

<https://doi.org/10.1038/s41524-024-01295-6>

Machine learning-aided first-principles calculations of redox potentials

Check for updates

Ryosuke Jinnouchi¹ ✉, Ferenc Karsai² & Georg Kresse^{2,3}

We present a method combining first-principles calculations and machine learning to predict the redox potentials of half-cell reactions on the absolute scale. By applying machine learning force fields for thermodynamic integration from the oxidized to the reduced state, we achieve efficient statistical sampling over a broad phase space. Furthermore, through thermodynamic integration from machine learning force fields to potentials of semi-local functionals, and from semi-local functionals to hybrid functionals using Δ -machine learning, we refine the free energy with high precision step-by-step. Utilizing a hybrid functional that includes 25% exact exchange (PBE0), this method predicts the redox potentials of the three redox couples, $\text{Fe}^{3+}/\text{Fe}^{2+}$, $\text{Cu}^{2+}/\text{Cu}^+$, and $\text{Ag}^{2+}/\text{Ag}^+$, to be 0.92, 0.26, and 1.99 V, respectively. These predictions are in good agreement with the best experimental estimates (0.77, 0.15, 1.98 V). This work demonstrates that machine-learned surrogate models provide a flexible framework for refining the accuracy of free energy from coarse approximation methods to precise electronic structure calculations, while also facilitating sufficient statistical sampling.

Green energy and a circular economy are some of the key paradigms that our human society needs to realize in the next few decades. This implies that we need to give up on the combustion of fossil fuels. A key element to achieve this paradigm shift is the use of electrochemistry, be it for batteries and fuel cells, to convert electrical energy to hydrogen or other valuable chemicals, or to convert hydrogen back to energy without direct combustion in air.

The redox potential of electron transfer (ET), $\text{Ox} + ne^- \rightarrow \text{Red}$ in liquids, is an essential property for a variety of electrochemical energy conversion devices, such as batteries, fuel cells, and electrochemical fuel synthesis. It determines the alignment of redox levels relative to the Fermi level of a metal, or valence band maximum (VBM) and conduction band minimum (CBM) of semiconductor and insulator electrodes. It also determines the stability windows of ions and molecules in solutions, that is the range of voltages within which a specific ion or molecule can undergo electrochemical reactions. This information is vital to designing redox species and solvent molecules, such as redox couples for redox-flow batteries¹, solvents and additives for Li-ion batteries^{2–4}, radical scavengers for fuel cells⁵ and electrocatalysts for fuel synthesis^{6,7}.

Unfortunately, to date, accurate first-principles (FP) predictions of this crucial property remain challenging, with typical prediction errors around 0.5 V. Sprik and co-workers developed a thermodynamic integration (TI) method utilizing the computational standard hydrogen electrode (CSHE)^{8,9} and applied this method to several redox reactions in aqueous solutions^{10,11}.

They discovered that the use of a semi-local functional leads to errors exceeding 0.5 V. This discrepancy arises because the functional inaccurately yields the shallow valence band edge and the deep conduction band edge, resulting in incorrect hybridization with the redox levels. Similar magnitudes of errors have also been observed in other FP calculations that employ semi-local approximations^{12,13}. As a result, Sprik and co-workers opted for a hybrid functional. Nonetheless, they observed a significant spread of values for two metal ion couples, with the $\text{Cu}^{2+}/\text{Cu}^+$ couple ranging from -1.13 to -0.20 V (experimental value 0.16 V) and the $\text{Ag}^{2+}/\text{Ag}^+$ couple ranging from 0.90 to 1.72 V (experimental value 1.98 V)^{11,14}. These variations were attributed to differences in the pseudopotential and the computational code base (CMPD versus CP2K). While the “best” values obtained using the hybrid functional and highly accurate pseudopotentials are relatively close to experimental values (-0.20 V for Cu, and 1.72 V for Ag), the agreement is still far from being quantitative. Due to the high computational cost of hybrid functionals, most calculations have been performed using approximated methods, such as continuum solvation models^{15–18} and QM/MM models^{19,20}. Although these models can reproduce the experimental redox potentials of ions and molecules with convenient accuracy, the computational results heavily rely on many approximations, making it unclear which predictions are strictly correct. Here, we briefly note that these FP and approximated methods have been extended to electrochemistry at liquid-solid interfaces^{21–29}. Nowadays, these methods have become indispensable for elucidating electrochemical interfacial phenomena and designing

¹Toyota Central R&D Labs., Inc., Nagakute 480-1192 Aichi, Japan. ²VASP Software GmbH, Berggasse 21, A-1090 Vienna, Austria. ³Computational Materials Physics, Faculty of Physics, University of Vienna, Kolingasse 14, A-1090 Vienna, Austria. ✉e-mail: e1262@mosk.tytlabs.co.jp

advanced materials^{30–35}. However, even in the calculation of redox reactions at interfaces, approximations are made in most applications, such as representing the motion of atomic nuclei with simple statistical models like the harmonic oscillator model^{12,22,30,31,33,36,37}, modeling solvents by reference interaction site model based on the integral equation theory²⁶, or modeling by continuum mediums^{23–25,27,29}. A rigorous FP method that eliminates these approximations is also desired in the field of interfacial electrochemistry.

The main goals of the present work are three-fold: First, we want to accurately calculate the redox potential of metal ions in water for three prototypical cases: Ag, Cu, and Fe. Ag²⁺ ions are among the most aggressive oxidants with a large redox potential, whereas the redox potential of Cu²⁺ ions is fairly shallow, and the Fe³⁺/Fe²⁺ reaction lies in between. The first two redox reactions involve large changes in the ion water coordination, which makes the calculation challenging, whereas the redox reaction of Fe is a so-called simple outer sphere ET reaction and has been the subject of numerous experimental and theoretical studies³⁸. The Fe ions are conceived to be particularly challenging for density functional theory. Second, we want to establish a computationally feasible pathway that yields statistically accurate results. Last but not least, we want to systematically explore different density functionals to set a guideline for future studies.

Results

Free energy change of electron transfer reaction

We begin with an overview of the used theory and modeling. Further details can be found in the Methods section. The reactions evaluated in this study

are electron transfer reactions in water: Fe³⁺ + e⁻ ↔ Fe²⁺, Cu²⁺ + e⁻ ↔ Cu⁺, and Ag²⁺ + e⁻ ↔ Ag⁺. We assume that other side reactions do not occur, and only the valency of redox species changes due to the reaction similar to the previous study¹¹. The redox potential U_{redox} is determined by the free energy difference ΔA between the reduced and oxidized states as

$$U_{\text{redox}} = -\frac{\Delta A}{en}, \quad (1)$$

where e is the elementary charge and n is the number of electrons involved in the reaction. Here, we also assume that the change in volume during the one electron-transfer half reaction is negligible similar to the previous studies^{8,11}. Then, the Gibbs free energy is replaced by the Helmholtz free energy (A). The free energy difference ΔA can be exactly determined by thermodynamic integration (TI)^{39,40}:

$$\Delta A = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda. \quad (2)$$

Here, $\langle X \rangle_\lambda$ denotes the expectation value of X for an ensemble created by the Hamiltonian at coupling λ . The integral seamlessly connects the oxidized state ($\lambda = 0$) to the reduced state ($\lambda = 1$) along a coupling path^{41,42}. The potential energy surface upon which atoms move is described by the grand potential Ω of the system opened for electrons⁴³. Consequently, the

