Exploring deep-learning and equivariant machine-learning models for molecular dynamics of aqueous Li–Mn–Cl battery electrolytes

Samuel D. Young*

Battery Science Branch, US Army DEVCOM Army Research Laboratory, 2800 Powder Mill Dr., Adelphi, MD 20783, USA

E-mail: samuel.d.young29.ctr@army.mil

This annual report summarizes the work I have done so far during my tenure at ARL as a Distinguished Postdoctoral Fellow under Contract no. W911NF-19-2-0186 in the Battery Sciences Branch. My work so far has centered on learning the theory and tools of my new field and conducting preliminary computational studies of aqueous zinc battery electrolytes. I first present my initial efforts to develop machine-learning (ML) models to accelerate molecular dynamics (MD) simulations of aqueous $Li_2ZnCl_4 \cdot x\,H_2O$ and $Li_2MnCl_4 \cdot x\,H_2O$ electrolytes. These initial exploratory ML models predicted unphysical behavior and produced unstable MD simulations. I then discuss how I simplified my approach to focus on models of bulk water in order to better diagnose issues with the ML training procedures. I share initial results from my studies of bulk water and investigate additional families of ML models that lead to more stable MD simulations. My ultimate goal is to develop a reproducible protocol for training ML models that function as fast and generalizable interatomic potentials for electrolyte systems like $Li_2MnCl_4 \cdot x\,H_2O$, or those with even more complex salt compositions.

1 Efforts during my first year of tenure

1.1 Background

Establishing domestic energy security and maintaining the readiness posture of our armed forces demand batteries with high energy densities, long cycling lifetimes, and high electrochemical stability. Lithium-ion batteries (LIBs) are widely adopted for their excellent energy and power density, but remain susceptible to anode–cathode dendrite formation, fires and explosions due to combustion of organic electrolytes, poor low-temperature performance, and other safety and efficacy concerns.[1,2] Additionally, there are geopolitical concerns affecting the long-term Li, Co, and Ni supply chains as well as economic and lifecycle considerations of extracting and refining Li.[3,4] These challenges motivate investigation of beyond-Li battery chemistries and aqueous-phase electrolytes.

One such family of battery chemistries involves pairing Zn electrodes with an aqueous electrolyte. Zn is favored for its high levels of mine production and reserves compared to Li, Pt-group metals, and other rare earth metals, as well as its high volumetric energy density.[5–7] On a Pourbaix diagram, Zn/Zn^{2+} plating/stripping processes at approximately -0.76 V vs. SHE are stable in acidic conditions (**Figure 1(a)**), which promotes the stability of the Zn anode during repeated charge cycling.

However, some issues complicate the development of aqueous electrolytes for Zn anodes. One primary issue is that the applied potential range favoring Zn/Zn^{2+} plating/stripping coincides with the potential range for the competing hydrogen evolution

reaction (HER). Zn^{2+} within a $[Zn(OH_2)_6]^{2+}$ ion is reduced via charge transfer through the water ligands, which destabilizes O–H bonds and encourages hydrolysis.[8–10] Thus, water is not a stable coordinating ligand for aqueous Zn^{2+} electrolytes. Additionally, in alkaline conditions, surface Zn that oxidizes to ZnO does not properly passivate the electrode surface, meaning that a ZnO surface–electrolyte interphase (SEI) may not protect the anode surface from undesired stripping at more positive cell potentials. Designing a SEI that adds reasonably low resistivity to the circuit, is stable in aqueous solution at the operating conditions of the battery and over large temperature ranges, and has a reduction potential that prevents water ligands from solvating Zn is an ongoing challenge.[11,12]

One approach to address the issue of water decomposition prevalent around Zn^{2+} ions issues is to add supporting salts to form a concentrated electrolyte. Such salts, especially chloride salts, change the coordination environment around Zn^{2+} by excluding water as a ligand, which lowers the rate of HER and thus stabilizes water in the electrolyte. Addition of supporting salts can also lower the minimum temperature at which the electrolyte can operate by preferentially coordinating free water molecules, disrupting the hydrogen bonding network and suppressing crystallization. For example, adding LiCl to an aqueous $ZnCl_2$ electrolyte, so as to maintain a $Li_2ZnCl_4 \cdot 9\,H_2O$ composition, was shown to maintain reversible Zn/Zn^{2+} plating/stripping and stable cycling at 0.4 mA cm^{-2} between $-60\,°C$ and $+80\,°C$.[1] The increased stability of water in the electrolyte is attributed to chloride ions, which displace water ligands in the Zn^{2+} solvation shells. Additionally, supporting salts containing bis(trifluoromethanesulfonyl)imide (TFSI) can also increase reversibility of a Zn-ion battery. An aqueous Zn electrolyte supported with 1 m $Zn(TFSI)_2$ and 20 m LiTFSI was shown to enable reversibility in a Zn/Zn^{2+} battery with $LiMn_2O_4$ or open-air cathodes, as well as reduced Zn dendrite formation.[10] Using salt additives to form high-density ionic electrolytes seems important to making Zn-ion batteries feasible.

Another challenge facing Zn-ion batteries is the interplay of electrolyte and cathode stability. In a $Li_2ZnCl_4 \cdot x\,H_2O$ electrolyte in acidic conditions, Cl has an oxidation potential of approximately $+1$ V vs. SHE to form $ClO_2(s) + LiClO_4(s)$, entraining all of the free Cl^- from solution. This is very close to the potential at which Li forms $LiClO_4(s)$ while leaving some Cl^- free, as part of the cathode reaction in an open-air cell. Overconsumption of free Cl^- at applied potentials that are too positive could again promote parasitic HER among the water molecules in the electrolyte, which is already a competing reaction at the pH and applied potential conditions favorable for Zn/Zn^{2+} plating/stripping. This motivates the search for cathode materials

which support charge transfer without causing the oxidation or reduction of other spectator ions in the electrolyte.

One promising class of such materials are multivalent Mn compounds, which can potentially undergo a dissolution–precipitation process in response to charge transfer. In particular, various polymorphs of MnO_2 have been studied as a versatile cathode material for Zn-ion batteries.[15–19] Mn exhibits a wide range of oxidation states, with Mn oxides containing Mn in every oxidation state from +2 to +7.[20] Thus, in an aqueous electrolyte, it may be possible for cathodic charge transfer to involve just the interconversion of different Mn oxide species without also impacting the oxidation state of Cl^- ions. A computed Pourbaix diagram[13,14] of Li_2MnCl_4 at pH = 6 and ion concentrations of 10^{-2} M demonstrates this idea (**Figure 1(b)**). Moving upward from the large $Mn^{2+} + Li^+ + Cl^-$ region to areas of the map at higher applied potentials, the system goes through both a $Mn_2O_3(s) + Li^+ + Cl^-$ phase (with Mn in the +3 oxidation state) and then a $MnO_2(s) + Li^+ + Cl^-$ phase (with Mn in the +4 oxidation state) before finally reaching the $MnO_2(s) + LiClO_4(s) + Cl^-$ phase that starts to remove Cl^- ions from solution at approximately +0.9 V vs. SHE. On the corresponding Pourbaix diagram for Li_2ZnCl_4 (**Figure 1(a)**) at pH = 6, the $Zn^{2+} + Li^+ + Cl^-$ phase ends at an applied potential of +0.9 V vs. SHE, above which the system immediately begins precipitating Cl^- out of solution to form the $Zn^{2+} + LiClO_4(s) + Cl^-$ phase.

My current research explores several critical unanswered questions about the Mn-based battery chemistry, including the solvation structure and dynamics of Mn-containing electrolytes; the mechanism of the MnO_2 dissolution–precipitation electrode and dissolution–precipitation electrodes generally; and the in-teraction of dissolution–precipitation electrodes with the electrolyte. Foremost among these concerns is my aim to understand the local solvation environments around Mn^{2+} ions, and whether the composition of solvation shells around Mn^{2+} differ from analogous solvation shells around Zn^{2+} ions. I hypothesize that the solvation structure and dynamics of $Li_2MnCl_4 \cdot x\,H_2O$ electrolytes resemble those of $Li_2ZnCl_4 \cdot x\,H_2O$ electrolytes, such that the MD simulation methodology of our prior work[1] to analyze the latter can be applied to the former.

To test this hypothesis at a high level, I proposed training machine-learned force fields (MLFFs) to conducted accelerated MD simulations of representative $Li_2MnCl_4 \cdot x\,H_2O$ electrolytes. Prior work in our group studied the solvation structures and dynamics of $Li_2ZnCl_4 \cdot x\,H_2O$ electrolytes using both Born–Oppenheimer molecular dynamics (BOMD) and force-field-based molecular dynamics (FF–MD) to understand the solvation structures of $[ZnCl_{4-m}^{2-m}]_n$ ionic networks that form.[1] One finding in our work was that FF–MD calculations predicted a strongly aggregated $[ZnCl_{4-m}^{2-m}]_n$ network of long $ZnCl_2$ chains, corresponding to sluggish transport kinetics. BOMD calculations led to similar structures but also predicted the existence of some smaller $ClZnCl – ZnCl_2$ clusters. These smaller clusters may undergo faster exchange with water in the Zn first solvation shell, potentially increasing the predicted electrolyte conductivity to more closely match experiment. However, testing this theory is limited by the high expense of and consequently short timescales accessible to BOMD simulations, even at artificially elevated simulation temperatures. Additionally, Mn salts may exhibit even more complex behavior, with multiple populations of $Mn(H_2O)_x Cl_y$ species coexisting. The
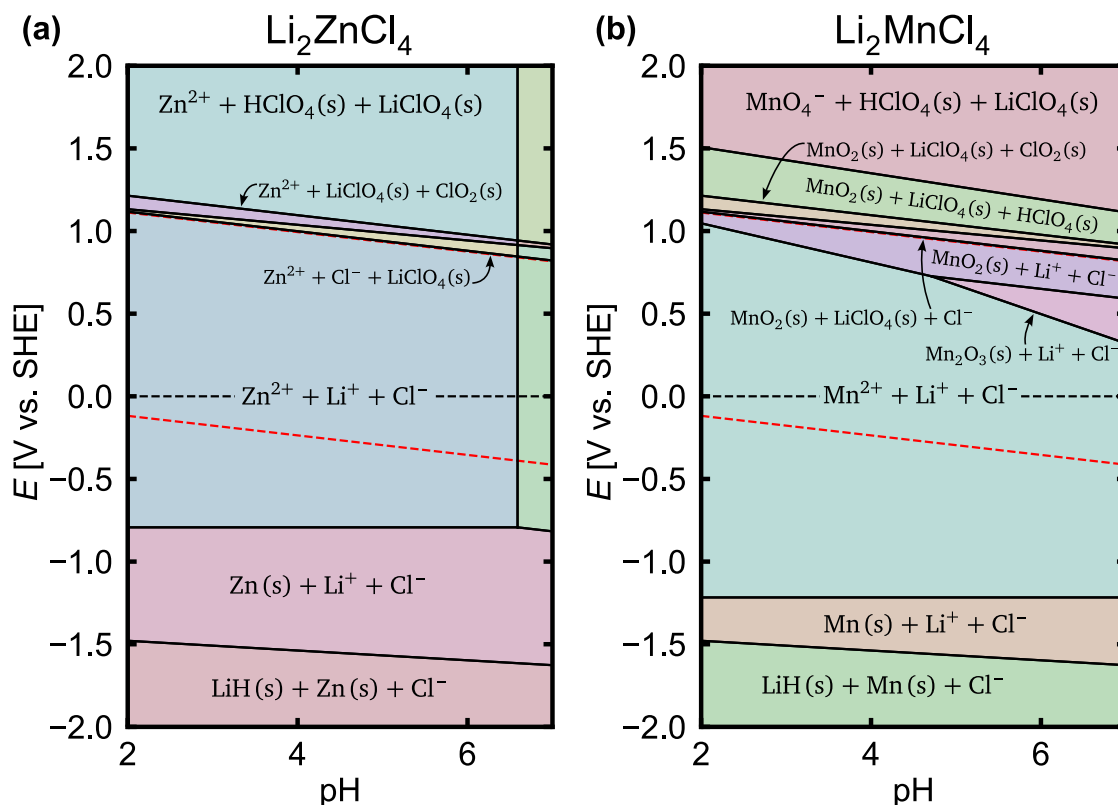

Figure 1. Computational Pourbaix diagrams[13,14] for the **(a)** $Li_2ZnCl_4 \cdot x\,H_2O$ and **(b)** $Li_2MnCl_4 \cdot x\,H_2O$ systems. Diagrams are shown for Li, Zn (or Mn), and Cl in a 2:1:4 molar composition ratio and for ion concentrations of 10^{-2} M.

need for DFT-fidelity simulations to properly treat the solvation interactions of complex ionic species over long simulation timescales (i.e., on the order of 1 ms) motivates the development of accurate and computationally efficient machine-learned force fields.[21] I seek to understand which types of ML models and training protocols will properly treat the charged anions present in battery electrolytes when used in a MD simulation.

With access to long-timescale MD simulations of $Li_2MnCl_4 \cdot x\,H_2O$, I can calculate the solvation shell compositions and thus determine whether adding LiCl salts to the electrolyte to suppress HER will work for a Li_2MnCl_4 electrolyte as it does for a Li_2ZnCl_4 electrolyte. My research will both (*i*) evaluate $Li_2MnCl_4 \cdot x\,H_2O$ as a potential electrolyte for wide-temperature-range Zn-ion batteries and (*ii*) develop design guidelines for training and deploying machine-learned interatomic potentials for simulation of dense ionic liquids more generally. Beyond the current application of Zn-ion secondary batteries, the models and methodology I develop could be used in other fields where the dynamics and environment of charged particles is important, such as in semiconductors, optoelectronics, and other solid-state electronic devices.

