JCTC Journal of Chemical Theory and Computation

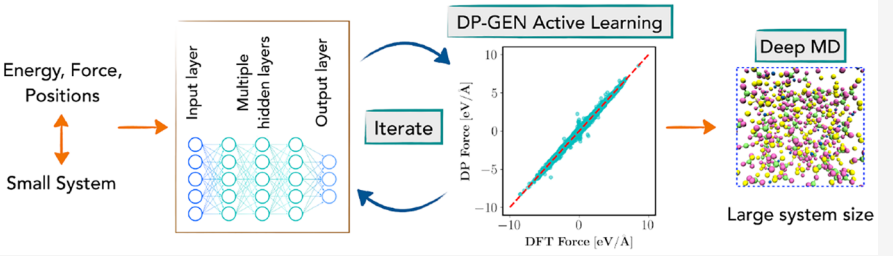
pubs.acs.org/JCTC Article

# Modeling Chemical Reactions in Alkali Carbonate−Hydroxide Electrolytes with Deep Learning Potentials

Anirban Mondal,*[,¶] Dina Kussainova,[¶] Shuwen Yue, and Athanassios Z. Panagiotopoulos*

Cite This: https://doi.org/10.1021/acs.jctc.2c00816 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information




**ABSTRACT:** We developed a deep potential machine learning model for simulations of chemical reactions in molten alkali carbonate-hydroxide electrolyte containing dissolved $CO_2$, using an active learning procedure. We tested the deep neural network (DNN) potential and training procedure against reaction kinetics, chemical composition, and diffusion coefficients obtained from density functional theory (DFT) molecular dynamics calculations. The DNN potential was found to match DFT results for the structural, transport, and short-time chemical reactions in the melt. Using the DNN potential, we extended the time scales of observation to 2 ns in systems containing thousands of atoms, while preserving quantum chemical accuracy. This allowed us to reach chemical equilibrium with respect to several chemical species in the melt. The approach can be generalized for a broad spectrum of chemically reactive systems.



## 1. INTRODUCTION

Molten carbonate salts are a relatively extended class of molecular ionic compounds that are receiving increasing attention in fundamental and applied fields. They are considered in the development of technological devices (e.g., fuel cells)[1−3] as well as for providing an improved understanding of geochemical processes (e.g., the role of carbonate melts in the geodynamics of the Earth's mantle).[4,5] Their high ionic conductivity, low viscosity, chemical stability, and low environmental impact make molten carbonates essential in high-temperature technology, with applications in heat transport and thermal energy storage. Perhaps the most promising set of potential applications are molten carbonate fuel cells (MCFCs)[1] and high capacity rechargeable batteries.[6] In MCFCs, power generation is coupled with $CO_2$ capture which enhances the efficiency and environmental appeal of these devices. As the solvent, the electrolyte helps stabilize intermediates for the electrochemical reactions taking place in MCFCs. The electrolyte consists of a mixture of $Li^+/K^+$ or $Li^+/Na^+$ carbonates near the eutectic composition, and the cathode feed gas contains $CO_2$, $H_2O$, and $O_2$. As a result of chemical reactions between carbonate and dissolved species, the high-temperature ionic melt consists of additional molecular or ionic moieties[7,8] besides the dominant carbonate ions. Previous studies[9−11] have confirmed the presence of significant concentrations of hydroxide ions, which can affect the carbon capture efficiency of fuel cells as well as cause electrolyte loss.[12,13]

As a result of high operating temperatures (above 600 °C) of molten carbonate fuel cells, direct characterization of the electrolyte structure and properties is experimentally challenging, requiring special instruments usually adapted to the measurement of a single observable. Computational approaches can potentially overcome these issues and be used for reliable predictions of the physicochemical properties of molten salts. However, representing the potential energy surfaces of molten salts poses interesting challenges considering the complexities in modeling atomic/molecular structures, multiple phases, and transformations that can significantly influence the salt properties.[14] In the past few decades, several classical interatomic potentials have been employed in modeling molecular simulations of molten carbonates, where such potentials have been developed using *ab initio* force-fitting methods.[15−22] Assumptions about the relevant interactions and an explicit definition of the potential energy functional form

**Special Issue:** Machine Learning for Molecular Simulation

**Received:** August 8, 2022

are often essential. Although classical models have been successfully applied to compute thermodynamic and transport properties of pure and mixtures of salts,[23] these potentials are subject to limitations of accuracy and transferability. Furthermore, most classical force fields cannot capture effects such as charge transfer, speciation, and chemical reactivity. An attractive alternative to this approach is *ab initio* molecular dynamics (AIMD), where forces acting on the atoms are computed using an explicit description of the electrons via quantum chemistry methods, e.g., density functional theory (DFT). AIMD simulations are thus free of potential parametrization and offer the advantage of transferability and describing chemistry, specifically bond breaking and forming. For example, by employing AIMD simulations, Corradini et al. reported the behavior of $CO_2$ in molten $CaCO_3$ where they observed the formation of a pyrocarbonate anion via a spontaneous reaction between $CO_2$ and a carbonate anion.[24] Recently, we investigated an MCFC electrolyte consisting of $CO_2$ and molten carbonate-hydroxide via an AIMD approach.[25] We observed formation of several ionic species, including pyrocarbonate, bicarbonate, etc., primarily driven by the gas-phase concentrations of $CO_2$, $H_2O$, and $O_2$.[7,8] Our results demonstrated the presence of water at concentrations double or more than that of $CO_2$. However, AIMD simulations are computationally expensive and are limited with respect to both system size (a few hundred atoms) and time scales (picoseconds). Therefore, many practically and physically relevant properties of molten carbonates, such as the diffusivity, viscosity, ionic conductivity, and equilibrium composition of a reactive mixture, cannot be computed accurately.

In addressing the trade-off between accuracy and computational expense, machine learning (ML)-based potentials are a promising solution for development of next-generation molecular simulation force fields.[26,27] Among the most common approaches are high-dimensional neural networks (NNs)[26,28−36] and kernel-based methods[27,37−40] trained on potential energy surfaces generated from accurate quantum chemical methods such as DFT. Most recently, the introduction of equivariant NNs[41−44] has led to highly data-efficient methods of potential training. The great advantage of ML-based potentials is their ability to combine the efficiency comparable to that of classical force fields and the accuracy of high-level *ab initio* methods. The past decade has seen a significant rise in the development of ML potentials, covering a broad spectrum of applications.[28−35,45] Jackson and co-workers developed an automated workflow to obtain ML potentials for molten salts and applied the derived potential to examine the structural and dynamic properties of molten LiCl.[46] Li et al.[47] and Tovey et al.[48] have also developed neural network potentials for molten NaCl that account for many-body polarization. Lam et al.[49,50] have demonstrated the efficiency of atom-centered neural network interaction potentials in predicting the thermodynamic and transport properties of LiF and FLiBe molten salts. A comparative study of the structural and transport properties of a prototypical molten salt, LiF-NaF-KF (FLiNaK), using machine-learned NN potential and a reparametrized classical force field was reported by Lee et al.[51] They showed the ability of the former potential to predict structural and dynamic properties of the molten salt with quantum chemical accuracy and computational efficiency of classical force fields. Nguyen et al.[52] have also adopted an ML approach to study the thermophysical properties of actinide molten salts.

Despite these successes, the major obstacle in applying ML potentials to the broader chemical space of molten salts, including molten carbonate and hydroxides, is the difficulty of potential parametrization. In ML potential development, there must be a balance between the computational cost of training set generation with the need to sample diverse atomic configurations in order to optimize ML parameters to obtain models of adequate predictive accuracy and transferability. To the best of our knowledge, except for the very recent report by Feng et al.,[53] no other study is reported in the literature devoted to developing ML-based models for molten carbonates and hydroxides. In ref 53, an ML-based model was employed to investigate the local structure and transport properties of molten $(Li_nK_{1-n})_2CO_3$ $[n = 0.4, 0.5, 0.6]$ binary salts. Chemical reactions involving $CO_2$ in the electrolyte melt were not investigated; these reactions are the main focus of our present work.

Specifically, we aim to construct and apply an ML-based model to study chemical reactions and characterize the formation/dissociation of different chemical species in a molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$. In addition, our goal in this work is to investigate the physicochemical properties of molten carbonates, hydroxides, and their mixtures. Toward this purpose, we trained a potential using the Deep Potential Molecular Dynamics (DPMD)[54] method which relies on deep neural networks (DNNs) to fit and evaluate high dimensional potential energy surfaces. DPMD simulations have been performed to model millions of atoms on time scales of tens of nanoseconds[55] and were successfully applied to study liquid−liquid phase transition in supercooled water,[56] chemical reactions at the interface of water and $TiO_2$,[57] gas-phase reactive systems,[58] in modeling metallic systems,[33] etc. with an accuracy matching to that of underlying *ab initio* models. The accuracy and transferability of DPMD models have been improved by including the Deep Potential GENerator (DP-GEN) scheme.[59] This approach involves an active learning algorithm to generate models and aims to reduce the overall computational cost for data generation and model training.

