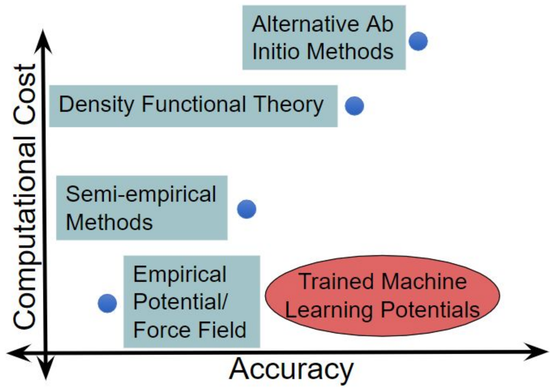
Modifying the Atomistic Machine-learning Package for Real-time Atomic Simulations with DFT Accuracy

Megan Le, Graham Stockton, Juliana Duncan, Graeme Henkelman

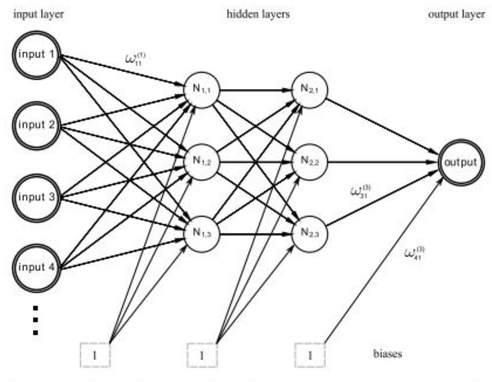
Department of Chemistry and the Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas, 78712-0165, USA

Contact: Graham Stockton, grahamstockton50@gmail.com

Motivations



Ab-initio methods such as Density Functional Theory (DFT) and Perturbation theory have high computational costs that make them unfeasible for real-time calculations, especially on standard laptop or desktop hardware.



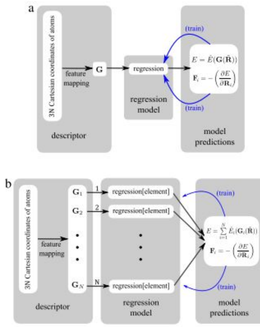
Khorshidi & Peterson, "Amp: A modular approach to machine learning in atomistic simulations", Computer Physics Communications 207:310-324, 2016. DOI:10.1016/j.cpc.2016.05.010

Through the use of Peterson and Khorshidi's Atomistic Machine-learning Package (AMP), neural networks can be trained to approximate potential energy surfaces given atomic positions with corresponding energy and forces, without the scalability or computational cost issues that occur using DFT.

Creating a ML-trained PES has significant impacts for chemical simulations, including speeding up local and global optimizations, as well as molecular dynamic simulations. AMP can be integrated into a software package called 'feel-the-force,' which translates the ML-trained forces and energies to a haptic control device, allowing the users to interact and construct chemical systems. However, one drawback of the standard AMP code is that it uses a database management method that involves writing and reading individual files for each entry.

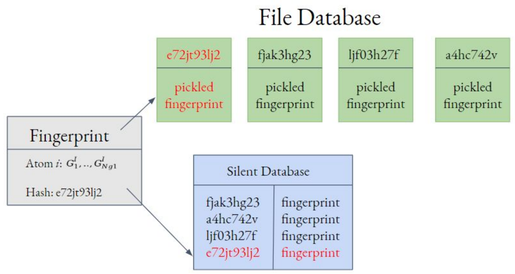
Methodology

The 'feel-the-force' software calculates potential energy and forces using the positions of the atoms, then scales the forces to be translated to a haptic device. The haptic device allows the user to control the movement of atoms while relaying the feel of the forces acting on the controlled atom. The Atomistic Machine-learning Package (AMP) was integrated with the 'feel-the-force' software to allow for the use of machine-learning generated potential energy surfaces.



Khorshidi & Peterson, "Amp: A modular approach to machine learning in atomistic simulations", Computer Physics Communications 207:310-324, 2016. DOI:10.1016/j.cpc.2016.05.010

AMP is unique in the way that it uses "fingerprints" as training images for atom clusters, instead of taking all positions into account for each atom. These fingerprints focus on one atom in a cluster and describe all other atoms within a certain radius of that atom. By using these fingerprints for each atom in a cluster, not only are more fingerprints generated per sample cluster, but the fingerprints generated are less unique and thus are easier for the model to be trained with.



