AUTOMATION IN PROCESS- AND DEVICE-SIMULATORS

W.Jüngling, P.Pichler, S.Selberherr, H.Pötzl

Institut für Allgemeine Elektrotechnik und Elektronik
Technical University of Vienna
Gußhausstr.27-29, 1040-Vienna, AUSTRIA

Abstract - Up-to-date process and device simulation requires sophisticated physical models and a secure numerical environment to reveal accurate simulations and to keep pace with the increasing demands of new process techniques. We have developed one- and two-dimensional coupled process and device simulators which cover these demands. The high level of automation in the programs frees the user from many laborious tasks like the linearization and discretization of the differential equations and the installation of the spatial and transient grids. The wide range of physical models which can be implemented and the fully adaptive spatial and transient grids which enable accurate simulations are explained by process and device simulations.

Introduction

The development of models will remain the task of engineers and scientists. However, the efficient evaluation of process and device models can be supported by software tools which offer an optimal numerical environment. The numerical support includes the simple and userfriendly implementation of physical models as well as the automatic detection and resolution of critical processes and domains in space and time. The automation of most of the numerical computations, e.g. linearization and discretization of the partial differential equations and the installation of the spatial and transient grid frees the user from laborious and repeatitive works.

With one-dimensional simulations a good choice of the grid for appropriate resolution of critical domains determines the accuracy of the simulation. In many cases a similar accuracy can be obtained only by an exhaustively memory and time consuming simulation with a rigid grid. In two or even three dimensions the optimal choice of the spatial and transient grids determines whether the simulation can be performed at all. Our strategies for fully adaptive spatial and transient grids are therefore of interest for the development of two- or three-dimensional process or device simulators and the development of new physical models.

We present results obtained by the one- and two-dimensional coupled process- and device-simulators 'ZOMBIE' and 'PROMIS'. As a simple one-dimensional example we simulate the fabrication of a n^+-p-p^+ diode and its transient electric characteristic. The process and device simulation of a MOS transitor is a typical two-dimensional example.

The numerical environment which enables the evaluation of the critical process and device simulations is demonstrated and explained by the examples.

Process Simulation

The process parameters are chosen in such a way that critical simulation steps occur. The starting material is a silicon wafer with a thickness of 50 microns which is also the simulation domain of all succeeding simulations. Arsenic is implanted with a dose of $2 \cdot 10^{16} cm^{-2}$ and an energy of 140keV. The wafer is annealed at 1000°C for 20min. The ramping from 800°C to 1000°C and back within 5min is not included in the 20min annealing. The p-contact fabrication at the low doped p-domain is simulated by a boron predeposition with a boundary value of $10^{18} cm^{-3}$ for 20min at 1200°C. A ramping down to 800°C within 10min and a final annealing at 800°C for 55min is performed to avoid crystal defects.

The annealing of the high dose As-implantation requires the use of a dynamic arsenic cluster model since the solubility limit of the electrically active arsenic is exceeded for all process temperatures. During the annealing the arsenic splits up into active arsenic and clustered arsenic, which is electrically inactive and is assumed not to diffuse. The simulation will show that the assumption of equilibrium between the two arsenic phases is not permitted during the fabrication steps.

In our simulation we use the cluster

require therefore a fine resolution during all processes. A similar problem occurs during the simulation of the p-contact fabrication. In the very beginning a fine grid is necessary close to the surface, but it turns out to be of no importance afterwards.

A fully adaptive spatial grid in space and time is necessary to fulfil all specified requirements. A rigid equidistant mesh can by no means fulfil all requirements at all times without consuming an enourmous amount of memory. A rigid problem oriented grid requires the knowledge of all critical domains in advance, i.e. the solution must be known before the simulation.

