Adaptive multigrid applied to a bipolar transistor problem

J. Molenaar *

Centre for Mathematics and Computer Science, P.O. Box 4079, 1009 AB Amsterdam, Netherlands

Abstract

In this paper an adaptive nonlinear multigrid method is presented for the solution of the steady 2D semiconductor equations. The discretization is made on an adaptive grid by means of the (hybrid) mixed finite element method on rectangles. The integrals involved are approximated by means of the trapezoidal rule in order to obtain a generalization in 2D of the well-known Scharfetter–Gummel scheme. We show that the use of the trapezoidal rule does not influence the accuracy of the discretization.

The discrete equations thus obtained are solved by means of the dual version of the FAS-FMG algorithm. A Vanka-type relaxation is used as a smoother, and a local damping of the restricted residual is introduced in order to be able to use very coarse grids. Consistent with the FAS-FMG algorithm, we use the relative truncation error between coarse and fine grids as a refinement criterion for constructing adaptive grids. We study the relative truncation errors for the semiconductor equations in detail and show how they can be incorporated into a practical grid adaptation scheme. Results are shown for a realistic bipolar transistor problem.

Keywords: Semiconductor equations; Mixed finite element method; Adaptive grids; Multigrid methods

1. Introduction

The electric behavior of semiconductor devices is usually described by the classical drift-diffusion model, which consists of a system of three nonlinear partial differential equations. In this paper we only consider the steady state equations in two space dimensions. The usual approach for the discretization of the semiconductor device equations is the application of a box method (finite volume method), where the fluxes between the control volumes are approximated by the one-dimensional Scharfetter–Gummel scheme (cf. [1,19]). Brezzi et al. [4] introduced a two-dimensional exponential fitting method for the semiconductor equations using a (hybrid) mixed finite element method. We also use a mixed finite element scheme, but by using quadrature in the evaluation of integrals involved we retain the Scharfetter–Gummel

---

* Current address: Technical University Delft, Fac. TWI, P.O. Box 5031, 2600 GA Delft, Netherlands. E-mail: hansmo@twi.tudelft.nl.

scheme in the interior of the domain. The boundary treatment differs because our grid is cell-centered rather than vertex-centered. The advantage of this procedure is that standard error estimates are available for the mixed finite element method.

To solve the system of nonlinear equations obtained after discretization, we use a nonlinear multigrid method. Multigrid methods are optimal in the sense that the computation time needed is proportional to the number of grid cells. Linear multigrid methods are used for the iterative solution of the decoupled device equations (Gummel's iteration). However if the equations are strongly coupled, convergence is slow. The nonlinear multigrid method does not suffer from this drawback, as the equations are solved simultaneously.

The mixed finite elements bring about a cell-centered multigrid method, and the grid transfer operators are chosen in accordance with the discretization. As the semiconductor equations are strongly nonlinear and badly scaled, it is not straightforward to apply the multigrid method: special attention has to be paid to the formulation of the coarse grid problem. On the coarse grid we need some approximation of the solution. However, as we use the same discretization on all grids, this approximation implicitly determines the coefficients of the coarse grid problem. In this paper we discuss several possibilities of constructing a coarse grid approximation. In order to admit very coarse grids, it appears necessary to introduce a local damping of the restricted residual on the coarse grids. The amount of damping is determined by comparing corresponding diagonal elements of the coarse and fine grid Jacobian matrices. With this precaution we obtain an efficient multigrid algorithm for solving the discrete equations [18].

Another difficulty of the semiconductor equations is that they are singularly perturbed (cf. [16]), so we may expect that the dependent variables vary rapidly in small parts of the domain. Therefore it is desirable to have a fine mesh in parts of the domain where large variations of the solution occur. Several refinement criteria have been proposed for the semiconductor equations: estimates of the local truncation error, taking the singularly perturbed nature of the equations into account [16,25], the second derivative of the electrostatic potential [23], or estimates of the error in the electric field and the current densities [6]. Our adaptive mesh refinement scheme is based on the equidistribution of the relative truncation errors between the coarse and fine grids that are approximations of the coarse grid truncation errors; as there are three equations to be solved, we merge the different relative truncation errors into a single value. Using the relative truncation error as a refinement criterion is fully consistent with the multigrid algorithm used (cf. [11]).

Although many adaptive schemes have been proposed for the device equations, and also some nonlinear multigrid algorithms (cf. [5,7,18]), the combination of these two features has hardly been studied (see [18]). In this paper we analyze a nonlinear multigrid method on adaptively refined grids. This approach appears to be very attractive: the multigrid method is optimal with respect to the number of grid cells, whereas the number of grid cells is minimized by the adaptive gridding. Moreover the multigrid algorithm provides a natural grid refinement criterion.

An outline of this paper is as follows. In Sections 2 and 3 we present the equations solved and their discretization. The influence of the use of quadrature in the discretization is analyzed in Section 4. The multigrid method and the Vanka-type relaxation are presented in Sections 5 and 6. In Section 7 we analyze the different relative truncation errors and indicate how they

can be used in a practical mesh refinement strategy. In Section 8 we report numerical results for a bipolar transistor problem. In the final section our conclusions are summarized.

2. The equations

The equations to be solved are:

$$
\begin{aligned}
\operatorname{div} \boldsymbol{D} &= q(p - n + D), & \boldsymbol{D} &= -\varepsilon \operatorname{grad} \psi, \\
\operatorname{div} \boldsymbol{j}_n &= +qR, & \boldsymbol{j}_n &= +q\mu_n(U_T \operatorname{grad} n - n \operatorname{grad} \psi), \\
\operatorname{div} \boldsymbol{j}_p &= -qR, & \boldsymbol{j}_p &= -q\mu_p(U_T \operatorname{grad} p + p \operatorname{grad} \psi).
\end{aligned}
\tag{2.1}
$$

Here the dependent variables are ψ, n and p, which denote the electrostatic potential and the concentrations of electrons and holes, respectively, and $\boldsymbol{D}$, $\boldsymbol{j}_n$ and $\boldsymbol{j}_p$, the displacement current and the current densities of electrons and holes, respectively. The first equation is Poisson's equation: ε is the dielectric constant, q the elementary charge, and D the (given) dope function. The other two equations are continuity equations for electrons and holes with zero time derivatives; μ_n and μ_p are the electron and hole mobilities, U_T is the thermal voltage and R models the recombination rate of electrons and holes. In this paper we assume μ_ψ, μ_n and μ_p to be constant, and we only consider the Shockley–Read–Hall model for the recombination rate R,

$$
R = \frac{np - 1}{\tau_p(n + 1) + \tau_n(p + 1)},
\tag{2.2}
$$

where τ_n and τ_p are the electron and hole lifetimes, respectively.

In calculations we use the quasi-Fermi potentials ϕ_n and ϕ_p as dependent variables; these are related to n and p by

$$
n = n_i \mathrm{e}^{(\psi - \phi_n)/U_T}, \qquad p = n_i \mathrm{e}^{(\phi_p - \psi)/U_T},
\tag{2.3}
$$

with n_i the intrinsic density of free charge carriers.

To simplify the notation we use the following scaling:

Symbol	Scaling factor
ψ, ϕ_n, ϕ_p	U_T
n, p, D	n_i
R	q^{-1}
μ_n, μ_p	$(U_T n_i q)^{-1}$

and introduce

$$
j_\psi = \frac{D}{n_i q}, \qquad \mu_\psi = \frac{U_T}{n_i q} \varepsilon.
$$

The set of equations (2.1) expressed in the variables (ψ, ϕ_n, ϕ_p) is strongly nonlinear, but the range of the values assumed by (ψ, ϕ_n, ϕ_p) is of the same order as the voltages applied to the device. This makes them better suited for numerical computation than e.g. (ψ, n, p) (cf. [19]). In the discretization of (2.1) we use the Slotboom variables

$$\Phi_n = e^{-\phi_n}, \qquad \Phi_p = e^{+\phi_p}, \tag{2.4}$$

for which the equations appear in symmetric positive-definite form:

$$-\operatorname{div}(\mu_\psi \ \operatorname{grad} \ \psi) = e^{-\psi}\Phi_p - e^\psi\Phi_n + D,$$

$$-\operatorname{div}(\mu_n e^{+\psi} \ \operatorname{grad} \ \Phi_n) = -R, \tag{2.5}$$

$$-\operatorname{div}(\mu_p e^{-\psi} \ \operatorname{grad} \ \Phi_p) = -R.$$

The boundary conditions are of Dirichlet type $(\psi, \Phi_n, \Phi_p$ given) at the Ohmic contacts; at the remaining parts of the boundary homogeneous Neumann conditions are given: $n \cdot j_\psi = n \cdot j_n = n \cdot j_p = 0$, with n the outward normal unit vector at the boundary.

3. The mixed finite element discretization

The set of equations (2.5) can be written as

$$\sigma - A \ \operatorname{grad} \ u = 0, \quad \text{in } \Omega, \tag{3.1a}$$

$$\operatorname{div} \ \sigma = f, \quad \text{in } \Omega, \tag{3.1b}$$

$$u = g, \quad \text{on } \partial\Omega_D, \tag{3.1c}$$

$$n \cdot \sigma = 0, \quad \text{on } \partial\Omega_N, \tag{3.1d}$$

where $\partial\Omega_D$ and $\partial\Omega_N$ denote the parts of the boundary with Dirichlet or homogeneous Neumann boundary conditions, respectively. The sign is chosen such that $A > 0$.

