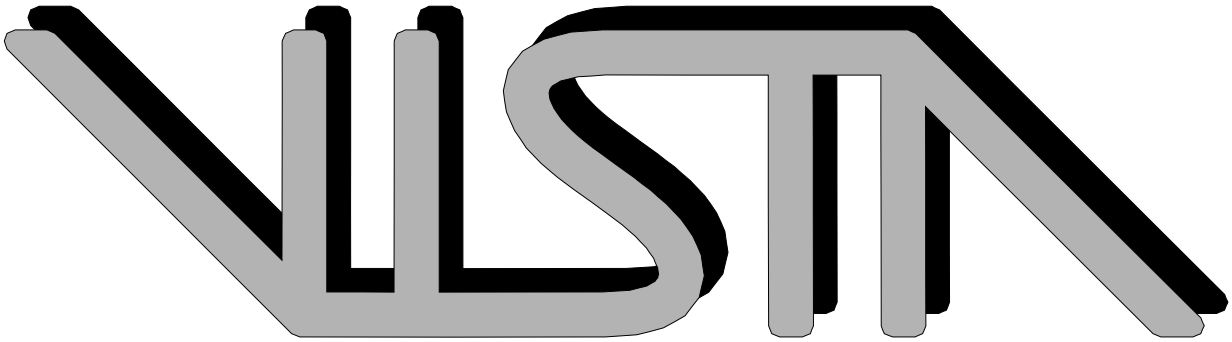

VISTA Status Report
June 2006

S. Dhar, L. Li, M. Pourfath,
M. Spevak, V. Sverdlov, S. Selberherr

Institute for Microelectronics
Technical University Vienna
Gußhausstraße 27-29
A-1040 Wien, Austria

Contents

1 Numerical and Analytical Modeling of the High-Field Electron Mobility in Strained Silicon 1

1.1 Introduction 1

1.2 Monte Carlo Results 1

1.3 Low field mobility 2

1.4 high field model 2

1.5 Conclusion 5

2 An Analytical Model for Organic Thin Film Transistors 6

2.1 Introduction 6

2.2 Variable Range Hopping Transport in Organic Semiconductors 6

2.3 Sheet Conductance of the OTFT 7

2.4 Drain Current 8

2.5 Results and Discussion 9

2.6 Conclusion 10

3 Optimal Design for Carbon Nanotube Transistors 11

3.1 Introduction 11

3.2 Approach 11

3.2.1 Static Response 11

3.2.2 Dynamic Response 13

3.3 Simulation Results 13

3.4 Conclusion 15

4 Process and Device Simulation with a Generic Scientific Simulation Environment 16

4.1 Introduction 16

4.2 The Generic Scientific Simulation Environment 17

4.3 Device Simulation 18

4.4 Process Simulation 19

4.5 Runtime Efficiency 19

4.6 Conclusion 20

5 Current Transport in Nanoelectronic Semiconductor Devices 21

5.1 Introduction 21

5.2 Semiclassical Transport 22

5.3 Quantum-Ballistic Transport 23

5.4 Dissipative Quantum Transport 25

5.5 Conclusions 26

1 Numerical and Analytical Modeling of the High-Field Electron Mobility in Strained Silicon

We have performed a detailed analysis of the electron transport at high electric field in strained Si for different field directions and stress/strain conditions using Full-band Monte Carlo simulations. A phenomenological model describing the velocity-field relationship for electrons in biaxially or uniaxially strained Si has been developed. The model is suitable for incorporation into any device simulator for performing TCAD tasks.

1.1 Introduction

The introduction of strain in Si channels, resulting in significant mobility enhancements, is becoming increasingly important for developing next generation CMOS technologies. The strain in Si can be generated either globally, using an epitaxial layer on a relaxed substrate [1][2][3], by mechanical stressing [4][5], or can be induced during the processing steps [6][7]. Biaxially strained Si layers (two surface dimensions being stressed) using relaxed SiGe substrates have shown large enhancements of electron mobility but suffer from several integration issues. There has thus been a interest towards uniaxially straining Si (one surface dimension stressed) which not only enhances the electron mobility but also exhibits superior hole mobility enhancements for PMOS devices [8]. In this work, we studied electron high-field transport in strained Si using Full Band Monte Carlo (FBMC) simulations. The band structure of strained Si was calculated using the empirical pseudopotential method [9] for biaxial tensile strain and uniaxial stress along the substrate normal. A [100] substrate is considered. We propose an empirical model to describe the velocity-field relationship as obtained from FBMC simulations.

1.2 Monte Carlo Results

Figure 1 and Figure 2 show the velocity-field characteristics as obtained from FBMC simulations for biaxially strained Si grown on a relaxed SiGe substrate for different Ge content and field along [100] and [001] directions, respectively. The figures show an increase (decrease) in the total velocity with increasing strain for field along [100] ([001]) direction. For the [100] field direction, the total velocity shows a region of small negative differential mobility.

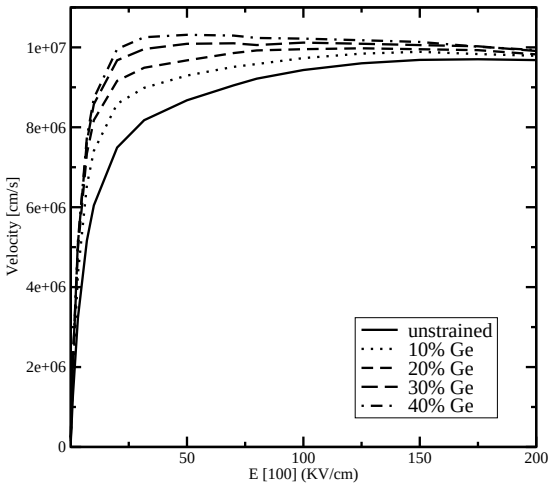

Figure 1: Electron velocity versus field (along [100]) in strained Si on SiGe with increasing Ge content obtained from MC simulations.

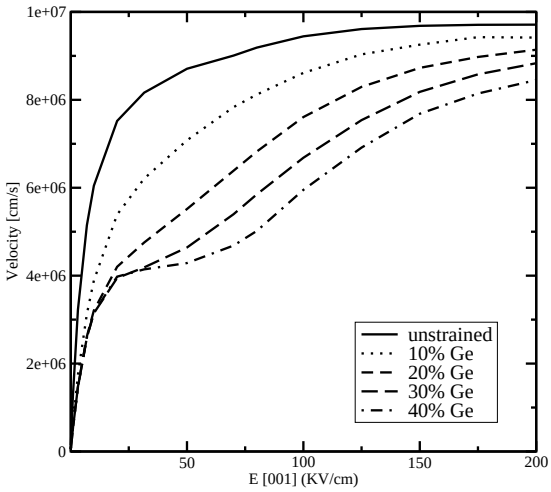

Figure 2: Electron velocity versus field (along [001]) in strained Si on SiGe with increasing Ge content obtained from MC simulations.

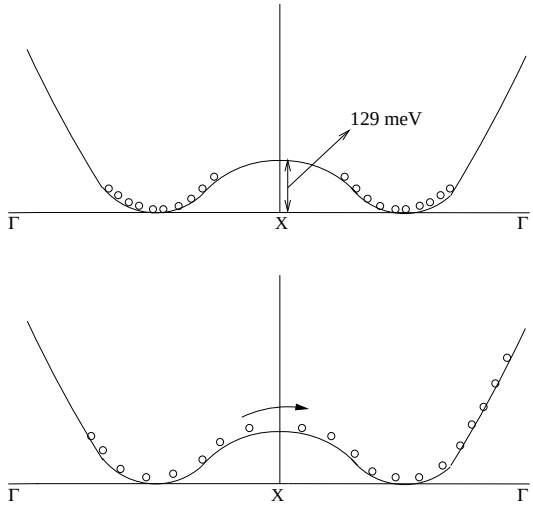

Figure 3: Repopulation effect within the double valley in the equilibrium state (top) and at high-field (bottom)

The velocity-field characteristics shown in Figure 2 exhibit an untypical form for high strain levels. This phenomenon can be explained by the repopulation of valleys induced by the field. The velocity-field curve in Figure 2 is associated with the Δ_2 valleys which move down in energy with increasing strain and have the longitudinal mass in the field direction. These valleys are located at a scaled distance of 0.85 and 1.15 from the center of the first Brillouin zone and are separated by an energy barrier of 129 meV at the X-point as shown in Figure 3. Repopulation effects within this double valley together with the high population of the double valley, lead to the particular shape of the $v(E)$ curve.

1.3 Low field mobility

A key parameter in the modeling of the field dependent mobility is the accurate determination of the low-field mobility. Strain induced enhancement of the low-field mobility can be attributed to two complementary effects. Firstly, the inter-valley phonon scattering is reduced due to decreased number of final available states. Secondly, due to the energy lowering of the Δ_2 valleys, the electrons prefer to occupy this valley and therefore experience a lower in-plane effective conductivity mass.

The low-field mobility for electrons in strained Si can be calculated using the model developed in [10]. It describes the mobility tensor for electrons in strained Si layers as a function of strain. It includes the effect of strain-induced splitting of the conduction band valleys in Si, inter-valley scattering, doping dependence and temperature dependence.

The variation of the in-plane mobility with increasing strain for biaxially strained Si, as calculated from MC simulations and the analytical model is shown in Figure 4.

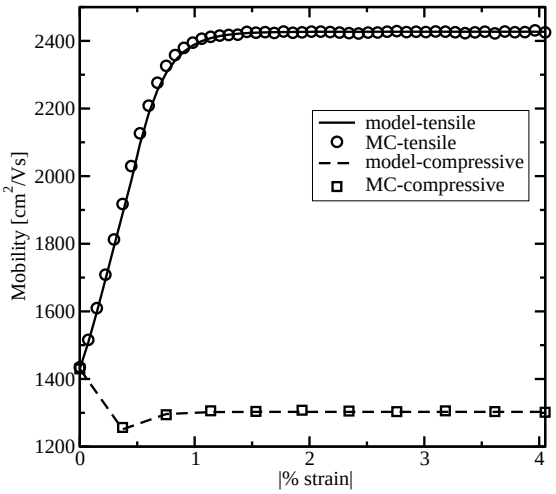

Figure 4: Variation of in-plane mobility for biaxial tensile and compressively strained Si with increasing strain.

Tensile strained Si enhances the electron mobility whereas compressive strain results in a reduced mobility. There is a minimum in the mobility occurring for the case of compressively strained Si. This can be explained by the interplay between the increased conductivity mass of the Δ_4 valleys (which move down in energy for compressive strain) in the transport direction, and the decreasing inter-valley scattering.

1.4 high field model

The high field behavior is modeled using direct fits of empirical expressions to the MC data. This approach has been chosen after a more physically-based, three-valley model had been investigated.

However, one problem of such model would be its complexity. For instance, a nonlinear system for nine unknowns, namely the valley populations, valley velocities and valley temperatures, would have to be solved numerically. The peculiar shape of the $v(E)$ curves in Figure 2 would pose another problem. While the physically-based high-field model could deal with arbitrary strain conditions, the chosen empirical model can not. Therefore, we restricted our study to such strain conditions where only one pair of X-valleys is shifted with respect to the four-fold degenerate X-valleys. These conditions include biaxial stress and uniaxial stress applied along the principal axes of the crystal ($\{100\}$ directions for Si).

If the velocity is governed by the transversal mass m_t, the velocity along the field direction is modeled phenomenologically using the expression

$$v(E) = \frac{2\mu_0 E}{1+\left[1+\left(\frac{2\mu_0 E}{v_s - m\cdot E}\right)^{\beta}\right]^{1/\beta}}. \quad (1)$$

This expression is an extension of the standard mobility model used in [11]. Here μ_0 is the low field mobility and v_s is the saturation velocity. The parameter β governs the velocity transition from low to high fields whereas m accounts for the small negative differential mobility occurring in strained Si for higher strain levels.

If the velocity is governed by the longitudinal mass, m_l,the velocity along the field direction is modeled using the expression

$$v(E) = \frac{2\mu_0 E}{1+\left[1+\left(\frac{2\mu_0 E}{v_s(1-\xi)}\right)^{\beta}\right]^{1/\beta}} + v_s\xi\frac{(E/E_0)^{\gamma}}{1+(E/E_0)^{\gamma}} \quad (2)$$

The additional term incorporated in (2) models the velocity kink occurring due to repopulation effects within the X-valleys. The parameter $\xi < 1$ is used to signify the velocity plateau occurring approximately at $v_s(1-\xi)$. E_0 and γ are fit parameters. The empirical dependences of all parameters on the valley splitting are described by simple analytic expressions. In all cases a linear or quadratic expression is sufficient, except for ξ, which is modeled by a rational expression.

$$\xi = \frac{\Delta E/\xi_0}{1+(\Delta E\cdot\xi_1/\xi_0)^2} \quad (3)$$

Here, ξ_0 and ξ_1 are constants. The parameter variation for biaxially tensile strained Si and [001] field direction are plotted in Figure 5.

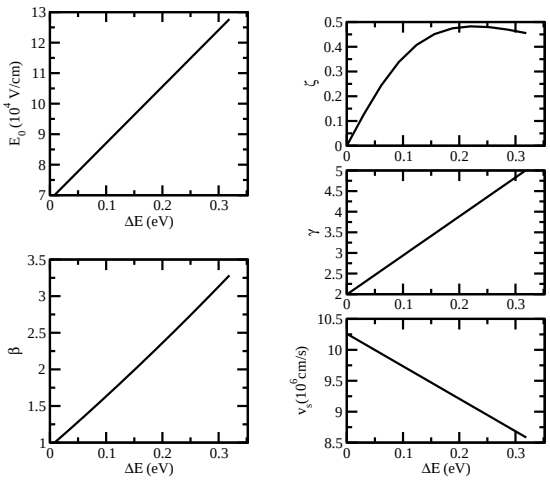

Figure 5: Variation of parameters with strain induced valley splitting for biaxially tensile strained Si and field along [001] direction.

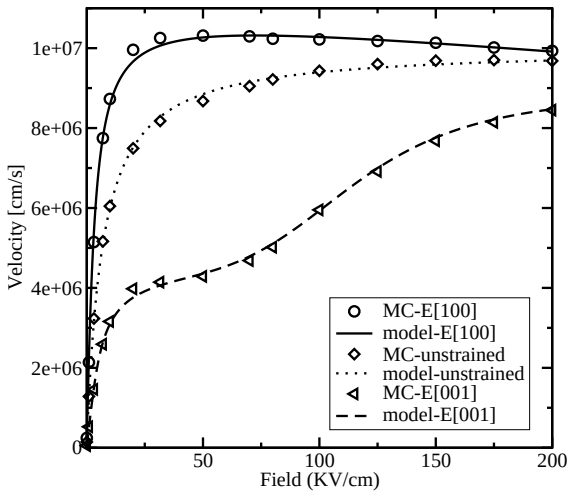

Figure 6: Electron velocity versus field characteristics for unstrained Si and strained Si on $Si_{0.6}Ge_{0.4}$.