The article is organized as follows: In Section 2, we discuss the model development and assessment methods. The analyses of model training errors, results obtained, and overall performance of the potential in predicting thermodynamic properties are presented in Section 3. The last section provides a summary of the work.

## 2. METHODS

**2.1. *Ab Initio* Data Set Generation.** Training a robust DNN model capable of capturing a range of local atomic environments requires a data set that sufficiently covers the breadth of configurations that may manifest during finite-temperature MD simulations. As such, data for molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$ configurations were acquired from previously obtained AIMD data with simulation details as described in ref 25. Briefly, the first-principles calculations were performed at the DFT level[60,61] as implemented in the CP2K simulation program[62] using the Quickstep module.[63] Exchange-correlation potentials were treated within the generalized gradient approximation (GGA) employing the Perdew−Burke−Ernzer-


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

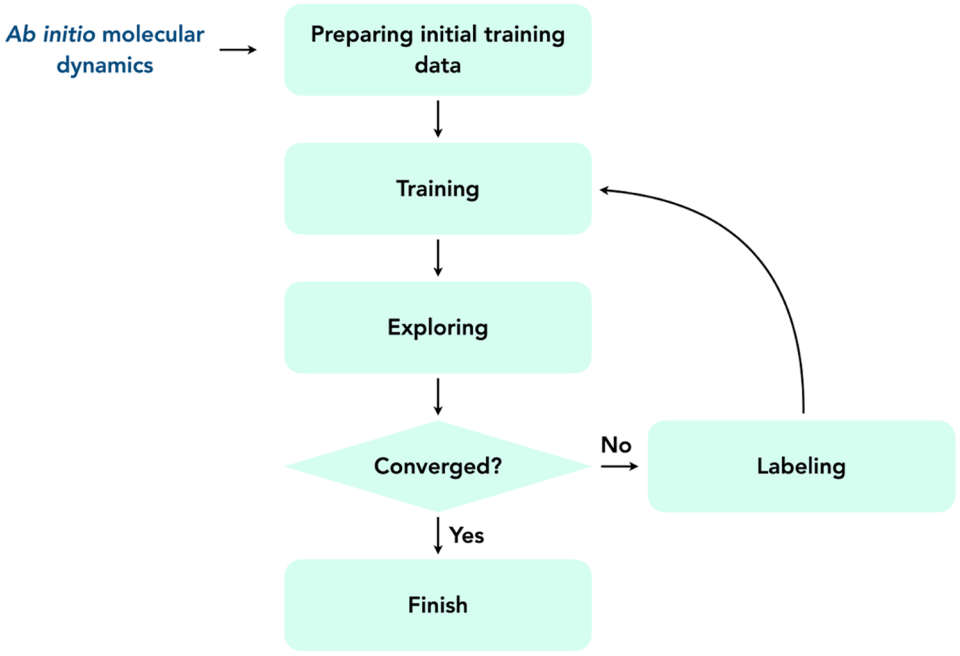



**Figure 1.** Flowchart of the training procedure of the DNN model using the active learning approach.

hof (PBE) functional.[64] The GGA formalism has been shown to yield reasonably accurate results for physicochemical properties of molten salts.[46−52] A double-$\zeta$ valence plus polarization (DZVP) basis set was adopted to expand the wave functions. The energy cutoff was set to 400 Ry for the auxiliary plane wave expansion of the charge density. Valence electrons were modeled explicitly, whereas core electrons have been treated with norm-conserving Goedecker−Teter−Hutter (GTH) pseudopotentials[65,66] with 1, 3, 4, 6, and 9 valence electrons for H, Li, C, O, and K, respectively. Dispersive interaction corrections were included by utilizing the empirical dispersion correction (D3) from Grimme,[67] with a cutoff of 40 Å. Born−Oppenheimer MD simulations were performed with a time step of 0.5 fs to integrate the equations of motion. The initial system was equilibrated for 8 ps in the isothermal−isobaric ($NpT$) ensemble at 1 bar using an isotropic unit cell according to the scheme of Martyna et al.,[68] with a barostat time constant of 250 fs. The simulation cell size after the $NpT$ equilibration was 21.773 Å. Equilibration was followed by a production period of 65 ps in the canonical ($NVT$) ensemble. We set the temperature at 923.15 K, controlled by a chain of six Nosé−Hoover thermostats[69] with a time constant of 100 fs. Three-dimensional periodic boundary conditions were employed in all simulations. We extracted ∼6000 configurations of the $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$ from the AIMD trajectory where each configuration consisted of 630 atoms [120 $Li^+$, 80 $K^+$, 80 $CO_3^{2-}$, 40 $OH^-$, and 10 $CO_2$].

**2.2. Deep Neural Network Model Development.** We employed the DP-GEN active machine learning approach of Zhang et al.[59] to construct a DNN model for our system. The procedure consists of multiple iterations as displayed in Figure 1 and is summarized as follows.

**Step 0: Preparing Initial Training Data.** The AIMD-generated configurations were subjected to single point energy/force calculations within the DFT framework at the same level of theory as described above. The gradients on the wave functions were optimized with convergence criteria of $10^{-6}$ a.u. Because of the large supercells, the integration of the Brillouin zone was performed with a reciprocal space mesh consisting of only the gamma point. The energy $E$ and force $\mathbf{F}_i$ on each atom $i$ together with the atomic positions and box dimension of these snapshots were used as the initial training data.

**Step 1: Training.** We employed the DeepMD-kit package[70] to train four independent DNN models based on the same input training data, but different random seeds were utilized in initialization of the neural network parameters. In the DPMD approach, for every snapshot in the training set, a local coordinate frame is set up for each atom $i$ with respect to its neighbors within a cutoff distance $r_c$. In this work, $r_c$ was set to 6 Å. This treatment accounts for the preservation of translational and rotational symmetries.[70] Summation over all possible permutations of atoms of the same type was performed to preserve the permutational invariance. These representations were fed to a DNN as inputs to yield atomic energies $E_i$ for each atom $i$, so that the total potential energy of a given configuration is $E = \sum_i E_i$.[70] Atomic forces $\mathbf{F}_i$ were computed from the gradient of the system energy. Then, parameters of the DNN model were optimized by minimizing the loss function $L$, defined as a sum of squared errors normalized by the number of atoms,

$$L = p_e \Delta E^2 + \frac{p_f}{3N} \sum_i \Delta \mathbf{F}_i^2 \quad (1)$$

where $\Delta E^2$ and $\Delta \mathbf{F}_i^2$ represent the squared errors of system energy and atomic component-wise force, respectively, between the training data and the current DPMD prediction, and $N$ is the number of atoms. To optimize the training efficiency, the terms of the loss function in eq 1 were weighted with prefactors $p_e$ and $p_f$ where $p_e$ was progressively increased from 0.2 to 8, while $p_f$ was progressively decreased from 1000 to 1. The number of training steps was set to $1.5 \times 10^6$. Throughout the training, we employed three layers of NN, where each layer consisted of 240 neurons.


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

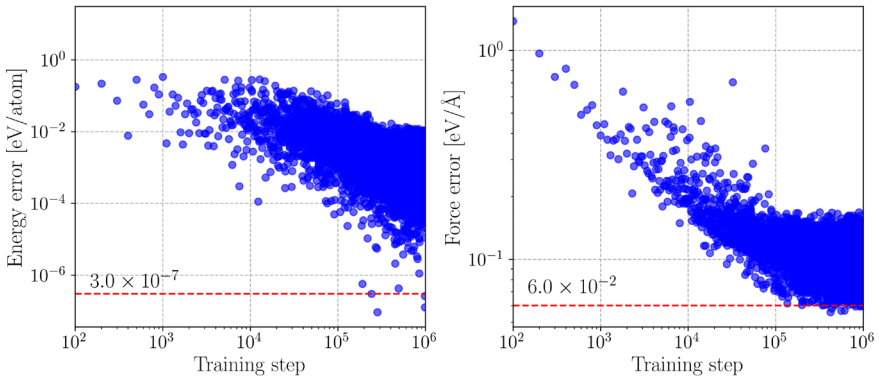



**Figure 2.** Learning curves from training the DPMD model. Left panel: system energies. Right panel: atomic forces.

**Step 2: Exploring.** To systematically improve the DNN model, additional training data was collected via iterative exploration of configurational space with short DPMD simulations at different thermodynamic conditions. These simulations were performed in the *NpT* ensemble considering a range of initial configurations. Molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$ was simulated at a pressure of 1 bar and temperatures $T$ = 850, 900, 925, 950, 975, and 1225 K. The pure carbonate systems, $Li_2CO_3$ and $K_2CO_3$, were simulated at 1 bar and $T$ = 1025 and 1200 K, whereas pure hydroxide systems, LiOH and KOH, were simulated at 1 bar and $T$ ranging from 825 to 1075 K with an interval of 50 K. In these simulations, all the trajectories were propagated using one of the DNN models randomly selected from the four models produced in Step 1, while the force acting on each atom at each time step was evaluated by employing all four models. The metric for convergence was the maximum standard deviation of the predicted atomic forces, calculated as $\zeta = \max_i \sqrt{\langle \|\mathbf{F}_i - \bar{\mathbf{F}}_i\|^2 \rangle}$ for each snapshot, where $\bar{\mathbf{F}}_i = \langle \mathbf{F}_i \rangle$ is the average force on atom $i$ predicted by the four different DNN models.

