculate the potential at each grid point. Consequently the electron and hole density at each grid point are known.

4.2.3 The momentum conservation equation

The equilibrium solution is used to initialize the non-equilibrium calculations. Voltages are applied to the source and drain contacts and the new potential distribution is calculated. Using the E_{eff}-W_n relation, as obtained from Monte Carlo simulations (2.5.10), the transport parameters m_n^*, τ_{pn} and τ_{wn} are calculated as a function of W_n.

For given potential Ψ, electron density n, and transport parameters, next the electron current density is calculated by solving equation (1.4.18). A highly efficient discretization scheme given by Scharfetter and Gummel [1969] to solve the drift diffusion equation (1.4.25) is used. This method assumes a linear variation of Ψ along each interval and the current density $\vec{J}_n$ is assumed to be constant along each interval. This discretization scheme has been dealt with extensively in the literature. An extension of this method, that makes it possible also to include the drift term due to a gradient in the electron temperature was presented by Tang [1984]. The extra assumption needed is a linear variation of T_n along each interval. A further extension by Rudan and Odeh [1986] allows the inclusion of the velocity convection, which is taken to be constant along each interval. An alternative way to include the velocity convection term will be presented in this work. Generally, if the interval sizes are small enough and if abrupt changes in doping profile are avoided, all assumptions are acceptable.

To obtain equation (1.4.18) the assumption was made that the conduction band is parabolic. Section 4.2.3.1 deals with the discretization of (1.4.18). To model a non-parabolic conduction band an extra term has to be included. This is dealt with in section 4.2.3.2.

4.2.3.1 Parabolic band structure

The momentum conservation equation (1.4.18) is repeated here for the one-dimensional situation, assuming the conductivity effective mass to have a constant value

$$J_n + n\,\tau_{pn}\frac{\partial}{\partial t}\left(\frac{J_n}{n}\right) - \frac{\tau_{pn}}{q}\left(J_n\cdot\frac{\partial}{\partial x}\right)\frac{J_n}{n} = q\,\mu_n n\,E + \mu_n\frac{\partial}{\partial x}\left(nk_BT_n\right) \qquad (4.2.5)$$

Now consider the current density at the i-th interval between mesh points i-1 and i. The equation (4.2.5) is split in two, using the interval label i as an extra suffix

$$J_{ni} + n_i\,\tau_{pni}\frac{\partial}{\partial t}\left(\frac{J_{ni}}{n_i}\right) = J_{0i} \qquad (4.2.6)$$

$$J_{0i} = q\mu_{ni}\left[n_i\left(E_i + V_{ni}\right) + \frac{\partial}{\partial x}\left(\frac{n_ik_BT_{ni}}{q}\right)\right] \qquad (4.2.7)$$

$$V_{ni} = \frac{m^*_{ni}}{q}\,v_{ni}\frac{\partial}{\partial x}v_{ni} \qquad (4.2.8)$$

Equation (4.2.6) is solved using an implicit Euler time discretization scheme with time step Δt. This yields

$$J^{new}_{ni} = \frac{\tau_{pni}n^{new}_iJ^{old}_{ni} + \Delta t\,n^{old}_iJ_{0i}}{n^{old}_i\left(\tau_{pni} + \Delta t\right)} \qquad (4.2.9)$$

Equation (4.2.7) is solved using the modified Gummel-Scharfetter algorithm. The discretization scheme used is the one given by Forghieri et al. [1988]. In this scheme not the mobility, as is usual, but the diffusion coefficient is assumed to be constant within each interval. As shown in section 2.6, D_n is a less sensitive function of electron energy than the mobility, so this assumption certainly is allowed and even better. Using a constant D_n, the energy conservation equation can be discretized using a scheme analogous to the Gummel-Scharfetter scheme. This cannot be done if the mobility is assumed to be constant. The gradient of the electron temperature T_n is assumed to have a constant value β_{Ti} over each interval and is defined by

$$\beta_{Ti} = \frac{\partial}{\partial x}\frac{k_BT_n}{q} = \frac{k_BT_{ni} - k_BT_{ni-1}}{q\,\Delta x_i} \qquad (4.2.10)$$

Combining equations (4.2.7) and (4.2.10) yields

$$J_{0i} = qD_{ni} \left(\frac{q\,n}{k_B T_n} \left(E_i + V_i + \beta_{Ti} \right) + \frac{\partial\,n}{\partial\,x} \right) \tag{4.2.11}$$

Defining

$$\gamma = \frac{E_i + V_i}{\beta_{Ti}} \tag{4.2.12}$$

The parameter γ is a constant since E_i, V_i and β_{Ti} are constants on each interval. Next multiplying equation (4.2.11) with a factor $T_n^{\gamma+1}$ yields

$$J_{0i}\, T_n^{\gamma+1} = qD_{ni} \left[\frac{q\,n}{k_B} T_n^{\gamma} \left(\gamma + 1 \right) \beta_{Ti} + T_n^{\gamma+1} \frac{\partial\,n}{\partial\,x} \right] \tag{4.2.13}$$

This equation is rewritten using definition (4.2.10)

$$\frac{k_B J_{0i}}{q\,\beta_{Ti}} \quad T_n^{\gamma+1} \frac{\partial\,T_n}{\partial\,x} = qD_{ni} \left[n \left(\gamma + 1 \right) T_n^{\gamma} \frac{\partial\,T_n}{\partial\,x} + T_n^{\gamma+1} \frac{\partial\,n}{\partial\,x} \right] \tag{4.2.14}$$

The integration of (4.2.14) over the interval (x_{i-1}, x_i) is straightforward and yields

$$\frac{k_B J_{0i}}{q\,\beta_{Ti}} \frac{T_n^{\gamma+2}}{\gamma + 2} \Bigg|_{x_{i-1}}^{x_i} = qD_{ni}\; n\, T_n^{\gamma+1} \Bigg|_{x_{i-1}}^{x_i} \tag{4.2.15}$$

Then J_{0i} can be solved

$$J_{0i} = qD_{ni}(\gamma + 2)\, \beta_{Ti} \left(\frac{q\,n_i}{k_B T_{ni}} \frac{1}{1 - \left(\frac{T_{ni-1}}{T_{ni}} \right)^{\gamma+2}} - \frac{q\,n_{i-1}}{k_B T_{ni-1}} \frac{1}{\left(\frac{T_{ni}}{T_{ni-1}} \right)^{\gamma+2} - 1} \right) \tag{4.2.16}$$

Define

$$\Delta_i = (\gamma + 2)\, \log \left(\frac{T_{ni-1}}{T_{ni}} \right) \tag{4.2.17}$$

Then (4.2.16) can be rewritten

$$J_{0i} = - \frac{q\,(\gamma + 2)\, D_{ni}\beta_{Ti}}{\Delta_i} \left(\frac{q\,n_i}{k_B T_{ni}} \mathcal{B}(\Delta_i) - \frac{q\,n_{i-1}}{k_B T_{ni-1}} \mathcal{B}(-\Delta_i) \right) \tag{4.2.18}$$

Here $\mathcal{B}(\Delta)$ represents the Bernoulli function, given by

$$\mathcal{B}(\Delta) = \frac{\Delta}{\exp(\Delta) - 1} \tag{4.2.19}$$

If $T_{ni} = T_{ni-1}$ then γ and Δ_i are not defined. However in the limit $T_{ni} \rightarrow T_{ni-1}$ J_{0i} can be calculated. If everywhere in the device $T_n = T_0$, then the original Scharfetter-Gummel scheme as used in the drift-diffusion model, is obtained.

A problem is the calculation of the velocity convection V_i. This term very easily gives rise to instabilities, and an upstream discretization scheme is necessary. If $v_{ni} > 0$ this gives

$$v_{ni} \frac{\partial}{\partial x} v_{ni} = v_{ni} \frac{v_{ni} - v_{ni-1}}{\frac{1}{2}\,(\Delta x_i + \Delta x_{i-1})} \tag{4.2.20}$$

If $v_{ni} < 0$ this gives

$$v_{ni} \frac{\partial}{\partial x} v_{ni} = v_{ni} \frac{v_{ni+1} - v_{ni}}{\frac{1}{2}\,(\Delta x_i + \Delta x_{i+1})} \tag{4.2.21}$$

If J_{ni} is known, v_{ni} can be calculated

$$v_{ni} = - \frac{J_{ni}}{\frac{1}{2}\,q\,(n_i n_{i-1})^{1/2}} \tag{4.2.22}$$

4.2.3.2 Non-parabolic band structure

The momentum conservation equation including a non-parabolicity correction was given by Bordelon et al. [1990]. An equation very similar to (1.4.18) was obtained. The difference is the inclusion of the parameter H_n in the diffusion term. However our calculations show the expression of H_n given by Bordelon to be incorrect. The correct expression for the momentum conservation equation is (if the velocity and the energy convection are omitted)

$$J_n + n\,\tau_{pn} \frac{\partial}{\partial t}\left(\frac{J_n}{n}\right) = q\,\mu_n n\,E + \mu_n\,(1+2\alpha W_n)^3 \frac{\partial}{\partial x}\left[n k_B T_n \frac{(1+\alpha W_n)}{(1+2\alpha W_n)^2} \right] \tag{4.2.23}$$

The procedure to solve this equation uses again a Gummel-Scharfetter scheme identical to the one followed in section 4.2.3.1. In this scheme it is necessary to assume the non-parabolicity factor to be constant at each interval. In analogy to equation (4.2.11) it has been found

$$J_{0i} = qD_{ni} \left[\frac{q\,n}{k_B T_n} \left(E_i + V_i + H_n \beta_{Ti} \right) + H_n \frac{\partial}{\partial x}\, n \right] \tag{4.2.24}$$

Here H_n is defined by

$$H_n = (1 + \alpha W_n) \cdot (1 + 2\alpha W_n) \tag{4.2.25}$$

Now define

$$\gamma = \frac{E_i + V_i}{H_n \beta_{Ti}} \tag{4.2.26}$$

Combining equations (4.2.24) and (4.2.26) yields

$$J_{0i} = qH_{ni} D_{ni} \left[\frac{q\,n}{k_B T_n} \left(\gamma + 1 \right) \beta_{Ti} + \frac{\partial\, n}{\partial\, x} \right] \tag{4.2.27}$$

The final result is nearly identical to equation (4.2.18) with the exception of an extra factor H_{ni} at the RHS and a different definition of γ.

4.2.4 The energy conservation equation

For Ψ_i, n_i and J_{ni} given everywhere in the device, the energy conservation equation (1.4.21) has been solved, yielding values for T_{ni} and W_{ni}. For convenience the energy conservation equation is repeated here for the one-dimensional case

$$\frac{\partial\, (nW_n)}{\partial\, t} + \frac{\partial\, S_n}{\partial\, x} = E \cdot J_n - n\, \frac{W_n - W_0}{\tau_{wn}} + GR \cdot W_n \tag{4.2.28}$$

Here W_n is defined by

$$W_n = \frac{3}{2}\, k_B T_n + \frac{1}{2}\, m_n^*\, v_n^2 \tag{4.2.29}$$

And S_n is defined by

$$S_n = -\, Q_n - (\, W_n + k_B T_n\,)\, \frac{J_n}{q} \tag{4.2.30}$$

First the energy flow S_n is discretized. At the i-th interval between mesh points i-1 and i the label i is used as an extra suffix. Assuming S_{ni} to be constant over each interval, S_{ni} can be given by the general equation

$$S_{ni} = A_i\, k_B T_{ni} - B_i\, k_B T_{ni-1} + C_i \tag{4.2.31}$$

However the coefficients A_i, B_i and C_i depend on the way the heat flow Q_n is modeled. Three models are considered here

1. The adiabatic model which assumes $Q_n = 0$.
2. The model following the Wiedemann-Franz law which defines

$$Q_n = -\kappa_n \frac{\partial}{\partial x} k_B T_n \text{ and } \kappa_n = (\tfrac{5}{2} + c)\, D_n\, n \tag{4.2.32}$$

The value of the factor c is nearly independent of the material and is given a value of $c = 0.8$.

3. The model presented by Bordelon et al. [1990] who deduced that both the Q_n and the term $k_B T_n J_n / q$ can be neglected. Then equation (4.2.30) reads

$$S_n = - W_n J_n / q \tag{4.2.33}$$

Combining the three models mentioned with the general equation (4.2.31) the following is obtained:

Ad 1 For the adiabatic model the parameters A_i, B_i and C_i are defined by

$$A_i = 0; \quad B_i = \frac{5}{2} \frac{J_{ni}}{q}; \quad C_i = -\frac{1}{2} m^*_{ni} \frac{J^3_{ni}}{q^3 n^2_i} \tag{4.2.34}$$

Ad 2 For the Wiedemann-Franz model, equation (4.2.30) in combination with (4.2.32) can be rewritten in the form

$$\frac{\partial}{\partial x} (k_B T_n) + \frac{c_1}{n} k_B T_n = -\frac{c_2}{n} - \frac{c_3}{n^3} \tag{4.2.35}$$

In this equation the parameters c_1, c_2 and c_3 read

$$c_1 = \frac{\frac{5}{2} J_{ni}}{(\frac{5}{2} + c)\, q\, D_{ni}} \qquad c_2 = \frac{S_{ni}}{(\frac{5}{2} + c)\, D_{ni}} \qquad c_3 = \frac{\frac{1}{2} m^*_{ni} J^3_{ni}}{(\frac{5}{2} + c)\, q^3 D_{ni}} \tag{4.2.36}$$

In the discretization, an exponential variation of n along the interval is assumed

$$n = n_i \exp(\alpha x_i) \quad \text{where} \quad \alpha = \frac{1}{\Delta x_i} \log(\frac{n_i}{n_{i-1}}) \tag{4.2.37}$$

Solving the differential equation (4.2.35), combined with (4.2.37), yields an expression of the form (4.2.31) with the definitions for the parameters A_i, B_i and C_i

$$A_i = -\frac{(\frac{5}{2} + c)\, D_{ni}\, \tilde{n}_i}{\Delta x_i} \mathcal{B}(\Omega_i) \tag{4.2.38}$$

$$B_i = -\frac{(\frac{5}{2}+c)\, D_{ni}\tilde{n}_i}{\Delta x_i}\, \mathcal{B}(-\Omega_i) \tag{4.2.39}$$

$$C_i = -\frac{(\frac{5}{2}+c)}{\frac{5}{2}}\, \frac{1}{2}\, m^*_{ni}\, \frac{J^2_{ni}}{q^2}\, \frac{D_{ni}\tilde{n}_i}{\Delta x_i}\, [\mathcal{B}(\Omega_i)\, F_i - \mathcal{B}(-\Omega_i)\, F_{i-1}] \tag{4.2.40}$$

In these equations the parameters Ω_i, $\tilde{n}_i$ and F_i are defined by

$$\Omega_i = -\frac{\frac{5}{2}\,\Delta x_i J_{ni}}{(\frac{5}{2}+c)\, q\, D_{ni}\tilde{n}_i} \qquad \tilde{n}_i = \frac{n_i n_{i-1}}{n_i - n_{i-1}} \log\left(\frac{n_i}{n_{i-1}}\right) \tag{4.2.41}$$

$$F_i = \frac{1}{n_i^2} + \frac{2\,\alpha}{c_i n_i} + \frac{2\,\alpha^2}{c_i^2} \tag{4.2.42}$$

Ad 3 The Bordelon model is very similar to the adiabatic model. The difference is only a factor $k_B T_n$. This yields

$$A_i = 0; \quad B_i = \frac{3}{2}\,\frac{J_{ni}}{q}\,; \quad C_i = -\frac{1}{2}\, m^*_{ni}\, \frac{J^3_{ni}}{q^3 n_i^2} \tag{4.2.43}$$

For a given S_{ni} the electron temperature is solved from the energy conservation equation (4.2.28) by an implicit Euler time differentiation, and by using the tridiagonal matrix solver DGTSL.

4.2.5 Device currents

Two kinds of currents are defined:

1. The source and drain currents: The source and drain contacts are assumed to be ohmic and there is a voltage drop at the contacts, as described in section 3.2.7, the value of which is given by equation (3.2.8). Charge neutrality is assumed, which implies $n = N_D$ (as we are dealing only with electron transport). Generally the source and drain currents J_S and J_D consist of a transport and a displacement component, defined by

$$J_{S,D} = J_n + \epsilon\, \frac{\partial E}{\partial t} \tag{4.2.44}$$

2. Avalanche current: In the model no hole transport is taken into account. However at high fields weak-avalanche effects occur, and electrons and holes are created. In the model only electrons are generated and the total electron generation is called the avalanche current.

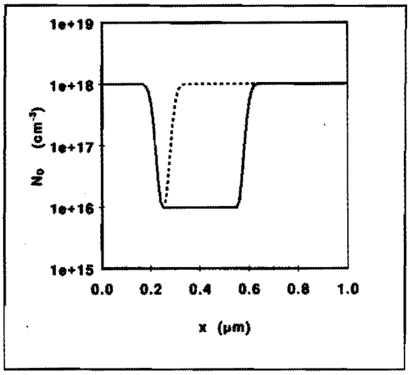

Fig 4.1 Doping profiles of different drift devices, one with a 100 nm drift region (dashed line) and one with a 400 nm drift region (solid line).

