A Coupled Electrostatic-Quantum Transport Framework for Exa-scale Systems

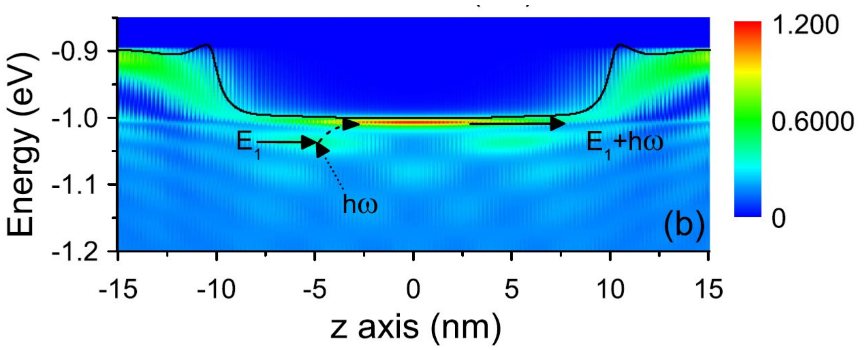
Saurabh Sawant*, François Léonard[†], Jackie Yao*, Andrew Nonaka*

***Center for Computational Sciences and Engineering**
Lawrence Berkeley National Laboratory

[†] **Nanoelectronics and Nanophotonics Group, Sandia National Laboratories**

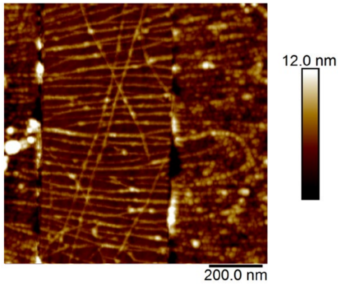
IWCN 2023
Session: Quantum Electron Transport II

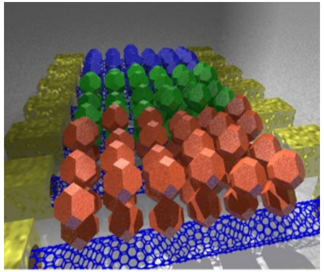
Quantum Transport Simulations Have Become More Sophisticated



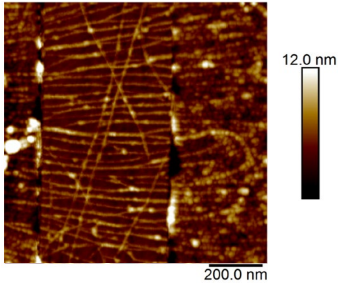
Ab initio
Scattering processes
External fields
Self-consistency
Time-dependent
Many-body interactions

3D architectures experimentally important, but difficult to simulate when no symmetry





Need for *Scalable* Tools for Self-Consistent Modeling of 3D Structures



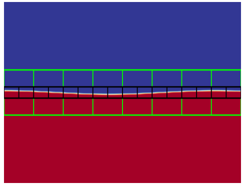
Electrostatics

$$\nabla \cdot (\epsilon \nabla U) = -\Delta \rho$$

Quantum Transport

Non-**E**quilibrium **G**reen's **F**unction method

AMReX



github.com/AMReX-Codes/amrex

GPU-capability Portability
High-scalability Opensource

AMR: Adaptive Mesh Refinement

AMReX is developed as a part of DOE's **E**xascale **C**omputing **P**rojects (ECP).