**Step 3: Labeling.** The snapshots that had $\zeta > 0.1$ eV/Å were extracted. The energy $E$ and force $\mathbf{F}_i$ on each atom $i$ of these snapshots were evaluated within the DFT framework and added to the training data set. Steps 1 through 3 were then repeated until the configurations that had $\zeta > 0.1$ eV/Å accounted for less than 0.007% of the total configurations. At the end of the DP-GEN process consisting of 28 iterations, our training data set contained 15 840 structures, spanning a temperature range of 825−1225 K.

*2.2.1. Targeted Learning.* We employed the DNN model generated at the end of the DP-GEN active learning exercise to perform DPMD simulations using the LAMMPS program.[71] A detailed description of such simulations is described in the following subsection. Specifically, we considered molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$, an identical system as discussed in Section 2.1. However, we observed some nonphysical results upon analyzing the trajectories. We observed a discrepancy between the AIMD and DPMD predictions in the percentage of C−O covalent bond retained in $CO_2$ molecules. All $CO_2$ molecules are found to be intact throughout the AIMD trajectory. In contrast, almost 4% of the C−O covalent bonds in $CO_2$ molecules were broken within 80 ps in the DPMD simulations using a DNN model generated at the end of the DP-GEN active learning cycle. This suggests a shortcoming of the initially trained DNN potential in accurately capturing the underlying chemistry of the system as observed in the reference AIMD simulations. The observed discrepancy is attributed to the inability of the active learning procedure to access configurations representative of the C−O bond-breaking process. On the infrequent occasions of breaking C−O bonds in the AIMD simulations, this is a fast process occurring on a time scale of ∼20 fs. Therefore, it is difficult for our active learning process, which samples configurations every 0.1 ps, to capture snapshots of this very fast, rare process. The bond-breaking reaction coordinate is not sampled comprehensively. As a result, our model is trained on mostly intact $CO_2$ and fails to capture the accurate time scale, frequency, and energetic barriers of C−O bond breaking. The original model seems to underestimate the barriers of C−O bond breaking.

In order to address this deficiency in the original trained model, we performed another round of training with additional configurations (not obtained from active learning) where we manually fixed C−O bonds over a range of distances to systematically and homogeneously sample the bond dissociation pathway. In this "targeted learning" approach, we considered all the training data sets generated via the active learning process plus the initial training data. Specifically, we took a system identical to the AIMD simulation (see Section 2.1) consisting of 630 atoms, including 10 $CO_2$ molecules. New configurations were created by fixing the C−O bond lengths in $CO_2$ molecules from 1.0 to 2.0 Å, where bonds were manually stretched in a step size of 0.2 Å. After the distances were fixed, equilibrium MD was performed on all molecules except for the manually fixed C−O bond of interest; this bond length is fixed while all other molecules are equilibrated to ensure we are sampling this bond length. The newly constructed snapshots were equilibrated using classical potential[22] and GROMACS MD engine.[72,73] These equilibrium simulations were performed in the *NpT* ensemble for a short duration of 500 ps using standard MD protocols. The equilibrated snapshots were used to obtain energies and atomic forces within a DFT framework employing the same parameters as described in Section 2.1. This strategy resulted in a total of 1350 configurations with energies and atomic forces closely resembling the potential energy surface of $CO_2$ dissociation and was added to the training data set. We then retrained the model based on the updated training data set. The number of training steps was set to $1 \times 10^6$. All other training parameters were kept similar as described in Section 2.2. The DNN potential obtained at the end of this targeted


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

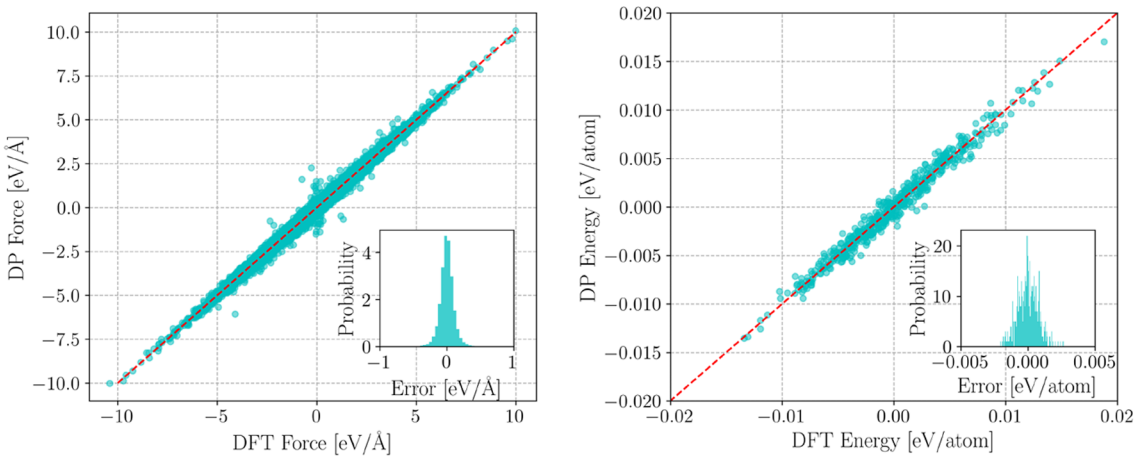



**Figure 3.** Forces and energies from the DFT calculations and DPMD model. Left: atomic forces. Right: potential energy per atom. Insets in both plots are the probability distributions of the absolute difference in force and energy between the two methods.

learning is the final model used for all calculations from this point forward.

**2.3. Deep Potential Molecular Dynamics Simulations.** The converged DNN potential after training was employed to conduct DPMD simulations using the LAMMPS code[71] interfaced with the DeepMD-kit.[70] Specifically, the trained DNN model was used as a pair style in LAMMPS to compute energy and force during MD steps. We considered two different types of systems to investigate chemical reactions in molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$. Type I, a system of the same size as for the AIMD simulations (see Section 2.1), consisted of 630 total atoms, to validate the DNN potential by comparing against the short-time AIMD data. In addition, we performed simulations in a larger system (Type II), taking advantage of the linear scaling of the DNN potential with simulation system size. This was 2.5 times the size of the Type I system, containing 1575 atoms. For both Type I and II systems, 10 independent simulations were performed starting from different initial configurations to improve statistical accuracy. DPMD simulations were performed within the *NpT* ensemble at 923.15 K and 1 bar. The integration time step was set to 1.0 fs. Type I systems were simulated for 100 ps, whereas for Type II systems, the total simulation time was 2 ns. Periodic boundary conditions along all three spacial dimensions were employed in the simulations.

## 3. RESULTS AND DISCUSSION

**3.1. Training Accuracy.** Figure 2 shows the learning curves of the final DNN model, where the squared energy and force errors are defined in eq 1. The fitting errors of the DNN model for energies and forces are of the order of $3 \times 10^{-7}$ eV and $6 \times 10^{-2}$ eV/Å, respectively, after $10^6$ training steps. These numbers indicate that the typical DPMD training accuracy has been achieved.[74] We observed significant improvement in training accuracy by incorporating more data points into the training data set as we performed more DP-GEN iterations. It can be seen by comparing the maximum standard deviation of the predicted atomic forces ($\zeta$, defined in Section 2.2) by the four different DNN models during each DP-GEN iteration. A comparison between iteration 6 and iteration 28 is displayed in Figure S1 of the Supporting Information (SI). In iteration 6, the magnitude of $\zeta$ is as high as 300 eV/Å and the peak of the distribution occurs at $\zeta$ = 20 eV/Å. In contrast, $\zeta$ values are distributed all below 1 eV/Å, and the peak of the distribution is 0.14 eV/Å, at the end of iteration 28. Such an improvement in the training accuracy can be attributed to the careful selection of training data sets and demonstrates the overall efficiency of the DP-GEN active learning process. In addition, to confirm the absence of overfitting in the trained model, the force error as a function of training steps is compared for both the training set and the validation set, which comprises 20% of the whole data set. As shown in Figure S2 of the SI, the train and validation profiles fall on each other, confirming no overfitting.

We randomly selected 150 configurations from the DPMD production run trajectories of Type I systems. Density functional theory calculations were performed to obtain energy and atomic forces for these configurations, using the same DFT parameters as for the preparation of the DNN training data set. Figure 3 demonstrates the comparisons of energies and atomic forces predicted by the DNN model to those predictions by DFT. Similar comparisons for pure $Li_2CO_3$ and KOH systems are depicted in Figures S3 and S4 of the SI. For both energies and forces, the DNN potential exhibits excellent agreement with the *ab initio* data. Specifically, the root-mean squared errors of the energy and atomic force predicted by the DNN model with respect to DFT were $2.28 \times 10^{-4}$ eV/atom and $1.04 \times 10^{-2}$ eV/Å, respectively. These values clearly manifest that the trained DNN potential has achieved typical DPMD accuracy.[74] Moreover, these errors are comparable in magnitude to the accuracy limits of DFT calculations as a result of **k**-point sampling and approximations in the selection of a finite plane-wave basis set. In the test set, the good agreement of forces suggests that the potential energy surface is not overfitted since the gradients are computed accurately in the test set which was not employed in training the DNN model.