4.3 Simulation results

A series of control simulations have been performed on bulk material. Results were in close agreement with the Monte Carlo results described in 2.5. Next, a series of simulations using the n^+-n-n^+ drift device already described in section 3.4 has been performed. In figure 4.1 the doping profile of the devices with a drift region of 100 nm (dashed line) and 400 nm (solid line), respectively, is given. The simulation results will be used to obtain answers to the following questions:

1. Does, compared to the Monte Carlo method the hydrodynamic method describe hot electron behavior adequately.
2. How important are the velocity convection $(\vec{J}_n \cdot \nabla)\,\vec{v}_n$ and the energy convection $\frac{1}{2}\, m_n^* |\vec{v}_n|^2$.
3. Which heat-flow model is the best.

The bias voltage is varied from 0 to 5 V. The mesh lines are distributed uniformly over the device and the spatial resolution has been varied between 10 nm and 20 nm. It has been found that over the whole range of bias voltages the resolution mainly effects the transient behavior of the avalanche current. The effect on the steady-state currents is less then 2 % for both the avalanche and the device current. From these results it is concluded that a spatial resolution of 10 nm is sufficient. This value has been used in all further simulations described in this section.

Since band gap narrowing is not implemented in the Monte Carlo program, it also has not been taken into account in the hydrodynamic simulations. From the simulation results it has been concluded that the velocity convection and energy convection are of minor importance in Si devices and can be neglected. Their effect on any parameter never exceeds the 2 % margin. This in contrast with e.g. GaAs simulations where both terms are important due to the higher electron velocity in the material.

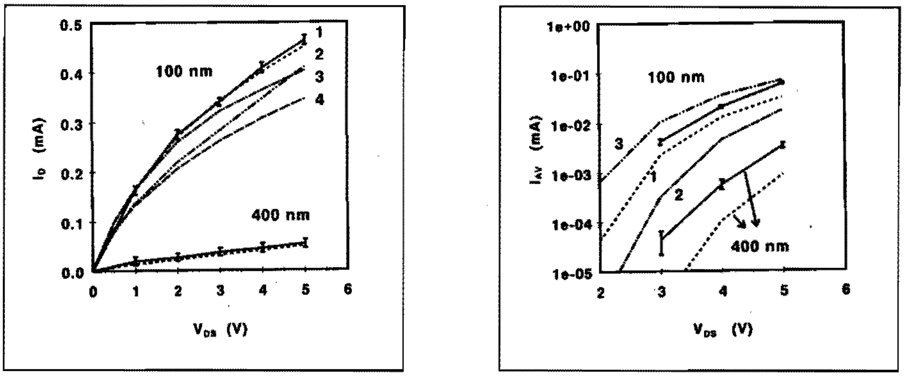

Fig 4.2 Steady-state drain current (a) and avalanche current (b) as a function of the applied voltage. Plotted are the Monte Carlo results (solid line), various hydrodynamic results (1. Adiabatic (dashed line), 2. Wiedemann-Franz (dash-dot), 3. Bordelon (dash-dot-dot)), and 4. the drift-diffusion model (large dashes).

In figures 4.2a and 4.2b the total drain current and the avalanche current in the steady state have been plotted as a function of the applied field. For the 100 nm device the Monte Carlo model (solid line with error bars) is compared to the drift-diffusion model (dashed line, index 4), and to three versions of the hydrodynamic model, using different heat-flow models: (1) the adiabatic model (dashed line), (2) the Wiedemann-Franz model (dash-dot line) and (3) the Bordelon model (dash-dot-dot). The transport parameters are calculated by a spline interpolation between various data points obtained by the Monte Carlo bulk simulator. No extra non-parabolicity factor is included. For the 400 nm device only the results from Monte Carlo simulations and the adiabatic hydrodynamic model have been plotted.

From figure 4.2a it can be seen that the adiabatic model gives the best fit to the Monte Carlo data (Elias et al. [1990]). The small discrepancy can be accounted for by the larger avalanche current calculated by the Monte Carlo method, which has been plotted in figure 4.2b. The rather large discrepancy between hydrodynamic and Monte Carlo results shown in figure 4.2b is due to a different model of the avalanche generation.
In the hydrodynamic simulations the avalanche parameters have been chosen in accordance with experimental data. The Monte Carlo program however, uses the Keldysh model. As has been discussed in chapter 2, there is good agreement between Monte Carlo and experimental results at high fields (see figure 2.8), but at low fields the Monte Carlo method overestimates the avalanche generation.

In figures 4.3 to 4.6 the internal energy and velocity distribution in the 100 nm and the 400 nm device have been plotted.

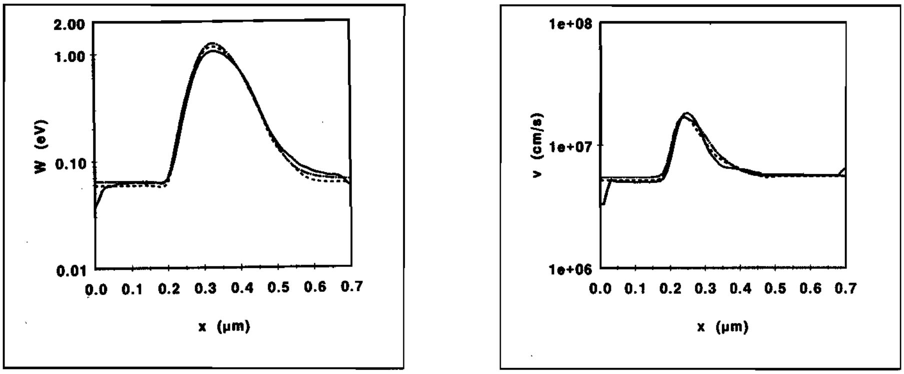

Fig 4.3 Electron energy (a) and electron velocity (b) distribution in the steady state in the 100 nm drift device. The drift region is defined from x=0.2 to x=0.3 μm, and the device is driven at 5 V. Monte Carlo results (solid line) and results from the adiabatic hydrodynamic model (dashed lines) are shown. The transport parameters have been defined in two different ways, either by a spline approximation (simple dashed line) or by an analyical function (dash-dots).

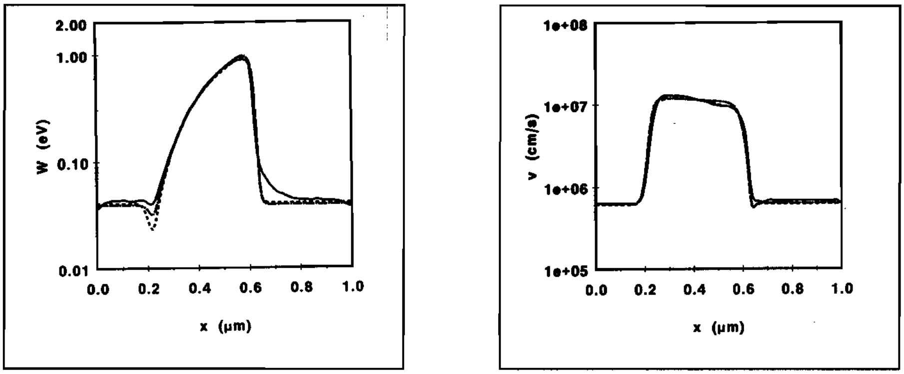

Fig 4.4 Plot similar to 4.3, but now the 400 nm device is taken into consideration; the active region is defined from 0.2 to 0.6 μm.

In figures 4.3a and 4.3b results obtained by the Monte Carlo (solid line) and by the adiabatic model (dashed lines) are compared for the 100 nm device. The voltage applied is 5 V. This yields a peak field of 400 kV/cm. The simple dashed line assumes a spline interpolation of the transport parameters, the dash-dot line the analytical approximations given by (2.5.5) and (2.5.10). Figures 4.4a and 4.4b give similar results for the 400 nm device. Here a peak field of 250 kV/cm has been found.

These figures show that the results obtained by the adiabatic model are in good agreement with the Monte Carlo results on a microscopic level. Furthermore, it is confirmed that the analytical expressions (2.5.5) and (2.5.10) give adequate models of the transport parameters. At the given voltages no significant effect on the drain current has been found. However, since the avalanche rate is very sensitive to the electron energy, due to the exponential relation in (2.5.6), a small difference in the energy gives rise to substantial differences in the avalanche current. This has been observed in our simulations. If the transport parameters are defined by (2.5.5) and (2.5.10) the avalanche current in general will be decreased substantially (in the worst case the value is about one third compared to the spline definition of the transport parameters). The main discrepancies between the adiabatic hydrodynamic and the Monte Carlo approach are the following:

1. In the 400 nm device the adiabatic model clearly gives a stronger cooling at the source side of the drift region. Furthermore a discrepancy in the lower energy tail occurs at the drain contact. However these have hardly any effect on the device characteristics or the impact ionization rate.
2. In the velocity distribution plots it can be noted that both methods show velocity overshoot, especially in the 100 nm device, and that the Monte Carlo method gives a slightly higher velocity maximum.

In figures 4.5 and 4.6 the similar results are presented as in figures 4.3 and 4.4, but now the Monte Carlo model (solid line) is compared to the Wiedemann-Franz model (dashed line) and the Bordelon model (dash-dot line).
As can be seen in figure 4.6a, in the 400 nm device these two energy models give fairly good agreement in the energy distribution compared to the Monte Carlo results. However large discrepancies occur in the velocity distribution: the Bordelon model gives values that are much too low, and the Wiedemann-Franz model gives a large velocity overshoot peak at the drain side. For the 100 nm device the discrepancy between the different models is even larger, for both velocity and energy distribution. Assuming a non-parabolic band structure by giving the non-parabolicity parameter α a value not equal to zero, and using the equation given in section 4.2.3.2, does not improve the fit between hydrodynamic and Monte Carlo results. This can be explained by considering the valley model used in the EMC method. At low energies the non-parabolicity of the X-valleys is hardly important. But at high energies a substantial part of the electrons will be situated in the L-valleys, and they have been assumed to be parabolic (see table 2.1).

Further, some simulations have been performed by varying the factor c in the Wiedemann-Franz heat-flow model. But again, no significant improvements have been obtained.

In figure 4.7 the transient behavior is shown for both the 100 nm and the 400 nm device when the voltage is abruptly switched on from 0 to 5 V. The solid line represent the Monte Carlo results, the dashed lines the

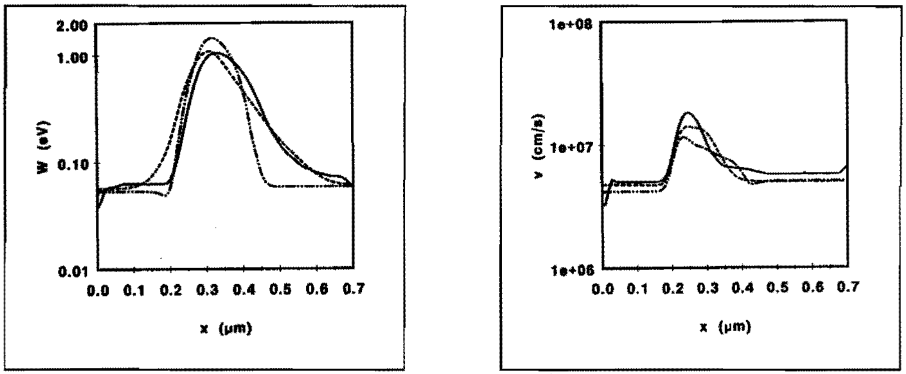

Fig 4.5 Plot similar to 4.3, but now the Wiedemann-Franz (dashed line) and the Bordelon hydrodynamic results (dash-dot line) are plotted together with the Monte Carlo results (solid line).

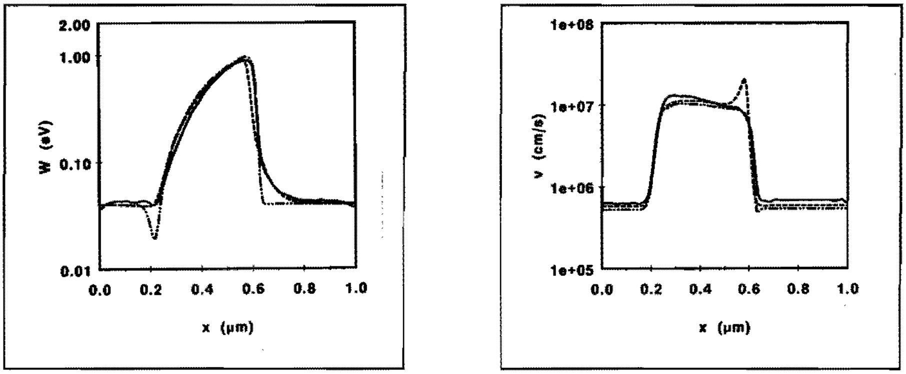

Fig 4.6 Plot similar to 4.5, but now the 400 nm device is taken into consideration.

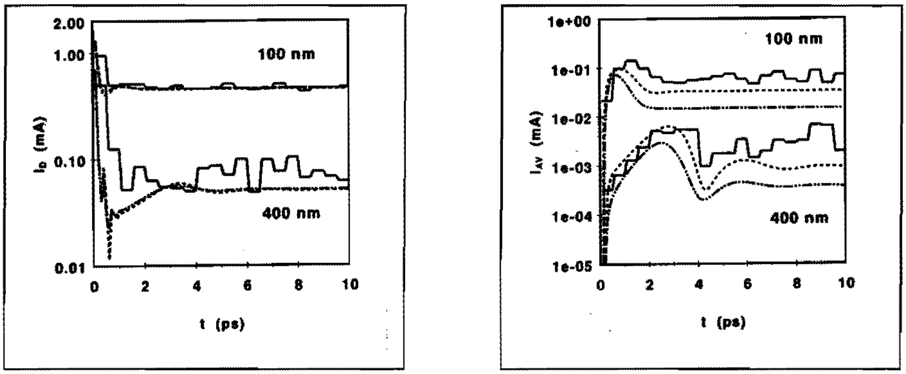

Fig 4.7 Transient behavior of the device current (a) and the avalanche current (b) in both the 100 and 400 nm drift device, driven at 5 V. Monte Carlo results (solid lines) and results from the adiabatic model (dashed lines) are given. The transport parameters have been defined in two different ways, either by a spline approximation (simple dashed line) or by an analyical function (dash-dots).

hydrodynamic results using the adiabatic heat-flow model (with a difference in the definition of the transport parameters). For the drain current (figure 4.7a) excellent agreement has been obtained, also when transient behavior is dominant, just after the voltage step has been applied. The fit for the avalanche current (4.7b) is less and, as noted before, rather sensitive to the exact transport parameter model. The discrepancy amounts to a factor of two.

4.4 Summary and conclusions

Summarizing it can be concluded that:

1. Both the velocity convection and the energy convection can safely be neglected, both in the steady state and in transient situations.
2. A spatial resolution of the mesh of 10 nm is adequate.
3. The non-parabolic model given by Bordelon et al. [1990] is not correct.
4. To represent the heat flow the adiabatic model has to be preferred over the Wiedemann-Franz and the Bordelon models.
5. If the adiabatic model is used, the hydrodynamic model gives good agreement with the Monte Carlo model, even for very small devices and high energy. This agreement holds for steady-state and transient simulations.
6. Neither the energy and velocity distributions, nor the drain current is very sensitive to the exact transport parameter definition. However, the effect on the avalanche current is large.

In this chapter all conclusions have been drawn by comparing different models. In order to get a better impression of the validity of the hydrodynamic model, simulation and experimental results are compared in the next chapter.

Chapter 5

QUASI TWO-DIMENSIONAL HYDRODYNAMIC MOS DEVICE SIMULATOR

5.1 Introduction

Solving the transport equations in a MOS device is essentially a two-dimensional problem. Contrary to drift-diffusion simulators, two-dimensional hydrodynamic simulators are not very robust, need much computing time, and give easily rise to numerical instabilities. Therefore in this chapter a more robust, quasi two-dimensional hydrodynamic MOS simulator is described, in which the electron transport is considered to be a one-dimensional problem in the direction parallel to the Si-SiO_2 interface. Since in a MOS device the charge transport is concentrated in a potential well (the channel) at the Si-SiO_2 interface, this is a reasonable first-order approximation. However, in calculating the electric potential also the vertical structure of the device is taken into account. The electron transport in the MOS is dealt with in a similar way as is described in the previous section.

In the simulator quantum effects (due to the potential well induced by the gate voltage) are accounted for by introducing a correction term in the electric potential of the gate contact, as was suggested by Van Dort and Woerlee [1991]. This is discussed in detail in section 5.2.

