

RESEARCH Open Access

3D finite element modeling and simulation of industrial semiconductor devices including impact ionization

Aurelio Mauri^{1*}, Andrea Bortolossi², Giovanni Novielli³ and Riccardo Sacco⁴

*Correspondence: aureliogianc@micron.com ¹ Process R & D, Micron Technology, via Olivetti 2, Agrate Brianza, MB 20864, Italy Full list of author information is available at the end of the article

Abstract

In this article we address the numerical study of 3D semiconductor devices for applications in electronics industry including charge generation phenomena due to impact ionization. With this aim, we propose two novel 3D finite element (FE) models (methods A and B), for electron and hole Drift-Diffusion (DD) current densities. Method A is based on a primal-mixed formulation of the DD model as a function of the quasi-Fermi potential gradient, while method B is a modification of the standard DD formula based on the introduction of an artificial diffusion matrix. Method A is a Galerkin FE approximation of the density current (written in generalized ohmic form) where the harmonic average of the electrical conductivity is used instead of the standard average while method B is a genuine 3D extension of the classic 1D Scharfetter-Gummel difference formula interpreted as a stabilized Galerkin FE approximation with the use of an 'optimal' artificial diffusion. The proposed methods are compared in the 3D simulation of a p-n junction diode and of a p-MOS transistor in the on-state regime. Results show that method A outperforms method B in physical accuracy and numerical stability. Method A is then used in the 3D simulation of a *n*-MOS transistor in the off-state regime including impact ionization. Results demonstrate that the model is able to accurately compute the I-V characteristic of the device until drain-to-bulk junction breakdown.

PACS Codes: 85.30 **MSC:** 35Jxx; 65N30

Keywords: Drift-Diffusion model; finite element method; numerical simulation;

impact ionization; semiconductor device design

1 Introduction and motivation

Semiconductor technology is undergoing a continuously increasing advancement in the aggressive scaling of device length [1]. In this scenario, three-dimensional (3D) device modeling and numerical simulation techniques play a critical role in the prediction of the electrical performance of the system under investigation. In the case of novel memory devices, due to the different undergoing physical phenomena, a self-consistent multiphysics approach is preferred with respect to the simulation of independent phenomena such as chemical reactions, electrical conduction and material properties modification. To respond to this need the software FEMOS (Finite Element Modeling Oriented Simulator) has been designed: FEMOS is a general-purpose modular numerical code based



on the use of the Galerkin Finite Element Method (GFEM) implemented in a fully 3D framework through shared libraries using an objected-oriented programming language (C++). The first results of this new approach have been recently presented in [2, 3] and [4]. In the present work we extend the FEMOS computational platform in the study of the *Drift Diffusion model* (DD) [5, 6] and focus on the issue of endowing the simulation tool of a consistent, stable and accurate procedure for the approximation of electron and hole current densities in the device. This is of utmost importance in: (i) visualization and post-processing; (ii) evaluation of conduction currents at device terminals; and (iii) inclusion in the DD model of generation phenomena due to Impact Ionization (II). Here, our attention is devoted to (iii), because of the critical role of II in the convergence and numerical stability of the iterative algorithm used to solve the DD system (see [7], Chapter 3 and [8]), although the methods we propose for the treatment of (iii) can also be profitably employed for (i) and (ii).

To allow a consistent treatment of the generation term due to II within the FE procedure, we propose two novel discrete models for electron and hole current densities over the computational grid. The two methods provide a constant approximation of the current density inside each mesh element and for this reason they can be easily implemented in any simulation environment not necessarily employing the GFEM but utilizing, instead, the Box Integration Method (BIM) that is widely employed in contemporary device simulation tools. We refer to [9, 10] for an introduction to the mathematical structure of the BIM; we refer also to [5], Chapter 6, for a detailed discussion of the implementation of the BIM in the context of a finite box geometrical discretization, and to [11] for a general overview of the BIM in the solution of a general transport model including electro-thermal effects.

The first scheme (method A) is based on the use of a primal-mixed formulation of the DD model written as a function of the quasi-Fermi potential gradient (see [12] and [13]). The second scheme (method B) is a modification of the standard DD formula based on the introduction of an artificial diffusion matrix (see also [14]). Method A can be classified as a Galerkin FE approximation of the density current (written in generalized ohmic form) where the harmonic average of the electrical conductivity is used instead of the standard average (cf. [15, 16]), while method B is a genuine 3D extension of the classic 1D Scharfetter-Gummel (SG) difference formula for the computation of the current density over a 1D element [17], having interpreted this latter formula as the result of a stabilized Galerkin FE approximation with the use of an 'optimal' artificial diffusion (cf. [18]). The proposed finite element models are validated in the numerical study of 3D device structures (p-n junction diode and MOS transistors) under on-state and off-state regimes. Results clearly indicate that method A provides the best performance in terms of physical accuracy and numerical stability, and demonstrate the ability of the simulation model in accurately computing the I-V characteristic of the device until the onset of drain-to-bulk iunction breakdown.

The outline of the article is as follows. In Section 2 we review the DD model and the algorithms for its finite element discretization. In Section 3 we illustrate the novel methods proposed to calculate the current densities in the device. In Section 4 we carry out their extensive validation by comparing the obtained results with a reference simulation suite in the study of $3D \, p$ -n and MOS structures including the II generation term. Concluding remarks and perspectives are addressed in Section 5.

2 Physical models, solution map and finite element discretization

In this section we present the mathematical model of the problem, the functional iteration for its decoupled solution and the finite element approximation of the linearized subproblems obtained through decoupling.

2.1 Drift-Diffusion model

The Drift Diffusion model (DD) is described by the Poisson equation coupled with the continuity equations for the mobile carriers, electrons and holes (see [5–7]):

$$\nabla \cdot \mathbf{D} = q(p - n + D),\tag{1a}$$

$$-q\frac{\partial n}{\partial t} + \nabla \cdot \mathbf{J}_n = q(R - G),\tag{1b}$$

$$q\frac{\partial p}{\partial t} + \nabla \cdot \mathbf{J}_p = -q(R - G),\tag{1c}$$

$$\mathbf{D} = \epsilon \mathbf{E},\tag{1d}$$

$$\mathbf{E} = -\nabla \varphi,\tag{1e}$$

$$\mathbf{J}_n = q\mu_n n\mathbf{E} + qD_n \nabla n,\tag{1f}$$

$$\mathbf{J}_{p} = q\mu_{p}\mathbf{E} - qD_{p}\nabla p,\tag{1g}$$

where \mathbf{D} , \mathbf{E} , \mathbf{J}_n and \mathbf{J}_p are the electric displacement, electric field and electron and hole current densities, while ϵ is the dielectric constant, φ the electrostatic potential, p and n the hole and electron concentrations, D the total net doping, q the elementary charge, μ_n and μ_p the electron and hole mobilities, D_n and D_p the electron and hole diffusivities, while R and G are the net recombination and generation rates. Adopted models for μ_n and μ_p account for an appropriate description of the scattering mechanism from lattice and impurities and the degradation due to high electric fields. Adopted models for R account for two-particle (Shockley Read and Hall model) and three-particle net recombination mechanisms (Auger model) while the function G is a pure generation term and accounts for particle rate per unit volume due to Impact Ionization (Van Overstraeten-De Man model) with the projection of electric field along current density direction as a driving force. We refer to [5], Chapter 4 for an exhaustive discussion of recombination/generation mechanisms in semiconductors and associated model parameters used in numerical simulations.