Before we discuss the discretization of (3.1) we introduce some notation. Let $L^2(\Omega)$ be the space of square integrable functions on Ω with inner product $(\cdot, \cdot)$ and norm $\| \cdot \|_0$, and let $H^{BC}(\operatorname{div}, \Omega)$ be defined by

$$H^{BC}(\operatorname{div}, \Omega) = \left\{ \sigma \mid \sigma \in \left(L^2(\Omega)\right)^2, \operatorname{div} \ \sigma \in L^2(\Omega), \ n \cdot \sigma = 0 \text{ on } \partial\Omega_N \right\},$$

with norm $\| \cdot \|_H$ defined by

$$\| \sigma \|_H^2 = \| \sigma \|_0^2 + \| \operatorname{div} \ \sigma \|_0^2.$$

For ease of notation we write $V = H^{BC}(\operatorname{div}, \Omega)$ and $W = L^2(\Omega)$. The bilinear forms $a : V \times V \to \mathbb{R}$ and $b : V \times W \to \mathbb{R}$ are defined by

$$a(\sigma, \tau) = \int_\Omega A^{-1}\sigma \cdot \tau \ d\Omega, \qquad b(\sigma, t) = \int_\Omega t \ \operatorname{div} \ \sigma \ d\Omega.$$

Using this notation we write (3.1) in its weak form: find $(\sigma, u) \in V \times W$:

$$a(\sigma, \tau) + b(\tau, u) = \langle g, \tau \rangle_{\partial \Omega_D}, \quad \forall \tau \in V,$$
$$b(\sigma, t) = (f, t), \qquad\qquad \forall t \in W, \tag{3.2}$$

with

$$\langle g, \tau \rangle_{\partial \Omega_D} = \int_{\partial \Omega_D} g\tau \cdot n \, ds, \qquad (f, t) = \int_\Omega ft \, d\Omega.$$

For the discretization of (3.2) we assume that Ω can be divided by a regular partitioning in open disjoint, rectangular cells Ω^i, $\overline{\Omega} = \cup \overline{\Omega}^i$. On this partitioning of the domain we define lowest-order Raviart–Thomas elements [21]. On each cell Ω^i we have the characteristic function e^i, and for each edge E^j, $E^j \not\subset \partial \Omega_N$, of a cell Ω^i we define the "tent function" ε^j, i.e. the vector function of which each component is piecewise linear on each Ω^i and which satisfies $\varepsilon^j \cdot n^k = \delta_{jk}$, where n^k denotes the unit vector normal on the edge E^k in the positive coordinate direction; δ_{jk} is the Kronecker delta. Furthermore we introduce the function $\bar{e}^j \in L^2(\Omega)$, defined on the cell edges E^j, $E^j \not\subset \partial \Omega_D$,

$$\bar{e}^j(x) = \begin{cases} 1, & x \in E^j, \\ 0, & x \notin E^j, \end{cases}$$

and the "half tent" function $\varepsilon^{i,j} = e^i \varepsilon^j \in (L^2(\Omega))^2$ for $E^j \not\subset \partial \Omega_N$. Our discrete approximation spaces are defined by

$$V_h = \text{span}(\varepsilon^j) \subset V, \tag{3.3}$$
$$W_h = \text{span}(e^i) \subset W,$$
$$M_h = \text{span}(\bar{e}^j),$$
$$H_h = \text{span}(\varepsilon^{i,j}).$$

The mixed finite element discretization of (3.1) is: find $(\sigma_h, u_h) \in V_h \times W_h$ such that

$$a(\sigma_h, \tau_h) + b(\tau_h, u_h) = \langle g, \tau_h \rangle_{\partial \Omega_D}, \quad \forall \tau_h \in V_h, \tag{3.4a}$$
$$b(\sigma_h, t_h) = (f, t_h), \qquad\qquad \forall t_h \in W_h. \tag{3.4b}$$

Assuming a piecewise constant approximation for g, the integrals $\langle g, \tau_h \rangle$ and $b(\tau_h, t_h)$ are easily evaluated. The integrals $a(\sigma_h, \tau_h)$ and (f, t_h) in (3.4) are approximated by a repeated, weighted trapezoidal rule for rectangles:

$$\int_\Omega w(x)z(x) \, d\Omega = \sum_i \int_{\Omega^i} w(x)z(x) \, d\Omega \approx \sum_i \sum_{\nu = 1,2,3,4} z(x^{i,\nu}) \int_{\Omega^{i,\nu}} w(x) \, d\Omega, \tag{3.5}$$

where $x^{i,\nu}$ are the four vertices of Ω^i, and $\Omega^{i,\nu}$ the four quarter rectangles, parts of Ω^i, associated with these vertices, respectively.

If (f, t_h) is approximated by (3.5) with $w = t_h$ and $z = f$, we obtain from (3.4b)

$$\forall \Omega^i: \quad \sum_j h^j d^{i,j} \sigma^j = \text{area}(\Omega^i) \frac{1}{4} \sum_{\nu = 1,2,3,4} f(x^{i,\nu}), \tag{3.6}$$

with

$$d^{i,j} = \begin{cases} +1, & \text{if } E^j \text{ is an N- or E-edge of } \Omega^i, \\ -1, & \text{if } E^j \text{ is an S- or W-edge of } \Omega^i, \\ 0, & \text{otherwise}, \end{cases} \tag{3.7}$$

and h^j the length of edge E^j. We notice that (3.6) and (3.7) imply discrete current conservation.

For the approximation of $a(\sigma_h, \tau_h)$ we use (3.5) with $w = A^{-1}$ and $z = \sigma_h \cdot \tau_h$. Here the use of the trapezoidal rule is also called lumping, because the matrix $a(\varepsilon^j, \varepsilon^k)$ is approximated by a matrix $\tilde{a}(\varepsilon^j, \varepsilon^k)$ that is diagonal:

$$\tilde{a}(\varepsilon^j, \varepsilon^k) = \delta_{jk} \int_{\Delta_E^k} A^{-1} \, d\Omega.$$

Here $\Delta_E^k = \bigcup \{\Omega^{i,\nu} \mid \overline{\Omega}^{i,\nu} \cap E^k \neq \emptyset\}$; i.e., Δ_E^k is the dual box related to the edge E^k. If we consider (3.1) as a Poisson equation ($A = 1$), it is advantageous to use lumping: the non-lumped form of the discretization does not yield an M-matrix after elimination of σ, which is the case if lumping is used (cf. [20]). Moreover for the continuity equations we are able to retain the Scharfetter–Gummel scheme [22] by using the quadrature rule (3.5). For the continuity equations A^{-1} is the exponentially varying function $A^{-1} = e^{\pm \psi}$. If we approximate ψ in Δ_E^k by a linear function, interpolating ψ from its values ψ^R and ψ^L in the neighboring cells Ω^i, $i = $ R,L we obtain

$$\int_{\Delta_R^k} e^{\psi} \, d\Omega = \text{area}(\Delta_E^k) \, \text{Bexp}^{-1}(\psi^R, \psi^L),$$

with

$$\text{Bexp}(x, y) = \frac{x - y}{e^x - e^y}.$$

For the discretization of (2.5) we apply the above scheme for the discretization of (3.1) with $u = (\psi, \Phi_n, \Phi_p)$, $\sigma = (j_\psi, j_n, j_p)$ and $A = (-\mu_\psi, +\mu_n \exp(+\psi), -\mu_p \exp(-\psi))$, respectively. In order to obtain the usual definition of (j_ψ, j_n, j_p) (cf. [19]) we allow negative values for A. For an edge E^j with adjacent cells Ω^i, $i = $ R,L, we obtain

$$j_\psi^j = -\frac{h^j}{a^j} \mu_\psi (\psi^R - \psi^L),$$

$$j_n^j = +\frac{h^j}{a^j} \mu_n (\Phi_n^R - \Phi_n^L) \text{Bexp}(-\psi^R, -\psi^L), \tag{3.8}$$

$$j_p^j = -\frac{h^j}{a^j} \mu_p (\Phi_p^R - \Phi_p^L) \text{Bexp}(+\psi^R, +\psi^L),$$

with h^j the length of E^j, and $a^j = \text{area}(\Delta_E^j)$. We observe that after lumping the fluxes j^j may be eliminated to yield a scheme that is equivalent to the usual box scheme (see e.g. [19]) in the

interior of the domain Ω. However the geometry of our discretization is cell-centered as opposed to the usual box scheme that is vertex-centered.

In our calculations on adaptive grids (Section 6) we will need approximations for the potentials u at edges of cells. It is known that the Lagrange multipliers that appear in the hybrid mixed finite element discretization are a good approximation for these values (cf. [3]). In the hybrid mixed finite element method the half tent functions $\varepsilon^{i,j}$ are used as test functions in (3.4a), and a Lagrange multiplier λ_h is introduced to enforce sufficient continuity of σ_h.

