

ARTICLE OPEN



Atomistic learning in the electronically grand-canonical ensemble

Xi Chen¹, Muammar El Khatib¹, Per Lindgren¹, Adam Willard², Andrew J. Medford³ and Andrew A. Peterson¹✉

A strategy is presented for the machine-learning emulation of electronic structure calculations carried out in the electronically grand-canonical ensemble. The approach relies upon a dual-learning scheme, where both the system charge and the system energy are predicted for each image. The scheme is shown to be capable of emulating basic electrochemical reactions at a range of potentials, and coupling it with a bootstrap-ensemble approach gives reasonable estimates of the prediction uncertainty. The method is also demonstrated to accelerate saddle-point searches, and to extrapolate to systems with one to five water layers. We anticipate that this method will allow for larger length- and time-scale simulations necessary for electrochemical simulations.

npj Computational Materials (2023)9:73 | <https://doi.org/10.1038/s41524-023-01007-6>

INTRODUCTION

In recent years, two new techniques have emerged that have changed the way that atomistic simulations at electrochemical interfaces are approached. First, atomistic machine-learning methods have been developed^{1–6} that use machine-learning surrogate models to quickly emulate and predict the results of expensive electronic structure calculations. These methods offer numerous advantages, including fast evaluation times and linear scaling with system size, at least when the number of atomic elements does not change. This approach can allow systematic exploration at much larger length- and time-scales, while maintaining accuracies similar to those of electronic structure calculations. Crucial to their successful use is the intelligent choice of training data and the ability to estimate the uncertainty associated with each predicted quantity.

Second, grand-canonical electronic structure methodologies have been developed that allow electrochemical simulations to be run at constant simulated voltage^{7–11}. This is accomplished by adding or removing fractions of an electron to the simulation, while achieving charge neutrality by various approaches which typically involve a screened countercharge. For example, in the solvated jellium method (SJM)⁷, the countercharge is placed in a jellium slab above both the slab and explicit solvent, an implicit solvent is used to screen the otherwise large field, and a dipole correction ensures the excess charge localizes on the correct side of the slab, as shown schematically in Fig. 1. Over the course of an elementary reaction, the number of electrons in a simulation changes in order to hold the work function of the simulated electrode constant. Thus, these calculations are electronically grand-canonical—they control the potential by allowing the number of electrons to be a free variable. These simulation approaches have allowed, for example, the calculation of reaction barriers at a systematic series of electrode potentials, leading to an understanding of the potential dependence of reaction barriers and free energy diagrams.

The system sizes that are studied with electrochemical simulations are typically too small to fully capture the role of the environment in a particular reaction. Solvent molecules and

ions at the interface and within the double-layer region can make significant contributions to the rates and mechanisms of electrochemical processes. In order to properly identify and characterize these contributions, the simulated system size must exceed the correlation length and time scales that emerge at the interface. These scales are vastly larger for solution-phase reactions than they are in the gas phase, often spanning nanometer length scales and nanosecond time scales¹². Electronic structure-based simulations, which are well-suited for addressing gas-phase systems, thus tend to fall short in describing reactions at solid-liquid interfaces.

While atomistic machine-learning approaches can accelerate traditional, canonical electronic structure calculations, they do not natively operate in the electronically grand-canonical ensemble. That is, in typical atomistic learning approaches, the energy is considered to be a unique function of the nuclear positions. In the electronically grand-canonical ensemble, the energy is a function of both the atomic positions *and* the electrode potential. It is the purpose of this work to develop a practical approach to atomistic learning in the electronically grand-canonical ensemble.

ALGORITHM

Background

In 2007, Behler and Parrinello¹ proposed a machine-learning scheme to model the potential energy surface of atomistic simulations, in an ansatz where the total system energy can be decomposed into per-atom terms. In this scheme, the atomic coordinates are first transformed into feature vectors, known as atomic fingerprints or symmetry functions. The symmetry functions distinguish the local environment around a central atom due to its neighboring atoms. A neural network exists for each atom type, which takes the descriptor as the input and gives the corresponding atomic energy as the output. The total energy is calculated as the sum of all atomic energies. The neural network for each atomic element is identical, and the parameters of each are tuned simultaneously by minimizing the difference between the predicted and true energies of a training set of images. Once

¹School of Engineering, Brown University, Providence, RI 02912, USA. ²Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. ³School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30318, USA. ✉email: andrew.peterson@brown.edu

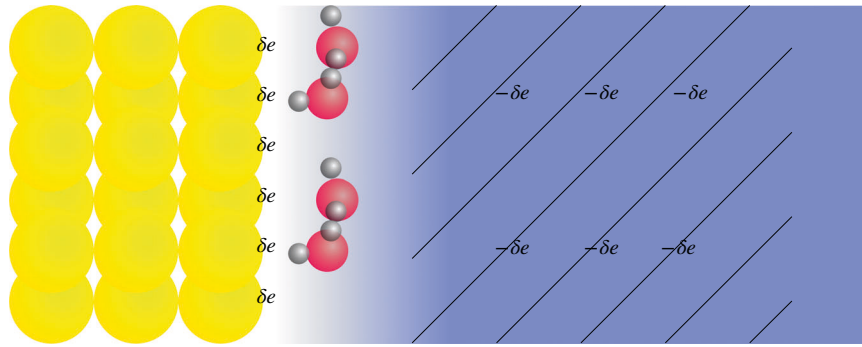


Fig. 1 Schematic of the Solvated Jellium Method. The unit cell is charge-neutral with excess charge localizing on the metal surface, balanced by charge of opposite sign in the jellium. The total number of electrons is adjusted until the work function meets the target value.

trained to appropriate data, the machine-learning method can decrease the computational resources by orders of magnitude, and scales linearly with the number of atoms being modeled (whereas density functional theory, DFT, typically scales cubically with the number of atoms, although efforts are underway to reduce this¹³). Recent work has shown that atomic energies localized by the Behler–Parrinello scheme are in good agreement with those decomposed from DFT, where the energy density is first decomposed into the kinetic contribution, the classic Coulomb energy, the exchange–correlation interactions, and the nonlocal pseudopotential contribution and then integrated over the whole space to satisfy the DFT total energy^{14,15}. In the Behler–Parrinello scheme and similar methods inspired from it, the energy is a unique function of atomic positions. Therefore, these schemes intrinsically cannot work for electronically grand-canonical calculations, where the energy is also a function of the electric potential of the system.

Goedecker and colleagues introduced a machine-learning scheme to allow long-range charge transfer, in which a neural-network model predicts the per-atom electronegativity, rather than the per-atom energy³. In their scheme, the total charge of the system is constrained through a Lagrange factor and the charge distribution (comprised of per-atom charges) is found by solving a linear system to minimize the system energy, which is found through a charge–electronegativity expansion. This system energy is used as a training target to adjust the neural-network parameters that predict atomic electronegativities.

