Constant-Potential Machine Learning Force Field for Electrochemical Interface

Anonymous Author(s)
Affiliation
Address
email

Abstract

Better understanding and prediction of electrochemical interface requires large-scale atomistic simulations. Machine learning force field (MLFF) has proven to be an effective approach. However, current MLFFs typically do not account for the effect of electrode potential, which requires treating interface electrons with grand canonical ensemble. Here we develop a constant potential MLFF (CP-MLFF) based on equivariant graph neural network and implement it into MACE. Specifically, we design an architecture which can take the number of electrons as input and accurately predict the Fermi level. The CP-MLFF allows us to examine the convergency of electrochemical barrier with respect to sampling, which we demonstrate through the example of CO_2 reduction on Ni-N-C catalyst. Our work provides a useful method and tool enabling accurate and efficient large-scale simulation of electrochemical interface.

1 Introduction

Electrochemical interface is at the center of many technologies to address energy and environmental challenges, such as water splitting, CO_2 conversion, nitrate reduction, and oxygen reduction [1–5]. Having an accurate atomic level understanding of the interface is critical to further developing these technologies. Density functional theory (DFT) has been widely used to simulate the interface at atomic level [6–12]. However, its high computational cost has limited the simulations to a small number of atoms and configurations, which are often insufficient to realistically represent the system and adequately sample its vast phase space. Particularly, the liquid electrolyte near the interface requires extensive sampling to capture different atomic configurations and their interactions with reaction species [10, 13, 14]. Therefore, it is important to develop methods for large-scale (both spatially and temporarily) simulations without sacrificing accuracy.

Machine learning force field (MLFF) emerges as a promising approach for large-scale simulation [15–20]. By learning from DFT results, MLFF can efficiently predict the forces for new atomic structures with accuracy comparable to DFT, thereby accelerating the simulations. However, the existing MLFFs are not able to describe the behavior of electrons at electrochemical interface. As depicted in Fig. 1a, the interface system is connected to an electrode with certain potential controlled externally (U_{ext}). The electrons in the system can exchange with those in the electrode, adjusting its Fermi level (E_{F}) according to U_{ext}. During elementary reaction steps, U_{ext} often remains constant. Therefore, the electrons in the interface system should be treated by grand canonical ensemble (GCE) [10, 21–29]. An important consequence is that the number of electrons in the interface system (N_{e}) is no longer a constant, instead, it fluctuates at a finite temperature and evolves along reaction path (Fig. 1b). This feature is very different from most thermal reactions/processes, which in contrast have a constant N_{e}. To emphasize that U_{ext} is fixed rather than Ne, this GCE treatment is often also called constant (electrode) potential (CP) approach.

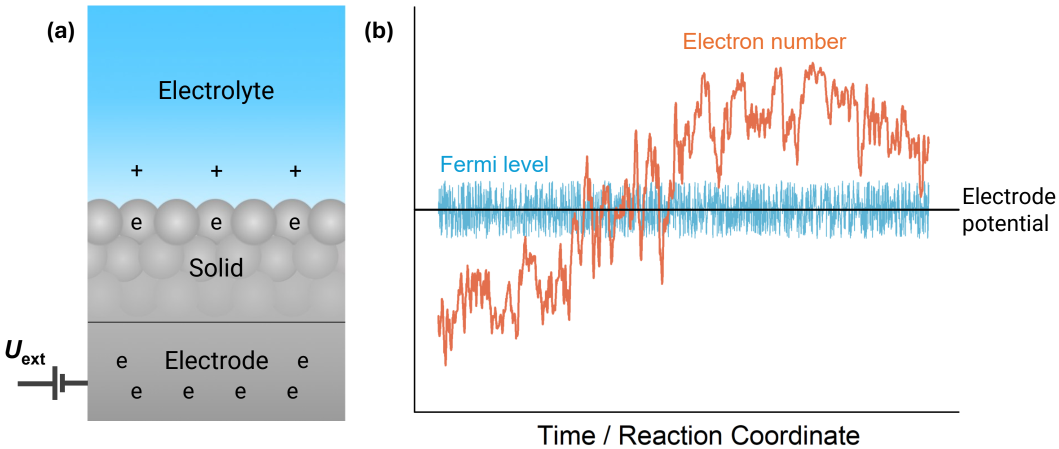

Figure 1: (a) Schematic illustration of an electrochemical interface connected to an electrode. (b) evolutions of electron number and Fermi level under constant electrode potential at finite temperature.

In CP simulations, the N_{e} within the interface system is adjusted according to the E_{F}, following equations such as [23]:

$$\dot{N}_{\mathrm{e}} = \frac{P_n}{M_n}, \tag{1}$$

$$\dot{P}_n = E_{\mathrm{F}} - |e|U_{\mathrm{ext}} - \frac{P_\xi}{M_\xi}P_n, \tag{2}$$

$$\dot{\xi} = \frac{P_\xi}{M_\xi}, \tag{3}$$

$$\dot{P}_\xi = \frac{P_n^2}{M_n} - k_{\mathrm{B}}T\xi, \tag{4}$$

where M_n and P_n are the fictitious mass and momentum for the N_{e} degree of freedom; ξ represents the coupling between the system and the electrode, and P_ξ, M_ξ and T_ξ are its associated momentum, mass, and temperature, respectively. Out of these quantities, (M_n, M_ξ, T_ξ) are independent parameters. These equations of motion are derived from an extended Lagrangian and analogous to Nosé–Hoover thermostat except the degree of freedom is N_{e} instead of atomic positions and the "force" is given by the difference between the instantaneous Fermi level and its target value (plus the contribution from the coupling with the electrode). There exist various schemes for CP simulation, and they all require determining both the forces (F) and E_{F} for a given atomic structure (R) and N_{e}. The F and E_{F} are then used to update both R and N_{e} for subsequent simulation step. Note that even if R remains fixed, the F still varies with N_{e}. Therefore, we need to establish a mapping from (R, N_{e}) to (F, E_{F}), as illustrated in Fig. 2a. Such mapping is straightforward in DFT, and the CP-DFT/CP-AIMD [30–34] has been widely used to improve the understandings and predictions of various electrochemical interface systems. However, this mapping is not available in typical MLFFs, as they focus on the relation between R and F with the assumption that N_{e} is a constant. Therefore, to enable the CP simulations with MLFFs, a new type of MLFF is needed, namely CP-MLFF that can predict (F, E_{F}) from (R, N_{e}). Very recently, several efforts have been made to incorporate the electrode potential into the MLFFs [35–38], however, its incorporation into the state-of-the-art MLFF model — equivariant graph neural network (EGNN) [39–43] — is still missing. Moreover, some prior works do not have N_{e} as input/output [36], or constrain the Fermi level to a fixed value and do not allow it to fluctuate [35, 36] as required by the GCE.

To fill this gap, in this work, we develop a CP-MLFF based on EGNN and implement it into MACE [43], a widely used software for generating MLFFs. Applying this CP-MLFF, we investigate a critical question for electrochemical interface simulations: what simulation duration is necessary to sufficiently sample the phase space and reliably converge the calculation of activation energies? This question was challenging to answer with CP-DFT as the high computational cost limits the simulation to short durations that are not enough to see convergence. However, our CP-MLFF substantially reduces computational cost, permitting much longer simulations and thereby enabling us to answer this question. This is exemplified by CO_2 reduction on a single nickel atom embedded in nitrogen doped graphene (Ni-N-C), which has attracted wide interest due to its promising performance and the importance of CO_2 reduction in advanced energy and manufacturing technologies [44–47]. Our work provides a useful method and tool enabling accurate and efficient large-scale simulation of electrochemical interface.

2 Methods

As mentioned earlier, to enable the CP simulation, the MLFF model needs to take both R and N_{e} as input, and output both F and E_{F}. Here we use the EGNN as implemented in MACE as our base model to demonstrate two approaches, which can be also extended to other models though. EGNN represents the state-of-the-art MLFF model. It represents the atomic system as graph with nodes and edges. Through successive iterations of message passing, the updated node features effectively capture complex many-body atomic interactions, allowing for accurate predictions of the properties.