The primary change to AMP was the introduction of the SilentDatabase (SDB) class. This option allows AMP to store fingerprint data in dictionaries, preventing the accumulation of files during execution and speeding up overall runtime. Since this application of AMP is intended for use with a pre-made and pre-configured AMP calculator, it is not necessary to keep a textual record of the fingerprints.

Results

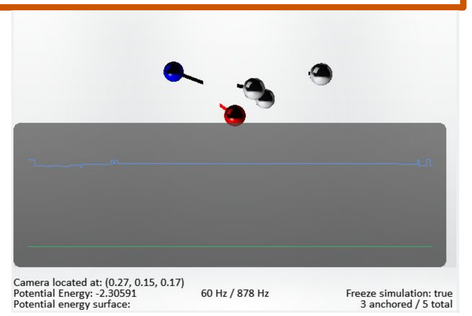


Fig 1. "Feel-the force" software modeling a 5-atom AMP cluster. Blue graph function represents AMP potential energy over time, with a green reference line for the global minimum.

The addition of AMP functionality to the "feel-the-force" software is useful both for visualizing accurate potentials much faster than would be possible using DFT, as well as for developing and troubleshooting AMP calculators. The "feel-the-force" software, shown above, shows the magnitudes and directions of the forces on each atom, and graphs the potential energy of the cluster (blue) and the predetermined global minimum for a Lennard-Jones cluster of that size (green).

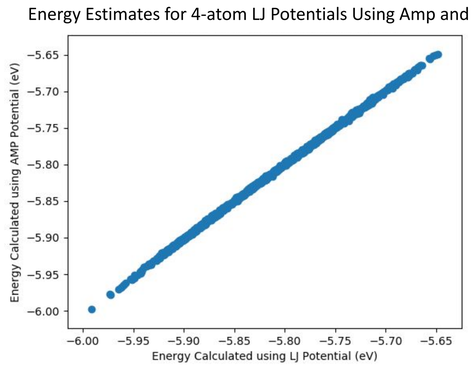


Fig 2. A comparison of potential energy estimates made by AMP and the Lennard-Jones potential energy function for a 4-atom cluster.

AMP potentials can approximate the potential energy surfaces of molecular systems to a degree of accuracy close to the method used for training. For small clusters, AMP potentials can be easily trained to within an rms error of .001. These models work particularly well with clusters near their optimized form, as these arrangements do not have high variance in the positions of atoms.

AMP potentials can also be trained with a variety of elements, allowing for energies to be calculated without complexity arising from electronic structure methods or combined potentials. However, AMP potentials also behave less uniformly for configurations outside their training set, so training must be conducted thoughtfully.

Results cont.

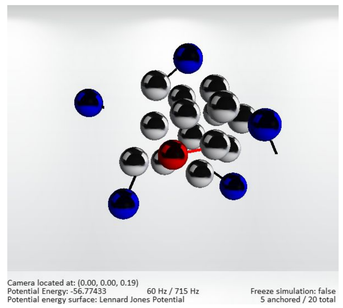


Fig 3. "Feel-the force" software modeling a 20-atom cluster and using the Lennard-Jones potential to calculate potential energy and forces.

The "feel-the-force" software now has options for Lennard-Jones, Morse, and machine-learning generated potentials available at launch. Clusters of atoms can be randomly generated or loaded from files created by other software, then manipulated using the haptic device, mouse, or keyboard. Atoms can be frozen in place and moved, and then released to observe their motion.

Conclusions

Attempts to implement and optimize AMP calculators with the "feel-the-force" software were successful, and the modified version of AMP and 'feel-the-force' software can be found publicly online. AMP was modified to prevent writing unnecessary files and is in continuing development with the goal of further optimizing speed. Accuracy comparisons between LJ and AMP potentials give promising initial results, which suggest that AMP potentials may be similarly able to approximate DFT calculations for fast simulation. Future work includes further optimization of the machine-learning integration with the 'feel-the-force' software and continued measurement and investigation of machine-learning generated potential energy surfaces.

References

1) Alireza Khorshidi and Andrew Peterson. "*Amp*: A modular approach to machine learning in atomistic simulations." *Computer Physics Communication* 207 (October 2016): 310-324.

2) Silent AMP: github.com/megan-le/silent-amp

3) 'Feel-the-force': github.com/styu99/haptic-device