Semiconductor Device Simulation Approaches for Massively Parallel Computing Architectures

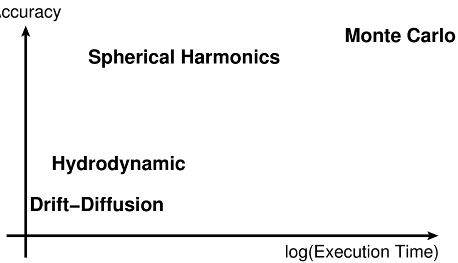
Karl Rupp[1,2], Tibor Grasser[2], Ansgar Jüngel[3]

[1] Freelance Computational Scientist
[2] Institute for Microelectronics
[3] Institute for Analysis and Scientific Computing
TU Wien, Austria

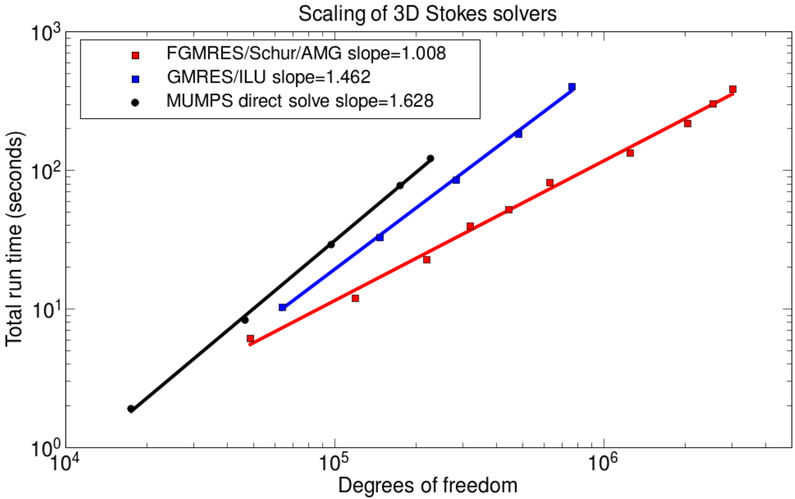
MS: Computational and Numerical Methods in Electronics
SIAM CSE, Atlanta, February 27th, 2017

Selection of Semiclassical Models

- Drift-diffusion model
- Hydrodynamic model
- Energy transport model
- ...



Scalability of Linear Solvers

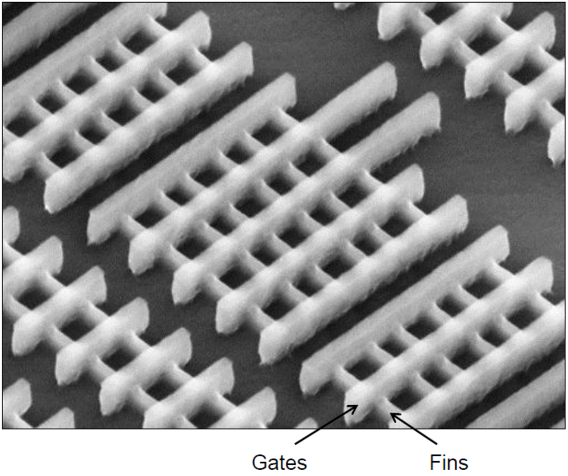


(courtesy of Jed Brown)

