Simulation of Critical IC Fabrication Processes Using Advanced Physical and Numerical Methods

WERNER JÜNGLING, PETER PICHLER, SIEGFRIED SELBERHERR, SENIOR MEMBER, IEEE, EDGAR GUERRERO, AND HANS W. PÖTZL, MEMBER, IEEE

Abstract—Critical steps of IC fabrication are simulated by one- and two-dimensional computer programs using advanced physical models. Our codes deal with an arbitrary number of physical quantities such as concentrations of dopants, vacancies, interstitials and clusters, the electrostatic potential, and so on. Furthermore, they easily permit the exchange or variation of the physical models under consideration. As typical applications phenomena of coupled diffusion in one and two dimensions and dynamic arsenic clustering are investigated. The differences caused by the models of the zero space-charge approximation and the solution of the exact Poisson equation are studied by examples of As–B diffusion with various doping concentrations at different temperatures. A dynamic cluster model developed for the simulation of thermally annealed As implantations is compared to measured data of laser annealing experiments. A short outline of the mathematical and the numerical problems is given to show the amount of sophistication necessary for up-to-date process simulation.

I. Introduction

THE RECENT advances in device miniaturization have called for a better understanding of physical processes and their effect upon the dopant distribution during device and circuit fabrication, especially in VLSI. To date, a fair amount of work has been done in terms of experimentation and in the development of advanced physical models (e.g., dynamic cluster/precipitation models, enhancement of diffusion by vacancies, interstitials and the electrical field, etc.). This research requires efficient computer programs to evaluate, to compare, to verify, or to reject the various models. In addition, accurate simulations of process steps are often desirable to compute unknown parameters or to calculate more precisely roughly estimated parameters of physical models.

The complicated structure of advanced physical models often exceeds the capabilities of commonly used simulation programs. The implementation of models often requires simplifications which diminish the expected improvements by the models.

To counteract this situation we have developed a general-purpose simulation program which treats an arbitrary number of physical quantities such as concentrations of donors, acceptors, vacancies, interstitials and clusters, and the electrostatic potential, etc. In the first step, the application is mainly scientific and the code is used to study advanced models. We investigate the validity range of current models and estimate, as closely as possible, if improvements of more complicated models are justified by the additional amount of computation time and memory resources. The program has been developed to quickly vary physical models (e.g., varying the number and kind of physical quantities, modifying some of the physical parameters, etc). Since the main purpose of the program is concentrated on the physical problem, the mesh necessary for the discretization is created automatically and adapted with respect to the specific problem. This feature frees the user from mathematical and numerical problems.

In Section II we explain the capabilities of our code and discuss the differential equations which can be solved by our programs. In Section III we use our code to study some phenomena of coupled diffusion and compare the results of the quasi-neutral approximation to those using the exact Poisson equation. Section IV compares dynamic cluster models for shallow highly doped arsenic implantations.

A short outline of the numerical methods and capabilities is given in Section V.

II. Specifications of the Programs

On specifying the capabilities of our code, we had to find a compromise between a wide range of applications and an acceptable amount of computation time and memory requirement. The specifications to be described in the following are concerned with ion implantation and redistribution of dopants by diffusion.

Ion implantation is one of the most applied doping techniques in IC fabrication, especially in VLSI. The simulation of implantation must be a basic capability of any process simulation program. Advanced simulation programs tend to use distribution functions which are specified by a higher number of moments to describe the impurity concentrations, e.g., Pearson IV distributions. Since our code is laid out to have mainly scientific applications, we have implemented all commonly used distribution functions. This allows for investigations of differences caused by the assumption of different impurity distribution

Manuscript received October 18, 1984. This work was sponsored by the Siemens Research Laboratories, Munich, Germany, and the Fond zur Forderung der wissenschaftlichen Forschung, Project S 43/10.

The authors are with the Institut für Allgemeine und Elektronik, Abteilung für Physikalische Elektronik, 1040-Wein, Austria.

0018-9200/85/0200-0076$01.00

functions and for comparison with any published simulations. In the first step, we have installed the simple Gaussian, the joined half Gaussian, and the Pearson IV distributions. Ion implantation can be simulated in single layer as well as in multilayer structures. The layers may consist of arbitrary ordering of silicon, silicon nitride, and silicon dioxide. A summary of the distribution functions and their mathematical background can be found in [11].

Sometimes it is necessary to use experimental data, obtained by SIMS, $n(\alpha)$, or spreading resistance measurements as an initial solution for further process simulation. It should be noted that measured data often need a mathematical pretreatment to yield sufficiently continuous distribution functions for the simulation of a following diffusion step. An example may be a short time annealing step using CW or pulsed laser. Monte Carlo methods have been omitted at present because they lead to very time consuming computations.

This level of implementation permits the computation of nearly all impurity distributions which may be the initial solution for a following diffusion step.

Impurity redistribution during IC fabrication will be the main application of the first version of our code described in this paper. The physical model on which all considerations are based is a "general continuity equation" (1) and a "general flux relation" (2). The number n of physical quantities in (1) and (2) is arbitrary and is only restricted by computer resources. The ordering and the physical meaning of the quantities must be specified by the user through a few parameter definitions. This feature offers the possibility to optimize the program for a specific problem (e.g., omitting unnecessary quantities or modifying some physical parameters). With respect to a fixed amount of memory, the user can choose between additional physical quantities or additional grid points to obtain a higher accuracy of the simulation.

If the physical quantity is a concentration, (1) and (2) state that the concentration can be modified by a diffusion flux, a field flux, or a generation or a recombination process. The first two effects are included in Fick's law and are commonly used to describe dopant redistribution whereas the other effects become essential when time dependent transformations between physical quantities occur (e.g., vacancies—interstitials, electrically active arsenic—clustered arsenic, etc.). Certainly any of these effects can be included or excluded if necessitated by the corresponding physical model. It can be easily seen that (1) and (2) also include the Poisson equation.

The program solves a system of n partial differential equations (PDE) of the form

$$\sum_{j=1}^{n} \alpha_{ij} \cdot \frac{\partial C_j}{\partial t} = - \sum_{j=1}^{n} \beta_{ij} \cdot \operatorname{div} J_j + G_i \tag{1}$$

$$J_i = - D_i \cdot \operatorname{grad} C_i + / - \mu_i \cdot C_i \cdot \operatorname{grad} \psi \tag{2}$$

where α_{ij}, β_{ij}, D_i, μ_i, and G_i are functions of space, time and any of the quantities C_k and ψ represents the electrostatic potential. The index i runs from 1 to n and indicates the ith equation of the differential equation system. It is clear that unusual functions, which would change the type of the differential equations, must be excluded. This constraint, however, does not affect the commonly used modeling of the parameters.

This set of PDE's includes the possibility to simulate a simple diffusion step, a Poisson equation, coupled diffusion of dopants with same or different charge, dynamic clustering of dopants, the effect of precipitation and enhanced diffusion by vacancies or interstitials. The variables C_k in (1) and (2) represent not only a concentration or a potential, but can also represent the radius of precipitates or any other physical quantity.

By changing the functions α, β, D, μ, and G we can simply vary physical models and apply the main part of our code unchanged to very many problems.

The initial solution for the PDE's can be computed from the modeling of ion implantation, can be given by the background impurities of the wafer or can be computed from measured data. It is worthwhile now to discuss the boundary conditions in more detail. In the first level of our code we constrain ourselves to rigid boundaries, i.e., we exclude the simulation of oxidations with moving boundaries. The boundary conditions for each of the differential equations are given by (3).

$$\sum_{j=1}^{n} \xi_{ij} \cdot C_j + \sum_{j=1}^{n} \eta_i \psi \cdot J_j + B_i = 0 \tag{3}$$

where ξ_{ij}, η_{ij} are functions of space and time and B_i is a function of space, time, and any of the quantities C_k. It is worth noting that ξ_{ij} and η_{ij} are independent from α_{ij} and β_{ij}.