Below, I discuss my initial MD simulations of the Mn^{2+} solvation shell within a $Li_2MnCl_4 \cdot x\,H_2O$ electrolyte. I then describe my efforts so far to develop and test ML interatomic potentials to accelerate these MD simulations to usable timescales.

1.2 Preparation of initial configurations using the AMOEBA polarizable force field

My initial simulations sought to address the question of the local solvation environment around Mn^{2+} ions in $Li_2MnCl_4 \cdot x\,H_2O$ electrolytes compared to Zn^{2+} ions in $Li_2ZnCl_4 \cdot x\,H_2O$ electrolytes. We hypothesized that the Mn^{2+} coordination environment would be similar to that of Zn^{2+} in that Mn^{2+} would preferentially coordinate Cl^- over H_2O. To test this hypothesis, I began by simulating the solvation shell of Mn^{2+} ions within several $Li_2MnCl_4 \cdot x\,H_2O$ systems using BOMD. These replicates were prepared by my collaborator and informal mentor, Dr. Travis Pollard,[22] who used the PACKMOL[22] software package and the AMOEBA polarizable force field[23] in order to prepare some systems in which Mn^{2+} ions start off highly coordinated by water molecule ligands, as well as other systems in which Mn^{2+} ions start off highly coordinated by Cl^- ion ligands instead. The reason for preparing biased systems like these is to create a varied training dataset that samples very different parts of the configurational phase space. Because Mn^{2+} likely prefers to coordinate one of these ligands more than the other, the hope is that running these different initial systems through BOMD will create MD trajectories with high-energy configurations of ligands exchanging to allow the more preferred ligand to occupy the solvation shell. Using training data with these high-energy configurations in addition to configurations near local minima in the potential energy surface helps avoid developing MLFFs that incorrectly predict dynamics when computing the kinetics behavior of electrolytes.

To create these biased initial systems, a base system containing Li^+, Cl^- and Mn^{2+} ions along with H_2O molecules underwent classical MD simulations using the AMOEBA polarizable force field. Within the AMOEBA framework, the parameter $R_{ij,Mn-Cl}$ controls the equilibrium distance between Mn^{2+} and Cl^- ions. Because the AMOEBA force field does not have a default value of this parameter for the Mn^{2+}–Cl^- ion pair, we based our simulations on the analogous value for Fe and Cl ions ($R_{ij,Fe-Cl}$). In some simulations, $R_{ij,Mn-Cl}$ was set to a lower value (i.e., $R_{ij,Mn-Cl}/R_{ij,Fe-Cl} < 1$) to produce configurations where Mn^{2+} is mostly coordinated by Cl^-, while in others it was set to a higher value (i.e., $R_{ij,Mn-Cl}/R_{ij,Fe-Cl} > 1$) to find configurations in which Mn^{2+} is mostly coordinated by H_2O. These classical MD simulations were done using the GPU version of Tinker9 (CUDA, Nvidia A40 GPUs).[24] Ewald and real-space cutoffs were truncated to 7 Å with a $20 \times 20 \times 20$ particle mesh Ewald grid. Time integration was carried out with a multi-step RESPA integrator[25] with a 2 fs outer timestep under the NVT ensemble. **Table 1** shows the target $R_{ij,Mn-Cl}$ values used to produce each replicate. R_{ij} values for Li^+, Cl^-, and H_2O were reused from our previous work on $Li_2ZnCl_4 \cdot x\,H_2O$ electrolytes.[1]

Table 1. AMOEBA equilibrium Mn–Cl distances used to prepare initial systems with different Mn coordination environments. There was originally a system "R3" with $R_{ij,Mn-Cl} = R_{ij,Fe-Cl}$ (0% difference), but an inadvertent error made that system unusable, so it is omitted from further calculations.

Replicate	$R_{ij,Mn-Cl}$ compared to $R_{ij,Fe-Cl}$	$R_{ij,Mn-Cl}$ [Å]
R1	−10.0%	3.3325
R2	−5.0%	3.5176
R4	+7.5%	3.9804
R5	+10.0%	4.0730
R6	+12.5%	4.1656

These five replicates were generated for each of three water compositions, in which each Li_2MnCl_4 formula unit is paired with either 6, 9, or 12 H_2O molecules, producing nominal compositions of $Li_2MnCl_4 \cdot 6\,H_2O$, $Li_2MnCl_4 \cdot 9\,H_2O$, or $Li_2MnCl_4 \cdot 12\,H_2O$. Renderings of these five replicates for the $Li_2MnCl_4 \cdot 6\,H_2O$ composition appear in **Figure 2**. **Table 2** shows the actual numbers of atoms used for each composition.

I reason that if Mn^{2+}, like Zn^{2+}, preferentially coordinates one of these two ligands, then running MD simulations of these systems in parallel will quickly reveal which ligand Mn^{2+} prefers to coordinate. For example, if Mn^{2+} preferentially coordinates Cl^- instead of H_2O, I would expect that BOMD simulations of replicates R1 and R2 would keep Cl^- in the Mn^{2+} first solvation shell, while simulations of replicates R5 and R6 would start having mostly H_2O ligands around Mn^{2+} but eventually transition to having Cl^- ligands coordinated instead. Initial calculations of all five replicates supported this idea. Therefore, to save computational resources, I focused on the $6\,H_2O$, $9\,H_2O$, and $12\,H_2O$ versions of the R4 replicate (**Figure 3**), which is the replicate currently biased to have H_2O in the Mn^{2+} first solvation shell, but which should most quickly undergo replacement of those H_2O molecules with Cl^- ions. If Mn^{2+} prefers to coordinate Cl^-, then running BOMD calculations on R4 replicates should reveal ligand exchanges (i.e., substitutions of H_2O ligands with Cl^- ligands) more quickly than using the R5 or R6 replicates.

1.3 BOMD simulations of electrolytes to generate MLFF training data

I performed BOMD simulation on these three R4 systems (i.e., the R4 replicates for $Li_2MnCl_4 \cdot 6\,H_2O$, $Li_2MnCl_4 \cdot 9\,H_2O$, and $Li_2MnCl_4 \cdot 12\,H_2O$ compositions) using Kohn–Sham density

(a) R1   **(b)** R2   **(c)** R4   **(d)** R5   **(e)** R6

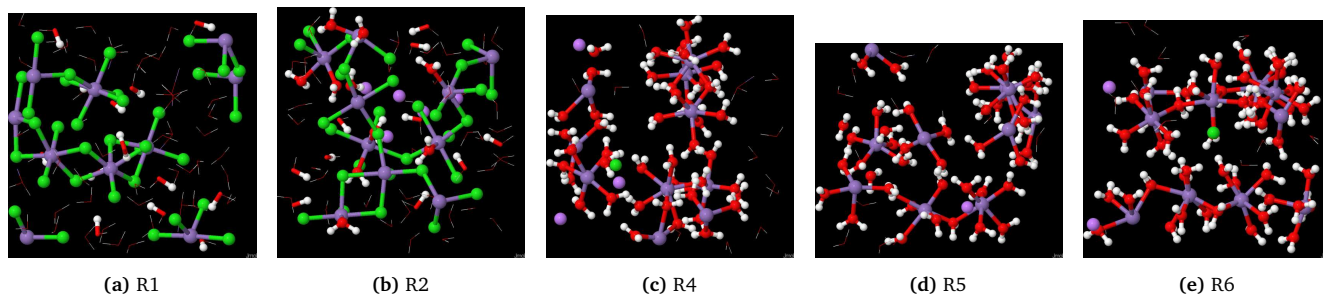**Figure 2.** Atom renderings for the five $Li_2MnCl_4 \cdot 6H_2O$ replicates. Color legend: $\bullet$ = Mn, $\bullet$ = Li^+, $\bullet$ = Cl^-, $\circ$ = H, $\bullet$ = O. Atoms within 3.6 Å of a Mn atom are shown as CPK spheres; all others are shown as wireframes.

Table 2. Effective compositions and box lengths for each nominal $Li_2MnCl_4 \cdot xH_2O$ composition.

Nominal composition	Effective composition	Total atoms	Cubic box length [Å]
$Li_2MnCl_4 \cdot 6H_2O$	$Li_{20}Mn_{10}Cl_{40} \cdot 60H_2O$	250	14.2
$Li_2MnCl_4 \cdot 9H_2O$	$Li_{20}Mn_{10}Cl_{40} \cdot 90H_2O$	340	15.3
$Li_2MnCl_4 \cdot 12H_2O$	$Li_{20}Mn_{10}Cl_{40} \cdot 120H_2O$	430	16.3

functional theory (DFT)[26] as implemented in the Vienna Ab Initio Simulation (VASP) software package[27–31] (version 6.4.3, built for CPU). Electron orbitals were expanded in a plane-wave basis up to a cutoff of 500 eV. Initial magnetic moments were set to +6.0 for Mn^{2+} and 0.0 for all other atoms. Exchange and correlation were computed using the revPBE functional,[32] and dispersion corrections were handled using Grimme's DFT–D3 scheme[33] with Becke–Johnson damping.[34] Calculations used Gaussian smearing with a smearing constant of 0.05 eV and spin polarization. The molecular dynamics time integration was performed by a velocity–Verlet algorithm[35] with a timestep of 0.5 fs. The simulation was carried out in the NVT ensemble using a Langevin thermostat with a damping coefficient of 50 ps^{-1} for atomic degrees of freedom. Deuterium mass was used for H atoms.

Each of the three systems was thermalized from 50 K to 450 K under the NVT ensemble at a rate of 50 K/ps, then stabilized under the NVT ensemble at 450 K for an additional 10 ps. The systems were then switched to the NPT ensemble, with Parinello–Rahman dynamics used to enforce an isobaric system. The temperature was maintained at 450 K. Lattice cell walls with a fictitious mass of 10 amu and with Langevin friction coefficients of 10 ps^{-1} were allowed to relax against an external pressure of 1 atm. The same Langevin friction factor was also used for each atom's degrees of freedom. I then ran the simulation for 10 ps in the NPT ensemble.

To evaluate whether Cl^- displaces H_2O in the Mn^{2+} first coordination shell, I calculated the mean coordination number for each of these ligands over the final 10 ps timeframe corresponding to the NPT portion of the simulation. The evolution in coordination number, arising from the radial pair distribution functions of Mn^{2+} with H_2O and Cl^-, was evaluated using the MDAnalysis[36,37] and SolvationAnalysis[38] software packages. The resulting coordination number evolution appears in **Figure 4**.

For all three replicates, the total coordination number of Mn^{2+} remains near 6, indicating qualitatively reasonable Mn^{2+} solvation shells with octahedrally coordinated ligands. With increasing time, I qualitatively observe ligand exchanges in which

Cl^- replaces H_2O in the first solvation shell, consistent with our hypothesis that Mn^{2+} preferentially coordinates one ligand species over another, and with the same preference for Cl^- that Zn^{2+} has. The ligand exchange is pronounced for the $6H_2O$ system, which has a relatively high concentration of Cl^- ligands compared to the $9H_2O$ and $12H_2O$ systems. For $12H_2O$, the rise in Cl^- coordination number is depressed compared to systems with fewer H_2O molecules, but this finding is also influenced by the overall loss in total ligand coordination number. More investigation is needed to understand if other, unexpected dynamics might be at play in the $12H_2O$ system. The Mn^{2+}–H_2O average first solvation shell size appears to decrease as the concentration of H_2O in the system increases, with $12H_2O$ having the smallest shell radius of 3.05 Å, reflective of the shorter average Mn^{2+}–O bonds in which more water is entrained around Mn^{2+} ions. No clear trend is apparent for the Mn^{2+}–Cl^- average first solvation shell size, although the relatively small Mn^{2+}–Cl^- average shell size for the $6H_2O$ system may reflect the high substitution of complexed H_2O with Cl^- observed in the temporally later portion of that NPT simulation.

The coordination number evolution also points to a challenge: the H_2O–Cl^- ligand exchanges appear to happen relatively slowly. Even for the $6H_2O$ system, the Mn^{2+}–Cl^- coordination number rises by only approximately 1.0 over the 10 ps period of NPT simulation. This timing indicates that much longer time scales will be needed to study the dynamics of Mn^{2+} solvation than can be feasibly achieved using current BOMD methods. Given that 1 ps of BOMD simulation time for a $Li_{20}Mn_{10}Cl_{40} \cdot 120H_2O$ system can require a full day of compute time on a modern, 4-GPU compute node, exploring the ligand exchange dynamics for this system will be expensive and relatively untenable using BOMD. This motivates training ML interatomic potentials to accelerate MD calculations and access longer simulation timescales.