Fig.4A-B-C show the grid modifications during the annealing steps in domains of interest. A continous line represent an existing gridpoint, a beginning or terminating line indicates a modifiaction in the grid. The vertical bars at the top of the figures indicate the transient grid used for the simulation. Fig.4A summarizes the modification of the grid caused by the 1000°C annealing step and the beginning of the 1200°C step. The time steps are small in the very beginning when active arsenic is converted into clustered arsenic and during the high temperature annealing step. Fig.4B shows the grid modifications during the second high temperature annealing step. The spatial grid moves with the spreading arsenic profile and provides a fine resolution of the steep arsenic gradient and the p-n junction during the whole simulation. The transient grid resolves the fast spreading of the arsenic accurately whereas the step width increases significantly as the final 800°C annealing begins. Fig.4C shows the grid modifications in the vicinity of the upper boundary during the predeposition step.

The two-dimensional MOS transistor fabrication step shows the capabilities of our two-dimensional codes. The source and drain regions have been obtained by an implantation of arsenic with doses of $2 \cdot 10^{15} cm^{-2}$ and an energy of 150keV. Fig.5A shows the doping profile after the implantation and the corresponding spatial grid. The grid used for the simulations is identical to the grid used for the figures. The gridlines are accumulated close to the gradients of the arsenic. Two annealing steps are performed after the implantation. The first one for 60min at 1000°C, the second one for 15min at 950°C. Fig.5D shows the final arsenic profile and the corresponding grid. The comparison of Fig.5A and Fig.5D indicates the modifications of the spatial grid which are similar to the one-dimensional simulations. After the first annealing the chanal implantations are performed with doses of $5 \cdot 10^{12} cm^{-2}$ and $2 \cdot 10^{13} cm^{-2}$ and energies of 25keV and 100keV. Fig.5B shows the boron distribution after the doping, Fig.5C the boron distribution after the 950°C annealing. Hereagain, the comparison of the grid before and after the annealing indicates the modifications.

Device Simulations

The transient device characteristic of the n^+p-p^+ diode and the MOS transistor have been simulated using the same programs, i.e. the same solvers and the same strategies for the transient and spatial grid. We have just changed the physical models, i.e. the variables changed from active arsenic, clustered arsenic and boron to potential, electrons and holes.

Fig.6A-B-C show the switching of the diode from -4.5V reverse bias to +0.5V forward bias and the corresponding grid modifications in the vicinity of the space charge layer. In the beginning Fig.6C shows two domains of fine spatial resolution which coincide with the boundaries of the space charge layer and the strong decrease of the electron and hole concentrations. As the reverse voltage decreases the space charge layer gets smaller, especially in the low doped p-domain. At 0.9μsec the external voltage gets zero and the space charge layer disappears. Now electrons are injected from the n^+-domain into the p-domain and a fine resolution close to the contacts gets important. The spatial grid as well as the transient grid resolve the critical domains carefully and provide an accurate simulation.

Fig.7A-B show the electron density and the electric field for a drain voltage of 3V and a gate voltage of 0V, Fig.8A-B for a drain voltage of 3V and a gate voltage of 3V. The obtained results coincide well with the simulations of well established simulators /7/ and reveal that our grid strategies can handle semiconductor equations too.

How to Create Adaptive Grids

The strategies for the creation and modifications of grids must be independent from the physical model under consideration. The design of a spatial grid can be split up into two independent parts. Computation of the position of the maximum discretizaton error and refinement of the grid in the vicinity of this position. The computation of the discretization error depends on the discretization method, e.g. finite differences, finite boxes or finite elements. We use the method of finite differences and therefore the deviation of the distribution of the variables from a local polynomal of second order can be taken as a measure of the spatial discretization errors. The computation of

modeling of /1/. Three continuity equations Eq.(1)-(3) describe the redistribution of the dopants.

$$\frac{\partial C_{As}}{\partial t}=\mathrm{div}(D_{As}\cdot \mathrm{grad}C_{As}+C_{As}\cdot\mu_{As}\cdot\mathrm{grad}\Psi)+ k_D\cdot C_{Cl}-k_C\cdot C_{As}^{m}\cdot n^{k} \quad (1)$$

$$m\cdot\frac{\partial C_{Cl}}{\partial t}=-k_D\cdot C_{Cl}+k_C\cdot C_{As}^{m}\cdot n^{k} \quad (2)$$

with m=3 and n=1.

$$\frac{\partial C_{B}}{\partial t}=\mathrm{div}(D_{B}\cdot \mathrm{grad}C_{B}-C_{B}\cdot\mu_{B}\cdot\mathrm{grad}\Psi) \quad (3)$$

$$\Psi = U_t\cdot \mathrm{arsinh}\left(\frac{C_{As}+(m-k)\cdot C_{Cl}-C_B}{2\cdot n_i}\right) \quad (4)$$

The electric potential is computed by the quasineutral assumption, Eq.(4), which turns out to be a sufficiently accurate approximation for the solution of the exact Poisson equation /2/.

