An Iterative Method for Adaptive Finite Element Solutions of An Energy Transport Model of Semiconductor Devices ⋆

Ren-Chuen Chen [a]

[a] *Department of Applied Mathematics, National Chiao Tung University, Hsinchu, Taiwan.*

Jinn-Liang Liu [b,*]

[b] *Department of Applied Mathematics, National Chiao Tung University, Hsinchu, Taiwan.*

Abstract

A self-adjoint formulation of the energy transport model of semiconductor devices is proposed. This new formulation leads to symmetric and monotonic properties of the resulting system of nonlinear algebraic equations from an adaptive finite element approximation of the model. A node-by-node iterative method is then presented for solving the system. This is a globally convergent method that does not require the assembly of the global matrix system and full Jacobian matrices. An adaptive algorithm implementing this method is described in detail to illustrate the main features of this paper, namely, adaptation, node-by-node calculation, and global convergence. Numerical results of simulations on deep-submicron diode and MOSFET device structures are given to demonstrate the accuracy and efficiency of the algorithm.

1. INTRODUCTION

Computer-aided simulation is one of the important processes in developing semiconductor devices. Numerical methods for the fundamental semiconductor equations play a significant role in this development. For most practical device structures, the electrostatic potential, carrier concentrations, and carrier temperatures exhibit extreme layers or peaks, particularly in the neighborhood of p-n junctions and the oxide [7,24]. Presence of this kind of singular

⋆ This work was supported by NSC under grant 92-2115-M-009-006, Taiwan.

* Corresponding author
Email address: jinnliu@math.nctu.edu.tw (Jinn-Liang Liu).

phenomena implies that adaptive mesh generation of unstructured grids is inevitable if an accurate and efficient device simulation platform is required.

To obtain numerical solutions of semiconductor equations, one must solve a system of nonlinear algebraic equations resulting from a discretization by, for example, the finite element method (FEM). The standard method for the solution is Newton's method or its variant. Newton's method is a local method that converges quadratically in a sufficiently small neighborhood of the exact solution. Although Newton's method has been dominantly used in device simulations [4,36], it is very sensitive to initial guesses due to its local convergence property. In practical simulation, the device terminal characteristics of I-V curves (i.e., I-V points) is usually of interest. A conventional approach to obtain these curves is by the continuation method from lower to higher biases by Newton's method, which can be very costly in terms of computing time and human work load associated with the convergence problems of the method.

We propose here a global iterative method for the energy transport model for which a new self-adjoint formulation of carrier energy balance equations is introduced. This formulation is purely mathematical rather than physical and is motivated by the transformation of carrier densities to the Slotboom variables. It leads to symmetric and monotonic properties of the resulting system of nonlinear algebraic equations from FE approximation. These properties provide several advantageous features for device simulation. First of all, the iterative method is globally and monotonically convergent with simple upper or lower solutions of the self-adjoint semilinear PDEs as initial guesses. This allows us to have a simultaneous (parallel) computing of multiple I-V points with various biasing conditions and with independent constant initial guesses for each I-V point calculation. The computational effort can thus be dramatically reduced [20]. Secondly, the solution procedure can be performed in a node-by-node manner that does not require the assembly of the global matrix system and full Jacobian matrices. Thirdly, the method does not produce non-physical negative values for the minority carrier concentration under heavy recombination [32,36] since the stiffness matrices are diagonally dominant. Finally, its implementation is considerably simpler than that of Newton's method since it essentially depends only on one crucial component, i.e., the monotone parameter matrix which is a simplified diagonal Jacobian. An adaptive algorithm implementing this method is also described in detail to illustrate the main features of this paper, namely, adaptation, node-by-node calculation, and global convergence. The algorithm is based on the general framework proposed in [21] and on the object-oriented programming (OOP) prototype developed in [22].

This paper is organized as follows. The energy transport model considered herein is stated in Sec. 2. In Sec. 3, we introduce the self-adjoint formulation

of carrier energy balance equations. In Sec. 4, we first analyze the structure of the stiffness matrix of adaptive finite element systems for the Laplace equation that leads to diagonal dominance feature of the resulting matrices of the model problem. Starting with the upper and lower solutions as initial guesses, it is also shown in this section that the iterative method for the solution of the nonlinear algebraic systems is globally convergent. We then summarize in Sec. 6 our implementation of a complete solution process into the adaptive algorithm which consists mainly of 1-irregular mesh refinement, Gummel's decoupling scheme, the iterative method, and error estimation. Sec. 7 represents a part of our extensive numerical experiments on various deep-submicron diode and MOSFET device structures to demonstrate the accuracy and efficiency of the algorithm.

2. THE ENERGY TRANSPORT MODEL

A variety of energy transport models have been developed in the literature [36]. In particular, we use the following model which is also considered in [1,2,9,12,13,19]

$$\Delta\phi = \frac{q}{\varepsilon_s}(n - p + N_A^- - N_D^+), \tag{1}$$

$$\frac{1}{q}\nabla \cdot \mathbf{J}_n = R, \tag{2}$$

$$\frac{1}{q}\nabla \cdot \mathbf{J}_p = -R, \tag{3}$$

$$\nabla \cdot \mathbf{S}_n = \mathbf{J}_n \cdot \mathbf{E} - n\left(\frac{\omega_n - \omega_0}{\tau_{n\omega}}\right), \tag{4}$$

$$\nabla \cdot \mathbf{S}_p = \mathbf{J}_p \cdot \mathbf{E} - p\left(\frac{\omega_p - \omega_0}{\tau_{p\omega}}\right), \tag{5}$$

where ϕ is the electrostatic potential, n and p are the electron and hole concentrations, q is the elementary charge, ε_s is the permittivity constant of semiconductor, N_A^- and N_D^+ are the densities of ionized impurities, $\mathbf{J}_n$ and $\mathbf{J}_p$ are the current densities, R is the function describing the balance of generation and recombination of electrons and holes, $\mathbf{S}_n$ and $\mathbf{S}_p$ are the energy fluxes for carriers, $\mathbf{E}$ is the electric field, $\tau_{n\omega}$ and $\tau_{p\omega}$ are the carrier energy relaxation times, ω_0 is the thermal energy, and ω_n and ω_p are the carrier average energies. These physical variables are tightly coupled together with the following auxiliary relationships

$$\mathbf{E} = -\nabla\phi, \tag{6}$$

$$\mathbf{J}_n = -q\mu_n n\nabla\phi + qD_n\nabla n = -qn\mathbf{v}_n, \tag{7}$$

$$\mathbf{J}_p = -q\mu_p p\nabla\phi - qD_p\nabla p = qp\mathbf{v}_p, \tag{8}$$

$$\mathbf{S}_n = \frac{\mathbf{J}_n}{-q}\omega_n + \frac{\mathbf{J}_n}{-q}k_B T_n + \mathbf{Q}_n, \tag{9}$$

$$\mathbf{S}_p = \frac{\mathbf{J}_p}{+q}\omega_p + \frac{\mathbf{J}_p}{+q}k_B T_p + \mathbf{Q}_p, \tag{10}$$

$$\omega_0 = \frac{3}{2}k_B T_L, \tag{11}$$

$$\omega_n = \frac{3}{2}k_B T_n + \frac{1}{2}m_n^* \left|\mathbf{v}_n\right|^2, \tag{12}$$

$$\omega_p = \frac{3}{2}k_B T_p + \frac{1}{2}m_p^* \left|\mathbf{v}_p\right|^2, \tag{13}$$

$$\mathbf{Q}_n = -\kappa_n \nabla T_n, \tag{14}$$

$$\mathbf{Q}_p = -\kappa_p \nabla T_p, \tag{15}$$

$$\kappa_n = 2(\frac{k_B}{q})^2 nq\mu_n T_L, \tag{16}$$

$$\kappa_p = 2(\frac{k_B}{q})^2 pq\mu_p T_L, \tag{17}$$

$$R = \frac{np - n_i^2}{\tau_n^0 \left(p + p_T\right) + \tau_p^0 \left(n + n_T\right)} \tag{18}$$

where $\mathbf{Q}_n$ and $\mathbf{Q}_p$ are the heat fluxes for carries, k_B is Boltzmann's constant, T_n, T_p, and T_L are the electron, hole and lattice temperatures, μ_n and μ_p are the field-dependent electron and hole mobilities, D_n and D_p are the electron and hole diffusion coefficients expressed by the Einstein relation with the mobilities, m_n^* and m_p^* are the electron and hole effective masses, $\mathbf{v}_n$ and $\mathbf{v}_p$ are the electron and hole velocities, κ_n and κ_p are the electron and hole heat conductivities, and (18) is the Shockley-Read-Hall generation-recombination model with n_i being the intrinsic carrier concentration, τ_n^0 and τ_p^0 the electron and hole lifetimes, and p_T and n_T the electron and hole densities associated with energy levels of the traps. In the above equations, vectors are denoted by bold letters.

Based on Boltzmann statistics [16,35], the convectional drift-diffusion (DD) model (1)-(3) can be written as

$$\Delta\phi = F\left(\phi, u, v\right), \tag{19}$$

$$\nabla \cdot (D_n n_i \exp(\frac{\phi}{V_T})\nabla u) = R\left(\phi, u, v\right), \tag{20}$$

$$\nabla \cdot (D_p n_i \exp(\frac{-\phi}{V_T})\nabla v) = R\left(\phi, u, v\right), \tag{21}$$

where $V_T = (k_B T_L)/q$ is the thermal voltage,

$$u = \exp(\frac{-\varphi_n}{V_T}), \tag{22}$$

$$v = \exp(\frac{\varphi_p}{V_T}), \tag{23}$$

are the Slotboom variables in which the quasi-Fermi potentials φ_n and φ_p are expressed as

$$n = n_i \exp(\frac{\phi - \varphi_n}{V_T}), \tag{24}$$

$$p = n_i \exp(\frac{\varphi_p - \phi}{V_T}), \tag{25}$$

$$\mathbf{J}_n = qD_n n_i \exp(\frac{\phi}{V_T})\nabla u = -qn\mu_n \nabla\varphi_n, \tag{26}$$

$$\mathbf{J}_p = -qD_p n_i \exp(\frac{-\phi}{V_T})\nabla v = -qp\mu_p \nabla\varphi_p, \tag{27}$$

and

$$F(\phi, u, v) = \frac{qn_i}{\varepsilon_s}(u\exp(\frac{\phi}{V_T}) - v\exp(\frac{-\phi}{V_T})) + \frac{q(N_A^- - N_D^+)}{\varepsilon_s}, \tag{28}$$

$$R(\phi, u, v) = \frac{n_i^2(uv - 1)}{\tau_n^0\left(n_i v\exp(\frac{-\phi}{V_T}) + p_T\right) + \tau_p^0\left(n_i u\exp(\frac{\phi}{V_T}) + n_T\right)}. \tag{29}$$

The system (19)-(21) is subject to some appropriate conditions on the boundary of a bounded domain denoted by $\Omega \subset \Re^2$. The boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ is piecewise smooth consisting of Dirichlet $\partial\Omega_D$ and Neumann $\partial\Omega_N$ parts. The Dirichlet part corresponds to the ohmic contacts on the device. By assuming the charge neutrality condition and the mass-action law [37], the Dirichlet boundary conditions of the model in terms of the variables ϕ, u, and v are described as follows:

$$\phi = V_O + V_b, \tag{30}$$

$$u = \exp(\frac{-V_O}{V_T}), \tag{31}$$

$$v = \exp(\frac{V_O}{V_T}) \ , \tag{32}$$

$$T_n = T_p = 300 \ , \tag{33}$$

where V_O denotes the applied voltage and V_b represents the built-in potential [37]. The Neumann part corresponds to the artificial boundary conditions $\partial\phi/\partial n = \partial u/\partial n = \partial v/\partial n = \partial T_n/\partial n = \partial T_p/\partial n = 0$ for the state variables.

Several remarks on the model are in order.

Remark 2.1. The carrier current densities (7) and (8) are derived from those of the standard hydrodynamic model, namely, from

$$\mathbf{J}_n - \frac{\tau_{pn}}{q}(\nabla \cdot \mathbf{J}_n)\frac{\mathbf{J}_n}{n} = q\mu_n[\frac{k_B T_n}{q}\nabla n + n\nabla(\frac{k_B T_n}{q} - \phi)]$$

(and a similar equation for the hole) by neglecting the second (convective) term of this equation and by assuming the electron to be in thermal equilibrium with the lattice, i.e., $T_n = T_p = T_L$ [12,19,23,31]. The same assumption is also used for the heat conductivities from which (16) and (17) are thus implied by the Wiedemann-Franz law [18].

Remark 2.2. In our numerical simulations, the kinetic energy in the carrier average energies (12) and (13) is neglected. Nevertheless, it does not affect the self-adjoint transformation as shown in the next section.

Remark 2.3. Following the Caughey-Thomas expression [8,33], we use the field-dependent mobility model

$$\mu_{n,p} = \frac{\mu^0_{n,p}}{(1 + (\frac{\mu^0_{n,p}|\mathsf{E}|}{v^{sat}_{n,p}})^{\beta_{n,p}})^{1/\beta_{n,p}}},$$

where $\mu^0_n = 1500cm^2V^{-1}s^{-1}$ and $\mu^0_p = 600cm^2V^{-1}s^{-1}$[37,33] are lattice mobility constants, $v^{sat}_n = 1.1 \times 10^7$ and $v^{sat}_p = 9.5 \times 10^6$ [8,33] are saturated velocities, and $\beta_n = 2$ and $\beta_p = 1$. Typically, the electron mobility varies between $50cm^2V^{-1}s^{-1}$ and $1500cm^2V^{-1}s^{-1}$ while the hole mobility varies between $50cm^2V^{-1}s^{-1}$ and $600cm^2V^{-1}s^{-1}$ for silicon at room temperature. Moreover, the linear proportionality of the electric field strength and the magnitudes of the drift velocities only holds at relatively low electric fields. The drift velocities saturate at high electric fields due to carrier heating. This effect must be accounted for by field-dependent mobilities if high field effects are to be analyzed.

