SAND2017-2541C

Exceptional service in the national interest

Sandia National Laboratories

SNAP: Automated Generation of High-Accuracy Interatomic Potentials using Quantum Data

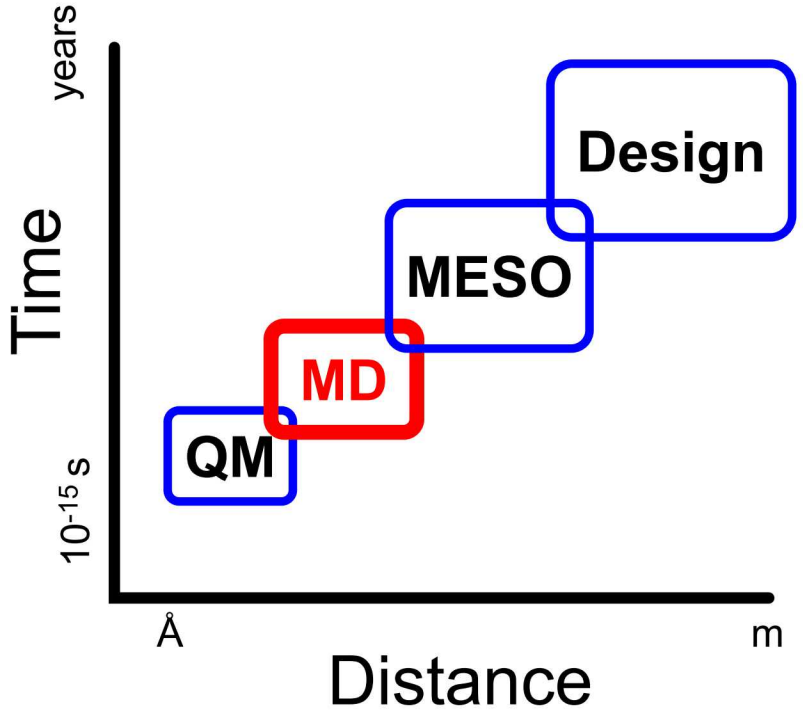
AIDAN THOMPSON
MITCH WOOD
(Sandia National Laboratories)
SIMON PHILLPOT
(University of Florida)

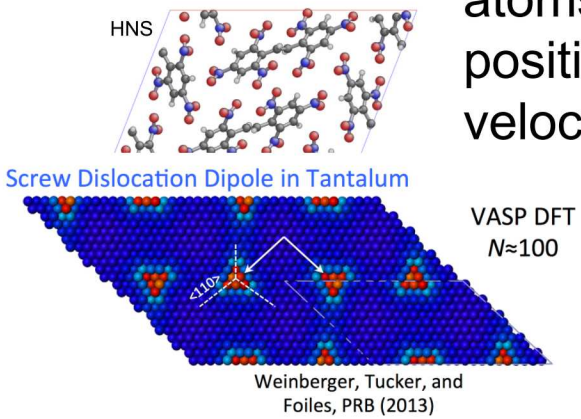
U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

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.

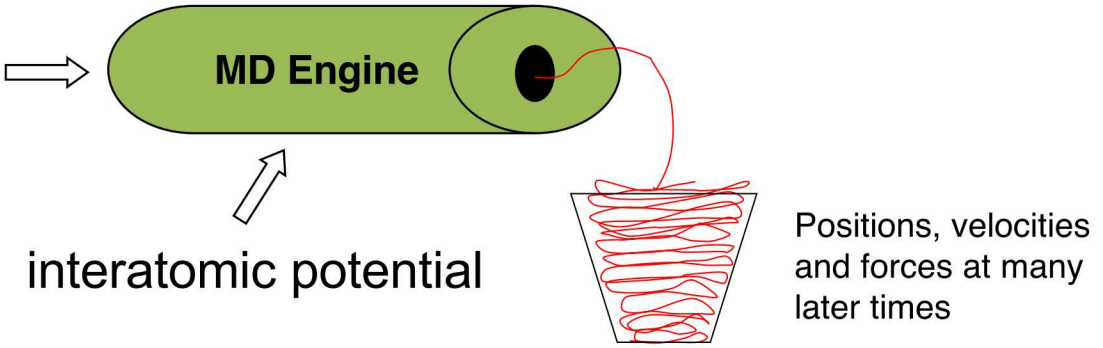
What is Molecular Dynamics Simulation?

