RAIRO
MODÉLISATION MATHÉMATIQUE
ET ANALYSE NUMÉRIQUE

J. J. H. MILLER
SONG WANG

An analysis of the Scharfetter-Gummel box method for the stationary semiconductor device equations

RAIRO – Modélisation mathématique et analyse numérique, tome 28, nº 2 (1994), p. 123-140.
<http://www.numdam.org/item?id=M2AN_1994__28_2_123_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol. 28, n° 2, 1994, p. 123 à 140)

AN ANALYSIS OF THE SCHARFETTER-GUMMEL BOX METHOD FOR THE STATIONARY SEMICONDUCTOR DEVICE EQUATIONS(*)

by J. J. H. MILLER and SONG WANG (1)

Communicated by R. TEMAM

Abstract. — *An exponentially fitted box method, known as the Scharfetter-Gummel box method, for the semiconductor device equations in the Slotboom variables is analysed. The method is formulated as a Petrov-Galerkin finite element method with piecewise exponential basis functions on a triangular Delaunay mesh. No restriction is imposed on the angles in the triangulation. The stability of the method is proved and an error estimate for the Slotboom variables in a discrete energy norm is given. When restricted to the two continuity equations the error estimate depends only on the first-order seminorm of the exact flux and the approximation error of the zero order and inhomogeneous terms. This is in contrast to standard error estimates which depend on the second order seminorm of the exact solution. The evaluation of the ohmic contact currents is discussed and it is shown that the approximate ohmic contact currents are convergent and conservative.*

Résumé. — *On étudie ici la méthode de « box Scharfetter-Gummel ». Cette méthode est formulée comme une méthode de Petrov Galerkin pour les éléments finis avec comme fonctions de base des fonctions exponentielles par morceaux sur un maillage Delaunay triangulaire. On n'impose aucune restriction sur les angles du maillage. La stabilité de la méthode est établie et on donne une estimation d'erreur pour les variables de Slotboom dans une norme dérivée d'une énergie discrète. Quand on restreint cette estimation aux deux équations de continuité, elle ne dépend que de la semi-norme du premier ordre du flux exact, de l'erreur d'approximation des termes d'ordre zéro et des termes non homogènes. Ce résultat contraste avec les estimations d'erreur standards qui dépendent de la semi-norme d'ordre 2 de la solution exacte. En outre on calcule les courants de contacts ohmiques et on montre qu'on a une propriété de convergence.*

1. INTRODUCTION

Solutions of the semiconductor device equations display interior layers due to the abrupt change in doping profile. Applications of classical discretisation

(*) Manuscript received August 1992, revised February 1993.

(1) *Present address :* School of Mathematics, The University of New South Wales, Kensington, NSW 2033, Australia.

Department of Mathematics, Trinity College, Dublin 2, Ireland.

methods such as the central difference or the linear finite element method to these equations often yield results with non-physical properties such as spurious oscillations. To overcome such difficulties Scharfetter and Gummel [15] proposed a novel method in the one dimensional case. Extensions of the Scharfetter-Gummel method to higher dimensions have been presented by different authors [2, 3, 8, 9, 10]. One of them is the Scharfeter-Gummel box method (*cf.*, for example, [3, 9]) which has been widely used in semiconductor device simulation. Although the Scharfetter-Gummel box method works well in practice, an understanding of the underlying mathematics is still very limited. In [12] and [13] Mock analysed the method using discrete subspaces. His approach, however, gives error estimates for only the approximate flux. In [1] Bank and Rose gave an error estimate for the method when it is applied to a linear Poisson equation. In this case the Scharfetter-Gummel box method reduces to the standard central difference box method (*cf.* [7]). Based on a mixed finite element formulation Miller and Wang [10] proposed and analysed a method similar to the Scharfetter-Gummel box method. This approach could also be used to provide an analysis for the Scharfetter-Gummel box method, but the resulting error estimate would still be not satisfactory.

In this paper we analyse the Scharfetter-Gummel box method in the finite element framework. This error analysis is more satisfactory because the only assumption on the mesh is that it is a triangular Delaunay mesh, and so no restriction on the angles (such as no obtuse angles) is needed. This property has been well known to engineers in practice, but it has not been analyzed satisfactorily before now. Moreover, the present analysis can be extended without difficulty to three dimensions and to a general Delaunay mesh consisting of a mixture of triangles and rectangles in two dimensions and tetrahedra, pentahedra and hexahedra in three dimensions. A similar analysis is given in [11] for a singularly perturbed problem. The paper is organised as follows.

In the next section we give a mathematical description of the semiconductor device problem. For the sake of mathematical simplicity the original equations in the variables of electron and hole densities are transformed into equations in the Slotboom variables. It should be noted that the former variables are physically more interesting. In Section 3 we reformulate the method as a Petrov-Galerkin finite element method with exponential basis functions. In Section 4 the stability of the method is proved and an $O(h)$ error estimate for the approximate solution in the Slotboom variables in a discrete energy norm is given. When applied to the two continuity equations the error estimate depends only on the first order seminorm of the exact flux and the approximation error of the zero order and inhomogeneous terms. This is in contrast to the standard error estimate for the piecewise linear finite element method which depends on the second order seminorm of the exact solution. It is likely that the flux is physically better behaved than the exact solution, although mathematically this has not been proved. It should be noted that the error constant in this estimate still depends exponentially on the maximum or minimum values of the exact electrostatic potential. The evaluation of ohmic contact currents is discussed in Section 5 where the computed ohmic contact currents are shown to be convergent and conservative. The error estimate for these currents depends on the first order seminorm of the exact flux and the approximation error of the inhomogeneous term.

