On the Feasibility of Spherical Harmonics Expansions of the Boltzmann Transport Equation for Three-Dimensional Device Geometries

Karl Rupp°,*, Tibor Grasser°, and Ansgar Jüngel*
°Institute for Microelectronics, TU Wien, Austria.
*Institute for Analysis and Scientific Computing, TU Wien, Austria.

Abstract

Accurate simulation of carrier transport requires the solution of Boltzmann's transport equation (BTE), which can be obtained by higher-order spherical harmonics expansion (SHE) techniques. Unfortunately, the high computational effort of the SHE method has so far prevented its application to 3D geometries. We refine the SHE method by suggesting and evaluating numerical techniques which allow for efficient solution of higher-order expansions even in the unchartered 3D regime.

Introduction

In order to overcome the limitations of the Monte-Carlo (MC) method, like the inherent noise and the long simulation times, deterministic solution methods for the BTE using a first-order spherical harmonics expansion (SHE) have been suggested (1),(2). Since a first-order expansion has been found to be insufficient (3), the SHE method has been extended to include higher expansion orders (3),(4). While the SHE method is considerably faster than the MC technique and allows for accurate resolution of the lowly populated tails of the distribution function, the computational costs are much higher than other established macroscopic transport models, especially when higher expansion orders are considered. Unfortunately, particularly for highly scaled devices, such higher expansion orders are required to obtain a level of accuracy comparable to MC simulations (3),(4).

With fully 3D semiconductor device layouts hitting the market, simulations using a 2D device layout are unable to reflect all physical processes. So far, however, the memory requirements of the SHE method have prevented its use for this case. In order to overcome this limitation, we employ adaptive variable-order expansions to better resolve sensitive areas of the device, use unstructured grids in 3D space, and use a preconditioner which allows for parallelization on modern multi-core and many-core CPUs and GPUs (graphics processing units). We demonstrate that these extensions allow for the simulation of state-of-the-art devices in three spatial dimensions in reasonable simulation times.

Adaptive Variable-Order SHE

The SHE method is based on an expansion of the carrier distribution function

$$f(\boldsymbol{x},\boldsymbol{p},t) \cong \sum_{l=0}^{L}\sum_{m=-l}^{l} f_{l,m}(\boldsymbol{x},H,t)Y_{l,m}(\theta,\varphi) \ ,$$

where $\boldsymbol{x}$ denotes the spatial point, $\boldsymbol{p}=\hbar\boldsymbol{k}$ the semiclassical momentum, t time, H total energy and $Y_{l,m}$ the spherical harmonic of order l with angular components θ and φ. So far, a certain fixed expansion order L, typically $L=1$ or $L=3$, and more recently $L=5$ or $L=7$, has been employed uniformly over the simulation domain (3),(4).

While higher expansion orders considerably improve the accuracy, the total number of unknowns depends quadratically on L. Thus, the use of the required expansion orders 5 to 7 becomes very demanding in 2D and virtually intractable in 3D using the existing methods (3)–(5). We have recently suggested an adaptive expansion-order approach and evaluated it on 1D *nin* structures (7). Here, we extend this method to the 3D case. We expand the distribution function adaptively by setting the expansion order individually at each grid point in the simulation domain, i.e. $L=L(\boldsymbol{x}_i,H_i)$, where $\boldsymbol{x}_i$ denote the spatial grid points and H_i denote the total energy grid points.

It is impractical for every-day TCAD purposes to specify the expansion orders manually at each point of the device. Therefore, expansion orders are assigned by a fully-adaptive scheme: Starting with a uniform first-order expansion, expansion orders are automatically increased based on an error estimator. This procedure is repeated until the desired accuracy is obtained, or a user-provided maximum order is reached.