The augmented variational equations read: find $(\sigma_h^*, u_h^*, \lambda_h) \in H_h \times W_h \times M_h$ such that

$$\sum_i \int_{\Omega^i} A^{-1}\sigma_h^* \cdot \tau_h \, d\Omega + \sum_i \int_{\Omega^i} u_h^* \, \mathrm{div} \, \tau_h \, d\Omega = \sum_i \oint_{\partial \Omega^i} \lambda_h \tau_h \cdot n \, ds, \quad \forall \tau_h \in H_h,$$

$$\sum_i \int_{\Omega^i} t_h \, \mathrm{div} \, \sigma_h^* \, d\Omega = \sum_i \int_{\Omega^i} f t_h \, d\Omega, \quad \forall t_h \in W_h, \qquad (3.9)$$

$$\sum_j \int_{E^j} \mu_h \sigma_h^* \cdot n^j \, ds = 0, \quad \forall \mu_h \in M_h.$$

The third equation guarantees that $\sigma_h^* \in H^{\mathrm{BC}}(\mathrm{div}, \Omega)$, hence in the interior of the domain the solution of (3.9) coincides with the solution of (3.4). The values λ^j are the coefficients in $\lambda_h = \sum_j \lambda^j \bar{e}^j \in M_h$; on the interiors edges E^j the λ^j can be expressed in the values of u^i in the adjacent cells Ω^i, $i = \mathrm{R,L}$,

$$\lambda^j = u^{\mathrm{L}} \frac{\int_{\bar{\Omega}^{\mathrm{R}} \cap \Delta_j} A^{-1} \, d\Omega}{\int_{\Delta_j} A^{-1} \, d\Omega} + u^{\mathrm{R}} \frac{\int_{\bar{\Omega}^{\mathrm{L}} \cap \Delta_j} A^{-1} \, d\Omega}{\int_{\Delta_j} A^{-1} \, d\Omega}. \qquad (3.10)$$

If the weighted trapezoidal rule (3.5) is used to evaluate the integrals in (3.10), this comes down to linear interpolation for the Poisson equation and exponential interpolation for the continuity equations as was used for the one-dimensional case in [12].

4. Influence of quadrature

In this section we study the influence of the quadrature rule (3.5) on the accuracy of the discrete solution of (3.1). We assume that A and f are given functions and that the boundary conditions are homogeneous ($g = 0$).

The Sobolev spaces $W^{m,p}(\Omega)$ are, as usual, defined by

$$W^{m,p}(\Omega) = \{u \mid D^\alpha u \in L^p(\Omega), 0 \leq |\alpha| \leq m\}, \qquad H^m(\Omega) = W^{m,2}(\Omega),$$

with seminorm

$$|u|_{m,p,\Omega} = \left(\sum_{|\alpha|=m} \int_\Omega |D^\alpha u|^p \, d\Omega \right)^{1/p}$$

and norm

$$\| u \|_{m,p,\Omega} = \left(\sum_{0 \leqslant l \leqslant m} | u |^p_{l,p,\Omega} \right)^{1/p},$$

$$\| u \|_{m,\infty,\Omega} = \max_{0 \leqslant |\alpha| \leqslant m} \sup_{\Omega} | D^\alpha u |,$$

$$\| u \|_m = \| u \|_{m,2,\Omega},$$

where D^α denotes the distributional derivative of order α. Furthermore we introduce a norm on the partitioning of the domain Ω,

$$\| u \|_{m,p,\Delta} = \left(\sum_i \| u \|^p_{m,p,\Omega^i} \right)^{1/p}, \tag{4.1}$$

$$\| u \|_{m,\Delta} = \| u \|_{m,2,\Delta}.$$

The projection operators $\Pi_h^\sigma : V \to V_h$ and $\Pi_h^u : W \to W_h$ are defined by (cf. [21])

$$\langle \bar{e}^j, \sigma \rangle_{E^j} = \langle \bar{e}^j, \Pi_h^\sigma \sigma \rangle_{E^j}, \quad \forall E^j,$$

$$(e^i, u) = (e^i, \Pi_h^u u), \quad \forall \Omega^i,$$

with

$$\langle \bar{e}^j, \sigma \rangle_{E^j} = \int_{E^j} \bar{e}^j \sigma \cdot n^j \, ds.$$

For these projections on the lowest-order Raviart–Thomas elements the divergence operator and the projection operator commute (cf. [8]):

$$\text{div } \Pi_h^\sigma = \Pi_h^u \text{ div}. \tag{4.2}$$

Furthermore we have (cf. [10])

$$\| \Pi_h^\sigma \sigma \|_V \leqslant C \| \sigma \|_V, \quad \forall \sigma \in V. \tag{4.3}$$

For the mixed finite element discretization of (3.1) the following result is known [8].

Theorem 4.1. *If the domain Ω is such that the problem $\text{div}(\text{grad } u) = w$ with homogeneous Dirichlet boundary conditions is regular, i.e. for each $w \in L^2(\Omega)$ there exists a unique solution $u \in H^2(\Omega)$ with $\| u \|_2 \leqslant C \| w \|_0$, and if $a(\cdot, \cdot)$ is such that*

$$\exists C > 0: \quad a(\sigma, \tau) \geqslant C \| \sigma \|_0 \| \tau \|_0, \quad \forall \sigma, \tau \in V,$$

then problem (3.2) has a unique solution (σ, u) and problem (3.4) has a unique solution (σ_h, u_h). Moreover there exists a $C > 0$, independent if h, such that

$$\| \sigma - \sigma_h \|_0 \leqslant Ch \| u \|_2,$$

$$\| \text{div}(\sigma - \sigma_h) \|_0 \leqslant C \| u \|_2,$$

$$\| u - u_h \|_0 \leqslant Ch \| u \|_2.$$

Proof. For a proof see [8]. □

The accuracy of the quadrature rule (3.5) is stated in the following lemma.

Lemma 4.2. *For functions* $w, z \in C^2(\overline{\Omega}^i)$ *we have*

$$\left| \int_\Omega wz \, d\Omega - \sum_i \sum_\nu z(x^\nu) \int_{\Omega^{i,\nu}} w(x) \, d\Omega \right| \leqslant Ch^2 \| wz \|_{2,1,\Delta}. \tag{4.4}$$

Proof. The quadrature is exact on Ω^i for all constant functions z, and also for constant functions w and linear functions z, so by using Taylor expansions around the centre of Ω^i we obtain

$$\left| \int_\Omega wz \, d\Omega - \sum_i \sum_\nu z(x^\nu) \int_{\Omega^{i,\nu}} w(x) \, d\Omega \right| \leqslant \sum_i \left| \int_{\Omega^i} wz \, d\Omega - \sum_\nu z(x^\nu) \int_{\Omega^{i,\nu}} w(x) \, d\Omega \right|$$

$$\leqslant Ch^2 \sum_i \| wz \|_{2,1,\Omega^i} = Ch^2 \| wz \|_{2,1,\Delta}. \qquad \square$$

We are now ready to estimate the influence of the use of quadrature in the discretization. We split this influence in errors that are induced by the use of quadrature for the right-hand side (f, t_h) (cf. Theorem 4.3), and errors that are induced by the use of quadrature for $a(\varepsilon_h^j, \varepsilon_h^k)$ (cf. Theorem 4.5). As t_h is piecewise constant, the use of (3.5) is equivalent to replacing f by a piecewise bilinear interpolation function f_h and using exact integration. Therefore we write the approximation of (f, t_h) by the quadrature rule (3.5) as (f_h, t_h).

Theorem 4.3. *Let* (σ_h, u_h) *be the solution of*

$$\begin{aligned} a(\sigma_h, \tau_h) + b(\tau_h, u_h) &= 0, \quad &\forall \tau_h \in V_h, \\ b(\sigma_h, t_h) &= (f, t_h) \quad &\forall t_h \in W_h, \end{aligned} \tag{4.5}$$

and let $(\hat{\sigma}_h, \hat{u}_h)$ *be the solution of*

$$\begin{aligned} a(\hat{\sigma}_h, \tau_h) + b(\tau_h, \hat{u}_h) &= 0, \quad &\forall \tau_h \in V_h, \\ b(\hat{\sigma}_h, t_h) &= (f_h, t_h), \quad &\forall t_h \in W_h. \end{aligned} \tag{4.6}$$

Under the conditions of Theorem 4.1 we have

$$\| \sigma_h - \hat{\sigma}_h \|_0 \leqslant Ch^2 \| f \|_{2,\Delta}, \tag{4.7}$$

$$\| u_h - \hat{u}_h \|_0 \leqslant Ch^2 \| f \|_{2,\Delta}.$$

Proof. First we prove that $\| \sigma_h - \hat{\sigma}_h \|_0^2 \leqslant Ch^2 \| f \|_{2,\Delta} \| u_h - \hat{u}_h \|_0$:

$$\| \sigma_h - \hat{\sigma}_h \|_0^2 \leqslant C \left| a(\sigma_h - \hat{\sigma}_h, \sigma_h - \hat{\sigma}_h) \right| = C \left| b(\sigma_h - \hat{\sigma}_h, u_h - \hat{u}_h) \right|$$

$$= C \left| (f - f_h, u_h - \hat{u}_h) \right| \leqslant Ch^2 \| f(u_h - \hat{u}_h) \|_{2,1,\Delta}$$

$$\leqslant Ch^2 \| f \|_{2,\Delta} \| u_h - \hat{u}_h \|_0. \tag{4.8}$$

Suppose that ψ is the solution of $\mathrm{div}(\mathrm{grad}\ \psi) = u_h - \hat{u}_h$. Let $\tau_h = \Pi_h^\sigma \mathrm{grad}\ \psi$, then

$$\mathrm{div}\ \tau_h = \mathrm{div}\ \Pi_h^\sigma \mathrm{grad}\ \psi = \Pi_h^u \mathrm{div}\ \mathrm{grad}\ \psi = u_h - \hat{u}_h,$$

so

$$\| u_h - \hat{u}_h \|_0^2 = | b(\tau_h, u_h - \hat{u}_h) | = | a(\sigma_h - \hat{\sigma}_h, \tau_h) |$$

$$\leqslant C \| \sigma_h - \hat{\sigma}_h \|_0 \| \Pi_h^\sigma \text{ grad } \psi \|_0$$

$$\leqslant C \| \sigma_h - \hat{\sigma}_h \|_0 \| \text{grad } \psi \|_V \leqslant C \| \sigma_h - \hat{\sigma}_h \|_0 \| u_h - \hat{u}_h \|_0 .$$

Hence

$$\| u_h - \hat{u}_h \|_0 \leqslant C \| \sigma_h - \hat{\sigma}_h \|_0 . \tag{4.9}$$

Combining (4.8) and (4.9) proves the theorem. □

Before we state a theorem on the influence of the use of quadrature in the evaluation of $a(\cdot, \cdot)$, we remark that the space V_h has the following inverse property:

$$h \| \sigma_h \|_V \leqslant C \| \sigma_h \|_0 , \quad \forall \sigma_h \in V_h . \tag{4.10}$$

In the proof of the theorem we need the following estimate for the norm $\| \cdot \|_{2,1,\Delta}$ that appears in the right-hand side of (4.4).