This formulation includes a Dirichlet, a Neumann, and a mixed boundary condition. From the physical point of view, these implementations offer the possibility to simulate an inert diffusion step (e.g., $J_i = 0$), a predeposition step ($C_i = C_0$), an ideal sink for vacancies or interstitials ($C_I = C_I^{eq}$), a concentration dependent source of interstitials $J_I = J_I(C_k)$, etc.

III. Coupled Diffusion

Field enhancement strongly affects the migration of dopants in diffusion steps. The effect dominates in the vicinity of p-n junctions, especially when the carrier concentrations near the junction exceed the intrinsic number. Neglecting heavy doping effects, the exact formulation of the field enhancement is given by (4)–(6).

$$\frac{\partial C_{\mathrm{As}}}{\partial t} = \operatorname{div}\left(D_{\mathrm{As}} \cdot \operatorname{grad} C_{\mathrm{As}} + \mu_{\mathrm{As}} \cdot C_{\mathrm{As}} \cdot \operatorname{grad} \psi \right) \tag{4}$$

$$\frac{\partial C_B}{\partial t} = \operatorname{div}\left(D_B \cdot \operatorname{grad} C_B - \mu_B \cdot C_B \cdot \operatorname{grad} \psi \right) \tag{5}$$

$$\operatorname{div} \operatorname{grad} \psi = \frac{q}{\epsilon} \cdot \left(2 \cdot n_i \cdot \sinh\left(\frac{e\psi}{kT} \right) + C_B - C_{\mathrm{As}} \right). \tag{6}$$

This correct implementation is, for the sake of simplicity,

seldomly used in simulation programs. When only equally charged dopants are present the influence of the internal field is modeled by the field-enhancement factor of the diffusion coefficient (7).

$$h = 1 + \frac{C}{2n_i} \cdot \left(\frac{C^2}{4n_i^2} + 1 \right)^{-0.5}. \tag{7}$$

If p-n junctions are present, the analytic solution of the quasi-neutral approximation, (3), replaces the Poisson equation.

$$\psi = U_T \cdot \operatorname{arsinh}\left(\frac{C_{As} - C_B}{2 \cdot n_i} \right). \tag{8}$$

Both methods reduce the amount of computation time at least by a factor two since the amount of computation time is proportional to about the square of the number of the physical quantities. The use of the field-enhancement factor, however, additionally simplifies the structure of the continuity equation, causing an additional nonlinear term in the diffusivities. Considering a simulation of arsenic, boron, and the electrostatic potential we can reduce the amount of computation time and memory by a factor 4/9 using (8) instead of (6). These facts have supported the implementation of the quasi-neutral approximation into simulation programs for engineering application where high performance and a small amount of memory are powerful arguments. Some publications, e.g., [1]–[3], have dealt with coupled diffusion and the question if the quasi-neutral approximation may replace the exact Poisson equation or not. The results presented in these papers are not satisfying. In [3] the exact Poisson equation was replaced by the depletion assumption, e.g., set $n = 0$, $p = 0$, and ψ computed by (6) in the vicinity of the junction. In the other domains the quasi-neutral approximation has been used to replace the Poisson equation. The boundary between the two domains was posed at the end of the space-charge region. In [2] an exact solution of the Poisson equation is presented; however, the computations do not involve a p-n junction.

Correct implementation of field-enhanced diffusion is not the only point of interest. Sometimes it is claimed that the influence of the electrostatic potential on the charge state of vacancies necessitates a precise knowledge of these physical quantities or at least an estimation of the differences caused by various models. As it is often proposed that the diffusion of, e.g., boron, arsenic, etc. occurs mainly by charged vacancies, the electrostatic potential can influence the enhancement of diffusion via point defects.

Since recently developed models for diffusion have become more advanced and accurate, a comparison between the results of the two models seems necessary. As the space charge causes the difference between the two models, we will without loss of generality concentrate our investigations to boron and arsenic p-n junctions with fairly high concentrations at the junctions.

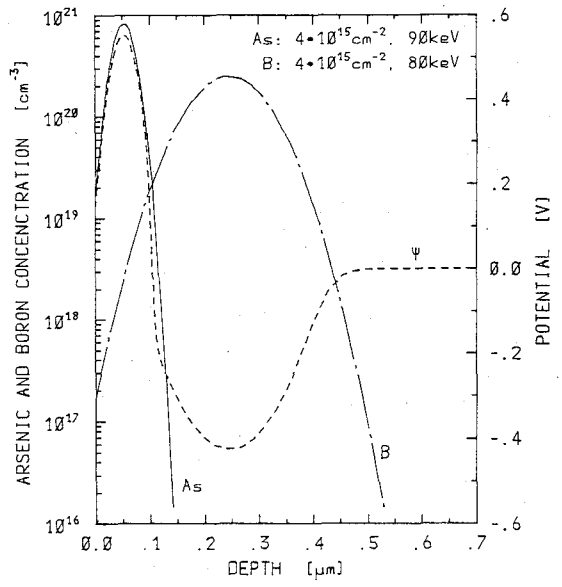

Fig. 1. Arsenic and boron profile after ion implantations. As: 4×10^{15} cm^{-2} with 90 keV, LSS, Gaussian; B: 4×10^{15} cm^{-2} with 80 keV, LSS, Gaussian. The electrostatic potential is computed from the exact Poisson equation at 950°C.

The first example demonstrates a worst case estimation. Ion implantations, As 4×10^{15} cm^{-2} with 90 keV and boron 4×10^{15} cm^{-2} with 80 keV, lead to boron dopant concentrations which exceed those commonly used. Furthermore, we neglect any clustering effects of boron and arsenic to obtain an extreme electrical field and considerable space charges. The diffusivities are defined by (9) and (10). They are modeled temperature dependent and include enhanced diffusion by charged point defects.

$$D_B = 0.037 \frac{\text{cm}^2}{s} \cdot \exp\left(\frac{-3.46\ \text{eV}}{kT} \right) + 0.72 \frac{\text{cm}^2}{s} \cdot \exp\left(\frac{-3.46\ \text{eV} - q \cdot \psi}{kT} \right) \tag{9}$$

$$D_{As} = 0.066 \frac{\text{cm}^2}{s} \cdot \exp\left(\frac{-3.44\ \text{eV}}{kT} \right) + 12.0 \frac{\text{cm}^2}{s} \cdot \exp\left(\frac{-4.05\ \text{eV} + q \cdot \psi}{kT} \right). \tag{10}$$

Fig. 1 shows the initial distribution of boron and arsenic and the solution of the potential due to (6) for a temperature of 950°C. Fig. 2 reveals the final distribution and the electrostatic potential after an inert diffusion of 30 min at 950°C. The transient behavior of the dopants is worth being investigated more carefully. Figs. 3–5 reveal a survey of the time dependent migration of the boron and arsenic dopants in the vicinity of the junction. They indicate that the diffusion consists of two periods, a short field-controlled period (ca. 100 s) followed by an ambipolar diffusion. The strong field-controlled migration of both dopants in the vicinity of the junction leads to a steepening of the dopant profiles. This causes the transition from the n-domain to the p-domain to become narrower and increases, therefore, the electrical field. These two effects

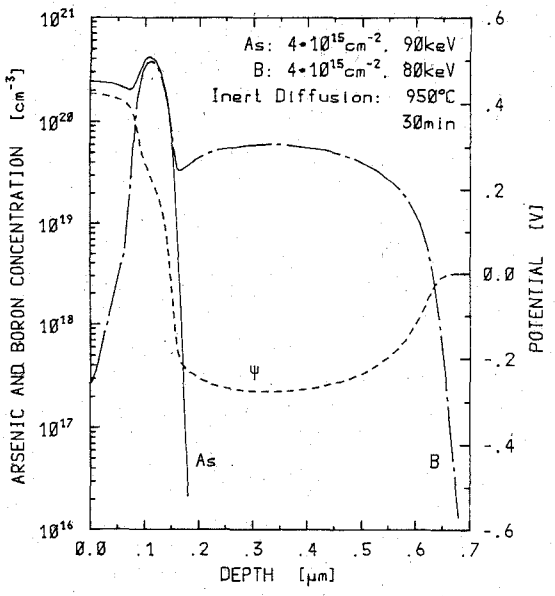

Fig. 2. Arsenic and boron distribution and electrostatic potential after an inert diffusion step of 30 min at 950°C. The initial distribution is given in Fig. 1.

