Machine Learning: Science and Technology

PAPER • OPEN ACCESS

Enabling robust offline active learning for machine learning potentials using simple physics-based priors

To cite this article: Muhammed Shuaibi *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 025007

IOP Publishing Mach. Learn.: Sci. Technol. **2** (2021) 025007 https://doi.org/10.1088/2632-2153/abcc44

MACHINE LEARNING Science and Technology

RECEIVED
22 August 2020

REVISED
30 October 2020

ACCEPTED FOR PUBLICATION
19 November 2020

PUBLISHED
29 December 2020

PAPER

Enabling robust offline active learning for machine learning potentials using simple physics-based priors

Muhammed Shuaibi, Saurabh Sivakumar, Rui Qi Chen and Zachary W Ulissi

Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States of America

E-mail: zulissi@andrew.cmu.edu

Keywords: machine learning potentials, molecular simulations, neural networks, active learning, physics priors

Supplementary material for this article is available online

Abstract

Machine learning surrogate models for quantum mechanical simulations have enabled the field to efficiently and accurately study material and molecular systems. Developed models typically rely on a substantial amount of data to make reliable predictions of the potential energy landscape or careful active learning (AL) and uncertainty estimates. When starting with small datasets, convergence of AL approaches is a major outstanding challenge which has limited most demonstrations to online AL. In this work we demonstrate a Δ-machine learning (ML) approach that enables stable convergence in offline AL strategies by avoiding unphysical configurations with initial datasets as little as a single data point. We demonstrate our framework's capabilities on a structural relaxation, transition state calculation, and molecular dynamics simulation, with the number of first principle calculations being cut down anywhere from 70%–90%. The approach is incorporated and developed alongside AMP*torch*, an open-source ML potential package, along with interactive Google Colab notebook examples.

1. Introduction

The last decade has seen a surge in machine learning (ML) applications to material science, physics, and chemistry [1–7]. Characterizing a molecular system's potential energy surface (PES) has been a crucial step to the development of new catalysts and materials. Structure relaxation, molecular dynamics (MD), and transition state calculations rely almost entirely on an accurate PES to serve their functions. Machine learning potentials (MLPs) have demonstrated chemical accuracy at orders of magnitude faster computation times than traditional *ab-initio* methods including density functional theory (DFT) and coupled cluster single double triple [8]. However, these demonstrations have generally required large datasets and careful uncertainty estimates. More importantly, the models developed have struggled to generalize to new systems and faced convergence issues when adding data, making the practicality of their day-to-day applications challenging [5, 9–11]. The potential of active learning (AL) in molecular simulations has not been fully realized due to convergence and implementation challenges.

The careful curation of training datasets for accurate molecular simulations has recently given way to AL [12–15]. AL is the branch of ML concerned with systematically querying data points to be be part of the training set [16]. The iterative process queries new data, trains a model, and repeats until a model performance is achieved. AL methods are particularly useful when the cost of querying data is substantial—as in the case of computing DFT. There are two main classes of strategies with relevance to molecular simulations. In Online-AL, configurations are generated sequentially using a MLP and for each a decision is made whether to accept the estimate, perhaps using an uncertainty estimate. In Offline-AL, a pool of candidates is generated and a decision is made which of the pool to add to the training set.

Although there are many strategies available for both Online-AL and Offline-AL, both commonly assume that all generated candidates are feasible to be queried and that adding data will not reduce accuracy on previous training data. Both of these assumptions are difficult with MLP: DFT often fails to converge on far-from-equilibrium structures, and many MLPs suffer if even a small number of configurations with large

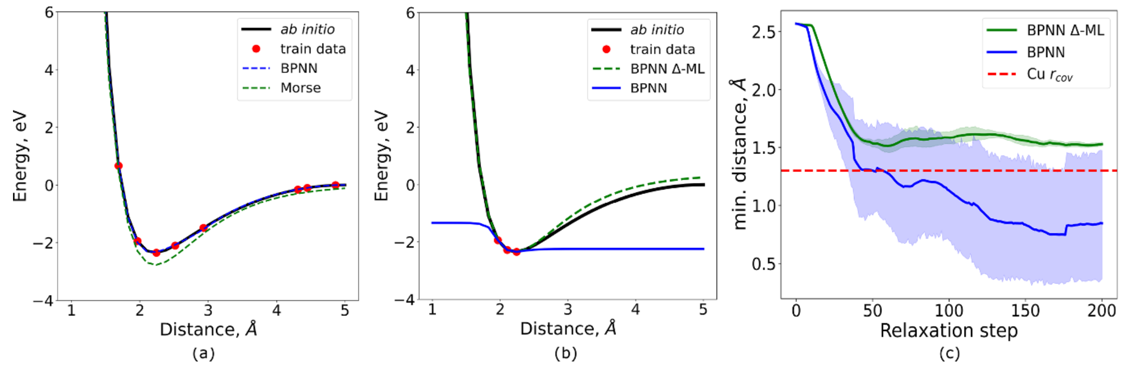

Figure 1. A traditional Behler–Parinello neural network (BPNN) trained to replicate the potential energy surface (PES) of a Cu–Cu bond with (a) a dataset spanning the PES and (b) a limited dataset trained with and without a Morse prior. (c) The minimum pair-wise distance of a structure relaxation carried out with a BPNN model, with and without the Morse prior. Relative to the covalent radius of Cu, our model consistently predicts more physically-consistent configurations as compared to the more unstable BPNN. Error bars represent the 95% confidence interval.