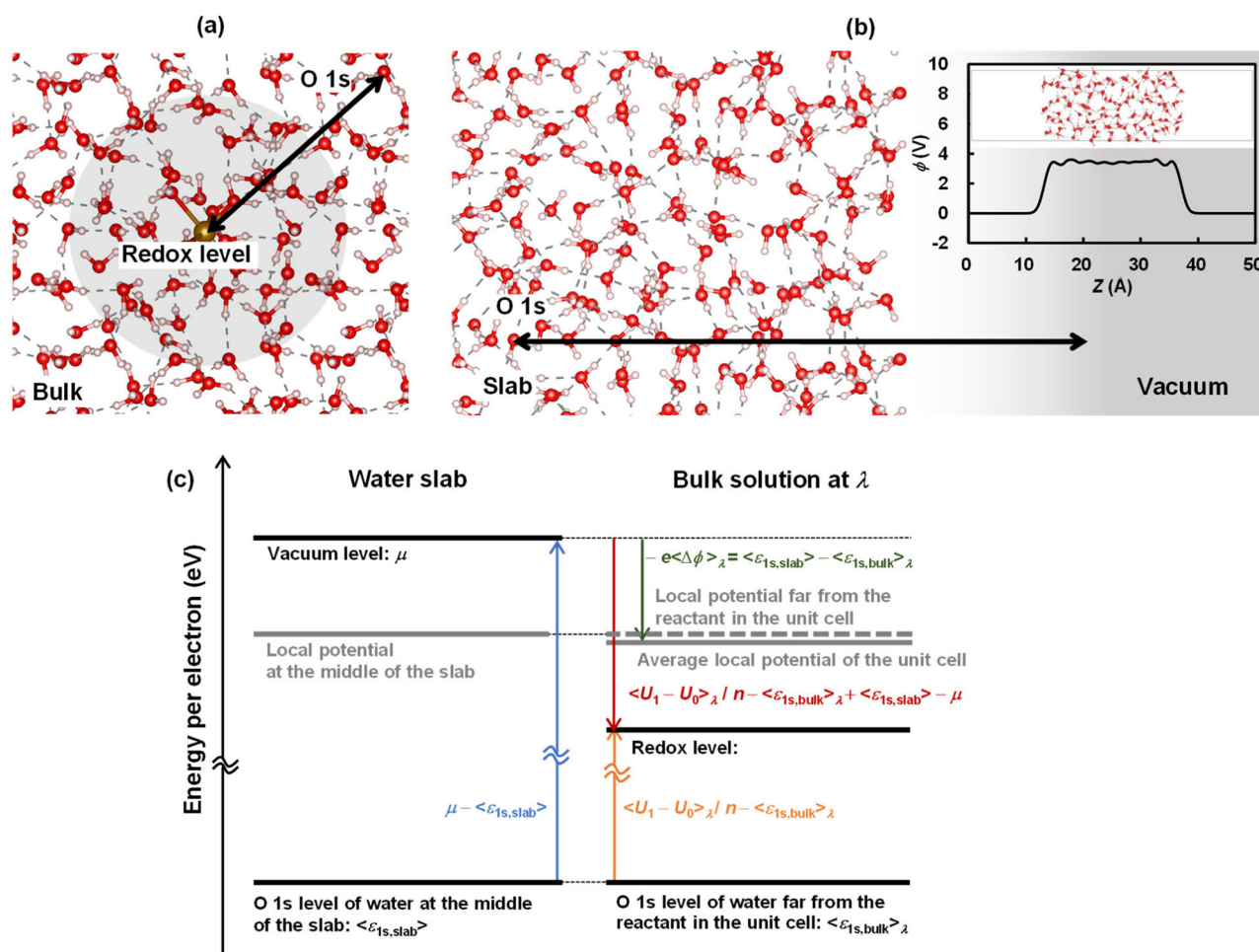


Fig. 1 | Aligning energy levels based on the O 1s level of water molecules. a Aligning the redox level based on the O 1s level of water molecules far from the redox species in the bulk solution model under a periodic boundary condition, (b) aligning the O 1s level of water molecules at the middle of the slab based on the local

potential at the middle of the vacuum layer in the slab model, and (c) schematic of the alignment. The figure inset in (b) shows the snapshot of the water slab and computed local potential profile across the water slab. The graphics showing bulk and interfacial models are made by VESTA¹⁰⁵.

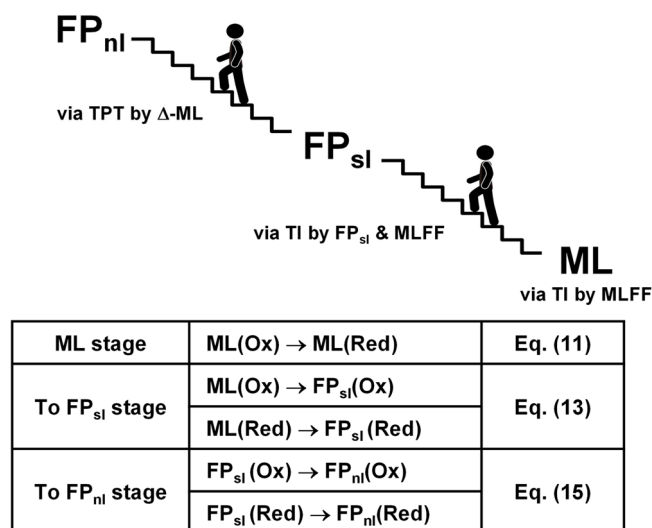


Fig. 2 | Schematic of ML-aided TI and TPT to compute the free energy difference ΔA . The notations ML, FP_{sl} and FP_{nI} mean the machine-learning force field, FP method with semi-local functional and FP method with non-local hybrid functional, respectively. See details of equations in the Methods section.

Hamiltonian of the system is described as follows:

$$H = \sum_{i=1}^{N_a} \frac{|\mathbf{p}_i|^2}{2m_i} + \Omega, \quad (3)$$

$$\Omega = U - \mu N, \quad (4)$$

where N_a is the number of atoms, m_i and \mathbf{p}_i are the mass and momentum vector of the i -th atom, and μ and N are the chemical potential and the number of electrons. The chemical potential μ is fixed at the reservoir level, whereas N varies by n along the coupling path. U represents the potential energy surface at λ , equating to the sum of the Helmholtz free energy of the electronic subsystem and the electrostatic interactions among nuclei. Following previous studies^{41,42}, U can be described as

$$U = \lambda U_1 + (1 - \lambda)U_0, \quad (5)$$

where U_0 and U_1 are the potential energies of the oxidized and reduced states, respectively. Hence, the free energy difference ΔA is written as

$$\Delta A = \int_0^1 \langle U_1 - U_0 \rangle_\lambda d\lambda - \mu n. \quad (6)$$

If the structural changes are significant from the oxidized to the reduced species — recall that this is the case for Ag and Cu — many integration steps are required to accurately determine the energy difference. The application of this approach entails two difficulties. (i) Clearly, it implies huge computational cost if applied directly to hybrid functionals; if 100,000 timesteps using a complete plane wave basis set are required to obtain good statistical accuracy, several 10 mio core hours are necessary. (ii) Second, during the reaction one electron needs to be transferred from the reservoir, characterizing the chemical potential of the electrons. The vacuum level is the best-suited reference chemical potential that allows one to align the redox levels and band edges of the electrode in the absolute potential scale. However, in FP calculations of bulk systems under periodic boundary conditions, the vacuum level is a quantity that cannot be directly accessed during simulations.

Chemical potential of electrons

We will address the second point (ii) first. Jiao and co-workers⁴⁴ suggested using the average electrostatic potential as a suitable reference point, and Leung⁴⁵ calculated the position of the average electrostatic potential with respect to the vacuum level in a second independent calculation involving a water slab. We refine this approach in a conceptually easy-to-understand way that simultaneously reduces finite-size errors. As a reference, instead of using the vacuum level, we employ the O 1s level of water, which is fixed relative to the vacuum level and can be conveniently calculated with the FP code used in this study. Our approach is schematically illustrated in Fig. 1. In FP calculations of a solution system under a periodic boundary condition, the energy contribution $\langle U_1 - U_0 \rangle_\lambda$ in Eq. (6) is equal to the negative electron affinity of the oxidized species scaled to the average local potential of the system. The same calculation can also determine the O 1s level $\langle \epsilon_{1s,bulk} \rangle_\lambda$ of water, sufficiently far from the redox species and unaffected by the reactant, scaled to the average local potential. Therefore, measuring the redox level using the O 1s level as a reference results in $\langle U_1 - U_0 \rangle_\lambda / n - \langle \epsilon_{1s,bulk} \rangle_\lambda$, as highlighted in orange letters in Fig. 1c. In practice, $\langle \epsilon_{1s,bulk} \rangle_\lambda$ may slightly vary along the coupling path due to finite size effects (refer to Supplementary Table 4). By aligning the potentials between the 'defect' and the 'host' within the same supercell in this manner, the finite size effects can be mitigated^{46,47}. The vacuum level referenced to the O 1s level can be calculated using a slab model. As depicted in Fig. 1b, when referencing the O 1s level of water molecules located in the middle layer of the water slab, the vacuum level can be expressed as $\mu - \langle \epsilon_{1s,slab} \rangle$, as indicated in blue letters in Fig. 1c. The difference between the redox level and vacuum level scaled to the O 1s level results in the redox level scaled to the vacuum level, as shown in red letters in Fig. 1c. Consequently, the free energy difference ΔA on an absolute scale is written as

$$\Delta A = \int_0^1 \langle U_1 - U_0 \rangle_\lambda d\lambda - ne\Delta\bar{\phi}, \quad (7)$$

$$e\Delta\bar{\phi} = \int_0^1 \langle \epsilon_{1s,bulk} \rangle_\lambda d\lambda - \langle \epsilon_{1s,slab} \rangle, \quad (8)$$

where the vacuum level μ is set to zero. As illustrated by the green letters in Fig. 1c, $\Delta\bar{\phi}$ accounts for the difference between the local potential at the vacuum in the slab model and the one in the bulk solution model.

Equation (7) is similar to the approach used in the CSHE method described in previous studies^{48,49}. In these studies, the electrostatic potential of water was employed for alignment instead of the O 1s level. As shown by the gray dashed line in Fig. 1c, using the electrostatic potential, referred to here as the local potential, away from the redox species yields a result that is consistent with the use of the O 1s level within the statistical error bar (refer to Supplementary Note 3 for the case of pure water). However, our method differs in two key ways from the previous approach. First, we calculate the absolute vacuum reference rather than using SHE as a reference, which allows for the assessment of absolute potentials in half-cell reactions. Second, machine-learned (ML) force fields (FFs) can create many statistically independent configurations for the water slab. We do this by on-the-fly learning an H₂O force field for the bulk and then for the surface and performing finally extensive million-step (total 1.5 ns) ML molecular dynamics for the surface. From this simulation, we draw 3000 statistically independent snapshots. Only for these snapshots, FP calculations are performed to determine the average O 1s level with respect to the vacuum level. This substantially reduces the required computational time from 1 mio core hours for brute force runs using the semi-local functional to only 2200 core hours for the FP calculations on 3000 structures, including the ML simulations and training runs, while retaining statistical accuracy, as demonstrated by the local potential profile shown in the inset of Fig. 1b (see details of the estimation of compute time in Supplementary Note 2).

