Chapter 1
Numerical Power/HV Device Modeling

Oliver Triebl and Tibor Grasser

Abstract Modern semiconductor devices have to fulfill many requirements in terms of performance, reliability, and costs. The structures have become very complex and have undergone major optimizations compared to the original proposals half a century ago. This complexity almost always requires Computer Aided Design (CAD) tools for the design of electric and electronic units. Usually different engineering levels have to be considered in the design process. At the circuit level CAD tools like SPICE can be used to adjust and test electronic circuits. Analysis of a device itself can be considered one step down on the engineering hierarchy levels. Dopant and carrier distributions become important and the spatial distribution of quantities has to be considered. Software tools supporting the design of devices at this level are known as Technology Computer Aided Design (TCAD) tools. This chapter will give an introduction in the drift-diffusion method which is probably the most important carrier transport model used in TCAD. Modeling of mobility and of generation and recombination will also be considered. Finally, a discussion on numerical considerations on solving the problems will be given.

Keywords Tcad · Device simulation · High-voltage · Drift-diffusion · Mobility modeling

1 Introduction

Modern semiconductor devices have to fulfill many requirements in terms of performance, reliability, and costs. Structures have become very complex and have undergone major optimizations compared to the semiconductor structures originally proposed half a century ago. This complexity almost always requires Computer

O. Triebl (✉) and T. Grasser
Christian Doppler Laboratory for Technology Computer Aided Design,
Institute for Microelectronics, TU Wien, 1040 Wien, Austria
e-mail: triebl@iue.tuwien.ac.at

W. Grabinski and T. Gneiting (eds.), *Power/HVMOS Devices Compact Modeling*,
DOI 10.1007/978-90-481-3046-7_1,

Aided Design (CAD) tools for the design of electric and electronic units. Different engineering levels are considered in the design process. At the circuit level CAD tools like SPICE can be used to adjust and test electronic circuits. Here, the single devices are commonly simulated using calibrated compact models. Analysis of the device itself can be considered one step down on the engineering hierarchy levels. Dopant and carrier distributions become important and the spatial distribution of quantities has to be considered. Software tools supporting the design of devices at this level are known as Technology Computer Aided Design (TCAD) tools.

This chapter gives an introduction on how TCAD tools for device simulation work and how they can be used. The basic equations needed to perform drift–diffusion simulations are presented together with some mobility and generation/recombination estimations. Sample simulations show the influence of crucial physical models and demonstrate how TCAD gives the device engineer an insight into the device behavior and how that information can be used for further device optimization. Some considerations on solving the equation systems by applying discretization schemes and using iterative solution techniques will be given.

As the focus of this introductive chapter is on the modeling and simulation of high-voltage devices, the validity of drift–diffusion framework will be assumed. These drift–diffusion equations are also the starting point for the derivation of most compact models. More advanced and computationally demanding transport models such as energy-transport, six moment models, and Monte Carlo simulation have gained some interest during the last decade and are briefly summarized.

1.1 TCAD – Technology Computer Aided Design

Tools used to simulate semiconductor manufacturing processes and semiconductor device behavior belong to the group of TCAD tools. These tools aim to reproduce and especially predict the physical mechanisms and to determine the resulting device structure and device behavior. For this, models describing the best physical phenomena are required. Proper models have to be found and have to be included in the simulation tools resulting in differential equation systems. Since the resulting mathematical problem generally cannot be solved explicitly, numerical techniques are applied. Solutions can be found using statistical methods, the Monte Carlo method for example, or iterative techniques. For efficient computation in engineering environments, most often simplified models have to be used.

1.2 Benefits of Numerical Modeling

In numerical device simulation the structure of a device is represented using distributed quantities. With distributed we mean that in the most general case important quantities such as the electrostatic potential ψ and the carrier concentrations n and

p are calculated in the three-dimensional space $r = (x, y, z)$. Since the physical processes in the semiconductor device are modeled as realistically as possible, the device behavior directly results from the simulation. This is fundamentally different compared to compact modeling. There, every type of device requires its own compact model, whereas in device simulation proper modeling of semiconductor physics delivers results for all types of devices. This implies that the geometry and the doping profile of the real device are considered implicitly in the simulation as accurately as required. On the other hand, one has to consider that the use of any of the physically based approaches addressed in this chapter requires a considerable computational effort compared to compact modeling. For an LDMOS structure, for example, it might be required to calculate quantities on 10,000 or even more mesh points. Application of the drift–diffusion model with three unknown quantities for each point results for a two-dimensional problem in an equation system with more than 30,000 unknowns which has to be solved for every step in the iterative solution process.

Nevertheless, numerical device simulation gives the device engineer the possibility to understand and analyze the inner life of the device which is completely hidden in compact modeling. The insight helps to find possible failure mechanisms. Failures in real world devices can be reproduced in the simulation and solutions can be derived and tested. This gives the possibility to reduce the number of test devices that have to be fabricated to a minimum.

TCAD tools also allow to estimate the performance of completely new structures which is especially helpful to predict any parasitic effects in integrated circuits. Possible failures due to parasitic transistors can be discovered by analyzing two or more neighboring devices together in a single simulation. This has also become important for high-voltage and power devices, since integration with other technologies as is used in so called SMART power devices, has become popular and introduces many potential failures. Reliability demands and lifetime predictions are also important for many applications. The analysis of aging and degradation can be supported using numerical modeling approaches.

1.3 Sample Device

For all simulations performed in this chapter the device simulation tool MINIMOS-NT [1] has been used. The sample device used is based on a high-voltage sample device fabricated by austriamicrosystems [2]. It is an n-channel lateral DMOSFET with a gate oxide thickness of 7 nm, fabricated in a 0.35 µm CMOS-based technology. The gate length is 0.5 µm, the width is 40 µm, and the specified application voltage is 50 V. The doping profile and the geometry depicted in Fig. 1.1 were simulated using the Synopsys process simulation tools. The domain used for simulation has a depth of 15 µm, while most figures depict only the upper 3.5 µm of the device.

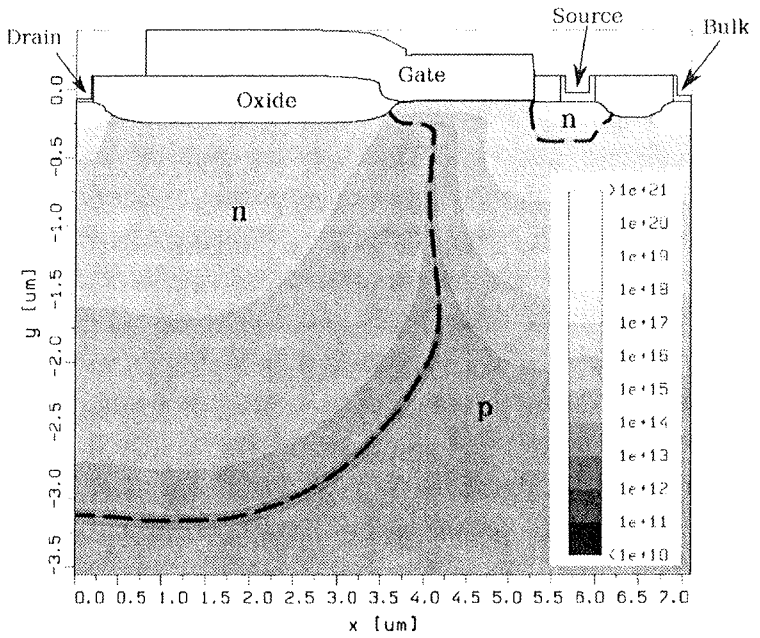

Fig. 1.1 Structure of the sample n-channel LDMOSFET used for simulations in this chapter. The transitions between n- and p-doped regions are marked with dashed lines and the absolute value of the net doping concentration is shown in $1/cm^3$. The simulated structure of the device continues into a depth of 15 µm, while only the upper part is depicted

2 Device Modeling

TCAD simulations require appropriate physical and mathematical models describing the device behavior in a self consistent way. The fundamental equations needed to model semiconductor device structures are the Poisson equation, the continuity, and transport equations, which in their simplest form have become known as the drift–diffusion model [3]. In the following sections, models for different physical properties relevant to high-voltage devices are presented. The high fields particularly influence the mobility within the device leading to high differences of the mobility in the drift and the channel region. These mobility variations strongly depend on the operating point. The peaks of the electric field require the consideration of impact ionization, whereas Shockley–Read–Hall generation and recombination dominates the currents in space charge regions for lower fields. Due to the thick oxides used in high-voltage devices quantum-mechanical tunneling currents are normally of minor importance. Also, consideration of quantum confinement is usually not required for the relatively thick oxides used in these devices. Thermal modeling definitely plays an important role. Since changes of the device temperature due to power dissipation have a strong impact on nearly all device properties.

2.1 Semiconductor Equations

In macroscopic semiconductor device modeling, Poisson's equation and the continuity equations play a fundamental role. Poisson's equation, one of the basic

equations in electrostatics, is derived from the Maxwell equation $\nabla \cdot \mathbf{D} = \rho$ and the material relation $\mathbf{D} = \hat{\epsilon}\mathbf{E}$. Using the electrostatic potential ψ with $\mathbf{E} = -\nabla\psi$, Poisson's equation reads

$$\nabla \cdot (\hat{\epsilon}\nabla\psi) = -\rho. \tag{1.1}$$

Since even for high frequencies the wavelength is typically much smaller than the device dimension, the quasi stationary approximation used to derive Poisson's equation are justified. In semiconductors the charge density ρ is commonly split into fixed charges which are in particular ionized acceptors N_A^- and donors N_D^+ and into free charges which are electrons n and holes p. The permittivity tensor $\hat{\epsilon}$ is considered time invariant in the derivation of Poisson's equation. In isotropic materials like silicon the permittivity can be additionally approximated by the scalar value ϵ. Furthermore the permittivity is often considered to be constant within a material segment, therefore Poisson's equation used in device simulation tools looks like

$$\nabla^2\psi = -\frac{\mathrm{q}}{\epsilon}(n - p + N_A^- - N_D^+). \tag{1.2}$$

The second important equation, the continuity equation, can also be derived from Maxwell's equations and reads

$$\nabla \cdot \mathbf{J} + \frac{\partial\rho}{\partial t} = 0. \tag{1.3}$$

The current density $\mathbf{J}$ is split into $\mathbf{J}_n$ and $\mathbf{J}_p$, for the contributing carrier types n and p, respectively. By introducing the recombination rate R, two separate continuity equations, one for each carrier type, can be written as

$$\nabla \cdot \mathbf{J}_n - \mathrm{q}\frac{\partial n}{\partial t} = +\mathrm{q}R \qquad \text{and} \tag{1.4}$$

$$\nabla \cdot \mathbf{J}_p + \mathrm{q}\frac{\partial p}{\partial t} = -\mathrm{q}R. \tag{1.5}$$

Alternatively, Eqs. 1.4 and 1.5 can be derived from the Boltzmann transport equation using the method of moments [4]. The separation into two equations allows independent transport modeling of the carrier types. Generation and recombination rates of electrons and holes are expressed using physically or empirically based recombination models [3], some of which are described in Section 2.3. The rate R represents the net rate only, which is zero in thermal equilibrium where generation and recombination are balanced.

2.2 *Carrier Transport Equations*

A semiclassical description of carrier transport is given by Boltzmann's transport equation (BTE) which describes the evolution of the distribution function in the six-

dimensional phase space (x, y, z, p_x, p_y, p_z). Unfortunately, analytical solutions exist only for very simple configurations. One popular approach for solving the BTE is the Monte Carlo method [5] which is highly accurate but also very time consuming.

Simulation on an engineering level requires simpler transport equations which can be solved for complex structures within reasonable time. One method to perform this simplification is to consider only moments of the distribution function [6]. Depending on the number of moments considered for the model, different transport equations can be evaluated. Use of the first two moments results in the drift–diffusion model, a widely applied approach for modeling carrier transport.

