Learning Atomistic Force Fields On-the-Fly with Bayesian Inference

Jonathan Vandermause,[1,2] Steven B. Torrisi,[1] Simon Batzner,[3] and Boris Kozinsky[2]

[1] Department of Physics, Harvard University, Cambridge, MA 02138, USA
[2] John A. Paulson School of Engineering and Applied Sciences,
Harvard University, Cambridge, MA 02138, USA
[3] Center for Computational Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
(Dated: March 25, 2019)

Ab initio molecular dynamics is a powerful tool for accurately probing the dynamics of molecules and solids, but it is limited to system sizes on the order of 1000 atoms and time scales on the order of 10 ps. We present a scheme for rapidly training a machine learning model of the interatomic force field that approaches the accuracy of *ab initio* force calculations but can be applied to larger systems over longer time scales. Gaussian process models are trained on-the-fly, with density-functional theory calculations of the atomic forces performed whenever the model encounters chemical configurations outside of the training set. We demonstrate the flexibility of our approach by testing it on vacancy diffusion in bulk metals.

ORGANIZATION

- Figure 1: Correlation between prediction error and out-of-sample variance. Train on a low temperature crystal and test at higher temperatures to systematically steer the model away from the training set. Establishing a correlation motivates the use of predictive variance for on-the-fly learning.

- Algorithm 1: On-the-fly algorithm, including GP, MD, and DFT updates.

- Figure 2: Number of DFT calls vs simulation time. A sharp increase should be observed when new phases are introduced.

- Table 1: Performance of FLARE compared with state-of-the art EAM and ML models. Models tested on independent *ab initio* MD runs. MAEs and mean absolute force components reported in eV/A.

- Figure 3: Validation of the force field. Parity plots, radial distribution function, and activation profiles should agree with *ab initio* calculations.

- Supplementary Figure 1: Numerical demonstration of energy conservation. Plot the RMS energy fluctuations against the time step squared and check that the resulting graph is linear [1].

- Supplementary Figure 2: Justification of cutoff selection. Plot likelihood, noise parameter, and out-of-sample error vs. cutoffs.

- Supplementary Table 1: Table of formulas for calculating kernels.

- Supplementary Table 2: Justification of quadratic cutoff. Record log likelihood for different cutoffs (cosine, hard, quadratic).

SUPPLEMENTARY INFORMATION

Covariant kernels for direct force prediction

The total energy E of a system of atoms in a periodic cell is modelled as a sum over two- and three-body contributions,

$$E = \sum_{ij} \varepsilon_{ij} + \sum_{ijk} \varepsilon_{ijk}, \qquad (1)$$

where the sums range over all unique pairs and triplets of atoms containing at least one atom from the unfolded primary cell. In practice, the sums are truncated by considering local atom-centered environments surrounding each atom in the primary cell and neglecting contributions from atoms beyond a chosen cutoff distance from the central atom. The energy may then be expressed as

$$E = \sum_i \left(\frac{1}{2} \sum_{j \in \rho_i} \varepsilon_{ij} + \frac{1}{3} \sum_{j,k \in \rho_i} \varepsilon_{ijk} \right), \qquad (2)$$

where ρ_i denotes the local environment of atom i containing all atoms within the cutoff sphere and the fractional factors take care of multiple counting due to the repeated appearance of bonds and triplets in neighboring environments. This may be written more compactly as

$$E = \sum_i \varepsilon_i, \qquad (3)$$

where $\varepsilon_i \equiv \frac{1}{2} \sum_{j \in \rho_i} \varepsilon_{ij} + \frac{1}{3} \sum_{j,k \in \rho_i} \varepsilon_{ijk}$ may be viewed as the local energy of atom i.

In Gaussian process models, the covariance between targets is set equal to a kernel or similarity measure between inputs. The covariance between total energy observations E_l, E_m of two distinct structures σ_l, σ_m may be written as

$$\langle E_l E_m \rangle = \sum_{i \in \sigma_l} \sum_{j \in \sigma_m} \langle \varepsilon_i \varepsilon_j \rangle, \qquad (4)$$

where the covariance between local energies is

$$\langle \varepsilon_i \varepsilon_j \rangle = \frac{1}{4} \sum_{n \in \rho_i} \sum_{p \in \rho_j} \langle \varepsilon_{in} \varepsilon_{jp} \rangle + \frac{1}{9} \sum_{n,q \in \rho_i} \sum_{p,r \in \rho_j} \langle \varepsilon_{inq} \varepsilon_{jpr} \rangle.$$