Table 1 | Redox potentials U_{redox} of three reported in ref RPBE +D3, PBE0 (0.25), PBE0 (0.50), PBE0+D3 (0.25) and PBE0+D3 (0.50) using MLFF and Δ -ML

XC functional	Fe	Cu	Ag	RMSE
RPBE+D3	0.80	0.66	1.88	0.29
PBE0 (0.25)	0.92	0.26	1.99	0.11
PBE0 (0.50)	0.79	-0.34	2.12	0.30
PBE0+D3 (0.25)	0.94	0.24	2.02	0.11
PBE0+D3 (0.50)	0.83	-0.38	2.13	0.32
HSE06 ^[a]		-0.20	1.72	0.31
Exp. ^[b]	0.77	0.15	1.98	

Here, values in the parenthesis are the fraction of the exact exchange. The results for 64 water molecular systems are tabulated. The potential of SHE is set to 4.44 V reported in ref. 102. The root means square errors (RMSE) compared to the experimental redox potentials are also shown. The results of the HSE06 functional reported by Sprick and co-workers^[a] are also listed.

^[a] From ref. 11. The calculations were done by the CP2K¹⁰³ code with using the Goedecker-Teter-Hutter dual Gaussian norm-conserving pseudopotential¹⁰⁴ with the electronic configuration of 3s²3p⁶3d¹⁰4s¹ for Cu and 4s²4p⁶4d¹⁰5s¹ for Ag and localized Gaussian basis sets for orbital.

^[b] From ref. 14.

Thermodynamic integration

To address the problem of computing the free energy difference, i.e. point (i), we propose the ML-aided scheme as depicted in Fig. 2. Here, we use the abbreviations $\text{FP}_{\text{nl}}(\text{Ox/Red})$, $\text{FP}_{\text{sl}}(\text{Ox/Red})$, and $\text{ML}(\text{Ox/Red})$ to denote calculations using a *non-local* hybrid functional, a *semi-local* functional and *machine-learned* force field for the oxidized and reduced cases, respectively. Naively, one could just perform the required TI using ML surrogate models. As we will show later in this article, this yields only acceptable accuracy. Errors in ΔA resulting from inaccuracies in the trajectory and the energy predictions by the ML potential can be corrected by performing TI from the ML potential to the FP potential for both the oxidized and reduced states. We will adopt this strategy for the FP_{sl} method. So this involves two calculations: TI from the oxidized to the reduced species using ML surrogate models via Eq. (11) in the Methods section, $\text{ML}(\text{Ox}) \rightarrow \text{ML}(\text{Red})$, and then for each oxidation state, TI from MLFF to the FP_{sl} Hamiltonian via Eq. (13), $\text{ML}(\text{Ox}) \rightarrow \text{FP}_{\text{sl}}(\text{Ox})$ and $\text{ML}(\text{Red}) \rightarrow \text{FP}_{\text{sl}}(\text{Red})$. This two-step integration has three advantages as summarized below:

- The integration $\text{ML}(\text{Ox}) \rightarrow \text{ML}(\text{Red})$ using the MLFF takes into account most of the non-linear components of the integrand in the TI (see Supplementary Figs. 8, 9). Excellent statistical accuracy can be reached for this step.
- The MLFFs also provide well-equilibrated initial structures required for other calculational steps.
- The integrands in $\text{ML}(\text{Ox}) \rightarrow \text{FP}_{\text{sl}}(\text{Ox})$ and $\text{ML}(\text{Red}) \rightarrow \text{FP}_{\text{sl}}(\text{Red})$ are small and almost linear in the coupling parameter (see Supplementary Fig. 10) owing to the accurate reproduction of the FP_{sl} structures by the MLFF (see Overview of results). Hence, these demanding integrals (evaluation of FP_{sl} calculation in every MD step) converge using a few tens of picosecond MD simulations.

There is one final obstacle though: performing TI to a potential calculated by a hybrid functional that includes non-local exchange (FP_{nl}) is still exceedingly demanding and challenging. So in this specific case, as depicted in Fig. 2, we have decided to apply the Δ -machine learning (Δ -ML)^{48–55} which learns the *difference* ΔU between the FP_{sl} potential and the FP_{nl} potential. Due to the very smooth energy difference between the FP_{sl} functional and the FP_{nl} functional, it is possible to learn an extremely accurate ML representation of ΔU with just a few tens of FP_{nl} calculations, with errors significantly smaller by an order of magnitude or more compared to those associated with MLFF models (see details in Supplementary Figs. 2 to 4 and Figs. 16 and 17). In the current implementation, the TI integration has been replaced with thermodynamic perturbation theory

(TPT),

$$\Delta A = A_1 - A_0 = -\frac{1}{\beta} \langle e^{-\beta \Delta U} \rangle_0 = -\frac{1}{\beta} \langle e^{\beta \Delta U} \rangle_1, \quad (9)$$

where β is the inverse temperature, and the symbol ΔU denotes the potential energy difference between the two end points. Although Eq. (9) is in principle exact, the potential energy difference might need to be evaluated for thousands or even many ten thousand configurations if the ensembles generated by the two potentials are too distinct. This implies the significantly expensive FP_{nl} calculations. The Δ -ML scheme allows for the circumvention of this issue, enabling the reduction of the required FP_{nl} calculations to merely tens. Thanks to the remarkable accuracy of the Δ -ML models, it is possible to obtain exceedingly accurate free energy differences between different FP methods without further correction (see Supplementary Fig. 12). This is one of the key advances of the present work. The computational cost is ultimately only limited by generating sufficient configurations using the FP_{sl} . Thus, the required compute time for direct TI using the FP_{nl} method is reduced from 20 mio core hours to 16800 core hours for the FPMD simulations that generate configurations using the FP_{sl} method (see details of the estimation in Supplementary Note 2).

Overview of results

We now detail our results and will show that the adopted procedure yields statistically highly accurate results. The calculations were performed using VASP^{56,57} and the projector-augmented wave (PAW) method^{58,59}. For the ML force fields (MLFFs) the implementation detailed in previous publications is used^{60–62}. Similar to the pioneering ML approaches^{50,63,64}, the potential energy in our MLFF method is approximated as a summation of local energies [see Eq. (20)]. The local energy is approximated as a weighted sum of kernel basis functions [see Eq. (21)]. A Bayesian formulation allows to efficiently predict energies, forces and stress tensor components as well as their uncertainties. The predicted uncertainty enables the reliable sampling of the reference structures on the fly during the FPMD simulation. Details of the equations, parameters and training conditions are summarized in the Methods section and Supplementary Note 1. As in the previous studies^{60,65–67}, the MLFFs trained on a semi-local functional with dispersion corrections achieve root mean square errors (RMSEs) of 1–5 meV atom⁻¹ and 0.04–0.11 eV Å⁻¹ for energies and forces (see error distributions in Supplementary Figs. 1 to 4). The three ET reactions are examined in water by using a semi-local functional⁶⁸ with a dispersion correction^{69,70} (RPBE+D3) and hybrid functionals^{71,72} with and without a dispersion correction (PBE0 and PBE0+D3). Systematic comparisons of different functionals help us to study the effects of the exact exchange as well as dispersion corrections. As shown in Table 1 [see lines of PBE0 (0.25) and PBE0+D3 (0.25)] good agreement with experiment is achieved using the hybrid functional with 1/4 exact exchange, regardless of whether dispersion corrections are used or not.

Water surface calculations

For RPBE+D3, the present MLFF provides a surface tension of 79 ± 5 mN m⁻¹ for the 128 molecular system and 84 ± 5 mN m⁻¹ for the 1024 molecular system at 298 K. Here, the surface tension was computed as⁷³

$$\gamma = \frac{L_z}{2} \left(p_{zz} - \frac{p_{xx} + p_{yy}}{2} \right), \quad (10)$$

where x and y define the directions parallel to the macroscopic interface, z defines the direction normal to the interface, L_z is the length of the unit cell in the z -direction, and p_{ij} is the pressure tensor. The results are slightly larger than the value of 68 ± 2 mN m⁻¹ calculated by a neural network potential⁷³ and experimental value of 72 mN m⁻¹⁷⁴ while it is within the range (50–90 mN m⁻¹) of previous MD results by FP_{nl} and classical force fields^{76,77}. Distributions of interfacial water dipole moments for both, 128 and 1024 molecular systems, are shown in Supplementary Fig. 5. They consistently

indicate that the orientation of interfacial water molecules is bimodal as reported in previous MD studies employing the classical SPC/E force field⁷⁶. The distributions are also consistent with the results of sum frequency generation (SFG) analyses⁷⁸.