2.2.1 The Drift–Diffusion Model

The drift–diffusion current equation can be derived from the BTE using the method of moments [3] or, alternatively, from the basic principles of irreversible thermodynamics [7]. The resulting electron and hole current relations contain at least two components caused by carrier drift and carrier diffusion. Inclusion of the driving force caused by the lattice temperature gradient (∇T_L) [8] leads to

$$\mathbf{J}_n = \mathrm{q}n\mu_n\mathbf{E} + \mathrm{q}D_n\nabla n + \mathrm{q}nD_n^T\nabla T_\mathrm{L} \tag{1.6}$$

$$\mathbf{J}_p = \mathrm{q}p\mu_p\mathbf{E} - \mathrm{q}D_p\nabla p - \mathrm{q}pD_p^T\nabla T_\mathrm{L}. \tag{1.7}$$

μ_ν (ν stands for n and p) represents the carrier mobility, D_ν^T the thermal diffusion coefficients, and D_ν the diffusivity which is often expressed via the mobility invoking the Einstein relation

$$D_\nu = \mu_\nu \frac{\mathrm{k_B}T_\mathrm{L}}{\mathrm{q}}, \tag{1.8}$$

where $\mathrm{k_B}$ is the Boltzmann constant. The Einstein relation is strictly valid only in equilibrium [9].

The Eqs. 1.6 and 1.7 together with (1.4), (1.5), and (1.2) form the drift–diffusion model which was first presented by Van Roosbroeck in [10]. Rigorous derivations from the BTE show that many simplifications are required to obtain the drift–diffusion equations as shown. Simplifications are, for example, the single parabolic band structure or the cold Maxwellian carrier distribution function which assumes the carrier temperature equal to the lattice temperature. Nevertheless, due to its simplicity and excellent numerical properties, the drift–diffusion equations have become the workhorse for most TCAD applications. It also forms the basis for the bulk of compact models.

The following example illustrates that the drift–diffusion model as presented above can be directly applied to semiconductor devices. In this first approach the example LDMOSFET in Fig. 1.1 was simulated, where the recombination rate R was neglected and the mobility assumed to be constant. Note that these are standard

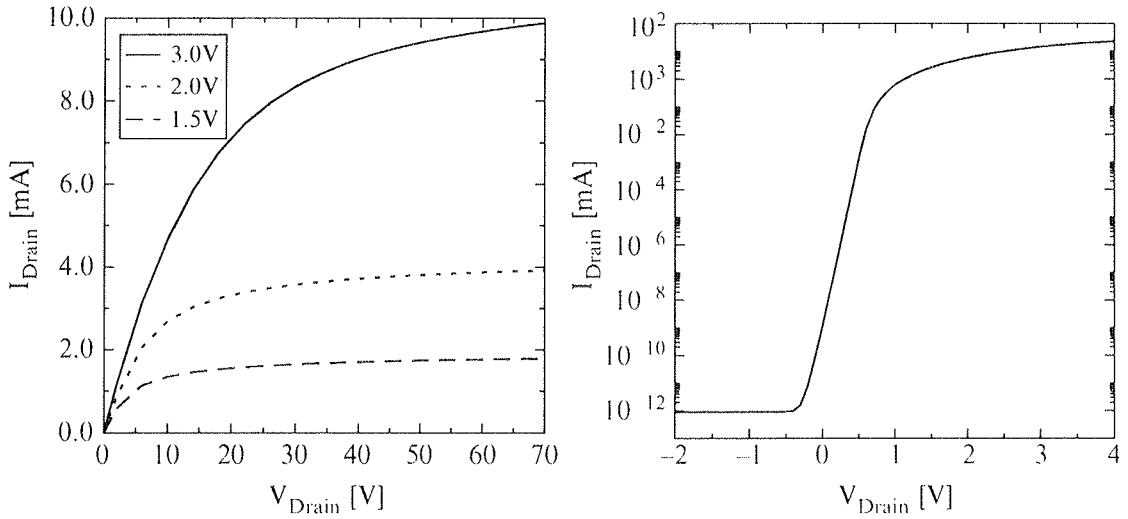

Fig. 1.2 Simulation of output (*left*) and transfer (*right*) characteristic of the sample device using the drift–diffusion model only. Generation and recombination were neglected and the mobility was assumed to be constant

assumptions for the derivation of compact models. This basic simulation gives the output characteristics shown in Fig. 1.2, no other models are required. Despite these simplifications, the basic characteristic of MOSFET devices can already be seen. However, neglecting all generation and recombination effects and assuming a constant mobility is a strong oversimplification resulting in quantitative and qualitative errors. However, the spatial variation of the mobility can be easily included in TCAD simulations since the mobility and other distributed parameters can be evaluated individually for every position within the device. The device geometry and the doping profiles are therefore considered implicitly and directly influence the device behavior.

2.2.2 Higher-Order Transport Models

To obtain a better approximation of the BTE, higher-order transport models can be derived using the method of moments, for example. The most prominent models beside the drift–diffusion model are the energy-transport/hydrodynamic models which use three or four moments. These models are based on the work of Stratton [11] and Bløtekjær [12], a detailed review is given in [4]. In addition to the quantities used in the drift–diffusion model, the energy flux and the carrier temperatures with corresponding equations are introduced, which require additional transport parameters. Modeling of carrier mobility and impact-ionization benefits from more accurate models based on the carrier temperatures rather than the electric field. This advantage is caused by the non-local behavior of the average energy with respect to the electric field and becomes especially relevant for small device structures (Fig. 1.3 left). However, it shows that describing the energy distribution function using only the carrier concentration and the average carrier temperature is still not sufficient for specific problems depending on high energy tails (Fig. 1.3 right). Hot carrier mod-

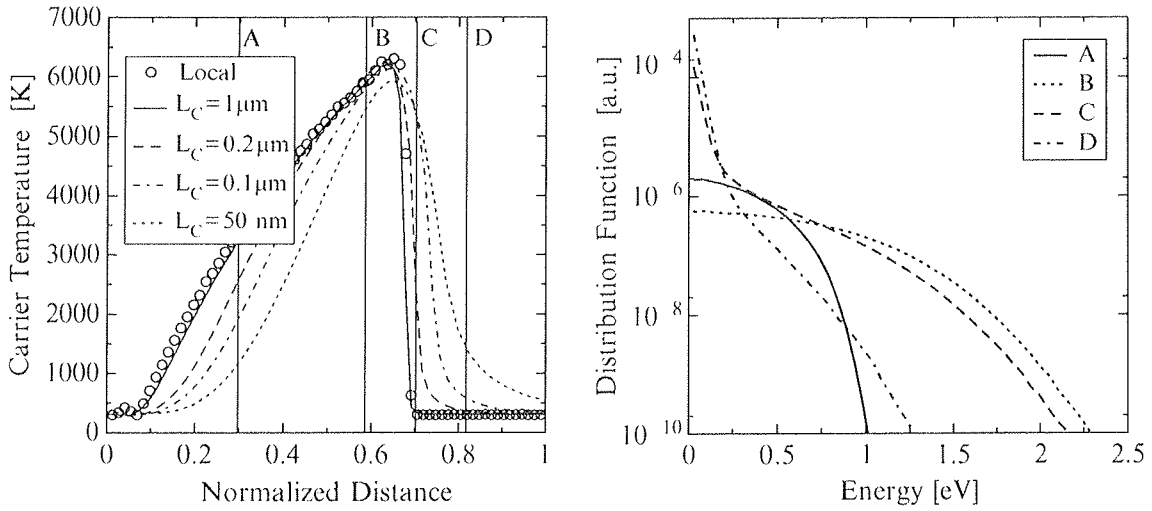

Fig. 1.3 Electron temperature (*left*) and distribution functions (*right*) of comparable $n^+ - n - n^+$ structures with varying channel lengths L_C. The spatial coordinates have been normalized to get an overlapping electric field. It can be seen that the local electric field approach is sufficient for larger structures but gives poor results for small structures. The distribution functions on the right are shown for $L_C = 200\,\text{nm}$ at the positions A to D. Note that the average energies at the points A and C are the same, whereas the distribution function looks completely different. Also note the high-energy tail at point D where the carrier temperature is already close to the lattice temperature with a value of 370 K [13]

eling in small structures, for example, which is based on accurate modeling of the high-energy part of the distribution function would require more complex models. The six moments method [14] is one possibility to improve the approximation of the distribution function.

Due to their better approximation of the BTE, higher-order transport models often give better results than the drift–diffusion model [15]. This effect is especially relevant for small structures where non-local effects gain importance (Fig. 1.3). However, since high-voltage device structures are relatively large, the drift–diffusion model is usually sufficient. The drawbacks of using more complex transport equations are the higher computation time and the increase of numerical instabilities and convergence issues. This issue is even more relevant for a high number of mesh points which are necessary for large devices. If required, different transport equations can be used within one device. This allows to model critical areas using higher-order transport models whereas the drift–diffusion model is used for the rest of the device. Areas of interest might be the channel area of an MOS device or areas having spatially strongly varying electric fields.

2.3 *Parameter Modeling*

The semiconductor equations discussed in the former section show the basic relations between carrier distribution and the electrostatic potential. Two parameters,

the mobility and the recombination rate were introduced, which require appropriate modeling. The physical phenomena influencing these parameters are manifold and will be discussed in the following.

2.3.1 Mobility

The derivation of the mobility originates from carrier relaxation times. The mobility is influenced by the lattice and its thermal vibrations, impurity atoms, surfaces and interfaces to neighboring materials, the carriers themselves, the energy of the carriers, and other effects like lattice defects. Mobility models are used to make an estimation considering these effects and make simulations in continuous systems possible. Since exact derivations are too complex or just do not exist, empirical approaches are often used. Some of the commonly used approaches will be discussed here.

A common method for modeling the mobility is the hierarchically encapsulation of the physical mechanisms. In this approach, the most fundamental mechanism is considered to be the lattice scattering dependence (μ^{L}) followed by the ionized impurity dependence (μ^{I}). Especially in MOS devices, a surface correction (μ^{S}) is of special importance. These three contributions classify the low-field mobility models. Modeling of high-field effects is introduced with a field dependence model (μ^{F}). These contributions may be combined as it is for example done in the MINIMOS mobility model [16] which looks like

$$\mu^{\mathrm{LISF}} = \mu^{\mathrm{LISF}}\left(\mu^{\mathrm{LIS}}\left(\mu^{\mathrm{LI}}\left(\mu^{\mathrm{L}}\right)\right)\right). \tag{1.9}$$

The individual mechanisms are assumed to be independent of each other. All values resulting from mobility calculations are obviously different for electrons and holes. To demonstrate the impact of lattice, impurity, and surface scattering on the mobility, the electron mobility distribution in the sample device is shown in Fig. 1.4.

In contrast to the encapsulation approach (1.9) for the mobility calculation used in the MINIMOS model, the Lombardi model [17] combines three carrier mobility components using Matthiessen's rule. The components are derived from surface acoustic phonon scattering, from bulk carrier mobility, and from surface-roughness scattering. A similar expression has been used by Agostinelli [18] for holes, additionally accounting for interface charge and screened Coulomb scattering (1.10).

$$\frac{1}{\mu_\nu^{\mathrm{LIS}}} = \frac{1}{\mu_\nu^{\mathrm{ph}}} + \frac{1}{\mu_\nu^{\mathrm{sr}}} + \frac{1}{\mu_\nu^{\mathrm{c}}} \tag{1.10}$$

Here, the phonon scattering component μ_ν^{ph} combines scattering with bulk phonons, surface phonons, and fixed interface charges. μ_ν^{sr} includes the dependence of the surface-roughness scattering on the electric field orthogonal to the interface and μ_ν^{c} models the screened Coulomb scattering. Modeling the influence of interface charges is of special interest in reliability modeling. Effects like negative bias temperature instability [19] or hot carrier degradation [20] generate interface

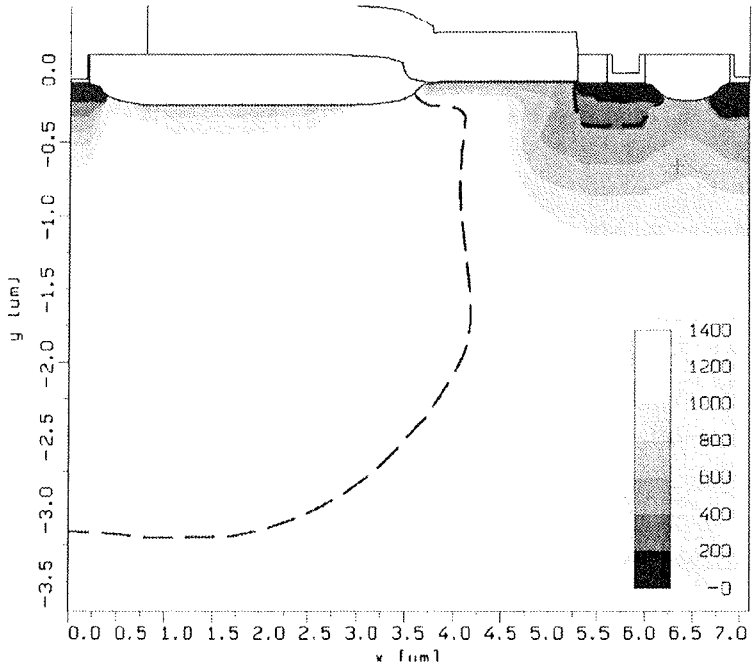

Fig. 1.4 Electron mobility distribution (in cm^2/Vs) in the sample device taking into account the lattice, impurity, and surface scattering models

traps leading to interface charges. The prediction of this device degradation is of crucial interest for semiconductor manufactures. Other mobility models based on Matthiessen's rule have been developed, for example, by Darwish [21] or Neinhüs [22], the latter additionally includes quantum confinement influences on the mobility.