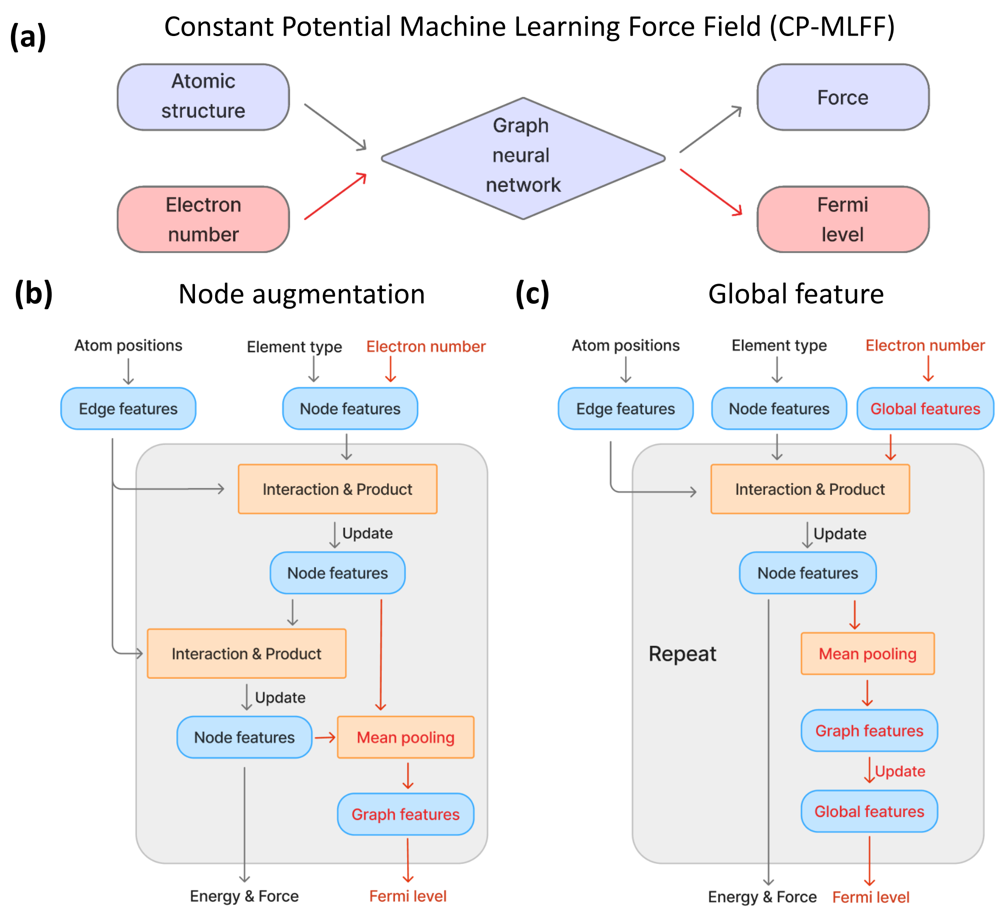

Figure 2: (a) CP-MLFF enables the additional input of electron number and output of Fermi level. (b-c) Two implementations of CP-MLFF model. The red highlights the modifications introduced to MACE.

Node Augmentation: In the first approach (Fig. 2b), we append the N_{e} as an additional feature to each atomic node in the graph. After several layers of message passing, we extract the E_{F} from the updated node features. Specifically, the state of each node i at layer t ($0 \leq t \leq T$) is represented as:

$$\sigma_i^{(t)} = \left(\mathbf{r}_i, \mathbf{h}_i^{(t)}\right), \tag{5}$$

where $\mathbf{r}_i$ is the coordinate of atom i, and $\mathbf{h}_i^{(t)}$ are the learnable features of node i. The initial node feature is constructed by the element type (z_i) and the normalized electron number $\hat{N}_\mathrm{e}$ as:

$$h_{i,k00}^{(0)} = \sum_z W_{kz}^{\mathrm{ele}} \delta_{zz_i} + W_k^{N_\mathrm{e}} \hat{N}_\mathrm{e}, \tag{6}$$

where W_{kz}^{ele} and $W_k^{N_\mathrm{e}}$ are the weights. The node features are then updated by taking information from its neighbors:

$$h_{i,klm}^{(t+1)} = U_t^{kl} \left(\sigma_i^{(t)}, \bigoplus_{j \in \mathcal{N}(i)} M_t \left(\sigma_i^{(t)}, \sigma_j^{(t)} \right) \right), \tag{7}$$

where k, l and m are the indices for the individual feature component, M_t represents a learnable message passing function, U_t^{kl} is a learnable update function, and $\bigoplus_{j \in \mathcal{N}(i)}$ denotes a learnable permutation-invariant pooling operation over all neighbors of atom i. After several message-passing iterations, a mean pooling operation is applied to each layer to generate a graph-level representation that reflects the entire system's property. A final multi-layer perceptron (MLP) uses these graph-level features to predict the Fermi level:

$$E_\mathrm{F} = \mathrm{MLP}\left(\left[h_{i,k00}^{(1)}; \ldots; h_{i,k00}^{(T)} \right] \right). \tag{8}$$

Here, only the invariant features $h_{i,k00}^{(t)}$ is used, which ensures the invariance of the predicted E_F.

Global Feature: Alternatively, the second approach treats the N_e and E_F as a distinct global attribute of the graph. As shown in Figure 2c, the model initializes the global features $\mathbf{g}^{(0)}$ with the electron number:

$$\mathbf{g}^{(0)} = \theta(\hat{N}_\mathrm{e}), \tag{9}$$

where θ is an invariant encode function which maps the $\hat{N}_\mathrm{e}$ to a higher-dimensional latent space. These global features are used to update the node features:

$$\tilde{h}_{i,k00}^{(t)} = \mathrm{MLP}\left(\left[h_{i,k00}^{(t)}; \mathbf{g}^{(t)} \right] \right), \tag{10}$$

where the node features are concatenated with the global features before MLP. It is then used in Eq. (5) and (7) to get $h^{(t+1)}$. The $h^{(t+1)}$ are aggregated to update the global features in turn:

$$\mathbf{g}^{(t+1)} = \mathrm{MLP}\left(\left[\left\{ h_{i,k00}^{(t+1)} \right\}_{\mathrm{pooling}}; \mathbf{g}^{(t)} \right] \right), \tag{11}$$

where average pooling compresses the node features into a graph-level vector. The global features of the last layer serve as input to a readout function that outputs the E_F:

$$E_\mathrm{F} = \mathrm{MLP}\left(\mathbf{g}^{(T)} \right). \tag{12}$$

This global-feature method keeps the system-level information in a separate channel that interacts with the atomic nodes.

To test these methods, we consider a representative electrochemical interface system — $\mathrm{CO_2}$ reduction on single nickel atom embedded in nitrogen doped graphene. As shown in Fig. 3a, the system consists of a single Ni atom coordinated with 1 N atom and 3 C atoms embedded in graphene, 44 water molecules, and $^*\mathrm{COOH}$ on Ni. This site structure has been shown to have the best performance for $\mathrm{CO_2R}$ compared with other possible site structures [44]. We focus on a key step in $\mathrm{CO_2R}$: $^*\mathrm{COOH} + \mathrm{e}^- \rightarrow {}^*\mathrm{CO} + \mathrm{OH}^-_{(\mathrm{aq})}$ and use C–O bond length as the reaction coordinate to describe the reaction.

3 Training set

To generate the initial training set, CP-AIMD simulation are performed with DFT as implemented in CP-VASP [30, 31]. The electrode potential is set to be -0.827 V vs RHE (i.e -3.36 V vs electrolyte), with temperature of 300 K and time step of 1 fs. Further details of the CP-AIMD can be found in the appendices. The system is first roughly equilibrated through 2 ps free MD simulation (without any constraint/bias), followed by 2 ps slow-growth constrained MD (which gradually increases the reaction coordinate: C-O bond length) at a rate of 1 Å/ps to quickly sample the reaction path. Figure 3b depicts the evolutions of N_e (relative to the N_e at charge-neutral state) and E_F during the simulations. As expected, the E_F fluctuates around $-|eU_\mathrm{ext}|$, while N_e gradually increases during the slow growth, indicating the gain of electrons which is driven by the formation of OH- species. Note that due to the high computational expense associated with calculating E_F (which requires implicit solution to set up a potential reference; see Appendix A), we evaluate the E_F and update the N_e only every 5 steps. Therefore, in total, 4000 structures are generated, out of which 800 have E_F calculated. Half of these structures (i.e. 400) are used to construct the initial training dataset.

