JOURNAL OF COMPUTATIONAL PHYSICS **119**, 46–61 (1995)

Current–Voltage Characteristics Simulation of Semiconductor Devices Using Domain Decomposition*

S. MICHELETTI, A. QUARTERONI,† AND R. SACCO

Department of Mathematics, Politecnico di Milano, Via Bonardi 9, 20133, Milan, Italy

Received February 2, 1994; revised October 24, 1994

We study the current–voltage characteristics of one-dimensional semiconductor devices by numerical approximation based on finite elements of the steady-state semiconductor device equations. A block nonlinear Gauss–Seidel procedure is employed to decouple the full system. Then, at each iteration, a Neumann–Neumann domain decomposition method is applied to solve the linearized equations. Numerical examples will be given, with special emphasis on charge generation effects due to impact ionization. © 1995 Academic Press, Inc.

1. INTRODUCTION

The aim of this paper is to study the current–voltage (I-V) characteristics of a one-dimensional *p-n* diode, which is a typical test device in semiconductor modeling. In Section 2 we will introduce the well-known *drift-diffusion* equations (see [13, 17]) that describe charge flow in a semiconductor device at steady-state conditions. The mathematical problem consists of a set of three highly nonlinearly coupled equations in the unknowns (ψ, n, p), which are respectively electric potential and carrier concentrations (electrons and holes).

In order to reduce the computational effort, a block nonlinear Gauss–Seidel algorithm known in semiconductor literature as Gummel's map [12] is considered to decouple the full system in Section 3. The three resulting linearized equations are suitably formulated and then successively solved in Section 4 by a Neumann–Neumann domain decomposition method [1]. This strategy is motivated by the strongly varying nature of the solutions, which usually exhibit very sharp interior layers across a thin region at the *p-n* junction while behaving smoothly in the remaining part of the device domain.

Concerning the spatial discretization we will respectively employ piecewise linear finite elements for the electric potential equation and exponentially fitted finite elements (à la Scharfetter and Gummel [16]) to handle the convection–diffusion equations for both electron and hole densities (see also [2, 3, 14]). It is well known that the latter elements provide a stable discretization scheme which can be proved to be equivalent to an *optimal upwinding* method, thus being ideally suited for reproducing the sharp layers described above.

In the concluding section 6 we will discuss several numerical results relative to the study of the diode (I-V) curve at some working points of remarkable interest. In particular, we will concentrate on the solution of the problem at the onset of *avalanche breakdown* (see, e.g., [19]). Under these conditions a blowup of carrier concentrations gives rise to an unbounded current flow throughout the device, which in turn dramatically slows down the convergence of Gummel's map. This behaviour is confirmed by a spectral analysis of the map iteration matrix. A suitable strategy to accelerate the original algorithm is therefore proposed, based on a preconditioned BI-CGSTAB iterative method [20]. Our experimental analysis outlines that this method improves substantially the convergence properties of the classical Gummel's map.

2. THE DRIFT-DIFFUSION MODEL FOR SEMICONDUCTORS

Charge flow throughout a semiconductor device at steady state is commonly modeled (see, e.g., [13, 17]) by a set of elliptic partial differential equations that, after a suitable scaling [7], read

$$-\Delta\psi = \rho \tag{1}$$

$$\operatorname{div} \mathbf{J}_n = R(\psi, n, p) \tag{2}$$

$$\operatorname{div} \mathbf{J}_p = -R(\psi, n, p), \tag{3}$$

where

$$\rho = \rho(\psi, n(\psi), p(\psi)) = (p - n + C) \tag{4}$$

$$\mathbf{J}_n = \mu_n(\nabla n - n\nabla\psi) \tag{5}$$

$$\mathbf{J}_p = -\mu_p(\nabla p + p\nabla\psi). \tag{6}$$

Note that ρ depends on ψ through p and n, as will be pointed

---

* This research has been supported in part by Fondi MURST 40% and partly by the Sardinia Regional Authorities.

† Also CRS4, Via N. Sauro 10, 09123 Cagliari, Italy.

0021-9991/95 $6.00

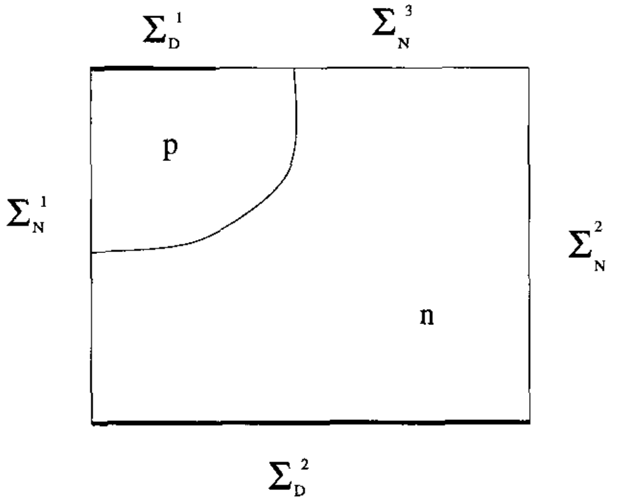

FIG. 1. Simulation geometry of a planar *p-n* diode.

out later on. The equations above are known as the *drift-diffusion* model for semiconductors and can be solved in any open bounded set Ω of $\mathbf{R}^m$ ($m = 1, 2$ in this paper) having a Lipschitz boundary Γ. The unknowns (ψ, n, p) are respectively electric potential and free carrier densities (electrons and holes) so that (1) is a classical Poisson equation which relates electric potential to the space charge density ρ, while (2) and (3) are two continuity equations for electron and hole current densities $\mathbf{J}_n$ and $\mathbf{J}_p$. The constitutive laws (5) and (6) describe current flow throughout the semiconductor as a superposition of a *drift* (or *convective*) term proportional to the electric field $\mathbf{E} = -\nabla\psi$ and a *diffusion* component proportional to the gradient of the concentrations. The doping profile $C(\mathbf{x})$ at the right-hand side of (1) is a known function while the net recombination/generation rate R and the carrier mobilities $\mu_{n,p}$ depend nonlinearly on the unknowns (ψ, n, p), according to the relations (see, e.g., [17, Chap. 4])

$$R = R_{\mathrm{SHR}} + R_{\mathrm{AU}} + R_{\mathrm{II}}, \tag{7}$$

where

$$R_{\mathrm{SHR}} = \frac{pn - 1}{\tau_n(p + 1) + \tau_p(n + 1)} \tag{8}$$

$$R_{\mathrm{AU}} = (pn - 1)(C_n n + C_p p) \tag{9}$$

$$R_{\mathrm{II}} = -(\alpha_n|\mathbf{J}_n| + \alpha_p|\mathbf{J}_p|) \tag{10}$$

and

$$\mu_{n,p} = \left(\mu^0_{n,p} \Big/ \left(1 + \left(\frac{\mu^0_{n,p}|\nabla\psi|}{v^{\mathrm{sat}}_{n,p}}\right)^{\beta_{n,p}}\right)\right)^{1/\beta_{n,p}}. \tag{11}$$

Subscripts in (8)–(10) stand for the corresponding recombination/generation mechanisms, namely, SHR and AU respectively mean Shockley–Hall–Read and Auger recombination/generation, while II denotes impact ionization generation. The function R provides a measure of how much the semiconductor device deviates from thermodynamical equilibrium (in which case $R \equiv 0$). In particular, $R > 0$ ($R < 0$) denotes net charge recombination (generation) per unit time and unit volume. Concerning the physical models employed in our numerical computations, we will assume the carrier lifetimes $\tau_{n,p}$ as in [6], the Auger coefficients $C_{n,p}$ as in [8], the parameters $v^{\mathrm{sat}}_{n,p}$ and $\beta_{n,p}$ as in [4], and the ionization coefficients $\alpha_{n,p}$ according to Chynoweth's law [5],

$$\alpha_{n,p} = \alpha^{\infty}_{n,p} \exp(-\mathrm{E}^{\mathrm{crit}}_{n,p}/\mathrm{E}_{n,p}), \tag{12}$$

where the constants $\alpha^{\infty}_{n,p}$, $\mathrm{E}^{\mathrm{crit}}_{n,p}$ are taken as in [21] and $\mathrm{E}_{n,p}$ are the electric field components in the direction of electron and hole current densities.

Boundary conditions of mixed type will be assumed to hold for system (1)–(6); to this aim we will divide the boundary Γ into disjoint Dirichlet and Neumann segments Σ^i_{D} and Σ^k_{N}, such that $\Gamma_{\mathrm{D}} = \cup_{i=1}^{r_{\mathrm{D}}} \Sigma^i_{\mathrm{D}}$, $\Gamma_{\mathrm{N}} = \cup_{k=1}^{r_{\mathrm{N}}} \Sigma^k_{\mathrm{N}}$, and $\Gamma = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}$ (see Fig. 1). Uniqueness of the solution demands that Γ_{D} is not empty.

Dirichlet boundary data (ψ_{D}, n_{D}, p_{D}) are then assigned at each Dirichlet segment Σ^i_{D} ($i = 1, ..., r_{\mathrm{D}}$) as

$$\psi^i_{\mathrm{D}} = V^i_a + \phi^i_{bi} \quad \text{with } \phi^i_{bi} = \sinh^{-1}(C_i/2) \tag{13}$$

$$n^i_{\mathrm{D}} = (C_i + \sqrt{C_i^2 + 4})/2 \tag{14}$$

$$p^i_{\mathrm{D}} = (-C_i + \sqrt{C_i^2 + 4})/2, \tag{15}$$

where C_i is the (constant) value of doping $C(\mathbf{x})$ at the Dirichlet segment Σ^i_{D}, ϕ^i_{bi} is the so-called "built-in" potential, and V^i_a is the external applied potential. Relations (13) and (14)–(15) represent respectively *flat band approximation* for potentials and *space charge neutrality* and they define the mathematical model for the *ideal ohmic contact* Σ^i_{D} (see also [17, pp. 28–30). Concerning the Neumann boundary conditions, we will assume that the semiconductor device is "self-contained"; namely, all currents flow only throughout the ideal ohmic contacts Σ^i_{D}. This means that the normal components of both the electric field and the current densities vanish along each Neumann segment Σ^k_{N} ($k = 1, ..., r_{\mathrm{N}}$),

$$\nabla\psi \cdot \nu_k = \mathbf{J}_n \cdot \nu_k = \mathbf{J}_p \cdot \nu_k = 0, \tag{16}$$

where ν_k is the unit outward oriented normal on Σ^k_{N}.