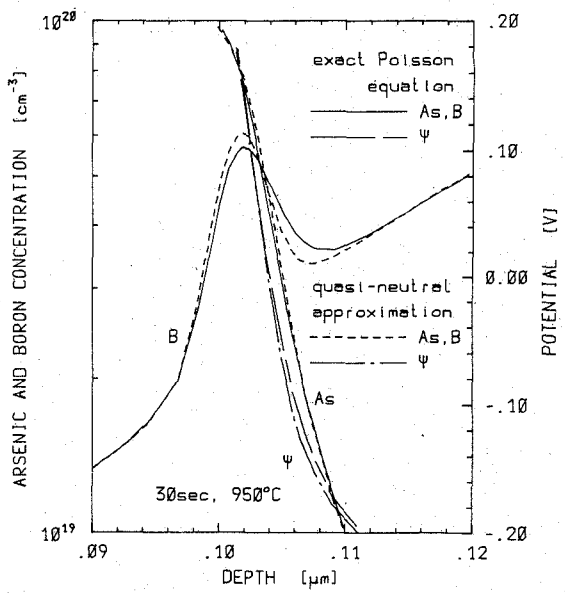

Fig. 4. Arsenic and boron distributions after 30-s inert diffusion at 950°C due to the models *P* and *Q*.

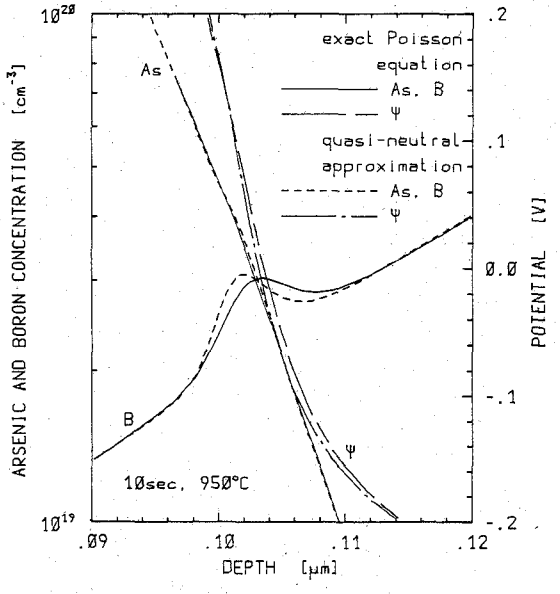

Fig. 3. Arsenic and boron distributions after 10-s inert diffusion at 950°C due to the models *P* and *Q*.

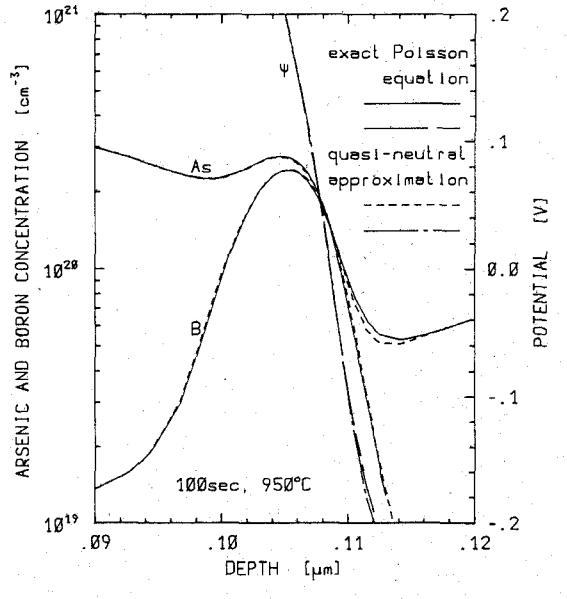

Fig. 5. Arsenic and boron distributions after 100-s inert diffusion at 950°C due to the models *P* and *Q*.

enhance each other and lead to a steadily increasing flux of dopants, especially of boron atoms which have a higher mobility. The process will weaken as soon as the dopant migration will have reduced the abrupt transient in the junction and will have replaced it by a smooth junction. The decreasing electrical field weakens the field-controlled flux. Figs. 3–5 show the As and B concentrations, and ψ for both models at 10, 30, and 100s, respectively. The potential, due to the models, differs from the beginning of the diffusion step. The steeper slope of ψ computed by model *Q* (quasi-neutral approximation (4),(5),(8)) leads to a stronger electrical field. Therefore the field-controlled migration of the boron and arsenic due to model *Q* is more distinct than that of model *P* (exact Poisson equation (4)–(6)). The boron profiles reveal the differences more clearly because the mobility is larger and the gradient is smaller than that of the arsenic profiles. Fig. 6 reveals the transient dependence of the space charge and confirms our findings. Since the space charge, causing the main difference between the two models *Q* and *P*, disappears after 100 s, the differences between the two models will certainly not increase. On the contrary, Figs. 3–5 indicate that the small differences which have been built up in the first period are too small to resist the following diffusion step. Fig. 7 supports the suggestion. Fig. 7 shows the time dependence of the quantities at a depth of 0.099 μm. This depth coincides with the n-side of the p-n junctions close to the junction. Differences of the dopant concentrations

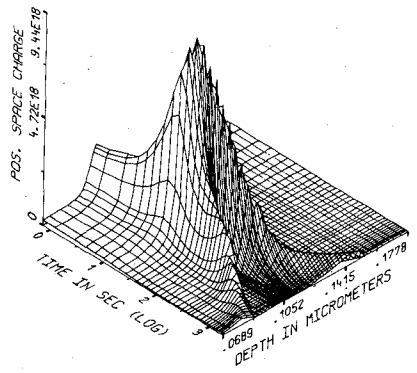

Fig. 6. Time dependent survey of the positive space charge density which is defined as the maximum of 0 and $n-p+N_A-N_D$ in the vicinity of the p-n junction for an inert diffusion step at 950°C. Initial and final distribution are given by Figs. 1 and 2, respectively.

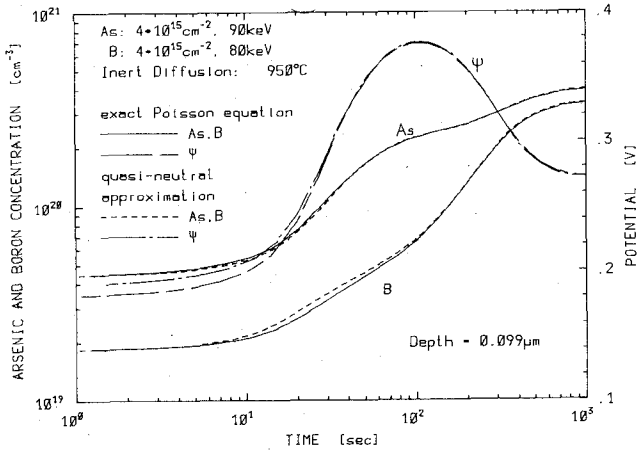

Fig. 7. Transient behavior of arsenic, boron and the electrostatic potential at a depth of 0.099 μm for model P and model Q.

begin to build up after 10 s, remain stable until 200 s, and then slowly disappear.

