VISTA Status Report
December 2007

O. Baumgartner, W. Gös, A. Nentchev,
F. Stimpfl, V. Sverdlov, S. Selberherr

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

Contents

1 Adaptive Energy Integration of Non-Equilibrium Green's Functions

To obtain the physical quantities of interest within the non-equilibrium Green's function formalism, numerical integration over energy space is essential. Several adaptive methods have been implemented and tested for their applicability. The number of energy grid points needed and the convergence behavior of the Schrödinger-Poisson iteration have been evaluated. An adaptive algorithm based on a global error criterion proved to be more efficient than a local adaptive algorithm.

1.1 Introduction

For nanoscaled devices, numerical simulations based on the non-equilibrium Green's functions (NEGF) formalism are commonly performed [1–4]. A very efficient implementation of this method has been achieved by means of a recursive algorithm [5]. Proper numerical integration methods are vital for the stability and accuracy of NEGF simulations.

1.2 The NEGF Formalism

The retarded and advanced Green's functions are determined by the equation

$$G^{\mathrm{R}}(\mathbf{r}, \mathbf{r}', \mathcal{E}) = G^{\mathrm{A}\dagger}(\mathbf{r}, \mathbf{r}', \mathcal{E})$$
$$= [\mathcal{E}I - H(\mathbf{r}, \mathbf{r}', \mathcal{E}) - \Sigma^{\mathrm{R}}(\mathbf{r}, \mathbf{r}', \mathcal{E})]^{-1}, \quad (1)$$

where $H(\mathbf{r}, \mathbf{r}', \mathcal{E})$ is the Hamiltonian of the system and $\Sigma^{\mathrm{R}}(\mathbf{r}, \mathbf{r}', \mathcal{E})$ is the retarded self-energy. The less-than Green's function is calculated as

$$G^{<}(\mathbf{r}, \mathbf{r}', \mathcal{E}) = G^{\mathrm{R}}(\mathbf{r}, \mathbf{r}', \mathcal{E})\Sigma^{<}(\mathbf{r}, \mathbf{r}', \mathcal{E})G^{\mathrm{A}}(\mathbf{r}, \mathbf{r}', \mathcal{E}). \quad (2)$$

The lesser self energy of the left and right contact is given by $\Sigma_{\mathrm{l,r}}^{<}(\mathcal{E}) = \imath \Im \left\{ \Sigma_{\mathrm{l,r}}^{\mathrm{R}}(\mathcal{E}) \right\} f_{\mathrm{l,r}}(\mathcal{E})$ with the occupation function $f_{\mathrm{l,r}}(\mathcal{E})$ of the left and right lead, respectively. The Green's functions allow the calculation of physical quantities of interest such as the local density of states, $N(\mathbf{r}, \mathbf{r}, \mathcal{E}) = -\frac{1}{\pi}\Im \left\{ G^{\mathrm{R}}(\mathbf{r}, \mathbf{r}, \mathcal{E}) \right\}$, and the electron and current density

$$n(\mathbf{r}) = -2\imath \int G^{<}(\mathbf{r}, \mathbf{r}, \mathcal{E})\frac{d\mathcal{E}}{2\pi}, \quad (3)$$

$$j(\mathbf{r}) = -\frac{\hbar q}{m^{*}} \int \left[(\nabla - \nabla')G^{<}(\mathbf{r}, \mathbf{r}', \mathcal{E}) \right]\Big|_{\mathbf{r}'=\mathbf{r}} \frac{d\mathcal{E}}{2\pi}. \quad (4)$$

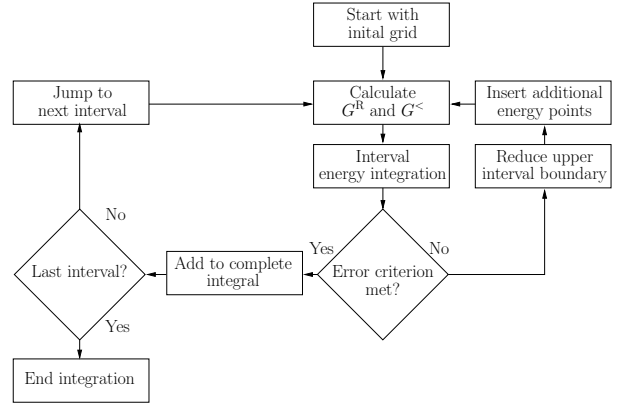For these quantities integration over energy is required.

Figure 1: Illustration of the adaptive integration algorithm with grid refinement. Grid points are added as long as the error criterion is not met.

1.3 Integration Methods

The numerical evaluation of these quantities require a discretization of the energy space. A simple approach using an equidistant energy grid suffers from two problems. A small number of grid points will not correctly resolve narrow resonances, whereas a vast number can lead to an unpredictable summation of numerical errors and to intractable memory requirements. These effects can yield instability or poor convergency of a self-consistent iteration loop [6]. Therefore adaptive energy integration (AEI) on a non-equidistant grid is required to increase accuracy, numerical stability, and memory efficiency. The following section outlines the different approaches that were implemented and tested for the applicability within the NEGF formalism.

1.3.1 Simpson's Rule

Simpson's rule is a closed Newton-Cotes rule of second order. The integral of a function $f(x)$ over an interval $[a, b]$ is given by

$$I_1 = \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]. \quad (5)$$

One strategy to decrease the interpolation error is to subdivide the interval into two equal parts and to apply Simpson's rule on each subinterval. This leads to the composite Simpson rule which, for five grid points, writes as

$$I_2 = \frac{b-a}{12}\left[f(a) + 4f\left(\frac{a+b}{4}\right) + 2f\left(\frac{a+b}{2}\right) \right.$$
$$\left. + 4f\left(3\frac{a+b}{4}\right) + f(b) \right]. \quad (6)$$

To obtain an error criterion for the adaptive integration algorithm (Fig. 1), the electron concentration within the current integration interval is calculated using the 3-point and the 5-point composite Simpson rule. This leads to a local error which is compared to the desired error tolerance τ

$$\left| \frac{I_1 - I_2}{I_2} \right| < \tau. \tag{7}$$

If this condition is satisfied, the integral is considered accurate enough and the grid in the given interval is not further refined.

1.3.2 Polynomial Interpolation

Simpson's rule is based on equidistant grid points and an interpolation polynomial of second order. As an alternative, a more general approach with non-equidistant grid points and polynomials of arbitrary degree can be considered. For a monomial power basis the interpolation polynomial on N nodes takes the form

$$p(x) = \sum_{i=1}^{N} a_i x^{i-1}. \tag{8}$$

To obtain the coefficient vector $\mathbf{a} = [a_1, a_2, a_3, \ldots, a_N]^T$ an equation system of rank N needs to be solved

$$\underbrace{\begin{bmatrix} 1 & x_1 & \cdots & x_1^{N-1} \\ 1 & x_2 & \cdots & x_2^{N-1} \\ \vdots & & \ddots & \vdots \\ 1 & x_N & \cdots & x_N^{N-1} \end{bmatrix}}_{\underline{V}} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \tag{9}$$

where $\underline{V}$ is called the Vandermonde matrix. Unfortunately this system is often ill-conditioned and its solution may become numerically unstable. Björck and Pereyra [7] developed an algorithm that is able to calculate the coefficient vector $\mathbf{a}$ in a fast and stable manner.

After the coefficients of the polynomial are obtained the integral of the interpolation function in the interval $[x_1, x_N]$ can be calculated. For an arbitrary odd number N of grid points, a subset of $(N+1)/2$ grid points may be used to obtain a second polynomial and consequentially a second approximation of the integral. These two results are then compared to yield the error criterion for the adaptive integration algorithm. Unfortunately, polynomial interpolation functions on equidistant points suffer from Runge's phenomenon for a higher degree. This can be avoided by using non-equidistant grid points as done by the Clenshaw-Curtis Rule described in the succeeding section.

1.3.3 Clenshaw-Curtis Integration

Fejér [8] proposed to use the zeros of the Chebyshev polynomial $T_n = \cos(n \arccos x)$ in the interval $]-1, 1[$ as quadrature points of the integral of $f(x)$,

$$\int_{-1}^{1} f(x)\mathrm{d}x = \sum_{k=0}^{n} w_k f(x_k). \tag{10}$$

For Fejér's second rule, the $n-1$ extreme points of T_n are used. Clenshaw and Curtis [9] extended this open rule to a closed form which includes the boundary points $x_0 = -1$ and $x_n = 1$ of the interval. The $n + 1$ quadrature points are

$$x_k := \cos(\vartheta_k), \quad \vartheta_k := k\frac{\pi}{n}, \quad k = 0, 1, \ldots, n. \tag{11}$$

The weights w_k in equation (10) are to be obtained by an explicit expression or by means of discrete Fourier transforms [10]. The explicit expressions of the Clenshaw-Curtis weights are:

$$w_k = \frac{c_k}{n} \left(1 - \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{b_j}{4j^2 - 1} \cos(2j\vartheta_k) \right). \tag{12}$$

The coefficients b_j and c_k are given by

$$b_j = \begin{cases} 1, & \text{if } j = n/2 \\ 2, & \text{if } j < n/2 \end{cases}, \tag{13}$$

$$c_k = \begin{cases} 1, & \text{if } k = 0 \mod n \\ 2, & \text{otherwise.} \end{cases} \tag{14}$$

A useful property of the Clenshaw-Curtis rule is the option to create subsets of the quadrature nodes. To move from $n + 1$ to $2n + 1$ points only n new function values need to be evaluated.

1.3.4 Extended Doubly Adaptive Quadrature Routine

So far the presented methods used a local error criterion for adaptive energy integration. A different approach, which comprises two refinement strategies, has been presented by Espelid [11]. A global error criterion is used to find the most erroneous subinterval. This interval is then treated locally either by subdivision and applying a smaller order Newton-Cotes rule, or by inserting additional energy grid points and using a higher order rule, depending on the estimated error. The local integral and error of the superior method for a given subinterval are then added to the global values. This procedure is repeated until the global error is below a given tolerance as depicted in Fig. 3.
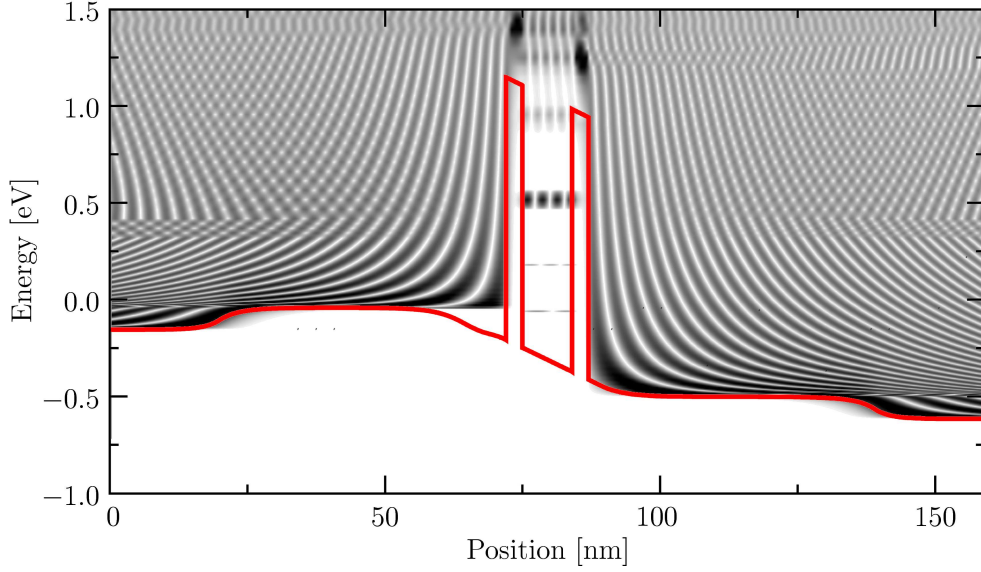
Figure 2: Local density of states of the resonant tunneling diode. The penetration into classically forbidden regions, the reflections at the barriers, as well as the resonant tunneling states within the quantum well can be seen clearly.

1.4   Results

The implemented energy integration algorithms were applied to an unbiased double barrier structure. The number of energy grid points needed to meet a given relative error τ is plotted in Figure 4. For the polynomial interpolation and the Clenshaw-Curtis method a relative error $\tau < 10^{-4}$ was needed to correctly resolve the resonance in the quantum well. Using Simpson's rule even $\tau < 10^{-5}$ is required to set enough grid points. Comparing the local adaptive procedures, the polynomial interpolation performs best considering the number of grid points. To evaluate the convergence behavior of a self-consistent band edge calculation as seen in Figure 2, the square of the potential update norm is plotted over the iteration number (Fig. 5). For the polynomial and the Clenshaw-Curtis method a local error criterion of $\tau = 10^{-6}$ has been chosen. For the global adaptive algorithm the relative error has been set to $\tau = 10^{-3}$ and $\tau = 10^{-5}$, respectively. All methods show similar good convergence, whereas the number of energy grid points differs considerably. The global adaptive method requires about half of the points of the polynomial interpolation and a third of the Clenshaw-Curtis integration.
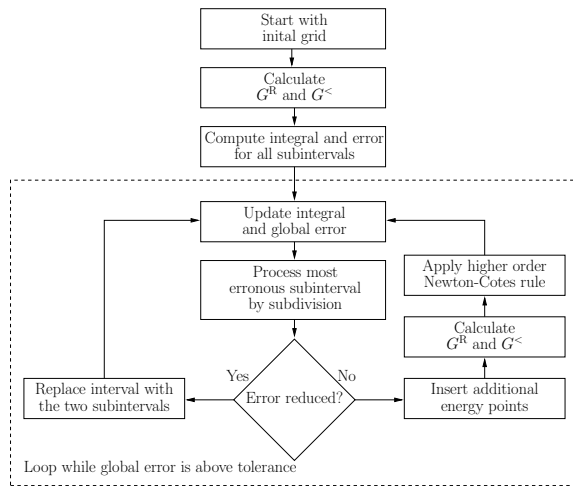

Figure 3: Illustration of the extended doubly adaptive quadrature routine.
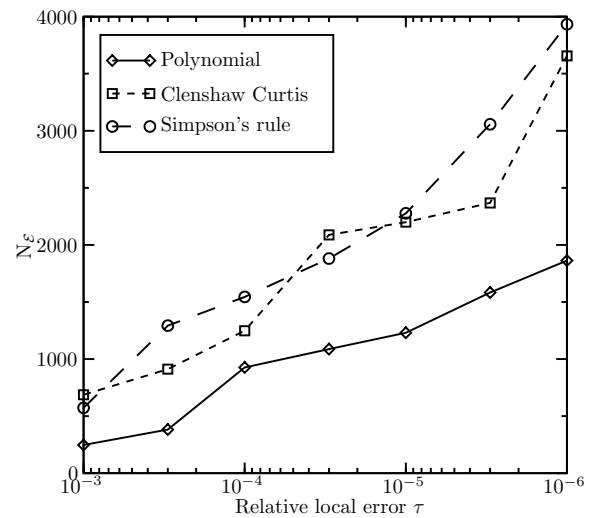

Figure 4: Number of energy grid points needed for a given relative error tolerance.

Figures 6 and 7 show the distribution of the grid points versus energy for the self-consistent calculation of the bandedge of a resonant tunneling diode under bias. The histograms give the number of grid points in energy intervals of 1 meV width. For both applied methods, the significant energies at resonant levels or the contact chemical potentials can be distinguished. At these energies, many more grid points are placed by the algorithms. The global adaptive procedure requires approximately half of the grid points to properly resolve a resonance as compared to the polynomial interpolation.