energies/forces are added to the training dataset [14]. These concerns are especially problematic when dealing with little to no initial data. The most common approach to address these challenges is to carefully monitor uncertainty in the AL process and prevent extrapolation to unphysical regions. This strategy is relatively straightforward to implement in Online-AL: if the uncertainty estimate is below a threshold, accept the prediction, otherwise run the DFT calculations. If the step size is small enough, the new configuration should not be so different from configurations in the training set. However, in Offline-AL it is difficult to ensure the queried configurations will converge with DFT and will not contaminate the dataset once added. Instead of solving this problem, we show that it is possible to mostly fix the underlying issues leading to unrealistic configurations.

In this work, we demonstrate that stable convergence in Offline-AL with MLP is possible by adding simple repulsive potentials and robust training procedures. This approach is implemented for the common combination of Behler–Parinello MLP fingerprints with neural network atomic energy models [17]. We show that a Δ-ML approach with a base pairwise Morse potential and linear mixing rules is capable of sufficiently capturing the repulsive interactions between atoms that lead to DFT errors. Since this Morse potential is not responsible for capturing the full potential, the parameterization only needs to be done once for each element. We demonstrate this approach for several types of calculations common in catalysis: structure relaxations, MD, and transition state calculations. In each case, convergence with the addition of training data is essentially impossible with the base potential and well behaved with the Δ-ML approach. In most cases this process allows for a reduction of 70%–90% in the number of DFT single-point evaluations necessary. This process is further improved using standard neural network training approaches in the ML community to reduce the impact of random initial weights on small datasets. All of these are demonstrated in the open-source and accessible AMP*torch* GitHub repository with Google Colab ASE examples [18, 19].

2. Methods

The ML community continues to make advancements in the optimization and implementation of neural network-based models [20–22]. To leverage some of these approaches, we employ a Behler–Parinello neural network (BPNN). BPNNs construct element-specific neural networks with the energy of the system the sum of atomic energy contributions. Per-atom forces are directly obtained from the negative gradient of the energy with respect to the atomic positions. We refer the readers to several reviews for a more detailed discussion on the BPNN model [5, 6, 17]. Additionally, neural network-based models do not suffer the same kernel selection and scalability challenges that can come with Gaussian processes (GP) and other Bayesian models [23]. Training neural networks, however, can be an extremely challenging task we hope to address in this work.

In the presence of an abundance of data, BPNN-like models have shown great success in replicating the PES of various systems [4, 6, 24]. In the small data limit, however, neural network-based models are unable to successfully characterize the energy surface, figure 1(b). More notably, model predictions are entirely 'physics-free', such that simple repulsive interactions are only ever learned by the model once enough data has been provided. As a result, a considerable amount of time may be wasted learning simple, widely understood characteristics of the PES. Hybrid physics-based ML models can provide an important path forward to making reliable, physically-consistent discoveries in the sciences [25, 26]. To address this, we incorporate a Δ-ML approach [27, 28] to learn the correction, E_{NN}, between a simple Morse potential, $\Delta E_{\mathrm{morse}}$, and *ab-initio* level theory—namely, DFT, $\Delta E_{\mathrm{DFT}}(\mathbf{x})$:

$$\begin{aligned}
\Delta E_{\mathrm{DFT}}(\mathbf{x}) &= E_{\mathrm{DFT}}(\mathbf{x}) - E_{\mathrm{DFT}}(x_{\mathrm{ref}}) \\
\Delta E_{\mathrm{morse}}(\mathbf{x}) &= E_{\mathrm{morse}}(\mathbf{x}) - E_{\mathrm{morse}}(x_{\mathrm{ref}}) \\
E_{\mathrm{NN}}(\mathbf{x}) &= \Delta E_{\mathrm{DFT}}(\mathbf{x}) - \Delta E_{\mathrm{morse}}(\mathbf{x}) \\
E_{\mathrm{morse+NN}}(\mathbf{x}) &= \Delta E_{\mathrm{morse}}(\mathbf{x}) + E_{\mathrm{DFT}}(x_{\mathrm{ref}}) + E_{\mathrm{NN}}(\mathbf{x})
\end{aligned}$$

where $E_{\mathrm{DFT}}(x_{\mathrm{ref}})$ and $E_{\mathrm{morse}}(x_{\mathrm{ref}})$ correspond to the reference energies necessary to correct for differences in their absolute energies. Reference energies are computed from a same arbitrary structure, x_{ref}; the dataset's first structure was used in our applications. Per-element parameters of the Morse potential, D_e, r_e, and a, are fitted to DFT data *a priori*. A more detailed description of the fitting procedure is included in the (SI). By leveraging the Morse potential as the backbone to the model, the ML component is allowed to learn the remaining functional form while still capturing physics-based repulsive interactions previously missed. Additionally, learning a correction can allow the neural network to learn a much smoother function than the underlying PES, improving training stability and convergence.