1.4 Motivating the use of the Allegro framework for creating MLFFs

MD simulations consist of a system of atoms or particles which interact with each other and move through space as the simula-
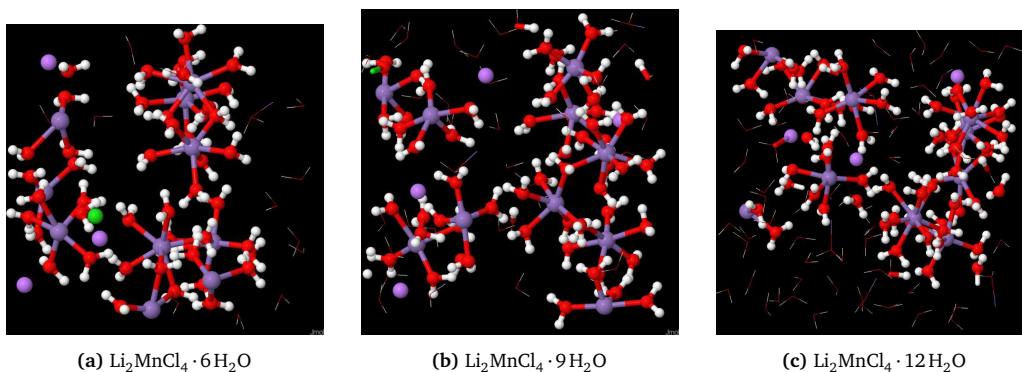
(a) $Li_2MnCl_4 \cdot 6\,H_2O$ **(b)** $Li_2MnCl_4 \cdot 9\,H_2O$ **(c)** $Li_2MnCl_4 \cdot 12\,H_2O$

Figure 3. Atom renderings for the three R4 replicate systems. Color legend: ● = Mn, ● = Li^+, ● = Cl^-, ○ = H, ● = O. Atoms within 3.6 Å of a Mn atom are shown as spheres; all others are shown as wireframes.
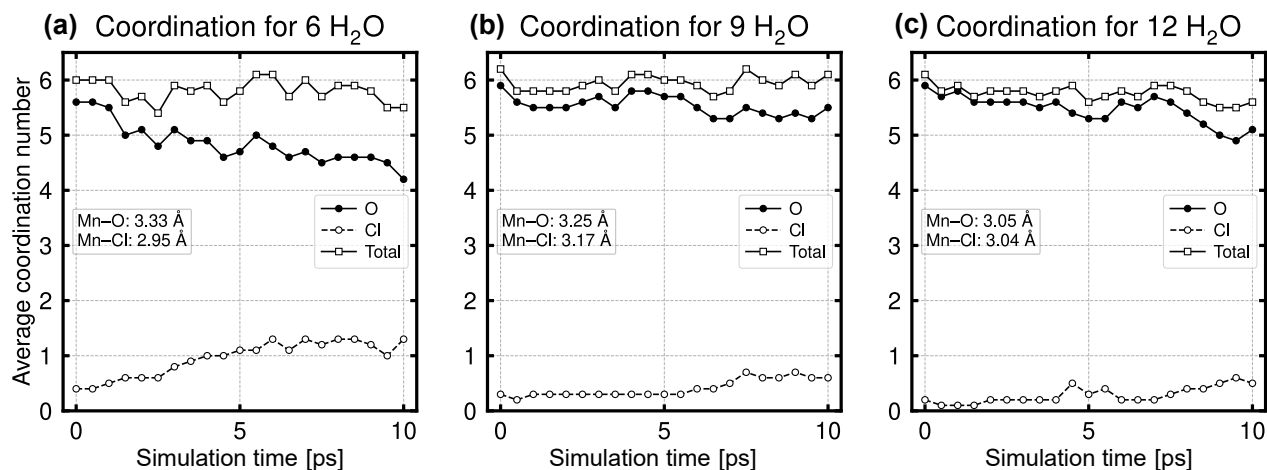

Figure 4. Evolution of first-shell coordination number of Mn^{2+} with either O (representing H_2O) or Cl^- for the **(a)** $Li_2MnCl_4 \cdot 6\,H_2O$, **(b)** $Li_2MnCl_4 \cdot 9\,H_2O$, and **(c)** $Li_2MnCl_4 \cdot 12\,H_2O$ R4 replicates at 450 K and 1 atm. Coordination numbers are sampled every 0.5 ps over the final 10 ps of NPT equilibration. Also listed on the left side of each panel are the average radii of the Mn^{2+}–O and Mn^{2+}–Cl^- first solvation shells. These radii are used as the cutoff for determining whether an H_2O or Cl^- ligand is coordinated or not.

tion moves forward in time. The critical calculation enabling MD simulations is the computation of directional forces on each atom, which tells the computer how to move each atom as the simulation progresses. Interatomic potentials are algorithms used to compute the forces on each atom and vary in complexity and computational cost. BOMD uses density functional theory as the interatomic potential to compute atomic forces but is very expensive because it must explicitly treat the ground-state electron density of the atomic system. Other interatomic potentials, such as the Lennard–Jones model,[39] are computationally very cheap but do not compute accurate atomic forces and produce MD simulations that inaccurately predict macroscopic material properties.

To accelerate MD simulations of my $Li_2MnCl_4 \cdot x\,H_2O$ systems, I surveyed ML models that could be trained to act as interatomic potentials. The catalysis and surface science communities have developed several ML models focused on systems of small molecules adsorbed on heterogeneous metal catalyst surfaces, including crystal graph convolutional neural networks, message-passing networks, and other deep-learning networks.[40–44] These models enable relatively cheap but high-accuracy predictions of total system energy and atomic forces, suitable for ML-accelerated geometry relaxation of slab–adsorbate catalyst systems, but may not offer this accuracy in the context of bulk

liquids like a battery electrolyte. Additionally, a ML model serving as an interatomic potential for a MD simulation should ideally be E(3)-equivariant with respect to symmetry operations like translation, rotation, and mirroring. This broadly means that predicted scalar quantities like the system energy should not change when the system's atoms are uniformly translated, rotated, or mirrored, but that predicted vector quantities like atomic forces and dipole moments should, as is the case physically.

One family of models designed both for bulk materials simulation generally and with E(3) equivariance built into the model architecture includes the NequIP[45] and Allegro[46] models developed by Boris Kozinsky's group at Harvard University. NequIP follows the lead of crystal graph convolutional neural networks in representing atoms in a bulk material as nodes with learnable parameters interconnected via graph edges with more learnable parameters.[47] However, NequIP extends this concept by learning parameters that are higher-order geometric tensors rather than simple scalars and by encoding angular information about the local chemical environment around an atom rather than just scalar interatomic distances. This allows NequIP to achieve energy and force predictions, achieving respective mean absolute errors (MAEs) on the order of 5 meV/atom or 5 meV/Å and outperforming similar deep-learning models. The Allegro

model architecture approximates the NequIP architecture and is designed to encode the same local chemical environment, but using only information within a strict radial cutoff around each atom. This strictly localized Allegro architecture enables efficient parallelization across an atomic system's physical subdomains, greatly increasing the speed of Allegro MD simulations on massively parallel computers without significantly compromising accuracy.

1.5 Initial training and validation of equivariant Allegro MLFF models on $Li_2ZnCl_4 \cdot xH_2O$ data

I intended to train Allegro ML models on my existing $Li_2MnCl_4 \cdot xH_2O$ BOMD data in order to enable accelerated MD simulations for this system. However, my $Li_2MnCl_4 \cdot xH_2O$ BOMD data indicated slow ligand exchange of H_2O and Cl^- within the Mn^{2+} first solvation shell (**Figure 4**). Understanding ligand exchange is critical to understanding macroscopic properties of battery electrolytes, such as ionic conductivity. The 10 ps of data I generated for each R4 replicate system likely does not adequately sample the region of configurational space in which H_2O–Cl^- ligand exchanges occur, and training a MLFF on this data will lead to MD calculations that don't yield useful information on ligand exchange processes. Additionally, I found that DFT calculations on $Li_2MnCl_4 \cdot xH_2O$ systems are significantly more challenging to converge and take much longer than the corresponding calculations on $Li_2ZnCl_4 \cdot xH_2O$ systems in our previous work. [1] It would likely take months of BOMD calculations on the $Li_2MnCl_4 \cdot xH_2O$ system to generate enough data to observe the types of ligand exchanges already available in our previous $Li_2ZnCl_4 \cdot xH_2O$ BOMD data. Therefore, I chose to instead train Allegro MLFFs on the available $Li_2ZnCl_4 \cdot xH_2O$ data to better understand the MLFF training process and typical challenges we might encounter.

I collated existing $Li_2ZnCl_4 \cdot xH_2O$ BOMD data generated by my colleague, Dr. Travis Pollard, into a shuffled collection of images covering NPT-ensemble simulations of $Li_2ZnCl_4 \cdot (6,8,10,15)H_2O$ systems at 1 atm. As with my $Li_2MnCl_4 \cdot xH_2O$ BOMD data, the effective compositions in each box are ten times the nominal compositions, as shown in **Table 3**. This data is comprised mostly of MD trajectories from two tasks: (*i*) NPT annealing of $Li_2ZnCl_4 \cdot xH_2O$ systems from 450 K to 298 K at 1 atm; and (*ii*) NPT equilibration of the annealed systems at 298 K and 1 atm for up to 24 ps. The frames I selected from these datasets are summarized in **Table 4**. To bias the MLFF training towards learning intermolecular interactions from the higher-temperature portions of the annealing trajectory, I sampled images from the first half of each annealing trajectory (going from 450 K to 374 K). To avoid spurious system shocks or artefacts related to the beginning of equilibration, I sampled images from only the second half (from 12 ps to 24 ps) of each of the NPT equilibration trajectories. I sampled only one of every 50 images in these equilibration trajectories to again focus the MLFF on high-temperature, more repulsive conditions.

Initial calculations indicated that the Allegro MLFF might benefit from even more examples of repulsive forces between atoms, so I additionally included some shrunken images in the training dataset as well. I randomly selected 25% of the annealing and equilibration images (or 360 images), made copies of each image scaled down to both 92% and 80% of

its original lattice constant, and performed single-point DFT evaluations of the system energy and atomic forces for each image. This resulted in a total of 720 images with significantly higher interatomic repulsive forces. These images were then added to the original 1200 annealing and 240 equilibration images for a total of 2160 images in the dataset.

I then trained an Allegro MLFF on this dataset, using a 90/10 train/validation split, a training batch size of 1 image (stochastic gradient descent), a radial cutoff of 4 Å, spherical harmonics of order 2, one tensor product layer, and equal weights on force and per-atom system energy losses in the loss function. Training proceeded until the energy and force MAEs on the validation set were approximately 5 meV/atom and 50 meV/atom, respectively, which is in line with the reported MAEs for the Li_3PO_4 solid electrolyte system used an Allegro model benchmark. [48] However, even with the bias in training data towards interatomic repulsion through using higher-temperature or shrunken images, my Allegro model predicted unexpected atomic forces in a MD simulation of $Li_2MnCl_4 \cdot 6H_2O$ that led to unphysical particle aggregation of H and O atoms (**Figure 5**).

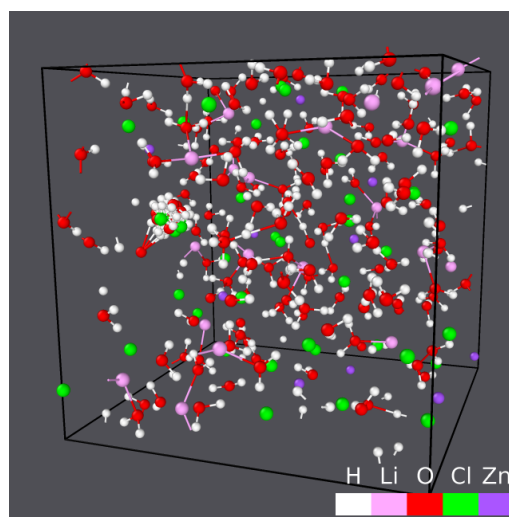**Figure 5.** Unphysical particle aggregation arising from a classical MD simulation of $Li_2ZnCl_4 \cdot xH_2O$ using an Allegro MLFF. The Allegro model is trained using spherical harmonics of order 2 and a radial cutoff of 4 Å. The classical MD simulation is done under NVT at 450 K and this frame is taken at 1.72 ps into the simulation. Aggregation is visible near the left of the frame in which several H and some Cl atoms surround a core of O atoms.

I reasoned that adding a short-range Ziegler–Biersack–Littmark (ZBL) screening interaction [49] might help compensate for the lack of repulsive forces between H and O atoms in the training data. However, ZBL corrections only delayed and did not prevent this unphysical aggregation.

1.6 Switching to bulk water to better understand MLFF model training protocols

The challenges with training an Allegro MLFF to predict the behavior of $Li_2ZnCl_4 \cdot xH_2O$ electrolytes could have arisen for several reasons. For example, the MLFF may not have been trained for a sufficiently high number of epochs, and the model hyperparameters used above also may not be optimal for the $Li_2ZnCl_4 \cdot xH_2O$ system. Alternatively, the MLFF might itself be incapable of representing a charged ionic liquid no matter

Table 3. Effective compositions and box lengths for each nominal $Li_2ZnCl_4 \cdot x\,H_2O$ compositions. Box lengths were taken from the beginning of each NPT equilibration trajectory at 298 K and 1 atm.

Nominal composition	Effective composition	Total atoms	Cubic box length [Å]
$Li_2ZnCl_4 \cdot 6\,H_2O$	$Li_{20}Zn_{10}Cl_{40} \cdot 60\,H_2O$	250	14.2
$Li_2ZnCl_4 \cdot 8\,H_2O$	$Li_{20}Zn_{10}Cl_{40} \cdot 80\,H_2O$	310	14.9
$Li_2ZnCl_4 \cdot 10\,H_2O$	$Li_{20}Zn_{10}Cl_{40} \cdot 100\,H_2O$	370	15.7
$Li_2ZnCl_4 \cdot 15\,H_2O$	$Li_{20}Zn_{10}Cl_{40} \cdot 150\,H_2O$	520	17.1

Table 4. Composition of the training dataset used to train the Allegro model. Each row represents a different subset of BOMD training data. All subsets are shuffled together in the final training dataset.