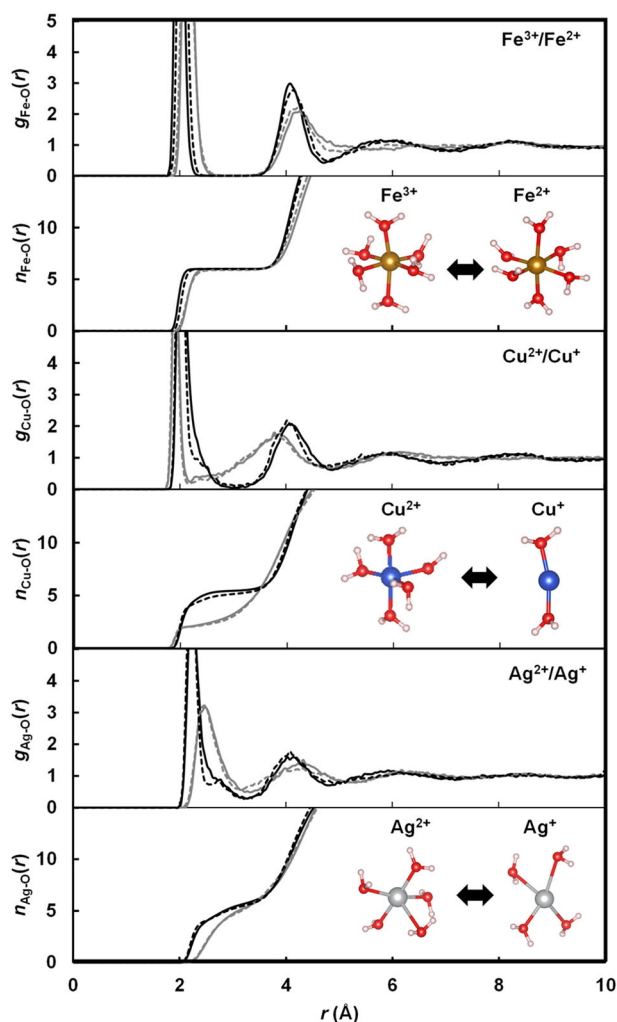


Fig. 3 | Metal-oxygen radial distribution functions (g_{X-O}) and running integration numbers (n_{X-O}) provided by 100 ps MLFF-MD and 10 ps FPMD simulations. Gray and black lines are for the reduced and oxidized states, respectively. Solid and dashed lines are results obtained by the MLFF and FP_{sl}, respectively. Graphics in the insets show first solvation structures at the reduced and oxidized states.

Metal water coordination

Figure 3 shows metal-oxygen radial distribution functions (RDFs) and running integration numbers (RINs) at the reduced and oxidized states calculated by the MLFF and FP_{sl} methods. The MLFFs well reproduce RDFs and RINs of the FP_{sl} method. Both methods show that the coordination number of Fe ions is 6 independent of the charge state. In contrast, the value for Cu changes from 5-6 in the oxidized state (Cu²⁺) to 2-3 in the reduced state (Cu⁺). The coordination number of Ag also changes from 5-6 in the oxidized state (Ag²⁺) to 4-5 in the reduced state (Ag⁺). These hydration structures agree with the ones reported in previous MD studies using FPMD methods⁷⁹⁻⁸¹ and empirical force fields⁷⁹. Although there are slight deviations in the Fe-O distance and shoulders for Cu-O and Ag-O in the RDFs likely related to the short FPMD simulation time and errors in the MLFFs, overall, our MLFFs reproduce the first-principles energies and structures of the hydrated metal cations with good accuracy.

Redox potentials

After verifying the size effect on the redox potentials obtained at the FP_{sl} level using unit cells containing 32, 64 and 96 water molecules (see U_{redox} in Supplementary Fig. 13), calculations were conducted on the bulk solutions containing 64 water molecules in the unit cell. The computed redox potentials are compared with the experimental ones in Fig. 4. All relevant data ($\langle U_1 - U_0 \rangle_\lambda$ and $\Delta\phi$), as well as results of other functionals with error bars, are summarized in Supplementary Notes 3, 4 and 5. The MLFFs trained on FP_{sl} (RPBE+D3) (see ML in Fig. 4) lead to non-negligible deviations of 30-250 mV from the values of full FP_{sl} calculations without any MLFF (see FP_{sl} w/o ML) depending on the training data size (see Supplementary Note 9). The deviations can be corrected by two TI integrations [ML(Ox) → FP_{sl}(Ox) and ML(Red) → FP_{sl}(Red) in Fig. 2] as shown by FP_{sl} w/ ML. However, the semi-local functional results in fairly large and non-systematic errors. For Ag, the redox potential is underestimated, whereas for Cu it is significantly overestimated compared to experiment.

The errors can be significantly decreased to 0.11 V on average using hybrid functionals with one-quarter exact exchange. As tabulated in Table 1, U_{redox} for the Cu²⁺/Cu⁺ couple decreases with increasing fractions of the exact exchange, whereas the redox potential for the Ag²⁺/Ag⁺ couple increases with increasing fractions. For Fe³⁺/Fe²⁺, the trend is not so obvious (first increase then slight decrease). Overall the present trends agree with the results obtained using semi-local and hybrid functionals as reported by Liu and co-workers¹¹. Finally, the effects of Grimme's dispersion correction are small for all redox couples. This implies that changes in the electronic properties (such as water valence band maximum and conduction band minimum) are most relevant, whereas all the functionals give a similar and good account of the solvation structure. It remains unclear, however, why one-quarter of exact exchange results in balanced accuracy. The functional form of PBE0 was rationalized by the adiabatic connection from the

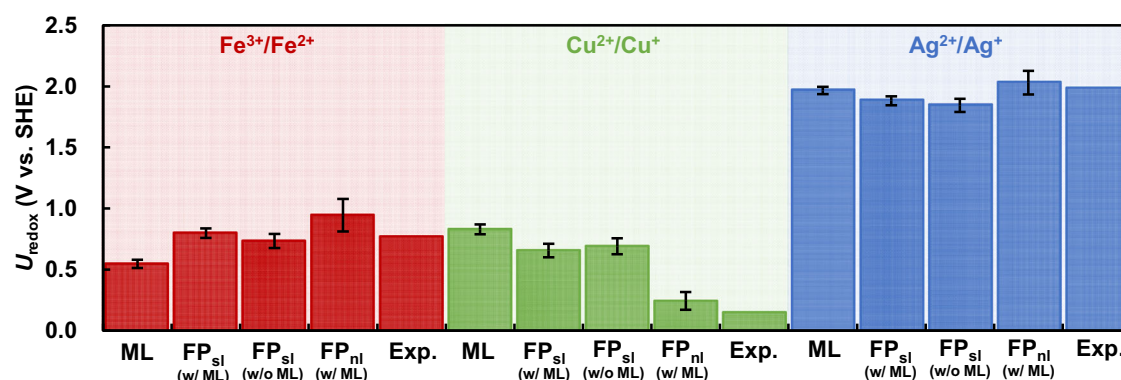


Fig. 4 | Computed and experimental redox potentials. ML means the results obtained by the MLFF trained on FP_{sl} (RPBE+D3) without any correction. The small letters 'w/ML' under FP_{sl} mean that the ML result was corrected by the scheme

shown in Fig. 2. The letters 'w/o ML' mean the result calculated by FP_{sl} using Eqs. (18) and (19) without MLFF. Here, FP_{nl} means PBE0+D3 (0.25) result obtained by the scheme shown in Fig. 2. Experimental values are taken from ref. 14.

uncorrelated exact exchange to the fully interacting energy, which is approximated by the PBE functional^{71,82}. Nonetheless, the ratio of exact exchange continues to be a parameter. One-quarter of exact exchange is known to achieve balanced accuracy for the geometries, thermochemistry, and spectroscopic properties of molecules. However, as reported in previous studies⁸³, this functional underestimates the band gap of liquid water, even though it provides a more accurate prediction than the HSE06 functional. The mechanism behind this remains an open question.

Another key observation in this study lies in the relationship between the error of the ML surrogate model and the error in the redox potential. Our MLFF models achieve an RMSE of a few meV atom⁻¹ for energy and tens of meV Å⁻¹ for force. These accuracies can be considered standard level compared to ML models generated in past research^{50,60,63,64,67,84–86}, yet they yield non-negligible deviations in the redox potential from the FP method. In comparison, Δ-ML models, which attained an RMSE substantially lower by more than an order of magnitude, markedly diminished the deviation in the redox potential to below 10 mV (refer to Supplementary Fig. 12). These results suggest that in aiming for an accuracy of 10 mV in reproducing the redox potential of the FP method, an RMSE at least an order of magnitude smaller than that shown by standard MLFFs is required. Achieving this level of accuracy is highly challenging for MLFFs, even if they are trained on larger training datasets, as demonstrated in the previous study on liquid water⁸⁷. While the accuracy of emerging MLFFs continues to improve^{88,89}, there is always a risk that machine learning models may produce errors concerning the structure of extrapolation regions outside the training data. Even in the future where machine learning models have further advanced, our ML correction schemes will serve as a powerful method for quantifying errors and providing results from accurate FP calculations.

