VISTA Status Report
June 2004

H. Ceric, S. Holzer, A. Sheikholeslami,
T. Ayalew, R. Wittmann, S. Selberherr

Institute for Microelectronics
Technische Universität Wien
Gusshausstrasse 27-29
A-1040 Wien, Austria
# Contents

1 The Resistance and Current Density During Electromigration  
   1.1 Introduction ......................................................... 1  
   1.2 Modelling Approach ............................................... 1  
      1.2.1 Average Current Density Calculation ...................... 1  
      1.2.2 Velocity of the Evolving Void Surface ................. 2  
   1.3 Simulation Results and Discussion ............................. 2  

2 Optimization and Inverse Modeling for TCAD Applications  
   2.1 Introduction ......................................................... 4  
   2.2 SIESTA ................................................................. 4  
      2.2.1 Inverse Modeling .............................................. 4  
      2.2.2 Optimization Methods ....................................... 4  
   2.3 Applications ........................................................ 5  
   2.4 Application Examples ............................................. 6  
      2.4.1 Topography Optimization .................................... 6  
      2.4.2 Parameter Extraction ....................................... 6  
   2.5 Conclusion .......................................................... 7  

3 3D Topography Simulation for Deposition and Etching Processes  
   3.1 Introduction ........................................................ 8  
   3.2 Initialization ....................................................... 8  
   3.3 Extending the Speed Function and Narrow Banding Using a Fast Marching Method ....... 9  
   3.4 Stability and the Courant-Friedrichs-Levy (CFL) Condition .......... 9  
   3.5 Transport Models .................................................. 9  
      3.5.1 Particle Distribution for Deposition and Etching ........ 10  
      3.5.2 Visibility Determination ................................... 10  
      3.5.3 Iterative Solver .............................................. 11  
   3.6 Simulation Results ................................................. 11  
   3.7 Conclusion .......................................................... 11
4 Improving SiC lateral DMOSFET Reliability under High Field Stress 13
  4.1 Introduction ......................................................... 13
  4.2 Device Structure and Operation ................................. 13
  4.3 Physical Models .................................................. 14
  4.4 Device Simulation ................................................ 15
  4.5 Result and Discussion ........................................... 16
  4.6 Conclusion ......................................................... 18

5 Calibration for the Simulation of Ion Implantation in Relaxed Sige 19
  5.1 Introduction ......................................................... 19
  5.2 The Simulator .................................................... 20
  5.3 Monte Carlo Implantation in SiGe Alloys ......................... 20
    5.3.1 Principle of the Monte Carlo Method ...................... 20
    5.3.2 Modelling of the SiGe Crystal .......................... 21
    5.3.3 Nuclear Stopping ........................................ 21
    5.3.4 Electronic Stopping .................................... 22
  5.4 Calibration ....................................................... 22
    5.4.1 Arsenic Implantation in SiGe ............................ 22
    5.4.2 Boron Implantation in SiGe ............................ 24
  5.5 Two-Dimensional MOSFET Application .......................... 24
  5.6 Conclusion ....................................................... 25
1 The Resistance and Current Density During Electromigration

We investigate the behavior of the resistance of a three-dimensional via during the evolution of an intrinsic void. The presented method is eligible for both copper and aluminum interconnects. An analysis of the mean current density over the void surface is also carried out.

1.1 Introduction

Electromigration is the main reliability issue in IC designs, which can trigger a system failure at some undefined future time [1]. The phenomenon is particularly likely to afflict thin, tightly spaced interconnect lines of deep-submicron designs. Electromigration is an atomic transport process which results from momentum transfer to the constituent metal atoms due to collisions with the current conducted electrons. As atoms electromigrate, there is a depletion of material “upstream” and an accumulation “downstream” at sites of flux divergence. This can lead to the formation and growth of voids at points of material depletion, causing a large increase in electrical resistance. On the other hand, accumulation of material may cause dielectric cracking and the formation of an extrusion, resulting in a short circuit between adjacent lines. The development of intrinsic voids, which leads to interconnect failure goes through two distinctive phases. These phases exhibit not only different influence on the operating ability of the interconnect but are also based on different physics. The first phase is the void nucleating phase. In this phase no electromigration generated voids are present and there is no significant resistance change. The second phase begins when a void is nucleated and visible in SEM pictures [2]. This is the rapid phase of the failure development. The void expands from its initial position (nucleation site) to a size which can significantly change the resistance or completely sever the connection. 

1.2 Modelling Approach

An initial void with some small radius \( r_0 \) is placed on some characteristic position inside the interconnect (Fig. 1). Since most of the fatal voids are nucleated in the vicinity or in the area of interconnect vias we consider in particular these cases. The configururable initial void volume is \( V_0 \) which is smaller than \( 4\pi r_0^3/3 \) because the void area is confined by sphere and boundary of the interconnect (Fig. 1). Starting from the initial void radius \( r_0 \), the void radius is gradually incremented \( r_0, r_1 = r_0 + \Delta r_1, r_2 = r_1 + \Delta r_2, ... \), with \( \Delta r_1 \geq \Delta r_2 \geq ... \geq \Delta r_n \). For each void radius the electrical field in the interconnect structure is calculated by means of the finite element method using a diffuse interface approach. To obtain the distribution of the electrical potential inside the interconnects the Poisson equation has to be solved

\[ \nabla \cdot (\sigma_{Cu}(\phi) \nabla \phi) = 0. \]  

(1)

To imply correct boundary behavior of the electrical field on the spherical void surface a diffuse interface approach has been applied [3]. In this approach the electrical field in the metal and the void is calculated on the same mesh. The electrical conductivity depends on the scalar field \( \phi(x, y, z) \)

\[ \rho(\phi) = \frac{\sigma_{Cu}[1 + \phi(x, y, z, t)]}{2}. \]  

(2)

In order to obtain sufficient accuracy the scalar field \( \phi(x, y, z, t) \) must be resolved on a locally refined mesh (Fig. 2). For an electrical field calculated in such a way, the resistance of the interconnect via is also calculated [4, 5]. With growing void size the resistance increases. The whole process is stopped when a void radius is reached for which \( 100 \times (R_{\text{actual}}/R_{\text{initial}} - 1) > 20\% \).

1.2.1 Average Current Density Calculation

The primary driving force of material transport at the void surface is electromigration proportional to the tangent component of the vector current density. Since the diffuse interface approach for the calculation of the current density ensures physical behavior of the electrical field in the vicinity of the isolating void, the normal component of the current density on the void surface is always zero.
and we can apply the formula
\[
J_{m,i} = 2 \int_V \frac{\|J\| [1 - \phi_i^2(x, y, z)]}{\|1 - \phi_i^2(x, y, z)\|} dV,
\]
for the average current density over a void with radius \( r_i \). (3) expresses the averaging of the current density weighted with finite element volume inside the interconnect. Since \( \phi_i(x, y, z) = 1 \) in metal and \( \phi_i(x, y, z) = -1 \) in the void area, the term \( 1 - \phi_i^2(x, y, z) \) is non-zero only in void-metal interface area.

### 1.2.2 Velocity of the Evolving Void Surface

The evolution of the void is caused by material transport on the void surface and in the vicinity of the void surface. The mass conservation law gives the mean propagation velocity \( v_i \) of the evolving void-metal interface
\[
v_i = \frac{D_v}{kT} e^{Z^*} J_{m,i},
\]
here \( D_v \) is the vacancy diffusivity and \( Z^* \) effective charge number of vacancies. The (4) is valid for all void shapes.

### 1.3 Simulation Results and Discussion

As we can see from Fig. 4, the average current density on the void surface increases with the void size. Both, current density and resistance, exhibit a very similar dynamic behavior. The dynamical resistance increase is in accordance with the measurement results presented in [1]. Compared with the earlier result [6], which assumes cubical void shapes, our approach enables more realistic simulations.

