A Matrix Compression Scheme for Spherical Harmonics Expansions of the Boltzmann Transport Equation

Karl Rupp[1], Tibor Grasser[1] and Ansgar Jüngel[2] (Faculty Mentor)
[1]Institute for Microelectronics, [2]Institute for Analysis and Scientific Computing
Vienna University of Technology
Vienna, Austria
Email: {rupp,grasser}@iue.tuwien.ac.at, juengel@asc.tuwien.ac.at

Abstract — *We investigate the numerics behind the numerical solution of the Boltzmann transport equation using an expansion of the distribution function in spherical harmonics for the purpose of electronic device simulation. Traditional implementations of higher order spherical harmonics expansions suffer from huge memory requirements especially for two and three dimensional devices. To overcome these complexity limitations, a compressed matrix storage scheme based on Kronecker products is proposed, which reduces the memory requirements for the storage of the system matrix such that the total memory requirements are asymptotically dominated by the memory required for the unknowns.*

I. INTRODUCTION

While in the early years of the semiconductor industry macroscopic models have been sufficient for device simulation, this is not the case anymore for the small feature sizes used today. As long as quantum mechanical effects are not dominant, the microscopic behavior of electrons is best described by a distribution function $f(\boldsymbol{x}, \boldsymbol{k}, t)$ that depends on the spatial coordinate $\boldsymbol{x} = (x, y, z)$, the wave vector $\boldsymbol{k} = (k_x, k_y, k_z)$ 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 method is the deterministic spherical harmonics expansion (SHE) method.

The major challenge of the SHE method is the huge memory consumption reported even for two-dimensional devices [1] at moderate expansion orders, which has so far prohibited an application of the SHE method to three-dimensional devices. To overcome these limitations, we present a new system matrix compression scheme that reduces the memory requirements by orders of magnitude and paves the way for three-dimensional device simulations using the SHE method.

II. THE PROJECTED EQUATIONS

Instead of an expansion of the electron distribution function into spherical harmonics, it is for reasons of numerical stability of advantage to expand the generalized energy distribution function [2]

$$g(\boldsymbol{x}, \varepsilon, \theta, \varphi, t) = 2Z(\varepsilon, \theta, \varphi)f^\nu(\boldsymbol{x}, \boldsymbol{k}(\varepsilon, \theta, \varphi), t) \,,$$

with the generalized density of states Z, into orthonormal and real valued spherical harmonics $Y_{l,m}(\theta, \varphi)$ and truncate the series after $(L + 1)^2$ terms:

$$g(\boldsymbol{x}, \varepsilon, \theta, \varphi, t) \approx \sum_{l=0}^{L} \sum_{m=-l}^{l} g_{l,m}(\boldsymbol{x}, \varepsilon, t)Y_{l,m}(\theta, \varphi) \,.$$

Partial differential equations for the coefficients $g_{l,m}$ are directly obtained from projections of the BTE, which results (using Einstein's summation convention) in

$$\begin{aligned}
\frac{\partial g_{l,m}}{\partial t} &+ \frac{\partial \boldsymbol{F} \cdot \boldsymbol{v}_{l,m}^{l',m'} g_{l',m'}}{\partial \varepsilon} \\
&+ \boldsymbol{v}_{l,m}^{l',m'} \cdot \nabla_{\boldsymbol{x}} g_{l',m'} - \boldsymbol{F} \cdot \boldsymbol{\Gamma}_{l,m}^{l',m'} g_{l',m'} \\
&= s_{l,m}^{l',m';\mathrm{in}} g_{l',m'}(\boldsymbol{x}, \varepsilon \mp \hbar\omega_\eta, t) \\
&\quad - s_{l,m}^{l',m';\mathrm{out}} g_{l',m'} \,.
\end{aligned}$$

for all $l = 0, \ldots, L$ and $m = -l, \ldots, l$.

It has been shown that the scattering terms $s_{l,m}^{l',m';\mathrm{in}}$ and $s_{l,m}^{l',m';\mathrm{out}}$ do not couple different expansion coefficients in the case of spherical energy bands [2]. We show that the coupling by the velocity terms $\boldsymbol{v}_{l,m}^{l',m'}$ and the angular coupling terms $\boldsymbol{\Gamma}_{l,m}^{l',m'}$ is again sparse:

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

1. *If $\boldsymbol{v}_{l,m}^{l',m'}$ is nonzero, then $l \in \{l' \pm 1\}$ and $m \in \{\pm|m'| \pm 1, m'\}$.*

