

I. SUPPLEMENTARY INFORMATION

A. High-temperature MD

In a similar manner to Figure 4 of the main text, we demonstrate our framework’s ability to successfully converge to an accurate high-temp MD despite the highly perturbed sampled configurations. This same experiment was unsuccessful without the inclusion of the more potential prior we have introduced in this work. Beginning with a dataset containing a single structure, we run the proposed framework over a 2ps MD simulation of CO on Cu(100) in a 800K NVT ensemble. We illustrate our results in Figure S.1, with good agreement as early as the 3rd iteration - suggesting that the sampled highly-perturbed structures aid in reaching a converged simulation in much fewer iterations.

B. Interactive examples

Several interactive Google Colab notebooks have been prepared to allow readers to conveniently explore the proposed methods. Accelerated structural relaxations and transition state calculations can be found at Ref [1]. Random query strategies are used to demonstrate the effectiveness of even the simplest of strategies. We encourage users to explore query and termination strategies that best suits their application of choice. DFT calculations are performed directly in the notebook examples via a GPU-enabled Quantum Espresso package.

C. Morse parameters fitting

The Morse potential was selected for our primarily bonded, catalytic systems. Parameters of the Morse potential, D_e , r_e , and a , corresponding to well depth, equilibrium distance, and well width were computed in the following manner for a given element, X:

1. Lone atomic energies, E_X , obtained through singlepoint DFT calculations;
2. Diatomic atoms relaxed to obtain a relaxed state energy, E_{X_2} , and equilibrium distance, r_e .

3. Well depth, D_e , is calculated as follows:

$$D_e = -(E_{X_2} - 2 * E_X) \quad (\text{S.1})$$

4. Diatomic bond stretched and corresponding DFT points fit to Morse potential functional form (S.2) to obtain a Figure (S.2).

$$E_{morse} = D_e(e^{-2a(r-r_e)} - 2e^{-a(r-r_e)}) \quad (\text{S.2})$$

To make use of the Morse potential for multi-element systems, linear mixing rules are utilized to compute element pair parameters. Adapted from Yang, et al.[2] the Morse potential is rewritten and parameter combinations applied accordingly (S.3-S.6)

$$E_{morse} = D_e(\exp[-\frac{2C}{\sigma}(r - r_e)] - 2\exp[-\frac{C}{\sigma}(r - r_e)]) \quad (\text{S.3})$$

$$D_{AB} = \sqrt{D_A D_B} \quad (\text{S.4})$$

$$r_{e,AB} = \frac{r_{e,A} + r_{e,B}}{2} \quad (\text{S.5})$$

$$\sigma_{AB} = \frac{\sigma_A + \sigma_B}{2} \quad (\text{S.6})$$

Where $C = \ln 2/(r_e - \sigma)$ and σ corresponds to $E_{morse}(\sigma) = 0$. Although more sophisticate combination rules exist [2], the accuracy of our Morse potential is not crucial for the success of our framework as it is meant to provide some guidance to the model.

D. Convergence

The convergence of the Offline-AL loop can be accelerated through the use of a learning rate scheduler. Figure S.3 compares the learning curves of AL frameworks with and without a learning rate scheduler, ceteris paribus. We demonstrate that a cosine annealing scheduler with warm restarts [3] was able to assist the convergence by smoothing out the learning curve and requiring fewer training images to reach a similar level or error.

[1] (2020), <https://github.com/ulissigroup/OfflineAL-for-MLPs-manuscript>.
 [2] L. Yang, L. Sun, and W. Q. Deng, Journal of Physical Chemistry A (2018), 10.1021/acs.jpca.7b11252.

[3] I. Loshchilov and F. Hutter, “Sgdr: Stochastic gradient descent with warm restarts,” (2016), arXiv:1608.03983 [cs.LG].