Lemma 4.4. *Let* $\sigma_h, \tau_h \in V_{hi}$, *then*

$$\| \sigma_h \cdot \tau_h \|_{2,1,\Delta} \leqslant C \| \sigma_h \|_V \| \tau_h \|_V , \tag{4.11}$$

with C independent of h.

Proof. By repeated use of Cauchy's inequality we obtain

$$\| \sigma_h \cdot \tau_h \|_{2,1,\Delta} = \sum_i \sum_{|\alpha| \leqslant 2} \int_{\Omega^i} | D^\alpha (\sigma_h \cdot \tau_h) | \, d\Omega$$

$$\leqslant C \sum_i \sum_{\substack{|\alpha| \leqslant 2 \\ \alpha_1 + \alpha_2 = \alpha}} \sum_{d=x,y} \int_{\Omega^i} | D^{\alpha_1} \sigma_{h,d} | \, | D^{\alpha_2} \tau_{h,d} | \, d\Omega$$

$$\leqslant C \sum_i \sum_{\substack{|\alpha| \leqslant 2 \\ \alpha_1 + \alpha_2 = \alpha}} \sum_{d=x,y} \left[\int_{\Omega^i} | D^{\alpha_1} \sigma_{h,d} |^2 \, d\Omega \right]^{1/2} \left[\int_{\Omega^i} | D^{\alpha_2} \tau_{h,d} |^2 \, d\Omega \right]^{1/2}$$

$$\leqslant C \left[\sum_i \sum_{|\alpha| \leqslant 2} \sum_{d=x,y} \int_{\Omega^i} | D^\alpha \sigma_{h,d} |^2 \, d\Omega \right]^{1/2}$$

$$\times \left[\sum_i \sum_{|\alpha| \leqslant 2} \sum_{d=x,y} \int_{\Omega^i} | D^\alpha \tau_{h,d} |^2 \, d\Omega \right]^{1/2} .$$

For $\sigma_h \in V_h$ we have

$$\sum_i \sum_{|\alpha| \leqslant 2} \sum_{d=x,y} \int_{\Omega^i} | D^\alpha \sigma_{h,d} |^2 \, d\Omega$$

$$= \sum_i \int_{\Omega^i} \left(| \sigma_{h,x} |^2 + | \sigma_{h,y} |^2 + | \partial_x \sigma_{h,x} |^2 + | \partial_y \sigma_{h,y} |^2 \right) d\Omega = \| \sigma_h \|_V^2 . \quad \Box$$

Theorem 4.5. *Let* $(\hat{\sigma}_h, \hat{u}_h)$ *be as in Theorem 4.3, and let* $(\tilde{\sigma}_h, \tilde{u}_h)$ *be the solution of*

$$\tilde{a}(\tilde{\sigma}_h, \tau_h) + b(\tau_h, \tilde{u}_h) = 0, \quad \forall \tau_h \in V_h,$$
$$b(\tilde{\sigma}_h, t_h) = (f_h, t_h), \quad \forall t_h \in W_h. \tag{4.12}$$

If $\tilde{a}(\cdot, \cdot)$ *is such that*

$$\exists a_h > 0: \quad \tilde{a}(\sigma_h, \sigma_h) \geqslant a_h \| \sigma_h \|_0^2, \quad \forall \sigma \in V_h \tag{4.13}$$

and the conditions of Theorem 4.1 hold, then

$$\| \tilde{\sigma}_h - \hat{\sigma}_h \|_0 \leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} a_h^{-1} \| \hat{\sigma}_h \|_V, \tag{4.14}$$

$$\| \tilde{u}_h - \hat{u}_h \|_0 \leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \left(a_h^{-1} \| \hat{\sigma}_h \|_V + \| \tilde{\sigma}_h \|_V \right). \tag{4.15}$$

Moreover, if $\| \sigma \|_V$ *is bounded then* $\| \hat{\sigma}_h \|_V$ *and* $\| \tilde{\sigma}_h \|_V$ *are bounded.*

Proof. Subtracting (4.12) from (4.6) yields

$$b(\hat{\sigma}_h - \tilde{\sigma}_h, t_h) = 0, \quad \forall t_h \in W_h, \tag{4.16}$$

so

$$b(\hat{\sigma}_h - \tilde{\sigma}_h, \operatorname{div}(\hat{\sigma}_h - \tilde{\sigma}_h)) = \| \operatorname{div}(\hat{\sigma}_h - \tilde{\sigma}_h) \|_0^2 = 0. \tag{4.17}$$

By using (4.4), (4.11), (4.13) and (4.17) we obtain

$$a_h \| \tilde{\sigma}_h - \hat{\sigma}_h \|_0^2 \leqslant \tilde{a}(\tilde{\sigma}_h - \hat{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h)$$

$$\leqslant | \tilde{a}(\tilde{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h) - a(\hat{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h)| + | a(\hat{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h) - \tilde{a}(\hat{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h)|$$

$$= | a(\hat{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h) - \tilde{a}(\hat{\sigma}_h, \tilde{\sigma}_h - \hat{\sigma}_h)|$$

$$\leqslant Ch^2 \| A^{-1} \hat{\sigma}_h \cdot (\tilde{\sigma}_h - \hat{\sigma}_h) \|_{2,1,\Delta}$$

$$\leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \| \hat{\sigma}_h \cdot (\tilde{\sigma}_h - \hat{\sigma}_h) \|_{2,1,\Delta}$$

$$\leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \| \hat{\sigma}_h \|_V \| \tilde{\sigma}_h - \hat{\sigma}_h \|_V$$

$$= Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \| \hat{\sigma}_h \|_V \| \tilde{\sigma}_h - \hat{\sigma}_h \|_0.$$

This proves the first part of the theorem. Next suppose that ψ is the solution of $\operatorname{div}(\operatorname{grad} \psi) = \tilde{u}_h - \hat{u}_h$. Let $\tau_h = \Pi_h^\sigma \operatorname{grad} \psi$, then

$$\| \tilde{u}_h - \hat{u}_h \|_0^2 = | b(\tau_h, \tilde{u}_h - \hat{u}_h)| = | a(\hat{\sigma}_h, \tau_h) - \tilde{a}(\tilde{\sigma}_h, \tau_h)|$$

$$\leqslant | a(\hat{\sigma}_h - \tilde{\sigma}_h, \tau_h)| + | a(\tilde{\sigma}_h, \tau_h) - \tilde{a}(\tilde{\sigma}_h, \tau_h)|$$

$$\leqslant C \left(\| \hat{\sigma}_h - \tilde{\sigma}_h \|_0 \| \tau_h \|_0 + h^2 \| A^{-1} \tilde{\sigma}_h \cdot \tau_h \|_{2,1,\Delta} \right)$$

$$\leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \left(a_h^{-1} \| \hat{\sigma}_h \|_V \| \tau_h \|_0 + \| \tilde{\sigma}_h \|_V \| \tau_h \|_V \right)$$

$$\leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \left(a_h^{-1} \| \hat{\sigma}_h \|_V + \| \tilde{\sigma}_h \|_V \right) \| \operatorname{grad} \psi \|_V$$

$$\leqslant Ch^2 \| A^{-1} \|_{2,\infty,\Omega} \left(a_h^{-1} \| \hat{\sigma}_h \|_V + \| \tilde{\sigma}_h \|_V \right) \| \tilde{u}_h - \hat{u}_h \|_0.$$

Now it remains to prove that $\| \hat{\sigma}_h \|_V$ and $\| \tilde{\sigma}_h \|_V$ are bounded. It follows from Theorem 4.1 that $\| \sigma_h \|_V$ is bounded; from (4.7), (4.14) and (4.10) we obtain

$$\| \sigma_h - \hat{\sigma}_h \|_V \leqslant Ch \| f \|_{2,\Delta}$$

and

$$\| \tilde{\sigma}_h - \hat{\sigma}_h \|_V \leqslant Ch \| A^{-1} \|_{2,\infty,\Omega} a_h^{-1} \| \hat{\sigma}_h \|_V,$$

therefore we conclude that $\| \hat{\sigma}_h \|_V$ and $\| \tilde{\sigma}_h \|_V$ are bounded. □

From Theorems 4.1, 4.3 and 4.5 we conclude that the quadrature rule (3.5) does not spoil the order of accuracy of the discretization, because the errors introduced by the quadrature are of higher order than the discretization errors.

5. Vanka-type relaxation

The efficiency of any multigrid algorithm depends on the choice of the relaxation procedure. Several procedures have been proposed for the system of equations that arises from the mixed finite element discretization. Block-wise relaxation with current conservation has been used by Schmidt and Jacobs [24] for the solution of Poisson's equation with Neumann boundary conditions; Maitre et al. [15] give an analysis of Uzawa relaxation. The relaxation method we use is related with the one proposed by Vanka [26] for the solution of the incompressible Navier–Stokes equations.