Figure 6 shows the $v(E)$ characteristics for unstrained Si and strained Si on $Si_{0.6}Ge_{0.4}$ as obtained from the MC simulations and from our

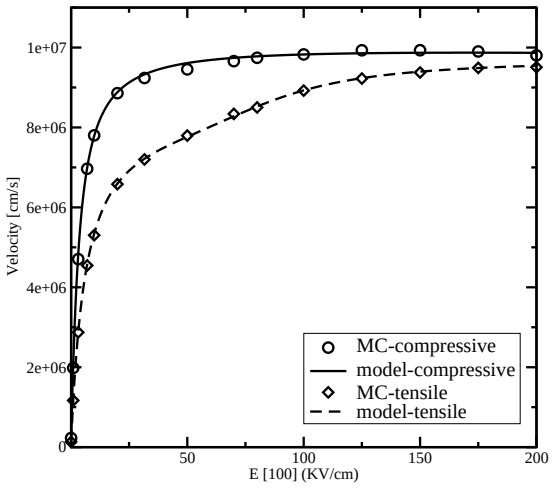

Figure 7: Electron velocity versus field characteristics for Si under uniaxial stress (1GPa) along [001] and field along [100].

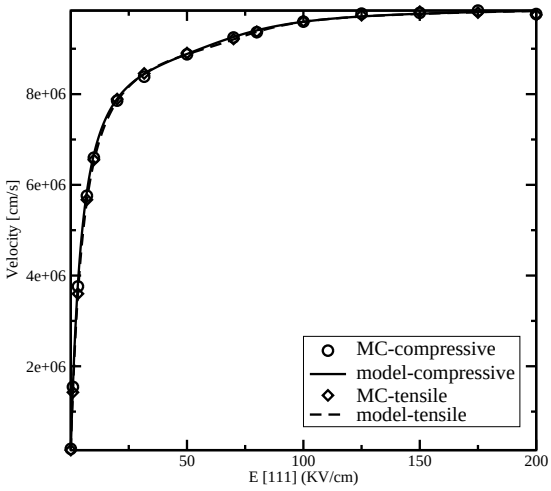

Figure 9: Electron velocity versus field characteristics for Si under uniaxial stress (1GPa) along [001] and field along [111].

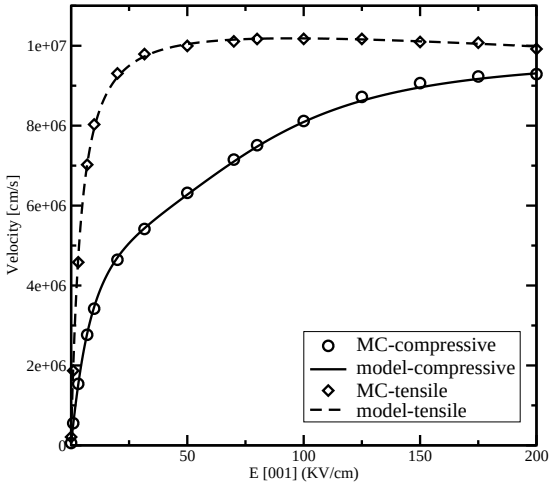

Figure 8: Electron velocity versus field characteristics for Si under uniaxial stress (1GPa) along [001] and field along [001].

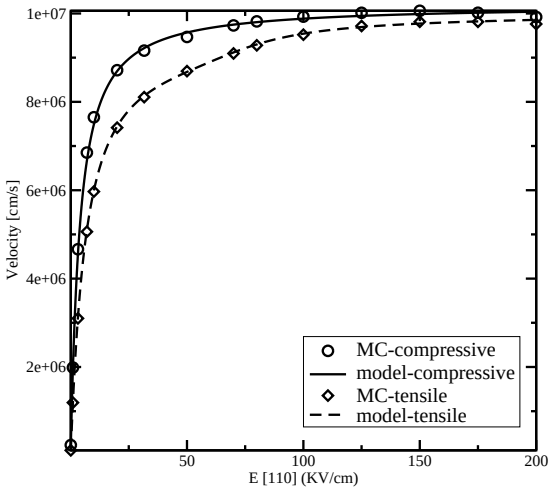

Figure 10: Electron velocity versus field characteristics for Si under uniaxial stress (1GPa) along [001] and field along [110].

model. Biaxial tensile strain increases (decreases) the in-plane (out-of-plane) electron velocity for the complete field range shown. Figure 7 shows the $v(E)$ characteristics for a 1GPa stressed Si layer for field along [100]. From this figure it can be seen that the application of a uniaxial compressive stress also enhances the in-plane velocity. Conversely, applying uniaxial tensile stress results in an enhanced out of plane velocity, as shown in Figure 8.

The $v(E)$ characteristics for the [111] field directions is shown in Figure 9 and . From the figure it is observed that for [111] field direction, there is almost no change in the $v(E)$ characteristics for tensile and compressively strainer Si. Figure 10 shows the $v(E)$ characteristics for [110] field

direction. It is observed that this field direction gives the highest electron velocity in comparison to [100] and [111] directions.

1.5 Conclusion

We have investigated the high-field transport behavior in strained Si for different field directions and stress/strain conditions using FBMC simulations. A phenomenological approach to describe the velocity-field characteristics has been proposed which shows good agreement with the simulations results. The explicitness of the model enables its implementation into any conventional device simulator for performing TCAD tasks.

2 An Analytical Model for Organic Thin Film Transistors

An analytical model that describes the DC characteristics of organic thin film transistors (OTFTs) is presented. The model is based on the variable range hopping theory, i.e. thermally activated tunneling of carriers between localized states. As verified by published data, the model provides an accurate and efficient prediction for transfer characteristics and output characteristics of OTFT via simple formulations.

2.1 Introduction

In recent years, organic thin transistors (OTFTs) have found important application in large-area, low performance and low-cost integrated circuits. Such applications include driving devices for active matrix flat panel displays, organic light identification tags, sensors, etc. The key traits distinguishing field effect transistors with organic active layer from conventional FETs are their potential for low-cost and low-temperature processing, and their compatibility with flexible substrates. As OTFT applications increase, a more accurate and yet simple model of device characteristics is necessary for understanding, improving, and applying these devices. Up to now, many of the numerical or analytical OTFT models use the same expressions as used for for MOS crystalline devices. However, OTFTs present several differences with respect to crystalline MOSFETs because of the low conductivity of organic semiconductors. Furthermore, OTFTs are primarily operated as accumulation field effect transistors as opposed to the usual inversion mode of crystalline MOSFETs. OTFTs are normally conducting at zero gate voltage, and the field-effect mobility usually increases with the gate voltage [12].

In this paper, we derive a basic expression for the sheet conductance based on the variable range hopping (VRH) theory. This theory describes thermally activated tunneling of carriers between localized states (electrons in conduction states) around the Fermi level in the tail of a Gaussian distribution. It has been used to calculate the mobility of OTFTs successfully. After some simplification for the surface potential, simple and efficient analytical expressions for the transfer characteristics and output characteristics are obtained. The model does not require as input parameters the explicit definition of the threshold and saturation voltage, which are rather difficult to evaluate for this kind of device. The obtained results are in good agreement with experimental data.

2.2 Variable Range Hopping Transport in Organic Semiconductors

Because most organic films have an amorphous structure, and disorder is dominating the charge transport, variable-range-hopping in positionally and energetically disordered systems of localized states is widely accepted as the conductivity mechanism in organic semiconductors. Different from hopping, where the charge transport is governed by the thermally activated tunneling of carriers between localized states, rather than by the activation of carriers to the extended-state transport level, the concept of variable range hopping means that a carrier may either hop over a small distance with high activation energy or hop over a long distance with a low activation energy. In an organic thin film transistor with a typical structure shown in Figure 11, an applied gate voltage gives rise to an accumulation of charge carriers in the region of the organic semiconductors close to the insulator. As the charges in the accumulation layer fill the low-energy states of organic semiconductors, any additional carrier in the accumulation layer will require less activation energy to hop to a neighboring site. This results in a higher mobility with increasing gate voltage. Connecting percolation theory, Vissenberg [13] studied the influence of temperature and the influence of the filled states on the conductivity based on the variable range hopping theory. The expression of the conductivity as a function of the temperature and carrier concentration is

$$\sigma(\delta,T) = \\ = \sigma_0 \left[\frac{\pi N_t \delta (T_0/T)^3}{(2\alpha)^3 B_c \Gamma(1-T/T_0)\Gamma(1+T/T_0)} \right]^{T_0/T}, \quad (4)$$

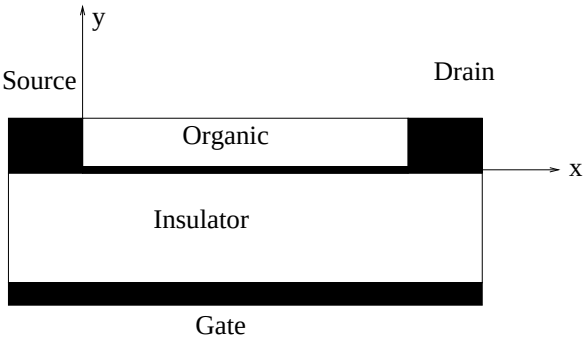

Figure 11: Schematic structure of an organic thin film transistor.

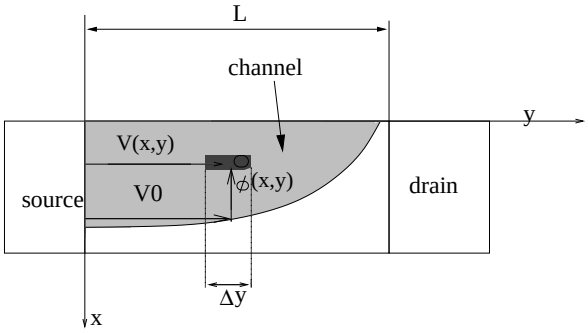

Figure 12: Geometric definition.

where σ_0 is a prefactor, α is an effective overlap parameter, which governs the tunneling process between two localized states, and $B_c \cong 2.8$ is the critical number of bonds per site in the percolating network [14], T_0 is the effective temperature, N_t is the number of states per unit volume and δ is the fraction of the localized states occupied by a carrier.

The carrier concentration is δN_t, and can be expressed in equilibrium as

$$\rho(V) = N_t\delta(V) = N_t\delta_0 \exp\left(\frac{q\Phi}{K_BT_0}\right), \quad (5)$$

where Φ is the electrostatic potential and K_B is the Boltzmann constant, and the δ_0 is the carrier occupation far from the organic-insulator interface.

2.3 Sheet Conductance of the OTFT

From the developed model for amorphous TFT, the drain current I_D can be expressed as

$$I_D = \frac{W}{L}\int_{V_G-V_{FB}-V_D}^{V_G-V_{FB}} G_S(V)dV, \quad (6)$$

where W is the channel width, L is the channel length, V_{FB} is the flat-band voltage, and G_S is the sheet conductance of the channel with $V_D = 0$. The potential V is defined as $V_G - V_{FB} - V_0(y)$, where $V_0(y)$ is the potential at the edge of the space-charge layer where there is no band bending. The basic definition of channel configuration and the variables for the OTFT investigated are illustrated in Figure 12.

The electrostatic potential in the space charge layer at the point (x,y) in the channel is expressed as $V(x,y) = V_0(y) + \Phi(x,y)$, where the $\Phi(x,y)$ is the amount of the band bending in the channel. The conductance for an element of channel length Δy and the width W can be written as

$$G_s = \frac{W}{\Delta y}\int_0^t \sigma dx = \frac{W}{\Delta y}\frac{\sigma(\delta_0,T)}{\sigma_0}\int_0^t \exp\left(\frac{q\Phi}{K_BT}\right)dx \quad (7)$$

where t is the thickness of organic layer. Changing the variable of integration yields

$$G_S = A\int_{\Phi(t(y))}^{\Phi_s(y)} \frac{\exp(q\Phi/k_BT)}{\partial\Phi/\partial x}d\Phi \quad (8)$$

where $\Phi_s(y)$ is the surface band bending and $A = \sigma(\delta_0,T)$. With the identity $\Gamma(1+x)\Gamma(1-x) = \pi x/\sin(\pi x)$ we obtain

$$A = \sigma_0\left[\frac{N_t\delta_0(T_0/T)^4\sin(\pi T/T_0)}{B_c(2\alpha)^3}\right]^{T_0/T} \quad (9)$$

In order to solve (8), we need to get an expression for $\Phi(x)$. By solving Poisson's equation in the gradual channel approximation

$$\frac{\partial^2\Phi(x)}{\partial x} = -\frac{\rho(x)}{\varepsilon_0\varepsilon_s}, \quad (10)$$

we obtain the electric field.

$$-F_x = \frac{\partial\Phi}{\partial x} \approx \sqrt{\frac{2k_BT_0N_t\delta_0}{\varepsilon_0\varepsilon_s}}\exp\left[\frac{q\Phi(x)}{2k_bT_0}\right] \quad (11)$$

From (8) and (2.3), we obtain

$$G_s = A\int_{\Phi(t)}^{\Phi_s}\exp\left[\frac{q\Phi}{k_B}\left(\frac{1}{T}-\frac{1}{2T_0}\right)\right]d\Phi \quad (12)$$

An expression for Φ_s is required. The surface charge density Q_s is related to Φ_s by

$$\begin{aligned} Q_s &= -\varepsilon_0\varepsilon_s F_s = \\ &= \sqrt{2k_BT_0N_t\varepsilon_0\varepsilon_s\delta_0}\exp\left(q\Phi_s/2k_BT_0\right). \end{aligned} \tag{13}$$

The surface band bending is related to the applied gate voltage by

$$V_G = V_{FB} + V_i + \Phi_s \tag{14}$$

where V_i is the voltage drop across the insulator,

$$V_i = \frac{Q_s}{C_i} \tag{15}$$

where $C_i = \varepsilon_i/d_i$ is the insulator capacitance per unit area. From the equations above, an expression for Φ_s is obtained

$$V_G - V_{FB} - \Phi_s = \gamma\exp(q\Phi_s/(2k_BT_0)) \tag{16}$$

For an accumulation mode OTFT, the surface potential is negative, $\Phi_s \leq 0$, corresponding to $V_G \leq 0$.

$$V_G = V_{FB} + \Phi_s - \frac{\sqrt{2k_BT_0\delta_0N_t\varepsilon_0\varepsilon_s}}{C_i}\exp\left(-\frac{q\Phi_s}{2k_BT_0}\right) \tag{17}$$

In order to reduce computation time, an explicit yet accurate relation between surface and gate voltage is preferable. In (16), we can get Φ_s using a numerical approach. However, in the accumulation mode, it holds $\exp(-\Phi_s) \gg \Phi_s$, so that an approximate expression of surface potential can be obtained as

$$\Phi_s = -\frac{2k_BT_0}{q}\ln\left[\frac{(V_{FB}-V_G)C_i}{\sqrt{2K_bT_0\delta_0N_t\varepsilon_0\varepsilon_s}}\right] \tag{18}$$

A comparison between numerical calculation and approximate calculation are shown in Figure 14 and Figure 15. As can be seen, the agreement is very satisfactory. Parameters are from [15, 16].

