A Deterministic Boltzmann Equation Solver for Two-Dimensional Semiconductor Devices

Sung-Min Hong, Christoph Jungemann
EIT4
Bundeswehr University
85577 Neubiberg, Germany
Email: hi2ska2@gmail.com, Jungemann@ieee.org

Matthias Bollhöfer
Institut Computational Mathematics
Technical University of Braunschweig
38106 Braunschweig, Germany

***Abstract*— We have developed a Boltzmann equation solver for two-dimensional (2D) semiconductor devices based on the spherical harmonics expansion and the maximum entropy dissipation scheme for stabilization. The large system of equations is partitioned according to the order of the spherical harmonics and solved by a memory efficient blockwise Gauss-Seidel method. Results are presented for a 2D NPN Si bipolar junction transistor.**

I. Introduction

Semiclassical transport in semiconductor devices can be described by the Boltzmann equation (BE). The usual approach for solving the BE, the Monte Carlo (MC) method, has many disadvantages due to its stochastic nature. For example, small currents entail excessive CPU times and small-signal analysis is exceedingly difficult. Therefore, a deterministic approach to the BE is preferable for such calculations.

Among alternative nonstochastic methods for the BE is the spherical harmonics expansion (SHE) [1], [2], where the electron distribution function is expanded with spherical harmonics, $Y_{l,m}$ in the wavevector space reducing the high dimensionality of the phase space.

Since the necessary order of SHE increases with decreasing device size as the transport becomes more and more ballistic [3], a SHE solver with an arbitrary number of spherical harmonics is required. For 1D devices, it has been successfully applied to the calculation of transport [3], [4] and noise [5] including many spherical harmonics. However, the application of SHE to 2D devices has been restricted to the lowest possible number of spherical harmonics, l_{max}=1, so far [1], [6]. The main obstacle which prevents inclusion of higher number of spherical harmonics is the huge memory requirement.

In this work, we present an implementation of the SHE solver for 2D devices with an arbitrary number of spherical harmonics. In Sec. II, the two-dimensional implementation is presented. An interpolation scheme for the total energy of the valley minima in the 2D real space is described. In Sec. III, a partitioning scheme for the memory-efficient Gauss-Seidel method [7] is explained. In Sec. IV, a 2D NPN Si bipolar junction transistor is investigated as a numerical example. Conclusions are drawn in Sec. IV.

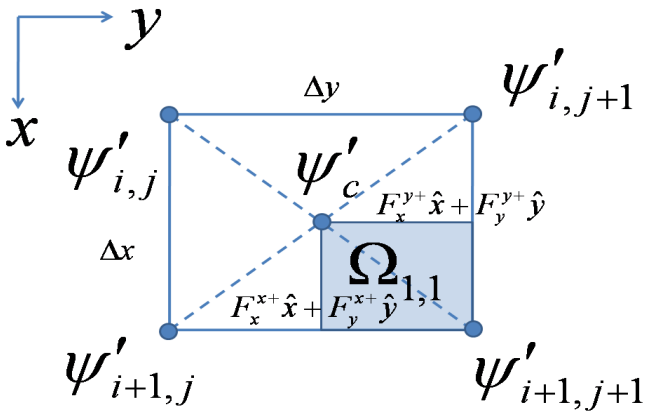

Fig. 1. Cartesian tensor product grid for the 2D real space. The four grid nodes at the corners of the rectangle belong to the direct grid and the center node to the adjoint one.

II. Two-Dimensional Implementation

The stabilization scheme based on the maximum entropy dissipation scheme [8], previously applied to 1D devices [3], has been extended to 2D. A staggered grid in the 2D real space and the energy space is used, where the even components of the distribution function are defined on the nodes of the direct grid, and the odd ones on the adjoint grid. In Fig. 1, the Cartesian tensor product grid for the 2D real space is shown. The control volume of a certain adjoint node is denoted as Ω.

In order to obtain a stabilized expression for the odd components of the distribution function (fluxes), the following integrals over a part of the control volume of a certain adjoint node, $\Omega_{\alpha,\beta}$, are required [3].

$$HV_{\nu,\alpha,\beta} = \int_{\Omega_{\alpha,\beta}} d\mathbf{r} \exp(-\psi'_\nu(\mathbf{r})), \tag{1}$$

$$H\mathbf{A}_{\nu,\alpha,\beta} = \int_{\Omega_{\alpha,\beta}} d\mathbf{a} \exp(-\psi'_\nu(\mathbf{r})), \tag{2}$$

978-1-4244-1753-7/08/$25.00 ©2008 IEEE

where $-\psi'_\nu(\mathbf{r})$ is the (normalized) total energy of the minimum of the ν-th valley and $d\mathbf{a}$ contains only the surface which overlaps with the surface of Ω. Since the electrostatic potential is defined on the nodes of the direct grid, the total energy of the valley minimum, which is the sum of the conduction band edge, valley shift and electrostatic potential, is also defined on those nodes. Therefore, we need an interpolation formula to compute the valley minimum over Ω from the values defined on the direct nodes.