The dopant migration during the following time period is mainly characterized by an "ambipolar" diffusion in the vicinity of the p-n junction. This means that the dopants migrate in such a way that neither significant space charge nor a significant electric field can be built up. The ambipolar diffusion seems controlled by the diffusivity of the arsenic atoms which leads to a strongly retarded diffusion of the boron atoms. Away from the junction the diffusion of the majority dopants reveals normal results whereas the migration of minority dopants is strongly retarded. This is caused by the extreme values of the potential which enhances majority diffusivities but retards the minority diffusivities. A comparison of the initial profile (Fig. 1) and the final profile (Fig. 2) reveals the impression that the two dopant distributions moved towards each other, a fact which is not unexpected, since the dopants have a different charge. It is worth noting that although the models lead to different dopant concentrations in the vicinity of the junction, the junction depth is not affected by these differences. It might be interesting if the small differences at the p-n junctions which occur at the beginning of the diffusion step are sufficiently large to cause a different electrical characteristic of the device.

Our second example shows a more realistic process. Two implantations, As 4×10^{15} cm^{-2} with 90 keV and boron 7×10^{13} cm^{-2} with 40 keV, are followed by an 1000°C annealing step for 40 min. In this example the boron concentration is far too small to influence the electrostatic potential at the n-doped side. Since the peak value of the boron concentration is of the same order of magnitude as the intrinsic number (6×10^{18} cm^{-3} at 1000°C), we can almost neglect the influence of the boron concentration on the computation of the potential away from the peak value. The influence of the field can clearly be seen by a "valley" in the boron profile (Fig. 8) caused by the field controlled flux. The electrical field pushes the boron atoms towards the surface against the diffusion flux. Certainly the "valley" in the boron profile moves as the steep gradient of the arsenic profile shifts into the bulk. The space charge, Fig. 9, indicates again two periods. A short time domain characterized by remarkable space charge and electrical field and a following period dominated by normal diffusion of the arsenic and the retarded diffusion of the boron atoms. In contrary to the former example we cannot speak of ambipolar diffusion. The arsenic migration is hardly influenced by the boron concentration since arsenic exceeds the intrinsic number and the boron concentration by some orders of magnitude. On the other hand, the boron migration is controlled by the arsenic in two ways. Firstly, the movement of the boron atoms is field controlled in the vicinity of the p-n junction and secondly strongly retarded in the regions of high arsenic concentration. (The positive electrostatic potential suppresses the second term in the diffusion coefficient of (9).) Here again, the quasi-neutral approximation computes a steeper slope of the electrostatic potential, causing a stronger electrical field. Therefore, the "valley" in the boron profile computed by model Q is more distinct than that of model P. The maximum difference between the two models can be expected when the space charge reaches its maximum, which occurs at 30 s. Fig. 10 shows the time dependence at a depth of 0.12 μm. The differences caused by the two models appear as soon as the junction moves into the vicinity of the depth under consideration and disappears as soon as the junction moves away. Fig. 11 shows the p-n junction at a time of 30 s. The stronger electrical field causes a deeper "valley" in the boron profile for model Q than for model P. Away from the p-n junction no differences can be seen. The depth of the p-n junction is not dependent on the model because the steep gradient of the arsenic compensates the different boron profiles.

The two examples justify the use of the quasi-neutral approximation instead of the exact Poisson equation when the duration of the diffusion step is long enough to exceed the first, field-controlled period of dopant migration. The first example reveals that the differences diminish after at least 200 s at a temperature of 950°C. We could not observe any significant differences for typical diffusion steps in IC fabrication at temperatures beyond 950°C. At lower temperatures one has to be more careful. To estimate the temperature dependence of the coupled diffusion we have simulated the dopant migration of example one at a

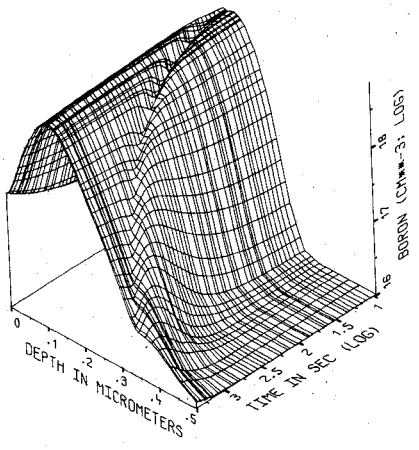

Fig. 8. Time-dependent boron profile for an inert diffusion step of 40 min at 1000°C The initial distribution has been obtained by ion implantations. As: 4×10^{15} cm^{-2}, 90 keV and B: 7×10^{13} cm^{-2}, 40 keV.

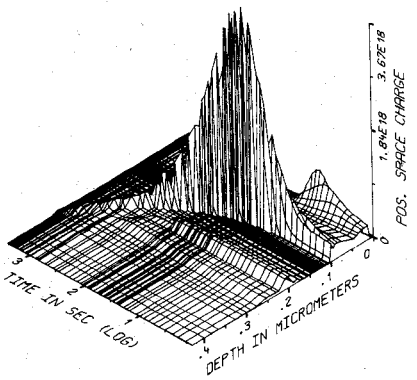

Fig. 9. Transient behavior of the positive space charge density which is defined as the maximum of 0 and $n-p+N_A-N_D$ for the process of Fig. 8.

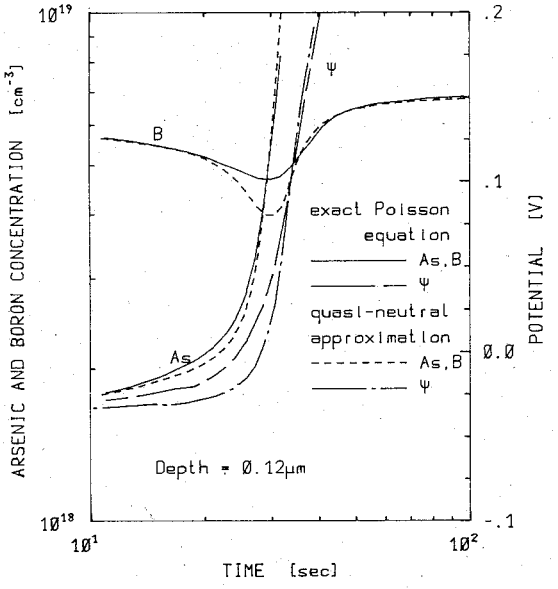

Fig. 10. Time dependence of arsenic, boron and the electrostatic potential due to the models P and Q for an inert diffusion at 1000°C at a depth of 0.12 μm.

temperature of 800°C. The diffusivities at 800° and 950°C differ typically by a factor of 100

$$\exp\left(\frac{-3.44\text{ eV}}{kT}\right)=\begin{cases}6.92\times10^{-17}, & \text{for } T=800°\text{C}\\ 6.64\times10^{-15}, & \text{for } T=950°\text{C}.\end{cases}$$

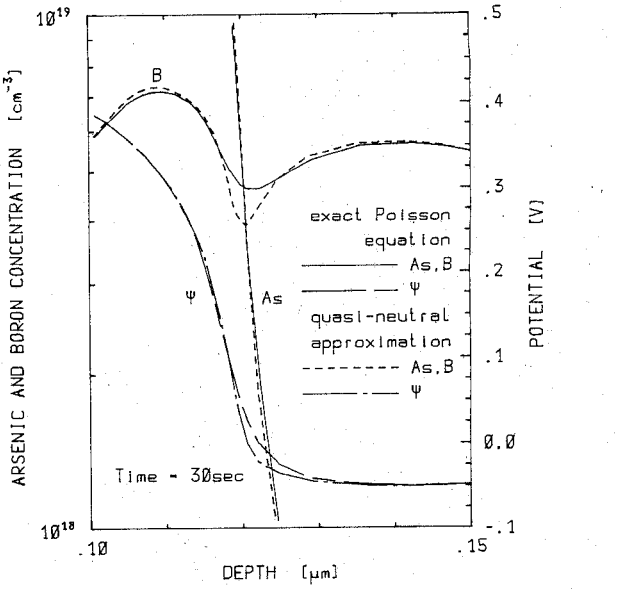

Fig. 11. Arsenic, boron, and the electrostatic potential due to the models P and Q for an inert diffusion at 1000°C after 30 s.