2. STATEMENT AND REFORMULATION OF THE PROBLEM

The stationary behaviour of semiconductor devices may be described by the following (scaled) nonlinear system of second-order elliptic equations ([19])

$$\nabla^2 \psi - n + p = -N \tag{2.1}$$

$$\nabla \cdot (\nabla n - n \nabla \psi) - R(\psi, n, p) = 0 \tag{2.2}$$

$$\nabla \cdot (\nabla p + p \nabla \psi) - R(\psi, n, p) = 0 \tag{2.3}$$

with appropriate boundary conditions, where ψ is the electrostatic potential, n is the electron concentration, p is the hole concentration, N denotes the doping function and R denotes the recombination/generation rate which is assumed to be monotone with respect to n and p, i.e.

$$\frac{\partial R}{\partial n} \geqslant 0 \,, \quad \frac{\partial R}{\partial p} \geqslant 0 \,.$$

Using Gummel's method ([6]) and Newton's method we can decouple and linearise the above system so that at each iteration step we sequentially solve a Poisson equation and two continuity equations. We assume that the Dirichlet boundary conditions for ψ, n and p are homogeneous. The inhomogeneous case can be transformed into the homogeneous case by substracting a special function satisfying the boundary conditions. We consider the following decoupled linearised continuity equation for the electron concentration n and the appropriate boundary conditions

$$-\nabla \cdot \mathbf{f} + Gu = F \quad \text{in} \quad \Omega \tag{2.4}$$

$$\mathbf{f} = \nabla u - u \nabla \psi \tag{2.5}$$

$$u \big|_{\partial \Omega_D} = 0 \,, \quad \mathbf{f} \cdot \mathbf{n} \big|_{\partial \Omega_N} = 0 \tag{2.6}$$

where $\Omega \in \mathbb{R}^2$, $\partial \Omega = \partial \Omega_D \cup \partial \Omega_N$ is the boundary of Ω, $\partial \Omega_D \cap \partial \Omega_N = \emptyset$, $\mathbf{n}$ denotes the unit outward normal vector on $\partial \Omega$, $G \in C^0(\bar{\Omega}) \cap H^1(\Omega)$,

$F \in L^2(\Omega)$ and the flux $\mathbf{f}$ is the electron current. The coefficient function $G = \frac{\partial R}{\partial n}$ and thus we have $G \geq 0$. Similar results for the corresponding equations for ψ and p follow immediately on replacing $\nabla\psi$ by 0 and $-\nabla\psi$ respectively in (2.4).

In what follows $L^2(S)$, $L^\infty(S)$ and $W^{m,s}(S)$ denote the usual Sobolev spaces with norms $\|\cdot\|_{0,S}$, $\|\cdot\|_{\infty,S}$ and $\|\cdot\|_{m,p,S}$, respectively, for any measurable open set $S \subset \mathbb{R}^n$ $(n = 1, 2)$. The inner product on $L^2(S)$ and $(L^2(S))^2$ is denoted by $(\cdot,\cdot)_S$ and the kth order seminorm on $W^{m,p}(S)$ by $|\cdot|_{k,p,S}$. The Sobolev space $W^{m,2}(S)$ is written $H^m(S)$ with corresponding norm and kth order seminorm $\|\cdot\|_{m,S}$ and $|\cdot|_{k,S}$, respectively. When $S = \Omega$, we omit the subscript S in the above notation. We put $\mathbf{L}^2(\Omega) = (L^2(\Omega))^2$, $\mathbf{L}^\infty(\Omega) = (L^\infty(\Omega))^2$ and $H_D^1(\Omega) = \left\{v \in H^1(\Omega) : v\big|_{\partial\Omega_D} = 0\right\}$. We use $|\cdot|$ to denote absolute value, Euclidean length, or area depending on the context.

We now reformulate (2.4-6) by introducing the Slotboom variable w (*cf.* [16]) defined by

$$w = e^{-\psi} u . \tag{2.7}$$

In terms of w the equations and boundary conditions have the form

$$-\nabla \cdot \mathbf{f} + Ge^{\psi} w = F \quad \text{in} \quad \Omega \tag{2.8}$$

$$\mathbf{f} = e^{\psi} \nabla w \tag{2.9}$$

$$w\big|_{\partial\Omega_D} = \mathbf{f} \cdot \mathbf{n}\big|_{\partial\Omega_N} = 0 . \tag{2.10}$$

The Bubnov-Galerkin variational problem corresponding to (2.8-10) is

PROBLEM 2.1 : *Find $w \in H_D^1(\Omega)$ such that for all $v \in H_D^1(\Omega)$*

$$(e^{\psi} \nabla w, \nabla v) + (Ge^{\psi} w, v) = (F, v) . \tag{2.11}$$

Since $G \geq 0$, using standard arguments, we know that there exists a unique solution to Problem 2.1.

3. FORMULATION OF THE SCHARFETTER-GUMMEL BOX METHOD

To discuss the method we first define some meshes on Ω. Let $\mathcal{T}$ denote a family of triangulations of Ω

$$\mathcal{T} = \{T_h : 0 < h \leq h_0\}$$

where T_h denotes a triangulation of Ω with each triangle t having diameter less than or equal to h and h_0 is a positive constant which is smaller than the diameter of Ω. For each $T_h \in \mathcal{T}$, let $X_h = \{x_i\}_1^N$ denote the vertices of T_h and $E_h = \{e_i\}_1^M$ the edges of T_h. We assume that the nodes in X_h and the edges in E_h are numbered such that $X_h' = \{x_i\}_1^{N'}$ and $E_h' = \{e_i\}_1^{M'}$ are respectively the set of nodes in X_h not on $\partial\Omega_D$ and the set of edges in E_h not on $\partial\Omega_D$.

DEFINITION 3.1 : *T_h is a Delaunay triangulation if, for every $t \in T_h$, the circumcircle of t contains no other vertices in X_h (cf.* [4]).

We assume henceforth that each $T_h \in \mathcal{T}$ is a Delaunay triangulation.

DEFINITION 3.2 : *The Dirichlet tessellation D_h, corresponding to the triangulation T_h, is defined by $D_h = \{d_i\}_1^N$ where the tile*

$$d_i = \{x \in \Omega : |x - x_i| < |x - x_j|, x_j \in X_h, j \neq i\} \tag{3.1}$$

for all $x_i \in X_h$ (cf. [5]).

We remark that for each $x_i \in X_h$, the boundary ∂d_i of the tile d_i is the polygon having as its vertices the circumcentres of all triangles with the common vertex x_i. Each segment of ∂d_i is perpendicular to one of the edges sharing the vertex x_i.