With the simplified surface potential and (11), we can get the simplified sheet conductance as

$$G_s = \beta\left[\left(\frac{V_G - V_{FB}}{\gamma}\right)^{2T_0/T-1} - 1\right] \tag{19}$$

For a thick organic semiconductor layer, $\Phi(t) = 0$ and the constants β is

$$\beta = \sigma_0\sqrt{\frac{2\varepsilon_0\varepsilon_sK_BT_0}{\delta_0N_t}}\frac{k_BT}{q(T-2T_0)}$$

$$\gamma = \frac{(2\alpha)^3B_c2k_BT_0\varepsilon_0\varepsilon_s}{C_i^2(T_0/T)^3\sin(\pi T/T_0)}$$

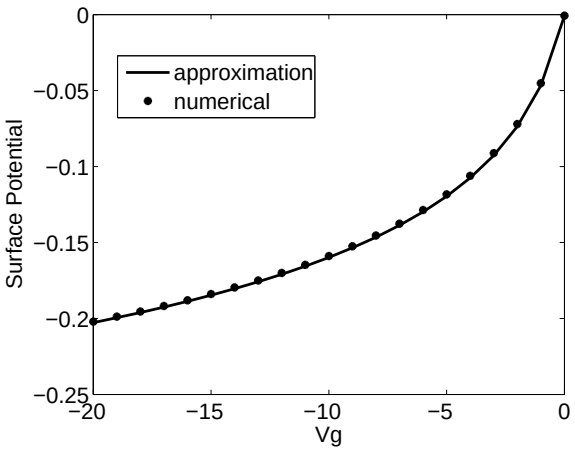

Figure 13: The electrostatic surface potential as a function of gate voltage obtained by the implicit relation (16) and the approximation (17) (solid line).

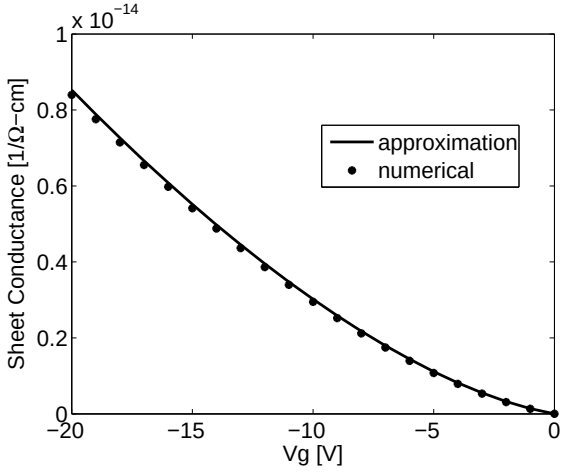

Figure 14: Sheet conductance from numerical calculation (symbols) and the approximation.

2.4 Drain Current

The drain current can be calculated by substituting the expression of G_s into (6), and we can get

$$\begin{aligned} I_D = \beta\frac{W}{L}\Bigg[&\left(\frac{V_G - V_{FB}}{\gamma}\right)^{2T_0/T} - \\ &- \left(\frac{V_G - V_{FB} - V_D}{\gamma}\right)^{2T_0/T}\Bigg] \end{aligned} \tag{20}$$

in the triode region ($V_{GS} - V_{FB} \geq V_{DS}$) and

$$I_D = \beta \frac{W}{L} \left(\frac{V_{GS} - V_{FB}}{\gamma} \right)^{2T_0/T} \qquad (21)$$

in saturation ($V_{GS} - V_{FB} \leq V_{DS}$).

2.5 Results and Discussion

This model has been confirmed by comparisons between experimental data and simulation results.

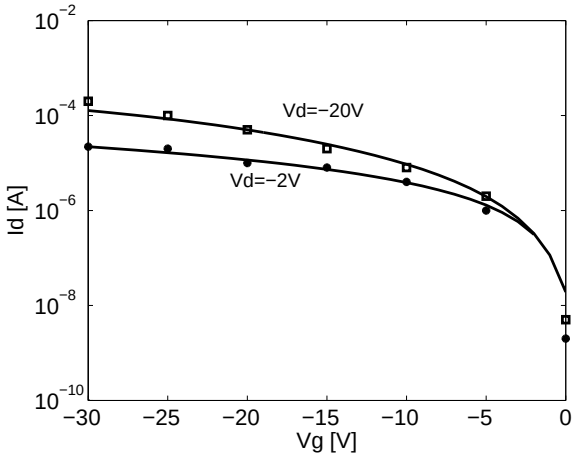

Figure 15: Measured (symbols) and calculated currents of a pentacence OTFT at room temperature.

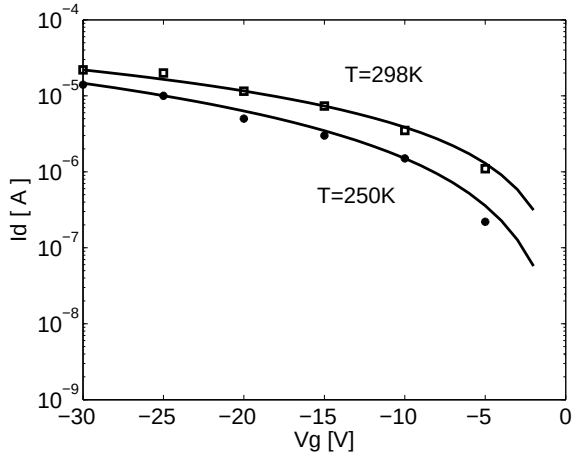

Figure 16: Measured (symbols) and calculated transfer characteristics of a pentacence OTFT at different temperature at $V_D = -2V$.

Input parameters are taken from [15]: $W = 20,000\mu m$, $L = 10\mu m$, $\varepsilon_s = 3$, $C_i = 17\mathrm{F}/(\mu m)^2$, δ_0 =3.5 S/m, $\alpha^{-1} = 3.1 \times 10^{-10}$m, T_0 = 385 K.

In Figure 15 and Figure 16, the transfer characteristics of a pentacene OTFT are given for $V_{FB} = 1V$ and different gate voltage and different temperature. Both figures show a good agreement between the analytical model and experimental data. Here we also model the transfer characteristics of a PTF OTFT, where some parameters are different from those for pentacence $T_0 = 382, \delta_0 = 5.6S/m; \alpha^{-1} = 1.5 \times 10^{-10}m$.

The modeled output charateristics of the pentacene OTFT is shown in Figure 18.

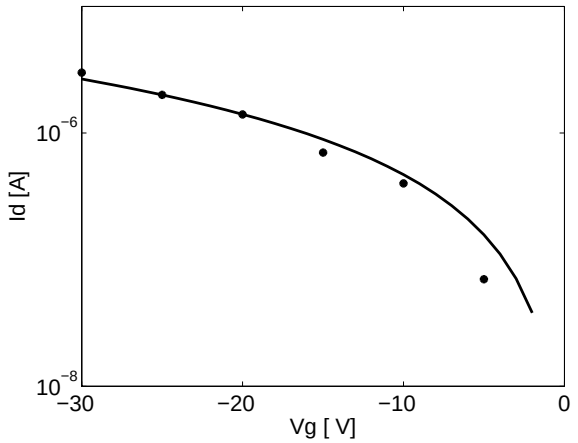

Figure 17: Measured (symbols) and calculated transfer characteristics of a PTV OTFT at room temperature at $V_D = -2V$.

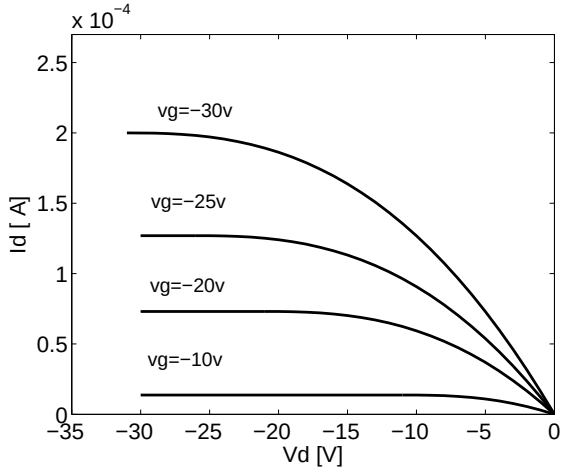

Figure 18: Modeled output characteristics of a pentacene OTFT.

2.6 Conclusion

An analytical expression has been derived for the sheet conductance of organic thin film transistors on the basis of the variable range hopping transport theory. With a suitable approximation for the surface potential a simple analytical model for the DC characteristcs is derived. The derivation is similar to that presented in [16], however, the present work focuses on the surface potential approximation. A simplified model, obtained in [16] as a reduction of a more complex model, is derived in a direct way. The model shows close agreement with experimental data. The resultant equations can be employed in CAD circuit analysis programs.

3 Optimal Design for Carbon Nanotube Transistors

A numerical study of carbon nanotube field effect transistors is presented. To investigate transport phenomena in such devices the non-equilibrium Green's function formalism was employed. Phenomena like tunneling and electron-phonon interactions are rigorously taken into account. The effect of geometrical parameters on the device performance was studied. Our results clearly show that device characteristics can be optimized by appropriately selecting geometrical parameters.

3.1 Introduction

A carbon nanotube (CNT) can be viewed as a rolled-up sheet of graphite with a diameter of a few nano-meters. Depending on the chiral angle the CNT can be either metallic or semiconducting. Semiconducting CNTs can be used as channels for field-effect transistors (FETs). CNTFETs have been studied in recent years as potential alternatives to CMOS devices because of their capability of ballistic transport.

Depending on the work function difference between the metal contact and the CNT, carriers at the metal-CNT interface encounter different barrier heights. Devices with positive [17] and zero [18] barrier heights were fabricated. The barrier height is defined as the potential barrier which is seen by carriers at the Fermi level in the metal. Therefore, in a device with zero barrier height, carriers with energies above the Fermi level of the metal reach the channel by thermionic emission and carriers with energies below the Fermi level have to tunnel to reach the channel.

Devices with positive barrier heights have lower on-current and also suffer from ambipolar behavior [19, 20], while devices with zero barrier height theoretically [21] and experimentally [22] show better performance. In this work we focus on devices with zero barrier height for electrons. The barrier height for holes is given by the band gap of the CNT. Since the dispersion relations for electrons and holes are the same, our discussions are valid for holes as well.

Using the non-equilibrium Green's function (NEGF) formalism quantum phenomena like tunneling, and scattering processes can be rigorously modeled. Here we extended our previous work [23] by including the effect of electron-phonon interaction in the calculations, considering large signal dynamic response, and investigating the influence of geometrical parameters. In the next section our methodology is described. Then the effect of different geometrical parameters on the device characteristics is analyzed, and methods for performance optimization are suggested.

3.2 Approach

In this section the models used to study the static and dynamic response of CNTFETs are explained.

3.2.1 Static Response

Based on the NEGF formalism we investigated the effect of device geometry on the performance of carbon nanotube field-effect transistors. We have solved the coupled system of transport and Poisson equations numerically. Due to quantum confinement along the tube circumference, carrier have bound wave functions around the CNT and can propagate along the tube axis. Under the assumption that the potential profile does not vary around the circumference of the CNT, sub-bands will be decoupled. In this work we assume bias conditions for which the first sub-band contributes mostly to the total current. In the mode-space approach [24] the transport equation for each sub-band can be written as:

$$G^{\mathrm{R,A}}_{\mathbf{r},\mathbf{r}'}(E) = [EI - H_{\mathbf{r},\mathbf{r}'}(E) - \Sigma^{\mathrm{R,A}}_{\mathbf{r},\mathbf{r}'}(E)]^{-1} \quad (22)$$

$$G^{<,>}_{\mathbf{r},\mathbf{r}'}(E) = G^{\mathrm{R}}_{\mathbf{r},\mathbf{r}'}(E)\Sigma^{<,>}_{\mathbf{r},\mathbf{r}'}(E)G^{\mathrm{A}}_{\mathbf{r},\mathbf{r}'}(E) \quad (23)$$

In (22) an effective mass Hamiltonian was assumed. All our calculations assume a CNT with a band gap of $\mathrm{E_g} = 0.6$ eV corresponding to a CNT with a diameter of $\mathrm{d_{CNT}} = 1.6$ nm, and $m^* = 0.05m_0$ for both electrons and holes.

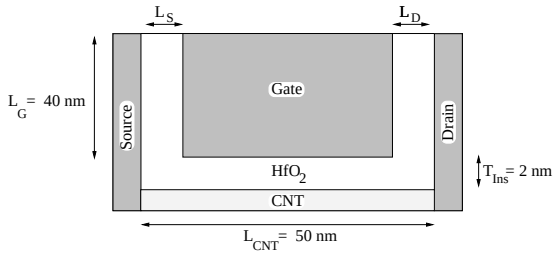

Figure 19: The device structure. The device is 250 nm extended into the third dimension. $\varepsilon_r = 15$.

A recursive Green's function method is used for solving (22) and (23) [25]. The total self-energy in (22) consists of the self-energies due to the source contact, drain contact, and electron-phonon interaction, $\Sigma^R = \Sigma^R_S + \Sigma^R_D + \Sigma^R_{el-ph}$. The self-energy due to electron-phonon interaction consists of the contribution of elastic and inelastic scattering mechanisms, $\Sigma^{<,>}_{e-ph} = \Sigma^{<,>}_{el} + \Sigma^{<,>}_{inel}$. Assuming a single sub-band the electron-phonon self-energies are simplified to (24)-(27).

$$\Sigma^{<,>}_{el,(\mathbf{r},\mathbf{r})}(E) = D_{el} G^{<,>}_{\mathbf{r},\mathbf{r}}(E) \tag{24}$$

$$\begin{aligned}\Sigma^{<}_{inel,(\mathbf{r},\mathbf{r})}(E) = &\sum_{\nu} D^{\nu}_{inel} \\ &[(n_B(\hbar\omega)+1)G^{<}_{\mathbf{r},\mathbf{r}}(E+\hbar\omega) \\ &+n_B(\hbar\omega)G^{<}_{\mathbf{r},\mathbf{r}}(E-\hbar\omega)]\end{aligned} \tag{25}$$

$$\begin{aligned}\Sigma^{>}_{inel,(\mathbf{r},\mathbf{r})}(E) = &\sum_{\nu} D^{\nu}_{inel} \\ &[(n_B(\hbar\omega)+1)G^{>}_{\mathbf{r},\mathbf{r}}(E-\hbar\omega) \\ &+n_B(\hbar\omega)G^{>}_{\mathbf{r},\mathbf{r}}(E+\hbar\omega)]\end{aligned} \tag{26}$$

$$\Im m[\Sigma^R(E)] = \frac{1}{2i}[\Sigma^{>} - \Sigma^{<}] \tag{27}$$

where n_B is given by the Bose-Einstein distribution function. In general electron-phonon interaction parameters ($D_{el,inel}$) depends on the diameter and the chirality of the CNT. The calculation of these parameters is presented in [26]. The imaginary and real parts of the self-energy broadens and shifts the density of states, respectively. We neglected the real part of the self-energy.