1.5 Conclusion

Local as well as a global adaptive integration strategies have been used in NEGF simulations. The Simpson rule does not suffice the demands of the diverse energy spectrum of a nano-electronic device. Although the polynomial interpolation and the Clenshaw-Curtis method combined with a local error criterion prove suitable for the numerical energy integration, a global adaptive approach is superior due to less grid points.
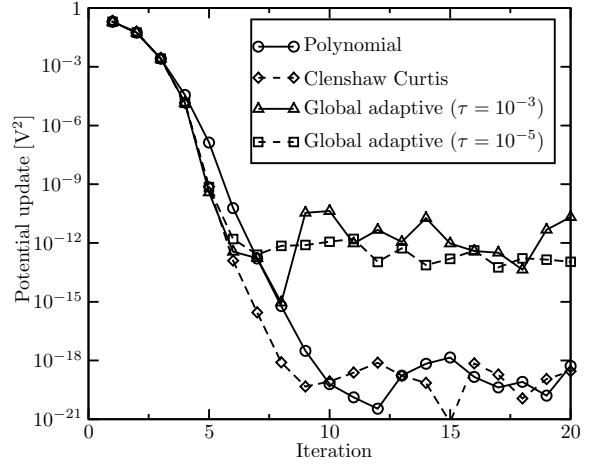

Figure 5: The evolution of the squared potential update norm is given for a self-consistent bandedge calculation of a double barrier structure.
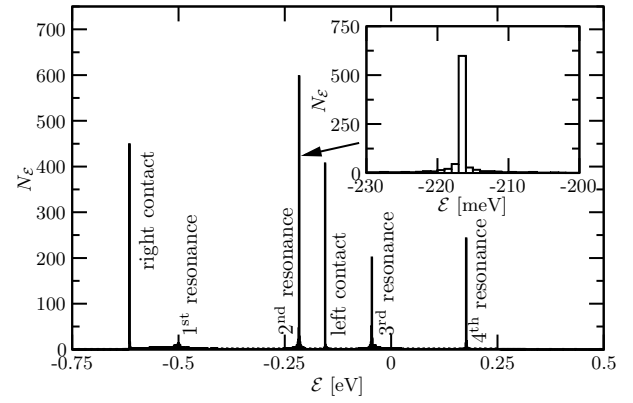

Figure 6: Histogram of energy grid points for the polynomial interpolation.
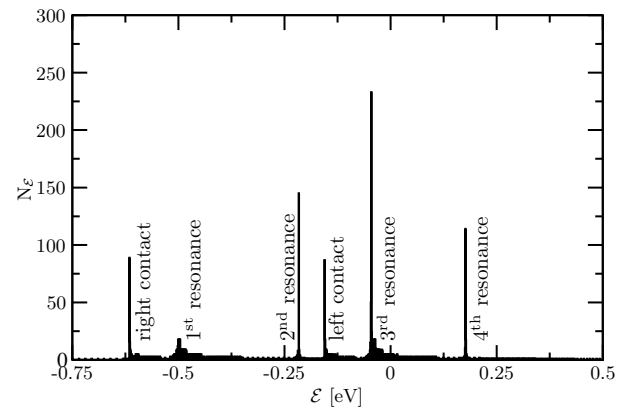

Figure 7: Histogram of energy grid points for the doubly global adaptive integration.

2 First-Principles Investigation on Oxide Trapping

We conduct a thorough investigation of the tunneling dynamics of oxide traps in a-SiO$_2$, in particular of the E$'_\delta$ center, the E$'_\gamma$ center, their hydrogenated counterparts, and the H atom. Based on these findings their behavior in the context of tunneling can be deduced. It is found that an E$'_\gamma$ center can exchange electrons with the Si bulk. The E$'_\delta$ center shows two distinct behaviors induced by a spread in its tunneling levels. The H atom is not affected by the presence of an interface, whereas a H bridge may occur in every charge state.

2.1 Introduction

Defect levels of traps have long been the subject of numerous investigations. Several experimental studies provide convincing indications of electron- or hole trapping in a-SiO$_2$ during bias temperature stress or exposure to irradiation [12]. A large number of defects have been studied in literature [12, 13], either identified by their electro-paramagnetic-resonance (EPR) signal or by their atomic-scale configurations. However, the fact that defects undergo structural relaxation accompanied by a shift in their energy levels has not received much attention to date, but has important implications on the tunneling kinetics in a-SiO$_2$. This issue can be tackled within the framework of first-principles simulations. Although a similar study for c-SiO$_2$ has been published [13], a detailed investigation addressing the variations of tunneling levels due to deviations of the local atomic arrangement in a-SiO$_2$ has not been performed up to date. The need to mimic the oxide with a-SiO$_2$ has been demonstrated by [14].

2.2 Method

For this purpose, simulations were done employing density functional theory (DFT) based on gradient corrections for the exchange-correlation functional and the projector augmented waves method (PAW) representing ion cores [15]. The plane wave cut-off energy was expanded up to $400\,\mathrm{eV}$. Structural optimization was controlled through a conjugate gradient algorithm which limits the force on each atom to be below $0.3\,\mathrm{eV/\mathring{A}}$. Charged supercells were calculated introducing a homogeneous compensating background charge to ensure neutrality. The use of a plane wave code implicitly involves periodic boundary conditions and implies interactions between periodically arranged defects. To minimize this effect, large supercells of a-SiO$_2$ ($\sim 11.79\,\mathring{A}$) comprising of 36 SiO$_2$ units (108 atoms) were chosen. The

Brillouin zone was sampled at the Γ-point. For production of a-SiO$_2$ an initial random configuration respecting exclusion radii between atoms was created applying an empirical potential method (BKS-potential). An equilibration step at $3000\,\mathrm{K}$ for $30\,\mathrm{ps}$ with a timestep of $1\,\mathrm{fs}$ was followed by a quench to $0\,\mathrm{K}$ for $30\,\mathrm{ps}$ with a timestep of $1\,\mathrm{fs}$. Subsequently, a structural optimization using DFT was carried out and the resulting silica network was classified by pair-correlation functions, angle distributions and the ring-distribution, which are in perfect agreement with published data [16]. Defect levels for tunneling mechanisms (also referred to as switching levels or transition states) are calculated as total energy differences of supercells in their initial and final charge state. The atomic structure is kept fixed according to the Frank-Condon principle.

$$\varepsilon^{+/0}[\mathrm{X}^+] = E_\mathrm{f}^0[\mathrm{X}^+] - E_\mathrm{f}^+[\mathrm{X}^+]$$
$$\varepsilon^{0/+}[\mathrm{X}^0] = E_\mathrm{f}^0[\mathrm{X}^0] - E_\mathrm{f}^+[\mathrm{X}^0]$$
$$\varepsilon^{-/0}[\mathrm{X}^-] = E_\mathrm{f}^-[\mathrm{X}^-] - E_\mathrm{f}^0[\mathrm{X}^-]$$
$$\varepsilon^{0/-}[\mathrm{X}^0] = E_\mathrm{f}^-[\mathrm{X}^0] - E_\mathrm{f}^0[\mathrm{X}^0] \tag{15}$$

The sign in E_f^q corresponds to the formation energy in the charge state q, where $[\mathrm{X}^q]$ denotes the equilibrium configuration.

After a tunneling process, a defect is not in its equilibrium configuration any more due to changes in the electrostatics. This gives rise to structural relaxation and different energetics including a new transition state. In consequence, different levels for tunneling into a defect and out of a defect arise, as depicted in Fig. 10. Concerning the bandgap alignment, we used the same procedure as proposed in [13], but due to the amorphous nature of a-SiO$_2$ we found a bandgap offset of approximately $2.6\,\mathrm{eV}$ consistent with valence band offsets extracted from [17].

2.3 Results

The O-vacancy can be envisioned by removing a bridging O atom from the silica network, where the adjacent Si atoms reestablish a common bond. In the neutral charge state, the bond length of about $2.3 - 3.0\,\mathring{A}$ comes close to the typical Si bond length in c-Si. For the positively charged variant, which is also referred to as E$'_\delta$ centers, both Si atoms repel each other without breaking their bond. Therefore this variant exhibits a longer bond length which expands to values varying from $2.7\,\mathring{A}$ to $3.2\,\mathring{A}$. Irrespective of their spread associated with variations in their configurations, these structural parameters are in good agreement with their crystalline counterparts published in [13, 14]. The bistable partner of the E$'_\delta$, the E$'_\gamma$ center (see Fig. 8), is confirmed by a long range of theoretical studies [13, 14, 17]. One part of this configuration simply persists as a normal Si dangling bond carrying only one electron. The other part

undergoes puckering, that is, the Si atom binds to a back O to stabilize this configuration. In the neutral charge state, the additional electron occupies the Si dangling bond. In the hydrogenated variant of the O-vacancy (also termed H bridge), one H atom is located directly in-between both Si atoms, moving them apart. As illustrated in Fig. 10, this defect shows a strong asymmetry ($d_{1,\text{Si}-\text{Si}} = 1.5\,\text{Å}$ and $d_{2,\text{Si}-\text{Si}} = 12.2 - 2.8\,\text{Å}$) with the H being attached closer to one of the Si atoms in the neutral case. For the positively charged Si atom, this asymmetry ($d_{\text{Si}-\text{H}} = 1.6 - 1.8\,\text{Å}$) nearly vanishes, which is in agreement with investigations in c-SiO$_2$ [13].

Interstitial atomic H was introduced by [13] to be a candidate for trapping in a-SiO$_2$. In the absence of any other type of defect, its bonding configurations are well established and can be explained in terms of electronegativity. Due to the polarity of the Si-O bond, more charge is distributed near the O atom. The proton is therefore attracted to the nearest network O atom with a bond length of $1.00\,\text{Å}$, where the negatively charged H is attached to a Si network atom with a bond length of $1.48 - 1.52\,\text{Å}$. Both values are in good agreement with values obtained in c-SiO$_2$ [13,18]. In the neutral charge state, the H atom prefers a position in the middle of a void.

Now we address the defect levels for tunneling in and tunneling out of the defects, which are summarized in Tab. 1 and illustrated in Fig. 10. For the positively charged E$'_\delta$ center, the energy level for neutralization is located $1.7 - 3.6\,\text{eV}$ above the Si oxide valence band maximum. When an electron is allowed to tunnel into the trap, its energy level shifts down to the valence band of a-SiO$_2$. As a consequence, the electron does not find a hole in the Si to escape the defect. Therefore, this trap remains neutral and is annealed permanently. This result is similar to the trap level behavior observed in c-SiO$_2$ [13]. In some cases, the trap level for tunneling in of an electron is found in the silicon bandgap. In consequence, this defect can also act as a fixed positive charge, if a hole is provided in the a-SiO$_2$ valence band.

As opposed to the E$'_\delta$ center, the E$'_\gamma$ center has a trap level $+/0$ slightly above the Si conduction band minimum. The trap level after relaxation allows hole capture from the Si valence band. As far as the energetics of hydrogenated E$'_\delta$ are concerned, the trap levels are located near the Si band edges so that this defect strongly interacts with the Si. Except for the case of charging the defect negatively, the H atom exhibits trap levels far from the Si band edges so that an exchange of electrons with the interface will not occur in a relevant extent.

2.4 Conclusion

We presented defect levels for tunneling in and tunneling out of a trap for several well-known defects and mimicked the oxide by a-SiO$_2$. The E$'_\delta$ center turned out to show a different behavior depending on the exact position of the transition energy $\varepsilon_{0/+}$. The E$'_\gamma$ center as well as the H bridge is found to easily exchange electrons with silicon. In contrast, the H atom is revealed not to interact with the interface.

	E$'_\delta$ eV	E$'_\gamma$ eV	E$'_\delta$H eV	H eV
$+/0$	$1.7 - 3.6$	$3.9 - 4.1$	$4.4 - 4.9$	$4.9 - 5.1$
$0/+$	$0.3 - 0.5$	$2.1 - 2.4$	$1.7 - 2.3$	$0.3 - 0.9$
$0/-$			$4.1 - 4.3$	$3.2 - 4.4$
$-/0$			$1.2 - 2.6$	$0.9 - 1.2$

Table 1: Energy trap levels ($+/0$, $0/+$, etc.) relative to the theoretical oxide valence band maximum of a-SiO$_2$. The first sign denotes the equilibrium configuration of the defect in the corresponding charge state. The second sign gives the charge state of the defect for a given configuration.
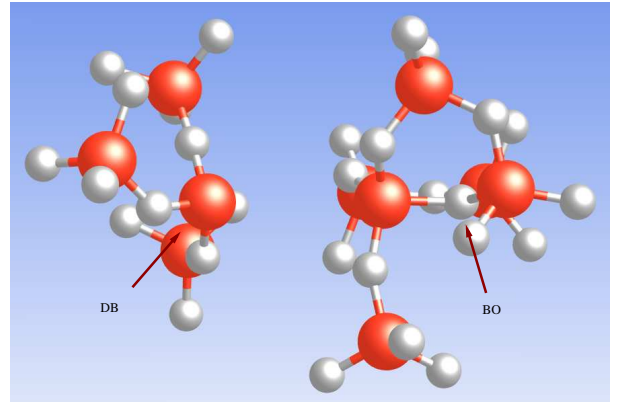

Figure 8: Atomic structure of an E$'_\gamma$ center. DB: dangling bond, BO: back O, red spheres: Si atoms, white spheres: O atoms
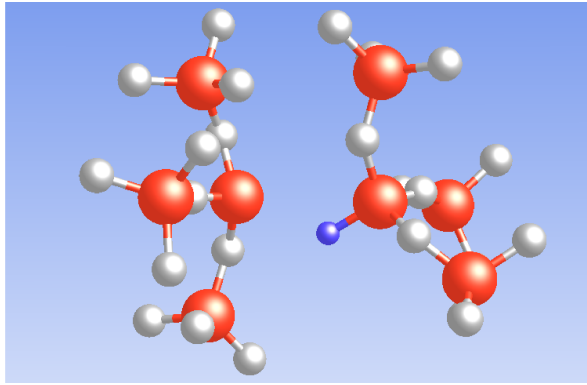
Figure 9: Atomic structure of a neutral hydrogenated E'_δ center. This representation clearly demonstrates the asymmetry of the H bridge.
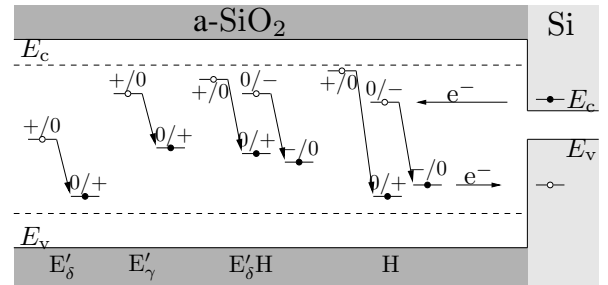

Figure 10: Band scheme including the calculated trap levels. The filled circles denote occupied defect levels, whereas open circles are unoccupied defect levels.

3 Three-Dimensional On-Chip Inductance and Resistance Extraction

An efficient three-dimensional finite element frequency domain method to extract the inductance and resistance of small structures like on-chip interconnects or on-chip inductors is presented. Skin effect and proximity effect are taken into account. The parameters are obtained from the field energy calculated from the magnetic field distribution in the simulation domain. The small dimensions of the domain of interest provide the opportunity of using the optimized model of dominant magnetic field even at very high operating frequencies. Vector and scalar shape functions are used for finite element equation system assembling. Series of simulations for an instancing on-chip inductor at frequencies between 1 MHz and 100 GHz are performed to extract the parameters and to visualize the field distributions in the simulation area.

