Modeling LiF and FLiBe molten salts with Robust Neural Network Interatomic Potentials

Stephen T. Lam[1,2]*, Qing-Jie Li[2], Ronald Ballinger[2,3], Charles Forsberg[2], Ju Li[2,3]

[1]Department of Chemical Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA

[2]Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

* Corresponding author: stephen_lam@uml.edu

Abstract

Lithium-based molten salts have attracted significant attention due to their applications in energy storage, advanced fission reactors and fusion devices. Lithium fluorides and particularly 66.6%LiF-33.3%BeF$_2$ (Flibe) are of considerable interest in nuclear systems, as they show an excellent combination of desirable heat-transfer and neutron-absorption characteristics. For nuclear salts, the range of possible local structures, compositions, and thermodynamic conditions presents significant challenges in atomistic modeling. In this work, we demonstrate that atom-centered neural network interatomic potentials (NNIP) provide a fast and accurate method for performing molecular dynamics of molten salts. For LiF, these potentials are able to accurately model dimer interactions, crystalline solids under deformation, semi-crystalline LiF near the melting point and liquid LiF at high temperatures. For Flibe, NNIPs accurately predicts the structures and dynamics at normal operating conditions, high temperature-pressure conditions, and in the crystalline solid phase. Furthermore, we show that NNIP-based molecular dynamics of molten salts are scalable to reach long timescales (e.g., nanosecond) and large system sizes (e.g., 10^5 atoms), while maintaining ab initio accuracy and providing more than three orders of magnitude of computational speedup for calculating structure and transport properties.

Keywords: nuclear energy, molten salt, lithium fluoride, atomistic simulation, neural networks

1   Introduction

Molten salts are important high-temperature liquids for many industrial applications such as waste oxidation, catalytic coal gasification, concentrated solar power, and advanced nuclear reactors [1][2][3][4][5]. Generally, the modeling of ionic liquids poses interesting challenges considering the complexities in modeling various atomic structures, multiple phases (solid, liquid, vapor and glass) and transformations that can dramatically change salt's properties [6]. For advanced fission and compact fusion reactors, lithium-based fluoride salts are the material of choice owing to their high actinide solubility and desirable heat transfer characteristics [7]. Specifically, $LiF-BeF_2$ mixtures have been identified as important prototype salts in reactor systems in the thermal neutron spectrum. This is because a) $LiF-BeF_2$ salts show low neutron absorption, and b) LiF is a common constituent used to depress a mixture's melting point, thus reducing the chance of salt freezing. Furthermore, Li is desirable in fusion systems due to its ability to generate nuclear fuel as tritium.

In the past, physics-based interatomic potentials have been developed for molten salts using ab initio force-fitting. Such empirical potentials often require an explicit definition of the potential energy functional form, and assumptions about the relevant interactions in a given salt mixture. For example, to better reproduce thermophysical properties for salts containing multiply charged cations [8], one needs to fit both forces and multipoles to capture electronic polarization effects. Therefore, constructing such models is time-consuming and eventually intractable with increasing complexity in coordination environments and the number of species in a salt mixture. This is particularly salient in high-temperature molten salt systems, where operating conditions can vary greatly, and selecting from salt mixtures of many possible constituent combinations are of interest. On the other hand, machine-learning based methods such as artificial neural networks are considered universal function approximators, thus representing a new paradigm for modeling molten salt atomic systems. In this work, we trained two NNIPs using LiF and Flibe (66.6%LiF-33.3%BeF_2) as our model systems, due to their relevance in technological application and their intrinsically interesting properties. The LiF system is included since it is relatively simpler and has well-defined

crystalline solid structure that can be studied to test and validate the NNIP method. In Flibe, the mixture of monovalent Li and divalent Be results in heterogeneous ion transport behavior, causes a number of chemical reactions with impurities, and forms extended structures like BeF_2 corner-sharing polyhedral chains [9]. Thus, the combination of these two salts provides good prototypes for testing NNIPs in molten salt applications.

2   Methods

2.1 Materials and Salt Systems

In order to train a robust neural network capable of capturing a range of local atomic environments, a variety of different systems were included. For the LiF potential, the total dataset includes 3600 images of atomic pairs (1400 F-F, 1400 Li-F and 1400 Li-Li) at different separation distances, 1400 solid crystalline configurations at 0K under $\pm$ 20% deformation, 6000 solid configurations near the LiF melting point at 1050K, 6000 semi-crystalline LiF near the melting point at 1120K, and 8000 liquid LiF configurations at 1200K. For Flibe, all configurations were sampled with molecular dynamics with 8000 configurations at 973K, 4000 configurations at 973K under 15% compression, 4000 configurations at a high temperature of 5000K. All finite temperature configurations were subsampled from ab initio molecular dynamics trajectories as described by the following section.