The transport equations (22) to (27) are iterated to achieve convergence of the electron-phonon self-energies, resulting in a self-consistent Born approximation. Then the coupled system of transport and Poisson equation is solved iteratively. The carrier concentration and the current density at some point $\mathbf{r}$ of the device can be calculated as (28) and (29).

$$n_{\mathbf{r}} = -4i \int G^{<}_{\mathbf{r},\mathbf{r}}(E) \frac{dE}{2\pi} \tag{28}$$

$$j_{\mathbf{r}} = \frac{4q}{\hbar} \int \mathrm{Tr}[\Sigma^{<}_{\mathbf{r},\mathbf{r}} G^{>}_{\mathbf{r},\mathbf{r}}(E) - \Sigma^{>}_{\mathbf{r},\mathbf{r}} G^{<}_{\mathbf{r},\mathbf{r}}(E)] \frac{dE}{2\pi} \tag{29}$$

In CNTs elastic scattering is caused by acoustic phonons and inelastic scattering occurs due to zone boundary (ZB), optical (OP), and radial breathing (RBM) phonon modes. In CNTs with diameters in the range $d_{CNT} = 1-2$ nm, the energies of the these phonon modes are $\hbar\omega_{ZB} \approx 160$ and 180 meV, $\hbar\omega_{OP} \approx 200$ meV, and $\hbar\omega_{RBM} \approx 30$ meV respectively [27, 28]. Due to small occupation number of high energy phonons, such as OP and ZB phonon modes, they do not degrade the performance considerably, whereas the RBM phonon mode can have a detrimental effect. However, due to weak electron-phonon coupling the RBM mode has a negligible effect at room temperature. The electron-phonon coupling is also weak for acoustic phonon (AP) modes. Therefore, short CNTFETs can operate close to the ballistic limit. Figure 20 shows excellent agreement between simulation results

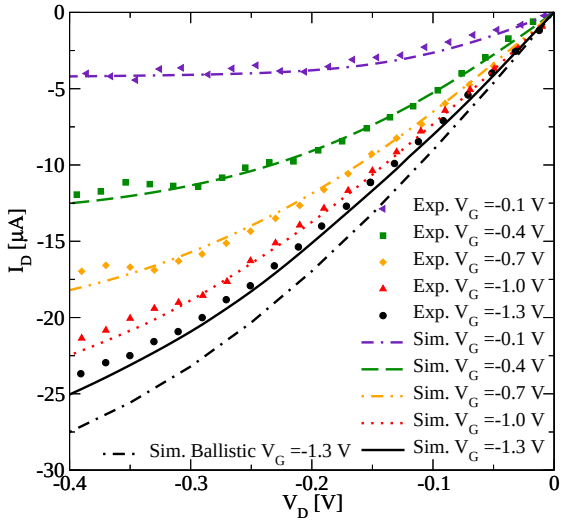

Figure 20: Comparison of the simulation results and experimental data [22] for the output characteristics. The results for the bias point $V_G = -1.3$ V are compared with the ballistic limit.

and experimental data [22]. The result for the bias point $V_G = -1.3\,V$ is compared with the ballistic limit, which confirms the validity of nearly ballistic transport in short CNTFETs.

3.2.2 Dynamic Response

To investigate the dynamic response of the device we consider the device delay time defined as:

$$\tau = \frac{C_G V_{DD}}{I_{on}} \tag{30}$$

Here, $C_G = C_{GS} + C_{GD} + C_{GG}$ with ${C_{GG}}^{-1} = {C_{Ins}}^{-1} + {C_Q}^{-1}$. The quantum capacitance is given by $C_Q = 8q^2/h\nu_F \approx 400aF/\mu m$, including the twofold band and spin degeneracy [29, 30]. The insulator capacitance, occurring between the tube and a plane, is given by [31]:

$$C_{Ins} = \frac{2\pi\kappa\varepsilon_0}{\cosh^{-1}(T_{Ins}/R_{CNT} + 1)} \tag{31}$$

For the geometry parameters given in Figure 19 $C_{Ins} \approx 400aF/\mu m$. For a device with 50 nm channel length $C_{GG} \approx 10aF$. To calculate the gate-source and gate-drain parasitic capacitances we assumed the capacitance of two parallel plates, $C_{GS,GD} = \kappa\varepsilon_0 A/L_{S,D}$, (see Figure 19). Even with a small total area of $A = 250\,nm \times 40\,nm$ and a large spacer width of $L_{GS,GD} = 10\,nm$ the parasitic capacitances $C_{GS} + C_{GD} \approx 260\,aF$ are much bigger than C_{GG}. As a result, $C_G \approx C_{GS} + C_{GD} = \kappa\varepsilon_0 A(1/L_S + 1/L_D)$.

3.3 Simulation Results

In this section the effects of the gate-source spacer, gate-drain spacer, insulator thickness, and the insulator dielectric constant on the device characteristics are studied.

Due to ambipolar behavior, in the off-regime the drain current of CNTFETs starts to increase [19, 22, 32]. To reduce this effect we have proposed to increase the gate-drain spacer [23]. When increasing L_D, the off-current decreases, while the on-current remains nearly unchanged, such that the I_{on}/I_{off} ratio increases. By increasing L_D the gate-drain parasitic capacitance decreases,

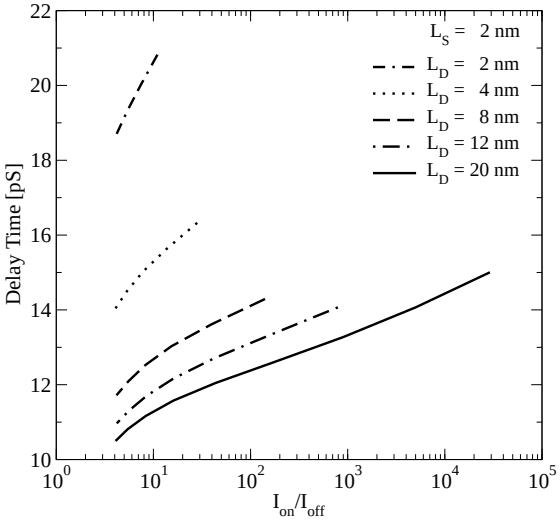

Figure 21: The effect of L_D on the device delay time versus I_{on}/I_{off} ratio. $L_S = 2$ nm and $V_{DD} = 0.8$ V.

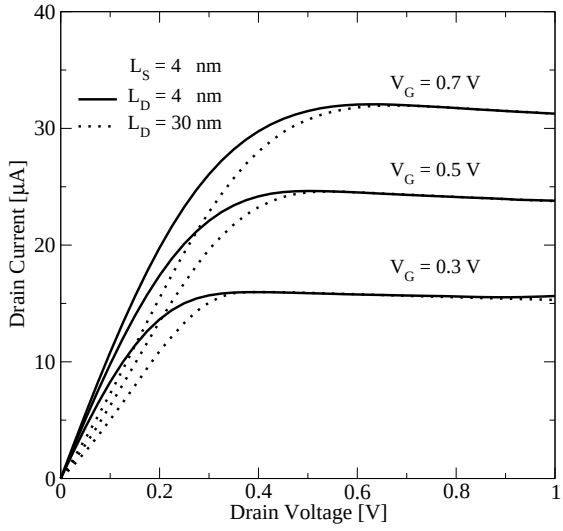

Figure 22: Output characteristics at different gate biases for devices with $L_D = 4$ nm and $L_D = 20$ nm. $L_S = 4$ nm.

which results in reducing the device delay time. Figure 21 shows the effect of L_D on the device delay time versus I_{on}/I_{off}. As shown, a significant performance improvement is achieved. The disadvantage of this method is that at low drain biases electrons have to tunnel through a thicker barrier to reach the drain contact, resulting in a smaller drain current (Figure 22).

When increasing L_S, the gate-source parasitic capacitance is reduced, and so is the on-current. The band edge profile near the source contact

plays an important role in controlling the total current. Increasing L_S reduces the gate control of the band-edge profile near the source contact. Both the tunneling current and thermionic emission current contribute to the total current. Electrons with energies lower than the barrier height have to tunnel through the source-sided metal-CNT interface barrier to reach the channel while electrons with energies higher than the barrier height are injected by thermionic emission. Since the tunneling probability decreases exponentially with the barrier width, the tunneling current decreases with increasing L_S. However, the thermionic emission current is independent of the barrier width. The contribution of the tunneling current decreases with decreasing barrier height, while that of thermionic emission increases. Since τ is proportional to the parasitic capacitance and inversely proportional to the on-current (30), there is an optimal value for L_S, which minimizes τ. As shown in Figure 23 the optimal value of L_S for the given material and geometrical parameters results in optimized device characteristics. It can be easily shown that the optimal value L_{S0}, where $\frac{\partial \tau}{\partial L_S}|_{L_{S0}} = 0$, is achieved when $\frac{1}{C_G}\frac{\partial C_G}{\partial L_S}|_{L_{S0}} = \frac{1}{I_{on}}\frac{\partial I_{on}}{\partial L_S}|_{L_{S0}}$. Considering the expression derived for C_G in Section II.B, we have $\frac{1}{C_G}\frac{\partial C_G}{\partial L_S} = [L_S(1+L_S/L_D)]^{-1}$.

Figure 24 shows the sensitivity of the on-current to L_S. However, the mentioned sensitivity is not zero due to the contribution of the tunneling current from states below the Fermi level. Since at positive gate biases the conduction band-edge is pushed below the source Fermi level, even in devices with zero barrier height the tunneling current can contribute to the total current. For thinner insulators the width of the source-sided barrier decreases, resulting in a higher tunneling current contribution to the total current and a higher sensitivity of the on-current to L_S. The optimal spacer width is $L_S \approx 6$ nm at $T_{Ins} = 2$ nm and $L_D = 20$ nm. Note that the optimal value for L_S depends on L_D. For small values of L_D the gate-drain parasitic capacitance dominates the gate-source parasitic capacitance, therefore any further decrease of L_S does not improve the delay time.

Electron-phonon interaction reduces the on-current, both, directly and indirectly [33, 34].

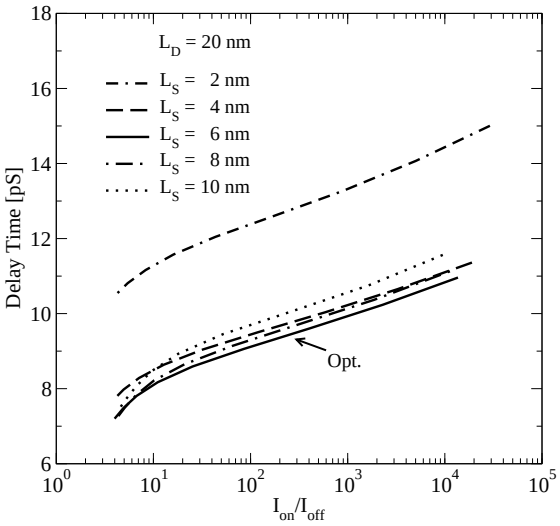

Figure 23: The effect of L_S on the device delay time versus I_{on}/I_{off} ratio. $V_{DD} = 0.8$ V. The optimal L_S for both device types are shown.

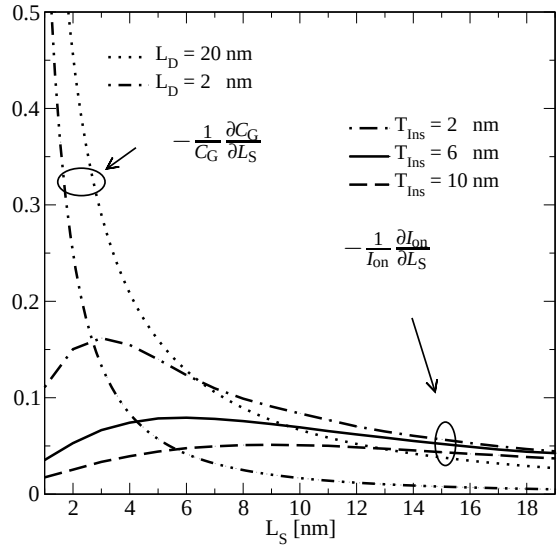

Figure 24: The sensitivity of the parasitic capacitance and the on-current to L_S for different insulator thicknesses. The intersection of the curves gives the optimal L_S, which minimizes τ.

The direct effect is due to backscattering of carriers, but scattering also redistributes the carrier concentration profile along the device. This redistribution affects the band-edge profile so that it reduces the total current. To reduce the indirect effect one should increase the gate-CNT coupling. If thin and High-κ insulators are used then $C_{Ins} \gg C_Q$ and $C_{GG} \approx C_Q$, implying that

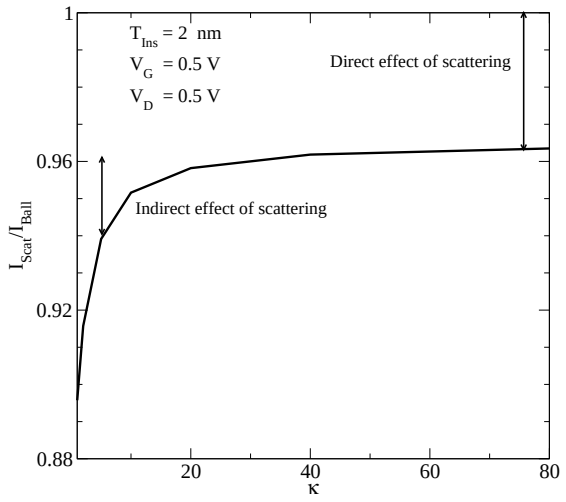

Figure 25: The ratio of the drain current in the presence of scattering to the ballistic limit for different κ. The proportions due to direct and indirect effect of scattering on the on-current are shown. For high-κ the indirect part reduces.