Maxwell–Boltzmann statistics is assumed to hold for carrier concentrations

$$n = e^{\psi - \varphi_n} \tag{17}$$

$$p = e^{\varphi_p - \psi}, \tag{18}$$

where $\varphi_{n,p}$ are electron and hole quasi-Fermi potentials. By plugging (17)–(18) into (4), space charge density reads

$$\rho(\psi, n(\psi), p(\psi)) = \rho_p e^{-\psi} - \rho_n e^{\psi} + C(\mathbf{x}). \tag{19}$$

Two different choices of the dependent variables, alternative to the set (ψ, n, p), are suggested by the relations above. These are respectively the sets $(\psi, \varphi_n, \varphi_p)$ and (ψ, ρ_n, ρ_p), where $\rho_{n,p}$ are usually known as *Slotboom* variables [18] and are related to (n, p) by

$$\rho_n = e^{-\varphi_n} = ne^{-\psi} \tag{20}$$

$$\rho_p = e^{\varphi_p} = pe^{\psi}. \tag{21}$$

We will discuss in the next two sections how the new variables can be coupled with (ψ, n, p) for an efficient abstract and numerical solution of system (1)–(3).

3. GUMMEL'S MAP FOR THE ITERATIVE SOLUTION OF SEMICONDUCTOR DEVICE EQUATIONS

In this section we will recall a block nonlinear procedure for the iterative solution of Eqs. (1)–(3) which is long established in semiconductor device modeling and has proved to work quite effectively in most of the numerical experiments performed so far. The algorithm, commonly known as *Gummel's map* [12], forces a decoupling among the equations, thus leading to the successive solution of one *nonlinear* Poisson problem and two *linearized* convection-diffusion equations.

Gummel's map reads as:

given $(\psi^k, n^k, p^k) \equiv \mathbf{w}^k, \quad k = 1, 2, \ldots,$

solve

$$\begin{aligned} -\Delta\psi^{k+1} &= \rho(\mathbf{x}, \psi^{k+1}) = \\ &(\rho_p^k e^{-\psi^{k+1}} - \rho_n^k e^{\psi^{k+1}} + C(\mathbf{x})) = \mathcal{H}(\psi^{k+1}) \end{aligned} \tag{22}$$

(by a suitable linearization method, e.g., the Newton method. This introduces an inner loop)

Then set $\mathbf{w}^{k+1/2} \equiv (\psi^{k+1}, n^k, p^k)$

and solve

$$-\operatorname{div} \mathbf{J}_n(n^{k+1}, \psi^{k+1}) = -R(\psi^{k+1}, n^k, p^k) \tag{23}$$

$$\operatorname{div} \mathbf{J}_p(p^{k+1}, \psi^{k+1}) = -R(\psi^{k+1}, n^k, p^k) \tag{24}$$

Now update $\mathbf{w}^{k+1} \equiv (\psi^{k+1}, n^{k+1}, p^{k+1})$

and repeat until convergence

We will refer in the following to this iterative procedure on k as the *outer loop* of the global algorithm; Gummel's map is precisely the one that furnishes $\mathbf{w}^{k+1}$ in terms of $\mathbf{w}^k$.

As can be seen, the intermediate nonlinear step (22) provides a new electric potential ψ^{k+1} which is plugged into the two convection–diffusion equations (23) and (24) to be solved for the carrier concentration n^{k+1} and p^{k+1}, respectively; note that the gradient of ψ^{k+1} is employed to evaluate the carrier mobilities (11), the current densities (5) and (6), and the ionization rates (12) which are eventually used to make up the impact ionization generation term (10). The procedure is stopped as soon as $\|\mathbf{w}^{k+1} - \mathbf{w}^k\|$ (for a suitable norm) falls below a fixed tolerance.

Among the few results existing in literature on the convergence properties of this map we will make reference in the following to a couple of papers by Kerkhoven [9, 10], where both abstract and spectral analyses of the decoupling algorithm are performed in the case of vanishing recombination/generation mechanisms, i.e., $R \equiv 0$.

A first goal of our work is to generalize the Gummel's map so that it is able to converge in those situations where *all* the recombination/generation terms (8)–(10) are being properly considered. In this case the decoupling algorithm is shown to be inadequate as the applied "biasing potentials" come close to the *breakdown* voltage V_{BR}, where an unbounded increase of the current densities and of the carrier concentrations takes place in the physical device as the result of a corresponding growth of the impact ionization generation term R_{II}.

The mathematical counterpart of this physical picture is the strong positive feedback due to the coupling between the source term and the left-hand side of the current continuity equations (2)–(3). It is indeed possible to show that the exact solutions of these latter may become as large as infinity for some particular values of the ionization coefficients α_n, α_p, which in turn depend on the local electric field strength $|\nabla\psi(\mathbf{x})|$ according to (12).

There are two main approaches for tackling the problem of the quick worsening of convergence rate of Gummel's map at breakdown conditions.

First, one has to build suitable variants of the standard decoupling algorithm characterized by stronger convergence properties; a typical procedure is to solve two *nonlinear* convection–diffusion problems at each global iteration, usually by means of the Newton's method (cf. [17, Chap. 7]). A second possibility consists in using vectorial acceleration methods that produce a superlinear asymptotic rate of convergence without excessively increasing the CPU time and the memory resources [11].

Alternatively, one can switch to a numerical method that solves the whole coupled system (1)–(3) by nonlinear iterations (e.g., the Newton–Raphson method); the latter approach, however, has several drawbacks, primarily, its very large computational costs. Furthermore, the Jacobian matrix is in general neither symmetric nor positive definite, in contrast with what happens for the simple Gummel's algorithm where the coefficient matrices for Poisson and convection–diffusion problems are respectively a symmetric positive definite matrix and two M-matrices.

We will discuss the analysis of our variant of Gummel's map based on the use of BI-CGSTAB method [20] in the concluding section of this paper.

4. THE DOMAIN DECOMPOSITION METHOD FOR THE ITERATIVE SOLUTION OF SEMICONDUCTOR DEVICE EQUATIONS

In this section we will develop some motivations based on physical grounds for employing the *domain decomposition method* (DDM) as an efficient strategy to solve (22)–(24) at each iteration. To this aim, we will briefly describe the qualitative behaviour of the solutions of a typical test problem in semiconductor device modeling, namely, a one-dimensional *p-n* diode with an abrupt doping profile $C(x)$.

Let the device domain be the interval $\Omega \equiv [0, L]$ and let the doping $C(x)$ in (1) be the piecewise constant function

$$C(x) = \begin{cases} -|N_a|, & 0 \le x \le \dfrac{L}{2} \\ N_d, & \dfrac{L}{2} < x \le L. \end{cases} \tag{25}$$

Supposing $|N_a|, N_d \gg 1$, the Dirichlet boundary conditions read in this case

$$\begin{aligned} &\phi_b(0) = -\log|N_a|, \quad \phi_b(L) = \log N_d \\ &n(0) = \frac{1}{|N_a|}, \quad n(L) = N_d, \\ &p(0) = |N_a|, \quad p(L) = \frac{1}{N_d}, \\ &\psi(0) = V_{ap} + \phi_b(0), \quad \psi(L) = V_{an} + \phi_b(L). \end{aligned} \tag{26}$$

Depending on the sign of the voltage difference $V_a = V_{ap} - V_{an}$ applied across the device, we will mean respectively *forward bias* if $V_a > 0$ and *reverse bias* if $V_a < 0$; the case $V_a \equiv 0$ defines the condition of thermodynamical equilibrium. A qualitative behaviour of the solutions of (1)–(3) may be deduced by setting $R \equiv 0$ and working out the regional approximations typical of semiconductor device physics (see, e.g., [19]), leading to the graphs shown in Fig. 2.

As can be seen, all of the unknowns exhibit sharp interior layers in the neighbourhood of the *p-n* junction ($x = L/2$); the potential layer is due to the jump of the doping profile, whilst the layer in the carrier distributions is due to the convection-dominated nature of the drift-diffusion equations in the narrow *depleted region* across $x = L/2$.

The extension of this simple physical picture to the study of more realistic device geometries is straightforward, since we exactly know *a priori* the position of each single interior layer, which typically occurs at the *p-n* junctions, i.e., where the doping profile $C(x)$ attains its steepest gradients. This makes the choice of a suitable partition of the device domain an easy matter, so that DDM may be profitably employed for an efficient solution of the semiconductor device equations.

4.1. A Neumann–Neumann Domain Decomposition Algorithm

The Neumann–Neumann (NN for short) domain decomposition method allows the solution of self-adjoint boundary-value problems in regions partitioned into subdomains through an iterative procedure among subdomains. At each step the updating is achieved by solving independent subproblems with Dirichlet conditions; this phase is followed by a correction yielding subproblems with Neumann conditions at the interfaces among the subdomains.

The method is described in [1], where an acceleration procedure relying on conjugate gradient (CG) iterations for the so-called Steklov–Poincare's interface operator is also proposed.

In order to describe the NN algorithm in detail, it is useful to introduce some notations; in this section Ω will be an open bounded set of $\mathbf{R}^2$ with a Lipschitz boundary Γ, partitioned in subdomains Ω_i with boundaries Γ_i as illustrated in Fig. 3.

Let us introduce the boundaries

$$\begin{aligned} &\Gamma = \Gamma_D \cup \Gamma_N \\ &\Gamma_i^D = \Gamma_i \cap \Gamma_D \\ &\gamma_{ij} = \gamma_{ji} = \Gamma_i \cap \Gamma_j \\ &S = \bigcup_{i,j} \gamma_{ij} \\ &S_i = \bigcup_j \gamma_{ij}, \end{aligned}$$

where S is the "skeleton" made by all the subdomain interfaces and the spaces

$$\begin{aligned} H^1(\Omega) &= \{v \in L^2(\Omega), \nabla v \in L^2(\Omega)\} \\ V &= \{v \in H^1(\Omega), v = 0 \text{ on } \Gamma_D\} \\ Y &= \{v|_S, v \in V\} \\ V_i &= \{v \in H^1(\Omega_i), v = 0 \text{ on } \Gamma_i^D\} \\ V_{0,i} &= \{v \in H^1(\Omega_i), v = 0 \text{ on } \Gamma_i^D \cup S_i\}. \end{aligned}$$

We then consider the following abstract variational problem

$$\textit{find } u \in V \textit{ such that} \qquad (\mathcal{P})$$
$$a(u, v) = (f, v) \quad \forall v \in V,$$

where $a(u, v) : V \times V \to \mathbf{R}$ is a symmetric, continuous, and coercive bilinear form and (f, v) is the usual inner product in $L^2(\Omega)$ for any given function $f \in L^2(\Omega)$.