(5)

Letting $F_{i\xi} = -\frac{dE}{d\xi_i}$ denote the force on atom i along Cartesian component ξ, the covariance between force observations may be written as

$$\langle F_{i\xi} F_{j\chi} \rangle = \sum_{n \in \rho_i} \sum_{p \in \rho_j} \frac{\partial^2}{\partial \xi_i \partial \chi_j} \langle \varepsilon_{in} \varepsilon_{jp} \rangle + \sum_{n,q \in \rho_i} \sum_{p,r \in \rho_j} \frac{\partial^2}{\partial \xi_i \partial \chi_j} \langle \varepsilon_{inq}$$

(6)

where here the fractional factors do not appear as the sums are restricted to the local environments of atoms i and j. This is convenient in practice, as it allows local information about individual atoms to be used at test time without having to store the atomic environments of all the other atoms in the structure.

In order to infer energies from force observations, it is necessary to consider the covariance between energies and forces.
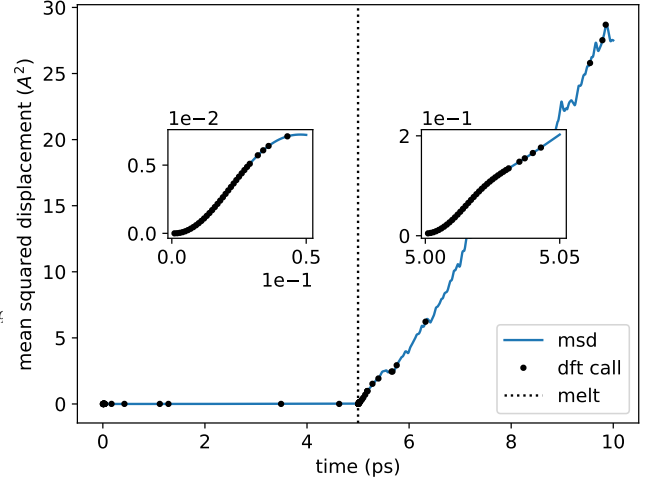

FIG. 1. On-the-fly learning of an aluminum force field at multiple temperatures.
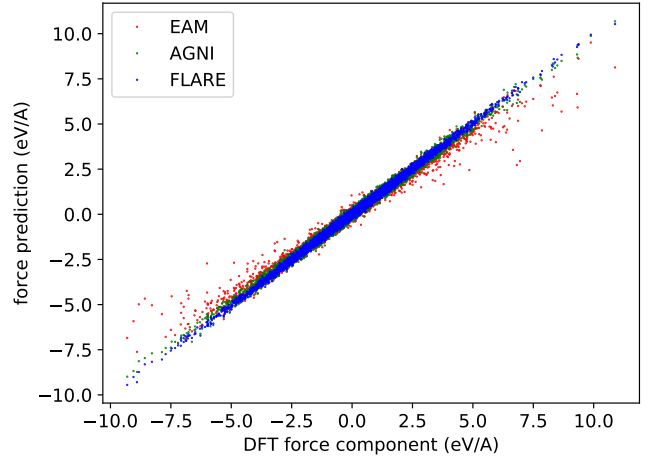

FIG. 2. Mean square displacement of aluminum melt.

		Solid	Liquid	Slab	Vacancy
OTF					
EAM					
AGNI					

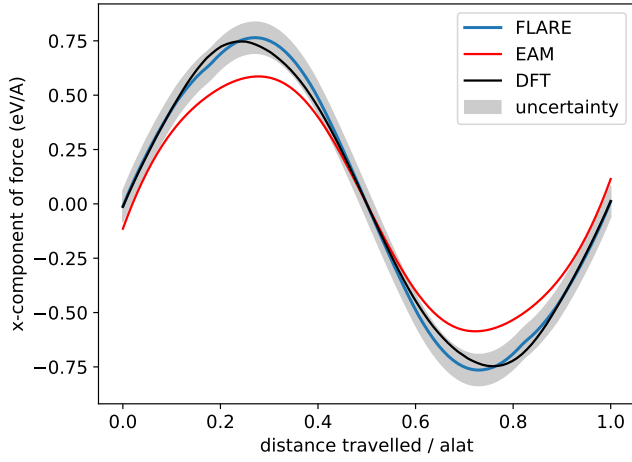TABLE I. On-the-fly force field error compared to a recent EAM potential.