The valley minimum on the node of the adjoint grid at the center of the control volume is assumed to be the average of the values on the four direct nodes.

$$\psi'_c = \frac{\psi'_{i,j} + \psi'_{i+1,j} + \psi'_{i,j+1} + \psi'_{i+1,j+1}}{4}. \tag{3}$$

Then, the control volume is splitted into four triangles along its diagonals. After that, the valley minimum is linearly interpolated in each triangle. The two integrals in (1) and (2) for $\Omega_{1,1}$ in Fig. 1, can be solved exactly

$$HV_{\nu,1,1} = \exp(-\psi'_c)\frac{\Delta x}{2}\frac{\Delta y}{2} \times \left\{B_2^{-1}(-F_x^{x+}, -F_y^{x+}) + B_2^{-1}(-F_y^{y+}, -F_x^{y+})\right\}, \tag{4}$$

$$H\mathbf{A}_{\nu,1,1} = \frac{\Delta y}{2}\exp(-\psi'_{i+1,j+1})B^{-1}(-\frac{\psi'_{i+1,j+1} - \psi'_{i+1,j}}{2})\hat{x} + \frac{\Delta x}{2}\exp(-\psi'_{i+1,j+1})B^{-1}(-\frac{\psi'_{i+1,j+1} - \psi'_{i,j+1}}{2})\hat{y}, \tag{5}$$

where $B_2^{-1}(x,y) = \frac{B^{-1}(x+y) - B^{-1}(x)}{y}$, $B^{-1}(x) = \frac{\exp(x)-1}{x}$, and the forces F_x^{x+}, F_y^{x+}, F_x^{y+}, and F_y^{y+} are defined in Fig. 1. Similar expressions can be obtained for the other regions.

III. Blockwise Gauss-Seidel Method

The inclusion of a second dimension in real space results in the loss of the cylindrical symmetry of the electron distribution function and the number of unknowns grows quadratically with l_{max}. In addition, the number of entries per equation increases resulting in a huge increase in required memory. Although ILUPACK [9], the linear solver used in this work, is very memory efficient, a further reduction in memory space is required for a large l_{max}.

In this work, we adopt the blockwise Gauss-Seidel method [7], and a partitioning scheme suitable for the Gauss-Seidel method is developed. Since the zeroth order component yields the electron density and the first order the current density, we can conclude that important information is contained in the lowest order components. Therefore, a partitioning scheme according to the order of spherical harmonics is a natural choice.

As a first step, the odd equations are eliminated by a pre-solver. The resultant compressed Jacobian matrix is arranged and divided into many small blocks according to the order of the spherical harmonics. For example, in the case of $l_{max} = 5$ three diagonal blocks of the order 0, 2, and 4 appear (where the zeroth block contains also the Poisson equation). The number of unknowns for the block of order l, N_l, is given by

$$N_l = N_{xy} \times N_\epsilon \times N_\nu \times (l+1), \tag{6}$$

where N_{xy} is the number of the spatial nodes, N_ϵ the number of nodes in energy, N_ν the number of distinguishable valleys in the Si conduction band, and the factor $(l+1)$ is introduced to consider different m values.

Since we neglect the Pauli principle, there is no coupling between different valleys for blocks with $l > 0$. This observation allows us to further divide each block with positive l into N_ν independent subblocks. For example, when we consider $l_{max} = 9$, the number of unknowns for the biggest subblock is just three times larger than the smallest, zeroth block.

Because only the diagonal blocks are exactly solved by ILUPACK, the memory requirement is greatly reduced. The incomplete factorization for the diagonal blocks, which is time-consuming for 2D devices, needs to be performed only once in each Newton step. Moreover, each block can be handled by a different computer and the memory restriction is further alleviated. The Gauss-Seidel method is used to solve the complete linear system by iteration.

IV. Numerical Example

Results are presented for a 2D NPN Si BJT. The electron model is based on the analytical six valley band structure and phonon scattering mechanisms developed by the Modena group [10]. The elliptical valleys are mapped onto spherical ones by the Herring-Vogt transform [11]. Due to the translational symmetry along the third dimension in the real space, all coefficients of the distribution function with negative m vanish.