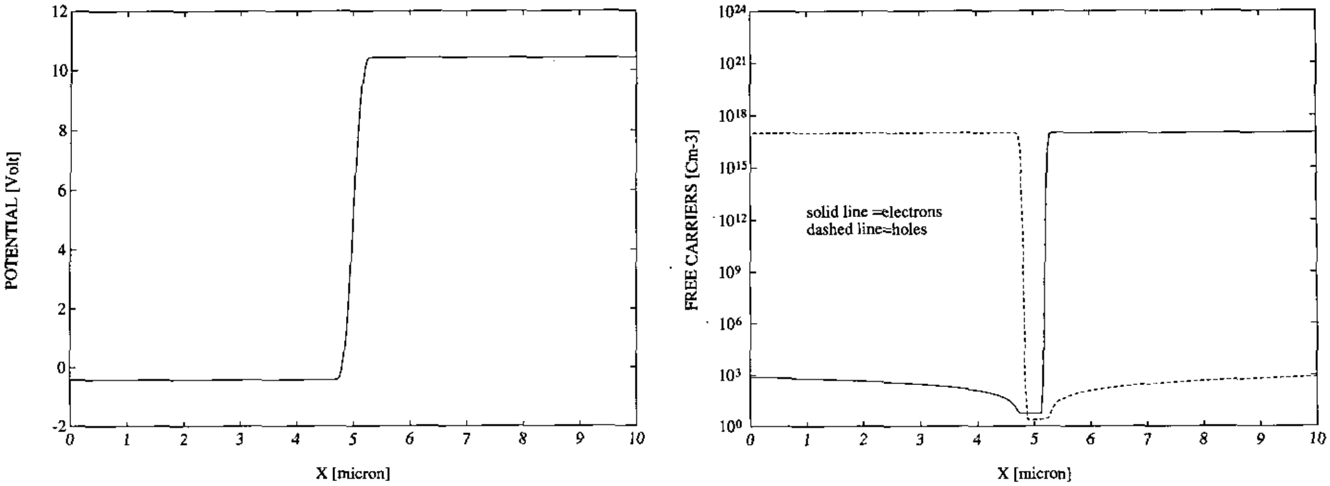

FIG. 2. *Potential and carrier distributions under reverse bias.*

It is also possible to define a suitable inner product on the spaces V_i introduced before, which will be denoted in the following by $\tilde{a}_i(u, v)$ for any $u, v \in V_i$. Moreover, we let $Tr_i^{-1}(\lambda)$ any element $z \in V_i$ whose trace on S_i equals λ. Denoting by $a_i(u, v)$ the restriction of $a(u, v)$ to each subdomain Ω_i, problem $(\mathcal{P})$ takes the form

$$\text{find } u \in V \text{ such that} \qquad (\mathcal{P}')$$
$$\sum_{i=1}^{N} a_i(u, v) = (f, v) \quad \forall v \in V,$$

where N is the total number of subdomains Ω_i.

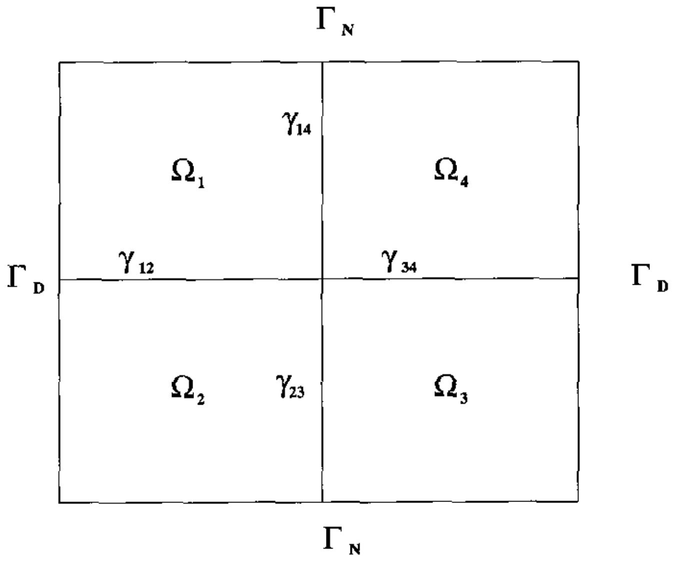

FIG. 3. Definition of the subdomains and boundaries.

We report in the following the main steps of the NN algorithm [1] that solves problem $(\mathcal{P}')$.

Initialization

For λ given in Y solve the Dirichlet problems in any Ω_i

$$u_{i,0} \in V_i$$
$$a_i(u_{i,0}, v) = (f, v) \quad \forall v \in V_{0,i}$$
$$u_{i,0} = \lambda \quad \text{on } S_i$$

Compute $\mathcal{L}_i(v) = \Sigma_j(a_j(u_j, \mathrm{Tr}_j^{-1}(\alpha_i v)) - (f, \mathrm{Tr}_j^{-1}(\alpha_i v))$. Solve the Neumann problems

$$\varphi_{i,0} \in V_i$$
$$\tilde{a}_i(\varphi_{i,0}, v) = \mathcal{L}_i(v) \quad \forall v \in V_i$$

Compute

$$d_0 = \sum_i \tilde{a}_i(\varphi_{i,0}, \varphi_{i,0}) = \sum_i \mathcal{L}_i(\varphi_{i,0})$$
$$w_{i,0} = \varphi_{i,0}, \quad R_{i,0}(v) = \mathcal{L}_i(v)$$

Loop on n: For n = 0 Until Convergence Do

Compute $\lambda = \Sigma_j \alpha_j w_{j,n}$
Solve the Dirichlet problems

$$z_i \in V_i$$
$$a_i(z_i, v) = 0 \quad \forall v \in V_{0,i}$$
$$z_i = \lambda \quad \text{on } S_i$$

Compute $\mathcal{L}_i(v) = \Sigma_j(a_j(z_j, \mathrm{Tr}_j^{-1}(\alpha_i v)) \quad \forall v \in V_i$

Solve the Neumann problems

$$\psi_i \in V_i$$
$$\tilde{a}_i(\psi_i, v) = \mathcal{L}_i(v) \quad \forall v \in V_i$$

Descent Step

Set

$$r_n = \sum_j \tilde{a}_j(w_{j,n}, \psi_j) = \sum_j \mathcal{L}_j(w_{j,n})$$
$$\theta_n = d_n / r_n$$
$$u_{i,n+1} = u_{i,n} - \theta_n z_i$$
$$\varphi_{i,n+1} = \varphi_{i,n} - \theta_n \psi_i$$
$$R_{i,n+1}(v) = R_{i,n}(v) - \theta_n \mathcal{L}_i(v)$$

Computation of the New Descent Direction

Compute

$$d_{n+1} = \sum_j \tilde{a}_j(\varphi_{j,n+1}, \varphi_{j,n+1}) = \sum_j R_{j,n+1}(\varphi_{j,n+1})$$

or, alternatively,

$$d_{n+1} = d_n - \theta_n \sum_j \mathcal{L}_j(\varphi_{j,n} + \varphi_{j,n+1})$$

If

$$\sqrt{d_{n+1}/d_0} < 10^{-5}$$

then stop
else let

$$w_{i,n+1} = \varphi_{i,n+1} + \frac{d_{n+1}}{d_n} w_{i,n}.$$

As for the Neumann problems, it must be noted that the best choice would be to set $\tilde{a}_i(u, v) \equiv a_i(u, v)$, since in this way the Neumann step is the exact inverse of the Dirichlet step with respect to the Steklov–Poincare's operator. By so doing, the condition number of the algebraic problem at the interfaces turns out to be very small.

The algorithm above may be easily implemented in a way suitable for a finite element method (FEM), where the spaces V_i are approximated by finite dimensional spaces and the trace operators are defined at each interface node by the relationships

$$\alpha_i(v)(P_k) = \frac{a_i(\phi_k, \phi_k)}{\sum_j a_j(\phi_k, \phi_k)} v(P_k),$$

where ϕ_k is the weighting function at the node P_k, i.e., the finite element test function. Going back to the original problem at the continuous level, we may say that, starting from an initial guess u^0, the NN method generates a sequence of approximants $\{u^n\}$ of the exact solution; such functions are continuous over the device domain Ω but, in general, will have discontinuous derivatives just at the interfaces. The Neumann step does the job of smoothing these irregularities (*smoothing phase*) by distributing the jumps at each interface of the electric field and of the discrete current densities all over the subdomains.

4.2. A Discussion on the Convenient Choice of the Dependent Variables

There are two kinds of problems that need to be faced at each step of the outer loop. The former concerns a Poisson equation arising from a suitable linearization of (22) and the latter from the current continuity equations (23) and (24) which are nothing but advection–diffusion equations. A *naive* application of the NN algorithm described above for the solution of (22), (23), and (24) presents several drawbacks. Each of them will be discussed and motivated by means of some examples and a concluding possible approach, which has proved to work quite effectively in our numerical experiments, will then be proposed.

As for Eq. (22), the simplest approach consists in using Picard iterations. Denoting by m the inner iteration counter, the generic stage of the algorithm may be written as

$$-\psi''^{k+1}_{m+1} = \mathcal{H}(\psi^{k+1}_m), \quad m = 0, 1, \ldots. \tag{27}$$

Taking $m = 0$ in (27) and $\psi^{k+1}_0 \equiv \psi^k$, a single Picard iteration is obtained to update ψ^{k+1}. This amounts to solve the following linear problem

$$\begin{aligned} -\psi''^{k+1}(x) &= (p^k - n^k + C(x)) = \rho^k(x), \quad 0 < x < L, \\ \psi^{k+1}(0) &= \psi_0, \quad \psi^{k+1}(L) = \psi_L. \end{aligned} \tag{28}$$

The solution of (28) by a NN procedure does not present any difficulty; indeed, the numerical experiments show that the rate of convergence is independent of the initial guess (even if it is very far from the exact solution), of the source term ρ^k and of the mesh size h, being only a function of the number of interfaces. This agrees with the theoretical convergence properties of the CG method if we point out that we are actually solving a linear system coming from the discretization of the Steklov–Poincare's operator and whose order equates the number of subdomain interfaces, since a unique mesh node is defined at each interface.