In section 5.3 it is shown how the electrical potential is calculated. Two important assumption have been made. The first assumption states that underneath the gate contact the horizontal field is much smaller than the vertical field. In that case the calculation of the potential can be considered to be a one-dimensional problem in the direction perpendicular to the Si-SiO_2 interface, and the charge induced at the gate is equal to the total charge induced in the semiconductor according to Gauss' law. By a double integration of the Poisson equation, taking into account the various boundary conditions, the potential at the Si-SiO_2 interface is calculated. The second assumption made is that the MOSFET operates in strong inversion. This means that the electron density at the surface is high, and the voltage drop over the channel in the direction perpendicular to the interface is small. Then the potential in the channel equals the potential at the Si-SiO_2 interface.

Simulation results are compared to experimental data on submicron devices. These data were obtained from Philips Research Laboratories (transistors with

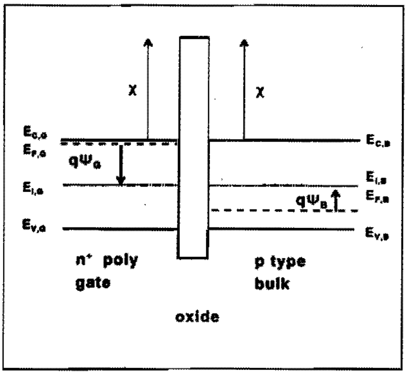

Fig 5.1 Band diagram for a n^+-poly gate on a p-substrate at flat-band condition.

channel lengths down to 0.25 μm) and from the literature (devices developed at MIT with channel lengths down to 0.09 μm, Shahidi et al. [1988]). In section 5.4 the devices and experimental results are described.

In 5.5 simulation results are compared to the experimental data. Special attention has been paid to avalanche generation and velocity overshoot in the devices. The chapter ends with a summary of the results in 5.6.

5.2 The electric potential at the gate contact

In most practical devices the gate consists of highly doped poly-Si material. In the devices considered, the gate is heavily n-type doped. This means that the Fermi-level in the gate is situated at the bottom of the conduction band. Since the potential at each position in the device is defined at mid gap, relative to the Fermi-level in the substrate, and since both crystalline and poly-Si have the same work function, it can be easily seen (figure 5.1) that the electric potential at the gate contact ψ_G can be obtained from

$$\psi_G = V_G + \frac{1}{2}\frac{U_{gap}}{q} \tag{5.2.1}$$

Here U_{gap} represents the energy gap between conduction and valence band.

In (5.2.1) V_G represents the external voltage applied to the gate contact. However (5.2.1) is only valid if the conducting electrons are situated at the bottom of the conduction band and if all quantization effects can be neglected. As was described in the introduction section 1.3.1, due to the high channel implant,this is not the case in deep-submicron MOS devices.

As the electrons are confined in a narrow potential well at the Si-SiO_2 interface, the motion of the electrons is quantized in the direction perpendicular to the interface and the energy difference Δu between the lowest energy level in the well and the bottom of the conduction band has to be accounted for. Now an effective gate potential can be defined

$$\Psi_G = \psi_G - \frac{\Delta u}{q} \tag{5.2.2}$$

Two different ways to calculate Δu are given:

1. Assuming the quantum well to be triangular, the following expression has been given by Stern and Howard [1967]

$$\Delta u \approx \frac{3}{2} \left(\frac{3\pi^2}{16} \; \frac{q^2 \hbar^2 E_y^2}{m_n^*} \right)^{1/3} \tag{5.2.3}$$

   E_y represents the electric field in the well and m_n^* represents the effective electron mass.

2. In the electric quantum limit approximation only the lowest energy level is occupied, and the following has been given by Pals [1972]

$$\Delta u \approx \left\{ \frac{45}{14\, a_{min}} \left(1 - \frac{1431\, \alpha}{21504} \right) + \frac{15}{7\, a_{min}} \right\} \left(\frac{q^2 \hbar^2 E_y^2}{m_n^*} \right)^{1/3} \tag{5.2.4}$$

   In this equation the parameter a_{min} is defined by

$$a_{min} = 5^{1/3} \left(1 + \frac{2581}{7168}\, \alpha \right)^{1/3} \tag{5.2.5}$$

   where α is the ratio of the total charge in the inversion layer and the total charge in the depletion layer.

Our simulations have shown that in MOSTs both approximations give nearly identical results.

5.3 The channel potential

The part of the device underneath the gate is considered to be a 4-layered structure (see figure 5.2) consisting of the oxide layer with thickness d_{ox}, the channel region with thickness Δ, the depletion region with thickness W_D-Δ, and the bulk. In the oxide layer a uniform charge density ρ_{ox} is assumed. Furthermore, a box-like doping profile is assumed, having a value N_{As} in the channel region, a value N_{Ad} in the depletion region, and a value N_{Ab} in the bulk. The depletion region is considered to be fully depleted, so

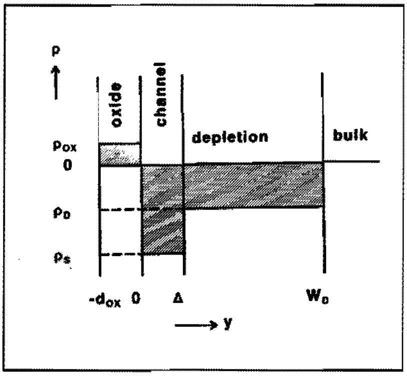

Fig 5.2 Model of the charge density ρ in a MOSFET in the vertical cross section directly underneath the gate.

the charge density ρ_d equals $- q N_{Ad}$. In the simulator the doping profile, oxide properties and geometrical structure of the device have to be given a priori. Unknown parameters are the channel potential $\Psi(0)$, the channel charge density ρ_s, the channel thickness Δ, the thickness of the depletion layer W_D-Δ, and the gate-voltage correction Δu.

The device simulator operates in two modes. First, in the initialization mode, the source-drain voltage and the source-bulk voltage are set to zero, and the channel thickness Δ and the gate-voltage correction Δu are solved in a self-consistent way. This is described in section 5.3.1. In the non-equilibrium mode the external voltages are applied and the device currents and internal distributions in the channel are calculated (see section 5.3.2).

5.3.1 Initialization mode

Directly underneath the gate the device is considered to be a one-dimensional vertical structure consisting of the layers described before. Now Poisson's equation is integrated over the various layers. This yields three equations

$$\frac{\partial \Psi}{\partial y}\Big|_{W_D^-} - \frac{\partial \Psi}{\partial y}\Big|_{\Delta^+} = \frac{q N_{Ad}}{\epsilon_{Si}} (W_D - \Delta) \qquad (5.3.1)$$

$$\frac{\partial \Psi}{\partial y}\Big|_{\Delta^-} - \frac{\partial \Psi}{\partial y}\Big|_{0^+} = - \frac{\rho_s}{\epsilon_{Si}} \Delta \qquad (5.3.2)$$

$$\frac{\partial \Psi}{\partial y}\Big|_{0^-} - \frac{\partial \Psi}{\partial y}\Big|_{-d_{ox}} = - \frac{\rho_{ox}}{\epsilon_{ox}} d_{ox} \qquad (5.3.3)$$

Double integration of the Poisson equation yields three more equations

$$\Psi(W_D^-) - \Psi(\Delta^+) = \frac{1}{2}\frac{q\,N_{Ad}}{\epsilon_{Si}}\,(\,W_D - \Delta\,)^2 + \frac{\partial\,\Psi}{\partial\,y}\Big|_{\Delta^+}(\,W_D - \Delta\,) \tag{5.3.4}$$

$$\Psi(\Delta^-) - \Psi(0^+) = -\frac{1}{2}\frac{\rho_s}{\epsilon_{Si}}\,\Delta^2 + \frac{\partial\,\Psi}{\partial\,y}\Big|_{0^+}\Delta \tag{5.3.5}$$

$$\Psi(0^-) - \Psi(-d_{ox}) = -\frac{1}{2}\frac{\rho_{ox}}{\epsilon_{ox}}\,d_{ox}^2 + \frac{\partial\,\Psi}{\partial\,y}\Big|_{-d_{ox}} d_{ox} \tag{5.3.6}$$

Several boundary conditions have to be satisfied:

1. It is assumed that no charge sheets are present, so between two layers the voltage and the electric displacement are continuous functions

$$\Psi(0^+) = \Psi(0^-) = \Psi(0) \text{ and } \Psi(\Delta^+) = \Psi(\Delta^-) = \Psi(\Delta) \tag{5.3.7}$$

$$\epsilon_{ox}\frac{\partial\,\Psi}{\partial\,y}\Big|_{0^-} = \epsilon_{Si}\frac{\partial\,\Psi}{\partial\,y}\Big|_{0^+} \quad\text{and}\quad \frac{\partial\,\Psi}{\partial\,y}\Big|_{\Delta^-} = \frac{\partial\,\Psi}{\partial\,y}\Big|_{\Delta^+} = \frac{\partial\,\Psi}{\partial\,y}\Big|_{\Delta} \tag{5.3.8}$$

2. The potential at the top of the oxide $\Psi(-d_{ox})$ equals the effective gate potential Ψ_G given by (5.2.2)

$$\Psi(-d_{ox}) = V_G + \frac{1}{2}\frac{U_{gap}}{q} - \frac{\Delta u}{q} \tag{5.3.9}$$

3. Assuming a triangular well, Δu is given by (5.2.4), and the electric field E_y in the well is defined by

$$E_y = -\frac{\partial\,\Psi}{\partial\,y}\Big|_{0^+} \tag{5.3.10}$$

4. The potential at the bottom of the depletion layer $\Psi(W_D^-)$ equals the bulk potential, and this gives, with V_B the externally applied voltage to the substrate contact

$$\Psi(W_D^-) = \Psi_B = V_B - \frac{k_B T_L}{q}\log\frac{N_{Ab}}{n_I} \tag{5.3.11}$$

5. Furthermore, it is assumed that in the bulk no field is present

$$\frac{\partial\,\Psi}{\partial\,y}\Big|_{W_D^-} = 0 \tag{5.3.12}$$

After combining equations (5.3.1) to (5.3.6) with the given boundary conditions, and after elimination of all the derivative terms, the following three equations are left:

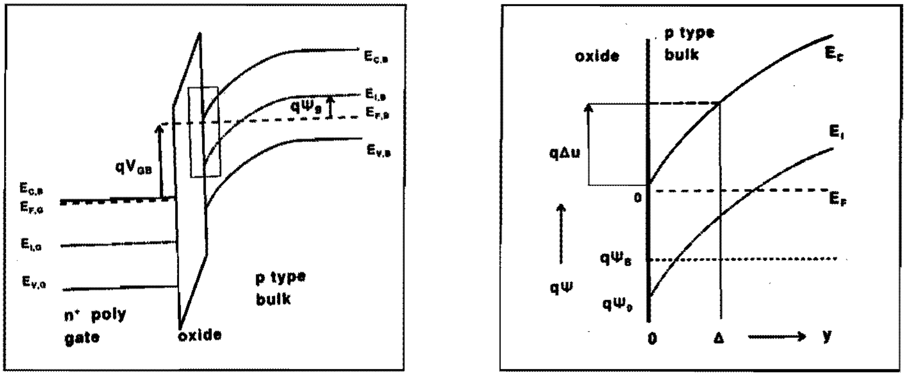

Fig 5.3 Band diagram for a n^+ poly-gate on a p-type substrate (a). The conduction band near the interface is shown in more detail (b), also showing the channel thickness Δ.

$$\Psi(\Delta) = \Psi_B + \frac{1}{2}\frac{q\,N_{Ad}}{\epsilon_{Si}}(\,W_D - \Delta\,)^2 \tag{5.3.13}$$

$$\Psi(0) = \Psi_B + \frac{1}{2\,\epsilon_{Si}}\{\,q\,N_{Ad}(\,W_D^2 - \Delta^2) - \rho_s\Delta^2\,\} \tag{5.3.14}$$

$$\Psi(0) = V_G + \frac{1}{2}\frac{U_{gap}}{q} - \frac{\Delta u}{q} + \frac{d_{ox}}{\epsilon_{ox}}\{\,\frac{1}{2}\rho_{ox}d_{ox} + \rho_s\Delta - q\,N_{Ad}(\,W_D - \Delta)\,\} \tag{5.3.15}$$

These three equations contain six unknown parameters $\Psi(\Delta)$, $\Psi(0)$, W_D, Δ, ρ_s and Δu. To solve this set of equations, three more conditions are needed. The first of these extra conditions is given by equation (5.2.4). The electric field E_y is defined by (5.3.10) and, in combination with (5.3.1), (5.3.2) and the boundary conditions, this yields

$$E_y = -\frac{\rho_s}{\epsilon_{Si}}\Delta + \frac{q\,N_{Ad}}{\epsilon_{Si}}(\,W_D - \Delta\,) \tag{5.3.16}$$

The second of these extra equations is the definition of the charge density in the channel. In the initialization mode no current is flowing and the quasi Fermi-levels for both electrons and holes are flat. Assuming a Boltzmann distribution in the channel, the following equation holds

$$\rho_s = -\,q\,N_{As} + q\,n_I\,\{\,\exp(-\frac{q\,\Psi(0)}{k_BT_L}) - \exp(\frac{q\,\Psi(0)}{k_BT_L})\,\} \tag{5.3.17}$$

Combining equations (5.3.13) to (5.3.17) leaves only one more condition to be defined. The thickness of the inversion layer Δ has to be given. Assuming that the channel thickness Δ is determined by the position of the lowest

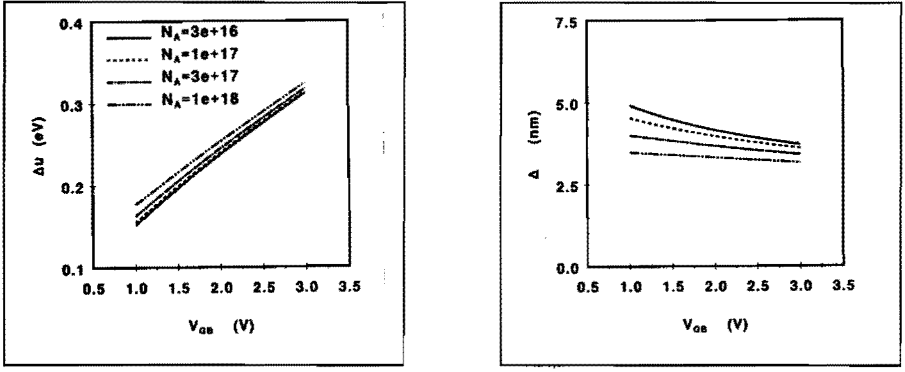

Fig 5.4 (a) Gate-voltage correction and (b) channel thickness as a function of the applied gate voltage. The doping concentration is the parameter.

energy level in the quantum well, yields (see also figure 5.3)

$$\Psi(\Delta) = \Psi(0) - \frac{\Delta u}{q} \tag{5.3.18}$$

Now the set of equations is complete and all parameters can be calculated. This is done using a Newton-Raphson method. The results are used in the second mode of the simulator, the non-equilibrium mode.
In figure 5.4a the calculated gate-voltage correction Δu is given as a function of applied gate voltage, and in figure 5.4b the calculated channel thickness Δ has been plotted with the doping concentration N_A as parameter.

In these data a triangular quantum well has been assumed. However, this assumption is not very important, since results obtained using the electric quantum limit approximation given by equation (5.2.5) are nearly identical to the data presented.

The simulation data in figure 5.4a agree qualitatively very well with the expectations. An increase in the gate voltage, which increases the slope of the quantum well, also increases the energy difference Δu. Furthermore, the higher the value of N_A, the narrower and steeper the quantum well. Consequently Δu increases. In spite of this higher value of Δu, due to the narrowing of the well, the channel thickness Δ decreases with increasing gate voltage or doping concentration, as can be seen in figure 5.4b. Quantitatively there is good agreement with Van Dort and Woerlee [1991], who find a value of $\Delta u = 0.226$ eV if $N_A = 3.3\ 10^{17}$ cm^{-3}. Using two-dimensional drift-diffusion simulations, an estimate of the channel thickness can be made which gives a value of 5 nm. This is also in good agreement with our results.

5.3.2 Non-equilibrium mode

In the second step of the simulator, the non-equilibrium mode, a voltage difference is applied between the source and the drain contacts. The results from the initialization mode are used. Optionally, either both Δ and Δu are assumed to retain the same constant values that have been calculated in the initialization mode (section 5.3.2.1), or the above assumption is dropped, and local changes in Δ and Δu due to the changes in the surface potential are allowed (section 5.3.2.2).

5.3.2.1 Approach assuming constant Δ and Δu

In the non-equilibrium mode, the values for Δ, Δu, obtained from the initialization mode are assumed to be constant, while the values for $\Psi(0)$, W_D and ρ_s are used as initial values. The non-equilibrium problem is solved iteratively. Given the surface potential $\Psi(0)$ at each position in the device, the electron density n at each position in the device can be calculated by solving the hydrodynamic equations as described in section 4.2. If n and the doping concentration N_{As} at each position are known and if the hole density can be neglected everywhere in the channel (which is the case in n-MOS devices), then the charge density can be calculated

$$\rho_s = - q \left(N_{As} + n \right) \tag{5.3.19}$$

Given this charge density the new surface potential is calculated. This process is repeated until the steady state is reached. The results during the iteration also represent the transient behavior of the device as a function of time. During each iteration, the potential at each position in the device has to be solved. Therefore the device is divided into three different regions: the n-type contact regions outside the gate, the channel region underneath the gate that is p-type, and the intermediate region between contact and channel that is underneath the gate but is n-type.