Inspired by these previous works, we introduce the following dual-learning scheme in the electronically grand canonical ensemble.

Dual-learning scheme

We consider an image—that is, a single configuration of atoms at a specified potential—with N atoms, which is calculated in the electronically grand-canonical ensemble. Its potential is set to a value ϕ by varying the net number of electrons in the simulation to N_e , which we define as the number of electrons relative to a neutral system. The net system charge is then $Q = -N_e$. Since Q is an observable of an electronic structure calculation, it is a quantity that can be learned as a function of the atomic positions $\{\vec{R}_i\}$ and the potential ϕ .¹ Analogous to the Behler–Parrinello approach—in which the system energy is learned by decomposing it into per-atom energies, each of which is predicted by a regression model—we decompose the predicted system charge

into per-atom charges,

$$\hat{Q} = \sum_i^N \hat{q}_i(\{\vec{R}_i\}, \phi) \quad (1)$$

where the per-atom charges \hat{q}_i are predicted based upon the local atomic environment and the target potential. (Throughout this work, we will use a hat, $\hat{\cdot}$, to denote a predicted quantity.) We note that per-atom charges (q_i) are not a direct output of an electronic structure calculation, so in our scheme training takes place by minimizing the difference between \hat{Q} and Q for each image of the training set in a combined loss function, which we describe later. Doing this requires a different feature vector to be developed that incorporates the potential, which we describe in the next section.

Now that we have a framework to predict the per-atom charges, \hat{q}_i , we set out to predict the image’s grand-canonical energy, $\hat{\Omega} \equiv E - N_e \mu_e$, where E is the canonical energy, N_e is the net number of electrons (relative to a neutral system) and μ_e is the electron chemical potential, which is typically taken to be equivalent to the work function in grand-canonical electronic structure treatments¹⁶. We also take this convention here. $\hat{\Omega}$ is the energy returned by many grand-canonical calculators, such as the Solvated Jellium⁷ method implemented in GPAW^{17,18}. In the grand-canonical ensemble, the forces are conservative with $\hat{\Omega}$, not E ¹⁶. We use an ansatz that \hat{E} can be predicted from a summation of atomic terms \hat{E}_i , each of which can be predicted from a truncated charge expansion, similar to the expression from Rappe and Goddard¹⁹. Upon converting to $\hat{\Omega}$ we have

$$\begin{aligned} \hat{\Omega} &= \left(\sum_{i=1}^N \hat{E}_i \right) - N_e \mu_e = \sum_{i=1}^N (E^{0,Z(i)} + \hat{\chi}_i \hat{q}_i + \frac{1}{2} J^{Z(i)} \hat{q}_i^2 + \dots) + \sum_{i=1}^N \hat{q}_i \mu_e \\ &= \sum_{i=1}^N (E^{0,Z(i)} + (\hat{\chi}_i + \mu_e) \hat{q}_i + \frac{1}{2} J^{Z(i)} \hat{q}_i^2 + \dots) \end{aligned} \quad (2)$$

where $Z(i)$ is the element type of atom i , $E^{0,Z(i)}$ is a reference energy of element $Z(i)$, and $\hat{\chi}_i$ and $J^{Z(i)}$ are referred to as the electronegativity and hardness of element $Z(i)$ respectively. A full derivation of Eq. (2) is contained in Supplementary Note 1.1.1.

In our implementation, we take both $E^{0,Z(i)}$ and $J^{Z(i)}$ to be element-specific trainable parameters, while $\hat{\chi}_i$ is an environment-dependent per-atom electronegativity predicted from a machine-learning model, as proposed by Goedecker and colleagues³; that is,

$$\hat{\chi}_i = \hat{\chi}_i(\{\vec{R}_i\}) \quad (3)$$

The specific structure of the $\hat{\chi}_i$ machine-learning model is described in the section “Implementation”. The per-atom charges \hat{q}_i in Eq. (2) are those predicted from the charge-learning scheme of Eq. (1). In this way, both per-atom charges and per-atom

¹We are using $\{\vec{R}_i\}$ as a shorthand for the set of atomic positions of an image; that is, $\{\vec{R}_i | i \in \mathbb{Z}, 1 \leq i \leq N_{\text{atoms}}\}$.

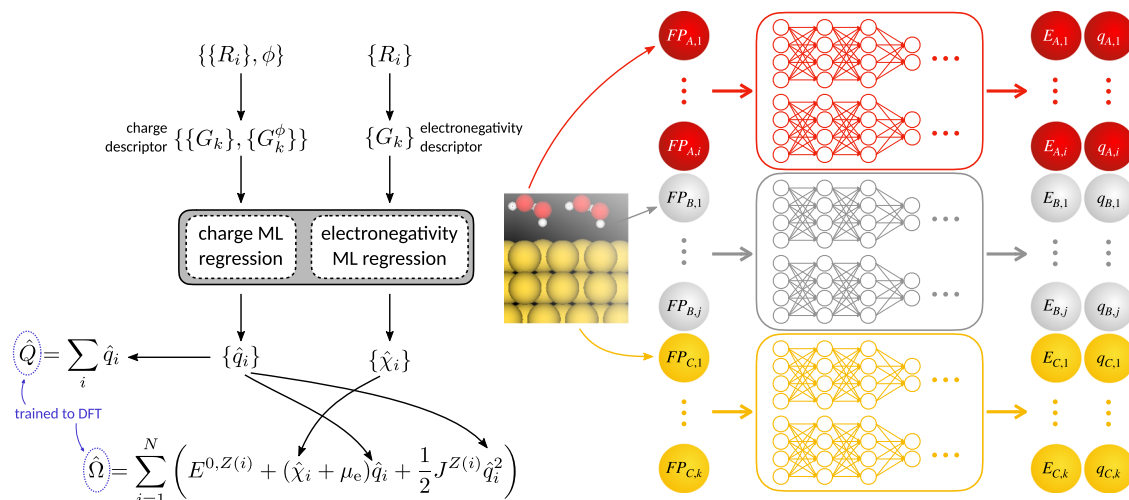


Fig. 2 Scheme for grand-canonical machine learning. Atoms with the same element type share the same neural-network structures and parameters. Two sub-neural networks—one predicts atomic charges from the geometry and electrode potential, the other predicts atomic electronegativities from solely geometry information—calculate the total energies and total charges at a fixed electrode potential.

electronegativities are deduced in a dual-learning scheme, leading to predictions of the per-image charge and the per-image energy.