For a diffusion-controlled process we would expect the 800°C process to be retarded by a factor 100 compared to the 950°C process. But, because the first period is field-controlled we have to modify our retardation factor. The relation between the field flux and the diffusion flux can be chosen to estimate the time dependence in the first period. If we confine our interest to the vicinity of the p-n junction and assume that $C_{As}=C_B$, $D_{As}\ll D_B$ and $C_{As}-C_B\ll 2\cdot n_i$, some algebra leads to J_F: $J_D=C_{As,B}$: n_i. Since n_i evaluates to 4.4×10^{18} cm^{-3} and 1.5×10^{18} cm^{-3} at 950° and 800°C, respectively, we, therefore, expect all diffusion processes to be retarded by a factor 100 whereas the field controlled processes are only retarded by a factor of 100/3 = 33. Furthermore, the effects of field-controlled dopant migration should be more distinct.

Our simulations confirm our rough estimations. For 800°C the space charge shows its maximum at 1700 s in contrast to the 50 s of Fig. 6 for the 950°C step. The dopant distributions in the vicinity of the junction at 300, 1000, and 3000 s for the 800°C process are similar to the distributions at 10, 30, and 100 s for the 950°C step. The qualitative behavior of the two processes differs little. The last example confirms that even at low temperatures, the quasi-neutral approximation describes the dopant migration qualitatively well in the vicinity of a p-n junction and excellently in the other domains. The small differences at the junction are probably not sufficiently large to cause considerable effects on the electrical behavior of the device.

As a simple example of our two-dimensional code we have simulated a coupled arsenic boron diffusion. Equations (4) and (5) are the continuity equations for boron and arsenic, respectively, and (6) represents Poisson's equation. The diffusion coefficients are modeled as given by (7) and (8).

We simulate an inert annealing step at 1000°C temperature. As initial state we assume a homogeneously doped boron substrate (10^{15} cm^{-3}) implanted with arsenic (10^{15} cm^{-2}, 130 keV, LSS, Gaussian) through a mask (35 nm)

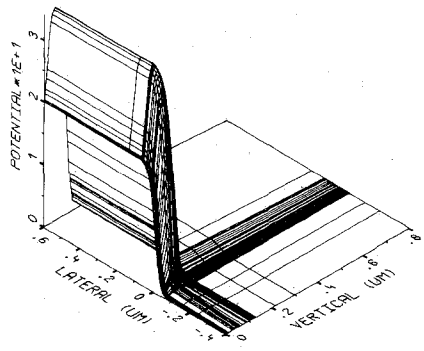

Fig. 12. Two-dimensional distribution of the electrostatic potential after an inert diffusion of 10 s at 1000°C.

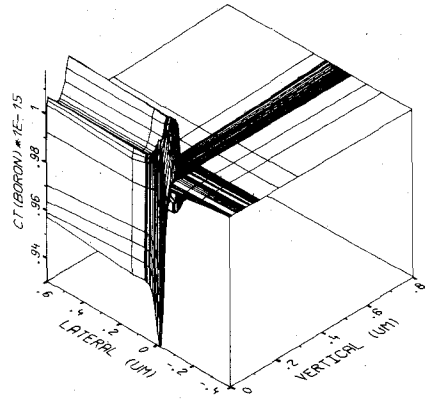

Fig. 13. Two-dimensional distribution of the boron dopants after an inert diffusion of 10 s of 1000°C.

with an infinitely steep edge to field oxide (at the origin of the coordinate axis in the following figures).

The influence of the electrical field on the boron profile is the main purpose of this simulation. Since the arsenic concentration exceeds the intrinsic number and the boron concentration by several orders of magnitude we do not expect noticable effects of the electrical field on the shape of the arsenic profile. The diffusion of the arsenic profile is, therefore, of minor interest and we concentrate on the discussion of the electrostatic potential and the boron profile.

Fig. 12 shows the electrostatic potential at the beginning of the diffusion. The field term in the continuity equation for boron causes a flux of doping atoms in the direction of positive gradients of the potential. Fig. 13 indicates that in the beginning of the diffusion the boron profile forms a U-shaped maximum of increased concentrations at the surface and in the bulk.

The maximum near to the surface is smaller than the maximum in the bulk. This is caused by the less pronounced gradient of the electrostatic potential at the surface. Since the ambient condition at the surface ($J_{\text{Boron}} = 0$) inhibits a flux of boron perpendicular to the surface the concentration of boron is relatively strongly depleted in this region. At the edge of the U-shaped maximum the boron profile is depleted. The minima of the boron profile are caused by the gradients of the potential in lateral and vertical direction at the origin of the coordinate system.

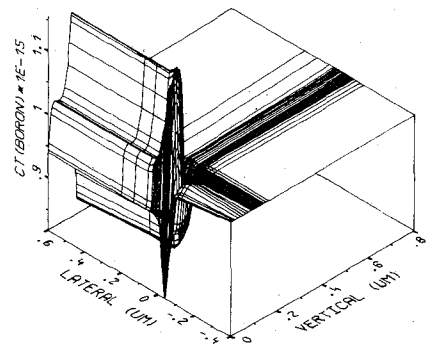

Fig. 14. Two-dimensional boron distribution after an inert diffusion of 100 s at 1000°C.

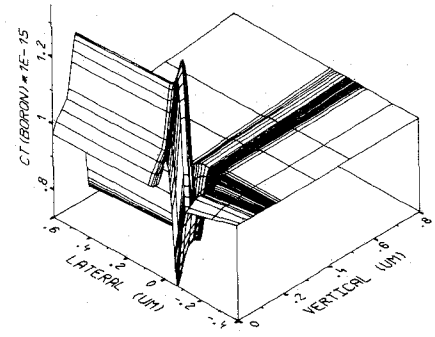

Fig. 15. Two-dimensional boron distribution after an inert diffusion of 500 s at 1000°C.

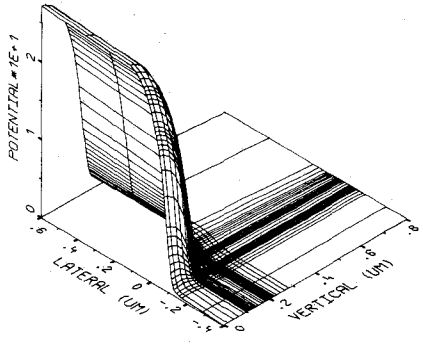

Fig. 16. Two-dimensional distribution of the electrostatic potential after an inert diffusion of 1500 s at 1000°C.