We illustrate the benefits of this simple Morse potential by running a structure relaxation of carbon on copper (C/Cu) with our model trained on a single image (figure 1(c)). The minimum pair-wise distances of the resulting trajectory are compared to those not employing a Morse potential. Our model consistently predicts configurations above the covalent radius of copper, a good indication repulsive forces are being captured. On the other hand, a traditional BPNN shows wide variations while on average predicting configurations well below the more stable covalent radius.

The fitting of MLPs is an important process in our AL framework, as they are responsible for generating candidates for training data. A poorly fit MLP may generate unfeasible candidates that DFT can not converge upon. This is especially true when working without a physics-based potential. Working within a small data regime allows us to leverage quasi-Newton optimizers, namely limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS). LBFGS and other second-order optimizers provide us with improved convergence of our model training over standard first-order methods such as stochastic gradient descent (SGD) and Adam. This advantage, however, is only really feasible in the small data limit where the computational cost of such methods can be afforded. Additionally, we incorporate a cosine annealing learning rate scheduler with warm restarts [29] to aid in the convergence of the Offline-AL framework. A more detailed comparison can be found in the supplementary information (which is available online at stacks.iop.org/MLST/2/025007/mmedia).

Similar to previous works [12, 13], our Online-AL framework begins with little to no data and must identify the right points to query and improve the model over the course of a molecular simulation (figure 2). Rather than relying on kernel-based models, our Online-AL framework utilizes the proposed physics-coupled BPNN. We incorporate bootstrap-ensembling, or bagging, in order to quantify our model's uncertainty. Bagging involves training multiple, randomly initialized, independent models with training sets randomly sampled, *with* replacement, from an original dataset [30]. Predictions and uncertainty estimations are then calculated from the ensemble statistics.

An Offline-AL can offer model and computational advantages over Online-AL frameworks. Rather than making query decisions in a dynamic process, we present a method to select from a pool of generated candidates. Prior works have incorporated Offline-AL to various extents. Sivaraman *et al* [31] used AL to downselect from an existing hafnium dioxide *ab initio* MD simulation to train a GP model. Novikov *et al* [32] used AL and moment tensor potentials to run atomistic simulations. We show, however, that a standard neural network is unable to follow a similar framework without careful modifications. Rossi *et al* [33] used an ensemble of neural networks to estimate uncertainty along an atomistic simulation. Having begun from an extensively sampled training dataset, their need for retraining was avoided, a problem we address for neural networks in the small data regime. While the use of AL has shown incredible success in training models with fractions of the dataset, it assumes such datasets exist to begin with. We propose a framework to enable accurate atomistic simulations beginning with as little as a single data point. We accomplish this by iteratively running an ML-driven molecular simulation. After each iteration, a querying strategy samples from the generated trajectory. Queried points are then evaluated with DFT, added to an original dataset, and the ML model retrained (figure 2). The process is repeated until a defined convergence criteria is met. Despite the ML model resulting in inaccurate simulations early on, diverse, informative configurations are generated to train the ML model. In dealing with a pool of query candidates, the framework allows us to explore alternative querying strategies other than uncertainty estimates of Online-AL [16]. The reliance on

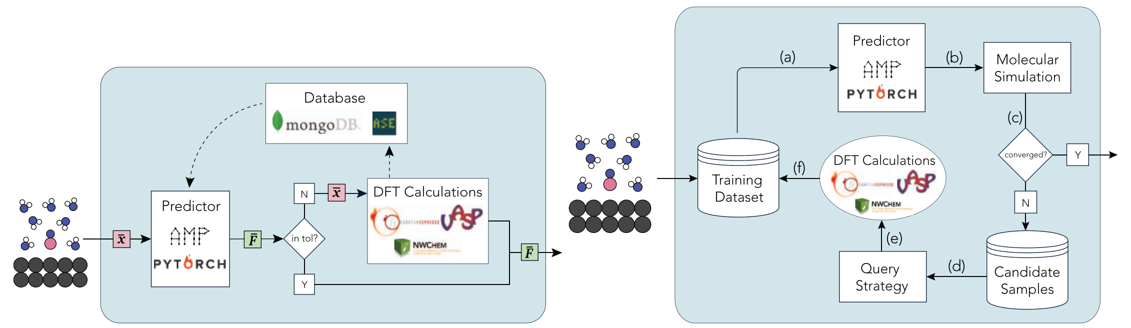

Figure 2. Online and Offline active learning frameworks to accelerate molecular simulations. Left: Online-AL. At each time step, our ML model makes a prediction of the energy and forces and assesses the uncertainty of its estimate. If confident, the ML results are used to take a step in the molecular simulation. Otherwise, a DFT call is made, added to a database, and the model retrained. Right: Proposed Offline-AL. (a) An initial training dataset is used to train the ML model; (b) the trained ML model runs the atomistic simulation of interest; (c) termination if converged, otherwise, the generated data is stored as a pool of potential candidates; (d) a query strategy identifies what points to be added to the training set; (e) *ab-initio* calculations are performed on selected candidates; (f) queried points are added to the training set. Repeat until convergence is reached.