**3.2. Bond Dissociation Kinetics and Composition Evolution.** A key motivation for this work was to investigate the ability of the DNN model to describe chemical reactions and characterize the formation and dissociation of chemical moieties in molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte containing dissolved $CO_2$. The scalability of the DNN model offers the advantage to examine such chemical reactions on a much longer time scale and large system sizes over the conventional AIMD approach, which typically involves trajectories of a few hundreds of picoseconds using moderate-size systems.[25] Our


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

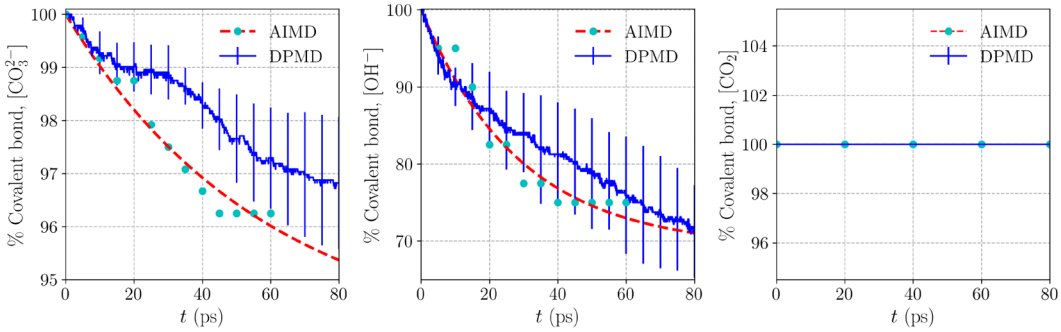



**Figure 4.** DPMD prediction of covalent bond retention [%] in molecules and ions as a function of time compared against AIMD reference data for an identical system (Type I). The employed DNN potential was trained using the potential energy cross section data representing the $CO_2$ dissociation.

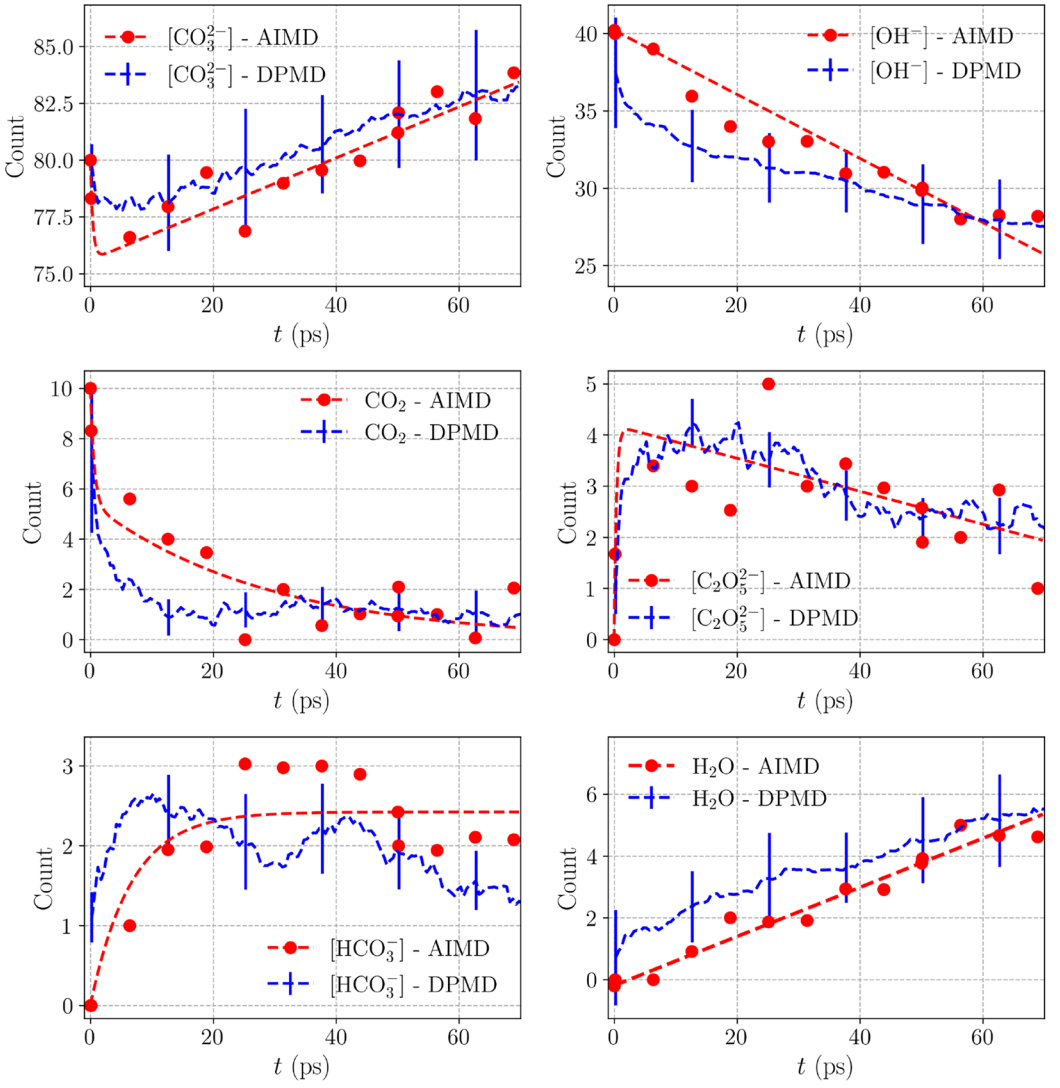



**Figure 5.** Number of different chemical species present as a function of time during the DPMD trajectory compared against AIMD reference data for an identical system. Red points are the actual observed numbers in the simulation box, while the red lines are drawn to guide the eye.

first step was to perform DPMD simulations starting from ten different initial configurations for the smaller (Type I) system, which consisted of 630 atoms. We obtained the fraction of covalent bond retention in molecules and ions as a function of time averaged over these 10 trajectories and compared them against the AIMD reference data in Figure 4. As shown in Figure 4, we observed satisfactory agreement between the DPMD and AIMD predictions. The C−O covalent bonds in $CO_2$ molecules are retained fully, and all $CO_2$ molecules are intact throughout the 80 ps DPMD trajectory, according to the AIMD results. This observation indicates that previous discrepancies in the DNN potential (discussed in Section


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

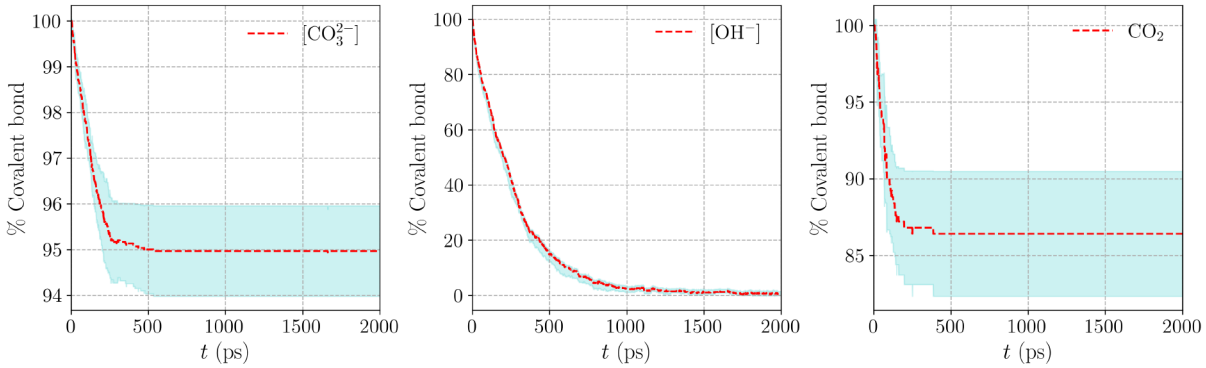



**Figure 6.** DPMD prediction of covalent bond retention [%] in molecules and ions as a function of time for a large system (Type II). The shaded area represents standard deviations.

2.2.1) have been rectified through the help of targeted learning by incorporating more training data representing the potential energy surface of $CO_2$ dissociation. Moreover, one of the limiting factors of the active learning approach is the short time scale which prevents access to events that may appear at a longer time scale enabled by NNs. Also, there could be circumstances where trajectories from the *ab initio* dynamics are unavailable. Ideally, the reference data set should contain diverse local atomic environments. However, traditional methods are prone to sampling similar configurations frequently, mainly due to the Boltzmann statistics. A solution to overcome this barrier would be introducing a sampling method optimized for semiautomatically collecting manifold yet appropriate configurations. One possible approach would be to employ enhanced sampling techniques such as metadynamics with the descriptor for the local atomic environment as a collective variable.[75] Consequently, the simulation is expected to be steered toward an unvisited local environment space so that each atom explores diverse chemical environments. In the case of $OH^-$, the overall profile almost reproduces the AIMD curve within the limits of the error bar. However, there are some systematic differences in the percentage of covalent bond retention in $CO_3^{2-}$ between AIMD and DPMD, as seen in the left panel of Figure 4. Despite these differences, the overall mixture composition is represented well, as discussed in the following paragraph.