- Continuum models require underlying models of the materials behavior
- Quantum methods can provide very complete description for 100s of atoms
- Molecular Dynamics acts as the "missing link"
  - Bridges between quantum and continuum models
  - Moreover, extends quantum accuracy to continuum length scales; retaining atomistic information

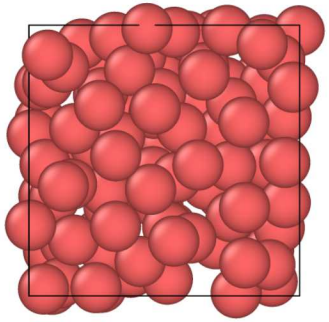




atoms,
positions,
velocities



SNAP: Spectral Neighbor Analysis Potentials



Geometric descriptors of atomic environments

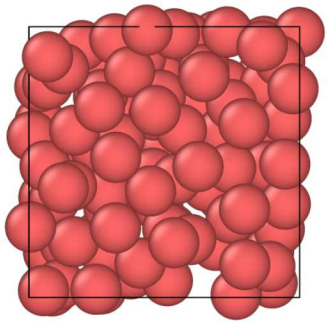
Energy as a function of geometric descriptors

$$E^{SNAP} = \sum_{i=1}^{N} E_i^{SNAP} + \sum_{j<i}^{N} \phi_{ij}^{rep}\left(r_{ij}\right)$$

$$E_i^{SNAP} = \beta_0 + \sum_{k \in \{J<J_{\max}\}} \beta_k B_k^i$$

- **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):** Our SNAP approach uses GAP's neighbor bispectrum, but replaces Gaussian process with **linear regression**.
  - More robust
  - Lower computational cost
  - Decouples MD speed from training set size
  - Enables large training data sets, more bispectrum coefficients
  - Straightforward sensitivity analysis

SNAP Quantum-Accurate Interatomic Potentials



Geometric descriptors of atomic environments

Energy as a function of geometric descriptors

$$E^{SNAP} = \sum_{i=1}^{N} E_i^{SNAP} + \sum_{j<i}^{N} \phi_{ij}^{rep}\left(r_{ij}\right)$$

$$E_i^{SNAP} = \beta_0 + \sum_{k \in \{J<J_{\max}\}} \beta_k B_k^i$$

- **GAP (Gaussian Approximation Potential):** Bartok, Csanyi et al., *Phys. Rev. Lett, 2010.* Uses 3D neighbor density bispectrum
- **SNAP (Spectral Neighbor Analysis Potential):** Our SNAP approach (Thompson *et al., J.Comp.Phys, 2015*) combines the neighbor bispectrum with weighted linear regression
  - Robust
  - Decouples MD speed from training set size
  - Enables large training data sets, more bispectrum coefficients
  - Straightforward sensitivity analysis

Bispectrum Components

- Neighbors of each atom are mapped onto unit sphere in 4D
- Expand density around each atom in a basis of ***4D hyperspherical harmonics***,
- Bispectrum components of the 4D hyperspherical harmonic expansion are used as the geometric descriptors of the local environment
- Preserves universal physical symmetries
- Rotation, translation, permutation
- Size-consistent

$$u^j_{m,m'} = U^j_{m,m'}(0,0,0) + \sum_{r_{ii'}<R_{cut}} f_c(r_{ii'}) w_i U^j_{m,m'}(\theta_0,\theta,\phi)$$

$$B_{j_1,j_2,j} = \sum_{m_1,m_1'=-j_1}^{j_1} \sum_{m_2,m_2'=-j_2}^{j_2} \sum_{m,m'=-j}^{j} (u^j_{m,m'})^* H^{jmm'}_{j_1 m_1 m_1' \; j_2 m_2 m_2'} \; u^{j_1}_{m_1,m_1'} u^{j_2}_{m_2,m_2'}$$

		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

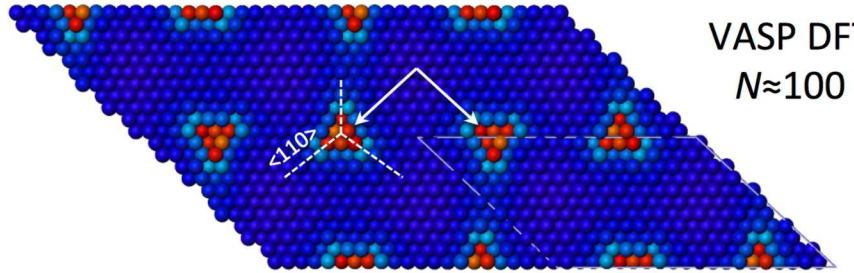
SNAP potentials predict correct Peierls barrier for Ta screw dislocations