Table 1. Comparison of different offline active learning batching scenarios on the structural relaxation of C/Cu(100). At each iteration, a varying number of queries are randomly made from the generated relaxation. A tradeoff in performance and the number of samples per iteration is observed for a fixed total number of DFT calls = 20. All models trained here incorporated the proposed Morse prior.

Batching scenario			
Iterations	Samples per iteration	Energy MAE (eV)	Structure MAE (Å)
20	1	0.0063	0.0037
10	2	0.0069	0.0063
5	4	0.0080	0.0067

uncertainty estimates can pose more fundamental questions surrounding energy conservation from a retrained potential [32] and how trustworthy a model's estimates really are [34].

We demonstrate the proposed framework on several common catalysis applications: structure relaxations, transition-state calculations, and MD with system sizes between ∼12 and 30 atoms. A random sample query strategy is introduced in the Offline-AL schemes to demonstrate the effectiveness of even the simplest of query strategies over Online-AL. More problem-specific query strategies are proposed for structural relaxations and transition-state calculations, further improving the convergence. To show the generality of this approach in small-data applications, we also use two common DFT packages: the Vienna *ab initio* Simulation Package (VASP) and Quantum Espresso (QE) [35–37]. The use of QE allows for interactive and open demonstrations of this approach. Several Google Colab notebooks have been included in the supplementary information, allowing users to easily experiment and explore new systems with AMP*torch* and QE without needing to locally install and manage dependencies.

3. Results and discussion

A structural relaxation is performed for C/Cu(100) with cell size $2 \times 2 \times 3$. An initial guess of 3 Å from the surface is made for the adsorbate. Periodic boundary conditions are applied in the x and y directions and the last slab layer is fixed from relaxations. The training dataset begins with a single initial structure.

Performance is measured by the final structure and energy mean-absolute-errors (MAEs). A random sample query strategy selects configurations from the generated relaxations to be queried. We run the Offline-AL framework under a variety of batching scenarios, terminating after N iterations, sampling M configurations per iteration, for an arbitrary total of $NM = 20$ DFT calls. Results are summarized in table 1.

Under the above random query strategy, systematic termination of the Offline-AL loop is quite heuristic. To address this, we incorporate alternative query and termination strategies—quasi-random and uncertainty sampling. In quasi-random, at each iteration, in addition to a random configuration, the predicted relaxed structure is also queried. Similarly, uncertainty sampling samples the k-most uncertain points in addition to the relaxed structure. In both strategies, if the predicted relaxed structure's max per-atom force, as evaluated by DFT, is below the optimizer's convergence criteria, the AL loop is terminated. Otherwise, the configurations are added to the original dataset, and the framework cycles. In querying the model's predicted relaxed structure we are assured in our framework's ability to accurately reach a local minima.

Table 2. Summary of the various strategies' performance on the structural relaxation of C/Cu(100). The effects of the Morse prior on the convergence of both the offline and online active learning are also shown. The querying strategy employed by the Offline-AL framework relies on a quasi-random strategy, additionally sampling and assessing convergence on the framework's generated relaxed structure.

Framework (tolerance)	MLP	Energy MAE (eV)	Structure MAE (Å)	DFT calls
DFT	—	—	—	51
Offline-AL (0.03 eV Å^{-1})	BPNN Δ-ML	0.0039	0.0032	17
Offline-AL (0.05 eV Å^{-1})	BPNN Δ-ML	0.0049	0.0059	15
Offline-AL (0.05 eV Å^{-1})	BPNN only		does not converge	
Online-AL (0.05 eV Å^{-1})	BPNN Δ-ML	0.0073	0.0107	30
Online-AL (0.05 eV Å^{-1})	BPNN only	0.2884	0.0263	22

Table 3. Comparison of different offline active learning query strategies on the structural relaxation of C/Cu(100). All models trained here incorporated the proposed Morse prior.

Query strategy	Energy MAE (eV)	Structure MAE (Å)	DFT calls
Random	0.0063	0.0037	20
Quasi-random	0.0049	0.0059	15
Uncertainty sampling	0.0061	0.0050	19