The Dirichlet tessellation D_h is a non-triangular mesh dual to the Delaunay mesh T_h. We define $\mathcal{D} = \{D_h : 0 < h \leq h_0\}$ to be the family of all such meshes. The subset of D_h corresponding to X_h' is denoted by $D_h' = \{d_i\}_1^{N'}$.

A second non-triangular mesh, dual to T_h, is defined as follows. With each edge $e_k \in E_h$ we associate an open box b_k which is the interior of the polygon having as its vertices the two end-points of e_k and the circumcentres of the triangles having e_k as a common edge. If e_k is not on $\partial\Omega$ the region b_k consists of two triangles. The set $B_h = \{b_k\}_1^M$ forms a box mesh which is also dual to T_h. We let $\mathcal{B} = \{B_h : 0 < h \leq h_0\}$ denote the family of all such meshes.

Corresponding to the two meshes T_h and D_h, we now construct two finite-dimensional spaces $L_h \subset L^2(\Omega)$ and $M_h \subset L^2(\Omega)$, respectively, each of dimension N'.

To construct M_h we define a set of piecewise constant basis functions ξ_i, $(i = 1, 2, \ldots, N)$ corresponding to the mesh D_h as follows

$$\xi_i = \begin{cases} 1 & \text{on } d_i \\ 0 & \text{otherwise}. \end{cases}$$

We then define $M_h = \text{span } \{\xi_i\}_1^{N'}$.

For each $e_{i,j} \in E_h'$ connecting x_i and x_j we define an exponential function $\phi_{i,j}$ on $e_{i,j}$ by

$$\begin{aligned} &\frac{d}{d\mathbf{e}_{i,j}}\left(e^{\psi}\frac{d\phi_{i,j}}{d\mathbf{e}_{i,j}}\right) = 0 \\ &\phi_{i,j}(x_i) = 1\,, \quad \phi_{i,j}(x_j) = 0 \end{aligned} \tag{3.2}$$

where $\mathbf{e}_{i,j}$ is the unit vector from x_i to x_j. From the definition it follows that $\mathbf{e}_{i,j} = -\mathbf{e}_{j,i}$. We then extend $\phi_{i,j}$ to $b_{i,j}$ by defining it to be constant along directions perpendicular to $e_{i,j}$. This exponential function can be extended to Ω as follows

$$\phi_i = \begin{cases} \phi_{i,j} & \text{on} \quad b_{i,j} \quad \text{if} \quad j \in I_i \\ 0 & \text{otherwise} \end{cases}$$

where

$$I_i = \{j : e_{i,j} \in E_h\} \tag{3.3}$$

denotes the index set of all neighbour nodes of x_i. The support of ϕ_i is star-shaped. We put $L_h = \text{span}\ \{\phi_i\}_1^{N'}$. Obviously we have $L_h \subset L^2(\Omega)$.

For simplicity we make the following assumption, which implies that ψ is piecewise linear on Ω.

ASSUMPTION 3.1 : *For every $t \in T_h$ the function ψ is linear on t.*

We comment that since (2.1-3) are solved iteratively using Gummel's method, ψ used in (2.2) is the numerical solution of (2.1) and thus we can always use the piecewise linear interpolant of this numerical solution.

For any sufficiently smooth function w we can easily show that for each $e_{i,j} \in E_h'$ the L_h-interpolant w_I of w satisfies

$$\begin{aligned} &\frac{d}{d\mathbf{e}_{i,j}}\left(e^{\psi}\frac{dw_I}{d\mathbf{e}_{i,j}}\right) = 0 \quad \text{on} \quad \mathbf{e}_{i,j} \\ &w_I(x_i) = w(x_i)\,, \quad w_I(x_j) = w(x_j)\,. \end{aligned}$$

From this it follows that

$$f_{i,j} \equiv e^{\psi}\frac{dw_I}{d\mathbf{e}_{i,j}} = e^{\psi_i}B(\psi_i - \psi_j)\frac{w_j - w_i}{|e_{i,j}|} \tag{3.4}$$

where $B(x)$ denotes the Bernoulli function defined by

$$B(x) = \begin{cases} \dfrac{x}{e^x - 1} & x \neq 0 \\ 1 & x = 0 \end{cases} \tag{3.5}$$

$\psi_i = \psi(x_i)$ and $w_i = w(x_i)$. It is easy to see that $f_{i,j} = -f_{j,i}$. Let $\mathbf{f}_{i,j} = \mathbf{f} \cdot \mathbf{e}_{i,j}$. When restricted to $e_{i,j}$ it is easy to see that $f_{i,j}$ is the projection of $\mathbf{f}_{i,j} \in \mathbf{L}^2(e_{i,j})$ on to the space of all zero-order polynomials on $e_{i,j}$ with respect to the weighted inner product $\int_{e_{i,j}} e^{-\psi} fg\, ds$ for any $f, g \in L^2(e_{i,j})$. Thus, from the conventional projection theorem (*cf.* [14, Theorem 6.8]) we have

$$\int_{e_{i,j}} |\mathbf{f}_{i,j} - f_{i,j}|^2\, ds \leq C\, |e_{i,j}|^2 |\mathbf{f}_{i,j}|^2_{1, e_{i,j}}$$

where C denotes a generic positive constant, independent of h. Using a Taylor expansion we obtain from the above inequality

$$\left[\|\mathbf{f}_{i,j} - f_{i,j}\|_{\infty, \bar{e}_{i,j}} + |e_{i,j}| \frac{d\mathbf{f}(\eta)}{d\mathbf{e}_{i,j}} \right]^2 \leq C\, |e_{i,j}|\, |\mathbf{f}_{i,j}|^2_{1, e_{i,j}}$$

for some $\eta \in \bar{e}_{i,j}$. Taking the square root on both sides, and using the relation $|a| - |b| \leq |a+b|$ for any real numbers a and b, we get

$$\begin{aligned} \|\mathbf{f}_{i,j} - f_{i,j}\|_{\infty, \bar{e}_{i,j}} &\leq C\, |e_{i,j}|^{1/2} |\mathbf{f}_{i,j}|_{1, e_{i,j}} + |e_{i,j}|\, |\mathbf{f}_{i,j}|_{1, \infty, e_{i,j}} \\ &\leq C\, |e_{i,j}|\, |\mathbf{f}_{i,j}|_{1, \infty, e_{i,j}} \end{aligned}$$

where we have used the inequality $|\mathbf{f}_{i,j}|_{1, e_{i,j}} \leq |\mathbf{f}_{i,j}|_{1, \infty, e_{i,j}} |e_{i,j}|^{1/2}$. Finally, since $e_{i,j} \subset b_{i,j}$, it is easy to see that

$$\|\mathbf{f} \cdot \mathbf{e}_{i,j} - f_{i,j}\|_{\infty, b_{i,j}} \leq Ch\, |\mathbf{f}|_{1, \infty, b_{i,j}} \tag{3.6}$$

where C is a positive constant, independent of h and w.