An open question is how to use the obtained average current density (Fig. 4) for the estimation of the void growing time \( t_E \) up to the critical void size. In [6] a simple formula is applied
\[
t_E = \frac{V_c - V_0}{v_m A_s},
\]
(5)
In this equation \( V_c \) is the critical void size, \( V_0 \) is the initial void size, \( A_s \) is the cross section of the interconnect in the vicinity of the growing void, and \( v_m \) is the mean velocity of the evolving void-metal interface. However, this formula is only valid in the case of the cubical void which is a very rough approximation of the real situation. According to the newer experimental results [7] the real void shape is significantly better approximated by a spherical sector. In this case \( t_E \) can be estimated as
\[
t_E = \sum \frac{\Delta r_i + 1}{v_i},
\]
(6)
assuming that for sufficiently small \( \Delta r_i + 1 \), the void radius grows from \( r_i \) to \( r_i + 1 \) with a constant velocity \( v_i \).

As we can see from (4), the velocity \( v_i \) depends on vacancy diffusivity \( D_v \) which itself has
\[ D_v = D_b + D_{gb} \left( \frac{\delta_{gb}}{d} \right) + D_{Cu/b} q_{Cu/b} + D_{Cu/N} q_{Cu/N}. \]  

Figure 3: Typical current density distribution picture in the vicinity of the spherical void. The grey area marks peak values of the current density.

Figure 4: Change of the average current density and via resistance depending on the void radius.

significantly varying values depending on the diffusion path. The electromigration assisted self-diffusion of copper is a complex process which includes simultaneous diffusion through the crystal bulk, along grain boundaries, along the copper/barrier interfaces, and along the copper/cap-layer interface. Therefore, the diffusivity applied in (4) must be a cumulative value as used in [2]

\[ D_{gb}, D_{gb}, D_{Cu/b}, \text{ and } D_{Cu/N}, \] represent the diffusivity through the bulk, along the grain boundaries, copper/barrier interfaces, and copper/caplayer interfaces, respectively. \( \delta_{gb} \) is the width of the grain boundary and \( d \) the average length of a grain boundary. Coefficients \( q_{Cu/b} \) and \( q_{Cu/N} \) depend only on the layout geometry. For the feasible estimation of \( t_E \), reliable, experimentally determined values for all relevant diffusivities are needed and this is until now not the case [2, 7].
2 Optimization and Inverse Modeling for TCAD Applications

We present the capabilities and some applications for the framework SIESTA (Simulation Environment for Semiconductor Technology Analysis). This framework supports a wide range of simulators, optimizers, and strategies to optimize different properties such as speed, geometry, power, or reliability of electronic devices. We present examples for optimizing the topography of a memory cell and for extracting material parameters of a polycrystalline silicon fuse.

2.1 Introduction

With the ongoing shrinking of device structures parasitic effects have to be considered in addition to the main properties of the structure. In state-of-the-art process technology nodes in the submicron regime, effects of grains boundaries, variations of geometry, or fluctuations of process parameters can cause serious problems. For instance, a device under mechanical or thermal stress due to operational conditions may change its behavior significantly.

For these reasons devices have to be designed very carefully and tested in simulation environments. The simulation framework SIESTA was developed in order to optimize device parameters and to investigate the sensitivity of the output characteristics on the input parameters.

2.2 SIESTA

State-of-the-art simulation environments like the framework SIESTA [8, 9, 10, 11] support a wide range of simulators, optimizers, and optimization strategies. Contrary to commercially available software such as [12] and [13] our framework provides an open architecture for numerous types of simulators and optimizers which can be individually chosen for a particular problem. To achieve this goal this simulation environment offers modular and flexible interfaces by which external tools can be integrated with minor changes as outlined in [14].

2.2.1 Inverse Modeling

Fig. 5 shows an abstracted view of the internal data flow of SIESTA. The optimization procedure consists of a loop that terminates when the result has reached the required accuracy which can be determined by, e.g., the derivative of the score function. Another possible termination criterion is that a local optimum has been detected which cannot be improved any more with a gradient based optimizer [15]. For genetic or evolutionary approaches [16] the loop is terminated if the maximum number of genomes has been reached.

At start-up SIESTA generates the first parameter set based on the initial values within the user-defined constraints for each simulation branch. After applying post processing tools the simulation result is parsed in order to compare it with reference data. A score value is determined which indicates how well these two data sets match. The optimizers use this score value to generate the parameter set for the next simulation run in order to improve the score value that will be evaluated after each simulation run based on the currently generated parameter set. A typical example for the mentioned reference data are measurements with additional requirements to meet specific physical constraints, a global minimization, or a maximization condition.

2.2.2 Optimization Methods

SIESTA supports different optimization modes in order to achieve appropriate capabilities for a specific problem, such as the optimization for a CVD (chemical vapour deposition) process as shown in Section 2.4. In this example some input parameters depend implicitly on other input parameters, and, e.g., the constraints may change which can be seen for a special radiosity model [17]. Generally, an optimum with constraints for representing the
physics is hard to obtain, because some of the conditions may change with a new optimization state.

Another big challenge are optimizations with restrictions and boundaries for the output characteristics. Additionally, most of the optimization algorithms used in SIESTA do not support such restrictions. Therefore, only a limited number of optimizers can be used for this specific task. Thus, the user has to carefully design the specification of the problem in order to obtain a smooth behavior, otherwise the optimization could become instable and an interaction with the user is necessary.

The available optimization modes for SIESTA are called optimization, calibration, design of experiment (DOE), and genetic. Each of them can be set to minimization or maximization. In all these modes the score function which indicates the quality of the simulation result will be minimized or maximized, respectively.

The default mode is optimization which optimizes targets of single values like leakage currents of transistors and side effects like parasitic capacitances of different structures and the on-resistance of transistors. Furthermore, the power consumption, maximum temperature of a device, the maximum value of the electric field can be minimized. Additionally, also combinations of different device parameters like the gain of amplifier devices, different sorts of yield, or the ratio of capacitance per area for memory cells can be optimized.

The simulation mode DOE [18] is used to investigate the system behavior with only a small number of simulations. The main goal is to get output characteristic with as few simulations as possible with reasonable accuracy. Therefore, the layout (operational mode) [14] of DOE can be specified to choose the selected subspace best for the input parameters.

The calibration mode can be used for inverse modeling. SIESTA optimizes the parameters in order to fit specific requirements which can consist of output characteristics or constraint equations for some optimizers. Section 2.4 shows examples how process as well as material parameters can be extracted. These techniques can be used to search for parameter sets which best meet the requirements. This enables to predict the behavior also for new devices and structures before fabrication. Thus, valuable time can be saved and the costs of feasibility studies are reduced compared to investigations on hardware.

The genetic mode can be used for optimization problems even if no information about the correlation between input parameter and target value is known, for instance when the target function shows many local minima. In such a case gradient based algorithms cannot be used and a genetic algorithm is required.

2.3 Applications

The open software architecture of SIESTA enables to add simulators and optimizers with only minor changes to the configuration. Thus, many software components can be easily added and combined with each other. However, the different file formats of the simulators have to be considered. Different simulations tools of vendors make it necessary to use pre- and post-processing tools to transform the currently available input format to a format readable for the next simulator. In the current stage SIESTA does not automatically check the consistency of the input parameters for the different simulators.

Interfaces to the following simulation tools exist: The three-dimensional device simulator MINIMOS-NT [19] solves the non-linear semiconductor equations together with equations which give corrections according to special effects and different materials. Furthermore, the interconnect simulator STAP from the Smart Analysis Package (SAP) [20, 21] investigates coupled electro-thermal interconnect problems including the extraction of resistances, capacitances, and inductances by means of stationary
and transient simulations. The three-dimensional Finite Element Diffusion and Oxidation Simulator (FEDOS) [22] solves complex problems on diffusion and oxidation. In additional to these simulation topics also topography simulations can be included with the three-dimensional topography simulator Enhanced Level Set Applications (ELSA) [17] for the simulation of etching and deposition processes. Moreover, simulation tools from vendors are supported as well, e.g., the device simulator DESSIS [12] from ISE and the process simulator TSUPREM [23] from Synopsys. The procedure of adding new programs to the open architecture of the simulation environment SIESTA is described in detail in [14].

2.4 Application Examples

We will give a brief overview of the capabilities of SIESTA by presenting a topography optimization for a TEOS CVD process and the calibration and identification of material parameters for a polycrystalline silicon fusing structure.