The boundary conditions during the inert annealing steps are summarized in Eq.(5) and Eq.(6).

$$J_{As}(x=0\mu m)=0,\ J_{As}(x=50\mu m)=0,$$
$$J_B\ (x=0\mu m)=0,\ J_B\ (x=50\mu m)=0 \quad (5)$$

$$J_{As}(x=0\mu m)=0,\ J_{As}(x=50\mu m)=0,$$
$$J_B\ (x=0\mu m)=0,\ C_B\ (x=50\mu m)=10^{18}cm^{-3} \quad (6)$$

Fig.1 shows the temperature as a function of time during the diode fabrication steps and summarizes the process parameters. Fig.2A-B-C-D show the distribution of the dopants after the implantation, the first annealing step and after the contact fabrication at both boundaries.

Fig.3 shows the transformation from active to clustered arsenic and back. Since a cluster consists of three arsenic atoms, the changes in the active arsenic are three times the changes in the clustered arsenic. The total amount of arsenic remains constant. Whenever the temperature increases the active dose increases too. Fig.3 shows that the arsenic dose is not constant during a process step with a constant temperature since the equilibrium is not obtained within the duration of the process step. (The effects cannot be explained by arsenic diffusion only).

During the diffusion the arsenic concentration varies by some orders of magnitude, especially close to the steep gradient of the profile. The spreading of the arsenic profile and the p-n junction