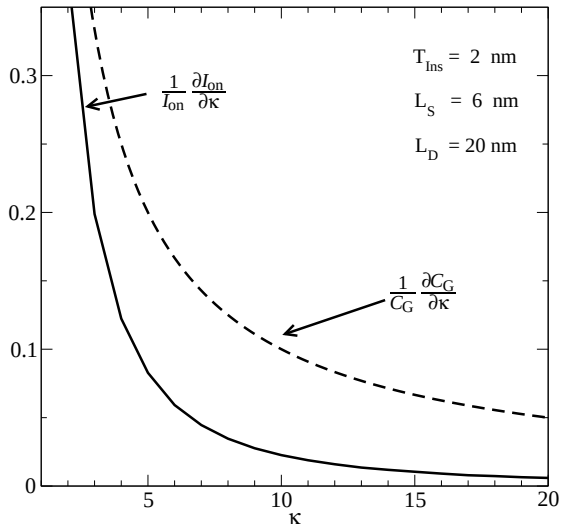

Figure 26: The sensitivity of the parasitic capacitance and the on-current to κ. Since the curves do not intersect at high values of κ lower values of κ minimizes τ.

the potential on the tube becomes the same as the gate (perfect coupling). This regime is called quantum capacitance limit in which the device is potential-controlled rather than charge-controlled [35]. Figure 25 compares the ratio of the current in the presence of scattering to the ballistic limit for different insulators. For the given material and geometrical parameters a $\kappa > 20$ maximizes the performance of the device. But, with using high-κ materials not only the on-current but also the parasitic capacitances increase. Therefore, there is a κ which optimizes the delay time. It can be shown that the optimized value is achieved when $\frac{1}{C_G}\frac{\partial C_G}{\partial \kappa}|_{\kappa_0} = \frac{1}{I_{on}}\frac{\partial I_{on}}{\partial \kappa}|_{\kappa_0}$. Considering the expression derived for C_G in Section II.B, we have $\frac{1}{C_G}\frac{\partial C_G}{\partial \kappa} = \frac{1}{\kappa}$. Figure 26 shows the sensitivity of the on-current and parasitic capacitances to κ. Since the curves do not intersect at high values of κ, lower values minimizes τ. Therefore, there is a trade-off between device delay time and the on-current. For a specific application this parameter can be optimized.

3.4 Conclusion

We showed that the device characteristics can be optimized by appropriately selecting the geometrical parameters. With increasing the gate-drain spacer, the off-current and the gate-drain parasitic capacitance reduce at the cost of a drain current reduction at low bias voltages. With increasing the gate-source spacer, the drain current and gate-source parasitic capacitance decrease. Since the device delay time is proportional to the parasitic capacitances and inversely proportional to the on-current, there is a value for the gate-source spacer which minimizes the device delay time. The optimal point is where the sensitivity of these quantities are equal. By using high-κ insulators the gate-CNT coupling increases which results in higher on-current, but the parasitic capacitances increase and as a result the device delay time increases.

4 Process and Device Simulation with a Generic Scientific Simulation Environment

We present a high performance environment for scientific simulation applications (GSSE). This environment is based on the three orthogonal concepts of topology, geometry, and quantities. Lambda calculus is used in order to assemble various partial differential equations for TCAD and other physical equations. We present examples from device and process simulation which show the applicability of our environment. Despite the high abstraction level we can archieve high performance.

4.1 Introduction

A general environment for TCAD equations has to provide methods for the solution of different physical phenomena such as carrier transport, diffusion, electromagnetic wave propagation, heat transfer, mechanical deformation, fluid flow, and quantum effects. Due to the wide range of applications it is not trivial to develop an environment which is capable of handling all equations within a homogenous environment. In the field of TCAD coupled partial differential equations and multi quantity equations are often employed. For the solution of these equations we use discretization schemes such as the finite element method, the finite volume method, the finite difference method or the boundary element method.

Each of these schemes has its merits and shortcomings and is therefore more or less suited for different classes of equations. All of these methods have in common that they require a proper tessellation and adaption of the simulation domain [36, 37]. Due to the diversity of the mathematical structures themselves, combined with efficiency considerations, in particular in three dimensions, the development of simulation software is quite difficult. Supporting libraries for numerical simulations exist, but no complete environment for scientific computing to tackle the following issues:

- Support for different geometries and topologies
- Complete and tested discretization schemes
- Support for mathematical modeling
- High performance calculation
- (Real-time) visualization

Hence, the simulation tools are typically written by experts specialized in other fields. In the extreme case all areas of simulator development, like software design, programming, testing, and evaluation are done by one person. In the last decade our institute has developed different simulators and libraries, like SAP [38], Wafer-State-Server [39], and Minimos-NT [11]. However, none of them has shown to be perfect for the rapid progress in scientific software development. Even the reuse of simple code parts is difficult, due to the non-generic library approach.

Therefore a scientific environment with high flexibility and adaptivity of meshes combined with great flexibility in numerical treatment and discretization schemes in all dimensions is called for. It should be possible to use a common code base which is easily adaptable to special requirements but does not require specialized features for different discretization schemes such as element matrices like many other specialized finite element simulation environments.

On that account, we have extracted the main concepts from all of our simulation tools and developed the generic and lightweight environment GSSE which suits scientific requirements. On the one side, generic library means that each part of GSSE can be used separately. The complete GSSE is based on header files only and therefore the required mechanisms can be included without incurring additional dependencies. On the other hand, generic means that all data types are parameterized and can be exchanged easily, for instance the numerical data type for quantity storage. GSSE is designed for rapid development of simulation software. One of the most important facts is that errors can be easily found and even prevented. As errors are often not obvious to detect it may already take a lot of experience to decide if the result

from a simulation is erroneous or not. If the result is not correct, it might be reduced to a programming bug, a logical error in the program flow, or a badly chosen parameter. Due to this reason it is necessary to locate an error quickly. Each of the data structures which we provide can itself be tested before compound data structures are tested. Thus the development effort for the final code can be reduced enormously.

In the following we will present some examples of our simulations from the field of semiconductor device and process simulation. First we solve a simple device simulation example. Then we benchmark a finite element example with the electromagnetic simulation tool SAP. Both examples will show that our environment can perform fast calculations.

4.2 The Generic Scientific Simulation Environment

Due to the large variety of available models and differential equations, discretization schemes, and simulation domains we had to extract the base concepts of a simulation environment. The main aspect of software design is orthogonality as well as modularity. Each component should be usable without any dependences to another. If a higher structure combines two base structures (e.g. quantities on a topology) it does not depend on the lower structures but it only relies on the concept.

Topology. Within a scientific simulation environment it is crucial to have neighborhood information of vertices, edges, faces, and cells available within a constant time. For this reason we have implemented a data structure which covers only topological information of vertex cell incidence and cell vertex incidence. Even though storing one of these incidence functions is redundant it is necessary to guarantee constant time for traversal. Based on this incidence information we generate inter-dimensional objects such as edges and faces. The incidence information of edges and faces does not need to be stored explicitly because it can be derived from the base traversal functions and the archetype information.

The archetype concept implies that each cell of the tessellation has the same topological shape or very few different shapes. Therefore we need not store all edge and face information but we can derive it from an archetype. For this reason the unstructured topology is highly flexible. We can use archetypes of different dimensions and shapes.

Geometry. Even though we have a convenient method for describing the topology of a simulation domain this does not imply that we store the coordinate information on the vertices and cells. Topology tells us for instance if two vertices are connected by an edge. It however does not provide any information about the real geometry of the curve. This is the task of the geometry concept. The basic geometry concept is the point list. The point list contains the geometrical point coordinates associated to the topological vertices. From this information we can derive the geometry of all edges, faces and cells. We can perform orientation tests, volume measurements, and the calculation of the voronoi information.

Quantity. Quantities are numerical attributes which can be stored on all topological elements using their handles. Using an associative storage concept we can store values with respect to an associative key of the topology and a quantity descriptor key (which might be a string). The value types of the quantities can be chosen differently. The simplest case is a scalar floating point value. We also provide vector and tensor quantities as well as string quantities. The quantity data type can be parameterized on the data type so that it is possible to use any type as data type for the quantity.

Based on these core concepts, a mathematical function layer was implemented to provide easy development of all different modeling issues on the one side and a high performance computing on the other side. This layer includes all necessary functions for a convenient numerical analysis as well as accumulation functions which will be discussed later on in the example section. We even have the possibility to work with linearized functions which provide direct access to the system matrix for line-wise entry as well as finite element stencils.

As we have parameterized data types for numerical calculations it is possible to introduce abstract matrix data types with lazy evaluation concepts which can reduce the execution time as well as the number of temporary objects. For this reason we use expression templates [40, 41] as well as lambda expressions [42].

A solver interface is integrated in this environment for the use of all different kinds of state-of-the-art solvers called TRILINOS. For the important visualization purpose within scientific computing, IBM's data explorer [43] is integrated with a few modifications to make a real-time visualization possible.

4.3 Device Simulation

The field of device simulation requires the use of many different numerical techniques. Macroscopic transport [44] models are among the most widely employed calculations. These models can be derived by applying the moment method on the Boltzmann equation. Together with the Poisson equation they form a system of partial differential equations which are capable of describing the behavior of semiconductor devices. We use the simplest macroscopic model, the drift diffusion transport equations. In the following we show how the equations can be discretized in our environment by the means of the finite volume method. The discretization formula for the Poisson equation yields

$$\sum_{\text{edge vertex}} (\Psi_j - \Psi_i) \frac{A}{d} = V (n - p + N_\text{A} - N_\text{D}) \, \text{q} \quad (32)$$

Based on this discretization formula we obtain the final discretization routine (32) we can write the following statement in the GSSE.

```
eqn = (sum<vertex_edge>[
  diff<edge_vertex>(0.0)[potential]*
  A/d] - q * volume *
  (n - p + N_A - N_D))(v);
```

The same assembly routine can be applied to the current relations using the Scharfetter-Gummel discretization [45] (generation and recombination rates are omitted here).

$$J_w = \frac{1}{\lambda} \left(n_j \mathcal{B}(\lambda) - n_i \mathcal{B}(-\lambda) \right), \quad (33)$$

where $\mathcal{B}$ is the Bernoulli function. Using the method of finite volumes we have discretize the current integral on the boundary of the control volumes. If we consider the stationary case as well as zero recombination we obtain the following code snippet for our discretization scheme.

```
eqn = sum<vertex_edge>(0.0)[
 area/distance * diff<edge_vertex>()
        [n * bern(d_psi/u_th) ]
        [n * bern(-d_psi/u_th)]](v);
```

Both discretization terms need quantities which are located on topological elements (Figure 27). The potential and charge terms as well as the box volume are stored on the vertices. The distance (d) and area (a) as well as the potential difference d_ψ are stored on the edges. The sum as well as the difference operations change the locality. The expression within the brackets gives the kind of traversal; `vertex_edge` iterates over all edges which cover a vertex, whereas `edge_vertex` iterates over all vertices which are covered by an edge. The formula in the square brackets are evaluated on all elements of the traversal and accumulated. `bern` denotes the Bernoulli function $\mathcal{B}$ which is used in the Scharfetter-Gummel current relation. From this specification the Jacobian matrix and the right hand side are assembled automatically.

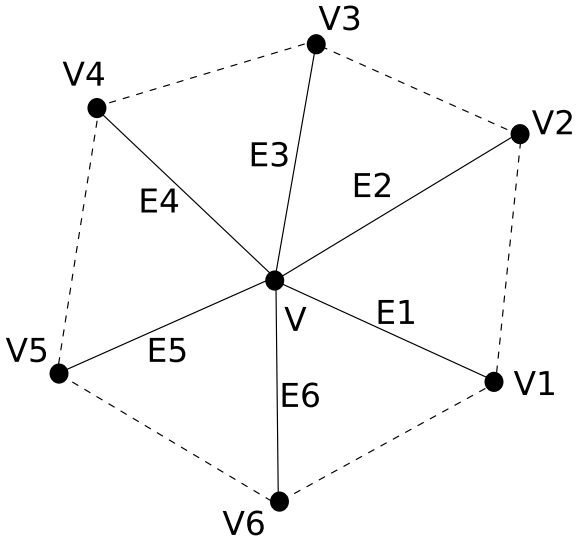

Figure 27: The local patch of a vertex. We outlined the vertices as well as the edges.

An automatic derivation of the linearized function, which is very error prone if done by hand, can be achieved by using an associative data structure. This data structure contains matrix entries as well as a right hand side value. For structures which are called linearized equations, we have implemented basic

operations such as addition and multiplication as well as other numerical functions such as the already demonstrated Bernoulli function. This method can be used in order to assemble linear as well as nonlinear equations which include discrete couplings between single topological elements. The complete application does not need more than 100 lines of code, the core is only about 25 of these lines.

4.4 Process Simulation

In general, three-dimensional process simulation steps need special surface treatment and must provide the ability to handle surface elements of arbitrary complexity containing degenerated or even faulty elements.

For the solution of problems in process simulation as well as interconnect simulation finite element methods are commonly used. In the following example we apply the finite element method to the Laplace equation in order to calculate capacitances as well as resistances of interconnect structures. In the following we will calculate two simple structures in order to evaluate the correctness of our simulation and to show the performance of the environment. The example is a single interconnect line. Even though it is very simple it shows the applicability of the GSSE as well as the performance.

4.5 Runtime Efficiency

To compare the runtime efficiency of the generic scientific environment we used the fastest (Poisson only) simulation tools from our institute (SAP) and run different application benchmarks with an automatic benchmark system.

As our environment is on a high semantic level and also does not impose a high abstraction penalty it is easy to make special optimizations if simpler partial differential equations have to be discretized:

- There is only one solution quantity
- We only apply one kind of differential equation

Even though these conditions are not always met we can gain performance if any of them is fulfilled. The genericity of our environment allows us to use these simplifications for all kinds of differential equations.

For the testing of the finite element code we divide the simulation time into the following parts. Each of the parts is measured independently in order to show the differences.

- Preparation (I/O, memory allocation)
- Assembly (Jacobian matrix and RHS)
- Solution

Our performance test example has to be very simple because we have to eliminate implication which result from problems with complicated geometries. Our test sample is a simple via line with a length of 10μm and a square cross-section of 1μm^2. For the test case we apply a potential difference of 1V. For both examples we use meshes which have been generated with `vgmodeler` [36] with 8.700, 14.000 and 97.000 tetrahedra. The comparison of the run time on an Intel Pentium 4 with 2.80GHz shows the following results (Figure 28, Figure 29, Figure 30):

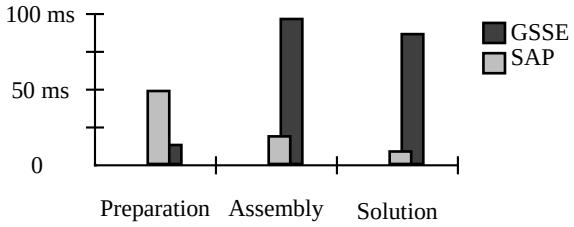

Figure 28: Structure with 8.700 tetrahedra and 2.300 vertices.

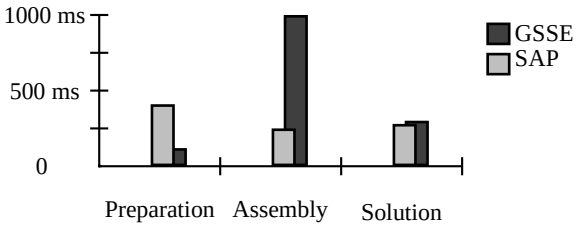

Figure 29: Structure with 64.000 tetrahedra and 14.000 vertices.