2.4.1 Topography Optimization

The ongoing shrinking of memory devices results mainly in miniaturization of capacitances in the memory cell. The smaller the capacitor, the more devices per area can be integrated. Hence, it is very important to compare the different process parameters to see which settings meet the requirements and can operate within the proposed tolerance band.

Fig. 6 shows an optimization result for CVD process parameters using the etching and deposition simulator ELSA. For optimization we investigated the sticking coefficients in order to obtain a good agreement with the measurements. This result allows to apply the model to future device structures. Starting from this point one can use other simulators to follow the fabrication line in order to investigate or optimize a complete series of steps for a complete device structure. Additionally, we can analyze for instance the sensitivity of the output geometry on the sticking parameters for a TEOS CVD process. Fig. 7 compares the original simulation result with the results of slightly changed input parameters. Within the range of the two solid lines we can verify that the results with the tolerance band meet the given requirements according to the internal fabrication specifications.

2.4.2 Parameter Extraction

With shrinking of critical dimensions small non-volatile memory cells based on fuses become a very interesting alternative in terms of production costs, area saving, and efficiency.
Therefore, the geometrical and thermal design becomes very important. For instance in fusing structures the electro-thermal transient behavior determines the shortening of the fuse. The faster the structure heats up the shorter is the fusing time. Therefore, a fully three-dimensional electro-thermal investigation is necessary in order to predict the material reaction. As seen in Fig. 8, the resistance shows a significant increase with time. After a certain time the resistance falls rapidly due to thermal run-away until the fuse shortens.

To improve the fusing procedure we investigate the temperature distribution during such a programming cycle in order to obtain improved designs. For these electro-thermal analyses we use the transient mode of the three-dimensional interconnect simulator STAP from the SAP package. The input parameters are measurements of the current through the device with an an applied voltage ramp. The simulator shows the resistance of the complete structure as well as the internal temperature distribution at particular time steps.

After extracting the material parameters we have varied one of the first order thermal coefficients in order to show the significance of proper calculation. The difference between the original simulation and the changed one are shown in Fig. 8. If the temperature is not calculated correctly, thermal run-away starts at a different point, which can cause serious problems. Fig. 9 shows the temperature distribution of the fusing structure at the point of highest resistance at approximately 65 µs, which can be used to analyze the heat flux.

2.5 Conclusion

We have shown a wide range of different applications for our simulation environment SIESTA. Different purposes of the simulation environment, for instance optimizations, calibrations, and DOE (design of experiments) have been used for basic investigations on device and process analysis. These optimization procedures can be performed with already existing systems and devices to verify the developed models and the currently used optimization setup. New devices, models, and systems can be automatically evaluated, optimized, calibrated, and investigated according to the specified requirements. With future simulation software we can optimize a complete process and investigate sensitivities on various process and device parameters.

![Figure 8: Small variation of a first order temperature coefficient](image1)

![Figure 9: Temperature distribution in [K] at the point with the highest resistance](image2)
3D Topography Simulation for Deposition and Etching Processes

We present the application of level set and fast marching methods to the simulation of surface topography of a wafer in three dimensions for deposition and etching processes. These simulations rest on many techniques, including a narrow band level set method, fast marching for the Eikonal equation, extension of the speed function, transport models, visibility determination, and an iterative equation solver.

3.1 Introduction

To understand the influence of edge topography on device characteristics, which is important for highly integrated ICs, an accurate three-dimensional topography simulator is required. However, topography simulation is still faced with many challenges which limit its general applicability and usefulness. In addition, three-dimensional topography simulation tends to be very CPU and memory intensive to date.

Various surface representation algorithms have been used for the development of three-dimensional topography simulators [24]. Roughly speaking, these algorithms fall into three categories.

- Segment-based models, such as the facet motion model [25, 26]: In this model a nodal triangularization of the interface is used. The position of the nodes is then updated by determining front information about the normals and curvature of surface facets. Because interstices or duplications between neighboring surface facets occur during their advance along the normal, area-readjustment procedures are needed. However, these procedures induce significant computational error into the simulation result in proportion to the complexity of the process geometry.

- Cell-based models, such as the cellular model [27, 28]: These models can easily handle topological changes and can be extended to three dimensions, whereas the determination of geometric quantities such as surface normals and curvature can be inaccurate.

- Level set method-based models [29, 30, 31]: In this method the interface extraction is based on the solution of a hyperbolic partial differential equation. The location of an interface is the zero level set of a higher dimensional function called level set function. This model provides an interesting alternative method for solving the above mentioned problems.

Based on an efficient and precise level set method including narrow banding and extending the speed function in a sophisticated algorithm, we have developed a general three-dimensional topography simulator for the simulation of deposition and etching processes. The simulator is called ELSA (Enhanced Level Set Applications) and works efficiently concerning computational time and memory consumption, and it simultaneously ensures high resolution.

The outline of this paper is as follows. First, we present an optimized method to obtain the initial level set function. Second, we describe briefly how to extend the speed function combined with narrow banding using a fast marching method. Third, the stability and the complexity of the simulator is discussed. Fourth, we present the transport models. Finally, simulation results are shown.

3.2 Initialization

The basic idea of the level set method is to view the curve or surface in question at a certain time as the zero level set (with respect to the space variables) of a certain function called level set function. Each point on the surface is moved with a certain speed normal to the surface and this determines the time evolution of the surface. For points on the zero level set the speed function is usually determined by physical models and in our case by
the etching and deposition processes, or more precisely by the fluxes of certain gas species and subsequent surface reactions. The speed function at grid points has to be extended from the known values of the speed function on the surface. We will discuss this more in detail in the next section.

In order to apply the level set method a suitable initial function has to be determined. A good choice is the signed distance function of a point from the given surface. This function is the common distance function multiplied by minus or plus one depending on which side of the surface a point lies. Since we later apply the level set algorithm only in narrow bands, it is sufficient to calculate the signed distance function only in this narrow band. This method reduces the computational effort of initialization from $O(n^3)$ to $O(n^2)$, where $n$ is the grid resolution in each direction.

### 3.4 Stability and the Courant-Friedrichs-Levy (CFL) Condition

For advancing the level set function we have used a second order space convex finite difference scheme [32]. Consider $\Delta x$, $\Delta y$, $\Delta z$, and $\Delta t$ as discretization steps in space and in time, respectively. A necessary condition for the stability of this scheme is the Courant-Friedrichs-Levy (CFL) condition which requires that

$$\Delta t \cdot F_{\text{max}} \leq \min(\Delta x, \Delta y, \Delta z).$$

The CFL condition guarantees that the front can cross no more than one grid cell during each time step. In order to have a stable simulator based on the finite difference method, the CFL condition must definitely be satisfied. However, there is a problem stemming from the CFL condition, which limits the simulator performance. If we increase the spatial resolution by $\lambda$, assuming that $F_{\text{max}}$ remains constant, we have to reduce the maximum $\Delta t$ by the same factor $\lambda$, which increases the number of simulation steps by $\lambda$ for reaching the same thickness. Furthermore, an increase in spatial resolution by $\lambda$ increases approximately the number of extracted surface elements by $\lambda^2$ and then the computational effort of the visibility determination by $\lambda^3$. In summary, an increase in spatial resolution by $\lambda$ leads to an increase in simulation time by a factor $\lambda^5$, if one uses the most precise visibility determination.

### 3.5 Transport Models

The transport of the particles above the wafer surface specifies the deposition and etch rate. Assume that within a feature the frequency of particle-particle collisions is negligible relative to particle-surface collisions, that is, we are in the molecular or Knudsen regime [33]. In this case the transport of the particles can be simulated using the radiosity model. In the other case the collision of single particles plays a major role and

---

**3 3D Topography Simulation for Deposition and Etching Processes**
their concentration is determined by the diffusion equation.

### 3.5.1 Particle Distribution for Deposition and Etching

For modeling deposition it is assumed that the distribution of the particles coming from the source obeys a cosine function around the normal vector of the plane in which the source lies [33, 34]. This implies that the flux at a surface element is proportional to the cosine of the angle between the connecting line between the center of mass of a surface element and the source and the normal vector of the source plane.