A universal dependence of the effective minority carrier mobility in inversion channels on the effective vertical field ($E_{\perp,\mathrm{eff}}$) has been shown in [23]. The effective vertical field in the inversion layer is modeled using the bulk (depletion) Q'_B and mobile (inversion) Q'_I charge layer densities,

$$E_{\perp,\mathrm{eff}} = \frac{1}{\epsilon}\left(Q'_\mathrm{B} + \frac{1}{2}Q'_\mathrm{I}\right). \tag{1.11}$$

and the effective mobility can be approximated using the empirical fit [24]

$$\mu_\mathrm{eff}(E_{\perp,\mathrm{eff}}) = \mu_\mathrm{max}\left(\frac{E_\mathrm{C}}{E_{\perp,\mathrm{eff}}}\right)^{C_\mathrm{I}}. \tag{1.12}$$

In this model, fixed oxide charges and the channel doping are used to model the maximum effective mobility μ_max as well as the quantities E_C and C_I. This concept has been used among others by Huang [25], to derive a mobility model. Here, the surface-roughness and phonon scattering mobility contributions are modeled using the effective field (1.11). For the calculation of the mobility, the charges Q'_B and Q'_I have to be extracted. This cannot be included into the TCAD concept straight forwardly, since for the calculation of Q'_B and Q'_I, integration of the charge along the normal direction to the interface is necessary. The estimated minority carrier

mobility is the same along this cut. Apart from the necessity of an additional preprocessing step to extract the charges, the ability of estimating the mobility for every location independently is lost. This approach also results in numerical difficulties since the integration introduces a lot of dependencies in the equation system, which leads to a poor solver performance. Approaches based on the effective vertical field are therefore used rather for compact modeling than for device simulation.

Up to this point, the discussion on mobility did not consider the electric field in the direction of the current flow and are therefore also called low-field mobility models. However, the carrier mobility strongly depends on the distribution function. Since the distribution function used in the transport models is strongly simplified and the detailed shape is not available, models have to be based on other quantities. In the case of the drift–diffusion model the electric field is commonly used and models are therefore called high-field mobility models. Simulation tools often differ between low- and high-field mobility and let the user select the models independently. Modeling of the high-field mobility dependence is often accomplished using the approach presented by Caughey and Thomas [26]. A slightly different version, suggested by Jaggi [27, 28], is used in the MINIMOS model:

$$\mu_\nu^{\mathrm{LISF}} = \frac{2\mu_\mu^{\mathrm{LIS}}}{1 + \left(1 + \left(\dfrac{2\mu_\mu^{\mathrm{LIS}}|\mathbf{F}_\nu|}{v_\nu^{\mathrm{sat}}}\right)^{\beta_\nu}\right)^{1/\beta_\nu}}. \tag{1.13}$$

$\mathbf{F}_\nu$ describes the driving force, which is the gradient of the quasi-Fermi level, v_ν^{sat} is the saturation velocity, and the coefficient β_ν is 1 for holes and 2 for electrons. The impact of this field dependence can be seen in the mobility distribution of the sample device shown in Fig. 1.5.

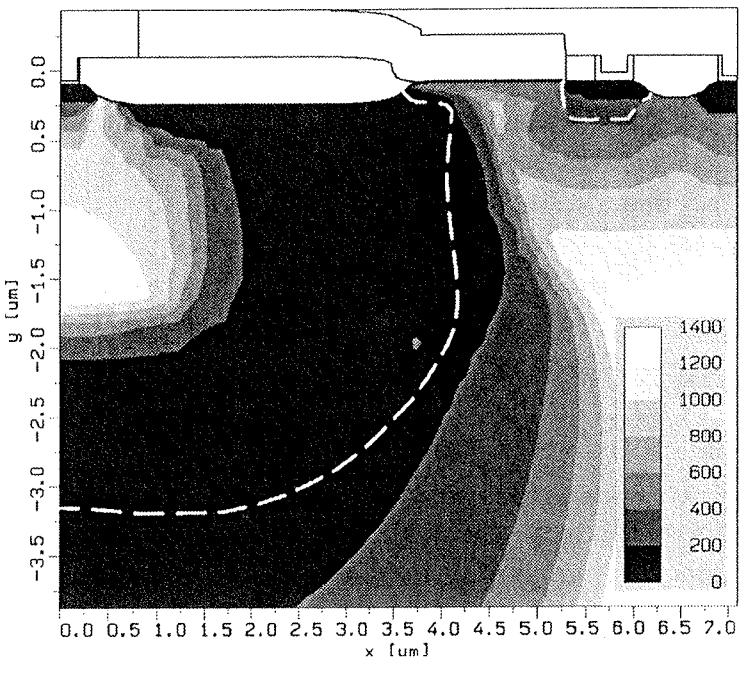

Fig. 1.5 Electron mobility distribution (in cm^2/Vs) in the sample device considering additionally to the lattice, impurity, and surface scattering models a field dependence model

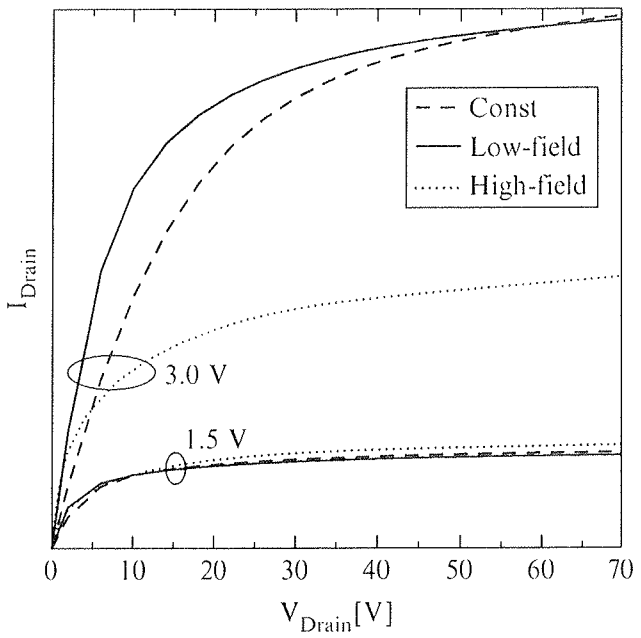

Fig. 1.6 Output characteristic of the sample device for different gate voltages using a constant mobility, a low-field mobility, and a high-field mobility model (1.13). For visualization the curves were scaled along the current axis so that the output for the gate voltage of 1.5 V overlaps

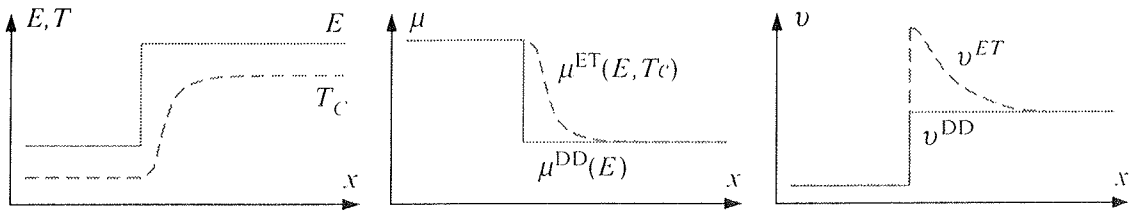

Fig. 1.7 Carriers traversing an abruptly changing electric field do not gain the resulting energy immediately and the carrier temperature therefore increases with a delay (*left*). High-field mobility models for drift–diffusion are based on the electric field and therefore react instantly to changes of the electric field while models based on the carrier temperature capture that delay (*center*). As a consequence the velocity overshoot ($v = \mu E$) cannot be observed in drift–diffusion (*right*)

To illustrate the impact of the high-field mobility a comparison of simulation results with and without active high-field mobility are shown in Fig. 1.6. Qualitative and quantitative differences can be seen. Both the constant mobility and the low-field mobility model only depend on fixed quantities so that the mobility does not change with the operating point which results in a similar transconductance. The fact that the low-field dependent model shown has spacial varying mobility values leads to the change of the shape of the output characteristic compared to the constant mobility model. The high-field mobility model leads to a considerable reduction of the mobility with increasing fields and therefore to a substantial reduction of the output current for increasing gate voltages.

Carrier mobility modeling has been investigated since the beginning of semiconductor engineering and there are still new models published. However, all models in drift–diffusion incorporating the influence of carriers that are not in thermal equilibrium basically rely on the electric field. Changes in the electric field therefore directly change the calculated mobility (see Fig. 1.7), whereas the carrier temperature does not increase immediately. Mobility models in energy-transport which

are based on the carrier temperature capture this effect. As a consequence velocity overshoot can be observed. In larger device structures, such as the high-voltage devices discussed here, this effect is normally unimportant and can be neglected.

2.3.2 Carrier Generation and Recombination

The recombination rate R was formally introduced in the drift–diffusion equations by splitting the continuity equation into two individual equations for electrons and holes. From a physical point of view this term includes the generation and the recombination of electron-hole pairs. In thermal equilibrium carrier generation and recombination are balanced and the carrier concentrations are given by their equilibrium values n_0 and p_0 ($n_0 p_0 = n_i^2$). The net recombination rate therefore vanishes. An excess number of carriers leads to an increased recombination, a low carrier concentration leads to an increased generation. The generation and recombination processes contributing to the total effective net generation rate are based on different physical effects which are modeled independently. Each model is evaluated separately and the total net recombination rate is calculated by adding the individual rates. The resulting rate is used to complete the continuity equations (1.4) and (1.5).

One important generation/recombination process is the well-known Shockley–Read–Hall (SRH) mechanism [29,30] which describes a two-step phonon transition. One trap level which is energetically located within the band-gap is utilized. There are four partial processes considered: the capture and the emission of both, electrons and holes, on the trap level. Balance equations can be formulated for the trap occupancy function. In the stationary case the rates for electrons and holes are equal. The trap occupancy function can be eliminated and the SRH generation rate can be calculated using

$$R^{\mathrm{SRH}} = \frac{np - n_i^2}{\tau_p (n + n_1) + \tau_p (p + p_1)}. \tag{1.14}$$

n_1 and p_1 are auxiliary concentrations depending on the energy level of the traps, τ_n and τ_p are carrier lifetimes for electrons and holes.

In MOS devices SRH generation especially influences the bulk current. In an n-channel device, for example, holes generated at the pn-junction are attracted by the low bulk potential which leads to an increased bulk current. This can be easily observed in device simulation since models can be switched on or off allowing to deactivate SRH. Figure 1.8 shows the hole current flow and the SRH generation rate in the sample device and in Fig. 1.9 the current components on the bulk contact are compared with and without SRH enabled.