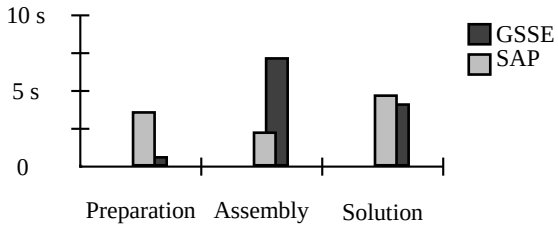

Figure 30: Structure with 520.000 tetrahedra and 97.000 vertices.

The benchmarks (Figure 28, Figure 29, Figure 30) show that the GSSE does not have a high time consumption for the preparation of the quantities. However in assembly time as well as in solution time the highly optimized code is faster but within the same order of magnitude. The solution time shows that the TRILINOS solver package is well designed for large matrices.

4.6 Conclusion

A generic environment for scientific computing has been presented. It can handle a large variety of differential equations which can be specified with different discretization schemes such as finite elements, finite volumes and finite differences. We have shown that a high semantic level does not necessarily imply an abstraction penalty so that the performance is comparable to highly optimized programs.

Even though we have shown the applicability of our environment on very simple structures it is possible to extend the features very conveniently. First the simulation domain can be taken from any meshing output. The partial differential equations can be extended to more complex models using a C++ embedded language.

Using this environment it is possible for scientists to formulate PDE problems with a full topological and geometrical support for the development of applications with minimal in-depth knowledge of internal data structures.

5 Current Transport in Nanoelectronic Semiconductor Devices

An overview of models used for the simulation of current transport in nanoelectronic devices within the framework of TCAD applications is presented. Modern enhancements of semiclassical transport models based on microscopic theories as well as quantum mechanical methods used to describe coherent and dissipative quantum transport are specifically addressed. This comprises the incorporation of quantum correction and tunneling models up to dedicated quantum-mechanical simulators, and mixed approaches which are capable to account for both, quantum interference and scattering. Specific TCAD requirements are discussed from an engineer's perspective and an outlook on future research directions is given.

5.1 Introduction

The breathtaking increase in computational power and speed of integrated circuits in the past decades has been supported by the aggressive size reduction of semiconductor devices. This trend is expected to continue in the coming decade as predicted and institutionalized by the International Technology Roadmap for Semiconductors [46]. Today, when the 90 nm technology node with physical transistor gate lengths in the range of 40 nm is in mass production, the challenge is to introduce the 65 nm technology node already in a year. A new technology node is introduced every 3 years, with a long-term projection of the 22 nm node to be in mass production by the year 2016. A possibility to build metal-on-insulator field effect transistors (MOSFETs) with even shorter gate lengths has been successfully established after the 6 nm gate length transistor has been demonstrated in research labs [47, 48].

From a theoretical viewpoint even a few nm gate length device has been predicted to be functional [49, 50]. Nevertheless, emerging outstanding technological challenges related to different aspects of MOSFET fabrication and reliability in mass production, as well as the rapidly increasing power dissipation may slow down the so far exponential scaling of Complimentary MOSFETs (CMOS). Besides, with the ongoing search for new technological solutions vital for CMOS downscaling, developing conceptually new devices and architectures is becoming increasingly important. New nanolelectronic structures, such as carbon nanotubes, nanowires, and molecules, are considered to be the most prominent candidates for the post-CMOS era. Since conventional MOSFETs are already operating in the sub-100 nm range, new nanoelectronic devices are expected to complement and substitute some of the current CMOS functions after being integrated into CMOS technology.

Technology CAD (TCAD) tools are designed to assist in development and engineering at all stages ranging from process simulation to device and circuit optimization. The main purpose of TCAD is the technology-development related cost reduction which currently amounts to 35% and is expected to rise to 40%, according to ITRS [46]. Due to the aggressive downscaling of CMOS device feature sizes and newly emerging nanoelectronic devices, various shortcomings of presently applied TCAD tools appear. These tools are frequently based on semiclassical macroscopic transport models. From an engineering point of view, classical models like the drift-diffusion model, have enjoyed an amazing success due to their relative simplicity, numerical robustness, and the ability to perform two- and three-dimensional simulations on large unstructured grids [51]. Hot-carrier effects have motivated the development of higher-order transport models such as the hydrodynamic, energy-transport and six-moments models [44]. However, inaccuracies originate from the non-local nature of carrier propagation in ultra-scaled devices [52].

Non-local effects may be of classical or quantum-mechanical nature, depending on the underlying physics relevant to the transport process. Classical non-localities appear when the mean-free path is comparable to the device feature size. Quantum mechanical non-local effects start to determine the transport properties when the device size is of the order of the De-Broglie electron wave length. Size quantization of carrier motion in inversion layers

of MOSFETs and in ultra-scaled multi-gate devices as well as the tunneling current, including the gate leakage current, are the most important examples of quantum effects in MOSFETs.

Figure 31 shows the hierarchy and mutual interrelation of models currently used for the description of current transport. Semiclassical transport models are based on the Boltzmann equation which includes scattering integrals describing realistic microscopic processes. These semiclassical models, augmented with quantum corrections, are still of great importance due to their relative computational simplicity, numerical stability, and an ability to provide reasonable quantitative results within seconds even for devices with gate length as short as 50 nm. A brief overview of the currently developed semiclassical transport models will be presented in Section 2.

Quantum ballistic transport models describe a coherent propagation of carriers. They are based on the solution of the SCHRÖDINGER equation for the wave function, supplemented with the corresponding boundary conditions. This approach is efficient and provides accurate results when carrier scattering is irrelevant and can be neglected. The method will be illustrated in Section 3 with an example of transport in carbon nanotubes [53].

Finally, dissipative quantum transport theory represents the most complete description of transport, which combines the coherent carrier motion between the scattering events with coherence (or phase) breaking due to carrier scattering. Different formalisms are currently used, based on the Dyson equation for the non-equilibrium Green's functions, the Liouville/Von-Neumann equation for the density matrix, or the Wigner transport equation. Section 4 deals with quantum transport characterized by both scattering and quantization. A conclusion will summarize the main findings and give directions for future research.

5.2 Semiclassical Transport

After the ground-breaking work of Scharfetter and Gummel [54], who first proposed a robust discretization scheme for the drift-diffusion equation, computer programs like MINIMOS [55] and PISCES [56] played a pioneering role in numerical simulation of current transport properties of semiconductor devices. Since then, numerous transport models of increasing complexity have been introduced. The semiclassical transport description is based on the Boltzmann equation for the distribution of carriers $f(\mathbf{r},\mathbf{k},t)$ in the phase space. The Boltzmann equation includes carriers' scattering with phonons, impurities, interfaces, and other scattering sources through the corresponding collision integrals.

Although the solution of the Boltzmann equation can be found numerically by means of Monte Carlo (MC) methods, TCAD models based on moments of the distribution function are highly desirable. Being computationally significantly less expensive than the MC method, these higher-order moments' methods provide a reasonable quantitative answer for devices as short as 50 nm within seconds. The fairly new six moments model [44], based on non-Maxwellian distribution takes naturally into account the hot-carrier effects such as avalanche generation, hot carrier induced gate currents, or hot-carrier diffusion, which typically take place in Silicon-On-Insulator (SOI) floating body MOSFETs. For the purpose of calibration the full-band MC method is often accepted, since it can precisely account for the various scattering processes [57].

Another important development of transport models is related to the MC methods for solving the Boltzmann equation. After the pioneering work of Kurosawa in 1966 [58], who was the first to apply the MC method to simulate carrier transport in semiconductors, the significantly improved MC method was successfully applied to transport description in a variety of semiconductors [59]. For electrons in silicon, the most thoroughly investigated case, it is believed that a satisfactory understanding of the band structure and of the basic scattering mechanisms has been achieved giving rise to a "standard model" [60]. Nowadays, an accurate MC evaluation of carrier transport properties in inversion layers is of primary importance for predicting performance of modern CMOS bulk devices. Due to the strong confinement of carriers in the inversion layer of bulk MOSFETs or due to the geometric confinement in multi-gate FETs the carrier motion is quantized in one or two confinement directions giving rise to the formation of subbands. One possibility to address the effect of quantum confinement on the electron concentration is to use an effective potential. This can be achieved by a convolution of the electrostatic potential with a GAUSSian function, which leads to a smoothing of the original potential [61], [62], [63]. Another option is to use the self-consistent Poisson-SCHRÖDINGER-based quantum corrected potential [64], [65], which suppresses the carrier concentration close to the interface, mimicking the real quantum-mechanical behavior. These approaches combine advantages of full-band structure and flexibility of scattering processes of three-dimensional classical MC simulations with the generality of material composition and transport peculiarities due to quantum confinement and may also address the strain effects.

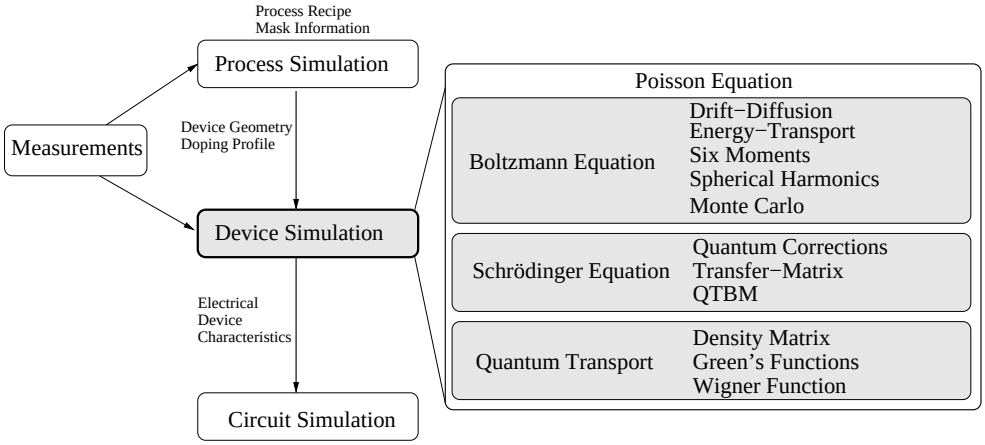

Figure 31: Schematic classification of approaches used in semiconductor current transport modeling.

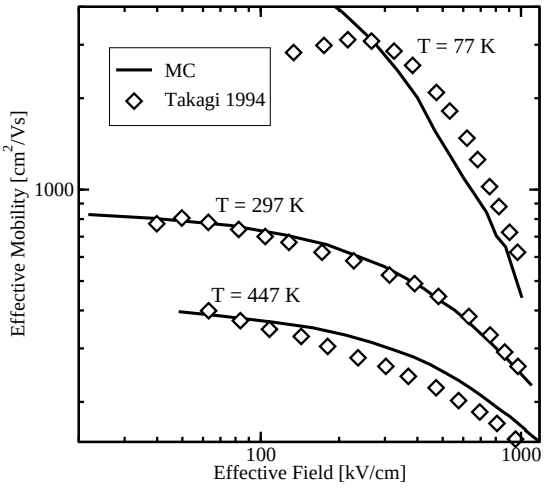

Figure 32: Comparison of subband MC simulations with the experimental [66] universal mobility of surface layer in silicon. The deviation of the experimental mobility from simulations at low effective fields is due to Coulomb scattering not included in the MC simulations.

The MC approach may incorporate the quantized carrier motion in the direction orthogonal to the current exactly. The quantum-mechanical motion of carriers in the confined direction is addressed by the self-consistent solution of the corresponding SCHRÖDINGER and Poisson equation, leading to the formation of subbands. The carrier motion within each subband may still be considered semiclassical and therefore can be well described by the corresponding Boltzmann equation written for the subband distribution function $f_n(\mathbf{r},\mathbf{k},t)$. Because of possible carrier transitions between different subbands due to scattering, the collision integrals on the right-hand-side of the Boltzmann equation should include the terms responsible for the intersubband scattering processes. The transport in the inversion layer of a MOSFET is finally described by a set of Boltzmann equations for every subband, coupled to each other via the intersubband scattering integrals. The set of the subband Boltzmann equations for $f_n(\mathbf{r},\mathbf{k},t)$ is conveniently solved by a MC method. This approach therefore combines the advantages of a quantum description in confinement direction with a semiclassical description in transport direction and represents a transition between semiclassical and quantum-mechanical pictures. An example of the simulation of the low-field surface mobility in inversion layers of silicon, when the transport in the current direction may be treated semiclassically is shown in Figure 32, together with the experimental "universal mobility" curve [66]. In order to reproduce the universal mobility curve, up to 40 unprimed and 20 primed subbands formed at a (100) silicon interface were taken into account, with realistic electron-phonon and surface roughness scattering included [67].

5.3 Quantum-Ballistic Transport

With the aggressive downscaling of MOSFET dimensions continuing, the classical description of carrier motion in transport direction is gradually loosing its validity. When the characteristic scale of the potential variation along the channel is comparable to the De-Broglie wave length of a carrier, a TCAD transport model must include the quantum effects in transport direction. If scattering processes can be ignored and particle propagation in the device is coherent, the carrier motion is determined by the solution of the SCHRÖDINGER equation,

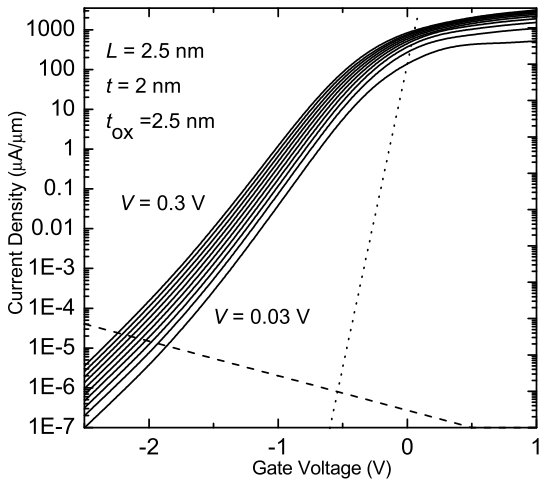

Figure 33: Subthreshold characteristics for a double-gate MOSFET with silicon thickness of 2 nm, gate length of 2.5 nm and the oxide thickness of 2.5 nm [49]. The dotted line corresponds to the ideal 60 mV/decade subthreshold slope. The dashed line shows the leakage current.

supplemented with open boundary conditions. In order to determine the current density J, it is enough to know the transmission coeffi cient $TC(\mathcal{E})$ as well as the supply function $N(\mathcal{E}_x)$ from the electrodes [68]:

$$J = \frac{4\pi m_{\text{eff}} q}{h^3} \int_{\mathcal{E}_{\min}}^{\mathcal{E}_{\max}} TC(\mathcal{E}_x) N(\mathcal{E}_x)\, \mathrm{d}\mathcal{E}_x. \qquad (34)$$

A similar approach can also be used to determine the gate leakage current [69]. The solution of the SCHRÖDINGER equation with open boundary conditions can be achieved by means of the quantum transmitting boundary method [70, 71]. An established alternative framework for these calculations is the non-equilibrium Green's Function method [72] in its reduced coherent version. It is conveniently used for one-dimensional studies of resonant tunneling diodes or carbon nanotubes. Simulators accounting for a full two-dimensional solution of the open-boundary SCHRÖDINGER equation have been reported and applied to the simulation of 10 nm double-gate MOSFETs [73, 74].