3. A SELF-ADJOINT FORMULATION FOR CARRIER TEMPERATURES

In order to have a self-adjoint expression of (4)-(5), we rewrite (9)-(10) more precisely as

$$\mathbf{S}_n = \frac{5\mathbf{J}_n}{-2q}k_B T_n - \kappa_n \nabla T_n + \frac{\mathbf{J}_n}{-q}(\frac{1}{2}m^*_n|\mathbf{v}_n|^2), \tag{34}$$

$$\mathbf{S}_p = \frac{5\mathbf{J}_p}{+2q} k_B T_p - \kappa_p \nabla T_p + \frac{\mathbf{J}_p}{+q}(\frac{1}{2} m_p^* |\mathbf{v}_p|^2). \tag{35}$$

Introduce new variables g_n and g_p for carrier temperatures $T_n = g_n \exp(c\varphi_n)$ and $T_p = g_p \exp(-c\varphi_p)$ where c is a constant to be determined. With these two variables, Eqs. (34)-(35) are reformulated as

$$\begin{aligned}\mathbf{S}_n = &\frac{5\mathbf{J}_n}{-2q} k_B g_n \exp(c\varphi_n) - \kappa_n \left[\exp(c\varphi_n)\nabla g_n + c g_n \exp(c\varphi_n)\nabla\varphi_n\right] \\ &+ \frac{\mathbf{J}_n}{-q}(\frac{1}{2} m_n^* |\mathbf{v}_n|^2), \end{aligned} \tag{36}$$

$$\begin{aligned}\mathbf{S}_p = &\frac{5\mathbf{J}_p}{+2q} k_B g_p \exp(-c\varphi_p) - \kappa_p \left[\exp(-c\varphi_p)\nabla g_p - c g_p \exp(-c\varphi_p)\nabla\varphi_p\right] \\ &+ \frac{\mathbf{J}_p}{+q}(\frac{1}{2} m_p^* |\mathbf{v}_p|^2). \end{aligned} \tag{37}$$

An instrumental choice of the unknown constant c is to eliminate the gradient of the quasi-Fermi potentials leading to a divergence formulation for the new state variables. We thus solve the following two equations for c:

$$\frac{5\mathbf{J}_n}{-2q} k_B g_n \exp(c\varphi_n) - \kappa_n c g_n \exp(c\varphi_n)\nabla\varphi_n = 0, \tag{38}$$

$$\frac{5\mathbf{J}_p}{+2q} k_B g_p \exp(-c\varphi_p) + \kappa_p c g_p \exp(-c\varphi_p)\nabla\varphi = 0. \tag{39}$$

Substituting (26)-(27) into (38)-(39) we have

$$\frac{5}{2} n\mu_n k_B = c\kappa_n = 2cT_L(\frac{k_B}{q})^2 nq\mu_n, \tag{40}$$

$$\frac{5}{2} p\mu_p k_B = c\kappa_p = 2cT_L(\frac{k_B}{q})^2 pq\mu_p, \tag{41}$$

which yield $c = \frac{5q}{4k_B T_L} = \frac{5}{4V_T}$. The new state variables g_n and g_p can therefore be defined as

$$T_n = g_n \exp(\frac{5\varphi_n}{4V_T}), \tag{42}$$

$$T_p = g_p \exp(-\frac{5\varphi_p}{4V_T}). \tag{43}$$

Note that these two expressions are purely mathematical and very similar to that of the Slotboom variables (22) and (23). It is unknown to us whether these formulas have been developed in the literature or have any significant

and interesting physical relationships between the quasi-Fermi potentials and carrier temperatures. Nevertheless, as shown below, they provide an intriguing alternative to the energy transport model in terms of mathematical properties, especially from the computational viewpoint. The energy fluxes (9)-(10) are then reduced to

$$\mathbf{S}_n = -\kappa_n \exp(\frac{5\varphi_n}{4V_T})\nabla g_n + \frac{\mathbf{J}_n}{-q}(\frac{1}{2}m_n^* |\mathbf{v}_n|^2), \tag{44}$$

$$\mathbf{S}_p = -\kappa_p \exp(-\frac{5\varphi_p}{4V_T})\nabla g_p + \frac{\mathbf{J}_p}{+q}(\frac{1}{2}m_p^* |\mathbf{v}_p|^2). \tag{45}$$

Substituting $\mathbf{v}_n$ and $\mathbf{v}_p$ for $\mathbf{J}_n$ and $\mathbf{J}_p$ in (22)-(23), we obtain the following self-adjoint system in terms of the new variables

$$\nabla \cdot (\kappa_n \exp(\frac{5\varphi_n}{4V_T})\nabla g_n) = R_n(g_n), \tag{46}$$

$$\nabla \cdot (\kappa_p \exp(-\frac{5\varphi_p}{4V_T})\nabla g_p) = R_p(g_p), \tag{47}$$

where

$$R_n(g_n) = n(\frac{\omega_n - \omega_0}{\tau_{n\omega}}) - \mathbf{J}_n \cdot \mathbf{E} - \frac{1}{q}\nabla \cdot (\frac{1}{2}m_n^* \frac{|\mathbf{J}_n|^2}{q^2 n^2}\mathbf{J}_n), \tag{48}$$

$$R_p(g_p) = p(\frac{\omega_p - \omega_0}{\tau_{p\omega}}) - \mathbf{J}_p \cdot \mathbf{E} + \frac{1}{q}\nabla \cdot (\frac{1}{2}m_p^* \frac{|\mathbf{J}_p|^2}{q^2 n^2}\mathbf{J}_p). \tag{49}$$

The boundary conditions are changed accordingly to

$$g_n = \frac{300}{\exp(\frac{5V_O}{4V_T})} \text{ and } g_p = \frac{300}{\exp(-\frac{5V_O}{4V_T})} \text{ on } \partial\Omega_D, \tag{50}$$

$$\frac{\partial g_n}{\partial n} = \frac{\partial g_p}{\partial n} = 0 \text{ on } \partial\Omega_N. \tag{51}$$

Remark 3.1. It should be noted that effective approximation of the gradient of current densities in formulas (48)-(49) is in general very difficult to acquire. Simplified models for these formulas based on physical consideration are possible. For example, by assuming that the drift energy is only a small part of the total kinetic energy [9], (48)-(49) can be reduced to

$$R_n(g_n) = n(\frac{\omega_n - \omega_0}{\tau_{n\omega}}) - \mathbf{J}_n \cdot \mathbf{E},$$

$$R_p(g_p) = p(\frac{\omega_p - \omega_0}{\tau_{p\omega}}) - \mathbf{J}_p \cdot \mathbf{E},$$

which will be used in our numerical simulations.

4. FINITE ELEMENT APPROXIMATION

Let $\mathcal{T}$ be a finite element partition of the domain Ω of Laplace's equation $\Delta u = 0$ with Dirichlet and Neumann boundary conditions such that $\mathcal{T} = \{ \tau_j; j = 1, ..., M, \bar{\Omega} = \cup_{j=1}^{M} \bar{\tau_j}\}$ and $S_h(\mathcal{T})$ denote a finite element subspace on $\mathcal{T}$ for the Laplace problem. The FE approximation of the problem in $S_h(\mathcal{T})$ is then to find $u_h \in S_h(\mathcal{T})$ such that

$$B_h(u_h, v_h) = 0 \, \forall v_h \in S_h(\mathcal{T})$$

with

$$B_h(u_h, v_h) \equiv \sum_{\tau \in \mathcal{T}} \int_{\tau} \nabla u_h \cdot \nabla v_h dxdy.$$

We consider particularly that the partition $\mathcal{T}$ is generated by the 1-irregular mesh refinement scheme [11,22]. Let N^a be a set of $\mathcal{N}$ indices that are assigned to active degrees of freedom (i.e., regular nodes) and N^c assigned to constrained degrees of freedom (irregular nodes). By an *active* degree of freedom, we mean one that defines a parameter associated with the global stiffness matrix whereas a *constrained* degree of freedom is a linear combination of active degrees of freedom that are associated with the constrained node by element connectivity. For each $i \in N^c$, there exists a set $A(i) \subset N^a$ of corresponding active degrees of freedom such that the resulting finite element space $S_h(\mathcal{T})$ consists of continuous functions. If rectangular elements are used, then u_h is of the following form [11,22]:

$$\begin{aligned} u_h &= \sum_{i \in N^a} u_i \hat{b}_i + \sum_{j \in N^c} u_j \hat{b}_j, \\ &= \sum_{i \in N^a} u_i \hat{b}_i + \sum_{j \in N^c} \sum_{k \in A(j)} \frac{1}{2} u_k \hat{b}_j, \end{aligned}$$

where u_i are scalars and $\hat{b}_i$ are *unconstrained bilinear bases* which can be constructed via the following four shape functions

$$\begin{aligned} s_1 &= (1-\xi)(1-\eta)/4, \\ s_2 &= (1+\xi)(1-\eta)/4, \\ s_3 &= (1+\xi)(1+\eta)/4, \\ s_4 &= (1-\xi)(1+\eta)/4, \end{aligned}$$

defined on the reference element $\hat{\tau} = \{(\xi, \eta) : |\xi| \leq 1, |\eta| \leq 1\}$. For every $i \in N^a$, let $C(i) = \{j \in N^c \mid i \in A(j)\}$. We rewrite u_h in the form

$$
\begin{aligned}
u_h &= \sum_{i\in N^a} u_i \hat{b}_i + \sum_{k\in N^a} \sum_{j\in C(k)} \frac{1}{2} u_k \hat{b}_j, \\
&= \sum_{i\in N^a} u_i (\hat{b}_i + \sum_{j\in C(i)} \frac{1}{2} \hat{b}_j).
\end{aligned}
$$

Thus, the functions

$$
b_i = \hat{b}_i + \sum_{j\in C(i)} \frac{1}{2} \hat{b}_j \ \forall i \in N^a,
$$

form *constrained bilinear bases.*

Let $(x_i, y_i) \in \overline{\Omega}$ be a mesh point in $\mathcal{T}$. For each $i \in N^a$, there exists a set $V(i) \subset N^a$ of active degrees of freedom such that $B_h(b_k, b_i) \neq 0, \forall k \in V(i), k \neq i$. Using the standard notation $u_i \approx u(x_i, y_i)$ as unknown scalars, the approximation results in a system of linear equations

$$
\xi_i u_i - \sum_{k\in V(i)} \xi_k u_k = f_i^*, \tag{52}
$$

where $\xi_i = B_h(b_i, b_i)$, $\xi_k = -B_h(b_k, b_i)$, and $f_i^* = 0$ if $(x_i, y_i) \in \Omega$ and is associated with the boundary condition if $(x_i, y_i) \in \partial\Omega$.

Theorem 4.1. The matrix induced by (52) is diagonally dominant, i.e.,

$$
\begin{aligned}
&\xi_i \geq \sum_{k\in V(i)} \xi_k, \\
&\xi_k \geq 0 \ \ \forall k \in V(i).
\end{aligned} \tag{53}
$$

Furthermore, the strict inequality in (53) holds for at least one $i \in N^a$.

Proof. Without loss of generality, we construct the proof in the region $\{(x, y) : 0 \leq x \leq 2, 0 \leq y \leq 2\}$.

Type 0: Without subdivided elements. This is a 9-point stencil grid. For simplicity, we enumerate the unknown scalars $u_1, u_2, \ldots, u_9$ from the top left to the bottom right. Assume that all nodes are unconstrained. For $i = 5$, we have $V(i) = \{1, 2, 3, 4, 6, 7, 8, 9\}$ and $\xi_5 = B(b_5, b_5) = \sum_{i=1}^{4} \int_{-1}^{1} \int_{-1}^{1} \nabla s_i \cdot \nabla s_i d\xi d\eta = 8/3$. Similarly, we have $\xi_k = 1/3$ for all $k \in V(5)$. Hence (52) reads as

$$
\begin{aligned}
(8/3)u_5 &-(1/3)u_1 -(1/3)u_2 -(1/3)u_3 -(1/3)u_4 \\
&-(1/3)u_6 -(1/3)u_7 -(1/3)u_8 -(1/3)u_9 = 0,
\end{aligned}
$$

which obviously satisfies (53). Suppose that exactly one node is constrained, say node 8. Then $V(5) = \{1, 2, 3, 4, 6, 7, 9\}$, $A(8) = \{7, 9\}$, and

$$
\begin{aligned}
-\xi_7 &= B_h(b_5, b_7) \\
&= B_h(\hat{b}_5, \hat{b}_7 + \frac{1}{2}\hat{b}_8) \\
&= \int_{-1}^{1}\int_{-1}^{1} \nabla s_1 \cdot \nabla s_3 d\xi d\eta \\
&\quad + \frac{1}{2}\left[\int_{-1}^{1}\int_{-1}^{1} \nabla s_2 \cdot \nabla s_3 d\xi d\eta + \int_{-1}^{1}\int_{-1}^{1} \nabla s_1 \cdot \nabla s_4 d\xi d\eta\right] \\
&= (-1/6) + (1/2)(-1/6 - 1/6) = -1/2.
\end{aligned}
$$

Other $\xi_k' s$ are computed in a similarly way. Hence (52) reads as

$$
\begin{aligned}
(8/3)u_5 &-(1/3)u_1 -(1/3)u_2 -(1/3)u_3 -(1/3)u_4 \\
&-(1/3)u_6 -(1/2)u_7 \qquad\qquad -(1/2)u_9 = 0,
\end{aligned}
$$

which again satisfies (53). If $V(5) = \{1, 2, 3, 6, 7, 9\}$, i.e., nodes 8 and 4 are constrained, then we have

$$
\begin{aligned}
(8/3)u_5 &-(1/2)u_1 -(1/3)u_2 -(1/3)u_3 \\
&-(1/3)u_6 -(2/3)u_7 \qquad\qquad -(1/2)u_9 = 0.
\end{aligned}
$$

All other cases with different constrained nodes can be computed similarly and (53) holds for this type of finite elements.

