A Wavelet Method to Solve High-dimensional Transport Equations in Semiconductor Devices

SISPAD 2011, Osaka, Japan
Vincent Peikert and Andreas Schenk,
Integrated Systems Laboratory, ETH Zurich, Switzerland
Multi-Wavelet Basis

- Our goal: to solve the Boltzmann and Wigner transport equations in 2 and 3 dimensional devices numerically (5 + dimensional phase spaces).

- Problem: exponential increase of the computational cost with dimension. Example: 100 coefficients per dimension result in $100^6 = 10^{12}$ coefficients.

- Status quo:
  - Even the accurate simulation of simple 1-dimensional structures (with 3-dimensional phase spaces) are computationally very expensive (e.g. Jungemann 2006: 667,660 DOF, Y.Cheng 2009: 1,036,800 DOF).
  - Simple 2-dimensional devices (with a 5-dimensional phase space) can only be simulated with very coarse meshes. Even conventional adaptive meshing methods could not cope with the “curse of dimensionality” of phase spaces with more than 4 dimensions.

- Our idea: Hierarchical multiscale compression techniques could cope with the curse of dimensionality in 6-dimensions.

- Our special suggestion as a basis for semiconductor transport equations: “Multi-Wavelets”.
Multi-Wavelet Basis

- MWs build a hierarchical orthonormal basis for piecewise polynomials
- MWs enable hp-wavelet adaptivity, -compression, -preconditioning and matrix compression techniques.
- MWs offer the maximum possible number of vanishing moments (high compression rate).
- Legendre polynomials are special cases of MWs (the popular Legendre expansion belongs to pure MW p-refinement).
- Piecewise constant functions are special cases of MWs (Finite Volume, WENO, SH-expansion in energy-real space belong to pure MW h-refinement).
- Contemporary particle-conserving weak formulations like SH expansion or Discontinuous Galerkin method can directly be combined with MWs.
### One-dimensional Multi-Wavelet Basis

<table>
<thead>
<tr>
<th></th>
<th>h=0</th>
<th>h=1</th>
<th>h=2</th>
</tr>
</thead>
<tbody>
<tr>
<td>p=0</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
</tr>
<tr>
<td>p=1</td>
<td><img src="image4" alt="Graph" /></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>p=2</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
<td><img src="image9" alt="Graph" /></td>
</tr>
<tr>
<td>p=3</td>
<td><img src="image10" alt="Graph" /></td>
<td><img src="image11" alt="Graph" /></td>
<td><img src="image12" alt="Graph" /></td>
</tr>
</tbody>
</table>
One-dimensional Adaptivity with Haar-Wavelets

\[ \Phi(x) = \sum_{i=0}^{\infty} \alpha_i \psi_i(x) \]
\[ \alpha_i = \int_{\text{supp}(\psi_i)} \Phi(x) \psi_i(x) \, dx \]
\[ |\Phi(x)|^2 = \sum_{i=0}^{\infty} |\alpha_i|^2 \]

Remove smallest, refine largest!
Multi-dimensional Haar-Wavelet Basis
Multi-dimensional Adaptivity with Haar Wavelets

\[
\begin{align*}
\psi(0,0) & \quad \psi(1,0) & \quad \psi(0,1) & \quad \psi(1,1) \\
\psi(0,0) - \psi(1,0) &= \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \\
\psi(0,1) - 0.25 \times \psi(0,0) &= \begin{bmatrix} 0.75 & 1.25 \\ -1.25 & -0.75 \end{bmatrix} \\
-0.25 \times \psi(1,0) - \psi(0,1) &= \begin{bmatrix} -0.75 & 1.25 \\ -1.25 & 0.75 \end{bmatrix} \\
\psi(1,0) - \psi(1,1) &= \begin{bmatrix} 1.25 & 0.75 \\ -1.25 & -0.75 \end{bmatrix}
\end{align*}
\]
The “Multiwavelet Discontinuous Galerkin Method” (for the Boltzmann Transport Equation)

- BTE is a conservation law:
  \[ \nabla (x, \mu, \omega) \cdot [\tilde{g}(x, \mu, \omega) \Phi] = C(\Phi) \]

- Weak formulation, Greens theorem:
  \[ \int_{\partial \Omega} \bar{n} \cdot [\tilde{g}(x, \mu, \omega) \Phi] d\sigma = \int_{\Omega} C(\Phi) v d\Omega \quad \forall v(x, \mu, \omega) \in L^2 \]

- Cheng 2009: Piecewise polynomials as basis functions. → DG

- We: Multiwavelets as basis functions and apply wavelet adaptivity/compression. → MWDG

\[ \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \beta_j \int_{\partial \Omega} \bar{n} \cdot \tilde{g}(x, \mu, \omega) \psi_i \psi_j d\sigma = \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \beta_j \int_{\Omega} C(\psi_i) \psi_j d\Omega \]

**Without compression:** MWDG equivalent to DG.
- Unique and local conservative.

**With compression:** MWDG **not local conservative** in general (but global conservative.)
- Does it work anyways?
- Conservation violation is controllable. New unique adaptivity criterion?