In the channel region the equations (5.3.14) and (5.3.15) are solved in combination with (5.3.19) at each position. Because Δ and Δu are defined this can be done self-consistently yielding values for the surface potential and the depletion width at each position. In the contact regions Poisson's equation is solved directly using a three-point iteration scheme. The vertical dimension of the contacts Δ_{con} has to be given a priori. In the intermediate region it is assumed that no depletion layer is present, so $W_D - \Delta$ equals zero, while Δ and Δu are assumed to equal the values given for the channel region. Then only equation (5.3.15) remains relevant and is used to calculate the surface potential in the intermediate regions.

5.3.2.2 Approach allowing local changes in Δ and Δu

In the non-equilibrium model discussed here, the channel thickness Δ and the gate-voltage correction Δu are no longer assumed to have a constant value during the simulation. At each iteration step the new channel potential is calculated by solving the full set of equations (5.3.13) to (5.3.15) in combination with (5.2.4), (5.3.18) and (5.3.19). However, the channel thickness Δ is not constant everywhere in the device, and an additional problem arises in the electron transport calculations because the problem is not one-dimensional anymore. This problem is solved in the following way: given the current densities J_{ni} and J_{ni+1}, the previous electron density n_{iold}, and the channel thickness Δ_i at the i-th grid point and Δ_{i+1} at the i+1-th grid point, the new electron density is calculated by a discretization of the electron conservation equation (1.4.16), which yields

$$n_i = n_{iold} + \Delta t \, \frac{\left(J_{ni+1} \dfrac{\Delta_{i+1}}{\Delta_i} - J_{ni} \right)}{\frac{1}{2} \left(\Delta x_i + \Delta x_{i+1} \right)} + GR_i \Delta t \qquad (5.3.20)$$

A similar expression is used to calculate the electron energy.

5.4 Experimental data of a 0.17 μm MOSFET technology

A series of scaled submicron devices have been processed at Philips Research Laboratories Eindhoven using a 0.17 μm process technology. The oxide thickness of these devices is 5.9 nm. Electrical measurements for devices with gate lengths of 0.35, 0.6 and 1.0 μm respectively have been made available. Detailed knowledge of the doping profiles of these devices was obtained from SIMS measurements. A full two-dimensional plot of the doping profile of the 0.35 μm device is given in figure 5.5. The devices have a high surface doping underneath the gate, so quantum effects are to be expected, and a correction on the gate potential as described in the previous section is needed. In figure 5.6a the horizontal profile at the Si-SiO$_2$ interface is given; figure 5.6b shows a detail of the vertical profile directly underneath the gate.

In figure 5.7 the drain current is given as a function of the source-gate voltage at very low source-drain voltage (V_{DS}= 0.1 V). By extrapolation this plot yields the low-field mobility μ_{n0} in the channel. A value μ_{n0}= 320 cm^2/Vs has been found. Also the value of the effective channel length L_{eff} can be calculated from these data. It has been found that L_{eff} is 80 nm shorter than the gate length for all devices. The full characteristics of the 1.0 μm and the 0.35 μm device have been plotted in figures 5.8a and

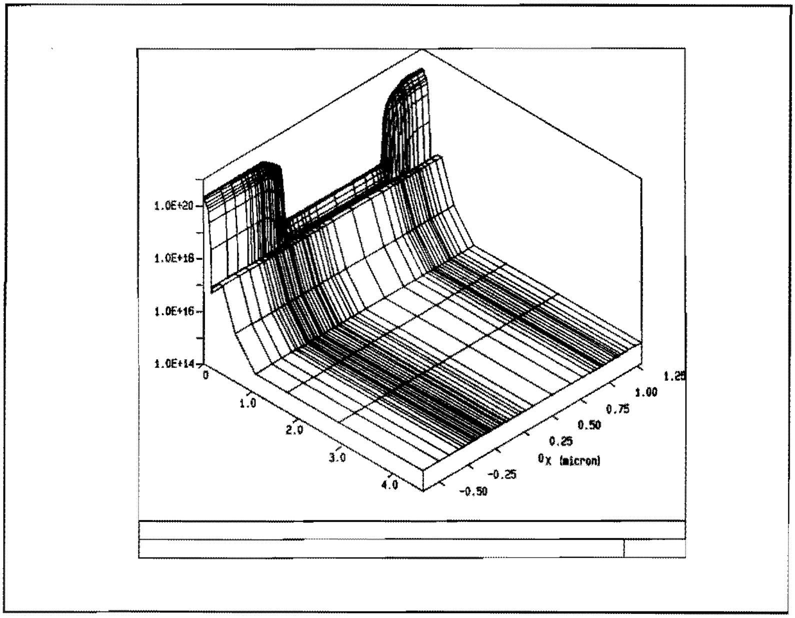

Fig 5.5 2-D doping profile of the 0.35 µm n-MOSFET processed at Philips Research Laboratories.

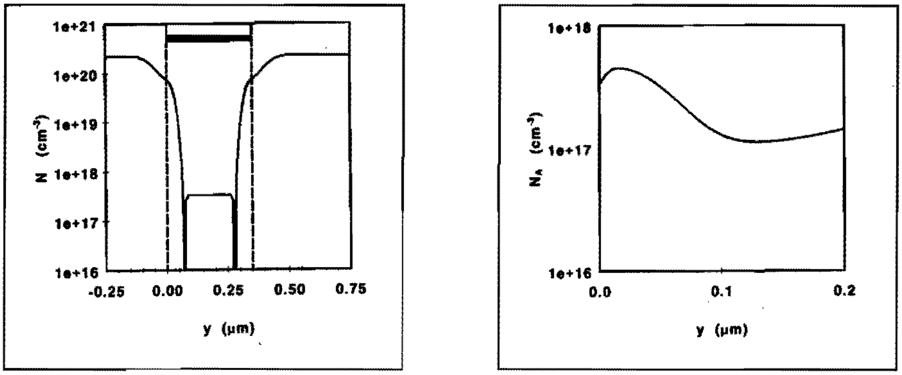

Fig 5.6 Doping profile of the 0.35 µm n-MOSFET: (a) horizontal cross section along the Si-SiO$_2$ interface; (b) vertical cross section directly underneath the gate (detail). The position of the gate contact is shown in (a) by the black bar at the top.

5.8b (solid lines). They are discussed in more detail in the next chapter. In figure 5.10a and 5.10b the substrate current of the same devices has been plotted (solid lines). The dashed lines in figures 5.8 and 5.10 represent simulation results that are discussed in the next section.

5.5 Simulation results

The final device and simulator description is given in section 5.5.1. In the

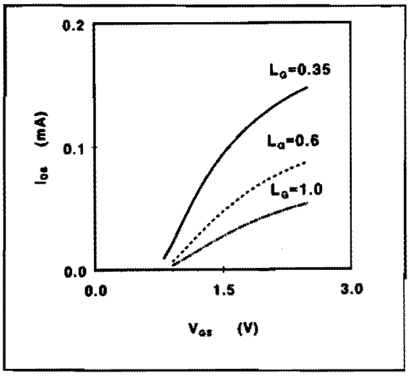

Fig 5.7 Drain current of the n-MOSFETs processed at Philips Research Labs as a function of gate voltage driven at V_{ds}=0.1 V.

simulations the following questions are dealt with. First, how well describes the quasi two-dimensional simulator the devices in the steady state. This is discussed in more detail in section 5.5.2. Secondly, attention is paid to the question of the limits of applicability of the drift-diffusion method, compared to the hydrodynamic method. Important in this validity problem is how well the electron energy and the avalanche generation are calculated. This is dealt with in section 5.5.3. In addition attention has been paid to a third major question, whether in MOS transistors the avalanche effects are adequately described by the impact ionization parameters given by Van Overstraeten for bulk material, or whether a new set of parameters is needed. In section 5.5.4 velocity overshoot effects are discussed. Finally in section 5.5.5 two heat-flow models (Adiabatic and Wiedemann-Franz) are compared.

5.5.1 Device geometry and simulator mode

The hydrodynamic transport model in the MOST simulator is identical to the one described in the previous chapter and is essentially one-dimensional. Some parameters have been set:

1. The mobility in the channel is described by (2.5.3) and (2.5.4), the Selberherr model. The surface scatter parameters have been given such a value that the low-field mobility is in accordance with experimental data.
2. The energy relaxation time is described by (2.5.5) and the field-energy relation by (2.5.10). The results presented in chapter 2 have been preferred to the data given by Fischetti [1991] for reasons stated in that chapter.
3. The impact ionization rate is described by the Chynoweth formula (2.5.6).
4. In the simulations local changes in Δ and Δu are allowed (see section 5.3.2.2).
5. The triangular well approximation (5.2.4) has been used.
6. The adiabatic heat-flow model is used in most of the simulations, because that gives the best fit to Monte Carlo results as was shown in chapter 4.

	Philips device	*MIT device*	*unit*
d_{ox}	5.5	7.5	nm
Δ_{con}	80	10	nm
N_{Dc}	$2.2\ 10^{20}$	$3.0\ 10^{20}$	cm^{-3}
N_{As}	$3.3\ 10^{17}$	$1.5\ 10^{16}$	cm^{-3}
N_{Ad}	$4.0\ 10^{17}$	$5.7\ 10^{16}$	cm^{-3}
N_{Ab}	$1.0\ 10^{17}$	$1.0\ 10^{17}$	cm^{-3}
μ_{n0}	320	450	cm^2/Vs

Table 5.1 Simulation parameters of various devices.

The simulator is applied to the devices described in the previous section. In table 5.1 the corresponding process parameters are given. Additionally, in this table similar data is given of the devices from MIT, described by Shahidi et al. [1988] is given.

Regarding the Philips process, the following assumptions have been made:

1. The surface scatter parameters have been given such a value that the low-field mobility equals 320 cm^2/Vs in accordance with experimental data.
2. The exact doping profile is approximated by a box profile (see figure 5.6b). The doping concentration in the channel is given a value $N_{As} = 3.3\ 10^{17}\ cm^{-3}$. Considering the fact that in the devices discussed the depletion region width is approximately 50 nm, the assumption that the depletion concentration N_{Ad} has a constant value of $N_{Ad} = 4.0\ 10^{17}\ cm^{-3}$ seems appropriate. The active device is concentrated in the first 100 nm underneath the oxide, and therefore the bulk layer also is assumed to begin in this region. A reasonable value for the bulk concentration N_{Ab} appears to be $N_{Ab} = 10^{17}\ cm^{-3}$. Control simulations show that small variations in N_{Ad} and N_{Ab} only have minor effects on both the drain and the avalanche currents.
3. The contact dope N_{Dc} equals $2.2\ 10^{20}\ cm^{-3}$ and the contact thickness Δ_{con} is taken to be 80 nm. This gives a good agreement in the active charge density, if compared to the results obtained by MINIMOS.
4. The distance between grid lines has a finest value of 3 nm at the channel-drain junction. Control simulations show that if the latter grid resolution is halved, a maximum deviation of 1 % in the drain current is found. In the avalanche current the deviation is larger but still small.

In a similar way the device parameters for the MIT devices have been obtained, although the doping profile of these devices was less accurately known. Corresponding data are also presented in table 5.1.

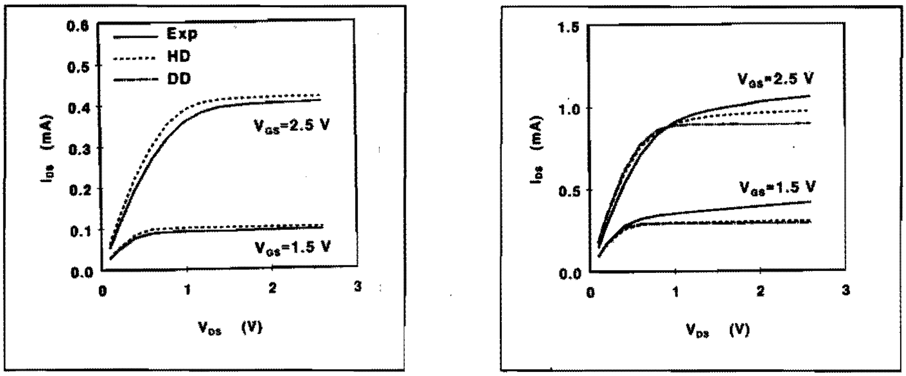

Fig 5.8 Drain current as a function of applied V_{ds}. The gate lengths are 1.00 μm (a) and 0.35 μm (b) resp.; the gate voltage is the parameter; the solid lines represent the experimental results, the dashed lines the hydrodynamic simulation results, and the dash-dot lines in (b) the drift-diffusion results.

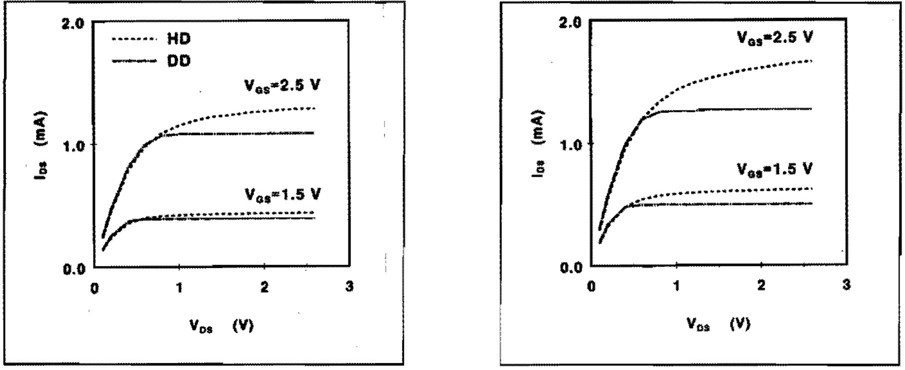

Fig 5.9 These plots are similar to 5.8 but now the device currents have been plotted for a 0.25 μm device (a) and a 0.20 μm device (b). The simple dashed lines represent hydrodynamic simulation results, and the dash-dot lines the drift-diffusion results.

5.5.2 Device characteristics

The device characteristics of the devices described in section 5.5 are calculated and compared to experimental data. The drain current of the devices with gate lengths of 0.35 μm and 1.0 μm (which means $L_{eff} \approx$ 0.27 and 0.92 μm, respectively) have been plotted in figures 5.8a and 5.8b (simple dashed lines) together with the experimental data (solid lines).

The simulations describe both the 1.0 μm and the 0.35 μm device well. The resulting discrepancies in the saturation region are a result of two-dimensional effects that have been neglected in our simulations.

Additionally, drift-diffusion simulations using the same process parameters have been carried out and have been plotted in the same figures (dash-dot lines). For the 1.0 μm device, the hydrodynamic and the drift-diffusion simulation results are indistinguishable. However, for the 0.35 μm device

significant differences occur: in the drift-diffusion model the current saturates for high drain voltages, while in the hydrodynamic model there is an increase of the current due to energy effects. This effect is stronger for increasing gate voltages.
Similar effects have been found for smaller devices, using the same process parameters. In figures 5.9a and 5.9b the device currents for a 0.25 μm and a 0.20 μm device have been plotted (which means an effective channel length of 0.17 μm and 0.12 μm respectively). Again, the simple dashed lines represent the hydrodynamic results, the dash-dot lines represent the drift-diffusion results (no experimental data of these devices was available).

These results show that for low gate voltages the drift-diffusion method gives a good first-order approximation of the 0.35 μm device currents, notwithstanding the fact that the electrons have high energies in a considerable part of the channel.
For increasing gate voltages and for decreasing channel lengths the discrepancy between the hydrodynamic and the drift-diffusion approach increases. However, the difference in the drain current calculated in the drift-diffusion mode, compared to the results obtained by the hydrodynamic mode never exceeds a factor of 25% for devices with effective channel lengths down to 100 nm.

Although from a fundamental point of view the drift-diffusion method is not correct for deep-submicron devices, from a practical point of view the use of the drift-diffusion method, combined with a post-routine that calculates energy and avalanche generation afterwards, is very attractive. The use of full two-dimensional hydrodynamic simulators still causes great problems: the simulators available are not very robust and easily give rise to numerical instabilities. Furthermore, the computing time needed is, roughly speaking, one order of magnitude larger than the time needed for a drift-diffusion solution. For this reason it is interesting to see how well the results of a full hydrodynamic simulator agree with the results of a drift-diffusion simulator extended with a post-processing routine for the energy distribution and avalanche generation. This is dealt with in the next section.

5.5.3 Energy distribution and avalanche generation

The avalanche currents for the 0.35 and 1.0 μm devices have been plotted in figures 5.10a and 5.10b. The data are obtained both from experiment (solid lines) and from simulations (dashed lines). These simulations over-estimate the values of the avalanche current by one order of magnitude.