Type 1: With one subdivided element. For a typical case of this type of partition, we refer to Fig. 1. Other cases can be shown in a similar way. All nodes are enumerated as that shown in the figure. Note that the coordinates of node 10 is (1.5,0.5), 11 is (1.5,1), and 12 is (1,0.5) etc. Nodes 11 and 12 are constrained. Consider the case $V(5) = \{1, 2, 3, 4, 6, 7, 8, 10, 13, 14\}$, $A(11) = \{5, 6\}$, $A(12) = \{5, 8\}, C(5) = \{11, 12\}$. We have

$$
\begin{aligned}
b_5 &= \hat{b}_5 + \frac{1}{2}\hat{b}_{11} + \frac{1}{2}\hat{b}_{12} \\
b_8 &= \hat{b}_8 + \frac{1}{2}\hat{b}_{12} \\
b_6 &= \hat{b}_6 + \frac{1}{2}\hat{b}_{11}
\end{aligned}
$$

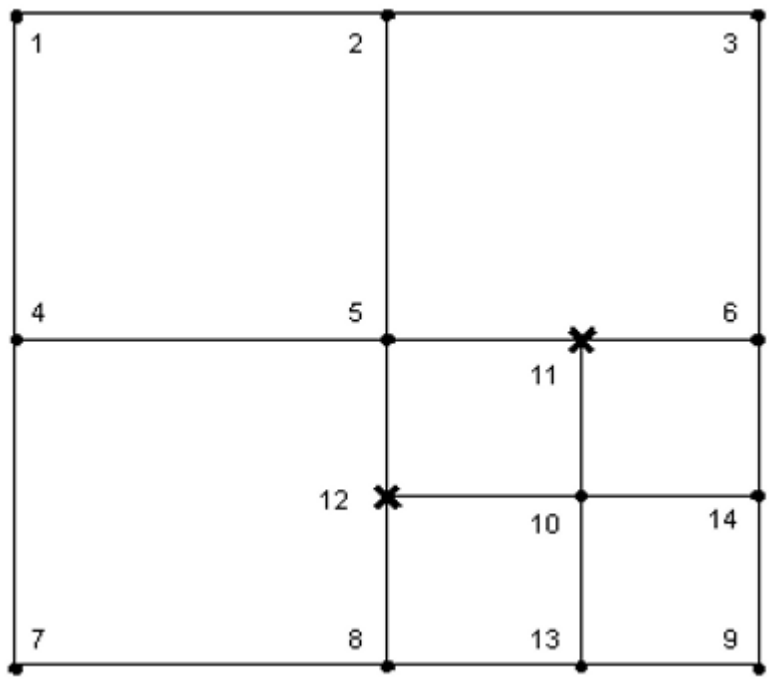

Fig. 1. A 1-irregular mesh with one subdivided element.

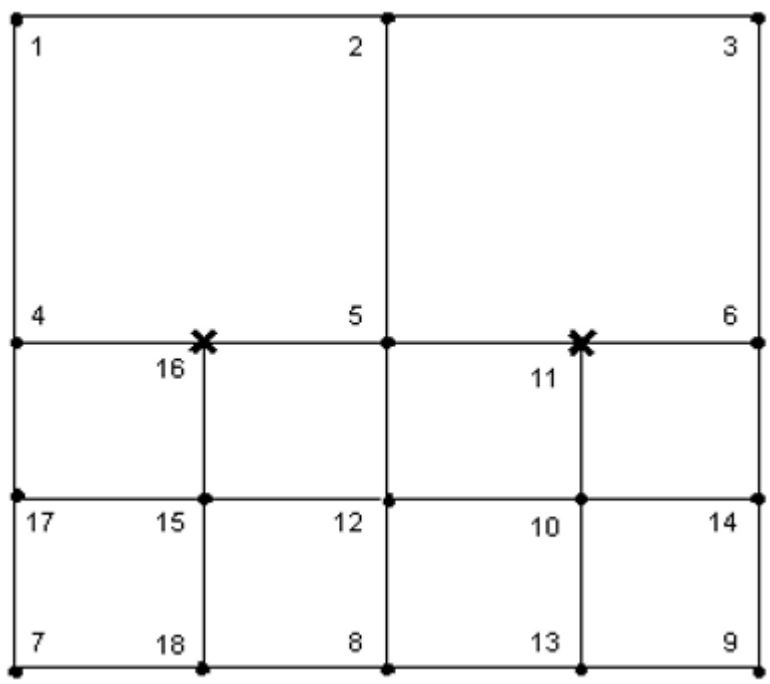

Fig. 2. A 1-irregular mesh with two subdivided elements.

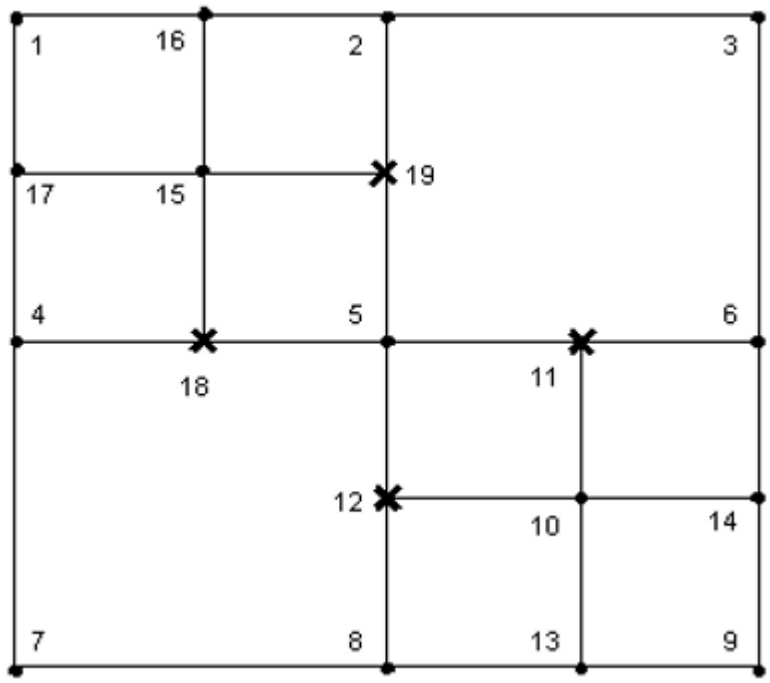

Fig. 3. A 1-irregular mesh with two subdivided elements.

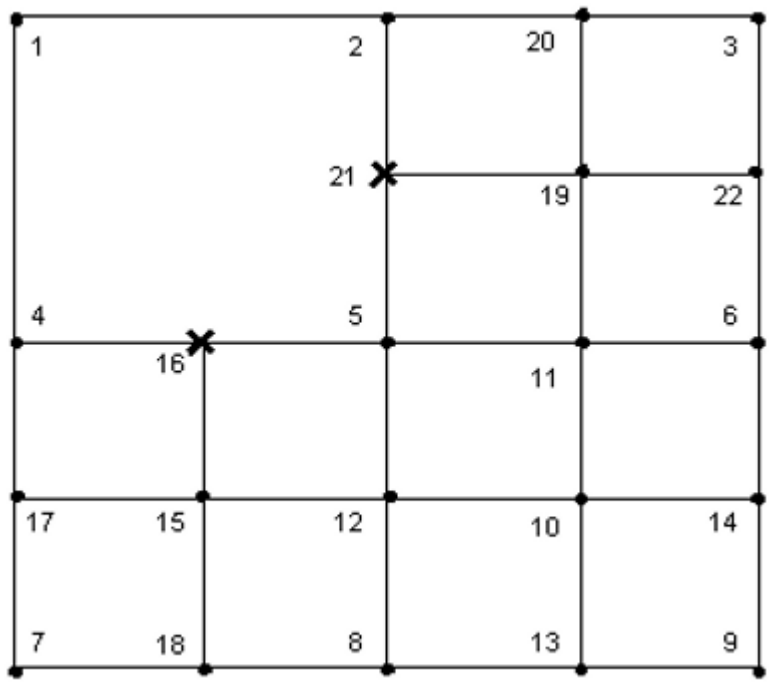

Fig. 4. A 1-irregular mesh with three subdivided elements.

where

$$\hat{b}_5 = \begin{cases} (2-x)(2-y) & \text{if } 1 \le x \le 2, 1 \le y \le 2 \\ x(2-y) & \text{if } 0 \le x \le 1, 1 \le y \le 2 \\ xy & \text{if } 0 \le x \le 1, 0 \le y \le 1 \\ 4(1.5-x)(y-0.5) & \text{if } 1 \le x \le 1.5, 0.5 \le y \le 1 \end{cases}$$

$$\hat{b}_{11} = \begin{cases} 4(x-1)(y-0.5) & \text{if } 1 \le x \le 1.5, 0.5 \le y \le 1 \\ 4(2-x)(y-0.5) & \text{if } 1.5 \le x \le 2, 0.5 \le y \le 1 \end{cases}$$

$$\hat{b}_{12} = \begin{cases} 4(1.5-x)(1-y) & \text{if } 1 \le x \le 1.5, 0.5 \le y \le 1 \\ 4(1.5-x)y & \text{if } 1 \le x \le 1.5, 0 \le y \le 0.5 \end{cases}$$

Other basis functions can be constructed analogously. Hence, for example,

$$\begin{aligned}
\xi_5 &= \int_0^2 \int_0^2 \nabla b_5 \cdot \nabla b_5 d\xi d\eta \\
&= \int_0^2 \int_0^2 \nabla \hat{b}_5 \cdot \nabla \hat{b}_5 d\xi d\eta + \frac{1}{2} \sum_{j \in C(5)} \int_0^2 \int_0^2 \nabla \hat{b}_5 \cdot \nabla \hat{b}_j d\xi d\eta \\
&+ \frac{1}{2} \sum_{j \in C(5)} \int_0^2 \int_0^2 \nabla \hat{b}_j \cdot \nabla \hat{b}_5 d\xi d\eta + \frac{1}{4} \sum_{j \in C(5)} \int_0^2 \int_0^2 \nabla \hat{b}_j \cdot \nabla \hat{b}_j d\xi d\eta \\
&+ \frac{1}{4} \int_0^2 \int_0^2 \nabla \hat{b}_{11} \cdot \nabla \hat{b}_{12} d\xi d\eta + \frac{1}{4} \int_0^2 \int_0^2 \nabla \hat{b}_{12} \cdot \nabla \hat{b}_{11} d\xi d\eta \\
&= (8/3) + (1/2)(-1/6 - 1/6) + (1/2)(-1/6 - 1/6) \\
&+ (1/4)(4/3 + 4/3) + (1/4)(-1/3) + (1/4)(-1/3) \\
&= 17/6
\end{aligned}$$

and we have

$$\begin{aligned}
(17/6)u_5 &-(1/3)u_1 \ -(1/3)u_2 \ -(1/3)u_3 \ -(1/3)u_4 \ -(1/12)u_6 \\
&-(1/3)u_7 \ -(1/12)u_8 \ -(4/6)u_{10} \ -(1/6)u_{13} \ -(1/6)u_{14} = 0,
\end{aligned}$$

for (52) which also satisfies (53). It can be shown similarly that (53) holds for various cases of $A(k)$ and $V(i)$ for this type of elements.

Type 2: With two subdivided elements. Again, for simplicity, we refer to two typical cases as shown in Figs. 2 and 3. The proof of (53) proceeds analogously as that of Type 1. For instance, for the case of $V(5) = \{1, 2, 3, 10, 12, 14, 15, 17\}$, $A(11) = \{5, 6\}$, $A(16) = \{5, 4\}, C(5) = \{11, 16\}$ in Fig. 2, we have

$$\begin{aligned} 3u_5 &-(1/3)u_1 -(1/3)u_2 -(1/3)u_3 -(1/2)u_{10} -(2/3)u_{12} \\ &-(1/6)u_{14} -(1/2)u_{15} -(1/6)u_{17} = 0. \end{aligned}$$

And, for $V(5) = \{2, 3, 4, 6, 7, 8, 10, 13, 14, 15, 16, 17\}$, $A(18) = \{4, 5\}$, $A(11) = \{5, 6\}$, $A(12) = \{5, 8\}$, $A(19) = \{2, 5\}$ and $C(5) = \{11, 12, 18, 19\}$ in Fig. 3, we have

$$\begin{aligned} 3u_5 &-(1/12)u_2 -(1/3)u_3 -(1/12)u_4 -(1/12)u_6 -(1/3)u_7 -(1/12)u_8 \\ &-(2/3)u_{10} -(1/6)u_{13} -(1/6)u_{14} -(2/3)u_{15} -(1/6)u_{16} -(1/6)u_{17} = 0. \end{aligned}$$

Type 3: With three subdivided elements. Finally, we prove for a typical case of this type of elements by referring to Fig. 4, i.e., $V(5) = \{1, 10, 11, 12, 15, 17, 19, 20\}$, $A(16) = \{4, 5\}$, $A(21) = \{2, 5\}$, $C(5) = \{16, 21\}$, which gives

$$\begin{aligned} 3u_5 &-(1/3)u_1 -(1/3)u_{10} -(1/2)u_{11} -(1/2)u_{12} -(1/2)u_{15} \\ &-(1/6)u_{17} -(1/2)u_{19} -(1/6)u_{20} = 0. \end{aligned}$$

The strict inequality in (53) holds for at least one $i \in N^a$ since the boundary condition is not of pure Neumann type. This completes the proof.

We are now concerned with the matrix properties of the FE approximation of the model system, namely, Eqs. (19), (20), (21), (46) and (47) associated with their corresponding boundary conditions. In device simulations, the coupled system of nonlinear PDEs is usually solved one by one in a manner of Gummel's decoupling. For each Gummel's iteration and after the discretization, each one of these equations will result in a system of nonlinear algebraic equations for which we only have to examine the property of the stiffness matrix. Obviously, the matrix corresponding to the problem (19) is exactly that of (52) and hence Theorem 4.1 applies to this problem. However, for the rest of equations, it is well known that the Scharfetter-Gummel discretization induces non-physical diffusion in the direction normal to drift velocity for multidimensional problems, which has led to various modifications of the method [4,26,30,34,39,40]. In order to obtain the same matrix property, we extend in particular the method proposed in [39] to the 1-irregular mesh refinement

scheme. It suffices to consider only the electron energy transport equation (46). Analysis for other equations is completely analogous to what follows.