eXstatic

eXascale electro**static** solver

https://github.com/AMReX-Microelectronics/eXstatic

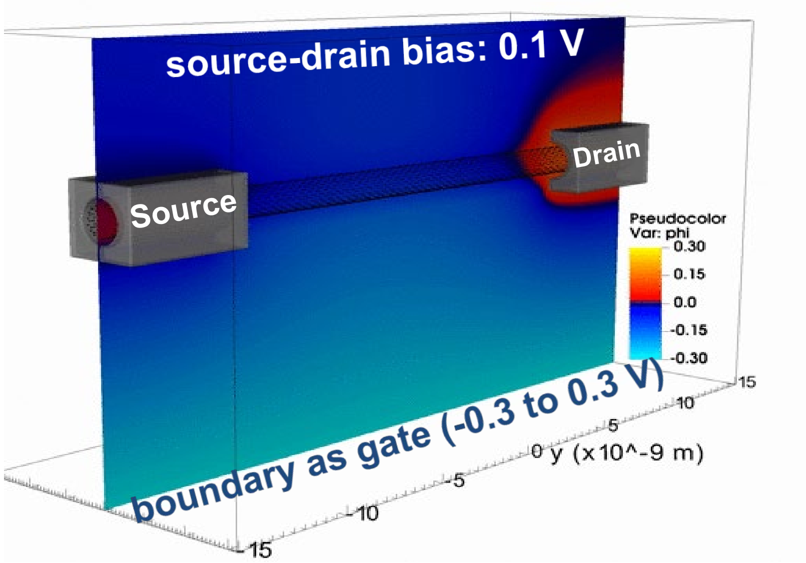
Complex geometries

3D embedded boundaries for contacts

Various boundary conditions (Robin, Neumann, Dirichlet)

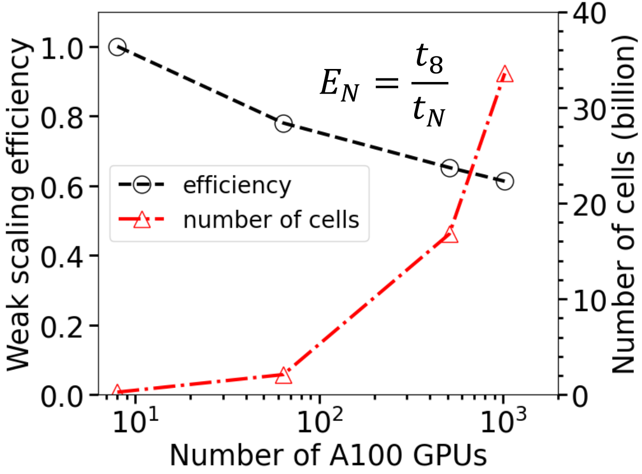
Application to Carbon Nanotube Devices

Self-Consistent NEGF and Poisson



Tight-binding, one band
All atoms included

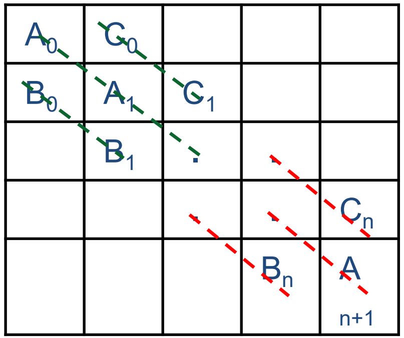
Weak-scaling on Perlmutter using Nvidia A100 GPUs



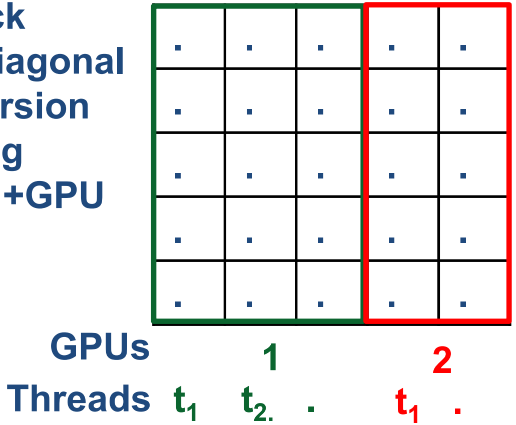
2.7 s for calculating potential for a given gate voltage in a domain with **33.5 b cells** using **1024 GPUs**. (56 µm CNT)