The potential profile spreads with increasing diffusion time corresponding to the spreading arsenic profile as indicated in Figs. 12 and 16. The maximum of the boron profile close to the surface is, therefore, reduced (Fig. 14) and vanishes finally (Figs. 15 and 17), whereas the maximum in the bulk increases and spreads. It should be mentioned that the minimum of the boron profile remains at the surface next to the mask edge (origin of the coordinate system) caused by the lateral field controlled flux of boron. Affected by the field in vertical and lateral direction, the boron profile forms a distinct peak value with increasing diffusion time. This extreme accumulation of boron atoms can only be explained by a two-dimensional simulation. The shape of the potential causes lateral and vertical fluxes of boron which accumulate the dopants in a center and cause the peak values. After 1500 s diffusion

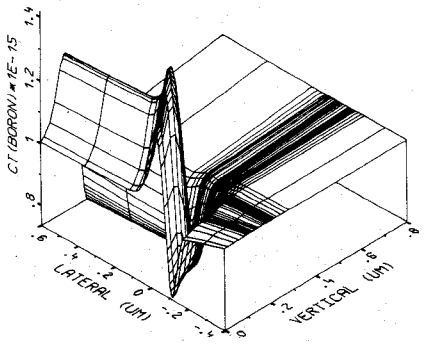

Fig. 17. Two-dimensional distribution of the boron concentration after an inert diffusion of 1500 s at 1000°C.

time no significant change in the qualitative behavior of diffusion will take place.

An experimental verification of this fairly simple example is, most unfortunately, not possible, since the resolution of even the most modern, e.g., SIMS equipment is too coarse for that purpose. However, we have certainly verified the numerical accuracy of our program with problems whose solutions are known.

IV. Arsenic Cluster Models

Arsenic impurities in silicon have a temperature dependent solubility, e.g., [4] or [5]. Arsenic concentrations below this solubility limit are electrically active. Beyond this limit they form clusters and/or precipitate and become electrically inactive. Therefore, the electrically measured profiles (e.g., spreading resistance measurements) differ from atomistic measurements (e.g., SIMS measurements). The usage of shallow high dose As implantations for the fabrication of MOS source and drain regions in VLSI devices necessitates dynamic cluster models for process simulation.

A fair amount of work has been done to evaluate equilibrium solubilities but relatively few experiments have been done recently to determine the transient behavior of clustering and declustering. We have investigated the experiments published in [5] and [6].

In [5] measurements and simulations of typical MOS source fabrication processes are simulated: An implantation of 2×10^{16} cm^{-2} arsenic with an energy of 140 keV through a 25-nm oxide followed by an annealing step of 20 min at 1000°C (+a further annealing step at 800°C for 60 min). The physical model used for the simulation is well described in [5], ((11)–(14) and (6)) allowing us to proceed with the simulation and test the abilities of our code. The only difference from the model of [5] is the use of the exact Poisson equation instead of the quasi-neutral approximation. The exchange of the arsenic diffusivity from (10) to (14) offers no problems and causes only minimal modifications of the code.

$$m\cdot \mathrm{As}^{+} + k\cdot e^{-} \leftrightharpoons \mathrm{Cl}^{(m-k)+} \tag{11}$$

$$\frac{\partial C_{\mathrm{As}}}{\partial t} = \mathrm{div}\left(D_{\mathrm{As}}\cdot \mathrm{grad}\, C_{\mathrm{As}} + \mu_{\mathrm{As}}\cdot C_{\mathrm{As}}\cdot \mathrm{grad}\,\psi\right) + k_D\cdot C_{\mathrm{Cl}} - k_C\cdot C_{\mathrm{As}}^{m}\cdot n^{k} \tag{12}$$

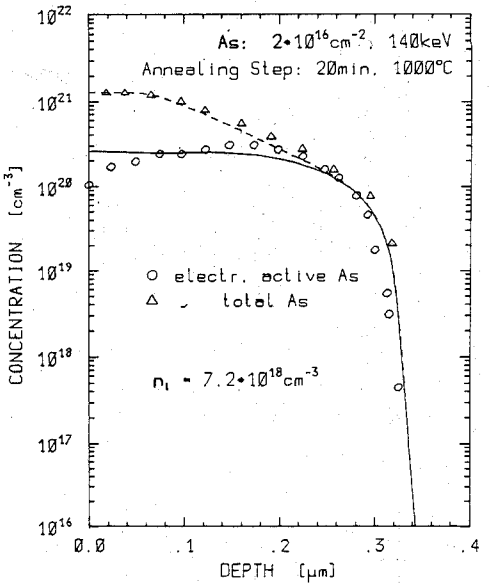

Fig. 18. Measured and computed data of the total and electrically active arsenic concentration after a 20-min annealing step at 1000°C.

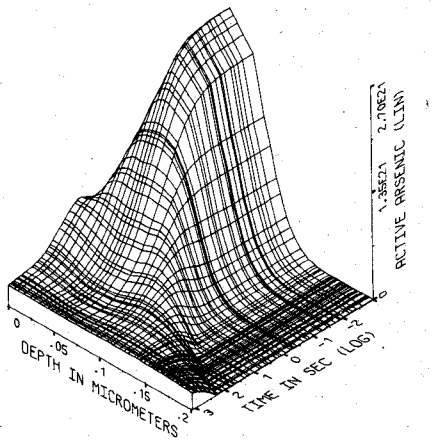

Fig. 19. Transient dependence of the electrically active arsenic during a 20 min annealing step at 1000°C.

$$m\cdot\frac{\partial C_{\mathrm{Cl}}}{\partial t} = -k_D\cdot C_{\mathrm{Cl}} + k_C\cdot C_{\mathrm{As}}^{m}\cdot n^{k} \tag{13}$$

$$D_{\mathrm{As}} = 22.9\cdot\exp\left(\frac{-4.1\ \mathrm{eV}}{kT}\right)\cdot\frac{1+100\cdot n/n_i}{1+100}\ \mathrm{cm}^2\cdot\mathrm{s}^{-1}. \tag{14}$$

In (13) and (14) k_D and k_C have to be corrected by a factor m to be consistent with the formulas in [5]. The only parameter to be matched has been the intrinsic number which equals 7.2×10^{18} cm^{-3} at 1000°C in our calculation. Fig. 18 shows the result after the first annealing step which coincides well with the measured data and the published results in [5].

Figs. 19 and 20 show the transient dependence of the active and the clustered arsenic. Clustering and declustering are very fast processes compared to diffusion at 1000°C. Fig. 19 reveals that the equilibrium between the active and the clustered arsenic has well been obtained after 20 min. The cluster concentration shows a significant maximum

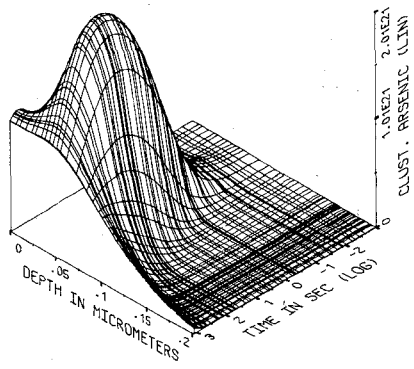

Fig. 20. Transient dependence of the clustered arsenic concentration during a 20-min annealing step at 1000°C.

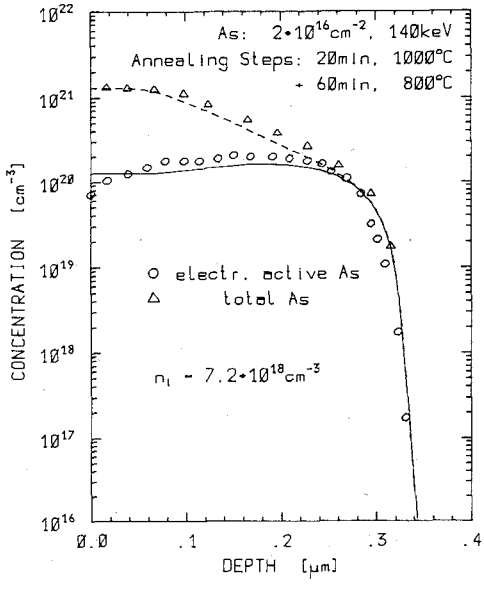

Fig. 21. Total and electrically active arsenic concentrations after two annealing steps.

after 30 s of annealing and can easily be explained. In the first 30 s, clustering is the dominating process and results in increasing cluster concentrations. Then the diffusion of the active arsenic leads to decreasing active concentrations which are immediately compensated by declustering. This leads to a decrease of the cluster concentration.