2.2 Geometrical notation and boundary conditions

The DD equation system (1a)-(1g) is solved in a 3D computational domain Ω that represents the geometrical model of the semiconductor device under investigation. In our analysis stationary conditions are always assumed (i.e., $\partial n/\partial t = \partial p/\partial t = 0$), without compromising the validity of the proposed approximation methods.

The device domain Ω is the union of two open disjoint subsets, $\Omega_{\rm Si}$ (doped silicon part), and $\Omega_{\rm ox}$ (oxide part) which it is assumed to be a perfect insulator. The device boundary $\partial\Omega$ is divided into three disjoint subsets, Γ_D , Γ_N and $\Gamma_{\rm int} = \partial\Omega_{\rm Si} \cap \partial\Omega_{\rm ox}$. The subset Γ_D includes the ideal ohmic contacts of the device, i.e., equipotential surfaces where Dirichlet boundary conditions for the dependent variables φ , n and p are applied. On the subset

 Γ_N homogeneous Neumann boundary conditions are applied to prevent any exchange of electric and current flux with the surrounding environment. On the interface subset $\Gamma_{\rm int}$ continuity of electric potential and of the normal component of the electric displacement is assumed, while tangentially flowing electron and hole current fluxes are enforced. This latter condition reflects the fact that the oxide region is assumed to behave as a perfect insulator and prevents from simulating current injection from the semiconductor into the oxide region.

The complete set of boundary conditions is:

$$\varphi = \varphi_D,$$
 $\mathbf{D} \cdot \mathbf{n} = 0,$ $[\![\mathbf{D} \cdot \mathbf{n}]\!] = 0,$
 $n = n_D \text{ on } \Gamma_D,$ $\mathbf{J}_n \cdot \mathbf{n} = 0 \text{ on } \Gamma_N,$ $\mathbf{J}_n \cdot \mathbf{n} = 0 \text{ on } \Gamma_{\text{int}},$ (2)
 $p = p_D,$ $\mathbf{J}_p \cdot \mathbf{n} = 0,$ $\mathbf{J}_p \cdot \mathbf{n} = 0,$

where φ_D , p_D and n_D are boundary data for the dependent variables (computed as described in [5], Chapter 5), \mathbf{n} is the outward unit normal vector over $\partial \Omega = \Gamma_D \cup \Gamma_N$ and over $\Gamma_{\rm int}$, while $[\![\cdot]\!]$ is the jump operator across the interface $\Gamma_{\rm int}$.

2.3 The Gummel solution map

The highly nonlinear coupled nature of the DD system makes its analytical treatment very difficult, if not impossible. For this reason, numerical schemes must be used to compute an approximate solution. To this purpose, we introduce the Maxwell-Boltzmann (MB) statistics for electrons and holes:

$$n = n_i \exp((\varphi - \varphi_n)/V_{\text{th}}), \tag{3a}$$

$$p = n_i \exp((\varphi_p - \varphi)/V_{\text{th}}), \tag{3b}$$

where n_i is the intrinsic concentration in the semiconductor material, φ_n and φ_p are the quasi-Fermi potentials for electrons and holes while $V_{\rm th}$ is the thermal voltage. The Gummel's decoupled algorithm for the iterative solution of the DD system (2.1) consists of solving first the Poisson equation (1a) with respect to the dependent variable φ , having expressed n and p with the Maxwell-Boltzmann statistics (3a) and (3b). This makes (1a) exponentially nonlinear with respect to the electric potential so that to update this latter dependent variable the Newton method is employed [19]. The continuity equations (1b)-(1c) are then successively solved in linear form with respect to the dependent variables nand p inside the adopted decoupled Gummel map algorithm [20]. The lagging procedure introduced and analyzed in [6] is used to treat the recombination/generation terms in such a way that each decoupled continuity equation enjoys a continuous maximum principle. This ensures the well-posedness of each subproblem and the strict positivity of the carrier densities as physically required. Then, once a step of the Gummel iteration is completed, the MB relations (3a) and (3b) are inverted to update the quasi-Fermi potentials using the computed iterates for φ , n and p, until convergence. Linear convergence of the Gummel algorithm is theoretically proved in [6], although practical computational experience often shows superlinear convergence behavior.

2.4 Finite element approximation

The simulation domain Ω is divided in a discrete partition \mathcal{T}_h made by elements K, each element K being a tetrahedron such that $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} \overline{K}$. Then, each differential problem

in the Gummel decoupled algorithm of Section 2.3 is written in weak form (see [21], Chapter 5) and numerically approximated using the displacement-based Galerkin Finite Element Method (GFEM) with piecewise linear conforming elements for potential and carrier densities (see [21], Chapter 6). In the particular case of the linearized continuity equations, to avoid numerical instabilities due to possible dominance of the drift term, the variant of the GFEM denoted Edge Averaged Finite Element (EAFE) method proposed and analyzed in [22–25] is used in the numerical approximation. The EAFE scheme enjoys several good properties (numerical stability, robustness and convergence with respect to the mesh discretization parameter). In particular, in the 2D case if \mathcal{T}_h is a Delaunay triangulation, then the stiffness matrix associated with (1c) and (1b) is a M-matrix [24]. As a consequence, the computed solution (electron and hole density) satisfies the Discrete Maximum Principle (DMP) and is strictly positive over the computational domain. Unfortunately this property is no longer valid in a 3D framework because the Delaunay condition on the mesh is not enough to guarantee the M-matrix property, so that a more restrictive condition, sufficient to ensure the DMP, is presented in [22] (cf. Eq. (2.5), Lemma 2.1). In the numerical experiments shown in Section 4 we have used this latter condition to check the geometrical quality of the triangulation \mathcal{T}_h in the critical case where the generation term due to impact ionization phenomena becomes the dominating driving force in the simulation of charge transport in the device.