*1. In parallel to us one preprint (without simulation figures) from 03/2010 exists that mixes DG and wavelets (Archibald, Oakridge National Laboratory).
2. Schwab (ETH) is the first who applied wavelets to a transport equation (radiative transfer) with a least-squares ansatz in 2008.
NIN Device Simulation at 1 V Drain
NIN 1V Simulation Conclusion
Conclusions

- Multi-Wavelets have been proposed as a new basis for the numerical solution of high-dimensional semiconductor transport equations.
- Multi-Wavelets can span the basis of the most important contemporary Boltzmann solvers but additionally have multi-dimensional wavelet properties which could cope with the “curse of dimensionality”.
- Multi-Wavelets have been combined with a Discontinuous Galerkin formulation of the Boltzmann equation.
- NIN simulations in a 3-dimensional phase space with first order Multi-Wavelets (Haar Wavelets) show that the “Multiwavelet Discontinuous Galerking method” is robust under wavelet compression.
- The compression rate in the NIN simulation is up tp 99% compared to a “manual adaptive” nodal Discontinuous Galerkin approach.
- Due to more vanishing moments, even higher compression rates are expected for higher order Multi-Wavelets...
Thank you!
NIN Density at 1 V Drain

1e+17
1e+16
1.8e+16
1.6e+16
1.4e+16
1.2e+16
0.4
0.5
0.55
0.6
0.65
0.7
x [10^{-6} m]
Density [cm^{-2}]

1e+17
1e+16
1.8e+16
1.6e+16
1.4e+16
1.2e+16
0
0.2
0.4
0.6
0.8
1
x [10^{-6} m]
Density [cm^{-2}]

- Nodal DG Optimum (1.147,000)
- Wavelet L2 Compression (17,000)
- Wavelet DG L2 Adaptive (17,000)
- Nodal DG Optimum (111,872)
- Nodal DG Optimum (52,800)
NIN Velocity at 1 V Drain
NIN Energy at 1 V Drain

Graphs showing energy distribution with different compression methods and nodal DG optimisations.
NIN Current at 1 V Drain
NIN Current at 1 V Drain
NIN Density at 1 V Drain (How Far Can We Go?)
NIN Velocity at 1 V Drain (How Far Can We Go?)
NIN Energy at 1 V Drain (How Far Can We Go?)

Graphs showing energy vs. distance for different compression algorithms.
NDG Solutions, Different Refinements, 1 V, x = 0.5

~1M:

~100K:

~50K:
NDG Solutions, Different Refinements, 1 V, x = 0.6

~1M:

~100K:

~50K:
NDG Solutions, Different Refinements, 1 V, x = 0.7

~1M:

~100K: TOOK WRON

~50K:
NDG Solutions, Different Refinements, $1 \ V, \ x = 0.8$

~1M:
Wrong
Axis: $P$

~100K:

~50K:
NDG vs. Compressed MWDG Solutions, 1 V, x = 0.5

NDG:

MDG:

Errors:
NDG vs. Compressed MWDG Solutions, 1 V, x = 0.6

NDG:

MDG:

Errors:
NDG vs. Compressed MWDG Solutions, 1 V, $x = 0.7$

NDG:

MDG:

Errors:
NDG vs. Compressed MWDG Solutions, 1 V, x = 0.8

NDG:

MDG:

Errors:
Future Plans

• This year: Write paper(s) about the Multiwavelet Discontinuous Galerkin Method and its applications for the Boltzmann Transport Equations (eventually split theoretical aspects of the MWDG method to a separate paper)

• Improve the algorithms and implementation to speed up the coefficient adding and removing from matrix and simulate more devices. Test iterative solvers. (1 month).

• Direct performance benchmarks with spherical harmonics solvers (Jungemann, Synopsis) (1 month)

• Software Architecture of an efficient 2D solver (2 weeks)

• Efficient Implementation of 2D solver (Eventually in Phython/C++) (1.5 month)

• Simulate 2D devices with less than 100,000 coefficients

• ...
NIN Device Simulation

**Graphs:**
- Electric Field at 1V
- Doping Concentration [cm⁻³]

**Table:**
<table>
<thead>
<tr>
<th>X [μm]</th>
<th>Y [μm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.03125</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0078125</td>
</tr>
<tr>
<td>0.2</td>
<td>0.00390625</td>
</tr>
<tr>
<td>0.3</td>
<td>0.0009765625</td>
</tr>
<tr>
<td>0.4</td>
<td>0.03125</td>
</tr>
<tr>
<td>0.5</td>
<td>0.015625</td>
</tr>
<tr>
<td>0.6</td>
<td>0.00390625</td>
</tr>
<tr>
<td>0.7</td>
<td>0.0078125</td>
</tr>
<tr>
<td>0.8</td>
<td>0.03125</td>
</tr>
<tr>
<td>0.9</td>
<td>0.0078125</td>
</tr>
<tr>
<td>1</td>
<td>0.03125</td>
</tr>
</tbody>
</table>