We evaluated the overall composition of the electrolyte medium as a function of time for the small (Type I) system over the period of time for which AIMD results are available. We used two distance criteria of 1.8 and 1.3 Å for C−O and H−O covalent bonds, respectively, based on their respective radial distribution functions (RDFs), to define a molecular or ionic entity from the MD trajectory. We used a similar breadth-first search algorithm as described in our previous study[25] to define a chemical species based on a particular distance criterion. Once again, DPMD predicted profiles are averaged over ten independent trajectories and are compared against the AIMD data in Figure 5. We observed a satisfactory agreement between DPMD and AIMD predictions for the numbers of molecules and ions present in the electrolyte, as can be seen from Figure 5. It is encouraging to find such consistency between the two approaches in the quantitative estimations of molecule/ion counts. These results further corroborate the significance of the DNN potential in exhibiting accuracy and transferability similar to a DFT-level calculation, with a much lower computational cost. Nevertheless, the profiles displayed in Figure 5 do not reveal the equilibrium composition of the electrolyte system, as these correspond to short-time features present well below the 100 ps time scale.

**3.3. Long-Time Behavior.** We used the revised DNN potential to perform DPMD simulations of Type II systems at the same thermodynamic conditions, i.e., 923.15 K and 1 bar, for a much longer time scale (2 ns). The bond dissociation kinetics for ions and $CO_2$ molecules was evaluated and averaged over 10 independent starting configurations and is shown in Figure 6. We can conclude that the bond dissociation processes have reached equilibrium from these profiles. The percentage C−O covalent bond retention in both $CO_3^{2-}$ and $CO_2$ exhibits a saturation within the 500 ps time scale. The saturated values are 95% and 86.5% in $CO_3^{2-}$ and $CO_2$, respectively. In our previous AIMD study,[25] we observed the presence of dynamic equilibrium in pyrocarbonate ($C_2O_5^{2-}$) ion formation and dissociation from a reaction between $CO_3^{2-}$ and $CO_2$.

$$CO_2 + CO_3^{2-} \rightleftharpoons C_2O_5^{2-} \tag{2}$$

Our investigation also revealed that the formation of the $C_2O_5^{2-}$ ion is a kinetically favored process. However, due to the limitation of computational expenses, we could not reach an equilibrium point based on the AIMD trajectory. Interestingly, these equilibrium numbers obtained from DPMD simulations strongly suggest that the C−O covalent bond formation/dissociation process has stopped after 500 ps. In other words, the dynamic equilibrium shown in eq 2 has also reached an end point within the same time scale. This observation confirms that the $C_2O_5^{2-}$ ion is a transient species in the system. On the other hand, the H−O covalent bond retention percentage in $OH^-$ ions takes a longer time to reach an equilibrium, as shown in Figure 6. As shown in the following equations, the longer equilibration time is attributed to complex reaction dynamics involving hydroxide ions.

$$CO_2 + OH^- \rightleftharpoons HCO_3^- \tag{3}$$

$$HCO_3^- + OH^- \rightleftharpoons CO_3^{2-} + H_2O \tag{4}$$

$$H_2O + OH^- \rightleftharpoons OH^- + H_2O \tag{5}$$

In principle, $OH^-$ ions participate in all of these reactions, and eventually, all the H−O covalent bonds originally present in the hydroxide ions are broken within 1.5 ns, which is once again following our previous prediction based on the AIMD results.[25] Furthermore, this observation substantiates the


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

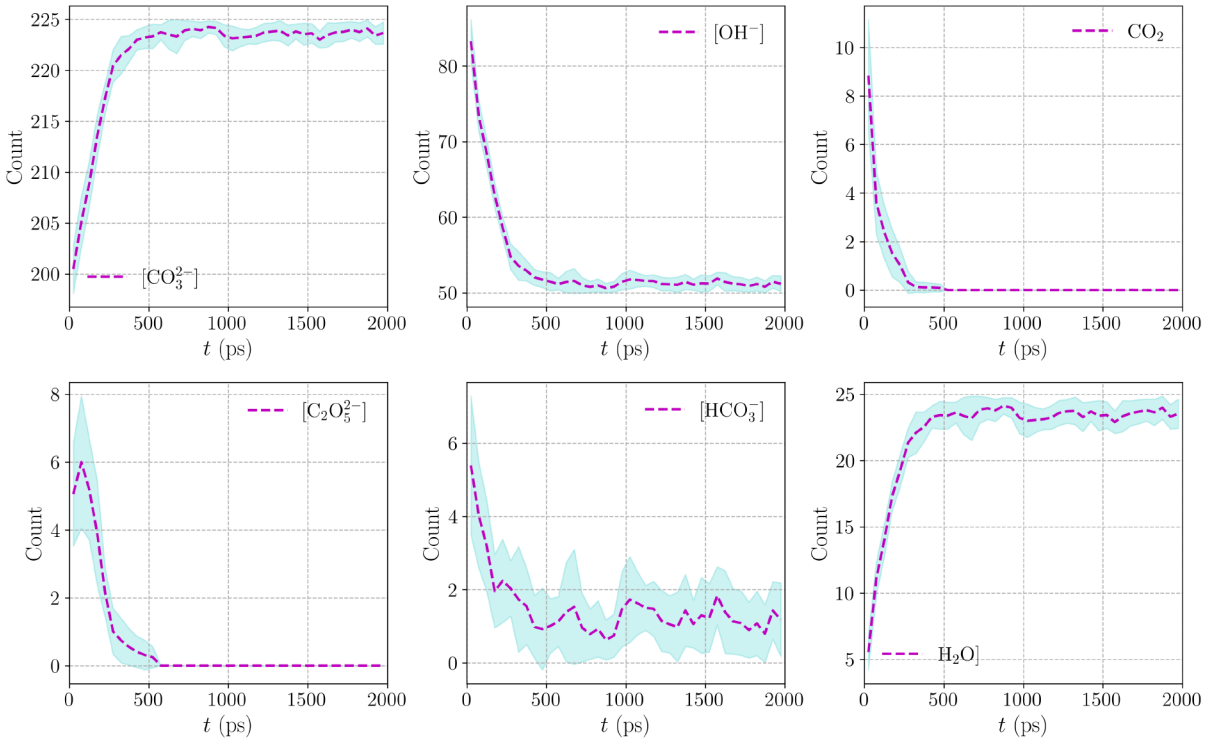



**Figure 7.** Number of different chemical species present as a function of time during the DPMD trajectory of a large system (2.5 times larger than the training system, Type II). The shaded area represents standard deviations.

AIMD anticipated thermodynamic nature of water and bicarbonate ($HCO_3^-$) ion formation in the electrolyte.

We analyzed the trajectories obtained from DPMD simulations of Type II systems to investigate the equilibrium composition of the electrolyte. The numbers of molecules and ions present in the electrolyte were evaluated and averaged from 10 independent trajectories and are shown in Figure 7. It is evident from Figure 7 that within 1 ns, the system has attained an equilibrium composition. At equilibrium, the fraction of $CO_3^{2-}$ anion has increased from its initial concentration primarily due to $HCO_3^-$ dissociation reaction with the $OH^-$ ion as shown in eq 4. On the other hand, the concentration of hydroxide ions at equilibrium is reduced by half from its starting fraction. Once again, this observation can be justified based on the transformation of $OH^-$ ions to either $HCO_3^-$ anions or water molecules via eq 3 and eq 4. It is interesting to find that, at 500 ps, there is no trace of $CO_2$ in the system—it is converted to either $C_2O_5^{2-}$ or $HCO_3^-$ anions following one of the reactions between eq 2 or eq 3. These observations involving ions and molecules present in the initial configuration are hand in hand with our previous speculations based on AIMD simulations.[25] Next, we analyzed the equilibrium fractions of the entities, which are the products of chemical reactions involving $CO_2$ and $CO_3^{2-}$ or $OH^-$. First, the pyrocarbonate ($C_2O_5^{2-}$) ion concentration exhibits a sharp rise at a very short time scale followed by a steep decay and eventually becomes zero around 500 ps. Bicarbonate anions ($HCO_3^-$) also show their presence in significant amounts, almost similar to $C_2O_5^{2-}$, in the initial time scale, and then decay faster and finally converge to a number close to two. In contrast, water molecules exhibit a steep surge in their numbers until 500 ps before converging to an equilibrium. These observations further substantiate that water and $HCO_3^-$ ions are thermodynamically more stable products than $C_2O_5^{2-}$, which is more like a transient species. Another important finding from Figure 7 is the contrasting behavior in $CO_2$ and $H_2O$ fractions at the equilibrium—even though we started our simulations with a finite number of $CO_2$ molecules, there is no trace found after 500 ps. An opposite scenario is observed for water molecules where we started with zero concentration and ended up with a large number of $H_2O$ molecules at equilibrium. This particular feature cements the fact that, in this kind of eutectic mixture, water is a more stable species compared to $CO_2$, as concluded based on our previous results from chemical reaction equilibrium simulations[11] and also inferred from AIMD trajectories.[25] In summary, these quantitative estimations provide us with an excellent insight into such a complex chemical environment. Furthermore, it is only possible because of the DNN potential to study the chemical reactivity involving large system sizes and a longer time scale to reach an equilibrium.