In summary, our approach enables efficient statistical sampling that is indispensable for accurate computations of the free energies of aqueous systems. The TI and TPT calculations allow us to improve the accuracy from the ML model to the semi-local functional and from the semi-local functional to the hybrid functional step-by-step. Combining TPT and Δ-machine learning is particularly promising since this allows us to obtain statistically highly accurate results even for expensive functionals in very little compute time. For instance, it is well conceivable that one could also use methods beyond density functional theory for the final step. Our final results reproduce the redox potentials of the three transition metal cations with excellent accuracy using a standard hybrid functional. The integration pathways chosen here are generalizable to a wide variety of electron transfer reactions. We believe that the scheme will pave the way to first-principles electrochemistry predicting the key properties of redox reactions in energy conversion devices.

Methods

TI and TPT

The TI and TPT shown in Fig. 2 in the main text are conducted by using the equations listed below.

- ML(Ox) → ML(Red)

$$\Delta A^{\text{ML}} = \int_0^1 \left\langle \frac{\partial H^{\text{ML}}}{\partial \lambda} \right\rangle_{\lambda} d\lambda, \quad (11)$$

$$H^{\text{ML}} = \sum_{i=1}^{N_a} \frac{|\mathbf{p}_i|^2}{2m_i} + \lambda U_1^{\text{ML}} + (1 - \lambda)U_0^{\text{ML}} - Ne\Delta\bar{\phi}. \quad (12)$$

- ML(Ox) → FP_{sl}(Ox) and ML(Red) → FP_{sl}(Red)

$$\Delta A_{\kappa}^{\text{FP}_{\text{sl}}-\text{ML}} = \int_0^1 \left\langle \frac{\partial H_{\kappa}^{\text{FP}_{\text{sl}}-\text{ML}}}{\partial \eta} \right\rangle_{\eta} d\eta, \quad (13)$$

$$H_{\kappa}^{\text{FP}_{\text{sl}}-\text{ML}} = \sum_{i=1}^{N_a} \frac{|\mathbf{p}_i|^2}{2m_i} + \eta U_{\kappa}^{\text{FP}_{\text{sl}}} + (1 - \eta)U_{\kappa}^{\text{ML}}. \quad (14)$$

- FP_{sl}(Ox) → FP_{nl}(Ox) and FP_{sl}(Red) → FP_{nl}(Red)

$$\Delta A_{\kappa}^{\text{FP}_{\text{nl}}-\text{FP}_{\text{sl}}} \simeq \langle \Delta U_{\kappa}^{\Delta\text{ML}} \rangle_{\text{FP}_{\text{sl}}} - \frac{\beta}{2} \left\langle \left(\Delta U_{\kappa}^{\Delta\text{ML}} - \langle \Delta U_{\kappa}^{\Delta\text{ML}} \rangle_{\text{FP}_{\text{sl}}} \right)^2 \right\rangle_{\text{FP}_{\text{sl}}}. \quad (15)$$

The symbols $U_{\kappa}^{\text{FP}_{\text{nl}}}$, $U_{\kappa}^{\text{FP}_{\text{sl}}}$ and U_{κ}^{ML} are the potential energies for the oxidized ($\kappa = 0$) and reduced ($\kappa = 1$) states calculated by the non-local functional, semi-local functional and MLFF trained on the semi-local functional, respectively. The symbol $\Delta U_{\kappa}^{\Delta\text{ML}}$ denotes the potential energy difference calculated by the Δ-ML model trained on the potential energy difference $U_{\kappa}^{\text{FP}_{\text{nl}}} - U_{\kappa}^{\text{FP}_{\text{sl}}}$ between the non-local and semi-local functionals. In Eq. (15), the second-order cumulant expansion is employed. The expansion is exact if the probability distribution of $\Delta U_{\kappa}^{\Delta\text{ML}}$ is Gaussian (see the derivation in Supplementary Note 8). The condition is reasonably satisfied as shown in Supplementary Fig. 11. Preliminary TI and TPT simulations using the MLFFs also indicate that the TPT calculation reproduces TI results as shown in Supplementary Note 6.

The free energy differences of the FP_{sl} and FP_{nl} methods are obtained as

$$\Delta A^{\text{FP}_{\text{sl}}} = \Delta A^{\text{ML}} + \Delta A_1^{\text{FP}_{\text{sl}}-\text{ML}} - \Delta A_0^{\text{FP}_{\text{sl}}-\text{ML}}, \quad (16)$$

$$\Delta A_{\text{nl}}^{\text{FP}} = \Delta A^{\text{FP}_{\text{sl}}} + \Delta A_1^{\text{FP}_{\text{nl}}-\text{FP}_{\text{sl}}} - \Delta A_0^{\text{FP}_{\text{nl}}-\text{FP}_{\text{sl}}}. \quad (17)$$

To validate the MLFF-aided computations of the free energy difference $\Delta A^{\text{FP}_{\text{sl}}}$, the same property was also computed by the TI without using the ML method:

$$\Delta A^{\text{FP}_{\text{sl}}} = \int_0^1 \left\langle \frac{\partial H^{\text{FP}_{\text{sl}}}}{\partial \lambda} \right\rangle_{\lambda} d\lambda, \quad (18)$$

$$H^{\text{FP}_{\text{sl}}} = \sum_{i=1}^{N_a} \frac{|\mathbf{p}_i|^2}{2m_i} + \lambda U_1^{\text{FP}_{\text{sl}}} + (1 - \lambda)U_0^{\text{FP}_{\text{sl}}} - Ne\Delta\bar{\phi}. \quad (19)$$

The TI calculation in Eq. (2) can be decomposed into the two terms on the right-hand side of Eq. (7). The integrand in the first term nonlinearly varies along the coupling path (see Supplementary Figs. 8 and 9) while the integrand in Eq. (8), which is relevant to the second term in Eq. (7), varies only slightly (see Supplementary Table 4). To perform the integration of the first term in Eq. (7), Simpson's rule with equidistant five points was used following the previous FP study by Blumberger and co-workers⁴¹. For the integration in Eq. (8), the average of the O 1s levels in the fully reduced and oxidized states was used based on the trapezoidal rule. For each point, the ensemble average was taken over an 80-ps-NVT-ensemble MD simulation at 298 K after a 20 ps equilibration. Similar to the MLFF calculations, Simpson's rule with equidistant five points was used for the TI calculation in Eq. (18). For each grid, the ensemble average was taken over a 20-ps-MD simulation starting from the final structure of the TI calculation using the MLFF at the same grid point. Each initial structure of the MD simulations was prepared by annealing the system from 400 K to 298 K by a 100-ps-NVT-ensemble MD simulation using the MLFF after annealing the same system from 1000 K to 400 K by a 1-ns-NVT ensemble MD simulation using the polymer consistent force field (PCFF)⁹⁰ implemented in a homemade MD program⁹¹. Supplementary Figs. 8 and 9 show the integrands of Eqs. (11) and (18), respectively, as functions of the coupling parameter λ . In the same figures, probability distributions of $\Delta U^{\text{ML}} = U_1^{\text{ML}} - U_0^{\text{ML}}$ and $\Delta U^{\text{FP}_{\text{sl}}} = U_1^{\text{FP}_{\text{sl}}} - U_0^{\text{FP}_{\text{sl}}}$ at each λ are also shown. For all redox couples, the variance of the distribution varies with changing λ , and thus, the integrand is non-linear with respect to λ [see Supplementary

Eq. (14)]. Hence, the second cumulant expansion Supplementary Eq. (12) is not applicable to the whole integration from the oxidized state to the reduced state.

The TI calculations in Eq. (13) were conducted using the trapezoidal rule with three equidistant points. At each point, a 10-ps-NVT-ensemble MD simulation at 298 K was performed. The integrands shown in Supplementary Fig. 10 are smaller than the ones shown in Supplementary Figs. 8 and 9, respectively. They are also nearly proportional to the coupling parameter η .

In the TPT calculations using the Δ -ML models, the ensemble average in Eq. (15) was taken over 1400 configurations selected randomly from 70-ps-NVT-ensemble FPMD simulations using the FP_{sl} method. Although these FPMD simulations are expensive, the overall computational time is still much smaller than full FP simulations. To ensure the applications of the second-order cumulant expansion, we show the probability distributions of the energy difference $\Delta U_{\kappa}^{\Delta ML}$ in Supplementary Fig. 11. The distribution is well-fitted by Gaussian functions, indicating that Eq. (15) is a reasonable approximation.

The MD simulations were performed using a Langevin thermostat⁹². For efficient sampling, the mass of hydrogen and time step were set to 2 amu and 1 fs.