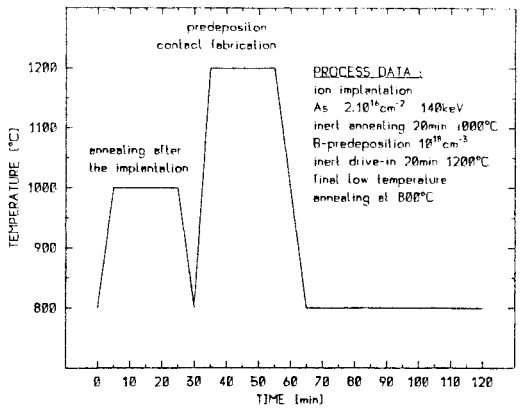


Fig.1

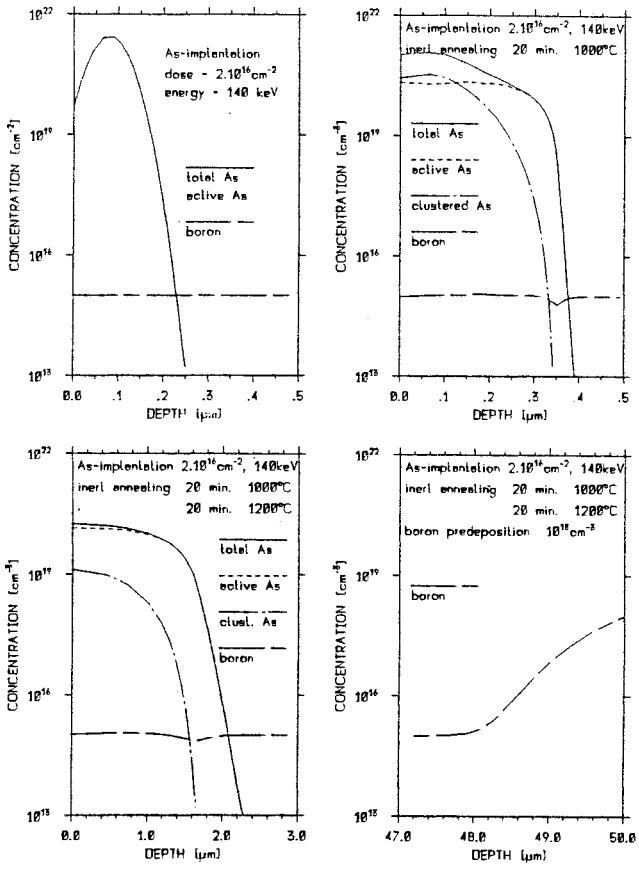


Fig. 2A-B

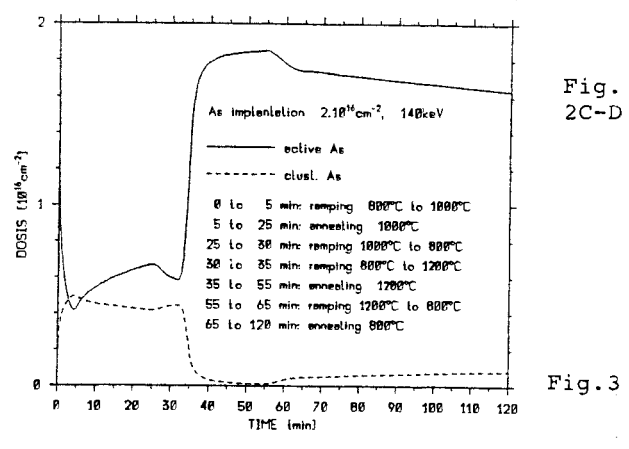


Fig. 2C-D

Fig.3

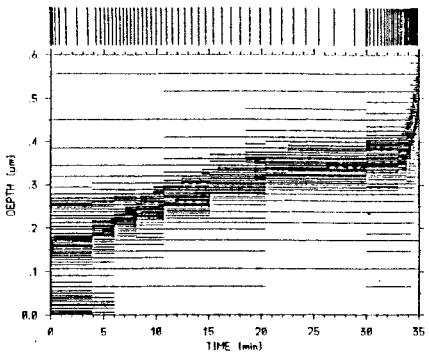


Fig.4A

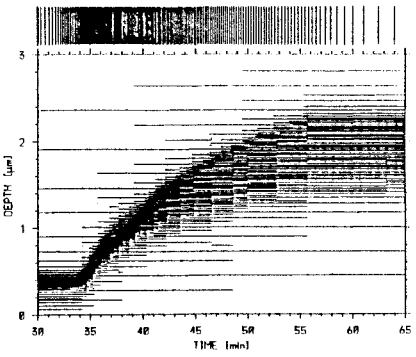


Fig.4B

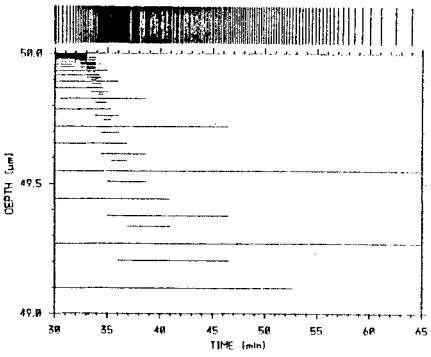


Fig.4C

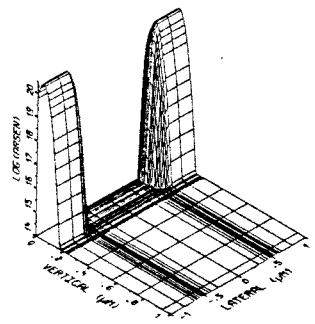


Fig.5A

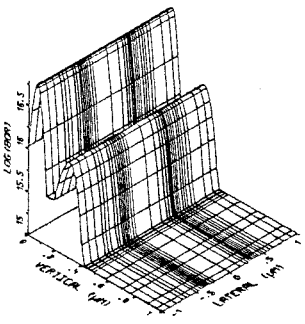


Fig.5B

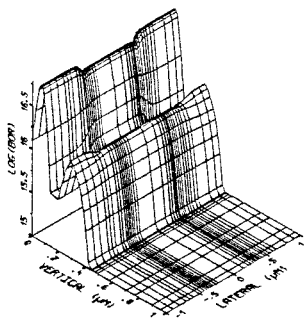


Fig.5C

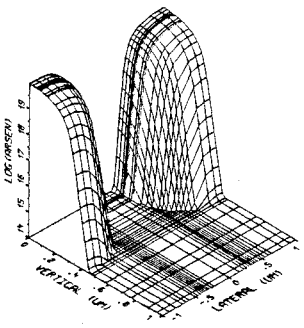


Fig.5D

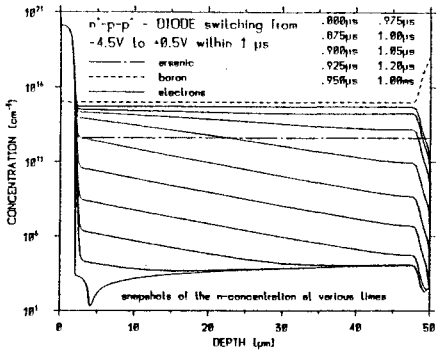


Fig.6A

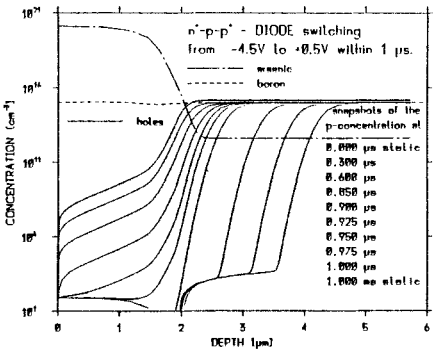