The flat profiles in the active arsenic concentration results from the fact that the change of C_{As} with time is proportional to the deviations from the equilibrium concentration.

The simulation of the following 800°C annealing step for 60 min is shown in Figs. 21 and 22. The results coincide again well with the experimental data and the results obtained by [5].

The descending slope of C_{As} in Fig. 22 indicates that the equilibrium between clustered and active arsenic has not yet been achieved. The use of a static instead of a dynamic cluster model would have resulted in too small active arsenic concentration.

One difference between the code presented in [5] and our code is worth being discussed more carefully. In [5] it is stated that "clustering effects are computed individually"

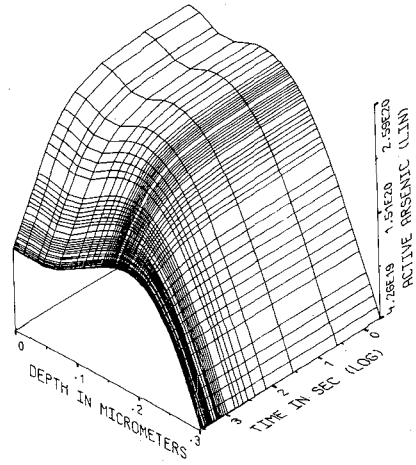

Fig. 22. Transient dependence of the electrically active arsenic during a 60-min annealing step at 800°C.

and "concentrations are updated with small time intervals." This indicates that (12)–(14) are evaluated in two steps (a diffusion step followed by a clustering step). To the contrary, our code solves (12) to (14) simultaneously. Clearly the differences between the two methods become negligible if the time and space discretization is sufficiently small. Nevertheless, we believe our method to be the better one from a physical and mathematical point of view. The differences will occur when the grid width in space and time is increased, in order to obtain a faster and less memory intensive program. We also want to pronounce that a dynamic phase transformation, such as the arsenic clustering, has not to be "implemented" into the code but represents a typical physical effect for which our code has been developed.

The model of Tsai is based on experiments using only thermal treatment of the wafer to investigate the clustering of arsenic. It seems interesting to compare this model to experiments which have been published in [6]. These experiments use laser annealing after the implantation to be sure that all arsenic rests on lattice sites and is electrically active. Further thermal treatment by a CW CO_2-laser and conventional furnace is performed to investigate the transient behavior of the clustering.

The experimental results are shown in Fig. 23. The x-axis shows the annealing time in a logarithmic scale, the y-axis the sheet carrier concentration per cm^{-2} surface. We have chosen the impurity distribution after the laser annealing as the initial solution for our calculation and have simulated the annealing experiments taking into account diffusion as well as clustering. The dashed lines show the simulated results using the model of [5] for 900°, 800°, and 700°C. Unfortunately measurements and simulation differ greatly. This may be caused by the different annealing methods applied to the wafers after implantations and/or by the different thermal post treatment. The different total arsenic concentration of the two experiments (ca. 2×10^{21} cm^{-3} in [5] and 4×10^{20} cm^{-3} in [6]) will cause different values of the stable active arsenic concentration. This is strongly pronounced in the linear scale of the sheet carrier

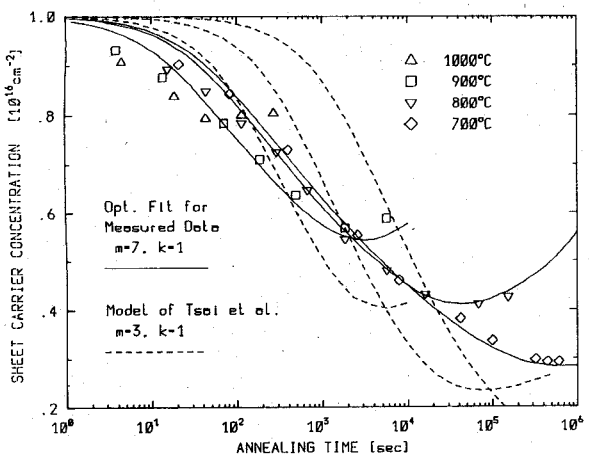

Fig. 23. Measured data of the sheet carrier concentration of [6] at four temperatures. Results due to the cluster model of [5] and an optimal fit to measured data.

concentration in Fig. 23. The differences reveal that the dynamic properties of arsenic clustering are not yet fully understood in terms of developing a model which covers a wide range of processes and can be installed in general process simulation programs. We have used our code to find values for k_C and k_D to agree with measurements optimally. Our computations reveal that the value of the minimum is mainly affected by the ratio of k_C/k_D and that the time dependence of the minimum can be controlled by modifying k_C and k_D in the same way by a factor. It is worth noting that the slope is nearly independent of k_C and k_D and yields far too steep lines for $m=3$ and $k=1$. After several attempts $m=7$ and $k=1$ yields the best results. The values of k_D and $k_{eq}=k_C/k_D$ used for the optimal fit of the measured data in Fig. 23 are listed below.

T (°C)	k_D (s^{-1})	k_{eq} (cm^{21})
700	1.089×10^{-6}	5.500×10^{-142}
800	3.366×10^{-5}	4.000×10^{-143}
900	6.335×10^{-4}	4.792×10^{-144}

The temperature dependence of the values of k_{eq}, k_D, and k_C can be described by Arrhenius laws.

The 1000°C process could not be optimized due to the shortage of experimental data. The poor agreement at the beginning of the annealing step is most likely caused by a very simple initial solution. We have assumed that all of the arsenic was activated by the laser annealing step.

Our computations, using the measured data of Fig. 23, indicate a trend to larger cluster sizes. Since the values of n nearly equals that of C_{As}, we could have also set $m=6$, $k=2$; $m=5$, $k=3$ or $m=4$, $k=4$ without changing the term $C_{As}^m \cdot n^k$ and, therefore, the dynamics of (12) or (13).

V. Some Mathematical and Numerical Aspects

A high level of sophistication on the numerical side is necessary to produce a code which is useful for the investigations of physical problems. Since the main purpose of our code is the investigation of physical models, we try to free the user from as many mathematical problems as possible, e.g., the mesh in space and time is created and adapted automatically. The creation and modifications of the grid are mainly influenced by the idea to minimize the truncation errors. In the space domain this implies optimal distribution of the gridpoints, in the time domain the steps have to be sufficiently small. The decisions in both domains are critical and represent a compromise between accuracy and the computer capabilities. The distribution of the grid points is controlled by the following three criteria.

a) Quasi-uniformity: (only in the 1-D code) In [11] it is shown that the spatial truncation error of a quasi-uniform mesh reduces proportional to the square of the grid distances if the maximum ratio between two adjacent distances minus unity is small compared to unity.

b) The minimization of the second derivative of the physical quantities.

c) The maximum ratio/difference criterion.