MLFF and Δ -ML

Similar to previous machine-learning approaches^{63,64}, the potential energy U of a structure with N_a atoms in our MLFF method is approximated as a summation of local energies U_i written as

$$U = \sum_{i=1}^{N_a} U_i. \quad (20)$$

Following the Gaussian approximation potential pioneered by Bártok and co-workers⁶⁴, the local energy U_i is approximated as a weighted sum of functions $K(\mathbf{x}_i, \mathbf{x}_{i_b})$ centered at reference points $\{\mathbf{x}_{i_b} | i_b = 1, \dots, N_B\}$

$$U_i = \sum_{i_b=1}^{N_B} w_{i_b} K(\mathbf{x}_i, \mathbf{x}_{i_b}). \quad (21)$$

The coefficients $\{w_{i_b} | i_b = 1, \dots, N_B\}$ are optimized to best reproduce the FP energies, forces, and stress tensor components as obtained by the FPMD simulations. The descriptor \mathbf{x}_i used in this study is a vector containing two and three body contributions⁶⁷:

$$\mathbf{x}_i^T \rightarrow \left(\sqrt{\beta^{(2)}} \mathbf{x}_i^{(2)T}, \sqrt{\beta^{(3)}} \mathbf{x}_i^{(3)T} \right), \quad (22)$$

Here, $\beta^{(2)}$ and $\beta^{(3)} (= 1 - \beta^{(2)})$ are the weights on the two and three body descriptors, $\mathbf{x}_i^{(2)}$ and $\mathbf{x}_i^{(3)}$, respectively. The vectors $\mathbf{x}_i^{(2)}$ and $\mathbf{x}_i^{(3)}$ collect the expansion coefficients of two and three body distribution functions with respect to the orthonormal radial and angular basis sets^{60,67}:

$$\rho_i^{(2)}(r) = \frac{1}{\sqrt{4\pi}} \sum_{n=1}^{N_R^0} c_n^i \chi_{n0}(r) \quad (23)$$

$$\rho_i^{(3)}(r, s, \theta) = \sum_{l=0}^{L_{\max}} \sum_{n=1}^{N_R^l} \sum_{\nu=1}^{N_R^l} \sqrt{\frac{2l+1}{2}} P_{\nu l}^j \chi_{nl}(r) \chi_{\nu l}(s) P_l(\cos \theta). \quad (24)$$

The two and three body distribution functions $\rho_i^{(2)}$ and $\rho_i^{(3)}$ are defined as:

$$\rho_i^{(2)}(r) = \frac{1}{4\pi} \int \rho_i(r\hat{\mathbf{r}}) d\hat{\mathbf{r}}, \quad (25)$$

$$\rho_i^{(3)}(r, s, \theta) = \iint d\hat{\mathbf{r}} d\hat{\mathbf{s}} \delta(\hat{\mathbf{r}} \cdot \hat{\mathbf{s}} - \cos \theta) \rho_i(r\hat{\mathbf{r}}) \rho_i(s\hat{\mathbf{s}}), \quad (26)$$

$$\rho_i(\mathbf{r}) = \sum_{j=1}^{N_i} f_{\text{cut}}(|\mathbf{r}_j - \mathbf{r}_i|) g(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)) \quad (27)$$

The function g is the smoothed δ -function, and f_{cut} is a cutoff function that smoothly eliminates the contribution from atoms outside a given cutoff radius R_{cut} . For χ_{nl} and P_l normalized spherical Bessel functions $\chi_{nl} = j_l(q_n r)$ and Legendre polynomials of order l are used in this work, respectively. For the kernel basis functions, the smooth overlap of atomic positions (SOAP) kernel⁵⁰ is employed

$$K(\mathbf{x}_i, \mathbf{x}_{i_b}) = (\hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_{i_b})^\zeta. \quad (28)$$

The hat symbol $\hat{\mathbf{x}}_i$ denotes a normalized vector of \mathbf{x}_i . The normalization and exponentiation in Eq. (28) introduce non-linear terms that mix two- and three-body contributions.

The same formulation is used for the Δ -ML method. In the Δ -ML method, differences in potential energies and forces between two FP methods, semi-local and non-local functionals in this study, are used as the training data.

Parameter sets of the descriptors and kernel basis functions used in previous publications were employed in this study^{60,62,67}. The parameters are tabulated in Supplementary Table 1.

Bulk solutions containing the redox species were modeled by systems as shown in Fig. 1. The number of water molecules was set to 32, 64, and 96. Three different model sizes were examined to clarify the system size effect. The sizes of the unit cells were set to obtain a water density of 0.99 g cm⁻³. The size of the unit cell for the 32 water molecules is the same as the one used in previous FPMD studies^{10,11,41,80}. For each of the reduced and oxidized states, MLFF and Δ -ML models were constructed. All MLFF models were generated on the fly during a 100-ps-NVT-MD simulation at 400 K by using the active-learning algorithm developed in our previous study⁶⁰. The temperature for the training runs was set to a value higher than the target one of 298 K for production runs, to ensure that training data and kernel basis functions were provided in a wider phase space. A Langevin thermostat⁹² was used to control the temperature. Exchange-correlation interactions between electrons were described by the semi-local RPBE functional⁶⁸ with Grimme's D3 dispersion corrections^{69,70}. Probability distributions of the errors of the constructed MLFFs for energies and forces on test data are shown in Supplementary Figs. 1 to 4. The RMSEs are similar to those of MLFFs used in previous studies^{60,65-67}.

After examining the system size effect using the semi-local functional (see results in Supplementary Fig. 13), Δ -ML models were constructed on systems containing 64 water molecules. Each Δ -ML model was trained on FP energies and forces of 40 structures selected randomly from a trajectory of a 20 ps NVT-ensemble FPMD simulation at 298K. The FPMD simulation was performed using the RPBE+D3 functional. Differences in energies and forces between the non-local and semi-local functionals for these 40 structures were used as training data. PBE0⁷² with and without Grimme's D3 dispersion correction^{69,70} was employed as the non-local functional because the functional is known to accurately predict properties of water⁹³⁻¹⁰¹. The fraction of the exact exchange was set to 0.25 and 0.50 to determine how this influences the redox potentials. Error distributions of the Δ -ML models on test structures are shown in Supplementary Figs. 2 to 4. The RMSEs are one to two orders of magnitude smaller than the errors of the RPBE+D3 MLFFs.

The vacuum-water interface for the production run was modeled by a pure water slab without the redox species composed of 128 water molecules per unit cell. Following the previous study⁴⁵, a rectangular cell of 12.5 × 12.5 × 50 Å³ was employed. Similar to the MLFFs for the bulk

solution systems containing the redox species, the MLFF for the interface was also generated by using the active-learning scheme. The systems used for the training were a pure water bulk composed of 64 water molecules in a $12.4 \times 12.4 \times 12.4 \text{ \AA}^3$ cubic cell and a pure water slab composed of 64 water molecules in an $8.8 \times 8.8 \times 40.8 \text{ \AA}^3$ rectangular cell. Training simulations for both the bulk and slab were performed by NVT-ensemble MD simulations at 300, 400, 600 and 800 K. As shown in Supplementary Fig. 1, the constructed MLFF realizes small errors on test data taken from a 100-ps-MD simulation of a water slab composed of 128 water molecules at 298 K.

The annealing procedure used for the production runs explained in the previous subsection was also used to prepare for the initial structures for the training runs. All FP calculations were performed using VASP^{56,57}. A $2 \times 2 \times 2$ \mathbf{k} -point mesh was used for the bulk systems containing 32 water molecules. For other systems, Γ -point was used. Plane-wave cutoff energy was set to 520 eV. The PAW^{58,59} distributed in VASP 5.4 was used in all FP calculations. The PAW atomic reference configuration was $1s^1$ for H, $2s^2 2p^4$ for O, $3d^7 4s^1$ for Fe, and $4d^{10} 5s^1$ for Ag. The comparison of two atomic configurations for Cu, specifically $3d^{10} 4p^1$ and $3p^6 3d^{10} 4p^1$, was conducted to examine the impact of semi-core electron relaxations on the redox potential. Upon verification that these effects are minimal within the PAW framework in VASP, as detailed in Supplementary Note 7, we employed the less computationally demanding $3d^{10} 4p^1$ electronic configuration. The parameters for the MD simulations are the same as the ones described in the previous subsection.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The VASP code is distributed by the VASP Software GmbH. The machine learning modules will be included in the release of vasp.6.3. Prerelease versions are available from G.K. upon reasonable request.