One notable limitation of this scheme is the neglect of explicit inter-atomic Coulombic interactions in the formulation of the learned per-image energy. As such, any contribution of these interactions to the system energy will be accounted for implicitly within the learned electronegativities, $\hat{\chi}_i$. Since $\hat{\chi}_i$ only sees atoms within a specified cutoff radius (typically 5–8 Å), this can account for only short-range charge-interaction effects. Hence, this scheme may be best suited for application to systems with minimal longer-range charge interaction effects, such as homogeneous systems comprised of neutral and non-polar constituents. For all of the test systems included in this work, we found excellent emulation using only the charge–electronegativity formalism, and thus did not develop this aspect. If necessary, our scheme can be extended to include inter-atomic Coulombic interactions. For finite systems, Ω can be formulated to include the Coulomb energy, $\sum_{i<j} q_i q_j / r_{ij}$, such as described in the literature^{3,20–23}. For periodically replicated systems, the Coulomb interactions could be formulated in terms of an Ewald summation²⁴. However, care would need to be taken to account for the net charge in each unit cell, which if implemented naïvely would lead to infinite energies. Such an approach, while less practical, would ensure that learned energetics are physically interpretable.

Charge-predicting fingerprints

As described in Eq. (1), per-atom charges are predicted as a function of positions $\{\vec{R}_i\}$ and potential ϕ . Thus, any feature vector that enters a regression model (e.g., a neural network) must contain these two quantities. Standard feature vectors, such as the symmetry functions of Behler¹, depend only on $\{\vec{R}_i\}$. There are many conceivable ways to add potential-dependence to such vectors. Here, we simply extend a standard feature vector, $\{G_k\}$ as

$$\{G_k^{\text{charge}}\} = \{G_k\} \cup \{G_k^\phi\} \quad (4)$$

where we give each of the fingerprints in $\{G_k^\phi\}$ an exponential form:

$$G_{i,k}^\phi = \begin{cases} 0, & \text{if } r_{iz} \leq r_{z,\text{surface}} \\ \phi e^{-\eta_k \cdot (z_i - z_{\text{surface}})}, & \text{if } r_{iz} > r_{z,\text{surface}} \end{cases} \quad (5)$$

where ϕ is the electric potential; in practice, we use the work function of the simulation. z_i is the vertical position of the atom i being fingerprinted; this is aligned with the direction of the

electrostatic field. z_{surface} in this implementation indicates where the electrode–electrolyte interface is. In the current implementation, the surface is defined as the z coordinate of the top-most metal atom plus the van der Waals radius, with a custom-defined correction.

η_k is a hyper-parameter specific to the fingerprint element; in this work, we use a set of values of η_k of {2, 1, 0.5, 0.1}.

In the electrochemical simulations we emulate, the periodic boundary conditions along x - and y -directions model the infinite slab and solvents. The electrode is modeled as a finite-depth slab, as shown in Fig. 1 for a three-layer thick slab; however, this is a practical representation of a semi-infinite system. That is, the back side of this slab represents bulk-like atoms, and excess charge cannot be allowed to localize at this surface in properly constructed electrostatic simulations such as the SJ method⁷. Our charge fingerprints in this implementation set the electrostatic terms of the non-surface metal atoms to be zeros indicating the backside of the electrode cannot be affected by the external electrostatic field, while the charge fingerprints follow the exponential decay mimicking the strength of the electric field as the Gouy–Chapman model describes. In other words, the effect from the electric field on the atoms above the metal surface decays exponentially with the distance $z_i - z_{\text{surface}}$ to the surface, while the metal atoms below the surface representing the backside of electrode have a zero contribution from the electrostatic interactions with the field. This is also the case in reality where the back side of electrode is field-free⁷. These charge fingerprints could measure the interactions not only between atoms but also with the electric field. As such, we anticipate that this approach can be applied to any surface which can be approximated with a plane, including stepped surfaces; we show data on a stepped 211 surface in the Supplementary Figures. Of course, this may have to be re-developed for more elaborate surface geometries in future implementations.

The first part of the charge-predicting fingerprint (that is, $\{G_k\}_i$ in Eq. (4)) can be any fingerprint that describes chemical environments. In this paper, we used Gaussian symmetry functions, including G^{II} and G^{V} types as suggested in ref. ². Specific values of the fingerprint hyperparameters are contained in the section “Implementation”.

These charge-predicting fingerprints bypass the problem that the number of electrons in a simulation is unknown a priori, but rather is determined over the course of an electronic structure calculation from the atomic positions and the work function.

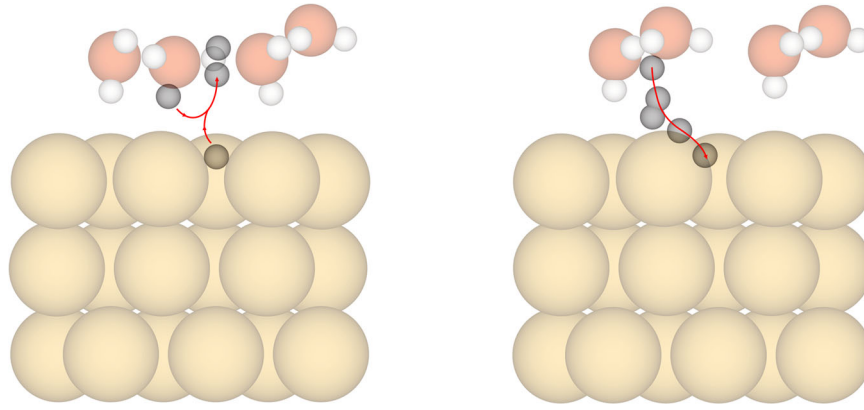


Fig. 3 Reaction geometries. Atomistic figure corresponding to Volmer (right) and Heyrovsky (left) reactions, respectively.

Loss functions and forces

Thus, the scheme is split into two sub-regression models, as shown in Fig. 2, to predict the charge \hat{Q} and energy $\hat{\Omega}$ of each image. The first regression model takes the atomic positions, $\{\mathbf{R}_i\}$, and the electrode potential, ϕ , as inputs and predicts per-atom charges $\{\hat{q}_i\}$, using an ML regression model. The sum of these charges, $\hat{Q} = \sum_i \hat{q}_i$, is compared to the total charge Q of the parent calculation. The second regression model takes as input the atomic positions $\{\mathbf{R}_i\}$ and predicts per-atom, environment-dependent electronegativities, $\{\hat{\chi}_i\}$. The electronegativities are used along with the set of charges predicted by the other model, $\{\hat{q}_i\}$ to predict the system energy $\hat{\Omega}$ by Eq. (2). $\hat{\Omega}$ is then compared to the actual Ω calculated by the parent calculator for each image. A combined loss function is assembled that simultaneously can optimize the parameters of the two ML models (along with the parameters E_i^0 and J_i):

$$L = \frac{1}{2} \sum_{j=1}^M \left[\left(\frac{\Omega_j - \hat{\Omega}_j}{N_j} \right)^2 + a_{\text{charge}} \cdot \left(\frac{Q_j - \hat{Q}_j}{N_j} \right)^2 \right] \quad (6)$$

where the summations are over the M images in the training set. a_{charge} is a hyper-parameter that lets us tune the relative importance of predicting charges versus energies. We also note that an alternate strategy is to first train the charge-predicting model, and then use the resulting per-atom charges to train the energy-predicting model. In practice, we found both strategies to work fine, but we implemented the combined loss function such that training could be accomplished in a single step.