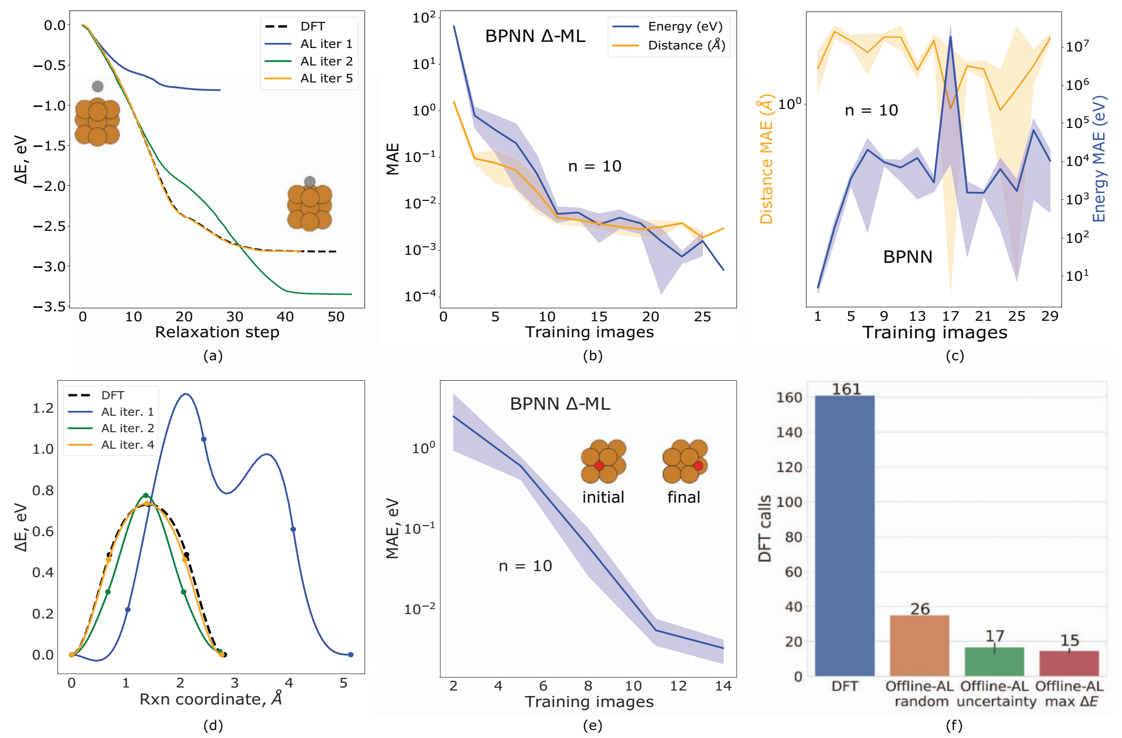

Figure 3. Offline-AL applications to structural relaxations and transition state calculations. (a) Evolution of the structural relaxation of C on Cu(100) over a few cycles of the Offline-AL. (b) Relaxed structure and energy learning curves of the Offline-AL framework, using the BPNN Δ-ML model. (c) Convergence instability associated with not incorporating the Morse potential prior in an Offline-AL context. (d) Evolution of the transition state calculation for the surface diffusion of O on Cu(100). Despite the poor performance of the first iteration, the framework is able to recover and converge to an accurate prediction. (e) Learning curve associated with the energy barrier of the oxygen diffusion example of (d). (f) Total number of DFT calls queried by the Offline-AL under different querying strategies for the energy barrier associated with the diffusion of oxygen on copper. Error bars represent the 95% confidence interval.

We compare the performance of this Offline-AL scheme and Online-AL with and without the Δ-ML in table 2. Offline-AL and Online-AL tolerances correspond to the max per-atom force termination criteria and max force variance tolerated by the ensemble, respectively. Force termination criteria of 0.03 and 0.05 eV Å^{-1} are compared to explore the tradeoff between accuracy and number of DFT calls. Online-AL was empirically set to query a DFT call when the ensemble based force uncertainty reached above a threshold of 0.05 eV Å^{-1}. The energy and structure MAE associated with the system's initial structure is 2.82 eV and 0.15 Å, respectively. Our best-performing framework, Offline-AL with Δ-ML (0.03 eV Å^{-1}), reported average energy and structure MAEs of 0.0039 eV and 0.0032 Å with 17 total DFT calls—a 66.7% reduction. Without