Sandia National Laboratories

- Peierls barrier is the activation energy to move a screw dislocation
- Many simple interatomic potentials incorrectly predict a metastable state
- Leads to erroneous dynamics
- SNAP potential agrees well with DFT calculations

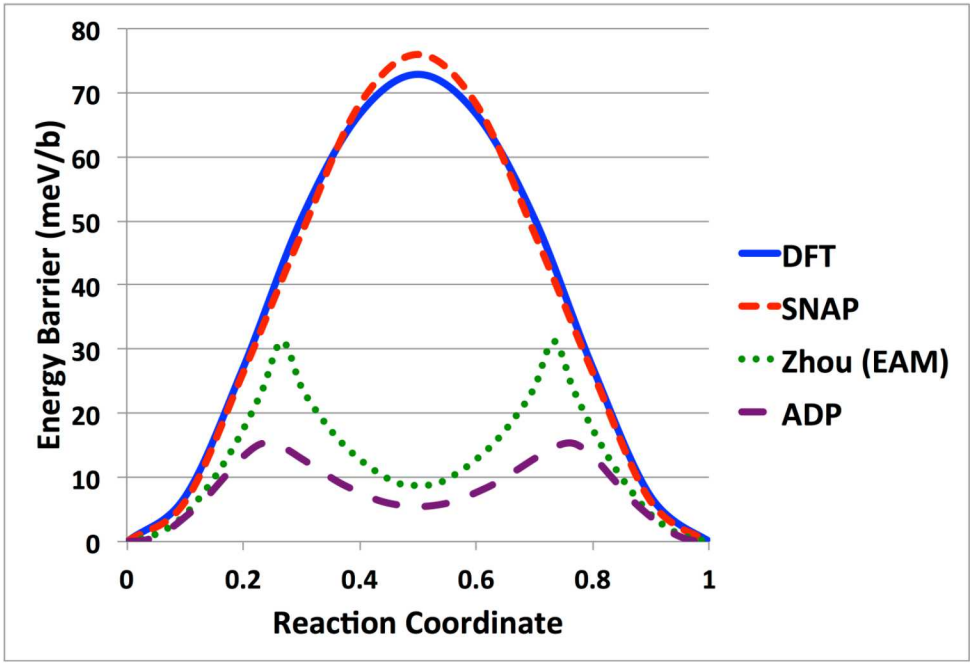
Thompson et al. arxiv.org/abs/1409.3880
J. Comp. Phys. (2015)

Screw Dislocation Dipole in Tantalum



VASP DFT
$N \approx 100$

Weinberger, Tucker, and Foiles, PRB (2013)