A serious drawback of the linearization above is the strong dependence of the solution on the right-hand side, i.e., of ψ^{k+1} on ρ^k. In particular, when the space charge region, i.e., the interval where $\rho^k(x)$ remarkably differs from zero, attains varia-

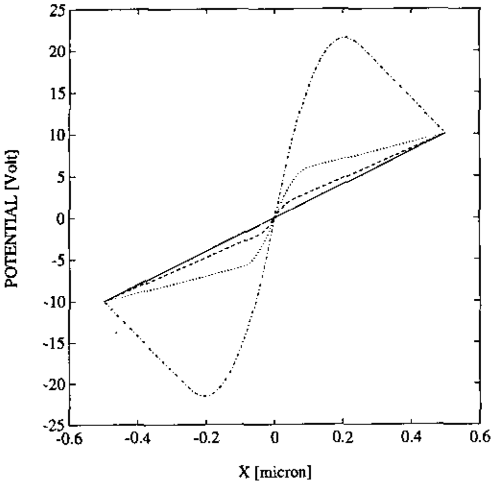

FIG. 4. Solution of the linear Poisson problem as a function of ρ^k (through x_n).

tions of the order of h during the numerical solution phase, the corresponding variations of electric potential ψ^{k+1} in the neutral region, i.e., the zone where $\rho^k(x)$ is $\simeq 0$, are no longer acceptable. In fact, the potential ψ^{k+1} assumes unrealistic slopes which may strongly compromise the global stability of the iterative map when the output from the Poisson equation (28) is plugged into the next two drift-diffusion equations.

An explicit example of this undesirable phenomenon may be given by solving the boundary-value problem (28) over the symmetric domain $\Omega \equiv [-L/2, L/2]$ with

$$\rho^k(x) = \begin{cases} 0, & -\dfrac{L}{2} \leq x \leq -x_n \equiv \Omega_1, \\ -N, & -x_n < x \leq 0 \equiv \Omega_2, \\ +N, & 0 < x \leq x_n \equiv \Omega_3, \\ 0, & x_n < x \leq \dfrac{L}{2} \equiv \Omega_4, \end{cases}$$

where $N > 0$ and $\psi^{k+1}(L/2) = -\psi^{k+1}(-L/2) = \overline{\psi}$. The right-hand side is consistent with the so-called *depletion approximation* (see, e.g., [19]), which is the usual one in semiconductor device physics.

The corresponding $\psi(x)$ is respectively linear in Ω_1 and Ω_4 and parabolic in $\Omega_2 \cup \Omega_3$. By imposing continuity of $\psi^{k+1}(x)$ and $\psi'^{k+1}(x)$ at $x = \pm x_n$ and the boundary data, the slope A of the solution in the regions $|x| > x_n$ is readily obtained as

$$A = \frac{2\overline{\psi} - N x_n^2}{L}.$$

Electric potential in the case $N = 1000$, $L = 1$, $\overline{\psi} = 10$, $x_n = \{0, L/20, L/10, L/4\}$ is reported in Fig. 4, clearly showing its nonlinear sensitivity to the increasing values of the parameter x_n. It is worth noting that this effect is strictly connected to the peculiar nature of the space charge density ρ^k, namely, its implicit nonlinear dependence on potential ψ^{k+1}, and it must not be ascribed to the NN method. Indeed, plugging any of the potentials shown in Fig. 4 into the drift-diffusion equations produces completely unphysical carrier distributions that may become excessively over/under estimated according to the corresponding value of x_n.

A correct and stable formulation of the global iterative map may be obtained by taking into account the explicit dependence of ρ on ψ in the Poisson problem, as it will be pointed out later on.

As for current continuity equations, our first effort has been to provide a self-adjoint equivalent formulation, as demanded by the NN method.

We point out that, for each k, Eq. (23) yields (in its one-dimensional version)

$$\begin{aligned} -J_n' &\equiv -(\mu_n(n' - n\psi'))' = -R, \quad 0 < x < L, \\ n(0) &= n_0, \quad n(L) = n_L, \end{aligned} \tag{29}$$

where the dependence on the iteration counter $(k + 1)$ for both n and ψ is understood. Since the two drift-diffusion equations are of the same nature, we will restrict our analysis only to the electron current continuity law.

We operate the change of variables (17)–(18) on (29) which introduces the new set of unknowns $(\psi, \varphi_n, \varphi_p)$. Owing to (17), electron current density may be written as

$$J_n = -\mu_n e^{\psi - \varphi_n} \varphi_n',$$

so that we can replace (29) by

$$\begin{aligned} J_n' &\equiv -(\mu_n e^{\psi - \varphi_n} \varphi_n')' = R, \quad 0 < x < L, \\ \varphi_n(0) &= V_{ap}, \quad \varphi_n(L) = V_{an}, \end{aligned} \tag{30}$$

where the boundary data are the values of the voltages applied to each ohmic contact.

Despite the fact that the functions ψ, μ_n, and R are known, problem (30) turns out to be *nonlinear* in φ_n, thus making the Neumann–Neumann algorithm described in Section 4.1 not directly applicable to it. In order to overcome this drawback one may "freeze" the coefficient $e^{\psi - \varphi_n}$ by setting it equal to the value of electron concentration computed at the previous global iteration step, thus leading to solution of the elliptic boundary value problem

$$\begin{aligned} -(\mu_n n^k \varphi_n')' &= R, \quad 0 < x < L, \\ \varphi_n(0) &= V_{ap}, \quad \varphi_n(L) = V_{an}. \end{aligned} \tag{31}$$

Approximation of (31) by standard Galerkin piecewise linear finite elements yields a symmetric positive definite linear system, for which the NN subdomain iterative procedure is ideally well suited. Unfortunately, on this problem the NN method is shown to be unsuccessful. Precisely, the corresponding solution turns out not to be acceptable in the neutral p-side of the diode, where the electron quasi-Fermi level presents a cusp right at the interface position placed in that zone, whilst it is correct in the neighbourhood of the interface in the n-side of the device. This behaviour is due to the extremely large variability of the coefficient $a(x) \equiv \mu_n n^k$, whose values run through the range $10^{-10} \div 10^{10}$ and attain their minimum just in the neutral p-side of the pn junction. This reflects negatively on the residual evaluations required by the CG method, which are indeed proportional to the product $a(x)\varphi'_{nh}$, i.e., to the discrete current density. These residuals may be very small in practice, even with a great jump of φ'_{nh}, because of the smallness of the coefficient $a(x)$.

Moreover, we remark that, even if the obtained solution is reasonably close to the expected one, its successive plugging into the Poisson equation (22) produces a completely unphysical space charge density ρ, which, as we have already pointed out, is dangerous for the stability of the global map. This is due to the exponential dependence of ρ on φ_n (17); it is indeed possible to express this dependence by the following sensitivity relation which holds true for all $x \in (0, L)$,

$$\frac{\delta n}{n} = \frac{1}{V_{\text{th}}} \varphi_n \frac{\delta \varphi_n}{\varphi_n},$$

where δ means "variation," V_{th} is the thermal voltage (at room temperature it equals 25 mV), and φ_n may assume in practical computations values of the order of tens or hundreds of volts. A second way to get a self-adjoint form is based on the use of the Slotboom variable ρ_n. Owing to (20), the current density J_n in this case becomes

$$J_n = \mu_n e^{\psi} \rho'_n$$

and the related differential problem may be formulated as

$$\begin{aligned} -J'_n = -(\mu_n e^{\psi} \rho'_n) &= -R, \quad 0 < x < L, \\ \rho_n(0) = n(0)e^{-\psi_0}, &\quad \rho_n(L) = n(L)e^{-\psi_L}, \end{aligned} \tag{32}$$

which can be recognized as a linear elliptic boundary value problem, since ψ is given at the $(k + 1)$th stage. Due to the presence of the exponential factor e^{ψ}, the actual solution of (32) turns out to be feasible only for small values of the Dirichlet data ψ_0 and ψ_L, i.e., close to the thermodynamical equilibrium (equivalently, $V_{ap} = V_{an} = 0$ V). In all other biasing conditions, underflow/overflow problems due to the enormous range of variation of the coefficient e^{ψ} make the ρ_n-variable formulation (32) unsuitable for practical computations. This negative feature is very well stressed in semiconductor device literature (see, e.g., [17, 13], and for the computational aspects [2]).

The third (and successful!) way exploits the Scharfetter–Gummel [16] approximation of the current density J_n which is assumed to be *piecewise constant* over the device domain. It is indeed possible to set up two perfectly equivalent numerical schemes based respectively on the n and ρ_n variables. This means that we can formally implement the NN method on a discretized version of problem (32) as will be pointed out in the next section. However, when computing the residuals (involving the jumps of J_n at subdomain interfaces), we will make use of the equivalent nonexponential expression $J_n = \mu_n(n' - n\psi')$.

We wish to focus now on those steps in the Neumann–Neumann algorithm described in detail in Section 4.1, where the change of variable $\rho_n \to n$ comes out explicitly. Each single Dirichlet and Neumann problem may be directly solved in terms of electron concentration, since the change of variable simply acts as a scaling of the stiffness matrix.

The evaluations of the $\mathscr{L}_i(v)$ terms is analogously transparent while possible numerical difficulties may arise in the computation of the d_n and r_n residuals, since these latter are weighted by the nodal values of the φ-functions that are the solutions of the single Neumann problem and act in this context as Slotboom variables. (We are referring to the notations introduced in Section 4.1).

A simple example will be considered now in order to better explain what happens in the numerical computations. Let us suppose that the domain $\Omega \equiv [0, L]$ is partitioned into three subdomains Ω_i, where the interfaces have been respectively denoted by γ_{12} and γ_{23}. The computation of the d_0 term gives

$$\begin{aligned} d_0 = \sum_i \mathscr{L}_i(\varphi_{i,0}) &= \mathscr{L}_1(v)\varphi_1(\gamma_{12}) + \mathscr{L}_2^1(v)\varphi_2(\gamma_{12}) \\ &\quad + \mathscr{L}_2^2(v)\varphi_2(\gamma_{23}) + \mathscr{L}_3(v)\varphi_3(\gamma_{23}) \end{aligned}$$

which, by invoking the transformation $\varphi = \tilde{\varphi} e^{-\psi}$, may be expressed in terms of the variable $\tilde{\varphi}$ (the solution of the Neumann problem in the drift-diffusion form) as

$$\begin{aligned} d_0 = \mathscr{L}_1(v)e^{-\psi(\gamma_{12})}\tilde{\varphi}_1(\gamma_{12}) &+ \mathscr{L}_2^1(v)e^{-\psi(\gamma_{12})}\tilde{\varphi}_2(\gamma_{12}) \\ &+ \mathscr{L}_2^2(v)e^{-\psi(\gamma_{23})}\tilde{\varphi}_2(\gamma_{23}) + \mathscr{L}_3(v)e^{-\psi(\gamma_{23})}\tilde{\varphi}_3(\gamma_{23}). \end{aligned}$$

A similar expression may be obtained for the generic term d_n and for the residual r_n, where in this case the φ-functions must be substituted by the w-functions (and, similarly, $\tilde{\varphi}$ by $\tilde{w}$); it is clear that the exponential terms in the relation above are a potential source of trouble. In any case, since the *ratio* between residuals r_n and d_n must be computed, in the two particular cases in which either $\psi(x) \equiv$ const over the device domain Ω (*pure diffusion problem*), or there is one single interface γ, these exponentials are common to both numerator and denominator in the formula above so that they cancel each other out and their computation may thus be avoided.