System	Procedure	Time, conditions	Images	Effective total images included in training set
$Li_2ZnCl_4 \cdot (6,8,10,15)\,H_2O$	Annealing: NPT @ 450 K → 298 K	12 ps (ramp −12.67 K/ps)	2.4k	1.2k
$Li_2ZnCl_4 \cdot (6,8,10,15)\,H_2O$	NPT equilibration @ 298 K	12 ps (12 → 24 ps)	24k	240
Shrunken images: $Li_2ZnCl_4 \cdot (6,8,10,15)\,H_2O$	Annealing: NPT @ 450 K → 374 K (300 images); NPT equilibration @ 298 K (60 images)	All 360 images scaled to both 92% and 80% of original lattice constant.	360	720

which hyperparameters are used or for how long it trains. To isolate the cause of failure and to better understand how to train such models, we proposed first training an Allegro MLFF on bulk water. Successfully training Allegro MLFFs on bulk water data and reproducing the known system density, self-diffusion coefficient, and viscosity of bulk water gives us confidence in applying Allegro to more complex systems.

We proposed training Allegro MLFFs on systems of increasing complexity: first on bulk water, then on $LiCl \cdot x\,H_2O$, then on $Li_2ZnCl_4 \cdot x\,H_2O$, then finally on $Li_2MnCl_4 \cdot x\,H_2O$. Each stage of training should yield insight on how to extend MLFFs to the next more complex system. Successfully training the bulk water system should inform how we train a $LiCl \cdot x\,H_2O$ system; successfully training the $LiCl \cdot x\,H_2O$ system should inform how we deal with Cl^- ions in a $Li_2ZnCl_4 \cdot x\,H_2O$ system; and so on.

To begin this process, I switched to using MD trajectories of molecular water in a box under the NPT ensemble at 298 K and 1 atm as the training data. I took five random frames of bulk water from a NPT trajectory used as training data in the work of Cheng et al.[50] (https://github.com/BingqingCheng/ab-initio-thermodynamics-of-water/blob/master/training-set/dataset_1593.xyz). I then used these five frames as initial frames to generate five parallel MD trajectories of liquid water under NPT conditions at 298.15 K and 1 bar. These classical BOMD calculations were performed with CP2K[51] (version 2024.1, CPU). Forces and energies were evaluted by DFT at revPBE level of theory with DFT–D3 Grimme dispersion corrections. Molecular orbitals were expanded in a Gaussian-and-plane-wave approach, using local double-ζ basis sets (DZVP–MOLOPT–SR–GTH) and Goedecker–Teter–Hutter pseudopotentials (GTH–PBE) parameterized for the PBE family of functionals. Molecular dynamics were conducted in the NPT ensemble with an isotropic cell for at least 20 ps at a timestep

of 0.5 fs and with every 4 frames written to disk, for a total of 56.9k frames. For each of the five resulting trajectories, the first 2.5 ps of the trajectory was discarded and the rest used as equilibrated data. These remaining frames were shuffled and divided into chunks of 2.5k to provide 13 separate datasets for validation and testing tasks. A similar procedure was used to prepare the hydrogen-mass-repartitioned MD training data discussed later.

1.7 Challenges with using Allegro MLFFs for inference during bulk water MD simulations

I conducted basic hyperparameter exploration and trained approximately 30 Allegro MLFFs in hopes of finding one that could reproduce known macroscopic properties of liquid water. I found that most Allegro models predict either system energy or atomic forces well, but usually do not predict both quantities well simultaneously. The velocity–Verlet algorithm depends more on accurate forces than on having an accurate total system energy. Thus, I preferred models with accurate force predictions and ultimately settled on one with a force MAE of approximately 30 meV/Å, on par with the benchmark Li_3PO_4 system in the Allegro paper[48] and reasonably cheap to evaluate. When using this model for accelerated MD calculations of water, I did not observe particle aggregation, but did observe a large deviation in the system density from the expected water density of 1 g/mL.

In fact, Allegro models appear to systematically underpredict the density of liquid water. This underprediction happens despite hyperparameter optimization of the spherical harmonic function order, the number of tensor product layers, the radial cutoff, or the dimensions of the environment-embedding, latent, or two-body network layers. For hyperparameter com-

binations that didn't cause the simulation box to immediately blow up or collapse, the system density converges to approximately 0.80 g/mL (**Figure 6**). Counterintuitively, I found that specifying a weight on the losses of cell stresses in the loss function actually made the predicted water-box density under NPT less accurate than omitting stress from the loss function altogether. Additionally, the units of cell stress in the training data, which from the available documentation could be either bar or eV/Å^3, did not seem to affect the accuracy of the predicted system density. I also added ZBL screening corrections with inner and outer radial cutoffs of 0.1 Å and 1.0 Å, reasoning that a lack of repulsive forces between H and O atoms might affect macroscopic density change, but these corrections did not meaningfully improve the system density.
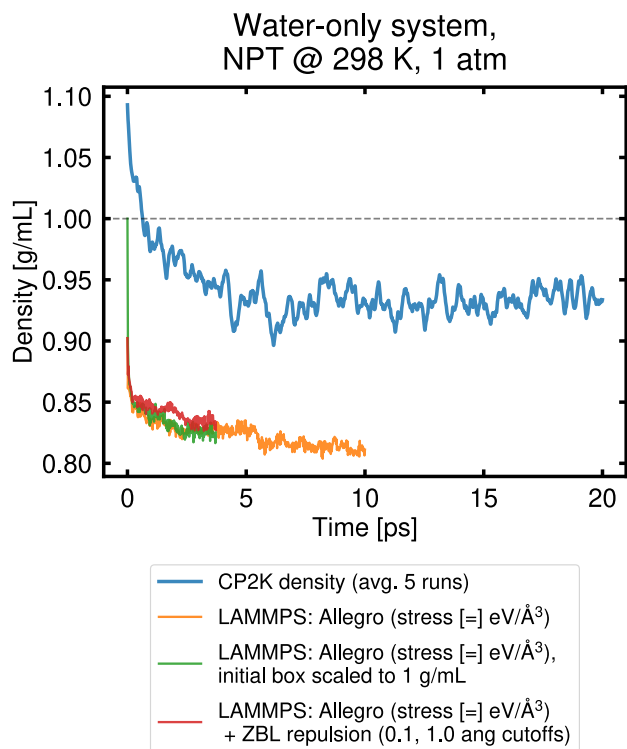
Water-only system, NPT @ 298 K, 1 atm

Figure 6. Density evolution of liquid water (1728 H$_2$O) during a NPT simulation as calculated by BOMD (CP2K) versus LAMMPS with an Allegro potential. The dotted line indicates a density of 1 g/mL.

One possible explanation for the underpredicted density might be that my original BOMD training data of water molecules used a H weight of 1 amu, which may overrepresent intramolecular H–O vibrations in the atomic forces in the training data. Fast H–O vibrations across many water molecules correspond to a very noisy potential energy surface, making system energies and atomic forces harder to predict correctly. These fast vibrations may also artificially inflate repulsive forces between water molecules, which might explain why even the CP2K-predicted density in **Figure 6** is less than experimentally expected. To address this problem, I generated new BOMD trajectories in which I set the respective masses of H and O atoms to 3 and 12 amu. Moving some mass from the O atom of a water molecule to its neighboring H atoms is a technique called hydrogen mass repartitioning. Hydrogen mass repartitioning damps some of the more noisy H–O vibrations while still

conserving the total mass of each water molecule. The resulting BOMD trajectories can improve ML model predictions of system energies and forces, but impact water transport properties less than if one was to simply deuterate or tritiate each H atom without adjusting the O masses.[52,53] Unfortunately, hydrogen mass repartitioning did not resolve the problem of artificially low system density: the average water-box density predicted by DFT still holds near 0.93 g/mL over 5 BOMD replicates and 20 ps of simulation, and the resulting Allegro models still predicted even lower densities near 0.80 g/mL over 5 classical MD replicates and 20 ps of simulation.

With the Allegro model seemingly unable to reproduce even the density of water, I concluded that this family of MLFFs may not be currently appropriate for my research because liquid electrolyte density and viscosity are critically important properties for battery electrolyte development.

1.8 Motivation for exploring DeePMD-kit MLFFs

I next turned to exploring the DeePMD-kit ML framework, which offers another family of deep-neural-network-style ML models for learning interatomic potentials from system energies, atomic forces, and system stresses.[54,55] I initially avoided DeePMD-kit models because they require more training data than NequIP/Allegro models do[45] and, like the ML framework built into VASP,[56–58] also may not be equivariant with respect to mirror plane inversions. However, the challenges in training NequIP/Allegro models reproduce even the density of water made it compelling to investigate alternative MLFFs.

A DeePMD-kit model combines one of of several available atomic system descriptors with a fitting net. Descriptors are algorithms that examine the local chemical environment around a particular atom and convert this information to tensors of numbers. Such tensors specifically encode the types of neighboring atoms, their distances from the particular atom, and their orientation around the particular atom. The generated tensors are a machine-readable representation of the atomic system that make it easier for a ML model to learn output properties. DeePMD-kit comes with several descriptors ranging in complexity from encoding just position data about neighboring atoms[59] or two-/three-body interactions,[60,61] to larger neural-net, attention-based algorithms[62] not unlike those of NequIP. Recent versions of DeePMD-kit have newer, "smooth" descriptors that are continuous and differentiable at their cutoff radii, which facilitates model training and simulation stability. The fitting net is a neural network model that learns to predict the system energy, atomic forces, or atomic cell stresses given the output of the descriptor. Both the descriptors and fitting nets have several tunable hyperparameters. The DeePMD-kit framework has some support for long-range electrostatic interactions,[63] which would be useful for modeling charged ions an electrolyte liquid, although these descriptors are more complex to train and use than the local descriptors are. For the purposes of this preliminary DeePMD-kit investigation, I focused only on the smooth, local descriptors discussed below.

1.9 Initial training and validation of DeePMD-kit MLFFs on bulk water data

As with my previous work using Allegro models, I first attempted to train DeePMD-kit interatomic potentials for systems of liq-

uid water before moving to more complex systems of solvated ions. I considered the `se_e2_a`, `se_e2_r`, `se_e3`, `se_a_tpe`, `se_a_ebd_v2`, and `se_atten_v2` descriptors as implemented in DeepMD-kit 2.2.11. I began by establishing reasonable default values for each descriptor's hyperparameters based on the documentation, then enumerated different possible values for each tunable hyperparameter, leading to 255 different trainable DeepMD-kit models. The general DeepMD model architecture makes these models individually cheaper to train and evaluate than NequIP/Allegro models, allowing wider exploration of the hyperparameter space. Each of these models was trained on H-mass-repartitioned water-box BOMD data, with a training data split of 2025 frames taken from a 81/9/10 train/validation/test split of 2500 frames. Each frame contains 64 water molecules (or 192 atoms total). Training was done using stochastic gradient descent (i.e., a training batch size of 1), as I suspected that this setting will help improve model performance on predicting individual atomic forces as it does for NequIP/Allegro models. Of these initial 255 DeepMD-kit models, 189 of them trained without crashing due to memory or hyperparameter compatibility errors. Initial hyperparameter exploration showed that with default hyperparameters, the `se_a_ebd_v2` descriptor performed the best of the available descriptors. This best-performing model (which I named `hpe-aaad`) has a cutoff radius of 6 Å as well as a fitting net of 3 layers of 240 neurons each, achieving validation set root-mean-squared errors (RMSEs) of approximately 55 meV/Å (force) and 0.09 meV/atom (energy) after training for about 25k gradient descent steps, again on par with the accuracy required for an Allegro model to achieve a stable MD simulation.[48]

1.10 Challenges with using DeepMD-kit MLFFs for inference during bulk water MD simulations

While the training process appeared to succeed, model inference indicated more challenges. When I attempted classical MD using model `hpe-aaad` under NPT conditions at standard conditions (velocity–Verlet dynamics with a timestep of 0.5 fs), I observed particle aggregation as soon as 0.5 ps and contraction of the simulation box to a system density above 1 g/mL. Adding a ZBL screening interaction with respective inner and outer cutoffs of 0.5 Å and 1.0 Å to increase repulsion appears to delay the onset of particle aggregation, but particle aggre-

gation still dominates the simulation by approximately 10 ps (Figure 7), along with causing the system density to drop below 1 g/mL. Extending the ZBL inner and outer cutoffs outward to 0.75 Å and 1.5 Å to temper the repulsive interactions between molecules causes drastic contraction of the simulation box to a density of 4.55 g/mL. Even when running under NVT conditions and enforcing a system density of 1 g/mL, significant particle aggregation occurs within 7 ps, much sooner than the timeframe required to obtain statistically reliable results. As with Allegro models, adding a ZBL repulsion interaction to the DeepMD-kit interatomic potential during classical MD only delays but does not prevent particle aggregation under either the NVT or NPT ensembles.