A function which has been used for ions in plasma systems for etching processes is the normal distribution 

$$f(\theta) = (2\pi\sigma)^{-1/2} \cdot \exp(-\theta^2/2\sigma^2)$$

where $\theta$ is the angle around the normal vector of the source plane and the angular width of the distribution is specified by $\sigma$. For the reflections of particles diffuse and specular reflection are assumed for deposition and etching processes, respectively [33].

### 3.5.2 Visibility Determination

Most of the computation time for simulating the transport of the particles above the wafer by the radiosity model is consumed in determining the visibility between the surface elements which is an $O(m^2)$ operation, where $m$ is the number of surface elements growing approximately like $O(n^2)$. If the connecting line between the center of mass of two surface elements does not intersect the surface, i.e., the zero level set, those surface elements are visible from each other. In order to decrease the computational effort related to determining the visibility between the surface triangles, we have assumed that two triangles are visible from each other if the center point of the grid cells in which the triangles are located, are visible from each other. Since there are at least two triangles in each grid cell, considerable time is saved. The radiosity model assumes that the total flux depends on the flux directly from the source, as well as an additional flux due to the particles which do not stick and are re-emitted. After discretizing the problem the flux vector whose elements are the total flux at different surface elements can be expressed by a matrix equation. There are two numerical
approaches for solving this problem. The first one is to use a direct solver for the matrix equation. Although this is practical in two dimensions [29], it becomes impractical due to the computational effort needed by calculating the inverse matrix for three-dimensional problems. In three dimensions we solve the equation iteratively.

3.5.3 Iterative Solver

The iterative solution [31] consists of a series expansion in the interaction matrix. Suitably interpreted, it can be viewed as a multi-bounce model, in which the number of terms in the series expansion corresponds to the number of bounces that a particle can undergo before its effects are negligible. This approach allows to check the error remainder term to determine how many terms must be kept. Since most of the particles either stick or leave the domain after a reasonable number of bounces, this is an efficient approach. By constructing the remainder term, we can measure the convergence of the expansion and keep enough terms to bound the error below a user-specified tolerance.

3.6 Simulation Results

In this section we present some simulation results for deposition and etching processes. We begin with a source deposition into a rectangular trench shown in Fig. 10. Fig. 11 shows the simulation result of a source deposition from a plane located above the trench leading to void formation including the visibility and shading effects. The particle distribution is a cosine distribution around the normal vector of source plane.

Fig. 12 shows a straightforward simulation of isotropic etching of the same trench from which material is being isotropically etched. As expected, the sides of the trench are cleanly etched away and are rounded.

Finally Fig. 13 shows directional etching of the same trench. The incoming flux at the surface element is a cosine function of the angle between the surface normal and the normal vector of the source plane without visibility effects. The reflection effects are also ignored which is approximately the case for directional etching of ions by a plasma etching process. The trench has been etched less than the selective isotropic etching at the sides and tends to be etched more in vertical direction.

Table 1 shows a comparison between the simulation times of these different simulation processes for different grid resolutions. The most time consuming simulation is the deposition simulation because all of the time expensive steps, e.g., visibility determination, extending the speed function, and the iterative solver are required. By directional etching, extending the speed function is the only time consuming part of the simulator. Therefore, the simulation time is smaller than for the deposition process. For isotropic etching neither visibility determination, the iterative solver, nor extension the speed function contribute. Thus the simulation time is very small compared to the other simulations. In the third column of Table 1 the simulation times for a grid resolution two times that of the second column are presented. As an example, for deposition the time has been increased by about a factor of 32.

3.7 Conclusion

State of the art algorithms for surface evolution processes like deposition and etching processes in three dimensions have been implemented. A
Table 1: Simulation times for different simulations presented in this paper

<table>
<thead>
<tr>
<th>Grid resolution</th>
<th>30·30·30</th>
<th>60·60·60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposition time/step</td>
<td>1.54s</td>
<td>47.4s</td>
</tr>
<tr>
<td>Isotropic etching time/step</td>
<td>0.53s</td>
<td>2.77s</td>
</tr>
<tr>
<td>Directional etching time/step</td>
<td>0.97s</td>
<td>21.1s</td>
</tr>
</tbody>
</table>

general simulator called ELSA was developed based on the level set method combining the narrow banding and fast marching method for extending the speed function. The speed of simulation was improved in several steps, e.g., in initialization, visibility determination, and solving the radiosity matrix. A comparison between the simulation time of different simulation processes was presented that shows how time consuming the different parts of the simulator are. Furthermore, the effect of increasing the grid resolution on the simulation time was shown.
4 Improving SiC lateral DMOSFET Reliability under High Field Stress

We propose a new device structure for a SiC lateral DMOSFET, which improves the reliability of oxides under high field stress. A numerical simulation in order to get an insight into the physics and the characteristic of the device has been carried out. The key parameters that alter the device performance and reliability have been optimized using the device simulator MINIMOS-NT. The relationship between blocking and driving capability of our structure was closely examined. The peak surface electric field has been kept below 1.5 MV/cm at a breakdown voltage of 1460 V. Excellent transfer characteristics with significantly reduced gate bias voltage, and a fairly large advantage in terms of electrical performance and device reliability have been achieved.

4.1 Introduction

The physical and electronic properties of silicon carbide (SiC) make it the foremost semiconductor material for high temperature, radiation resistant, and high-power/high-frequency electronic devices [35]. SiC-based electronic devices can operate at extremely high temperatures without suffering from intrinsic conduction effects (10-30 orders lower than Si) because of the wide energy bandgap of 3-3.3 eV and high thermal conductivity of 4.9 W/cm-K [36]. Devices formed in SiC can withstand an electric field of 2.5-3.0 MV/cm (8-10 times higher than Si) without undergoing avalanche breakdown [37]. SiC devices can operate at high frequencies (RF and microwave) because of the larger saturated electron drift velocity of 2-3 times compared to Si [38].

SiC is the only semiconductor material besides silicon on which a thermal oxide can be grown, thus enabling MOS devices. In the particular area of power devices, theoretical appraisals have indicated that SiC power MOSFETs would operate over higher voltage and temperature ranges, have superior switching characteristics, and yet have die sizes nearly 20 times smaller than correspondingly rated Si-based devices [39]. SiC vertical DMOSFETs have been demonstrated with a specific on-resistance almost ten order of magnitude lower compared to the theoretical lower limit of Si MOSFETs of similar breakdown voltage [40].

SiC lateral DMOSFETs are attractive for monolithic integration with low voltage logic components in form of power IC; however, their design is more challenging due to the presence of a high surface field in SiC. The lack of material development and design is a major cause to most SiC lateral DMOSFETs surface problems such as step-bunching and non-uniform doping density, which leads to poor inversion layer electron mobility and oxide reliability [41].

The off-state operation is hampered by the possibility of gate oxide breakdown before avalanching occurs in SiC. According to Gauss’ law, the field in the oxide is approximately 2.5 times larger than the peak field in the SiC bulk. At this field the mean-time-before-failure (MTBF) of the gate oxide can be significantly reduced. This raises concerns about the reliability of oxides on SiC under high field stress.

In order to minimize these problems a new design of an accumulation-mode structure for a LD-MOSFET (Lateral DMOSFET) is proposed. The key parameters that alter the device performance and reliability have been optimized using the device simulator MINIMOS-NT [19]. A numerical simulation in order to get an insight into the physics and the characteristics of the device has been carried out. The relationship between blocking and driving capability was closely examined. The peak surface electric field has been kept below 1.5 MV/cm at a breakdown voltage of 1460 V. Excellent transfer characteristics with significant improvement on the reduction of the gate bias voltage, and a fairly large advantage on electrical performance and device reliability have been achieved.