Actually, the Neumann–Neumann iterative algorithm tested in the case of more than two subdomains has always exhibited problems similar to those already described for the choice of set (ψ, φ_n, φ_p). In fact, the values of the concentrations tend to stagnate in the minority regions, where the solutions present a cusp as in the case of the quasi-Fermi variables discussed before.

4.3. The Solution Algorithm

On the grounds of the conclusions drawn in Section 4.1 and on the basis of the numerical experiments performed, we propose the following solution algorithm.

Outer iteration on k

- Costruct Newton iterates $\{\psi_m^k\}$ for the nonlinear elliptic boundary value problem

$$\begin{aligned} -\psi''^k &= (\rho_p^{k-1}e^{-\psi^k} - \rho_n^{k-1}e^{\psi^k} + C(x)) \\ &= \rho(x, \psi^k(x)), \quad 0 < x < L, \\ \psi^k(0) &= \psi_0, \quad \psi^k(L) = \psi_L, \end{aligned} \tag{33}$$

  such that

$$\lim_{m\to\infty} \psi_m^k = \psi^k$$

- for each m (*inner loop*) solve the linearized problem by the NN multidomain method.
- Solve the linear drift-diffusion problem

$$\begin{aligned} -(\mu_n^k(n'^k - n^k\psi'^k))' &= -R(\psi^k, n^{k-1}, p^{k-1}), \quad 0 < x < L, \\ n^k(0) &= n_0, \quad n^k(L) = n_L. \end{aligned} \tag{34}$$

Convergence of the approximants $\{\psi_m^k\}$ may be proved by associating a minimum variational problem to Eq. (33) (see, e.g., [15]). The functional minimized by the solution of the quasilinear Poisson problem is

$$\mathcal{J}(\psi) = \int_0^L \left(\frac{1}{2}(\psi')^2 + \rho_p e^{-\psi} + \rho_n e^{\psi} - C\psi\right) dx + \mathcal{K},$$

where $\mathcal{K}$ is a constant and the dependence on the iteration counter k is understood.

Concerning the linear convection–diffusion problem (34) (and its twin for the p variable), it is necessary to modify the standard Neumann–Neumann method in such a way as to solve a series of differential problems having *just one* interface each and, consequently, only one degree of freedom. This kind of subproblem may be in fact easily solved in a *single* iteration by the original domain decomposition procedure. Actually, the starting coupled problem partitioned in N subdomains is *recursively* led to $N/2$ subdomains problems by an algorithm that systematically eliminates the interfaces until *single-interface* problems are reached and eventually solved. This strategy is illustrated in Fig. 5 in the case of $N = 4$.

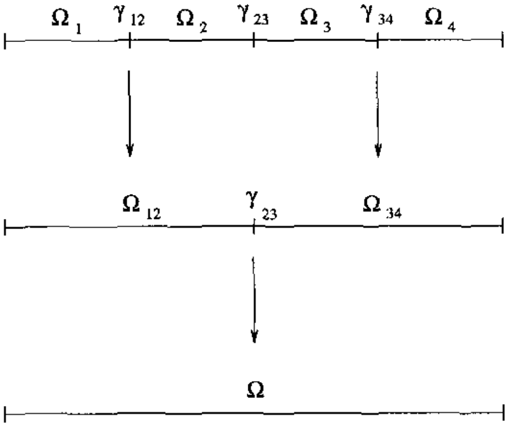

FIG. 5. Recursive partitioning algorithm to reduce to single interface problems for n^k by the NN multidomain method and proceed similarly for p^k.

When extended to either two- or three-dimensional configurations, the above procedure needs to be modified accordingly. The remarkable feature of being driven by a single parameter at the interface in one-dimensional devices is unfortunately lost when considering higher dimensional semiconductors. In the latter case, the number of free parameters does coincide with the overall number of finite element nodes lying at subdomain interfaces. Nonetheless, the outer iterative procedure yielded by the Neumann–Neumann domain decomposition algorithm can be generalized straightforwardly and the conjugate gradient approach can still take care of selecting optimal acceleration parameters. In the specialized literature it is extensively demonstrated that this domain decomposition method can provide a convergence rate which does not depend on the finite element mesh size (and, therefore, on the number of nodes on interfaces). However, its efficiency can deteriorate at a linear rate, as the number of subdomains gets larger. The latter circumstance does not necessarily occur for the problem at hand, which suggests a coarse subdomain partition that fits the ''regional'' behaviour of the solutions, as already stressed in the introductory part of Section 4. On the other hand, if one wishes to resort to fine partitions in order to exploit machine parallelism, as every other domain decomposition method, the Neumann–Neumann procedure could be modified through the introduction of auxiliary coarse-level partitions that make the overall method almost insensitive to the number of subdomains.

5. FINITE ELEMENT DISCRETIZATION OF THE DRIFT-DIFFUSION SYSTEM

In this section we deal with the spatial discretization of the three differential subproblems that are to be solved in sequence at the kth step of Gummel's decoupled algorithm. Again we confine our analysis to the one-dimensional case. The explicit dependence of the variables on the iteration counter will be omitted for the sake of clarity in the following.

We first introduce some useful notations; let $\mathcal{T}_h \equiv \{T_j\}_{j=1}^{N_h+1}$ be a nonuniform partition of the device domain $\Omega \equiv (0, L)$, where the element $T_j \equiv (x_{j-1}, x_j)$ has size $h_j = x_j - x_{j-1}$, and we denote by $h = \max h_j$, $\sigma_j \equiv T_j \cup T_{j+1}$ the support of the jth shape function and by $\mathbf{P}_k$ the space of polynomials of degree $\leq k$.

Let us now consider the finite element approximation of the quasilinear Poisson equation (33); this latter is first written as

$$-\psi''(x) - \rho(x, \psi(x)) \equiv \mathcal{F}(\psi) = 0 \tag{35}$$

and then solved iteratively by Newton's method (*inner loop*), according to the procedure

$$\begin{aligned} -\delta\psi''_{m+1} + (p_m + n_m)\,\delta\psi_{m+1} &= -\mathcal{F}(\psi_m) \\ \psi_{m+1} &= \psi_m + \delta\psi_{m+1}, \end{aligned} \tag{36}$$

where $\delta\psi(0) = \delta\psi(L) = 0$ and m is the inner iteration counter.

The weak formulation of (36) reads

$$\begin{aligned} &\text{find } \delta\psi_{m+1} \in H_0^1(\Omega) \quad \text{such that } \forall v \in H_0^1(\Omega) \\ &\int_\Omega \delta\psi'_{m+1} v'\,dx + \int_\Omega (p_m + n_m)\,\delta\psi_{m+1} v\,dx \\ &\quad = -\left(\int_\Omega \psi'_m v'\,dx - \int_\Omega \rho(x, \psi_m) v\,dx\right). \end{aligned} \tag{37}$$

Its approximation by a pure Galerkin piecewise linear finite element scheme leads to the linear system

$$(A + D)\,\delta\mathbf{w}_{m+1} = -A\mathbf{w}_m + \mathbf{b}_m, \tag{38}$$

where we have respectively denoted by $\delta\mathbf{w}_{m+1}$ the solution vector, by $\mathbf{w}_m$ the nodal potentials at the previous step of Newton's method, by A the usual Laplacian stiffness matrix, by D the positive diagonal matrix obtained by evaluating the second integral in the left-hand side of (37) by a trapezoidal quadrature, and by $\mathbf{b}_m$ the vector obtained by approximating the second integral in the right-hand side by the same rule. Use of trapezoidal quadrature introduces a strong positive diagonal perturbation in the Laplacian stiffness matrix so that the resulting matrix $M \equiv (A + D)$ turns out to be symmetric strongly diagonally dominant (and thus positive definite) and then the linear system (38) may be successfully solved by the NN method.

We address now the second item in Gummel's decoupled algorithm, namely, the numerical solution of the two linear drift-diffusion equations. The heavily convection-dominated nature of these latter demands a suitable choice of the finite element basis functions in order to get a stable and accurate approximation of the carrier distributions. Such a goal is fully achieved by assuming as trial space for electron (respectively hole) concentration the set of the functions deduced by the Scharfetter–Gummel (SG) approximation [16] of the current density field J_h which is assumed to be piecewise constant (i.e., $J_h|_{T_j} = \text{const}\ \forall j$). The divergence-free assumption on J_h gives rise to a stiffness matrix which is an M-matrix with a strictly positive inverse (cf. [2, 3, 14]). This property is quite desirable in semiconductor device simulation, since it guarantees positiveness of the free carrier distributions obtained by solving the discretized current continuity equations.

The resulting finite element scheme can be regarded as a Petrov–Galerkin one and it is characterized by a $O(h^2)$ convergence rate. Besides, it may be proved that it is equivalent to an *optimal upwinding* finite element method [2, 14].

6. NUMERICAL RESULTS

In this concluding section we discuss some numerical examples relative to the simulation of a one-dimensional p-n diode with an abrupt doping profile $C(x)$ and subject to some different values of the biasing potential $V_a = V_{ap} - V_{an}$. The choice of such a model problem makes the algorithmic effort relatively easier while retaining a wide range of generality, since the basic physical properties of the solutions in more complex and realistic geometries are well reproduced by the one-dimensional approximation.

The goal of the numerical experiments is the study of the p-n diode current-voltage (I-V) characteristic. This latter is the curve describing the variation of current measured at the device's ohmic contacts as a function of the applied bias voltage V_a and turns out to be all one needs to know in order to characterize the device's performances. Denoting by W the cross section of the real three-dimensional device, current is defined as $I \equiv (J_n + J_p)W$.

Three zones may be distinguished on the (I-V) curve in Fig. 6, where the values $V_a = (V_{FW}, V_{BR})$ respectively denote the "built-in" and the breakdown voltages. Accordingly, it turns out that solutions of problem (22)–(24) belonging to the same zone exhibit similar properties and behaviour as the parameter V_a changes with continuity. We thus only report the numerical results relative to the study of some specific working points of particular interest on the (I-V) curve.

We assume the device length $L = 10\ \mu\text{m}$ (junction at $x = L/2$) and the doping concentrations $|N_a| = N_d = 10^{17}\ \text{cm}^{-3}$, so that the builtin voltage $V_{\text{FW}} \simeq 0.84$ V and the breakdown voltage may be roughly overestimated as $|V_{\text{BR}}| \simeq 25$ V. The drift-diffusion system has then been solved, supported by the following voltage boundary conditions:

(i) $V_a = 0$ V (thermal equilibrium),
(ii) $V_a = -10$ V (moderate reverse bias),
(iii) $V_a = 0.8$ V (high forward bias),
(iv) $V_a = -24$ V (breakdown).