3.1 Introduction

High frequencies in an integrated circuit (IC) affect both, the resistance and the inductance of the on-chip interconnects. These often as parasitics treated parameters cause longer signal rise, fall, and delay times and limit the maximum allowed frequency of modern ICs. However, as the operating frequencies increase, small inductors of high speed circuits can be also actively used. They can be even constructed on the chip. Thus the inductance of an on-chip interconnect line can be a disadvantage or very useful depending on the application. Of course the collateral resistance must also be considered. In each case it is necessary to investigate the structure of interest to obtain its inductance and resistance in order to estimate the impact on the entire electric circuit [19]. In the case of applications in radio frequency (RF) ICs such as voltage controlled oscillators or low noise amplifiers the inductance and the resistance of the on-chip inductors must be extensively investigated for the RF circuit design, performance optimization, and inductor quality factor. The frequency dependent inductance and resistance of wide on-chip interconnects must be captured to obtain the impact on power supply stability and signal delay.

Currently there are two major techniques for modeling of on-chip inductors — analytical compact modeling and numerical field calculation based modeling. In the case of a spiral inductor, where the models can be restricted to specific geometry classes, closed-form analytical models are very well suited for fast designs typical for the very early stage of the developing process [20, 21]. However, analytical modeling of arbitrary shaped three-dimensional structures is very complicated, if possible at all. Thus, analytical parameter extraction methods have only limited applicability. For final analysis prior to fabrication and for irregular inductor geometries numerical simulation methods normally based on solving the Maxwell equations provide the most accurate characterization. Moreover, the investigated interconnect structure can often be embedded in a small simulation region for which the optimized model of the dominant magnetic field (DMF) can be used even at very high frequencies. The distributed vector and scalar fields must be extracted in structures which may consist of different inhomogeneous complex shaped three-dimensional regions, like splittings, widenings, and vertical connections. As a consequence, the vector and scalar finite element method (FEM) [22] in the frequency domain on unstructured tetrahedral grid [23,24] needs to be addressed.

In this work an optimized model for inductance and resistance analysis of an on-chip inductor at different frequencies is proposed. The model describes the proximity effect and the skin effect typically arising at higher frequencies as well. The three-dimensional FEM simulation software **SAP** (**S**mart **A**nalysis **P**rograms) [25] is extended to implement the developed model. Simulation results demonstrate the physical plausibility of the applied model and numerical methods, as well as the necessity of three-dimensional simulations.

3.2 The Theoretical Background

If the characteristic lengths of the analyzed structures are much smaller than the considered wave lengths and if for conducting areas $\epsilon/\gamma \ll \tau$, then the displacement current can be neglected [26]. τ is the characteristic period of the time change rate, ϵ is the characteristic permittivity, and γ is the characteristic conductivity. In this case the Maxwell equations can be considerably simplified and for time independent μ the model DMF is achieved:

$$\vec{\nabla} \times \vec{E} = -\mu \partial_t \vec{H} \qquad (16)$$

$$\vec{\nabla} \cdot (\mu \vec{H}) = 0 \qquad (17)$$

$$\rho(\vec{\nabla} \times \vec{H}) = \vec{E}, \qquad (18)$$

where $\vec{E}$ is the electric field intensity, $\vec{H}$ is the magnetic field intensity, μ is the material's permeability, and ρ is the material's electric resistivity. The derivative with respect to time is shortly notated as ∂_t instead of $\partial/\partial t$.

3.2.1 The Problem Description

Applying the rotor operator to (18) and substituting the right hand side of (18) by (16) yields the following

second-order differential equation

$$\vec{\nabla}\times(\rho\vec{\nabla}\times\vec{H}) + \mu\partial_t\vec{H} = 0. \qquad (19)$$

For the stationary case ($\partial_t\vec{H} = 0$) the finite element analysis of (19) leads to an overdetermined linear equation system [27]. The matrix of this system is positive semi-definite and its zero eigenvalues correspond to the number of tree edges in the graph spanned by the finite element mesh edges [28].

Since $\vec{\nabla}\times\vec{\nabla}\varphi = 0$, (19) may be written as

$$\vec{\nabla}\times[\rho\vec{\nabla}\times(\vec{H}+\vec{\nabla}\varphi)] + \mu\partial_t\vec{H} = 0, \qquad (20)$$

which leads with the aid of the substitution $\vec{H}_1 = \vec{H} + \vec{\nabla}\varphi$ to

$$\vec{\nabla}\times(\rho\vec{\nabla}\times\vec{H}_1) + \mu\partial_t(\vec{H}_1 - \vec{\nabla}\varphi) = 0, \qquad (21)$$

where φ is an arbitrary scalar field and $\vec{\nabla}\varphi$ must exist. Equation (21) contains two unknown functions — the vector field $\vec{H}_1$ and the scalar field φ. Thus an additional independent criterion is needed. For numerically stable and unique calculation of $\vec{H}$ it is natural to impose the divergence condition (17), which has not been used till now

$$\vec{\nabla}\cdot[\mu(\vec{H}_1 - \vec{\nabla}\varphi)] = 0. \qquad (22)$$

Thus the unknown fields $\vec{H}_1$ and φ are the solution of the partial differential equation system consisting of (21) and (22), which is a boundary value problem numerically calculated by FEM in a simulation domain $\mathcal{V}$ enclosing the investigated structures. The fields are approximated by the following expressions:

$$\vec{H}_1(\vec{r},t) = \sum_{j=1}^{n} c_j(t)\,\vec{N}_j(\vec{r}) + \vec{H}_K(\vec{r},t) \qquad (23)$$

$$\varphi(\vec{r},t) = \sum_{j=n+1}^{m} c_j(t)\,\lambda_j(\vec{r}) + \varphi_K(\vec{r},t) \qquad (24)$$

Due to the FEM domain discretization the region of interest $\mathcal{V}$ and its surface $\partial\mathcal{V}$ are subdivided into smaller mesh elements — tetrahedrons and triangles consisting of edge connected nodes. The boundary of the simulation area $\partial\mathcal{V}$ is divided into Dirichlet boundary $\partial\mathcal{V}_{D1}$ and Neumann boundary $\partial\mathcal{V}_{N1}$ for $\vec{H}_1$ and into Dirichlet boundary $\partial\mathcal{V}_{D2}$ and Neumann boundary $\partial\mathcal{V}_{N2}$ for φ, respectively ($\partial\mathcal{V}=\partial\mathcal{V}_{D1}+\partial\mathcal{V}_{N1}$ and $\partial\mathcal{V}=\partial\mathcal{V}_{D2}+\partial\mathcal{V}_{N2}$). The edges and the nodes in the simulation area are labeled with a set of integers. The non-Dirichlet edges are indexed from 1 to n and the non-Dirichlet nodes are indexed from $n+1$ to m ($m>n$). The non-Dirichlet edges are the edges which do not belong to $\partial\mathcal{V}_{D1}$, and the non-Dirichlet nodes are the nodes which do not belong to $\partial\mathcal{V}_{D2}$, respectively. The known function $\vec{H}_K$ satisfies the Dirichlet boundary conditions on $\partial\mathcal{V}_{D1}$ and does not have a tangential component on the non-Dirichlet edges of the simulation area. Analogously the known function

φ_K satisfies the Dirichlet boundary condition on $\partial\mathcal{V}_{D2}$ and vanishes in the remaining simulation area. $\vec{N}_j^{(e)}$ is the Whitney 1-Form vector basis function [29] in a single mesh element and is associated with the j-th edge. For example, in a tetrahedral element, if the j-th edge belongs to the nodes p and q, the corresponding element edge function is:

$$\vec{N}_j^{(e)} = \vec{N}_{pq}^{(e)} = l_{pq}(\lambda_p^{(e)}\vec{\nabla}\lambda_q^{(e)} - \lambda_q^{(e)}\vec{\nabla}\lambda_p^{(e)}),$$

where l_{pq} is the length of the corresponding edge. The Lagrange interpolation polynomial $\lambda_j^{(e)}$ at vertex j in the tetrahedral mesh element is given by:

$$\lambda_j^{(e)} = \frac{1}{4} - (\vec{r}-\vec{r}_B)\cdot\frac{F_j\vec{n}_j}{3V},$$

where $\vec{r}_B=(\sum_{i=1}^{4}\vec{r}_i)/4$ is the barycenter of the tetrahedral mesh element, F_j is the area of the face opposite to the j-th node, $\vec{n}_j$ is normal to this face and points outward, and V is the volume of the tetrahedron. The global edge function $\vec{N}_j$ consists of all element edge functions belonging to the global edge j and does not have a tangential component on the remaining edges. The global node function λ_j consists of all element node functions belonging to the global node j and vanishes in the remaining nodes. Consequently the calculated coefficients c_j represent the tangential component of the vector field $\vec{H}_1$ on the j-th edge or the value of the scalar field φ on the j-th node. The global node or edge functions are "active" only in the neighbor elements of the corresponding node or edge, respectively. Applying the Galerkin method to (21) and (22)

$$\int_{\mathcal{V}}[\vec{\nabla}\times(\rho\vec{\nabla}\times\vec{H}_1) + \mu\partial_t(\vec{H}_1-\vec{\nabla}\varphi)]\cdot\vec{N}_i\,dV = 0,$$
$$i = 1\ldots n \qquad (25)$$

$$\int_{\mathcal{V}}\vec{\nabla}\cdot[\mu(\vec{H}_1-\vec{\nabla}\varphi)]\lambda_i\,dV = 0,\ i = n+1\ldots m \quad (26)$$

and using the first vector Green's theorem and the first scalar Green's theorem the following system is obtained:

$$\int_{\mathcal{V}}(\vec{\nabla}\times\vec{N}_i)(\rho\vec{\nabla}\times\vec{H}_1)\,dV + \int_{\mathcal{V}}\mu\partial_t\vec{H}_1\cdot\vec{N}_i\,dV -$$
$$-\int_{\partial\mathcal{V}}\vec{n}\cdot[\vec{N}_i\times(\rho\vec{\nabla}\times\vec{H}_1)]\,dA - \int_{\mathcal{V}}\mu\partial_t\vec{\nabla}\varphi\cdot\vec{N}_i\,dV = 0,$$
$$i = 1\ldots n \qquad (27)$$

$$\int_{\mathcal{V}}\vec{\nabla}\lambda_i\cdot(\mu\vec{H}_1)\,dV - \int_{\mathcal{V}}\vec{\nabla}\lambda_i\cdot(\mu\vec{\nabla}\varphi)\,dV -$$
$$-\int_{\partial\mathcal{V}}\lambda_i\vec{n}\cdot[\mu(\vec{H}_1-\vec{\nabla}\varphi)]\,dA = 0,\ i = n+1\ldots m$$
$$(28)$$

The correct gradient-field-free magnetic field intensity $\vec{H}$ and the current density distribution $\vec{J}$ are calculated from:

$$\vec{J} = \vec{\nabla}\times\vec{H} = \vec{\nabla}\times\vec{H}_1,\ \vec{H} = \vec{H}_1 - \vec{\nabla}\varphi \qquad (29)$$

3.2.2 Boundary Conditions

The boundary term in (27) can be expressed as follows:

$$\int_{\partial\mathcal{V}} \vec{n}\cdot[\vec{N}_i\times(\rho\vec{\nabla}\times\vec{H}_1)]\, dA = \int_{\partial\mathcal{V}} \vec{n}\cdot(\vec{N}_i\times\vec{E})\, dA =$$

$$\int_{\partial\mathcal{V}} (\vec{n}\times\vec{N}_i)\cdot\vec{E}\, dA = \int_{\partial\mathcal{V}} (\vec{E}\times\vec{n})\cdot\vec{N}_i\, dA$$

(30)

with $\vec{E}=\rho\vec{J}=\rho\vec{\nabla}\times\vec{H}_1$. $\vec{N}_i$ has a tangential component only on the i-th edge. For $i=1\dots n$ the i-th edge is a non-Dirichlet edge. Thus $\vec{N}_i$ has no tangential component on the edges of the Dirichlet boundary $\partial\mathcal{V}_{D1}$ and must be perpendicular to $\partial\mathcal{V}_{D1}$ (or parallel to $\vec{n}$). As clearly shown by the third member of (30) the boundary integral in (27) has a contribution only for the Neumann boundary $\partial\mathcal{V}_{N1}$: $\int_{\partial\mathcal{V}_{N1}}(\vec{E}\times\vec{n})\cdot\vec{N}_i\, dA$. Only the supply parts of the wire, which are used to force the electric current, lie directly on the boundary faces of the simulation area. The remaining parts of the wire are surrounded by dielectric material. In this work the dielectric environment enclosing the wire is assumed to be sufficiently thick so that $\vec{E}$ could be neglected on the dielectric outer bounds of the simulation domain. On the other hand the electric current density is forced in a direction perpendicular to the conductor boundary faces. Thus $\vec{E}$ will be also perpendicular to these faces and the boundary term in (27) will be zero.

The supplied total current in the inductor wire is considered by the Dirichlet condition for $\vec{H}_1$:

$$I = \oint_{\partial\mathcal{A}} \vec{H}_1\cdot d\vec{r} = \oint_{\partial\mathcal{A}} \vec{H}\cdot d\vec{r} \simeq H_t \sum_{i=0}^{p} l_i, \qquad (31)$$

where $\partial\mathcal{A}$ is an arbitrary closed loop around the conducting wire, p is the number of Dirichlet edges, which build this loop, and l_i is the length of the i-th loop edge. If the Dirichlet edges are labeled with integers from $m+1$ to k ($k>m$), the Dirichlet boundary function $\vec{H}_K$ from (23) is expressed as $\vec{H}_K=\sum_{j=m+1}^{k} c_j\vec{N}_j$ with $c_j=H_t$ for $j=m+1\dots k$ and $p=k-m$.
Analogously the boundary term in (28) can be expressed as

$$\int_{\partial\mathcal{V}} \lambda_i\vec{n}\cdot[\mu(\vec{H}_1-\vec{\nabla}\varphi)]\, dA = \int_{\partial\mathcal{V}} \lambda_i\vec{n}\cdot\vec{B}\, dA = \int_{\partial\mathcal{V}_{N2}} \lambda_i\vec{n}\cdot\vec{B}\, dA$$

with $\vec{B}=\mu\vec{H}=\mu(\vec{H}_1-\vec{\nabla}\varphi)$. $\vec{B}$ is assumed to be perpendicular to $\vec{n}$ on the Neumann boundary $\partial\mathcal{V}_{N2}$. Thus the boundary integral in (28) vanishes.

Dirichlet boundary conditions must be given also for φ. The Dirichlet nodes are labeled with numbers from $k+1$ to l ($l>k$). The function φ_K from (24) can analogously be defined as $\varphi_K=\sum_{j=k+1}^{l} c_j\lambda_j$. For the calculations

it is sufficient that one node of the simulation domain is forced as Dirichlet node. Thus, in this specific case $\varphi_K = c_{k+1}\lambda_{k+1}$. As far as only $\vec{\nabla}\varphi$ is required and not φ, the coefficient c_{k+1} of this Dirichlet node can be set to an arbitrary value.