Received: 26 October 2023; Accepted: 11 May 2024;

Published online: 20 May 2024

References

- Weber, A. Z. et al. Redox flow batteries: A review. *J. Appl. Electrochem.* **41**, 1137–1164 (2011).
- Ong, S. P., Andreussi, O., Wu, Y., Marzari, N. & Ceder, G. Electrochemical windows of room-temperature ionic liquids from molecular dynamics and density functional theory calculations. *Chem. Mater.* **23**, 2979–2986 (2011).
- Xu, K. Electrolytes and interphases in li-ion batteries and beyond. *Chem. Rev.* **114**, 11503–11618 (2014).
- Haregewoin, A. M., Wotango, A. S. & Hwang, B.-J. Electrolyte additives for lithium ion battery electrodes: progress and perspectives. *Energy Environ. Sci.* **9**, 1955–1988 (2016).
- Zatoń, M., Roziere, J. & Jones, D. J. Current understanding of chemical degradation mechanisms of perfluorosulfonic acid membranes and their mitigation strategies: a review. *Sustain. Energy Fuels* **1**, 409–438 (2017).
- Pinaud, B. A. et al. Technical and economic feasibility of centralized facilities for solar hydrogen production via photocatalysis and photoelectrochemistry. *Energy Environ. Sci.* **6**, 1983–2002 (2013).
- Morikawa, T., Sato, S., Sekizawa, K., Suzuki, T. M. & Arai, T. Solar-driven CO₂ reduction using a semiconductor/molecule hybrid photosystem: From photocatalysts to a monolithic artificial leaf. *Acc. Chem. Res.* **55**, 933–943 (2022).
- Costanzo, F., Sulpizi, M., Valle, R. G. D. & Sprik, M. The oxidation of tyrosine and tryptophan studied by a molecular dynamics normal hydrogen electrode. *J. Chem. Phys.* **134**, 244508 (2011).
- Le, J., Iannuzzi, M., Cuesta, A. & Cheng, J. Determining potentials of zero charge of metal electrodes versus the standard hydrogen electrode from density-functional-theory-based molecular dynamics. *Phys. Rev. Lett.* **119**, 016801 (2017).
- Adriaanse, C. et al. Aqueous redox chemistry and the electronic band structure of liquid water. *J. Phys. Chem. Lett.* **3**, 3411–3415 (2012).
- Liu, X., Cheng, J. & Sprik, M. Aqueous transition-metal cations as impurities in a wide gap oxide: The Cu²⁺/Cu⁺ and Ag²⁺/Ag⁺ redox couples revisited. *J. Phys. Chem. B* **119**, 1152–1163 (2015).
- Caro, M. A., Lopez-Acevedo, O. & Laurila, T. Redox potentials from ab initio molecular dynamics and explicit entropy calculations: Application to transition metals in aqueous solution. *J. Chem. Theory Comput.* **13**, 3432–3441 (2017).
- Bouid, A. & Pasquarello, A. Redox levels through constant fermi-level ab initio molecular dynamics. *J. Chem. Theory Comput.* **13**, 1769–1777 (2017).
- Bard, A., Parsons, R. & Jordan, J. *Standard Potentials in Aqueous Solution*. Monographs in Electroanalytical Chemistry and Electrochemistry (Taylor & Francis, 1985).
- Baik, M.-H. & Friesner, R. A. Computing redox potentials in solution: density functional theory as a tool for rational design of redox agents. *J. Phys. Chem. A* **106**, 7407–7412 (2002).
- Jaque, P., Marenich, A. V., Cramer, C. J. & Truhlar, D. G. Computational electrochemistry: The aqueous Ru³⁺/Ru²⁺ reduction potential. *J. Phys. Chem. C* **111**, 5783–5799 (2007).
- Jinnouchi, R. & Anderson, A. B. Aqueous and surface redox potentials from self-consistently determined gibbs energies. *J. Phys. Chem. C* **112**, 8747–8750 (2008).
- Neugebauer, H., Bohle, F., Bursch, M., Hansen, A. & Grimme, S. Benchmark study of electrochemical redox potentials calculated with semiempirical and DFT methods. *J. Phys. Chem. A* **124**, 7166–7176 (2020).
- Vaissier, V. & Van Voorhis, T. Adiabatic approximation in explicit solvent models of RedOx chemistry. *J. Chem. Theory Comput.* **12**, 5111–5116 (2016).
- Nicholson, M. I. G., Bueno, P. R. & Feliciano, G. T. Ab initio QM/MM simulation of ferrocene homogeneous electron-transfer reaction. *J. Phys. Chem. A* **125**, 25–33 (2021).
- Taylor, C. D., Wasileski, S. A., Filhol, J.-S. & Neurock, M. First principles reaction modeling of the electrochemical interface: Consideration and calculation of a tunable surface potential from atomic and electronic structure. *Phys. Rev. B* **73**, 165402 (2006).
- Skúlason, E. et al. Density functional theory calculations for the hydrogen evolution reaction in an electrochemical double layer on the pt(111) electrode. *Phys. Chem. Chem. Phys.* **9**, 3241–3250 (2007).
- Jinnouchi, R. & Anderson, A. B. Electronic structure calculations of liquid-solid interfaces: Combination of density functional theory and modified poisson-boltzmann theory. *Phys. Rev. B* **77**, 245417 (2008).
- Letchworth-Weaver, K. & Arias, T. A. Joint density functional theory of the electrode-electrolyte interface: Application to fixed electrode potentials, interfacial capacitances, and potentials of zero charge. *Phys. Rev. B* **86**, 075140 (2012).
- Mathew, K., Sundararaman, R., Letchworth-Weaver, K., Arias, T. A. & Hennig, R. G. Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways. *J. Chem. Phys.* **140**, 084106 (2014).
- Nishihara, S. & Otani, M. Hybrid solvation models for bulk, interface, and membrane: Reference interaction site methods coupled with density functional theory. *Phys. Rev. B* **96**, 115429 (2017).
- Hörmann, N. G. et al. Absolute band alignment at semiconductor-water interfaces using explicit and implicit descriptions for liquid water. *Npj Comput. Mater.* **5**, 100 (2019).
- Groß, A. & Sakong, S. Modelling the electric double layer at electrode/electrolyte interfaces. *Curr. Opin. Electrochem.* **14**, 1–6 (2019).
- Islam, S. M. R., Khezeli, F., Ringe, S. & Plaisance, C. An implicit electrolyte model for plane wave density functional theory exhibiting

- nonlinear response and a nonlocal cavity definition. *J. Chem. Phys.* **159**, 234117 (2023).
30. Nørskov, J. K. et al. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *J. Phys. Chem. B* **108**, 17886–17892 (2004).
31. Nørskov, J. K. et al. Trends in the exchange current for hydrogen evolution. *J. Electrochem. Soc.* **152**, J23 (2005).
32. Nørskov, J. K., Bligaard, T., Rossmeisl, J. & Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* **1**, 37–46 (2009).
33. Jinnouchi, R., Kodama, K., Hatanaka, T. & Morimoto, Y. First principles based mean field model for oxygen reduction reaction. *Phys. Chem. Chem. Phys.* **13**, 21070–21083 (2011).
34. Kulkarni, A., Siahrostami, S., Patel, A. & Nørskov, J. K. Understanding catalytic activity trends in the oxygen reduction reaction. *Chem. Rev.* **118**, 2302–2312 (2018).
35. Nitopi, S. et al. Progress and perspectives of electrochemical CO₂ reduction on copper in aqueous electrolyte. *Chem. Rev.* **119**, 7610–7672 (2019).
36. Man, I. C. et al. Universality in oxygen evolution electrocatalysis on oxide surfaces. *ChemCatChem* **3**, 1159–1165 (2011).
37. Kortlever, R., Shen, J., Schouten, K. J. P., Calle-Vallejo, F. & Koper, M. T. M. Catalysts and reaction pathways for the electrochemical reduction of carbon dioxide. *J. Phys. Chem. Lett.* **6**, 4073–4082 (2015).
38. Marcus, R. A. Electron transfer reactions in chemistry. theory and experiment. *Rev. Mod. Phys.* **65**, 599–610 (1993).
39. Zwanzig, R. W. High-temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.* **22**, 1420–1426 (1954).
40. Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **3**, 300–313 (1935).
41. Blumberger, J., Tavernelli, I., Klein, M. L. & Sprik, M. Diabatic free energy curves and coordination fluctuations for the aqueous Ag⁺/Ag²⁺ redox couple: A biased born-oppenheimer molecular dynamics investigation. *J. Chem. Phys.* **124**, 064507 (2006).
42. Dörner, F., Sukurma, Z., Dellago, C. & Kresse, G. Melting Si: Beyond density functional theory. *Phys. Rev. Lett.* **121**, 195701 (2018).
43. Mermin, N. D. Thermal properties of the inhomogeneous electron gas. *Phys. Rev.* **137**, A1441–A1443 (1965).
44. Jiao, D., Leung, K., Rempe, S. B. & Nenoff, T. M. First principles calculations of atomic nickel redox potentials and dimerization free energies: A study of metal nanoparticle growth. *J. Chem. Theory Comput.* **7**, 485–495 (2011).
45. Leung, K. Surface potential at the air-water interface computed using density functional theory. *J. Phys. Chem. Lett.* **1**, 496–499 (2010).
46. Lany, S. & Zunger, A. Assessment of correction methods for the band-gap problem and for finite-size effects in supercell defect calculations: Case studies for ZnO and GaAs. *Phys. Rev. B* **78**, 235104 (2008).
47. Freysoldt, C. et al. First-principles calculations for point defects in solids. *Rev. Mod. Phys.* **86**, 253–305 (2014).
48. Balabin, R. M. & Lomakina, E. I. Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies. *J. Chem. Phys.* **131**, 074104 (2009).
49. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Big data meets quantum chemistry approximations: The Δ -machine learning approach. *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
50. Bartók, A. P., Kondor, R. & Csányi, G. On representing chemical environments. *Phys. Rev. B* **87**, 184115 (2013).
51. Chmiela, S., Sauceda, H. E., Müller, K.-R. & Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **9**, 3887 (2018).
52. Sauceda, H. E., Chmiela, S., Poltavsky, I., Müller, K.-R. & Tkatchenko, A. Molecular force fields with gradient-domain machine learning: Construction and application to dynamics of small molecules with coupled cluster forces. *J. Chem. Phys.* **150**, 114102 (2019).
53. Liu, P., Verdi, C., Karsai, F. & Kresse, G. Phase transitions of zirconia: Machine-learned force fields beyond density functional theory. *Phys. Rev. B* **105**, L060102 (2022).
54. Verdi, C., Ranalli, L., Franchini, C. & Kresse, G. Quantum paraelectricity and structural phase transitions in strontium titanate beyond density functional theory. *Phys. Rev. Mater.* **7**, L030801 (2023).
55. Liu, P. et al. Combining machine learning and many-body calculations: Coverage-dependent adsorption of CO on Rh(111). *Phys. Rev. Lett.* **130**, 078001 (2023).
56. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
57. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15 (1996).
58. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
59. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
60. Jinnouchi, R., Karsai, F. & Kresse, G. On-the-fly machine learning force field generation: Application to melting points. *Phys. Rev. B* **100**, 014105 (2019).
61. Jinnouchi, R., Miwa, K., Karsai, F., Kresse, G. & Asahi, R. On-the-fly active learning of interatomic potentials for large-scale atomistic simulations. *J. Phys. Chem. Lett.* **11**, 6946–6955 (2020).
62. Jinnouchi, R., Minami, S., Karsai, F., Verdi, C. & Kresse, G. Proton transport in perfluorinated ionomer simulated by machine-learned interatomic potential. *J. Phys. Chem. Lett.* **14**, 3581–3588 (2023).
63. Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **98**, 146401 (2007).
64. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
65. Jinnouchi, R., Lahnsteiner, J., Karsai, F., Kresse, G. & Bokdam, M. Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on the fly with bayesian inference. *Phys. Rev. Lett.* **122**, 225701 (2019).
66. Jinnouchi, R., Karsai, F. & Kresse, G. Making free-energy calculations routine: Combining first principles with machine learning. *Phys. Rev. B* **101**, 060201 (2020).
67. Jinnouchi, R., Karsai, F., Verdi, C., Asahi, R. & Kresse, G. Descriptors representing two- and three-body atomic distributions and their effects on the accuracy of machine-learned inter-atomic potentials. *J. Chem. Phys.* **152**, 234102 (2020).
68. Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7421 (1999).
69. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT - D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
70. Grimme, S. Density functional theory with london dispersion corrections. *WIREs Comput. Mol. Sci.* **1**, 211–228 (2011).
71. Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982–9985 (1996).
72. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **110**, 6158–6170 (1999).
73. Wohlfahrt, O., Dellago, C. & Sega, M. Ab initio structure and thermodynamics of the RPBE - D3 water/vapor interface by neural-network molecular dynamics. *J. Chem. Phys.* **153**, 144710 (2020).

