SNAP: Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials for LAMMPS

Aidan Thompson
Computing Research Center
Sandia National Laboratories, New Mexico
August 2015 LAMMPS Users' Workshop and Symposium
SAND 2014-1297C
Why Use 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

![Diagram showing time-distance relationship between MD, MESO, and Design](image)

- MD Engine produces interatomic potential
- Positions, velocities and forces at many later times

Screw Dislocation Dipole in Tantalum

VASP DFT
N=100

Weinberger, Tucker, and Foiles, PRB (2013)
Example: Plasticity in BCC Metals

Screw Dislocation Motion in BCC Tantalum

VASP DFT
$N \approx 100$

Weinberger, Tucker, and Foiles, PRB (2013)

Polycrystalline Tantalum Sample

LAMMPS MD
$N \approx 10^8$
SNAP: Spectral Neighbor Analysis Potentials


- **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

\[
E_{SNAP} = \sum_{i=1}^{N} E_i^{SNAP} + \sum_{j<i}^{N} \phi_{ij}^{rep}(r_{ij})
\]

\[
E_i^{SNAP} = \beta_0 + \sum_{k \in \{J<J_{max}\}} \beta_k B_k^i
\]
Bispectrum Components as Descriptor

- Neighbors of each atom are mapped onto unit sphere in 4D

\[(\theta_0, \theta, \phi) = \left(\theta_0^{\text{max}} \frac{r}{r_{\text{cut}}}, \cos^{-1}(z/r), \tan^{-1}(y/x)\right)\]

- Expand density around each atom in a basis of 4D hyperspherical harmonics,

- Bispectrum components of the 4D hyperspherical harmonic expansion are used as geometric descriptors of the local environment:
  - Preserves universal physical symmetries
  - Rotation, translation, permutation
  - Size-consistent

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

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

Symmetry relation:

\[\frac{B_{j_1,j_2,j}}{2j+1} = \frac{B_{j,j_2,j_1}}{2j_1+1} = \frac{B_{j_1,j,j_2}}{2j_2+1}\]
SNAP Fitting Process

FitSnap.py

“Hyper-parameters”
• Cutoff distance
• Group Weights
• Number of Terms
• Etc.

Dakota
optimization, sensitivity

fitsnap.py
Communicate with LAMMPS; weighted regression to obtain SNAP coefficients

LAMMPS

Metrics
• Force residuals
• Energy residuals
• Elastic constants
• Etc.

QUEST
QDFT
Training Data

Bispectrum components & derivatives, reference potential
Ta SNAP potential was fit to a DFT-based training set containing ‘usual suspects’

For each configuration in training set, fit total energy, atomic forces, stress
- Equilibrium lattice parameter
- Elastic constants ($C_{11}$, $C_{12}$, and $C_{44}$) and bulk modulus (B)
- Free surface energies: (100), (110), (111), and (112)
- Generalized planar stacking fault curves: {112} and {110}
- Energy-Volume (Contraction and Dilation) - BCC, FCC, HCP, and A15
- Lattices with random atomic displacements
- Liquid structure

Example: DFT-based Generalized Stacking Fault Energies

![Graphs showing energy vs. normalized displacement for (110) and (112) fault energies.](image)
Effect of Higher-order Bispectrum Components

- Liquid force errors decrease with increasing J
- Diminishing returns beyond J = 7/2

<table>
<thead>
<tr>
<th>$2J$</th>
<th>$N$</th>
<th>Ferr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2.09</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>1.39</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>0.66</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>0.53</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>0.44</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>0.35</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>0.30</td>
</tr>
</tbody>
</table>
SNAP potential yields good agreement with DFT results for some standard properties