I reasoned that despite the relatively low validation RSMEs of my DeepMD-kit model for atomic forces and system energy, perhaps my model was not adequately learning atomic forces near the end of training, as the default loss function in DeepMD-kit models involves the gradual shift of weight from force loss to energy loss as the simulation proceeds, with no weight on virial tensor component losses. I hypothesized that a greater weight on virial tensor component loss along with constant weights on the force, energy, and virial tensor component losses might allow the model to more accurately learn intermolecular interactions, avoiding particle aggregation and blowup/collapse of the system simulation box under NPT. To test this hypothesis, starting with my existing `hpe-aaad` model, I further explored different weights of energy, force, and virial tensor component losses in the loss function as well as the initial learning rate for the fitting net. This produced an additional 179 models to train. Training proceeded until validation set RMSEs were again near 50 meV/Å (for forces) and 5 meV/atom (for energies). All 179 models completed training for approximately 24 hours, or about 19k gradient descent steps. I chose six models (named `hpe-25a8`, `hpe-1cb6`, `hpe-a010`, `hpe-ba1f`, `hpe-0620`, and `hpe-ba44`) that appeared to perform best on both system energy and atomic force predictions. These models still tend to weight force loss over energy loss by a factor of 1000, with weights on energy losses typically a factor of 2 greater than weights on virial tensor component losses. Furthermore, these models typically have validation RMSEs of 0.015 meV/atom (for energy) and 60 meV/Å, again within the typical magnitude of error corresponding to stable MD simulations for NequIP/Allegro models.[48]
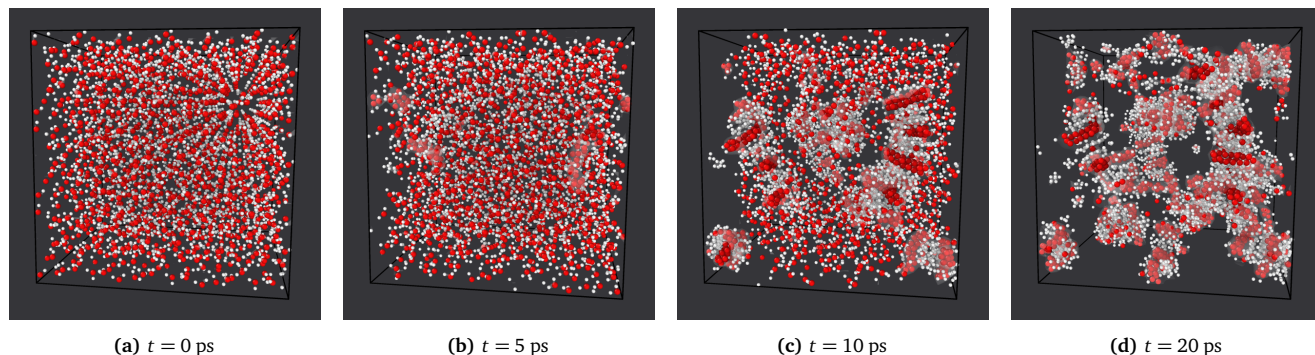

(a) $t = 0$ ps   **(b)** $t = 5$ ps   **(c)** $t = 10$ ps   **(d)** $t = 20$ ps

Figure 7. Classical MD simulation of $1728\,H_2O$ using DeepMD-kit model `hpe-aaad` based on the `se_a_ebd_v2` descriptor with a 6 Å radial cutoff and fitting net dimensions of $[240, 240, 240]$ neurons. Simulation is done in the NVT ensemble at 298.15 K and with a ZBL repulsion screening interaction at respective inner and outer cutoffs of 0.5 Å and 1.0 Å. Despite using the ZBL repulsion interaction, significant particle and phase aggregation still occurs. Semitransparent white regions surround areas of particle aggregation. ○ = H, ● = O.

However, these models unexpectedly performed even worse than the `hpe-aaad` model with respect to NPT simulation of bulk water. Instead of the box density only slowly converging to a density other than 1 g/mL (perhaps with some gradual particle aggregation), the simulation box immediately balloons or contracts drastically within the first 2.5 ps of simulation (**Figure 8**). All models except `hpe-25a8` cause the simulation box to expand towards a water density of 0 g/mL, leading to isolated water molecules or clusters. Model `hpe-25a8` behaves in the opposite fashion, causing the box to collapse and eventually reach a system density of approximately 5.5 g/mL. Further

Density evolution using loss-weight-optimized derivatives of model hpe − aaad

Mean density, model hpe-25a8
Mean density, model hpe-1cb6
Mean density, model hpe-a010
Mean density, model hpe-ba1f
Mean density, model hpe-0620
Mean density, model hpe-ba44
Mean density, pretrained se_e2_a model
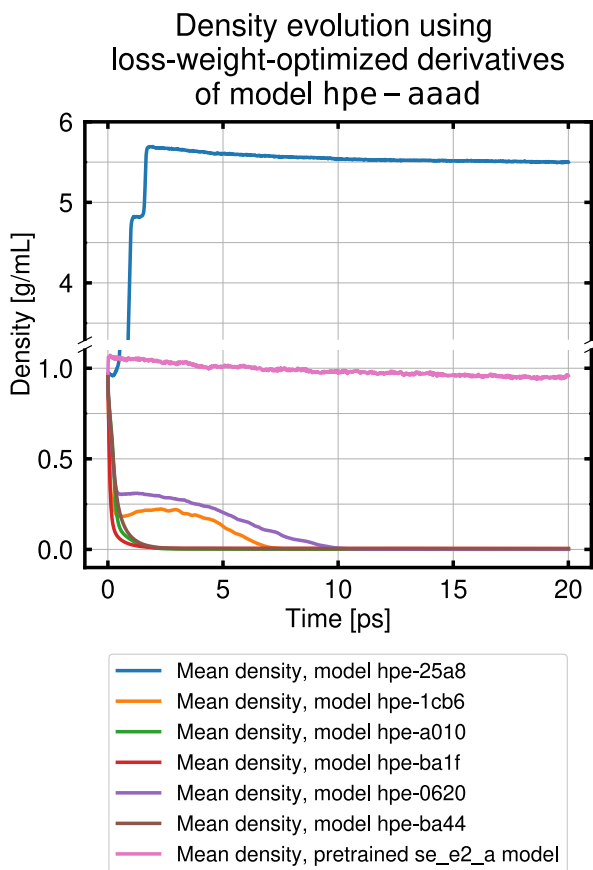
Figure 8. Evolution of water system density under standard conditions (NPT ensemble) for classical MD simulations using DeePMD-kit derivatives of the `hpe-aaad` interatomic potential. An additional MD simulation using a pretrained `se_e2_a` DeePMD-kit model is included for comparison.

investigation of each model's performance on predicting energies, forces, and virial tensor components show that these six models universally predict atomic forces very well, but system energy predictions have much more variance and tend to be overpredicted by 1 to 3 eV in most cases. I also unexpectedly observed that most of these models accurately predict off-diagonal virial tensor components well, but tend to over- or underpredict on-diagonal components by 0.001 to 0.005 eV/Å^3 (0.16 to 0.80 kbar). However, no consistent correlation exists between whether the on-diagonal components are over- or underpredicted and whether the simulation box will respectively expand or contract under NPT conditions.

To assess whether the inaccuracies of my DeePMD-kit model predictions arise from the way I trained the models or are sys-

temic to the DeePMD-kit framework itself, I ran simulations using a pretrained DeePMD-kit model based on water-box data which uses the `se_2_a` descriptor and which is included in the supporting information of the DeePMD-kit V2 paper.[55] The classical MD system density evolution using this model is shown in **Figure 8** for comparison. Like the `hpe-aaad` model I trained, this model also uses a radial cutoff of 6 Å, fitting network layer dimensions of [240, 240, 240], and the same default ramping energy, force, and virial-tensor-component loss function weights (0.02 → 1 for energy, 1000 → 1 for forces, and 0 → 0 for virial tensor components). Surprisingly, using this pretrained model as an interatomic potential in a classical MD simulation of my water-box data produces the behavior I would expect: a consistent system density of near 1 g/mL without particle or phase aggregation through the entire 20 ps simulation period. While the density does begin to drop closer to 0.9 g/mL at 15 ps, this model performs much better than any of the DeePMD-kit models I have trained. Some differences in training protocol may explain why the pretrained model performs better than my models. In particular, the pretrained model was trained using an automatic batch size, indicating that stochastic gradient descent approaches may not offer the same likelihood of accurate force predictions for DeePMD-kit models as they do for NequIP/Allegro models. Additionally, the DeePMD-kit model architecture, like NequIP/Allegro models, seems to more accurately reproduce water system density when virial tensor component losses are excluded from the loss function. These findings suggest that rules of thumb for training one deep-learning, tensor-based MLFF may not extend to another MLFF. Additionally, for liquid systems, there may currently be a trade-off between accurately capturing stresses on the simulation cell (which are needed to compute macroscopic viscosity) and accurately predicting the system density, limiting the accuracy to which one can predict the system's kinematic viscosity.

It is also possible that my models were trained for too short a time and on too small a dataset of BOMD images, and an ongoing effort is to retrain DeePMD-kit models on larger training datasets, for greater numbers of gradient descent steps. As a non-E(3)-equivariant family of models, DeePMD-kit models also likely require more training data to achieve comparable accuracy in system energy and atomic force predictions.[45,64] One possibility is to train the DeePMD-kit models on the original dataset of 300k non-hydrogen-mass-repartitioned water-box MD frames, which is about an order of magnitude larger than the hydrogen-mass-repartitioned dataset. Additionally, I would train DeePMD-kit models using larger batch sizes as this was the training protocol for the more accurate pretrained model discussed above.

A final possibility is to investigate whether my DeePMD-kit models used too many parameters and thus exhibited high variance that led to the unphysical predicted behavior of water. The fitting network layer dimensions of [240, 240, 240] nodes may be too large to represent the my water-box system, and limiting the size of the embedding and fitting networks could lower the amount of training data required. A DeePMD-kit model trained on as few as 100 frames of coarse-grained water MD data with embedding network layer dimensions of [8, 16, 32] nodes and energy fitting network layer dimensions of [32, 32, 32] nodes was shown to acceptably reproduce the O–O radial distribution functions and three-body angular distribution functions

of atomistic TIP3P and SPC/E water.[65] Carefully limiting the size of the neural networks in a DeePMD-kit model may both reduce model variance with respect to test data and make the model cheaper to evaluate, directly translating to more stable MD simulations that extend to longer timescales.

2 Ongoing and future efforts

While I have not completely abandoned hope that NequIP/Allegro and DeePMD-kit models might eventually be trainable for my systems of interest, I am focusing my efforts on other approaches to constructing a performant MLFF. These include the active learning capability built into recent (as of version 6.3.0) versions of VASP, as well as the machine-learning polarized atomic orbital (PAO–ML) framework of CP2K. The VASP ML framework differs from the ML frameworks I previously investigated because it strategically constructs the training dataset as it trains rather than requiring a training dataset to be completely generated at the start. The CP2K PAO–ML framework is also interesting because it internally learns atomic orbital structure rather than atomic forces and system energies, enabling calculation of properties beyond system energy and atomic forces. These efforts are ongoing and no clear comparison to the other ML approaches is available yet.

2.1 Active learning with stream-based sampling to construct ML force fields

The ML approach built into VASP is an application of active learning that uses primarily stream-based sampling (**Figure 9**). Active learning is an approach to training supervised ML models when it is expensive to calculate the labels on unlabeled training data.[66] In my systems, the unlabeled training data is the configuration of atomic positions, atom types, and cell dimensions for each frame of MD trajectory. The labels are system energy, atomic forces, and cell stresses, and are generally expensive to calculate. Supervised learning approaches, such as those implemented in NequIP/Allegro and DeePMD-kit models, assume that all MD trajectory frames that comprise the training data have already had their system energies, atomic forces, and cell stresses calculated from the start. In this scenario, training is mostly an iterative process that seeks to minimize the error between model predictions and known label values. In contrast, active learning approaches begin model training with little or no training data and use one or more sampling approaches to decide which pieces of unlabeled training data to add to the model at each training step (i.e., by invoking the expensive calculation of the system energy, atomic forces, and cell stresses). The hope is that with an adequate sampling approach, the ML model can learn how to predict the correct system energy, atomic forces, and cell stresses while minimizing the number of expensive calculations that must be done to obtain this training data in the first place.

VASP's on-the-fly ML scheme uses a molecular dynamics engine to generate new training configurations of atoms (i.e., the unlabeled data represented as frames of atomic positions and cell dimensions). Either electronic structure calculations or the current version of the ML model is used to calculate forces on those atoms and propagate them in space, generating the next
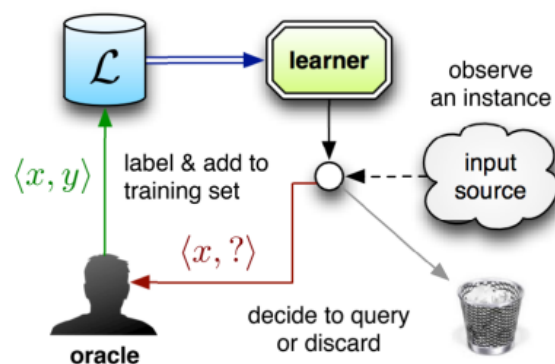

Figure 9. Strategy for stream-based (or selective) sampling. The oracle represents an expensive computation call, like a DFT evaluation of the atomic forces of a system in VASP. Reproduced from Settles[66], Figure 1.5(a).

frame of unlabeled data. At each step, a Bayesian error estimate is used to calculate the model uncertainty with respect to predictions of forces on each atom. Predicted Bayesian errors above a specified threshold trigger an expensive first-principles evaluation of the system energy, atomic forces, and cell stresses, with this now-labeled MD frame added to the pool of labeled training data that is used to periodically update the ML model weights. The sorting of each frame of training data for inclusion into or exclusion from the labeled training data individually and in sequence, rather than mass sorting of a pool of frames at once, indicates a stream-based sampling approach. VASP's ML framework appears to be designed primarily for accelerating MD calculations. Stream-based sampling is used to gradually expose the ML training process to atomistic configurations that resemble the phase space that MD would normally explore, with the ability to save time by retraining the ML model only when a new frame is evaluated with DFT and added to the dataset. The training data fed to the model can then be adapted in real time to better balance exploration and exploitation by biasing the simulation in the same ways one would apply during MD to explore a larger region of phase space. For example, the VASP documentation recommends on-the-fly training using a thermalization ramp of gradually increasing MD simulation temperature. As the model improves at predicting atomic forces for low-temperature structures of low atomic repulsion (i.e., high exploitation of local phase space), it can be forced to explore higher-temperature ranges with more chaotic and repulsive interactions (i.e., more exploration of non-local phase space).