3 Finite element models for the current density

The construction of a stable and accurate approximate current density field in a primalbased FE formulation is not a trivial task because of possible numerical problems arising from differentiation and cancellation in the DD transport relations (1f)-(1g). Throughout the remainder of this section we assume that the solution of the DD system (1a)-(1g) is carried out using the Gummel decoupled algorithm described in Section 2.3. We first introduce a simple finite element formula for the representation of the current density over each mesh element which is a straightforward implementation of (1f)-(1g). Then, we propose two novel finite element methodologies for current density discretization. The novel approaches have a much simpler implementation than other more sophisticated formulations (such as the dual-mixed hybridized FEM with exponential fitting, see [16, 26]) and are completely compatible for use in the classical Box Integration Method (BIM) [9, 10]. Of the three examined methods, one reveals to be particularly robust and accurate, as demonstrated by several tests discussed in Section 4. In the remainder of the article, for a given element K in \mathcal{T}_h we denote the volume of K by vol(K); moreover, with the subscript K we refer to a quantity defined in the interior of K while the subscript h refers to a quantity defined at the vertices of K. For a given function $f: K \to \mathbb{R}$ we define $\langle f \rangle_K := \int_K f \, dK / \operatorname{vol}(K)$ the mean integral value of f over K. We also assume that carrier mobility (and the associated diffusivity through Einstein's relation) is constant in K, while the electric potential is linear (so that the associated electric field is a constant vector in *K*).

3.1 The DD formula

The simplest FE model for the current densities (method DDFE) is obtained by substituting into the transport equations (1f)-(1g) the functions n, p and φ with their corresponding FE approximations n_h , p_h and φ_h , and then by computing the integral average of the re-

sulting expressions over the element K. This yields

$$\mathbf{J}_{n,K} = q\mu_n \langle n \rangle_K \mathbf{E}_K + q D_n \nabla n_{h,K}, \tag{4a}$$

$$\mathbf{J}_{p,K} = q\mu_p \langle p \rangle_K \mathbf{E}_K - qD_p \nabla p_{h,K}. \tag{4b}$$

It is immediate to see that the discrete current densities (4a) and (4b) automatically reproduce the limiting cases of pure diffusive flow ($\mathbf{E}_K = \mathbf{0}$) and pure ohmic flow ($n_{h,K} = \text{constant}$ and $p_{h,K} = \text{constant}$). In the case of thermal equilibrium ($\mathbf{J}_{n,K} = \mathbf{J}_{p,K} = \mathbf{0}$), we can anticipate computational difficulties with the use of (4a) and (4b) because of exact cancellation of drift and diffusive current contributions. Thus, by a continuity argument, we also see that method DDFE does not seem appropriate in the numerical treatment of the subthreshold current regime, where currents are not exacty equal to zero but are very small. The two following formulations are designed to overcome this limitation.

3.2 Method A

To describe method A we consider the case of electron continuity equation because a completely similar treatment holds for the hole continuity equation. Using (3a) into (1f) yields the equivalent ohmic form of the electron current density

$$\mathbf{J}_{n}=\sigma_{n}\mathbf{E}_{n},\tag{5a}$$

where:

$$\sigma_n = q\mu_n n = q\mu_n n_i \exp((\varphi - \varphi_n)/V_{\text{th}}), \tag{5b}$$

$$\mathbf{E}_{n} = -\nabla \varphi_{n} \tag{5c}$$

are the electrical conductivity and the effective electric field experienced by electrons, respectively. We notice that the use of the ohmic representation yields a nonlinear gradient form of the DD current density. To construct the finite element model for J_n as in (5a), we use the primal-mixed (PM) FEM introduced and analyzed in [12] and recently extended to the case of advective-diffusive operators in [13]. In the PM FEM of lowest order, the approximate current density is constant over each $K \in \mathcal{T}_h$ while the approximate quasi-Fermi potential is piecewise linear and continuous over \mathcal{T}_h . Let us introduce the finite element spaces of piecewise constant and piecewise linear continuous functions over \mathcal{T}_h :

$$Q_h = \left\{ w \in L^2(\Omega) : w|_K \in \mathbb{P}_0(K) \ \forall K \in \mathcal{T}_h \right\},\tag{5d}$$

$$V_h = \left\{ u \in C^0(\overline{\Omega}) : u|_K \in \mathbb{P}_1(K) \ \forall K \in \mathcal{T}_h \right\} \tag{5e}$$

and the approximate electrical conductivity $\sigma_{n,h} := q\mu_n n_i \exp((\varphi_h - \varphi_{n,h})/V_{\text{th}})$ where $\varphi_{n,h} \in V_h$ and $\varphi_h \in V_h$ are the finite element discrete analogues of φ_n and φ , respectively. Then, the PM-FE approximation of (5a) reads: find $J_{n,h} \in [Q_h]^3$ such that

$$\int_{\Omega} \sigma_{n,h}^{-1} \mathbf{J}_{n,h} \cdot \mathbf{q}_h \, d\Omega + \int_{\Omega} \nabla \varphi_{n,h} \cdot \mathbf{q}_h \, d\Omega = 0 \quad \forall \mathbf{q}_h \in [Q_h]^3, \tag{5f}$$

where $J_{n,h} \in [Q_h]^3$ is the finite element discrete analogue of J_n . Since functions in $[Q_h]^3$ are discontinuous, relation (5f) amounts to

$$\int_{K} \sigma_{n,h}^{-1} \mathbf{J}_{n,h} \cdot \mathbf{q}_{h} dK + \int_{K} \nabla \varphi_{n,h} \cdot \mathbf{q}_{h} dK = 0 \quad \forall \mathbf{q}_{h} \in \left[\mathbb{P}_{0}(K) \right]^{3}.$$
 (5g)

Using in (5g) the standard basis functions for $[\mathbb{P}_0(K)]^3$ we obtain

$$[\mathbf{J}_{\mathbf{n},\mathbf{h}}]_i = -\mathcal{H}_K(\sigma_{n,h}) \frac{\partial \varphi_{n,h}}{\partial x_i}, \quad i = 1, 2, 3, \forall K \in \mathcal{T}_h,$$
 (5h)

where $\mathcal{H}_K(\sigma_{n,h}) := (\langle \sigma_{n,h}^{-1} \rangle_K)^{-1}$ is the harmonic average of $\sigma_{n,h}$ over the element K.

We notice that the approximation of the current density with a constant vector may appear inadequate because of the exponential dependence of the electron conductivity on the quasi-Fermi potential (cf. (5b)), except in the case of a constant electron concentration. It can be proved (see [12]) that $\|\mathbf{J}_n - \mathbf{J}_{n,h}\|_{L^2(\Omega)} = \mathcal{O}(h)$, where h denotes the mesh parameter, so that an appropriate reduction of the grid size may be required to achieve a desired fixed tolerance.