4.2 Device Structure and Operation

The principal difference between our structure depicted in Fig. 14 and the conventional inversion-
4 Improving SiC lateral DMOSFET Reliability under High Field Stress

(layer structure is the presence of a thin n-channel region (accumulation-layer) below the gate oxide using a buried p-well region formed by ion-implantation. The thickness, length, and n-doping of this accumulation-layer is carefully chosen so that it is completely depleted by the built-in potential of the p/n junction. This causes a potential barrier between the n+ source and the n-drift regions, resulting in a normally-off device with the entire drain voltage supported by the n-drift region. Thus it can block high forward voltages at zero gate bias with low leakage currents. When a positive gate bias is applied, an accumulation channel of electrons at the SiO$_2$-SiC interface is created and hence a low resistance path for the electron current flow from the source to the drain can be achieved. This structure utilizes the buried p-well region as a shield to the influence of a high SiC bulk electric field on the gate oxide, in consequence of that improving the reliability of oxides under high field stress. The structure also offers the possibility of moving the channel away from the oxide interface, thereby removing the effect of the poor interface quality on the channel mobility.

4.3 Physical Models

The choice of appropriate physical models is fundamental for any comparative study that involves numerical simulation. Among the SiC polytypes commercially available, for this work 6H-SiC is preferred owing to its higher breakdown field strength [42]. A model that takes into account the mobility degradation due to surface scattering has been incorporated with [43] and implemented in our simulator MINIMOS-NT:

\[
\mu_{n,p}^{low} = \mu_{n,p}^{min} + \mu_{n,p}^{T_0} \cdot \left( \frac{T}{T_0} \right)^{\frac{\delta_{n,p}}{\beta_{sat n,p}}} \cdot \frac{1 - M(y)}{1 + M(y) \cdot \left( \frac{N_D + N_A}{N_{n,p}} \right) \frac{\gamma_{n,p}}{\mu_{n,p}}}
\]

where the function \( M(y) \) depends on the surface distance \( y \), and the parameter \( y_{ref} \) describes a critical length.

At high electric field the drift velocity \( v_{n,p} \) of the carriers saturates due to increasing optical phonon scattering and finally reaches the saturation velocity \( v_{n,p}^{sat} \), leading to the field dependent mobility as described by [44]

\[
\mu_{n,p}^{low} = \frac{1}{\left(1 + \left( \mu_{n,p}^{low} \frac{v_{n,p}^{sat}}{v_0} \right) \frac{\gamma_{n,p}^{sat}}{\mu_{n,p}} \right)}.
\]

We take the component of the electric field parallel to the electron motion as driving force. The temperature dependence of \( v_{n,p}^{sat} \) has been modeled by

\[
v_{n,p}^{sat} = v_0^{sat} \cdot \left( \frac{T}{T_0} \right)^{\frac{\gamma_{n,p}^{sat}}{\beta_{n,p}^{sat}}}.
\]

and

\[
\beta_{n,p}^{sat} = \beta_0^{sat} \cdot \left( \frac{T}{T_0} \right)^{1 - \frac{\gamma_{n,p}^{sat}}{\beta_{n,p}^{sat}}}
\]

Several temperature-dependent Hall measurements have been reported, regarding the carrier concentration, and hence, the donor (\( E_D \)) and acceptor (\( E_A \)) energy levels in SiC. A function to describe ionized shallow donor and acceptor substitutional impurities is given by [45]

\[
N_D^+ = \frac{N_D}{1 + g_D \frac{n}{N_C} \exp \left( \frac{E_D}{kT} \right)}
\]

\[
N_A^- = \frac{N_A}{1 + g_A \frac{p}{N_V} \exp \left( \frac{E_A}{kT} \right)}
\]

Figure 14: Schematics of the proposed SiC accumulation-mode LDMOSFET.
Because of these deep levels, the dopants are not fully ionized even at higher temperatures so that we obtain an explicit relation for the ionization degree of a single donor level in n-type material

$$\xi_D = \frac{N_D^+}{N_D} = \frac{-1 + \sqrt{1 + 4g_D N_D n_{ie}} \exp\left(\frac{E_{ip}}{kT}\right)}{2g_D N_D n_{ie} \exp\left(\frac{E_{ip}}{kT}\right)} \tag{15}$$

and similarly in p-type material

$$\xi_A = \frac{N_A^-}{N_A} = \frac{-1 + \sqrt{1 + 4g_A N_A p_{ie}} \exp\left(\frac{E_{ip}}{kT}\right)}{2g_A N_A p_{ie} \exp\left(\frac{E_{ip}}{kT}\right)} \tag{16}$$

One of the most important parameters of a SiC device is its breakdown voltage. In order to obtain a clear understanding of its breakdown characteristics, it is important to have an exact knowledge of the impact ionization coefficients for SiC, which are modeled according to [46], where the dependence of the impact ionization rate on the electric field and temperature is given by

$$\alpha_{n,p} = \alpha_{n,p} \gamma_a \exp\left(-\frac{b_{n,p} \gamma_a}{E_{||}}\right) \tag{17}$$

$$\gamma_a = \frac{\tanh\left(\frac{\hbar \omega_{op}}{2kT_0}\right)}{\tanh\left(\frac{\hbar \omega_{op}}{2kT}\right)} \tag{18}$$

In these expressions $\alpha_n$ and $\alpha_p$ are the impact ionization coefficients for electrons and holes, respectively. The factor $\gamma_a$ as a function of the optical phonon energy $\hbar \omega_{op}$ expresses the temperature dependence of the phonon gas against which the carriers are accelerated.

Models which account for generation and recombination have been employed. The Shockley-Read-Hall recombination is given by

$$GR_{SRH} = \frac{n_{ie}^2 - np}{\tau_p (n + n_{ie}) + \tau_n (p + n_{ie})} \tag{19}$$

where the life time $\tau_{n,p}$ can depend on a doping level as experimentally observed in Si technology [47] and is empirically modeled by the so-called Scharfetter relation:

$$\tau_{n,p} = \frac{\tau_{n0,p0}}{1 + \left(\frac{N_D + N_A}{N_{SRH} n_{p}}\right)^{\frac{1}{\gamma_{SRH}}}} \tag{20}$$

Additionally, the Auger recombination rate is given by [48]

$$R_{Au} = \left(C_n n + C_p p\right) \left(np - n_{ie}^2\right) \tag{21}$$

Here, $C_n$ and $C_p$ denotes the Auger coefficient of electrons and holes, respectively.

### 4.4 Device Simulation

For device simulation we have utilized published material data listed in Table 2. Six parameters that alter the device performance and reliability have been investigated: The doping concentration of the n-drift region; the depth and the concentration of the implanted p-well; the doping concentration and the thickness of the n-channel (accumulation-layer); and the gate oxide overlap.
length. The p-well region has a Gaussian profile buried between 0.3 and 1.0 μm, which has to be optimized because it determines the thickness of the accumulation-layer region which in turn affects the gate oxide field, breakdown voltage, and on-resistance.

In order to achieve acceptable device reliability, the maximum field in the oxide may need to be limited. If this occurs, the high-field capability of SiC cannot be fully utilized. For the desired breakdown voltage of 1500 V, the proposed structure is optimized to have a 33 μm cell pitch, a 10 μm thick n-drift region doped at $5.0 \times 10^{15} \text{ cm}^{-3}$ and an n+ polysilicon gate electrode with a 50 nm thick gate oxide.

When the buried p-well depth is larger, the built-in potential is unable to fully deplete the n-channel which causes high leakage currents, resulting in the degradation of the performance and reliability of the device at high temperature. Therefore, its depth and implanted peak concentration of 0.5 μm (between 0.3 – 0.8 μm) and $1.0 \times 10^{18} \text{ cm}^{-3}$ respectively was found to give the optimum accumulation layer thickness at which the criterion for the device optimization (figure of merit, FOM) [49] can be satisfied.

$$\left. \frac{V_B^2}{R_{on,sp}} \right|_{opt} = \mu_n \epsilon_s \left( \frac{2E_c}{3} \right)^3$$  \hspace{1cm} (22)

where $V_B$ is the breakdown voltage, $R_{on,sp}$ is the specific on resistance, $E_c$ is the critical electric field, $\mu_n$ is electron mobility parallel to the c axis and $\epsilon_s$ is the SiC dielectric constant.

High breakdown voltages provide for improvements in high power device performance, but this conflicts with the need for high switching devices to have a low on-resistance. The breakdown voltage is limited by the breakdown of the gate oxide which depends on the electric field in the oxide. From Gauss' law, the field in the oxide is approximately 2.5 times greater than the peak field in the SiC. Since the peak field in SiC can be almost 10 times higher than in Si, the fields in the oxide on SiC will tend to be 10 times higher than the oxide fields in Si devices.