The SRH model is not restricted to the description of capture and emission of carriers in the bulk, it can also be extended to determine the occupancy of interface traps [31]. Modeling of interface states is especially important for reliability modeling in MOS devices. Simulations of charge pumping measurements [32], for example, which are used to determine interface trap distributions, require appropriate modeling of trapping and de-trapping effects of carriers in interface traps. A simulation replicates the measurement procedure, by performing a transient

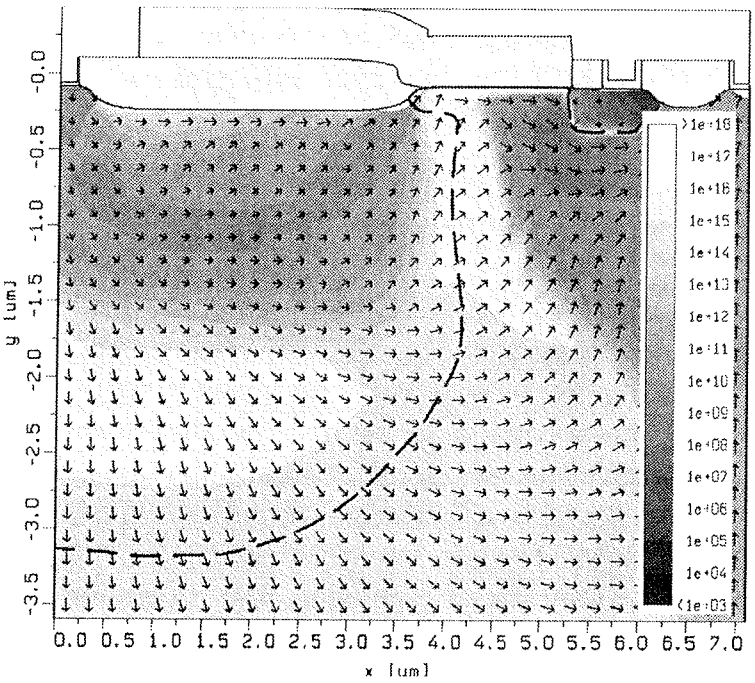

Fig. 1.8 SRH generation rate (in $s^{-1}\,cm^{-3}$) in the sample device with a drain voltage of 50 V and a gate voltage of 2 V. The *arrows* show the hole current flow

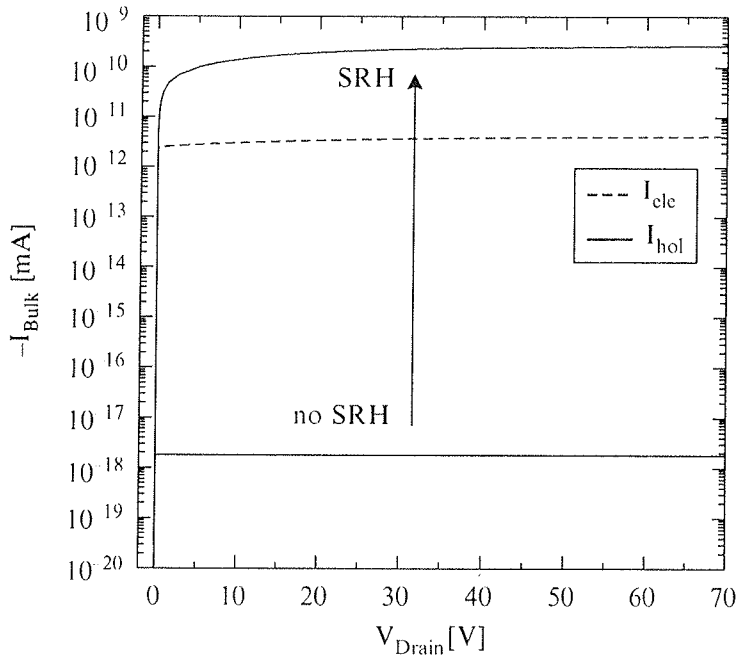

Fig. 1.9 Bulk current for the sample device using a gate voltage of 2 V with and without the SRH model activated. While the electron current stays the same, the hole current increases due to the SRH generated holes in the space charge region

simulation for every gate pulse level (Fig. 1.10). In contrast to the stationary SRH formulation shown in (1.14), time dependent simulations require to capture the transient behavior of the occupancy function [33]. The charge pumping curve can be constructed by extracting the mean current of the simulations for every single gate pulse (Fig. 1.11).

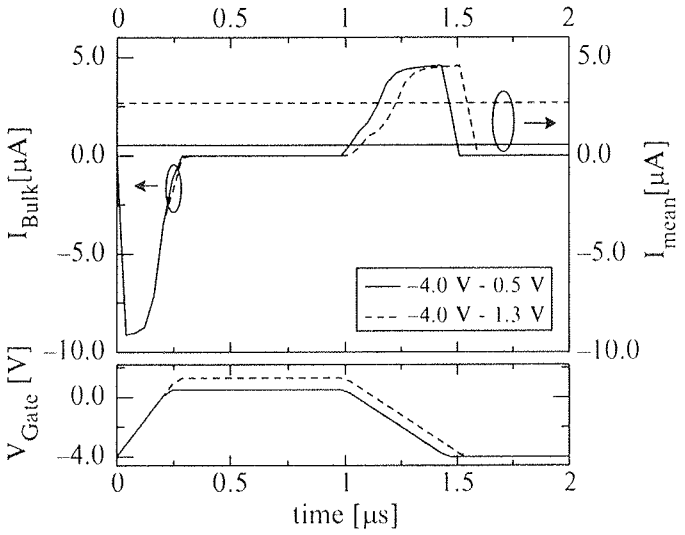

Fig. 1.10 Bulk current during two different gate voltage pulses. The mean currents are the charge pumping currents which are transferred to the charge pumping curve in Fig. 1.11

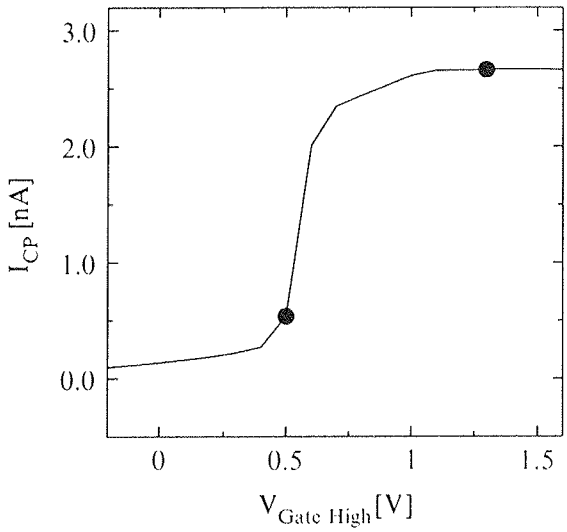

Fig. 1.11 Charge pumping curve of the sample device using gate pulses with constant low level of −4 V and changing high level. Values taken from Fig. 1.10 are highlighted

Beside the two-particle SRH mechanism there are two important three-particle generation/recombination mechanisms to mention: the Auger and the impact-ionization process, whereas the latter is a pure generation process. There, the energy required for generation is delivered by a third high-energetic electron or hole. The excess energy available after a recombination process is transferred to a third electron or hole. Modeling of this process can be achieved by defining rates for each partial process. Assuming a stationary case the rate evaluates to [3]

$$R^{\mathrm{AUG}} = \left(nC_n^{\mathrm{AUG}} + pC_p^{\mathrm{AUG}}\right)\left(np - n_i^2\right). \tag{1.15}$$

The coefficients C_n^{AUG} and C_p^{AUG} have only a weak dependence on the temperature [34] and are therefore often assumed to be constant, however, also temperature dependent models have been suggested [35].

The second three-particle mechanism mentioned, the impact-ionization carrier generation, describes electrons or holes which gain high energy from the electric field when traveling through the semiconductor. A scattering event between such

a high-energetic carrier and an electron or hole in the valence or conduction band, respectively, creates a new electron-hole pair. Impact-ionization is a pure generation process. Since there is no information on the carrier energy in the drift–diffusion model, as a compromise modeling is usually based on the current densities and on ionization coefficients which depend on the electric field. Many different approaches to model the coefficients have been proposed. Most device simulation tools [1, 36] include variations of the proposal of Chynoweth [37]:

$$G^{\mathrm{II}} = -\alpha_n(\mathbf{E})\frac{|\mathbf{J}_n|}{q} - \alpha_p(\mathbf{E})\frac{|\mathbf{J}_p|}{q} \tag{1.16}$$

A commonly used formulation applies as parameter the electric field in the direction of the current flow (1.17).

$$\alpha_\nu(\mathbf{E}, \mathbf{J}_\nu) = \alpha_\nu^\infty \exp\left(-\left(\frac{E_\nu^{\mathrm{crit}}\,|\mathbf{J}_\nu|}{\mathbf{E}\cdot\mathbf{J}_\nu}\right)^{\beta_\nu}\right) \tag{1.17}$$

α_ν^∞ is the temperature dependent maximum generation rate for high-fields. TCAD simulations using this model help to locate areas with high generation rates, see Fig. 1.12, and show the impact on the output characteristics Fig. 1.13.

Lower impact ionization rates have been observed for surface near currents [38] and models have been developed describing the transition between surface and bulk impact ionization [39]. However, Monte Carlo simulations have shown that there are no or only minor differences between surface and bulk impact ionization [40].

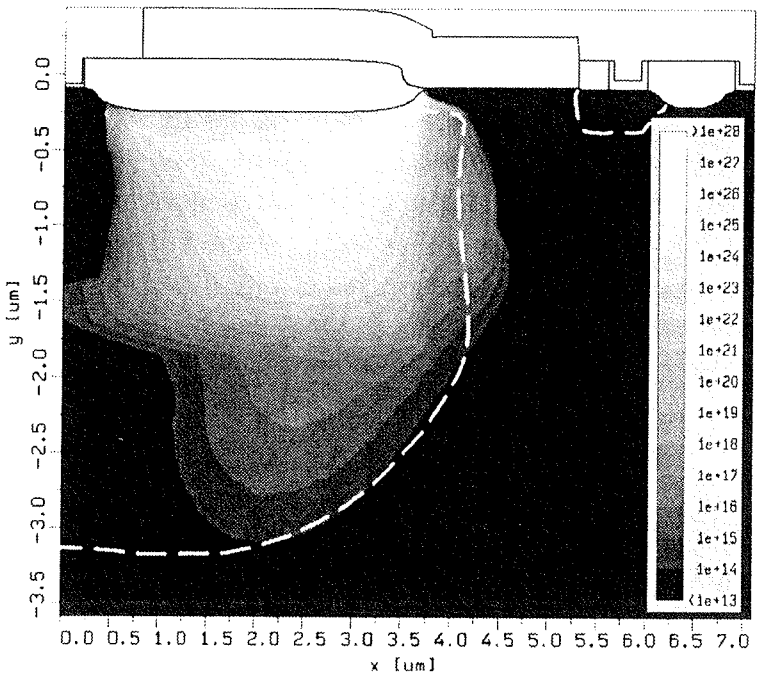

Fig. 1.12 Impact ionization generation rate in the sample device with a drain voltage of 40 V and a gate voltage of 2 V

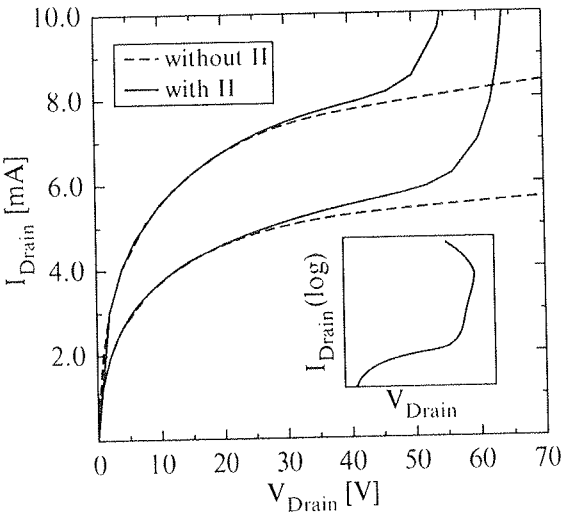

Fig. 1.13 Comparison of the output characteristic with and without impact-ionization. The inset shows a simulated snap-back curve

This means that there is no physical evidence of different rates near the surface and that these models are based on artifacts resulting from the approximate ionization rates based on the electric field.

