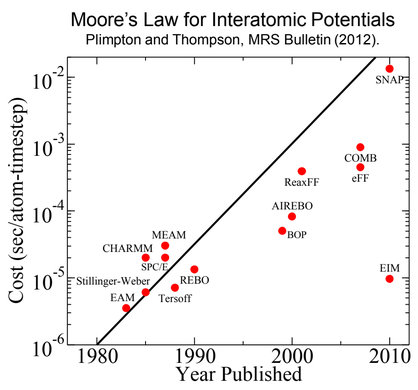
SAND2013-4097P

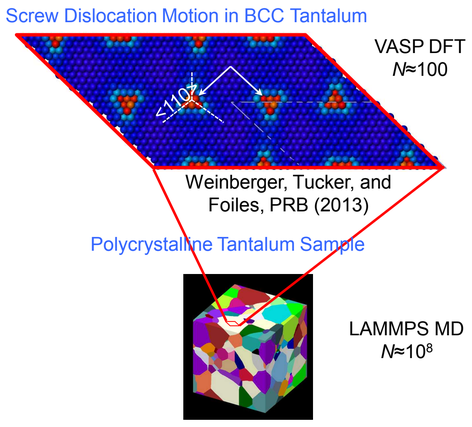
Quantum-Accurate LAMMPS SNAP Simulations on Petascale Platforms

Sandia National Laboratories
Aidan Thompson, Christian Trott, Laura Swiler, Garritt Tucker, Stephen Foiles, and Steve Plimpton
Sandia National Laboratories, New Mexico 87185

Problem

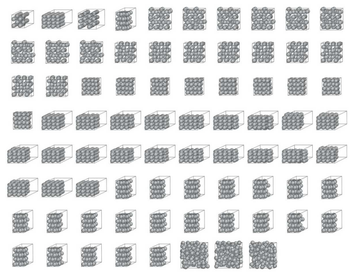
- **Problem**: Cost and complexity of molecular dynamics (MD) potentials are growing exponentially. Effort required to develop these potentials and implement them in LAMMPS is becoming unsustainable. The proliferation of new hardware designs further exacerbates the problem
- **Driver**: widespread availablity of quantum calculations (QM) for small systems: exposes errors in existing potentials and provides data for fitting new potentials
- **Solution**: Automated machine-learning framework for both fitting potentials to QM data and running large-scale MD simulations in LAMMPS





Approach

- Combination of fundamental physics and machine-learning
- GAP (Gaussian Approximation Potential): Bartok, Csanyi et al., *Phys. Rev. Lett, 2010.* Uses 3D neighbor density bispectrum and **Gaussian process regression**.
- SNAP (Spectral Neighbor Analysis Potential): Uses GAP's neighbor bispectrum, but replaces Gaussian process with **linear regression**.
  - More robust
  - Decouples MD speed from training set size
  - Allows large training data sets, more bispectrum coefficients
  - Straightforward sensitivity analysis

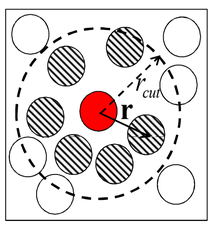


$$c^{j}_{m,m'} = U^{j}_{m,m'}(0,0,0) + \sum_k f(r_k) U^{j}_{m,m'}(\theta_0(r_k), \theta_k, \phi_k)$$

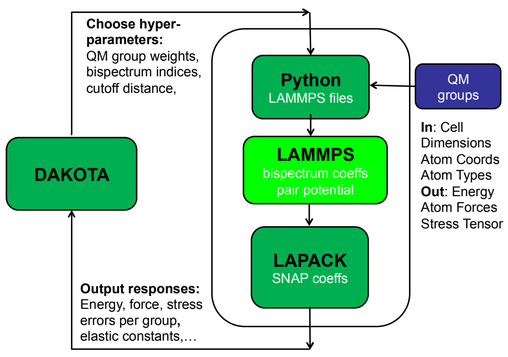
$$B_{j_1,j_2,j} = \sum_{m,m'=-j}^{j} \sum_{m_1,m_1'=-j_1}^{j_1} \sum_{m_2,m_2'=-j_2}^{j_2} c^{j*}_{m'm} C^{jm}_{j_1 m_1 j_2 m_2} C^{jm'}_{j_1 m_1' j_2 m_2'} c^{j_1}_{m_1' m_1} c^{j_2}_{m_2' m_2}$$