To numerically compute in a simple and accurate manner the harmonic average of the electrical conductivity, we use the following quadrature rule (see also [27, 28])

$$\left(\frac{\int_K \sigma_{n,h}^{-1} dK}{\operatorname{vol}(K)}\right)^{-1} \simeq \left(\frac{\int_{e*} \sigma_{n,h}^{-1} de}{|e^*|}\right)^{-1} \tag{5i}$$

 e^* being the edge of ∂K where the maximum drop of $\sigma_{n,h}$ occurs and $|e^*|$ its Euclidean length. Upon introducing the linear dimensionless potential $\Phi_n := (\varphi_h - \varphi_{n,h})/V_{\text{th}}$ and the two vertices of K: \mathbf{x}_m s.t. $\Phi_n(\mathbf{x}_m) = \Phi_{n,m} := \min_K(\Phi_n)$ and \mathbf{x}_M s.t. $\Phi_n(\mathbf{x}_M) = \Phi_{n,M} := \max_K(\Phi_n)$, we define e^* to be the edge of ∂K which connects \mathbf{x}_m and \mathbf{x}_M , so that (5i) yields

$$\int_{K} \sigma_{n,h}^{-1} dK \simeq q \mu_n n_i \exp(\Phi_{n,m}) \mathcal{B}(\Phi_{n,m} - \Phi_{n,M}), \tag{5j}$$

where $\mathcal{B}(Z) := Z/(\exp(Z) - 1)$ is the inverse of the Bernoulli function, such that $\mathcal{B}(0) = 1$. Using MB statistics (3a), relation (5j) can be written as

$$\int_{K} \sigma_{n,h}^{-1} dK \simeq q \mu_n n_m \mathcal{B}(\Phi_{n,m} - \Phi_{n,M}), \tag{5k}$$

where $n_m := n_i e^{\Phi_{n,m}}$. Expression (5j) can be also written as

$$\int_{K} \sigma_{n,h}^{-1} dK \simeq q \mu_{n} n_{i} \exp(\Phi_{n,m}) \exp(\Phi_{n,M}) \exp(-\Phi_{n,M}) \mathcal{B}(\Phi_{n,m} - \Phi_{n,M})$$

which, using again (3a) and the property $\exp(Z)\mathcal{B}(Z) = \mathcal{B}(-Z)$, becomes

$$\int_{K} \sigma_{n,h}^{-1} dK \simeq q \mu_n n_M \mathcal{B}(\Phi_{n,M} - \Phi_{n,m}), \tag{51}$$

where $n_M := n_i e^{\Phi_{n,M}}$. Combining (5k) and (5l), we find

$$\mathbf{J}_{n,K} = -q\mu_n \left[\frac{n_m \mathcal{B}(-\Delta \Phi_{n,\max}) + n_M \mathcal{B}(\Delta \Phi_{n,\max})}{2} \right] \nabla \varphi_{n,h}, \tag{5m}$$

where $\Delta\Phi_{n,\max} := \Phi_{n,M} - \Phi_{n,m}$. In the case where $\Delta\Phi_{n,\max}$ is sufficiently large, we see that the 1D harmonic average of n_h along the edge e^* has the effect of 'favoring' the smallest nodal concentration in the element K rather than the largest one, so that we approximately have

$$\mathbf{J}_{n,K} \simeq -q\mu_n n_m \max_{\partial K} \frac{(\varphi_h - \varphi_{n,h})}{2V_{\text{th}}} \nabla \varphi_{n,h}. \tag{5n}$$

If, instead, $\Delta\Phi_{n,\max}$ is small then the 1D harmonic average behaves appreciably as the standard average and we obtain

$$\mathbf{J}_{n,K} \simeq -q\mu_n \frac{n_m + n_M}{2} \nabla \varphi_{n,h}. \tag{50}$$

Following the same procedure as above also for the hole current density we have

$$\mathbf{J}_{p,K} = -q\mu_p \left[\frac{p_m \mathcal{B}(-\Delta \Phi_{p,\text{max}}) + p_M \mathcal{B}(\Delta \Phi_{p,\text{max}})}{2} \right] \nabla \varphi_{p,h}, \tag{5p}$$

where $\Phi_p := (\varphi_{p,h} - \varphi_h)/V_{\text{th}}$ and $\Delta \Phi_{p,\text{max}} := \max_K \Phi_p - \min_K \Phi_p$.

3.3 Method B

In the previous section the discrete form of the current density is constructed by starting from the equivalent 'ohmic' representation in terms of the quasi Fermi potential, and then by performing a suitable approximation of the electrical conductivity over the finite element K. Here, we continue along the same line of thought, but starting from the classic DD format (1f)-(1g), with the intent of using the method of Streamline Upwind artificial diffusion proposed in [18] for the advective-diffusive model to stabilize the computation in the presence of a high electric field.

3.3.1 The 1D SG method as an artificial diffusion scheme

In the 1D setting the artificial diffusion technique consists of replacing the electron diffusion coefficient D_n with the modified quantity

$$D_{n,h} = D_n + D_n \Phi(\mathbb{P}e|_K), \tag{6a}$$

where Φ is a suitable nonnegative stabilization function of the 1D local Pèclet number $\mathbb{P}e|_K = (h|\partial_x\varphi_h|)/(2V_{\text{th}}) = |\Delta\varphi|/(2V_{\text{th}})$, h and $\Delta\varphi$ being the length of the 1D interval and the potential drop over the interval, respectively. The local Pèclet number gives a measure of how much the drift term dominates over the diffusion term in the transport mechanism. If $\mathbb{P}e|_K > 1$ the problem is locally drift (advection)-dominated and in such a case we need introduce an extra amount of diffusion in (6a) (given by $D_n\Phi(\mathbb{P}e|_K)$) to prevent the occurrence of unphysical spurious oscillations in the computed solution, which might even lead to a negative electron concentration. If $\mathbb{P}e|_K < 1$ the problem is locally diffusion-dominated and there is no need of adding an extra diffusion, so that the standard GFEM is enough for obtaining an accurate and numerically stable solution. Based on the last observation, the function Φ has to satisfy the property of consistency

$$\lim_{\mathbb{P}e|_K\to 0} \Phi(\mathbb{P}e|_K) = 0 \quad \forall K \in \mathcal{T}_h.$$
 (6b)

The 1D approximation of the electron current density to be used in a stabilized GFEM is thus given by the following relation

$$J_{n,h}(n_h)|_K = q\mu_n \langle n \rangle_K E_{h,K} + qD_n (1 + \Phi(\mathbb{P}e|_K)) \partial_x n_h \quad \forall K \in \mathcal{T}_h.$$
 (6c)

To design in a physically sound and consistent manner the optimal stabilization function Φ , we assume the modified method to exactly satisfy some limiting cases that often occur in practical important electronic applications. Using a 3D notation, for sake of generality and because this will be used in later extension, the considered cases are:

- C1. Constant carrier concentrations (only drift contribution): $\mathbf{J}_n = q\mu_n n\mathbf{E}$.
- C2. Constant potential (**E** = 0, only diffusive contribution): $\mathbf{J}_n = qD_n \nabla n$.
- C3. Constant quasi Fermi potential (no current flow): $\mathbf{J}_n = -q\mu_n n \nabla \varphi_n = \mathbf{0}$. Notice that case C3 implies that

$$n = Ce^{\varphi/V_{\rm th}},\tag{6d}$$

where C is an arbitrary constant such that $C = \exp(-\overline{\varphi}_n/V_{th})$, $\overline{\varphi}_n$ being a given constant value. Thanks to assumption (6b) the stabilized current (6c) automatically satisfies cases C1 and C2. Case C3 is recovered by imposing (in the 1D setting)

$$J_{n,h}(\Pi_1^K(Ce^{\varphi/V_{\text{th}}})) = 0, \tag{6e}$$

where Π_1^K is the \mathbb{P}_1 -interpolant over the element K. Using (6e) in (6c), noting that $E_K = -\partial_x \varphi_h$ and using Einstein's relation, we obtain the following relation for the stabilization function

$$\Phi(\mathbb{P}e|_K) = \frac{\langle n \rangle_K}{V_{\text{th}}} \frac{\partial_x \varphi_h}{\partial_x \Pi_1^K (e^{\varphi/V_{\text{th}}})} - 1. \tag{6f}$$

Enforcing relation (6d) at the two vertices x_i of the interval, i = 1, 2, yields

$$\Phi(\mathbb{P}e|_K) = \sigma \mathbb{P}e|_K \frac{e^{\varphi_1/V_{\text{th}}} + e^{\varphi_2/V_{\text{th}}}}{e^{\varphi_2/V_{\text{th}}} - e^{\varphi_1/V_{\text{th}}}} - 1 = \sigma \mathbb{P}e|_K \frac{e^{2\sigma \mathbb{P}e|_K} + 1}{e^{2\sigma \mathbb{P}e|_K} - 1} - 1, \tag{6g}$$

where n_1 and n_2 are the two nodal values of n_h while $\sigma := \text{sign}(\Delta \varphi)$. Setting for brevity $X := 2\sigma \mathbb{P}e|_K$, relation (6g) becomes

$$\Phi(X) = \frac{X}{2} \left(\frac{e^X}{e^X - 1} + \frac{1}{e^X - 1} \right) - 1 = \frac{1}{2} \left(\mathcal{B}(-X) + \mathcal{B}(X) \right) - 1$$
$$= \frac{1}{2} \left(X + \mathcal{B}(X) + \mathcal{B}(X) \right) - 1 = \mathcal{B}(X) + \frac{X}{2} - 1,$$

and replacing the definition of *X* we obtain for both $\Delta \varphi > 0$ and $\Delta \varphi < 0$

$$\Phi(\mathbb{P}e|_K) = \mathcal{B}(2\mathbb{P}e|_K) + \mathbb{P}e|_K - 1 \tag{6h}$$

which, upon substitution into (6c), recovers the well known 1D Scharfetter-Gummel scheme [17]. This latter method was originally proposed as an 'exponential fitting' difference scheme by Allen and Southwell in [29] and subsequently referred to as an artificial diffusion stabilized scheme with 'optimal viscosity' by Brooks and Hughes in [18].

3.3.2 The 3D SG Artificial Diffusion method

In a 3D framework a straightforward extension of the 1D Scharfetter-Gummel stabilization (6g)-(6h) can be obtained by introducing a 3×3 diagonal stabilizing tensor $\underline{\underline{\Phi}}^K$ defined on each element $K \in \mathcal{T}_h$ as follows

$$\underline{\underline{\Phi}}_{ii}^{K} = \frac{\langle \Pi_{1}^{K} (e^{(\varphi_{h} - \varphi_{M})/V_{\text{th}}}) \rangle_{K} \partial_{x_{i}} \varphi}{\partial_{x_{i}} \Pi_{1}^{K} (e^{(\varphi_{h} - \varphi_{M})/V_{\text{th}}}) V_{\text{th}}} - 1, \quad i = 1, 2, 3,$$
(6i)

where we have used, as a reference value for the potential reference to avoid overflow exceptions in the machine evaluation of (6i), the maximum φ_M of φ_h in K.

The 3D approximate electron current density is then

$$\mathbf{J}_{n,K} = q\mu_n \langle n \rangle_K \mathbf{E}_K + q D_n (\underline{\mathcal{I}} + \underline{\mathbf{\Phi}}_K) \nabla n_h, \tag{6j}$$

where $\underline{\underline{\mathcal{I}}}$ is the 3 × 3 identity tensor. In a completely similar manner we have the relation for the hole current density

$$\mathbf{J}_{p,K} = q\mu_p \langle p \rangle_K \mathbf{E}_K - qD_p(\underline{\underline{\mathcal{I}}} + \underline{\underline{\Phi}}_K) \nabla p_h, \tag{6k}$$

where

$$\underline{\underline{\Phi}}_{ii}^{K} = \frac{\langle \Pi_{1}^{K} (e^{(\varphi_{m} - \varphi_{h})/V_{\text{th}}}) \rangle_{K} \partial_{x_{i}} \varphi}{\partial_{x_{i}} \Pi_{1}^{K} (e^{(\varphi_{m} - \varphi_{h})/V_{\text{th}}}) V_{\text{th}}} - 1, \quad i = 1, 2, 3,$$
(61)

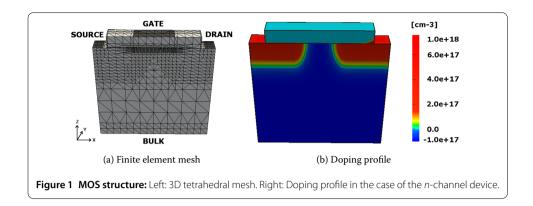
 φ_m being the minimum of φ over K. It is immediate to check that the two proposed novel approximations (6j) and (6k) satisfy all cases C1, C2 and C3 in Section 3.3.1.

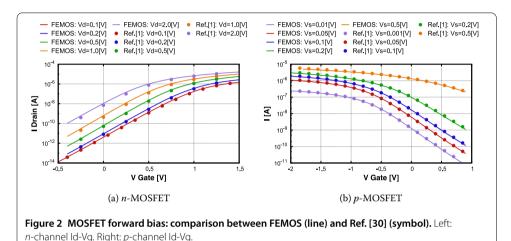
4 Numerical experiments

This section is devoted to the illustration of several numerical experiments. Specifically, Section 4.1 is aimed to check the correctness of the implementation of the 3D-EAFE method within the FEMOS code and to compare the performance of this finite element approach with a reference-commercial tool based on the BIM [30]. Section 4.2 is aimed to compare the performance of the various methods for current calculation of Section 3 against physical expectation. The conclusion of these tests is that method A turns out to be the most stable and accurate, so that Section 4.3 employs such method in a self-consistent simulation of semiconductor devices in the presence of II generation terms. 3D ideal and realistic devices commonly employed in semiconductor industry are taken into consideration: (1) a p-n ideal diode; (2) a p and n-MOS ideal devices; and (3) a p-MOS device resulting from a realistic 2D-process simulation accounting for non ideal doping profiles. Device (3) represents a severe benchmark to highlight the accuracy and the stability of the different current density calculation methods.