Concerning the impact-ionization model presented, one has to note that carriers do not gain the energy from the electric field instantly. From a physical point of view the dependence of the impact-ionization rate on the local electric field is therefore not correct and only valid under homogenous conditions. A rigorous modeling should be based on the energy distribution function of the carriers, since only the high energetic carriers are relevant for impact-ionization [41]. Modeling of this high energy tail requires higher order transport models like the six moments model (see Fig. 1.3). Since the average carrier temperature which is available in the hydrodynamic model gives no information on the shape of the high energy tail, models based on this average temperature often overestimate the actual rate (compare Fig. 1.14). However, for larger structures the electric field based approximations give good results and can be used in the drift–diffusion model. As empirical alternatives to models based on the high energy tail, non-local impact-ionization models for the drift–diffusion model have been suggested [42].

Beside the generation and recombination mechanisms discussed here, many additional mechanisms are possible, which are more important for other device types. Some of the typical mechanisms considered in modeling include direct recombination which is important for direct bandgap semiconductors, direct [44] and trap assisted [45] band-to-band tunneling in high field regions, and optical generation [3].

2.4 *Thermal Modeling*

Physical properties of semiconductor devices strongly depend on the lattice temperature. Due to self-heating effects in the devices and due to changing ambient

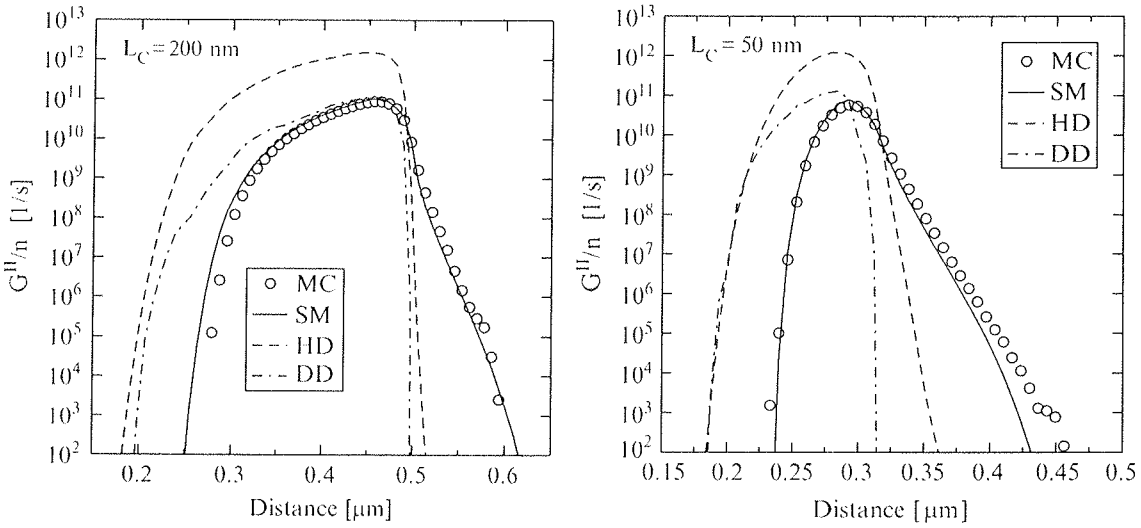

Fig. 1.14 Impact ionization rates in two comparable n^+-n-n^+ structures with a channel length of 200 nm (*left*) and 50 nm (*right*) [43]. The rates are calculated using the drift–diffusion (DD), hydrodynamic (HD), and six moments (SM) models. The Monte Carlo (MC) date serves as reference solution. Above 200 nm channel length the empirical drift–diffusion approximation fits reasonably well, whereas the hydrodynamic solution overestimates the generation rate. Only the six moments model is able to reproduce the impact-ionization rate in short channel devices

temperatures, the temperature distribution within a device is needed to estimate the device behavior at operating conditions. Modeling of the temperature requires some reference temperature, which might be a heat sink or the ambient temperature. The boundary condition for the temperature is commonly modeled with a thermal resistance to this fixed reference temperature. An important issue in thermal simulations is the size of the simulation domain. The heat flow in a semiconductor device extends to areas that are electrically less important and the simulation domain usually has to be extended compared to iso-thermal simulations.

The lattice temperature distribution T_L is modeled using the heat conduction equation [7]

$$c_\mathrm{tot}\frac{\partial T_\mathrm{L}}{\partial t} = \nabla\cdot(\kappa_\mathrm{tot}\nabla T_\mathrm{L}) + H, \tag{1.18}$$

where c_tot is the total heat capacity and κ_tot the total thermal conductivity. Both parameters include contributions from the lattice, the electron, and the hole subsystem. The temperature differences in the lattice lead to an additional driving force on the carriers which has to be considered in the current equations. For a non iso-thermal simulation, the last terms in (1.6) and (1.7) have to be considered in the solution of the drift–diffusion model.

Different approaches of modeling the heat generation rate H have been proposed. The most simple approach considers only the Joule heat $\mathbf{J}\cdot\mathbf{E}$ [46]. A more complex model according to Adler [47] describes the generated heat using

$$H = \mathbf{J}_n\cdot\nabla\frac{E_\mathrm{c}}{\mathrm{q}} + \mathbf{J}_p\cdot\nabla\frac{E_\mathrm{v}}{\mathrm{q}} + R\,(E_\mathrm{c}-E_\mathrm{v})\,. \tag{1.19}$$

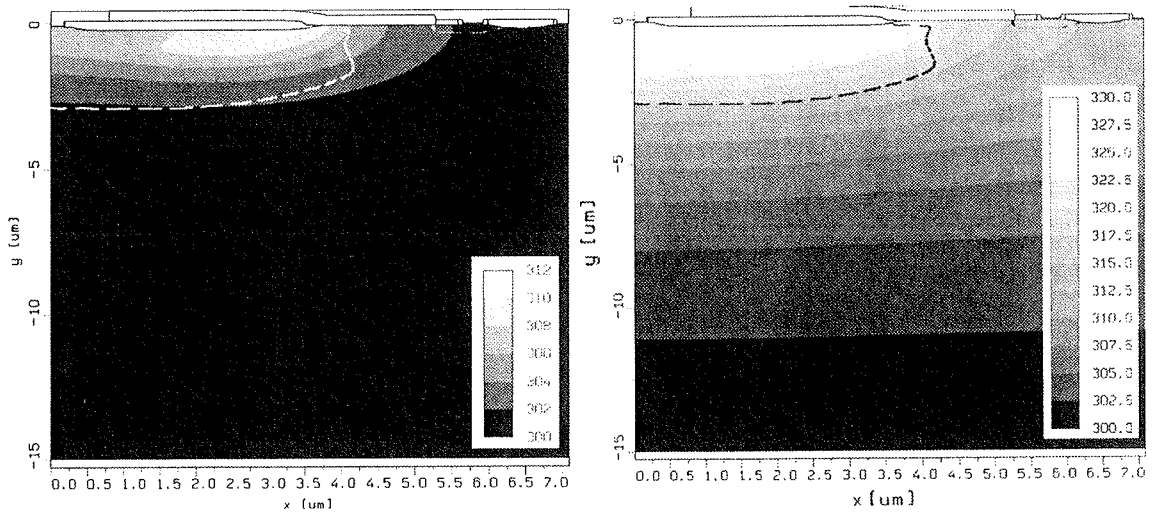

Fig. 1.15 Temperature distribution in Kelvin in the sample device with the gate biased to 2 V at the end of a linearly increase of the drain voltage from 0 to 50 V in 50 ns (*left*) and 500 ns (*right*)

Here, the energy dissipation due to recombination is considered. A more rigorous approach to thermal generation is given by Wachutka [7].

Transient simulations including thermal modeling were performed using the sample device. The lower bulk contact is linked with a thermal resistance to the ambient temperature of 300 K. In this simulation, the drain voltage is raised linearly from 0 to 50 V using two different slopes. The temperature distributions at the end of the two voltage ramps are depicted in Fig. 1.15. At the end of the 50 ns slope a rapid increase of the temperature near the birds beak can be observed.

2.5 *Additional Physical Effects*

In addition to the physical mechanisms addressed so far, there are obviously many other relevant modeling issues for semiconductor devices. For most of them, well established approaches are available in TCAD device simulation environments. Band-structure physics, for example, requires modeling for the bandgap energy and bandgap narrowing [48]. At low temperatures, incomplete ionization becomes important [49]. Also, semiconductor-metal contacts require appropriate treatment. The most common models for that include the well-known ohmic contact model where charge neutrality and equilibrium are assumed at the electrodes [3] and the Schottky contact models [50].

Especially in highly down-scaled MOS devices, tunneling and quantum effects have to be considered. For direct tunneling, which is most interesting for thin oxides, typically the Tsu–Esaki [51] or the Fowler–Nordheim [52] models are used. Herrmann and Schenk [53] proposed models for trap assisted tunneling, which has also been extended to multi-trap assisted tunneling models [54], especially interesting for highly degraded devices.

The inclusion of quantum confinement effects becomes especially important in SOI or double gate devices [55]. One modeling proposal is the modified local density approach [56] which is used in the model of Hänsch [57]. Here, a local correction of the effective density of states near the gate oxide is used to contribute to the quantum effects. An empirical correction approach has been presented by Van Dort et al. [58] which models the quantum confinement by increasing the band-gap near the interface.

3 Numerical Issues

In TCAD environments, the physical properties of a semiconductor device are described using models based on differential equation systems. Generally, the problems cannot be solved explicitly and numerical solution techniques are required. The system is solved at discrete points in space which are represented in terms of a mesh. The differential equations are then approximated using difference equations which can be solved using iterative solution techniques, typically based on the Newton method. Solving the transport equations together with the equations of the models for the mobility and for generation/recombination in a self-consistent way is a very complex task and requires considerable computational power. Since also the accuracy of the solution has to be considered, a proper trade off between accuracy, solvability, and simulation time has to be found.

3.1 Meshing

The points in space on which solutions for quantities are calculated are represented using a mesh consisting of nodes, edges, and elements. Quantities like the electrostatic potential or the electron and hole carrier concentrations are assigned to nodes. Fluxes like the electron or hole current are modeled along edges between the nodes. The structure of the real device obviously expands in three dimensions. However, to decrease the complexity of the problem, most applications can be reduced to a two-dimensional problem by assuming an infinite homogenous extension in the third dimension, the width of the device. This can be done if the fields and currents along the omitted dimension can be neglected, which is often possible for MOS devices. This simplification does not account for inhomogeneities along the width, effects near corners or changes in the doping profile at the border of the device are neglected. Considering that the width of the example device in this chapter is much larger than the length, most border effects are of minor importance. Further reduction of the problem complexity can be accomplished by utilizing symmetries in the device as much as possible without loosing information.

The simplest meshes used in device simulation are orthogonal grids that consist of mesh lines aligned parallel to the rectangular simulation domain (see Fig. 1.16).