As with previous efforts, I started by training VASP ML models of bulk water. Because system density was an issue in previous attempts with NequIP/Allegro and DeePMD-kit models, I investigated a variety of exchange–correlation functionals and dispersion corrections when performing the BOMD evaluations of training data during the active learning process. These combinations included the following: revPBE + DFT–D3(BJ); r^2SCAN + DFT–D3(BJ); r^2SCAN + DFT–D4; r^2SCAN + rVV10; rSCAN + DFT–D3(BJ); rSCAN + DFT–D4; rSCAN + rVV10; and rSCAN with no dispersion correction. Using these combinations aims to explore whether more advanced meta-GGA functionals are necessary to capture the density of liquid water under ambient conditions and the self-diffusion coefficient of water molecules

in the liquid. I am also interested in understanding which dispersion corrections would be the most helpful in modeling liquid interactions in a $Li_2MnCl_4 \cdot x\,H_2O$ electrolyte.

I started with a box of 64 H_2O molecules at a system density of 1 g/mL as the initial image. I set up VASP ML active-learning training jobs for each of the eight combinations of exchange–correlation functional and dispersion correction. As recommended in the VASP documentation, I set the plane-wave energy cutoff to 910 eV to accommodate changes in cell size during the NPT MD training runs. I also constrained the cell to remain cubic in shape throughout the training process. Spin polarization was used for all calculations, as well as a Gaussian smearing constant of 0.02 eV. The MD training run took place in the NPT ensemble at an external pressure of 1 atm and a starting temperature of 50 K, which was increased to 450 K at a rate of 50 K/ps. NPT conditions were enforced by Langevin dynamics, with an simulation timestep of 0.5 fs, a lattice degrees-of-freedom friction coefficient of 10 ps^{-1} and fictitious lattice mass of 10 amu. The Langevin friction coefficients for all atoms was set to 50 ps^{-1}, and deuterium mass was used for all H atoms. Integration over the first Brillouin zone was done using a single-point, Γ-centered k point grid and the hard versions of PBE pseudopotentials were used to model core electrons for O and H atoms. For calculations using the Grimme DFT–D3(BJ) and DFT–D4 dispersion corrections, the s_6, s_8, a_1, and a_2 parameters were taken from Ehlert et al.[67] For calculations using the rVV10 kernel, the b and c parameters were taken from Ning et al.[68] (for r^2SCAN) and Peng et al.[69] (for rSCAN, as an approximation from the SCAN + rVV10 parameters). Additional ML hyperparameters in the VASP ML model architecture were left at their default values. In total, eight ML models were produced, corresponding to the eight combinations of XC functional and dispersion correction.

To test the quality of the ML models, ML simulations were run on the same initial trajectory frame of 64 water molecules replicated 3 times in each dimension, for a total of 1728 water molecules. The ensemble, thermostat, timestep, and other MD settings were kept the same as in the active learning training above, except that hydrogen atoms were restored to a mass of

1.008 amu instead of using deuterium mass and system temperature was maintained at a constant 298.15 K. The simulation ran for a total of 20 ps in the NPT ensemble, comprised of a 10 ps equilibration period followed by a 10 ps production period. To evaluate the relative performance of each ML model, I compared each ML model's prediction of the O–O and O–H radial distribution functions, the water self-diffusion coefficient, and the evolution of system density.

Radial distribution functions. To evaluate water structure, O–O and O–H radial distribution functions (RDFs) were computed from the NPT trajectory using the implementation in the SolvationAnalysis[38] software package and appear in **Figure 10**. For computational tractability, RDFs were sampled every 40 frames (every 0.02 ps) from the production period. For comparison to experiment, I also include in the plot experimentally measured O–O and O–H RDFs. The O–O RDF is taken from Figure 2 of Pettersson and Takahashi[70], which in turn is computed from the X-ray diffraction experiments of Skinner et al.[71] The O–H RDF is taken from the X-ray diffraction data of Hura et al.[72] As a first-pass evaluation of each model's fidelity to experiment and to compensate for differences in system density, I scaled and shifted the experimental RDF curves to match the general positions of peaks and troughs in the RDF curves from my ML model results. All non-experimental RDF curves were normalized to have a value of 1 at the longest distances.

For the O–O RDFs, ML models trained using rSCAN + DFT–D4 and rSCAN with no dispersion correction appear to avoid some intermolecular understructuring in the troughs near 3.2 Å and 4.3 Å, while revPBE + DFT–D3(BJ) significantly overstructures the system at all distances. r^2SCAN models and rSCAN models with DFT–D3(BJ) or rVV10 corrections seem to systematically understructure water at distances below 5 Å, although they agree well with experiment and with each other at distances over 5 Å.

For the O–H RDFs, all ML models predict severe overstructuring compared to experiment at distances near 1 Å due to intramolecular O–H interactions being included in the RDF calculation. All models also predict overstructuring at the peak
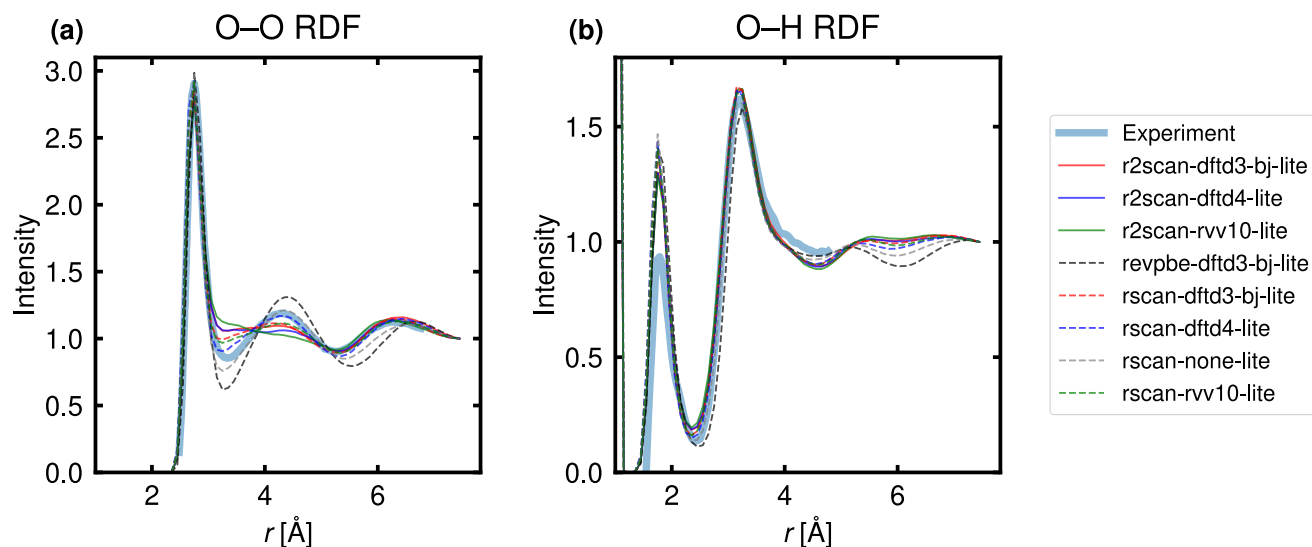

Figure 10. Calculated **(a)** O–O and **(b)** O–H RDFs for liquid water for all VASP ML models.

near 1.9 Å, but appear to more closely match the experimental RDF in the trough and peak at 2.4 Å and 3.2 Å, respectively. At longer distances (i.e., at 4.2 Å), it appears that revPBE + DFT–D3(BJ) and rSCAN with no dispersion correction most closely resemble the experimental RDF by not overstructuring the system, although the shape of the experimental curve at that distance is more closely matched by the three r^2SCAN models.

While these conclusions depend somewhat on the arbitrary way in which the experimental RDF curve was scaled and shifted, I generally see that no single ML model performs exceptionally well in reproducing the equilibrium structure of water. However, rSCAN models with DFT–D4 or without a dispersion correction may offer a good first approximation for the water structure at short-range distances (i.e., below 4.0 Å).

Self-diffusion coefficient. To evaluate how well each ML model predicts the dynamics of the water MD simulations, I computed the self-diffusion coefficients of water for each model. The self-diffusion coefficient was computed using the Einstein relation to the mean-squared displacement (MSD).[73] For each ML model, MSD curves were calculated over all particles and over the entire NPT trajectory (including both the equilibration and production periods). MSD curves were computed using the FFT implementation available in the MDAnalysis[74,75] software package and appear in **Figure 11**. Because the equilibration and production portions of the NPT trajectory were created in sequential but separate VASP jobs, a no-jump trajectory transformation[76] was applied to ensure undo automatic wrapping of atoms into the periodic cell boundary that would otherwise disrupt calculation of the MSD curve. **Figure 11** shows a subdiffusive region (approximately 0.02 ps to 0.06 ps), followed by a diffusive region beginning at approximately 10 ps for some curves. **Table 5** shows the computed self-diffusion coefficients, compared to an experimental value of 2.299×10^{-9} m^2/s (Holz et al. [77]). For each ML model, the water self-diffusion coefficient was determined by a linear fit to the corresponding curve over the approximate region in which that curve's log–log slope was $1.0 \pm 10\%$. All ML models have significant error in predicting water self-diffusivity, with revPBE + DFT–D3(BJ) the most egregious and r^2SCAN + DFT–D4 the most benign. While all models severely underpredict self-diffusivity, the r^2SCAN models still tend to outperform rSCAN and revPBE models, and within each r^2SCAN or rSCAN model class, DFT–D4 and rVV10 corrections outperform DFT–D3(BJ) corrections.

Correcting the severe underprediction of self-diffusivity is among the topics of my ongoing research. One possible cause
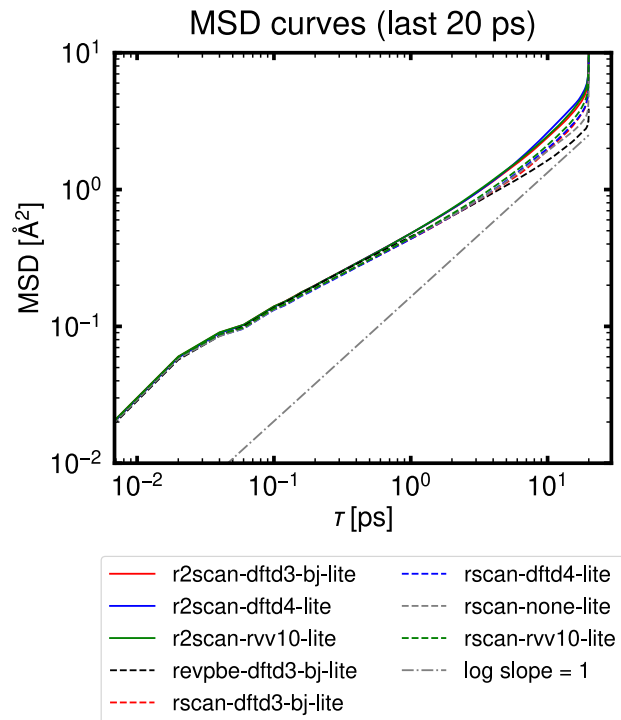

Figure 11. Computed MSD curves for all ML models. The gray dash-dotted line indicates a log–log slope of 1 for easier identification of the diffusive region.

could be that the ML models have not sufficiently learned repulsive forces that might facilitate greater transport of water and thus larger self-diffusion coefficients. Including trajectory frames from BOMD water simulations at higher temperatures might be necessary to allow the ML model to learn these additional interactions. Some studies include water simulated at temperatures up to 1000 K or even 2000 K.[78] While including training data at temperatures so high above the target temperature (298.15 K) is not recommended in the VASP documentation, I am investigating whether including a modest amount of high-temperature trajectory frames might improve prediction of water dynamics. To this end, I am currently extending the thermalization BOMD trajectory that serves as training data for the VASP ML models to higher temperatures: first to 600 K at NPT conditions, then to 800 K under NVT conditions (to avoid periodic cell deformation that happens at those higher temperatures).

Table 5. Computed water self-diffusion coefficient for each ML model, compared to experiment. t_{start} and t_{end} define the window over which linear fitting of the MSD curve is done to extract the self-diffusion coefficient via the Einstein relation. This window is the approximate window in which the log–log slope of the MSD curve is approximately 1.