**3.4. Transport Properties.** On the basis of the DPMD trajectory for Type II systems, we have computed the mean-squared displacements and subsequently the diffusion coefficients of $Li^+$ and $K^+$ ions in the melt. The results obtained from DPMD simulations were compared against the AIMD data and are summarized in Table S1 of the SI. Diffusion coefficients in the table (for finite systems) are expected to be approximately 10% lower than for the infinite system size. Overall, the DPMD predicted diffusion coefficient values show a satisfactory agreement with the AIMD data. Specifically, DPMD results underestimate the AIMD computed diffusion coefficients by a maximum of 5%. Most importantly, the trend observed in both approaches is the same—$K^+$ is the cation found to diffuse slightly faster than the $Li^+$ ion.

In addition to studying the reactive system, we have also applied the DNN model to investigate the pure molten


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

carbonates and hydroxide salts. Systems of pure molten carbonates and hydroxides were composed of 500 cations and 250 carbonate or 500 hydroxide anions, i.e., 1500 atoms. Simulations were carried out in the isothermal−isobaric ($NpT$) ensemble for 2 ns to obtain liquid phase densities. Temperature and pressure were maintained using the Nosé−Hoover thermostat[76,77] and barostat.[78] Five independent trajectories were generated, and densities were determined as the average from the last 1 ns trajectory. First, the system temperature was equilibrated in the $NVT$ ensemble for 1 ns using fixed densities obtained from $NpT$ simulations for the viscosity calculations. It was followed by a 4 ns production run in the microcanonical ($NVE$) ensemble. The $NVE$ ensemble was chosen to avoid any possible effects from the barostat or thermostat on the dynamics of the systems. Other details of the viscosity calculations are described in the SI.

Densities of molten carbonates and hydroxides obtained from DPMD simulations are compared against the available experimental measurements as well as AIMD computed results and are summarized in Table 1. We observed significant deviations between DPMD predicted densities and experimental or AIMD values. It is surprising that the density obtained from the DNN model consistently underestimates the AIMD density, irrespective of whether it is a carbonate or hydroxide melt. Since the DNN potential was trained against the AIMD reference potential energy surfaces containing pure carbonates and hydroxides, it would have been expected that the DPMD density would resemble the AIMD one. Such consistent underestimation of liquid phase density could be attributed to the limited configuration space used in the training of the DNN potential. To substantiate this, Figure S5 of the SI demonstrates a sampling of densities in pure salts during the active learning process. As can be seen, AIMD densities were not accurately sampled in these systems. One reason could be that the initial configurations used in the active learning were at low densities. Since these systems are high-temperature molten salts, sampling the density at lower $T$ is not physically meaningful. It must be noted here that, during the active learning process, all the reference potential energy surfaces, be it a carbonate−hydroxide mixture or pure salts, were obtained at $p$ = 1 bar pressure. Therefore, one must perform active learning at higher pressures to obtain snapshots with densities close to the AIMD ones. Moreover, the DNN model should be able to identify structures that may collapse. This information may be learned from example configurations with randomly displaced atoms and/or scaled distances between atoms.[82] Despite the observed discrepancies in the densities of pure salts, we believe it would not have any significant impact on the studied chemical reactions in the molten $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte system containing dissolved $CO_2$.

**Table 1. DPMD Computed Densities (kg m$^{-3}$) of Molten Carbonates and Hydroxides, Compared against the Experimental Measurements and AIMD Results**

| system | $T$ (K) | exptl | AIMD | DPMD |
|---|---|---|---|---|
| $Li_2CO_3$ | 1043.15 | 1813[79] | 1907.2 ± 4.6 | 1834.9 ± 0.2 |
| $K_2CO_3$ | 1213.15 | 1878[79] | 1932.1 ± 6.4 | 1809.0 ± 0.3 |
| $Li_{0.6}K_{0.4}CO_3$ | 923.15 | 1938[79] | 2004.0 ± 4.5 | 1791.0 ± 0.4 |
| LiOH | 923.15 | 1296[79,a] | 1534.4 ± 2.8 | 1451.5 ± 0.3 |
| | | 1362[80,a] | | |
| KOH | 923.15 | 1607[79,a] | 1807.2 ± 7.8 | 1520.3 ± 0.3 |
| | | 1642[80,a] | | |
| | | 1727[81,a] | | |

[a]Extrapolated above experimental temperature range.

Table 2 summarizes DPMD computed viscosities of molten carbonates and hydroxides and their comparison against the available experimental measurements as well as previous simulation results using a classical force field. As is evident, the DPMD predicted viscosity values compare well with the available experimental data. More importantly, when compared against the simulated estimates from an empirical potential,[22] the DNN model shows much better performance. Of course, a significant drawback in the applied classical force field in ref 22 was the usage of full charges (±1e) on the ions, neglecting the inevitable effects of electronic polarization and charge transfer between ions. Our viscosity results are also consistent with the results reported in ref 53 based on a DPMD potential for the molten $Li_2CO_3$−$K_2CO_3$ binary salts. However, ref 53 reports higher values of density for similar systems compared to our results, though our DPMD model overestimates the AIMD density of $Li_{0.6}K_{0.4}CO_3$ by 5%, while ref 53 compares the results only to experimental ones. These differences in densities may come from the different parameters used in the training of the DNN potential and *ab initio* data set generation. Despite this, the present results employing the DNN potential are pretty promising. They demonstrate the ability of such neural network potentials to predict transport quantities with a maximum deviation less than 15% from the experimental measurements. This is particularly a very positive outlook for the DNN potentials in terms of their applications in investigating transport quantities of molten salts in general, which is always a cumbersome task using classical empirical potentials.

**Table 2. DPMD Computed Viscosities (cP) of Molten Carbonates and Hydroxides, Compared against the Experimental Measurements[79] and Classical MD results[22a]**

| system | $T$ (K) | exptl | DPMD | MD |
|---|---|---|---|---|
| $Li_2CO_3$ | 1101 | 4.9 | 5.0 ± 1.0 (2%) | 7.3 ± 0.7 (50%) |
| $K_2CO_3$ | 1180 | 3.0 | 2.9 ± 0.8 (−5%) | 4.1 ± 0.2 (37%) |
| $Li_{0.6}K_{0.4}CO_3$ | 1103 | 3.5 | 3.0 ± 0.8 (−14%) | 7.1 ± 0.2 (103%) |
| LiOH | 805 | – | 2.6 ± 0.6 | 2.2 ± 0.1 |
| KOH | 839 | 0.9 | 0.8 ± 0.2 (−12%) | 1.7 ± 0.1 (78%) |

[a]Quantities in parentheses are the percent deviations from the experimental values.

## 4. CONCLUSIONS

In this work, the DPMD[54] framework was used to train a robust and versatile machine learning potential for the study of chemical reactions in $CO_2$ dissolved in $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte. Our goal was to investigate various chemical reactions and the equilibrium composition of chemical species present in the system with *ab initio* accuracy. The reference training data set was generated using DFT-based molecular dynamics simulations. With the aid of the appropriate design of the reference database and the minimization of a loss function involving both potential energy and atomic forces, the DNN potential achieved a DFT-level accuracy in energy and force predictions. Extensive validations including the local structure, reaction kinetics, and evolution of chemical moieties at a short time scale in the $[Li_{0.6}K_{0.4}]_3CO_3OH$ electrolyte system


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

containing dissolved $CO_2$ demonstrated remarkable and consistent agreement between DPMD and AIMD simulations. With the linear scaling of DNN potential, we explored the chemical reactivity, equilibrium composition, and dynamics of the molten salt mixture, including $CO_2$ molecules using a larger supercell (2.5 times) through MD with DFT accuracy. DPMD trajectories in the nanosecond time scale revealed that a reaction between $CO_2$ and $CO_3^{2-}$ (see eq 2) yielding a pyrocarbonate ($C_2O_5^{2-}$) ion was indeed a transient species in the system. Moreover, the reaction completely stopped within 500 ps. We observed a higher reactivity of $CO_2$ with $OH^-$ leading to the formation of bicarbonate ($HCO_3^-$) ion via eq 3. The bicarbonate ions followed a quick conversion to water and $CO_3^{2-}$ leading to a net increase in the water concentration in the system. As a result of the high reactivity of $CO_2$ with either $CO_3^{2-}$ or $OH^-$, there was no trace of the $CO_2$ molecule after 500 ps. The equilibrium composition of the mixture consisted of $CO_3^{2-}$, $OH^-$, $HCO_3^-$, and $H_2O$, further confirming the much higher thermodynamic stability of water than that of $CO_2$ in the melt. These observations reproduced our earlier predictions based on chemical reaction equilibrium simulations.[11] In addition, the diffusion coefficients of cations were predicted within the acceptable accuracy of earlier performed AIMD results.