In Vanka-type relaxation, all cells in a grid are scanned in some predetermined order. When a cell is visited the potentials related to that cell and the fluxes at its edges are relaxed simultaneously. In [17] a smoothing analysis is presented for Vanka relaxation applied to the mixed finite element discretization of Poisson's equation. For the semiconductor equations Vanka-type relaxation in a cell leads to a system of fifteen equations. The fluxes appear linearly in these equations so they are easily eliminated, reducing the system to be solved for each cell to a set of three nonlinear equations. To solve this small nonlinear system, we first try Newton's method combined with Schilders' correction transformation [19]. The advantage of Newton iteration is that it converges quadratically in the neighborhood of the solution. If Newton's method fails we resort point-wise to Gummel iteration, which is less efficient but more robust (cf. [18]). An analysis of point-wise Gummel iteration is presented in [13].

A consequence of the use of lumping in the discretization is that the equation related to an edge only depends on the flux at that edge and the potentials in the adjacent cells. Therefore it is a property of a Vanka-type relaxation that all equations related to edges are satisfied after a complete relaxation sweep. Underrelaxation of the fluxes, as proposed by Vanka [26], would spoil this property. Furthermore for our problem the use of a damping parameter (unequal to one) does not enhance the convergence properties of Vanka-type relaxation (cf. [17]), therefore we apply the Vanka-type relaxation without a damping parameter.

Vanka relaxation can also be used in a line-wise version. Again we can eliminate the fluxes to obtain a nonlinear system of which the Jacobian is a band matrix with 10 diagonals. If good initial estimates are available this system can be solved by Newton iteration. An experimental comparison of the efficiency of point- and line-wise Vanka relaxation is presented in Section 8.

6. Multigrid and adaptive grids

In our multigrid algorithm we use a hierarchy of grids. Starting from a coarsest grid, we construct finer grids by cell-wise refinement, i.e. the cells on the coarse grid can be split into four equal, smaller ones. The refinement of the cells need not be done uniformly: we may refine cells in part of the domain according to some refinement criterion. In this way adaptive grids are obtained.

Our algorithm for solving the discretized equations on the adaptive grids is the dual version of the FAS-FMG algorithm (cf. [2,11]). For the solution of the set of nonlinear equations, obtained after discretization,

$$N_h(q_h) = f_h,$$ (6.1)

we consider the nonlinear coarse grid correction stage of a two-grid algorithm

$$N_H(\tilde{q}_H) = N_H(R_H q_h) + \bar{R}_H(f_h - N_h(q_h)),$$ (6.2)

$$\tilde{q}_h = q_h + P_h(\tilde{q}_H - R_H q_h),$$ (6.3)

where N_H denotes the nonlinear coarse grid operator, P_h the prolongation operator for the solution, and R_H and $\bar{R}_H$ the restriction operators for the solution and the residual, respectively. The coarse grid operator N_H is constructed by discretization on the coarse grid, and on all grids we apply the same method of discretization.

In the dual version of FAS-FMG we rewrite (6.2) in the form

$$N_H(\tilde{q}_H) = \bar{R}_H f_h + \tau_{Hh},$$ (6.4)

where

$$\tau_{Hh} = N_H(R_H q_h) - \bar{R}_H N_h(q_h)$$ (6.5)

denotes the relative truncation error that can be used as an approximation of the local truncation error of the coarse grid discretization. In this way the fine grid is considered as a means of improving the right-hand side of the coarse grid equations; therefore a grid needs not to be refined locally if the relative truncation error is sufficiently small. If this is the case we put $\tau_{Hh} = 0$. In Section 7 we describe how the relative truncation error is used as a refinement criterion.

As usual, the coarse grid problem (6.4) is not solved exactly, but its solution is approximated by a combination of relaxation sweeps and coarse grid corrections on even coarser grids. Only on the coarsest grid the problem is solved accurately. Details on the coarse grid solution procedure are given in Section 8.

By using cell-wise refinement for all or part of the coarse grid cells, we get a nested sequence of approximating subspaces, $V_H \subset V_h$ and $W_H \subset W_h$. In this way the mixed finite element

method induces a natural set of grid transfer operators. For the operators $\bar{R}_H$ and P_h we use these natural prolongations and restrictions. This means that the prolongation for the potentials corresponds to piecewise constant interpolation, and for the fluxes to piecewise linear interpolation in one direction and piecewise constant in the other. The natural restriction $\bar{R}_H$ is the transpose of P_h because the spaces of test and trial functions are identical (cf. Section 3).

A priori there is no reason to assume that these transfer operators P_h and $\bar{R}_H$ are sufficiently accurate. As Vanka relaxation can be considered as a collective Gauss–Seidel procedure applied to the system of equations obtained after elimination of the fluxes (and hence the discrete equivalent of a second-order differential equation), the piecewise constant interpolation for the quantities related to cells seems too inaccurate. However in [17] we have shown by Fourier analysis that the natural transfer operators can be used in combination with Vanka-type relaxation indeed.

It remains to specify the restriction operator R_H for the solution. At coarse grid edges E_H^j that are split into $E_h^{j_1}$ and $E_h^{j_2}$ we require current conservation

$$\sigma_H^j = (R_H \sigma_h)_H^j = \tfrac{1}{2}\left(\sigma_h^{j_1} + \sigma_h^{j_2}\right). \tag{6.6}$$

This choice for R_H implies that we also have current conservation at the green edges: from (6.4) and (6.5) it follows that at convergence of the FAS-FMG algorithm we have $\sigma_H = R_H \sigma_h$. So all currents that flow out of the cells on the fine grid over a green edge, flow into the cells on the coarse grid.

The choice of a restriction operator for the potentials is less straightforward. In principle we could use the L^2-projection of any of the possible variable sets (ψ, ϕ_n, ϕ_p), (ψ, Φ_n, Φ_p) or (ψ, n, p). The use of a restriction based on the Slotboom variables (ψ, Φ_n, Φ_p) is suggested by the discretization. However in numerical experiments we observed that this may lead to coarse grid operators N_H of which the Jacobian matrix is ill-conditioned. Therefore we consider the other two possibilities.

For the semiconductor equations (2.5) without any scaling, the residual of the conservation law (3.1b) for the continuity equations corresponds to the rate-of-change in the carrier concentrations. Without row scaling this means that the size of the residuals varies widely in magnitude throughout the domain. It may also happen that the diagonal elements of the Jacobian matrices differ by orders of magnitude between a father cell and its four kid cells, especially if the transition between the n- and p-region is not properly resolved on the coarse grid. In this case a small residual (after row scaling) on the fine grid may result in a large correction on the coarse grid. For the 1D case, De Zeeuw [7] proposed to introduce a residual damping operator D_H in the coarse grid operator (6.2). This D_H is a diagonal matrix with entries in $[0,1]$ which are determined by comparing the diagonal elements of the coarse and fine grid Jacobian matrices. The modified coarse grid equation then reads (cf. 6.2):

$$N_H(\tilde{q}_H) = N_H(R_H q_h) + D_H(R_H q_h, q_h)\bar{R}_H(f_h - N_h(q_h)). \tag{6.7}$$

The elements of D_H differ from 1 only in small parts of the domain Ω (the transition regions), and the effect of damping is compensated in these regions by additional relaxation on the fine grid. This technique has been used successfully for 2D problems (cf. [18]).

This discussion makes clear that it is attractive to use a restriction that leads to coarse and fine grid Jacobians, the corresponding diagonal entries of which are of comparable magnitude.

The restriction based on the L^2-projection of the variables (ψ, n, p) seems to have this property: by rewriting (2.5) in the form

$$-\operatorname{div}(\mu_n n \operatorname{grad} \phi_n) = +R,$$
$$-\operatorname{div}(\mu_p p \operatorname{grad} \phi_p) = -R,$$

(6.8)

we see that the diagonal elements of the Jacobian matrices with respect to the variable set (ψ, ϕ_n, ϕ_p) should be of comparable magnitude on the coarse and the fine grid, because the concentrations (n, p) are of comparable magnitude in the corresponding cells on the coarse and the fine grid. Indeed in numerical experiments we observe that the diagonal elements of the damping operator D_H are all equal to 1. Unfortunately we also observe that the coarse grid matrix tends to ill-conditioning in cases close to thermal equilibrium. We think that this is due to the following: for thermal equilibrium we have the trivial fine grid solution $\phi^i_{n,h} = \phi^i_{p,h} = 0$. In a coarse grid cell $\Omega^I_H = \bigcup_i \Omega^i_h$, $i = 1, \ldots, 4$, which has some kid cells in the n-region and others in the p-region, we find

$$n^I_H p^I_H = \left(\tfrac{1}{4} \sum_i n^i_h\right)\left(\tfrac{1}{4} \sum_i p^i_h\right) \gg 1,$$

which implies that $\phi^I_{n,H} \neq \phi^I_{p,H}$, so we get nonzero values for $\phi^I_{n,H}$ and/or $\phi^I_{p,H}$ on the coarse grid which is unphysical.

Therefore we propose a restriction that is based on the L^2-projection of the variables (ψ, ϕ_n, ϕ_p). In numerical experiments we observe (Section 8) that this choice yields a multigrid algorithm that is both robust and efficient, although the introduction of the residual damping operator D_H is necessary. The precise construction of D_H is described in [18].