4.1 Accuracy of the 3D-EAFE discretization

Several tests have been performed in order to assess the correctness of the EAFE approach in the simulation of semiconductor devices in 3D framework. For reading benefits we just report a benchmark with a reference commercial tool based on n-p-channel MOSFET structure of Figure 1(a) (GateLength: 100 nm; GateOx: 30 nm and contacts S = Source,

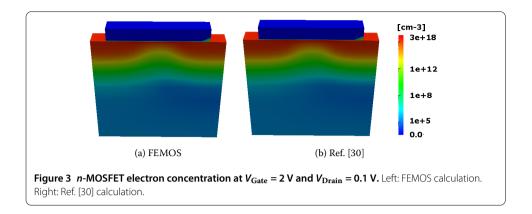


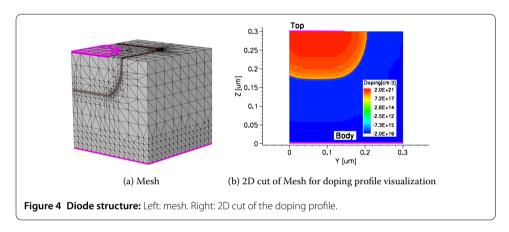


Gate, D = Drain, Body) with analytical doping profiles: Figure 1(b) reports the doping profile for the n-channel device (the p-channel device is obtained just by the inverting the doping type). The agreement with the reference simulation tool is very good as reported in Figure 2(a) for the n-channel device and in Figure 2(b) for the p-channel device for several Drain bias conditions (electron and hole mobilities have been calculated using the Canali model [31]). We point out that the calculation of the current at the contacts is carried out by extending to the 3D case the approach proposed in [32] for the 2D case: this novel formulation, known as residual method, is based on the approach proposed in [33]. We also notice that the 3D-profiles of carriers are indistinguishable between FEMOS and Ref. [30] as reported in Figure 3 in the case of electrons in the n-channel device at $V_{Gate} = 2$ V and $V_{Drain} = 0.1$ V.

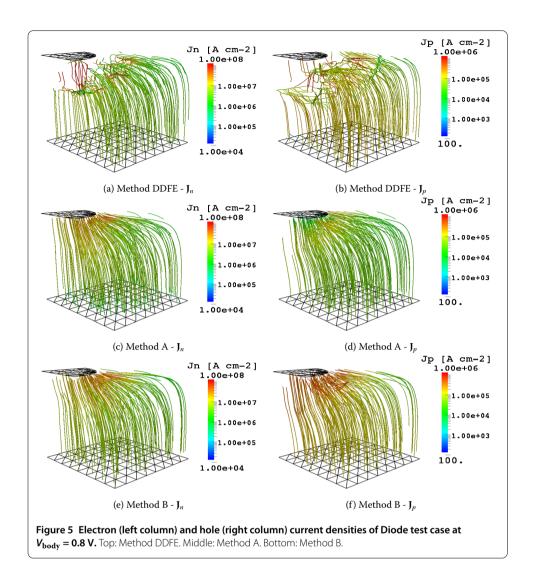
4.2 Evaluation of the performance of current calculation methods

To evaluate the simulation performance of the methods for current calculation introduced in Section 3 we have firstly considered a diode structure with a p-n junction geometrically represented by the cubic region $\Omega = (0,0.3)^3$ µm³ (see Figure 4(a)). A Gaussian implantation of donors with a peak of 10^{19} cm⁻³ and a depth of 0.15 µm (see Figure 4(b)) is made over a p-type region with a constant acceptor profile of 10^{18} cm⁻³ magnitude. One contact is defined for each of the doped regions: for the n-type part a rounded-shape contact is used (Top), while the p-type part is contacted at the bottom face (Body). As mentioned in Section 2.2, contacts are considered of pure Ohmic-type with appropriate Dirichlet





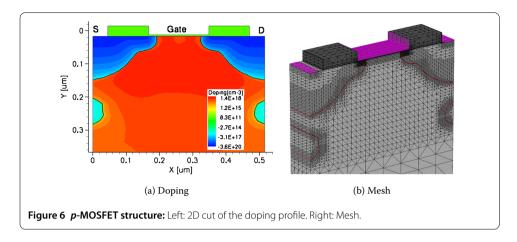
boundary conditions: the Top is maintained at ground while the Body is ramped at 0.8 V. In Figure 5 the results of the calculation of the current density vector field for electrons and holes in the semiconductor bulk are represented through streamlines connecting the Body with the Top contact. As expected the current calculation obtained with method DDFE (eq. (4a) and (4b)) is affected by a critical behavior, in particular close to and inside the *n*-junction as shown in Figure 5(a) and Figure 5(b) where instability has to be ascribed to numerical cancellation of the drift and diffusion contributions. Results get definitely better by employing method B using Eqs. (6j) and (6k) where the improvement can be appreciated in Figure 5(e) and Figure 5(f). However, a careful inspection of the hole current density reveals that some numerical instability is still evident inside the *n*-junctions. The extension of the 1D SG scheme to 3D provided by method A in Eq. (5h) and (5p) results in the streamlines presented in Figure 5(c) and Figure 5(d): no spurious instability can be observed anymore and our calculations are in excellent agreement with the results of a commercial code (not shown here). The comparison between the different current computational methods has also been carried out on a p-channel MOSFET. The doping profiles have been obtained by using a 2-D process simulator with implantation and diffusion steps [30] with the purpose to have a realistic doping as reported in Figure 6(a) with a GateLength: 180 nm and a GateOx: 4.5 nm. The presence of floating non-compensated p-type regions in the channel body increases the computational difficulties. The 2D doping profile has been then extruded in three spatial dimensions, and the generated mesh is shown in Figure 6(b) where the device contact has been highlighted with purple color (the body contact is not shown in the picture but is located in the bottom face of the sili-

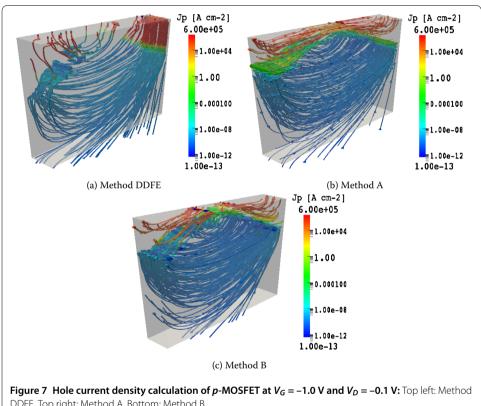


con region). Regarding the calculation of \mathbf{J}_p , the numerical difficulties found with method DDFE are still confirmed as clearly depicted in Figure 7(a) not only inside the p-type region but also around the floating regions present in the body (the visualization are referred to the bias conditions $V_G = -1$ V and $V_D = -0.1$ V). The marginality found using formula (6k) is increased as reported in Figure 7(c): this comes by the fact that the evaluation of the coefficient in (6i) is again undergoing numerical problems related to roundoff error. However method B is giving a much better current density evaluation with respect to the pure application of the Drift-Diffusion approach at the element level. We conclude this section by noting that, again, the best description of the expected physical behavior of the device is obtained by adopting method A, that turns out to provide an accurate and stable 3D extension of the 1D SG formula, as clearly demonstrated by Figure 7(b).