For durable reliability simulation results show that the accumulation layer parameters are key factors determined the peak electric field of the gate oxide remain below the practical limit of 4 MV/cm while utilizing the high-field capability of SiC. Values of the accumulation layer thickness, length and concentration of 0.3 μm, 4 μm and $5.0 \times 10^{15} \text{ cm}^{-3}$, respectively, have been established to achieve the desired on- and off-state characteristics. At these optimum values and room temperature, a specific on-resistance of 93.2 mΩ·cm$^2$ and a breakdown voltage of 1460 V with the corresponding small leakage current was achieved. The effect of the accumulation layer thickness on the maximum operating voltage, specific on-resistance and criterion for the device optimization obtained by simulation is illustrated in Fig. 15.

The gate oxide overlap was varied from 4 to 7 μm, and its influence on the surface field and operating voltage was analyzed. For the desired high stress voltage operation this overlap has to be kept as small as possible in order to minimize its parasitic capacitance. Simulation predicted that a gate oxide overlap length of 6 μm is optimal.

4.5 Result and Discussion

The proposed accumulation-mode LDMOSFET shows a fairly large advantage in terms of electrical performance and reliability compared to its standard inversion-mode LDMOSFET counter...
Improving SiC lateral DMOSFET Reliability under High Field Stress

Figure 16: Comparison of on-state characteristics.

Excellent I-V characteristics were obtained with good current saturation and gate control at high temperature operation of 470 K as depicted in Fig. 16.

One of the important areas of improvement for the SiC LDMOSFET device is the decrease in its conduction losses which is governed by its specific on-resistance. This on-resistance depends on the channel and the n-drift resistance of the device. An estimate of the on-resistance contribution indicates that 90% of the on-resistance is due to the large channel resistance, owing to the low inversion layer mobility. The proposed structure is able to minimize this resistance and improve the mobility. A simulated accumulation layer mobility of 120 cm$^2$/Vs compared to the 18 cm$^2$/Vs for the inversion-layer was observed, which is in good agreement with experimental results extracted at a different temperature [42]. Fig. 17 shows a significant improvement on the reduction of the gate bias voltage (a logic level gate bias of 5 V). The device is normally off with a threshold voltage of only 1 V compared to that of 3 V for the inversion-mode structure.

In addition to moving the channel away from the oxide interface and removing the influence of interface quality on the channel mobility, the proposed structure offers the possibility of serving as a shield to the influence of high SiC bulk electric field on the gate oxide. The influence of the accumulation-layer thickness on the SiC bulk electric field is illustrated in Fig. 18. The result clearly show that the proposed structure accumulation-mode n-channel thickness of 0.3 µm improves the electric field by 0.3 MV/cm compared to its inversion-mode counter part. At this optimum value and maximum operating voltage the peak surface electric field has been kept below 1.5 MV/cm as depicted in Fig. 19. That is equivalent to an oxide field of 3.75 MV/cm, and considerably lower than the practical limit of the electric field strength in the oxide. Therefore, the proposed structure improves the reliability of the device while utilizing the high breakdown electric field strength of SiC.

A breakdown voltage of 1460 V with a leakage current comparable to that of standard inversion-mode LDMOSFET was achieved as shown in Fig. 20. The off-state leakage current caused by the built-in potential of the p/n junction is ten orders of magnitude less than the on-state current for the same structure, but one order of magnitude larger than the inversion-mode structure. This can effectively be suppressed by calibrating parameters which enable a fully depleted accumulation-layer. High temperature causes an increase in the leakage current due to the increased intrinsic carrier concentration.

Figure 17: Comparison of transfer characteristics.
4 Improving SiC lateral DMOSFET Reliability under High Field Stress

Figure 18: Effect of the accumulation-layer thickness on the electric field in the SiC substrate at 470 K.

Figure 19: Profile of the electric field at the maximum operating voltage.

Figure 20: Comparison of off-state characteristics.

4.6 Conclusion

A new device structure for a SiC lateral DMOSFET that improves the reliability of oxides under high field stress has been proposed. Key parameters that alter the device performance and reliability were optimized and analyzed. Simulation based comparisons were conducted between the proposed device and the standard structure using the same condition and parameters. It may be concluded that the new structure offers a fairly large advantage in terms of electrical performance. A satisfactory improvement on the reduction of the gate oxide peak electric field to sustain durable reliability while utilizing the high-field capability of SiC has been achieved. A significant reduction on the gate bias voltage to obtain good on-state conduction and excellent transfer characteristics were obtained with the output current increased by two fold. The device also exhibits a blocking voltage of 1460 V with an oxide field below 3.75 MV/cm which is considerably lower than the practical limit of the oxide breakdown field.
5 Calibration for the Simulation of Ion Implantation in Relaxed SiGe

SiGe-based CMOS devices have significant performance enhancements compared to pure silicon devices. We have extended our Monte Carlo ion implantation simulator for Si$_{1-x}$Ge$_x$ targets in order to study the formation of shallow junctions. SiGe has a larger nuclear and electronic stopping power for ion implanted dopants compared to pure silicon due to the heavier and electron-rich germanium. It turned out that the Lindhard correction parameter of the electronic stopping model can be adjusted by a linear function of the germanium content to adopt the strength of the electronic stopping. The successful calibration for the simulation of arsenic and boron implantations in Si$_{1-x}$Ge$_x$ is demonstrated by comparing the predicted doping profiles with SIMS measurements. Thereby the non-linear shift towards shallower profiles with increasing germanium fraction is analyzed. Finally, the simulation result of source/drain implants for a MOS-transistor structure on a SiGe substrate is presented.

5.1 Introduction

While the first transistor was developed in 1947 by using germanium as the semiconductor material and GaAs devices have demonstrated high switching speed, it is silicon which completely dominates the present semiconductor market. This development has arisen due to the low cost of silicon CMOS technology. This mainstream technology offers the feasibility to produce billions of transistors on a single wafer, all with nearly identical properties. The fabrication processes and the device performance rely heavily on a number of natural properties of silicon, for instance, the availability of a good oxide. For alternative semiconductor materials much more expensive fabrication processes must be used, whereby the phenomenal yields achievable on a silicon CMOS line cannot be reached. The increase in packing density and performance of CMOS has been achieved by downscaling transistors and circuits over the years. One drawback of silicon is its relatively small carrier mobility. Since the device speed depends on how fast the carriers can be transported through the device under sustainable low operating voltages, silicon can be regarded as a relatively slow semiconductor. One of the most promising alternatives for the replacement of bulk silicon substrates in CMOS technology are silicon-germanium (SiGe) alloys.

SiGe alloys offer the possibility of bandgap engineering, enhanced carrier mobility, and a higher dopant solubility compared to pure silicon. The remarkable potential of the SiGe material technology arises from the possibility to modify its properties by altering the composition. For instance, the band gap decreases from 1.12 eV (pure silicon) to 0.66 eV (pure germanium) at room temperature. The band structure can also be tailored by strain. By building different kinds of Si–SiGe heterostructures various properties for device design can be optimized. An excellent example is the heterojunction bipolar transistor (HBT) which enables higher switching-speed performance compared to the conventional silicon based transistor. The HBT is now applied in high frequency applications competing with III–V technologies. Of great importance for the semiconductor industry is strained silicon CMOS technology based on relaxed SiGe, and associated with it, the heterojunction field effect transistor (HFET). Theoretical considerations predict that for a similar gate-length to CMOS technology, heterostructure CMOS technology has twice the speed and a factor of 4 to 6 lower power-delay product [50]. At present, the major challenge for the SiGe technology is the defect density of available virtual substrates or bulk SiGe substrates, which is still too large to achieve economic yields. With the amount of capital and knowledge invested in silicon based devices, the pressure is enormous to continue producing silicon based devices. The CMOS compatible SiGe material system which enables higher speed performance is a way to further use existing knowledge and manufacturing infrastructure.