We complete this section by describing the treatment of the "green" edges that appear on a partially refined grid. An edge is called green if it is not part of a physical boundary, and if the adjacent cells have a different level of refinement (cf. [18,24]). A straightforward approach is to impose inhomogeneous Neumann boundary conditions at the green edges on the fine mesh, as σ^j_H is given on the coarse mesh (cf. [24]). However this can lead to patches on the fine grid that have only Neumann boundary conditions, so that the solution u_h is only determined up to an arbitrary constant. As there is no way to fix this constant for the semiconductor device equations, we have to impose Dirichlet boundary conditions at (at least some of) the green interfaces.

The Lagrange multipliers λ_h, as introduced in Section 3, are a good approximation of the potentials u_h at the edges. So at a "green" edge E^j_h, which is part of the coarse grid edge E^J_H we use the Lagrange multiplier λ^J_H as a Dirichlet boundary condition on the fine grid. As noticed before, the flux σ^j_h at the green interface E^j_h is still a variable, so also on partially refined grids we have discrete current conservation due to the choice of the restriction operator for the fluxes σ_h.

7. Adaptive mesh generation

As was seen in the previous section, it is fully consistent with the FAS-FMG method to use the relative truncation error as a refinement criterion in an adaptive mesh refinement strategy.
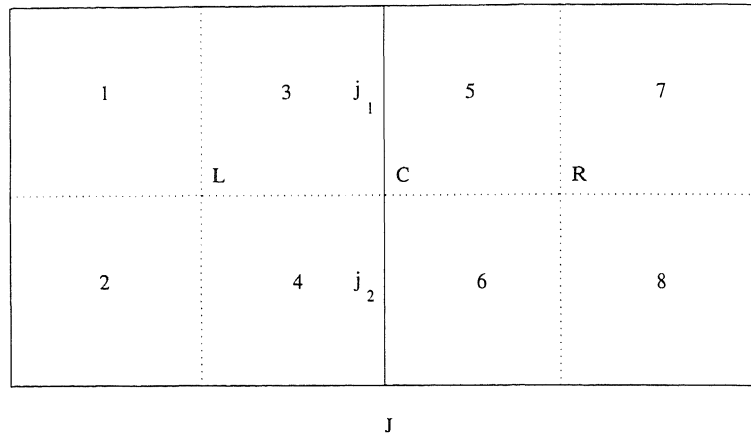
J

Fig. 1. Numbering of cells used in calculation of relative truncation error.

Therefore, before describing the adaptive mesh generation, we study the relative truncation error more closely.

Both for the equations related to the cells and to the walls relative truncation errors are defined. These are denoted by $\tau_{Hh}(\Omega_H^I)$ and $\tau_{Hh}(E_H^J)$, respectively. Using the definitions of the grid transfer operators given in Section 6, we find for the general problem (3.4), with a source f which is piecewise constant, that $\tau_{Hh}(\Omega_H^I)$ in a cell Ω_H^I, $\overline{\Omega}_H^I = \bigcup_{i=1,4}\overline{\Omega}_h^i$, is given by (cf. 3.6)

$$\tau_{Hh}(\Omega_H^I) = f(x_H^I) - \tfrac{1}{4}\sum_{i=1,4} f(x_h^i),\tag{7.1}$$

where x_h^i denotes the center of Ω_h^i. If lumping is used we obtain for $\tau_{Hh}(E_H^J)$ with $E_H^J = E_h^{j_1}\cup E_h^{j_2}$ (for the numbering of the cells see Fig. 1)

$$\tau_{Hh}(E_H^j) = \tilde{a}_H^j\frac{u_H^R - u_H^L}{2h} - \frac{1}{2}\left(\tilde{a}_h^{j_1}\frac{u_h^5 - u_h^3}{h} + \tilde{a}_h^{j_2}\frac{u_h^6 - u_h^4}{h}\right),\tag{7.2}$$

with

$$\tilde{a}_H^J = \tilde{a}(e_H^J,\ e_H^J),\quad \text{etc.}$$

$$u_H^L = \tfrac{1}{4}\sum_{i=1,4} u_h^i,\qquad u_H^R = \tfrac{1}{4}\sum_{i=5,8} u_h^i,\tag{7.3}$$

nd $2h$ the distance between x_H^L and x_H^R. Due to the choice of the restriction operator R_H for the fluxes (6.6), σ_h does not appear in the relative truncation errors.

The following two theorems state the order behavior of the relative truncation errors on uniform grids in the limit case of vanishing mesh width for the three semiconductor equations. We assume that $(\psi,\ \phi_n,\ \phi_p)\in(C^3(\overline{\Omega}))^3$ and that $(\psi^i,\ \phi_n^i,\ \phi_p^i)$ are the averages of $(\psi,\ \phi_n,\ \phi_p)$ over the cell Ω^i.

Theorem 7.1. *If the nonlinear source f in (3.1b) satisfies $f(x,\ \psi,\ \phi_n,\ \phi_p)\in C^2(\overline{\Omega}\times\mathbb{R}^3)$, then for all three equations we have for $h\to 0$*

$$|\tau_{Hh}(\Omega_H^I)|\leqslant Ch^2\,|f(x,\ \psi,\ \phi_n,\ \phi_p)|_{2,\infty,\Omega_H^I}.\tag{7.4}$$

Proof. Using a Taylor expansion around x_H^I, the center of Ω_H^I, we obtain from (7.1)

$$| \tau_{Hh}(\Omega_H^I) | \leq \left| f\left(x_H^I, \psi_H^I, \phi_{n,H}^I, \phi_{p,H}^I\right) \right.$$

$$- \left(f\left(x_H^I, \psi_H^I, \phi_{n,H}^I, \phi_{p,H}^I\right) + \tfrac{1}{4} \sum_{i=1,4} \left(x_h^i - x_H^I\right) \cdot \operatorname{grad} f\left(x_H^I\right) \right.$$

$$\left. \left. + \tfrac{1}{4} \sum_{\phi=\psi,\phi_n,\phi_p} \frac{\partial f}{\partial \phi} \sum_{i=1,4} \left(x_h^i - x_H^I\right) \cdot \operatorname{grad} \phi\left(x_H^I\right) \right) \right|$$

$$+ Ch^2 \, | \, f(x, \psi, \phi_n, \phi_p) \, | \,_{2,\infty,\Omega_H^I}$$

$$= Ch^2 \, | \, f(x, \psi, \phi_n, \phi_p) \, | \,_{2,\infty,\Omega_H^I}.$$

The last equality follows from a symmetry argument. □

We notice that the requirement for the source term f in Theorem 7.1 holds for the Shockley–Read–Hall recombination model, as well as for the Auger model, but it excludes the avalanche-generation term modeling impact ionization.

In the next theorem we give the order behavior of the truncation error at walls $\tau_{Hh}(E_H^J)$.

Theorem 7.2. *For Poisson's equation we have*

$$| \tau_{Hh}(E_H^J) | \leq Ch^2, \tag{7.5a}$$

and for the continuity equations for the electrons and holes

$$| \tau_{Hh}(E_H^J) | \leq Ch^2 \exp(\psi^C - \phi_n^C), \tag{7.5b}$$

$$| \tau_{Hh}(E_H^J) | \leq Ch^2 \exp(\phi_p^C - \psi^C), \tag{7.5c}$$

respectively, where C is bounded by the supremum norm of the first, second and third order derivatives of (ψ, ϕ_n, ϕ_p).

Proof. We only give a proof for the continuity equation for holes. The length of E_H^J is denoted by $2k$ and the mesh size perpendicular to E_H^J by $2h$. By using Taylor expansions of ψ and ϕ_p around x^C (see Fig. 1) and (7.2) and (7.3), we obtain for h small enough

$$| \tau_{Hh}(E_H^J) | = C \left| \frac{\psi_H^R - \psi_H^L}{\exp(\psi_H^R) - \exp(\psi_H^L)} \frac{\exp(\phi_{p,H}^R) - \exp(\phi_{p,H}^L)}{h} \right.$$

$$\left. - \left(\frac{\psi_h^5 - \psi_h^3}{\exp(\psi_h^5) - \exp(\psi_h^3)} \frac{\exp(\phi_{p,h}^5) - \exp(\phi_{p,h}^3)}{h} \right) \right.$$

$$+ \frac{\psi_h^6 - \psi_h^4}{\exp\!\left(\psi_h^6\right) - \exp\!\left(\psi_h^4\right)} \; \frac{\exp\!\left(\phi_{p,h}^6\right) - \exp\!\left(\phi_{p,h}^4\right)}{h} \Bigg) \Bigg|$$

$$\leqslant C \frac{\exp\!\left(\phi_p^C - \psi^C\right)}{h} \, \big| (1 + Ch^2)(1 + Ch^2)(1 + Ch^2)\!\left(h\partial_x \phi_p^C + Ch^3\right)$$

$$- \left(1 + \tfrac{1}{4}k\partial_y \phi_p^C + Ch^2\right)\!\left(1 - \tfrac{1}{4}k\partial_y \psi^C + Ch^2\right)(1 + Ch^2)$$

$$\times \left(\tfrac{1}{2}h\partial_x \phi_p^C + \tfrac{1}{4}hk\partial_{xy}^2 f_p^C + Ch^3\right)$$

$$- \left(1 - \tfrac{1}{4}k\partial_y \phi_p^C + Ch^2\right)\!\left(1 + \tfrac{1}{4}k\partial_y \psi^C + Ch^2\right)(1 + Ch^2)$$

$$\times \left(\tfrac{1}{2}h\partial_x \phi_p^C - \tfrac{1}{4}hk\partial_{xy}^2 \phi_p^C + Ch^3\right) \big|$$

$$\leqslant Ch^2 \, \exp\!\left(\phi_p^C - \psi^C\right). \qquad \square$$

Our adaptive mesh refinement scheme aims at equidistribution of the relative truncation errors; so from Theorems 7.1 and 7.2 we conclude that it makes sense to refine the mesh in areas where the relative truncation errors are large.