74. Vargaftik, N. B., Volkov, B. N. & Voljak, L. D. International tables of the surface tension of water. *J. Phys. Chem. Ref. Data* **12**, 817–820 (1983).
75. Ohto, T., Dodia, M., Imoto, S. & Nagata, Y. Structure and dynamics of water at the water–air interface using first-principles molecular dynamics simulations within generalized gradient approximation. *J. Chem. Theory Comput.* **15**, 595–602 (2019).
76. Taylor, R. S., Dang, L. X. & Garrett, B. C. Molecular dynamics simulations of the liquid/vapor interface of SPC/E water. *J. Phys. Chem.* **100**, 11720–11725 (1996).
77. Vega, C. & de Miguel, E. Surface tension of the most popular models of water by using the test-area simulation method. *J. Chem. Phys.* **126** (2007). 154707.
78. Du, Q., Superfine, R., Freysz, E. & Shen, Y. R. Vibrational spectroscopy of water at the vapor/water interface. *Phys. Rev. Lett.* **70**, 2313–2316 (1993).
79. Remsungnen, T. & Rode, B. M. Molecular dynamics simulation of the hydration of transition metal ions: the role of non-additive effects in the hydration shells of Fe²⁺ and Fe³⁺ ions. *Chem. Phys. Lett.* **385**, 491–497 (2004).
80. Blumberger, J., Bernasconi, L., Tavernelli, I., Vuilleumier, R. & Sprik, M. Electronic structure and solvation of copper and silver ions: a theoretical picture of a model aqueous redox reaction. *J. Am. Chem. Soc.* **126**, 3928–3938 (2004).
81. Bogatko, S. A., Bylaska, E. J. & Weare, J. H. First principles simulation of the bonding, vibrational, and electronic properties of the hydration shells of the high-spin Fe³⁺ ion in aqueous solutions. *J. Phys. Chem. A* **114**, 2189–2200 (2010).
82. Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993).
83. Pham, T. A., Zhang, C., Schwegler, E. & Galli, G. Probing the electronic structure of liquid water with many-body perturbation theory. *Phys. Rev. B* **89**, 060202 (2014).
84. Shapeev, A. V. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **14**, 1153–1173 (2016).
85. Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B* **99**, 014104 (2019).
86. Lysogorskiy, Y. et al. Performant implementation of the atomic cluster expansion (pace) and application to copper and silicon. *Npj Comput. Mater.* **7**, 97 (2021).
87. Montero de Hijes, P., Dellago, C., Jinnouchi, R., Schmiedmayer, B. & Kresse, G. Comparing machine learning potentials for water: Kernel-based regression and Behler-Parrinello neural networks. *J. Chem. Phys.* **160**, 114107 (2024).
88. Batzner, S. et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **13**, 2453 (2022).
89. Batatia, I., Kovacs, D. P., Simm, G., Ortner, C. & Csanyi, G. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Adv. Neural Inf. Process. Syst.* **35**, 11423–11436 (2022).
90. Sun, H. Force field for computation of conformational energies, structures, and vibrational frequencies of aromatic polyesters. *J. Comput. Chem.* **15**, 752–768 (1994).
91. Jinnouchi, R., Kudo, K., Kitano, N. & Morimoto, Y. Molecular dynamics simulations on O₂ permeation through nafion ionomer on platinum surface. *Electrochim. Acta* **188**, 767–776 (2016).
92. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids* (Oxford University Press, 2017).
93. Gaiduk, A. P., Gygi, F. & Galli, G. Density and compressibility of liquid water and ice from first-principles simulations with hybrid functionals. *J. Phys. Chem. Lett.* **6**, 2902–2908 (2015).
94. Warburton, R. E., Soudackov, A. V. & Hammes-Schiffer, S. Theoretical modeling of electrochemical proton-coupled electron transfer. *Chem. Rev.* **122**, 10599–10650 (2022).
95. Cheng, J. & Sprik, M. Aligning electronic energy levels at the TiO₂/H₂O interface. *Phys. Rev. B* **82**, 081406 (2010).
96. Cheng, J. & Sprik, M. Alignment of electronic energy levels at electrochemical interfaces. *Phys. Chem. Chem. Phys.* **14**, 11245–11267 (2012).
97. Hofer, T. S. & Hünenberger, P. H. Absolute proton hydration free energy, surface potential of water, and redox potential of the hydrogen electrode from first principles: QM/MM MD free-energy simulations of sodium and potassium hydration. *J. Chem. Phys.* **148**, 222814 (2018).
98. Prasetyo, N., Hünenberger, P. H. & Hofer, T. S. Single-ion thermodynamics from first principles: Calculation of the absolute hydration free energy and single-electrode potential of aqueous Li⁺ using ab initio quantum mechanical/molecular mechanical molecular dynamics simulations. *J. Chem. Theory Comput.* **14**, 6443–6459 (2018).
99. Andreussi, O., Dabo, I. & Marzari, N. Revised self-consistent continuum solvation in electronic-structure calculations. *J. Chem. Phys.* **136**, 064102 (2012).
100. Sakong, S., Naderian, M., Mathew, K., Hennig, R. G. & Groß, A. Density functional theory study of the electrochemical interface between a Pt electrode and an aqueous electrolyte using an implicit solvent method. *J. Chem. Phys.* **142**, 234107 (2015).
101. Tavernelli, I., Vuilleumier, R. & Sprik, M. Ab initio molecular dynamics for molecules with variable numbers of electrons. *Phys. Rev. Lett.* **88**, 213002 (2002).
102. Trasatti, S. The absolute electrode potential: an explanatory note (recommendations 1986). *Pure Appl. Chem.* **58**, 955–966 (1986).
103. Kühne, T. D. et al. CP2k: An electronic structure and molecular dynamics software package—quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **152**, 194103 (2020).
104. Goedecker, S., Teter, M. & Hutter, J. Separable dual-space gaussian pseudopotentials. *Phys. Rev. B* **54**, 1703–1710 (1996).
105. Momma, K. & Izumi, F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Crystallogr.* **44**, 1272–1276 (2011).

Acknowledgements

We thank Dr. Carla Verdi for helpful discussions and Dr. Jiabo Le for useful information on the computational standard hydrogen electrode. Finally, we are grateful to the referees for valuable comments.

Author contributions

R.J. developed the ML-aided TI and TPT simulation method. F.K. and G.K. developed the MLFF and VASP code. G.K. suggested the alignment method using the oxygen 1s levels. All authors participated in preparing and editing the manuscript.

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/s41524-024-01295-6>.

Correspondence and requests for materials should be addressed to Ryosuke Jinnouchi.

Reprints and permissions 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024