The 2D doping profile is shown in Fig. 2. The base of this transistor is 1 μm thick. A heavy p-type doping is included to form the base contact. The total number of spatial grid points is 800. When we consider $l_{max} = 9$, the total number of unknown variables in the uncompressed form is 6734400. In its compressed form, the number is 3060800.

Fig. 3 shows the electron current at $V_{BE} = 0.8$ V as obtained by SHE with $l_{max} = 1$ and $l_{max} = 3$. Early voltages calculated from $l_{max} = 1$ and $l_{max} = 3$ are -34.8 V and -34.5 V, respectively. It is noted that the Early voltage is very difficult to simulate by MC, due to its stochastic errors.

In Fig. 4, the error in collector current is shown. The error in collector current decreases exponentially with the maximum order. An expansion up to the third order yields an error which is sufficiently less than 1.0 %. However, as stated in Introduction, the necessary order of SHE increases with decreasing device size as the transport

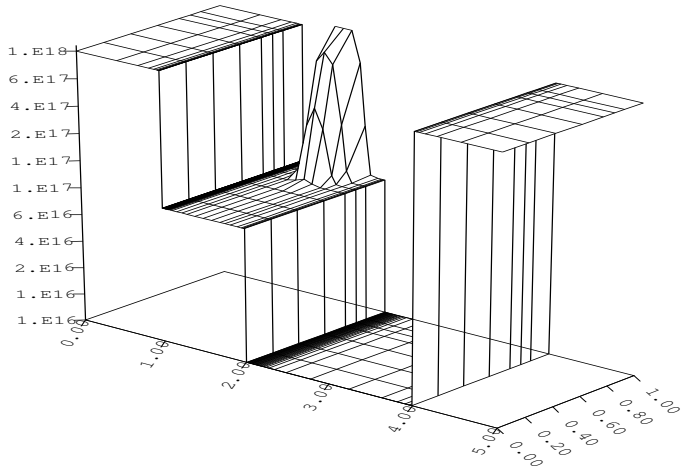

Fig. 2. Doping profile (absolute value) and nonequidistant grid of the Si BJT.

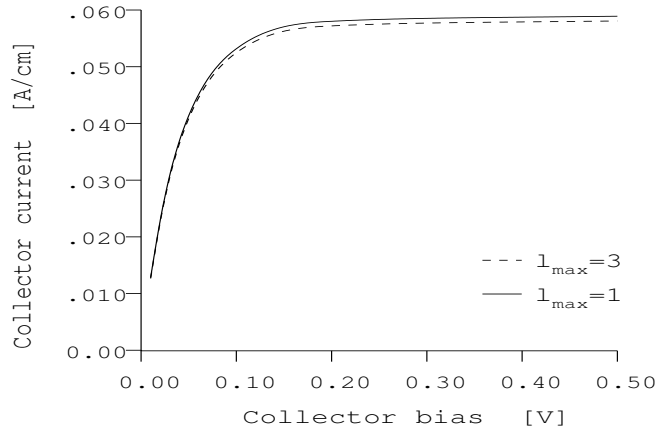

Fig. 3. Collector current at $V_{BE} = 0.8$ V as obtained by SHE. Early voltages calculated with $l_{max} = 1$ and $l_{max} = 3$ are -34.8 V and -34.5 V, respectively.

becomes more and more ballistic [3]. Fig. 5 shows the electron velocity along x-direction at $V_{BE} = 0.8$ V and $V_{CE} = 0.5$ V for different numbers of spherical harmonics. The convergence of SHE is clearly visible.

The linear response of the electron density, divided by the linear response of the collector current, is shown in Fig. 5. The bias condition is $V_{BE} = 0.8$ V and $V_{CE} = 0.5$ V, and a small voltage perturbation is applied to the base terminal. Thus, in contrast to MC, small-signal analysis, such as the cutoff frequency calculation in the quasi-stationary limit [12], is possible by SHE.

The convergence property of the blockwise Gauss-Seidel method for different Newton steps is shown in Fig. 7. The forcing term of the inexact Newton method [13] is set to be 10^{-4}. Up to the third Newton step, the norm of the residual vector decreases very rapidly, therefore, the blockwise Gauss-Seidel method takes only a few iterations. After the third newton step the convergence speed degrades, and approaches a lower constant value.

