Active Learning of Atomistic Force Fields

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: February 27, 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.

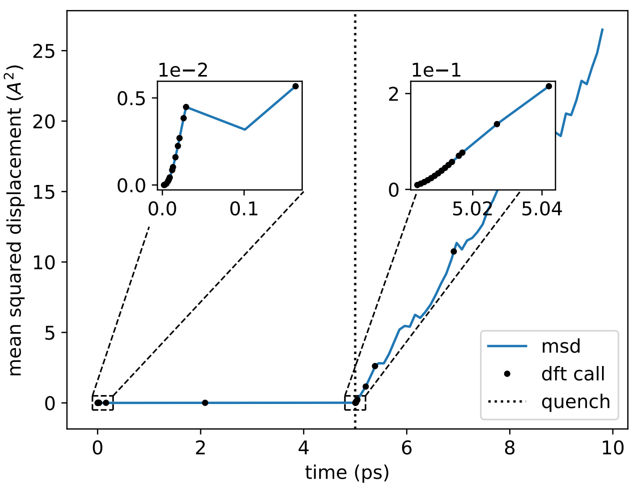

FIG. 1. Mean square displacement of aluminum quench.

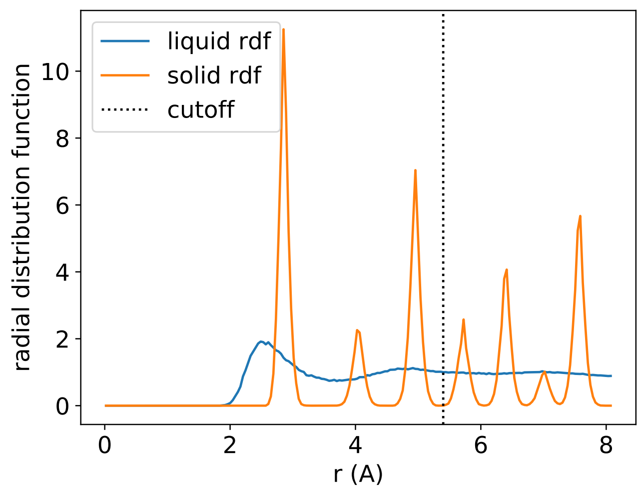

FIG. 2. RDF of Al quench.

	Solid	Liquid	Slab	Vacancy
OTF	**test**			
EAM				

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

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

```
Initialize time: t = 0
while t < T do
    predict forces and uncertainties with GP model
    if uncertainty above threshold then
        compute forces with DFT
        add highest uncertainty atom to training set
        update GP hyperparameters
        update structure with DFT forces
    else
        update structure with GP forces
    end if
    update time: t = t + Δt
end while
```

Energy Kernel	k_{inv}	$\sigma^2 \sum_{c,p} k f_{\text{cut}}(\vec{d}_c) f_{\text{cut}}(\vec{d}_p)$
-	k	$\exp\left(-\frac{\lvert\lvert \vec{d}_c - \vec{d}_p \rvert\rvert^2}{2\ell^2}\right)$
-	$\vec{d}^{(2)}$	(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	$\frac{\partial k}{\partial \xi_i} f_{\text{cut}}(\vec{d}_c) \frac{\partial f_{\text{cut}}(\vec{d}_p)}{\partial \chi_j}$
-	k_2	$\frac{\partial k}{\partial \chi_j} \frac{\partial f_{\text{cut}}(\vec{d}_c)}{\partial \xi_i} f_{\text{cut}}(\vec{d}_p)$
-	k_3	$\frac{\partial^2 k}{\partial \xi_i \partial \chi_j} f_{\text{cut}}(\vec{d}_c) f_{\text{cut}}(\vec{d}_p)$
-	$\frac{\partial k}{\partial \xi_i}$	$\frac{k B_1}{\ell^2}$
	B_1	$\sum_{q=1}^{N-1} \frac{(r_{i_q} - r_{j_q}) \xi_{i_q}}{r_{i_q}}$
-	$\frac{\partial k}{\partial \chi_j}$	$-\frac{k B_2}{\ell^2}$
-	B_2	$\sum_{q=1}^{N-1} \frac{(r_{i_q} - r_{j_q}) \chi_{j_q}}{r_{j_q}}$
-	$\frac{\partial^2 k}{\partial \xi_i \chi_j}$	$\frac{k}{\ell^4}\left(A\ell^2 - B_1 B_2\right)$
-	A	$\sum \frac{\xi_{i_q} \chi_{j_q}}{r_{i_q} r_{j_q}}$
ℓ Derivative	$\frac{\partial^3 k_{\text{inv}}}{\partial \ell \partial \xi_i \partial \chi_j}$	$\sigma^2 \sum_{c,p} \left(\frac{\partial k_0}{\partial \ell} + \frac{\partial k_1}{\partial \ell} + \frac{\partial k_2}{\partial \ell} + \frac{\partial k_3}{\partial \ell}\right)$
-	$\frac{\partial k}{\partial \ell}$	$\frac{k \lvert\lvert \vec{d}_c - \vec{d}_p \rvert\rvert^2}{l^3}$
-	$\frac{\partial^2 k}{\partial \ell \partial \xi_i}$	$B_1 \left(\frac{1}{\ell^2} \frac{\partial k}{\partial \ell} - \frac{2k}{\ell^3}\right)$
-	$\frac{\partial^2 k}{\partial \ell \partial \chi_j}$	$-B_2 \left(\frac{1}{\ell^2} \frac{\partial k}{\partial \ell} - \frac{2k}{\ell^3}\right)$
-	$\frac{\partial^3 k}{\partial \ell \partial \xi_i \partial \chi_j}$	$\left(A\ell^2 - B_1 B_2\right)\left(\frac{\partial k}{\partial \ell} \frac{1}{\ell^4} - \frac{4k}{\ell^5}\right) + \frac{2kA}{\ell^3}$
σ Derivative	$\frac{\partial^3 k_{\text{inv}}}{\partial \sigma \partial \xi_i \partial \chi_j}$	$2\sigma \sum_{c,p} (k_0 + k_1 + k_2 + k_3)$

TABLE II. Quantities used to calculate the smoothed N-body force kernel and its derivatives.