Let $C(\bar{\Omega})$ denote the space of all functions which are continuous on $\bar{\Omega}$. We introduce the mass lumping operator $P : C(\bar{\Omega}) \mapsto L_h$ such that

$$P(u)(x) = \sum_{x_i \in X_h} u(x_i)\, \xi_i(x) \quad \text{for all} \quad x \in \bar{\Omega}\,. \tag{3.7}$$

Using the two finite dimensional spaces L_h and M_h, we now define the following discrete Petrov-Galerkin problem corresponding to Problem 2.1.

PROBLEM 3.1 : *Find $w_h \in L_h$ such that for all $v_h \in M_h$*

$$a(w_h, v_h) + (P(Ge^{\psi} w_h), v_h) = (\hat{F}, v_h) \tag{3.8}$$

where $\hat{F}$ is an approximation to F and $a(\cdot\,,\,\cdot)$ denotes the bilinear form on $L_h \times M_h$ defined by

$$a(w_h, v_h) = -\sum_{d \in D_h} \int_{\partial d \setminus \partial \Omega} e^{\psi} \nabla w_h \cdot \mathbf{n} \gamma_0(v_h|_d)\, ds\,. \tag{3.9}$$

Here $v_h|_d$ denotes the restriction of v_h to d, $\gamma_0(v|_d)$ denotes the continuous extension of $v|_d$ to ∂d and $\mathbf{n}$ is the unit outward normal vector on ∂d.

Let $w_h = \sum_{i=1}^{N'} w_i \phi_i$, where $\{w_i\}_1^{N'}$ is a set of constants. Substituting this into (3.8) and taking $v_h = \xi_j$, we get, for $j = 1, 2, \ldots, N'$.

$$-\sum_{i=1}^{N'} \int_{\partial d_j} e^{\psi} w_i \nabla \phi_i \cdot \mathbf{n}\, ds + G_j e^{\psi_j} w_j |d_j| = \int_{d_j} \hat{F}\, dx \tag{3.10}$$

where $G_j = G(x_j)$. Let the line segment $l_{j,k} = \partial d_j \cap \partial d_k$. It is easy to check that $\partial d_j = U_{k \in I_j} l_{j,k}$ and $|l_{j,k}| = \dfrac{2|b_{j,k}|}{|e_{i,k}|}$. Thus, for $j = 1, 2, \ldots, N'$, we have from (3.10)

$$-\sum_{k \in I_j} \int_{\partial d_{j,k}} \left(e^{\psi} \frac{dw_h}{d\mathbf{e}_{j,k}} \right) \bigg|_{b_{j,k}} ds + G_j e^{\psi_j} w_j |d_j| = \int_{d_j} \hat{F}\, dx \tag{3.11}$$

where I_j is the index set defined in (3.3). Noticing that $\mathbf{n} = \mathbf{e}_{j,k}$ in (3.1) and using (3.4) we finally obtain from the above

$$\sum_{k \in I_j} e^{\psi_j} B(\psi_j - \psi_k) \frac{w_j - w_k}{|e_{j,k}|} |l_{j,k}| + G_j e^{\psi_j} w_j |d_j| = \int_{d_j} \hat{F}\, dx\,. \tag{3.12}$$

The coefficient matrix of (3.12) is a symmetric and positive-definite M-matrix, since it is diagonally dominant with positive diagonal elements and negative off-diagonal elements (*cf.* [20, p. 85]). Each element of this coefficient matrix depends exponentially on ψ_i for some i. This may cause the entries of the matrix to be computationally unbalanced (i.e., the values may vary by several orders of magnitude across an element). This drawback can be overcome by performing the inverse transformation to (2.7) at the discrete level, i.e., for $i = 1, 2, \ldots, N$, we put

$$w_i = e^{-\psi_i} u_i\,. \tag{3.13}$$

Substituting (3.13) into (3.12) we then have

$$\left(\sum_{k \in I_j} B(\psi_j - \psi_k)|l_{j,k}| + G_j|d_j|\right) u_j - \\ - \sum_{k \in I_j} e^{(\psi_j - \psi_k)} B(\psi_j - \psi_k)|l_{j,k}|\, u_k = \int_{d_j} \hat{F}\, dx\,. \quad (3.14)$$

From the definition (3.5) we have $B(-x) = e^x B(x)$. Therefore (3.12) reduces to

$$\left(\sum_{k \in I_j} B(\psi_j - \psi_k)|l_{j,k}| + G_j|d_j|\right) u_j - \\ - \sum_{k \in I_j} B(\psi_k - \psi_j)|l_{j,k}|\, u_k = \int_{d_j} \hat{F}\, dx\,. \quad (3.15)$$

For $j = 1, 2, \ldots, N'$. Obviously the entries of the coefficient matrix of (3.15) are more balanced than those of (3.12), although is not symmetric unless ψ is constant. However, it is diagonally dominant with respect to its columns. Furthermore, if we use $\mathbf{A}$ to denote the coefficient matrix of (3.15) and $\mathbf{D}$ to denote the diagonal matrix with ith diagonal entry $e^{-\psi_i}$ then we know that $\mathbf{AD}^{-1}$ is a positive definite M-matrix and hence $\mathbf{DA}^{-1} > 0$, (i.e. each element of $\mathbf{DA}^{-1}$ is greater than zero). Therefore, it is easy to show that $\mathbf{A}^{-1} > 0$. Combining this and the fact that $\mathbf{A}$ is a non-singular matrix with non-positive off-diagonal entries, we know that $\mathbf{A}$ is also an M-matrix (*cf.* [20, p. 85]). In practice (3.15) can be solved by a preconditioned conjugate gradient method, for example the CGS method (*cf.* [17]) or the Bi-CGSTAB (*cf.* [18]).

4. CONVERGENCE OF THE APPROXIMATE SOLUTION

In the previous section we showed that the method gives rise to a linear system having a coefficient matrix that is a symmetric and positive-definite M-matrix. This implies the existence and uniqueness of a solution to Problem 3.1. We now show that this approximate solution is stable with respect to a discrete energy norm and that it converges to the exact solution. We use the term error estimate in the sense that we estimate the error $w_h - w_I$ between the numerical solution w_h and the L_h interpolant w_I of w. This estimate is given in the form of an upper bound for $w_h - w_I$ in a discrete energy norm defined below on the discrete space containing w_h and w_I. It is important to note that this norm on L_h is not a norm on the solution space $H_D^1(\Omega)$ containing the exact solution w. We start with the following lemma :

LEMMA 4.1 : *The mass lumping operator defined in (3.7) is surjective from L_h to M_h.*

Proof : The proof is trivial and is omitted here. □

We let $b(.,.)$ be a bilinear form on $L_h \times L_h$ defined by

$$b(w_h, v_h) = a(w_h, P(v_h)) + (P(Ge^{\psi} w_h), P(v_h)). \tag{4.1}$$

We define the following Bubnov-Galerkin problem :

PROBLEM 4.1 : *Find $w_h \in L_h$ such that for all $v_h \in L_h$*

$$b(w_h, v_h) = (\hat{F}, P(v_h)). \tag{4.2}$$

We say that Problem 4.1 is equivalent to Problem 3.1 if any solution w_h of Problem 4.1 is also a solution of Problem 3.1, and *vice versa.*

LEMMA 4.2 : *Problem 4.1 is equivalent to Problem 3.1.*

Proof : The result is obvious since the operator P is surjective from L_h to M_h by Lemma 4.1. □

For any $w_h \in L_h$, we define a functional $\| . \|_h$ by

$$\|w_h\|_h^2 = \sum_{e_{i,j} \in E_h'} \left(\frac{w_j - w_i}{|e_{i,j}|} \right)^2 |b_{i,j}|. \tag{4.3}$$

We then have

LEMMA 4.3 : *The functional $\| . \|_h$ defined in (4.3) is a norm on L_h.*

Proof : The proof is trivial and is omitted here.

On L_h we define the discrete energy norm $\| . \|$ by

$$\|w_h\|^2 = \|w_h\|_h^2 + \sum_{i=1}^{N'} G_i e^{\psi_i} w_i^2 |d_i| \tag{4.4}$$

for each $w_h = \sum_{i=1}^{N'} w_i \phi_i \in L_h$. Let

$$\sigma = \min_{e_{i,j} \in E_h'} \frac{|e_{i,j}|}{\int_{e_{i,j}} e^{-\psi} ds}. \tag{4.5}$$

Since $|\psi|$ is bounded we have $e^{-\psi} > 0$ and thus $\sigma > 0$. It is also easy to verify that $\sigma = \min_{e_{i,j} \in E_h'} e^{\psi_i} B(\psi_i - \psi_j)$ because of Assumption 3.1. The following theorem shows the coercivity of the bilinear form $b(.,.)$ with respect to this norm.

THEOREM 4.4 : *For all* $w_h \in L_h$ *we have*

$$b(w_h, w_h) \geqslant \min \{1, 2\sigma\} \|w_h\|^2 . \tag{4.6}$$

Proof : If $w_h = 0$, then (4.6) holds. Let $w_h \neq 0$. Using the method for the derivation of (3.12) we have

$$\begin{aligned} a(w_h, P(w_h)) &= -\sum_{d \in D_h'} \int_{\partial d \setminus \partial \Omega} e^{\psi} \nabla w_h \cdot \mathbf{n} P(w_h)\, ds \\ &= \sum_{e_{i,j} \in E_h'} (w_j - w_i) \int_{l_{i,j}} e^{\psi} \nabla w_h \cdot \mathbf{e}_{i,j}\, ds \\ &= \sum_{e_{i,j} \in E_h'} (w_j - w_i)\, e^{\psi_i} B(\psi_i - \psi_j) \frac{w_j - w_i}{|e_{i,j}|} |l_{i,j}| \\ &\geqslant 2\sigma \|w_h\|_h^2 . \end{aligned}$$

In the above we used the relation $|l_{i,j}| = \dfrac{2|b_{i,j}|}{|e_{i,j}|}$. From (4.1) and (4.4) we finally have

$$\begin{aligned} b(w_h, w_h) &= a(w_h, P(w_h)) + (P(Ge^{\psi} w_h), P(w_h)) \\ &= 2\sigma \|w_h\|_h^2 + \sum_{i=1}^{N'} G_i\, e^{\psi_i} w_i^2 |d_i| \\ &\geqslant \min \{1, 2\sigma\} \|w_h\|^2 . \end{aligned}$$

□

Theorem 4.4 implies that the solution to Problem 4.1 is stable with respect to the norm $\| . \|$.

LEMMA 4.5 : *For any* $w_h \in L_h$, *there is a constant* $C > 0$, *independent of* h *and* w_h, *such that*

$$\|P(w_h)\|_0 \leqslant C \|w_h\|_h . \tag{4.7}$$

Proof : See [10, Lemma 3.4]. □

For any $\mathbf{p} \in (W^{1,\infty}(\Omega))^2$ we define the functional $| . |_{1,\infty,h}$ by

$$|\mathbf{p}|_{1,\infty,B_h} = \left(\sum_{e_{i,j} \in E_h'} |\mathbf{p}|^2_{1,\infty,b_{i,j}} \right)^{1/2} . \tag{4.8}$$

Obviously $|\,.\,|_{1,\infty,B_h}$ is only a seminorm on $(W^{1,\infty}(\Omega))^2$. The following theorem establishes the convergence of the approximate solution w_h to the L_h-interpolant of w.

THEOREM 4.6 : *Let w_h be the solution of Problem 4.1 and w_I be the L_h-interpolant of the solution of Problem 2.1. Then there is a constant $C > 0$, independent of h, w and ψ, such that*

$$\|w_h - w_I\| \leqslant \frac{C}{\min\{1, 2\sigma\}} (h|\mathbf{f}|_{1,\infty,B_h} +$$

$$+ \|Ge^{\psi} w - P(Ge^{\psi} w)\|_0 + \|F - \hat{F}\|_0). \quad (4.9)$$

Proof : Let C be a generic positive constant, independent of h and w. For any $v_h \in L_h$, multiplying (2.8) by $P(v_h)$ and integrating by parts, we get

$$a(w, P(v_h)) + (Ge^{\psi} w, P(v_h)) = (F, P(v_h)). \quad (4.10)$$

From (4.2) and (4.10) we have

$$a(w_h - w_I, P(v_h)) + (P(Ge^{\psi} w_h) - P(Ge^{\psi} w_I), P(v_h)) =$$

$$= a(w - w_I, P(v_h)) + (Ge^{\psi} w - P(Ge^{\psi} w_I), P(v_h))$$

$$+ (\hat{F} - F, P(v_h)). \quad (4.11)$$

Since $P(Ge^{\psi} w_I) = P(Ge^{\psi} w)$, using the definition of the bilinear form $b(.,.)$ and the Cauchy-Schwarz inequality, we have from (4.11)

$$|b(w_h - w_I, v_h)| \leqslant |a(w - w_I, P(v_h))| +$$

$$+ (\|Ge^{\psi} w - P(Ge^{\psi} w)\|_0 + \|F - \hat{F}\|_0)\|P(v_h)\|_0. \quad (4.12)$$

For the first term on the right side of (4.12) we have

$$a(w - w_I, P(v_h)) = -\sum_{d \in D_h'} \int_{\partial d \setminus \partial\Omega} e^{\psi} \nabla(w - w_I) . \mathbf{n} P(v_h)\, ds$$

$$= \sum_{e_{i,j} \in E_h'} (v_j - v_i) \int_{l_{i,j}} (\mathbf{f} . \mathbf{e}_{i,j} - f_{i,j})\, ds$$

where $f_{i,j}$ is as defined in (3.4). Using the Cauchy-Schwarz inequality and (3.6) we get from the above

$$
\begin{aligned}
|a(w - w_I, P(v_h))| &\leq \sum_{e_{i,j} \in E_h'} |v_j - v_i| \int_{l_{i,j}} |(\mathbf{f} \cdot \mathbf{e}_{i,j} - f_{i,j})| \, ds \\
&\leq \sum_{e_{i,j} \in E_h'} |v_j - v_i| \sup_{x \in b_{i,j}} |(\mathbf{f} \cdot \mathbf{e}_{i,j} - f_{i,j})| \frac{2|b_{i,j}|}{|e_{i,j}|} \\
&\leq Ch \sum_{e_{i,j} \in E_h'} \frac{v_j - v_i}{|e_{i,j}|} |\mathbf{f}|_{1, \infty, b_{i,j}} |b_{i,j}| \\
&\leq Ch \left(\sum_{e_{i,j} \in E_h'} \left(\frac{v_j - v_i}{|e_{i,j}|} \right)^2 |b_{i,j}| \right)^{1/2} \left(\sum_{e_{i,j} \in E_h'} |\mathbf{f}|^2_{1, \infty, b_{i,j}} |b_{i,j}| \right)^{1/2} \\
&= Ch \|v_h\|_h |\mathbf{f}|_{1, \infty, B_h} .
\end{aligned} \tag{4.13}
$$

Substituting (4.13) into (4.12) and using Lemma 4.5 we have

$$
\begin{aligned}
|b(w_h - w_I, v_h)| \leq C\,(h|\mathbf{f}|_{1, \infty, B_h} &+ \\
&+ \|Ge^{\psi} w - P(Ge^{\psi} w)\|_0 + \|F - \hat{F}\|_0) \|v_h\| .
\end{aligned} \tag{4.14}
$$

Letting $v_h = w_h - w_I$ in (4.14) and using (4.6) we obtain

$$
\begin{aligned}
\|w_h - w_I\| \leq \frac{C}{\min\{1, 2\sigma\}} (h|\mathbf{f}|_{1, \infty, B_h} &+ \\
&+ \|Ge^{\psi} w - P(Ge^{\psi} w)\|_0 + \|F - \hat{F}\|_0) .
\end{aligned}
$$

Thus we have proved the theorem. □

We remark that depending on the decoupling technique used for the two continuity equations, we may have $G = 0$ in (2.4) (as the case in Gummel's original work). In such cases the error bound (4.9) depends only on the seminorm of the flux $\mathbf{f}$ and the approximation error of the inhomogeneous term, while error bounds in the energy norm for classical linear finite element methods depend on $\|w\|_2$. It is remarkable that although our error estimate looks similar to that obtained in [8], there are some differences between the two. One is that w_I in (4.9) is the interpolant of w in the subspace spanned by the exponential basis functions, while w_I in [8] is the interpolant in the piecewise linear finite element space. In fact we have proved in [10] that the discrete energy norm $\|\cdot\|$ is equivalent to $\|\cdot\|_1$ on the piecewise linear finite element space. Another difference is that our results are based on arbitrary Delaunay triangulations, while those in [8] are based on triangulations with acute triangles only. Computationally, the Scharfetter-Gummel box method is simpler than other methods because it is essentially a weighted central difference method.

Finally we remark that the approximate flux $\mathbf{f}_h = e^{\psi} \nabla w_h$ does not converge to the exact flux $\mathbf{f} = e^{\psi} \nabla w$. This is because in each element $b_{i,j}$, $\mathbf{f}_h = f_{i,j} \mathbf{e}_{i,j}$ which converges locally only to $\mathbf{f} \cdot \mathbf{e}_{i,j}$. However, by post proccessing it is easy to define an approximate flux which converges to the exact one. For example, we can define

$$\mathbf{f}_h |_{bi,j} = f_{i,j} \mathbf{e}_{i,j} + \frac{\displaystyle\int_{l_{i,j}} \nabla w_h \cdot \mathbf{l}_{i,j} \, ds}{\displaystyle\int_{l_{i,j}} e^{-\psi} \, ds} \mathbf{l}_{i,j}$$

for all $e_{i,j} \in E_h'$, where $\mathbf{l}_{i,j}$ denotes the unit tangential vector along $l_{i,j}$. Moreover the computed ohmic contact currents are convergent, as is shown in the next section.

5. THE EVALUATION OF THE OHMIC CONTACT CURRENTS

We now consider the evaluation of the ohmic contact currents. This discussion is similar to that in [10]. For simplicity, we restrict our attention to a device with a finite number of ohmic contacts, and so $\partial \Omega_D$ is a finite set of separated contacts. For any $c \in \partial \Omega_D$, let $\{x_i^c\}_1^{N_c}$ denote the mesh nodes on c.

Let ξ^c be a piecewise constant function satisfying

$$\xi^c(x) = \begin{cases} 1 & x \in \bigcup_{i=1}^{N_c} d_i^c \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

where d_i^c denotes the element in D_h containing x_i^c. Taking $G = 0$ in (2.4) multiplying by ξ^c and integrating by parts we have

$$-\int_c \mathbf{f} \cdot \mathbf{n} \, ds - \sum_{i=1}^{N_c} \int_{\partial d_i^c \setminus c} \mathbf{f} \cdot \mathbf{n} \, ds = (F, \xi^c).$$

Thus the outflow current through c is

$$J_c \equiv \int_c \mathbf{f} \cdot \mathbf{n} \, ds = -\sum_{i=1}^{N_c} \int_{\partial d_i^c \setminus c} \mathbf{f} \cdot \mathbf{n} \, ds - (F, \xi^c). \tag{5.2}$$

Replacing $\mathbf{f}$ by the approximate flux $\mathbf{f}_h \equiv e^{\psi} \nabla w_h$ and F by the approximation $\hat{F}$, we obtain the following approximate outflow current through c

$$J_{c,h} = -\sum_{i=1}^{N_c} \int_{\partial d_i^c \setminus c} \mathbf{f}_h \cdot \mathbf{n}\, ds - (\hat{F}, \xi^c). \tag{5.3}$$

From (5.3), (5.1), (3.4) and the argument used in the derivation of (3.12), we obtain

$$\begin{aligned} J_{c,h} &= -\sum_{j=1}^{N_c} \left[\int_{\partial d_j^c \setminus c} \mathbf{f}_h \cdot \mathbf{n}\, ds + \int_{d_j^c} \hat{F}\, dx \right] \\ &= \sum_{j=1}^{N_c} \left[\sum_{k \in I_j,\, x_k \notin c} e^{\psi_j} B(\psi_j - \psi_k) \frac{2|b_{j,k}|}{|e_{j,k}|} \frac{w_j - w_k}{|e_{j,k}|} - \int_{d_j^c} \hat{F}\, dx \right] \end{aligned} \tag{5.4}$$

where I_j is the index set of neighbouring nodes of x_j as defined in Section 3.

The convergence and the conservation of the computed ohmic contact currents are established in the following theorem.

THEOREM 5.1 : *Let J_c and $J_{c,h}$ be respectively the exact and the computed outflow currents through $c \in \partial\Omega_D$. Then, there exists a constant $C > 0$, independent of h, ψ and w, such that*

$$|J_c - J_{c,h}| \le C\, (h |\mathbf{f}|_{1,\infty,B_h} + \|F - \hat{F}\|_0). \tag{5.5}$$

Furthermore

$$\sum_{c \in \partial\Omega_D} J_{c,h} = -\int_{\Omega} \hat{F}\, dx. \tag{5.6}$$

Proof: We follow the proof of Theorem 5.1 in [10]. Let C denote a generic positive constant, independent of h. From (5.11-2) we have

$$J_c - J_{c,h} = -\sum_{d \in D_h} \int_{\partial d \setminus c} \xi^c (\mathbf{f} - \mathbf{f}_h) \cdot \mathbf{n}\, ds - (F - \hat{F}, \xi^c). \tag{5.7}$$

Since ξ^c is constant on c we have from (5.7)

$$\begin{aligned} |J_c - J_{c,h}| &\le \left| \sum_{d \in D_h} \int_{\partial d \setminus c} \xi^c (\mathbf{f} - \mathbf{f}_h) \cdot \mathbf{n}\, ds \right| + |(F - \hat{F}, \xi^c)| \\ &\le C h \sum_{e_{i,j} \in E_h'} |\xi_j^c - \xi_i^c|\, |\mathbf{f}|_{1,\infty,b_{i,j}} \frac{2|b_{i,j}|}{|e_{i,j}|} + |(F - \hat{F}, \xi^c)| \end{aligned}$$

$$\leqslant Ch\left[\sum_{e_{i,j}\in E_h'}\left(\frac{\xi_j^c-\xi_i^c}{|e_{i,j}|}\right)^2|b_{i,j}|\right]^{1/2}\left[\sum_{e_{i,j}\in E_h'}|\mathbf{f}|^2_{1,\infty,b_{i,j}}|b_{i,j}|\right]^{1/2}$$
$$+\ \|F-\hat{F}\|_0\|\xi^c\|_0$$
$$\leqslant C\left(h|\mathbf{f}|_{1,\infty,B_h}+\|F-\hat{F}\|_0\right).$$

In the above we used Hölder's inequality and (3.6).

To prove (5.6), we first notice that

$$-\sum_{d\in D_h}\int_{\partial d\setminus\partial\Omega}\mathbf{f}_h\,.\,\mathbf{n}\,ds=0 \tag{5.8}$$

since, for all $e_{i,j}\in E_h'$, $\mathbf{f}_h$ is constant in each subregion $b_{i,j}$. Summing (5.3) over all the contacts we have

$$\sum_{c\in\partial\Omega_D}J_{c,h}=-\sum_{c\in\partial\Omega_D}\left[\sum_{d\in D_h}\int_{\partial d\setminus c}\xi^c\,\mathbf{f}_h\,.\,\mathbf{n}\,ds+(\hat{F},\xi^c)\right]$$
$$=-\sum_{d\in D_h}\int_{\partial d\setminus\partial\Omega}\xi\mathbf{f}_h\,.\,\mathbf{n}\,ds-(\hat{F},\xi) \tag{5.9}$$

where $\xi=\sum_{c\in\partial\Omega_D}\xi^c$. From (5.8) and (5.9) we obtain

$$\sum_{c\in\partial\Omega_D}J_{c,h}=-\sum_{d\in D_h}\int_{\partial d\setminus\partial\Omega}(\xi-1)\,\mathbf{f}_h\,.\,\mathbf{n}\,ds-(\hat{F},\xi-1)-(\hat{F},1)$$
$$=a(w_h,\xi-1)-(\hat{F},\xi-1)-(\hat{F},1)$$
$$=-\int_\Omega\hat{F}\,dx$$

where $a(\,.\,,\,.\,)$ is the bilinear form defined by (3.9). In the above we used (3.8) with $G=0$ since $\xi-1\in M_h$. □

6. CONCLUSION

In this paper we analysed the Scharfetter-Gummel box method for the advection-diffusion equations arising from the nonlinear system of equations governing the stationary behaviour of a semiconductor device. The method was reformulated as a Petrov-Galerkin finite element method with piecewise exponential basis functions on a triangular Delaunay mesh. No restriction is imposed on the angles in the triangulation. The stability of the method was proved and the rate of convergence in a discrete energy norm for the approximate solution in the Slotboom variables was shown to be $O(h)$. In contrast to the standard error estimates, our estimate depends only on the first order seminorm of the flux and the approximation error of the zero order and inhomogeneous terms. It is likely that the flux is physically better behaved than the exact solution, although mathematically this has not been proved. The error constant in the estimate still depends exponentially on the electrostatic potential ψ. The evaluation of the ohmic contact currents associated with this method was discussed and the resulting approximations were shown to be convergent and conservative.

REFERENCES

[1] R. E. BANK, D. J. ROSE, Some Error Estimates for the Box Method, *SIAM J. Numer. Anal.*, **24**, No. 4, 1987, pp. 777-787.

[2] F. BREZZI, P. MARINI, P. PIETRA, Two-dimensional exponentially fitting and applications to semiconductor device equations, *SIAM J. Numer. Anal.*, **26**, 1989, pp. 1342-1355.

[3] E. BUTURLA, P. COTTRELL, B. M. GROSSMAN, K. A. SALSBURG, Finite-Element Analysis of Semiconductor Devices : The FIELDAY Program, *IBM J. Res. Develop.*, **25**, No. 4, 1981, pp. 218-231.

[4] B. DELAUNAY, Sur la sphère vide, *Izv. Akad. Nauk. SSSR, Math. and Nat. Sci. Div.*, No. 6, 1934, pp. 793-800.

[5] G. L. DIRICHLET, Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen, *J. Reine Angew. Math.*, **40**, No. 3, 1850, pp. 209-227.

[6] H. K. GUMMEL, A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculation, *IEEE Trans. Elec. Dev., ED-11*, 1964, pp. 455-465.

[7] R. H. MACNEAL, An Asymmetrical Finite Difference Network, *Quart. Appl. Math.*, **11**, 1953, pp. 295-310.

[8] P. A. MARKOWICH, M. ZLÁMAL, Inverse-Average-Type Finite Element Discretisations of Selfadjoint Second-Order Elliptic Problems, *Math. Comp.*, **51**, No. 184, 1988, pp. 431-449.

[9] B. J. MCCARTIN, Discretization of the Semiconductor Device Equations from *New Problems and New Solutions for Device and Process Modelling*, ed. J. J. H. Miller, Boole Press, Dublin, 1985.

[10] J. J. H. MILLER, S. WANG, A Triangular Mixed Finite Element Method for the Stationary Semiconductor Device Equations, *M^2AN*, **25**, No. 4, 1991, pp. 441-463.

[11] J. J. H. MILLER, S. WANG, A New Non-conforming Petrov-Galerkin Finite Element Method with Triangular Elements for an Advection-Diffusion Problem, IMAJ. Num. Anal., to appear.

[12] M. S. MOCK, Analysis of a Discretization Algorithm for Stationary Continuity Equations in Semiconductor Device Models, *COMPEL,* **2,** No. 4, 1983, pp. 117-139.

[13] M. S. MOCK, Analysis of a Discretization Algorithm for Stationary Continuity Equations in Semiconductor Device Models, II, *COMPEL,* **3,** No. 3, 1984, pp. 137-149.

[14] J. T. ODEN, J. N. REDDY, *An Introduction to the Mathematical Theory of Finite Elements,* John Wiley & Son, New York-London-Sydney-Toronto, 1976.

[15] D. SCHARFETTER, H. K. GUMMEL, Large-signal analysis of a silicon read diode oscillator, *IEEE Trans. Elec. Dev., ED-16*, 1969, pp. 64-77.

[16] J. W. SLOTBOOM, Iterative Scheme for 1- and 2-Dimensional D. C.-Transistor, *IEEE Trans. Elect. Dev.,* **24,** 1977, pp. 1123-1125.

[17] P. SONNEVELD, CGS, A Fast Lanczos-Type Solver for Nonsymmetric Linear Systems, *SIAM J. Sci. Statist. Comput.,* **10,** 1989, pp. 36-52.

[18] H. A. VAN DER VORST, Bi-CGSTAB : A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems, *SIAM J. Sci. Stat. Comput.,* **13,** No. 2, 1992, pp. 631-644.

[19] W. V. VAN ROOSBROECK, Theory of Flow of Electrons and Holes in Germanium and Other Semiconductors, *Bell Syst. Tech. J.,* **29,** 1950, pp. 560-607.

[20] R. S. VARGA, *Matrix Iterative Analysis,* Prentice-Hall, Englewood Cliffs, New Jersey, 1962.