3.2.3 Assembling the Matrix and the Right Hand Side Vector

Under consideration of (23), (24) and the above specified boundary conditions the weighted equation system (27) and (28) corresponds in the frequency domain with

$$\begin{bmatrix} [A] & [B] \\ [B]^T & [C] \end{bmatrix} \{c\} = \{b\},$$

$$A_{ij} = \int_{\mathcal{V}}(\vec{\nabla}\times\vec{N}_i)\cdot(\rho\vec{\nabla}\times\vec{N}_j)\, dV + \jmath\omega\int_{\mathcal{V}}\vec{N}_i\cdot(\mu\vec{N}_j)\, dV,$$
$$i=1\dots n,\ j=1\dots n$$

$$B_{ij} = -\jmath\omega\int_{\mathcal{V}}\vec{N}_i\cdot(\mu\vec{\nabla}\lambda_j)\, dV,\ i=1\dots n,\ j=n+1\dots m$$

$$C_{ij} = \jmath\omega\int_{\mathcal{V}}\vec{\nabla}\lambda_i\cdot(\mu\vec{\nabla}\lambda_j)\, dV,\ i=n+1\dots m,\ j=n+1\dots m$$

$$b_i = \jmath\omega\sum_{j=k+1}^{l} c_j\int_{\mathcal{V}}\vec{N}_i\cdot(\mu\vec{\nabla}\lambda_j)\, dV - \jmath\omega\sum_{j=m+1}^{k} c_j\int_{\mathcal{V}}\vec{N}_i\cdot(\mu\vec{N}_j)\, dV -$$
$$-\sum_{j=m+1}^{k} c_j\int_{\mathcal{V}}(\vec{\nabla}\times\vec{N}_i)\cdot(\rho\vec{\nabla}\times\vec{N}_j)\, dV,\ i=1\dots n$$

$$b_i = \jmath\omega\sum_{j=m+1}^{k} c_j\int_{\mathcal{V}}\vec{\nabla}\lambda_i\cdot(\mu\vec{N}_j)\, dV - \jmath\omega\sum_{j=k+1}^{l} c_j\int_{\mathcal{V}}\vec{\nabla}\lambda_i\cdot(\mu\vec{\nabla}\lambda_j)\, dV,$$
$$i = n+1\dots m,$$

where the time convention $e^{\jmath\omega t}$ is used and suppressed. To obtain a symmetric equation system (28) has been multiplied by $-\jmath\omega$, where ω is the angular frequency.

3.2.4 Inductance and Resistance Extraction

The inductance and the resistance are calculated by the magnetic energy and by the electric power, respectively.

$$L = \frac{1}{I^2}\int_{\mathcal{V}} \mu H^2\, dV,\ R = \frac{1}{I^2}\int_{\mathcal{V}}(\vec{\nabla}\times\vec{H}_1)\cdot(\rho\vec{\nabla}\times\vec{H}_1)\, dV.$$

(32)

I is the total current in the inductor, which defines the Dirichlet boundary (31) for $\vec{H}_1$ and $\vec{H}$ is given by (29).

3.3   Domain Discretization

The domain discretization is the first step of the finite element analysis. For this purpose the entire domain is divided into smaller sub-domains, called elements. For three-dimensional problems the volume elements can be rectangular bricks or tetrahedrons, for instance. Then the boundary surface is broken up into rectangles or triangles, respectively. The rectangular elements are perfectly suited for discretizing rectangular domains with an uniform density. However, the physical models can not always be limited to specific regular geometries. The mesh density must be high enough to achieve sufficiently accurate solutions. Unfortunately a high number of mesh elements leads to many unknowns, causing high memory demands and long simulation times. For this purpose it is required to keep the element size as large as possible for a desired accuracy. It is desirable to use a finer mesh (or smaller elements) only in the regions where high field dynamic is anticipated. In this work the regions are described by tetrahedral or triangular elements, depending on the discretized object — volume or boundary surface. In addition, not only the density but also the quality of the tetrahedral elements affects directly the FEM accuracy and efficiency [30].

The example inductor geometry presented in the next section is described by the constructive solid geometry (CSG) format. It is discretized with the three-dimensional tetrahedron mesh generation software Netgen [31]. The CSG format is very convenient for the description of small or medium size structures like the spiral inductor presented in the example section. The geometry is defined by Eulerian operations (union, intersection, and complement) from primitives. The primitives are generic volume elements like cubes, cylinders, spheres, or even half-spaces defined by an arbitrary point in the boundary plane and an outward normal vector. If CSG input is used, Netgen starts with the computation of the corner points. Then the edges are defined and meshed into segments. Next, the faces are generated by an advancing front algorithm [32]. After optimization of the surface mesh the volume inside is filled with tetrahedrons by a fast Delaunay algorithm [33]. Finally the volume mesh is optimized.

3.4   Examples and Results

As example a typical on-chip spiral inductor structure as discussed in [34] is investigated. The simulation domain consists of a transparent insulating rectangular brick over an opaque substrate brick as shown in Fig. 11. The aluminum inductor is placed in the insulating environment about $5\,\mu$m above the substrate area. The substrate is assumed to have a constant resistiv-

Table 2: Calculated inductance and resistance.

$\mathbf{f}\,[GHz]$	$\mathbf{L}\,[nH]$ without substrate	$\mathbf{L}\,[nH]$ with substrate	$\mathbf{R}\,[\Omega]$
0.001	2.6887	2.6881	3.127
0.01	2.6887	2.6877	3.127
0.1	2.688	2.688	3.132
1	2.6516	2.6514	3.463
10	2.5501	2.5493	5.396
100	2.5458	2.5457	13.156

ity of $10\,\Omega\,cm$. The cross-section of the conductor is $20\,\mu$m $\times\,1.2\,\mu$m. The horizontal distance between the winding wires is $10\,\mu$m. The outer dimensions of the inductor are $300\,\mu$m $\times\,300\,\mu$m. The inductor is completely surrounded by the dielectric environment, except of the two small delimiting faces which lie directly in the boundary planes of the simulation domain. The conductor area, the dielectric, and the substrate area close to the conductor are discretized much finer then the remaining simulation domain. This is shown in Fig. 12 where a part of the dielectric environment is removed to visualize in detail the generated mesh inside the simulation domain. The variation of the fields in the finer discretized areas is expected to be much higher than in the coarser discretized domain. This special discretization reduces the number of generated nodes and edges, and the number of the linear equations respectively, even for big simulation environments which have to be used to satisfy the assumption of the zero Neumann boundary condition. Of coarse such a discretization is only possible, if an unstructured mesh is used.

The current density distribution depends heavily on the operating frequency in the analyzed frequency domain. It is unknown and arises from the simulation. At the beginning of the simulation only the total current in the inductor is known. As mentioned above it is set by the Dirichlet boundary condition for $\vec{H}_1$ which is given by the tangential component of the magnetic field intensity H_t on the element edges, surrounding one of the conductor faces lying on the outer bound of the simulation domain.

The resistance and inductance values of the structure of interest are calculated numerically at different frequencies with and without the substrate influence. The corresponding results are presented in Table 2. While the inductance decreases slowly with increasing operating frequency, the resistance rises dramatically, which matches well the observed current density distribution and the skin effect, respectively. A surface view of the

current density distribution in the conductor is shown in Fig. 13 and Fig. 14 for 100 MHz and 10 GHz, respectively. At 100 MHz the skin depth is about 6 μm and nearly the whole conductor cross-section is filled up by the current. At 10 GHz the skin depth is about 0.6 μm and the current is concentrated at the vertical side walls of the conductor. Fig. 15 shows the spatial current density distribution at 1 GHz as directed cones placed in the discretization nodes inside of the conductor area. The cone's size and darkness are proportional to the field strength. Fig. 16 depicts the corresponding spatial distribution of the magnetic field inside the dielectric environment around the inductor. Fig. 17 and Fig. 18 show the very different current density distribution close to the small via for 100 MHz and 10 GHz, respectively. Fig. 19 and Fig. 20 show the current density distribution and the corresponding magnetic field intensity in the substrate at 1 GHz. The underlying substrate does not influence the inductance and the resistance of the inductor, because of the relative high substrate resistivity. The values of the current density distribution in the substrate (Fig. 19) differ vastly to those in the inductor (Fig. 13 and Fig. 14). As shown in Table 2 the calculated inductance taking into account the influence of the substrate does not differ from the inductance without substrate influence.

As the Q-factor of an inductor is inversely proportional to its resistance, making the inductor wire thicker might decrease the resistance and increase the Q-factor. However, as the examples show this is not the case for high frequencies at which the skin effect is noticeable. In these cases the current flows only in the area very close to the vertical surface and a wider transversal conductor cross section would not change the situation. For the visualization VTK [35] is used.

3.5   Conclusion

In this work a new method for inductance and resistance parameter extraction is proposed. It is assumed that the operating frequencies and the characteristic lengths of the investigated structures are suitable to use the DMF model. This is typically the case for on-chip inductors and interconnect loops. The rotor operator which appears in the second order partial differential equation for the magnetic field $\vec{H}$ allows arbitrary unknown gradient fields $\vec{\nabla}\varphi$. The obtained partial differential equation system is solved numerically by FEM. Instead of solving the more complicated wave equation, considering DMF, it is sufficient to solve the diffusion equation to investigate the parameters and field distributions of interest, which is of course the more efficient way.
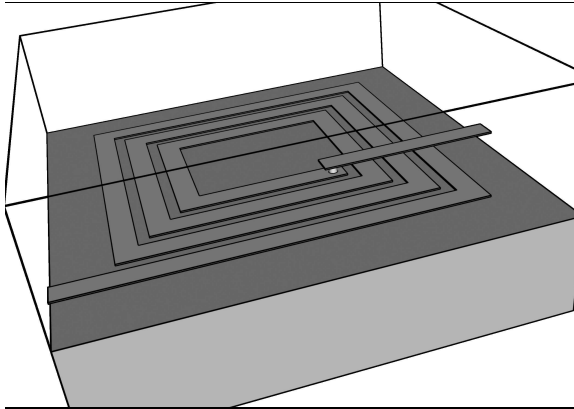
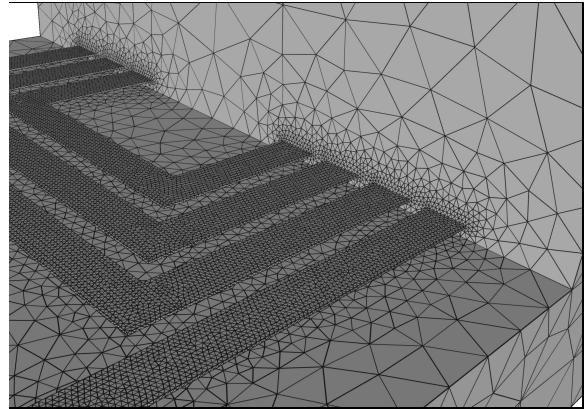**Figure 11:** The simulation domain.
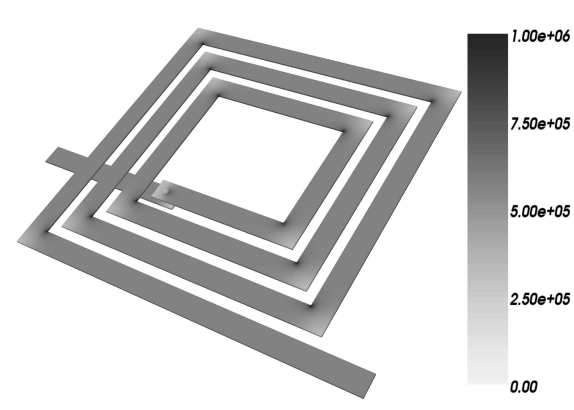
Figure 12: The generated grid.

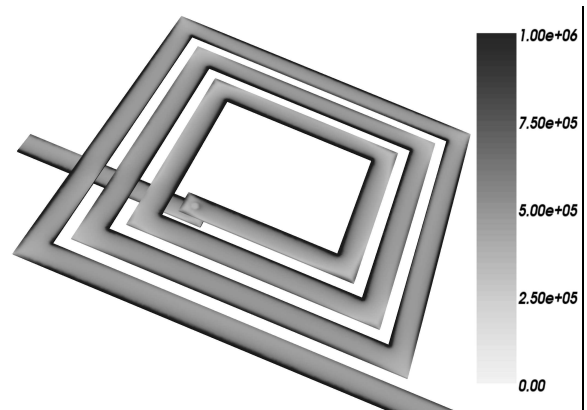**Figure 13:** Surface view of the current density $[A/m^2]$ distribution at 100 MHz.

Figure 14: Surface view of the current density $[A/m^2]$ distribution at 10 GHz.
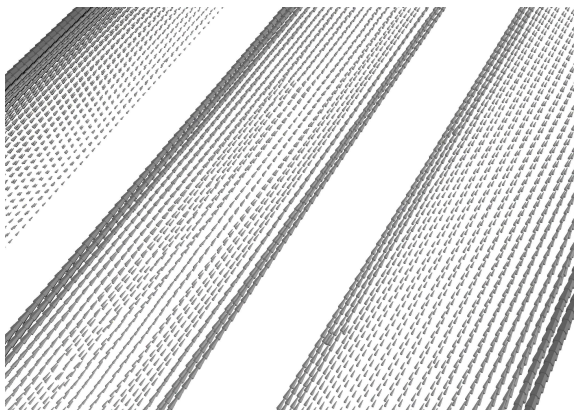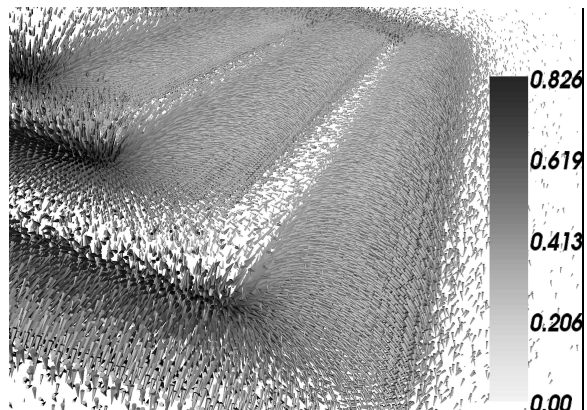
Figure 15: Current density distribution at 1 GHz.

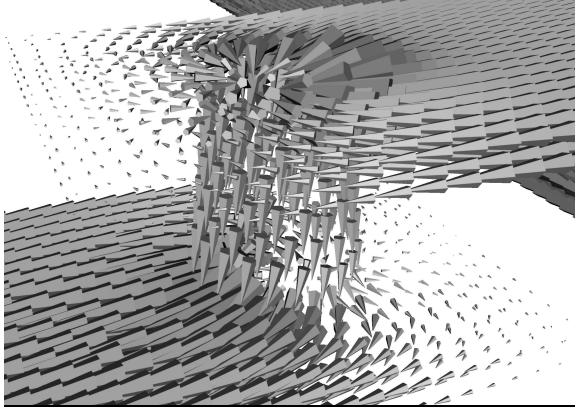**Figure 16:** Magnetic field intensity $[A/m]$ at 1 GHz.

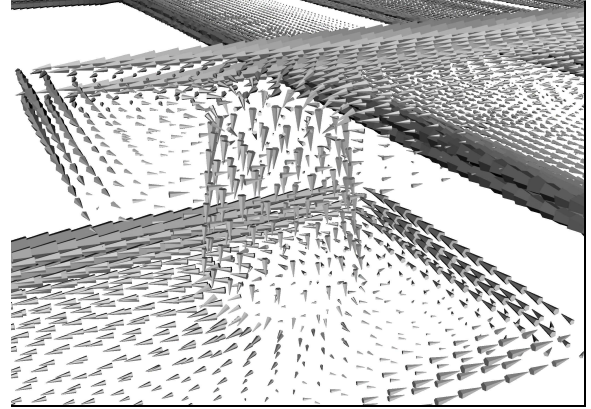**Figure 17:** Current density distribution in the via at 100 MHz.