The FE approximation of the electron energy transport equation is to find $g_n^h \in S_h(\mathcal{T})$ such that

$$\sum_{\tau\in\mathcal{T}} \int_\tau \mathbf{S}_h \cdot \nabla v_h dxdy = \sum_{\tau\in\mathcal{T}} \int_\tau R_n(g_n^h) v_h dxdy \quad \forall v_h \in S_h(\mathcal{T}), \tag{54}$$

where

$$\mathbf{S}_h \equiv S(\varphi_n)\nabla g_n^h, \tag{55}$$

$$S(\varphi_n) \equiv \kappa_n \exp(\frac{5\varphi_n}{4V_T}). \tag{56}$$

For each $i \in N^a$ with $g_i = g_n^h(x_i, y_i)$, let l_{ik} denote the line segment from node i to node k, $k \in V(i)$. On l_{ik}, we use the Scharfetter-Gummel scheme to approximate $S(\varphi_n)$ by

$$\begin{aligned}\widetilde{S}_{ik}(\varphi_n) &= \kappa_n|_{(k,i)} \left[\frac{1}{|l_{ik}|}\int_{l_{ik}} \frac{1}{\exp(\frac{5\varphi_n}{4V_T})} ds\right]^{-1} \\ &\approx \kappa_n|_{(k,i)} B(\frac{5\varphi_i - 5\varphi_k}{4V_T}) \exp(\frac{5\varphi_i}{4V_T}),\end{aligned} \tag{57}$$

where $\kappa_n|_{(k,i)} = (\kappa_n(x_k, y_k) + \kappa_n(x_i, y_i))/2$ and $B(t) = t/(e^t - 1)$ is the Bernoulli function for any real number t. For each element $\tau \in \mathcal{T}$ and taking $v_h = b_i$, we have

$$\begin{aligned}\int_\tau \mathbf{S}_h \cdot \nabla b_i dxdy &= \sum_{k\in V(i)\cup\{i\}} \int_\tau S(\varphi_n) g_k \nabla b_k \cdot \nabla b_i dxdy \\ &\approx \sum_{k\in V(i)} \widetilde{S}_{ik}(\varphi_n) \int_\tau (g_k \nabla b_k \cdot \nabla b_i + g_i \nabla b_i \cdot \nabla b_i)\, dxdy \\ &= \sum_{k\in V(i)} \widetilde{S}_{ik}(\varphi_n) \int_\tau (g_k \nabla b_k \cdot \nabla b_i - g_i \nabla b_k \cdot \nabla b_i)\, dxdy \\ &= \sum_{k\in V(i)} \xi_k \widetilde{S}_{ik}(\varphi_n)(g_k - g_i).\end{aligned} \tag{58}$$

Summing over all elements, the discrete form of (46) can thus be written as

$$\eta_i g_i - \sum_{k\in V(i)} \eta_k g_k = -R_n(g_i) + R_i^*, \tag{59}$$

for the i-th equation or in a more compact matrix form

$$\mathcal{A}U = -R_n(U) + R_n^*(U) =: -F(U), \tag{60}$$

where

$$\begin{aligned}
\eta_i &= \sum_{k\in V(i)} \eta_k, && (61)\\
\eta_k &= \xi_k d_k, && (62)\\
d_k &= \kappa_n|_{(k,i)} B(\frac{5\varphi_i - 5\varphi_k}{4V_T}) \exp(\frac{5\varphi_i}{4V_T}), && (63)\\
\xi_k &= -B_h(b_i, b_k), && (64)\\
U &= (g_1, \ldots, g_N), && (65)\\
R_n(U) &= (R_n(g_1), \ldots, R_n(g_N)), && (66)\\
R_n^*(U) &= (R_n^*(g_1), \ldots, R_n^*(g_N)). && (67)
\end{aligned}$$

Here $R_n(g_i)$ is defined by evaluating (54) at $(x_i, y_i) \in \Omega$, and R_i^* is associated with the boundary conditions if $(x_i, y_i) \in \partial\Omega$ and $R_i^* = 0$ if $(x_i, y_i) \in \Omega$. Note that, by the definition of the Bernoulli function and of the heat conductivity coefficient, the factors d_k in (63) are positive. We therefore conclude the following result.

Theorem 4.2. The matrix $\mathcal{A}$ in (60) is diagonally dominant, i.e.,

$$\begin{aligned}
\eta_i &\geq \sum_{k\in V(i)} \eta_k, && (68)\\
\eta_k &\geq 0 \ \ \forall k \in V(i).
\end{aligned}$$

Furthermore, the strict inequality in (68) holds for at least one $i \in N^a$.

The diagonal dominance of the resulting matrices of the model problems provides not only stability of numerical solutions (i.e., no non-physical oscillations) but also convergence of iterative procedures when the special properties of the nonlinearity in these problems are taken into account. Moreover, the existence and uniqueness of the solutions can also be guaranteed by means of the construction of lower and upper solutions which are defined as follows:

Definition. A vector $\tilde{U} \equiv (\widetilde{g}_1, \ldots, \widetilde{g}_N) \in \Re^N$ is called an upper solution of (60) if it satisfies the following inequality

$$\eta_i \widetilde{g}_i - \sum_{k\in V(i)} \eta_k \widetilde{g}_k \geq -F_i(\widetilde{g}_i) \tag{69}$$

and $\hat{U} \equiv (\widehat{g}_1, \ldots, \widehat{g}_N) \in \Re^N$ is a lower solution if

$$\eta_i \widehat{g}_i - \sum_{k \in V(i)} \eta_k \widehat{g}_k \leq -F_i(\widehat{g}_i), \tag{70}$$

for all $i \in N^a$.

It is obvious that every solution of (60) is an upper solution as well as a lower solution. We say that $\hat{U}$ and $\tilde{U}$ are ordered if $\hat{U} \leq \tilde{U}$. Given any ordered lower and upper solutions $\hat{U}$ and $\tilde{U}$, we define

$$\langle \hat{U}, \tilde{U} \rangle \equiv \{U \in \Re^N;\ \hat{U} \leq U \leq \tilde{U}\} \tag{71}$$
$$\langle \widehat{g}_i, \widetilde{g}_i \rangle \equiv \{g_i \in \Re;\ \widehat{g}_i \leq g_i \leq \widetilde{g}_i\}. \tag{72}$$

By (7), (12), (42), and (48), we can always choose the nonnegative scalars γ_i so that

$$\gamma_i \equiv \max \left\{ \frac{\partial R_n(g_i)}{\partial g_n};\ g_i \in \langle \widehat{g}_i, \widetilde{g}_i \rangle \right\} \tag{73}$$

or in matrix form

$$\Lambda \equiv diag(\gamma_i) \tag{74}$$

for all $i \in N^a$. Then by adding the term $\gamma_i g_i$ on both sides of (60) we obtain the equivalent system

$$\eta_i g_i + \gamma_i g_i = \sum_{k \in V(i)} \eta_k g_k + \gamma_i g_i - F_i(g_i). \tag{75}$$

Let $\overline{U}^{(0)} = \tilde{U}$ be an initial iterate. We construct a sequence $\{\overline{U}^{(m)}\}$ by solving the linear system

$$\eta_i \overline{g}_i^{(m+1)} + \gamma_i \overline{g}_i^{(m+1)} = \sum_{k \in V(i)} \eta_k \overline{g}_k^{(m)} + \gamma_i \overline{g}_i^{(m)} - F_i(\overline{g}_i^{(m)}), \tag{76}$$

for $m = 0, 1, 2, \ldots$ and $i \in N^a$. Similarly, by using $\underline{U}^{(0)} = \hat{U}$ as another initial iterate, we obtain a sequence $\{\underline{U}^{(m)}\}$ from the linear system

$$\eta_i \underline{g}_i^{(m+1)} + \gamma_i \underline{g}_i^{(m+1)} = \sum_{k \in V(i)} \eta_k \underline{g}_k^{(m)} + \gamma_i \underline{g}_i^{(m)} - F_i(\underline{g}_i^{(m)}), \tag{77}$$

for $m = 0, 1, 2, \ldots$ and $i \in N^a$. We refer to $\{\overline{U}^{(m)}\}$ and $\{\underline{U}^{(m)}\}$ as the maximal and minimal sequences. The following results are direct consequences of Theorem 4.2 for which a proof can be found, for instance, in [27]. These results insure that the iterative method (76) or (77) is monotonically, globally, and uniquely convergent.

Theorem 4.3. Let $\tilde{U}$, $\hat{U}$ be a pair of ordered upper and lower solutions of (60). Then the sequences $\{\overline{U}^{(m)}\}$ and $\{\underline{U}^{(m)}\}$ generated by solving (76) and (77) with $\overline{U}^{(0)} = \tilde{U}$ and $\underline{U}^{(0)} = \hat{U}$ converge monotonically to the solutions $\overline{U}$ and $\underline{U}$ of (76), respectively. That is

$$\hat{U} \leq \underline{U}^{(m)} \leq \underline{U}^{(m+1)} \leq \underline{U} \leq \overline{U} \leq \overline{U}^{(m+1)} \leq \overline{U}^{(m)} \leq \tilde{U}, \; m = 1, 2, \ldots. \quad (78)$$

Moreover, $\overline{U} = \underline{U}$ is the unique solution of (60).

Remark 4.1. A great deal of numerical methods have been developed for energy transport models, see e.g. [1,5,10,12,13,15,17,18,25,40]. Mixed FEM is one of the most frequently used methods to approximate the semiconductor equations [5,6,10,13,25], whereas our approximation is based on the standard FEM since only one linear FE space is used for all state variables. As shown in [6], the major difference between the standard FEM and mixed FEM is that the use of harmonic average in mixed FEM for the exponential function can extend the effective features of the Scharfetter-Gummel scheme to the two-dimensional problems whereas in the standard method one can only have one-dimensional harmonic average as that of (57). Nonetheless, our numerical experience suggests that sufficiently fine (adaptive) mesh can alleviate this drawback. Another frequently used method is the finite volume method (FVM) or the box method [4,12,14]. In view of the harmonic average (a simplification for the integral in (58)), the present FEM is essentially equivalent to FVM as proved in [4] and as numerically demonstrated in [20].

Remark 4.2. It is well-known that the Slotboom variables (22), (23), (42), and (43) can lead to catastrophic roundoff errors if these variables are not properly scaled during computations [4]. To improve matrix conditioning, we can divide (59) by $\kappa_n|_{(k,i)} \exp(\frac{5\varphi_i}{4V_T})$ in a node-by-node basis. The corresponding stiffness matrix is an M-matrix and Theorem 4.2 still holds with this scaling. Another way to improve the conditioning is to perform (at the discrete level) a change of the variable g_n back to its original variable (42) as suggested in [6]. The resulting stiffness matrix may not be an M-matrix. However, the mixed method can be used to recover the M-matrix property and furthermore to have the current conservation property [6]. However, the implementation of this method is more complicated than that of (59) since the discrete system is enlarged by this method and the matrix reduction by means of static condensation requires an element-wise inversion of the block-diagonal matrix

associated with the auxiliary variable. Moreover, a suitable numerical integration formula for the local and global matrices and for the right-hand side vector is required (see [6] for more details). The monotone parameters (73) will also be more involved with this method.

5. AN ADAPTIVE ALGORITHM

To illustrate the implementation of the iterative method (77), we briefly summarize our adaptive monotone-Gummel algorithm as follows. Here we use the notation l as Gummel's (outer) iteration index and m as the monotone (inner) iteration index.

Step 1. Initial Mesh: Create a coarse and structured mesh for which the number of nodes can be chosen as small as possible.

Step 2. Preprocessing: Since the initial mesh is usually very coarse, a sufficiently fine mesh created by solving the Poisson problem (19) with low biasing conditions proves to be an essential step for more effective refinement and faster convergence in the subsequent computations. With this step, a finer mesh can be generated to capture intrinsic irregularities caused by the junction layers of the doping profile and by the applied voltages. The problem is solved only for a few levels of refinement. Furthermore, the mesh near the interface between the oxide and the semiconductor will be refined several times by the refinement scheme as that in Step 6. We now have a nonuniform mesh with a better resolution in the vicinity of the interface and the junction.

Step 3. Gummel and Monotone Iterations on the Drift-Diffusion Model:

Step 3.1. Set $l = 0$ and do the following sub-steps:

Step 3.1.1. Set $m := 0$ and the initial guess $\phi_j^{(m)} = \widetilde{\phi}_j$ or $\widehat{\phi}_j$ $\forall\ j$ where $\widetilde{\phi}_j$ and $\widehat{\phi}_j$ are constant values for all $(x_j, y_j) \in \overline{\Omega}^h$ that can be easily verified to be an upper and lower solution of ϕ, respectively.

Step 3.1.2. Set $u^{(0)}$ and $v^{(0)}$ by the charge neutrality condition.