NEGF Quantum Transport Module: Strategy for GPUs, Scaling Studies

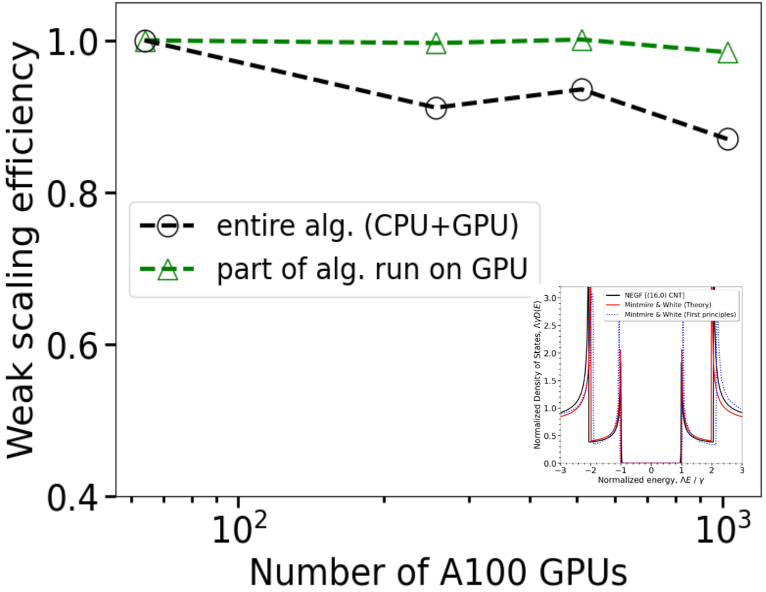
$\mathbf{G}^{-1}$ → $\mathbf{G}\ \&\ \mathbf{A} = G\,\Gamma\,G^{\dagger}$



Block Tridiagonal inversion using CPU+GPU



Weak-scaling (Nvidia A100 GPUs)

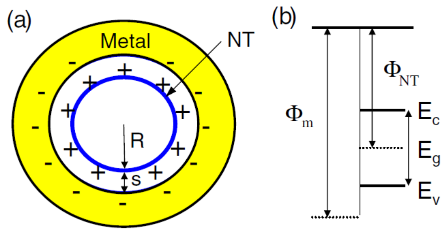


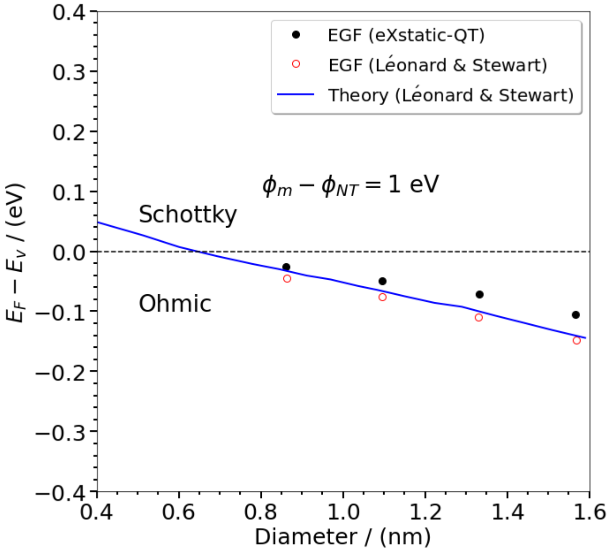
- Each CPU-GPU pair is responsible for computing specific columns of Green's and spectral functions
- Recursive part of the inversion algorithm is computed on CPUs; the rest is parallelized using GPU threads
- Charge density is computed using residue theorem and Gauss-Legendre polynomial mapping

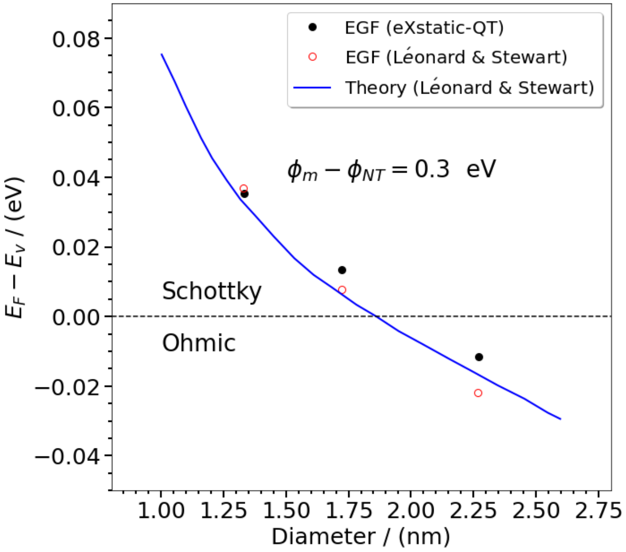
1.1 s for inverting **1.024 m x 1.024 m** matrix using 1024 GPUs, i.e. **109 micron nanotube** modeled with 1 subband