<table>
<thead>
<tr>
<th>Property</th>
<th>DFT</th>
<th>SNAP</th>
<th>Zhou (EAM)</th>
<th>ADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lattice Constant (Å)</td>
<td>3.320</td>
<td>3.316</td>
<td>3.303</td>
<td>3.305</td>
</tr>
<tr>
<td>B (Mbar)</td>
<td>1.954</td>
<td>1.908</td>
<td>1.928</td>
<td>1.971</td>
</tr>
<tr>
<td>(C' = \frac{1}{2}(C_{11} - C_{12})) (Mbar)</td>
<td>50.7</td>
<td>59.6</td>
<td>53.3</td>
<td>51.0</td>
</tr>
<tr>
<td>(C_{44}) (Mbar)</td>
<td>75.3</td>
<td>73.4</td>
<td>81.4</td>
<td>84.6</td>
</tr>
<tr>
<td>Vacancy Formation Energy (eV)</td>
<td>2.89</td>
<td>2.74</td>
<td>2.97</td>
<td>2.92</td>
</tr>
<tr>
<td>(100) Surface Energy (J/m²)</td>
<td>2.40</td>
<td>2.68</td>
<td>2.34</td>
<td>2.24</td>
</tr>
<tr>
<td>(110) Surface Energy (J/m²)</td>
<td>2.25</td>
<td>2.34</td>
<td>1.98</td>
<td>2.13</td>
</tr>
<tr>
<td>(111) Surface Energy (J/m²)</td>
<td>2.58</td>
<td>2.66</td>
<td>2.56</td>
<td>2.57</td>
</tr>
<tr>
<td>(112) Surface Energy (J/m²)</td>
<td>2.49</td>
<td>2.60</td>
<td>2.36</td>
<td>2.46</td>
</tr>
<tr>
<td>(110) Relaxed Unstable SFE (J/m²)</td>
<td>0.72</td>
<td>1.14</td>
<td>0.75</td>
<td>0.58</td>
</tr>
<tr>
<td>(112) Relaxed Unstable SFE (J/m²)</td>
<td>0.84</td>
<td>1.25</td>
<td>0.87</td>
<td>0.74</td>
</tr>
</tbody>
</table>
Liquid structure: SNAP and DFT are in excellent agreement

Liquid pair correlation function, $g(r)$ computed at 3250 K (~melting point) and experimental density

- DFT: 100 atoms, 2 picoseconds
- SNAP: 1024 atoms, 200 picoseconds
SNAP potentials predict correct Peierls barrier for Ta screw dislocations

- 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
  - Future work will explore dislocation dynamics based on this potential

SNAP Indium Phosphide

Additional Challenges
- Two elements
- Different atom sizes
- Diverse structures
- Defect formation energies
- Sensitive to curvature

Innovations
- Differentiate elements by: density weight, linear coefficients, neighbor cutoff
- Trained against relaxed defect structures
- Trained against deformed defect structures

Result (so far)
- Good overall fit
- Defect energy error > 1 eV
SNAP Silica: Promising Start
(Stan Moore, Paul Crozier, Peter Schultz)

Less than 3% error in predicted lattice parameters of 7 crystal polymorphs

Additional Challenges
- Electrostatics
- Started with no training data
- Goal: quantum-accurate prediction of Si/SiO₂ interface

Innovations
- Generated training data adaptively, on-the-fly
- Added fixed point charges, long-range electrostatics

Result (so far)
- Good agreement with QM for SiO₂ crystal polymorphs
- Good agreement with QM liquid structure for SiO₂

Good agreement with QM liquid structure for SiO₂
Conclusions

- SNAP is a new formulation for interatomic potentials
  - Geometry described by bispectrum components
  - Energy is a linear regression of bispectrum components
- Works well for Ta
  - Liquid structure
  - Peierls barrier for screw dislocation motion
- Ongoing work
  - Extension to binary systems: InP, SiO2, TaO_x
- SNAP Ta potential published
  - arxiv.org/abs/1409.3880
- SNAP Ta available in LAMMPS

Primary Collaborators
Laura Swiler
Stephen Foiles
Garritt Tucker

Additional Collaborators
Christian Trott
Peter Schultz
Paul Crozier
Stan Moore
Adam Stephens
FitSnap.py: Robust Software Framework

Key advantages of fitsnap.py
- Minimal file I/O
- Use of NumPy/SciPy
- Caching and reuse of data
- File-based input
- Supports parallel LAMMPS

Hyper-parameter Optimization

<table>
<thead>
<tr>
<th>Candidate SNAP Potentials</th>
<th>Error Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>150</td>
</tr>
<tr>
<td>4000</td>
<td>300</td>
</tr>
<tr>
<td>6000</td>
<td>450</td>
</tr>
<tr>
<td>8000</td>
<td>600</td>
</tr>
</tbody>
</table>

Defects Group Weight Sensitivity

- Energy Err. (eV) ➡️ MSD (Å) ➡️ Relaxation (eV) ➡️ Force Err. (eV/Å)

Log_{10}(Defects Group Weight) vs. Energy Err. (eV)