Fig.6B

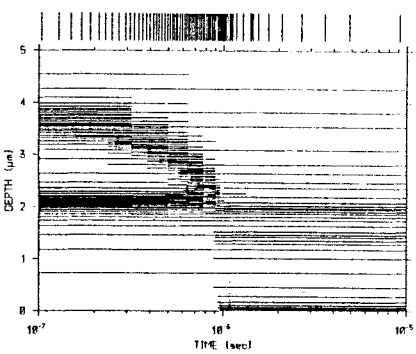


Fig.6C

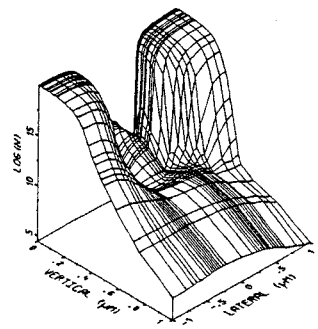


Fig.7A

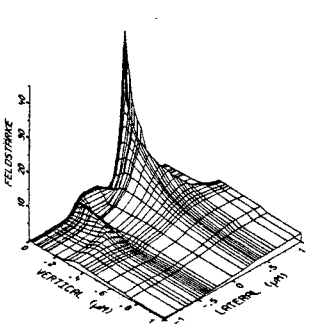


Fig.7B

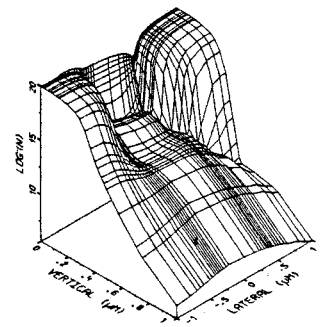


Fig.8A

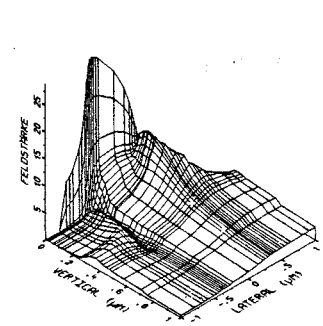


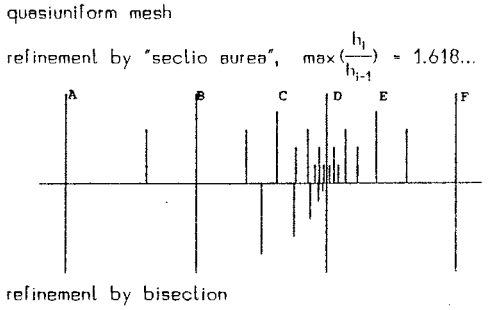
Fig.8B

the discretization error should be insensitive to slight trunaction errors and the error should decrease with decreasing mesh spacing. The use of numerical differentiation should be avoided therefore. All criteria which fulfil these demands may be used and will lead to similar spatial grids.