Step 3.1.3. Compute $\phi_j^{(m+1)}$ by solving the discrete potential system

$$\begin{cases} \xi_j \phi_j^{(m+1)} + \gamma_j(\phi)\, \phi_j^{(m+1)} = \sum_{k \in V(j)} \xi_k \phi_k^{(m)} \\ \quad -F(\phi_j^{(m)},\, u_j^{(l)},\, v_j^{(l)}) + \gamma_j(\phi)\, \phi_j^{(m)},\ \forall (x_j, y_j) \in \Omega^h, \\ \phi_j^{(m+1)} = V_O + V_b,\ \ \forall (x_j, y_j) \in \partial\Omega_D^h\ , \\ \frac{\partial \phi_j^{(m+1)}}{\partial n} = 0,\ \ \forall (x_j, y_j) \in \partial\Omega_N^h\ , \end{cases} \tag{79}$$

where

$$\gamma_j(\phi) = \max\left\{ \frac{\partial F(\phi_j)}{\partial \phi};\ \hat{\phi}_j \le \phi_j \le \tilde{\phi}_j \right\}, \tag{80}$$

$$\xi_k = -B_h(b_j, b_k).$$

Step 3.1.4. Set $\phi_j^{(m)} := \phi_j^{(m+1)}\ \forall\ j$ and $m := m+1$. Go to Step 3.1.3 until the stopping criteria of the inner iteration are satisfied.
Step 3.1.5. Set $\phi_j^{(l+1)} := \phi_j^{(m+1)}\ \forall\ j$.
Step 3.1.6. Set the initial guess $u_j^{(m)} = \widetilde{u}_j$ or $\widehat{u}_j \forall\ j$ where $\widetilde{u}_j$ and $\widehat{u}_j$ are constant values for all $(x_j, y_j) \in \overline{\Omega}^h$ that can be easily verified to be an upper and lower solution of u, respectively.
Step 3.1.7. Compute $u_j^{(m+1)}$ by solving the discrete electron system

$$\begin{cases} \zeta_j u_j^{(m+1)} + \gamma_j(u)\, u_j^{(m+1)} = \sum_{k\in V(j)}^{(m)} \zeta_k u_k^{(m)} \\ \quad -R(\phi_j^{(l+1)},\, u_j^{(m)},\, v_j^{(l)}) + \gamma_j(u)\, u_j^{(m)}, \quad \forall (x_j, y_j) \in \Omega^h, \\ u_j^{(m+1)} = \exp\left(\frac{-V_O}{V_T}\right), \quad \forall (x_j, y_j) \in \partial\Omega_D^h\ , \\ \frac{\partial u_j^{(m+1)}}{\partial n} = 0, \quad \forall (x_j, y_j) \in \partial\Omega_N^h, \end{cases} \tag{81}$$

where

$$\gamma_j(u) = \max\left\{\frac{\partial R(u_j)}{\partial u};\ \hat{u}_j \le u_j \le \tilde{u}_j\right\}, \tag{82}$$

$$\zeta_j = \sum_{k\in V(j)} \zeta_k, \zeta_k = \xi_k d_k,$$

$$D_n|_{(k,j)} = (D_n(x_k, y_k) + D_n(x_j, y_j))/2,$$

$$d_k = D_n|_{(k,j)} n_i B\left(\frac{\phi_j^{(l+1)} - \phi_k^{(l+1)}}{V_T}\right) \exp\left(\frac{\phi_j^{(l+1)}}{V_T}\right). \tag{83}$$

Step 3.1.8. Set $u_j^{(m)} := u_j^{(m+1)}\ \forall\ j$ and $m := m+1$. Go to Step 3.1.7 until the stopping criteria of the inner iteration are satisfied.
Step 3.1.9. Set $u_j^{(l+1)} := u_j^{(m+1)}\ \forall\ j$.
Step 3.1.10. The discrete hole continuity system is solved analogously as that in Step 3.1.7 for $v_j^{(l+1)}$.
Step 3.1.11. Set $l := l+1$.

Step 3.2. For $\phi_j^{(l+1)}$, do the following sub-steps:
Step 3.2.1. Set $m := 0$ and $\phi_j^{(m)} := \phi_j^{(l)}\ \forall\ j$.
Step 3.2.2. Compute $\phi_j^{(m+1)}$ by solving the discrete potential system (79).
Step 3.2.3. Set $\phi_j^{(m)} := \phi_j^{(m+1)}\ \forall\ j$ and $m := m+1$. Go to Step 3.2.2 until the stopping criteria of the inner iteration are satisfied.
Step 3.2.4. Set $\phi_j^{(l+1)} := \phi_j^{(m+1)}\ \forall\ j$.

Step 3.3 For $u_j^{(l+1)}$, do the following sub-steps:
Step 3.3.1. Set $m := 0$ and $u_j^{(m)} := u_j^{(l)}\ \forall\ j$.
Step 3.3.2. Compute $u_j^{(m+1)}$ by solving the discrete electron system (81).

Step 3.3.3. Set $u_j^{(m)} := u_j^{(m+1)}\ \forall\ j$ and $m := m+1$. Go to Step 3.3.2 until the stopping criteria of the inner iteration are satisfied.
Step 3.3.4. Set $u_j^{(l+1)} := u_j^{(m+1)}\ \forall\ j$.

Step 3.4. For $v_j^{(l+1)}$, the discrete hole continuity system is solved analogously as that in Step 3.3. Set $l := l+1$ and go to Step 3.2 until the stopping criteria of the outer iteration are satisfied.

Step 4. Monotone Iteration on Electron Energy Transport Model:

Step 4.1. Set $m := 0$.
Step 4.2. Set the initial guess $g_j^{(0)} = \tilde{g}_j$ or $\hat{g}_j \forall\ j$ where $\tilde{g}_j$ and $\hat{g}_j$ are constant values for all $(x_j, y_j) \in \overline{\Omega}^h$ that can be easily verified to be an upper and lower solution of g_n, respectively.
Step 4.3. Compute $g_j^{(m+1)}$ by solving the discrete electron system

$$\begin{cases} \eta_j g_j^{(m+1)} + \gamma_j\left(g\right) g_j^{(m+1)} = \sum_{k\in V(j)}^{(m)} \eta_k g_k^{(m)} \\ \quad -R_n(g_j^{(m)}) + \gamma_j\left(g\right) g_j^{(m)}, \quad \forall (x_j, y_j) \in \Omega^h, \\ g_j^{(m+1)} = 300/\exp(\frac{5V_Q}{4V_T}), \quad \forall (x_j, y_j) \in \partial\Omega_D^h\ , \\ \frac{\partial g_j^{(m+1)}}{\partial n} = 0, \quad \forall (x_j, y_j) \in \partial\Omega_N^h, \end{cases} \tag{84}$$

where

$$\gamma_j\left(g\right) = \max\left\{\frac{\partial R_n(g_j)}{\partial g};\ \hat{g}_j \le g_j \le \tilde{g}_j\right\}, \tag{85}$$

$$\begin{aligned} \eta_j &= \sum_{k\in V(j)} \eta_k, \eta_k = \xi_k d_k, \\ \kappa_n|_{(k,j)} &= \left(\kappa_n(x_k, y_k) + \kappa_n(x_j, y_j)\right)/2, \\ \varphi_k &= -\ln(u_k^{(l)}) V_T, \\ d_k &= \kappa_n|_{(k,j)} B(\frac{5\varphi_j - 5\varphi_k}{4V_T}) \exp(\frac{5\varphi_j}{4V_T}). \end{aligned} \tag{86}$$

Step 4.4. Set $g_j^{(m)} := g_j^{(m+1)}\ \forall\ j$ and $m := m+1$. Go to Step 4.3 until the stopping criteria of the inner iteration are satisfied.

Step 5. Error Estimation: For each element, we use the variation of the gradient of the computed potential, concentration, or temperature compared with that of all other elements as an local error indicator. Gradients are calculated with respect to every two nodes in the element, from which the largest one is chosen to be the error indicator. Error indicators are obtained on an element-by-element basis according to the hierarchical tree structure of the OOP data base. A set of criteria on such as global error estimators of approximated solutions, inner iteration, and outer iteration etc. will be verified. If none of

the stopping criteria is satisfied, the adaptive process will continue to Step 6, otherwise it will go to Step 7 for postprocessing the computed solutions.

Step 6. Refinement: Each one of the elements that are associated with error indicators greater than a preset error tolerance is divided into four subelements according the rules of the 1-irregular mesh refinement scheme (see, e.g., [22]). We then move on to Step 3.

Step 7. Postprocessing: All computed solutions are then postprocessed for further analysis of physical phenomena.

Note that the solution procedure in Step 3 consists of an outer loop associated with Gummel's iteration solving (19)-(21) consecutively and an inner loop associated with the monotone iteration for each nonlinear equation. The energy equation (46) is then solved by monotone iteration only. Moreover, as shown in the iterations of (79), (81) and (84), the assembly of global stiffness matrices of the resulting approximation is not required, that is, the solution of discretized nonlinear systems is performed on a node-by-node (regular node) basis. It is also readily seen that the implementation of these iterations is very simple since the monotone parameters (80), (82), and (85) can be easily evaluated. In general, they are evaluated by using the most recent available solution at the current grid point. For this case, the corresponding matrix (74) is simply the diagonal part of the Jacobian matrix used in Newton's method.

6. NUMERICAL EXAMPLES

To demonstrate the effectiveness and accuracy of the numerical algorithms presented in the previous sections, several numerical studies have been made for sample diode and MOSFET device structures. First of all, a benchmark model, namely, an abrupt n^+-n-n^+ silicon diode is used to verify our methods and formulation with the results reported in literature. Numerical experiments are performed on a 0.6-μm silicon diode at $300K$ with $n^+ = 5.0 \times 10^{17} cm^{-3}$ and $n = 2.0 \times 10^{15} cm^{-3}$. The length of the n-region is approximately $0.4\mu m$. The boundary conditions are given as follows:

(1) At the position $x = 0\mu m$: $\phi = V_b$, $u = 1.0$ and $g_n = 300$.
(2) At the position $x = 0.6\mu m$: $\phi = V_O + V_b$, $u = \exp\left(\frac{-V_O}{V_T}\right)$ and $g_n = 300/\exp(\frac{5V_O}{4V_T})$.

The steady state results for this problem are shown in Figs. 5-7 where the applied voltage V_O is taken as $1.5V$ (the solid line) or $2.0V$ (the dashed line). These results agree very well with that previously reported in the literature [1,15,28] and thus show the accuracy of the proposed self-adjoint formulation. This motivates us to further use the formulation in simulations for more complex device structures.

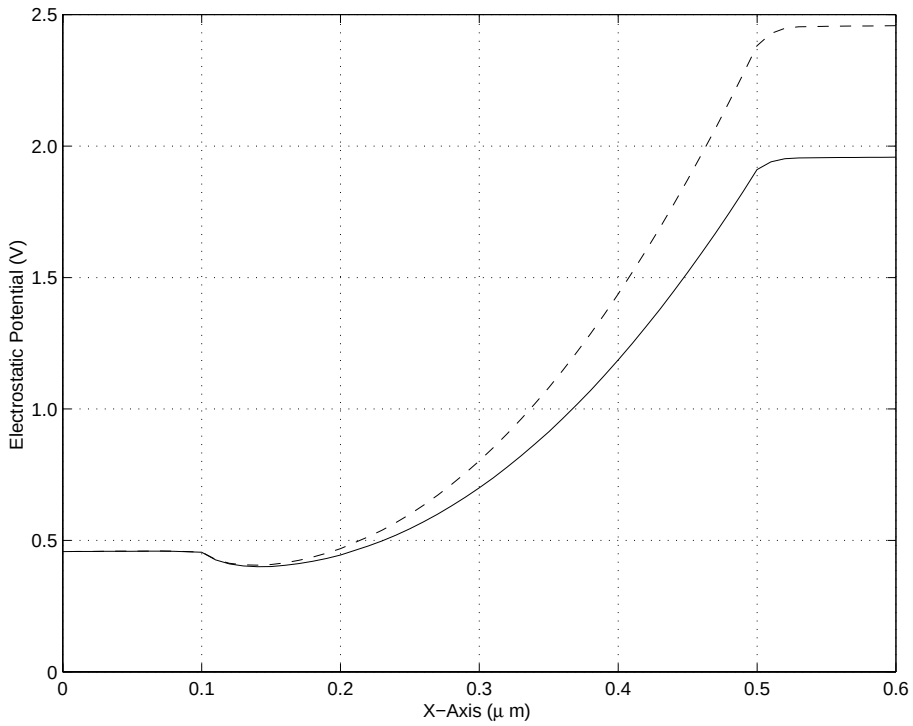

Fig. 5. Electrostatic potential.

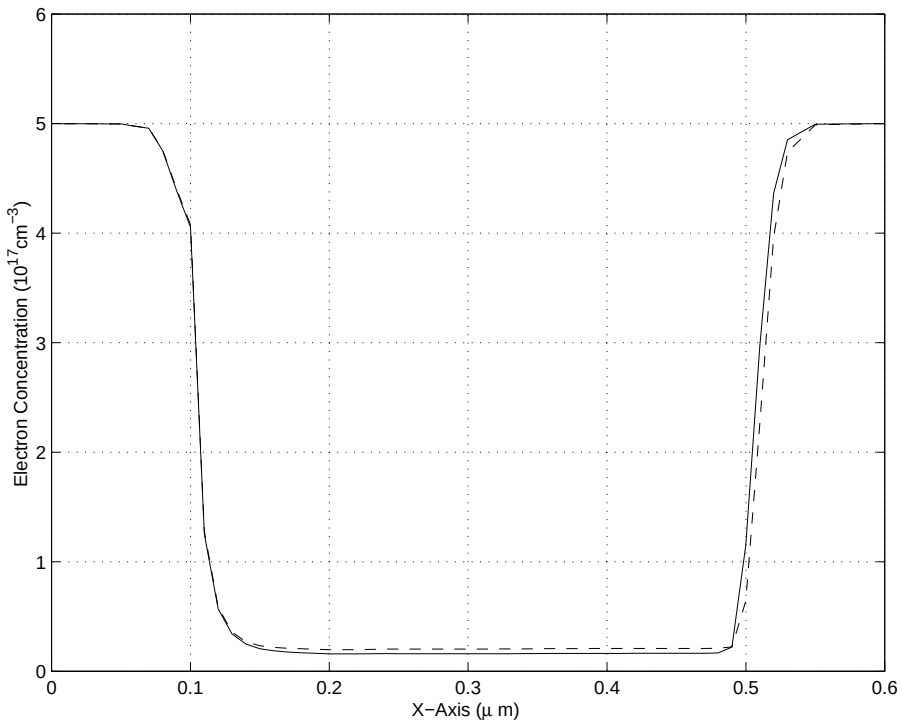

Fig. 6. Electron concentration.