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

III. DISCRETIZATION AND SYSTEM MATRIX COMPRESSION

In steady state, a discretization of the expansion coefficients of the generalized distribution function is obtained
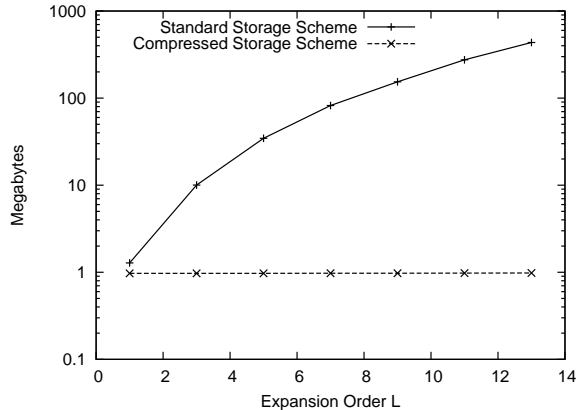
Figure 1: Memory used for the uncompressed and the compressed system matrix for different expansion orders L on a two-dimensional $(\boldsymbol{x}, \varepsilon)$-grid with 10.000 nodes.

Figure 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).

by a Galerkin method

$$g_{l,m} = \sum_{i=1}^{N} \alpha_{i;l,m}(t)\varphi_i(\boldsymbol{x}, \varepsilon) \,,$$

finally resulting in a system matrix $\boldsymbol{S}$ of size $N(L+1)^2 \times N(L+1)^2$ with a-priori $C_{\mathrm{sparse}}N(L+1)^4$ entries, where C_{sparse} is a constant that depends only on the regularity of the underlying mesh. With Theorem 1, we have shown that the number of entries in each row of $\boldsymbol{S}$ is at most $11C_{\mathrm{sparse}}$. Consequently, there are at most $11C_{\mathrm{sparse}}N(L+1)^2$ nonzero entries in $\boldsymbol{S}$. With the typical values $L = 9$ and $C_{\mathrm{sparse}} = 10$, the estimate becomes $11000N$, which still prohibits sufficiently fine discretizations for the simulation of three-dimensional devices.

We suggest a method to decouple the spherical harmonics expansion coefficients from the spatial discretization, such that the system matrix $\boldsymbol{S}$ can be written as

$$\boldsymbol{S} = \sum_{i=1}^{8} \boldsymbol{Q}_i \otimes \boldsymbol{R}_i \,, \qquad (1)$$

where $\otimes$ denotes the Kronecker product. The matrices $\boldsymbol{Q}_i$ are of size $N \times N$ and the matrices $\boldsymbol{R}_i$ are of size $(L+1)^2 \times (L+1)^2$ for $i = 1, \ldots, 8$. This allows a representation of $\boldsymbol{S}$ using only $32(L+1)^2 + 8C_{\mathrm{sparse}}N$ numbers. Since N is typically much larger than $(L+1)^2$, the full system matrix can for $C_{\mathrm{sparse}} = 10$ be stored with roughly $80N$ numbers, which means a reduction by a factor 137.5 compared to the uncompressed case.

IV. RESULTS

We have compared memory requirements for the storage of the system matrix at several expansion orders in a one-dimensional device simulation. The results in Figure 1 clearly demonstrate the asymptotic superiority of our approach: Already at an expansion order of $L = 5$, memory
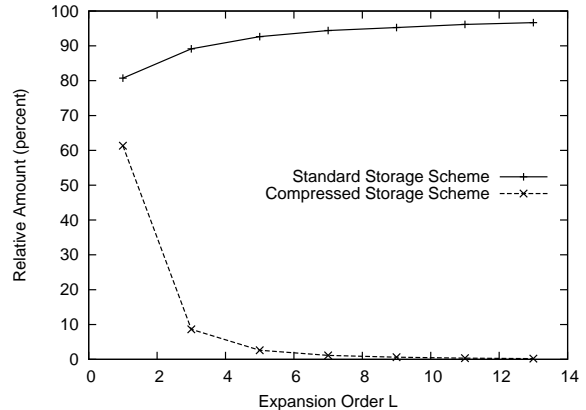
savings by a factor of 35 are observed, which increases to 442 at $L = 13$. Moreover, this leads to the situation that the memory required for the unknowns is much larger than the memory required for the representation of the system matrix, cf. Figure 2.

V. CONCLUSION

We have investigated the coupling structure of the SHE equations and shown a weak coupling of the expansion coefficients. This guarantees that the nonzero entries in the system matrix obtained from a discretization with N degrees of freedom in $(\boldsymbol{x}, \varepsilon)$-space and SHE order L are at most $11C_{\mathrm{sparse}}N(L+1)^2$ in contrast to $C_{\mathrm{sparse}}N(L+1)^4$ for the case of a dense coupling.

The proposed matrix compression scheme further reduces the memory requirements for the system matrix to $32(L+1)^2 + 8C_{\mathrm{sparse}}N \approx 8C_{\mathrm{sparse}}N$. While the huge memory requirements for the storage of the full system matrix prohibited the simulation of three-dimensional devices so far, our proposed scheme paves the way for such simulations especially for larger expansion orders L.

ACKNOWLEDGMENTS

REFERENCES

[1] 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*, 8:225–241, 2009.

[2] 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*, 100(2):024502–+, July 2006.