The error estimator is based on the fact that a zeroth order expansion is sufficient for the carrier distribution function in equilibrium. Also, it can be shown (6) that for a series expansion of an arbitrary function g in spherical harmonics the rate of decay of SHE coefficients is controlled by the number of continuous derivatives of g. Our indicator estimates the decay rate of the computed distribution function and increases the expansion order in regions where the decay is slow, indicating that f is far from equilibrium. More precisely, on each grid point $(\boldsymbol{x}_i,H_i)$ we use an estimator of the form $\eta(\boldsymbol{x}_i,H_i)=\eta_{\mathrm{d}}(\boldsymbol{x}_i,H_i)+\eta_{\mathrm{e}}(\boldsymbol{x}_i,H_i)$, where the two summands are given by

$$\eta_{\mathrm{d}}(\boldsymbol{x}_i,H_i) = \frac{\sum_{\mathrm{edges}}\sum_m |f_{L_i+1,m}|}{N_{\mathrm{edges}}(\boldsymbol{x}_i,H_i)\log(L_i+1)f_{0,0}(\boldsymbol{x}_i,H_i)},$$
$$\eta_{\mathrm{e}}(\boldsymbol{x}_i,H_i) = \alpha\log(f_{0,0}(\boldsymbol{x}_i,H_i)) \ ,$$

with the number of edges N_{edges} connected to the respective grid point and L_i denoting the maximum (even) expansion order at point $(\boldsymbol{x}_i,L_i)$. The function η_{d} estimates the decay of the spherical harmonics expansion at each grid point in the simulation domain. η_{e} is an empirical term accounting for the exponential decay of the

distribution function by adding extra weight closer to the band-edge, where higher contributions of the distribution function for the calculation of macroscopic quantities are obtained (7).

Unstructured Grids

A three-dimensional device simulation using macroscopic transport models such as the hydrodynamic model is commonly accomplished with the use of unstructured grids. Consequently, the use of unstructured grids with the SHE method is a key requirement for a state-of-the-art device simulator in order to be able to resolve complicated device geometries, to keep computational effort low, and to substantially ease the use with existing TCAD tools, for which SHE is applied as a post-processor for the electrostatic potential obtained by a simpler macroscopic model. However, no arbitrary-order SHE simulations using unstructured grids for spatially two- or three-dimensional device simulations have been reported so far.

We employ unstructured grids in two and three spatial dimensions using the box integration scheme proposed in (4). This allows for a resolution of even complicated device geometries while keeping the number of grid-points – and thus the computational effort – as low as possible. Although the spatial mesh is two- or three-dimensional, we treat the energy as an additional dimension, hence dealing with three- and four-dimensional simulation domains respectively.

Parallelization

The high computational effort for SHE requires the efficient utilization of modern multi-core and many-core architectures in CPUs and GPUs. This has not been possible for SHE so far, because good preconditioners are required to obtain reasonable convergence rates for the iterative linear solvers. The employed incomplete LU (ILU) factorization preconditioners used in recent publications (3), (4) are serial in nature and thus rely on a single-threaded execution model.

We have recently suggested a parallel block-preconditioner (8), which utilizes the fact that SHE of a reduced BTE, where only elastic scattering processes are taken into account, leads to a decoupled system of equations for each energy level. We use this simplified, decoupled system for the parallel construction of a preconditioner, which is then employed for the full, coupled system. This allows for a solution of the fully coupled problem with full accuracy at the high degree of parallelization of the decoupled problem.

Results

The adaptive variable-order expansion is demonstrated for the self-consistent simulation of a tri-gate transistor using our free open-source simulator `ViennaSHE` (9). Thanks to the lower number of points in the simulation domain outlined in Fig. 1, only a fifth of the number of unknowns compared to a structured grid enter the simulation. The