SNAP Potentials for W-He

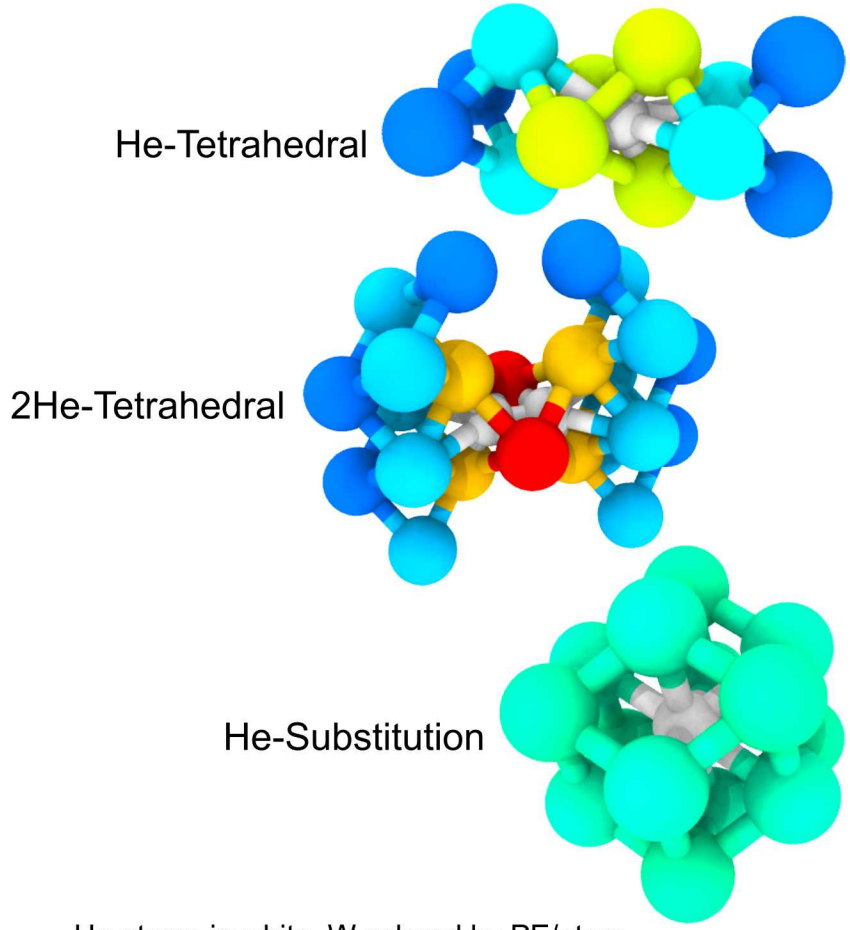
Transferrable, User Generated Training

2911 Pure W Training Data +

250 x Elastic Deformation w/ 1 octahedral He
250 x Elastic Deformation w/ 1 tetrahedral He
50 x EoS w/ 1 octahedral He
50 x EoS w/ 1 tetrahedral He
227 x Random displacements w/ 1,2 or 4 He

Objective Functions in Dakota:
- Energy Error w.r.t. DFT
- Force Error w.r.t. DFT
- Tungsten Structure Properties (a_0, elastic const.)
- Tungsten Defects (same as before)
- Helium substitution, Tetra, Octa and Tetra-Tetra binding

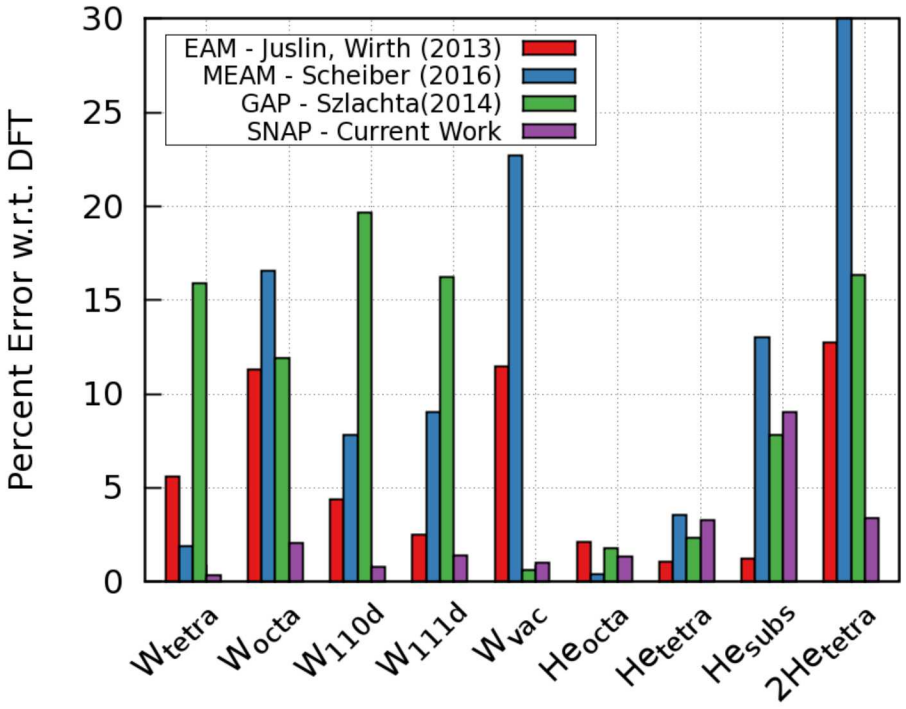
*Reference values taken from Brian's paper