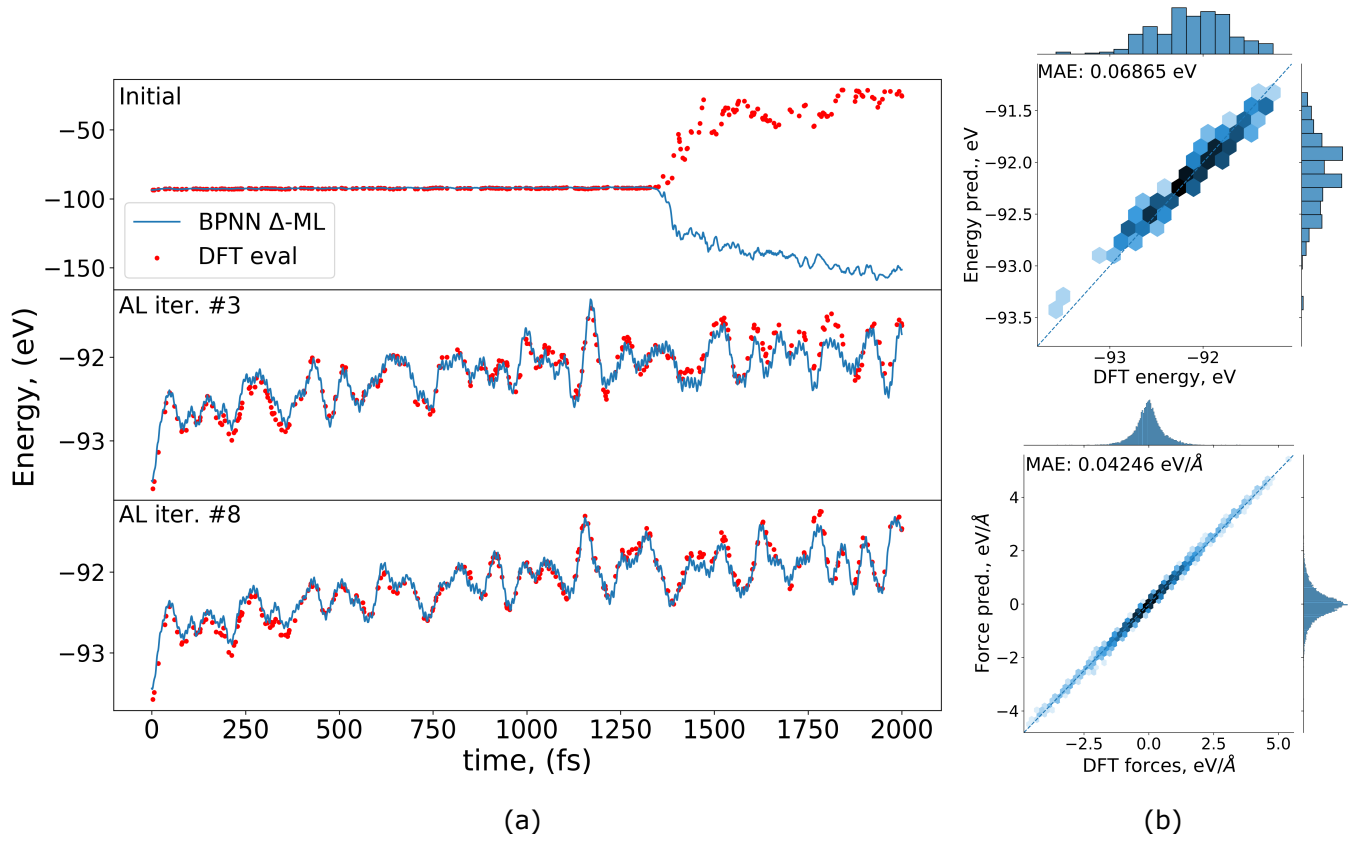


FIG. S.1. Offline-AL demonstration to a 2ps MD simulation of CO on Cu(100) at 800K **(a)** Evolution of the MD trajectory over several iterations of the active learning framework. We verify the effectiveness of our framework by randomly sampling configurations and comparing DFT evaluated energy and forces with that of our model's predictions. **(b)** Parity plots associated with the DFT evaluated configurations and our model's predictions on the 8th iteration, demonstrating good agreement. Shading was scaled logarithmically with darker shading corresponding to a higher density of points.