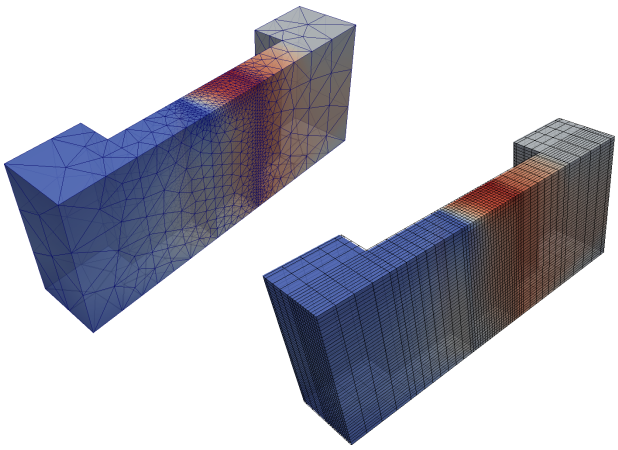
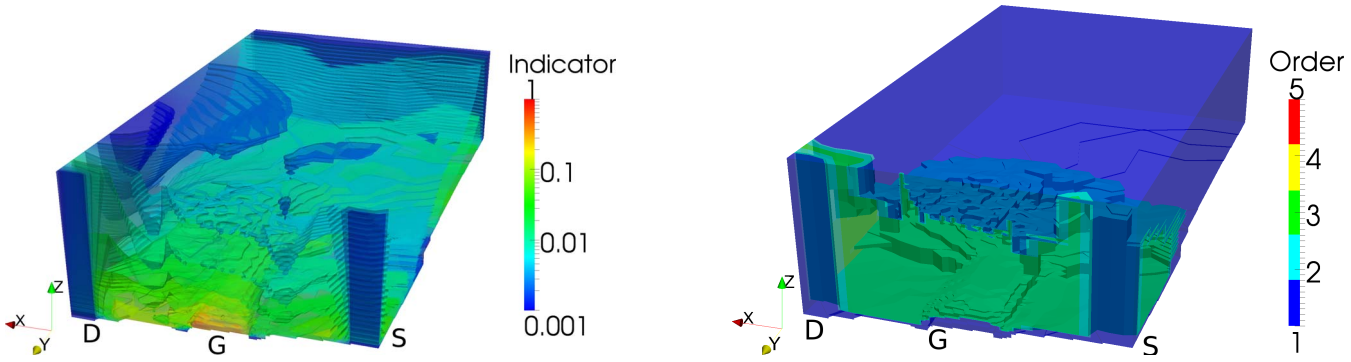


Figure 1: Comparison of the tetrahedral unstructured grid used for the simulation of a tri-gate transistor (no gate, oxide and body shown), and a structured grid with the same mesh size inside the channel. While the tetrahedral grid consists of 4838 points, the structured grid is made up of 27456 points, thus leading to considerably higher computational costs.

adaptive order expansion, as illustrated along the plane of symmetry of the device in Fig. 2, requires considerably less unknowns (Fig. 3), while excellent agreement with uniform expansions (Fig. 4) is preserved. Even without taking the savings due to unstructured grids into account, total SHE solver times are reduced by up to two orders of magnitude when combining the adaptive order strategy with parallelization, cf. Fig. 5. For accessing the power of multi-core CPUs and GPUs, our free open source library `ViennaCL` (10) is used. With the full distribution function at hand, quantities such as the density of carriers with energy above 1eV or the average carrier energy can be accurately extracted (Fig. 6) for the purpose of e.g. device degradation analysis, which is not available from momentum-based models.

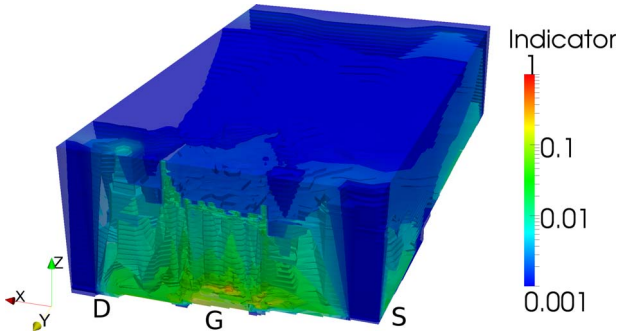
Conclusion

The proposed adaptive variable-order scheme allows us to solve the BTE with the accuracy of uniform high-order spherical harmonics expansions at considerably lower computational costs. The extension to unstructured grids and adaptive variable-order expansions renders the SHE method very attractive for TCAD tools, since it provides the accuracy of Monte Carlo methods at a fraction of the computational costs even on complicated 3D device geometries. For the case of the trigate transistor considered in this work, a reduction of memory requirements by one order of magnitude, and of execution times by two orders of magnitude over existing approaches are obtained. On the overall, our enhancements to the deterministic SHE method allows for the investigation of various phenomena such as impact ionization or hot carrier degradation at higher accuracy compared to established macroscopic models without an artificial restriction to simple device geometries.