He atoms in white, W colored by PE/atom

Comparison of W EAM, MEAM and SNAP + He pair

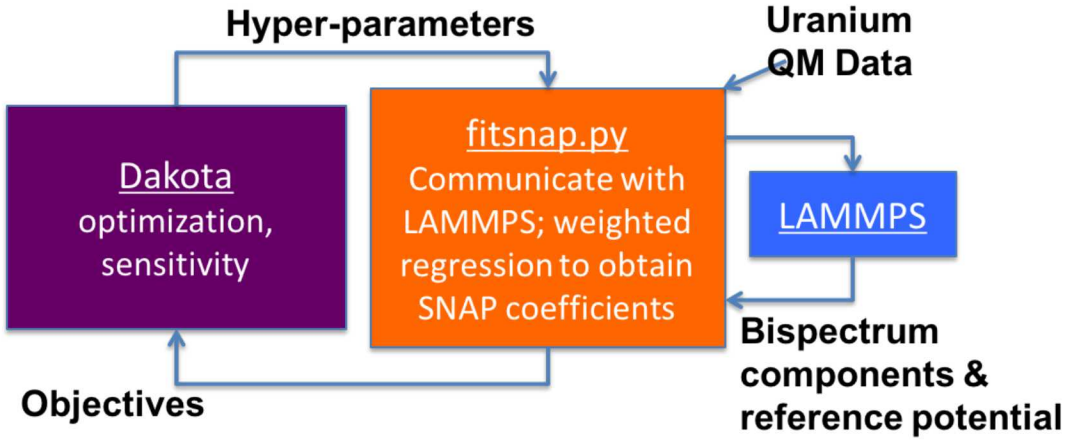
	DFT[a]	EAM[b]	MEAM[c]	GAP[d]	SNAP[e]
$E_{W-Tetra}$	11.1	10.4	11.3	12.8	11.1
E_{W-Octa}	11.7	10.4	13.6	13.1	11.5
$E_{W-[110]d}$	9.8	10.3	9.1	11.8	9.8
$E_{W-[111]d}$	9.6	9.8	10.4	11.1	9.7
$E_{W-vacancy}$	3.3	3.7	4.0	3.3	3.2
$E_{W-divacancy}$[f]	0.1	-0.4	-0.2	0.4	0.1
$E_{He-Tetra}$	6.2	6.2	6.4	6.3	6.4
$E_{He-Octa}$	6.4	6.2	6.4	6.3	6.3
$E_{He-Subs}$	4.7	4.8	5.3	4.3	4.3
$E_{2He-Tetra}$[g]	1.0	0.9	0.6	0.8	1.0



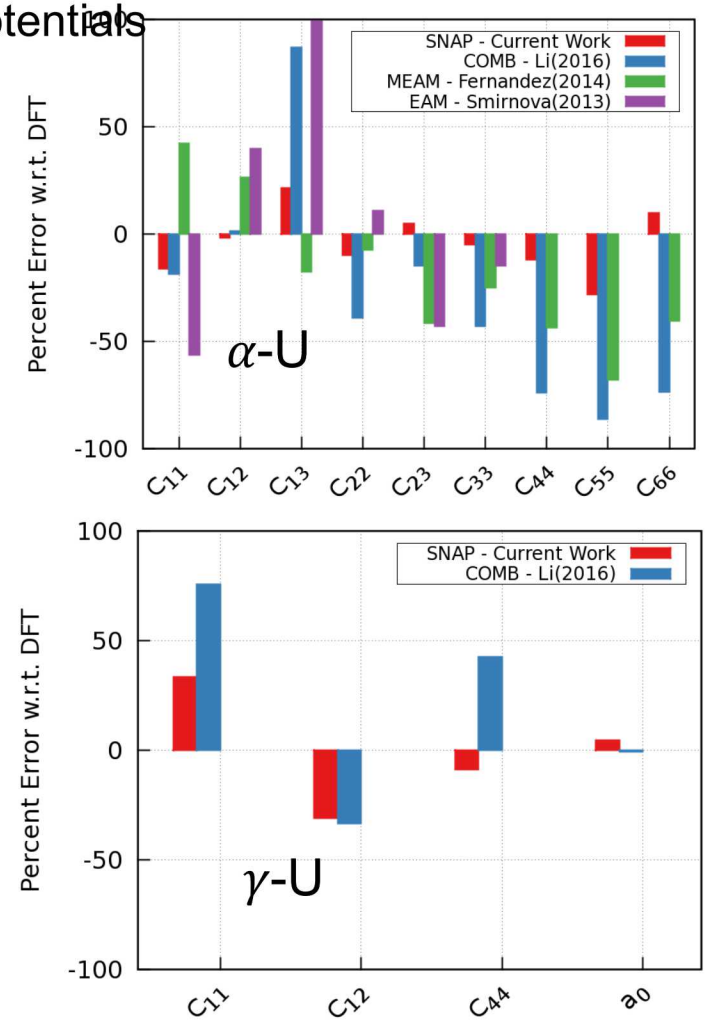
- Percent error is with respect to DFT values reported from *Becquart, Domain, Nucl. Instrum. Methods Phys. Res. B 255 (1) (2007) 23– 26.*
- Difference in He-Oct and He-Tet is tricky to handle in all potentials, should be -0.22eV