One of the key processes in the fabrication of state-of-the-art CMOS devices is ion implantation. Ion
implantation is the primary technology to introduce doping atoms into semiconductors to form devices and integrated circuits. The reason for the application of this technology is mainly the high accuracy in adjusting the doping concentration and the uniformity of the implantation across large wafers. A subsequent thermal annealing step often only repairs the induced crystal defects while it barely redistributes the dopant atoms. Therefore the distribution of the dopants in the final device is mainly determined by the ion implantation step, whereby the channeling effect caused by the anisotropy of the crystal plays a major role. Moreover effects resulting from non planar surfaces, can significantly influence the device behavior. The ion implantation process can effectively be simulated on computers. The capability of accurately predicting doping profiles can significantly reduce integrated process development and implementation time. In particular, the ongoing trend of scaling device feature sizes down into the sub-100nm regime puts high demands on the accuracy of simulation results.

Analytical ion implantation simulation tools which are often used due to their simplicity cannot accurately predict doping profiles for complex targets, for instance, multilayer targets or advanced devices with junction depths in the range of few nanometers. For a compound target like SiGe, the range predictions will be still worse, because the doping profiles additionally depend on the germanium fraction in a non-linear manner. The alternative are physics-based Monte Carlo methods which use an atomistic approach and, therefore, are able to simulate the channeling effect or the implantation induced point defects in crystalline targets as well. The accuracy of the simulation is mainly determined by the complexity of the models that describe the physical behavior. These models are applicable for a wide range of implantation conditions without the need for an additional calibration. One drawback of the Monte Carlo method are fairly long computing times, which is the main reason why the use of Monte Carlo simulation tools as standard ion implantation tools is usually avoided in technology optimization. However, the formation of ultra-shallow junctions by ion implantation technology is a prerequisite for the construction of sub-100nm transistors. Therefore exact knowledge of the as-implanted doping profile and of the ion implantation induced crystal damages is required in order to facilitate SiGe-based CMOS technology.

5.2 The Simulator

All Monte Carlo simulation experiments were performed with the object-oriented, multi-dimensional ion implantation simulator MCIMPL–II [51], [52]. The simulator is based on a binary collision approximation (BCA) and can handle arbitrary three-dimensional device structures consisting of amorphous and crystalline materials. In order to optimize the performance, the simulator uses cells arranged on an ortho-grid to count the number of implanted ions and of generated point defects. The final concentration values are smoothed and translated from the internal ortho-grid to an unstructured grid suitable for subsequent process simulation steps like finite element simulations for annealing processes.

5.3 Monte Carlo Implantation in SiGe Alloys

5.3.1 Principle of the Monte Carlo Method

The Monte Carlo method is based on imitating the random behavior of ions at an atomistic level. Particularly the position where an ion hits the crystalline target is calculated using appropriately scaled random numbers. Furthermore, the lattice atoms of the target are in permanent movement due to thermal vibrations. Thus, the actual positions of the vibrating atoms in the target are also simulated with random deviations. The ion implantation process is accurately simulated by computing a large number N of individual ion trajectories through a semiconductor material. The trajectory of each implanted ion is determined by the interactions with the atoms and electrons of the target material. The incoming doping atoms are slowed down due to the nuclear and electronic stopping power of the target material.
The final position of an implanted ion is reached where it has lost its kinetic energy. The Monte Carlo simulator uses an atomistic crystal model which enables to simulate the channeling effect of ions in crystalline targets. Additionally, the Kinchin-Pease model is used to calculate the vacancies and interstitials which are generated by an ion [53]. Being based on random numbers, the results obtained with the Monte Carlo method are never exact, but they converge to the used model characteristics by increasing the number \( N \) of simulated ions. The statistical error vanishes for \( N \to \infty \). The reduction of the statistical fluctuation of doping profiles is performed through a sophisticated smoothing algorithm based on the Bernstein polynomials (25). The main advantage of the Monte Carlo method is that it is a physically based method and therefore it is easily extendable to new target materials with the need for only calibrating the electronic stopping model for each dopant species.

5.3.2 Modelling of the SiGe Crystal

Silicon and germanium, which both crystallize in the diamond lattice structure, are completely miscible forming \( \text{Si}_{1-x}\text{Ge}_x \) solids with \( x \) ranging from 0 to 1. For \( \text{Si}_{1-x}\text{Ge}_x \) crystals the lattice parameter \( a(x) \) depends on the germanium fraction \( x \) and can be calculated according to the quadratic expression (23) with sufficient accuracy [54]. Vegard’s law determines the SiGe lattice parameter only by a linear interpolation of the parameters of the end-point elements Si and Ge. Whereas the relation takes the known small departure from Vegard’s law into account and approximates (23) the experimental data with a maximum deviation of about \( 10^{-3} \) Å.

\[
a(x) = 0.02733 \times^2 + 0.1992 \times + 5.431 \quad (\text{Å})
\]

(23)

While the ion moves through the target, a local crystal model is built up around the actual ion position for searching the next collision partner (Fig. 21). The selection of the target atom species in the crystal model is defined by probability \( x \) for germanium and \( 1 - x \) for silicon, respectively. This random choice of the atom species is acceptable because no ordering has been observed in bulk SiGe alloy crystals and ordering mechanisms in epitaxial grown layers are still under investigation.

5.3.3 Nuclear Stopping

The interaction of the moving ion with an atomic nucleus of the target (nuclear stopping) can be treated as an elastic collision process, whereas the interaction with the electrons can be treated as an inelastic process without any scattering effects (electronic stopping). The binary collision approximation assumes that only two particles, the ion (atomic number \( Z_1 \), mass \( M_1 \), energy \( E \)) and one target atom (atomic number \( Z_2 \), mass \( M_2 \)) are involved in one scattering process. While the moving particle passes and is deflected, the stationary particle recoils or at least activates thermal lattice vibrations. The final velocities and trajectories can be simply found from the conservation of energy and momentum of the system. For solving this two-body problem it is convenient to transform the scattering process from the laboratory coordinates to the center-of-mass coordinate (CM) system in which a single particle with transformed energy \( E_\text{c} \) moves in a stationary potential \( V(r) \). The scattering angle \( \Theta \) in the CM system is determined by (24) and depends on the energy.

![Figure 21: Si\(_{1-x}\)Ge\(_x\) crystal simulation model](image-url)
E_c, the interatomic potential V(r), and the impact parameter p [55]. In (24), r_0 is the distance of minimum approach between the particles and it is determined by the real root of the denominator.

$$\Theta(p, E_c) = \pi - 2p \int_{r_0}^{\infty} \frac{dr}{r^2 \sqrt{1 - \frac{V(r)}{E_c} - \frac{p^2}{r^2}}}$$  

(24)

The inverse transformation leads to equation (25) which determines the scattering angle $\vartheta$ of the ion in the laboratory system.

$$\tan \vartheta = \frac{\sin \Theta}{M_1 + \cos \Theta}$$  

(25)

From (25) it can be derived that if the ion is heavier than the target atom ($M_1 > M_2$) then a maximal scattering angle $\vartheta_{\text{max}} < 90^\circ$ exists according to (26).

$$\sin \vartheta_{\text{max}} = \frac{M_2}{M_1}$$  

(26)

An interesting conclusion can be drawn from (26) for the Monte Carlo implantation in SiGe target materials. For example, if an arsenic ion hits a silicon atom ($M_1/M_2 = 2.68$) then $\vartheta_{\text{max}} = 22^\circ$, and if the arsenic ion hits the heavier germanium atom ($M_1/M_2 = 1.07$) then a larger maximal scattering angle $\vartheta_{\text{max}} = 69^\circ$ is possible. Due to the fact that the angles of subsequent collisions have to be added up for a turn around from the incident direction, the backscattering probability for the dopant atoms increases with the germanium content in SiGe. Fig. 22 demonstrates the shift to shallower profiles by comparing SIMS measurements of 60 keV arsenic implantations into SiGe targets with a difference in the germanium fraction of 15%. This useful property of SiGe can be exploited to reduce the vertical junction depth needed to further scale down the MOS-transistor structure in the deep sub-100nm range.