The loss function above trains against the calculated charge and energy of each image. However, we can also fit to the atomic forces of a simulation, since atomic forces provide much more information than energy or charge alone. The force on each atom j can be calculated by a chain-rule derivation, since forces are equivalent to the negative gradient of Ω in the grand-canonical formalism¹⁶:

$$\hat{F}_j = - \sum_i \left(\hat{q}_i \frac{\partial \hat{\chi}_i}{\partial \mathbf{G}_i(\{\mathbf{R}\})} \cdot \frac{\partial \mathbf{G}_i(\{\mathbf{R}\})}{\partial \mathbf{R}_j} + (\hat{\chi}_i + J_i \hat{q}_i + \mu_e) \frac{\partial \hat{q}_i}{\partial \mathbf{G}_i(\{\mathbf{R}\})} \cdot \frac{\partial \mathbf{G}_i(\{\mathbf{R}\})}{\partial \mathbf{R}_j} \right) \quad (7)$$

This can be added to our loss function, so in the case of force training it becomes:

$$L = \frac{1}{2} \sum_{j=1}^M \left[\left(\frac{\Omega_j - \hat{\Omega}_j}{N_j} \right)^2 + a_{\text{charge}} \cdot \left(\frac{Q_j - \hat{Q}_j}{N_j} \right)^2 + \frac{a_{\text{force}}}{3N_j} \sum_{k=1}^3 \sum_{i=1}^{N_j} (F_{ijk} - \hat{F}_{ijk})^2 \right] \quad (8)$$

where here F_{ijk} indicates the k th Cartesian component of the force on atom i of image j .

Implementation

We implemented this scheme in our open-source machine-learning software AMP, the Atomistic Machine-Learning Package^{25,26}. An example script of the use of the method is contained in the Supporting information.

The machine-learning models of Eqs. (1) and (3) can, in principle, be any regression model; for generality and convenience we implemented them both as basic neural networks in the current work. All the neural networks were trained with the same structure and hyper-parameter settings. The machine learning descriptors were constructed following Behler and colleagues' suggestion^{1,2}. The cutoff radius was 6.5 Å. 36 symmetry functions were used in the descriptor. η 's were set to be {0.005, 4, 20, 80} in G^{II} -type symmetry functions and $\eta = 0.005$, $\zeta = \{1, 4\}$, $\gamma = \{+1, -1\}$ were chosen in G^{V} -type symmetry functions. The set of $\{\eta_k\}$ used in the charge-predicting fingerprints was {0.1, 0.2, 0.5, 2}. The neural-networks' hidden-layer structures were (20, 10) and (5, 5) for the charge and electronegativity networks, respectively. The convergence criterion was set to be 0.00005 (electrons) charge root-mean square error (RMSE) per atom and 0.0002 eV energy RMSE per atom.

Training and testing data

The bulk of the testing data reported in this work used a $2 \times 3 \times 3$ gold slab electrode which is periodic in the lateral (x, y) directions, whereas the z direction is normal to the electrode surface. The bottom layer is fixed during structure optimizations. Explicit water molecules were present near the surface (with implicit water above these), and extra protons/molecules were present in some simulations as described in the results. The bottom layer of the Au slab was fixed to represent the bulk of the electrode. All the grand-canonical DFT calculations were conducted in the solvated jellium method (SJM)⁷ in GPAW^{17,18}. A Monkhorst-Pack k -point grid of $4 \times 6 \times 1$ was employed and PBE²⁷ was used as the exchange–correlation functional. When structural optimization was employed, local optimization to below 0.03 eV Å⁻¹ and nudged elastic band (NEB)^{28,29} optimization to below 0.05 eV Å⁻¹ were used as targets.

The model was fit with a combination of images from both Volmer and Heyrovsky reactions. The convergence criterion for all systems in energy training was 0.2 meV RMSE per atom. All images in the training and testing sets were chosen randomly.

RESULTS

Here, we demonstrate the application of the grand-canonical machine-learning scheme to both replicate and predict the results of DFT calculations of electrified surfaces. We first report the ability

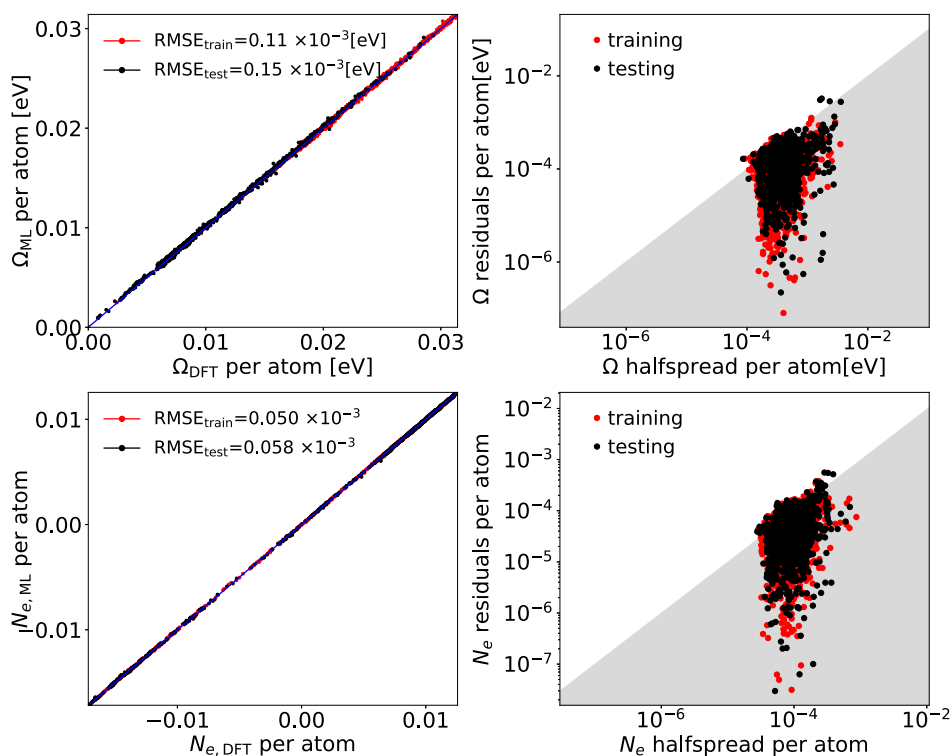


Fig. 4 Parity and uncertainty plots. The left panels describe the training and testing RMSE of the grand-canonical energy per atom (top) and the number of excess surface electrons (bottom) per atom in HER reactions. The energy parity plot is referenced to the minimum of energy per atom for plotting purpose. The right panels plot the ensemble half-spread (a predictor of uncertainty) versus the median error of the predictions from the bootstrap ensembles for HER reactions for the grand-canonical energy per atom (top) and the number of electrons per atom (bottom).

of the model on both training and testing data, and combine it with a bootstrap ensemble method⁴ to assess the usefulness of uncertainty predictions. We assess the model's ability to replicate both the energy and the charge of these simulations, using the energy/charge-predicting version of the model; that is, we did not train to forces in the first test.