**Colors:**
- Doping Concentration:
  - 6.0E+17
  - 1.7E+17
  - 5.5E+16
  - 1.8E+16
  - 6.0E+15
  - 2.0E+15
The “Multiwavelet Discontinuous Galerkin Method” (for the Boltzmann Transport Equation)


\[
\vec{n}(x,\mu,\omega) \cdot \left[ \vec{g}(x,\mu,\omega) \Phi \right] = C(\Phi)
\]

\[
\int_{\partial\Omega} \vec{n} \cdot \left[ \vec{g}(x,\mu,\omega) \Phi \right] \nu d\sigma = \int_{\Omega} C(\Phi) \nu d\Omega \quad \forall \nu(x,\mu,\omega) \in L^2
\]
The “Multiwavelet Discontinuous Galerkin Method”
Adaptive Mesh Coarsing for NDG Simulations

>1M
236x38x128
= 1,147,904

>100K
92x22x64
= 129,536

>50K
66x16x50
= 52,800
Motivation: Data Compression Techniques for the Boltzmann Equation are Needed

Spherical Harmonics Expansion

- **175,000 to 667,660 functions needed to simulate NIN device** (5th to 19th order, Jungemann, 2006)
  - Not conservative in momentum space
  - Not accurate if far away from equilibrium (non-localized, pure p-refinement)
  - Full matrix in momentum space in drift term (problem in 2D devices)
  - No hierarchical adaptivity possible in real- and energy space.

Nodal Solvers (e.g. Discontinuous Galerkin, Finite Volume, WENO,...)

- **Over 1,036,800 functions needed to simulate NIN device** (Y. Cheng 06/2009)
  - Full matrix in momentum space in scattering term (problem in 2D devices)

“Multiwavelet Discontinuous Galerkin” Method (this work)

- **7000 functions needed to simulate NIN device**
  - No full matrix neither in scattering term nor in drift term
  - Hierarchical hp-adaptivity in the whole phase space possible
One-dimensional Multi-Wavelet Basis

<table>
<thead>
<tr>
<th></th>
<th>h=0</th>
<th>h=1</th>
<th>h=2</th>
</tr>
</thead>
<tbody>
<tr>
<td>p=0</td>
<td><img src="image1" alt="Graph" /></td>
<td><img src="image2" alt="Graph" /></td>
<td><img src="image3" alt="Graph" /></td>
</tr>
<tr>
<td>p=1</td>
<td><img src="image4" alt="Graph" /></td>
<td><img src="image5" alt="Graph" /></td>
<td><img src="image6" alt="Graph" /></td>
</tr>
<tr>
<td>p=2</td>
<td><img src="image7" alt="Graph" /></td>
<td><img src="image8" alt="Graph" /></td>
<td><img src="image9" alt="Graph" /></td>
</tr>
</tbody>
</table>
Multi-Wavelet Basis

- MWs build a hierarchical orthonormal basis for piecewise polynomials.
  \[ S_m^k = S_0^k \oplus W_0^k \oplus W_1^k \oplus \ldots \oplus W_{m-1}^k \]

- MWs enable hp-wavelet adaptivity, -compression, -preconditioning and matrix compression techniques.

- MWs offer the maximum possible number of vanishing moments (high compression rate).
  \[ \int \psi_{k,j}(x)x^i\,dx = 0, \quad i = 0, 1, \ldots, j + k - 2 \]

- Legendre polynomials are special cases of MWs (the popular Legendre expansion belongs to pure MW p-refinement).

- Piecewise constant functions are special cases of MWs (Finite Volume, WENO, SH-expansion in energy-real space belong to pure MW h-refinement).

- Contemporary particle-conserving weak formulations like SH expansion or Discontinuous Galerkin method can directly be combined with MWs.
Uncertainties At The Beginning

• Stability of the nodal DG method to solve the Boltzmann equation? (No oscillations, solutions always positive, convergence with refinement, ...)

• Performance: Does the nodal DG method converge fast enough or is an unacceptable high refinement necessary?

• What happens if I remove some wavelets from the transformed basis?

• Is the matrix still invertible?

• Does the solution completely break down or does it change smoothly?

• How does the solutions change (shifts?, fluctuations?)

• How do the momenta behave (density, velocity, energy,...)

• Compression: How many of the coefficients can we save?

• Adaptivity: Does the adaptive algorithm find the correct coefficients?