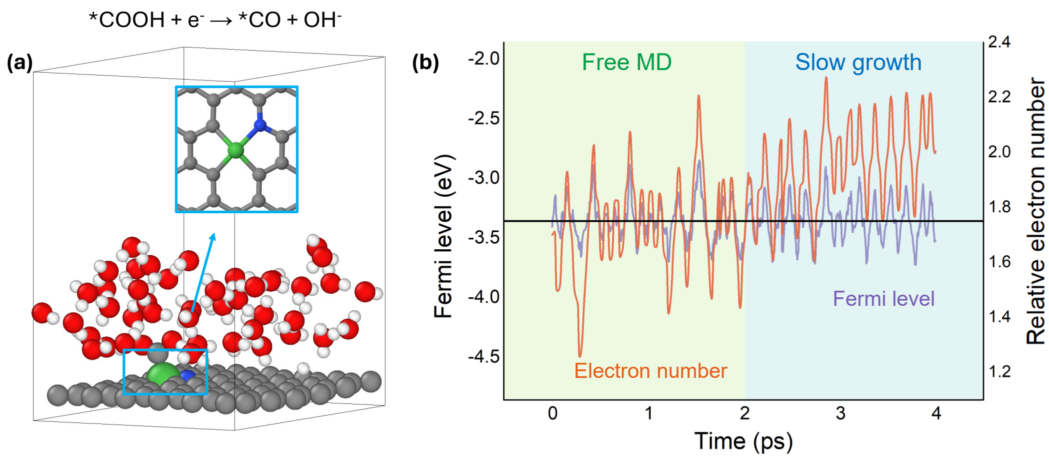

Figure 3: (a) Test system: electrochemical dissociation of *COOH to *CO and OH-, catalyzed by single Ni atom embedded in N-doped graphene in contact of water. (b) Time evolution of electron number (relative to that at charge-neutral state) and Fermi level during 2 ps free MD equilibration followed by 2 ps slow-growth MD simulation at an electrode potential marked by the horizontal line.

Using the initial dataset, 4 CP-FFs were trained independently with different random seeds and their average predictions were compared with the DFT results for the training set. Figure 4 presents the comparison for the CP-FFs generated by approach 1. The root mean square errors (RMSE) are 11.0 meV/Å for atomic forces and 22.1 meV for E_F, indicating that our model can accurately predict both types of quantities simultaneously. Fig. S2 shows the results for approach 2. We find that both approaches yield similar accuracy for E_F prediction, but approach 1 gives better force prediction for the present system. Therefore, in the following, we focus on approach 1, while it is worth noting that the relative performance between the two approaches may vary depending on the system.

To improve the robustness of CP-FFs, we adopted an active learning scheme based on a 4-model ensemble. During CP-FFMD, structures were labeled as “uncertain” if either (i) the standard deviation of force predictions exceeded 150 meV/Å for any atom, or (ii) the standard deviation of the predicted Fermi levels exceeded 40 meV. To avoid redundancy, 10% of the uncertain structures were selected for DFT labeling and added to the training set. This process was repeated until the ratio of uncertain structures fell below 1%. The final dataset comprised 1093 structures, with overall RMSEs of 8.4 meV/Å for forces and 10.8 meV for Fermi levels. More details of the active learning procedure are provided in the appendices.

4 Convergency of activation energy with respect to sampling

With the CP-FFs available, we can afford much longer simulation that DFT cannot afford, enabling us to examine how the calculated properties converge with increased sampling. Here we focus on the activation energy, which is critical to the reaction kinetics. The slow growth can calculate the free energy profile vs reaction coordinate, from which the activation energy can be extracted. In general, a slower rate will give a more accurate free energy profile and activation energy as it samples a larger phase space. Fig. 5 shows the slow growth free energy curves calculated with the same ensemble of

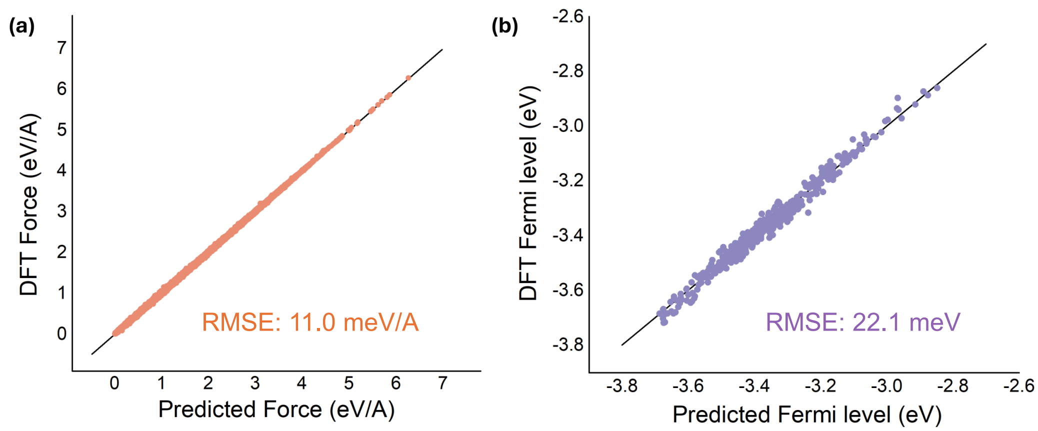

Figure 4: Predicted force (a) and Fermi level (b) compared against DFT reference values for the initial dataset.

CP-FFs (after the active learning) and U_{ext} (-0.827 V vs RHE), and from the same initial state (the last state of free MD in the active learning), but with different growth rates (1, 0.5, 0.1, 0.05, 0.01, and 0.005 ps). Note that the 1 Å/ps rate was used in AIMD to create part of the initial training set. The free energy curves calculated at this rate using AIMD and FFMD agree well (see Fig. S4), further validating the accuracy of the created FFs. As the rate decreases, the free energy curve generally decreases, and it is nearly identical for the 0.01 Å/ps and 0.005 Å/ps cases. The activation energy also generally decreases with lower rate, and converges after 0.01 Å/ps. Remarkably, the activation energy calculated at 1 Å/ps (a typical rate used for AIMD simulation in literature) is significantly higher than the converged value (0.894 eV vs 0.460 eV), underscoring the importance of sufficient sampling. At the converging rate (0.01 Å /ps), the slow growth samples 200,000 structures, which is very expensive for DFT calculations. This highlights the necessity of using our CP-MLFF model. It is worth noting that the proportion of highly uncertain structures remains to be low across different rates (2.35%, 1.90%, 0.92%, 1.38%, 2.17% and 1.14% from 1 Å/ps to 0.005 Å/ps respectively), further validating the effectiveness of active learning.

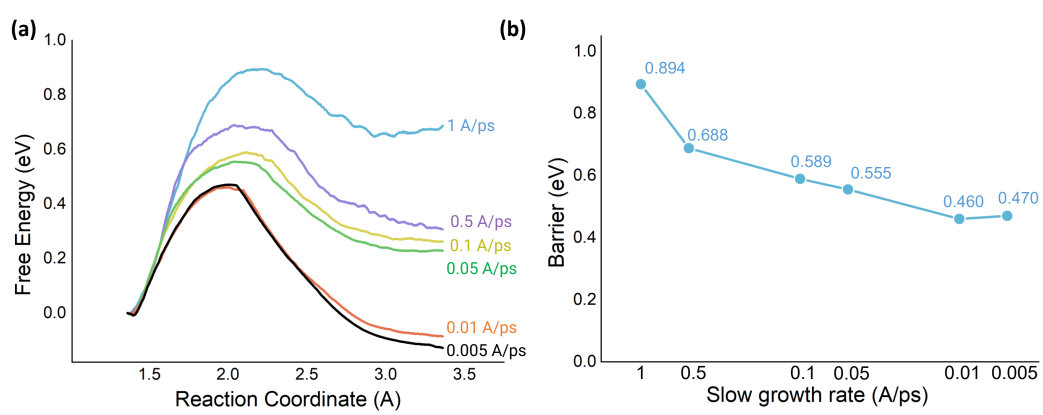

Figure 5: (a) Free energy profiles calculated using the CP-MLFF ensemble at different slow-growth rates. (b) Activation energy extracted from the free energy profile vs the slow-growth rate.

The activation energies may be different if starting from different initial states. To evaluate its effect, we run 9 independent slow-growth simulations with the same rate (0.01 Å/ps), but different initial states. Those states are selected from the last 2 ps of the 10 ps free MD trajectory, with a spacing of 0.25 ps. Fig. 6 shows the free energy profile for each run, and the distributions of the activation energies vs the number of runs. The statistics of the activation energies have converged within 9 runs, showing an average of 0.487 eV and standard deviation of 0.041 eV.