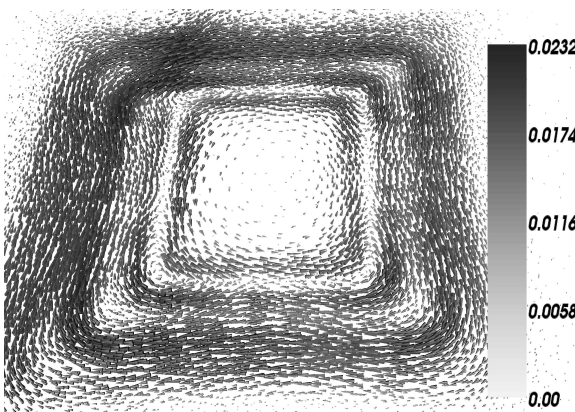**Figure 18:** Current density distribution in the via at 10 GHz.

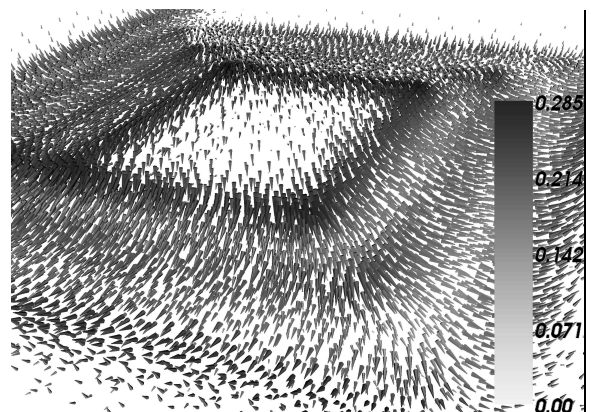**Figure 19:** Current density $[A/m^2]$ distribution in the substrate at 1 GHz.

Figure 20: Magnetic field intensity $[A/m]$ in the substrate at 1 GHz.

4 A Multi-Mode Mesh Generation Approach for Scientific Computing

With the steady evolution of software tools and programming languages, several programming paradigms have become available and new methodologies related to library centric application design have been developed. In addition to this paradigm shift, the increase of computing power requires an adaptation of well known techniques for parallel computing. Various problems of different disciplines appear, for which no efficient and definite solutions exist. This calls for new paradigms to be established and the development of new ways which fully exhaust the combination of sophisticated software and powerful hardware.

The area of spatial discretization, unstructured mesh generation, is one of these areas, which still exhibits various issues from fields such as modular software design, high performance, and most importantly robustness.

In this work problems for mesh generation are identified and solutions explored, which anticipate the given challenges. This approach is combined in a software library based on a generic scientific simulation environment and compared to existing tools.

4.1 Introduction and Motivation

The field of scientific computing as a whole, and Technology Computer Aided Design (TCAD) in particular, relies on either domain-specific closed applications, or components tied very closely to specific representations which were often built for very fixed purposes and then wrapped into a large and monolithic toolkit. To this date, data structures and algorithms are implemented in a heavily application specific way, making their reuse practically impossible.

One of the most fundamental issues in TCAD, spatial discretization, requires certain and diverse constraints during mesh generation. Many software tools exist which solve very narrow sets of problems in special and well defined settings, but unfortunately most of the solutions to these problems are implemented repeatedly. With the steady evolution of software engineering techniques and methodologies, new paradigms have emerged which offer great opportunities related to orthogonal module design and high performance approaches. Using a generic programming approach instead of object orientation, data structures and algorithms can be implemented orthogonally, where some languages, e.g. C++, offer additional high performance possibilities as well.

In addition to these basic software engineering issues, the calculating power of modern computer systems, which has been doubling every 18 months in the past, and the emerging CPU parallelism, require different programming paradigms and application design methodologies. These changes have to be taken into account as well as the upcoming shift to fully three-dimensional simulations [36]. This of course also implies, that three-dimensional mesh generation is getting more and more important in the next decade, especially in the field of TCAD where up to now, no comprehensive mesh generation application has emerged. Mesh generation is a quite difficult task to be implemented on computers, because a rigoros algorithmic formulation does not exist. The emerging paradigm of computational topology will support mesh generation in the coming years [37, 38]. Different fields of TCAD application impose a variety of different constraints and requirements on mesh generation, e.g. topography simulation requires a good approximation of surface elements, while ion implantation simulation requires high density near the surface, according to the gradient of the ion distribution. Diffusion simulations add a need for a fine mesh at interfaces in addition to a high mesh density near the surface. The complex field of device modeling even requires a completely different type of mesh, necessitating a remeshing step for the whole input structure. In summary, it can be observed that each simulation step has completely different requirements on the underlying spatial discretization. Therefore, meshing is still one of the major showstoppers in the field of TCAD. In this work we present a comprehensive modular, library centered approach which focuses on the problems described above. Already existing concepts and libraries in the field of scientific computing, which have proven to be successful, are combined on a new software basis.

4.2 Problems with Mesh Generation

Before introducing different techniques of mesh generation, common problems are sketched which impose various constraints, restrictions, and difficulties for the task of mesh generation in general.

- **Different element size:** The generated mesh can consist of elements of different size, meaning, that adjacent elements vary greatly in size, e.g. a thin layer of oxide within a three-dimensional device, which is more than three order of magnitudes different in size.
- **Boundary requirements:** Various boundary requirements cause problems for mesh generation, e.g. a tetrahedra configuration known as Schönhart polyhedra [39] or the inclusion of preset surface edges, which imposes additional difficulties, if the surface requires Delaunay conformity [40].

- **Cospherical point sets:** Points are said to be cospherical, if at least n + 2 point are located on the perimeter of an n-dimensional sphere S, where S does not contain any other points of the mesh. This problem only arises, when the Delaunay criterion has to be met.
- **Degenerated cells:** Degenerated cells may result due to several reasons, such as the requirement to adhere to a given boundary or that the generated mesh has to include a given point cloud. The problem with these degenerated cells is that they may be unsuitable for the subsequent simulation tasks, e.g. interpolation of physical quantities. This problem increases further, as degenerated cells cannot be refined without adding additional degeneration. Therefore several classes of degeneration are classified and various mesh adaptation techniques are used to minimize the number of these elements.

As already pointed out, the application design for mesh generation not only has to deal with the given difficulties based on mesh generation itself, but it also requires various software engineering methodologies to develop robust applications.

4.3   Meshing Methods

There are several methods for mesh generation, which are currently in use for scientific computing, such as:

- Giftwrapping
- Incremental Delaunay Triangulation
- Advancing Front
- Advancing Front with Delaunay Triangulation

To present the advantages and disadvantages of these different approaches, a brief overview of each is given in the following:

4.3.1   Giftwrapping

Starting with any element, a sphere is expanded around the element until the next point is found, where the extension of the sphere is controlled by a given direction vector. Figure 21 illustrates the idea behind the giftwrapping algorithm, which is suitable for use in multiple dimensions. Only convex hulls can be created using the giftwrapping algorithm and cospherical point sets impose additional difficulties.

4.3.2   Incremental Delaunay Triangulation

Many triangulations exist for a set of points in a plane. A triangulation which meets the so called circumcircle
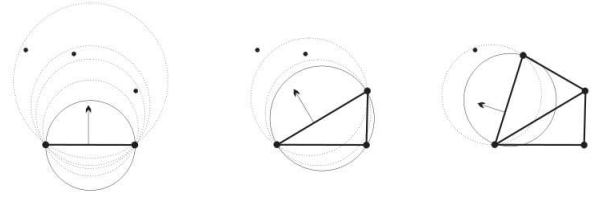

Figure 21: Giftwrapping example

or Delaunay property is called a Delaunay triangulation. This property states that the circumcircles of the triangles may contain no points beside the points forming the triangles (see Figure 22).

One of the main advantages of the Delaunay triangulation in two dimensions is the minimization of the maximum angles and the maximization of the minimum angles. In three dimensions this advantages cannot be easily accomplished due to non-unique flip constraints [41].
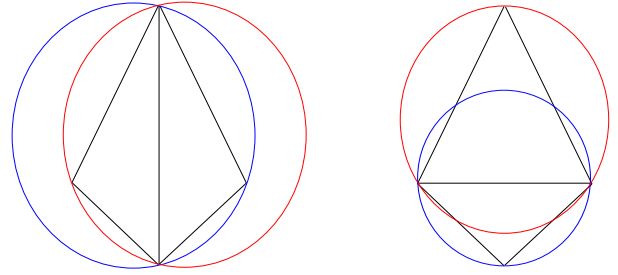

Figure 22: Delaunay property

An additional difficulty for this method is the incorporation of boundaries due to the fact that incremental Delaunay triangulation uses the convex hull of the input point set only. Given boundary representations have to be marked and reconstructed after the meshing process has completed.

Figure 22 shows two examples of a simple mesh. The left part violates the Delaunay criterion, because the cirumcircles include the points of both triangles. The right part is a correct Delaunay triangulation, because the circumcircles of both triangles contain only their respective three points.

4.3.3   Advancing Front

The advancing front algorithm starts with a set of boundary elements of a given dimension, e.g. edges. These edges form the initial front which is advanced into the simulation domain. An edge of this set is chosen to form a new triangle, either with an existing point or a newly created point. The current edge is then removed from the front and the two new edges are, depending on their vis-

ibility, added to the front. This process terminates when no edges remain within the front.

Figure 23 shows a simple advancing front triangulation. The figure sketches the advancing front mesh generation concept in two dimensions. The dotted line represents the current front. New triangles are inserted, by joining the two ends of a front edge to either a newly created point or an existing point, one at a time [42].
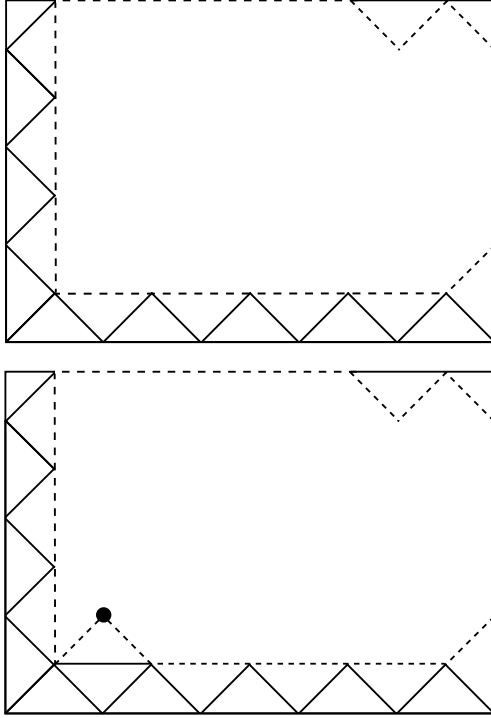

Figure 23: Advancing Front [42]

In contrast to the Delaunay triangulation, the advancing front method, easily incorporates the boundary. It starts from the given boundary representation and advances the front into the simulation domain. A major drawback of this method is that the quality of the generated elements heavily depends on the quality of the boundary elements. For different implementations of this type of mesh generation technique, the robustness issues are severe. The advantages of this method are the good control mechanism for the element sizes, the quality of the generated elements, and the not required Delaunay property which can optionally be incorporated easily. Parallelization can be quite challenging for surface creation, because new surface elements depend on the previously generated surface element. If the surface is partitioned, volume mesh generation, on the other hand, can be parallelized in order to distribute the generation process on several computers. Using the advancing front algorithm, it is no problem to divide the simulation domain into smaller domains and treat each domain on its own. However, boundary information, e.g. different materials, has to be available, otherwise the problem is similar as with sur-

face mesh generation. A method of incorporating the Delaunay property into the advancing front method has already been proposed [42], which adds new points ahead of the front and triangulates them according to the Delaunay criterion. This algorithm works like the conventional advancing front algorithm, only point insertion is extended to satisfy the Delaunay criterion.

Figure 24 shows the idea behind the algorithm. When inserting a new point, there are two possibilities:

- either the point is not inside any existing triangle circumcircle,
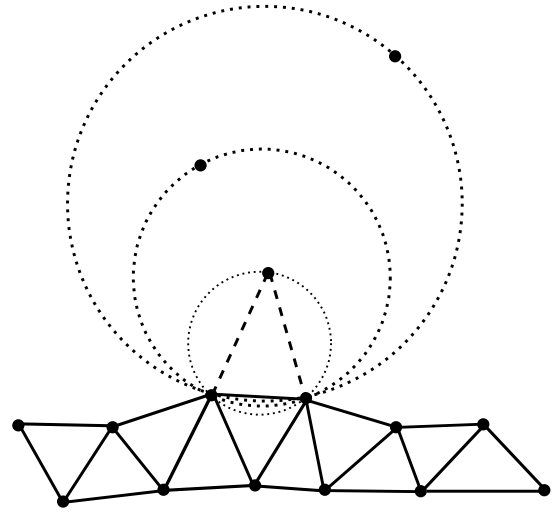- or there exists at least one triangle, whose circumcircle contains this new point.

Figure 24: Advancing Front Delaunay Triangulation [42]

A different implementation is available using an abstract rule set [43], which describes different methods of controlling the advancing front. Basically, a rule describes where points are to be generated. The rule also manages changes in the advancing front. Some boundary elements have to be added while others have to be deleted. If no matching rule is found, the quality class is decreased. This allows the application of a rule of a lower quality level, which may also delete the current element.

4.4   Software Design

The meshing methods, described in the previous section, have been included as modules in the Generic Scientific Simulation Environment (GSSE [44]). The software design of our meshing library can be seen in Figure 25. The following presents our solution of the previously mentioned meshing problems.

- **Different element size**
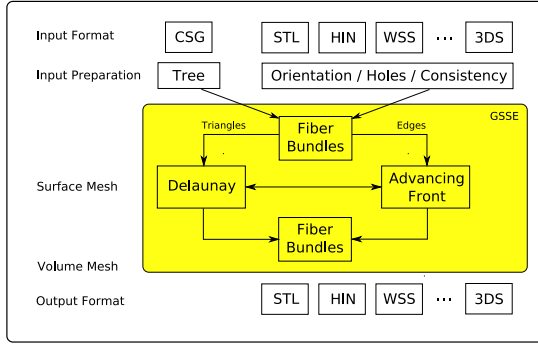 To cope with the problem of different element sizes we use the concept of a local feature size (lfs) [45]. The

Figure 25: Design of GSSE

lfs of a point p is defined as the radius of the smallest disk centered at p, which intersects two non-incident vertices or segments of the planar straight line graph. The lfs guarantees that there are no abrupt changes in the size of the elements, which is especially important with the advancing front algorithm.

- **Boundary requirements**
The application of the advancing front algorithm based on the introduced abstract rule mechanism easily incorporates the given boundary elements. Additionally, the Delaunay criterion can be used to create conforming boundary representations.

- **Cospherical point sets**
This problem arises, when a mesh is used for, e.g. the finite volume method. The Voronoi graph [41] is therefore used to model flux conservation. Due to the fact that a Voronoi graph is the dual graph of a Delaunay graph, the input mesh has to exhibit the Delaunay property. Our approach optionally guarantees Delaunay conformity for both the surface and the volume.