The use of a quasi-uniform mesh combines the advantage of the small discretization error of a uniform mesh and the possibility to accumulate gridpoints at domains of physical interest. The high dynamic range of the diffusion problem in device fabrication steps (dopant concentrations of interest are in the range from 10^{10} cm^{-3} to 10^{22} cm^{-3}) leads to problems in the design of a spatial mesh. On the one hand, dose constancy is required even at long time diffusions, on the other hand the dopant distribution in the vicinity of a p-n junction is of special interest for the electrical performance of the device. Calculations revealed that b) creates a grid which permits long time inert diffusions without any loss of dopants. Unfortunately this criterion ignores the low concentration domains which are very often the domains of electrical interest. Criterion c) sets points in the vicinity of steep gradients and close to the p-n junctions (max. difference criterion for the potential) and compensates the disadvantages of criterion b). A minimum mesh length has been invoked to avoid too fine grids near discontinuities (e.g., initial solution of a predeposition).

Since the dopant migration during a diffusion step leads to strongly varying profiles, a rigid grid seems to be unsuitable for an advanced simulation program. Our mesh is automatically modified during the diffusion step. Observations during simulations reveal that deviations from the initial doses mainly occur during grid modifications. If the grid contains enough points to fulfill all criteria the variations turn out to be statistical deviations and do not augment beyond the specified values.

For the computation of the time step we use backward differentiation formulas (BDF) similar to those proposed in [12]. In the one dimensional code we use BDF of sixth order to obtain an optimal transient behavior of our computation. In the two-dimensional program limited computer resources constrain us to use BDF of third order. Here again the wide dynamic range of interest in the physical quantities necessitates modifications of the commonly proposed model. The original error estimation in

[12] turns out to be far too restrictive to yield reasonable time steps.

In the following table the number of time steps, the number of newton iterations, the number of newton iterations per time step and the computation times are listed for the computations in Fig. 23. The spatial grid consists of approximately 120 points, the parameter is the accuracy for the transient integration.

number of time steps	number of Newton iterations	number of iterations per time step	computation time in sec
47	147	3.12	265
54	159	2.94	291
74	192	2.54	357
105	197	1.87	384
115	215	1.86	419

The numbers in the table indicate that there is no linear relation between the number of time steps and the computation time and that a less accurate transient integration does not significantly decrease the computation times. The larger time steps lead to less accurate predictions for the initial solution of the Newton iteration. This in turn necessitates more Newton iterations steps and compensates the intended reduction of computation time by using wider time steps.

For the discretization of the space operators we use the methods of finite differences. Equations identical to (1) to (3) are used in device simulation. We have, therefore, implemented a great deal of sophisticated numerical methods developed for device simulation into our code. The formulae for the discretization of (1)–(3) have been taken from [11]. An exponential fitting factor for the discretisation of the flux relation (2), which has been successfully applied in the device simulation, has also been used. The Newton linearisation of the discretization formulas leads to a block-tridiagonal (1-D) or a block-pentagonal (2-D) equation system which is solved by Gaussian elimination (1-D) and an iterative solver (2-D), respectively. An increment damping method developed by Deufelhard [13] is implemented to avoid detrimental overshooting effects during the iteration process.

VI. Summary and Conclusions

This paper presents a scientific general purpose program for process simulations in one and two dimensions. The number and kind of physical quantities to be simulated is specified by a few parameter definitions. This allows for a wide range of applications and simple modifications of the physical models under consideration.

First applications reveal that the quasi-neutral approximation seems to be an excellent approach to describe the field enhanced diffusion in engineering programs. The differences between the quasi-neutral approximation and the exact Poisson equation are small in the vicinity of p-n junctions and negligible elsewhere. The differences become smaller with increasing duration of the diffusion step and with increasing process temperature. We can, therefore, recommend the use of the quasi-neutral approximation whenever fast performance and limited memory resources necessitate simplifications of physical models. In our code we will, nevertheless, use the Poisson equation and investigate occasionally the differences between the two models at various applications.

A second application deals with the phenomenon of arsenic clustering. The comparison with published results reveals that our program can accurately simulate up-to-date models for dynamic arsenic clustering without exceeding its specifications. The strong deviations of time dependent measurements reveal that the physical background of arsenic clustering/precipitation is not yet understood and necessitates further investigations. The commonly proposed cluster models [5], [8], or [9] are well established but do not lead to a concentration independent C_{As} as it was pointed out in, e.g., [4]. On the contrary [10] proposes that precipitation is the physical effect responsible for the differences between electrically active and the total arsenic concentration. This model would lead to a doping independent active arsenic concentration. Furthermore, the influence of different annealing methods on the dynamics of arsenic has not been investigated satisfactorily. The time dependent measurements of [6] and [7] reveal contradictory results for the dynamics of clustering after laser annealing using different techniques. Our simulations of laser annealing process indicate that cluster sizes between 5 to 7 atoms lead to a better agreement between measured and computed data than a cluster size of 3 atoms.

From the numerical point of view the first applications of our code clearly reveal that a high level of numerical and mathematical sophistication is absolutely necessary for a scientific process simulation program. An optimal discretization of the partial differential equations and automatically generated grid in space and time guarantee minimal discretization errors and fine resolutions of spatial and transient domains of interest.

Acknowledgment

The authors would like to thank the Interuniversitäre Rechenzentrum for the generous amount of computer resources it provided.

References

[1] S M Hu, "Diffusion in heavily doped semiconductors," *J. Appl. Phys*, vol. 43, no. 4, pp. 2015–2108, 1972.
[2] R. Shrivastava and A. Marshak, "Charge neutrality and the internal electric field produced by impurity diffusion," *Solid-State Electron.*, vol. 23, pp 73–74, 1980.
[3] P J. Anthony, "Alteration of Diffusion Profiles in semiconductors due to p-n junctions," *Solid-State Electron.*, vol. 25, no. 10, pp. 1003–1009, 1982.

[4] A. Lietoila, J. F. Gibbons, and T. W. Sigmon, "The solid solubility and thermal behavior of metastable concentrations of As in Si," *Appl. Phys. Lett.*, vol. 36, no. 9, pp. 765–768, May 1, 1980.
[5] M. Y. Tsai, F. F. Morehead, J. E. E. Baglin, and A. E. Michel, "Shallow junctions by high-dose As implants in Si," *J. Appl. Phys.*, vol. 51, no. 6, pp. 3230–3235, June 1980.
[6] J. Goetzlich, P. H. Tsien, G. Henghuber, and H. Ryssel, "CO_2 laser annealing of ion-implanted silicon," *Springer Series in Electrophysics*, vol. 11, pp. 513–519, 1983.
[7] P. H. Tsien, H. Ryssel, D. Röschenthaler, and I. Ruge, "Nd:YAG laser annealing of arsenic-implanted silicon," *J. Appl. Phys.*, vol. 52(4), Apr. 1981.
[8] E. Guerrero, H. Pötzl, R. Tielert, M. Grasserbauer, and G. Stingeder, "Generalized model for the clustering of As dopants in Si," *J. Electrochem. Soc.*, vol. 139, no. 8, pp. 1826–1831, Aug. 1982.
[9] R. Fair, G. R. Weber, "Effect of complex formation on diffusion of arsenic in silicon," *J. Appl. Phys.*, vol. 44, no. 1, pp. 273–279, Jan. 1973.
[10] D. Nobili, A. Arabelos, G. Celotti, and S. Solmi, "Precipitation as the phenomenon responsible for the electrically inactive arsenic in silicon" *J. Electrochem. Soc.*, vol. 130, no. 4, pp. 922–928, 1983.
[11] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. New York: Springer-Verlag, 1984.
[12] R. K. Brayton, F. G. Gustavson, and G. D. Hachtel, "A new efficient algorithm for solving differential-algebraic systems using implicit backward differentiation formulas," *Proc. IEEE*, vol. 60, pp. 98–108, 1972.
[13] P. Deufelhard, "A modified Newton method for the solution of ill-conditioned systems of nonlinear equations with applications to multiple shooting," *Numer. Math.*, vol. 22, pp. 289–315, 1974.