Contents

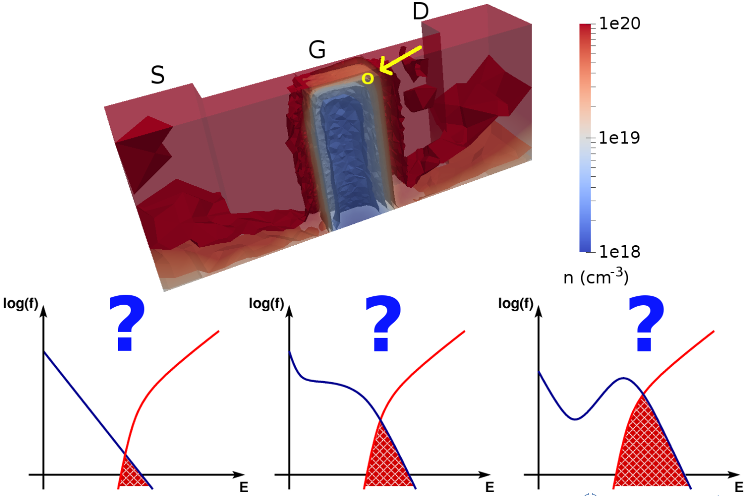
Semiconductor Devices in 3D: FinFET

Intel Trigate transistors

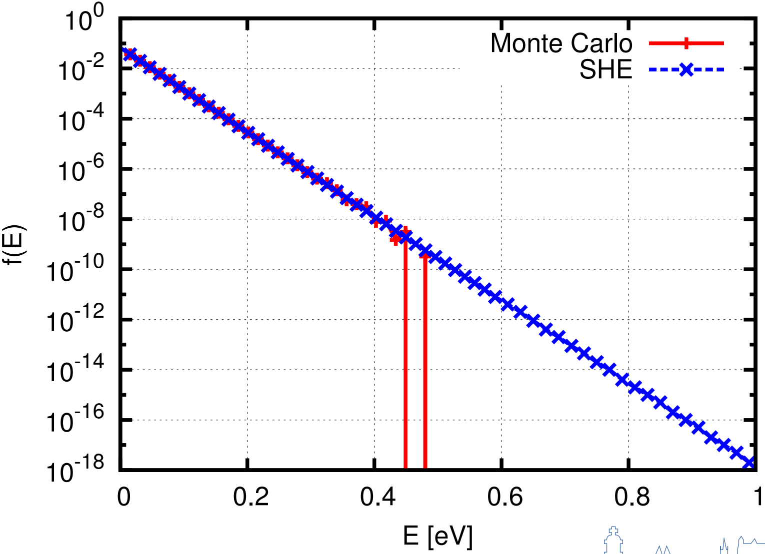


Electron Energy Distribution?

Distribution of electrons with respect to energy at x?



Electron Energy Distribution?

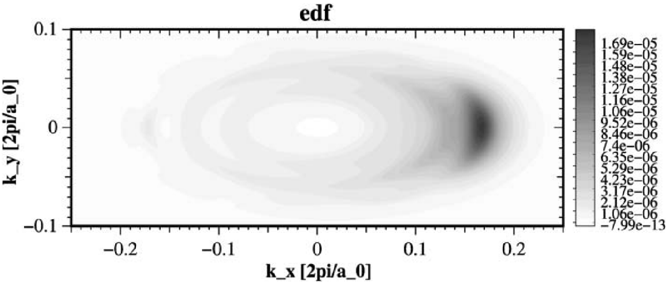


Spherical Harmonics Expansion Method

Spherical Symmetries

Maxwell distribution of carriers at equilibrium

Dispersion relation (Herring-Vogt transform, approx.)



S.-M. Hong and C. Jungemann (2009)

Spherical Harmonics Expansion Method

Spherical Symmetries

- Maxwell distribution of carriers at equilibrium
- Dispersion relation (Herring-Vogt transform, approx.)

Spherical Harmonics Expansion (SHE)

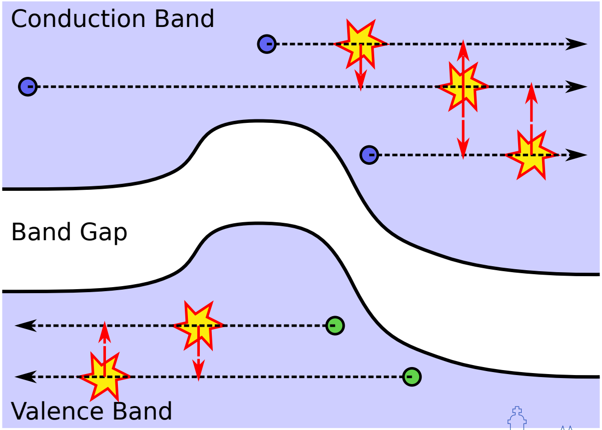
$$f(\boldsymbol{x}, \boldsymbol{k}, t) \simeq \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{l,m}(\boldsymbol{x}, E, t) Y_{l,m}(\theta, \varphi)$$