5.3.4 Electronic Stopping

The total stopping process of the ions in the target solid is modeled as a sequence of alternating nuclear and electronic stopping processes. The electronic stopping process is calculated by using the Hobler model which extends the Lindhard electronic stopping model (amorphous model) to crystalline silicon [56]. The only physical parameter required for this model is the impact parameter which is determined when selecting a collision partner. Due to the fact that the model implies a dependence on the charge and the mass of the atoms of the target material the electronic stopping power is averaged in the case of a compound material like SiGe. SiGe has a larger electronic stopping power than silicon, which is caused by the higher electron density of SiGe due to the electron-rich germanium atom [57]. In addition to the Lindhard correction parameter $k$ which adopts the strength of the electronic stopping, three other empirical parameters are necessary for each dopant species in crystalline silicon.

5.4 Calibration

5.4.1 Arsenic Implantation in SiGe

We are studying the implantation of arsenic as an n-type and boron as a p-type dopant in crystalline SiGe targets with different composition. Therefore, the Monte Carlo ion implantation simulator has been extended from silicon to $Si_{1-x}Ge_x$ by calibrating the empirical electronic stopping model used to accurately simulate the electronic stopping process in crystalline silicon. For this calibration it turned out to be most advantageous to arrange the Lindhard correction...
parameter $k$ as a linear function of the germanium fraction $x$ and let the other three parameters of the model unchanged. The equation (27) determines the parameter $k_{\text{As}}(x)$ for arsenic and it could be verified from pure silicon up to a germanium content of 50% by comparison with SIMS measurements (Fig. 23).

$$k_{\text{As}}(x) = 1.132 + 1.736x$$

Fig. 23 shows the simulated and experimental doping profiles of arsenic implantations into $\text{Si}_{1-x}\text{Ge}_x$ layers with a thickness of 150 nm on a silicon substrate. All implantations were simulated with an energy of 60 keV, a dose of $10^{11}\text{ cm}^{-2}$, a tilt of $7^\circ$, and a twist of $15^\circ$. The figure demonstrates the effect of the germanium fraction in $\text{Si}_{1-x}\text{Ge}_x$ targets on profiles from arsenic implants. Two effects can be observed in this figure. Firstly, with increasing germanium fraction there is a shift towards shallower arsenic profiles. Secondly, the germanium content produces a stronger decline of the arsenic concentration with increasing penetration depth compared to silicon. It has been pointed out by the interpretation of (26) that the impact of nuclear collision is significantly changed if the incoming ion hits the germanium atom which is heavier than the silicon atom. This causes an increased backscattering probability for the dopant atoms. The larger electronic stopping power of $\text{Si}_{1-x}\text{Ge}_x$ compared to pure silicon increases with the germanium fraction $x$ and causes a stronger decline of the concentration profiles in SiGe.

Fig. 24 demonstrates the successful calibration of the simulator which is valid for other energies too. It compares arsenic implants in a $\text{Si}_{0.65}\text{Ge}_{0.35}$ target, compared to SIMS measurements.

Fig. 25 presents simulated arsenic profiles performed with a lower energy and a higher dose. It again demonstrates the effect of the germanium content which facilitates the forming of shallow junctions but the trend to shallower profiles is considerably non-linear. The difference between $x = 0$ and $x = 0.25$ profiles is larger than the difference between $x = 0.5$ and $x = 0.75$ profiles, for instance. All implantations were performed
with an energy of 15 keV, a dose of $10^{15}$ cm$^{-2}$, a tilt of 7°, and a twist of 22°.

### 5.4.2 Boron Implantation in SiGe

For the calibration of boron implantations in Si$_{1-x}$Ge$_x$ a linearly rising function for the parameter $k_B(x)$ depending on $x$ according to (27) was used, whereas for the other three parameters of the model the values from crystalline silicon could be applied.

$$k_B(x) = 1.75 + 0.75 x$$  \hspace{1cm} (28)

Fig. 26 shows the simulated and experimental doping profiles of boron implantations into a Si$_{1-x}$Ge$_x$ layer with a thickness of almost 330 nm on a silicon substrate. All implantations were simulated with an energy of 50 keV, a dose of $10^{15}$ cm$^{-2}$, and a tilt of 7°. Additionally, a native oxide on the wafer surface with a layer thickness of 1 nm was taken into account for the simulation of the implantation of boron dopants.

**Figure 26**: Simulated 50 keV boron implantations in Si$_{1-x}$Ge$_x$ with $x = 0$, 10%, 20% compared to SIMS measurements

Fig. 26 points out that boron implants in Si$_{1-x}$Ge$_x$ show qualitatively the same characteristics as arsenic implants. Additionally, a larger effect of the germanium fraction for shifting the profiles towards the surface can be observed. Fig. 27 compares simulated boron profiles in targets with different germanium content. All simulations were performed with an energy of 5 keV, a dose of $10^{15}$ cm$^{-2}$, and a tilt of 7°. The effect of the germanium fraction on the low-energy boron profiles is extremely non-linear.

**Figure 27**: Simulated 5 keV boron profiles in Si$_{1-x}$Ge$_x$ with $x = 0$, 20%, 40%, 60%

### 5.5 Two-Dimensional MOSFET Application

Si$_{1-x}$Ge$_x$ alloys can be applied to construct strained silicon CMOS devices. One possible MOSFET structure is the surface channel HFET in which in-plane electron mobilities approaching 3000 cm$^2$/Vs have been reported [58]. The surface channel device has a single layer of thin strained silicon (typically 10 nm), grown on top of a thick, relaxed SiGe buffer layer. The biaxial tensile strain in the strained silicon layer can be tailored by the germanium content of the relaxed SiGe layer. This structure can be used for n- or p-MOSFETs depending on the implanted dopant type in the layers. The excellent properties of Si$_{1-x}$Ge$_x$ alloys for forming shallow vertical junctions are demonstrated with a two-dimensional MOSFET example application. We have simulated arsenic source/drain and extension implants for a 100 nm n-MOSFET structure on a Si$_{0.75}$Ge$_{0.25}$ substrate. Using scaling considerations, a source/drain vertical junction depth of 40 nm to 80 nm is recommended for processing of a 100 nm gate MOS transistor. Fig. 28 shows the Monte Carlo arsenic source/drain and extension implants for such a transistor. The simulation was performed with 2,000,000 simulated ions per
each implantation step. In the first implantation step the source/drain extensions were formed with an energy of 15 keV, a dose of $4 \cdot 10^{13}$ cm$^{-2}$, a tilt of 7°, and a twist of 22°. The source/drain implantation step was performed with an energy of 45 keV and a dose of $2 \cdot 10^{15}$ cm$^{-2}$. Although a relatively large energy of 45 keV was used, the required junction depth was met.

![Figure 28](image_url)

**Figure 28:** Simulated cross-section of a 100 nm gate n-MOSFET structure on a relaxed Si$_{0.75}$Ge$_{0.25}$ substrate for SiGe-based CMOS technology.

### 5.6 Conclusion

Relaxed SiGe layers strongly facilitate the forming of shallow vertical junctions which are a prerequisite to further scaling down MOSFET structures into the deep sub-100nm regime. The penetration depth for ion implanted dopant atoms in Si$_{1-x}$Ge$_x$ is reduced with the increase of the germanium content $x$ at a given implantation energy. This effect arises due to the larger nuclear and electronic stopping power of the germanium atom compared to the silicon atom of the target alloy. The heavier germanium atom leads to a significantly higher backscattering probability which has been derived from the scattering integral. This integral is evaluated by the Monte Carlo simulator to determine the scattering angle of the nuclear collision process. On the other hand, the larger electronic stopping power of germanium facilitates a stronger decline of the dopant concentration profiles. The calibration of the empirical electronic stopping model for the simulator is based on a linear relation between the Lindhard correction parameter $k$ and the germanium fraction $x$ for each dopant species. This assumption has been validated for arsenic and boron implantations into targets with different germanium fractions. An accurate agreement of the simulated doping profiles with the SIMS measurement data was found in all cases. Although a simple linear relation was used to include the effect of germanium on the electronic stopping power, the resulting doping profiles vary with increasing germanium fraction extremely non-linear. The inherent consideration of all involved atom species (dopant and target atoms) in the BCA approximation calculation and the existing accurate calibration for crystalline silicon has facilitated considerably the extension of the simulator MCIMPL-II to relaxed SiGe.
References