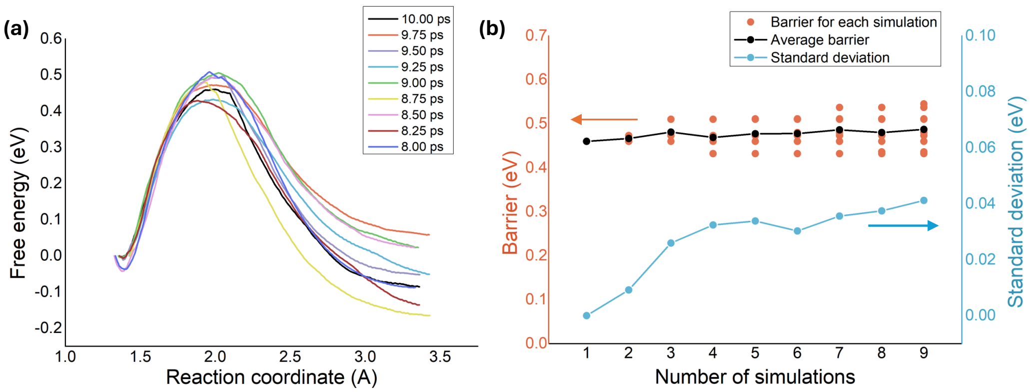

Figure 6: (a) Free energy profiles computed from different initial structures. The inset indicates the time point when the initial structure is extracted from the free MD trajectory. (b) Activation energy distribution along with the average value and standard deviation. The x-axis indicates the number of simulations included: x = 1 corresponds to the slow-growth simulation starting from the structure at 10 ps in the free MD; x = 2 includes both the simulation at x = 1 and an additional one starting from 9.75 ps, and so on.

Now we will discuss several additional capabilities of our CP-MLFF model. (1) Our CP-MLFF is able to extrapolate to new systems with different sizes and electrode potentials. To demonstrate this, the performance of the trained CP-MLFF is tested on a larger system that has a $\sqrt{43} \times \sqrt{43}$ graphene supercell (i.e. $6a + b$ lattice vectors, where a and b are the basis vectors of graphene primitive cell), 72 water molecules, and a different potential of $-|eU_{\text{ext}}| = -3.46$ eV. As detailed in Appendix F, the original CP-MLFF has a reasonable accuracy for this system and can be further improved by active learning with a small number of structures added to the training set. (2) There are different schemes for CP calculations. For example, instead of rigorously following the grand canonical ensemble, one may use simpler dynamic equations for N_{e} [23]. Our CP-MLFF is compatible with these alternative schemes as they need the same input and give the same output as Eq. 1-4. Particularly, some schemes do not allow the E_{F} to fluctuate, instead, it is set to be a constant for each structure during MD/relaxation. This can be simply realized by adjusting the N_{e} and thus EF to match the target value. Alternatively, one can also flip N_{e} and E_{F} in the training (i.e. using E_{F} as input and N_{e} as output) so that the created CP-FF can directly predict the force under the given E_{F}. (3) Although in this work each structure in the training set has its Fermi level available, the dataset can be expanded to include additional structures that contain only force or E_{F} information. Such expanded datasets can improve the accuracy of force/E_{F} prediction. Particularly, as mentioned earlier, when running CP-AIMD for building the initial training set, not every structure has its E_{F} calculated because of the high computational cost, leaving some structures with only forces available. As a test, we include those structures into the training set, and re-train the FFs. This effectively reduces the RMSE in force predictions (from 11 meV/Å to 7 meV/Å), without compromising the accuracy of E_{F} prediction (see Fig. S3). It is expected that the error in E_{F} can also be further reduced with more E_{F}-containing structures in the training set. This flexibility of our model to take various types of data is a useful feature for studying complex systems where different kinds of data may be present.

In addition to CP method, there exist other methods to simulate electrochemical interface, such as finite-field DFT-MD [48]. It would be useful to develop a MLFF for those methods as well.

5 Conclusion

In this work, we have developed a machine learning model that can create constant potential force field, which not only yields the forces for a given atomic structure but also predicts Fermi level as well as how the forces change with electron number. This force field enables large-scale simulation of electrochemical interface under constant electrode potential. As a demonstration, it is applied to study the *COOH dissociation — a critical step of CO_2 reduction — on single nickel atom embedded in nitrogen doped graphene in contact with water. We find that to converge the activation energy, the slow-growth rate needs to be as slow as 0.01 Å/ps, which is challenging for DFT and highlights the need of our method. Our work paves a step towards improved atomistic understanding of electrified interface at scale.

Limitation. Our model remains constrained to the systems included in training, and the accessible simulation timescales are still limited to the nanosecond regime. In addition,the ensemble method may not always be optimal in identifying the most informative configurations, which may limit the efficiency of active learning and the overall improvement of the model.

6 Experimental setting

For each iteration, we train the model with a learning rate of 1e-2 on 1 A100 GPU for 100 epochs. The batch size is 2 per GPU.

7 Acknowledgements

This work is supported by DARPA HR0011-24-2-0380 and NSF 2430894. The calculations are performed on ACCESS clusters

References

[1] Z. W. Seh, J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Norskov, and T. F. Jaramillo. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science*, 355(6321), 2017. doi: 10.1126/science.aad4998.

[2] J. Qiao, Y. Liu, F. Hong, and J. Zhang. A review of catalysts for the electroreduction of carbon dioxide to produce low-carbon fuels. *Chem Soc Rev*, 43(2):631–75, 2014. doi: 10.1039/c3cs60323g.

[3] V. R. Stamenkovic, D. Strmcnik, P. P. Lopes, and N. M. Markovic. Energy and fuels from electrochemical interfaces. *Nat Mater*, 16(1):57–69, 2016. doi: 10.1038/nmat4738.

[4] Isolda Roger, Michael A Shipman, and Mark D Symes. Earth-abundant catalysts for electrochemical and photoelectrochemical water splitting. *Nature Reviews Chemistry*, 1(1):0003, 2017. doi: 10.1038/s41570-016-0003.

[5] NM Marković and PN Ross Jr. Surface science studies of model fuel cell electrocatalysts. *Surface Science Reports*, 45(4-6):117–229, 2002. doi: 10.1016/S0167-5729(01)00022-X.

[6] O. M. Magnussen and A. Gross. Toward an atomic-scale understanding of electrochemical interface structure and dynamics. *J Am Chem Soc*, 141(12):4777–4790, 2019. doi: 10.1021/jacs.8b13188.

[7] Jens Kehlet Nørskov, Thomas Bligaard, Jan Rossmeisl, and Claus Hviid Christensen. Towards the computational design of solid catalysts. *Nature chemistry*, 1(1):37–46, 2009. doi: 10.1038/nchem.121.

[8] Jens Kehlet Nørskov, Jan Rossmeisl, Ashildur Logadottir, LRKJ Lindqvist, John R Kitchin, Thomas Bligaard, and Hannes Jonsson. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *The Journal of Physical Chemistry B*, 108(46):17886–17892, 2004. doi: 10.1021/jp047349j.

[9] Kathleen Schwarz and Ravishankar Sundararaman. The electrochemical interface in first-principles calculations. *Surface science reports*, 75(2):100492, 2020. doi: 10.1016/j.surfrep.2020.100492.

[10] Zachary Levell, Jiabo Le, Saerom Yu, Ruoyu Wang, Sudheesh Ethirajan, Rachita Rana, Ambarish Kulkarni, Joaquin Resasco, Deyu Lu, and Jun Cheng. Emerging atomistic modeling methods for heterogeneous electrocatalysis. *Chemical Reviews*, 124(14):8620–8656, 2024. doi: 10.1021/acs.chemrev.3c00735.

[11] X. Zhao, Z. H. Levell, S. Yu, and Y. Liu. Atomistic understanding of two-dimensional electrocatalysts from first principles. *Chem Rev*, 122(12):10675–10709, 2022. doi: 10.1021/acs.chemrev.1c00981.

[12] Jan Rossmeisl, Egill Skúlason, Mårten E. Björketun, Vladimir Tripkovic, and Jens K. Nørskov. Modeling the electrified solid–liquid interface. *Chemical Physics Letters*, 466(1):68–71, 2008. doi: 10.1016/j.cplett.2008.10.024.

[13] A. Gross and S. Sakong. Ab initio simulations of water/metal interfaces. *Chem Rev*, 122(12):10746–10776, 2022. doi: 10.1021/acs.chemrev.1c00679.

[14] L. Scalfi, M. Salanne, and B. Rotenberg. Molecular simulation of electrode-solution interfaces. *Annu Rev Phys Chem*, 72:189–212, 2021. doi: 10.1146/annurev-physchem-090519-024042.

[15] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schutt, A. Tkatchenko, and K. R. Muller. Machine learning force fields. *Chem Rev*, 121(16):10142–10186, 2021. doi: 10.1021/acs.chemrev.0c01111.