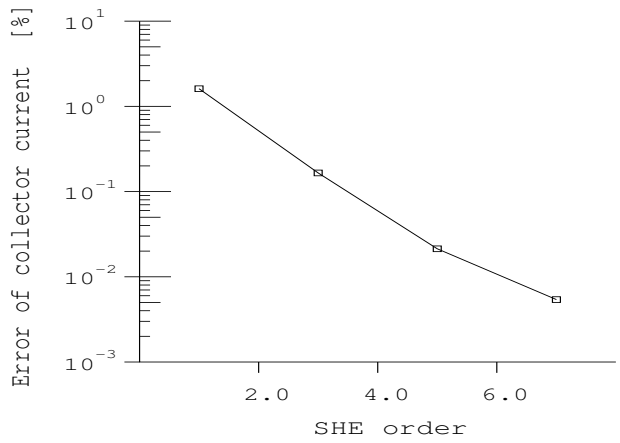

Fig. 4. Error of the collector current at $V_{BE} = 0.8$ V and $V_{CE} = 0.5$ V for different numbers of spherical harmonics relative to the ninth order.

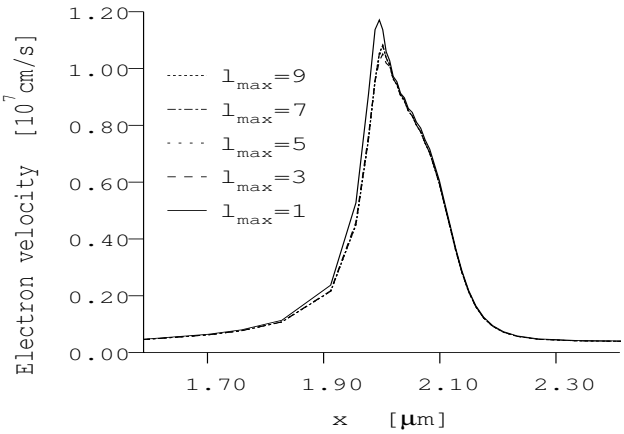

Fig. 5. Electron velocity along x-direction at $V_{BE} = 0.8$ V and $V_{CE} = 0.5$ V for different numbers of spherical harmonics. $y = 0$.

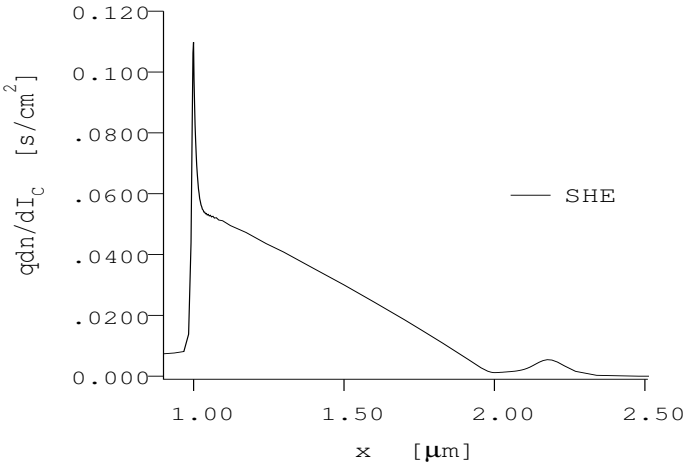

Fig. 6. Linear response of the electron density, divided by the linear response of the collector current. $V_{BE} = 0.8$ V and $V_{CE} = 0.5$ V. A small voltage perturbation is applied to the base terminal. $y = 0$.

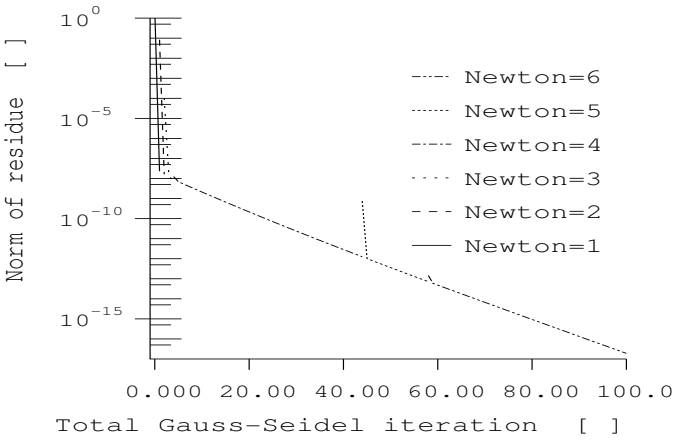

Fig. 7. Norm of the residue versus the number of Gauss-Seidel iterations summed over the Newton steps. V_{BE} = 0.8 V and V_{CE} = 0.5 V. l_{max} = 5.