Model	$\log_{10}(t_{\text{start}}/\text{ps})$	$\log_{10}(t_{\text{end}}/\text{ps})$	Diffusion coefficient [m^2/s]	Relative error
r^2SCAN + DFT–D3(BJ)	1.0	1.12	3.70×10^{-10}	–83.92%
r^2SCAN + DFT–D4	0.83	1.18	4.42×10^{-10}	–80.77%
r^2SCAN + rVV10	0.97	1.13	3.99×10^{-10}	–82.64%
revPBE + DFT–D3(BJ)	1.2	1.25	2.41×10^{-10}	–89.51%
rSCAN + DFT–D3(BJ)	0.96	1.16	2.97×10^{-10}	–87.07%
rSCAN + DFT–D4	1.06	1.19	3.31×10^{-10}	–85.58%
rSCAN (no correction)	1.19	1.23	2.72×10^{-10}	–88.19%
rSCAN + rVV10	1.0	1.15	3.35×10^{-10}	–85.45%

System density evolution. As an additional check of the stability and fidelity of the simulations, I computed the evolution of system density throughout the 20 ps of NPT simulation of bulk water under each ML model. **Figure 12** shows these computed densities as a function of simulation time.
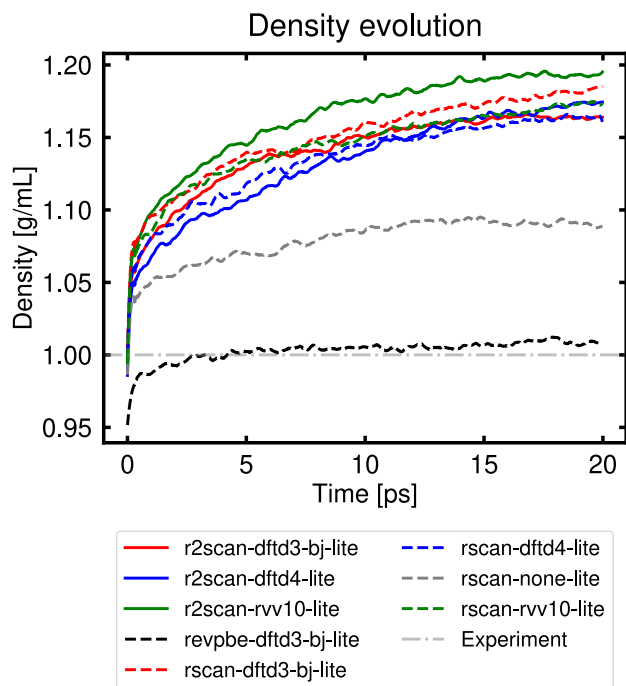

Figure 12. Computed system density over the entire 20 ps NPT trajectory at 298.15 K and 1 atm. The gray dash-dotted line shows the experimental density of 1 g/mL at 25 °C. Slight curve smoothing has been applied for readability.

The results are surprising given the relative performance of each ML model in other analyses. All ML models predict a sharp increase in density (collapse in box size) beyond what is expected for liquid water. r^2SCAN + rVV10 most overpredicts system density, followed by rSCAN with DFT–D3(BJ) or rVV10. r^2SCAN + DFT–D3(BJ) has the least severe density overprediction of the meta-GGA simulations that apply dispersion corrections. Surprisingly, revPBE + DFT–D3(BJ) maintains the density closest to experiment throughout all 20 ps, even as this model severely underperforms in predicting water structure and self-diffusivity.

This surprising ranking in model performance with respect to predicted system density is another reason I am investigating adding more high-temperature trajectory frames to each model. Including more repulsive forces may allow some of the dispersion-corrected models to approach the performance of revPBE + DFT–D3(BJ) while still maintaining their relatively superior prediction of water structure and self-diffusivity.

2.2 Accelerated BOMD using polarized atomic orbital optimization

Another current but nascent research effort involves a relatively new feature of the CP2K software package: the ability to create optimized, reduced basis sets representing the polarized atomic orbitals (PAOs) of atoms in a system. This framework, called PAO–ML, approaches the task of accelerated MD calculations from a perspective entirely different from the efforts described

previously. Most strategies to accelerate BOMD calculations using ML models focus on replacing the entire first-principles calculation (e.g., of electronic structure) with a surrogate calculator implemented as a ML model. The ML model is thus trained on atomic positions, atom types, and cell lattice vectors to directly and much more cheaply predict the system energy, atomic forces, and cell stresses than would normally result from a first-principles electronic structure calculation. In contrast, the PAO approach proposes accelerating rather than replacing DFT computation. This is accomplished by strategically reducing the basis set of functions representing the system's electrons such that the calculated electronic structure is substantially the same as that obtained during a normal DFT calculation, but at a much lower computational cost.[79] Accelerating BOMD calculations is then enabled by training a ML model to correctly but cheaply predict this reduced basis set for each frame's atomic positions, atom types, and cell lattice vectors (see **Figure 13**).

In theory, the PAO–ML approach is more versatile than traditional MLFFs because training a single ML model to predict the reduced basis sets immediately enables accelerated calculation of any system property that DFT can predict. Thus, a single ML model would enable quick DFT prediction of standard properties needed for MD, such as system energies, atomic forces, and cell stresses, as well as additional properties such as the Fermi energy or dipole moment. Additionally, PAO optimization may require less training data to become transferable to different chemical systems or different system geometries than would be required for a traditional MLFF.

In practice, however, the utility of the PAO–ML approach to my systems is unclear. The current implementation in CP2K requires the use of constrained Kohn–Sham equations, meaning that Mn and similar elements cannot receive proper spin-polarization treatment. Additionally, the PAO–ML framework requires multiple rounds of potentially expensive optimization. The first round of optimization produces the variationally optimized basis set that serves as the output labels for the ML model to be trained, and the second round of optimization is the training of the ML model itself to predict these variationally optimized basis sets. More significantly, training the ML model itself requires the user to define the loss function, for which the CP2K manual suggests using the variance in computed system energy or atomic forces when compared to the results one would get with a normal DFT calculation. But this approach encodes into the ML model a preference for greater accuracy of one or a few DFT-obtained properties over the rest. Such an approach undermines the claim that the PAO–ML approach produces a globally applicable DFT calculator that can be used to cheaply evaluate any DFT-computable property. Finally, the acceleration afforded by this method is unclear. While the claimed $50x$ acceleration of PAO–ML-enabled MD calculations (~ 0.05 ns/day) over BOMD is greater than the acceleration offered by VASP's ML models (~ 0.02 ns/day), it may still be much slower than a strategically trained NequIP/Allegro or DeePMD-kit model (~ 0.50 ns/day). While the PAO–ML approach is interesting, it currently seems less promising for modeling a $Li_2MnCl_4 \cdot x\,H_2O$ electrolyte and other similar systems of interest.
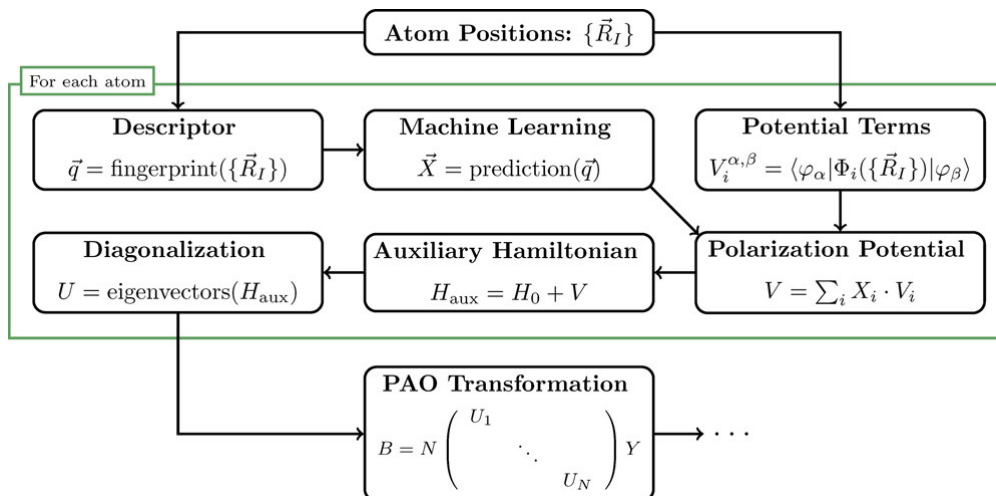
Figure 13. PAO–ML schematic, reproduced from Figure 1 in Schütt and VandeVondele [79]. The green box outlines the training task for the ML model: cheaply transforming an atomic geometry into a reduced basis set.

3 Conclusion

The first year of my postdoctoral appointment has been rewarding and enlightening. I have focused on learning the theory and practice of molecular dynamics simulations at length and time scales beyond those of the catalyst and surface science calculations that defined my graduate work. In the coming year, I hope to identify a ML-accelerated MD methodology that is successful in predicting the key bulk properties of liquid water, $LiCl \cdot x\,H_2O$, $Li_2ZnCl_4 \cdot x\,H_2O$, and ultimately $Li_2MnCl_4 \cdot x\,H_2O$. Results from accelerated MD calculations, coupled with experimental measurements of $Li_2MnCl_4 \cdot x\,H_2O$ electrolytes, will ultimately lead to one or more publications detailing design guidelines for these electrolytes and their evaluation as capable beyond-Li battery chemistries. I look forward to continuing this work and sharing further results in the next year of my tenure.

References

(1) Yang, C. et al. All-temperature zinc batteries with high-entropy aqueous electrolyte. *Nature Sustainability* **2023**, *6*, 325–335, DOI: 10.1038/s41893-022-01028-x.

(2) Song, Y.; Wang, L.; Sheng, L.; Ren, D.; Liang, H.; Li, Y.; Wang, A.; Zhang, H.; Xu, H.; He, X. The significance of mitigating crosstalk in lithium-ion batteries: a review. *Energy & Environmental Science* **2023**, *16*, 1943–1963, DOI: 10.1039/D3EE00441D.

(3) Tan, J.; Keiding, J. K. Mapping the cobalt and lithium supply chains for e-mobility transition: Significance of overseas investments and vertical integration in evaluating mineral supply risks. *Resources, Conservation and Recycling* **2024**, *209*, DOI: 10.1016/j.resconrec.2024.107788.

(4) Rentier, E. S.; Hoorn, C.; Seijmonsbergen, A. C. Lithium brine mining affects geodiversity and Sustainable Development Goals. *Renewable and Sustainable Energy Reviews* **2024**, *202*, DOI: 10.1016/j.rser.2024.114642.

(5) Idoine, N. E.; Raycraft, E. R.; Shaw, R. A.; Hobbs, S. F.; Deady, E. A.; Everett, P.; Evans, E. J.; Mills, A. J. *World Mineral Production: 2016–20*; Report, 2022.

(6) U.S. Geological Survey. Mineral Commodity Summaries 2024. **2024**, DOI: 10.3133/mcs2024.

(7) Rajput, N. N.; Seguin, T. J.; Wood, B. M.; Qu, X.; Persson, K. A. Elucidating Solvation Structures for Rational Design of Multivalent Electrolytes-A Review. *Top Curr Chem (Cham)* **2018**, *376*, 19, DOI: 10.1007/s41061-018-0195-2.

(8) Barnum, D. W. Hydrolysis of cations. Formation constants and standard free energies of formation of hydroxy complexes. *Inorganic Chemistry* **2002**, *22*, 2297–2305, DOI: 10.1021/ic00158a016.

(9) Sze, Y.-K.; Irish, D. E. Vibrational spectral studies of ion-ion and ion-solvent interactions. I. Zinc nitrate in water. *Journal of Solution Chemistry* **1978**, *7*, 395–415, DOI: 10.1007/BF00646111.

(10) Wang, F.; Borodin, O.; Gao, T.; Fan, X.; Sun, W.; Han, F.; Faraone, A.; Dura, J. A.; Xu, K.; Wang, C. Highly reversible zinc metal anode for aqueous batteries. *Nat Mater* **2018**, *17*, 543–549, DOI: 10.1038/s41563-018-0063-z.

(11) Pastel, G. R.; Pollard, T. P.; Liu, Q.; Lavan, S.; Zhu, Q.; Jiang, R.; Ma, L.; Connell, J.; Borodin, O.; Schroeder, M. A.; Zhang, Z.; Xu, K. Designing interphases for highly reversible aqueous zinc batteries. *Joule* **2024**, DOI: 10.1016/j.joule.2024.02.002.

(12) Ma, L.; Pollard, T. P.; Schroeder, M. A.; Luo, C.; Zhang, Y.; Pastel, G.; Cao, L.; Zhang, J.; Shipitsyn, V.; Yao, Y.; Wang, C.; Borodin, O.; Xu, K. Engineering a zinc anode interphasial chemistry for acidic, alkaline and non-aqueous electrolytes. *Energy & Environmental Science* **2024**, *17*, 2468–2479, DOI: 10.1039/D4EE00062E.

(13) Persson, K. A.; Waldwick, B.; Lazic, P.; Ceder, G. Prediction of solid-aqueous equilibria: Scheme to combine first-principles

calculations of solids with experimental aqueous states. *Physical Review B* **2012**, *85*, DOI: 10.1103/PhysRevB.85.235438.

(14) Patel, A. M.; Norskov, J. K.; Persson, K. A.; Montoya, J. H. Efficient Pourbaix diagrams of many-element compounds. *Phys Chem Chem Phys* **2019**, *21*, 25323–25327, DOI: 10.1039/C9CP04799A.

(15) Carlson, E. Z.; Chueh, W. C.; Mefford, J. T.; Bajdich, M. Selectivity of Electrochemical Ion Insertion into Manganese Dioxide Polymorphs. *ACS Appl Mater Interfaces* **2023**, *15*, 1513–1524, DOI: 10.1021/acsami.2c16589.

(16) Chen, C.; Shi, M.; Zhao, Y.; Yang, C.; Zhao, L.; Yan, C. Al-Intercalated MnO_2 cathode with reversible phase transition for aqueous Zn-Ion batteries. *Chemical Engineering Journal* **2021**, *422*, DOI: 10.1016/j.cej.2021.130375.

(17) Deng, S.; Wang, L.; Hou, T.; Li, Y. Two-Dimensional MnO_2 as a Better Cathode Material for Lithium Ion Batteries. *The Journal of Physical Chemistry C* **2015**, *119*, 28783–28788, DOI: 10.1021/acs.jpcc.5b10354.

(18) Wang, D.; Wang, L.; Liang, G.; Li, H.; Liu, Z.; Tang, Z.; Liang, J.; Zhi, C. A Superior delta-MnO_2 Cathode and a Self-Healing Zn-delta-MnO_2 Battery. *ACS Nano* **2019**, *13*, 10643–10652, DOI: 10.1021/acsnano.9b04916.