$$E^{SNAP}_i = \beta_0 + \sum_{k \in \{J < J_{\max}\}} \beta_k B^i_k$$

$$E = \sum_{i=1}^{N} E^{SNAP}_i + \sum_{j<i}^{N} \phi^{rep}_{ij}(r_{ij})$$



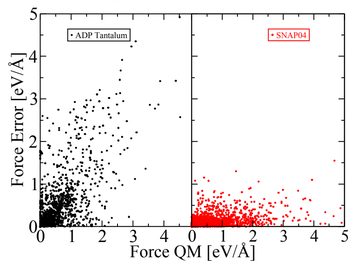
SNAP Fitting Process



Results

Accuracy

- Developed SNAP potential for tantalum, an important BCC metal
- Accuracy against QM training data 10x better than existing potentials
- Accuracy of standard physical properties (elastic constants, lattice constants) comparable to existing potentials
- Reproduced correct energy barrier for screw dislocation motion. Existing potentials are qualitatively wrong.



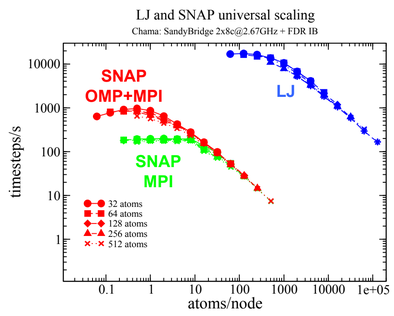
Reaction Coordinate

Computational Performance

A) SNL Chama (Intel SandyBridge CPU Cluster)

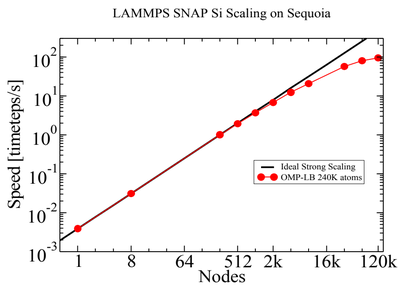
- Computational cost per atom 10,000x greater than simple potentials such as LJ and EAM
- Large computational cost offset by improved strong scaling
- Further improved strong scaling using micro-load balancing and thread parallelism over *i-j* neighbor interactions
- Reduced LJ/SNAP maximum speed ratio to 10x

		LJ	SNAP	SNAP/LJ
Data	kBytes/atom	1	1	1
Computation	MFlOps/atom-step	0.001	10	10,000
Min *N/P*	Atom/node	100	1	1/100
Max Speed	Step/Sec	10,000	1,000	1/10



B) LLNL Sequoia (IBM BG/Q)

- Ran fix-sized problem (240k atoms) on node counts from 1 to 122,880 nodes.
- Used 2 MPI tasks per node and 32 OMP threads per MPI task.
- On 120k nodes, the parallel efficiency was 20%, with only 2 atoms per node!



C) ORNL Titan (NVIDIA Kepler GPUs)

- Used Shannon (64 GPU testbed)
- Achieved 40,000 threads per atom by decomposing **within** each *i-j* pair
- GPU performance/watt similar to that of other high-end platforms
- Scaling tests indicate that 240k atoms on full Titan (18k nodes) will run 3x faster than full Sequoia (120k nodes)

Significance

This work demonstrates that by combining quantum-accurate potentials and petascale computation, it is now possible to perform **predictive** LAMMPS simulations of materials plasticity and other performance-critical atomistic phenomena that occur on mesoscopic lengthscales inaccessible to quantum methods.

National Nuclear Security Administration
U.S. DEPARTMENT OF ENERGY

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

SAND 2010-xxxP

Sandia National Laboratories