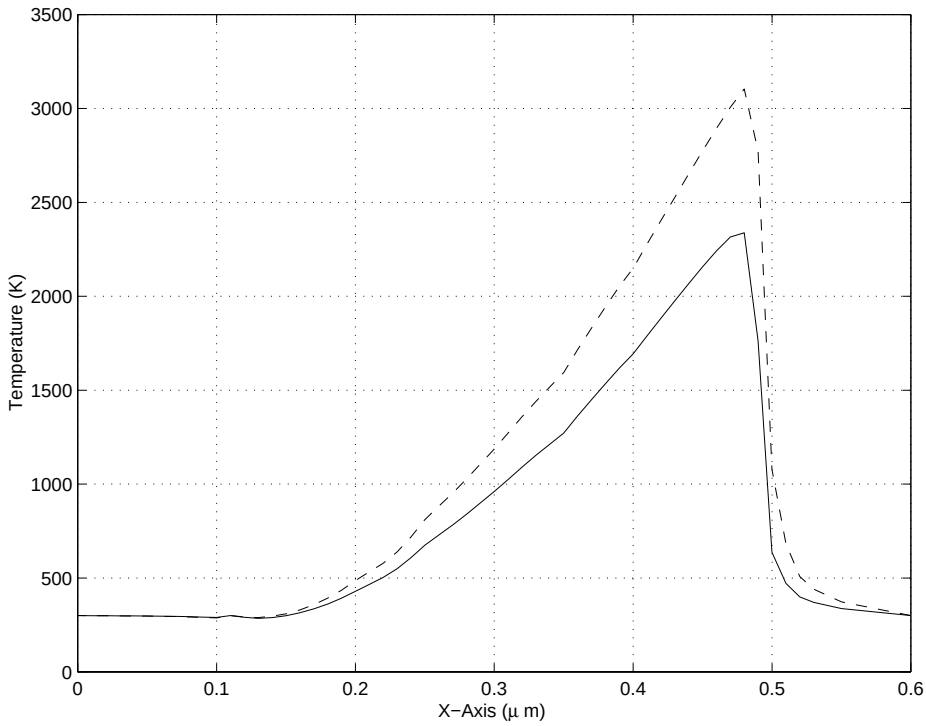

Fig. 7. Temperature.

The next model that we have verified is a typical 2D $n^+ - n - n^+$ deep-submicron diode illustrated in Fig. 8 [1]. The bold lines indicate the contact positions. Contacts A-B and A-F are terminated at a distance of $0.1\mu m$ from the top left corner. In order to simulate a realistic device, contacts are not extended to the full n^+ region near the top left corner as shown in Fig. 8. The channel length L is $0.18\mu m$. The doping profile in the highly doped regions is $5.0 \times 10^{17} cm^{-3}$ and in the lowly doped regions is $2.0 \times 10^{15} cm^{-3}$, i.e.,

$5.0 \times 10^{17} cm^{-3}$ for $0.0 \leq x \leq 0.4\mu m$ and $0 \leq y \leq 0.1\mu m$,
$5.0 \times 10^{17} cm^{-3}$ for $0.3 \leq x \leq 0.4\mu m$ and $0 \leq y \leq 0.4\mu m$,
$5.0 \times 10^{17} cm^{-3}$ for $0.0 \leq x \leq 0.12\mu m$ and $0.28 \leq y \leq 0.4\mu m$, and
$2.0 \times 10^{15} cm^{-3}$ elsewhere.

The associated boundary conditions are as follows:

(1) On the contacts A-B and A-F: $\phi = V_b$, $u = 1.0$ and $g_n = 300$.
(2) On the contacts D-C and D-E: $\phi = 1.0 + V_b$, $u = \exp\left(\frac{-1.0}{V_T}\right)$ and $g_n = 300/\exp(\frac{5}{4V_T})$.
(3) On the boundaries E-F and B-C: Neumann boundary conditions for ϕ, u and g_n.

The applied voltage is $1.0V$. The location of junction layers is evidently shown in the final adaptive mesh in Fig. 9. The corresponding potential, electron

concentration, and temperature distribution are shown in Figs. 10, 11 and 12, respectively.

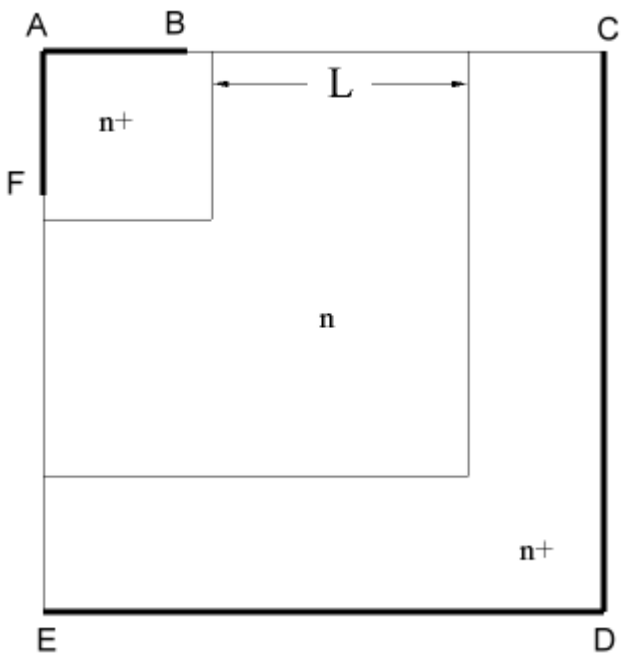

Fig. 8. A $0.4\mu m \times 0.4\mu m$ $n^+ - n - n^+$ silicon device. Contacts are denoted by bold lines.

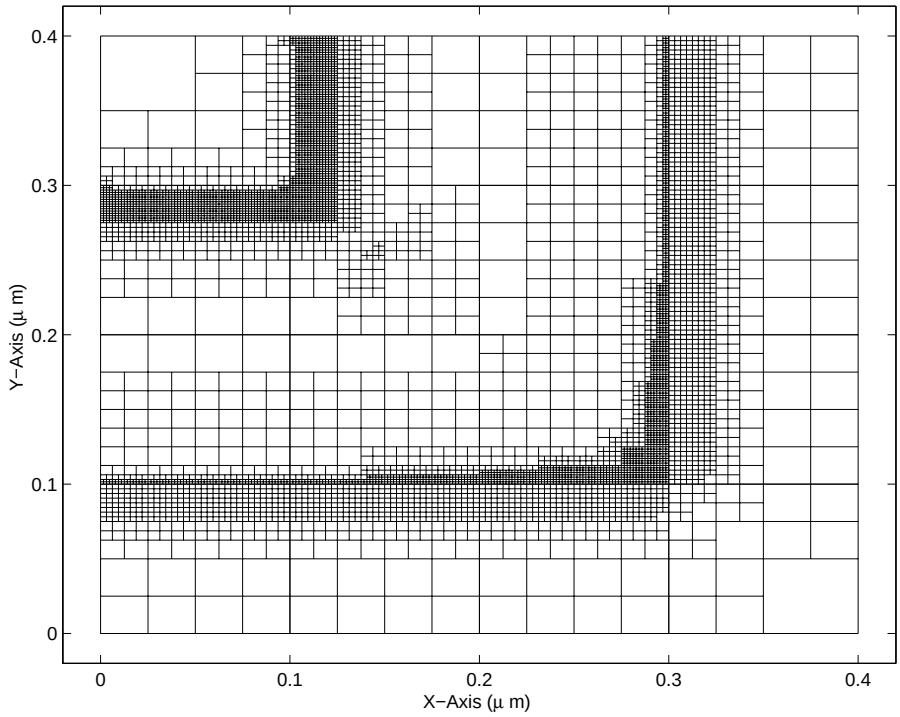

Fig. 9. The final adaptive mesh for a diode.

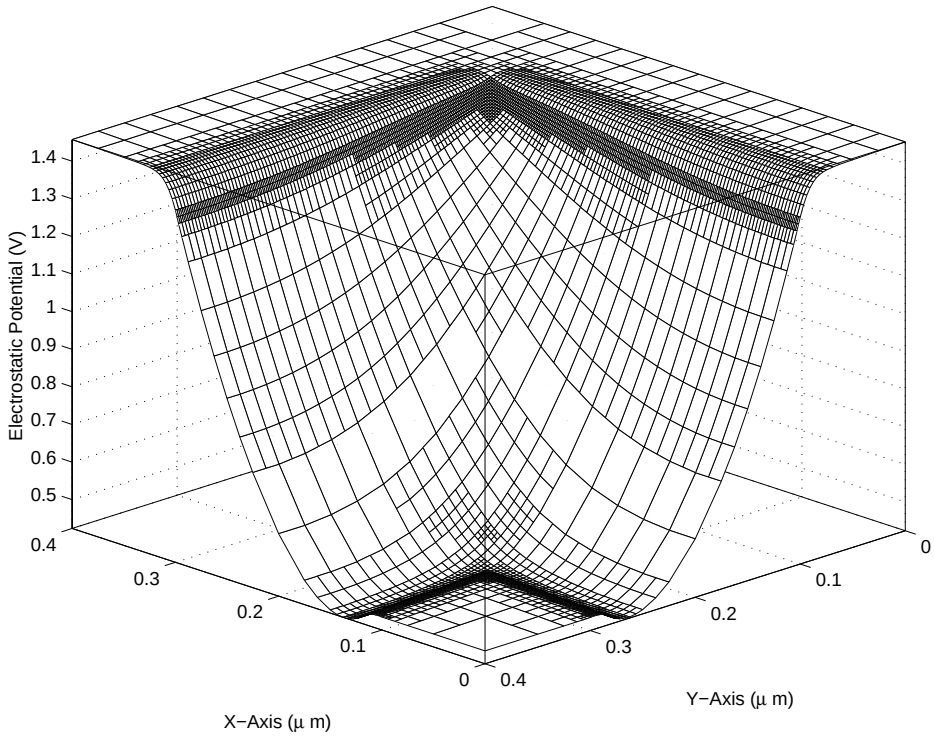

Fig. 10. Electrostatic potential.

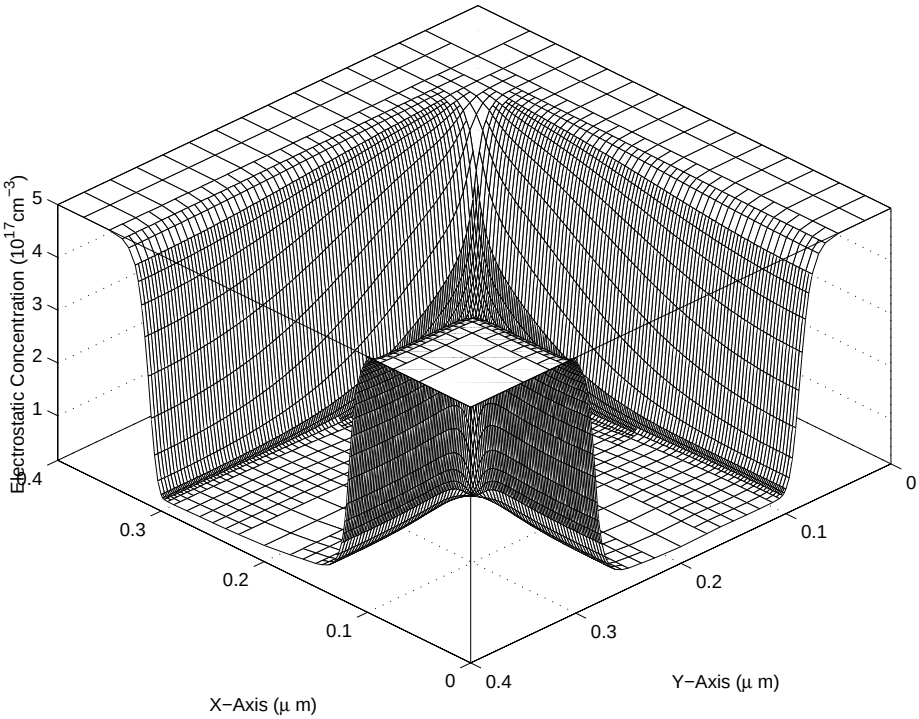

Fig. 11. Electron concentration.

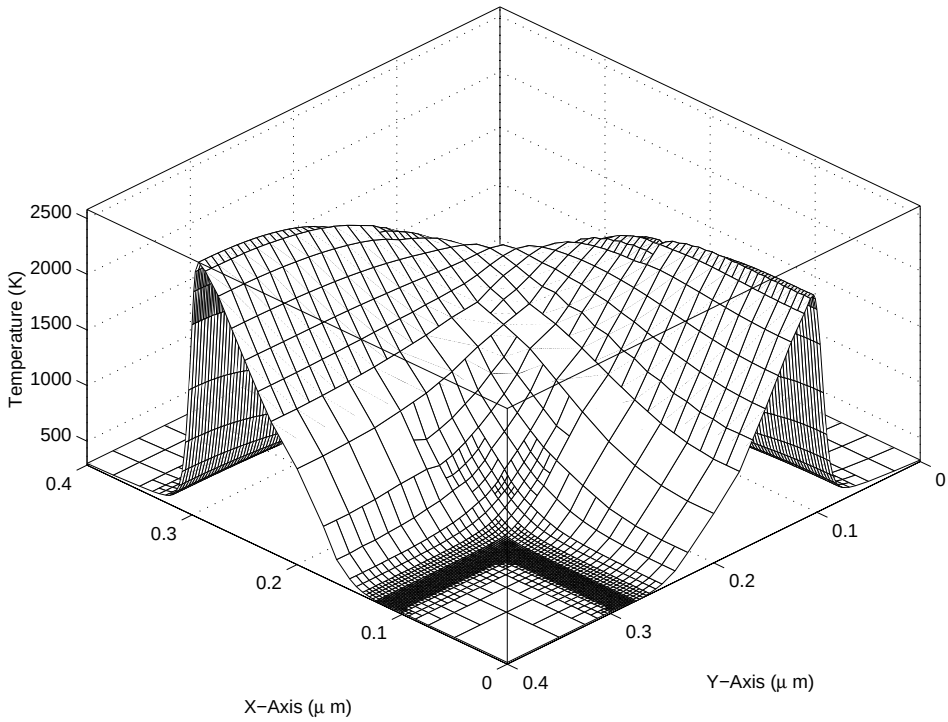

Fig. 12. Electron temperature.