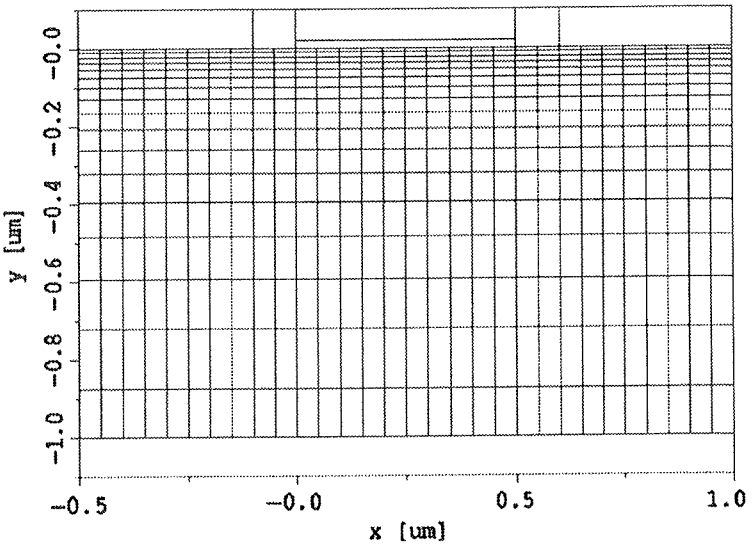

Fig. 1.16 Simple orthogonal grid of a planar MOS field effect transistor using a finer mesh near the channel

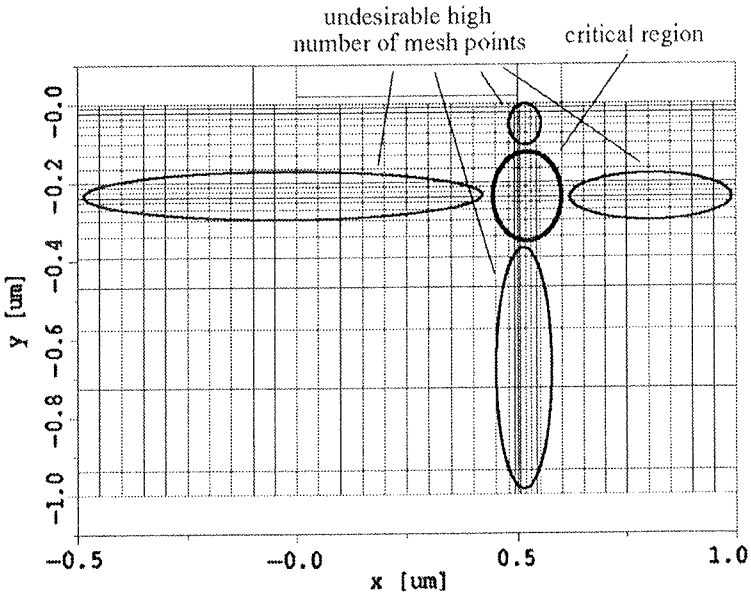

Fig. 1.17 Orthogonal grid with a refinement near the drain side of the channel. The number of grid points rises dramatically

This approach is easy to handle and not much effort is required to generate this type of mesh. In areas which require a higher spatial resolution a higher density of grid lines can be inserted. Adding grid lines is straight-forward but since grid lines are continued throughout the whole device, a high number of unnecessary mesh points in areas of low interest are created (see Fig. 1.17). This leads to long simulation times and poor convergence. Additionally, the rectangularly aligned mesh lines do not permit a smooth representation of non-planar surfaces (see Fig. 1.18), which is another major disadvantages of orthogonal grids.

Application of rectangular triangular elements allows to overcome the disadvantages mentioned above. The mesh used for the simulations on the sample device is shown in Fig. 1.19. It was created on the basis of an orthogonal mesh with refinements in areas of special interest. However, generation of triangular meshes suitable for device simulation is a cumbersome task. The box discretization (see Section 3.2) requires the mesh to fulfill the Delaunay criterion [59]. This criterion

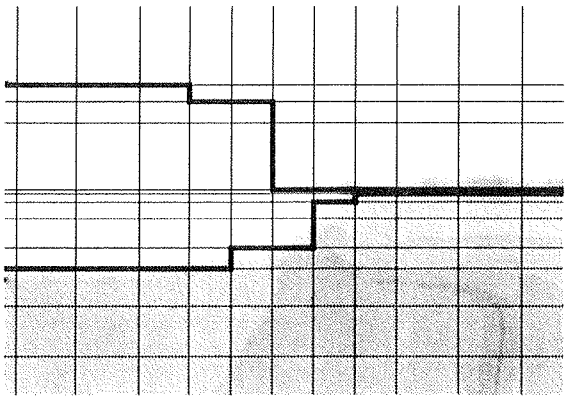

Fig. 1.18 Modeling of non-planar surfaces using orthogonal grids is not satisfying

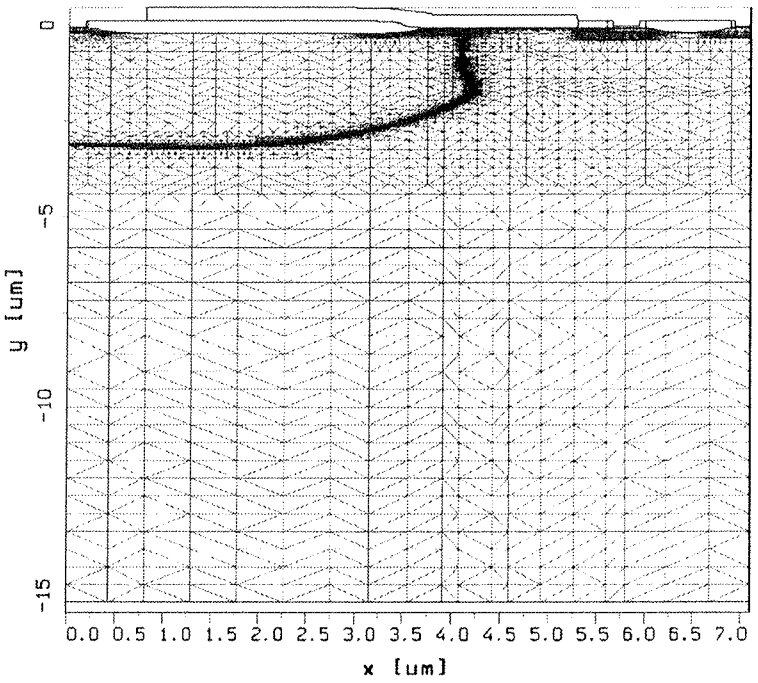

Fig. 1.19 The mesh of the sample device is based on an orthogonal grid which is triangulated and refined. Areas of special interest are the channel and the junction regions which thus require a denser mesh

describes triangular meshes constructed for a set of points such that no point is positioned inside the circumcircle of any triangle (see Fig. 1.20). This is always valid for triangulated rectangular grids but can be difficult to obtain in general. Other obstacles during mesh generation are the proper representation of borders and surfaces and the definition of mesh refinement criteria for areas of special interest. The details of the mesh refinement procedure often have to be specified manually, since mesh generation tools have no a priori information of the device behavior. Also the orientation of the elements to the current direction has to be considered, which is especially relevant in the channel region of MOS devices and additionally depends on the bias condition. Since fully automatic mesh generation would clearly simplify the work-flow, considerable efforts have been put into the development of

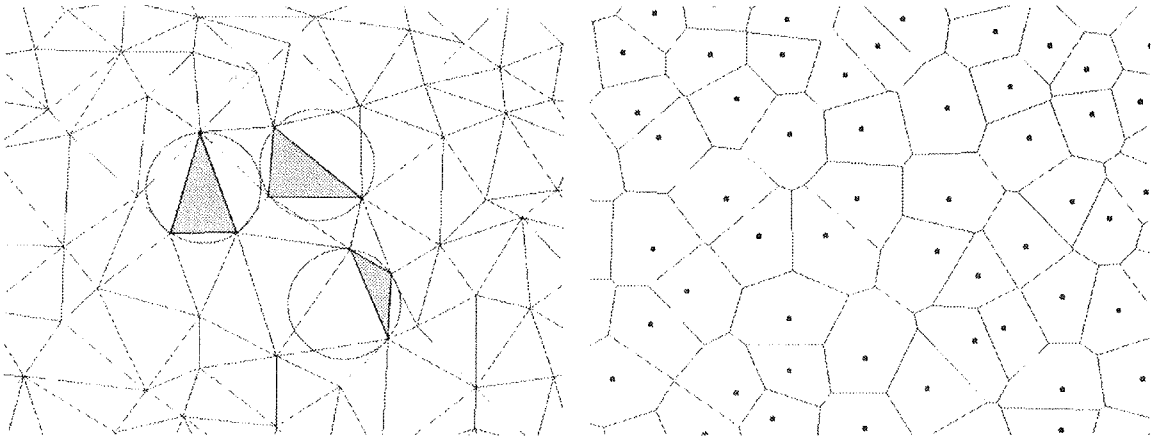

Fig. 1.20 A Delaunay mesh is shown at the left and the Voronoi tessellation of that mesh at the *right* (created using [61]). The circumcircles of three randomly selected triangles are emphasized. Every mesh point has an associated volume – every point in the domain has exactly one associated box volume

methods for automatic mesh generation. One approach applies error estimation algorithms on simulation results [60]. This information is recursively used to optimize the simulation mesh.

3.2 *Discretization*

The discretization of the partial differential semiconductor equations in space and time is needed to obtain difference equations which can be solved using numerical methods. A common approach for discretization of the differential equations is the box integration method [62, 63], also known as the finite volume method. For this the mesh has to fulfill the Delaunay criterion and can therefore be split into boxes using a Voronoi tessellation [59]. Doing this, every point in the domain is assigned to its closest mesh point and is therefore inside the box volume of that point (see Fig. 1.20). The basic method of the box discretization and how it is applied to the divergence operator is illustrated in the following using Poisson's equation,

$$\nabla \cdot \mathbf{D} = \rho, \tag{1.20}$$

where $\mathbf{D} = -\epsilon \nabla \psi$ and ϵ is considered to be a scalar and homogeneous. The equation can be transformed by integration over a volume V and by applying Gauss' law. With the assumption that ρ is constant inside the volume, one can write

$$\oint_{\partial V} \mathbf{D}\,\mathrm{d}\mathbf{A} = \rho V, \tag{1.21}$$

with $\mathbf{A}$ being the outwardly oriented surface area ∂V. The assumption that ρ is constant is very crude but is often used due to its simplicity [63, 64]. Other approaches

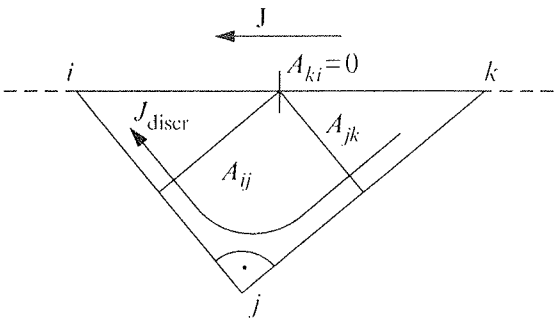

Fig. 1.21 A *triangle* aligned with its hypotenuse along the current direction. Due to the vanishing Voronoi surface A_{kj} there is no contribution along the current by the hypotenuse

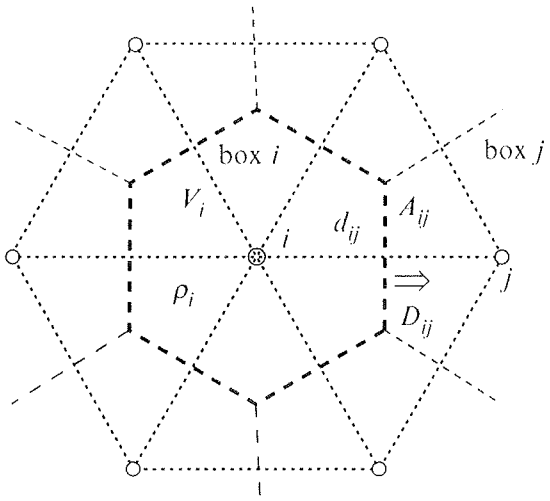

Fig. 1.22 Voronoi box i of mesh point i with connections to its neighboring mesh points. The flux D_{ij} from box i to box j through the area A_{ij} is depicted. ρ is assumed constant over the whole box volume V

estimate ρ element-wise which results in an implicitly increased resolution and a more accurate physical representation [65]. Equation 1.21 can be applied to each Voronoi box in the mesh (see Fig. 1.22). Since the boundary ∂V is a polygon (in 2D) or polyhedron (in 3D) it can be split into planar surface elements normal to the edges leading to the neighboring mesh points. For a mesh point i and its box volume V_i which has N_i neighboring boxes j, Eq. (1.21) can be approximated to

$$\sum_{j \in N_i} D_{ij} A_{ij} = \rho_i V_i . \tag{1.22}$$

D_{ij} is the dielectric flux density from box i to box j along the connecting edge through the common boundary area A_{ij}. The dielectric flux density D_{ij} can be approximated using the directional derivative of the electrostatic potential:

$$D_{ij} = \boldsymbol{\xi}_{ij} \cdot \mathbf{D} = -\epsilon \, \boldsymbol{\xi}_{ij} \cdot \nabla \psi = -\epsilon \frac{\partial \psi}{\partial \boldsymbol{\xi}_{ij}} \approx -\epsilon \frac{\psi_j - \psi_i}{d_{ij}} . \tag{1.23}$$