[16] I. Poltavsky and A. Tkatchenko. Machine learning force fields: Recent advances and remaining challenges. *J Phys Chem Lett*, 12(28):6551–6564, 2021. doi: 10.1021/acs.jpclett.1c01204.

[17] V. Botu, R. Batra, J. Chapman, and R. Ramprasad. Machine learning force fields: Construction, validation, and outlook. *The Journal of Physical Chemistry C*, 121(1):511–522, 2017. doi: 10.1021/acs.jpcc.6b10908.

[18] K. T. Schutt, H. E. Sauceda, P. J. Kindermans, A. Tkatchenko, and K. R. Muller. Schnet - a deep learning architecture for molecules and materials. *J Chem Phys*, 148(24):241722, 2018. doi: 10.1063/1.5019779.

[19] Han Wang, Linfeng Zhang, Jiequn Han, and Weinan E. Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228:178–184, 2018. doi: 10.1016/j.cpc.2018.03.016.

[20] Chi Chen, Weike Ye, Yunxing Zuo, Chen Zheng, and Shyue Ping Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9):3564–3572, 2019. doi: 10.1021/acs.chemmater.9b01294.

[21] N. G. Hormann, O. Andreussi, and N. Marzari. Grand canonical simulations of electrochemical interfaces in implicit solvation models. *J Chem Phys*, 150(4):041730, 2019. doi: 10.1063/1.5054580.

[22] Christopher D. Taylor, Sally A. Wasileski, Jean-Sebastien Filhol, and Matthew Neurock. First principles reaction modeling of the electrochemical interface: Consideration and calculation of a tunable surface potential from atomic and electronic structure. *Physical Review B*, 73(16):165402, 2006. doi: 10.1103/PhysRevB.73.165402.

[23] N. Bonnet, T. Morishita, O. Sugino, and M. Otani. First-principles molecular dynamics at a constant electrode potential. *Phys Rev Lett*, 109(26):266101, 2012. doi: 10.1103/PhysRevLett.109.266101.

[24] F. Deißenbeck, C. Freysoldt, M. Todorova, J. Neugebauer, and S. Wippermann. Dielectric properties of nanoconfined water: A canonical thermopotentiostat approach. *Physical Review Letters*, 126(13):136803, 2021. doi: 10.1103/PhysRevLett.126.136803.

[25] Sudarsan Surendralal, Mira Todorova, Michael W. Finnis, and Jörg Neugebauer. First-principles approach to model electrochemical reactions: Understanding the fundamental mechanisms behind mg corrosion. *Physical Review Letters*, 120(24):246801, 2018. doi: 10.1103/PhysRevLett.120.246801.

[26] Xi Chen, Georg Kastlunger, and Andrew A. Peterson. Fundamental drivers of electrochemical barriers. *Physical Review Letters*, 131(23):238003, 2023. doi: 10.1103/PhysRevLett.131.238003.

[27] Ravishankar Sundararaman, III Goddard, William A., and Tomas A. Arias. Grand canonical electronic density-functional theory: Algorithms and applications to electrochemistry. *The Journal of Chemical Physics*, 146(11):114104, 2017. doi: 10.1063/1.4978411.

[28] Zhaoming Xia and Hai Xiao. Grand canonical ensemble modeling of electrochemical interfaces made simple. *Journal of Chemical Theory and Computation*, 19(15):5168–5175, 2023. doi: 10.1021/acs.jctc.3c00237.

[29] Guoping Gao and Lin-Wang Wang. Substantial potential effects on single-atom catalysts for the oxygen evolution reaction simulated via a fixed-potential method. *Journal of Catalysis*, 391:530–538, 2020. doi: https://doi.org/10.1016/j.jcat.2020.08.032.

[30] Xunhua Zhao and Yuanyue Liu. Origin of selective production of hydrogen peroxide by electrochemical oxygen reduction. *Journal of the American Chemical Society*, 143(25):9423–9428, 2021. doi: 10.1021/jacs.1c02186.

[31] Saerom Yu, Zachary Levell, Zhou Jiang, Xunhua Zhao, and Yuanyue Liu. What is the rate-limiting step of oxygen reduction reaction on fe–n–c catalysts? *Journal of the American Chemical Society*, 145(46):25352–25356, 2023. doi: 10.1021/jacs.3c09193.

[32] Per Lindgren, Georg Kastlunger, and Andrew A. Peterson. Scaled and dynamic optimizations of nudged elastic bands. *Journal of Chemical Theory and Computation*, 15(11):5787–5793, 2019. doi: 10.1021/acs.jctc.9b00633.

[33] Assil Bouzid and Alfredo Pasquarello. Atomic-scale simulation of electrochemical processes at electrode/water interfaces under referenced bias potential. *The Journal of Physical Chemistry Letters*, 9(8): 1880–1884, 2018. doi: 10.1021/acs.jpclett.8b00573.

[34] Assil Bouzid, Patrick Gono, and Alfredo Pasquarello. Reaction pathway of oxygen evolution on pt(1 1 1) revealed through constant fermi level molecular dynamics. *Journal of Catalysis*, 375:135–139, 2019. doi: https://doi.org/10.1016/j.jcat.2019.05.025.

[35] Xi Chen, Muammar El Khatib, Per Lindgren, Adam Willard, Andrew J. Medford, and Andrew A. Peterson. Atomistic learning in the electronically grand-canonical ensemble. *npj Computational Materials*, 9(1):73, 2023. doi: 10.1038/s41524-023-01007-6.

[36] Jingwen Zhou, Yunsong Fu, Ling Liu, and Chungen Liu. Constant-potential machine learning molecular dynamics simulations reveal potential-regulated cu cluster formation on mos2. *The Journal of Physical Chemistry C*, 129(13):6414–6422, 2025. doi: 10.1021/acs.jpcc.4c08188.

[37] Menglin Sun, Bin Jin, Xiaolong Yang, and Shenzhen Xu. Probing nuclear quantum effects in electrocatalysis via a machine-learning enhanced grand canonical constant potential approach. *Nature Communications*, 16(1):3600, 2025. doi: 10.1038/s41467-025-58871-7.

[38] Letian Chen, Yun Tian, Xu Hu, Suya Chen, Huijuan Wang, Xu Zhang, and Zhen Zhou. A constant potential reactor framework for electrochemical reaction simulations. *arXiv preprint arXiv:2411.16330*, 2024.

[39] Ilyes Batatia, Simon Batzner, Dávid Péter Kovács, Albert Musaelian, Gregor N. C. Simm, Ralf Drautz, Christoph Ortner, Boris Kozinsky, and Gábor Csányi. The design space of e(3)-equivariant atom-centred interatomic potentials. *Nature Machine Intelligence*, 7(1):56–67, 2025. doi: 10.1038/s42256-024-00956-x.

[40] Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks, 2021.

[41] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, 2022. doi: 10.1038/s41467-022-29939-5.

[42] Kristof Schütt, Oliver Unke, and Michael Gastegger. Equivariant message passing for the prediction of tensorial properties and molecular spectra. In *International Conference on Machine Learning*, pages 9377–9388. PMLR, 2021.

[43] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.

[44] Xunhua Zhao and Yuanyue Liu. Unveiling the active structure of single nickel atom catalysis: Critical roles of charge capacity and hydrogen bonding. *Journal of the American Chemical Society*, 142(12):5773–5777, 2020. doi: 10.1021/jacs.9b13872.

[45] Wentuan Bi, Xiaogang Li, Rui You, Minglong Chen, Ruilin Yuan, Weixin Huang, Xiaojun Wu, Wangsheng Chu, Changzheng Wu, and Yi Xie. Surface immobilization of transition metal ions on nitrogen-doped graphene realizing high-efficient and selective co2 reduction. *Advanced Materials*, 30(18):1706617, 2018. doi: 10.1002/adma.201706617.

[46] Xiaogang Li, Wentuan Bi, Minglong Chen, Yuexiang Sun, Huanxin Ju, Wensheng Yan, Junfa Zhu, Xiaojun Wu, Wangsheng Chu, Changzheng Wu, and Yi Xie. Exclusive ni–n4 sites realize near-unity co selectivity for electrochemical co2 reduction. *Journal of the American Chemical Society*, 139(42):14889–14892, 2017. doi: 10.1021/jacs.7b09074.