Afterward, we test the scheme in a real-world application: namely the acceleration of nudged elastic band (NEB) calculations for saddle-point searching³⁰ in the electrochemical reactions at fixed electrode potentials. In this application, we employed the model with force training.

Finally, we test the ability of the method to extrapolate to larger-sized systems: specifically, we look at the ability of the model trained on fewer layers of water to predict energies and charges for more layers of water.

Accuracy of predictions, and uncertainty bounds

We tested two systems for both the accuracy of the prediction, as well as the reliability of the uncertainty bounds produced by a bootstrapping technique. The training and testing images were randomly chosen from DFT-calculated NEB iterations of the Volmer and Heyrovsky steps shown in Fig. 3 of the hydrogen evolution reaction, respectively:



Note that we chose the Heyrovsky as the second reaction instead of the Tafel reaction ($2\text{H} \rightarrow \text{H}_2 + 2*$), as it poses a greater test of constant-potential machine-learning methods due to the electron transfer. Particularly in the Heyrovsky step, the electron transfers rapidly when close to the saddle point—even though the

geometry only changes slightly; the steep barrier also leads to a large difference in potential energies at small geometric changes. The mixture of Volmer and Heyrovsky processes makes the training images more diverse.

The DFT reference data were calculated in the solvated jellium method⁷ as described earlier. The calculations were run on a Au fcc(111) surface at absolute electrode potentials ranging from 3.8 V to 4.6 V (about -0.6 V to 0.2 V on a standard hydrogen electrode (SHE) scale³¹) in 0.2 V increments. Sixty percent of the images were chosen randomly for training, with the rest reserved for testing.

Figure 4 compares the learned predictions to the electronic structure calculations; the top-left plot shows an excellent fit for the energies ($\Omega \equiv E - N_e\phi$) for both the training and testing images. The root-mean-square error (RMSE) is on the order of 0.1 meV/atom, which for a 50-atom system means the model is replicating and predicting the GC-DFT results to about 0.005 eV. For these systems, the precision of the learning representation ($\sim 10^{-3}$ eV) is safely within the accuracy range of typical DFT calculations ($\sim 10^{-1}$ eV)³². Charges (that is, the number of excess electrons) were replicated to an RMSE of 5×10^{-5} per atom, as shown in the bottom-left plot. The charge RMSEs for testing sets were in the range of 5×10^{-5} to 7×10^{-5} electrons.

The convergence criterion for all systems in energy training was 0.2 meV RMSE per atom, resulting in RMSEs in the testing set in the range of 0.25 to 0.50 meV atom⁻¹. All images in the training and testing sets were chosen randomly, and the number of images in the training sets was around 1500. We emphasize that the training and testing sets were randomly chosen from saddle-point searches for both the Volmer and Heyrovsky reactions with a wide range of work functions: $3.8, 4.0, 4.2, 4.4,$ and 4.6 eV (roughly spanning -0.6 to 0.2 V_{SHE}). It is notable that a single model could

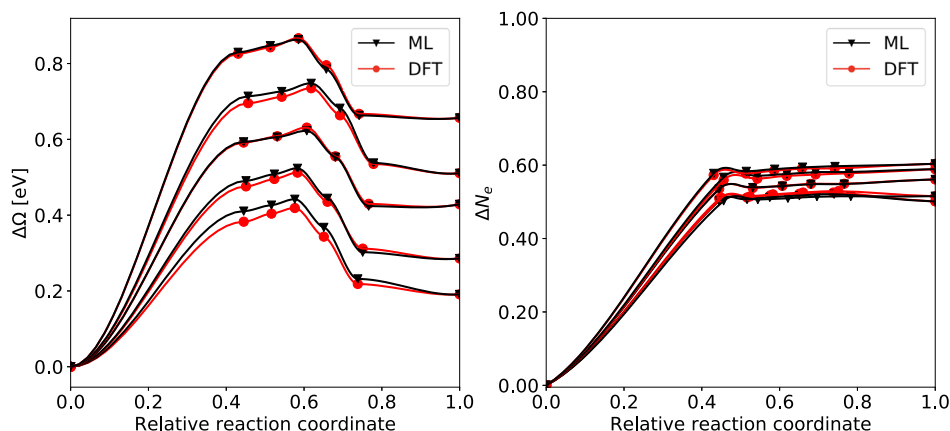


Fig. 5 Barrier-search acceleration. Comparison of reaction pathway (left) and charge transfer (right) along the band between pure DFT calculations and ML-assisted NEB for potentials (work functions) from 4.0 V to 4.8 V (from bottom to top).

be trained to fit all these potentials, and predictions could be made to any potential.

To assess the generality of the method across metal surfaces, we also tested on Pt, Cu, and Ag slabs; the results are similar and reported in the supporting materials.

In previous works with (canonical) atomistic learning^{4,33}, a bootstrap ensemble was shown to give a good estimate of the uncertainty inherent in predictions. Here, we test if this bootstrap-ensemble approach also can bracket the prediction uncertainty in the electronically grand-canonical ensemble. The right column of Fig. 4 shows runeplots for the halfspread (one-half the spread between the 5th and 95th percentile values of the model predictions) for these two systems; we expect the majority of points to lie below the parity line. We see just this for both the training and the test data, for both predictions of charge and energy. Thus, it appears that the bootstrap approach continues to work well in the grand-canonical framework.

For these systems, the ML training was about 100 and 800 times faster than the time of the parent DFT calculations (in CPU time) with and without force training, respectively. The resources required to call a trained model for calculation were negligible. Like most atomistic machine-learning approaches, the simulation time increases approximately linearly as the size of the system increases, in contrast to DFT which scales approximately cubically.

Saddle-point searches

The search for reaction barriers—first-order saddle points on the potential energy surface—is one of the most computationally demanding tasks in routine catalysis studies, and the ability to do so at a constant potential is one of the key advantages of electronically grand-canonical codes^{34,35}. We therefore expect one of the most common and demanding uses of a grand-canonical learning approach will be in saddle-point searches.