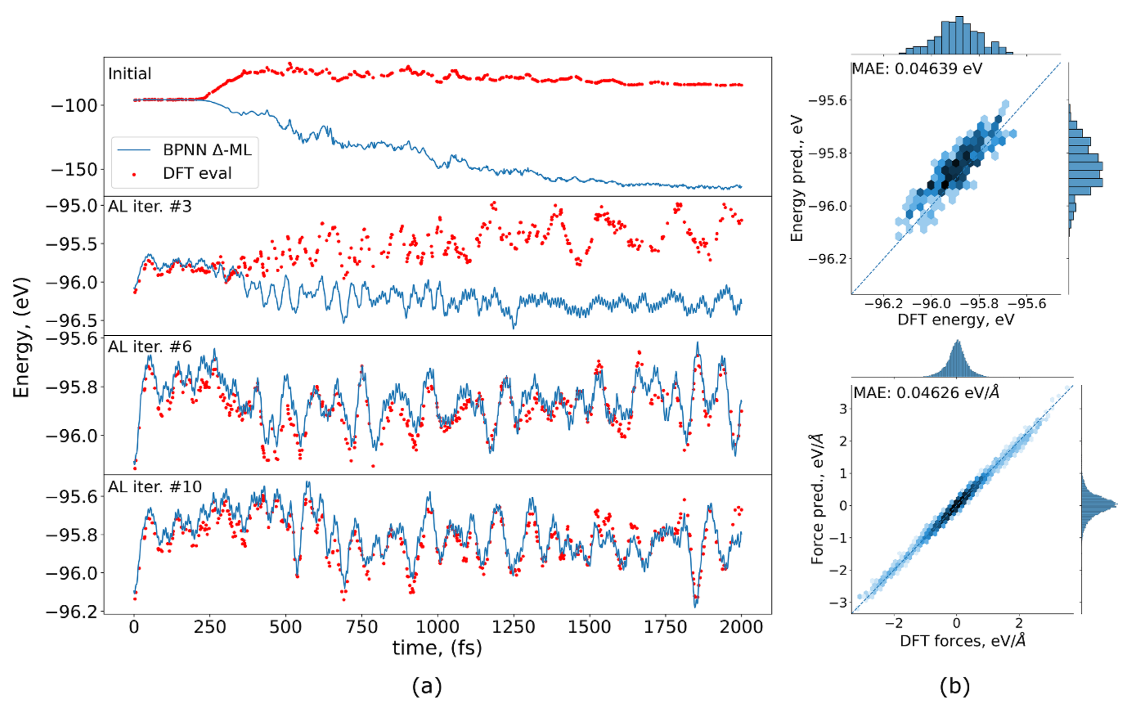

Figure 4. Offline-AL demonstration to a 2ps MD simulation of CO on Cu(100). (a) Evolution of the MD trajectory over several iterations of the active learning framework. We verify the effectiveness of our framework by randomly sampling configurations and comparing DFT evaluated energy and forces with that of our model's predictions. (b) Parity plots associated with the DFT-evaluated configurations and our model's predictions of the 6th iteration, demonstrating good agreement. Shading was scaled logarithmically with darker shading corresponding to a higher density of points.

the inclusion of the Morse prior, a standard BPNN was unable to converge, generating configurations that DFT was unable to evaluate in almost all our experiments. We compare several query strategies in table 3, demonstrating similar success in all scenarios.

Next, we demonstrate an application to transition state calculations, specifically nudge-elastic-band (NEB) methods [38, 39]. NEB calculations require defining the initial and final structures for the transition state to be calculated. Machine-learning accelerated NEB calculations have typically relied on *ab-initio* relaxed initial and final structures, a costly step of a NEB calculation [40]. In fixing the initial and final structures, the ML objective is simplified to an interpolation problem. We demonstrate our framework's ability to accelerate the complete NEB calculation, including initial and final structure relaxations, to find the surface diffusion energy barrier of oxygen on Cu(100). To illustrate our framework, we used five images to build the NEB, including the initial and final states which have not been relaxed previously. The initial training dataset includes only the unrelaxed initial and final structures.

The convergence evolution of our Offline-AL framework is illustrated in figure 3(d), approaching the true energy barrier after a few iterations. Similarly, convergence was not achieved, with often failing DFT evaluations, without the inclusion of the Morse prior. A simple random strategy is first employed. Here the images are randomly sampled from generated NEBs and evaluated using DFT before being added to the training data. Termination is achieved after a fixed number of iterations. Additionally, we compare the efficiency of two more systematic querying and early stopping methodologies. An uncertainty sampling strategy queries images with the highest uncertainty, which are then evaluated with DFT and added to the training data. Termination is reached when the difference between the predicted energies from ML and DFT at the saddle point are less than a tolerance. An additional strategy is also tailor-made for the NEBs, wherein the highest energy point, along with the initial and final points, are sampled at each iteration. The loop is terminated once the difference between the ML-predicted energies and DFT-evaluated energies of the three points is less than a specified threshold. All three cases demonstrate a significant reduction in the number of DFT calls required to construct the NEB as shown in (figure 3(f)).

ML surrogates to DFT are considerably favorable in the context of long time-scale simulations, namely MD. Unlike structural relaxations, MD simulations are typically carried out on orders of magnitudes more steps. Several works have addressed these challenges through GP-based Online-AL frameworks [12, 13]. We demonstrate that our proposed Offline-AL framework is capable of converging to an accurate MD simulation. A 2ps MD simulation of CO on Cu(100) in a 300K NVT ensemble is used for our demonstration.

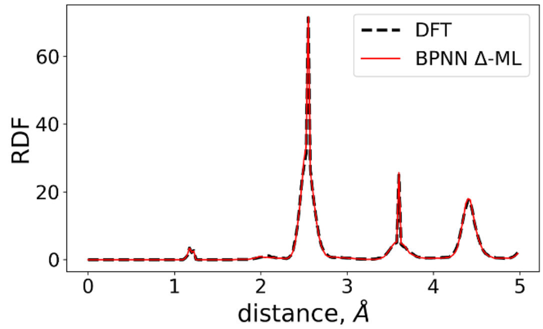

Figure 5. Radial distribution function (RDF) of the ground truth DFT and our framework's 6th iteration for the MD simulation of CO/Cu(100). Demonstrating good consistency even before the allotted number of iterations.