In all finest cells and walls we define error indicators $\eta(\Omega_h^i, \phi)$ and $\eta(E_h^j, \phi)$, with $\phi = \psi$, ϕ_n, ϕ_p, by the relative truncation error in the parent cell and wall, respectively. Next the three error indicators are merged by a summation of the normalized values, so we obtain single error indicators $\eta(\Omega_h^i)$, $\eta(E_h^j) \in [0, 1]$, for all cells and walls:

$$\eta\!\left(\Omega_h^i\right) = \frac{1}{3} \sum_\phi \frac{\left| \eta\!\left(\Omega_h^i, \phi\right) \right|}{\overline{\eta}_\Omega(\phi)}, \qquad \eta\!\left(E_h^j\right) = \frac{1}{3} \sum_\phi \frac{\left| \eta\!\left(E_h^j, \phi\right) \right|}{\overline{\eta}_E(\phi)}, \qquad (7.6)$$

with, respectively,

$$\overline{\eta}_\Omega(\phi) = \max_i \left| \eta\!\left(\Omega_h^i, \phi\right) \right|, \qquad \overline{\eta}_E(\phi) = \max_j \left| \eta\!\left(E_j^h, \phi\right) \right|. \qquad (7.7)$$

If the relative truncation errors are uniformly distributed then the error indicators are all equal to 1.

The actual mesh refinement procedure consists of two steps. In the first step we refine all cells and walls of which the error indicators $\eta(\Omega_h^i)$ and $\eta(E_h^j)$ are larger than user-defined parameters δ_Ω and δ_E, respectively; a wall is refined by refining both cells adjacent to it. In the second step we add some additional refinements in order to maintain a certain grid regularity: a cell is split if at least three of its neighbors are split, and the parent wall of a green wall is refined if it is also green.

In the next section we use this grid adaptation scheme in practical calculations.

8. Numerical experiments

As a test problem for our multigrid algorithm we use a bipolar *npn*-transistor from the CURRY example set [14]. Fig. 2 gives a schematic view of the geometry of the transistor and
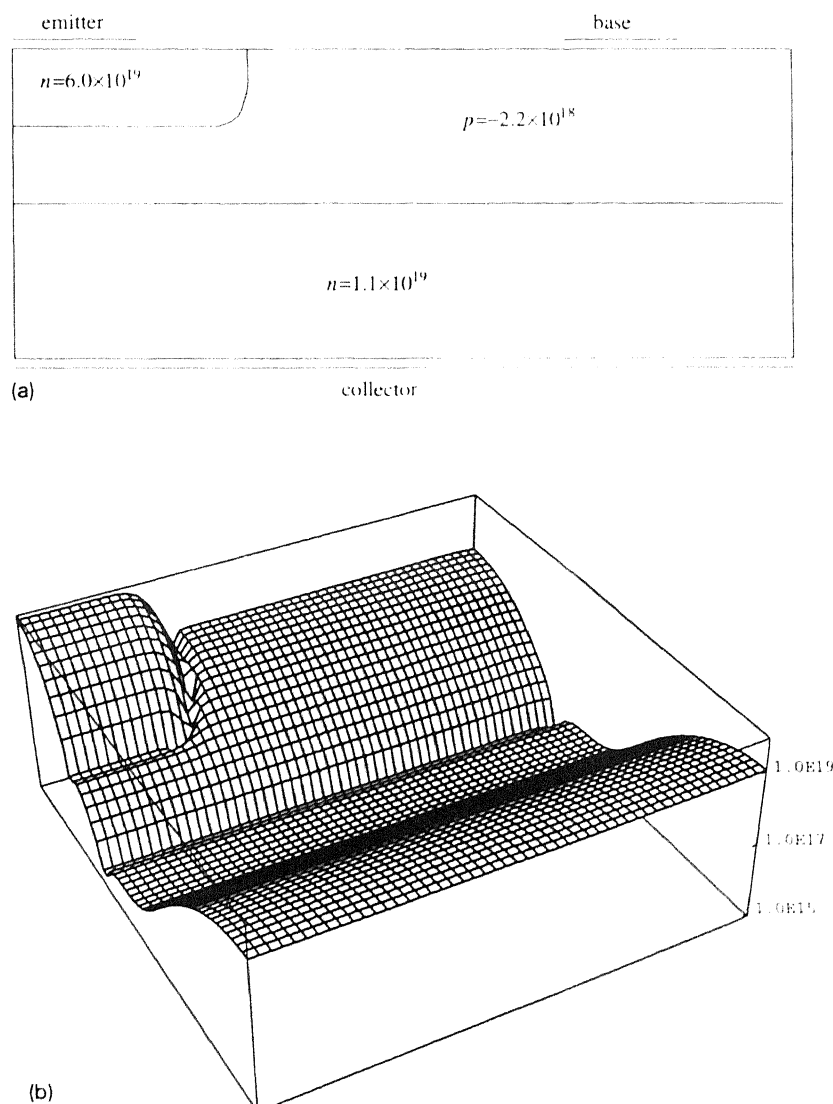
emitter   base

$n=6.0 \times 10^{19}$

$p=-2.2 \times 10^{18}$

$n=1.1 \times 10^{19}$

(a)   collector

(b)

Fig. 2. (a) Geometry and (b) doping profile of transistor.

the doping profile. The length of the device is 20 μm and the width is 8 μm; for a precise description of the doping profile we refer to [14]. The generation–recombination rate is given by the SRH model (2.2), with carrier lifetimes $\tau_p = \tau_n = 10^{-6}$ s. The applied voltages at the collector and the base are kept constant at $V_c = 1.0$ V and $V_b = 0.0$ V, respectively. Starting from $V_c = -0.50$ V, the applied voltage at the emitter is lowered during the simulation to $V_c = -0.80$ V in steps of 0.05 V.

The coarsest grid used in our calculation consists of 4×10 squares. In fact this mesh is too coarse to resolve the contacts properly. An obvious generalization of the discretization (Section 3) is used to treat these parts of the boundary. It is assumed that the currents through such a

Table 1
Collector currents (A/cm)

V_e	MFEM uniform mesh			Reference solution
	16×40	32×80	64×160	56×62
−0.50	9.5×10^{-5}	1.4×10^{-5}	1.0×10^{-5}	9.8×10^{-6}
−0.55	5.8×10^{-4}	9.5×10^{-5}	7.0×10^{-5}	6.7×10^{-5}
−0.60	3.4×10^{-4}	6.4×10^{-4}	4.8×10^{-4}	4.6×10^{-4}
−0.65	1.8×10^{-2}	4.3×10^{-3}	3.3×10^{-3}	3.1×10^{-3}
−0.70	8.4×10^{-2}	2.8×10^{-2}	2.2×10^{-2}	2.1×10^{-2}
−0.75	3.2×10^{-1}	1.7×10^{-1}	1.4×10^{-1}	1.3×10^{-1}
−0.80	$1.1 \times 10^{+0}$	7.9×10^{-1}	7.1×10^{-1}	6.9×10^{-1}

boundary edge, which are determined by the Dirichlet boundary conditions and the potentials in the adjacent cell, only flow through that part of the edge that is covered by the contact. This approach makes it possible to use rather coarse and regular coarsest grids in our calculations, even if tiny contacts are present. After a sufficient number of refinements is made on the finer grids these contacts are properly resolved.

The continuation of the boundary conditions happens on the coarsest mesh. We start by solving the thermal equilibrium case, i.e. no applied voltages. After changing the boundary conditions we solve the problem on the coarsest grid using the previously obtained solution as initial iterate. The solution procedure on the coarsest grid consists of a combination of Vanka-type relaxation sweeps and Newton steps; the relaxation sweeps are introduced to make the solution procedure more robust (cf. [12]). The new coarse grid solution is then interpolated to a next finer grid using the prolongation operator described in Section 6. This fine grid approximation is improved iteratively by a few W-cycles. Due to the robustness of the solution procedure on the coarsest grid we are able to take large steps in the continuation process. We have also tried the ramping procedure proposed by Edwards et al. [9,18], which aims at keeping the currents and the majority concentrations constant during the continuation of the applied voltage. This appeared to give no improvement, which—we believe—is due to the fact that the urrents do change when the applied voltages are altered.

Table 1 shows the collector currents that are computed on the different grids, together with a reference solution computed with the CURRY package on a non-uniform 56×62 grid. It appears that the collector currents converge at least linearly (cf. Section 4) if the mesh size decreases. To estimate the convergence rate of the multigrid algorithm we introduce the average reduction factor ρ,

$$\rho = \left(\frac{d^{(10)}}{d^{(0)}} \right)^{1/10}, \tag{8.1}$$

where $d^{(i)}$ denotes the maximum of the scaled residual after i FAS-sweeps. The residual is scaled point-wise by means of the diagonal 3×3 blocks of the Jacobian matrix: thus the scaled residual corresponds to corrections that would occur in a point-wise collective Jacobi relaxation. The maximum of this scaled residual is taken over the grid and over the three variables

Table 2
Average residual reduction factor ρ for W-cycles

V_e	point-Vanka			line-Vanka		
	16×40	32×80	64×160	16×40	32×80	64×160
-0.50	0.17	0.21	0.18	0.06	0.12	0.11
-0.55	0.18	0.20	0.18	0.07	0.13	0.12
-0.60	0.18	0.20	0.17	0.07	0.13	0.12
-0.65	0.17	0.20	0.17	0.08	0.13	0.12
-0.70	0.20	0.21	0.18	0.15	0.17	0.12
-0.75	0.31	0.24	0.17	0.21	0.21	0.12
-0.80	0.43	0.24	0.18	0.31	0.23	0.16

(ψ, ϕ_n, ϕ_p). Every FAS-sweep consists of a W-cycle: it appears that V-cycles are less robust for the semiconductor problem (cf. [12,18]).