We first describe our best practices—without machine learning—in searching for electrochemical reaction barriers in an efficient manner, such that we can fairly assess the ability of machine learning to further speed up such calculations. Our goal is to find the reaction barrier (that is, the peak of the minimum-energy pathway, MEP) for a single reaction over a range of potentials, for example the proton-deposition (or Volmer) reaction on gold. We first start with a single potential; in this example we start at a work function of 4.4 eV ($\sim 0 V_{\text{SHE}}$). Our initial guess of the minimum energy path is comprised of a linear interpolation of all atomic positions between the initial and final states. We run the simulation with the DyNEB³⁶ approach, which is particularly efficient for serial NEB calculations. Once this converges, we use the resulting MEP as the initial guess for DyNEB runs at the

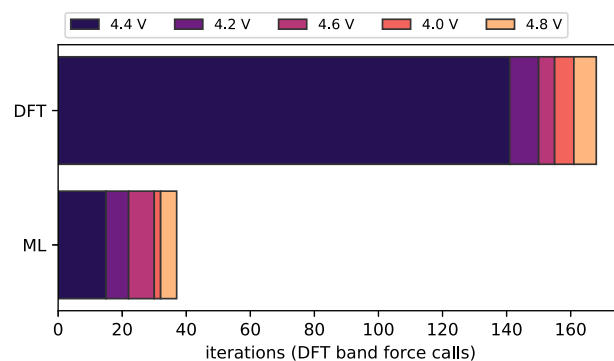


Fig. 6 Barrier-search acceleration. Comparison of number of DFT force calls of NEB band between DFT and ML-assisted DFT simulations.

neighboring potentials, of 4.2 and 4.6 eV; the good initial guess drastically reduces the number of NEB iterations required. We then use the 4.2 and 4.6 eV MEPs as initial guesses for the 4.0 and 4.8 eV runs, respectively. The number of band force calls in this approach is shown in Fig. 6; while the first potential takes 141 iterations, the subsequent potentials each take only 5–9 iterations.

To test the ability of atomistic learning to accelerate these grand-canonical barrier searches, we employ a simple method we published³⁰ for the acceleration of saddle-point searches in canonical calculations. We briefly describe the scheme; full details are in the reference. We start with an initial guess of the MEP, generated in an identical fashion to our conventional approach (either a linear interpolation or from a converged NEB at a nearby potential). We calculate each of these images in SJM-DFT, these ~ 7 images create a minimal training set on which we train a GC-ML calculator. We then use the trained ML as a surrogate calculator to run a NEB calculation to convergence, and we validate the predicted MEP by running each image in SJM-DFT. If the SJM-calculated forces indicate the NEB is converged, the algorithm is terminated. If not, we add these new images to our training set, re-train the calculator, and repeat until convergence.

To fairly assess the impact of machine learning, we adopt a similar strategy in the machine-learning scheme. We first converge the simulation at a work function of 4.4 eV, and use this converged MEP as the initial guess for the neighboring potentials, and so on. We further use the previous SJM-DFT images (for example, from 4.4 eV) as additional training data for the subsequent potentials; this is possible as our GC-ML approach is valid across differing potentials.

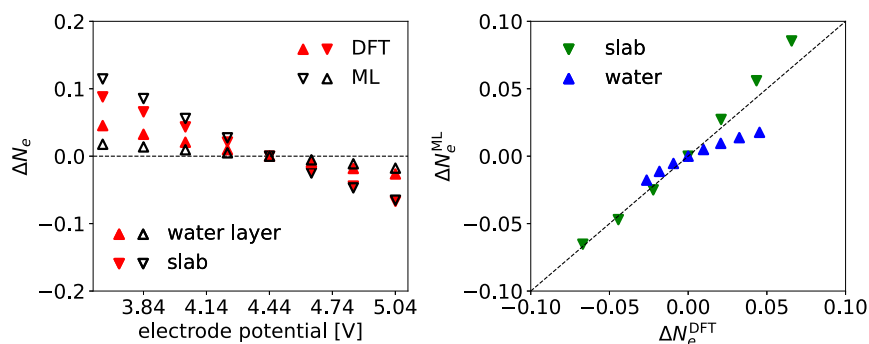


Fig. 7 Charge localization. Comparison between Bader analysis and ML on charge change in responding to electrode potentials varies on the Au fcc(111) surface with one layer of water and one extra proton. All electron changes are reference to potential 4.4V. Left: ΔN_e comparison between DFT (solid red) and ML (open black). The upper triangles represent the number of excess electrons summed from slab side and lower triangles represent that from water. Right: the parity comparison of the ΔN_e of water (blue) and slab (green) calculated from DFT (x-axis) and ML (y-axis), respectively.

Figure 5 compares the output of the ML-assisted scheme with the DFT-only scheme, in both energy and charge coordinates. We see that the ML-assisted scheme provides an excellent prediction of the MEPs on both metrics. We note that a slight deviation can be seen between the ML and pure-DFT approaches at some potentials. Within the tolerance of DFT-calculated NEBs (0.05 eV/Å), both are converged, so the difference is purely within the tolerance of the NEB method; that is, the machine-learning calculations are as valid as the DFT-alone calculations. From Fig. 4, we see that the GC-ML replicates the DFT results to within about 0.03 eV, much less than the difference in NEB tolerances.

The computational savings are shown in Fig. 6. We see a very large savings in the first potential (4.4 eV), when the system starts far away from the optimal reaction path. Since this is the bulk of the computational effort, this is the most important region for savings, and we see a reduction from 141 to 15 DFT-calculated band force calls. There is a much less dramatic reduction for the subsequent potentials, largely because the initial guess is so good in these bands already. On average, the standard DFT approach takes 6.75 band force calls for each of these conditions, while ML is slightly faster at 5.5 band force calls.

Charge partitioning

Although the partitioning of charge to atoms in an electronic structure calculation is arbitrary—that is, it requires assumptions about how to assign electron density to each atom—it is nevertheless useful to understand if the ML and the ab initio approaches are assigning charges in a similar manner. In the training scheme, we only require that the total charge, or excess electrons, of a system match the training data. Here, we examine if the charge partitioning in the ML scheme is reasonable.

To do so, we partition the charge into two regions: (1) the slab, or metal atoms, and (2) the water layer, including the excess proton if present. In the case of the ab initio data, we use a sum of per-atom Bader charges³⁷, although other partitioning schemes would also be valid. For the machine-learned implementation, per-atom charges are predicted directly, which we sum for each subsystem. Since atomic charge-partitioning schemes are inherently arbitrary, here we focus solely on the *differences* in per-atom charges as the potential is changed, which we expect to be less sensitive to the particular choice of charge-partitioning scheme. Examination of these charge differences can indicate if the machine-learning model and the ab initio model are localizing the excess electrons into similar locations within the simulation.

The initial state of the Volmer reaction on Au(111) was chosen for this test, covering several different electrode potentials. All results are reported in Fig. 7, referenced to the charge partitioning at a work function of 4.4 eV. Here, we see that the machine-

learned partitioning of excess electrons between the slab and the water layer matches the ab initio partitioning within reasonable deviations, suggesting the machine-learning model is localizing the charge in a similar manner to the ab initio model.