Validation for Self-Consistent eXstatic-Quantum Transport Computations

Band alignment at nanotube/metal contacts







Broyden's second algorithm optimized for memory usage[1] is used for faster convergence.

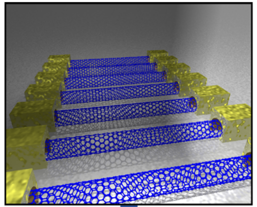
[1] G. P. Srivastava, *J. of Phys. A: Mathematical and General* 17.6 (1984): L317

[2] F. Léonard, and D. A. Stewart, *Nanotechnology* 17.18 (2006): 4699.

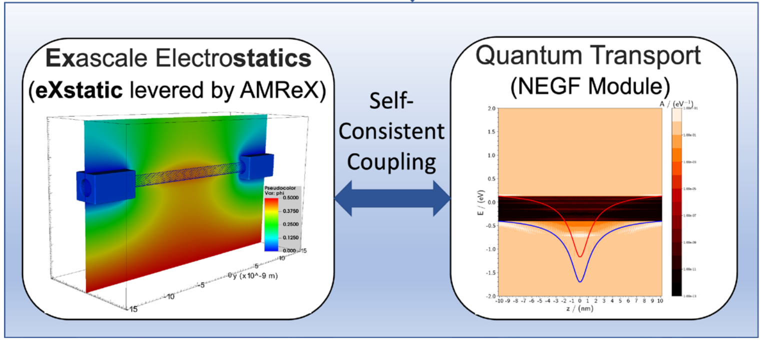
Key Takeaway

3D, GPU-enabled, open-source electrostatic—quantum transport framework for exascale systems

https://github.com/AMReX-Microelectronics/eXstatic



Scalable Modeling

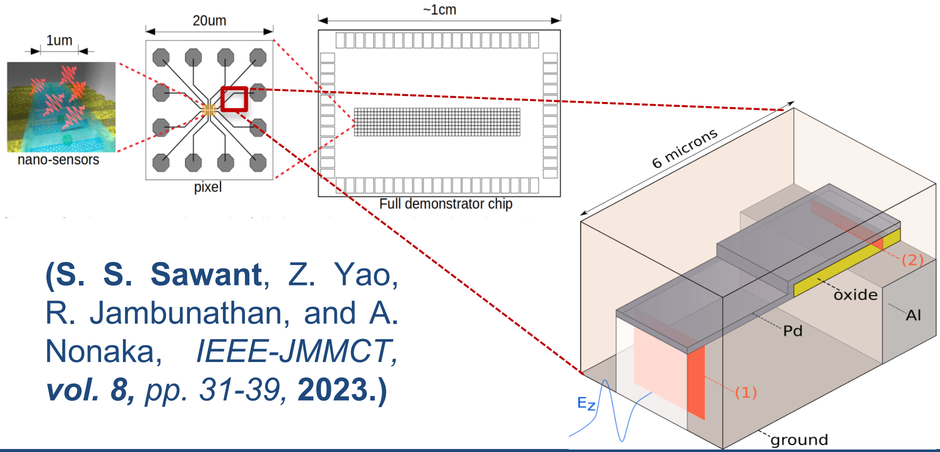


Compute potential in large domain (billions of cells)

Invert large matrices (million x million)

Future Work

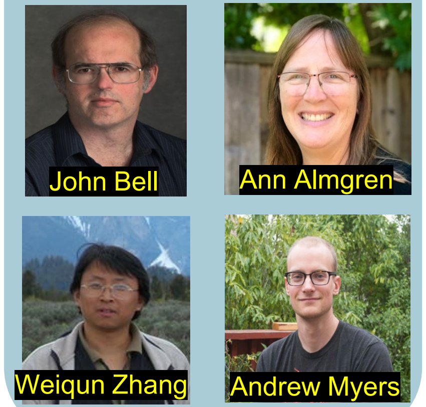
- Parallelize Broyden's algorithm &extend it for modeling **multiple CNTs**
- **Extend NEGF:** phonon scattering, time-dependent equations
- **Coupling to microscale EM** simulations.