Table 2 shows the average reduction factor ρ for different grids and both for symmetric point-Vanka and for alternating line-Vanka relaxation. We observe that the use of line-Vanka relaxation leads to a more efficient algorithm. The convergence behavior is not really grid-independent (in some cases it appears that the convergence is faster on finer grids!), but in all cases the convergence is fast, and only a few iterations are necessary to attain truncation error accuracy.

Finally, we demonstrate the use of adaptive grids. Starting from the coarsest 4×10 grid, we add a single level of uniform refinement, which is necessary to estimate the relative truncation error on the coarsest grid. After solving the discrete equations on the first two grids, we refine the grid adaptively. Fig. 3 shows an example of a grid generated by the adaptive procedure. The finest level corresponds to a uniform 64×160 grid. This adaptive grid is obtained for $V_e = -0.80$ V with $\delta_\Omega = \delta_E = 0.1$.

It is clear that small values for δ_Ω and δ_E give rise to relatively fine grids, which means that the solution is more accurate at the expense of more computational work. Table 3 shows this tradeoff between the number of cells in the grid and the accuracy of the discrete solution
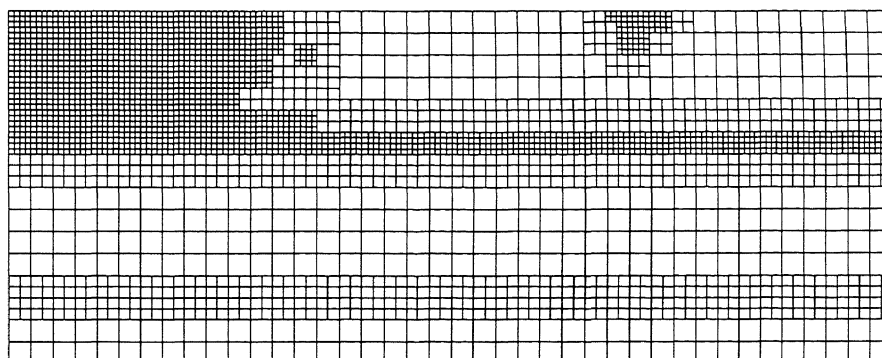

Fig. 3. Example of adaptive grid.

Table 3
Results for adaptive grids corresponding to a uniform 64×160 grid

$\delta_\Omega = \delta_E$	Number of fine cells	Relative error collector currents
0.01	5542	0.002
0.05	3007	0.012
0.10	1951	0.017
0.15	1909	0.073

calculated on the adaptive grid. The accuracy of the solution on the adaptive grid is measured by the relative error in the collector currents

$$\delta I_c(V_e) = \left| \frac{I_{c,U} - I_{c,A}}{I_{c,U}} \right|, \tag{8.2}$$

where $I_{c,A}$, $I_{c,U}$ denote the collector currents on the adaptive grid and the corresponding uniform grid, respectively. In fact Table 3 shows the maximum value of $\delta I_c(V_e)$, and the maximum number of fine cells in the adaptive grid for the series of applied voltages at the emitter. In all cases we took $\delta_\Omega = \delta_E$. From Table 3 we conclude that we do indeed save a substantial amount of work by using adaptive grids.

9. Conclusions

We have developed an adaptive multigrid algorithm for the iterative solution of the system of nonlinear equations resulting from the mixed finite element discretization of the 2D stationary semiconductor equations. We have changed the discretization by using a quadrature rule, in order to obtain the Scharfetter–Gummel discretization. However we prove that the use of quadrature does not spoil the accuracy of the discrete approximations.

Our multigrid algorithm for the iterative solution of the systems of nonlinear equations is based on a Vanka-type relaxation. In numerical experiments it appears that the line-wise version is more efficient than the point-wise relaxation. The use of extremely coarse grids in the multigrid procedure is problematic. However, we have shown that by using special grid transfer operators these problems can be overcome for rather coarse grids.

Our adaptive grid refinement procedure, based on the relative truncation error, is natural within the framework of the multigrid algorithm. The usefulness of this method is demonstrated by means of a numerical experiment.

References

[1] R.E. Bank, D.J. Rose and W. Fichtner, Semiconductor device simulation, SIAM J. Sci. Statist. Comput. 4 (1983) 391–435.
[2] A. Brandt, Guide to multigrid development, in: W. Hackbusch and U. Trottenberg, eds., Multigrid Methods Lecture Notes in Mathematics 960 (Springer-Verlag, Berlin, 1982) 220–312.

[3] F. Brezzi, On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers, *RAIRO Numer. Anal.* 8 (1974) 129–151.

[4] F. Brezzi, L.D. Marini and P. Pietra, Two-dimensional exponential fitting and applications to drift-diffusion models, *SIAM J. Numer. Anal.* 26 (1989) 1342–1355.

[5] R. Constaple, (1990). MAGIC—Ein zweidimensionaler Bauelemente-Simulator auf der Basis von Mehrgitter-Algorithmen, in: W. Joppich, ed., *Device Simulation*, GMD-Studien 177 (GMD, 1990) 63–75.

[6] K. Deljouie-Rakhshandeh, A self-adaptive approach for numerical device simulation, in: *Proceedings SISDEP-88*, Bologna, Italy (1988) 519–527.

[7] P.M. De Zeeuw, Nonlinear multigrid applied to a one-dimensional stationary semiconductor model, *SIAM J. Sci. Statist. Comput.* 13 (1992) 512–530.

[8] J. Douglas and J.E. Roberts, Mixed finite element methods for second order elliptic problems, *Mat. Apl. Comput.* 1 (1982) 91–103.

[9] S.P. Edwards, A.M. Howland and P.J. Mole, Initial guess strategy and linear algebra techniques for a coupled two-dimensional semiconductor equation solver, in: *Proceedings NASECODE IV* (Boole Press, Dublin, 1985) 272–280.

[10] M. Fortin, An analysis of the convergence of mixed finite element methods, *RAIRO Numer. Anal.* 11 (1977) 341–354.

[11] P.W. Hemker, On the structure of an adaptive multi-level algorithm, *BIT* 20 (1980) 289–301.

[12] P.W. Hemker, A nonlinear multigrid method for one-dimensional semiconductor device simulation: results for the diode, *J. Comput. Appl. Math.* 30 (1990) 117–126.

[13] P.W. Hemker and J. Molenaar, An adaptive multigrid approach for the solution of the 2D semiconductor equations, in: W. Hackbusch and U. Trottenberg, eds., *International Series of Numerical Mathematics* 98 (Birkhäuser, Basel, 1991).

[14] C. Lepoeter, CURRY example set, Technical Report No. 4322.271.6005, Philips, CAD Centre, Eindhoven, (1987).

[15] J.F. Maitre, F. Musy and P. Nigon (1985). A fast solver for the Stokes equation using multigrid with a Uzawa smoother, in: D. Braess, W. Hackbusch and U. Trottenberg eds., *Notes on Numerical Fluid Dynamic* 11 (Braunschweig, 1985) 77–83.

[16] P.A. Markowich, *The Stationary Semiconductor Device Equations* (Springer-Verlag, Wien, New York, 1986).

[17] J. Molenaar, A two-grid analysis of the combination of mixed finite elements and Vanka-type relaxation, Report NM-R9102, Centre for Mathematics and Computer Science, Amsterdam (1991).

[18] J. Molenaar and P.W. Hemker A multigrid approach for the solution of the 2D semiconductor equations, *IMPACT* 2 (1990) 219–243.

[19] S.J. Polak, C. den Heijer, W.H.A. Schilders and P. Markowich, Semiconductor device modelling from the numerical point of view, *Internat. J. Numer. Methods Engrg.* 24 (1987) 763–838.

[20] S.J. Polak, W.H.A. Schilders and H.D. Couperus, A finite element method with current conservation, in: *Proceedings SISDEP-88*, Bologna, Italy (1988) 453–462.

[21] P.A. Raviart and J.M. Thomas, A mixed finite element method for second order elliptic problems, in: *Mathematical Aspects of the Finite Element Method*, Lecture Notes in Mathematics 606 (Springer-Verlag, Berlin, 1977).

[22] D.L. Scharfetter and H.K. Gummel, Large-signal analysis of a silicon read diode oscillator, *IEEE Trans. Electron. Devices* 16 (1969) 64–77.

[23] W.H.A. Schilders, A novel approach to adaptive meshing for the semiconductor problem, in: *Proceedings SISDEP-88*, Bologna, Italy (1988) 519–527.

[24] G.H. Schmidt and F.J. Jacobs, Adaptive local grid refinement and multi-grid in numerical reservoir simulation, *J. Comput. Phys.* 77 (1988) 140–165.

[25] S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer-Verlag, Wien, 1984).

[26] S.P. Vanka, Block-Implicit Multigrid Solution of Navier-Stokes Equations in Primitive Variables, *J. Comput. Phys.* 65 (1986) 138–158.