Accomplishment: Ensemble of Uranium Potentials

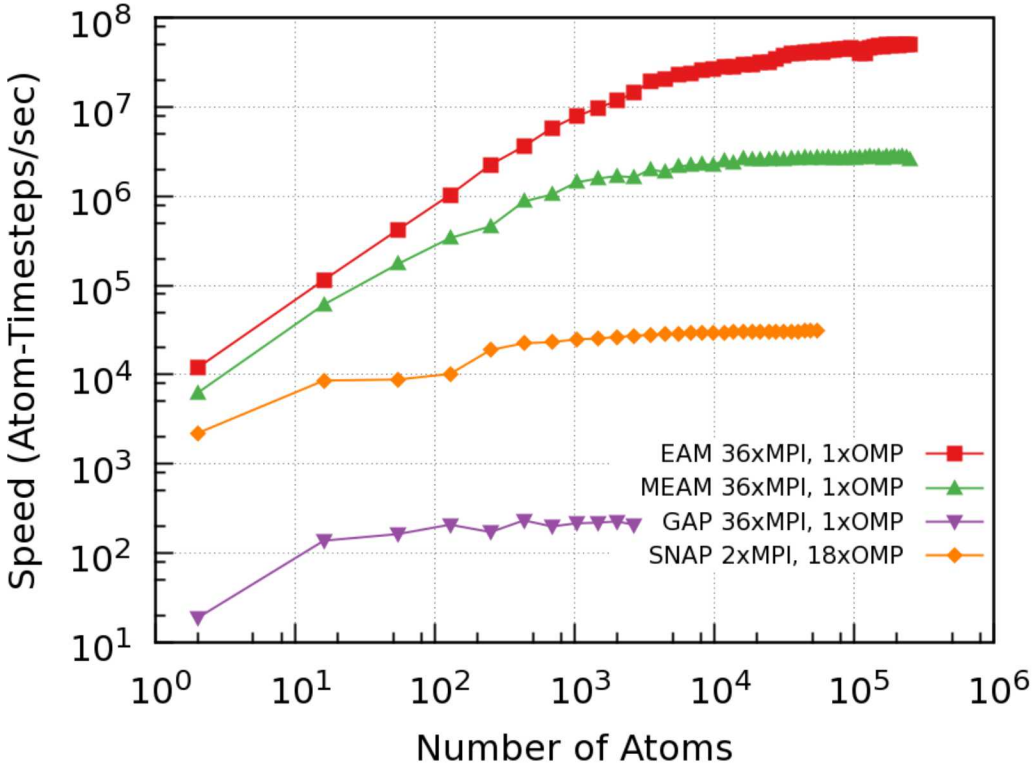
- **Generating interatomic potentials is essentially a black box, this is true for all potential types. We aim to improve the transparency and transferability of this process.**
- SNAP fitting process generates ensemble of potentials
- Evaluation against multiple objectives identifies Pareto-optimal potentials, subject to further screening
- **Sub-set of SNAP potential ensemble out-perform previous published potentials on elastic constants**



Comparison of elastic constants predictions for α-U and γ-U phases: representative SNAP potential against previous published potentials



Computational Cost of Tungsten Potentials



Sandia National Laboratories

Conclusions

- Highly complex potentials becoming more prevalent
- High Flops/atom allows excellent strong scaling
- Many-core and GPU platforms require exploiting low-level parallelism
- Extensive code modification required for each new platform
- Algorithmic improvements also important
- SNAP code and potentials available at http://lammps.sandia.gov