We have also applied the DNN potential to study pure carbonate and hydroxide melts based on $Li^+$ and $K^+$ cations. Although the predicted density of molten salts showed reasonable agreement with experimental results, the resulting densities were linked to the lack of training data at a wider range of densities and with unphysical structures and, thus, not sampled through the whole configuration space. Such deviations in the computed densities of molten salts obtained from machine learning potentials were also observed in the previous study of LiF and FLiBe molten salts.[50] Another prospect would be to include information about pressure tensors in the training of the DNN potential as well as long-ranged interactions beyond the cutoff of DPMD, which may also improve the resulting liquid density.

In summary, a deep neural network potential has demonstrated high adaptability to diverse local environments while retaining near quantum chemical accuracy. This feature is highly advantageous for practical applications since real systems involve multiple components and span wide thermodynamic conditions. It demonstrates a remarkable edge over classical empirical potentials, which are designed for specific applications and possess accuracy limitations, specifically in reactive systems and systems of growing complexity. Therefore, DNN-based MD simulations represent a promising pathway to address the long-standing accuracy-versus-cost trade-off. Furthermore, the predictive quality of DNN potentials can be systematically improved by incorporating information pertaining to various phases, states, chemical species, etc., in the training database. Such potentials will be developed and employed to investigate more computationally challenging systems in future work, facilitating substantial acceleration in screening and permitting interpolation across chemical space.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00816.

Deviations in predicted atomic forces, liquid phase structure, and transport properties (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Authors

**Anirban Mondal** − *Discipline of Chemistry, Indian Institute of Technology Gandhinagar, Gandhinagar, Gujarat 382355, India;* orcid.org/0000-0003-3029-8840; Email: amondal@iitgn.ac.in

**Athanassios Z. Panagiotopoulos** − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0002-8152-6615; Email: azp@princeton.edu

### Authors

**Dina Kussainova** − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States*

**Shuwen Yue** − *Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0002-1259-4649

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00816

### Author Contributions

¶(A.M. and D.K.) These authors contributed equally to this work.



### Funding

Financial support for this work was provided by the Office of Basic Energy Sciences, U.S. Department of Energy, under Award DE-SC0002128. Princeton Research Computing provided computing resources. A.M. acknowledges IIT Gandhinagar and the Centre for Development of Advanced Computing for the computational facility.


### Notes

The authors declare no competing financial interest.

## ■ REFERENCES


(1) Cassir, M.; McPhail, S.; Moreno, A. Strategies and new developments in the field of molten carbonates and high-temperature fuel cells in the carbon cycle. *Int. J. Hydrogen Energy* **2012**, *37*, 19345−19350.

(2) Chery, D.; Lair, V.; Cassir, M. Overview on $CO_2$ Valorization: Challenge of Molten Carbonates. *Front. Energy Res.* **2015**, DOI: 10.3389/fenrg.2015.00043.

(3) Chery, D.; Lair, V.; Cassir, M. $CO_2$ electrochemical reduction into CO or C in molten carbonates: a thermodynamic point of view. *Electrochim. Acta* **2015**, *160*, 74−81.

(4) Gaillard, F.; Malki, M.; Iacono-Marziano, G.; Pichavant, M.; Scaillet, B. Carbonatite Melts and Electrical Conductivity in the Asthenosphere. *Science* **2008**, *322*, 1363−1365.

(5) Dasgupta, R.; Hirschmann, M. M. The deep carbon cycle and melting in Earth's interior. *Earth Planet. Sci. Lett.* **2010**, *298*, 1−13.

(6) Xu, K. Electrolytes and Interphases in Li-Ion Batteries and Beyond. *Chem. Rev.* **2014**, *114*, 11503−11618.

(7) Hemmes, K.; Peelen, W. H. A.; de Wit, J. H. W. Study of the (Electro)Chemical Equilibria in Molten Carbonate Under the MCFC Cathode Gas Atmosphere. Part I: Selection of Independent Sets of Equilibria to Describe the (Electro)Chemical Equilibrium in Molten Carbonate Under the MCFC Cathode Gas Atmosphere. *Electrochim. Acta* **1998**, *43*, 2025−2031.

(8) Peelen, W. H. A.; van Driel, M.; Hemmes, K.; de Wit, J. H. W. Study of the (Electro)Chemical Equilibria in Molten Carbonate Under MCFC Cathode Gas Conditions. Part II: Non-Equilibrium

Study of (Electro)Chemical Reactions Involved in Oxygen Reduction in Molten Carbonate. *Electrochim. Acta* **1998**, *43*, 3313−3331.
(9) Evans, A.; Xing, W.; Norby, T. Electromotive Force (EMF) Determination of Transport Numbers for Native and Foreign Ions in Molten Alkali Metal Carbonates. *J. Electrochem. Soc.* **2015**, *162*, F1135−F1143.
(10) Rosen, J.; Geary, T.; Hilmi, A.; Blanco-Gutierrez, R.; Yuh, C.-Y.; Pereira, C. S.; Han, L.; Johnson, R. A.; Willman, C. A.; Ghezel-Ayagh, H.; Barckholtz, T. A. Molten Carbonate Fuel Cell Performance for $CO_2$ Capture from Natural Gas Combined Cycle Flue Gas. *J. Electrochem. Soc.* **2020**, *167*, 064505.
(11) Young, J. M.; Mondal, A.; Barckholtz, T. A.; Kiss, G.; Koziol, L.; Panagiotopoulos, A. Z. Predicting Chemical Reaction Equilibria in Molten Carbonate Fuel Cells via Molecular Simulations. *AIChE J.* **2020**, *67*, e16988.
(12) Audasso, E.; Bosio, B.; Bove, D.; Arato, E.; Barckholtz, T.; Kiss, G.; Rosen, J.; Elsen, H.; Gutierrez, R. B.; Han, L.; Geary, T.; Willman, C.; Hilmi, A.; Yuh, C. Y.; Ghezel-Ayagh, H. New, Dual-Anion Mechanism for Molten Carbonate Fuel Cells Working as Carbon Capture Devices. *J. Electrochem. Soc.* **2020**, *167*, 084504.
(13) Audasso, E.; Bosio, B.; Bove, D.; Arato, E.; Barckholtz, T.; Kiss, G.; Rosen, J.; Elsen, H.; Blanco Gutierrez, R.; Han, L.; Geary, T.; Willman, C.; Hilmi, A.; Yuh, C. Y.; Ghezel-Ayagh, H. The Effects of Gas Diffusion in Molten Carbonate Fuel Cells Working as Carbon Capture Devices. *J. Electrochem. Soc.* **2020**, *167*, 114515.
(14) Rollet, A.-L.; Salanne, M. Studies of the local structures of molten metal halides. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2011**, *107*, 88−123.
(15) Tissen, J.; Janssen, G. Molecular-dynamics simulation of molten alkali carbonates. *Mol. Phys.* **1990**, *71*, 413−426.
(16) Janssen, G. J. M.; Tissen, J. T. W. M. Pair Potentials from ab initio Calculations for use in MD Simulations of Molten Alkali Carbonates. *Mol. Sim.* **1990**, *5*, 83−98.
(17) Tissen, J.; Janssen, G.; van der Eerden, P. Molecular dynamics simulation of binary mixtures of molten alkali carbonates. *Mol. Phys.* **1994**, *82*, 101−111.
(18) Habasaki, J. Molecular dynamics simulation of molten $Li_2CO_3$ and $Na_2CO_3$. *Mol. Phys.* **1990**, *69*, 115−128.
(19) Costa, M. Molecular dynamics of molten $Li_2CO_3$−$K_2CO_3$. *J. Mol. Liq.* **2008**, *138*, 61−68.
(20) Vuilleumier, R.; Seitsonen, A.; Sator, N.; Guillot, B. Structure, equation of state and transport properties of molten calcium carbonate ($CaCO_3$) by atomistic simulations. *Geochim. Cosmochim. Acta* **2014**, *141*, 547−566.
(21) Corradini, D.; Coudert, F.-X.; Vuilleumier, R. Insight into the $Li_2CO_3$−$K_2CO_3$ eutectic mixture from classical molecular dynamics: Thermodynamics, structure, and dynamics. *J. Chem. Phys.* **2016**, *144*, 104507.
(22) Mondal, A.; Young, J. M.; Barckholtz, T. A.; Kiss, G.; Koziol, L.; Panagiotopoulos, A. Z. Genetic Algorithm Driven Force Field Parameterization for Molten Alkali-Metal Carbonate and Hydroxide Salts. *J. Chem. Theory Comput.* **2020**, *16*, 5736−5746.
(23) Mondal, A.; Young, J. M.; Barckholtz, T. A.; Kiss, G.; Koziol, L.; Panagiotopoulos, A. Z. Transport and Interfacial Properties of Mixed Molten Carbonate/Hydroxide Electrolyes by Molecular Dynamics Simulations. *J. Phys. Chem. C* **2020**, *124*, 23532−23540.
(24) Corradini, D.; Coudert, F. X.; Vuilleumier, R. Carbon dioxide transport in molten calcium carbonate occurs through an oxo-Grotthuss mechanism via a pyrocarbonate anion. *Nat. Chem.* **2016**, *8*, 454−460.
(25) Mondal, A.; Young, J. M.; Kiss, G.; Panagiotopoulos, A. Z. First-Principles Modeling of Transport Mechanisms in Carbonate−Hydroxide Electrolytes. *J. Phys. Chem. C* **2021**, *125*, 4412−4422.
(26) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
(27) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.
(28) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **2017**, *8*, 3192−3203.
(29) Smith, J. S.; Roitberg, A. E.; Isayev, O. Transforming Computational Drug Discovery with Machine Learning and AI. *ACS Med. Chem. Lett.* **2018**, *9*, 1065−1069.
(30) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* **2019**, *10*, 2903.
(31) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet − A deep learning architecture for molecules and materials. *J. Chem. Phys.* **2018**, *148*, 241722.
(32) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15*, 3678−3693.
(33) Zhang, Y.; Hu, C.; Jiang, B. Embedded Atom Neural Network Potentials: Efficient and Accurate Machine Learning with a Physically Inspired Representation. *J. Phys. Chem. Lett.* **2019**, *10*, 4962−4967.
(34) Artrith, N.; Morawietz, T.; Behler, J. High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide. *Phys. Rev. B* **2011**, *83*, 153101.
(35) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-chemical insights from deep tensor neural networks. *Nat. Commun.* **2017**, *8*, 13890.
(36) Yang, M.; Bonati, L.; Polino, D.; Parrinello, M. Using metadynamics to build neural network potentials for reactive events: the case of urea decomposition in water. *Catal. Today* **2022**, *387*, 143−149.
(37) Rowe, P.; Deringer, V. L.; Gasparotto, P.; Csányi, G.; Michaelides, A. An accurate and transferable machine learning potential for carbon. *J. Chem. Phys.* **2020**, *153*, 034702.
(38) Scherer, C.; Scheid, R.; Andrienko, D.; Bereau, T. Kernel-Based Machine Learning for Efficient Simulations of Molecular Liquids. *J. Chem. Theory Comput.* **2020**, *16*, 3194−3204.
(39) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121*, 10073−10141.
(40) Deringer, V. L.; Bernstein, N.; Csányi, G.; Ben Mahmoud, C.; Ceriotti, M.; Wilson, M.; Drabold, D. A.; Elliott, S. R. Origins of structural and electronic transitions in disordered silicon. *Nature* **2021**, *589*, 59−64.
(41) Nigam, J.; Willatt, M. J.; Ceriotti, M. Equivariant representations for molecular Hamiltonians and N-center atomic-scale properties. *J. Chem. Phys.* **2022**, *156*, 014115.
(42) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 2453.
(43) Batatia, I.; Batzner, S.; Kovács, D. P.; Musaelian, A.; Simm, G. N. C.; Drautz, R.; Ortner, C.; Kozinsky, B.; Csányi, G.Design Space of E(3)-Equivariant Atom-Centered Interatomic Potentials. *arXiv*, 2022, https://arxiv.org/abs/2205.06643.
(44) Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C. J.; Kornbluth, M.; Kozinsky, B.Learning Local Equivariant Representations for Large-Scale Atomistic Dynamics. *arXiv*, 2022, https://arxiv.org/abs/2204.05249.
(45) Pun, G. P. P.; Batra, R.; Ramprasad, R.; Mishin, Y. Physically informed artificial neural networks for atomistic modeling of materials. *Nat. Commun.* **2019**, *10*, 2339.
(46) Sivaraman, G.; Guo, J.; Ward, L.; Hoyt, N.; Williamson, M.; Foster, I.; Benmore, C.; Jackson, N. Automated Development of Molten Salt Machine Learning Potentials: Application to LiCl. *J. Phys. Chem. Lett.* **2021**, *12*, 4278−4285.
(47) Li, Q.-J.; Küçükbenli, E.; Lam, S.; Khaykovich, B.; Kaxiras, E.; Li, J. Development of robust neural-network interatomic potential for molten salt. *Cell Reports Physical Science* **2021**, *2*, 100359.