- **Degenerated simplices**
The use of the advancing front algorithm in combination with the already introduced quality rules and the corresponding point placement strategies reduces the degeneration of elements. The existing rules were extended to meet the Delaunay criterion.

Related to software engineering issues, the proposed solution uses a library centric application design paradigm [44] based on a multi-paradigm approach.

To incorporate various input formats and to solve the problems with different data models, the proposed approach uses a fiber bundle data model. Fiber bundles were first introduced by Butler and Pendley [46] and afterwards enhanced [47] to include cell complex properties. The basic approach of this data model is the separation of cells into a base space and the connectivity information storage within a fiber space. All current data models can be mapped to the fiber bundle data model. This gives the advantage to read and write various file formats. Due to the modular design it is also possible to

quickly write new input or output modules to accommodate an even wider range of file formats.

An advantage of our approach is, that the developed modules can be combined in various ways, depending on the requirements of the subsequent simulation. For example a mesh, initially generated using the advancing front algorithm, can be refined afterwards at certain points, which are interesting for the simulation, or the mesh can be modified to meet the Delaunay criterion. The module design can also be seen in Figure 25. Another advantage of the modularized development approach is the possibility to interface different mesh engines and thereby fully utilize the orthogonal application approach. Our library also contains a resampling step which consists of several substeps [43]. First the points are extracted from the input model. This extraction only applies to the 'points of intersection', which are the points where at least three surfaces intersect. To solve this problem, a bisection algorithm based on geometric tests is used. Afterwards, using this minimal mesh, a new surface is generated, with an advancing front algorithm [43].

4.5 Meshing Libraries

As it is our goal to focus on library centric design, various meshing libraries are incorporated and therefore briefly presented here. **TetGen:** is developed by Hang Si, from the Weierstrass Institute for Applied Analysis and Stochastics, and generates a Delaunay triangulation. Additionally, it supports the creation of Voronoi diagrams and convex hulls for three-dimensional point sets [48–50]. **Netgen:** uses the described abstract rules for advancing front mesh generation and offers a Delaunay point cloud generation module [43]. This module lacks the modeling of a guaranteed Delaunay volume mesh as well as the conforming Delaunay property for surfaces. **deLink:** was developed for TCAD applications based on an advancing front method combined with a Delaunay method which does not create new points automatically [51]. Therefore, additional point clouds created separately have to be used to refine structures. The created meshes satisfy the Delaunay criterion for surfaces and volumes.

4.6 Our Approach

It is important to highlight that TetGen and Netgen cannot deal with corrupted input representations, e.g. holes. deLink and GSSE use various methods to correct the input representations, where our approach can also use a separate surface remeshing step. GSSE also incorporates constructive solid modeling [44]. Table 3 presents an

overview of the features of each mesh generation module.

Features	TetGen	Netgen	deLink	GSSE	
Delaunay	+	x	+	+	
Advancing Front	-	+	-	+	
Remesh Steps	-	+	-	+	
Automatic Points	-	+	-	+	
CSG Input	-	+	-	+	
Polygon Input	-	-	+	+	
Repair Surface	-	-	+	+	
Multi Material	-	-	+	+	

Table 3: Results of the comparison

4.7   Examples

This section presents some examples of meshes, created by the four previously outlined tools. The examples are analyzed and the quality distribution is presented. To compare the meshing tools, we categorize the elements of the mesh using two main parts [52]. First four classes of quality types, defined by the number of small dihedral angles, as shown in Figure 26 are set up. Then, the tetra-
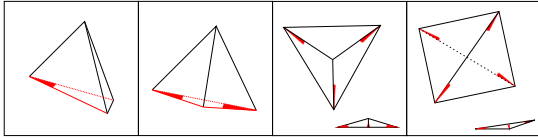

Figure 26: Four different classes of degenerated tetrahedra (wedge, spade, cap, sliver), sorted by the number of acute dihedral angles.

hedra are classified by the number of degenerated triangles, like daggers and blades. The dagger has one short edge and at least one small angle, where the blade has no short edge and therefore one large and two small angles. Figure 27 shows the needle (or spire) with three daggers (the short edges are marked in the figure), the slat (or splinter) with two opposite short edges and, therefore, four daggers. Finally, the spindle does not have short edges and has therefore four blades as triangles.
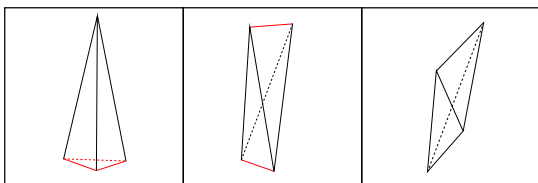

Figure 27: Three different types of degenerated tetrahedra (needle, slat, spindle).

The Figures 28 and 30 show the basic functionality and

the boundary representation of the different meshing kernels. The Figures 29 and 31 summarize the respecitve statistics. The pictures in Figure 32 and 33 present examples from TCAD / topography simulation.

4.8   Conclusion

We have identified several problems of mesh generation and presented our solution to these problems. A meshing library based on GSSE has been developed to collect these solutions. Finally, a comparison between GSSE and established meshing kernels was performed showing, that GSSE has more mesh generation possibilities than the other investigated tools. The analysis of the examples also shows that GSSE produces an output of at least the same performance and even better quality than the other tools.
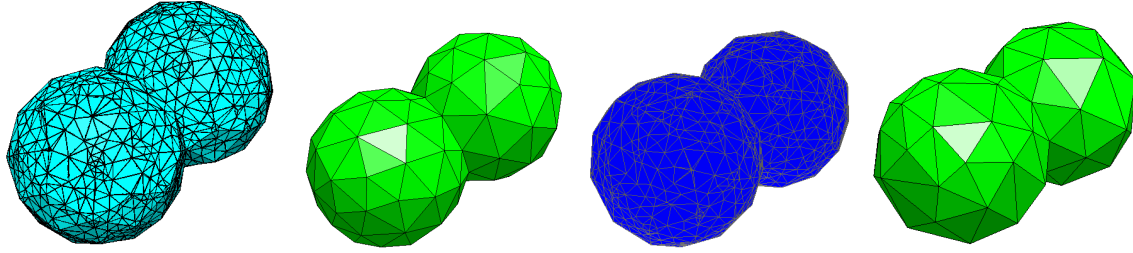
Figure 28: A simple sphere, but yet a complex task for a mesh generator, is used to highlight the different approaches related to the investigated types of meshing algorithms, from left to right: TetGen, Netgen, deLink, GSSE.
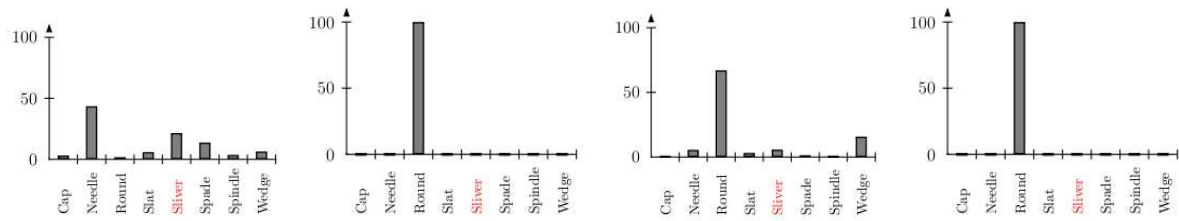

Figure 29: Statistics for the meshing example in Figure 28, from left to right: TetGen, Netgen, deLink, GSSE.
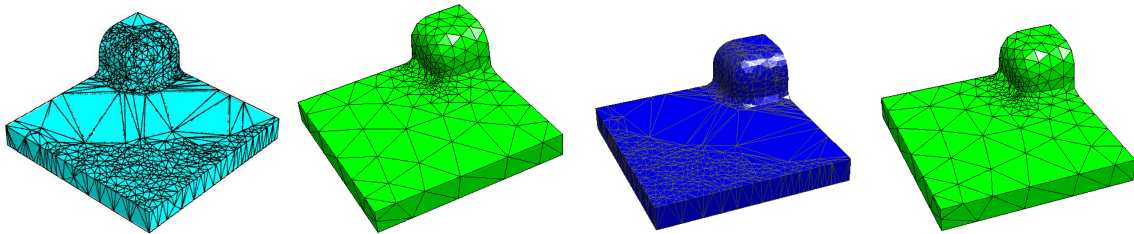

Figure 30: An example from topography simulation is used as example to illustrate the different types of meshing results, from left to right: TetGen, Netgen, deLink, GSSE.
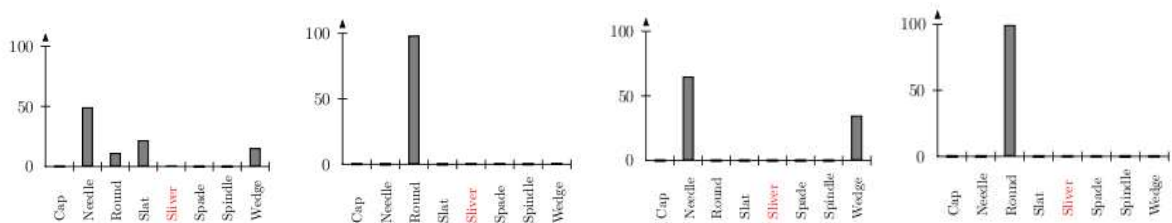

Figure 31: Statistics for the meshing example in Figure 30, from left to right: TetGen, Netgen, deLink, GSSE.
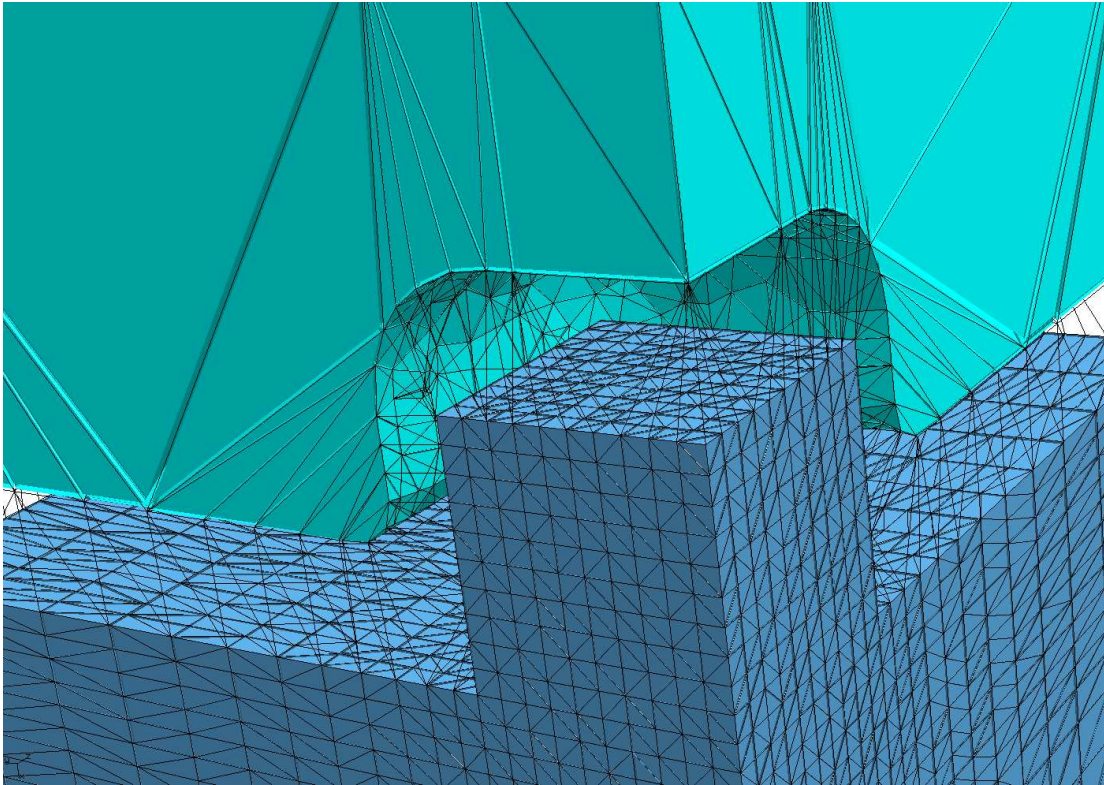
Figure 32: The complete example from topography simulation with the additional constraint of including a given point cloud as can be seen in the bottom part of the structure. The middle part is made transparent to expose the surface mesh. The top part is not required for subsequent simulation and is therefore meshed without quality constraints.
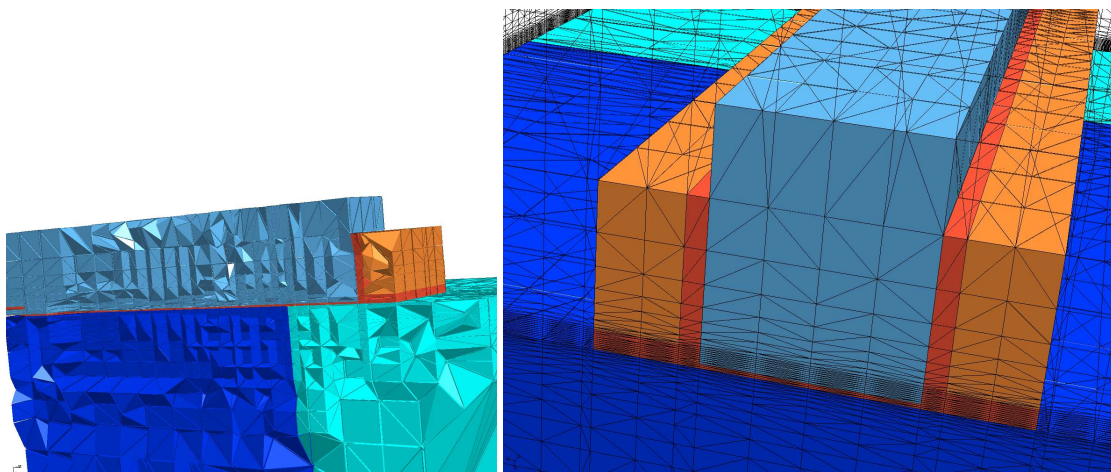

Figure 33: A three-dimensional device structure for a MOSFET with an additional externally supplied point cloud. The important part is the regularity of the elements in the channel region. The different aspect ratios, e.g. the thin red oxide part and the large blue silicon part, are also an additional complication for the mesh generation algorithm.

5 Effects of Shear Strain on the Conduction Band in Silicon: An Efficient Two-Band k·p Theory

We present an efficient two-band **k·p** theory which accurately describes the six lowest conduction band valleys in silicon. By comparing the model with full band pseudo-potential calculations we demonstrate that the model captures both the nonparabolicty effects and the stress-induced band structure modification for general stress conditions. It reproduces the stress dependence of the effective masses and the nonparabolicity parameter. Analytical expressions for the valley shifts and the transversal and longitudinal effective mass modifications induced by uniaxial [110] stress are obtained and analyzed. The low-field mobility enhancement in the direction of tensile [110] stress in {001} SOI FETs with arbitrary small body thickness is due to a modification of the conductivity mass and is shown to be partly hampered by an increase in nonparabolicity at high stress value.