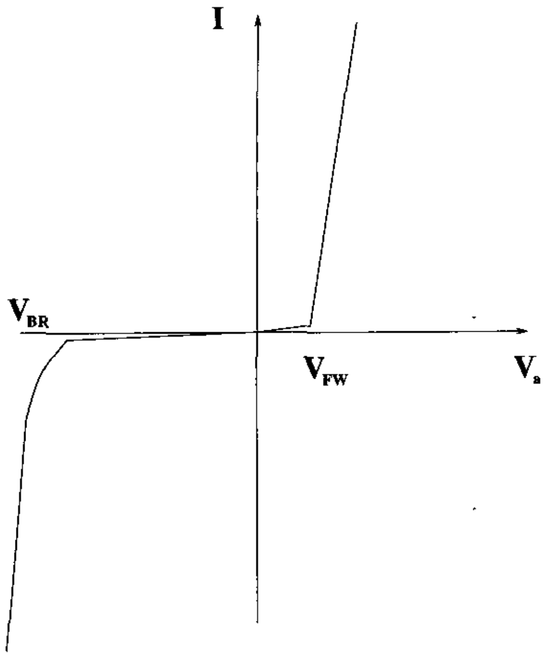

FIG. 6. *p-n* diode current–voltage characteristic.

Convergence of Gummel's decoupled algorithm has been monitored in the computer code by checking respectively the maximum absolute and relative nodal variations of the electric potential and carrier concentrations between two consecutive iterations, where the corresponding error tolerances have been both set equal to 10^{-3}. We show in Fig. 7 the numerical results relative to the cases $V_a = \{0, -10, 0.8\}$ V; the main parameters for each simulation are sketched in the headings, where IDOM is the number of subdomains, NP is the total number of internal nodes, and ITGLOB denotes the number of iterations on k needed to achieve convergence of Gummel's map. Note the highly nonuniform distribution of the mesh nodes over the device domain; the grid spacing has been taken constant within each subdomain, being quite coarse in the lateral quasi-neutral regions and much finer in the depleted zone across the junction, where the maximum variations of the solutions are expected.

As a first comment, we note that all the results show the quick convergence of the nonlinear block iterative procedure, which turns out to be reasonably independent of the applied voltage V_a. Further checks on the discrete distributions of the electric potential and carrier concentrations have also proved the accuracy of the method measured in the sup-norm. Moreover, we stress that the solution by the NN domain decomposition method of each subproblem (22)–(24) has always required a *number of iterations as low as the number of interfaces.*

From a more physical standpoint, it is interesting to outline the strong difference between the two biasing conditions $V_a > 0$ (*forward bias*) and $V_a < 0$ (*reverse bias*). In the first case the positive applied voltage allows for a large current flow throughout the device ($J_{n,p} \simeq 500$ Acm^{-2}), since the effective *potential barrier across the junction is reduced from the total* built-in voltage $V_{\rm FW}$ to the value $V_{\rm eff} \simeq V_{\rm FW} - V_a$.

On the other hand, current flow turns out to be almost negligible ($J_{n,p} \simeq 10^{-9}$ Acm^{-2}) under reverse bias, since in this case $V_{\rm eff} \simeq V_{\rm FW} + |V_a|$. It must be noted that this problem is heavily convection-dominated in the region across the junction, where the electric field attains a peak value of -400 KVcm^{-1} to be compared with a corresponding value of -2.5 KVcm^{-1} in the case $V_a = 0.8$ V. We remark on the effectiveness of the SG finite elements in reproducing the sharp interior layers resulting in the carrier distributions, as is further demonstrated in Fig. 8, showing the solutions of the same reverse-biased problem, where $R = 0$ has been assumed in the current continuity equations (23)–(24).

Note, in particular, the steep gradients of the free carrier populations which diminish from 10^{17} down to 10^{0} cm^{-3} over a region $\simeq 0.1$ μm wide; the high convective term due to the reverse biasing accounts for the above behaviour. We also remark that the thin layers in the minority carrier distributions in the forward bias case $V_a = 0.8$ V shown in Fig. 7 are only due to the choice of the logarithmic scale. This is clearly confirmed by looking at Fig. 8 (left), where the electron and hole concentrations are plotted in a linear scale. Indeed, the solution exhibits a linear variation in most of the device as a result of the very low value of the local Peclet number, $\Pi = h/(2\varepsilon)$ $(\varepsilon \equiv |\nabla\psi|^{-1}) = \mathcal{O}(10^{-2})$.

We conclude the study of the *p-n* diode characteristic by presenting some numerical results relative to the simulation of the device at the onset of *avalanche breakdown* (see, e.g., [19]). In this condition the convergence of Gummel's map quickly worsens, until it definitevely stops as V_a approaches the breakdown voltage $V_{\rm BR}$. This is the result of the strong positive feedback between the source term and the left-hand side in the current continuity equations, whose corresponding physical counterpart is the large and progressive, with respect to k, increase of current densities and carrier concentrations, which, if not properly prevented, eventually leads to the melting of the device itself due to Joule effect.

We show, respectively, in Figs. 9 and 10 the plots of carrier densities, electric field, the impact ionization generation term $R_{\rm II}$, and the Shockley–Hall–Read recombination/generation mechanism $R_{\rm SHR}$ relative to the reverse-bias voltage $V_a = -24$ V and, for a comparison, the two latter quantities relative to the case previously discussed, $V_a = -10$ V.

We must immediately outline that Gummel's iterative procedure does not converge within the maximum allowed number of iterations on k, which has been set equal to 100. Moreover, the maximum absolute and relative errors reported on the unknowns (ψ, n, p) at the end of the simulation are respectively 5.345×10^{-5} and 14.63%, where the latter quantity has been shown to be approximately constant over the last iterations. These results outline the good stability of the electric potential

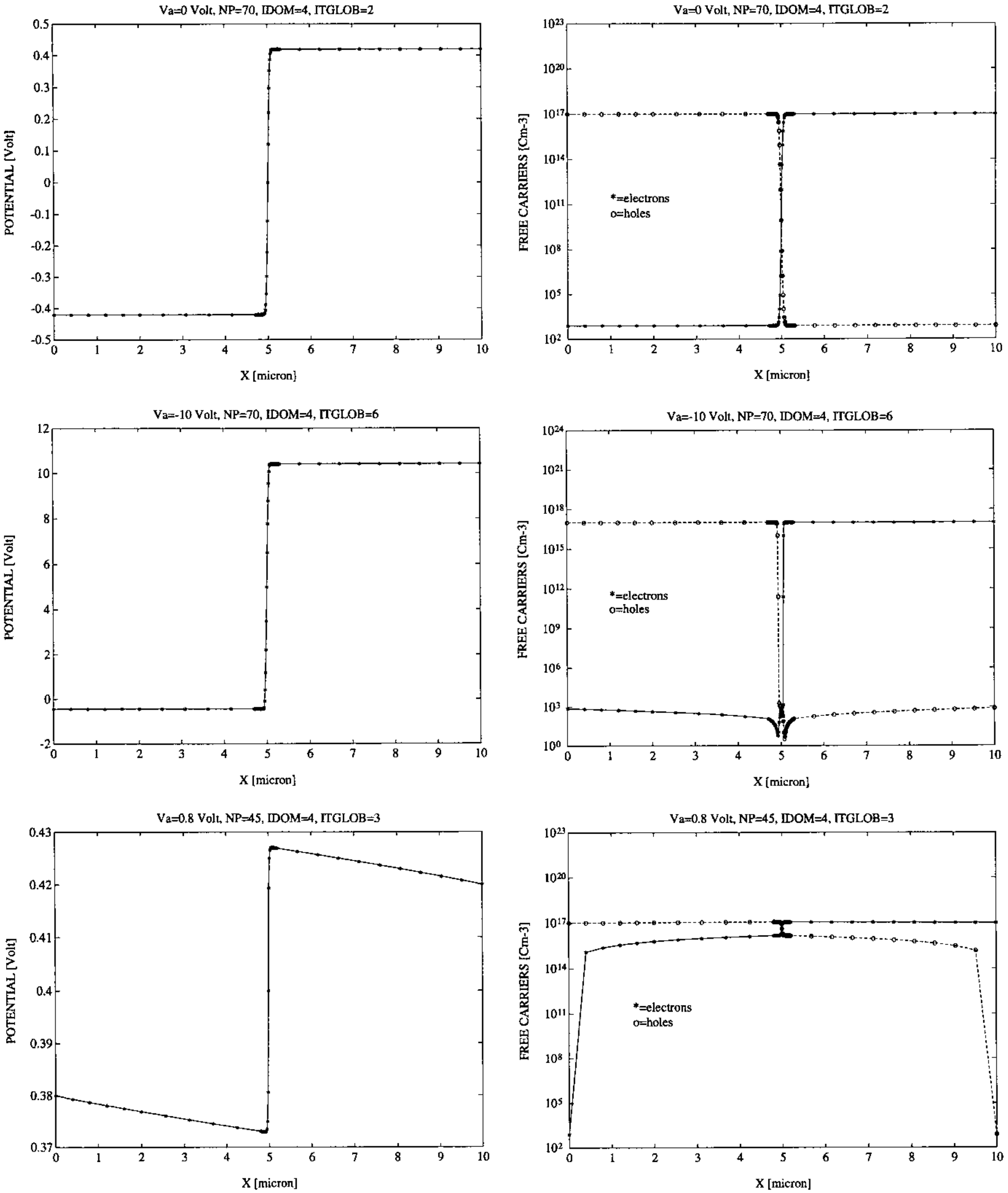

FIG. 7. Potential and carrier distributions at $V_a = (0, -10, 0.8)$ V.