(19) Zhang, Y.; Liu, Y.; Liu, Z.; Wu, X.; Wen, Y.; Chen, H.; Ni, X.; Liu, G.; Huang, J.; Peng, S. MnO_2 cathode materials with the improved stability via nitrogen doping for aqueous zinc-ion batteries. *Journal of Energy Chemistry* **2022**, *64*, 23–32, DOI: 10.1016/j.jechem.2021.04.046.

(20) Schmidt, M. *Anorganische Chemie II*; Wissenschaftsverlag, 1968; pp 100–109.

(21) Deringer, V. L.; Caro, M. A.; Csanyi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv Mater* **2019**, *31*, e1902765, DOI: 10.1002/adma.201902765.

(22) Martinez, L.; Andrade, R.; Birgin, E. G.; Martinez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J Comput Chem* **2009**, *30*, 2157–64, DOI: 10.1002/jcc.21224.

(23) Binninger, T.; Sarac, D.; Marsh, L.; Picard, T.; Doublet, M. L.; Raynaud, C. AMOEBA Polarizable Force Field for Molecular Dynamics Simulations of Glyme Solvents. *J Chem Theory Comput* **2023**, *19*, 1023–1034, DOI: 10.1021/acs.jctc.2c00926.

(24) Adjoua, O.; Lagardere, L.; Jolly, L. H.; Durocher, A.; Very, T.; Dupays, I.; Wang, Z.; Inizan, T. J.; Celerse, F.; Ren, P.; Ponder, J. W.; Piquemal, J. P. Tinker-HP: Accelerating Molecular Dynamics Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields Using GPUs and Multi-GPU Systems. *J Chem Theory Comput* **2021**, *17*, 2034–2053, DOI: 10.1021/acs.jctc.0c01164.

(25) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics* **1992**, *97*, 1990–2001, DOI: 10.1063/1.463137.

(26) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **1965**, *140*, A1133–A1138, DOI: 10.1103/PhysRev.140.A1133.

(27) Kresse, G.; Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys Rev B Condens Matter* **1993**, *47*, 558–561, DOI: 10.1103/physrevb.47.558.

(28) Kresse, G.; Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys Rev B Condens Matter* **1994**, *49*, 14251–14269, DOI: 10.1103/physrevb.49.14251.

(29) Kresse, G.; Hafner, J. Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements. *Journal of*

Physics: Condensed Matter **1994**, *6*, 8245–8257, DOI: 10.1088/0953-8984/6/40/015.

(30) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* **1996**, *6*, 15–50, DOI: 10.1016/0927-0256(96)00008-0.

(31) Kresse, G.; Furthmuller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B Condens Matter* **1996**, *54*, 11169–11186, DOI: 10.1103/physrevb.54.11169.

(32) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Simple". *Physical Review Letters* **1998**, *80*, 890–890, DOI: 10.1103/PhysRevLett.80.890.

(33) 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* **2010**, *132*, 154104, DOI: 10.1063/1.3382344.

(34) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J Comput Chem* **2011**, *32*, 1456–65, DOI: 10.1002/jcc.21759.

(35) Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* **1967**, *159*, 98–103, DOI: 10.1103/PhysRev.159.98.

(36) Gowers, R.; Linke, M.; Barnoud, J.; Reddy, T.; Melo, M.; Seyler, S.; Domański, J.; Dotson, D.; Buchoux, S.; Kenney, I.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. Proceedings of the Python in Science Conference. DOI: 10.25080/Majora-629e541a-00e.

(37) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **2011**, *32*, 2319–27, DOI: 10.1002/jcc.21787.

(38) Cohen, O. A.; Macdermott-Opeskin, H.; Lee, L.; Hou, T.; Fong, K. D.; Kingsbury, R.; Wang, J.; Persson, K. A. SolvationAnalysis: A Python toolkit for understanding liquid solvation structure in classical molecular dynamics simulations. *Journal of Open Source Software* **2023**, *8*, DOI: 10.21105/joss.05183.

(39) Lennard-Jones, J. E. Cohesion. *Proceedings of the Physical Society* **1931**, *43*, 461–482, DOI: 10.1088/0959-5309/43/5/301.

(40) Back, S.; Tran, K.; Ulissi, Z. W. Discovery of Acid-Stable Oxygen Evolution Catalysts: High-Throughput Computational Screening of Equimolar Bimetallic Oxides. *ACS Appl Mater Interfaces* **2020**, *12*, 38256–38265, DOI: 10.1021/acsami.0c11821.

(41) Chanussot, L. et al. Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catalysis* **2021**, *11*, 6059–6072, DOI: 10.1021/acscatal.0c04525.

(42) Gu, G. H.; Noh, J.; Kim, S.; Back, S.; Ulissi, Z.; Jung, Y. Practical Deep-Learning Representation for Fast Heterogeneous Catalyst Screening. *J Phys Chem Lett* **2020**, *11*, 3185–3191, DOI: 10.1021/acs.jpclett.0c00634.

(43) Sunshine, E. M.; Shuaibi, M.; Ulissi, Z. W.; Kitchin, J. R. Chemical Properties from Graph Neural Network-Predicted Electron Densities. *The Journal of Physical Chemistry C* **2023**, *127*, 23459–23466, DOI: 10.1021/acs.jpcc.3c06157.

(44) Tran, R.; Wang, D.; Kingsbury, R.; Palizhati, A.; Persson, K. A.; Jain, A.; Ulissi, Z. W. Screening of bimetallic electrocatalysts for water purification with machine learning. *J Chem Phys* **2022**, *157*, 074102, DOI: 10.1063/5.0092948.

(45) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accu-

rate interatomic potentials. *Nat Commun* **2022**, *13*, 2453, DOI: 10.1038/s41467-022-29939-5.

(46) Kozinsky, B.; Musaelian, A.; Johansson, A.; Batzner, S. Scaling the Leading Accuracy of Deep Equivariant Models to Biomolecular Simulations of Realistic Size. 2023; https://doi.org/10.1145/3581784.3627041.

(47) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys Rev Lett* **2018**, *120*, 145301, DOI: 10.1103/PhysRevLett.120.145301.

(48) Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C. J.; Kornbluth, M.; Kozinsky, B. Learning local equivariant representations for large-scale atomistic dynamics. *Nat Commun* **2023**, *14*, 579, DOI: 10.1038/s41467-023-36329-y.

(49) Ziegler, J. F.; Biersack, J. P. In *Treatise on Heavy-Ion Science: Volume 6: Astrophysics, Chemistry, and Condensed Matter*; Bromley, D. A., Ed.; Springer US: Boston, MA, 1985; pp 93–129, DOI: 10.1007/978-1-4615-8103-1_3.

(50) Cheng, B.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. Ab initio thermodynamics of liquid and solid water. *Proc Natl Acad Sci U S A* **2019**, *116*, 1110–1115, DOI: 10.1073/pnas.1815117116.

(51) Kuhne, T. D. et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J Chem Phys* **2020**, *152*, 194103, DOI: doi.org/10.1063/5.0007045.

(52) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *Journal of Computational Chemistry* **1999**, *20*, 786–798, DOI: 10.1002/(SICI)1096-987X(199906)20:8<786::AID-JCC5>3.0.CO;2-B.

(53) Lagardere, L.; Aviat, F.; Piquemal, J. P. Pushing the Limits of Multiple-Time-Step Strategies for Polarizable Point Dipole Molecular Dynamics. *J Phys Chem Lett* **2019**, *10*, 2593–2599, DOI: 10.1021/acs.jpclett.9b00901.

(54) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications* **2018**, *228*, 178–184, DOI: 10.1016/j.cpc.2018.03.016.

(55) Zeng, J. et al. DeePMD-kit v2: A software package for deep potential models. *J Chem Phys* **2023**, *159*, DOI: 10.1063/5.0155600.

(56) Jinnouchi, R.; Karsai, F.; Kresse, G. On-the-fly machine learning force field generation: Application to melting points. *Physical Review B* **2019**, *100*, DOI: 10.1103/PhysRevB.100.014105.

(57) 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* **2019**, *122*, 225701, DOI: 10.1103/PhysRevLett.122.225701.

(58) 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* **2020**, *152*, 234102, DOI: 10.1063/5.0009491.

(59) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys Rev Lett* **2018**, *120*, 143001, DOI: 10.1103/PhysRevLett.120.143001.

(60) Wang, X.; Wang, Y.; Zhang, L.; Dai, F.; Wang, H. A tungsten deep neural-network potential for simulating mechanical property

degradation under fusion service environment. *Nuclear Fusion* **2022**, *62*, DOI: 10.1088/1741-4326/ac888b.

(61) Zhang, L.; Han, J.; Wang, H.; Saidi, W. A.; Car, R.; Weinan, E. End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems* **2018**, 4441–4451, DOI: 10.5555/3327345.3327356.

(62) Zhang, D.; Bi, H.; Dai, F.-Z.; Jiang, W.; Liu, X.; Zhang, L.; Wang, H. Pretraining of attention-based deep learning potential model for molecular simulation. *npj Computational Materials* **2024**, *10*, DOI: 10.1038/s41524-024-01278-7.

(63) Zhang, L.; Wang, H.; Muniz, M. C.; Panagiotopoulos, A. Z.; Car, R.; E, W. A deep potential model with long-range electrostatic interactions. *J Chem Phys* **2022**, *156*, 124107, DOI: 10.1063/5.0083669.

(64) Maxson, T.; Szilvasi, T. Transferable Water Potentials Using Equivariant Neural Networks. *J Phys Chem Lett* **2024**, *15*, 3740–3747, DOI: 10.1021/acs.jpclett.4c00605.

(65) Loose, T. D.; Sahrmann, P. G.; Qu, T. S.; Voth, G. A. Coarse-Graining with Equivariant Neural Networks: A Path Toward Accurate and Data-Efficient Models. *J Phys Chem B* **2023**, *127*, 10564–10572, DOI: 10.1021/acs.jpcb.3c05928.

(66) Settles, B. *Active Learning*; Synthesis lectures on artificial intelligence and machine learning; Springer International Publishing: Cham, 2012; DOI: 10.1007/978-3-031-01560-1.

(67) Ehlert, S.; Huniar, U.; Ning, J.; Furness, J. W.; Sun, J.; Kaplan, A. D.; Perdew, J. P.; Brandenburg, J. G. r(2)SCAN-D4: Dispersion corrected meta-generalized gradient approximation for general chemical applications. *J Chem Phys* **2021**, *154*, 061101, DOI: 10.1063/5.0041008.

(68) Ning, J.; Kothakonda, M.; Furness, J. W.; Kaplan, A. D.; Ehlert, S.; Brandenburg, J. G.; Perdew, J. P.; Sun, J. Workhorse minimally empirical dispersion-corrected density functional with tests for weakly bound systems: r2SCAN+rVV10. *Physical Review B* **2022**, *106*, DOI: 10.1103/PhysRevB.106.075422.

(69) Peng, H.; Yang, Z.-H.; Perdew, J. P.; Sun, J. Versatile van der Waals Density Functional Based on a Meta-Generalized Gradient Approximation. *Physical Review X* **2016**, *6*, DOI: 10.1103/PhysRevX.6.041005.

(70) Pettersson, L. G. M.; Takahashi, O. The local structure of water from combining diffraction and X-ray spectroscopy. *Journal of Non-Crystalline Solids: X* **2022**, *14*, DOI: 10.1016/j.nocx.2022.100087.

(71) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A.; Benmore, C. J. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J Chem Phys* **2013**, *138*, 074506, DOI: 10.1063/1.4790861.

(72) Hura, G.; Russo, D.; Glaeser, R. M.; Head-Gordon, T.; Krack, M.; Parrinello, M. Water structure as a function of temperature from X-ray scattering experiments and ab initio molecular dynamics. *Physical Chemistry Chemical Physics* **2003**, *5*, DOI: 10.1039/B301481A.

(73) Maginn, E. J.; Messerly, R. A.; Carlson, D. J.; Roe, D. R.; Elliot, J. R. Best Practices for Computing Transport Properties 1. Self-Diffusivity and Viscosity from Equilibrium Molecular Dynamics. *Living Journal of Computational Molecular Science* **2020**, *2*, DOI: 10.33011/livecoms.1.1.6324.

(74) Calandrini, V.; Pellegrini, E.; Calligari, P.; Hinsen, K.; Kneller, G. R. nMoldyn - Interfacing spectroscopic experiments, molecular dynamics simulations and models for time correlation

functions. *École thématique de la Société Française de la Neutronique* **2011**, *12*, 201–232, DOI: `10.1051/sfn/201112010`.

(75) de Buyl, P. tidynamics: A tiny package to compute the dynamics of stochastic and molecular simulations. *Journal of Open Source Software* **2018**, *3*, DOI: `10.21105/joss.00877`.

(76) Kulke, M.; Vermaas, J. V. Reversible Unwrapping Algorithm for Constant-Pressure Molecular Dynamics Simulations. *J Chem Theory Comput* **2022**, *18*, 6161–6171, DOI: `10.1021/acs.jctc.2c00327`.

(77) Holz, M.; Heil, S. R.; Sacco, A. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate 1H NMR PFG measurements. *Physical Chemistry Chemical Physics* **2000**, *2*, 4740–4742, DOI: `10.1039/B005319H`.

(78) Wang, C.; Tian, W.; Zhou, K. Ab Initio Simulation of Liquid Water without Artificial High Temperature. *J Chem Theory Comput* **2024**, DOI: `10.1021/acs.jctc.4c00650`.

(79) Schütt, O.; VandeVondele, J. Machine Learning Adaptive Basis Sets for Efficient Large Scale Density Functional Theory Simulation. *J Chem Theory Comput* **2018**, *14*, 4168–4175, DOI: `10.1021/acs.jctc.8b00378`.

■