Extrapolation to larger solvent structures

A well-known limitation of DFT methods is the system size; it is typical to only use one or two layers of explicit water near the surface, although in the liquid phase many layers of explicit solvent may contribute to the observed kinetics. This is due to both the computational expense of DFT calculations ($\sim \mathcal{O}(N^3)$ scaling, where N is the number of electrons), and the many degrees of freedom of such systems, such that the computational expense precludes exploration of the potential energy surface. Thus, an intriguing use of GC-ML methods is for extending length scales; allowing both more layers of explicit water and greater exploration of the configurational space of those layers. Here, we explore the systematic ability of the GC-ML method to accommodate multiple water layers, and its ability to extrapolate.

Since electrochemical atomistic models are typically periodic along the slab planes (the x and y directions of the unit cell), it is natural that the ML calculator is transferable when the systems are enlarged by repeating along x and y directions, because both the chemical and electrostatic environments are similar for the corresponding atoms in repeating unit cells. However, when the systems are enlarged by adding more explicit water molecules—that is, extending the solvent in the z direction—the ML representation must be thoroughly tested.

We systematically explored the ability of GC-ML to make extrapolative predictions of extra water layers. We created a data set of 3647 images, comprised of images with 1, 2, 3, 4, and 5 water layers, at work functions of 3.8, 4.0, 4.2, 4.4, and 4.6 eV. These images additionally had geometrical variations of different layers of water and protonation of an arbitrary water molecule. Moreover, the configuration of water layers are also relaxed for more diverse in geometries. 124, 138, 197, 173, 30 training images with 1, 2, 3, 4, and 5 water layers respectively are sampled randomly with the remaining 250, 456, 741, 672, and 896 images for testing purposes.

We first trained the data on images that contain only a single water layer, and tested this model's ability to make predictions on systems with more water layers. The results are in the top row of Fig. 8, which uses the bootstrap-uncertainty approach to produce runeplots of the data. Here, we see that the one-layer testing structures fit similarly to the one-layer training structures. However, the bootstrap approach predicts a large uncertainty in the 2- to 5-layer structures, and we also see that the predictions are far from the true SJM-DFT results.

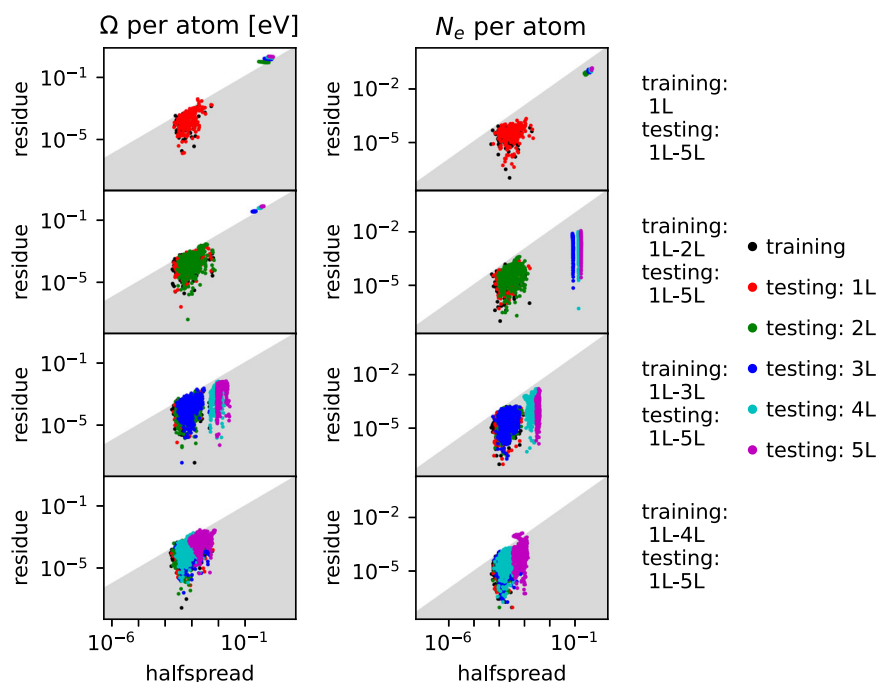


Fig. 8 Size extrapolation. Bootstrap uncertainty analysis for images with continuing adding more layers of water.

Next, we add training images from the two-layer system to our model, and use it to make predictions. As shown in the second row, now the one- and two-layer structures are fit well, but there are still deviations in the three- to five-layer structures.

As we repeat this systematic process, we see that by the time we reach four-layer training data, we are able to predict with reasonable certainty the five-layer structures. This is understandable, since the water molecules within the water layers take on three characteristic geometries: near the electrode surface, bulk-like, and near the top (implicit) surface. As the number of layers grows, examples of all of these water layers are present, and the system can reasonably be expected to extrapolate to larger numbers of water layers.

DISCUSSION

In the current work, we introduced a scheme for the machine-learning emulation of electronically grand-canonical electronic-structure calculations, of the type employed in studies of electrochemical interfaces. The framework we developed employs a dual-learning scheme where both the per-atom charge and the per-atom electronegativity are used to both predict the system charge and the system (grand-potential) energy. We showed that a single well-trained model could nicely predict both the charge and energy of a variety of structures from two different reactions across a range of potentials spanning 0.8 V. We also showed that a drop-in use of the existing bootstrap ensemble approach gives a good, yet conservative prediction on the uncertainty of each output quantity, as compared to the true deviations from grand-canonical density functional theory.

We further showed that this method works well to accelerate saddle-point searches within the grand-canonical framework, with the largest benefit coming in the regions where the system starts far from the optimal pathway. We also studied the ability to extrapolate to larger sizes, specifically more layers of water, and found that as more layers were added the extrapolation became more reliable, but in all cases the uncertainty predictions were able to identify unreliable predictions.

Like most machine-learning models, the current model has physical limitations compared to the true electronic-structure calculations and can be expected to fail in certain situations. In particular, the current model does not contain long-range Coulombic interactions, so if true long-range charge transfer occurs then this model may not capture it. It is conceivable that this model could be extended to account for such effects, but care must be taken to deal with the divergent nature of periodic finite charge summations.

DATA AVAILABILITY

A sample script for running *Amp* in the grand-canonical scheme is provided in the supporting information, as are parameters used within this study. Specific training sets are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The methods developed as part of the preparation of this manuscript are available as part of *Amp* (the Atomistic Machine-learning Package), which is publicly available at the following URL: <https://bitbucket.org/andrewpeterson/amp/>.