$\boldsymbol{\xi}_{ij}$ is the unit vector pointing from mesh points i to j and d_{ij} is the distance between the two mesh points. Discretization of the current continuity equations (1.4) and (1.5) is accomplished similarly. In the static case, at the left hand side of the discretized formulation (1.22) J_{ij} is used instead of D_{ij} and at the right hand side the

charge density ρ is replaced by the generation rate R. If the drift term dominates over the diffusion term numerical oscillation can result when a simple finite difference approach like in (1.23) is used [66]. Very fine meshes would be necessary to stabilize the system. However, a stable discretization can be obtained using the Scharfetter–Gummel method [67] instead of finite differences. Here, the drift–diffusion current equations (1.6) and (1.7) are used to solve the one-dimensional carrier concentration along the edge. The boundary conditions of the carrier concentration are given using the known values at the mesh points. The values of $J_{n,ij}$ and $J_{p,ij}$, E_{ij}, and μ_n and μ_p are considered constant along the edge. Solving this one-dimensional differential equation results in

$$J_{n,ij} = \frac{q\mu_n V_T}{d_{ij}} \left(n_j \mathcal{B}(\Delta_{ij}) - n_i \mathcal{B}(-\Delta_{ij}) \right) \tag{1.24}$$

for electrons and

$$J_{p,ij} = -\frac{q\mu_p V_T}{d_{ij}} \left(p_j \mathcal{B}(-\Delta_{ij}) - p_i \mathcal{B}(\Delta_{ij}) \right) \tag{1.25}$$

for holes, where $\Delta_{ij} = \left(\psi_j - \psi_i\right) / V_T$ and $\mathcal{B}$ is the Bernoulli function.

The box method is used in most device simulation environments as it has proven to deliver good results and is simple to implement. Problems arise when the Delaunay criterion is violated and obtuse elements degenerate the accuracy due to negative flux areas A_{ij} [68, 69]. Also, use of the one-dimensional Scharfetter–Gummel discretization to solve multiple dimensional problems leads to the crosswind diffusion effect resulting in artificial current components normal to the actual current direction [70]. The accuracy of the discretization also degrades if triangles are aligned with the hypotenuse along the current flow. As depicted in Fig. 1.21, a vanishing boundary area A_{ki} leads according to (1.22) to a vanishing contribution of the current along this edge. A zig-zag characteristic of the discretized current is the result. There have been many proposals for more accurate discretizations (e.g. Patil in [68]). Some focus on the extension of the one-dimensional to a two-dimensional Scharfetter–Gummel current discretization [71, 72]. But none of these extensions is as universal to use as the box integration method which is dimension independent and can be used for rectangular and triangular meshes alike.

3.3 *Vectors in Discretized Systems*

Some physical models such as the impact-ionization rate and the high-field mobility are approximate in the drift–diffusion framework as depending on vector quantities. Impact ionization generation models depend on the electric field projected on the current density vector. Within the box method the discretization of two or three-dimensional vector quantities is based on the one-dimensional discretization along the edges ($J_{ij} = \mathbf{J} \cdot \boldsymbol{\xi}_{ij}$). Since models that are field dependent are especially relevant

for high-voltage operation conditions, break-down and snap-back simulation results are influenced by the selected vector discretization approach.

As previously discussed, the calculation of the charge ρ in the discretized Poisson equation (1.22) can be calculated per box or per element [64]. The same holds for the generation rate in the continuity equation where the right hand side reads $R_i V_i$. The rate calculation for the rate per element approach requires discretized field vectors per element. For a constant electric field in a triangular mesh, this can be accomplished exactly by a linear combination of two of the three one-dimensional edge contributions. However, for non-constant fields or for a non-linear discretization like the Scharfetter–Gummel discretization, each edge combination delivers a different result and approximations have to be made. There are different approaches how to calculate the rates and vectors for each element. One approach is the edge pair method [73], other approaches calculate the rate for every edge in the element individually [74].

For the box based approach which corresponds to the formulation in (1.22), the calculation of rates per box requires vector quantities discretized for every box. The advantage of such an approach is that a single generation rate can be evaluated per box and the continuity equation can be solved directly as shown in (1.22). A scheme to estimate vector quantities per box has been presented in [64]. This scheme has the advantage that the same implementation is suitable for triangular and orthogonal meshes alike and can thus be used for two and three-dimensional problems.

3.4 Numerical Challenges Related to HV Devices

Application of the box discretization method to the drift–diffusion model and utilizing the Scharfetter–Gummel method for the current equations is an established method used in most TCAD simulation environments [1, 36, 65]. Together with the Newton procedure for solving the equation system, a numerically stable simulation environment can be built.

Numerical challenges usually originate from the models used for the mobility and for the generation rates. Especially field dependent models, like the high-field mobility (1.13) or the impact ionization rate (1.17) may lead to convergence difficulties. The dependence on vector quantities, especially the electric field, results in couplings to many neighboring mesh points which generate many dependencies in the Jacobian matrix. The high number of non-zero entries in the system matrix leads to poor performance or failure of iterative linear solvers [75]. Additionally, strong non-linear relations, for example the exponential dependence used to model the impact ionization rate, might lead to poor convergence. Many of the numerically problematic models are important in high-voltage and power devices and therefore have to be considered in the simulation.

An approach to overcome convergence issues is to calculate more intermediate solutions between the initial simulation step (equilibrium) and the desired operating point. This can be achieved by ramping up the contact potentials step by step until

the final value is achieved. Results from former steps can be used as initial guesses for the Newton method in the next step. Decreasing the step size therefore improves the quality of the initial guess for the next step and finally for the desired operating point. Obviously a good balance between step size, robustness, and simulation time has to be found. Other approaches to overcome convergence issues are to tune the Newton procedure, for example, by changing parameters of the damping algorithm [76]. This changes the calculation of the Newton update vectors and is often required to achieve convergence.

Numerical problems are frequently caused by the simulation mesh used. For mesh design a trade-off between accuracy and numerical stability has to be found. High convergence rates can be achieved having a moderate number of well shaped elements [77]. However, smaller elements usually lead to a higher resolution and therefore to a higher accuracy. On the other hand, the limited numerical accuracy in computer systems can result in numerical noise [78] which degrades the condition of the system matrix. For accurate results, a proper alignment of the elements regarding the direction of the current flow (see Fig. 1.21) is also important. Near the channel area this often leads to poorly shaped thin elements having small internal angles. Creating a mesh for the simulation of high-voltage devices that has good numerical properties and delivers accurate results is very challenging.

Applications with special demands on the numerics of TCAD simulation tools are break-down [77], electro-static discharge (ESD) [79], and snap-back simulations. Difficulties arise from the strongly field dependent behavior and the physically unstable operating points. Physical quantities undergo strong variations near break-down and snap-back processes as the device changes its operation state. This also impacts the convergence of the device due to the state transition leading to strong changes in conductivity, current path, and carrier concentration. The different states of the devices before and after snap-back additionally result in multivalued I/V curves. The boundary conditions therefore have to be selected appropriately to avoid unintentional transitions between the branches of the I/V curve. Special curve-tracing algorithms have been suggested to deal with these problems [80, 81].

4 Conclusion

TCAD in general and numerical device simulation software in particular can be used as a very powerful tool for device engineering in academic as well as in industrial environments. The possibility of obtaining an insight into the device behavior results in a better understanding of the physics and enables the device designer to tune the device performance. Consequences of changing the device design can be tested without the time consuming and expensive fabrication of test devices. To increase the performance of special parameters, automatic optimization procedures can be set up. Optimization goals might be the maximum power output, the ratio between the on- and off-resistance, or the ratio between the maximum blocking voltage and the on-resistance.

It is important that the user of TCAD simulation tools has a fundamental knowledge of semiconductor device physics. Many different physical mechanisms have to be considered and for most of them a variety of models exist. Some models are derived to accurately describe special operation conditions, others are for multi purpose usage and cover a broad range of different operation conditions. The models also differ in the complexity leading to different simulation times. It is also important to consider that the impact of the individual physical effects are more or less important for different devices and operation conditions. Together with the understanding of how the simulation environment works, the device engineer can select the proper models and meaningful simulation results can be produced.

Acknowledgements The authors would like to thank the industrial partners Infineon Technologies, Villach, and austriamicrosystems, Graz, Austria. Special thanks go to Rainer Minixhofer and Jong Mun Park from austriamicrosystems for providing us with a state of the art high-voltage device structure.

References

1. IμE, *MINIMOS-NT 2.1 User's Guide*, Institut für Mikroelektronik, Technische Universität Wien, Austria (2004), http://www.iue.tuwien.ac.at/software/minimos-nt
2. J. Park, H. Enichlmair, R. Minixhofer, Hot-carrier behaviour of a 0.35 μm high-voltage n-channel LDMOS transistor, in *Proceedings of SISPAD*, 2007, pp. 369–372
3. S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer, Wien New York, 1984)
4. T. Grasser, T.-W. Tang, H. Kosina, S. Selberherr, A review of hydrodynamic and energy-transport models for semiconductor device simulation, Proc. IEEE **91**(2), 251–274 (2003)
5. C. Jacoboni, L. Reggiani, The Monte Carlo method for the solution of charge transport in semiconductors with application to covalent materials, Rev. Mod. Phys. **55**(3), 645–705 (1983)
6. M. Vecchi, M. Rudan, Modeling electron and hole transport with full-band structure effects by means of the spherical-harmonics expansion of the BTE, IEEE Trans. Electron. Devices **45**(1), 230–238 (1998)
7. G. Wachutka, Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling, IEEE Trans. Comput. Aid. Des. Integr. Circ. Syst. **9**(11), 1141–1149 (1990)
8. R. Stratton, Semiconductor current-flow equations (diffusion and degeneracy), IEEE Trans. Electron. Devices **19**(12), 1288–1292 (1972)
9. C. Jungemann, B. Meinerzhagen, *Hierarchical Device Simulation: The Monte-Carlo Perspective* (Springer, Wien–New York, 2003)
10. W. VanRoosbroeck, Theory of flow of electrons and holes in germanium and other semiconductors, Bell Syst. Tech. J. **29**, 560–607 (1950)
11. R. Stratton, Diffusion of hot and cold electrons in semiconductor barriers, Phys. Rev. **126**(6), 2002–2014 (1962)
12. K. Bløtekjær, Transport equations for electrons in two-valley semiconductors, IEEE Trans. Electron. Devices **17**(1), 38–47 (1970)
13. T. Grasser, S. Selberherr, Limitations of hydrodynamic and energy-transport models, in *Proceedings of 11th International Workshop on the Physics of Semiconductor Devices*, 2001, pp. 584–591
14. T. Grasser, H. Kosina, C. Heitzinger, S. Selberherr, Characterization of the hot electron distribution function using six moments, J. Appl. Phys. **91**(6), 3869–3879 (2002)