4.3 Application of method A to the study of a reverse-biased MOS with II

In this section we implement method A for current calculation inside the Gummel map algorithm to evaluate the II generation term by projecting the electric field along the obtained current streamlines. Since we are using the Van Overstraeten-de Man model for II we have run the simulation not beyond the limits of validity of this model, i.e. when the





DDFE. Top right: Method A. Bottom: Method B.

avalanche field is in the range of $6 \cdot 10^5 \, \text{Vcm}^{-1}$. Computations are performed by keeping all the contacts at ground and applying a reverse bias to the Drain contact. Figure 8(a) and Figure 8(b) compare the FEMOS simulations with those obtained using [30] for both n and p channel device under condition of reverse bias. The current calculated at Drain and Source contact are quite aligned with the reference tool while the Bulk component is lower due to the different generation value of II terms (not shown here). Figure 9 and Figure 10 visualize the increase of the II generation term with the increasing of the reverse bias in the case of the n-MOS and p-MOS device. The drain-to-bulk breakdown is starting in the Drain-Gate overlap region and proceed until it reaches the drain contact where the current streamlines are more dense as shown in Figure 11(b). A final insight into

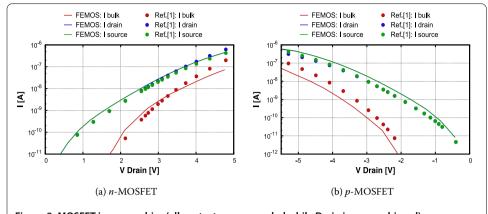


Figure 8 MOSFET in reverse bias (all contacts are grounded while Drain is reverse biased): comparison between FEMOS (line) and Ref. [30] (symbol). Left: *n*-channel device. Right: *p*-channel device.

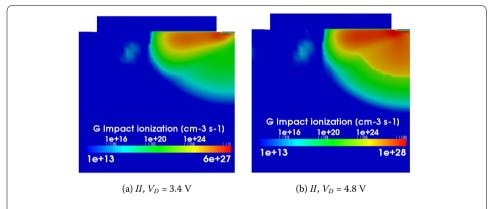
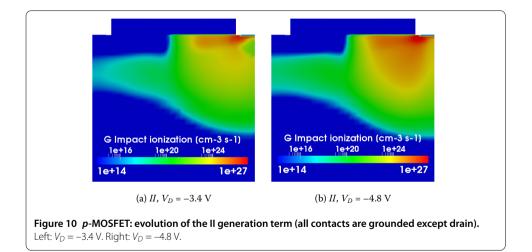


Figure 9 *n*-MOSFET: evolution of the II generation term (all contacts are grounded except drain). Left: $V_D = 3.4 \text{ V}$. Right: $V_D = 4.8 \text{ V}$.



the simulation of the n-MOS device using method A is offered by Figure 11 which reports the visualization of J_n in the off and on states: in accordance to physical expectation, in the on-state the electron current density is confined under the gate from drain-to-source while in the off-state it is flowing from source-to-drain and from source-to-body.

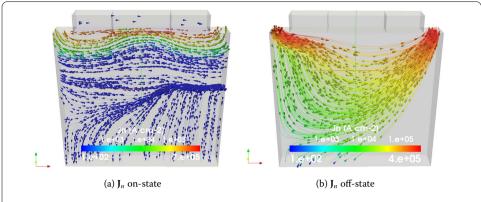


Figure 11 *n*-MOSFET: electron current density streamlines obtained with method A. Left: On-state at $V_G = 1 \text{ V}$ and $V_D = 0.1 \text{ V}$. Right: Off-state at $V_G = 0 \text{ V}$ and $V_D = 0.85 \text{ V}$.

5 Conclusions

In this article we have addressed the problem of representing in 3D the carrier current density by extending the beneficial properties emanating from the classic 1D Scharfetter-Gummel difference formula. To this purpose, we have adopted the Galerkin Finite Element Method for the numerical simulation of the Drift-Diffusion model in the 3D case, and we have proposed two novel methods for current density evaluation, denoted method A and B, to which, for comparison, we have added also the basic method DDFE using the DD formula. The three schemes compute a piecewise constant approximation of the current density over a tetrahedral partition of the device domain.

Method DDFE turns out to provide the worst results in the test experiments. Method B is a 3D extension of the method of optimal artificial diffusion and gives reasonably accurate results. Method A is based on a primal-mixed formulation endowed with a suitable quadrature formula for the approximate evaluation of the harmonic average of the electrical conductivity. It is by far the best of the three considered approaches, providing simulation results in excellent agreement with a commercial software. The three FE formulations are numerically verified in the study of realistic 3D device structures, also including the presence of Impact Ionization phenomena. Even in this latter case, method A is able to correctly describe the complex carrier flow patterns inside the device bulk and to track the I-V curve of the device up to the avalanche breakdown region.

Despite the fact that the proposed formulations are illustrated and validated in the study of the classic DD transport model in semiconductors, they can be applied in a straightforward manner to the numerical treatment of general conservation laws for advective-diffusive fluxes where the advective term is in gradient form, as is the case for ion electrodiffusion in electrochemistry and biology with the Poisson-Nernst-Planck model [34] and hydrodynamic and quantum-corrected charge transport in semiconductors [35–37].

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details

¹Process R & D, Micron Technology, via Olivetti 2, Agrate Brianza, MB 20864, Italy. ²IBM Italia, Via Circonvallazione Idroscalo, Segrate, MI 20090, Italy. ³STMicroelectronics, via Olivetti 2, Agrate Brianza, MB 20864, Italy. ⁴Dipartimento di Matematica, Politecnico di Milano, Piazza L. da Vinci 32, Milano, 20133, Italy.

Acknowledgements

Aurelio Mauri and Riccardo Sacco gratefully thank Drs. Claudio Lombardi and Chandra Mouli of Micron Technology, for supporting Andrea Bortolossi and Giovanni Novielli during their MSc thesis in the development of the program FEMOS.