(**S. S. Sawant**, Z. Yao, R. Jambunathan, and A. Nonaka, *IEEE-JMMCT, **vol. 8**, pp. 31-39,* **2023.)**

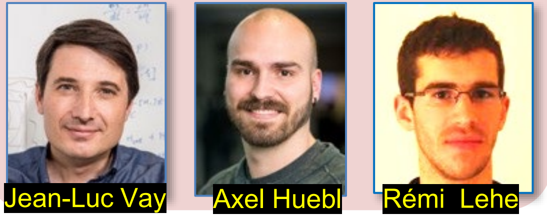
Thank You to Teams and Collaborators!

AMReX



AMReX-Microelectronics

WarpX



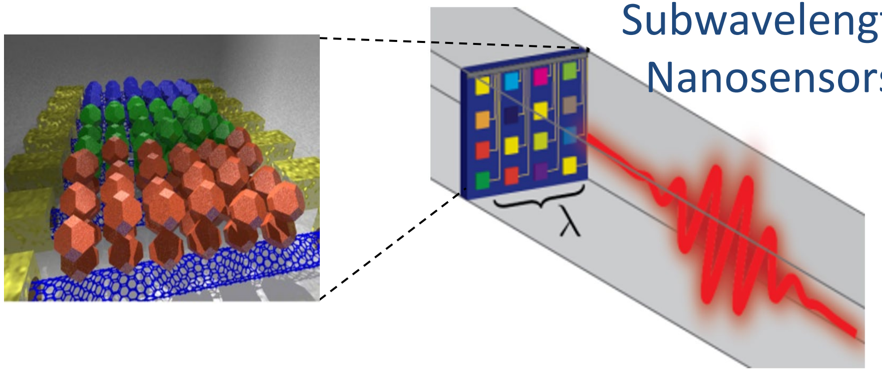
"Codesign and Integration of Nanosensors on CMOS" Team

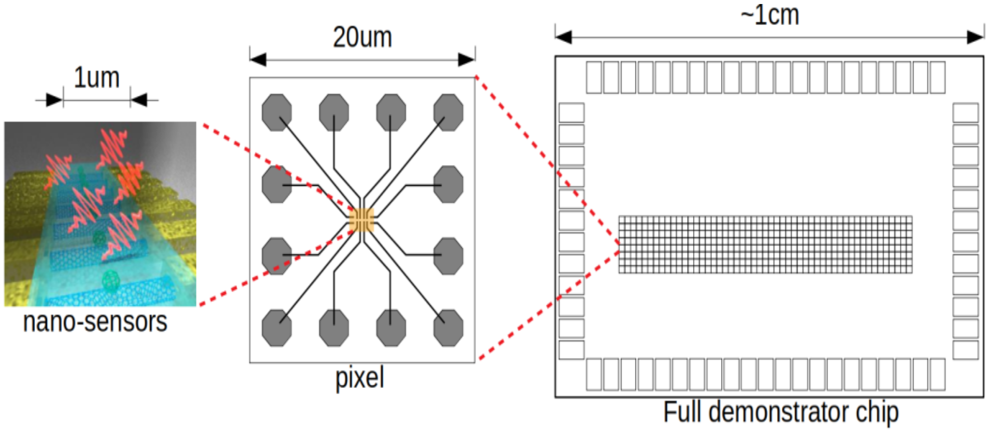
Carbon Nanotubes Field Effect Transistors (CNTFETs) for Photon Detection

CNTFETs in a advanced photodetector CMOS chip

DOE funded "Codesign and Integration of Nanosensors on CMOS"

Subwavelength Nanosensors





High surface to volume ratio makes carbon nanotubes **sensitive to their environment.**