15. T. Grasser, C. Jungemann, H. Kosina, B. Meinerzhagen, S. Selberherr, Advanced transport models for sub-micrometer devices, in *Proc. SISPAD*, 2004, pp. 1–8
16. S. Selberherr, W. Hänsch, M. Seavey, J. Slotboom, The evolution of the MINIMOS mobility model, Solid-State Electron. **33**(11), pp. 1425–1436 (1990)
17. C. Lombardi, S. Manzini, A. Saporito, M. Vanzi, A physically based mobility model for numerical simulation of nonplanar devices, IEEE Trans. Comput. Aid. Des. Integr. Circ. Syst. **7**(11), 1164–1171 (1988)
18. J. Agostinelli, V.M., J. Agostinelli, V.M., H. Shin, J. Tasch, A.F., A comprehensive model for inversion layer hole mobility for simulation of submicrometer MOSFET's, IEEE Trans. Electron. Devices **38**(1), 151–159 (1991)
19. T. Grasser, R. Entner, O. Triebl, H. Enichlmair, R. Minixhofer, TCAD modeling of negative bias temperature instability, in *Proceedings of SISPAD*, 2006, pp. 330–333
20. V.-H. Chan, J. Chung, Two-stage hot-carrier degradation and its impact on submicrometer LDD NMOSFET lifetime prediction, IEEE Trans. Electron. Devices **42**(5), 957–962 (1995)
21. M. Darwish, J. Lentz, M. Pinto, P. Zeitzoff, T. Krutsick, H. H. Vuong, An improved electron and hole mobility model for general purpose device simulation, IEEE Trans. Electron. Devices **44**(9), 1529–1538 (1997)
22. B. Neinhüs, C. Nguyen, C. Jungemann, B. Meinerzhagen, A CPU efficient electron mobility model for MOSFET simulation with quantum corrected charge densities, in *Proceedings of the 30th European Solid-State Device Research Conference*, 2000, pp. 332–335
23. A. Sabnis, J. Clemens, Characterization of the electron mobility in the inverted <100> Si surface, in *Proceedings of International Electron Devices Meeting*, vol. 25, 1979, pp. 18–21
24. S. Sun, J. Plummer, Electron mobility in inversion and accumulation layers on thermally oxidized silicon surfaces, IEEE Trans. Electron. Devices **27**(8), 1497–1508 (1980)
25. C.-L. Huang, G. Gildenblat, Measurements and modeling of the n-channel MOSFET inversion layer mobility and device characteristics in the temperature range 60-300 K, IEEE Trans. Electron. Devices **37**(5), 1289–1300 (1990)
26. D. Caughey, R. Thomas, Carrier mobilities in silicon empirically related to doping and field, Proc. IEEE **55**(12), 2192–2193 (1967)
27. R. Jaggi, H. Weibel, High-field electron drift velocities and current densities in silicon, Helv. Phys. Acta **42**, 631–632 (1969)
28. R. Jaggi, High-field drift velocities in silicon and germanium, Helv. Phys. Acta **42**, 941–943 (1969)
29. W. Shockley, W. Read, Statistics of the recombinations of holes and electrons, Phys. Rev. **87**(5), 835–842 (1952)
30. R. Hall, Electron-hole recombination in germanium, Phys. Rev. **87**(2), 387 (1952)
31. D. Fleetwood, H. Xiong, Z.-Y. Lu, C. Nicklaw, J. Felix, R. Schrimpf, S. Pantelides, Unified model of hole trapping, 1/f noise, and thermally stimulated current in MOS devices, IEEE Trans. Nucl. Sci. **49**(6), 2674–2683 (2002)
32. J. Brugler, P. Jespers, Charge pumping in MOS devices, IEEE Trans. Electron. Devices **16**(3), 297–302 (1969)
33. P. Habas, S. Selberherr, A closed-loop extraction of the spatial distribution of interface traps based on numerical model of the charge-pumping experiment, in *Proceedings of SSDM 92 Conference*, 1992, pp. 170–172
34. J. Dziewior, W. Schmid, Auger coefficients for highly doped and highly excited silicon, Appl. Phys. Lett. **31**(5), 346–348 (1977)
35. L. Huldt, N.G. Nilsson, K.G. Svantesson, The temperature dependence of band-to-band auger recombination in silicon, Appl. Phys. Lett. **35**(10), 776–777 (1979)
36. *ATLAS User's Manual*, SILVACO International, 2004
37. A.G. Chynoweth, Ionization rates for electrons and holes in silicon, Phys. Rev. **109**(5), 1537–1540 (Mar 1958)
38. J. Slotboom, G. Streutker, G. Davids, P. Hartog, Surface impact ionization in silicon devices, in *Proceedings of International Electron Devices Meeting*, 1987, ed. by G. Streutker, vol. 33, pp. 494–497

39. M. van Dort, J. Slotboom, G. Streutker, P. Woerlee, Lifetime calculations of MOSFETs using depth-dependent non-local impact ionization, Microelectron. J. **26**, 301–305 (1995)
40. C. Jungemann, S. Yamaguchi, H. Goto, Is there experimental evidence for a difference between surface and bulk impact ionization in silicon? in *Proceedings of International Electron Devices Meeting*, ed. by S. Yamaguchi, 1996, pp. 383–386
41. T. Grasser, H. Kosina, C. Heitzinger, S. Selberherr, Accurate impact ionization model which accounts for hot and cold carrier populations, Appl. Phys. Lett. **80**(4), 613–615 (2002)
42. J. Slotboom, G. Streutker, M. van Dort, P. Woerlee, A. Pruijmboom, and D. Gravesteijn, Non-local impact ionization in silicon devices, in *Proceedings of International Electron Devices Meeting Technical Digest*, 8–11 Dec. 1991, pp. 127–130
43. T. Grasser, H. Kosina, S. Selberherr, Hot carrier effects within macroscopic transport models, Int. J. High Speed Electron. Syst. **13**(3), 873–901 (Sept 2003)
44. G. Hurkx, D. Klaassen, M. Knuvers, A new recombination model for device simulation including tunneling, IEEE Trans. Electron. Devices **39**(2), 331–338 (1992)
45. H.-M. Lee, C.-J. Liu, C.-W. Hsu, M.-S. Liang, Y.-C. King, C.-H. Hsu, New trap-assisted band-to-band tunneling induced gate current model for p-channel metal-oxide-semiconductor field effect transistors with sub-3 nm oxides, Jap. J. Appl. Phys. **40**, 1218–1221 (2001)
46. S. Gaur, D. Navon, Two-dimensional carrier flow in a transistor structure under nonisothermal conditions, IEEE Trans. Electron. Devices **23**(1), 50–57 (1976)
47. M. Adler, Accurate calculations of the forward drop and power dissipation in thyristors, IEEE Trans. Electron. Devices **25**(1), 16–22 (1978)
48. J. Slotboom, H. de Graaff, Bandgap narrowing in silicon bipolar transistors, IEEE Trans. Electron. Devices **24**(8), 1123–1125 (1977)
49. S. Selberherr, E. Langer, Low temperature MOS device modeling, in *Proceedings of Workshop on Low Temperature Semiconductor Electronics*, 1989, pp. 68–72
50. S. Sze, *Physics of Semiconductor Devices*, 2nd edn. (Wiley, New York, 1981)
51. R. Tsu, L. Esaki, Tunneling in a finite superlattice, Appl. Phys. Lett. **22**(11), 562–564 (1973)
52. R.H. Fowler, L. Nordheim, Electron emission in intense electric fields, Proc. Roy. Soc. A **119**(781), 173– 181 (1928)
53. M. Herrmann, A. Schenk, Field and high-temperature dependence of the long term charge loss in erasable programmable read only memories: Measurements and modeling, J. Appl. Phys. **77**(9), 4522–4540 (1995)
54. R. Entner, T. Grasser, S. Selberherr, A. Gehring, H. Kosina, Modeling of tunneling currents for highly degraded CMOS devices, in *Proceedings of SISPAD*, 2005, pp. 219–222
55. M. Wagner, M. Karner, T. Grasser, Quantum correction models for modern semiconductor devices, in *Proceedings of the XIII International Workshop on Semiconductor Devices*, 2005, pp. 458–461
56. G. Paasch, H. Übensee, A modified local density approximation, Phys. Stat. Sol. (b) **113**(1), 165–178 (1982)
57. W. Hänsch, T. Vogelsang, R. Kircher, M. Orlowski, Carrier transport near the Si/SiO_2 interface of a MOSFET, Solid-State Electron **32**(10), 839–849 (1989)
58. M. van Dort, P. Woerlee, A. Walker, A simple model for quantisation effects in heavily-doped silicon MOSFETs at inversion conditions, Solid-State Electron. **37**(3), 411–414 (1994)
59. A. Okabe, B. Boots, K. Sugihara, *Spatial Tessellations* (Wiley, New York, 1992)
60. M. Spevak, R. Heinzl, P. Schwaha, T. Grasser, Simulation of microelectronic structures using a posteriori error estimation and mesh optimization, in *5th Mathmod Vienna Proceedings*, Wien, Feb 2006, pp. 5.1–5.8
61. L.P. Chew, Create a Voronoi diagram or Delaunay triangulation by clicking points (2007), http://www.cs.cornell.edu/home/chew/Delaunay.html
62. R. Bank, D. Rose, W. Fichtner, Numerical methods for semiconductor device simulation, IEEE Trans. Electron. Devices **30**(9), 1031–1041 (1983)
63. K. Kramer, G. Nicholas, W. Hitchon, *Semiconductor Devices, a Simulation Approach.* (Prentice Hall, Professional Technical Reference, 1997)
64. O. Triebl, T. Grasser, Investigation of vector discretization schemes for box volume methods, in *NSTI-Nanotech*, vol. 3, 2007, pp. 61–64

65. *ISE TCAD Release 9.5 - DESSIS*, Synopsys, Inc., originally published by ISE, 2003
66. S.J. Polak, C. den Heijer, and W. Schilders, "Semiconductor device modelling from the numerical point of view," *Intl.J.Numer.Methods Eng.*, vol. 24, pp. 763–838, 1987.
67. D. Scharfetter, H. Gummel, Large-signal analysis of a silicon read diode oscillator, IEEE Trans. Electron. Devices **16**(1), 64–77 (1969)
68. M. Patil, New discretization scheme for two-dimensional semiconductor device simulation on triangular grid, IEEE Trans. Comput. Aid. Des. Integr. Circ. Syst. **17**(11), 1160–1165 (1998)
69. J.Bürgler, R. Bank, W. Fichtner, R. Smith, A new discretization scheme for the semiconductor current continuity equations, IEEE Trans. Comput. Aid. Des. Integr. Circ. Syst. **8**(5), 479–489 (1989)
70. Y. He, G. Cao, A generalized Scharfetter-Gummel method to eliminate crosswind effects [semiconduction device modeling], IEEE Trans. Comput. Aid. Des. Integr. Circ. Syst. **10**(12), 1579–1582 (1991)
71. W. Allegretto, A. Nathan, H. Baltes, Numerical analysis of magnetic-field-sensitive bipolar devices, IEEE Trans. Comput. Aid. Design Integr. Circ. Syst. **10**(4), 501–511 (1991)
72. H. Kosina, O. Triebl, T. Grasser, Box method for the convection-diffusion equation based on exponential shape functions, in *Proceedings of SISPAD*, vol. 12, ed. by T. Grasser, S. Selberherr, 2007, pp. 317–320
73. H. Shao, in *Numerical analysis of meshing and discretization for anisotropic convection-diffusion equations with applications*, Dissertation, Duke University, Aug 1999
74. S. Laux, B. Grossman, A general control-volume formulation for modeling impact ionization in semiconductor transport, IEEE Trans. Electron. Devices **32**(10), 2076–2082 (1985)
75. O. Schenk, M. Hagemann, S. Rollin, Recent advances in sparse linear solver technology for semiconductor device simulation matrices, in *Proceedings of SISPAD*, 2003, pp. 103–108
76. P. Deuflhard, A modified Newton method for the solution of ill-conditioned systems of nonlinear equations with application to multiple shooting, Numer. Math. **22**, 289–315 (1974)
77. V. Axelrad, Grid quality and its influence on accuracy and convergence in device simulation, IEEE Trans. Comput. Aid. Des. Integr. Circ. Syst. **17**(2), 149–157 (1998)
78. N. Shigyo, H. Tanimoto, T. Enda, Mesh related problems in device simulation: Treatments of meshing noise and leakage current, Solid-State Electron. **44**, 11–16 (2000)
79. C. Salaméro, N. Nolhier, A. Gendron, M. Bafleur, P. Besse, M. Zécri, TCAD methodology for ESD robustness prediction of smart power ESD devices, IEEE Trans. Device Mater. Rel. **6**(3), 399–407 (2006)
80. R. Goossens, S. Beebe, Z. Yu, R. Dutton, An automatic biasing scheme for tracing arbitrarily shaped I-V curves, IEEE Trans. Comput. Aid. Design Integr. Circ. Syst. **13**(3), 310–317 (1994)
81. M. Bartels, S. Decker, B. Neinhus, B. Meinerzhagen, A robust curve tracing scheme for the simulation of bipolar breakdown characteristics with nonlocal impact ionization models, in *Proceedings of the 29th European Solid-State Device Research Conference*, vol. 1, 1999, pp. 492–495