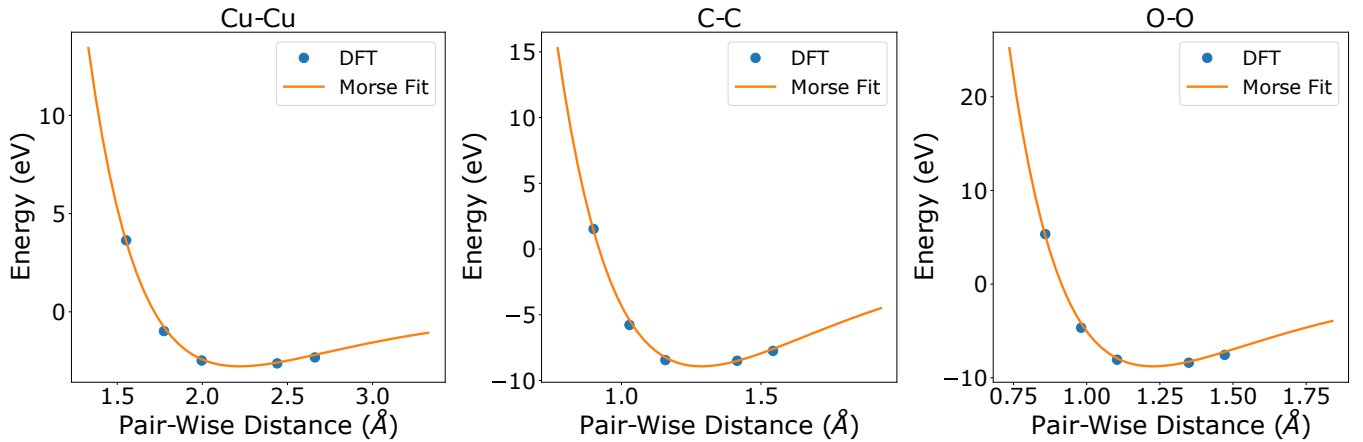


FIG. S.2. Morse parameters are obtained by fitting DFT points near the equilibrium distance to equation S.2. Sample fittings are illustrated for copper, carbon, and oxygen.

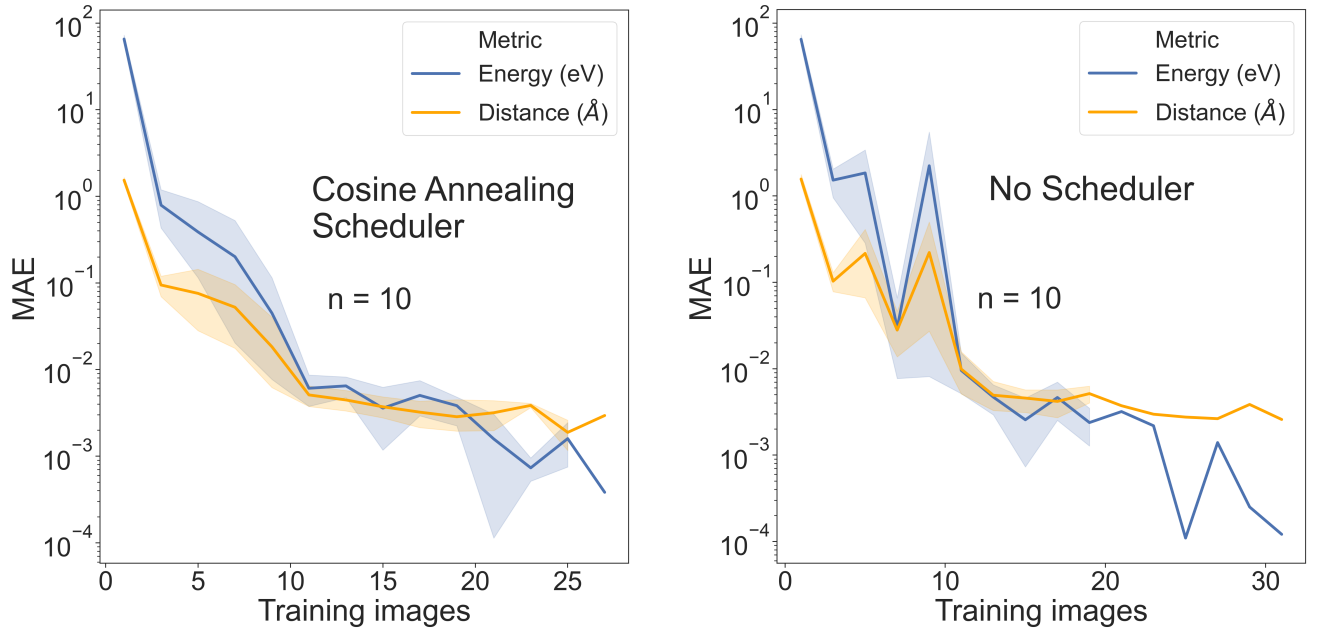


FIG. S.3. Offline-AL convergence of our BPNN Δ -ML is compared with and without a learning rate scheduler. The use of a scheduler, particularly with small data, enables our framework to converge more reliably to the local minima.