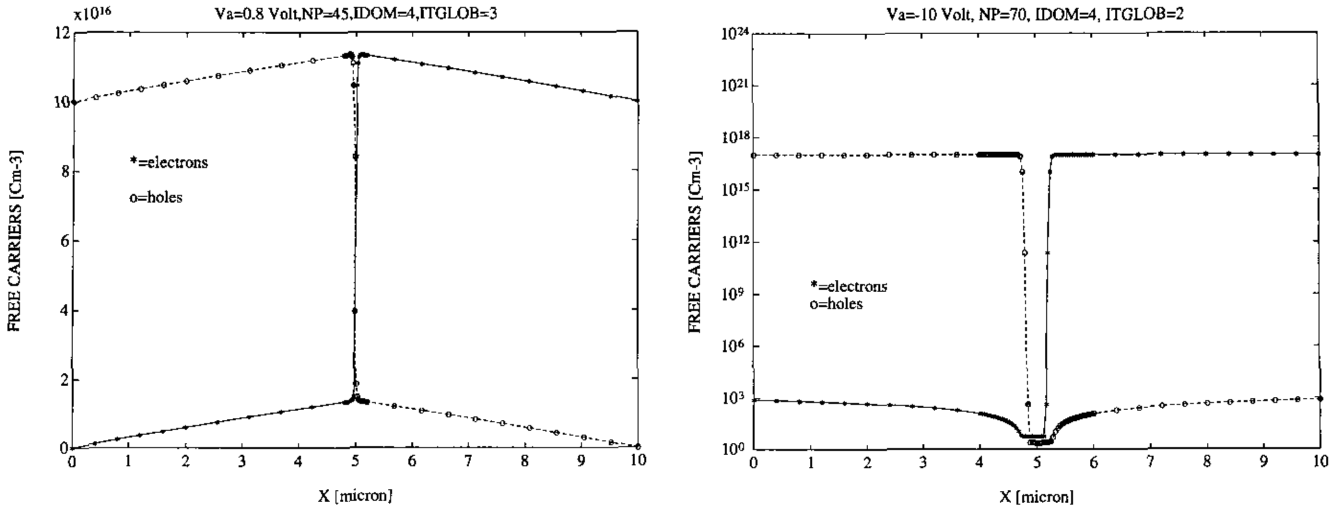

FIG. 8. Free carriers at $V_a = 0.8$ V (left) and $V_a = -10$ V ($R \equiv 0$).

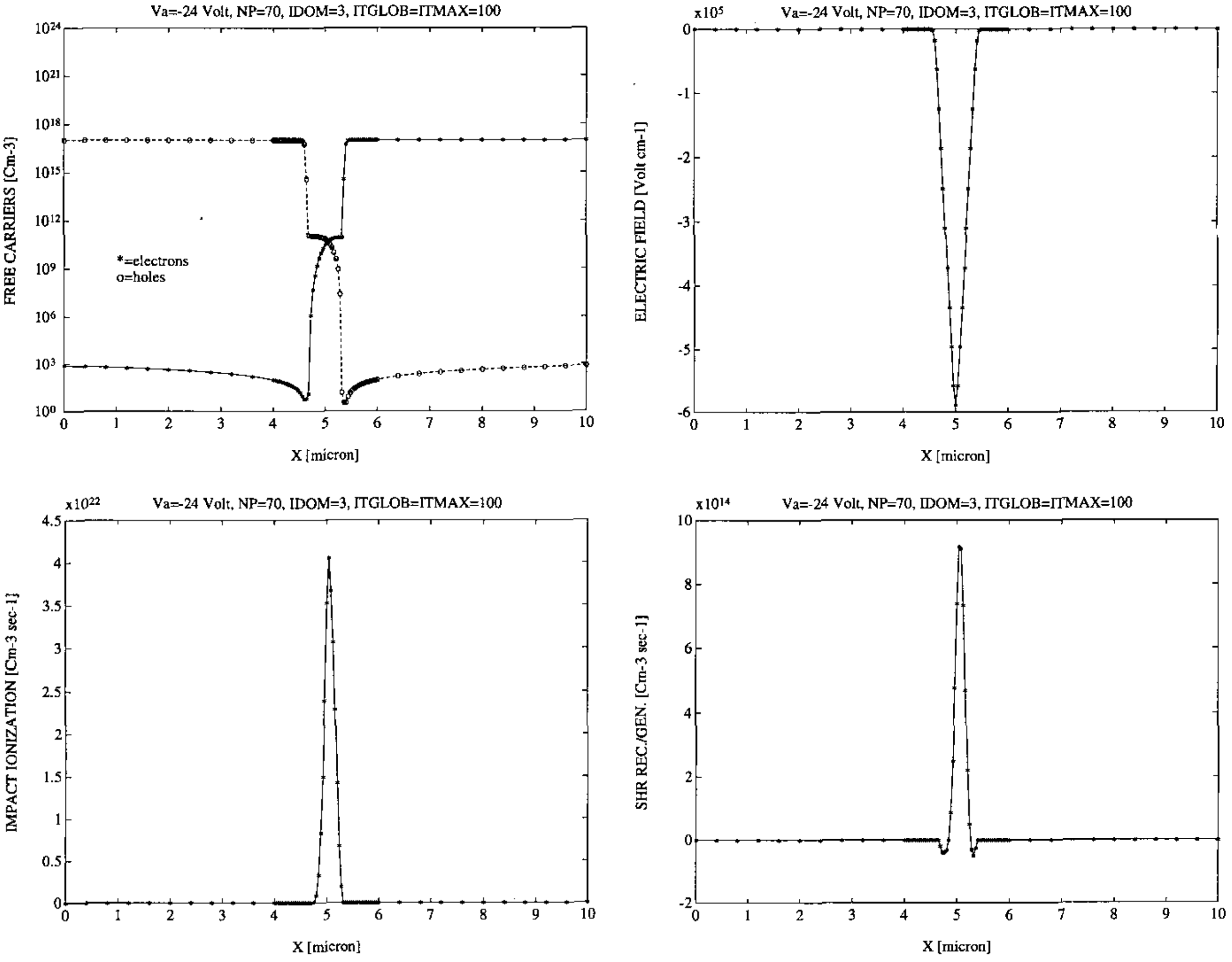

FIG. 9. Solutions at $V_a = -24$ V.

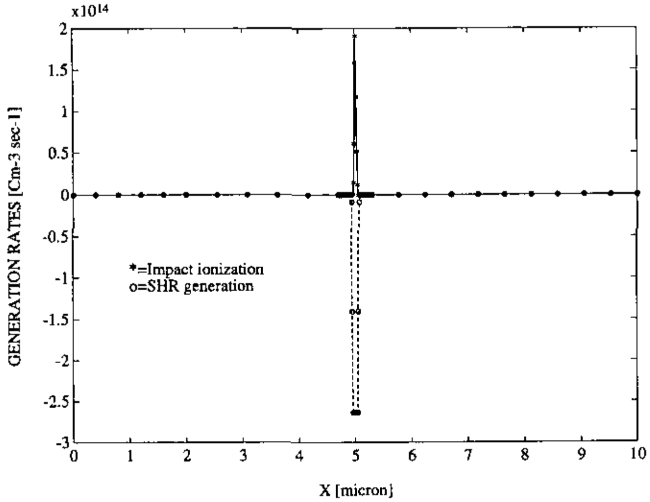

FIG. 10. Recombination/generation rate at $V_a = -10$ V.

and a trend of the carrier concentrations towards an indefinite increase, as k tends to infinity.

In order to better estimate this latter behaviour, it may be useful to compare the high values of the electron and hole densities across the junction with the corresponding ones at $V_a = -10$ V (in both the cases $R \neq 0$ and $R \equiv 0$) and, at the same time, plots of the recombination/generation mechanisms R_{II} and R_{SHR} (Auger generation R_{AU} has not been reported, since it is eight orders less than R_{SHR}). It clearly turns out that the increase of the carriers is to be ascribed only to the correspondingly large increase of the impact ionization generation term R_{II} as V_a is switched from -10 to -24 V. As a consequence, the electric field strength at the junction comes closer to its threshold value $E_{\mathrm{th}} = -600$ KVcm^{-1}, that is, the largest one for which model (12) holds true. This in turn produces the above large increase in R_{II} and explains the nonconvergence of Gummel's map at $V_a = -24$ V.

Starting from these numerical evidences, we will analyze Gummel's map in detail in the next subsection for all the reverse-biasing conditions, up to the breakdown voltage, and, eventually, we will propose a suitable variant that is shown to be successful.

6.1. Convergence of Gummel's Map at Breakdown

According to what has been shown by the numerical tests, we will make the following assumptions:

- the electric potential ψ is stabilized after few outer iterations ($k > 3$)
- the R_{SHR} term is saturated; i.e., it is independent of the applied voltage; and the R_{AU} term is negligible compared to R_{II}.

The whole differential problem may thus be written as

$$\begin{aligned} -J_n' + \alpha_n J_n + \alpha_p J_p &= f(x) \\ J_p' + \alpha_p J_p + \alpha_n J_n &= f(x), \end{aligned} \tag{39}$$

where we got rid of the Poisson equation, we included the R_{SHR} and R_{AU} terms in the function $f(x)$, and we used the fact that the current densities have a negative sign under reverse-bias conditions.

After the discretization procedure, we come to the linear system

$$M\mathbf{w} = \mathbf{f}, \tag{40}$$

where M is the 2×2 block matrix

$$M = \begin{pmatrix} M_{nn} & M_{np} \\ M_{pn} & M_{pp} \end{pmatrix}$$

and $\mathbf{w}, \mathbf{f} \in \mathbf{R}^{2N}$ are respectively the vectors

$$\mathbf{w} = (\mathbf{n}, \mathbf{p})^{\mathrm{T}}, \quad \mathbf{f} = (\mathbf{f}_n, \mathbf{f}_p)^{\mathrm{T}}$$

in which $\mathbf{n}, \mathbf{p} \in \mathbf{R}^N$ are respectively the nodal values of electron and hole concentrations and $\mathbf{f}_n, \mathbf{f}_p \in \mathbf{R}^N$ are the inner products of the function $f(x)$ with the finite element test functions v_{nh}, v_{ph}. The $N \times N$ blocks that make up the matrix M respectively approximate

$$\begin{aligned} M_{nn} &\rightarrow -J_n' + \alpha_n J_n \\ M_{np} &\rightarrow \alpha_p J_p \\ M_{pn} &\rightarrow \alpha_n J_n \\ M_{pp} &\rightarrow J_p' + \alpha_p J_p. \end{aligned}$$

We also define the block diagonal matrices,

$$D_0 = \begin{pmatrix} D_{0n} & 0 \\ 0 & D_{0p} \end{pmatrix}$$

$$P = \begin{pmatrix} M_{nn} & 0 \\ 0 & M_{pp} \end{pmatrix},$$

where D_{0n} and D_{0p} are $N \times N$ blocks respectively obtained by the finite element discretization of $-J_n'$ and J_p' in (39).