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(48) Tovey, S.; Narayanan Krishnamoorthy, A.; Sivaraman, G.; Guo, J.; Benmore, C.; Heuer, A.; Holm, C. DFT Accurate Interatomic Potential for Molten NaCl from Machine Learning. *J. Phys. Chem. C* **2020**, *124*, 25760−25768.
(49) Rodriguez, A.; Lam, S.; Hu, M. Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials. *ACS Appl. Mater. Interfaces* **2021**, *13*, 55367−55379.
(50) Lam, S. T.; Li, Q.-J.; Ballinger, R.; Forsberg, C.; Li, J. Modeling LiF and FLiBe Molten Salts with Robust Neural Network Interatomic Potential. *ACS Appl. Mater. Interfaces* **2021**, *13*, 24582−24592.
(51) Lee, S.-C.; Zhai, Y.; Li, Z.; Walter, N. P.; Rose, M.; Heuser, B. J.; Z, Y. Comparative Studies of the Structural and Transport Properties of Molten Salt FLiNaK Using the Machine-Learned Neural Network and Reparametrized Classical Forcefields. *J. Phys. Chem. B* **2021**, *125*, 10562−10570.
(52) Nguyen, M.-T.; Rousseau, R.; Paviet, P. D.; Glezakou, V.-A. Actinide Molten Salts: A Machine-Learning Potential Molecular Dynamics Study. *ACS Appl. Mater. Interfaces* **2021**, *13*, 53398−53408.
(53) Feng, T.; Yang, B.; Lu, G. Investigation on the local structure and properties of molten $Li_2CO_3$-$K_2CO_3$ binary salts by machine learning potentials. *J. Mol. Liq.* **2022**, *356*, 118979.
(54) 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.
(55) Lu, D.; Wang, H.; Chen, M.; Lin, L.; Car, R.; E, W.; Jia, W.; Zhang, L. 86 PFLOPS Deep Potential Molecular Dynamics simulation of 100 million atoms with ab initio accuracy. *Comput. Phys. Commun.* **2021**, *259*, 107624.
(56) Gartner, T. E.; Zhang, L.; Piaggi, P. M.; Car, R.; Panagiotopoulos, A. Z.; Debenedetti, P. G. Signatures of a liquid−liquid transition in an ab initio deep neural network model for water. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 26040−26046.
(57) Calegari Andrade, M. F.; Ko, H.-Y.; Zhang, L.; Car, R.; Selloni, A. Free energy of proton transfer at the water−$TiO_2$ interface from ab initio deep potential molecular dynamics. *Chem. Sci.* **2020**, *11*, 2335−2341.
(58) Zeng, J.; Zhang, L.; Wang, H.; Zhu, T. Exploring the Chemical Space of Linear Alkane Pyrolysis via Deep Potential GENerator. *Energy Fuels* **2021**, *35*, 762−769.
(59) Zhang, Y.; Wang, H.; Chen, W.; Zeng, J.; Zhang, L.; Wang, H.; E, W. DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* **2020**, *253*, 107206.
(60) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864−B871.
(61) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133−A1138.
(62) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. cp2k: atomistic simulations of condensed matter systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15−25.
(63) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103−128.
(64) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(65) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703−1710.
(66) Hartwigsen, C.; Goedecker, S.; Hutter, J. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **1998**, *58*, 3641−3662.
(67) 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.
(68) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé−Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(69) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177−4189.
(70) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178−184.
(71) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(72) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(73) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(74) Zhang, L.; Wang, H.; Car, R.; E, W. Phase Diagram of a Deep Potential Water Model. *Phys. Rev. Lett.* **2021**, *126*, 236001.
(75) Yoo, D.; Jung, J.; Jeong, W.; Han, S. Metadynamics sampling in atomic environment space for collecting training data for machine learning potentials. *npj Computational Materials* **2021**, *7*, 131.
(76) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255−268.
(77) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(78) Kamberaj, H.; Low, R. J.; Neal, M. P. Time reversible and symplectic integrators for molecular dynamics simulations of rigid molecules. *J. Chem. Phys.* **2005**, *122*, 224114.
(79) Janz, G. J.*Thermodynamic and Transport Properties for Molten Salts: Correlation Equations for Critically Evaluated Density, Surface Tension, Electrical Conductance, and Viscosity Data*; *Journal of Physical and Chemical Reference Data*; National Bureau of Standards: 1988; Vol. *17*.
(80) Claes, P.; Glibert, J. Electrical Conductivity and Specific Mass of the Molten LiOH - $LiNO_3$, NaOH - $NaNO_3$, and KOH - $KNO_3$ Mixtures. *J. Electrochem. Soc.* **1985**, *132*, 857−862.
(81) Ward, A.; Janz, G. Molten carbonate electrolytes: Electrical conductance, density and surface tension of binary and ternary mixtures. *Electrochim. Acta* **1965**, *10*, 849−857.
(82) Behler, J. Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem.* **2015**, *115*, 1032−1050.


https://doi.org/10.1021/acs.jctc.2c00816
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX