M2AN. MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS - MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

J. J. H. MILLER
S. WANG

A triangular mixed finite element method for the stationary semiconductor device equations

M2AN. Mathematical modelling and numerical analysis - Modélisation mathématique et analyse numérique, tome 25, nº 4 (1991), p. 441-463

<http://www.numdam.org/item?id=M2AN_1991__25_4_441_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol 25, n° 4, 1991, p 441 a 463)

A TRIANGULAR MIXED FINITE ELEMENT METHOD FOR THE STATIONARY SEMICONDUCTOR DEVICE EQUATIONS (*)

J. J. H. MILLER (1) and S. WANG (1)

Communicated by R TEMAM

Abstract — *A Petrov-Galerkin mixed finite element method based on triangular elements for a self-adjoint second order elliptic system arising from a stationary model of a semiconductor device is presented This method is based on a novel formulation of the corresponding discrete problem and can be regarded as a natural extension to two dimensions of the well-known Scharfetter-Gummel one-dimensional scheme Existence, uniqueness and stability of the approximate solution are proved for an arbitrary triangular mesh and an error estimate is given for an arbitrary Delaunay triangulation and its Dirichlet tesselation No restriction is required on the angles of the triangles in the mesh The associated linear system has the same form as that obtained from the conventional box method with an exponentially fitted approximation to the coefficient function on each element The evaluation of the terminal currents associated with the method is also discussed and it is shown that the computed terminal currents are convergent and conservative*

Résumé — *On présente ici une méthode d'éléments finis mixte, de type Petrov-Galerkin, basée sur des éléments triangulaires, pour un système elliptique auto-adjoint du second ordre, émanant d'un modèle stationnaire pour des semiconducteurs Cette méthode est basée sur une nouvelle formulation du problème discret correspondant et peut être considerée comme une extension bidimensionnelle naturelle de la methode bien connue de Scharfetter-Gummel L'existence, l'unicité et la stabilité de la solution approchée sont établies pour un maillage triangulaire arbitraire et une estimation de l'erreur est donnée pour une triangulation de Delaunay arbitraire et sa tesselation de Dirichlet Aucune restriction n'est imposée sur les angles des triangles du maillage Le système associe a la même forme que celle obtenue par la traditionnelle « box-method » avec une approximation du coefficient de type exponentiel sur chaque élément On discute aussi l'evaluation des courants à travers les terminaux associes à cette méthode et on démontre que les courants calculés sont convergents et conservatifs*

(*) Received December 1989, revised February 1990

(1) Numerical Analysis Group, Trinity College, Dublin 2, Ireland

1. INTRODUCTION

The stationary behaviour of semiconductor devices in two dimensions can be described by the following coupled system of nonlinear second-order elliptic partial differential equations [24].

$$\varepsilon \nabla^2 \psi = \eta e^{\psi} - \rho e^{-\psi} - N \tag{1.1}$$

$$\nabla . (\mu_n e^{\psi} \nabla \eta) = R(\psi, \eta, \rho) \tag{1.2}$$

$$\nabla . (\mu_p e^{-\psi} \nabla \rho) = R(\psi, \eta, \rho) \tag{1.3}$$

with appropriate interface and boundary conditions. Using Gummel's method [13] and Newton's method we can decouple and linearise the equations of this nonlinear system so that at each iteration step we have to solve a set of three linear equations of the form

$$- \nabla . (a(x) \nabla u) + G(x) u = F(x) \quad \text{in } \Omega \tag{1.4}$$

with the boundary conditions $u|_{\partial\Omega_D} = \gamma(x)$ and $\nabla u . \mathbf{n} |_{\partial\Omega_N} = 0$, where $\Omega \subset \mathbb{R}^2$, $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ is the boundary of Ω, $\partial\Omega_D \cap \partial\Omega_N = \emptyset$, $\mathbf{n}$ denotes the unit outward normal vector on $\partial\Omega$, $a \in C^0(\bar{\Omega})$, $a_1 \geq a(x) \geq a_0 > 0$, $G \in H^1(\Omega) \cap C^0(\bar{\Omega})$, $G_1 \geq G(x) \geq G_0 \geq 0$ and $F \in L^2(\Omega)$. Here a_0, a_1, G_0 and G_1 are constants.

In what follows we consider only homogeneous Dirichlet boundary conditions $\gamma(x) \equiv 0$. For the inhomogeneous case we can subtract a special function satisfying the boundary conditions and reduce the problem to a homogeneous one. We assume for simplicity that $\partial\Omega$ is polygonal.

To solve (1.4) with the given boundary conditions the box method [17, 8, 19] is often used. Analyses of this method can be found for example in [21, 4, 16 and 14]. More recently Markowich and Zlámal [18], presented a triangular finite element method for the solution of (1.4). Brezzi *et al.* [5, 6, 7] also presented some mixed finite element methods for the solution of (1.4). However, their methods are based on triangulations having acute angles only. In this paper we present a triangular finite element method for (1.4) under milder restrictions on the triangles. This method is based on a novel discrete formulation. The formulation of the method is discussed in the next section. The existence and uniqueness of the discrete solution are proved for an arbitrary triangular mesh in Section 3. In Section 4 we give an error estimate for the approximate solution under mild restrictions on the mesh. Finally, in Section 5 it is shown that the terminal currents computed by the method are convergent and conservative.

In what follows $L^2(\Omega)$ and $W^{m,p}(\Omega)$ denote the usual Sobolev Spaces with norms $\| . \|_0$ and $\| . \|_{W^{m,p}}$ respectively (*cf.* for example [1]). The inner product on $L^2(\Omega)$ and $(L^2(\Omega))^2$ is denoted by $(. , .)$ and the k-th order seminorm on $W^{m,p}(\Omega)$ is denoted by $| \cdot |_{k,p}$. The Sobolev space $W^{m,2}(\Omega)$ is denoted by $H^m(\Omega)$ and the corresponding norm and seminorm is denoted respectively by $\| \cdot \|_m$ and $\| \cdot \|_k$. We put $\mathbf{L}^2(\Omega) = (L^2(\Omega))^2$ and $H_D^1(\Omega) = \left\{ v \in H^1(\Omega) : v |_{\partial\Omega_D} = 0 \right\}$. We use $| \cdot |$ to denote absolute value, Euclidean length or area, depending on the context.

2. THE PETROV-GALERKIN MIXED FINITE ELEMENT FORMULATION

As in Miller *et al.* [20], by the introduction of a new variable $\mathbf{f} = a \nabla u$, we get from (1.4) a first order system of PDEs in the variables $[\mathbf{f}, u]$

$$\nabla u - a^{-1} \mathbf{f} = 0 \tag{2.1}$$
$$-\nabla . \mathbf{f} + Gu = F . \tag{2.2}$$

The corresponding variational problem is

PROBLEM 2.1 : *Find a pair* $[\mathbf{f}, u] \in \mathbf{L}^2(\Omega) \times H_D^1(\Omega)$ *such that for all* $[\mathbf{q}, v] \in \mathbf{L}^2(\Omega) \times H_D^1(\Omega)$

$$(\nabla u, \mathbf{q}) - (a^{-1} \mathbf{f}, \mathbf{q}) = 0 \tag{2.3}$$
$$(\mathbf{f}, \nabla v) + (Gu, v) = (F, v) . \tag{2.4}$$

The existence and uniqueness of the solution to Problem 2.1 have been proved (see, for example, [22]).

To discuss the finite element approximation to Problem 2.1 we first define some meshes on Ω. Let $\mathcal{T}$ denote a family of triangulations of Ω

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

where T_h denotes a triangulation of Ω with each triangle t having diameter h_t 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 $X_h = \{x_i\}_1^{N_V}$ denote the set of all vertices of T_h and $E_h = \{e_i\}_1^{N_E}$ the set of all edges of T_h. We denote by ρ_t the diameter of the incircle of t.

DEFINITION 2.1 : *The family of meshes* $\mathcal{T}$ *is regular if there exists a constant* $\sigma_1 > 0$, *independent of* h, *such that*

$$\max_{t \in T_h} \frac{h_t}{\rho_t} \leqslant \sigma_1 \quad \forall h \in (0, h_0] .$$

We assume henceforth that $\mathcal{T}$ is regular.

DEFINITION 2.2 : *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.* [10]*).*

DDEFINITION 2.3 : *The Dirichlet tessellation D_h corresponding to the triangulation T_h is defined by $D_h = \{D_i\}_1^{N_V}$ where*

$$D_i = \{x : |x - x_i| < |x - x_j|, x_j \in X_h, j \neq i\} \tag{2.5}$$

for all $x_i \in X_h$ (cf. [11]*).*

We now construct two new meshes associated with the triangulation T_h.

For each $x_i \in X_h$ we define the open region $\Omega(x_i)$ consisting of the union of all the triangles $t \in T_h$ with the common vertex x_i and an open region $b(x_i) \subset \Omega(x_i)$ constructed as follows : for each $t \subset \Omega(x_i)$, choose a point $p \in t$ arbitrarily and connect it to the midpoints of the two edges of t sharing x_i, as shown in figure 2.1. (We remark that $p \in t$ is not necessary. However, for simplicity, we assume it does. We also assume that the same $p \in t$ is chosen for each vertex of t.) The domain within the resulting polygon is $b(x_i)$. For the sake of convenience, we sometimes denote $b(x_i)$ simply by b. The set of all such $b(x_i)$ is denoted by B_h^V which we regard as a dual mesh to T_h. We put $\mathscr{B}^V = \{B_h^V : 0 < h \leq h_0\}$.

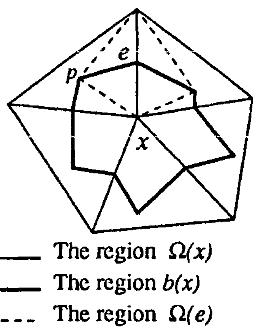

Figure 2.1. — The regions $\Omega(x)$, $b(x)$ and $\Omega(e)$ for the vertex x and edge e.

With each edge $e_i \in E_h$ we also associate an open region $\Omega(e_i)$ by connecting the two end-nodes of e_i with the two chosen points in the two triangles sharing e_i generated during the construction of B_h^V. This is shown in figure 2.1 by dashed lines. The corresponding quadrilateral mesh is denoted by B_h^E. We remark that B_h^E is determined uniquely by B_h^V, and *vice versa*, and that B_h^E divides each $t \in T_h$ into three triangles t_1, t_2, t_3. We put $\mathscr{B}^E = \{B_h^E : 0 < h \leq h_0\}$.

DEFINITION 2.4 : *The family of meshes $\mathscr{B}^E$ is regular if there exists a positive constant σ_2, independent of h, such that for any $i, j \in \{1, 2, 3\}$, $i \neq j$*

$$\max_{t \in T_h} \frac{|t_i| + |t_j|}{|t|} \geq \sigma_2 \quad \forall h \in (0, h_0]. \tag{2.6}$$

Regularity of $\mathscr{B}^E$ is equivalent to the condition that for all $t \in T_h$ and the chosen point $p \in t$, the minimal distance between p and the vertices of t has a positive lower bound. Note also that regularity of $\mathscr{B}^E$ implies that there is a positive constant σ_3 independent of h, such that

$$\min_{x_i \in X_h} \frac{|b(x_i)|}{|\Omega(x_i)|} \geq \sigma_3 \quad \forall h \in (0, h_0]. \tag{2.7}$$

We remark that if T_h is a Delaunay triangulation and the point $p \in t$ is chosen to be the circumcentre of t, for each $t \in T_h$, then the corresponding mesh B_h^V coincides with the Dirichlet tessellation dual to T_h, i.e. $B_h^V = D_h$.

Corresponding to the three meshs B_h^E, T_h and B_h^V, we now construct three finite-dimensional spaces $\mathbf{L}_h \subset \mathbf{L}^2(\Omega)$, $H_h \subset H_D^1(\Omega)$ and $L_h \subset L^2(\Omega)$ as follows. Without loss of generality, we assume that the edges and vertices are numbered so that $\{e_i\}_1^M$ is the set of all egdes in E_h not on $\partial\Omega_D$ and $\{x_i\}_1^N$ is the set of all nodes in X_h not on $\partial\Omega_D$.

For the mesh corresponding to B_h^E we define, for each $i = 1, 2, \ldots, N_E$, a piecewise constant vector-valued function with domain $\bar{\Omega}$ by

$$\mathbf{q}_i(x) = \begin{cases} \mathbf{e}_i & \text{if } x \in \Omega(e_i) \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbf{e}_i$ is the unit tangential vector along the edge e_i. Obviously we have $(\mathbf{q}_i, \mathbf{q}_j) = \delta_{ij} |\Omega(e_j)|$, where δ_{ij} is the Kronecker notation. We take $\mathbf{L}_h = \text{span } \{\mathbf{q}_i\}_1^M$.

Next, letting $\{\phi_i\}_1^{N_V}$ be the conventional piecewise linear basis functions for T_h, we define $H_h = \text{span } \{\phi_i\}_1^N$.

Finally, to construct L_h we define a set of piecewise constant basis functions ψ_i, $(i = 1, 2, \ldots, N)$ corresponding to the mesh B_h^V as follows:

$$\psi_i(x) = \begin{cases} 1 & \text{if } x \in b(x_i) \\ 0 & \text{otherwise} \end{cases}.$$

We then define $L_h = \text{span } \{\psi_i\}_1^N$.

We introduce the mass lumping operator $L : C^0(\bar{\Omega}) \mapsto L^2(\Omega)$ such that for any $u \in C^0(\bar{\Omega})$

$$L(u)(x) = \sum_{x_i \in X_h} u(x_i)\, \psi_i(x) \quad \forall x \in \bar{\Omega}\,. \tag{2.8}$$

Using the three finite dimensional spaces we now define the following discrete Petrov-Galerkin problem.

PROBLEM 2.2: *Find a pair* $[\mathbf{f}_h, u_h] \in \mathbf{L}_h \times H_h$ *such that for all* $[\mathbf{q}_h, v_h] \in \mathbf{L}_h \times L_h$

$$(\nabla u_h, \mathbf{q}_h) - (a^{-1}\mathbf{f}_h, \mathbf{q}_h) = 0 \tag{2.9}$$

$$A(\mathbf{f}_h, v_h) + (L(Gu_h), v_h) = (\hat{F}, v_h) \tag{2.10}$$

where $\hat{F}$ *is a (quadrature) approximation to* F *and* $A(\,.\,,\,.\,)$ *denotes the bilinear form on* $\mathbf{L}_h \times L_h$ *defined by*

$$A(\mathbf{f}_h, v_h) = - \sum_{b \in B_h^V} \int_{\partial b} \mathbf{f}_h \,.\, \mathbf{n}_{\partial b}\, \gamma_0(v_h|_b)\, ds\,. \tag{2.11}$$

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

Let $\mathbf{f}_h = \sum_{i=1}^{M} f_i\, \mathbf{q}_i$, $u_h = \sum_{i=1}^{N} u_i\, \phi_i$, where $\{f_i\}_1^M$ and $\{u_i\}_1^N$ are two sets of constants. Substituting these into (2.9) and taking $\mathbf{q}_h = \mathbf{q}_j$, we get

$$\sum_{i=1}^{M} f_i (a^{-1}\mathbf{q}_i, \mathbf{q}_j) - \sum_{i=1}^{N} u_i (\nabla\phi_i, \mathbf{q}_j) = 0$$

for each $j = 1, 2, \ldots, M$. This linear system of equations has the solution

$$f_j = \frac{1}{a_j^{-1}} \frac{u_{j2} - u_{j1}}{|e_j|} \quad j = 1, 2, \ldots, M$$

where $a_j^{-1} = \dfrac{1}{|\Omega(e_j)|} \displaystyle\int_{\Omega(e_j)} a^{-1}\, dx$ and $u_{jk} = u_h(x_{jk})$, $k = 1, 2$ (see *fig.* 2.2*a*). For the evaluation of a_j^{-1} we refer to the Appendix. Thus we have

$$\mathbf{f}_h = \sum_{i=1}^{M} \frac{1}{a_i^{-1}} \frac{u_{i2} - u_{i1}}{|e_i|} \mathbf{q}_i\,. \tag{2.12}$$

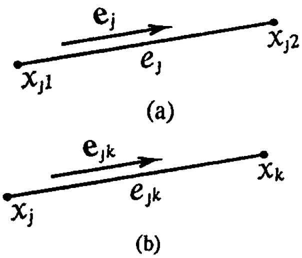

Figure 2.2. — Alternative notation for edges and their end-points.

Substituting (2.12) into (2.10) and letting $v_h = \psi_j$, we obtain

$$-\sum_{i=1}^{M} \frac{1}{a_i^{-1}} \frac{u_{i2} - u_{i1}}{|e_i|} \int_{\partial b(x_j)} \mathbf{q}_i \cdot \mathbf{n}_{\partial b(x_j)}\, ds + (L(Gu_h), \psi_j) = (\hat{F}, \psi_j) \quad (2.13)$$

for each $j = 1, 2, \dots, M$. From its definition (see *fig.* 2.1) we know that $\partial b(x_j)$ consists of a finite number of segments. Clearly $\partial b(x_j) \cap \Omega(e_i)$ consists of at most two segments which we denote by $\partial b_{ij,1}$ and $\partial b_{ij,2}$ if x_j is an end-point of e_i. Therefore, from (2.13), we have for $j = 1, 2, \dots, N$

$$-\sum_{i=1}^{M} \frac{1}{a_i^{-1}} \frac{u_{i2} - u_{i1}}{|e_i|} \sum_{k=1}^{2} \mathbf{e}_i \cdot \mathbf{n}_{ij,k} |\partial b_{ij,k}| + (L(Gu_h), \psi_j) = (\hat{F}, \psi_j) \quad (2.14)$$

where $\mathbf{n}_{ij,k}$ denotes the unit vector normal to $\partial b_{ij,k}$. It is easy to see that $|\mathbf{e}_i \cdot \mathbf{n}_{ij,k}|\,|\partial b_{ij,k}|$ is equal to the length of the projection of $\partial b_{ij,k}$ onto the line perpendicular to e_i. This length is actually the distance from the chosen point $p \in t_k$ to e_i, i.e. the height of the triangle $\Omega(e_i) \cap t_k$ with base e_i, where t_k denotes the triangle having e_i as one edge and containing $\partial b_{ij,k}$. From the formulae for evaluating the area of a triangle we have

$$\sum_{l=1}^{2} |\mathbf{e}_i \cdot \mathbf{n}_{ij,k}|\,|\partial b_{ij,k}| = \frac{2|\Omega(e_i)|}{|e_i|}\,. \quad (2.15)$$

Finally, with the notation in figure 2.2*b*, taking into account the sign of each $\mathbf{e}_i \cdot \mathbf{n}_{ij,k}$ and using (2.14) and (2.15) we find that for all $j = 1, 2, \dots, N$

$$\sum_{k \in I_j} \frac{1}{a_{jk}^{-1}} \frac{2|\Omega(e_{jk})|}{|e_{jk}|} \frac{u_j - u_k}{|e_{jk}|} + G_j u_j |b(x_j)| = (\hat{F}, \psi_j) \quad (2.16)$$

where $I_j = \{k : e_{jk} \in E_h\}$ is the index set of neighbouring nodes of x_j and $G_j = G(x_j)$.

The coefficient matrix of (2.16) is a symmetric, positive definite M-matrix. The latter property follows from the fact that it is a diagonally dominant matrix with all diagonal elements positive and all nonzero off-diagonal elements negative (see, for example, [25, § 3.5]). When T_h is a Delaunay triangulation and B_h^V is the Dirichlet tessellation dual to T_h, the term $2|\Omega(e_{jk})|/|e_{jk}|$ in (2.16) is the length between the circumcentres of the two triangles sharing e_{jk} and $|b(x_j)| = |D_j|$, where D_j is defined in (2.7). Therefore the expression (2.16) coincides with that obtained from the conventional box method with an inverse-average approximation to the coefficient function a.

3. EXISTENCE, UNIQUENESS AND STABILITY OF DISCRETE SOLUTIONS

In this section we prove that Problem 2.2 has a unique solution and that the solution is stable. Instead of proving directly that the mixed Problem 2.2 has a unique solution we consider an equivalent problem for which the existence and uniqueness of the solution can be established by standard finite element analysis.

Let Π_a be the operator from $\nabla H_h = \{\nabla u_h : u_h \in H_h\}$ to $\mathbf{L}_h$ determined by (2.9) with a as a parameter. Introducing the bilinear forms $\bar{A}(u_h, v_h) \equiv A(\Pi_a \nabla u_h, v_h)$ and $B(u_h, v_h) \equiv \bar{A}(u_h, v_h) + (L(Gu_h), v_h)$ on $H_h \times L_h$, we define the following problem :

PROBLEM 3.1 : *Find* $u_h \in H_h$ *such that for all* $v_h \in L_h$

$$B(u_h, v_h) = (\hat{F}, v_h) . \tag{3.1}$$

We say that problem 2.2 is equivalent to Problem 3.1 if the following two conditions hold :

(i) If $[\mathbf{f}_h, u_h]$ is a solution of Problem 2.2, then u_h is a solution of Problem 3.1.

(ii) If u_h is a solution of problem 3.1, then $[\Pi_a \nabla u_h, u_h]$ is a solution of Problem 2.2.

THEOREM 3.1 : *Problem* 2.2 *is equivalent to Problem* 3.1.

Proof : Assume that u_h is a solution of Problem 3.1. From (2.8) and the choice of $\mathbf{L}_h$ we know that there is an $\mathbf{f}_h \in \mathbf{L}_h$ such that for all $\mathbf{q}_h \in \mathbf{L}_h$

$$(\nabla u_h, \mathbf{q}_h) - (a^{-1}\mathbf{f}_h, \mathbf{q}_h) = 0$$

i.e. $\mathbf{f}_h = \Pi_a \nabla u_h$. Making use of the definition of $B(u_h, v_h)$ and the above equality we know that (3.1) reduces to (2.10) and therefore $[\Pi_a \nabla u_h, u_h]$ is a solution of Problem 2.2.

Conversely, if $[\mathbf{f}_h, u_h]$ is a solution of Problem 2.2, elimination of $\mathbf{f}_h$ in Problem 2.2 yields Problem 3.1, as shown in the previous section. □

LEMMA 3.2 : *Assume that that $\mathscr{B}^E$ is regular. Then there exist constants C_1, $C_2 > 0$, independent of h, such that for any $u_h \in H_h$*

$$C_1 \| \nabla u_h \|_0 \leq \| \Pi_1 \nabla u_h \|_0 \leq C_2 \| \nabla u_h \|_0 . \tag{3.2}$$

Proof: From Section 2 we know that for any $u_h \in H_h$, $\Pi_1 \nabla u_h = \sum_{i=1}^{M} (\nabla u_h \cdot \mathbf{e}_i)\big|_{e_i} \mathbf{q}_i$. Thus

$$\begin{aligned} \| \nabla u_h - \Pi_1 \nabla u_h \|_0^2 &= \sum_{i=1}^{M} \int_{\Omega(e_i)} | \nabla u_h - (\nabla u_h \cdot \mathbf{e}_i)\, \mathbf{e}_i |^2 \, dx \\ &= \sum_{i=1}^{M} \int_{\Omega(e_i)} | \nabla u_h \cdot \mathbf{e}_i^{\perp} |^2 \, dx \\ &= \sum_{i=1}^{M} \int_{\Omega(e_i)} | \nabla u_h |^2 \cos^2 \theta_{h,i} \, dx \end{aligned}$$

where $\mathbf{e}_i^{\perp}$ denotes the unit normal vector to e_i such that $\nabla u_h \cdot \mathbf{e}_i^{\perp} \geq 0$ and $\cos \theta_{h,i} = \dfrac{\nabla u_h \cdot \mathbf{e}_1^{\perp}}{|\nabla u_h|}$. Since ∇u_h is constant in each $t \in T_h$, summing over $t \in T_h$ we have

$$\begin{aligned} \| \nabla u_h - \Pi_1 \nabla u_h \|_0^2 &= \sum_{t \in T_h} | \nabla u_h |^2 \left[\cos^2 \theta_{h,1} |t_1| + \cos^2 \theta_{h,2} |t_2| + \cos^2 \theta_{h,3} |t_3| \right] \\ &\leq \alpha \| \nabla u_h \|_0^2 \end{aligned} \tag{3.3}$$

where t_1, t_2 and t_3 are the three triangles which form a partition of t, e_1, e_2 and e_3 are the three edges of t and

$$\alpha = \max_{t \in T_h} \frac{\cos^2 \theta_{h,1} |t_1| + \cos^2 \theta_{h,2} |t_2| + \cos^2 \theta_{h,3} |t_3|}{|t|} .$$

Since $\mathscr{B}^E$ is regular, from (2.6) we know that, for all $\mathbf{t} \in T_h$, the t_1, t_2, t_3 satisfy $|t| = |t_1| + |t_2| + |t_3|$ and

$$\max_i \frac{|t_i|}{|t|} \leq \sigma < 1$$

where σ is a positive constant, independent of h. Furthermore, since for each $i = 1, 2, 3$

$$\cos^2 \theta_{h,i} \leq \gamma$$

where $0 \leq \gamma \leq 1$ is a constant, independent of h, and $\gamma < 1$ for at least one $i \in \{1, 2, 3\}$, it is easy to see that $\alpha < 1$. Thus (3.3) implies (3.2). □

When restricted to H_h, the lumping operator L and its inverse L^{-1} are linear and bounded. This is established in the following lemma.

LEMMA 3.3 : *Assume that $\mathcal{B}^E$ is regular. Then there exist positive constants C_1 and C_2, independent of h, such that for all $u_h \in H_h$*

$$C_1 \| u_h \|_0 \leq \| L(u_h) \|_0 \leq C_2 \| u_h \|_0 . \tag{3.4}$$

Proof: See [15, p. 23] or [4]. □

Define a functional on L_h by $\| v_h \|_{1,L} = \| \nabla L^{-1}(v_h) \|_0$ for all $v_h \in L_h$. We have the following lemma.

LEMMA 3.4 : *Assume that $\mathcal{B}^E$ is regular. Then $\| \cdot \|_{1,L}$ is a norm on L_h and there exists a positive constant C, independent of h, such that $\| v_h \|_0 \leq C \| v_h \|_{1,L}$ for all $v_h \in L_h$.*

Proof: It is easy to see that the triangle inequality is satisfied by $\| \cdot \|_{1,L}$ and $\| \alpha v_h \|_{1,L} = |\alpha| \, \| v_h \|_{1,L}$ for all $v_h \in L_h$ and all $\alpha \in \mathbb{R}$. We now prove that if $\| v_h \|_{1,L} = 0$, then $v_h = 0$. In fact, if $\| v_h \|_{1,L} = 0$ for a $v_h \in L_h$, from its definition we have $\| \nabla L^{-1}(v_h) \|_0 = 0$. Since $L^{-1}(v_h) \in H_h$, we have $\| L^{-1}(v_h) \|_1 = 0$ (generalised Friedrichs' inequality, see, for example, [12, p. 25]) and so $\| L^{-1}(v_h) \|_0 = 0$. Using (3.4) we obtain $\| v_h \|_0 = 0$ and thus $v_h = 0$. Thus $\| \cdot \|_{1,L}$ is a norm.

Furthermore, since $L^{-1}(v_h) \in H_h$ for all $v_h \in L_h$, we have

$$\| v_h \|_{1,L} = \| \nabla L^{-1}(v_h) \|_0 \geq C \| L^{-1}(v_h) \|_1 \geq C \| L^{-1}(v_h) \|_0 \geq C \| v_h \|_0$$

where C denotes a generic positive constant, independent of h. In the above we used the generalised Friedrichs' inequality and Lemma 3.3. □

The existence and uniqueness of the solution to Problem 3.1 is contained in the following theorem :

THEOREM 3.5 : *Assume that $\mathcal{B}^E$ is regular. Then Problem* 3.1 *has a unique solution and the solution is stable.*

Proof: In fact the existence and uniqueness of the solution to Problem 3.1 has already been established since we have shown that the coefficient matrix of (2.16) is a symmetric and positive definite M-matrix. To prove the stability, we need to verify only the following coercivity inequality (see, for example [2])

$$\sup_{v_h \in L_h, \| v_h \|_{1,L} \leq 1} | B(u_h, v_h) | \geq a_0 \, \alpha \| u_h \|_1 \quad \forall u_h \in H_h \tag{3.5}$$

where α is a positive constant, independent of h, u_h, v_h and a, and constant a_0 is the lower bound of a defined in Section 1. In what follows we use C to denote a generic positive constant, independent of h, u_h, v_h and a.

If $u_h = 0$ then clearly (3.5) holds. When $u_h \neq 0$ we put $\bar{v}_h = L(u_h)/\gamma_{u_h}$, where $\gamma_{u_h} = \|\nabla u_h\|_{1,L}$. Then $\|\bar{v}_h\|_{1,L} = 1$ and

$$\begin{aligned}\bar{A}(u_h, \bar{v}_h) &= \bar{A}\left(u_h, \frac{L(u_h)}{\gamma_{u_h}}\right) \\ &= \frac{1}{\gamma_{u_h}} \sum_{b \in B_h^V} \int_{\partial b} \Pi_a \nabla u_h \cdot \mathbf{n}_{\partial b}\, L(u_h)\, ds \\ &= \frac{1}{\gamma_{u_h}} \sum_{b \in B_h^V} u_h(x_b) \int_{\partial b} \Pi_a \nabla u_h \cdot \mathbf{n}_{\partial b}\, ds \end{aligned} \tag{3.8}$$

where x_b is the mesh node contained in b. Summing (3.8) over $\partial b_\iota = \bigcup_{b \in B_h^V} (\partial b \cap \Omega(e_\iota))$ we obtain, using arguments similar to those in the derivation of (2.16)

$$\begin{aligned}\bar{A}(u_h, \bar{v}_h) &= \frac{1}{\gamma_{u_h}} \sum_{\iota=1}^{M} (u_{\iota 2} - u_{\iota 1}) \int_{\partial b_\iota} \Pi_a \nabla u_h \cdot \mathbf{n}'_{\partial b_\iota}\, ds \\ &= \frac{1}{\gamma_u} \sum_{\iota=1}^{M} \frac{1}{a_\iota^{-1}} \frac{(u_{\iota 2} - u_{\iota 1})^2}{|e_\iota|} \frac{2|\Omega(e_\iota)|}{|e_\iota|} \\ &\geqslant \frac{2\,a_0}{\gamma_u} \sum_{\iota=1}^{M} \left(\frac{u_{\iota 2} - u_{\iota 1}}{|e_\iota|}\right)^2 |\Omega(e_\iota)| \\ &= \frac{2\,a_0}{\gamma_{u_h}} \|\Pi_1 \nabla u_h\|_0^2 \\ &\geqslant 2\,a_0\, C \frac{\|\nabla u_h\|_0^2}{\|\nabla u_h\|_0} \\ &\geqslant a_0\, \alpha \|u_h\|_1\end{aligned}$$

where $\mathbf{n}'_{\partial b_\iota}$ denotes the unit vector normal to ∂b_ι chosen so that the angle between $\mathbf{n}'_{\partial b_\iota}$ and $\mathbf{e}_\iota$ is smaller than $\pi/2$. In the above we used (3.2) and the generalised Friedrichs' inequality (see, for example [12, p. 25]). It follows that for all $u_h \in H_h$

$$\sup_{v_h \in L_h, \|v_h\|_0 \leqslant 1} |B(u_h, v_h)| \geqslant \bar{A}(u_h, \bar{v}_h) + (L(Gu_h), \bar{v}_h) \geqslant a_0\, \alpha \|u_h\|_1$$

since $G \geqslant 0$ and

$$(L(Gu_h), \bar{v}_h) = \frac{1}{\gamma_{u_h}} (L(Gu_h), L(u_h)) = \frac{1}{\gamma_{u_h}} \sum_{i=1}^{N} G_i u_i^2 |b_i| \geqslant 0$$

where $u_i = u_h(x_i)$ This completes the proof of the theorem □

From Theorems 3 1 and 3 5 we have

COROLLARY 3 3 *Assume that $\mathscr{B}^E$ is regular Then Problem 2 2 has a unique solution*

We comment that Problem 3 1 can be regarded as a generalised finite element method which is closely related to a mixed finite element method For details, we refer the reader to, for example [3]

4 ERROR ESTIMATE

In this section we give an error estimate for the approximate solution to Problem 2 2 We first state the following lemma

LEMMA 4 1 *Assume that T_h is a Delaunay triangulation and that B_h^V is the corresponding Dirichlet tessellation Then there exist positive constants C_1 and C_2, independent of h, such that for all $w_h \in H_h$ and all $v_h \in L_h$*

$$\|L(w_h) - w_h\|_0 \leqslant C_1 h |w_h|_1 \tag{4 1}$$

and

$$|(Gw_h - L(Gw_h), v_h)| \leqslant$$
$$\leqslant C_2 h \|v_h\|_0 (|Gw_h|_1 + |(Gw_h)_I|_1 + \|G\|_{W^{2,\infty}} \|w_h\|_1) \tag{4 2}$$

where $(Gw_h)_I$ denotes the H_h-interpolant of Gw_h

Proof For the proof of (4 1) we refer to [9] or [15, p 23] We now prove (4 2)

For any $w_h \in H_h$, we have

$$|(Gw_h - L(Gw_h), v_h)| =$$
$$= |(Gw_h - (Gw_h)_I, v_h) + (L((Gw_h)_I) - L(Gw_h), v_h) + ((Gw_h)_I - L((Gw_h)_I), v_h)|$$

for all $v_h \in L_h$ Since $(Gw_h)(x_i) = (Gw_h)_I(x_i)$, $\forall x_i \in X$, using (4 1) we have from the above equality

$$|(Gw_h - L(Gw_h), v_h)| \leqslant$$
$$\leqslant |(Gw_h - (Gw_h)_I, v_h)| + |((Gw_h)_I - L((Gw_h)_I), v_h)|$$
$$\leqslant Ch \|v_h\|_0 |(Gw_h)_I|_1 + |(Gw_h - (Gw_h)_I, v_h)| \tag{4 3}$$

Using Cauchy-Schwarz inequality we have

$$
\begin{aligned}
|(Gw_h - (Gw_h)_I, v_h)| &= \left| \sum_{t \in T_h} \int_t (Gw_h - (Gw_h)_I)\, v_h \, dx \right| \\
&\leqslant \sum_{t \in T_h} \left(\int_t |Gw_h - (Gw_h)_I|^2 \, dx \right)^{1/2} \left(\int_t v_h^2 \, dx \right)^{1/2} \\
&\leqslant Ch \left(\sum_{t \in T_h} |Gw_h - (Gw_h)_I|_{0,t}^2 \right)^{1/2} \|v_h\|_0 \\
&\leqslant Ch \left(\sum_{t \in T_h} (|Gw_h|_{1,t}^2 + |Gw_h|_{2,t}^2) \right)^{1/2} \|v_h\|_0 \\
&\leqslant Ch(|Gw_h|_1 + |Gw_h|_2) \|v_h\|_0 . \qquad (4.4)
\end{aligned}
$$

In the above we used the estimate

$$
|Gw_h - (Gw_h)_I|_{0,t} \leqslant Ch(|Gw_h|_{1,t} + |Gw_h|_{2,t}) \quad \forall t \in T_h .
$$

For any $w_h \in H_h$ we have $|Gw_h|_2 \leqslant \|G\|_{W^{2,\infty}} \|w_h\|_1$. Substituting this estimate into (4.4) and then the result into (4.3) we obtain (4.2). □

Let $\bar{\mathbf{f}}$ and a_A be defined for each $e_i \in E_h$ by

$$
\bar{\mathbf{f}}|_{\Omega(e_i)} = \frac{1}{|\partial b_i|} \int_{\partial b_i} \mathbf{f}\, ds , \quad \partial b_i = \Omega(e_i) \cap \left(\bigcup_{b \in B_h^V} \partial b \right) \qquad (4.5)
$$

$$
a_A|_{\Omega(e_i)} = \left(\frac{1}{|\Omega(e_i)|} \int_{\Omega(e_i)} a^{-1} \, dx \right)^{-1} . \qquad (4.6)
$$

Define the norm $\|\cdot\|_a = (a^{-1}\cdot, \cdot)^{1/2}$. The error estimate for the solution to Problem 3.1 is given in the following theorem :

THEOREM 4.3 : *Let u and u_h be the solutions of Problems* 1.1 *and* 3.1 *respectively. Assume that T_h is a Delaunay triangulation and that B_h^V is the corresponding Dirichlet tessellation. Then there exists a positive constant $C = C(|\Omega|)$, independent of h, such that*

$$
\begin{aligned}
\|u_h - u_I\|_1 \leqslant \frac{C}{a_0} \Big(&\|\bar{\mathbf{f}} - a_A \nabla u_I\|_0 + G_1 \|u - u_I\|_0 + \|\hat{F} - F\|_0 \\
&+ h(|Gu_I|_1 + |(Gu_I)_I|_1 + \|G\|_{W^{2,\infty}} \|u_I\|_1) \Big) \qquad (4.7)
\end{aligned}
$$

$$
\|\mathbf{f} - \mathbf{f}_h\|_a \leqslant 2 \|\mathbf{f} - \mathbf{f}^I\|_a + \|a \nabla (u - u_h)\|_a \qquad (4.8)
$$

where u_I is the H_h-interpolant of u, $\mathbf{f}^I$ is the interpolant of $\mathbf{f}$ in $\mathbf{L}_h$ and a_0 and G_1 are respectively the lower bound of a and the upper bound of G, as defined in Section 1.

Proof: In what follows we let $C = C(|\Omega|)$ denote a generic positive constant, independent of h. Take $v_h \in L_h$ such that $\|v_h\|_{1,L} \leqslant 1$. Multiplying (2.2) by v_h and integrating by parts we get

$$A(\mathbf{f}, v_h) + (Gu, v_h) = (F, v_h). \tag{4.9}$$

Letting $\mathbf{f}_I = \Pi_a(\nabla u_I)$, where Π_a is the operator defined in the previous section, we have from (2.10) and (4.9)

$$\begin{aligned} A(\mathbf{f}_h - \mathbf{f}_I, v_h) + (L(Gu_h) - L(Gu_I), v_h) &= \\ &= A(\mathbf{f} - \mathbf{f}_I, v_h) + (Gu - L(Gu_I), v_h) + (\hat{F} - F, v_h) \\ &= A(\mathbf{f} - \mathbf{f}_I, v_h) + (Gu - Gu_I, v_h) \\ &\quad + (Gu_I - L(Gu_I), v_h) + (\hat{F} - F, v_h). \end{aligned} \tag{4.10}$$

By the definition of $B(.,.)$, from (4.10), using the Cauchy-Schwarz inequality we obtain

$$\begin{aligned} |B(u_h - u_I, v_h)| &\leqslant |A(\mathbf{f} - \mathbf{f}_I, v_h)| + |(Gu - Gu_I, v_h)| \\ &\quad + |(Gu_I - L(Gu_I), v_h)| + |(\hat{F} - F, v_h)| \\ &\leqslant |A(\mathbf{f} - \mathbf{f}_I, v_h)| + |(Gu_I - L(Gu_I), v_h)| \\ &\quad + C(G_1\|u - u_I\|_0 + \|F - \hat{F}\|)_0. \end{aligned} \tag{4.11}$$

In the above we used Lemma 3.4. The first term on the right side of (4.11) can be written in the form

$$A(\mathbf{f} - \mathbf{f}_I, v_h) = -\sum_{b \in B_h^V} \int_{\partial b} (\mathbf{f} - \mathbf{f}_I) \cdot \mathbf{n}_{\partial b}\, v_h\, ds.$$

Summing over ∂b_i we get

$$\begin{aligned} A(\mathbf{f} - \mathbf{f}_I, v_h) &= -\sum_{i=1}^{M} (v_{i2} - v_{i1}) \int_{\partial b_i} (\mathbf{f} - \mathbf{f}_I) \cdot \mathbf{n}'_{\partial b_i}\, ds \\ &= -\sum_{i=1}^{M} (v_{i2} - v_{i1})(\bar{\mathbf{f}} - \mathbf{f}_I)\big|_{\Omega(e_i)} \cdot \mathbf{n}'_{\partial b_i} |\partial b_i| \end{aligned} \tag{4.12}$$

where $\mathbf{n}'_{\partial b_i}$ denotes the unit normal vector on ∂b_i such that $\mathbf{e}_i \cdot \mathbf{n}'_{\partial b_i} > 0$ and v_{ij} ($j = 1, 2$) denote the nodal values of v_h at the two end-points of the edge e_i, as shown in figure 2.2(a). Since T_h is a Delaunay triangulation and B_h^V is the corresponding Dirichlet tessellation we have $\mathbf{n}'_{\partial b_i} = \mathbf{e}_i$. We get from (4.12)

$$A(\mathbf{f} - \mathbf{f}_I, v_h) = - \sum_{i=1}^{M} (v_{i2} - v_{i1})(\bar{\mathbf{f}} - \mathbf{f}_I)\big|_{\Omega(e_i)} \cdot \mathbf{e}_i \frac{2|\Omega(e_i)|}{|e_i|} .$$

Using Hölder's inequality we obtain

$$\begin{aligned} |A(\mathbf{f}_h - \mathbf{f}_I, v_h)| &\leqslant \\ &\leqslant 2 \left(\sum_{i=1}^{M} \left(\frac{v_{i2} - v_{i1}}{|e_i|} \right)^2 |\Omega(e_i)| \right)^{1/2} \times \\ &\quad \times \left(\sum_{i=1}^{M} \left((\bar{\mathbf{f}} - \mathbf{f}_I)\big|_{\Omega(e_i)} \cdot \mathbf{e}_i \right)^2 |\Omega(e_i)| \right)^{1/2} \\ &= 2 \, \| \Pi_1 \nabla L^{-1}(v_h) \|_0 \left\| \Pi_1(\bar{\mathbf{f}} - a_A \nabla u_I) \right\|_0 \\ &\leqslant C \, \|v_h\|_{1,L} \left\| \bar{\mathbf{f}} - a_A \nabla u_I \right\|_0 \\ &\leqslant C \, \left\| \bar{\mathbf{f}} - a_A \nabla u_I \right\|_0 \end{aligned} \tag{4.13}$$

since $\mathbf{f}_I = \Pi_a(\nabla u_I) = \Pi_1(a_A \nabla u_I)$. In the above we have made use of Lemmas 3.2 and 3.3. Substituting (4.13) into (4.11) and using Lemma 4.2 we obtain

$$\begin{aligned} |B(u_h - u_I, v_h)| \leqslant C \Big(&\left\| \bar{\mathbf{f}} - a_A \nabla u_I \right\|_0 + G_1 \|u - u_I\|_0 + \|F - \hat{F}\|_0 + \\ &+ h(|Gu_I|_1 + |(Gu_I)_I|_1 + \|G\|_{W^{2,\infty}} \|u_I\|_1) \Big) . \end{aligned} \tag{4.14}$$

Since (4.14) holds for all $v_h \in L_h$, $\|v_h\|_{1,L} \leqslant 1$, using the inf-sup condition (3.5) we finally obtain

$$\begin{aligned} a_0 \, \alpha \|u_h - u_I\|_1 &\leqslant \sup_{v_h \in L_h, \ \|v_h\|_{1,L} \leqslant 1} |B(u_h - u_I, v_h)| \\ &\leqslant C \Big(\left\| \bar{\mathbf{f}} - a_A \nabla u_I \right\|_0 + G_1 \|u - u_I\|_0 + \|F - \hat{F}\|_0 \\ &\quad + h(|Gu_I|_1 + |(Gu_I)_I|_1 + \|G\|_{W^{2,\infty}} \|u_I\|_1) \Big) . \end{aligned}$$

From the above inequality (4.7) follows immediately.

Since $\mathbf{f}_h - \mathbf{f}^I \in \mathbf{L}_h$, using (2.3) and (2.9) we have

$$(a^{-1}(\mathbf{f}_h - \mathbf{f}^I), \mathbf{f}_h - \mathbf{f}^I) = (a^{-1}(\mathbf{f}^h - \mathbf{f}), \mathbf{f}^h - \mathbf{f}^I) + (a^{-1}(\mathbf{f} - \mathbf{f}^I), \mathbf{f}_h - \mathbf{f}^I)$$
$$(\nabla(u_h - u), \mathbf{f}^h - \mathbf{f}^I) + (a^{-1}(\mathbf{f} - \mathbf{f}^I), \mathbf{f}^h - \mathbf{f}^I)\,.$$

Using Cauchy-Schwarz inequality w obtain from the above equation

$$\|\mathbf{f}_h - \mathbf{f}^I\|_a \leq \|\mathbf{f} - \mathbf{f}^I\|_a + \|a\,\nabla(u - u_h)\|_a\,. \tag{4.15}$$

Finally, (4.8) follows from the triangular inequality and (4.15). We thus have proved the theorem. □

We comment that (4.8) tell us nothing about the convergence of $\mathbf{f}_h$ to $\mathbf{f}$. This is because even $\mathbf{f}$ is constant, the first term on the right side of (4.8) becomes

$$\|\mathbf{f} - \mathbf{f}^I\|_a^2 \geq a_0^{-1} \sum_{i=1}^{M} \|\mathbf{f} - (\mathbf{f}\,.\,\mathbf{e}_i)\,\mathbf{e}_i\|_0^2 = a_0^{-1} \sum_{i=1}^{M} (\mathbf{f}\,.\,\mathbf{e}_i^{\perp})^2 |\Omega(e_i)|$$

where $\mathbf{e}_i^{\perp}$ denotes a unit vector perpendicular to $\mathbf{e}_i$. Thus, clearly, $\|\mathbf{f} - \mathbf{f}^I\|_a$ does not converge to zero as $h \to 0$. This shows that (4.8) does not imply that $\|\mathbf{f} - \mathbf{f}_h\|_a$ converges to zero as $h \to 0$. Nevertheless, as we will see in the next section, the computed terminal currents are convergent.

5. EVALUATION OF TERMINAL CURRENTS

The evaluation of terminal currents is of importance in practice. It is often the final goal of device modelling. We now discuss a method for evaluating terminal currents with the finite element method presented previously. Since in this section we are concerned only with the current continuity equations we have $G \equiv 0$. Furthermore, 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 segments.

For any $c \in \partial\Omega_D$, let $\{x_i^c\}_1^{N_c}$ denote the nodes on c. In what follows it is necessary to make some assumptions about the construction of the meshes T_h and B_h^V. These are contained in the following assumption.

ASSUMPTION 5.1 : *Assume that T_h is a Delaunay triangulation such that for each contact $c \in \partial\Omega_D$, the end-points of c belong to the set of vertices X_h of T_h and B_h^V is the Dirichlet tessellation dual to T_h.*

Let ψ^c be a piecewise constant function satisfying

$$\psi^c(x) = \begin{cases} 1 & x \in \bigcup_{i=1}^{N_c} b(x_i^c) \\ 0 & \text{otherwise}\,. \end{cases} \tag{5.1}$$

Multiplying (2.2) by ψ^c and integrating by parts we have

$$-\int_c \mathbf{f} \cdot \mathbf{n}\, ds - \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c \mathbf{f} \cdot \mathbf{n}\, ds = (F, \psi^c) .$$

Thus the outflow current through c is

$$J_c = \int_c \mathbf{f} \cdot \mathbf{n}\, ds = - \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c \mathbf{f} \cdot \mathbf{n}\, ds - (F, \psi^c) . \tag{5.2}$$

Replacing $\mathbf{f}$ by the finite element solution $\mathbf{f}_h$ and F by the quadrature approximation $\hat{F}$ in (5.2), we obtain the following approximate outflow current through c

$$J_c^h = \int_c \mathbf{f}_h \cdot \mathbf{n}\, ds = - \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c \mathbf{f}_h \cdot \mathbf{n}\, ds - (\hat{F}, \psi^c) . \tag{5.3}$$

From (5.3), (5.1) and the argument used in the derivation of (2.14), we obtain

$$\begin{aligned} J_c^h &= \sum_{j=1}^{N_c} \left[\int_{\partial b(x_j^c) \setminus (\partial b(x_j^c) \cap c)} \mathbf{f}_h \cdot \mathbf{n}\, ds - \int_{b(x_j^c)} \hat{F}\, dx \right] \\ &= \sum_{j=1}^{N_c} \left[\sum_{k \in I_j, x_k \notin c} \frac{2}{a_{jk}^{-1}} \frac{|\Omega(e_{jk})|}{|e_{jk}|} \frac{u_j - u_k}{|e_{jk}|} - \int_{b(x_j^c)} \hat{F}\, dx \right] \end{aligned} \tag{5.4}$$

where I_j is the index set of neighbouring nodes of x_j as defined in Section 2.

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

THEOREM 5.1 : *Let* $[\mathbf{f}, u]$ *and* $[\mathbf{f}_h, u_h]$ *be the solutions of Problems* 2.1 *and* 2.2 *respectively. Let* J_c *and* J_c^h *be respectively the exact and the computed outflow currents through* $c \in \partial\Omega_D$. *Under Assumption* 5.1, *there exists a constant* $\gamma > 0$, *independent of h, such that*

$$|J_c - J_c^h| \leq \gamma \|\psi^c\|_{1, L} \left(\left\| \bar{\mathbf{f}} - a_A \nabla u_h \right\|_0 + \left\| F - \hat{F} \right\|_0 \right) \tag{5.5}$$

where $\bar{\mathbf{f}}$ *and* a_A *are defined in* (4.5) *and* (4.6) *respectively. Furthermore*

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

Proof: The following proof of (5.5) is similar to that of (4.7). Let γ denote a generic positive constant, independent of h. From (5.2-3) we have

$$J_c - J_c^h = - \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c (\mathbf{f} - \mathbf{f}_h) \cdot \mathbf{n} \, ds + (\hat{F} - F, \psi^c) . \qquad (5.7)$$

Since ψ^c is constant on c, summing over ∂b_i, we obtain (5.5) as follows:

$$\begin{aligned}
|J_c - J_c^h| &\leqslant \left| \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c (\mathbf{f} - \mathbf{f}_h) \cdot \mathbf{n} \, ds \right| + \left| (\hat{F} - F, \psi^c) \right| \\
&= \left| \sum_{i=1}^{M} (\psi_{i2}^c - \psi_{i1}^c) \, (\bar{\mathbf{f}} - \mathbf{f}_h) \Big|_{\Omega(e_i)} \cdot \mathbf{e}_i \frac{2 \, |\Omega(e_i)|}{|e_i|} \right| + \left| (\hat{F} - F, \psi^c) \right|
\end{aligned}$$

$$\leqslant 2 \left[\sum_{i=1}^{M} \left(\frac{(\psi_{i2}^c - \psi_{i1}^c)}{|e_i|} \right)^2 |\Omega(e_i)| \right]^{1/2} \times$$

$$\times \left[\sum_{i=1}^{M} \left((\bar{\mathbf{f}} - \mathbf{f}_h) \Big|_{\Omega(e_i)} \cdot \mathbf{e}_i \right)^2 |\Omega(e_i)| \right]^{1/2}$$

$$+ \left\| F - \hat{F} \right\|_0 \| \psi^c \|_0$$

$$= 2 \, \| \Pi_1 (\nabla L^{-1} (\psi^c)) \|_0 \left\| \Pi_1 (\bar{\mathbf{f}} - a_A \nabla u_h) \right\|_0 + \left\| F - \hat{F} \right\|_0 \| \psi^c \|_0$$

$$\leqslant 2 \, \gamma \, \| \psi^c \|_{1, L} \left(\left\| \bar{\mathbf{f}} - a_A \nabla u_h \right\|_0 + \left\| F - \hat{F} \right\|_0 \right) .$$

In the above we used Hölder's inequality and Lemmas 3.2, 3.3 and 3.4.

To prove (5.6), we first notice that

$$A(\mathbf{f}_h, 1) = \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap \partial \Omega_D)} \mathbf{f}_h \cdot \mathbf{n} \, ds = 0 \qquad (5.8)$$

since, for all $e_i \in E_h$, $\mathbf{f}_h$ is constant in each subregion $\Omega(e_i)$. Summing (5.3) over all the contacts we have

$$\begin{aligned}
\sum_{c \in \partial \Omega_D} J_c^h &= - \sum_{c \in \partial \Omega_D} \left[\sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c \, \mathbf{f}_h \cdot \mathbf{n} \, ds - (\hat{F}, \psi^c) \right] \\
&= - \sum_{b \in B_h^V} \int_{\partial b \setminus (\partial b \cap c)} \psi^c \, \mathbf{f}_h \cdot \mathbf{n} \, ds - (\hat{F}, \psi) \\
&= A(\mathbf{f}_h, \psi) - (\hat{F}, \psi) \qquad (5.9)
\end{aligned}$$

where $\psi = \sum_{c \in \partial \Omega_D} \psi^c$ and $A(.,.)$ is the bilinear form defined in Section 2.

From (5.8) and (5.9) we obtain

$$\sum_{c \in \partial\Omega_D} J_c^h = A(\mathbf{f}_h, \psi - 1) - (\hat{F}, \psi - 1) - (\hat{F}, 1)$$
$$= - \int_\Omega \hat{F} \, dx .$$

In the above we used (2.10) with $G = 0$ since $\psi - 1 \in L_h$. □

6. CONCLUSION

A Petrov-Galerkin mixed finite element approach based on a novel formulation was used to approximate the self-adjoint system of second order elliptic PDEs describing a semiconductor device. The existence, uniqueness and stability of the approximate solution were proved for an arbitrary triangular mesh and an error estimate was obtained for an arbitrary Delaunay triangulation and corresponding Dirichlet tesselation. No restrictions need to be imposed on the angles of the triangles in the mesh. The resulting linar system coincides with that obtained from the conventional box scheme with an inverse-average approximation to the coefficient function. In the case of the semiconductor continuity equations this is in fact an exponentially fitted approximation to the coefficient functions. The evaluation of the approximate terminal currents associated with the method was discussed and the computed terminal currents were shown to be convergent and conservative. This method may be applied to the case of non-Delaunay triangulations if we introduce areas and lengths with negative weights. However the coefficient matrices of the resulting linear systems may not be M-matrices. We will discuss this further in a forthcoming paper.

ACKNOWLEDGEMENT

The authors would like to thank one of the referees for several helpful observations.

APPENDIX

Evaluation of the coefficients in the linear systems arising from the discretisation of the continuity equations

In this appendix we discuss in detail the evaluation of a_{jk}^{-1} in (2.16). From the definition we have

$$I(e_{jk}) = a_{jk}^{-1} = \frac{1}{|\Omega(e_{jk})|} \int_{\Omega(e_{jk})} a^{-1}\, dx \tag{A.1}$$

where, by definition, $\Omega(e_{jk})$ is the union of the two triangles obtained from the construction of the mesh B_h^V in Section 2 (*cf. fig.* 2.1). We omit the subscript jk in (A.1) and assume that $\Omega(e) = t_1 \cup t_2$, where $e \in E_h$ is an edge of the mesh and t_1 and t_2 are two triangles. Thus we have

$$\begin{aligned} I(e) &= \frac{1}{|\Omega(e)|} \int_{\Omega(e)} a^{-1}(\mathbf{x})\, dx \\ &= \frac{1}{|\Omega(e)|} \sum_{i=1}^{2} |t_i| \left(\frac{1}{t_i} \int_{t_i} a^{-1}(\mathbf{x})\, dx \right) \end{aligned} \tag{A.2}$$

where $\mathbf{x} = (x_1, x_2)$ and $dx = dx_1\, dx_2$. In the case of the continuity equations (1.2-3) $a(x)$ is equal to $\mu_n e^{\psi(\mathbf{x})}$ and $\mu_p e^{-\psi(\mathbf{x})}$ respectively. Therefore, if we assume that μ_n and μ_p are constant, from (A.2) we see that we need only to evaluate integrals of the form

$$I(t) = \frac{1}{|t|} \int_t e^{\phi(\mathbf{x})}\, dx \tag{A.3}$$

where t is a triangle with vertices $\mathbf{x}_i$ $(i = 1, 2, 3)$ and $\phi(\mathbf{x}) = \pm\, \psi(\mathbf{x})$. Since ψ is the computed solution to the Poisson equation, we assume that ψ is linear on t and $\psi(\mathbf{x}_i) = \psi_i$ $(i = 1, 2, 3)$. Let $\mathbf{s} = \mathbf{s}(\mathbf{x})$ be the linear transformation from t to the reference triangle $\hat{t}$ with vertices (0, 0), (1, 0) and (0, 1) in the (s_1, s_2). Using this transformation (A.3) can be written in the form

$$\begin{aligned} I(t) &= \frac{1}{|\hat{t}|\, \det(J)} \int_{\hat{t}} \det(J)\, e^{\phi(\mathbf{s})}\, ds_1\, ds_2 \\ &= \frac{1}{|\hat{t}|} \int_{\hat{t}} e^{\phi(\mathbf{s})}\, ds_1\, ds_2 \\ &= I(\hat{t}) \end{aligned} \tag{A.4}$$

where $J = \partial(x_1, x_2)/\partial(s_1, s_2)$ is the Jacobian of the transformation which is a matrix with constant entries and det (.) denotes the determinant. Since $\phi(\mathbf{x})$ is linear, $\phi(\mathbf{s})$ is also linear and

$$\phi(\mathbf{s}) = \alpha_0 + \alpha_1 s_1 + \alpha_2 s_2 \tag{A.5}$$

for some constants α_i $(i = 1, 2, 3)$. Substituting this into (A.4) we obtain

$$\begin{aligned} I(\hat{t}) &= \int_0^1 ds_1 \int_0^{1-s_1} e^{\alpha_0 + \alpha_1 s_1 + \alpha_2 s_2} ds_2 \\ &= \frac{e^{\alpha_0}}{\alpha_2} \int_0^1 e^{\alpha_1 s_1}(e^{\alpha_2(1-s_1)} - 1)\, ds_1 \\ &= \frac{e^{\alpha_0}}{\alpha_2}\left(e^{\alpha_2}\frac{e^{\alpha_1-\alpha_2}-1}{\alpha_1-\alpha_2} - \frac{e^{\alpha_1}-1}{\alpha_1}\right) \\ &= \frac{e^{\alpha_0}}{\alpha_2}\left(e^{\alpha_2}B^{-1}(\alpha_1-\alpha_2) - B^{-1}(\alpha_1)\right) \end{aligned} \tag{A.6}$$

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

$$B(x) = \begin{cases} x/(e^x - 1) & x \neq 0 \\ 1 & x = 0 \end{cases}.$$

Since $\psi(\mathbf{x})$ is linear and $\psi(\mathbf{x}_i) = \psi_i$ $(i = 1, 2, 3)$ we have

$$\psi(\mathbf{s}) = \psi_1 + (\psi_2 - \psi_1) s_1 + (\psi_3 - \psi_1) s_2 . \tag{A.7}$$

Furthermore, since $|\hat{t}| = 1/2$, from (A.3)-(A.7) we obtain

$$\frac{1}{|t|}\int_t e^{\psi(\mathbf{x})} dx = \frac{2}{(\psi_3 - \psi_1)}\left(e^{\psi_3} B^{-1}(\psi_2 - \psi_3) - e^{\psi_1} B^{-1}(\psi_2 - \psi_1)\right) \tag{A.8}$$

$$\frac{1}{|t|}\int_t e^{-\psi(\mathbf{x})} dx = \frac{2}{(\psi_1 - \psi_3)}\left(e^{-\psi_3} B^{-1}(\psi_3 - \psi_2) - e^{-\psi_1} B^{-1}(\psi_1 - \psi_2)\right) . \tag{A.9}$$

Substituting (A.8) (or (A.9)) into (A.2) we obtain an expression for the coefficient a_{jk}^{-1} in the linear system arising from the electron (or hole) continuity equation. This expression is a function of the nodal values of ψ. We remark that when the absolute values of the differences between the nodal values ψ_i $(i = 1, 2, 3)$ are small, it is necessary to use Taylor expansions about zero to obtain accurate evaluations of the right-hand sides of (A.8)-(A.9).

REFERENCES

[1] R. A. ADAMS, *Sobolev Spaces,* Academic Press, New York (1975).

[2] I. BABUŠKA, A. K. AZIZ, *Survey Lectures on the Mathematical Foundations of the Finite Element Method* from The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, Academic Press, New York (1972).

[3] I. BABUŠKA, J. E. OSBORN, *Generalized Finite Element Methods : Their Performance and Their Relation to Mixed Methods,* SIAM J. Numer. Anal., *20,* No. 3 (1983) 510-536.

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

[5] F. BREZZI, P. MARINI, P. PIETRA, *Méthodes d'éléments finis et schéma de Scharfetter-Gummel,* C.R. Acad. Sci. Paris, *305,* Série I (1987) 599-604.

[6] F. BREZZI, P. MARINI, P. PIETRA, *Two-Dimensional Exponential Fitting and Applications to Semiconductor Device Equations,* SIAM J. Numer. Anal. *26* (1989) 1342-1355.

[7] F. BREZZI, L. D. MARINI, P. PIETRA, *Numerical Solution of Semiconductor Devices,* Comp. Meth. Appl. Mech. Engin., *75* (1989) 493-514.

[8] 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) 218-231.

[9] P. G. CIARLET, P. A. RAVIART, *General Lagrange and Hermite Interpolation in R^n with Applications to Finite Element Methods,* Arch. Rat. Mech. Anal., *46* (1972) 177-199.

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

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

[12] V. GIRAULT, P. A. RAVIART, *Finite Element Approximation of the Navier-Stokes Equations,* Lect. Notes in Math., No. 749, Springer-Verlag (1979).

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

[14] B. HEINRICH, *Finite difference methods on irregular networks,* Birkhäuser Verlag, Basel-Boston-Stuttgart (1987).

[15] T. IKEDA, *Maximum Principle in Finite Element Models for Convection-Diffusion Phenomena,* North-Holland (1983).

[16] R. H. LI, *Generalized Difference Methods for a Nonlinear Dirichlet Problem,* SIAM J. Numer. Anal. Vol. 24, No. 1 (1987) 77-88.

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

[18] P. A. MARKOWICH, M. ZLÂMAL, *Inverse-Average-Type Finite Element Discretisations of Selfadjoint Second-Order Elliptic Problems,* Math. Comput., *51,* No. 184 (1988) 431-449.

[19] 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).

[20] J. J. H. MILLER, S. WANG, C. H. WU, *A Mixed Finite Element Method for the Stationary Semiconductor Continuity Equations,* Engin. Comput., *5,* No. 4 (1988) 285-288.

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

[22] J. T. ODEN, J. K. LEE, *Theory of Mixed and Hybrid Finite-Element Approximations in Linear Elasticity* from IUTAM/IUM Symp. Applications of Methods of Functional Analysis to Problems of Mechanics, Lect. Notes in Math. No. 503, Springer-Verlag (1976).

[23] J. T. ODEN, J. N. REDDY, *An Introduction to the Mathematical Theory of Finite Elements,* John Wiley & Sons, New York (1976).

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

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