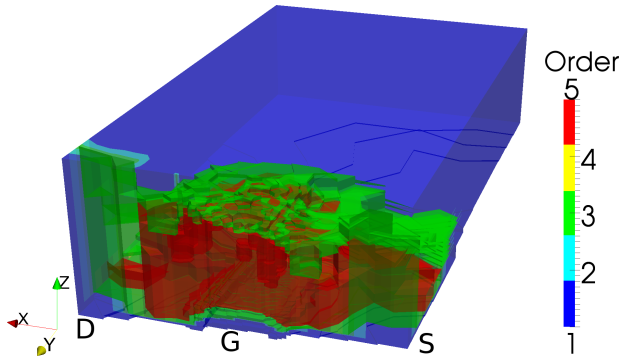


(a) Error Indicator used for the first adaption step.

(b) Expansion orders after the first adaption step.



(c) Error Indicator used for the second adaption step.



(d) Expansion orders after the second adaption step.

Figure 2: Based on the error indicator the expansion orders are locally increased. The bulk is kept at a fixed first order, since the contribution to transport is negligible. The z-axis denotes total energy.

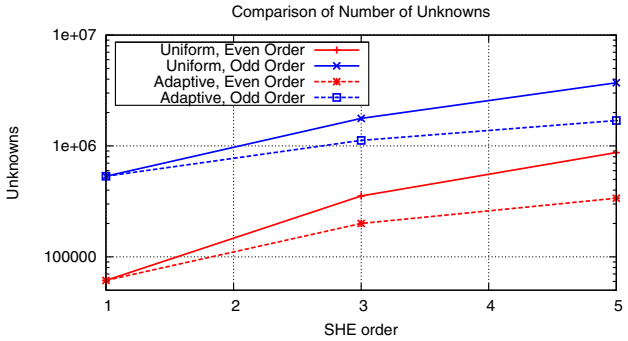


Figure 3: Comparison of the number of unknowns of the linear system on a coarse mesh in order to compare with high uniform expansion orders. The use of adaptive expansion orders leads to a reduction by a factor of more than two.

Acknowledgement

K. Rupp and A. Jüngel acknowledge partial support from the Austrian Science Fund (FWF), grants P20214, P22108, and I395; the Austrian-Croatian Project HR 01/2010; the Austrian-French Project FR 07/2010; and the Austrian-Spanish Project ES 08/2010 of the Austrian Exchange Service (ÖAD). Support by the Graduate School PDEtech at the TU Wien is gratefully acknowledged by the authors.

References

(1) N. Goldsman, L. Hendrickson, and J. Frey, A Physics-Based Analytical/Numerical Solution to the Boltzmann Transport Equation for the Use in Device Simulation, *Solid State Electronics*, vol. 34, p. 389–396 (1991).

(2) A. Gnudi, D. Ventura, G. Baccarani, F. Odeh, Two-Dimensional MOSFET Simulation by Means of a Multidimensional Spherical Harmonics Expansion of the Boltzmann Transport Equation, *Solid State Electronics*, vol. 36, no. 4, p. 575–581 (1993).

(3) C. Jungemann, A. T. Pham, B. Meinerzhaben, C. Ringhofer, 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, no. 2, p. 024502-+ (2006).

(4) S. M. Hong and C. Jungemann, A Fully Coupled Scheme for a Boltzmann-Poisson Equation Solver Based on a Spherical Harmonics Expansion, *Journal of Computational Electronics*, vol. 8, p. 225–241 (2009).