- New unknowns: $f_{l,m}(\boldsymbol{x}, E, t)$
- Solution in five-dimensional $(\boldsymbol{x}, E, t)$-space
- S.-M. Hong and C. Jungemann, *J Comput Electron* (2009):
  Fifth-order, three-dim. $(\boldsymbol{x}, E)$-space, 26 GB memory, 12 hours

Parallelization

Preconditioner for Iterative Linear Solvers

- No fast general-purpose parallel preconditioner available
- Physics-based parallel block preconditioner developed



Parallelization

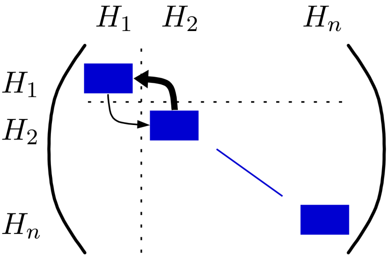
Scaling of Solution Variables

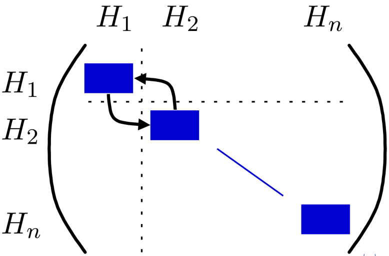
Exponential decay with energy: $f(E_i) \sim \exp(-\frac{E_i}{k_B T})$

Rescale unknowns: $\tilde{f}(E_i) = \exp(\frac{E_i}{k_B T}) f(E_i)$

New system: $\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{D}^{-1}\boldsymbol{x} = \boldsymbol{b}$

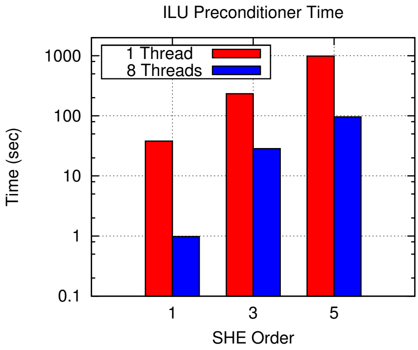
Row normalization: $\hat{\boldsymbol{A}}\tilde{\boldsymbol{x}} = \boldsymbol{P}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}} = \boldsymbol{P}\boldsymbol{b}$

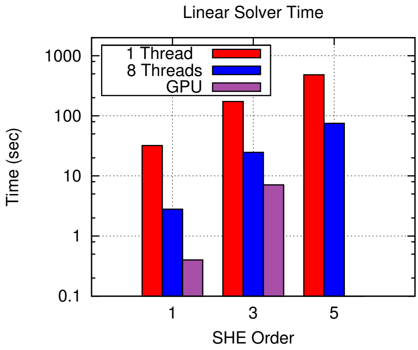




Parallelization

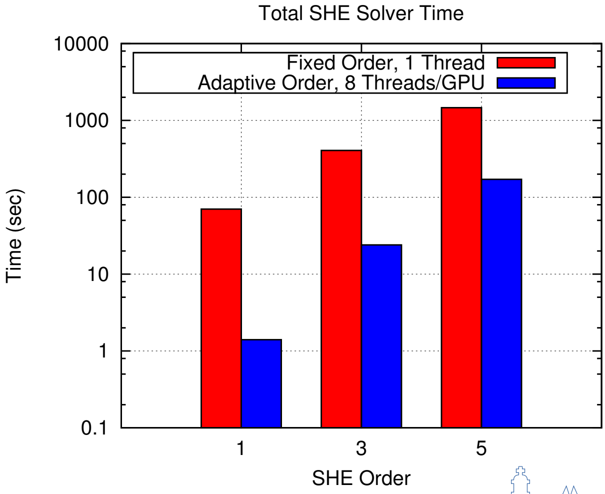
Benchmark results for a FinFET (INTEL Core i7 960, NVIDIA GTX 580)