[47] Song Liu, Hong Bin Yang, Sung-Fu Hung, Jie Ding, Weizheng Cai, Linghui Liu, Jiajian Gao, Xuning Li, Xinyi Ren, Zhichong Kuang, Yanqiang Huang, Tao Zhang, and Bin Liu. Elucidating the electrocatalytic co2 reduction reaction over a model single-atom nickel catalyst. *Angewandte Chemie International Edition*, 59(2):798–803, 2020. doi: 10.1002/anie.201911995.

[48] Linnéa Andersson, Michiel Sprik, Jürg Hutter, and Chao Zhang. Electronic response and charge inversion at polarized gold electrode. *Angewandte Chemie International Edition*, 64(1):e202413614, 2025. doi: https://doi.org/10.1002/anie.202413614.

[49] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, 1996. doi: 10.1103/PhysRevB.54.11169.

[50] P. E. Blöchl. Projector augmented-wave method. *Physical Review B*, 50(24):17953–17979, 1994. doi: 10.1103/PhysRevB.50.17953.

[51] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865–3868, 1996. doi: 10.1103/PhysRevLett.77.3865.

[52] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132(15):154104, 2010. doi: 10.1063/1.3382344.

[53] S. M. Rezwanul Islam, Foroogh Khezeli, Stefan Ringe, and Craig Plaisance. An implicit electrolyte model for plane wave density functional theory exhibiting nonlinear response and a nonlocal cavity definition. *The Journal of Chemical Physics*, 159(23):234117, 2023. doi: 10.1063/5.0176308.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Our abstract and introduction clearly explain how we develop a constant potential machine learning force field method based on the equivariant graph neural network.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss this issue in Section 5.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

Justification: Our method does not make strong theoretical contributions.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Detailed implementation information is provided in Section 4.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [NA]

Justification: At this stage, we do not release the code but provide the implementation details. We will open-source related code after publication.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Detailed experimental implementation information is provided in Sections 3 and 4.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Yes, please refer to the figures.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).

- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The details are thoroughly discussed in Section 6.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We have complied with the Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA]

    Justification: No direct societal impact is performed.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [Yes]

    Justification: We do not identify such a risk.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: Yes, we have added explanations in the relevant sections.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: We do not release any new assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: Our method does not rely on large language models (LLMs) at its core.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.

Appendix

A Details of DFT, CP-AIMD and CP-FFMD calculations

In this work, density functional theory (DFT) calculations are performed using the Vienna Ab Initio Simulation Package (VASP) [49]. The projector augmented wave (PAW) method is employed to describe the interaction between ions and electrons [50]. A plane-wave energy cutoff of 400 eV is used during MD simulations. Exchange–correlation interactions are treated with the Perdew–Burke–Ernzerhof (PBE) functional [51], combined with the D3 dispersion correction to account for van der Waals interactions [52]. The simulation supercell consists of a 6×6 graphene sheet, where six carbon atoms are replaced by an Ni–N complex, and 44 explicit water (H_2O) molecules are added. All calculations are carried out using a Γ-centered $3 \times 3 \times 1$ k-point mesh.

For the CP-AIMD simulations, we employ the CP-VASP with VASPsol++ implicit solvation environment [53]. The electrolyte concentration is set to 1 mol/L, and the effective ionic radius is specified as 4 Å. Grand canonical sampling is performed using `NESCHEME=5`, with $M_n = 660.74$ eV·fs^2, $T_\xi = 300$ K, and $M_\xi = k_\mathrm{B} T_\xi \times 81.27$, which is selected based on the vibrational frequency of the O–H bond.

For CP-FFMD simulations, the settings are mostly kept consistent with those used in CP-AIMD, except for the T_ξ, which is reduced to 60 K. This adjustment is made to reduce the amplitude of electron number fluctuations, and consequently to reduce Fermi level fluctuations—from approximately 1 eV in CP-AIMD to around 0.5 eV in CP-FFMD.

B Training details for the CP-MLFF ensemble

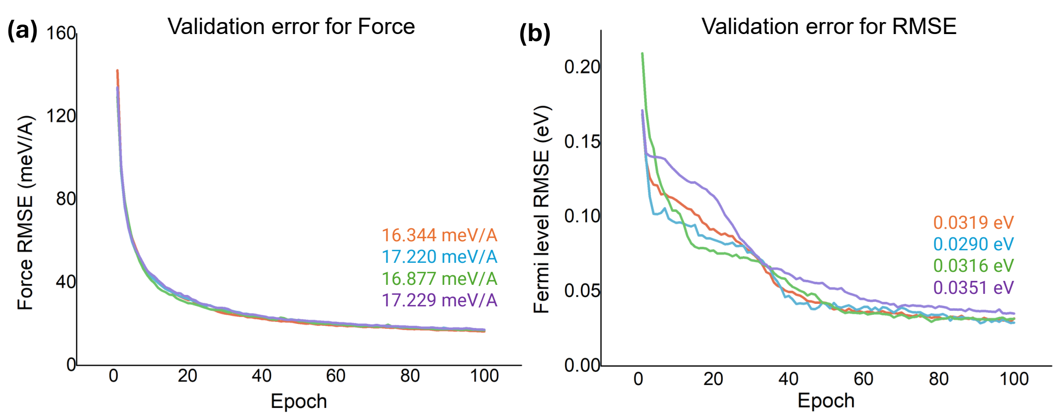

Figure 7: RMSE of predicted (a) forces and (b) Fermi level as a function of training epochs for four CP-MLFF models independently trained over the initial training set.

For the hyperparameters, the CP-MLFF models were trained using a validation fraction of 5% to monitor performance and prevent overfitting. The hidden layer of the network was constructed using irreducible representations consisting of 128 scalar features (0e) and 128 vector features (1o) (hidden_irreps='128x0e+128x1o'). A cutoff radius of 5.0 Å was employed to define the atomic neighborhood for message passing. During training, the batch size was set to 2. The loss function combined contributions from atomic forces, energies, and Fermi levels, with weighting factors set to 100.0 for atomic forces, 1.0 for energy, and 10.0 for the Fermi level predictions. These settings were chosen to balance the accuracy across the predicted physical properties.

The validation RMSE values for initial training as shown in Fig. 7 demonstrate consistent convergence and high accuracy for both atomic forces ($\sim$17 meV/Å) and Fermi levels ($\sim$0.03 eV) after 100 training epochs, highlighting the robustness and reliability of the trained ensemble.

B.1 Active learning

When MLFFs are used to run MD, new structures that significantly differ from those in the training set may be encountered, potentially leading to inaccurate predictions. To address this issue, we adopt an active learning strategy is adopted, utilizing an ensemble of FFs, as shown in Fig. 8a. As mentioned earlier, we independently train 4 CP-FFs and take the average of their predictions as the final prediction. These CP-FFs are then used to perform a 10 ps free MD simulation under the same U_{ext} as in the initial training dataset. For each structure in the CP-FFMD, we measure the prediction uncertainty by the standard deviation among the 4 CP-FFs, and mark a structure as "uncertain" if it satisfies either of the following criteria:

1. For any atom in the structure, the standard deviation of the force predictions across the ensemble exceeds 150 meV/Å.

2.The standard deviation of the predicted Fermi levels across the ensemble exceeds 40 meV.

Not all the uncertain structures are calculated by DFT, because many of them are similar to each other. To reduce the redundancy, we uniformly select 10% of the uncertain structures, which are ordered by their time of appearance in the MD trajectory. However, these structures can still be too many for DFT calculations, especially in the early iterations. Therefore, we set an upper limit and select at most 100 structures in each iteration. Those structures along with their N_{e} (as determined during the CP-FFMD) are calculated by DFT to obtain the accurate forces and E_{F}. These new data are incorporated into the training dataset to update the CP-FFs. Then the updated CP-FFs are used for the next iteration. We repeat this process and record the ratio of uncertain structures. As shown in Fig. 8b, it steadily decreases and reaches below 1% after 4 iterations.

Following the free MD simulations, an additional active learning cycle is carried out using a 20 ps slow-growth simulation at a rate of 0.1 Å/ps and under the same U_{ext}. The iterative procedure continues until the ratio of uncertain structures again dropped below the 1% threshold. This second active learning procedure requires 6 iterations to reach convergence.

To assess the effectiveness of active learning, CP-FF predictions are compared against DFT results for the new structures added to the training set at each iteration. Note that the predictions are made before updating the CP-FFs. As shown in Fig. 8c, the RMSEs generally decrease with the iteration, reaching 17 meV/Å for force and 56 meV for E_{F} for the new structures at the last iteration of slow growth. The final training dataset comprised 1093 structures, and the overall RMSEs across the entire dataset are only 8.4 meV/ Å for atomic forces and 10.8 meV for E_{F}.