It may appear that in the quantum-ballistic case the determination of the full wave function as a solution of the SCHRÖDINGER equation is not necessary and the knowledge of the transmission coeffi cient is enough for the current calculations. In the contact block reduction method [75] the transmission function is fully determined by the reduced contact part of the full Green's function. However, the carrier concentration alters the electrostatic potential in the device via the POISSON equation. The carrier concentration is proportional to the square of the wave function, implying that the accurate determination of the transmission coeffi cient and therefore the current requires a self-consistent solution of the SCHRÖDINGER and POISSON equation simultaneously. For quasi-one dimensional transport this can be achieved straightforwardly [49]. An example of the output characteristics simulated for an ultra-thin body double-gate MOSFET with a gate length L as short as 2.5 nm is shown in Figure 33. Surprisingly, even such a small device possesses an $I_{\text{on}}/I_{\text{off}}$ ratio suffi cient for logic applications and displays a reasonable short-channel effect and acceptable DIBL, a conclusion recently reached from more detailed atomistic calculations [50]. It should be noted that the sensitivity to small MOSFET dimension variations, control of doping as well as the whole manufacturing process development represent signifi cant challenges for multi-gate MOSFETs with gate length below 10 nm.

Self-consistent solution of the two- or three-dimensional SCHRÖDINGER equation together with the POISSON equation represents a signifi cant computational challenge [73]. Two- and three-dimensional quantum ballistic simulations can be performed by means of an approximate separation of the quantum motion in the confi nement direction y from the motion along the current direction x by means of the following ansatz for the wave function $\Psi_n(y,x)$:

$$\Psi(x,y) = \sum_n \xi_n(x)\psi_n(y,x). \qquad (35)$$

This method allows the independent solution of the SCHRÖDINGER equation for the subband wave function $\psi_n(y,x)$ at position x. Transport in the current direction is characterized by a system of one-dimensional SCHRÖDINGER equations with open boundary conditions for the wave functions $\xi_n(x)$. Each SCHRÖDINGER equation describes the transport inside the particular subband. Transport in each subband is independent from the one in other subbands, if the subband wave functions $\psi_n(y)$ do not depend on the position x in transport direction.

The SCHRÖDINGER equations describing the transport in each subband are decoupled from each other, when the potential $U(x,y)$ in the device is the sum of two contributions, each depending either on y or x coordinate alone. In a general case when the subband wave functions depend on the position x in transport direction, the transport in the subbands n and m is coupled, with the coupling described by the Hamiltonian $\delta H_{nm}(x)$. However, when the intersubband coupling Hamiltonian $\delta H_{nm}(x)$ is small and may be neglected, transport in the

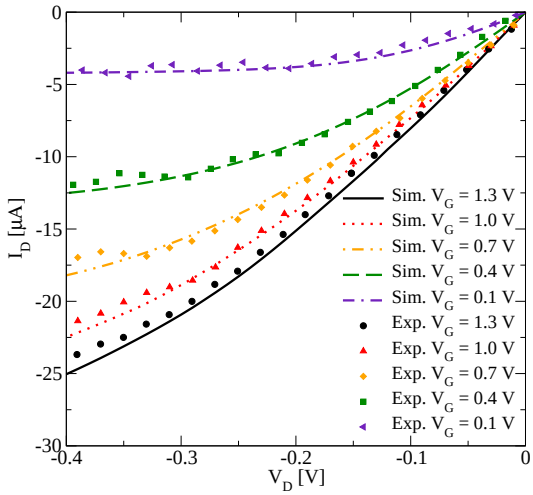

Figure 34: Output characteristics of CNT-FET with ohmic contacts [81] compared to experimental data [**?**].

subbands can still be considered as independent from each other. This approximation simplifies the calculations and reduces the computational effort significantly [76, 77, 78, 79]. The coupling Hamiltonian is expected to be small if the dependence of the subband wave function on x is weak. An example where the subband decomposition turns out to be an excellent approximation is the quantum transport in ultra-scaled SOI MOSFETs [79]. In the opposite limit of abrupt junctions between contact reservoirs and the channel, the intersubband coupling is expected to be the strongest. However, even in this case the current value calculated self-consistently was found to be only 10% lower as compared to the calculations with neglected intersubband coupling [80]. More study is needed to clarify the situations.

The coherent quantum transport description is justified if the size of the channel region is shorter that the phase-coherence length. In carbon nanotubes, where elastic scattering can be ignored and inelastic scattering has little effect on current [82], transport can be considered to be coherent and is therefore well described within the quantum-ballistic approach [53]. Similar methods can be applied to describe the output characteristics of FinFETs in the ballistic approximation [83]. In silicon MOSFETs, however, the mean-free path in the area close to the potential maximum at 300 K is only a few nm [84], and the full quantum description which includes dissipative processes must be adopted to simulate MOSFETs with gate length of around 10 nm. A consistent introduction of realistic scattering into simulators based on the coherent description alone creates outstanding computational difficulties ranging from a necessity to invert huge matrices in NEGF formalism [72] to calculations of nonlocal scattering rates in Pauli master equation approaches [85]. Besides the difficulties of introducing scattering into the simulators based on the coherent description, these simulators are often limited to specific geometries, grids and short length scales, which makes their integration into modern engineering TCAD tools problematic. Nevertheless, these simulation approaches are necessary for the estimation of upper bounds of current transport at the quantum limit.

5.4 Dissipative Quantum Transport

The methods described so far are either based on the assumption of semiclassical or pure quantum ballistic transport. Nevertheless, in modern microelectronic devices quantum effects are usually dominant in a small active region connected to large, heavily doped contact areas where the carrier dynamics is essentially classical. Therefore, modern TCAD simulators should be able to incorporate both classical and quantum-mechanical modeling approaches on equal footing. To a certain extent, various quantum corrections can account for the quantum effects, as already discussed.

The non-equilibrium Green's functions method addresses the problem in the most consistent and complete way. Due to its completeness, the method is computationally complex and is usually applied to one-dimensional problems [72] and for a restricted set of scattering mechanisms [86] only. The carbon nanotube (CNT) FET which is widely considered to be a potential alternative to the conventional MOSFETs, represents a good example where the nonequilibrium Green's function method provides accurate results and is successfully used. Simulated output characteristics of a CNT-FET with ohmic contacts [81] are compared in Figure 34 to experimental data [**?**], showing good agreement.

An alternative approach which handles both quantum-mechanical and dissipative scattering effects is based on the Wigner function formalism. Realistic scattering processes can be easily embedded into the Wigner equation via Boltzmann-like scattering integrals. The Wigner function is given by the density matrix in mixed representation [87, 88]. The kinetic equation for the Wigner function is similar to the Boltzmann equation, with the classical force term replaced by a non-local potential operator. The Wigner function formalism treats scattering and quantum mechanical effects on equal footing

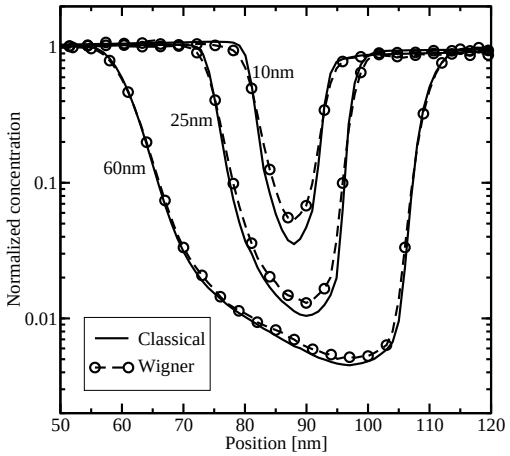

Figure 35: Comparison of Wigner and classical Monte Carlo results for electron concentration in double-gate MOSFETs, for different channel lengths [79].

through the corresponding scattering integrals. It is attractive to use the well established scattering models used in classical MC simulations and solve the quantum WIGNER transport equations by means of the MC technique. Such programs were recently realized in [89, 90, 79]. However, since the kernel of the quantum scattering operator is not positively defined, the numerical weight of a particle trajectory increases rapidly, and the numerical stability of a trajectory-based MC algorithm becomes a critical issue.

A multiple trajectories method was recently suggested [90] in order to overcome the difficulty. In this algorithm the problem of a growing statistical weight of a single trajectory is addressed by creating an increasing number of trajectories with constant weights, which may assume positive and negative values. Being formally equivalent to the former method, the algorithm allows the annihilation of trajectories with similar statistical properties, introducing a possibility to control the number of trajectories. This approach has the advantage that a seamless transition between classical and quantum-mechanical regions in a device is possible [90]. Following [91], one can introduce a spectral decomposition of the potential profile $V(x)$ into a slowly varying, classical component and a rapidly varying, quantum mechanical component. This decomposition is conveniently carried out by applying a low-pass filter with a cut-off wave number $q_c \ll \pi/\Delta x$, where Δx is a grid step size. This separation of the total potential into a classical and a quantum mechanical contribution significantly improves the Wigner Monte Carlo convergence. The method can be applied to the simulation of resonant tunneling diodes [90], and it was recently used for the simulation of double-gate MOSFETs [79]. An example of the carrier concentration simulated with the Wigner function method is shown in Figure 35. in comparison with the classical result.

5.5 Conclusions

Well established classical TCAD tools are gradually loosing their ability to predict accurately the characteristics of nanoscale devices, prompting for enhancement to meet the engineering demands. Classical models using higher moments are able to include the hot-carrier effects and can reproduce results of the full-band MC, while relevant quantum corrections may be incorporated into different MC schemes.

Full quantum description is required for nanoscale devices. Contrary to the carbon nanotubes, where the transport properties can be well predicted within the coherent picture, a dissipative quantum description may be required for transport calculations in ultra-scaled MOSFETs with the gate lengths ranging around 10 nm. One option is the Wigner function approach which naturally combines the advantages of quantum description with the accurate scattering models relevant for devices in the nanoscale range. All quantum-mechanical models must be adapted for engineering applications for which timely results are more valuable than accurate analyses [92]. Device simulators must allow a coupling with process simulators, since a detailed, physics-based transport model may be misleading if geometry and doping are not described correctly. Support of unstructured grids is necessary in oder to optimize the simulation time. Furthermore, the simulators should be general-purpose and not limited to specific geometries or simulation models. It is still not clear which of the outlined quantum transport approaches will finally be integrated into TCAD environments. Its further success will depend on the ability to model quantum effects efficiently with reasonable accuracy.

References

[1] J. Welser, J. L. Hoyt, and J. F. Gibbons. NMOS and PMOS Transistors Fabricated in Strained Silicon/Relaxed Silicon-Germanium Structures. In *Proc.IEDM Tech.Dig*, pp 1000–1002, 1992.

[2] K. Rim, J. L. Hoyt, and J. F. Gibbons. Transconductance Enhancement in Deep Submicron Strained Si N-MOSFETs. In *Proc.IEDM Tech.Dig*, pp 707–710, 1998.

[3] J. L. Hoyt et al. Strained Silicon MOSFET Technology. In *Proc.IEDM Tech.Dig*, pp 23–26, 2002.

[4] A. Lochtefeld and D. Antoniadis. Investigating the Relationship Between Electron Mobility and Velocity in Deeply Scaled NMOS via Mechanical Stress. *IEEE Electron Device Lett.*, 22(12):591–593, 2001.

[5] S. Maikap, C. Y. Yu, S. R. Jan, M. H. Lee, and C. W. Liu. Mechanically Strained Si NMOSFETs. *IEEE Electron Device Lett.*, 25(1):40–42, 2004.

[6] A. Shimizu et al. Local Mechanical-Stress Control (LMC) : A New Technique for CMOS Performance Enhancement. In *Proc.IEDM Tech.Dig*, pp 433–437, 2001.

[7] S. Ito et al. Mechanical Stress Effect of Etch-Stop Nitride and its Impact on Deep Submicron Transistor Design. In *Proc.IEDM Tech.Dig*, pp 247–251, 2000.

[8] K. Uchida, R. Zednik, C. H. Lu, H. Jagannathan, J. McVittie, P. C. McIntyre, and Y. Nishi. Experimental Study of Biaxial and Uniaxial Strain Effects on Carrier Mobility in Bulk and Ultrathin-Body SOI MOSFETs. In *Proc.IEDM Tech.Dig*, pp 229–232, 2004.

[9] M. M. Rieger and P. Vogl. Electronic-Band Parameters in Strained $Si_{1-x}Ge_x$ Alloys on $Si_{1-y}Ge_y$ Substrates. *Physical Review B*, 48(19):14276–14287, 1993.

[10] S. Dhar, H. Kosina, V. Palankovski, E. Ungersboeck, and S. Selberherr. Electron Mobility Model for Strained-Si Devices. *IEEE Trans.Electron Devices*, 52(4):527–533, 2005.

[11] Institut für Mikroelektronik, Technische Universität Wien, Austria. *MINIMOS-NT 2.1 User's Guide*, 2005. http://www.iue.tuwien.ac.at/software.

[12] A. R. Brown, C. P. Jarrett, D. M. de Leeuw, and M. Matters. Field-effect transistors made from solution-processed organic semiconductors. *Synthetic Metals*, 88(1):37–55, 1997.

[13] M. C. Vissenberg and M. Matters. Theory of the field-effect mobility in amorphous organic transistors. *Phys.Rev.*, 57(20):12964–12967, 1998.

[14] G. E. Pike and C. H. Seager. Percolation and conductivity: A computer study.I. *Phys.Rev.*, 10(4):1421–1434, 1974.

[15] E. J. Meijer, C. Tanase, P. W. Blom, E. van Veenendaal, B. H. Huisman, D. M. de Leeuw, and T. M. Klapwijk. Switch-on voltage in disordered organic field-effect transistors. *Appl.Phys.Let*, 80(20):3838–3840, 2002.

[16] E. Calvetti, L. Colalongo, and Zs. M. Kovacs-Vajna. Organic thin transistors: a DC/dynamic analytical model. *Synthetic Metals*, 49:567–577, 2005.

[17] J. Appenzeller, M. Radosavljevic, J. Knoch, and Ph. Avouris. Tunneling Versus Thermionic Emission in One-Dimensional Semiconductors. *Physical Review Letters*, 92:048301, 2004.

[18] A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai. Ballistic Carbon Nanotube Field-Effect Transistors. *Letters to Nature*, 424(6949):654–657, 2003.

[19] M. Pourfath, E. Ungersboeck, A. Gehring, B. Cheong, W. Park, H. Kosina, and S. Selberherr. Improving the Ambipolar Behavior of Schottky Barrier Carbon Nanotube Field Effect Transistors. In *Proc. ESSDERC*, pp 429–432, 2004.