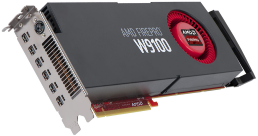
Parallelization

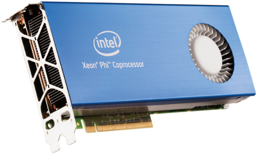
Benchmark results for a FinFET (INTEL Core i7 960, NVIDIA GTX 580)



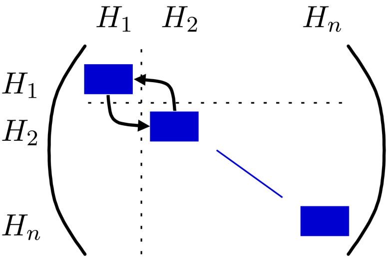
Part 2

Current Work:
Development of a Parallel Preconditioner
for (Heterogeneous) Distributed Memory Machines





Distributed SHE



Blueprint

- Keep block-Jacobi based on total energies
- Map total energies to MPI ranks
- Requires fine-grained parallel preconditioner per block

Parallel ILU

Sequential
```
for i=2..n
   for k=1..i−1, (i,k) in A
```
$a_{ik} = a_{ik}/a_{kk}$
```
      for j=k+1..n, (i,j) in A
```
$a_{ij} = a_{ij} - a_{ik}a_{kj}$

Parallel
```
for (sweep = 1, 2, ...)
   parallel for (i,j) in A
     if (i > j)
```
$l_{ij} = (a_{ij} - \sum_{k=1}^{j=1} l_{ik}u_{kj})/u_{jj}$
```
     else
```
$u_{ij} = a_{ij} - \sum_{k=1}^{=1} l_{ik}u_{kj}$

Fine-Grained Parallel ILU Setup

Proposed by Chow and Patel (SISC, vol. 37(2)) for CPUs and MICs

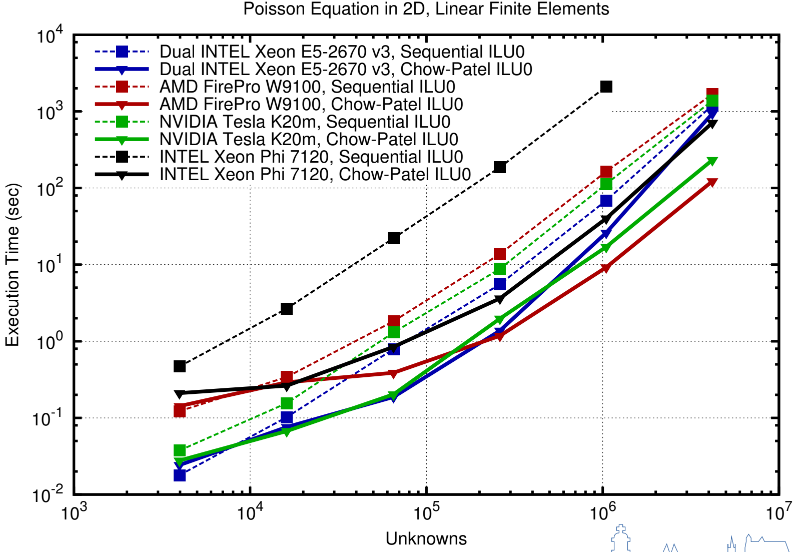
Massively parallel (one thread per row)

Preconditioner Application

Truncated Neumann series:

$$\mathbf{L}^{-1} \approx \sum_{k=0}^{K} (\mathbf{I} - \mathbf{L})^k, \quad \mathbf{U}^{-1} \approx \sum_{k=0}^{K} (\mathbf{I} - \mathbf{U})^k$$

Parallel ILU



Conclusion

SHE Method

- Viable alternative to Monte Carlo
- Full 3D device simulations possible
- Convergence behavior similar to drift-diffusion model
- Open source simulator: ViennaSHE

Large-Scale Solution

- Physics-based block-Jacobi preconditioner
- Replication of spatial mesh on all MPI ranks
- Fine-grained parallel ILU
- Combine functionality in PETSc and ViennaCL libraries