(5) K. Rupp, A. Jüngel, T. Grasser, Matrix Compression for Spherical Harmonics Expansions fo the Boltzmann Transport Equation for Semiconductors, *Journal of Computational Physics*, vol. 229, p. 8750–8765 (2010).

(6) W. Freeden, T. Gervens and M. Schreiner, *Constructive Approximation on the Sphere*, Clarendon Press (1998).

(7) K. Rupp, T. Grasser, and A. Jüngel, Adaptive Variable-Order Spherical Harmonics Expansion of the Boltzmann Transport Equation, *Proceedings of SISPAD*, p. 151-155 (2011).

(8) K. Rupp, T. Grasser, and A. Jüngel, Parallel Preconditioning for Spherical Harmonics Expansions of the Boltzmann Transport Equation, *Proceedings of SISPAD*, p. 147-150 (2011).

(9) ViennaSHE. URL: `http://viennashe.sourceforge.net/`

(10) ViennaCL. URL: `http://viennacl.sourceforge.net/`

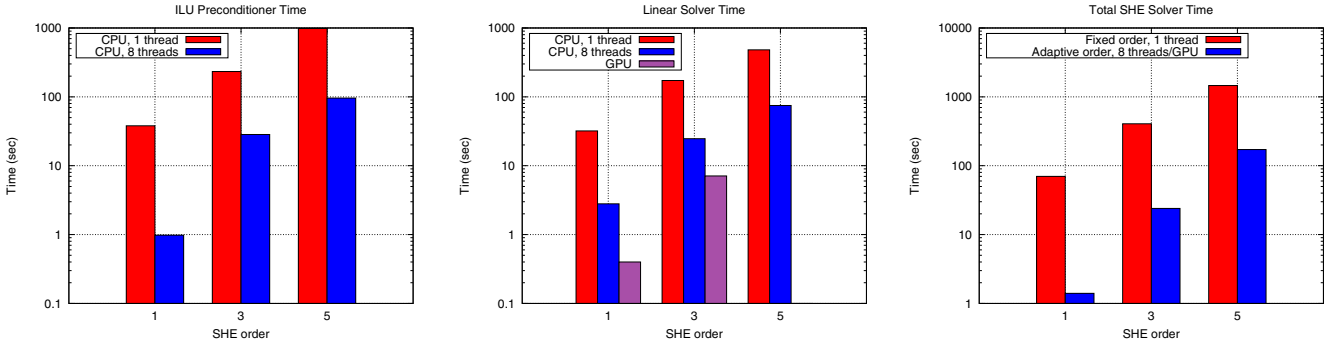
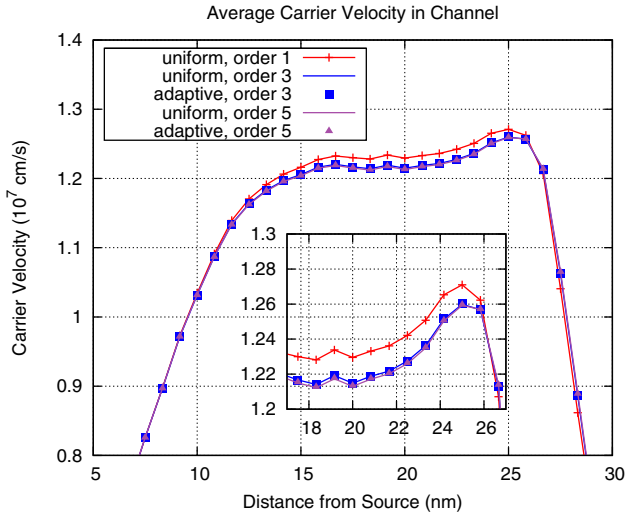
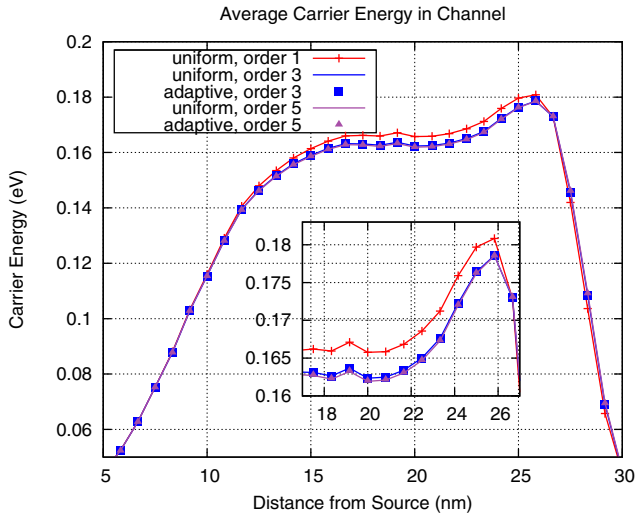