A quasiuniform spatial mesh can be recommended for all simulations /3/ and should be kept during all grid modifications. A mesh is called quasiuniform if the ratio between two adjacent grid spacings minus unity is small compared to unity. This mesh avoids abrupt transitions from a very coarse to a very fine mesh. Furthermore, it provides that the discretization errors reduce with the square of the mesh spacing /4/. We use the "sectio aurea" for the grid refinement, e.g. h_i/h_{i-1}=sqrt(1.25)-.5, 1 or sqrt(1.25)+.5. Fig.9 compares a quasiuniform mesh to an arbitray mesh. (The sectio aurea holds AB:BC=BC:CD and AB=BD). It should be noted that a quasiuniform mesh requires often the modification of a larger grid domain (C and E must be inserted together). The deletion of grid points is more difficult than the insertion of gridlines. Therefore, we prefer to create a completely new grid after a certain number of additional gridlines has been inserted or after a certain number of time steps has elapsed.

For the transient integration we use "backward difference formulae" of 6^{th} order /5/,/6/. This method permits the exact integration of polynomals of 6^{th} order. The comparison of a predicted with a computed solution permits the detection of critical simulation domains, the step width control and the order control. The large dynamic range of the semiconductor equations and the process variables permits a relative error control only. We check the error of every variable at every depth. The maximum error between the predicted and the computed solution determines the step width for the next iteration. The step width is computed for the actual order and the order minus and plus one. The order which permits the largest time step is chosen for the next iteration. If the error exceeds some limit the time step is rejected and the simulation repeats with a smaller time step.

Fig.9



Summary

We have presented simulations of critical process and device simulations using the one- and two-dimensional programs 'ZOMBIE' and 'PROMIS'. The simulators have been supported by fully adaptive spatial and transient grids. The adaptive spatial grid reduces the memory requirements by a factor up to 20 during process simulations and up to 5 during device simulations. The additional CPU-time used for the grid modifications is about 10% of the simulation time and by far compensated by the reduction of time during the solving of the partial differential equations with the decreased number of gridlines. The fully adaptive grid in time saves additional CPU-time.

A fully adptive grid will become important for process and device simulations in two ways. Firstly, it supports the development of new models for process and device simulation, because it frees the user from nearly all mathematical considerations about grid creation and step width control during the simulation. Secondly, it reduces the CPU-time and memory requirements of the simulations which will become important with two- or three-dimensional simulators.

References

/1/ Tsai M.Y., Morehead F.F., Baglin J.E.E., "Shallow Junctions by High-Dose As Implants in Si", J.Appl.Phys., Vol.51(6), pp.3230-3235, June 1980

/2/ Jüngling W., Pichler P., Selberherr S., Guerrero E., Pötzl H., "Simulation of Critical IC-Fabrication Processes Using Advanced Phys. and Num. Methods", IEEE-SSC-20(1), pp.156-167, Feb.1985

/3/ Pichler P., Jüngling W., Selberherr S., Guerrero E., Pötzl H., "Simulation of Critical IC-Fabrication Processes", IEEE-ED32(10), pp.1940-1953, Oct.1985

/4/ Selberherr S., "Analysis and Simulation of Semiconductor Devices", Springer Verlag, 1984, ISBN 3-211-81800-6

/5/ Brayton R.K., Gustavson F.G., Hachtel G.D., "A New Eff. Algorithm for Solving Diff.-Alg.Systems Using Implicit Backward Differentiation Formulas", Proc.IEEE, Vol.60, pp.98-108, 1972

/6/ Gear C.W., "The Automatic Integration of ordinary differential Equations", Information Processing 68, North Holland, 1968, pp.187-193

/7/ Selberherr S., Schütz A., Pötzl H., "MINIMOS - A Two-Dimensional MOS Transistor Analyzer", IEEE-ED27, No.8, August 1982, pp.1540-1549

This work has been sponsored by the SIEMENS Research Lab., Munich, W.Germany and the Fond zur Förderung der wissenschaftlichen Forschung, Project S43/10.