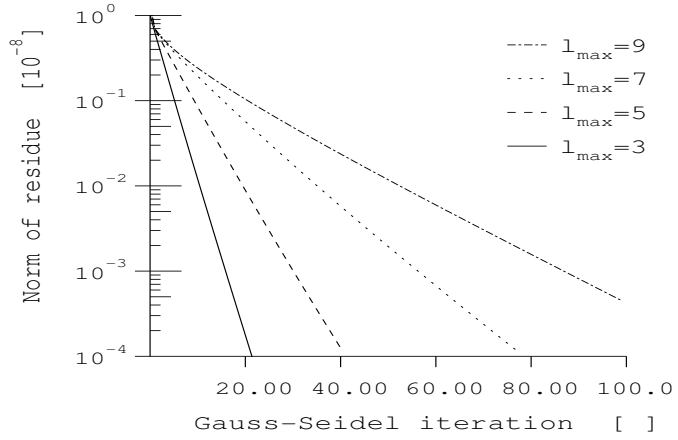

Fig. 8. Convergence property of the blockwise Gauss-Seidel method for different numbers of spherical harmonics. V_{BE} = 0.8 V and V_{CE} = 0.5 V. The norm for the fourth Newton iteration is shown.

The convergence property of the blockwise Gauss-Seidel method for different l_{max} is shown in Fig. 8. The norm for the fourth Newton iteration is shown. The convergence speed of the blockwise Gauss-Seidel method degrades as l_{max} and therewith the number of blocks increases.

V. Conclusion

In this work, we have presented the first implementation of a SHE solver for 2D devices with l_{max} larger than one. For the stabilization in the 2D real space, an interpolation scheme for the valley minimum was employed. To lessen the memory restriction, the memory-efficient blockwise Gauss-Seidel method was employed. Also a partitioning scheme suitable for the Gauss-Seidel method was developed. As an example, simulation results for a 2D NPN Si BJT with l_{max} up to 9, were shown.

Acknowledgment

The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG). S.-M. Hong's work was partially supported by Korea Research Foundation Grant by the Korean Government(MOEHRD). (KRF-2007-357-D00159)

References

[1] A. Gnudi, D. Ventura, G. Baccarani, and F. Odeh, "Two-dimensional MOSFET simulation by means of a multidimensional spherical harmonics expansion of the Boltzmann transport equation," Solid–State Electronics, vol. 36, pp. 575–581, 1993.

[2] N. Goldsman, L. Henrichson, and J. Frey, "A physics-based analytical/numerical solution to the Boltzmann transport equation for use in device simulation," Solid–State Electronics, pp. 575–581, 1993.

[3] C. Jungemann, A. T. Pham, B. Meinerzhagen, C. Ringhofer, and M. Bollhöfer, "Stable discretization of the Boltzmann equation based on spherical harmonics, box integration, and a maximum entropy dissipation principle," Journal of Applied Physics, vol. 100, p. 024502, 2006.

[4] K. Rahmat, J. White, and D. A. Antoniadis, "Simulation of semiconductor devices using a Galerkin/spherical harmonic expansion approach to solving the coupled Poisson-Boltzmann system," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 15, pp. 1181–1196, 1996.

[5] C. Jungemann, "A deterministic approach to RF noise in silicon devices based on the Langevin-Boltzmann equation," IEEE Transactions on Electron Devices, vol. 54, pp. 1185–1192, 2007.

[6] W. Liang, N. Goldsman, I. Mayergoyz, and P. J. Oldiges, "2-D MOSFET modeling including surface effects and impact ionization by self-consistent solution of the Boltzmann, Poisson, and hole-continuity equations," IEEE Transactions on Electron Devices, vol. 44, pp. 257–267, 1997.

[7] N. Black and S. Moore, "Gauss-Seidel Method," from MathWorld–A Wolfram Web Resource, created by Eric W. Weisstein. http://mathworld.wolfram.com/Gauss-SeidelMethod.html

[8] C. Ringhofer, "A mixed spectral-difference method for the steady state Boltzmann-Poisson system," SIAM Journal of Numerical Analysis, vol. 41, pp. 64–89, 2003.

[9] M. Bollhöfer and Y. Saad, ILUPACK, preconditioning software package, release 2.2.

[10] R. Brunetti, C. Jacoboni, F. Nava, L. Reggiani, G. Bosman, and R. J. J. Zijlstra, "Diffusion coefficient of electrons in silicon," Journal of Applied Physics, vol. 52, pp. 6713–6722, 1981.

[11] C. Herring and E. Vogt. Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering. *Phys. Rev.*, 101(3):944–962, 1956.

[12] H. K. Gummel, "On the definition of the cutoff frequency f_T," Proceedings of the IEEE, vol. 57, p. 2159, 1969.

[13] S. C. Eisenstat and H. F. Walker, "Choosing the forcing terms in an inexact Newton method," SIAM Journal of Scientific Computing, vol. 17, pp. 16–32, 1996.