5.1 Introduction

The **k·p** theory is a well established tool to describe the band structure analytically. After the pioneering work by Luttinger and Kohn [53] the six-band **k·p** method has become a standard approach to model the valence band in Si. However, the conduction band in Si is usually approximated by three pairs of equivalent minima located close to the X-points of the Brillouin zone. It is commonly assumed that close to the minima the electron dispersion is well described by the effective mass approximation. The nonparabolicity parameter $\alpha = 0.5 \, \text{eV}^{-1}$ is introduced to describe deviations in the density of states from the purely parabolic expression, which become pronounced at higher electron energies. In ultra-thin body (UTB) FETs, however, the band nonparabolicity affects the subband quantization energies substantially, and it was recently indicated that anisotropic, direction-dependent nonparabolicity can explain the mobility behavior at high carrier concentrations in a FET with (110) UTB orientation [54]. Therefore, a more refined description of the conduction band minima beyond the usual nonparabolic approximation is needed. Another reason to challenge the standard description of the conduction band based on a single-band nonparabolic approximation is its inability to properly address the band structure modification under stress.

Stress-induced mobility enhancement in Si has become a key technique to increase performance of modern CMOS devices. In biaxially stressed devices the electron mobility can be nearly doubled. The reason for the mobility enhancement lies in the stress-induced band structure

modification: the degeneracy between the six equivalent valleys is lifted due to stress-induced valley shifts. This reduces inter-valley scattering. In case of tensile biaxial stress applied in the {100} plane the four in-plane valleys move up in energy and become de-populated. The two populated out-of-plane valleys have favorable conductivity masses, which together with reduced inter-valley scattering results in the observed mobility enhancement [55].

Biaxial stress is naturally introduced by growing Si epitaxially on SiGe. This method, however, requires a substantial modification of the CMOS fabrication process and is not used in mass production. Instead, semiconductor industry is exploiting techniques compatible with existing CMOS process technologies when stress in the channel is created by local stressors and/or additional cap layers. Although already successfully used in mass production, the technologically relevant case of stress along [110] has received little attention within the research community. Only recently a systematic experimental study of the mobility modification due to [110] stress was performed [56]. It was shown that, contrary to [100] uniaxial stress, the electron mobility data for [110] stress suggest that the conductivity mass depends on stress. This conclusion was also supported by recent results of pseudo-potential band structure calculations [56, 57]. Shear strain modifies substantially both the longitudinal [58, 59] and transversal [56, 58–60] effective masses. Any dependence of the effective masses on stress is absent within the standard parabolic effective mass description of the conduction band and can only be introduced phenomenologically. In order to describe the dependence of the effective mass on stress a single-band description is not sufficient, and coupling to other bands has to be taken into account.

Recently, a 30 bands **k·p** theory was introduced [61]. Although universal, it cannot provide an explicit analytical solution for the energy dispersion. In this work we present an efficient two-band **k·p** theory. By comparing our results with predictions of the pseudo-potential band structure calculations we demonstrate that the theory accurately describes both the nonparabolicty effects and the stress induced band structure modification for general stress conditions. It accurately reproduces the stress dependence of the effective mass and of the nonparabolicity parameter. The analytical two-band **k·p** model allows one to study the influence of the conduction band structure on transport properties of stressed FETs for general UTB orientations.
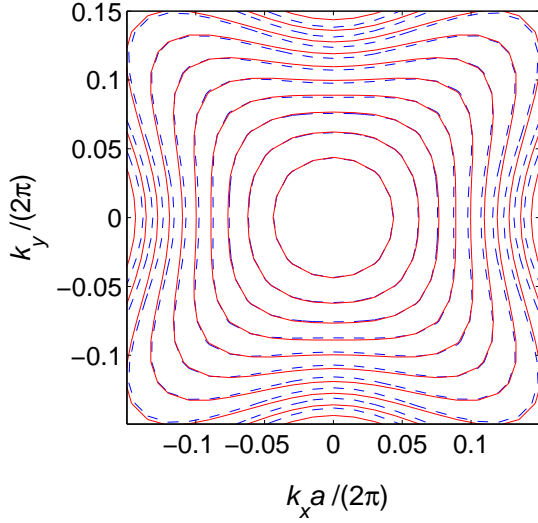
Figure 34: [001] valley band structure obtained from EPM calculations (lines) and the analytical expression (35) (dotted lines) at $k_z = -k_0$. The distance between contour lines is 50 meV. Nonparabolicity is strongly direction dependent.

5.2 Theory

We consider the pair of equivalent conduction band valleys along the [001] direction. Other valleys can be analyzed analogously.

5.2.1 Two-Band k·p Theory

The closest band to the first conduction band Δ_1 ($i = 1$), which we take into account, is the second conduction band $\Delta_{2'}$ ($i = 2$). These two bands become degenerate exactly at the X point. Since the minimum of the conduction band is only $k_0 = 0.15\frac{2\pi}{a}$ away from the X point, the dispersion around the minimum can be well described by degenerate perturbation theory, which only includes the two bands degenerate at the X point. Diagonal elements of the Hamiltonian $H_{ii}, i = 1, 2$ can be easily obtained using the standard k·p theory:

$$H_{ii}^0(k) = (-1)^{i-1}\frac{\hbar}{m_0}k_z p + \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 k_x^2}{2m_t} + \frac{\hbar^2 k_y^2}{2m_t}, \quad (33)$$

where m_0 is the free electron mass, m_t is the transversal, and m_l is the longitudinal effective mass. Here we took into account that the matrix elements $(p_z)_{ii}$ are different only in sign, which is positive for the lower band: $p = (p_z)_{11} = -(p_z)_{22}$. The values of k_z are negative since they are counted from the X point. In contrast to the 30 bands k·p theory, which is developed around the Γ point far away from the conduction band minimum [61], our perturbation analysis at the X point allows to get excellent results with only two bands. Taking into account the diagonal elements (33), we recover the commonly
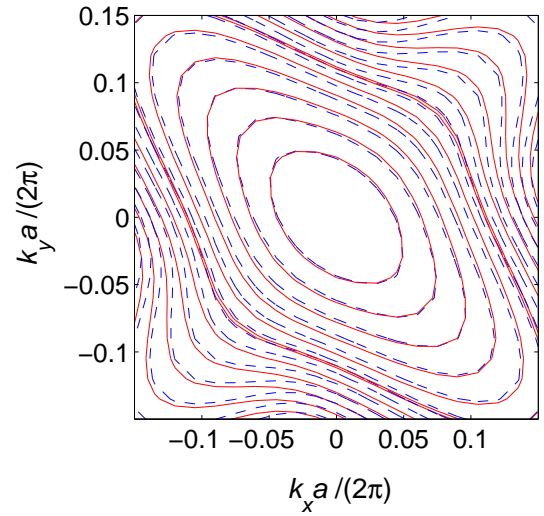

Figure 35: [001] valley band structure obtained from EPM calculations (lines) and the analytical expression (38) (dotted lines) at $k_z = k_{\min}$, for tensile [110] uniaxial stress of 3 GPa. The distance between the contour lines is 50 meV. Strong stress-induced anisotropy in the transversal mass is observed.

used dispersion for the conduction band (the linear term vanishes at the minimum $k_z = -k_0$). The coupling between the bands is described by the off-diagonal terms which up to the second order are:

$$H_{12}^0(k) = \frac{\hbar^2 k_x k_y}{M}. \quad (34)$$

The parameter M is obtained from k·p perturbation theory [60]:

$$\frac{1}{M} = \frac{2}{m_0^2}\left|\sum_{l \neq 1,2}\frac{(p_y)_{1l}(p_z)_{l2}}{E_k(X) - E_{\Delta_1}(X)}\right|.$$

Using degenerate perturbation theory, we find the following dispersion relation close to the minimum at $k_z = -k_0$:

$$
\begin{aligned}
E_0(\mathbf{k}) = \quad &\frac{\hbar^2(\delta k_z)^2}{2m_l} + \frac{\hbar^2(k_x^2 + k_y^2)}{2m_t} \\
&- \frac{\Delta}{2}\left(\left[1 + \left(\frac{2\hbar^2 k_x k_y}{M\Delta}\right)^2\right]^{1/2} - 1\right),
\end{aligned}
\quad (35)
$$

where $\delta k_z = k_z + k_0$, $\Delta = 2\hbar k_0 p/m_0$ is the gap between the Δ_1 and the $\Delta_{2'}$ conduction bands at $k_z = -k_0$. In Figure 34 this analytical expression (dotted contour lines) is compared to the numerical band structure obtained from the empirical pseudo-potential method (EPM) for $k_z = -k_0$. Excellent agreement is found up to an energy of 0.5 eV. Figure 34 demonstrates strong anisotropy in the nonparabolicity parameter, as anticipated in [54].

5.2.2   Stress

In order to account for stress in our model we consider again the valley along the [001] direction. For general stress conditions the following shift in energy is added to the diagonal matrix elements (33) [62]:

$$H_{ii} = H_{ii}^0 + \delta E_C, \tag{36}$$

where $\delta E_C = \Xi_d \left(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz} \right) + \Xi_u \, \varepsilon_{zz}$, with Ξ_d denoting the dilation and Ξ_u the uniaxial deformation potentials for the conduction band. The $\varepsilon_{ll}, l = x, y, z$ are the diagonal components of the strain tensor expressed in the principal coordinate system. The off-diagonal elements of the Hamiltonian are also modified by strain [60]:

$$H_{ij}(k) = H_{ij}^0 - D\varepsilon_{xy}, \tag{37}$$

where $D \geq 0$ denotes the deformation potential for the off-diagonal strain component. When the off-diagonal components in the Hamiltonian are ignored, the influence of the shear stress component is completely lost. The off-diagonal elements of the strain tensor are, however, generated by [110] uniaxial stress. Since this is exactly the stress direction used to enhance the performance of modern MOSFETs, shear strain must be taken into consideration. The dispersion relation of the [001] valleys including the shear strain component for the conduction band now reads as:

$$\begin{aligned} E(\mathbf{k}) = & \ \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 (k_x^2 + k_y^2)}{2m_t} + \delta E_C \\ & - \left[\left(\frac{\hbar}{m_0} k_z p \right)^2 + \left(D\varepsilon_{xy} - \frac{\hbar^2 k_x k_y}{M} \right)^2 \right]^{1/2}, \end{aligned} \tag{38}$$

where the value of ε_{xy} is positive for tensile stress in [110] direction. In Figure 35 the analytical band struc-
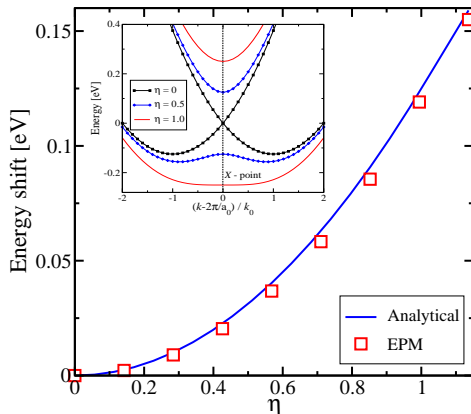

Figure 36:   [001] valley energy shift as function of the dimensionless off-diagonal component of the strain tensor, as predicted by (40) and by EPM calculations. Inset: conduction band profile along the [001] direction for different stress values.
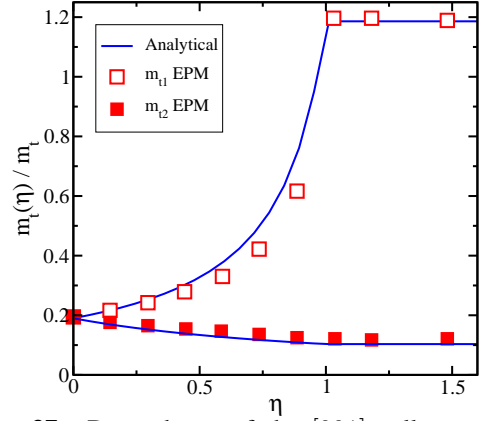

Figure 37:   Dependence of the [001] valley transversal effective mass on the dimensionless [110] uniaxial strain as predicted by (41,42) (lines) and EPM calculations (symbols). Shear stress generates strong anisotropy in the transversal masses
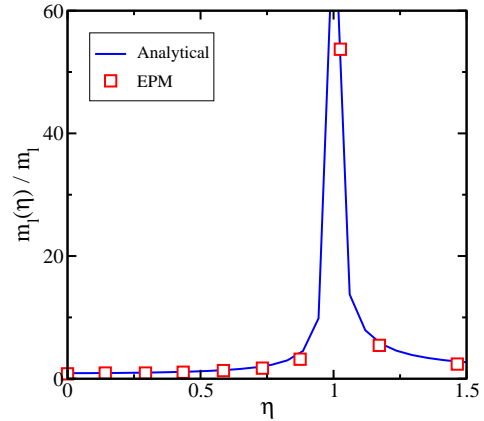

Figure 38:   Stress dependence of longitudinal effective mass in the [001] valleys due to [110] stress. Effective mass diverges at $\eta = 1$ suggesting that full-band theory must be used for such stress values.

ture (38) is compared with the results of the EPM calculations for uniaxial [110] tensile stress of 3 GPa. Even for such large stress values the agreement between the analytical model and the numerical EPM results is excellent up to 200 meV. The band structure shown in Figure 35 suggests a strong effective mass modification, which is analyzed in more details in the next section.

5.3   Conduction Band Modification Due to Shear Strain

The usually ignored off-diagonal strain component lifts the degeneracy between the two lowest conduction bands at the X points along the [001] axis in the Brillouin zone [60]. This lifting of degeneracy has a strong effect on the band structure. We investigate the shifts of

the valley minima, changes in the effective masses and in the nonparabolicity parameter.

5.3.1 Valley Shifts

Since the conduction band minimum along the $[001]$ axis is located near the X point, the gap opening at the X point affects the position of the minimum. First, the conduction band minimum $k_{\min}$ moves closer to the X point. From (38) we obtain

$$k_{\min} = -k_0\sqrt{1-\eta^2}. \qquad (39)$$

Here, the dimensionless off-diagonal strain $\eta = 2D\varepsilon_{xy}/\Delta$ is introduced. For $\eta \geq 1$ the conduction band minimum is located exactly at the X point.

The minima of the two $[001]$ valleys move down in energy with respect to the remaining four fold degenerate valleys. For $\eta \leq 1$ the strain dependence is quadratic, while it is linear for $\eta \geq 1$:

$$\Delta E_{\text{shear}} = \begin{cases} -\frac{\Delta}{4}\eta^2 & , \quad |\eta| < 1 \\ -(2|\eta|-1)\Delta/4 & , \quad |\eta| > 1 \end{cases} \qquad (40)$$

In Figure 36 the shifts predicted by (40) are compared with results from EPM calculations. Excellent agreement is found.

5.3.2 Stress Dependent Effective Masses

Shear strain modifies the effective masses in the $[001]$ valleys. Evaluating the corresponding second derivatives of (38) at the band minimum (39), we obtain two different branches for the effective mass across (m_{t1}) and along (m_{t2}) the stress direction:

$$m_{t1}(\eta)/m_t = \begin{cases} (1-\eta\frac{m_t}{M})^{-1} & , \quad |\eta| < 1 \\ (1-\text{sgn}(\eta)\frac{m_t}{M})^{-1} & , \quad |\eta| > 1 \end{cases} \qquad (41)$$

$$m_{t2}(\eta)/m_t = \begin{cases} (1+\eta\frac{m_t}{M})^{-1} & , \quad |\eta| < 1 \\ (1+\text{sgn}(\eta)\frac{m_t}{M})^{-1} & , \quad |\eta| > 1 \end{cases} \qquad (42)$$

Here, sgn denotes the sign function. The analytical expressions for the transversal masses (41) and (42) are compared with the masses obtained from EPM calculations in Figure 37. Strong anisotropy in the transversal masses generated by shear strain is predicted by the analytical model.
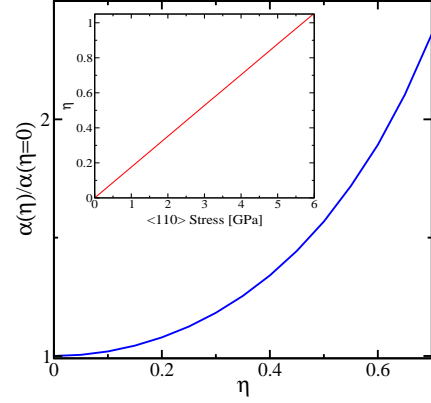

Figure 39: Nonparabolicity parameter in the $[001]$ valleys increases as function of $[110]$ tensile stress. Inset: Relation between dimensionless strain η and stress in GPa.