Finally, numerical results were also obtained for an n-MOSFET device with

the channel length of $0.34\mu m$ and with the gate oxide thickness of $7nm$. Fig. 13 illustrates the geometry of the MOSFET device structure. The boundary conditions are as follows:

(1) On B-C: $\phi = V_b$, $u = 1.0$ and $g_n = 300$.
(2) On I-J: $\phi = V_{GS} + V_b$.
(3) On D-E: $\phi = V_{DS} + V_b$, $u = \exp\left(-\frac{V_{DS}}{V_T}\right)$ and $g_n = 300/\exp(\frac{5V_{DS}}{4V_T})$.
(4) On A-F: $\phi = V_{BS} + V_b$, $u = \exp\left(-\frac{V_{BS}}{V_T}\right)$ and $g_n = 300/\exp(\frac{5V_{BS}}{4V_T})$.
(5) On A-B and E-F: Neumann boundary conditions for ϕ, u and g_n.
(6) On C-D: Interface boundary condition for ϕ. Neumann boundary conditions for u and g_n.

Note that the general formulation of the model in Sec. 2 is stated only in the silicon region, i.e., the solution domain Ω in (79) represents this region only. For MOSFET devices, we need to extend the solution domain to the oxide region to which Laplace's equation for the potential is applied. More precisely, the solution step in (79) is replaced by

$$\begin{cases} \xi_j \phi_j^{(m+1)} + \gamma_j(\phi)\,\phi_j^{(m+1)} = \sum_{k\in V(j)} \xi_k \phi_k^{(m)} - F_j(\phi_j^{(m)},\, u_j^{(g)},\, v_j^{(g)}) + \gamma_j(\phi)\,\phi_j^{(m)} \\ \quad \forall (x_j, y_j) \in \Omega^h \\ \phi_j^{(m+1)} = V_O + V_b \;\; \forall (x_j, y_j) \in \partial\Omega_D^h \\ \frac{\partial \phi_j^{(m+1)}}{\partial \nu} = 0 \quad \forall (x_j, y_j) \in \partial\Omega_N^h \\ \xi_j \phi_j^{(m+1)} = \sum_{k\in V(j)} \xi_k \phi_k^{(m)} \;\; \forall (x_j, y_j) \in \Omega_O^h \\ \phi_j^{(m+1)} = V_{GS} + V_b \;\; \forall (x_j, y_j) \in \;\Gamma_{IJ} \\ \frac{\partial \phi_j^{(m+1)}}{\partial \nu} = 0 \quad \forall (x_j, y_j) \in \;\Gamma_{CI} \cup \;\Gamma_{DJ} \\ \varepsilon_s \partial_y \phi_-^{(m+1)} = \varepsilon_d \partial_y \phi_+^{(m+1)} \;\; \forall (x_j, y_j) \in \partial\Gamma_{CD}, \end{cases} \tag{87}$$

where Ω_O denotes the gate oxide region, ε_s and ε_d are the permittivity constants of the silicon and the oxide, and the + and - signs refer to as the limits from the oxide and the silicon regions respectively to the interface.

The device has an elliptical $10^{20} cm^{-3}$ Gaussian doping profiles in the source and drain regions and $10^{16} cm^{-3}$ in the p-substrate region as shown in Fig. 14. The shallow implantation is needed to obtain a 'normal-off' device with positive threshold voltage and the deep implantation is necessary to avoid punchthrough. The junction depth is 0.2 μm and the lateral diffusion under gate is 0.08 μm. With $V_{BS} = 0V$, $V_{DS} = 1.5V$ and $V_{GS} = 1.0V$, Figs. 15, 16, 17, 18, and 19 present the final adaptive mesh, electrostatic potential, electron concentration, electron temperature distribution, and electron current density, respectively. Here the current density is illustrated in the longitudinal direction

since the transversal components are comparatively very small. Fig. 19 clearly shows that electrons are moving very close to the interface and jam to the intersection of the oxide, silicon, and drain contact.

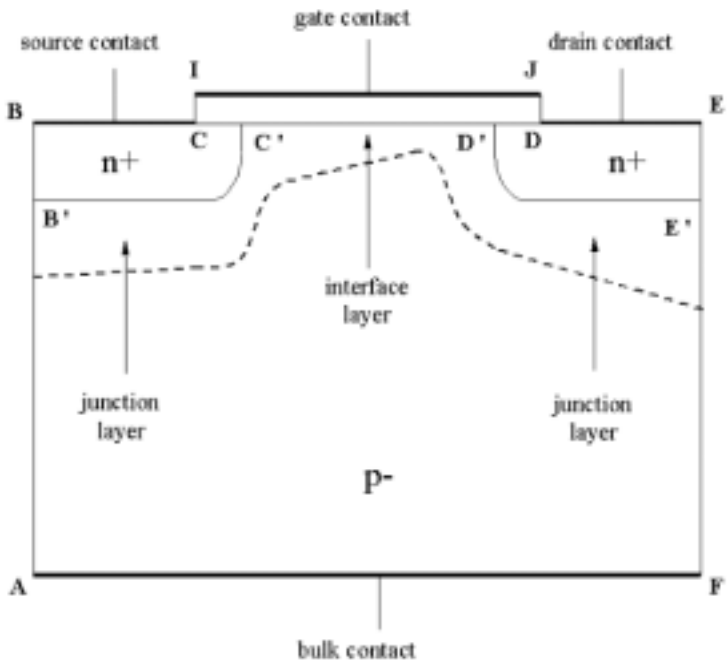

Fig. 13. Geometry of an N-MOSFET device.

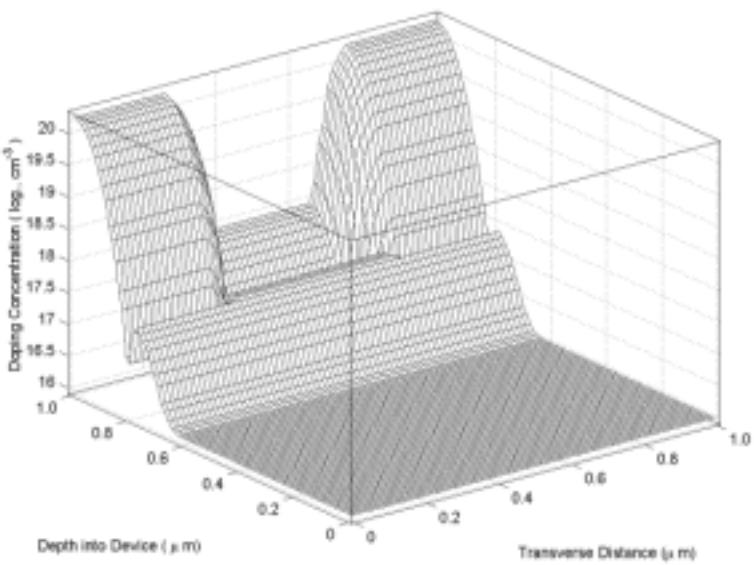

Fig. 14. Doping concentration.

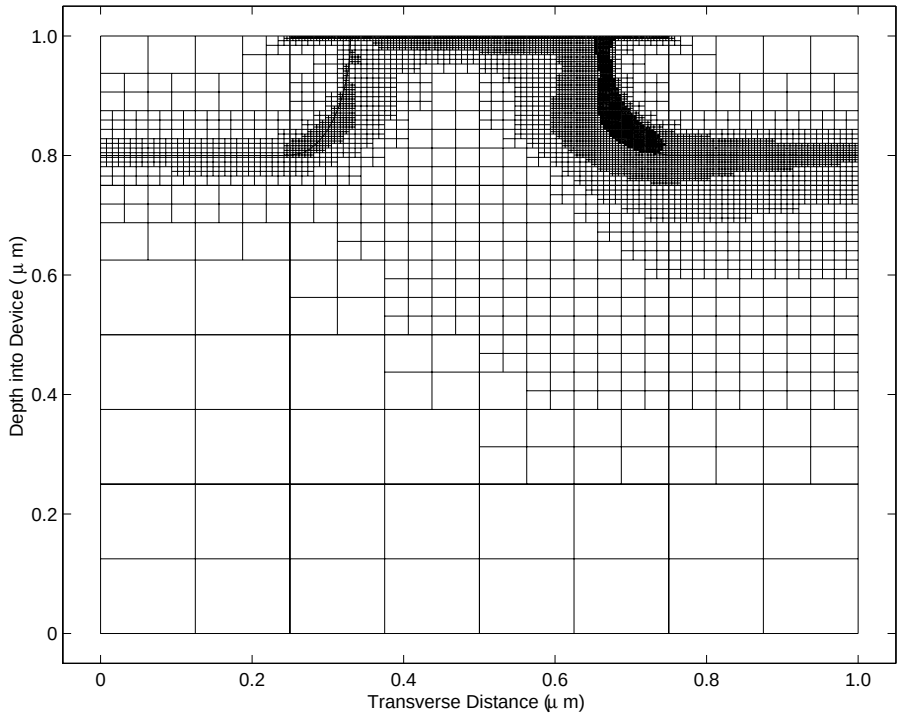

Fig. 15. The final adaptive mesh for an N-MOSFET.

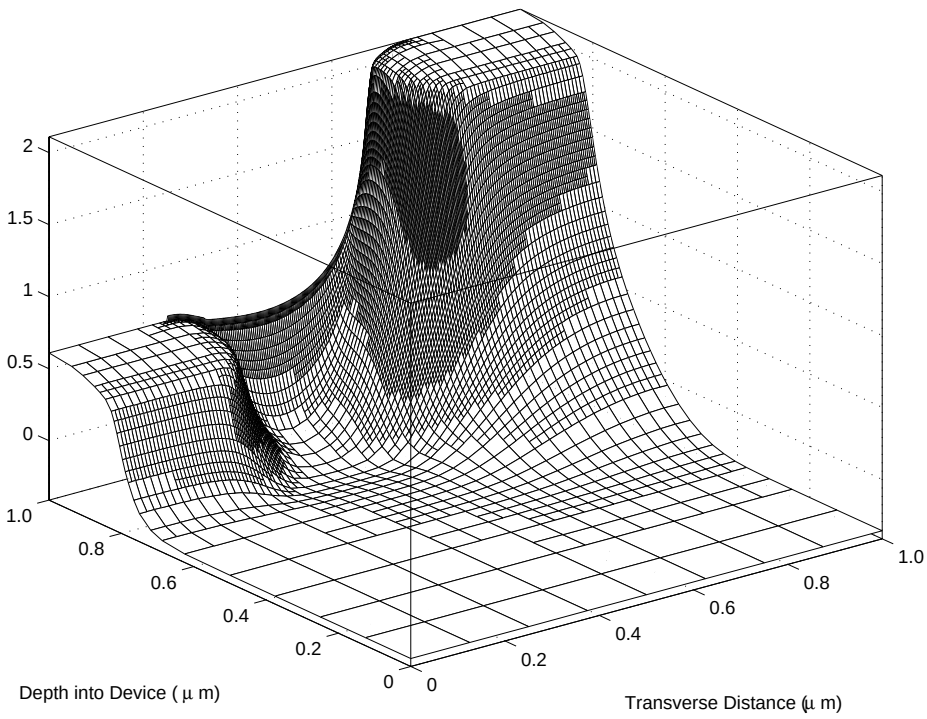

Fig. 16. Electrostatic potential.

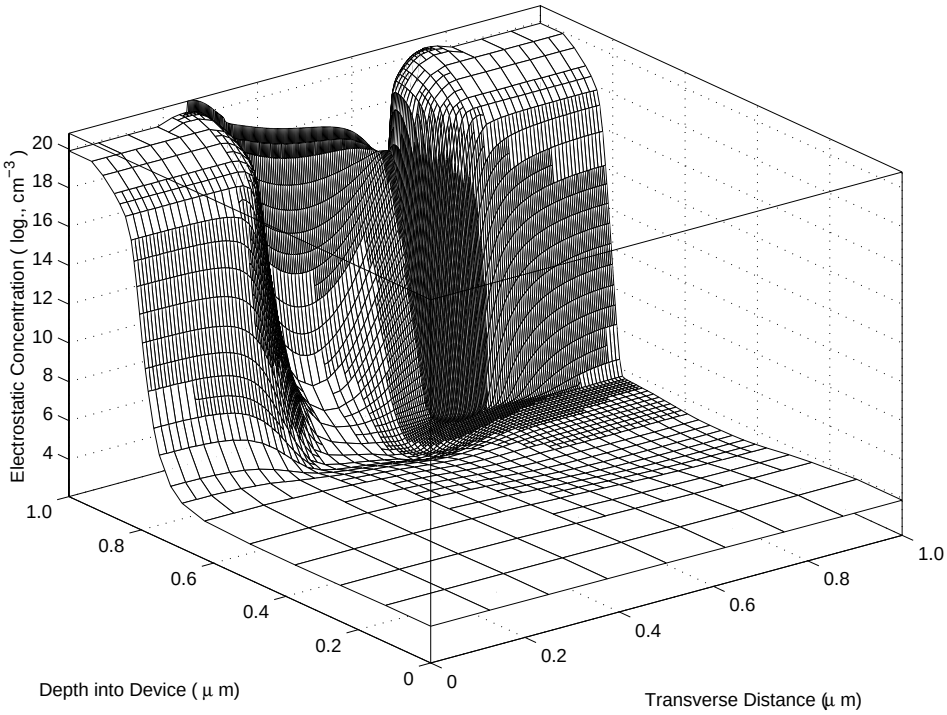

Fig. 17. Electron concentration.

