Deterministic Numerical Solution of the Boltzmann Transport Equation

Karl Rupp, Tibor Grasser and Ansgar Jüngel

Abstract Due to its deterministic nature, the spherical harmonics expansion method is an attractive alternative to the Monte Carlo method for the solution of the Boltzmann Transport Equation for the purpose of electronic device simulation. However, since the problem is posed in a six-dimensional problem space emerging from the three-dimensional space variable and the three-dimensional momentum variable, deterministic approaches typically suffer from huge memory requirements, which have prohibited their application to two and three-dimensional simulations. To reduce these high memory requirements, we first show that the coupling of the resulting system of partial differential equations is only weak and then propose a new scheme for the lossless compression of the resulting system of linear equations after discretization. This reduces the overall memory requirements significantly and paves the way for deterministic three-dimensional device simulations. Numerical experiments demonstrate the applicability of our method and confirm our theoretical results.

Karl Rupp
Christian Doppler Laboratory for Reliability Issues in Microelectronics at the Institute for Microelectronics, TU Wien, Gußhausstraße 27–29/E360, A-1040 Wien, Austria, e-mail: rupp@iue.tuwien.ac.at

Tibor Grasser
Institute for Microelectronics, TU Wien, Gußhausstraße 27–29/E360, A-1040 Wien, Austria, e-mail: grasser@iue.tuwien.ac.at

Ansgar Jüngel
Institute for Analysis and Scientific Computing, TU Wien, Wiedner Hauptstraße 8-10/E101, A-1040 Wien, Austria, e-mail: juengel@asc.tuwien.ac.at

1 Introduction

As long as quantum mechanical effects are negligible, the microscopic behavior of charge carriers in semiconductors is very well described by a distribution function $f(x,k,t)$ that depends on the spatial coordinate x, the momentum $\hbar k$ and time t, and fulfills the Boltzmann transport equation (BTE). The most commonly used method to solve the BTE is the Monte Carlo method, with the main disadvantage of its computational expense, especially when attempting to reduce the statistical noise in the low density tails of the distribution function. The most prominent alternative to the stochastic Monte Carlo (MC) method is the deterministic spherical harmonics expansion (SHE) method [1]. Recent results demonstrate that higher order expansions, e.g. of order nine, result in excellent agreement with MC simulations, while maintaining the performance benefit [3, 4].

The major challenge of the SHE method is the huge memory consumption reported even for two-dimensional devices at moderate expansion orders [3], which has so far prohibited an application of the SHE method to three-dimensional simulations. To overcome these limitations, we present a new system matrix compression scheme that reduces the memory requirements by up to two orders of magnitude and paves the way for three-dimensional device simulations using the SHE method.

2 The Projected Equations

After a truncated expansion of the distribution function into real-valued, orthonormal spherical harmonics $Y_{l,m}(\theta,\varphi)$ up to order L,

$$f(x,\varepsilon,\theta,\varphi,t) \approx \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{l,m}(x,\varepsilon,t) Y_{l,m}(\theta,\varphi) \,, \tag{1}$$

a spherical projection of the BTE and application of the H-transform [2], one obtains with Einstein's summation convention a system of coupled partial differential equations with shifted arguments [3]

$$\begin{aligned} &\frac{\partial f_{l,m} Z}{\partial t} + \nabla_x \cdot v_{l,m}^{l',m'} f_{l',m'} Z - F \cdot \Gamma_{l,m}^{l',m'} f_{l',m'} Z \\ &\quad = \sum_{\eta} s_{l,m}^{l',m';\mathrm{in}} f_{l',m'}(x, \varepsilon \mp \hbar\omega_\eta, t) Z(x, \varepsilon \mp \hbar\omega_\eta, t) - s_{l,m}^{l',m';\mathrm{out}} f_{l',m'} Z \end{aligned}$$

for all $l \in \{0,\dots,L\}$, $m \in \{-l,\dots,l\}$. The generalized density of states Z depends on the band structure, F is the force and $s_{l,m}^{l',m';\mathrm{in}}$ and $s_{l,m}^{l',m';\mathrm{out}}$ denote the in- and out-scattering coefficients. Function arguments are suppressed wherever appropriate to increase readability.

If all coupling coefficients $v_{l,m}^{l',m'}$, $\Gamma_{l,m}^{l',m'}$, $s_{l,m;\eta}^{l',m';\text{in}}$ and $s_{l,m}^{l',m';\text{out}}$ were multiples of the Kronecker delta $\delta_{l,l'}\delta_{m,m'}$, all equations would be decoupled and could be solved individually. Conversely, nonzero coupling coefficients for all quadruples (l,m,l',m') indicate a tight coupling, which usually complicates the solution process. This is in analogy to systems of linear equations: If the system matrix is diagonal, the solution is found immediately, but if the matrix is dense, typically a lot of computational effort is required to solve the system.