FIG. 3. Mean square displacement of aluminum melt.

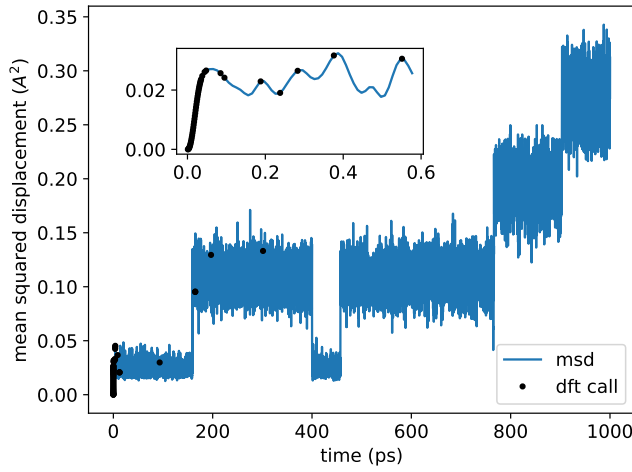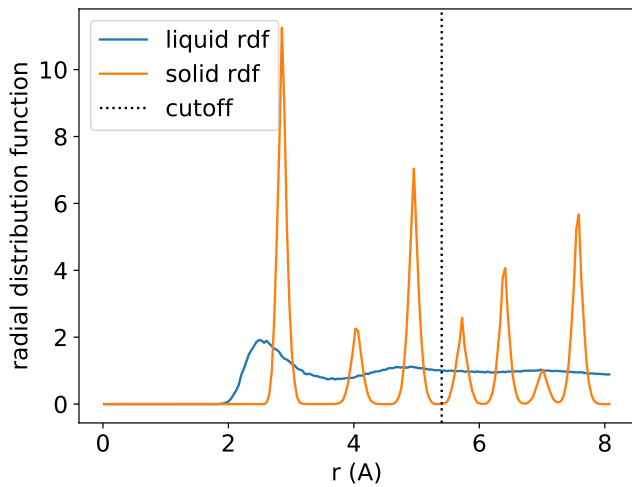FIG. 4. Mean square displacement of aluminum melt.

FIG. 5. RDF of Al melt.

Algorithm 1 Active Learning of Atomistic Force Fields

Require: initial structure (positions, velocities, periodic cell)
Require: initial GP model (kernel and hyperparameters)
Require: Δt: molecular dynamics time step
Require: T: total simulation time
Require: $\mathcal{U}$: initial uncertainty threshold
1: Initialize time: t = 0
2: **while** $t < T$ **do**
3:     predict forces and uncertainties with GP model
4:     **if** uncertainty above threshold **then**
5:         compute forces with DFT
6:         add highest uncertainty atom to training set
7:         update GP hyperparameters
8:         update structure with DFT forces
9:     **else**
10:         update structure with GP forces
11:     **end if**
12:     update time: $t = t + \Delta t$
13: **end while**

---

[1] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford university press, 2017).

Energy Kernel	k_{inv}	$\sigma^2 \sum_{c,p} k f_{\text{cut}}(\vec{d_c}) f_{\text{cut}}(\vec{d_p})$
-	k	$\exp\left(-\frac{\|\vec{d_c}-\vec{d_p}\|^2}{2\ell^2}\right)$
-	$\vec{d}^{(2)}$	(r_{i_1})
-	$\vec{d}^{(3)}$	$(r_{i_1}, r_{i_2}, r_{i_1,i_2})$
Force Kernel	$\frac{\partial^2 k_{\text{inv}}}{\partial \xi_i \partial \chi_j}$	$\sigma^2 \sum_{c,p} (k_0 + k_1 + k_2 + k_3)$
-	k_0	$k \frac{\partial f_{\text{cut}}(\vec{d_c})}{\partial \xi_i} \frac{\partial f_{\text{cut}}(\vec{d_p})}{\partial \chi_j}$
-	k_1	$\frac{\partial k}{\partial \xi_i} f_{\text{cut}}(\vec{d_c}) \frac{\partial f_{\text{cut}}(\vec{d_p})}{\partial \chi_j}$