We may recover the standard Gummel's map by splitting in system (40) the matrix M as $M = D_0 + K$, thus leading to

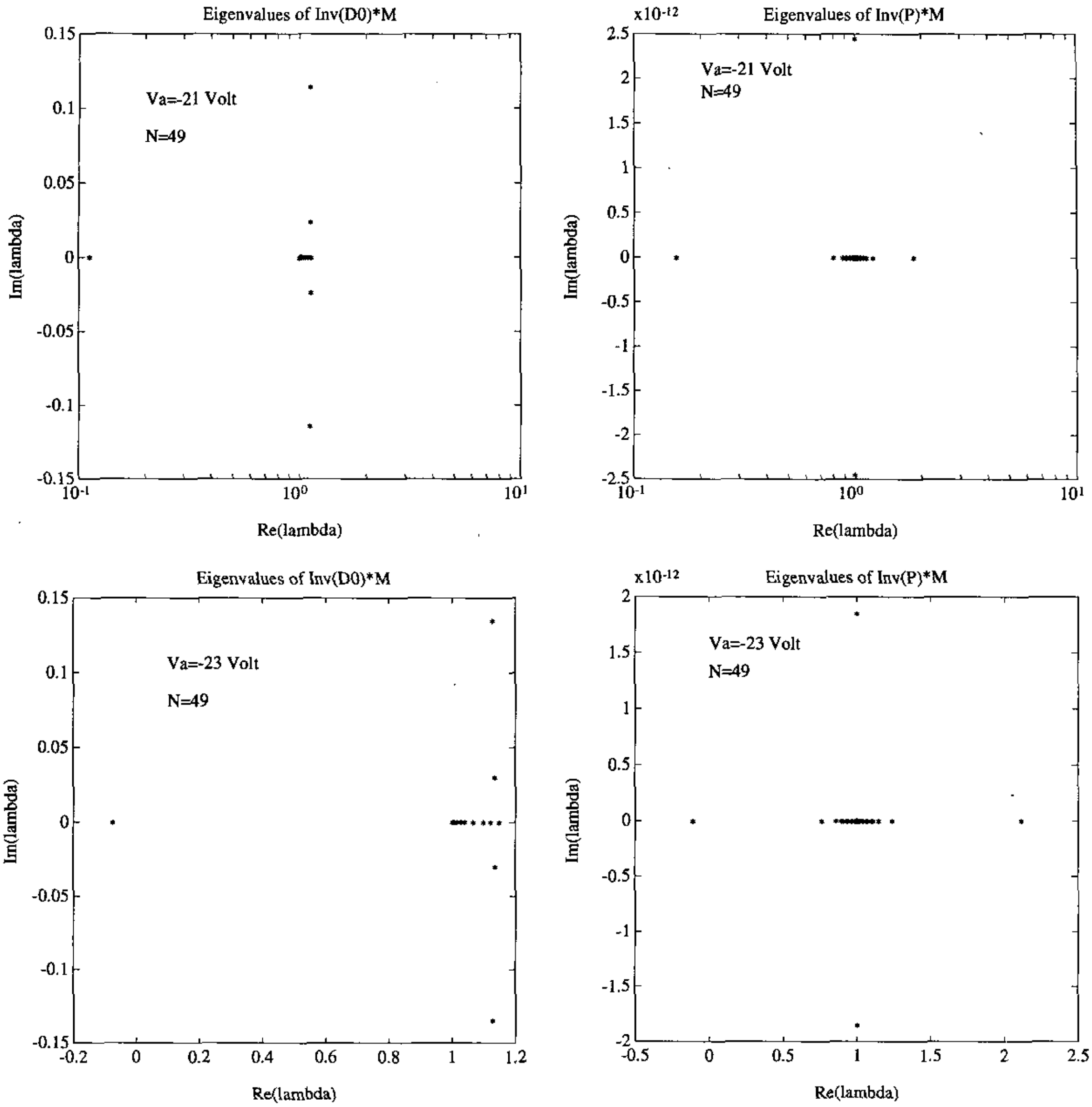

FIG. 11. Eigenvalues of the iteration matrices $D_0^{-1}M$ and $P^{-1}M$ at $V_a = (-21, -23)$ V.

$$D_0\mathbf{w} = -K\mathbf{w} + \mathbf{f} \tag{41}$$

and solving this system by Jacobi iterations

$$D_0\mathbf{w}^{k+1} = -K\mathbf{w}^k + \mathbf{f}. \tag{42}$$

The matrix D_0 therefore acts as a preconditioner of M.

The convergence condition for the approximants $\{\mathbf{w}^k\}$ is

$$\rho(I - D_0^{-1}M) < 1,$$

where $\rho(A)$ is the spectral radius of matrix A and I is the identity matrix. This requires

$$|1 - \lambda_j(D_1^{-1}M)| < 1 \quad \forall j = 1, ..., N,$$

that is satisfied iff all the eigenvalues of the matrix $D_0^{-1}M$ lie in the unit circle Γ with the center $C \equiv (1, 0)$ in the complex plane. The above result may be readily extended to the case of more general splittings of the form

$$M = S + (M - S),$$

where S is the preconditioner; the convergence condition becomes

$$\rho(I - S^{-1}M) < 1.$$

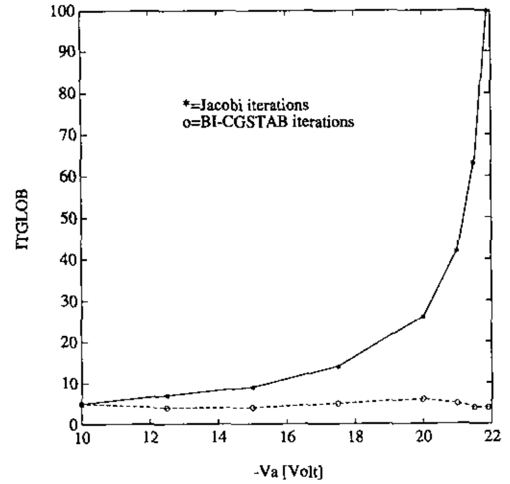

FIG. 12. Convergence rate of Jacobi and BI-CGSTAB iterations.

We consider, as a special case, $S \equiv P$, that corresponds to assigning the generation contributions due to each type of carrier to its associated drift-diffusion operator. The matrix P is closer to the matrix M than the standard preconditioner D_0 is, yet it preserves the decoupling between the current continuity equations. Use of such a preconditioner does not increase the overall computational cost and it actually improves the convergence rate of the classical iterative procedure. In the numerical experiments performed, the number of global iterations has been reduced up to a 10% factor.

In order to more precisely understand the slowing down of the convergence rate of Gummel's map as a function of the increasing reverse bias, we report in Fig. 11 a spectral analysis of both the iteration matrices $D_0^{-1}M$ and $P^{-1}M$ in the cases $V_a = -21$ and -23 V.

In the first case the eigenvalues of both matrices are strictly inside the unit circle Γ and turn out to be clustered around the center C; more specifically, we note that $\mathrm{Re}(\lambda_j(P^{-1}M))$ are symmetric with respect to C and that the imaginary parts are negligible.

An abrupt change of trend is exhibited by the two spectra in the second case, where some of the eigenvalues fall outside the convergence circle Γ, thus explaining the nonconvergence of the two iterative procedures.

The spectral analysis reported above and, in particular, the clustered dislocation of the eigenvalues suggests employing a different iterative procedure on the system (40), based on the use of the BI-CGSTAB method [20], which requires that the preconditioning matrix has eigenvalues with strictly positive real parts. This feature allows us to use the same preconditioner D_0 in such a way as to still preserve the decoupling between the two drift-diffusion equations, but in any case thoroughly changing the way of iterating. The actual implementation of this procedure for the resolution of (41) turns out to be very effective, with a spectacular improvement of the convergence rate in all the biasing conditions, up to the onset of the device breakdown. This is clearly shown in Fig. 12, where the number of global iterations is reported as a function of the increasing reverse applied voltage in both Jacobi and BI-CGSTAB algorithms.

The acceleration BI-CGSTAB procedure is started when the maximum nodal difference $|\delta\psi^k|$ is less than 10^{-10} V. This condition assures that electric potential is stabilized and in all the numerical experiments performed it is satisfied at $k = 3$. The total number of BI-CGSTAB iterations successively required to achieve convergence varies between 3 and 6.

The dramatic improvement with respect to the standard iterative procedure is clearly evident and this should make the use of BI-CGSTAB method very attractive for the simulation of multidimensional device geometries.

REFERENCES

1. J. F. Bourgat, R. Glowinski, P. Le Tallec, and M. Vidrascu, "Variational Formulation and Algorithm for Trace Operator in Domain Decomposition Calculations," in *Proceedings, 4th Symposium on Domain Decomposition Methods for Partial Differential Equations,* Vol. 3, edited by T. F. Chan *et al.* (SIAM, Philadelphia, 1990), p. 3.
2. F. Brezzi, L. D. Marini, and P. Pietra, *SIAM J. Numer. Anal.* **26,** 1347 (1989).
3. F. Brezzi, L. D. Marini, and P. Pietra, *Comput. Mech. Appl. Mech. Eng.* **75,** 493 (1989).
4. D. M. Caughey and R. E. Thomas, *Proc. IEEE* **52,** 2192 (1967).
5. A. G. Chynoweth, *Phys. Rev.* **109,** 1537 (1958).
6. D. C. D'Avanzo, M. Vanzi, and R. W. Dutton, Report G-201-5, Stanford University, 1979 (unpublished).
7. A. De Mari, *Solid-State Electron.* **11,** 33 (1968).
8. J. Dzewior and W. Schmid, *Appl. Phys. Lett.* **31,** 346 (1977).
9. T. Kerkhoven, *SIAM J. Numer. Anal.* **23,** 1121 (1986).
10. T. Kerkhoven, *SIAM J. Numer. Anal.* **25,** 1299 (1988).
11. R. Guerrieri, M. Rudan, P. Ciampolini, and G. Baccarani, "Vectorial Convergence Acceleration Techniques in Semiconductor Device Analysis," in *Proceedings, NASECODE IV Conference,* edited by J. J. H. Miller (Boole, Dublin, 1985), p. 293.
12. H. K. Gummel, *IEEE Trans. Electron Devices* **ED-11,** 455 (1964).
13. P. Markowich, *The Stationary Semiconductor Device Equations* (Springer-Verlag, Vienna/New York, 1986).
14. R. Sacco, *Appl. Num. Math.* **11,** 517 (1993).
15. R. Sacco, *Rend. Ist. Lomb. Accad. Sci. Lett. A* **125,** 265 (1991).
16. D. L. Scharfetter and H. K. Gummel, *IEEE Trans. Electron Devices* **ED-16,** 64 (1969).
17. S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer-Verlag, Vienna/New York, 1984).
18. J. W. Slotboom, *IEEE Trans. Electron Devices* **ED-20,** 669 (1973).
19. S. M. Sze, *Physics of Semiconductor Devices,* 2nd ed. (Wiley, New York, 1981).
20. H. Van Der Vorst, *SIAM J. Sci. Statist. Comput.* **13,** 631 (1992).
21. R. Van Overstraeten and H. De Man, *Solid-State Electron.* **13,** 583 (1970).