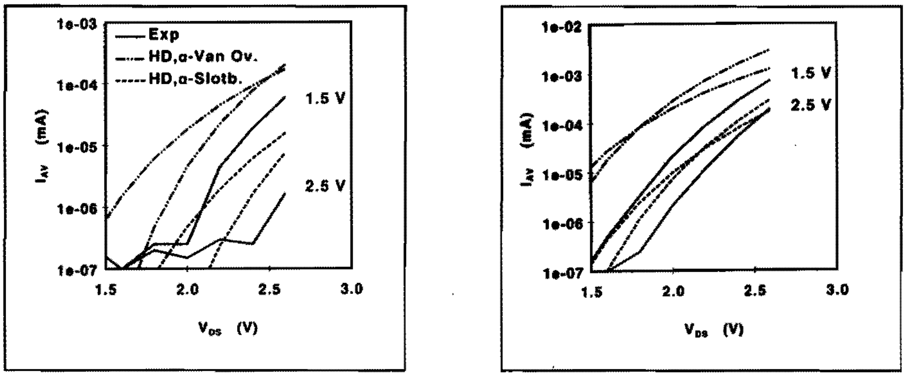

Fig 5.10 These plots are similar to 5.8 but now the substrate currents have been plotted. Using the Van Overstraeten parameters in the simulations yields the set of simple dashed curves; using the Slotboom parameters yields the dash-dot-dot curves.

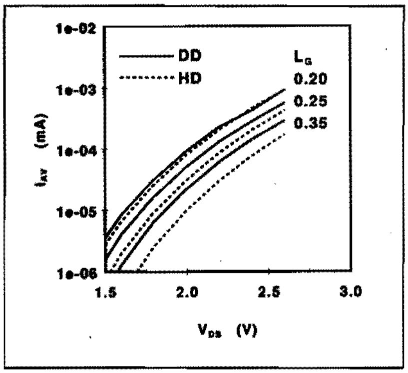

Fig 5.11 Avalanche currents of devices with different gate lengths (0.35, 0.25 and 0.20 μm). Plotted are the full hydrodynamic results (solid lines) and the results obtained by the drift-diffusion model combined with a post-routine to calculate energy and avalanche generation (dashed lines).

Essentially, the Van Overstraeten parameters were obtained from measurements on bulk material; however a MOS device operates at the Si-SiO_2 interface. According to Slotboom et al. [1987] this has important consequences for the impact parameters. Based on experimental data, they have presented for MOS devices a new set of values for the impact parameters: $\alpha_n^0 = 2.45\ 10^6\ cm^{-1}$ and $E_{crit} = 1.92\ 10^6$ V/cm. Simulations, using these parameters also have been carried out. The results have been plotted in figure 5.10 as well (dash-dot-dot curves). A decrease of the calculated avalanche current is found and the calculated values are more in agreement with the experimental data. However, the difference between calculations and experiments is still substantial, not only quantitatively, but also qualitatively: at high drain voltages the simulations give an increasing avalanche current for increasing gate voltage. This is not found in the experimental data.

In figure 5.11 the avalanche currents have been calculated for different gate lengths (in these devices the difference between gate and effective channel

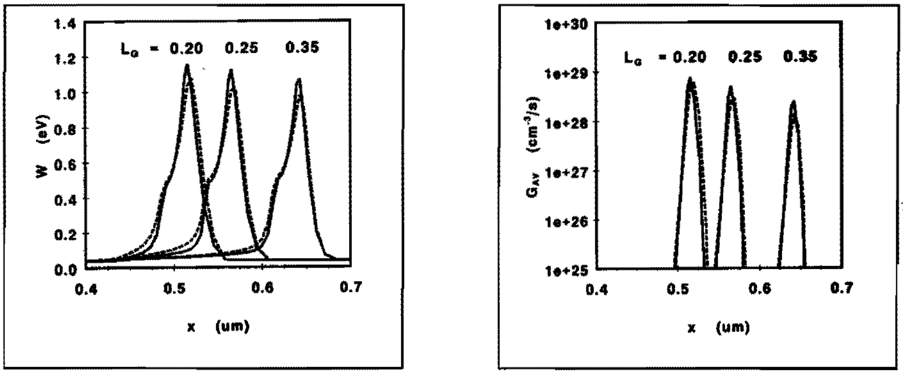

Fig 5.12 Energy distribution (a) and avalanche generation (b) in MOS devices with different gate lengths, driven at V_{gs}=1.5 V and V_{ds}= 2.6 V. The gate contact is defined from x=0.35 to x=0.55, x=0.60 and x=0.70 μm respectively. The solid line represents extended drift-diffusion results, the dashed line full hydrodynamic results.

length is approximately 80 nm). In these calculations at V_{GS}= 1.5 V the Slotboom avalanche parameters have been used.

The results obtained from full hydrodynamic calculations (dahed lines) and results from the drift-diffusion mode (solid lines) are compared. Considering the fact that, due to the exponential relation between avalanche and energy, the avalanche generation is very sensitive to slight variations in the energy distribution, the discrepancy observed in figure 5.11 is not very large. Even for devices with an effective channel length of only 120 nm the discrepancy between full hydrodynamic and the drift-diffusion method is at the most a factor three. This supports the work of Slotboom et al. [1991], who have shown that the avalanche currents in submicron MOS transistors calculated by a two-dimensional drift-diffusion simulator extended with an energy post-processing routine, give good agreement with experimental data.

Finally, in figures 5.12a and 5.12b the internal energy distribution and avalanche generation have been plotted for devices with different gate lengths (0.35, 0.25 and 0.20 μm). Again, the full hydrodynamic results (dashed lines) are compared to post-routine results (solid lines). The agreement is surprisingly well.

5.5.4 Velocity overshoot

According to transport theory, velocity overshoot occurs in small MOS devices at the drain side of the channel due to large local gradients in the electric field. In analogy with figures 5.12a and 5.12b the electron velocity for MOS devices with various gate lengths has been plotted in figure 5.13.

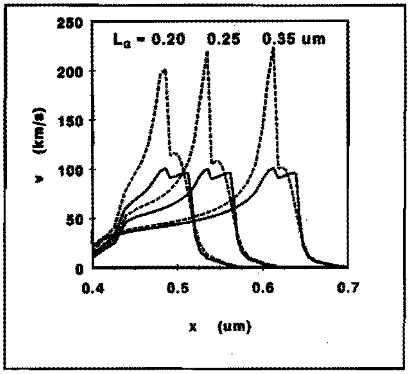

Fig 5.13 Plot similar to 5.12, but here the electron velocity is shown.

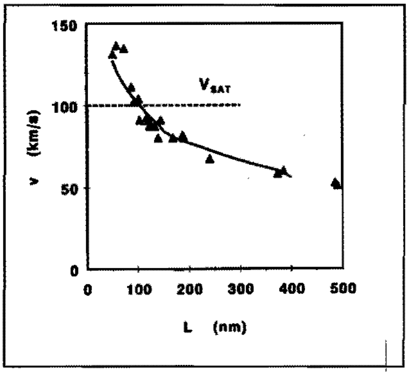

Fig 5.14 Calculated averaged channel velocity as a function of channel length in MOS devices processed at MIT. The triangles represent experimental data by Shahidi et al. [1988].

As can be seen, in the hydrodynamic mode velocity overshoot occurs (up to more than two times the saturation velocity), while in the drift-diffusion mode the electron velocity never exceeds the saturation velocity. Surprisingly, the maximum velocity overshoot depends only slightly on the channel length and does not cause the increase of the average velocity overshoot with decreasing gate length. By Shahidi et al. [1988] experimental data was obtained at MIT on the average channel velocity in MOS devices and their results confirmed that velocity overshoot indeed does occur in submicron MOS devices. These data are represented in figure 5.14 by triangular dots. The solid line represents the simulation data obtained by the simulator described in this chapter. Good agreement has been found.

5.5.5 Other simulation models

The choice of adiabatic heat-flow model and value of energy relaxation time was based on Monte Carlo simulation results. In combination with these, the use of the Slotboom parameters in stead of Van Overstraeten parameters does reduce the avalanche currents and improves the fit between simulation and experimental data. This was shown in the previous sections. However, it has to be noted that there are more ways of obtaining such an improvement:

1. Taking a higher value of energy relaxation time τ_{wn} also reduces the avalanche currents. A value $\tau_{wn} = 0.9$ ps, (found by Fischetti et al.) yields avalanche currents very similar to the results presented in figures 5.10.
2. The use of the Wiedemann-Franz heat-flow model results in a different temperature distribution. However, in combination with Van Overstraeten avalanche parameters, this model yields nearly the same avalanche currents as Slotboom parameters in combination with the adiabatic model. Only at high drain voltages the Wiedemann-Franz model yields notable lower results (factor 2 at $V_{DS} = 2.6$ V).

A final choice between these different models based on the experimental data available can not be made.

5.6 Summary and conclusions

A quasi two-dimensional hydrodynamic MOS device simulator has been developed. Device characteristics obtained by this simulator are in fair agreement with experimental data for devices with channel lengths down to 0.25 μm. The agreement between the calculated and measured avalanche currents is less good. Several possibilities to optimize the fit between experimental and simulation data have been found:

1. Adiabatic heat-flow model, combined with Slotboom avalanche parameters and a value $\tau_{wn} = 0.4$ ps.
2. Adiabatic heat-flow model, combined with Van Overstraeten avalanche parameters and a value $\tau_{wn} = 0.9$ ps.
3. Wiedemann-Franz heat-flow model, combined with Van Overstraeten avalanche parameters and a value $\tau_{wn} = 0.4$ ps.

A final choice between these different models based on the experimental data available can not be made.
Simulation results show that the use of a drift-diffusion simulator extended with a post-processing routine to calculate the electron energy and avalanche generation yields adequate results, even for deep-submicron devices, with the restriction that the avalanche current must be substantially lower than the device current. In the next chapter the same problem is studied again, but now for full two-dimensional simulators.

The maximum electron velocity exceeds the saturation velocity in submicron devices, and has a value up to approximately 2 10^7 cm/s. This value does only slightly depend on the channel length. The average channel velocity increases with decreasing channel length, in agreement with experimental findings by Shahidi et al. [1988].

Chapter 6

FULL TWO-DIMENSIONAL HYDRODYNAMIC MOST SIMULATIONS

6.1 Introduction

In this chapter various simulators and methods to calculate the electron temperature are compared. Table 6.1 gives an overview. In this introductory section this table is explained.

In the previous chapter, section 5.5, the quasi two-dimensional device simulations have been dealt with. The heat flow $\vec{Q}$ has been modeled either by the Wiedemann-Franz model (W-F) or the adiabatic heat-flow model $\vec{Q} = \vec{0}$. In all simulations an expression for the energy relaxation time (TW) has been used. Furthermore, both drift-diffusion results extended with an energy post-processing routine (DD+), and full hydrodynamic simulations (HD) have been carried out. This information is found systematically in table 6.1.

In this chapter mainly full two-dimensional simulations are discussed. For this purpose use has been made of the well known MOS simulator MINIMOS, developed by Selberherr et al. [1980], and of the general two-dimensional simulator TRENDY, developed at the University of Twente by Van Schie [1990] and Wolbert [1991]. Both simulators are discussed in more detail in sections 6.2.1 (MINIMOS) and 6.3.1 (TRENDY), respectively. The simulators have been applied to the devices described in the previous chapter. In section 6.2.2

Simulator	$\vec{Q}$	Relax	DD+	HD	Section
Quasi 2-D	W-F	TW	+	+	5.5.5
	Adiab	TW	+	+	5.5.2
MINIMOS	HOT		-	+	6.2.2
	Adiab	Analyt	+	-	6.4
TRENDY	W-F	TW	+	+	6.3.2
		LW	+	-	6.3.3
	Adiab	TW	*	*	
		LW	+	-	6.3.3
		Analyt	+	-	6.4

Table 6.1. Overview of all simulators and methods used to calculate the electron temperature. (+)=done; (-)=not done; (*)=not converging. This table is explained in detail in section 6.1.

simulation results obtained by MINIMOS, using a simplified set of hydrodynamic equations (HOT mode) are discussed. The results are not very satisfactory.

In section 6.3.2 results obtained by TRENDY, using the Wiedemann-Franz heat-flow model are discussed. Like all simulators that take into account the full hydrodynamic equations, TRENDY has two serious drawbacks:

1. The program is not very robust and numerical instabilities easily arise.
2. Hydrodynamic simulations require much computing time, typically one order of magnitude more than for the drift-diffusion mode.

To overcome these problems the use of extended drift-diffusion simulations is suggested, in analogy to section 5.5.3. In section 6.3.2 it is studied how appropriate this approach is, if applied to full hydrodynamic two-dimensional simulation results obtained by TRENDY, assuming the Wiedemann-Franz heat-flow model.

The use of the adiabatic heat-flow model did not yield converging results, neither in the hydrodynamic, nor in the extended drift-diffusion mode. However, by defining a relaxation length (LW) approach in stead of a relaxation time (TW), convergence can be achieved. This is discussed in section 6.3.3. Good agreement with experimental results has been obtained.

In section 6.4 an analytical solution to the energy post-processing problem, formulated by Slotboom et al. [1991], is given. This method is used in combination with MINIMOS and TRENDY results.

In section 6.5 the results obtained by the two-dimensional simulator are compared to the results obtained by the quasi two-dimensional MOS simulator, that was discussed in the previous chapter. Finally, in section 6.6 all conclusions are summarized.

6.2 Results with the simulator MINIMOS

The MINIMOS program developed by Selberherr et al. [1980] in particular is dedicated to the simulation of MOS transistors. Essentially MINIMOS is a drift-diffusion simulator. However, a hot-electron model is included, in which the electron temperature is calculated by a simple series expansion of the solution of the energy conservation law. This is described in section 6.2.1. In section 6.2.2 simulation results using the MINIMOS hot-electron model are presented.

6.2.1 The MINIMOS hot-electron model

In the calculation of the electron temperature in MINIMOS several simplifications have been made. This is described by Selberherr [1989]. Starting from the full energy conservation equation (1.4.21), these simplifications are:

1. The enrgy convection is neglected ($W_n = \frac{3}{2} k_B T_n$).
2. The energy flow is defined by $\vec{S}_n = k_B T_n \vec{J}_n / q$.
3. The gradient in the electron velocity is assumed to be small.
4. The generation/recombination term is neglected.

This yields in steady state

$$\vec{J}_n \cdot \left(\vec{E} + \frac{k_B}{q\,n} \nabla (n\,T_n) \right) = \frac{3}{2} n\,k_B \frac{T_n - T_L}{\tau_{wn}} \tag{6.2.1}$$

Applying these assumptions, also the momentum equation in steady state can be rewritten as

$$\vec{J}_n = q\,\mu_n\,n \left(\vec{E} + \frac{k_B}{q\,n} \nabla (n\,T_n) \right) \tag{6.2.2}$$

Combining (6.2.1) and (6.2.2) yields

$$\frac{k_B T_n}{q} = \frac{k_B T_L}{q} + \frac{2}{3} \tau_{wn}\,\mu_n\,E_n^2 \tag{6.2.3}$$

Here the driving force E_n is defined by

$$E_n = \left| \vec{E} + \frac{k_B}{q\,n} \nabla (n\,T_n) \right| \tag{6.2.4}$$

In analogy with equations (2.5.3) and (2.5.4) the mobility is defined by

$$\mu_n = \frac{2\,\mu_n^{\mathrm{LIS}}}{1 + \left(1 + \left(\frac{2\,\mu_n^{\mathrm{LIS}} E_n}{v_{\mathrm{sat}}} \right)^2 \right)^{1/2}} \tag{6.2.5}$$

In addition it can be shown that

$$\mu_n\,E_n^2 = v_{\mathrm{sat}}^2 \left(\frac{1}{\mu_n} - \frac{1}{\mu_n^{\mathrm{LIS}}} \right) \tag{6.2.6}$$

Furthermore, a definition for the energy relaxation time τ_{wn} is used

$$\tau_{\mathrm{wn}} = \gamma \, \frac{3}{2} \, \frac{k_B T_L}{q} \, \frac{\mu_n^{\mathrm{LIS}}}{v_{\mathrm{sat}}^2} \tag{6.2.7}$$

The value of the fudge factor γ is in the range [0,1].

The substitution of (6.2.6) and (6.2.7) in (6.2.3) finally yields the expression to calculate the electron temperature T_n used in MINIMOS

$$T_n = T_L \left(1 - \gamma + \gamma \, \frac{\mu_n^{\mathrm{LIS}}}{\mu_n} \right) \tag{6.2.8}$$

Several notes have to be made.

1. In both MINIMOS4.0 and MINIMOS5.1 the above equations were not implemented correctly. However this could be corrected quite easily.
2. The definition of $\vec{S}$ seems to be rather arbitrary, and is likely only taken to simplify the equations. No physical theory has been given.
3. The parameter τ_{wn} given by (6.2.7) has no energy (or field) dependence. However it is related to the doping concentration by the mobility. Using the default value $\gamma = 0.8$ gives in intrinsic material $\tau_{\mathrm{wn}} = 0.44$ ps, which is in fair agreement with the values found by our Monte Carlo simulations and in the literature. However, even at moderate doping concentrations, this value drops fast. At $N_I = 3.3 \; 10^{17}$ cm^{-3} a value $\tau_{\mathrm{wn}} = 0.15$ ps has been found, which is very low.