For the longitudinal effective mass one obtains the following expression from (38):

$$m_l(\eta)/m_l = \begin{cases} (1-\eta^2)^{-1} & , \quad |\eta| < 1 \\ (1-1/|\eta|)^{-1} & , \quad |\eta| > 1 \end{cases} \qquad (43)$$

Eq.(43) is compared with EPM results in Figure 38. The longitudinal mass diverges at $\eta = 1$ suggesting that a full-band description is necessary for such high stress values [58].

5.3.3 Stress and Nonparabolicity

Shear strain affects the value of the nonparabolicity parameter α. Proceeding as in [63], we arrive at an expression for the strain dependence of α:

$$\alpha(\eta) = \alpha_0 \frac{1+2(\eta m_t/M)^2}{1-(\eta m_t/M)^2} \qquad (44)$$

Expression (44) is plotted in Figure 39. The relative increase of $\alpha(\eta)$ is important at large stress values. Results of the mobility simulations in a strained ultra-thin body FET along the $[110]$ stress direction, with and without stress dependence of the the nonparabiolicity parameter taken into account, are shown in Figure 40. The stress dependence of the nonparabolicity parameter results in an almost 25% decrease to the mobility enhancement in a 3 nm thick SOI FET at a stress level of 3 GPa (Figure 40). For stress values larger than 3 GPa the energy difference from the minimum to the value at the X point becomes smaller than k_BT, and a full-band description is required [58].

5.4 Conclusion

An efficient two-band **k·p** model is presented, which accurately describes the conduction band minima in
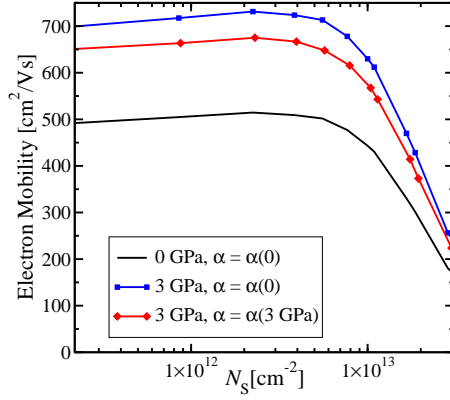
Figure 40: [110] channel mobility in a 3 nm thick UTB FET at 3 GPa tensile stress along the channel. Mobility is computed with $\alpha = \alpha_0$ and $\alpha = \alpha(\eta)$.

strained silicon. The model accurately describes stress dependences of the effective mass and of the nonparabolicity parameter. Analytical dependences of the valley shifts, transversal and longitudinal effective masses, and the nonparabolicity parameter on shear strain are obtained and analyzed. It is demonstrated that the enhancement of low-field mobility in uniaxially stressed UTB FETs is partly hampered by an increase in nonparabolicity at higher stress.

References

[1] Roger Lake, Gerhard Klimeck, R. Chris Bowen, and Dejan Jovanovic. Single and Multiband Modeling of Quantum Electron Transport Through Layered Semiconductor Devices. *J. Appl. Phys.*, 81(12):7845–1869, 1997.

[2] R. C. Bowen, G. Klimeck, R. K. Lake, W. R. Frensley, and T. Moise. Quantitative Simulation of a Resonant Tunneling Diode. *J. Appl. Phys.*, 81(7):3207–3213, 1997.

[3] A. Cresti, R. Farchioni, G. Grosso, and G. Pastori Parravicini. Keldysh-Green Function Formalism for Current Profiles in Mesoscopic Systems. *Phys.Rev.B*, 68(7):075306–1–075306–8, 2003.

[4] M. Pourfath, H. Kosina, and S. Selberherr. A Fast and Stable Poisson-Schrödinger Solver for the Analysis of Carbon Nanotube Transistors. *J. Comp. Electron.*, 5(2):155–159, 2006.

[5] S. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal. Two-Dimensional Quantum Mechanical Modeling of Nanotransistors. *J. Appl. Phys.*, 91(4):2343–2354, 2002.

[6] D. John, L. C. Castro, P. J. S. Pereira, and L. Pulfrey. A Schrödinger-Poisson Solver for Modeling Carbon Nanotube FETs. In *NSTI Nanotech*, volume 3, pp 65–68, 2004.

[7] A. Björck and V. Pereyra. Solutions of Vandermonde Systems of Equations. *Math. Comp.*, (24):893–903, 1970.

[8] L. Fejér. Mechanische Quadraturen mit positiven Cotesschen Zahlen. *Math. Z.*, 37:287–309, 1933.

[9] C.W. Clenshaw and A.R. Curtis. A Method for Numerical Integration on an Automatic Computer. *Num. Math.*, 2:197–205, 1960.

[10] J. Waldvogel. Fast Construction of the Fejr and ClenshawCurtis Quadrature Rules. *BIT Numerical Mathematics*, 46(1):195, 2006.

[11] T. O. Espelid. Doubly Adaptive Quadrature Routines Based on Newton-Cotes Rules. *BIT Numerical Mathematics*, 43:319–337, 2003.

[12] J.F. Conley, Jr. P.M. Lenahan, and P. Roitman. Room Temperature Reactions Involving Silicon Dangling Bond Centers and Molecular Hydrogen in Amorphous Si02 Thin Films on Si. *Appl. Phys. Lett.*, 62(40), 1993.

[13] P.E. Blöchl and J. H. Stathis. Hydrogen Electrochemistry and Stress-Induced Leakage Current in Silica. *Phys. Rev. Lett.*, 83(2):372–375, 1999.

[14] M. Boero, A. Pasquarello, J. Sarnthein, and R. Car. Structure and Hyperfine Parameters of E_{01} Centers in α-Quartz and in Vitreous SiO_2. *Phys. Rev. Lett.*, 78(5):887–890, 1997.

[15] G. Kresse and D. Joubert. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev.*, 59(3), 1999.

[16] R. M. Van Ginhoven, H. Jonsson, and L.R. Corrales. Silica Glass Structure Generation for Ab Initio Calculations Using Small Samples of Amorphous Silica. *Phys. Rev. B*, 71, 2005.

[17] S.T. Pantelides, S. N. Rashkeev, R. Buczko, D. M. Fleetwood, and R. D. Schrimpf. Reactions of Hydrogen with Si-SiO2 Interfaces. *IEEE Trans. Nucl. Sci.*, 47(6):2262–2268, 2000.

[18] A. Yokozawa and Y. Miyamoto. First-Principles Calculations for Charged States of Hydrogen Atoms in SiO2. *Phys. Rev. B*, 55(20):13783–13788, 1997.

[19] Y. Cao, X. Huang, D. Sylvester, T.-J. King, and C. Hu. Impact of On-Chip Interconnect Frequency-Dependent R(f)L(f) on Digital and RF Design. In *Proc. IEEE International ASIC-SoC Conference*, 2002.

[20] A. Nieuwoudt and Y. Massoud. Variability-Aware Multilevel Integrated Spiral Inductor Synthesis. 25(12):2613–2625, 2006.

[21] F. Y. Huang, J. X. Lu, D. M. Jiang, X. C. Wang, and N. Jiang. A Novel Analytical Approach to Parameter Extraction for On-Chip Spiral Inductors Taking into Account High-Order Parasitic Effect. *Solid-State Electronics*, 50(9-10):1557–1562, 2006.

[22] J. Jin. *The Finite Element Method in Electromagnetics*. John Wiley and Sons, New York, 2002.

[23] P. Fleischmann and S. Selberherr. A New Approach to Fully Unstructured Three-dimensional Delaunay Mesh Generation with Improved Element Quality. In *Proc. Simulation of Semiconductor Processes and Devices*, pp 129–130, 1996.

[24] P. Fleischmann, R. Sabelka, A. Stach, R. Strasser, and S. Selberherr. Grid Generation for Three-Dimensional Process and Device Simulation. In *Proc. Simulation of Semiconductor Processes and Devices*, pp 161–166, 1996.

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

[26] A. Prechtl. Vorlesungen über Theoretische Elektrotechnik. Institut für Grundlagen umd Theorie der Elektrotechnik, TU Wien, 1996. Zweiter Teil: Elektrodynamik.

[27] P. Meuris, W. Schoenmaker, and W. Magnus. Strategy for Electromagnetic Interconnect Modeling. 20(6):753–762, 2001.

[28] O. Biró, K. Preis, and K. R. Richter. On the Use of the Magnetic Vector Potential in the Nodal and Edge Finite Element Analysis of 3D Magnetostatic Problems. 32(3):651–654, 1996.

[29] H. Whitney. *Geometric Integration Theory*. Princeton, NJ:Princeton University Press, 1957.

[30] M. Dorica and D. Giannacopoulos. Impact of Mesh Quality Improvement Systems on the Accuracy of Adaptive Finite-Element Electromagnetics with Tetrahedra. 41(5):1692–1695, 2005.

[31] J. Schöberl. NETGEN - Automatic mesh generator. http://www.hpfem.jku.at/netgen.

[32] J. Schöberl. NETGEN - An Advancing Front 2D/3D-Mesh Generator Based On Abstract Rules. *Computing and Visualization in Science*, 1(1):41–52, 1997.

[33] H. Borouchaki and S. H. Lo. Fast Delaunay Triangulation in Three Dimensions. *Computer methods in applied mechanics and engineering*, 128(1):153–167, 1995.

[34] Y. Zhan, R. Harjani, and S. Sapatnekar. On the Selection of On-Chip Inductors for the Optimal VCO Design. In *Custom Integrated Circuits Conference*, pp 277–280, 2004.

[35] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit*. Kitware, Inc., USA, 2004.

[36] International Technology Roadmap for Semiconductors, 2005.

[37] A. J. Zomorodian. Topology for Computing. In *Cambridge Monographs on Applied and Computational Mathematics*, 2005.

[38] H. Edelsbrunner and D. Guoy. An Experimental Study of Sliver Exudation, 2001.

[39] E. Schönhart. Über die Zerlegung von Dreieckspolyedern in Tetraeder. In *Journal Mathematische Annalen*, 1928.

[40] P. Fleischmann. *Mesh Generation for Technology CAD in Three Dimensions*. Dissertation, Technische Universität Wien, 2000.

[41] J. R. Shewchuk. *Delaunay Refinement Mesh Generation*. Dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, 1997.

[42] D. Mavriplis. An Advancing Front Delaunay Triangulation Algorithm Designed for Robustness. *J. Comput. Phys.*, 117(1):90–101, 1995.

[43] J. Schöberl. NETGEN - An Advancing Front 2D/3D-Mesh Generator Based on Abstract Rules. *Comput. Visual. Sci.*, 1:41–52, 1997.

[44] R. Heinzl. *Concepts for Scientific Computing*. Dissertation, Technische Universität Wien, 2007.

[45] J. Ruppert. A Delaunay Refinement Algorithm for Quality 2-Dimensional Mesh Generation. *J. of Algorithms*, 18(3):548–585, 1995.

[46] D. Butler and M. Pendley. A Visualization Model Based on the Mathematics of Fiber Bundles. *Computers in Physics*, 3(5):45–51, 1989.

[47] W. Benger. *Visualization of General Relativistic Tensor Fields via a Fiber Bundle Data Model*. Dissertation, Freie Universität Berlin, 2004.

[48] H. Edelsbrunner and N. R. Shah. Incremental Topological Flipping Works for Regular Triangulations. *Algorithmica*, (15):223–241, 1996.

[49] H. Si and K. Gaertner. Meshing Piecewise Linear Complexes by Constrained Delaunay Tetrahedralizations. *14th Intl. Meshing Roundtable*, pp 147–163, 2005.

[50] H. Si. On Refinement of Constrained Delaunay Tetrahedralizations. *15th Intl. Meshing Roundtable*, 2006.

[51] P. Fleischmann and S. Selberherr. Enhanced Advancing Front Delaunay Meshing in TCAD. In *Proc. Conf. on Sim. of Semiconductor Processes and Devices*, pp 99–102, 2002.

[52] R. Heinzl and T. Grasser. Generalized Comprehensive Approach for Robust Three-Dimensional Mesh Generation for TCAD. In *Proc. Conf. on Sim. of Semiconductor Processes and Devices*, pp 211–214, Tokio, 2005.

[53] J. M. Luttinger and W. Kohn. Motion of Electrons and Holes in Perturbed Periodic Fields. *Physical Review*, 97(4):869–883, 1955.

[54] K. Uchida, A. Kinoshita, and M. Saitoh. Carrier Transport in (110) nMOSFETs: Subband Structure, Non-Parabolicity, Mobility Characteristics, and Uniaxial Stress Engineering. In *IEDM Techn. Dig.*, pp 1019–1021, 2006.

[55] S. I. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons. Comparative Study of Phonon-Limited Mobility of Two-Dimensional Electrons in Strained and Unstrained Si Metal-Oxide-Semiconductor Field-Effect Transistors. *J.Appl.Phys.*, 80(3):1567–1577, 1996.

[56] K. Uchida, T. Krishnamohan, K. C. Saraswat, and Y. Nishi. Physical Mechanisms of Electron Mobility Enhancement in Uniaxial Stressed MOSFETs and Impact of Uniaxial Stress Engineering in Ballistic Regime. In *IEDM Techn. Dig.*, pp 129–132, 2005.

[57] E. Ungersboeck, V. Sverdlov, H. Kosina, and H. Kosina. Electron Inversion Layer Mobility Enhancement by Uniaxial Stress on (001) and (100) Oriented MOSFETs. *International Conference on Simulation of Semiconductor Processes and Devices*, pp 43–46, 2006.

[58] E. Ungersboeck, S. Dhar, G. Karlowatz, V. Sverdlov, H. Kosina, and S. Selberherr. The Effect of General Strain on Band Structure and Electron Mobility of Silicon. *IEEE Trans.Electron Devices*, 2007.

[59] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr. Influence of Uniaxial [110] Stress on Silicon Band Structure and Electron Low-Field Mobility in Ultra-Thin Body SOI FETs. In *Proc. EUROSOI 2007*, pp 39–40, 2007.

[60] J. C. Hensel, H. Hasegawa, and M. Nakayama. Cyclotron Resonance in Uniaxially Stressed Silicon. II. Nature of the Covalent Bond. *Phys. Rev.*, 138(0):A225–A238, 1965.

[61] D. Rideau, M. Feraille, L. Ciampolini, M. Minondo, C. Tavernier, H. Jaouen, and A. Ghetti. Strained Si, Ge, and $Si_{1-x}Ge_x$ Alloys Modeled with a First-Principles-Optimized Full-Zone **k·p** Method. *Phys. Rev. B*, 74(19):195208, 2006.

[62] I. Balslev. Influence of uniaxial stress on the indirect absorption edge in Silicon and Germanium. *Physical Review*, 143:636–647, 1966.

[63] 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.