C Comparison between two approaches

Fig. 9 compares the RMSEs of (a) Fermi level and (b) atomic forces between the node augmentation method and the global feature method for the final training set. The averaged final RMSE values are 16.51 meV/Å (node augmentation) and 19.55 meV/Å (global feature) for atomic forces, and 40.05 meV (node augmentation) and 39.68 meV (global state feature) for E_{F}. These results indicate that both approaches yield comparable prediction accuracy for the E_{F}, while the node augmentation method provides better performance in predicting atomic forces for the studied system.

D Impact of including additional force-only data

The averaged final validation RMSE values for atomic forces are significantly reduced from 16.92 meV/Å (Fermi-level-containing structures only) to 8.67 meV/Å (all data), while RMSEs for E_{F} remain similar (31.90 vs. 33.28 meV). These results highlight the advantage of incorporating additional force-only data to substantially improve force predictions without compromising Fermi-level accuracy, demonstrating the flexibility and effectiveness of the CP-MLFF model for heterogeneous datasets.

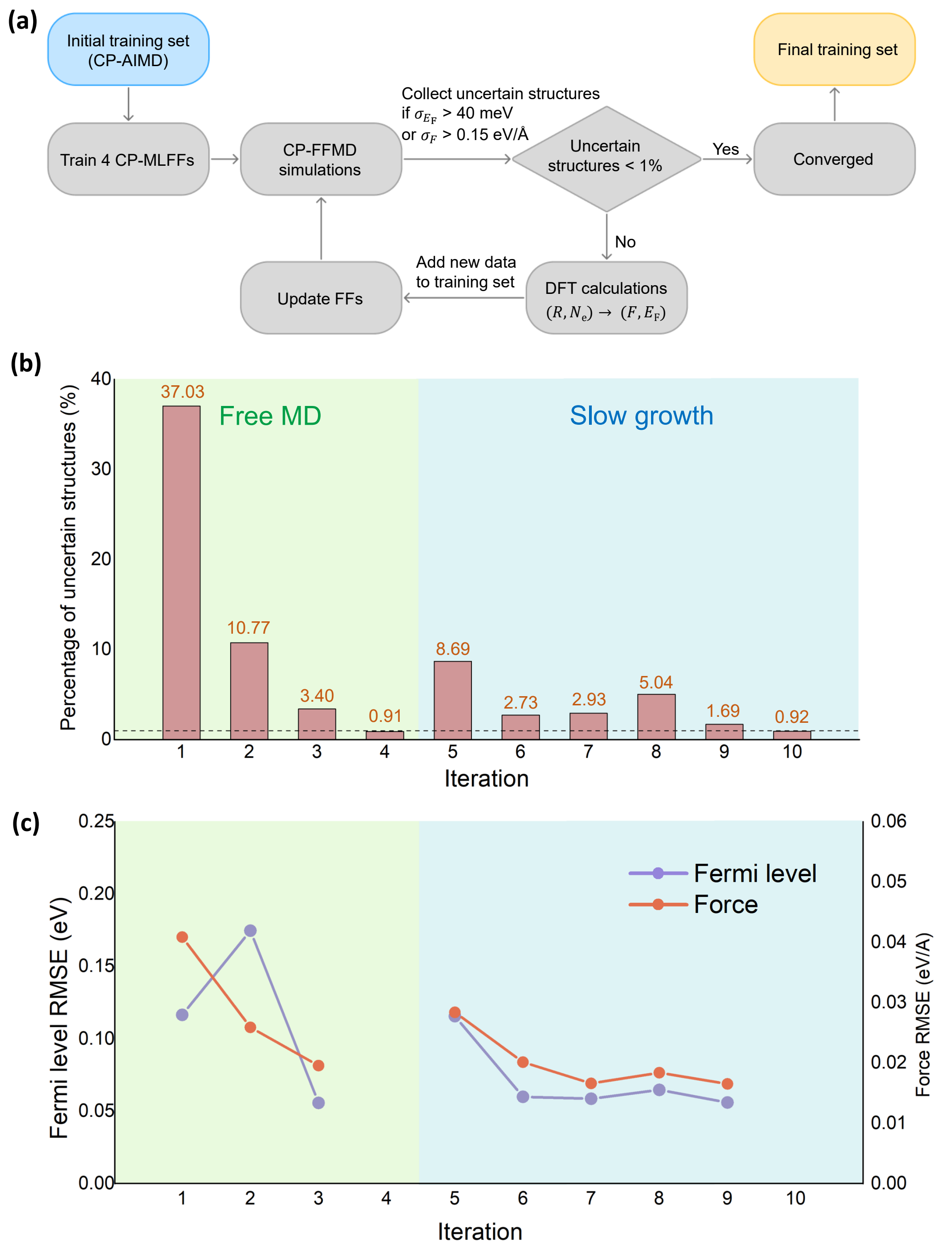

Figure 8: (a) Schematic of the active learning workflow used to iteratively improve the CP-MLFF. (a) Percentage of newly identified uncertain structures during each iteration of active learning. (b) RMSE of atomic forces and Fermi level for the downselected uncertain structures.

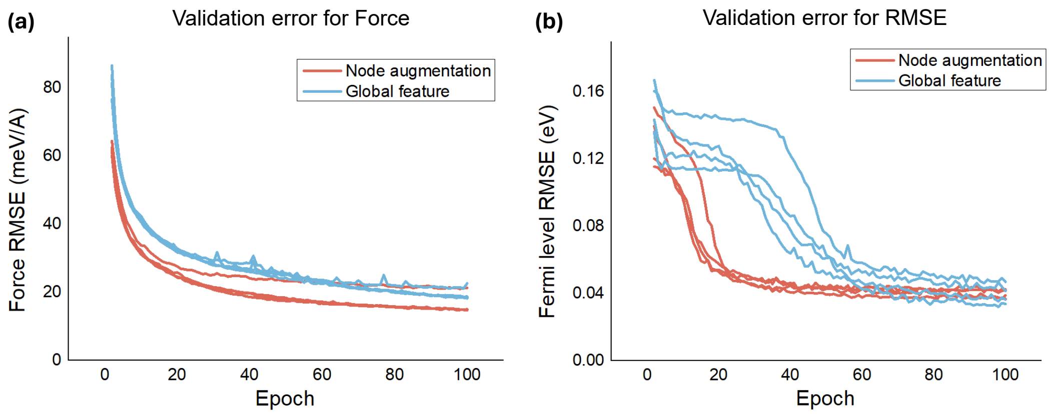

Figure 9: Validation errors of the two CP-MLFF approaches on the final training dataset.

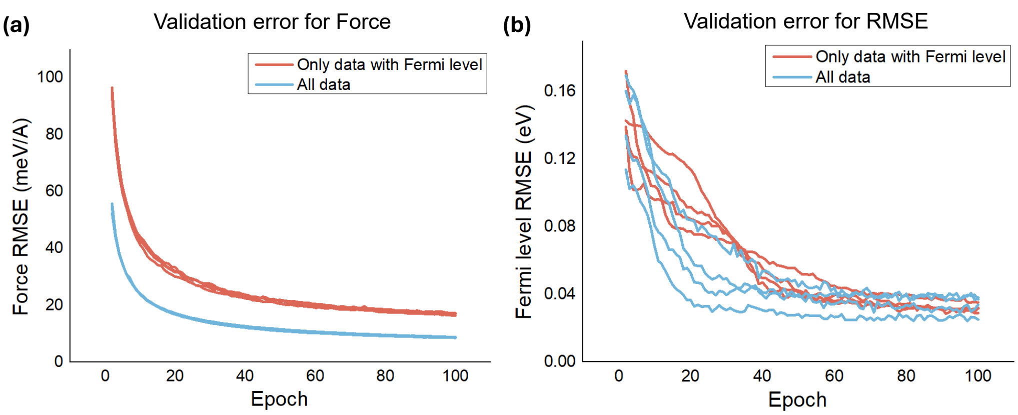

Figure 10: Comparison of validation RMSEs for two different training sets. “Only data with Fermi level” is the final training set used in the main text, where every structure has its Fermi level available. “All data” includes all the structures during CP-AIMD for building the initial training set, but some structures do not have Fermi level available.

E Comparison between CP-AIMD and CP-FFMD for slow-growth simulation

Despite the difference in the initial state and T_ξ, the two curves exhibit good agreement especially for the barrier. The activation barriers are 0.894 eV (CP-AIMD) and 0.884 eV (CP-FFMD), confirming the reliability of our model in reproducing DFT energetics under fast rates.