6.2.2 Results obtained using the hot-electron model

Simulation results were obtained for the devices described in section 5.5. The device characteristics for the 1.0 μm and the 0.35 μm devices have been plotted, respectively in figure 6.1a and 6.1b. The simulator was run either in the drift-diffusion mode (small dashes) or the hot-electron mode (large dashes). The solid lines give the experimental values.

Due to the limited knowledge of the exact doping profile, external resistors of 80 Ohm, both to the source and the drain contact have to be added to obtain a good fit in the low-field region at $V_{GS} = 1.5$ V. Also a gate-voltage correction $\Delta u = -0.27$ V has been added to account for the quantum well. In MINIMOS no automatic voltage correction was implemented, and therefore it has not been taken into account that Δu is in fact an increasing function of V_{GS} (see section 5.3). Therefore at high gate voltages the current calculated are expected to be too high. This also has been found. Figure 6.1a shows a very good agreement between simulated drift-diffusion results and experimental data for the 1.0 μm device, where hot-electron effects are not important. At

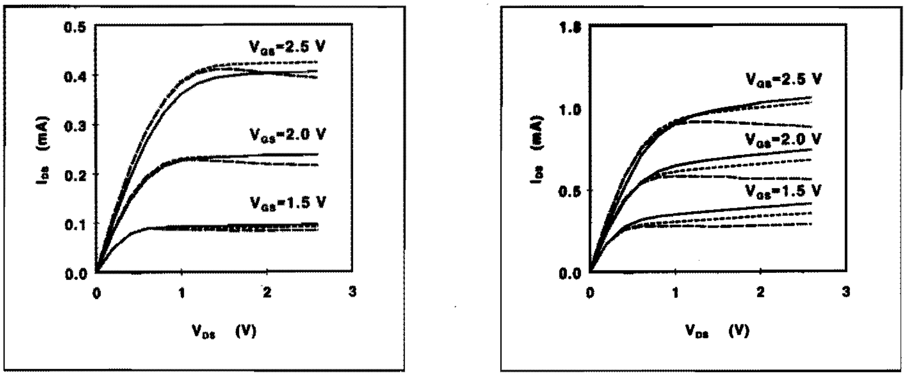

Fig 6.1 Primary device current for a 1.0 μm (a) and a 0.35 μm (b) device. Experimental data (solid lines) and MINIMOS5.1 simulation data are shown: the small dashes represent the 2-D mode, the large dashes the HOT mode.

V_{GS} = 2.5 V the calculated values of I_{DS} indeed are higher than the experimentally obtained values. For the 0.35 μm device (figure 6.1b) the calculated drift-diffusion currents are too low in the saturation region. The reason for this is that in such a small device the high electron temperature gradients cause an extra current component. This is not accounted for in the drift-diffusion model.

However, in MINIMOS, the hot-electron (HOT) mode yields unexpected results: strong negative differential resistances occur in the saturation region, even in the 1.0 μm device, where hardly any effect is expected. This effect is not observed in the experiments. Considering the theory given in 6.2.1, these negative slopes can be explained qualitatively: due to the increasing T_n at increasing applied voltages, the factor E_n increases. This results in a decreasing mobility and a decreasing source-drain current.

In figure 6.2 the electron temperature in the device, as calculated by MINIMOS, at a bias of V_{GS}= 1.5 V and V_{DS}= 2.6 V, is given. Qualitatively this results fulfills the expectations, showing a high peak of several thousands degrees Kelvin at the drain side of the channel. The peak is near the interface, extending 0.1 μm into the device.
The irregularities in e.g. the drain depletion layer, are caused by the fact that the resolution of the automatically generated mesh is not very good.

The avalanche generation rate GR_{AV} allows a more quantitative check on the temperature distribution. GR_{AV} is calculated as a function of the local electron temperature. The Chynoweth formula (2.5.6) has been used, in combination with an adequate T_n-E_{eff} relation. This relation has been given by (2.5.8), which yields, if the energy convection is neglected

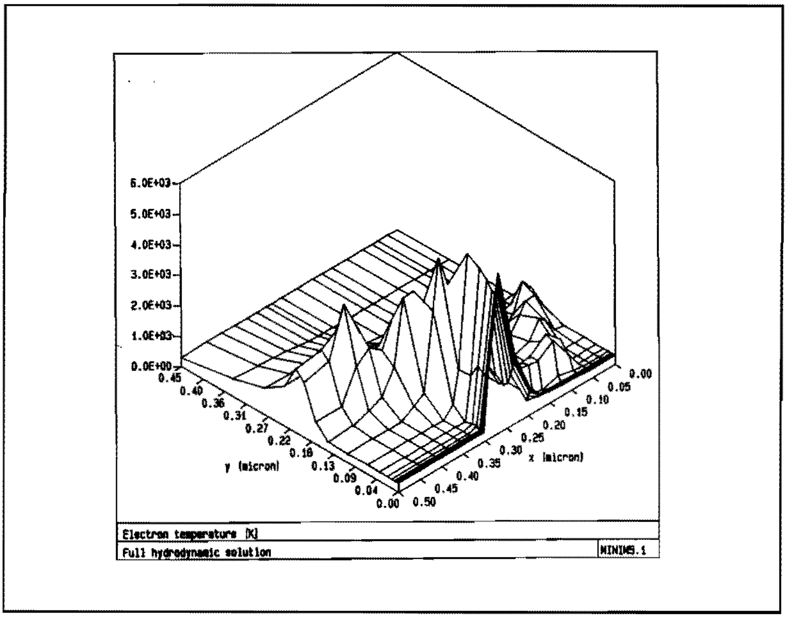

Fig 6.2 Electron temperature in a 0.35 μm device, biased at V_{gs}=1.5 V and V_{ds}=2.6 V, calculated by the MINIMOS5.1 HOT mode.

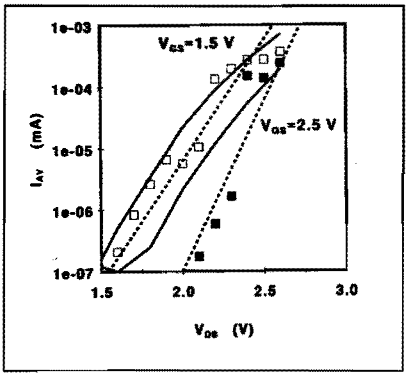

Fig 6.3 Avalanche current in a 0.35 μm device, biased at V_{gs}=1.5 and 2.5 V respectively. The solid lines represent experimental data and the squares MINIMOS5.1 simulation data. The dashed lines represent least squares fit of these simulated data.

$$E_{eff} = \frac{3}{2} \frac{k_B (T_n - T_0)}{q \; v_{sat} \; \tau_{wn}} \tag{6.2.9}$$

An acceptable parameter fit between calculated and experimental data for the 0.35 μm device has been found, using the Van Overstraeten avalanche parameters combined with a value τ_{wn} = 0.35 ps in equation 6.2.9. This is shown in figure 6.3. The data do not yield smooth curves due to the fact that the resolution of the mesh is not sufficient.

Further remarks to be made are the following:

1. More simulations show GR_{AV} to be very sensitive to the value τ_{wn}. A change

of 20% in τ_{wn} leads to an order of magnitude change in GR_{AV}.

2. The value of τ_{wn} used within the MINIMOS program is not consistent with the value of τ_{wn} needed to obtain a best fit with the experimental data.
3. MINIMOS calculations using a value $\gamma = 2.2$ (which yields a $\tau_{wn} = 0.4$ ps) were not converging. In addition, this higher γ also would yield a higher temperature, and thus a further decrease in the device currents.
4. It can easily be shown, that the use of the parameter values given by Slotboom et al. [1987], in combination with $\tau_{wn} = 0.25$ ps yields nearly identical results as presented in plot 6.3.

From these results it can be concluded that the HOT mode in MINIMOS does not work satisfactorily.

6.3 Results with the device simulator TRENDY

TRENDY has been developed and described by Van Schie [1990] and Wolbert [1991] at the University of Twente. The program allows hydrodynamic calculations. In the original version of TRENDY the transport parameters were defined as a function of local electric field. This is not correct and yields too high values for the avalanche currents. Therefore TRENDY has been modified. In this modified version the parameters are functions of local electron energy by using the principle of effective electric field as described before. The W_n-E relation used is given by equation (6.2.9).

A further modification concerning the heat flow was attempted. In the original version of TRENDY the Wiedemann-Franz heat-flow model was implemented.

An extra procedure has been added that allows the use of the adiabatic heat-flow model which is the better model as is shown in chapter 4 of this thesis. However we did not succeed in finding an adequate discretization scheme. None of the two-dimensional simulation runs, using the adiabatic heat-flow model, did converge.

6.3.1 The full hydrodynamic solution

With TRENDY the simulation of the previously described 0.35 μm device driven at the same bias has been attempted. Several serious problems became apparent:

1. The use of the hydrodynamic mode of TRENDY is extremely time-consuming.
2. There is no automatic grid generator available, and it takes an intelligent trial and error method to find an adequate grid.
3. The simulator has serious convergence problems that depend heavily on the mesh. Even drift-diffusion calculations sometimes do not converge.

0.35 μm device; Vgs=1.5 V; Vds =2.6 V							
I	Dim	Exper	$\vec{Q}$	Relax	Avalan	DD+	HD
I_{DS}	mA	0.413	W-F	TW		0.348	0.371
I_{AV}	nA	710	W-F	TW	Slotb	0.755	7.73
					Overst	40.6	170
				LW	Slotb	0.891	*
					Overst	45.4	*
			Adiab	LW	Slotb	267	*
					Overst	2340	*
				Analyt	Slotb	249	*
					Overst	2230	*

Table 6.2 Terminal current values obtained from full hydrodynamic and from various post-processing simulations.

4. Several bugs, e.g. the contact and interface definitions, in TRENDY had to be corrected.

Nevertheless, we did succeed in obtaining results at the bias V_{GS}= 1.5 V and V_{DS}= 2.6 V. Using the Wiedemann-Franz heat-flow model and the Van Overstraeten avalanche parameters, the simulation yields I_{DS} = 0.371 mA and I_{AV} = 170 nA. The value of I_{DS} is 10% lower than the experimental data. This may be due to the mobility model used, but it also suggests that the calculated electron temperatures are too low. This is also confirmed by the calculated avalanche currents, that are much lower than the experimental data (see table 6.2).

The value of I_{AV} calculated at the given bias is much lower than the value found experimentally (I_{AV} = 710 nA). If the avalanche parameters have the values given by Slotboom et al. [1987], even a much lower value I_{AV}= 7.73 nA is found. This is in agreement with section 5.5.4.

6.3.2 Comparison of the full hydrodynamic and the post-processing method

In this section the differences found between the results from the full hydrodynamic simulation with Wiedemann-Franz model, and from the drift-diffusion simulation combined with a post-processing routine that calculates the electron temperature and avalanche generation rate, are compared. The goal of this comparison is to be able to give a more founded opinion whether the use of post-processing methods is justified.

Table 6.2 shows that the hydrodynamic method gives a 10% higher source-drain current than the drift-diffusion method. This is due to the extra thermal currents. More striking is the difference in calculated avalanche currents. To analyze this phenomenon in more detail, the differences in calculated values of electron temperature, electron density, and avalanche generation rate is discussed in more detail.

First the electron temperature distribution is considered; compare figures 6.4a and 6.4b. At first sight, there is a large discrepancy between the two plots. An important difference is the high electron temperature ridge in the depletion layer in plot 6.4a. However this temperature ridge is situated in a region of low electron density and therefore has no relevance for the device characteristics. The same applies for all side peaks and the cool spot at mid-channel in figure 6.4b. Therefore, if the temperature ridge in the depletion layer and all side peaks are ignored, a great similarity between the two plots becomes obvious: the position along the channel of the temperature peak is identical, at 0.3 μm from the beginning of the channel. Both temperature peaks have a value of 4500 K at the Si-SiO_2 interface and this value increases with increasing depth, up to a final maximum of 5500 K in the extended drift-diffusion model and 6000 K in the full hydrodynamic model. This maximum lies at a depth of 0.08 μm, which is also the junction depth of the drain.

Next, the electron density distribution plotted in 6.5, is considered. The electron density distribution at the interface is in both plots very similar, with a slightly deeper dip at the drain side in the drift-diffusion model. However an obvious disagreement is found in the total current path; in the drift-diffusion situation, the current path is close to the Si-SiO_2 interface, and the channel does not extend beyond 25 nm from the interface. On the other hand, in the full hydrodynamic solution the current path is much more diffuse, and the channel extends more or less up to 80 nm. This is due to the fact that the high temperature peak gives rise to a strong diffusion term away from the hot spot, which forces electrons into the bulk of the semiconductor.
Furthermore, the high temperature peak also causes a strong velocity overshoot (up to a factor 3 of the saturation velocity), so the total current generated in the hydrodynamic situation is slightly higher.

Finally the avalanche generation rate GR_{AV} has been considered (see figure 6.6). GR_{AV} is a function of the electron temperature and of the electron current density. As expected from the above, the avalanche current is much more concentrated at the interface if the drift-diffusion model is used.

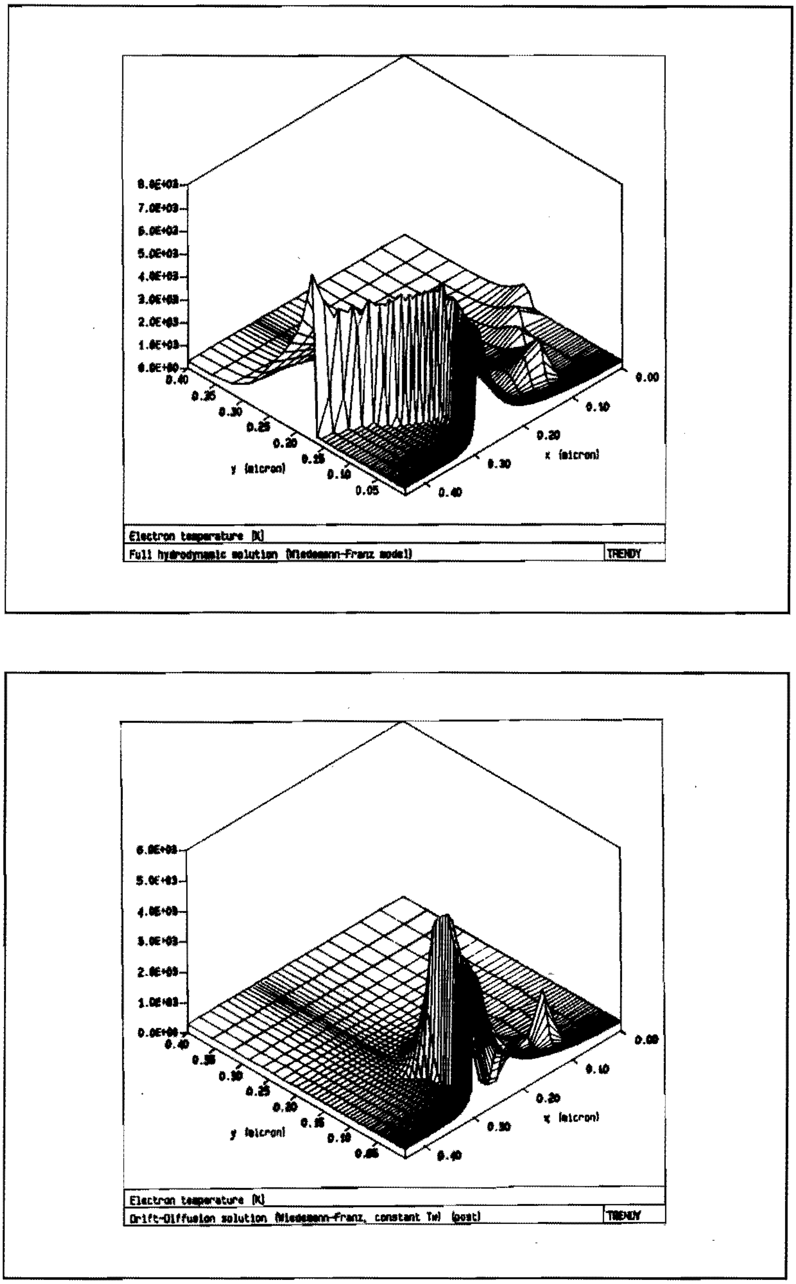

Fig 6.4 Electron temperature distribution in a 0.35 μm device, biased at V_{gs}=1.5 V and V_{ds}=2.6 V.
a. TRENDY full hydrodynamic results (Wiedemann-Franz heat flow model).
b. TRENDY drift-diffusion and energy post-processing.