Beginning with a dataset containing only the initial structure, our framework cycles for several iterations, randomly querying 50 configurations for a total of 500 DFT calls by the end of our experiment. Unlike structural relaxations with a well defined target, MD simulations are more stochastic in nature and are unlikely to follow an identical trajectory over multiple iterations. To demonstrate the effectiveness of our framework, we verify the performance, at each iteration, by randomly sampling 400 configurations from our ML predicted trajectory and validate their corresponding energy and force predictions with DFT. We illustrate the iterative convergence of our framework in figure 4. Despite the upper limit of ten iterations, we observe good agreement with DFT by iteration 6—a reduction of 85% in DFT calls. Additionally, we demonstrate consistency in the radial distribution function of our framework's generated simulation to that of the original DFT simulation (figure 5). Although our demonstration takes place at a moderate 300 K, the extremely limited data of our ML model results in highly perturbed configurations within the first few iterations of the simulation. Without the presence of our proposed Morse prior these configurations are far off equilibrium and often fail to converge by DFT. A similar demonstration under a more perturbed, higher temperature system is included in the supplementary information, with comparable success as early as the 3rd iteration—a 92% reduction in DFT calls.

4. Conclusion

The development of accurate and reliable MLP has been a challenging task for the community. The careful curation of datasets is especially difficult in trying to generalize to new systems. AL has provided promising results in accelerating molecular simulations while minimizing risks of extrapolation. Neural-network based models, however, have struggled with such demonstrations for their reliance on large amounts of data. As deep learning research continues to make significant strides, understanding how to better incorporate neural network-based MLPs into AL pipelines can help provide more accurate and robust models.

This manuscript presented a neural network-based Offline-AL framework to accelerate a variety of molecular simulations beginning with extremely limited data. We introduced a physics-based prior, Morse potential, into our model in a Δ-ML manner, to capture basic repulsive interactions crucial in the convergence of our framework. We demonstrate the framework's ability to accurately converge simulations including structural relaxations, MD simulations, and transition-state calculations. In each of these, the proportion of reduced DFT calls were 71%, 75%, and 91%, respectively. The framework presented is extremely flexible, allowing users to define their own querying strategies, termination criteria, and incorporate their own, more complex molecular simulations as they wish to accelerate with AMP*torch*. Similar to other works, the nature of our AL framework introduces assumptions and limitations surrounding the feasibility of DFT queries. While our framework helps in accelerating atomistic simulations, its applicability is limited by the time it takes to query DFT points. Systems in which DFT calls may be infeasible (10 000+ atoms or far-from-equilibrium) will fail under this and other AL strategies, leaving opportunities for the development of robust models trained on large datasets [41]. At this time we make no guarantees that the performance of the ML model will always improve when a queried data point is added to the dataset. Our experiments recognize this as a particular issue in the small data regime but this was often mitigated in our work by the presence of the Morse potential and more sophisticated learning rate schedulers, where otherwise it would have failed. Future directions will explore more systematic querying strategies and termination criteria to further accelerate the framework while being robust to larger, more complex systems still compute-feasible under DFT. Additionally, exploring alternative model priors and adversarial training techniques can help improve the performance, consistency, and generalizability of AL frameworks [42, 43].

5. Calculation settings

Single-point DFT calculations were performed in Quantum Espresso (QE) [37] implemented in ASE [44]; using the PBE exchange-correlation functional [45]; a plane wave basis set with an energy-cutoff of 500 eV; k-points of $4 \times 4 \times 1$; and the pseudopotentials provided by Garrity *et al* [46]. The same settings were also used for DFT calculations in fitting the Morse potential parameters. The following tools and settings were used for our DFT calculations: VASP 5.4.4.18 [35, 36]; using the PBE exchange-correlation functional; a plane wave basis set with an energy-cutoff of 400 eV; and k-points of $4 \times 4 \times 1$. VASP was used for all structure relaxation and MD examples and QE for the NEB examples. AMP*torch* [18] was used for all ML and AL components of the framework.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary information files).

Acknowledgment

We acknowledge the support from the U.S. Department of Energy, Office of Science, Basic Energy Sciences Award #DE-FOA-0001912. Additionally, we thank Andrew Peterson and A J Medford for their thoughtful discussions.