Received: 6 December 2022; Accepted: 16 March 2023;

Published online: 11 May 2023

REFERENCES

- Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
- Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **134**, 074106 (2011).
- Ghasemi, S. A., Hofstetter, A., Saha, S. & Goedecker, S. Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network. *Phys. Rev. B* **92**, 045131 (2015).
- Peterson, A. A., Christensen, R. & Khorshidi, A. Addressing uncertainty in atomistic machine learning. *Phys. Chem. Chem. Phys.* **19**, 10978–10985 (2017).
- Peterson, A. A. Acceleration of saddle-point searches with machine learning. *J. Chem. Phys.* **145**, 074106 (2016).
- Tran, K. & Ulissi, Z. W. Active learning across intermetallics to guide discovery of electrocatalysts for CO₂ reduction and H₂ evolution. *Nature Catal.* **1**, 696–703 (2018).

7. Kastlunger, G., Lindgren, P. & Peterson, A. A. Controlled-potential simulation of elementary electrochemical reactions: Proton discharge on metal surfaces. *J. Phys. Chem. C* **122**, 12771–12781 (2018).
8. Chan, K. & Nørskov, J. K. Electrochemical barriers made simple. *J. Phys. Chem. Lett.* **6**, 2663–2668 (2015).
9. Sundaraman, R., Goddard, W. A. & Arias, T. A. Grand canonical electronic density-functional theory: Algorithms and applications to electrochemistry. *J. Chem. Phys.* **146**, 114104 (2017).
10. Gauthier, J. A. et al. Unified approach to implicit and explicit solvent simulations of electrochemical reaction energetics. *J. Chem. Theory Comput.* **15**, 6895–6906 (2019).
11. Hutchison, P., Warburton, R. E., Soudackov, A. V. & Hammes-Schiffer, S. Multi-capacitor approach to interfacial proton-coupled electron transfer thermodynamics at constant potential. *J. Phys. Chem. C* **125**, 21891–21901 (2021).
12. Limmer, D. T., Willard, A. P., Madden, P. & Chandler, D. Hydration of metal surfaces can be dynamically heterogeneous and hydrophobic. *Proc. Natl Acad. Sci. USA* **110**, 4200–4205 (2013).
13. Bowler, D. R. & Miyazaki, T. $\mathcal{O}(N)$ methods in electronic structure calculations. *Rep. Prog. Phys.* **75**, 036503 (2012).
14. Huang, Y., Kang, J., Goddard, W. A. & Wang, L.-W. Density functional theory based neural network force fields from energy decompositions. *Phys. Rev. B* **99**, 064103 (2019).
15. Kang, J. & Wang, L.-W. First-principles Green-Kubo method for thermal conductivity calculations. *Phys. Rev. B* **96**, 020302 (2017).
16. Lindgren, P., Kastlunger, G. & Peterson, A. A. Electrochemistry from the atomic scale, in the electronically grand-canonical ensemble. *J. Chem. Phys.* **157**, 180902 (2022).
17. Enkovaara, J. et al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *J. Phys.: Condens. Matter* **22**, 253202 (2010).
18. Mortensen, J. J., Hansen, L. B. & Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B* **71**, 035109 (2005).
19. Rappe, A. K. & Goddard, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **95**, 3358–3363 (1991).
20. Artrith, N., Morawietz, T. & Behler, J. High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide. *Phys. Rev. B* **83**, 153101 (2011).
21. Siepmann, J. I. & Sprik, M. Influence of surface topology and electrostatic potential on water/electrode systems. *J. Chem. Phys.* **102**, 511–524 (1995).
22. Reed, S. K., Lanning, O. J. & Madden, P. A. Electrochemical interface between an ionic liquid and a model metallic electrode. *J. Chem. Phys.* **126**, 084704 (2007).
23. Nakano, H. & Sato, H. A chemical potential equalization approach to constant potential polarizable electrodes for electrochemical-cell simulations. *J. Chem. Phys.* **151**, 164123 (2019).
24. Wells, B. A. & Chaffee, A. L. Ewald summation for molecular simulations. *J. Chem. Theory Comput.* **11**, 3684–3695 (2015).
25. Khorshidi, A. & Peterson, A. A. Amp: a modular approach to machine learning in atomistic simulations. *Comput. Phys. Commun.* **207**, 310–324 (2016).
26. Peterson, A., El Khatib, M., Khorshidi, A. & Zeng, C. Amp: the atomistic machine-learning package v1.0 (2022). <https://doi.org/10.5281/zenodo.7035955>.
27. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
28. Henkelman, G., Uberuaga, B. P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000).
29. Sheppard, D., Terrell, R. & Henkelman, G. Optimization methods for finding minimum energy paths. *J. Chem. Phys.* **128**, 134106 (2008).
30. Peterson, A. A. Acceleration of saddle-point searches with machine learning. *J. Chem. Phys.* **145**, 074106 (2016).
31. Trasatti, S. The concept and physical meaning of absolute electrode potential: a reassessment. *J. Electroanal. Chem. Interfacial Electrochem.* **139**, 1–13 (1982).
32. Gillan, M. J., Alfè, D. & Michaelides, A. Perspective: How good is dft for water? *J. Chem. Phys.* **144**, 130901 (2016).
33. Zeng, C., Chen, X. & Peterson, A. A. A nearsighted force-training approach to systematically generate training data for the machine learning of large atomic structures. *J. Chem. Phys.* **156**, 064104 (2022).
34. Steinmann, S. N., Michel, C., Schwiedernoch, R., Filhol, J.-S. & Sautet, P. Modeling the HCOOH/CO₂ electrocatalytic reaction: When details are key. *ChemPhysChem* **16**, 2307–2311 (2015).
35. Lindgren, P., Kastlunger, G. & Peterson, A. A. A challenge to the $G \sim 0$ interpretation of hydrogen evolution. *ACS Catal.* **10**, 121–128 (2020).
36. Lindgren, P., Kastlunger, G. & Peterson, A. A. Scaled and dynamic optimizations of nudged elastic bands. *J. Chem. Theory Comput.* **15**, 5787–5793 (2019).
37. Bader, R. F. W. *Atoms in Molecules: A Quantum Theory* (Clarendon Press, 1990).

ACKNOWLEDGEMENTS

The authors thank Cheng Zeng and Georg Kastlunger, of Brown University, for many valuable technical discussions. The authors acknowledge support from the U.S. Department of Energy under Award DE-SC0019441 and the National Science Foundation under award 1553365. Calculations were undertaken at Brown University's Center for Computation and Visualization.

AUTHOR CONTRIBUTIONS

A.A.P. conceived the study. X.C. and M.E.K. developed the initial framework. P.L., A.W., A.J.M., A.A.P. contributed concepts such as grand-canonical transformations, potential-containing fingerprints, charge interactions, etc., which contributed to the final form. X.C. and M.E.K. modified the software. X.C. and A.A.P. drafted the manuscript, while all other authors provided input on writing.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-01007-6>.

Correspondence and requests for materials should be addressed to Andrew A. Peterson.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023