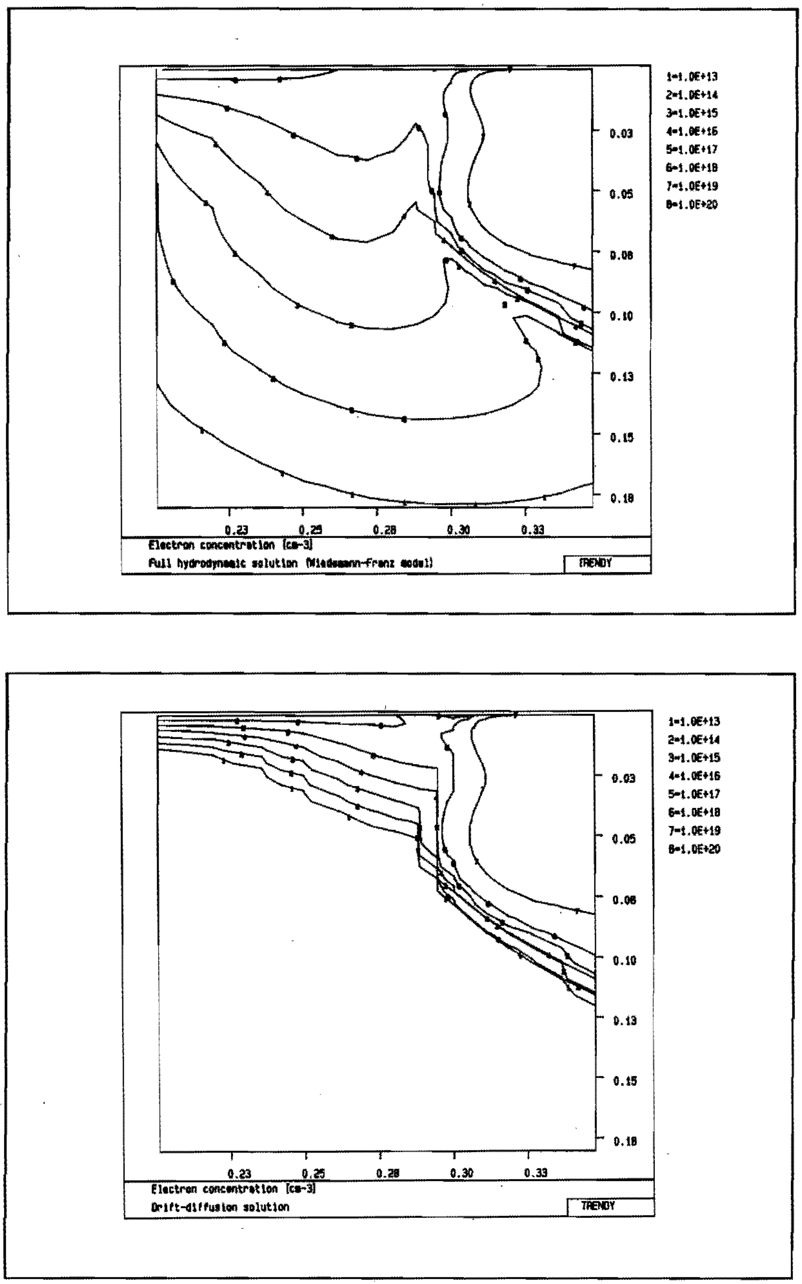

Fig 6.5 Electron density distribution in a 0.35 µm device, biased at V_{gs}=1.5 V and V_{ds}=2.6 V.
a. TRENDY full hydrodynamic results (Wiedemann-Franz heat flow model).
b. TRENDY drift-diffusion and energy post-processing.

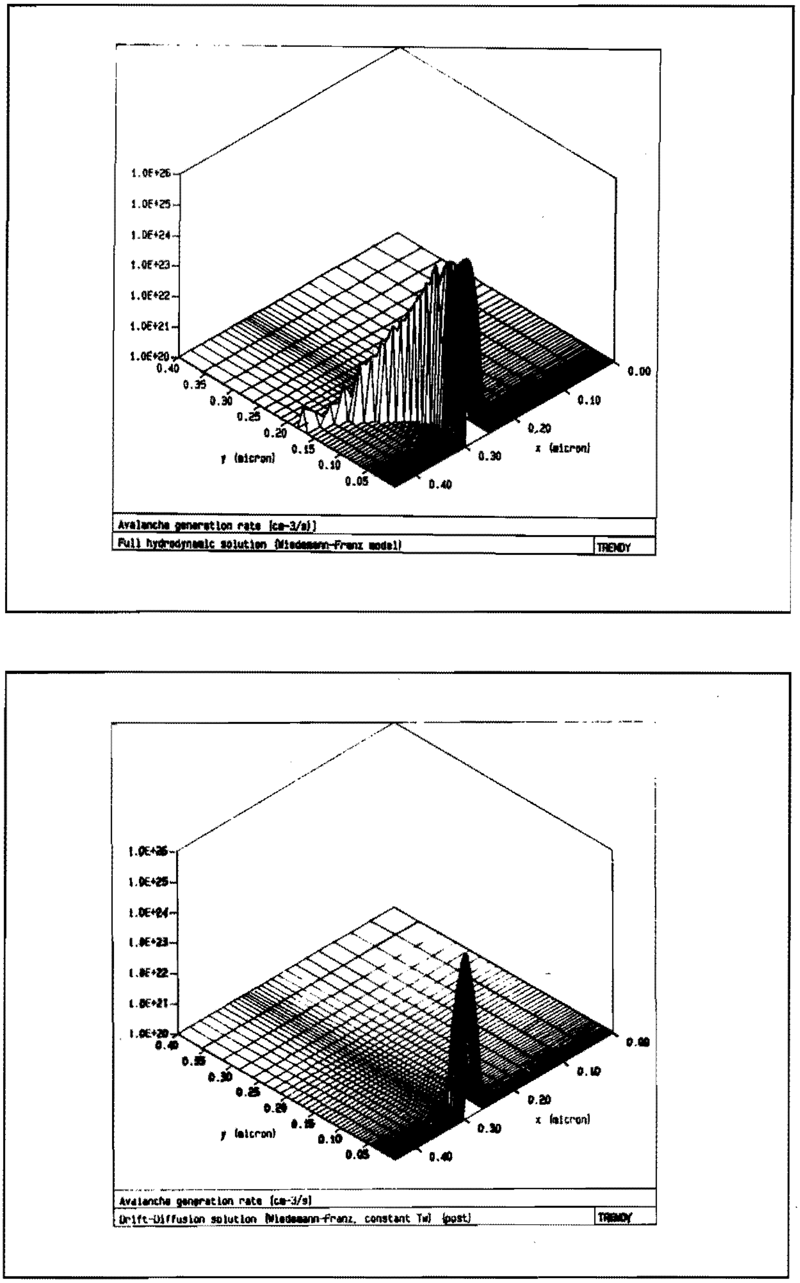

Fig 6.6 Avalanche generation rate in a 0.35 µm device, biased at V_{gs}=1.5 V and V_{ds}=2.6 V.

a. TRENDY full hydrodynamic results (Wiedemann-Franz heat flow model).

b. TRENDY drift-diffusion and energy post-processing.

Because the avalanche generation rate has an exponential relation with the electron temperature, minor differences in T_n result in major differences in GR_{AV}. Because both the electron temperature and the total current are slightly higher in the hydrodynamic model, the avalanche current is also higher. After integration we find the avalanche current calculated by the hydrodynamic model up to one order of magnitude higher than the value calculated by the extended drift-diffusion model.

From the above the following conclusions can be drawn:

1. The full hydrodynamic model yields higher currents and a more diffuse current path than the drift-diffusion model.
2. In both models the calculated avalanche current and the source-drain current are lower than the experimental data. This suggest that the electron temperatures calculated are too low.
3. Using the Wiedemann-Franz heat-flow model, the Van Overstraeten avalanche parameter values give better results than the Slotboom's parameters.
4. The post-processing method can be used to calculate the electron temperature distribution (under the restriction that the avalanche current is much lower than the drain current).
5. In calculating the avalanche currents, the post-processing method yields slightly lower values than the full hydrodynamic method.

6.3.3 Results obtained using the adiabatic heat-flow model

As was shown in chapter 4, the adiabatic heat-flow model yields higher electron temperatures than the Wiedemann-Franz model, and a better agreement with Monte Carlo results. Unfortunately, we did not succeed in constructing an adequate two-dimensional discretization scheme. Therefore no full hydrodynamic simulations using the adiabatic model could be executed, and a comparison between hydrodynamic and post-processing results is not possible. Actually, not even a converging two-dimensional post-processing routine based on the adiabatic heat-flow model could be developed. This is mainly due to the high fields and electron velocities occurring in the depletion layers. However a solution to these problems has been found by assuming a slightly different relaxation mechanism: not a constant relaxation time τ_{wn} is assumed, but a constant relaxation length λ_{wn}, as defined by

$$\lambda_{wn} = \frac{5}{3} v_n \tau_{wn} \tag{6.3.1}$$

Then a value $\lambda_{wn} = 67$ nm is equivalent to $\tau_{wn} = 0.4$ ps if v_n equals v_{sat}.

This method has been tested using the post-processing method with the Wiedemann-Franz heat-flow model. Simulations show that the use of a constant

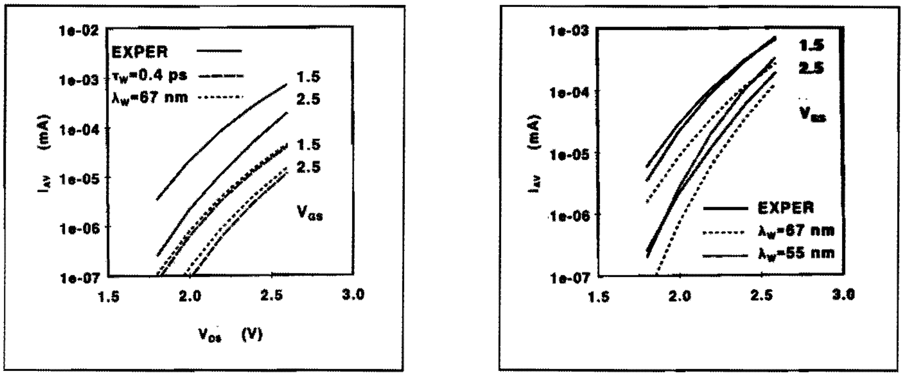

Fig 6.7 Avalanche currents. The solid lines represent experimental data, the dashed lines TRENDY extended drift-diffusion results.

(a) Wiedemann-Franz heat flow, combined with Van Overstraeten parameters. The large dashes represent results assuming a constant relaxation time, the small dashes results assuming a constant relaxation length.

(b) Adiabatic heat-flow model combined with Slotboom parameters, with two different values of the relaxation length.

λ_{wn} instead of a constant τ_{wn} yields slightly higher values for T_n and I_{AV}. The newly calculated value of I_{AV} is some 10% higher then the value obtained, using a constant τ_{wn}. This is shown in figure 6.7a.
This small increase of the peak temperature (only a few percent) is due to a higher relaxation time if the velocity is lower than the saturation velocity. From a pragmatic point of view the λ_{wn} approach is to be preferred above the τ_{wn} approach, both because of the increase of numerical stability, and because the small increase in I_{AV} partly compensates the decrease caused by the use of the post-processing method in stead of the full hydrodynamic solution.

In figure 6.7a the avalanche currents have been plotted using the Wiedemann-Franz model combined with Van Overstraeten parameters in the post-processing mode. Here the $\lambda_{wn} = 67$ nm approach has been used. The shape of the curves is very similar to the experimental curves, only the values are much lower. In accordance with experimental data it has been found that an increase in gate voltage yields a decrease in avalanche current.

In figure 6.7b the results of post-processing have been presented that assume an adiabatic heat-flow model, combined with a constant λ_{wn} and Slotboom avalanche parameters. In analogy with conclusion 4 of the previous section, it has been assumed that T_n can be calculated correctly if the post-processing routine is used.

Additional simulations show that the use of Van Overstraeten parameters yields avalanche currents that are much higher than experimental data. In combination with the adiabatic model the use of Slotboom's parameter values seems more adequate. The calculated avalanche currents are slightly lower than the experimental data. This was expected (see conclusion 5 of the previous section). Considering all results, the conclusion that the calculated electron temperature distribution is in good agreement with the real situation, seems justified.

It should be noted that the use of a smaller relaxation length $\lambda_{wn} = 55$ nm gives a better fit (figure 6.7b, large dashes). Although the difference between calculated and experimental data is better explained by the error made by the post-processing method, it can be compensated by a different value of λ_{wn}.

6.4 An analytical solution

Using the assumption of a constant relaxation length, an analytical solution to the calculation of the electron temperature was given by Slotboom [1991]. An important second assumption made is that the problem essentially is one dimensional along each current path. Furthermore several other assumptions were added to the energy conservation law (1.4.21):

1. The energy convection is neglected ($W_n = \frac{3}{2} k_B T_n$).
2. Using the adiabatic energy model ($\vec{Q} = \vec{0}$) and condition (1) yields an energy flow $\vec{S}_n = \frac{5}{2} k_B T_n \vec{J}_n / q$.
3. The velocity convection is neglected.
4. The generation/recombination term is neglected $GR = 0$.

Now (1.4.21) yields in the steady state

$$E_x J_{nx} + \frac{5}{2} \frac{k_B}{q} J_{nx} \partial_x T_n = \frac{3}{2} n \, k_B \frac{T_n - T_L}{\tau_{wn}} \qquad (6.4.1)$$

Rewriting (6.4.1) combined with the definition $J_{nx} = -qnv_{nx}$ and the definition (6.3.1) yields

$$\partial_x T_n + \frac{T_n - T_L}{\lambda_{wn}} + \frac{2}{5} \frac{q}{k_B} E_x = 0 \qquad (6.4.2)$$

Because λ_{wn} is assumed to have a constant value, equation (6.4.2) can be solved analytically, with the following expression for the electron temperature

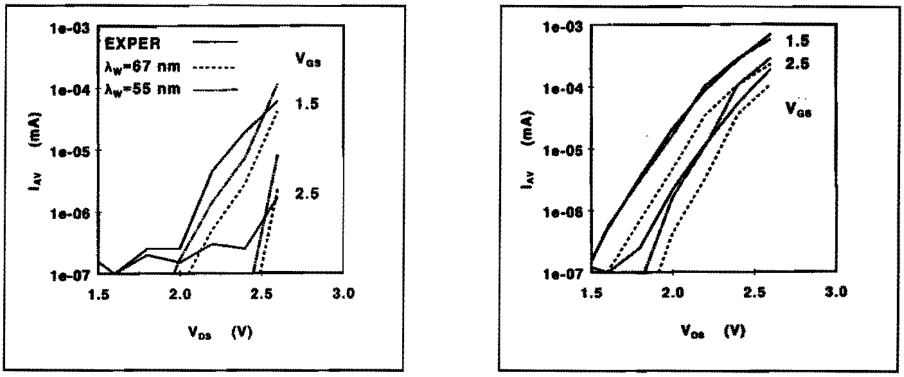

Fig 6.8 Avalanche currents. The simulation data are obtained from post-processing of MINIMOS drift-diffusion results using the analytical method: (a) 1.0 μm device, (b) 0.35 μm device. Results using two different values for the energy relaxation length are shown. The parameter is the gate voltage.

$$T_n(x) = T_L - \frac{2}{5}\frac{q}{k_B}\int_0^x E_x(u)\ \exp\left(\frac{u-x}{\lambda_{wn}}\right) du \tag{6.4.3}$$

The discretization of equation (6.4.3), assuming the electric field to have a constant value E_{xi} over the interval Δx_i between the grid points x_{i-1} and x_i (Gummel-Scharfetter), yields

$$T_n(x_i) = \exp\left(-\frac{\Delta x_{i-1}}{\lambda_{wn}}\right)\left(T_n(x_{i-1}) - H_i\right) + H_i \tag{6.4.4}$$

In (6.4.4) H_i is defined,

$$H_i = T_L - \frac{2}{5}\frac{q}{k_B}E_{xi}\lambda_{wn} \tag{6.4.5}$$

The results obtained by using this analytical method in post-processing are nearly indistinguishable from the results obtained by the two-dimensional post-process calculations that use the same constant relaxation length assumption. A great advantage of the analytical method is, that it gives no convergence problems at all, and can also be applied without any problem for MINIMOS results, which have, due to the smaller grid resolution, even more pronounced side peaks than TRENDY results. This way a fast, stable quasi-hydrodynamic simulator has been developed. Applying the method to MINIMOS drift-diffusion data yields figure 6.8a and 6.8b. An excellent fit between simulated and experimental data is found if $\lambda_{wn} = 55$ nm. This is in agreement with the results found by Slotboom et al. [1991].