It has been shown in [4] that the scattering terms $s_{l,m}^{l',m';\text{in}}$ and $s_{l,m}^{l',m';\text{out}}$ do not couple different expansion coefficients in the case of spherical energy bands. Moreover, the symmetry of the underlying processes yields for general band structures that

$$v_{2i,m}^{2i',m'} = v_{2i+1,m}^{2i'+1,m'} = 0, \quad \Gamma_{2i,m}^{2i',m'} = \Gamma_{2i+1,m}^{2i'+1,m'} = 0\ .$$

for all permissible integers i, i' and m, m' holds [4]. Therefore, all nonzero coupling coefficients possess different parities in the leading indices. This structural information about the coupling was already used in a preprocessing step for the solution of the discretized equations in [4].

Under the assumption of spherical energy bands, i.e. $\varepsilon(k) = \tilde{\varepsilon}(|k|)$, the velocity v, the modulus of the wave vector $|k|$ and the generalized density of states only depend on the energy ε, but not on the angles θ, φ. With this it can now be shown that the coupling induced by $v_{l,m}^{l',m'}$ and $\Gamma_{l,m}^{l',m'}$ is weak:

Theorem 1. *Under the assumption of spherical energy bands, the following holds true for indices $l,l' \in \{0,\dots,L\}$, $m \in \{-l,\dots,l\}$ and $m' \in \{-l',\dots,l'\}$:*

1. *If $v_{l,m}^{l',m'}$ is nonzero, then $l \in \{l' \pm 1\}$ and $m \in \{\pm|m'| \pm 1, m'\}$.*
2. *If $\Gamma_{l,m}^{l',m'}$ is nonzero, then $l \in \{l' \pm 1\}$ and $m \in \{\pm|m'| \pm 1, m'\}$.*

A proof is given in [5]. The theorem allows one to better eliminate those coefficients $v_{l,m}^{l',m'}$ and $\Gamma_{l,m}^{l',m'}$, which may not vanish in simulations due to numerical noise, even though they are analytically zero.

3 Discretization and System Matrix Compression

In steady state, a discretization of the expansion coefficients on a staggered grid (cf. [6]) with N grid points leads to a system of linear equations represented by a system matrix S of size $N(L+1)^2 \times N(L+1)^2$. The sparsity of S is a direct consequence of the finite difference or finite volume schemes used. Using the results of Theorem 1, it can be shown [5] that the resulting system matrix S can be written as

$$S = \sum_{i=1}^{8} Q_i \otimes R_i \tag{2}$$

where $\otimes$ denotes the Kronecker product. The Q_i only depend on the discretization in the (x,H)-space, while the R_i are determined by the coupling among spherical harmonics up to order L only. This allows for a representation of S using only $\mathcal{O}((L+1)^2+C_{\text{sparse}}N)$ numbers. Since N is typically much larger than $(L+1)^2$, the full system matrix can be stored for $C_{\text{sparse}} = 10$ with roughly $80N$ numbers, which means a reduction of two orders of magnitude compared to the uncompressed case.

4 Solution of the Linear System

The matrix compression scheme is of use only if the resulting system of linear equations can be solved without the need to recover the full matrix again. Such a reconstruction is, in principle, necessary if direct solvers are used, because during the solution process the matrix structure is altered in a way that destroys the block structure. For many popular iterative solvers from the family of Krylov methods, it is usually sufficient to provide matrix-vector multiplications [7].

Matrix-vector products for a matrix given as a Kronecker product of two smaller matrices can be carried out in a straightforward manner by decomposing the vector into blocks of suitable size. This allows for the realization of a very memory efficient matrix-vector multiplication if the system matrix is given in the form (2). However, numerical experiments indicate that the full system matrix for the even and odd order expansions coefficients is ill-conditioned. A substantial improvement of the system matrix condition number can be obtained if the unknowns for the odd order expansion coefficients are eliminated in a preprocessing step. However, a direct elimination by altering the system matrix is not possible without destroying the Kronecker product structure. This can be avoided by using the Schur complement, for which we refer the reader to [5]. Thus, the system matrix compression scheme can also benefit from the improved system matrix condition number after elimination of the odd order unknowns.

5 Results

We have compared memory requirements for the storage of the system matrix at several expansion orders in a two-dimensional device simulation. The results in Fig. 1 clearly demonstrate the asymptotic advantage of our approach: Already at an expansion order of $L = 5$, memory savings by a factor of 18 are observed, which increases to 146 at $L = 13$. With the compressed scheme, the memory required for the system matrix increases only by a few kilobytes as L increases, which is negligible.