ORCID iD

Zachary W Ulissi https://orcid.org/0000-0002-9401-4918

References

[1] Artrith N and Kolpak A M 2014 *Nano Lett.* **14** 5
[2] Rupp M, Ramakrishnan R and von Lilienfeld O A 2015 *J. Phys. Chem. Lett.* **6** 16
[3] Natarajan S K and Behler J 2016 *Phys. Chem. Chem. Phys.* **18** 28704–28725
[4] Peterson A A 2016 *J. Chem. Phys.* **145** 7
[5] Behler J 2016 *J. Chem. Phys.* **145** 17
[6] Khorshidi A and Peterson A A 2016 *Comput. Phys. Commun.* **207** 310–24
[7] Bartók A P, Payne M C, Kondor R and Csányi G 2010 *Phys. Rev. Lett.* **104** 136403
[8] Zuo Y *et al* 2020 *J. Phys. Chem.* A **124** 731–45 (arXiv:1906.08888)
[9] Chen C, Zuo Y, Ye W, Li X, Deng Z and Ong S P 2020 *Adv. Energy Mater.* **10** 8
[10] Mueller T, Hernandez A and Wang C 2020 *J. Chem. Phys.* **152** 5
[11] Schleder G R, Padilha A C M, Acosta C M, Costa M and Fazzio A 2019 *J. Phys.: Mater.* **2** 3
[12] Vandermause J, Torrisi S B, Batzner S, Xie Y, Sun L, Kolpak A M and Kozinsky B 2020 *npj Comput. Mater.* **6** 20
[13] Jinnouchi R, Lahnsteiner J, Karsai F, Kresse G and Bokdam M 2019 *Phys. Rev. Lett.* **122**
[14] Garrido Torres J A, Jennings P C, Hansen M H, Boes J R and Bligaard T 2019 *Phys. Rev. Lett.* **122** 156001
[15] Del Río E G, Mortensen J J and Jacobsen K W 2019 *Phys. Rev.* B **100** (arXiv:1808.08588)
[16] Settles B and Learning M 2010
[17] Behler J and Parrinello M 2007 *Phys. Rev. Lett.* **98** 146401
[18] Shuaibi M, Comer B, Sivakumar S, Lei X R, Chen R Q, Adams M and Musa E 2020 AMPtorch: Atomistic Machine Learning Package (AMP) - PyTorch (https://github.com/ulissigroup/amptorch)
[19] Shuaibi M and Sivakumar S 2020 Physics enabled convergence of offline active learning with machine learning potentials (https://github.com/ulissigroup/OfflineAL-for-MLPs-manuscript)
[20] Loshchilov I and Hutter F 2019 *5th Int. Conf. on Learning Representations, ICLR 2017 - Conf. Track Proc* (arXiv:1608.03983)
[21] Fey M and Lenssen J E 2019 (arXiv:1903.02428)
[22] Paszke A *et al* 2017 Automatic differentiation in PyTorch *Advances in Neural Information Processing Systems* (https://openreview.net/pdf/25b8eee6c373d48b84e5e9c6e10e7cbbbce4ac73.pdf)
[23] Bartõk A P and Csányi G 2015 *J. Quantum Chem.* **115** 16
[24] Schran C, Behler J and Marx D 2020 *J. Chem. Theory Comput.* **16** 1 (arXiv:1908.08734)
[25] Willard J, Jia X, Xu S, Steinbach M and Kumar V 2020 Integrating physics-based modeling with machine learning: a survey (arXiv:2003.04919)
[26] Karpatne A, Watkins W, Read J, and Kumar V 2017 (arXiv:1710.11431)
[27] Ramakrishnan R, Dral P O, Rupp M and Von Lilienfeld O A 2015 *J. Chem. Theory Comput.* **11** 5
[28] Zhu J, Vuong V Q, Sumpter B G and Irle S 2019 *MRS Commun.* **9** 3
[29] Loshchilov I and Hutter F 2016 Sgdr: Stochastic gradient descent with warm restarts (arXiv:1608.03983)
[30] Peterson A A, Christensen R and Khorshidi A 2017 *Phys. Chem. Chem. Phys.* **19**
[31] Sivaraman G, Krishnamoorthy A N, Baur M, Holm C, Stan M, Csányi G, Benmore C and Vázquez-Mayagoitia A 2020 *npj Comput. Mater.* **6** 104
[32] Novikov I S, Gubaev K, Podryabinkin E V and Shapeev A V 2020 The MLIP package: Moment tensor potentials with MPI and active learning (arXiv:2007.08555)

[33] Rossi K, Jurásková V, Wischert R, Garel L, Corminbœuf C and Ceriotti M 2020 *J. Chem. Theory Comput.* **16** 8

[34] Tran K, Neiswanger W, Yoon J, Zhang Q, Xing E and Ulissi Z W 2020 *Mach. Learn.: Sci. Technol.* **1** 2 (arXiv:1912.10066)

[35] Kresse G and Hafner J 1993 *Phys. Rev.* B **48** 17

[36] Kresse G and Furthmüller J 1996 *Comput. Mater. Sci.* **6** 1 (http://www.sciencedirect.com/science/article/pii/0927025696000080)

[37] Giannozzi P *et al* 2009 *J. Phys. Conden. Matter* **21** 39 (arXiv:0906.2569)

[38] Henkelman G, Uberuaga B P and Jónsson H 2000 *J. Chem. Phys.* **113** 22

[39] Henkelman G and Jónsson H 2000 *J. Chem. Phys.* **113** 22

[40] Ang S J, Wang W, Schwalbe-Koda D, Axelrod S and Gomez-Bombarelli R 2020 *ChemRxiv* (10.26434/chemrxiv.11910948.v2)

[41] Chanussot L *et al* 2020 The open catalyst 2020 (oc20) dataset and community challenges (arXiv:2010.09990)

[42] Fan J and Li W 2020 Adversarial training and provable robustness: a tale of two objectives (arXiv:2008.06081)

[43] Steinhardt J, Koh P W and Liang P 2017 Certified defenses for data poisoning attacks (arXiv:1706.03691)

[44] Hjorth Larsen A *et al* 2017 *J. Phys. Conden. Matter* **29** 27

[45] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 18

[46] Garrity K F, Bennett J W, Rabe K M and Vanderbilt D 2014 *Comput. Mater. Sci.* **81** 446–452 (arXiv:1305.5973) (http://www.sciencedirect.com/science/article/pii/S0927025613005077)