Received: 21 December 2014 Accepted: 18 February 2015 Published online: 14 March 2015

References

- Cogez P, Graef M, Huizing B, Mahnkopf R, Ishiuchi H, Ikumi N, Choi S, Choi JH, Diaz CH, See YC, Gargini P, Kingscott T, Steff I. The International Technology Roadmap for Semiconductors: 2013. Technical report. Jointly sponsored by ESIA, JEITA, KSIA, TSIA and SIA. 2013. http://www.itrs.net/Links/2013ITRS/Summary2013.htm.
- 2. Novielli G, Ghetti A, Varesi E, Mauri A, Sacco R. Atomic migration in phase change materials. In: Electron devices meeting (IEDM), 2013 IEEE international. 2013. p. 22.3.1-22.3.4. doi:10.1109/IEDM.2013.6724683.
- Mauri A, Sacco R, Verri M. Electro-thermo-chemical computational models for 3D heterogeneous semiconductor device simulation. Appl Math Model. 2014 in press, doi:10.1016/j.apm.2014.12.008.
- Benvenuti A, Ghetti A, Mauri A, Liu H, Mouli C. Current and future prospects of non-volatile memory modeling. In: Simulation and Semiconductor Process and Devices (SISPAD), 2014 IEEE international. 2014. p. 249-52.
- 5. Selberherr S. Analysis and simulation of semiconductor devices. Berlin: Springer; 1984.
- 6. Jerome JW. Analysis of charge transport. Berlin: Springer; 1996.
- Markowich PA. The stationary semiconductor device equations. Computational microelectronics. Berlin: Springer; 1986. ISBN 9783211818923.
- 8. Micheletti S, Quarteroni A, Sacco R. Current-voltage characteristics simulation of semiconductor devices using domain decomposition. J Comput Phys. 1995;119(1):46-61. doi:10.1006/jcph.1995.1115.
- 9. Bank RE, Rose D, Fichtner W. Numerical methods for semiconductor device simulation. SIAM J Sci Stat Comput. 1983;4(3):416-35. doi:10.1137/0904032.
- Bank RE, Rose D, Fichtner W. Numerical methods for semiconductor device simulation. IEEE Trans Electron Devices. 1983;30(9):1031-41. doi:10.1109/T-ED.1983.21257.
- Ciampolini P, Pierantoni A, Liuzzo A, Baccarani G. 3-D simulations of silicon devices: physical models and numerical algorithms. In: Baccarani G, editor. Process and device modeling for microelectronics. Amsterdam: Elsevier; 1993. p. 53-107.
- 12. Roberts J, Thomas JM. Mixed and hybrid methods. Handb Numer Anal. 1991;2:523-633.
- 13. Sacco R, Carichino L, de Falco C, Verri M, Agostini F, Gradinger T. A multiscale thermo-fluid computational model for a two-phase cooling system. Comput Methods Appl Mech Eng. 2014;282(1):239-68. doi:10.1016/j.cma.2014.08.003.
- Bank RE, Burgler JF, Fichtner W, Smith RK. Some upwinding techniques for finite element approximations of convection-diffusion equations. Numer Math. 1990;58:185-202.
- Babuška I, Osborn JE. Generalized finite element methods: their performance and their relation to mixed methods. SIAM J Numer Anal. 1983;20:510-36.
- Brezzi F, Marini L, Pietra P. Two-dimensional exponential fitting and applications to Drift-Diffusion models. SIAM J Numer Anal. 1989;26(6):1342-55. doi:10.1137/0726078.
- 17. Gummel HK, Scharfetter D. Large-signal analysis of a silicon Read diode oscillator. IEEE Trans Electron Devices. 1969;16:64-77.
- Brooks AN, Hughes TJR. Streamline Upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng. 1982;32:199-259. doi:10.1016/0045-7825(82)90071-8.
- Ortega JM, Rheinboldt WC. Iterative solution of nonlinear equations in several variables. New York: Academic Press; 1970.
- 20. Gummel HK. A self-consistent iterative scheme for one-dimensional steady state transistor calculations. IEEE Trans Electron Devices. 1964;11:455-65.
- 21. Quarteroni A, Valli A. Numerical approximation of partial differential equations. Berlin: Springer; 2008.
- 22. Xu J, Zikatanov L. A monotone finite element scheme for convection-diffusion equations. Math Comput. 1999;68:1429-46.
- 23. Zikatanov LT, Lazarov RD. An exponential fitting scheme for general convection-diffusion equations on tetrahedral meshes. Comput Appl Math. 2012;1:60-9.
- 24. Bank RE, Coughran W, Cowsar LC. The Finite Volume Scharfetter-Gummel method for steady convection diffusion equations. Comput Vis Sci. 1998;1:123-36.
- Gatti E, Micheletti S, Sacco R. A new Galerkin framework for the Drift-Diffusion equation in semiconductors. East-West J Numer Math. 1998;6:101-36.
- 26. Degond P, Juengel A, Pietra P. Numerical discretization of energy-transport models for semiconductors with nonparabolic band structure. SIAM J Sci Comput. 2000;22(3):986-1007. doi:10.1137/S1064827599360972.
- Brezzi F, Marini L, Pietra P. Numerical simulation of semiconductor devices. Comput Methods Appl Mech Eng. 1989;75:493-514.
- 28. Brezzi F, Marini LD, Markowich PA, Pietra P. On some numerical problems in semiconductor device simulation. In: Toscani G, Boffi V, Rionero S, editors. Mathematical aspects of fluid and plasma dynamics. Lecture notes in mathematics. vol. 1460. 1991. p. 31-42.
- 29. Allen DNG, Southwell RV. Relaxation methods applied to determine the motion, in two dimensions, of a viscous fluid past a fixed cylinder. Q J Mech Appl Math. 1955;8:129-45.
- 30. Sentaurus device user guide. Synopsis Inc. 2013.
- Canali C. Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature. IEEE Trans Electron Devices. 1975;22:1045-7.
- 32. Gusmeroli R, Spinelli AS. Accurate boundary integrals calculation in semiconductor device simulation. IEEE Trans Electron Devices. 2006;53:1730-3.
- Hughes TJR, Engel G, Mazzei L, Larson MG. The continuous Galerkin method is locally conservative. J Comput Phys. 2000:163:467-88.

- 34. Rubinstein I. Electrodiffusion of ions. Philadelphia: SIAM; 1990.
- 35. Forghieri A, Guerrieri R, Ciampolini P, Gnudi A, Rudan M, Baccarani G. A new discretization strategy of the semiconductor equations comprising momentum and energy balance. IEEE Trans Comput-Aided Des Integr Circuits Syst. 1988;7(2):231-42. doi:10.1109/43.3153.
- 36. de Falco C, Lacaita AL, Gatti E, Sacco R. Quantum-corrected Drift-Diffusion models for transport in semiconductor devices. J Comput Phys. 2005;204:533-61.
- 37. de Falco C, Jerome JW, Sacco R. Quantum-corrected drift-diffusion models: solution fixed point map and finite element approximation. J Comput Phys. 2009;228(5):1770-89. doi:10.1016/j.jcp.2008.11.010.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Immediate publication on acceptance
- \blacktriangleright Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com