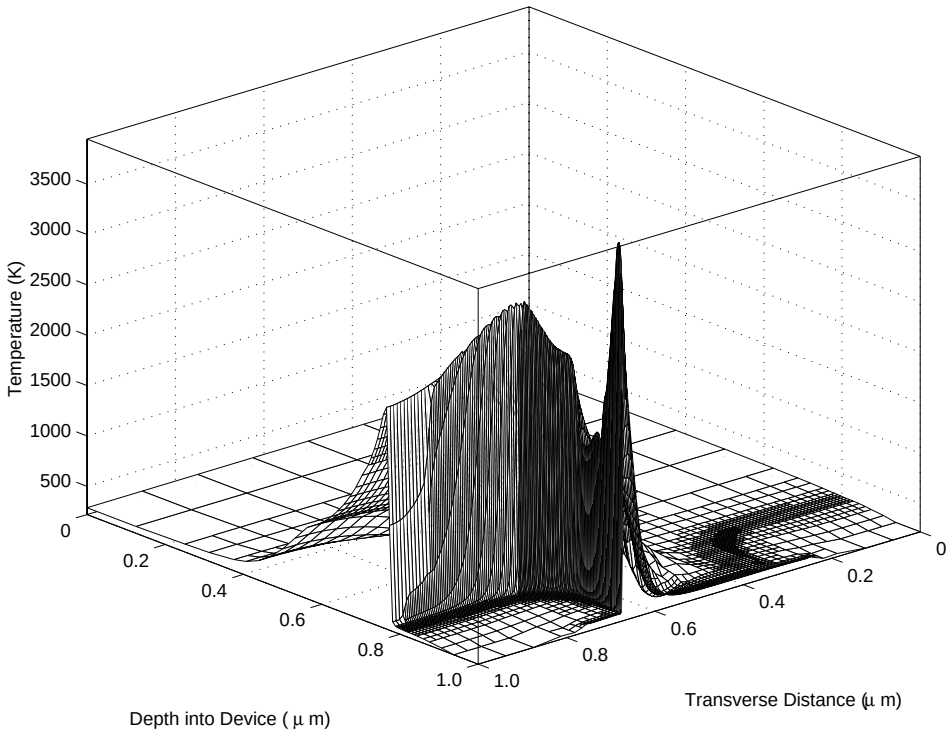

Fig. 18. Electron temperature.

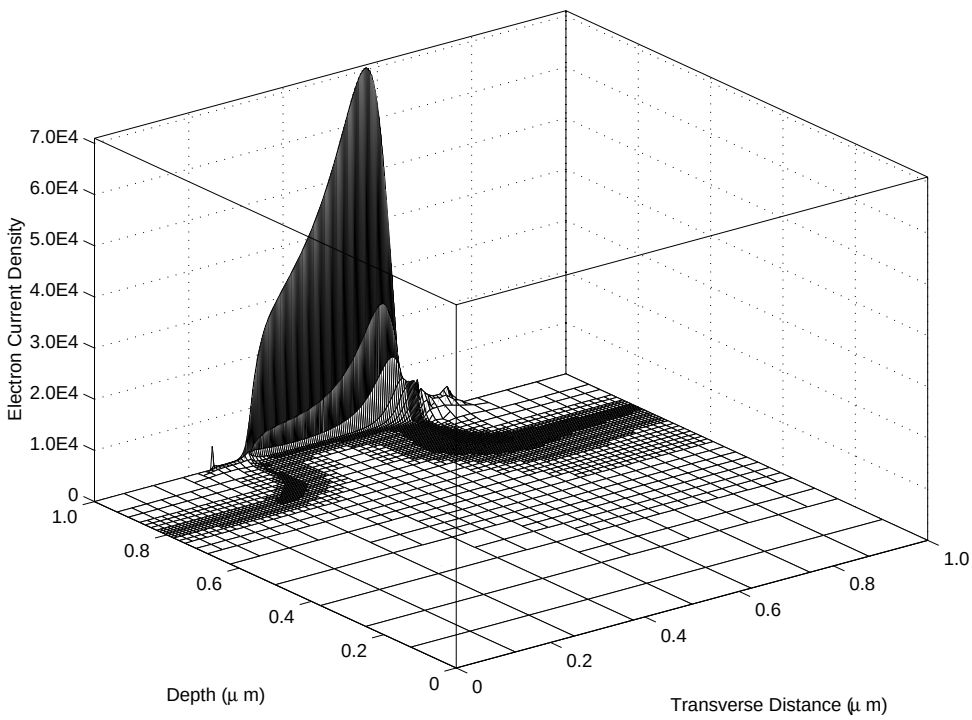

Fig. 19. Electron current density.

We make a remark on the convergence and uniqueness of the iterative method that has been verified by using the upper $\tilde{u} = 1$ and the lower $\hat{u} = 1.0e - 18$ solutions as initial guesses for the electron continuity equation with the voltage $V_{DS} = 1$. The lower solution is readily obtained by the charge neutrality condition $\exp(\frac{-V_{DS}}{V_T}) \simeq 1.7e - 17$. As mentioned above, the dynamic range of the numerical values of u and v is very large in computations. The worst case of the numerics for the Slotboom variables u and v that we have experienced during the course of the development of our code is about of order 10^{100} on our computing systems (Unix on DEC workstations and Linux on Pentium III) with the machine number of order 10^{300}. The range of applied voltages that have been tested with our code is -10V (the reverse bias) to 10V (the forward bias) for a diode and 0V to 5V (the drain bias) and 0V to 4V (the gate bias) for a MOSFET. Although the Slotboom variables are limited for larger bias conditions because of the exponential of the Slotboom variables, applied voltages are much less than $5V$ in submicron electron devices with the gate oxide thickness less than $22nm$ [38]. This may suggest that the Slotboom formulation for classical device properties may be revived in future simulations as the miniaturization of devices continues.

7. CONCLUSION

A self-adjoint model of the energy transport model is proposed in this paper. An iterative method is then developed for the solution of the resulting non-

linear algebraic equations of the model from adaptive finite element approximation. This method exhibits many appealing features for device simulations as follows:

Global convergence: The self-adjointness implies the diagonal dominance of the resulting matrices, which provides not only stability of numerical solutions but also convergence of iterative procedures. The method is shown to generate convergent and unique solutions with simple initial iterates (e.g., the charge neutrality condition). Moreover, a good approximation of the electrostatic potential by the iterative and adaptive methods can lead to the convergence of the solution for the Slotboom variables. The decoupling (Gummel) approach is hence shown to work well with the present iterative method for submicron devices.

Efficiency: The method is highly parallel due to it's Jacobi nature. The conventional approach to obtain an I-V curve is by some continuation procedure from lower to higher voltages due to the local property of Newton's method. This procedure can be very costly in terms of computing time and human work load. The present method can be used to simultaneously (in parallel) simulate device I-V points with independent initial guesses due to the global convergence. The parallelism is thus two folds: the method itself and the I-V computations. This greatly reduces the overall working time for device simulation practitioners.

Easy implementation: This is a constructive method that essentially depends only on one crucial element – the monotone parameter. It can be easily implemented on a node-by-node basis. No global stiffness matrix is necessary. The treatment of the monotone parameter matrix is much more simpler than of the standard Jacobian matrix.

However, it is unknown to us that the self-adjoint formulation for carrier temperatures represents any interesting properties in semiconductor physics. Our numerical results nevertheless show that the carrier temperature can be well approximated with this formulation.

References

[1] N. R. Aluru, A. Raefsky, P. M. Pinsky, K. H. Law, R. J. G. Goossens and R. W. Dutton, A finite element formulation for the hydrodynamic semiconductor device equations, *Comput. Methods Appl. Mech. Engrg.*, 107, 269 (1993).

[2] Y. Apanovich, E. Lyumkis, B. Polsky, A. Shur, and P. Blakey, Steady-state and transient analysis of submicron devices using energy balance and simplified hydrodynamic models, *IEEE Trans. CAD*, 13, 702 (1994).

[3] I. Babuška and J. E. Osborn, Generalized finite element methods: their performance and their relation to mixed methods, *SIAM J. Numer. Anal.,* 20, 510 (1983).

[4] R. E. Bank, D. J. Rose, and W. Fichtner, Numerical methods for semiconductor device simulation, *IEEE Trans. Electron Devices*, ED-30, 1031 (1983).

[5] F. Bosisio, E. Gatti, R. Sacco, F. Saleri, Exponentially fitted mixed finite volumes for energy balance models in semiconductor device simulation , ENUMATH97, Proceedings of the Second European Conference on Numerical Mathematics and Advanced Applications, 188 (1998).

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

[7] J. F. Bürgler, W. M. Coughran, Jr., and W. Fichtner, An adaptive grid refinement strategy for the drift-diffusion equations, *IEEE Trans. Computer-Aided Design*, 10, 1251 (1991).

[8] D. M. Caughey and R. E. Thomas, Carrier mobilities in silicon empirically related to doping and field, *Proc. IEEE,* 52, 2192 (1967).

[9] D. Chang and J. G. Fossum, Simplified energy-balance model for pragmatic multi-dimensional device simulation, *Solid-St. Electron.*, 41, 1795 (1997).

[10] P. Degond, A. Jüngel and P. Pietra, Numerical discretization of energy-transport models for semiconductors with non-parabolic band structure, *SIAM on Scien. Comput.,* 22, 986 (2000).

[11] L. Demkowicz, J. T. Oden, W. Rachowicz, and O. Hardy, Toward a universal *h-p* adaptive finite element strategy. Part 1. Constrained approximation and data structure, *Compt. Meth. Appl. Mech. Engng.*, 77, 79 (1989).

[12] A. Forghieri, R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan and G. Baccarani, A new discretization strategy of the semiconductor equations comprising momentum and energy balance, *IEEE Trans. on CAD,* 7, 231 (1988).

[13] M. Fortin, G. Yang and Y. Perréal, Simulation of the hydrodynamic model of semiconductor devices by a finite element method, *COMPEL,* 15, 4 (1996).

[14] A. F. Franz, G. A. Franz, S. Selberherr, C. Ringhofer, and P. Markowich, Finite boxes - a generalization of finite-difference method suitable for semiconductor device simulation, *IEEE Trans. Electron Devices*, ED-30, 1070 (1983).

[15] C. L. Gardner, J. W. Jerome and D. J. Rose, Numerical methods for the hydrodynamic device model: Subsonic flow, *IEEE Trans. on CAD,* 8, 501 (1989).

[16] J. W. Jerome, Consistency of semiconductor modeling: An existence/stability analysis for the stationary van Roosbroeck system, *SIAM J. Appl. Math.*, 45, 565 (1985).

[17] J. W. Jerome and C.-W. Shu, Transport effects and characteristic modes in the modeling and simulation of submicron devices, *IEEE Trans. CAD*, 14, 917 (1995).

[18] C. Lab and P. Caussignac, An energy-transport model for semiconductor hetrostructure devices: application to AlGaAs/GaAs modfets, *COMPEL,* 18, 61 (1999).

[19] W. Lee, S. E. Laux, M. V. Fischetti, G. Baccarani, A. Gnudi, J. M. C. Stork, J. A. Mandelman, E. F. Crabbé, M. R. Wordeman, F. Odeh, Numerical modeling of advanced semiconductor devices, *IBM J. Res. Develop.*, 36, 208 (1992).

[20] Y. Li, J.-L. Liu, S. M. Sze, and T.-S. Chao, A new parallel adaptive finite volume method for the numerical simulation of semiconductor devices, *Comput. Phys. Commun.*, 142, 285 (2001).

[21] J.-L. Liu, On weak residual error estimation, *SIAM J. Sci. Comput.*, 17, 1249 (1996).

[22] J.-L. Liu, I.-J. Lin, M.-Z. Shih, R.-C. Chen, and M.-C. Hsieh, Object oriented programming of adaptive finite element and finite volume methods, *Appl. Numer. Math.* 21, 439 (1996).

[23] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser, *Semiconductor Equations* (Springer-Verlag, 1990).

[24] P. A. Markowich, C. A. Ringhofer, S. Selberherr, and M. Lentini, A singular perturbation approach for the analysis of the fundamential semiconductor equations, *IEEE Trans. Electron Devices*, ED-30, 1165 (1983).

[25] A. Marrocco and P. Montarnal, Simulation des modeles energy-transport a l'aide des elements finis mixtes, C.R. Acad. Sci. Paris 323, Serie I, 535 (1996).

[26] C. C. McAndrew, K. Singhal, and E. L. Heasell, A consistent nonisothermal extension of the Scharfetter-Gummel stable difference approximation, *IEEE Trans. Electron Devices Lett.*, 6, 446 (1985).

[27] C. V. Pao, Block monotone iterative methods for numerical solutions of nonlinear elliptic equations, *Numer. Math.*, 72, 239 (1995).

[28] M. Rudan, F. Odeh and J. White, Numerical solution of the hydrodynamic model for a one-dimensional device, *COMPEL* 6, 151 (1987).

[29] R. Sacco and F. Saleri, Mixed finite volume methods for semiconductor device simulation, *Numer. Meth. Part. Diff. Eq.*, 13, 215 (1997).

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

[31] A. Schenk, *Advanced Physical Models for Silicon Device Simulation*, Springer-Verlag, 1998.

[32] T. I. Seidman and S. C. Choo, Iterative scheme for computer simulation of semiconductor devices, *Solid-St. Electron.*, 15, 1229 (1972).

[33] S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Wien-New York: Springer, 1984).

[34] S. Selberherr, A. Schutz, and H. W. Potzl, MINIMOS–a two-dimensional MOS transistor analyzer, *IEEE Trans. Electron Devices*, 278, 64 (1980).

[35] J. W. Slotboom, Computer-aided two-dimensional analysis of bipolar transistor, *IEEE Trans. Electron Devices*, 20, 669 (1973).

[36] A. W. Smith and K. F. Brennan, Hydrodynamic simulation of semiconductor devices, *Prog. Quant. Electr.*, 21, 293 (1998).

[37] S. M. Sze, *Physics of Semiconductor Devices* (2nd Ed., Wiley-Interscience, New York, 1981).

[38] S. M. Sze, *Modern Semiconductor Device Physics* (John Wiley & Sons, Inc., New York, 1998).

[39] G.-L. Tan, X.-L. Yuan, Q.-M. Zhang, W. H. Ku and A.-J. Shey, Two-dimensional semiconductor device analysis based on new finite-element discretization employing the S-G scheme, *IEEE Trans. CAD*, 8, 468 (1989).

[40] T. W. Tang, Extension of the Scharfetter-Gummel algorithm to the energy balance euqation, *IEEE Trans. Electron Devices*, 31, 64 (1984).