[20] M. Pourfath, A. Gehring, E. Ungersboeck, H. Kosina, S. Selberherr, B.-H. Cheong, and W. Park. Separated Carrier Injection Control in Carbon Nanotube Field-Effect Transistors. *J.Appl.Phys.*, 97:1061031–1061033, 2005.

[21] J. Guo, S. Datta, and M. S. Lundstrom. A Numerical Study of Scaling Issues for Schottky-Barrier Carbon Nanotube Transistors. *IEEE Trans.Electron Devices*, 51(2):172–177, 2004.

[22] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R.G. Gordon, M. Lundstrom, and H.J. Dai. Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays. *Nano Letters*, 4(7):1319–1322, 2004.

[23] M. Pourfath, H. Kosina, B. H. Cheong, W. Park, and S. Selberherr. Improving DC and AC Characteristics of Ohmic Contact Carbon Nanotube Field Effect Transistors. In *Proc. ESSDERC*, pp 541–544, 2005.

[24] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic. Simulation of Quantum Transport in Nanoscale Transistors: Real versus Mode-Space Approach. *J.Appl.Phys.*, 92(7):3730–3739, 2002.

[25] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal. Two-Dimensional Quantum Mechanical Modeling of Nanotransistors. *J.Appl.Phys.*, 91:2343–2354, 2002.

[26] G. D. Mahan. Electron-Optical Phonon Interaction in Carbon Nanotubes. *Physical Review B*, 68:125409, 2003.

[27] J.-S. Park, H. Shin, D. Connelly, D. Yergeau, Z. Yu, and R. W. Dutton. Ananlysis of 2-D Quantum Effects in the Poly-Gate and their Impact on the Short-Channel Effects in Double-Gate MOSFETs via the Density-Gradient Method. *Solid-State Electron.*, 48(7):1163–1168, 2004.

[28] R. Saito, G. D. Dresselhaus, and M. S. Dresselhaus. *Physical Properties of Carbon Nanotubes*. Imperial College Press, 1998.

[29] P. J. Burke. AC Performance of Nanoelectronics: Towards a Ballistic THz Nanotube Transistors. *Solid-State Electron.*, 48(10-11):1981–1986, 2004.

[30] D. L. John, L. C. Castro, P. J. S. Pereira, and D. L. Pulfrey. A Schrödinger-Poisson Solver for Modeling Carbon Nanotube FETs. In *Proc. Nanotech 2004*, 2004.

[31] P. J. Burke. An RF Circuit Model for Carbon Nanotubes. *IEEE Trans.Nanotechnology*, 2(1):55–58, 2003.

[32] M. Radosavljevic, S. Heinze, J. Tersoff, and Ph. Avouris. Drain Voltage Scaling in Carbon Nanotube Transistors. *Appl.Phys.Lett.*, 83(12):2435–2437, 2003.

[33] J. Guo and M. Lundstrom. Role of Phonon Scattering in Carbon Nanotube Field-Effect Transistors. *Appl.Phys.Lett.*, 86:193103, 2005.

[34] J. Guo. A Quantum-Mechanical Treatment of Phonon Scattering in Carbon Nanotube Transistors. *J.Appl.Phys.*, 98:063519, 2005.

[35] J. Guo and M. S. Lundstrom. A Computational Study of Thin-Body, Double-Gate, Schottky Barrier MOSFETs. *IEEE Trans.Electron Devices*, 49(11):1897–1902, 2002.

[36] R. Heinzl and T. Grasser. Generalized Comprehensive Approach for Robust Three-Dimensional Mesh Generation for TCAD. In *Proc. SISPAD*, pp 211–214, Tokyo, 2005.

[37] P. Schwaha, R. Heinzl, M. Spevak, and T. Grasser. Coupling Three-Dimensinoal Mesh Adaptation with an A Posteriori Error Estimator. In *Proc. SISPAD*, pp 235–238, Tokyo, 2005.

[38] R. Sabelka and S. Selberherr. SAP — A Program Package for Three-Dimensional Interconnect Simulation. In *Proc. Intl. Interconnect Techn. Conf.*, pp 250–252, Burlingame, California, 1998.

[39] A. Hössinger, R. Minixhofer, and S. Selberherr. Full Three-Dimensional Analysis of a Non-Volatile Memory Cell. In *Proc. SISPAD*, pp 129–132, Munich, 2004.

[40] T. L. Veldhuizen. C++ Templates as Partial Evaluation. In *Proc. of PEPM'99.*, pp 13–18. University of Aarhus, Dept. of Computer Science, 1999.

[41] J. G. Siek and A. Lumsdaine. The Matrix Template Library: A Unifying Framework for Numerical Linear Algebra. In *Proc. ECOOP Workshop*, pp 466–467, 1998.

[42] Corrado Böhm. Lambda-Calculus and Computer Science Theory. In *Springer Series Lecture Notes in Computational Sciene (LNCS)*, volume 37, Rome, 1975.

[43] IBM Corporation, Yorktown Heights, NY, USA. *IBM Visualization Data Explorer*, third edition, 1993.

[44] T. Grasser, T. W. Tang, H. Kosina, and S. Selberherr. A Review of Hydrodynamic and Energy-Transport Models for Semiconductor Device Simulation. *Proceedings of the IEEE*, 91(2):251–274, 2003.

[45] D. L. Scharfetter and H. K. Gummel. Large-Signal Analysis of a Silicon Read Diode Oscillator. *IEEE Trans.Electron Devices*, 16(1):64–77, 1969.

[46] International Technology Roadmap for Semiconductors - 2004 Update, 2004. http://public.itrs.net.

[47] B. Doris et al. Extreme Scaling with Ultra-Thin Si Channel MOSFETs. In *IEDM Techn. Dig.*, pp 267–270, 2002.

[48] H. Iwai. CMOS Downsizing Toward Sub-10 nm. *Solid-State Electron.*, 48(4):497–503, 2004.

[49] V. A. Sverdlov, T. J. Walls, and K. K. Likharev. Nanoscale Silicon MOSFETs: A Theoretical Study. *IEEE Trans.Electron Devices*, 50(9):1926–1933, 2003.

[50] L. Risch. Pushing CMOS Beyond the Roadmap. In *Proc. ESSDERC*, pp 63–68, Grenouble, 2005.

[51] S. Selberherr. *Analysis and Simulation of Semiconductor Devices*. Springer, 1984.

[52] F. O. Heinz, F. M. Bufler, A. Schenk, and W. Fichtner. Quantum Transport Phenomena and Their Modeling. In *Symposium on Nano Device Technology*, pp 2–8, Hsinchu, Taiwan, 2004.

[53] M. Pourfath, A. Gehring, E. Ungersbck, H. Kosina, S. Selberherr, B.-H. Cheong, and W. Park. Separated Carrier Injection Control in Carbon Nanotube Field-Effect Transistors. *J.Appl.Phys.*, 97(10):1061031–1061033, 2005.

[54] D. L. Scharfetter and H. K. Gummel. Large-Signal Analysis of a Silicon Read Diode Oscillator. *IEEE Trans.Electron Devices*, 16(1):64–77, 1969.

[55] S. Selberherr, W. Fichtner, and H. Pötzl. MINIMOS – A Program Package to Facilitate MOS Device Design and Analysis. In B. T. Browne and J. J. Miller, editors, *Numerical Analysis of Semiconductor Devices and Integrated Circuits*, volume I, pp 275–279, Dublin, 1979. Boole Press.

[56] M. R. Pinto. *PISCES IIB*. Stanford University, 1985.

[57] M. V. Fischetti and S. E. Laux. Monte Carlo Analysis of Electron Transport in Small Semiconductor Devices Including Band-Structure and Space-Charge Effects. *Physical Review B*, 38(14):9721–9745, 1988.

[58] T. Kurosawa. Monte Carlo Calculation of Hot Electron Problems. In *Proc. Intl. Conf. on Physics of Semiconductors*, pp 424–426, 1966.

[59] C. Jacoboni and L. Reggiani. The Monte Carlo Method for the Solution of Charge Transport in Semiconductors with Applications to Covalent Materials. *Reviews of Modern Physics*, 55(3):645–705, 1983.

[60] M.V. Fischetti and S.E. Laux. Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys. *J.Appl.Phys.*, 80(4):2234–2252, 1996.

[61] Y. Li, T.-W. Tang, and X. Wang. Modeling of Quantum Effects for Ultrathin Oxide MOS Structures with an Effective Potential. *IEEE Trans. Nanotechnology*, 1(4):238–242, 2002.

[62] K. Z. Ahmed, P. A. Kraus, C. Olsen, and F. Nouri. On the Evaluation of Performance Parameters of MOSFETs with Alternative Gate Dielectrics. *IEEE Trans.Electron Devices*, 50(12):2564–2567, 2003.

[63] D. Esseni P. Palestri, S. Eminente, C. Fiegna, E. Sangiorgi, and L. Selmi. An Improved Semi-Classical Monte-Carlo Approach for Nano-Scale MOSFET Simulation. *Solid-State Electron.*, 49:727–732, 2005.

[64] G. A. Kathawala, B. Winstead, and U Ravaioli. Monte Carlo Simulations of Double-gate MOSFETs. *IEEE Trans.Electron Devices*, 50(12):2467–2473, 2003.

[65] X.-F. Fan, X. Wang, B. Winstead, L. F. Register, U. Ravaioli, and S. K. Banerjee. MC Simulation of Strained-Si MOSFET With Full-Band Structure and Quantum Corrections. *IEEE Trans.Electron Devices*, 51(6):962–970, 2004.

[66] S.-I. Takagi, A. Toriumi, M. Iwase, and H. Tango. On the Universality of Inversion Layer Mobility in Si MOSFET's: Part I - Effects of Substrate Impurity Concentration. *IEEE Trans.Electron Devices*, 41(12):2357–2362, 1994.

[67] E. Ungersböck and H. Kosina. The Effect of Degeneracy on Electron Transport in Strained Silicon Inversion Layer. In *Proc. SISPAD*, pp 311–314, Tokyo, 2005.

[68] R. Tsu and L. Esaki. Tunneling in a Finite Superlattice. *Appl.Phys.Lett.*, 22(11):562–564, 1973.

[69] A. Gehring. *Simulation of Tunneling in Semiconductor Devices*. Dissertation, Technische Universität Wien, 2003.

[70] C. S. Lent and D. J. Kirkner. The Quantum Transmitting Boundary Method. *J.Appl.Phys.*, 67(10):6353–6359, 1990.

[71] W. R. Frensley. Numerical Evaluation of Resonant States. *Superlattices & Microstructures*, 11(3):347–350, 1992.

[72] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic. Single and Multiband Modeling of Quantum Electron Transport Through Layered Semiconductor Devices. *J.Appl.Phys.*, 81(12):7845–7869, 1997.

[73] S. E. Laux, A. Kumar, and M. V. Fischetti. Ballistic FET Modeling Using QDAME: Quantum Device Analysis by Modal Evaluation. *IEEE Trans. Nanotechnology*, 1(4):255–259, 2002.

[74] M. Sabathil, S. Hackenbuchner, J. A. Majewski, G. Zandler, and P. Vogl. Towards Fully Quantum Mechanical 3D Device Simulations. *J.Computational Electronics*, 1:81–85, 2002.

[75] D. Mamaluy, M. Sabathil, and P. Vogl. Efficient Method for the Calculation of Ballistic Quantum Transport. *J.Appl.Phys.*, 93(8):4628–4633, 2003.

[76] F. O. Heinz, A. Schenk, A. Scholze, and W. Fichtner. Full Quantum Simulation of Silicon-on-Insulator Single-Electron Devices. *J.Computational Electronics*, 1(1):161–164, 2002.

[77] G. Curatola, G. Fiori, and G. Iannaccone. Modeling and Simulation Challenges for Nanoscale MOSFETs in the Ballistic Limit. *Solid-State Electron.*, 48(4):581–587, 2004.

[78] J. Wang, E. Polizzi, A. Ghosh, S. Datta, and M. Lundstrom. Theoretical Investigation of Surface Roughness Scattering in Silicon Nanowire Transistor. 87:0431011–0431013, 2005.

[79] V. Sverdlov, A. Gehring, H. Kosina, and S. Selberherr. Quantum Transport in Ultra-Scaled Double-Gate MOSFETs: A Wigner Function-Based Monte Carlo Approach. *Solid-State Electron.*, 49(9):1510–1515, 2005.

[80] R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom. Quantum Mechanical Analysis of Channel Access Geometry and Series Resistance in Nanoscale Transistors. *J.Appl.Phys.*, 95(1):292–305, 2003.

[81] M. Pourfath, H. Kosina, B. Cheong, W. Park, and S. Selberherr. Improving DC and AC Characteristics of Ohmic Contact Carbon Nanotube Fiel Effect Transistors. In *Proc. ESSDERC*, pp 541–544, Grenouble, 2005.

[82] J. Guo and M. Lundstrom. Role of Phonon Scattering in Carbon Nanotube Field-Effect Transistors. *Appl.Phys.Lett.*, 86:193103–1–193103–2, 2005.

[83] X. Shao and Z. Yu. Nanoscale FinFET Simulations: A Quasi-3D Quantum MechanicalModel Using NEGF. *Solid-State Electron.*, 49:1435–1445, 2005.

[84] C. Jungemann, N. Subba, J.-S. Goo, C. Riccobene, Q. Xiang, and B. Meinerzhagen. Investigation of Strained Si/SiGe Devices by MC Simulation. *Solid-State Electron.*, 48(8):1417–1422, 2004.

[85] M. V. Fischetti. Theory of Electron Transport in Small Semiconductor Devices Using the Pauli Master Equation. *J.Appl.Phys.*, 83(1):270–291, 1998.

[86] A. Svizhenko and M. P. Anantram. Role of Scattering in Nanotransistors. *IEEE Trans.Electron Devices*, 50:1459–1466, 2003.

[87] E. Wigner. On the Quantum Correction for Thermodynamic Equilibrium. *Physical Review*, 40:749–759, 1932.

[88] H. Kosina and M. Nedjalkov. Wigner Function Based Device Modeling. In M. Rieth and W. Schommers, editors, *Handbook of Theoretical and Computational Nanotechnology*. Springer, 2005.

[89] L. Shifren, C. Ringhofer, and D. K. Ferry. A Wigner Function-Based Quantum Ensemble Monte Carlo Study of a Resonant Tunneling Diode. *IEEE Trans.Electron Devices*, 50(3):769–773, 2003.

[90] H. Kosina, M. Nedjalkov, and S. Selberherr. A Monte Carlo Method Seamlessly Linking Quantum and Classical Transport Calculations. *J. Computational Electronics*, 2(2–4):147–151, 2002.

[91] A. Gehring and H. Kosina. Wigner-Function Based Simulation of Quantum Transport in Scaled DG-MOSFETs Using the Monte Carlo Method. *J. Computational Electronics*, 4(1-2):67–70, 2005.

[92] M. Duane. TCAD Needs and Applications from a Users Perspective. *IEICE Trans.Electron.*, E82-C(6):976–982, 1999.