F Extrapolation to large system

We first use the CP-MLFFs trained from the original (small) system to run free-MD for the large system for 10 ps and uniformly pick 100 structures from the trajectory to perform DFT calculations. By comparing the DFT results with ML predictions, we obtain RMSEs of 19.3 meV/Å for the atomic forces and 125.9 meV for the Fermi level.

To further improve the accuracy, we performed active learning for the large system in the same way as we did for the small system. Specifically, we first conducted a 10 ps free MD simulation, followed by a 20 ps slow-growth simulation. As shown in Figure 12b, throughout the active learning process, the percentage of newly added uncertainty structures in each iteration remains very low. This indicates that only a small number of DFT calculations are needed, allowing the CP-MLFF to reach convergence efficiently. At the final iteration, the force and Fermi level RMSEs for the new structures

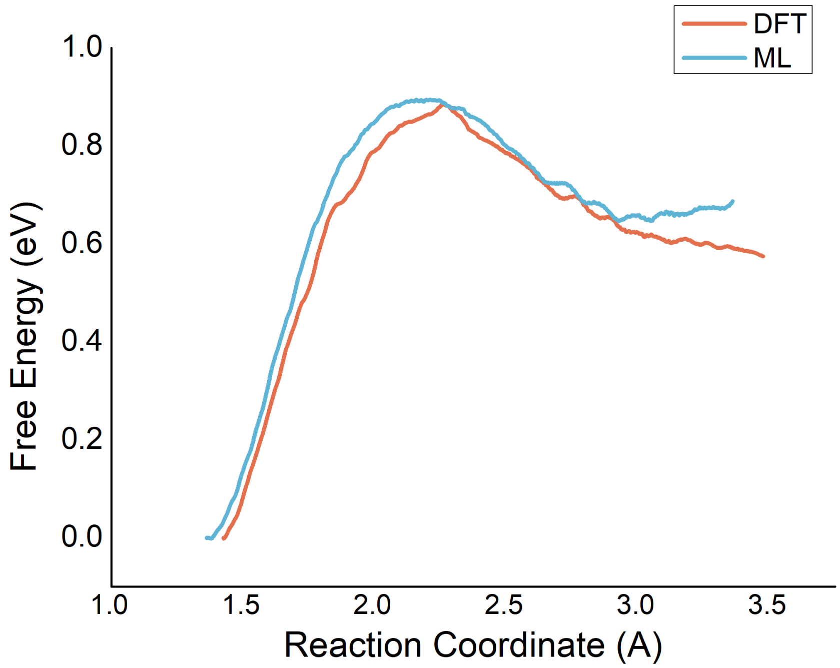

Figure 11: Comparison of free energy profiles from CP-AIMD and CP-FFMD simulations at 1 Å/ps slow growth rate.

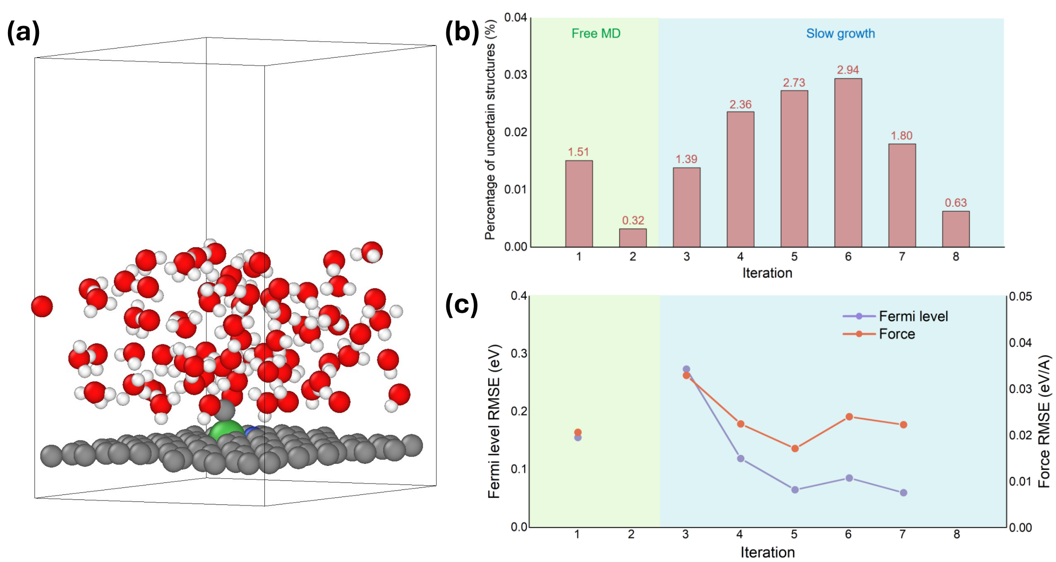

Figure 12: (a) Structure for the large system. (b) Percentage of newly identified uncertain structures during each iteration of active learning. (c) RMSE of atomic forces and Fermi level for the downselected uncertain structures.

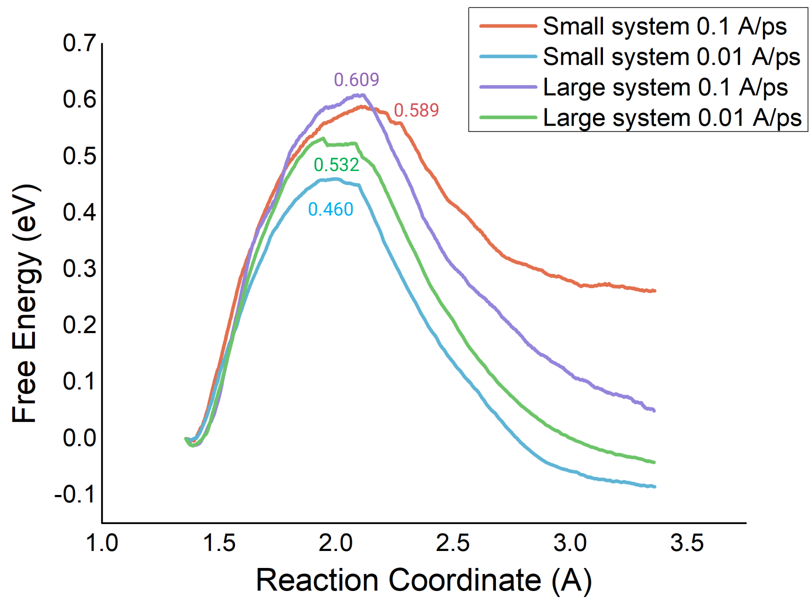

Figure 13: Free energy profiles comparison between large system and small system.

are 22.2 meV/Å and 59.8 meV, respectively, which are comparable to the accuracy achieved on the small system. Figure 13 compares the free energy profile obtained from slow-growth simulation for the large system and the original (small) system.

G Electron number and Fermi level during CP-MLFF slow-growth simulation

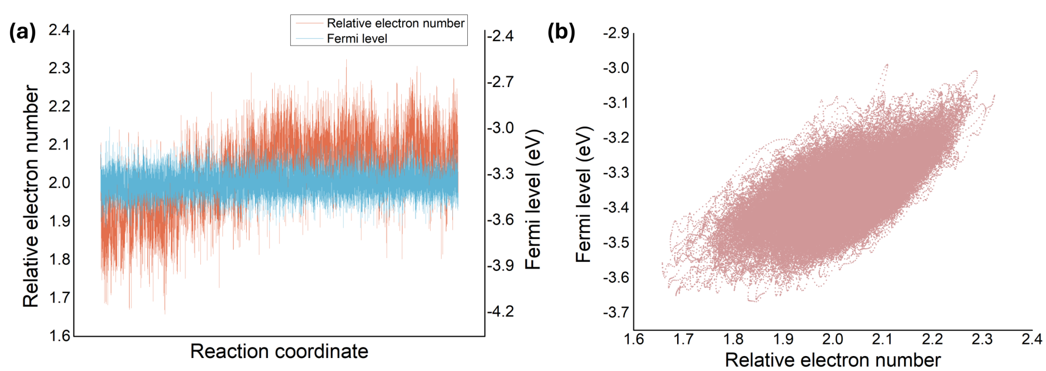

Figure 14: (a) Instantaneous relative electron number and Fermi level along the reaction coordinate for the slow-growth simulation with 0.01 Å/ps. (b) Fermi level versus relative electron number.

As shown in Figure 14a, both quantities fluctuate over time, with the electron number gradually increasing as OH^- species form. Figure 14b further demonstrates a clear positive correlation between the electron number and Fermi level, reflecting the expected physical trend that gaining electrons raises the system's Fermi level. Note that the structural change during the MD simulation also influences the Fermi level.