Since the memory required by the system matrix is of order $\mathcal{O}(N+L^2)$ and the memory for the unknowns is of order $\mathcal{O}(NL^2)$, the memory required for the unknowns is much larger than the memory required for the representation of the system matrix for large values of L, cf. Fig. 2. Therefore, the asymptotic memory require-

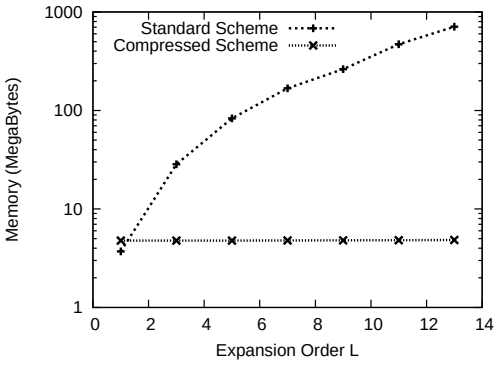

Fig. 1 Memory used for the uncompressed and the compressed system matrix for different expansion orders L on a three-dimensional grid with 12 500 nodes.

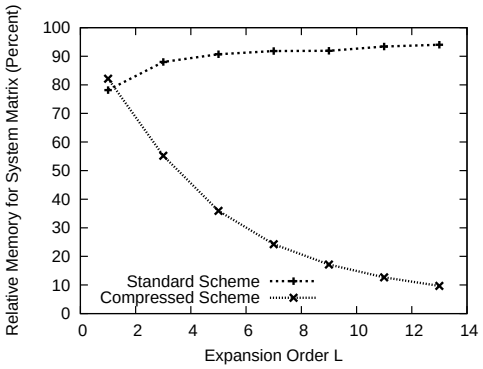

Fig. 2 Memory used for the system matrix in relation to the total amount of memory used (i.e. system matrix, unknowns and right hand side).

ments for the full simulation is still $\mathcal{O}(NL^2)$, but the constant of proportionality is of order one, while for the full system matrix it is around $11C_{\text{sparse}} \approx 100$, so a reduction of memory requirements by two orders of magnitude is obtained.

On a single CPU core, the minor price to pay for the dramatic reduction in memory consumption is that the execution times of matrix-vector products with the Schur complement increase by a factor of about two, cf. Fig. 3. However, the proposed matrix compression scheme is very well suited for parallel architectures, because the data required for the system matrix may even fit into the CPU caches, allowing for a very high performance. Moreover, since the system matrix can be written as a sum of Kronecker products, each summand can be computed on a separate core.

6 Conclusion

The matrix compression scheme presented in this work reduces the memory requirements for the system matrix arising from a SHE of the BTE from order $\mathcal{O}(NL^2)$ to $\mathcal{O}(N+L^2)$, which results in total memory savings for the full simulation run by up to two orders of magnitude. Therefore, our scheme paves the way for three-

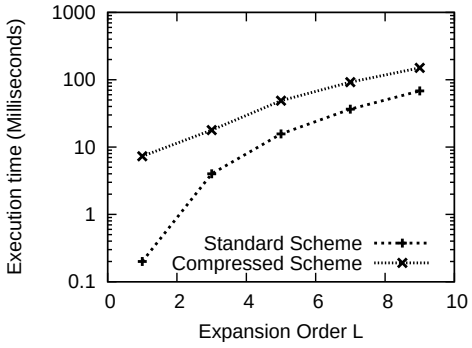

Fig. 3 Execution times for matrix-vector multiplication with the Schur complement on a single core of an Intel Core 2 Quad Q9550 CPU.

dimensional device simulations especially for larger expansion orders L. On a single CPU core, the small price to pay is a runtime penalty on matrix-vector multiplication of about a factor of two. However, the proposed method is expected to outperform the traditional storage scheme on parallel architectures.

Acknowledgments

Support by the Graduate School PDEtech at the TU Wien is gratefully acknowledged. A. Jüngel acknowledges partial support from the Austrian Science Fund (FWF), grants P20214, I395, and WK "Differential Equations", and the Austrian Exchange Service (ÖAD), grants ES 08/2010, FR 07/2010, HR 10/2010.

References

1. N. Goldsman et. al., 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 et. al., 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. S. M. Hong and C. Jungemann, A Fully Coupled Scheme for a Boltzmann-Poisson Equation Solver based on a Spherical Harmonics Expansion. *J. Comp. Elec.*, vol. 8, p. 225–241 (2009).
4. C. Jungemann et. al., Stable Discretization of the Boltzmann Equation based on Spherical Harmonics, Box Integration, and a Maximum Entropy Dissipation Principle. *J. Appl. Phys.*, vol. 100, no. 2, p. 024502-+ (2006).
5. K. Rupp et. al., Matrix Compression for Spherical Harmonics Expansions of the Boltzmann Transport Equation for Semiconductors. *J. Comput. Phys.*, vol. 229, p. 8750–8765, (2010).
6. C. Ringhofer, Numerical Methods for the Semiconductor Boltzmann Equation Based on Spherical Harmonics Expansions and Entropy Discretizations. *Transport Theory and Statistical Physics*, vol. 31, p. 431–452 (2002).
7. Y. Saad and M. H. Schultz, GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM J. Sci. Stat. Comput.*, vol. 7, no. 3, p. 856–869 (1986).