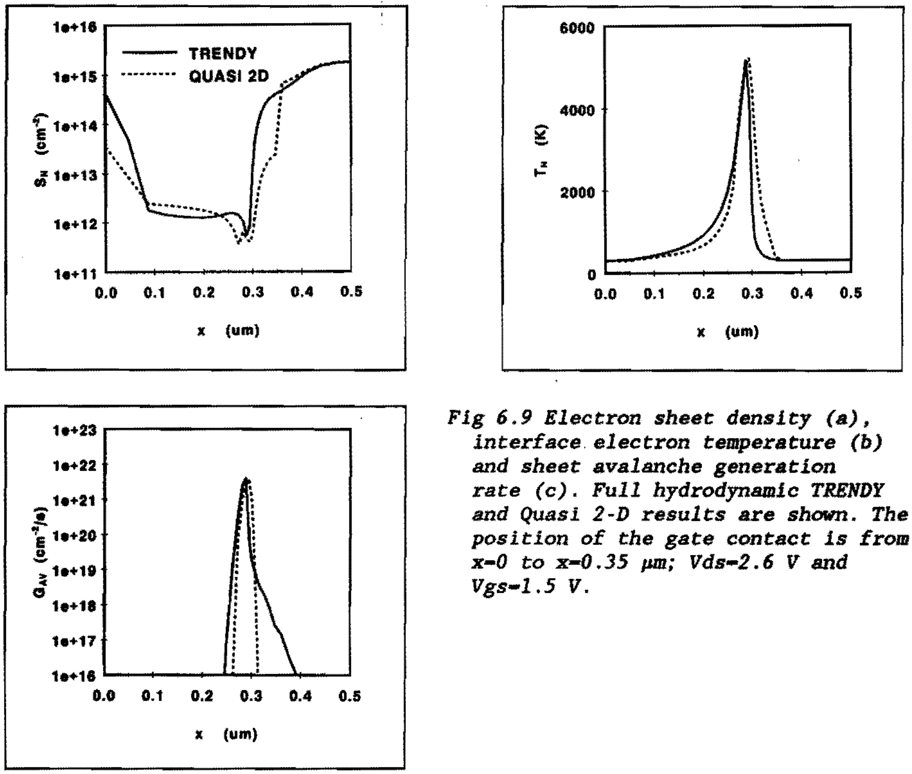

Fig 6.9 Electron sheet density (a), interface electron temperature (b) and sheet avalanche generation rate (c). Full hydrodynamic TRENDY and Quasi 2-D results are shown. The position of the gate contact is from x=0 to x=0.35 μm; Vds=2.6 V and Vgs=1.5 V.

6.5 Comparison of quasi and full two-dimensional results

In figures 6.9 full hydrodynamic results obtained by the quasi two-dimensional simulator and by TRENDY are compared. The Wiedemann-Franz heat-flow model has been used, and the 0.35 μm device has been biased at $V_{GS} = 1.5$ V and $V_{DS} = 2.6$ V. In this figure the gate contact is located from x= 0 μm to x= 0.35 μm. The solid line represents the full two-dimensional TRENDY solution and the dashed curve the quasi two-dimensional solution. In the quasi two-dimensional simulations the mobility is taken such that the device current is in agreement with the full two-dimensional results. In figure 6.9a the electron sheet densities have been plotted, that can be obtained from the electron density by integration in vertical direction.
A fair agreement is found at the position x= 0.3 μm, which is at the end of the channel, just in front of the drain p-n junction. For our purposes this is the most important region, because only there the electron temperature reaches high values, as can be seen in figure 6.9b.

In the ranges $0.3 < x > 0.35$ μm and $0 < x < 0.1$ μm the assumption that underneath the gate the vertical field is much larger than the horizontal

field is not fulfilled, and a discrepancy is found. However these areas have no significance for the device behavior.
A second discrepancy is found in the channel itself due to differences in the definition of the doping profile.

In figure 6.9b the electron temperature at the interface has been plotted. Again a good agreement is found at the position x= 0.3 μm. Because at this position the electron temperature reaches its maximum, also good agreement in the sheet avalanche generation rate is to be expected, which is confirmed by figure 6.9c.
From these results it can be concluded that the quasi two-dimensional model yields only a first order approximation of the device behavior, regarding both terminal behavior and internal distributions. Improvements can be added to the model, e.g. a full two-dimensional potential calculation instead of the layered structure model used, but whether these improvements will yield results comparable to the drift-diffusion method combined with an energy post-routine cannot be said a priori.

6.6 Summary and conclusions

In this chapter it has been shown that full hydrodynamic results of submicron n-channel MOSFETs using TRENDY can be obtained. However, much effort in defining an adequate grid and much computational effort was needed, because TRENDY has some serious convergence problems (like all hydrodynamic simulators) that have not been solved yet.
In all results obtained use was made of the Wiedemann-Franz heat-flow model. Using the adiabatic heat-flow model did not yield any results at all.

The above facts strongly limit the practical use of the TRENDY full hydrodynamic simulations in two-dimensions. This is in contrast with the TRENDY one-dimensional simulations, which do not give any problems.

Full two-dimensional hydrodynamic results have been compared to results obtained by applying a post-routine to drift-diffusion results. Although there are some serious discrepancies, e.g. in the electron density distribution, a fair agreement has been found in the electron temperature distribution. The peak temperature in the hydrodynamic mode is only some 10% higher than in the drift-diffusion mode. The same applies for the source-drain current: the hydrodynamic simulations yield 10% higher values than the drift-diffusion calculations. This is due to the extra thermal current components. However, these minor variations yield a major discrepancy in the avalanche current. Variations up to one order of magnitude have been observed. In this work the avalanche current always was much smaller than the source-drain current. If this is not the case, the discrepancy certainly will increase.

In this chapter we also looked in more detail at the analytical model proposed by Slotboom, who assumed an adiabatic model with a constant energy relaxation length λ_{wn}, in combination with the post-processing method, and the Slotboom avalanche parameters. The calculated avalanche currents are in fair agreement with experimental data. This can be explained by the fact that the error introduced by assuming a constant relaxation length in stead of a constant relaxation time, partly compensates the error introduced by the use of the post-processing method. Furthermore avalanche parameters that fit the experimental data have been used. In this way a fast and robust method to calculate good first-order approximations has been found, that may be useful for device designers. However, its limits of applicability have not been established.

In addition the hot electron mode of MINIMOS has been studied, but results obtained this way were not satisfactory.

Finally the quasi two-dimensional results, described in chapter 5, have been compared to full two-dimensional TRENDY results. The agreement is not very good, but the quasi two-dimensional model can be used as a first-order approximation.

Chapter 7

FUTURE WORK

The most important problems that demand further study are:

1. How to improve the convergence and stability of the full hydrodynamic model in two-dimensions. Promising in this respect is the method proposed by Schilders [1993].
2. Furthermore it is important to obtain a good discretization scheme in combination with the adiabatic heat-flow model, which, according to our Monte Carlo simulations, is the better model.

Only after these problems have been solved, the hydrodynamic model can be used as a useful tool for device designers.

Further work that is suggested as a continuation of the present work:

1. Improvement of the band-structure model and scatter mechanisms in the L-valley in the Monte Carlo simulators to model the high-field region more correctly.
2. Improvement of the ionized impurity and impact ionization models in the Monte Carlo method.
3. Further Monte Carlo studies of the interface behavior, including the impact parameters at the interface.
4. Extension of the Monte Carlo device simulator with hole transport.
5. Study of the transient behavior of the quasi two-dimensional model.
6. Improvement of the convergence of the quasi two-dimensional model by using a different iteration scheme. Maybe the model then can be used as a fast hydrodynamic device simulator in compact models.
7. Improvement of the quasi two-dimensional model by using a full two-dimensional potential calculation in stead of the layered structure used in this work.
8. Implemention of an automatic grid generator in TRENDY.
9. Implemention in both MINIMOS and in TRENDY of a subroutine to calculate automatically the gate-voltage correction due to the quantum well.

LIST OF REFERENCES

Aspnes D.A., 1976, Phys.Rev.B 14, p5331

Basu P.K., 1978, Solid-St.Comm. 27, p657

Bordelon T.J., Wang X.L., Maziar C.M. and Tasch A.F., 1990, Proc. IEDM, p353

Brooks H., 1955, Adv.Electronics and Electron Ph. 7, p85

Brunetti R., Jacoboni C., Matulionis A. and Dienys V., 1985, Physica 134B, p369

Canali C., Ottaviani G. and Quaranta A.A., 1971, J.Phys.Chem.Solids 32, p1719

Canali C., Jacoboni C., Nava F., Ottaviani G. and Alberigi-Quaranta A.A., 1975, Phys.Rev.B 12, p2265

Chattopadhyay D. and Queisser H.J., 1981, Rev.Mod.Phys. 53, p745

Chu-Hao, Zimmermann J., Charef M., Fauquembergue R. and Constant E., 1985, Solid-St.Elec. 8, p733

Chynoweth A.G., 1958, Phys.Rev. 109, p1537

Conwell E. and Weisskopf V.F., 1950, Phys.Rev. 77, p388

Elias P.J.H., Van de Roer Th.G. and Klaassen F.M., 1990, Proc. ESSDERC, p145

Fawcett W., Boardman A.D. and Swain S., 1970, J.Phys.Chem.Solids 31, p1963

Fischetti M.V. and Laux S.E., 1988, Phys.Rev.B , p9721

Fischetti M.V., 1991, IEEE ED-38, p634

Fischetti M.V. and Laux S.E., 1991, IEEE ED-38, p650

Forghieri A., Guerrieri R. Ciampolini P., Gnudi A., Rudan M. and Baccarani G., 1988, IEEE CAD-7, p231

Gradshteyn I.S. and Ryzhik I.M., 1980, *Table of integrals, series and products*, Academic, London

Graybeal W.J. and Pooch U.W., 1980, *Simulation: principles and methods*, Winthrop, Cambridge Mass.

Gummel H.D., 1964, IEEE ED-11, p455

Hammersley J.M. and Handscomb D.C., 1964, *Monte Carlo methods*, Methuen, London

Herring C. and Vogt E., 1956, Phys.Rev. 101, p944

Hockney R.W. and Eastwood J.W., 1981, *Computer simulation using particles*, McGraw-Hill, New York

Jacoboni C. and Reggiani L., 1983, Rev.Mod.Phys. 55,p645

Kane E.O., 1957, J.Phys.Chem.Solids 1, p249

Keldysh L.V., 1965, Sov.Phys.JETP 21, p1135

Kurosawa T., 1966, J.Phys.Soc.Japan Suppl. 21, p424

Laux S.E. and Fischetti M.V., 1988, IEEE EDL-9, p467

Lugli P. and Ferry D.K., 1983, Physica 117B & 118B, p251

McQuarrie D.A., 1976, *Statistical Mechanics*, Harper and Row

Nederveen K., 1989, *Electron Transport in AlGaAs/GaAs Heterostructures*, Ph.D. Thesis, University of Technology Eindhoven

Pals J.A., 1972, Phys.Lett. **39A**, p101

Perri F., 1986, *Sur un nouveau modèle de l'interaction electron-impurité dans les semiconducteurs*, Ph.D. Thesis, Université des Sciences et Techniques de Lille

Rees H.D., 1968, Phys.Lett. **26A**, p416

Rees H.D., 1969, J.Phys.Chem.Solids **30**, p643

Ridley B.K., 1977, J.Phys.C:Solid-St.Phys. **10**, p1589

Rode D.L., 1978, in: *Semiconductors and Semimetals, vol.* 10, Eds. Willardson R.K. and Beer A.C., Academic, London

Rudan M. and Odeh F., 1986, COMPEL **5**, p149

Scharfetter D.L. and Gummel H.D., 1969, IEEE **ED-16**, p64

Schilders W.H.A., 1993, Proc. SISDEP, to be published

Schoenmaker W., Vankemmel R., Cartuyvels R., Magnus W., Tijskens B., De Meyer K., 1991, COMPEL **10**, p621

Selberherr S., Schütz A. and Pötzl H., 1980, IEEE **ED-27**, p1540

Selberherr S., 1989, IEEE **ED-36**, p1464

Shahidi G.G., Antoniadis D.A. and Smith H.I., 1988, J.Vac.Sci.Techn.B **6**, p137

Slotboom J.W. and De Graaff H.C., 1976, Solid-St.Elec. **19**, p857

Slotboom J.W., Streutker G., Davids G.J.T. and Hartog P.B., 1987, Proc. IEDM, p1987

Slotboom J.W., Streutker G., Van Dort M.J., Woerlee P.H., Pruijmboom A. and Gravesteijn D.J., 1991, Proc. IEDM, p127

Stern F. and Howard W.E., 1967, Phys.Rev.B **163**, p816

Tang J.Y. and Hess K., 1983, J.Appl.Phys. **54**, p5139

Tang T.W., 1984, IEEE **ED-31**, p1912

Van Dort M.J. and Woerlee P.H., 1991, Proc. SISDEP, p451

Van de Roer T.G. and Widdershoven F.P., 1986, J.Appl.Phys. **59**, p813

Van Overstraeten R.J. and De Man H.J., 1970, Sol.St.Electr. **13**, p583

Van Schie E., 1990, *TRENDY, An integrated program for IC process and device simulation*, Ph.D. Thesis, Twente University of Technology

Widdershoven F.P. 1988, Phys.Rev.B **38**, p3391

Wiliams C.K., 1982, *Monte Carlo studies of the electron transport in III-V semiconductors and semiconductor devices*, Ph.D. Thesis, North Carolina State University

Wolbert P., 1991, *Modeling and simulation of semiconductor devices in TRENDY*, Ph.D. Thesis, Twente University of Technology

Zimmermann J., 1980, *Étude des phénomènes de transport électronique dans le silicium de type n en régimes stationnaires et non stationnaires par la méthode de Monte Carlo. Application à la simulation de composants submicroniques*, Ph.D. Thesis, Université des Sciences et Techniques de Lille

CURRICULUM VITAE

Peter Elias was born on March 15, 1958 in Tilburg, the Netherlands.

He graduated in 1976 at the Sint Odulphus Lyceum (Gymnasium β) in Tilburg.

On February 23, 1984 he received his M.Sc. degree in Electrical Engineering at the Eindhoven University of Technology, for theoretical work on planar multi-beam antennas. During his studies he also has been working for six months at the Philips Research Laboratories, Briarcliff Manor, New York, USA, studying passivation layers for high-voltage LDMOSTs.

From February 1984 to August 1986 he has been working as a teacher mathematics and information technology. In this same period he also spent half a year in an Israelian Kibbutz.

From August 1986 to December 1987 he has been working as an research associate at the Physics department of the Eindhoven University, modeling and measuring the small magnetic fields generated by the human heart and brain.

From January 1988 to June 1993 he was attached to the Electric Engineering department of the same University, working at the Electronic Device group on simulation studies of hot-electron effects in Si MOS devices.

Stellingen bij het proefschrift

Modeling of hot-electron effects in Si MOS devices

door

Peter Elias

1. Het adiabatische warmtestromingsmodel in hydrodynamische device-simulaties is te prefereren boven het Wiedemann-Franz model.
   Dit proefschrift, hfdst 4.

2. Het gebruik van een drift-diffusie-oplossing gecombineerd met een post-processor om de elektronentemperatuur en de vermenigvuldigingsfaktor te berekenen levert, indien vergeleken met een volledig hydrodynamische oplossing, een afwijking van ca 10% in de drainstroom en een afwijking van een faktor twee in de substraatstroom van MOS-transistoren met kanaallengten tot 0.3 μm.
   Dit proefschrift, hfdst 5 en 6.

3. Daar in de etymologie van het woord *okay* geen verbinding gelegd wordt met de nederlandse scheepvaart, mag verondersteld worden dat de uitdrukking *Op Koers* in het begin van de 19e eeuw nog niet gebezigd werd.
   Mencken [1952], The American language, Suppl.

4. Een verbetering van kwaliteit, produktiviteit en arbeidsvreugd in het wetenschappelijk onderzoek wordt ongetwijfeld verkregen, indien steeds twee AIO's aan eenzelfde projekt werken, en zij in hun mogelijkheden tot overwerk niet belemmerd worden.

5. Als de overheid niet in staat is op korte termijn serieus op te treden tegen het massaal overtreden van de regels voor verkeerslichten door fietsers en voetgangers, dan doet zij er verstandig aan deze regels af te schaffen.

6. Zolang in numerieke modellen, die gebruikt worden bij de Elektro- en Magneto-Encephalografie, uitsluitend bolsymmetrische beschrijvingen van het menselijk hoofd worden gehanteerd, zal de nauwkeurigheid van de bronlokalisatie niet beter worden dan circa 1 cm.
   Peters en Elias [1988], Med. & Biol. Eng. & Comp., vol 26.
   Cohen, Cuffin, e.a. [1990], Annals of Neurology, vol 28.

7. Programmatuur ontwikkeld aan nederlandse universiteiten zal in het algemeen een stille dood sterven, als ze niet overgedragen kan worden aan een commercieel software-huis.

8. Bij de berekening van de I-V karakteristieken van AlGaAs DBRT-diodes moet aandacht besteed worden aan de rol van de inhomogene verdeling van Aluminium in de barrières.
   Salemink en Albrektsen [1991], J. of Vacuum Sc. & Techn.B, vol 9, iss 2.