Figure 5: Comparison of execution times for the preconditioner setup (left), the linear solver run (center) and the gain of the complete SHE solver when using adaptive expansion orders. Benchmarks were carried out on a desktop machine equipped with a Intel Core i7 960 CPU, 12 GB RAM and a NVIDIA Geforce GTX 580 (out of RAM at order 5) on a coarse mesh in order to compare with high uniform expansion orders.

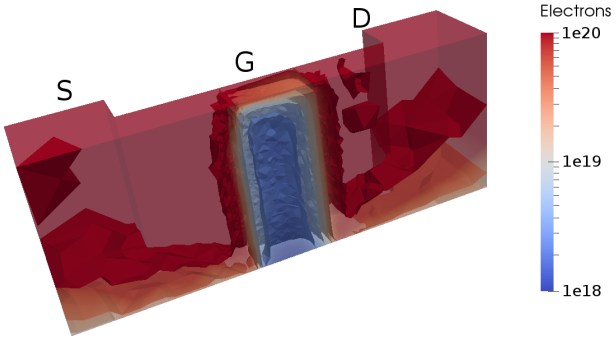


(a) Average carrier velocities.

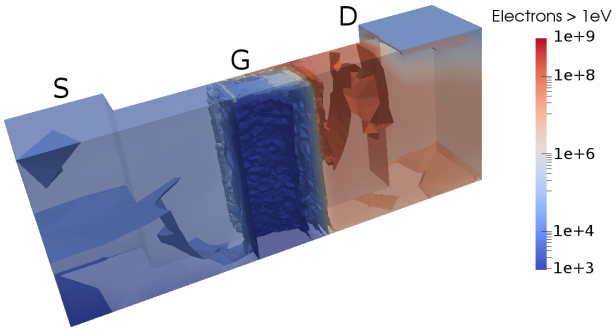


(b) Average carrier energies.

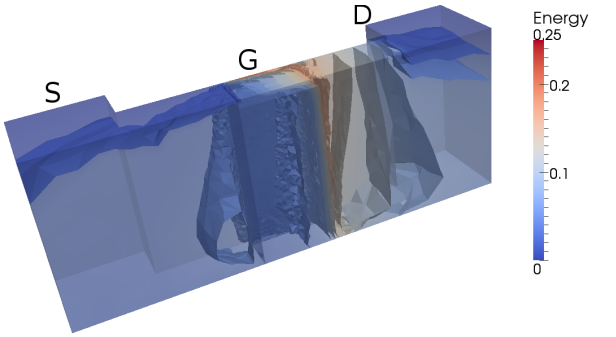
Figure 4: Average carrier velocities and energies for different uniform and adaptive SHE orders. Adaptive and uniform expansions yield identical results.



(a) Carrier density (cm^{-3}).



(b) Density of carriers with kinetic energy above 1eV (cm^{-3}).



(c) Average carrier energy (eV).

Figure 6: Carrier densities and average carrier energy in the simulated tri-gate transistor. The expected increase of carrier energy in the channel as well as the presence of heated carriers towards the drain contact is clearly visible.