2.2 Ab Initio Data Generation

Training and validation data were generated using density functional theory (DFT) [10] and Born-Oppenheimer Ab Initio Molecular Dynamics (AIMD) [11]. AIMD sampling was performed in the canonical (*NVT*) ensemble using a time step of 2 fs and a Nosé-mass corresponding to 80 fs. Configurations were uniformly sampled from AIMD trajectories every 10 fs to reduce correlation in the input data. Calculations were performed using the Vienna Ab-initio Simulation Package (VASP) [12] with plane-wave basis set and periodic boundary conditions in all cell directions. For LiF, pair interactions were calculated based on dimer configurations in a cubic cell with side length of 24 Å. For crystalline solid at 0 K, an 8-

atom unit cell is used with a 3×3×3 and Γ centered k-point grid. LiF AIMD simulations of solid and semi-crystalline systems were performed for 60 ps for systems containing 64 atoms, and liquid simulations were performed for 80 ps for a system containing 70 atoms. All AIMD systems used a 2×2×2 **k**-point sampling on a Γ-centered grid. In all cases, a plane wave energy cutoff of 550 eV was used. For Flibe, all configurations contain 91 atoms (26 Li, 52 F, and 13 Be) and the AIMD sampling was based on a plane wave energy cutoff of 600 eV, and Γ point only Brillouin zone sampling. Simulations at 973 K, under 15% compression, and at 5000 K were performed for 80, 40 and 40 ps, respectively. The **k**-point sampling density and plane wave energy cutoff was chosen to ensure an energy convergence of < 2 meV/atom between different calculations [10][9]. All calculations were performed using the Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation (GGA) exchange-correlation functional and projector-augmented wave (PAW) potentials for the nucleus and core electrons (Li_sv, Be, and F). These simulation protocols were previously validated for a variety of fluoride salts to produce accurate predictions for a variety of chemical and physical properties [9].

2.3 Neural Network Interatomic Potentials

The potential energy surface is modeled using a modified Behler-Parinello (BP) atom-centered approach. Here the total energy is considered as the sum of energy contributions of individual atoms i in the system $E_{\text{tot}} = \sum_{N_{\text{at}}} E_i(\mathbf{G}_i)$ where the local atomic environment is captured by a vector of symmetry functions $\mathbf{G}_i$. These modified BP functions were adopted from Smith *et al.* [13] and Lot *et al.* [14] and consist of a two-body and three-body descriptors, which capture the radial and angular environments, respectively. For a central atom i, the radial descriptor is defined as a sum of Gaussian functions:

$$G_i^R = \sum_{j \neq i}^{N_p} e^{-\eta(R_{ij}-R_S)^2} f_c(R_{ij}) \tag{1}$$

where η, and R_S represent the width and offset of the Gaussian, R_{ij} is the distance between atoms i and j, and the sum is taken over all ion pairs N_p within a cutoff that is defined by function f_c:

$$f_c(R_{ij}) = 0.5 \cos\left(\frac{\pi R_{ij}}{R_c}\right) + 0.5 \qquad \text{for} \qquad R_{ij} < R_c \qquad (2)$$

$$f_c(R_{ij}) = 0 \qquad \text{for} \qquad R_{ij} > R_c$$

This cutoff function presents a smooth decay to zero at the limit of a defined R_c, which was chosen to be 7 Å. The smooth cutoff ensures that the function is continuously differentiable, enabling the reliable calculation of forces for molecular dynamics simulations. This cutoff radius value was found to be a sufficient approximation for property prediction. The angular descriptor is defined:

$$G_i^A = 2^{1-\zeta} \sum_{j,k\neq i}^{N_t} \left(1 + \cos(\theta_{ijk} - \theta_S)\right)^\zeta e^{-\eta\left(\frac{R_{ij}+R_{ik}}{2} - R_S\right)^2} f_c(R_{ij}) f_c(R_{ik}) \qquad (3)$$

where ζ is the width of the gaussian function, θ_S is the cosine offset, and θ_{ijk} is the angle formed by atoms i, j, and k, and the sum is taken over all triplets within R_c. The angular and radial functions are calculated based on different sets of hyperparameters (ζ, θ_S, R_S), which are then concatenated into to neural network input vector $\mathbf{G}_i$. In this way, the resolution of the feature space can be arbitrarily increased by sampling the hyper-parameters at increasingly fine intervals. The general method is discussed in more detail in [13]. In this work, the hyperparameter sets were all combinations of $R_c = 7.0$ Å, $\eta = 12.0$, $\zeta = 50.0$, $R_{S\,(\text{radial})} = \{0.5, 0.75, \dots, 7.75\}$, $R_{S\,(\text{angular})} = \{2, 2.5, \dots, 5.5\}$, and $\theta_S = \left\{0, \frac{1}{6}\pi, \frac{1}{3}\pi, \frac{1}{2}\pi, \frac{2}{3}\pi, \frac{5}{6}\pi\right\}$. Thus, the size of $\mathbf{G}_i$ is 168 for LiF ($i = $ Li, F) and 294 for Flibe ($i = $ Li, Be, F), serving as input into a corresponding neural network based on the species of the central atom i.

Neural networks are trained using mini-batch gradient descent with the Adam optimization algorithm [15] using a decaying learning rate from 1E-3 to 1E-4. This was used to minimize the cost function J:

$$J(\mathbf{W}) = \sum_{i\in\text{batch}} \left[\left(E_i^{\text{DFT}} - E_i^{\text{NN}}(\mathbf{W})\right)^2 + \frac{c_f}{3N_i} \sum_{j=1}^{N_i} \sum_{k\in\{x,y,z\}} \left(F_{ij}^{\text{DFT},k} - F_{ij}^{\text{NN},k}(\mathbf{W})\right)^2\right] + C_\mathbf{W}\|\mathbf{W}\|^2 \qquad (4)$$

where **W** is the weight matrix that parameterizes the neural network, E_i is the energy of an atomic configuration i calculated by DFT or NN, F_{ij}^k is the forces of atom j in the configuration i in the direction k, c_f is a factor to weight the relative importance of obtaining accurate forces, and $c_\mathbf{W}$ is the regularization constant. Here, a batch size of 100 configurations is used, $c_\text{f} = 0.2$, and $c_\mathbf{W} = 0.1$ is used to prevent overfitting of the neural network. For LiF, all the atomic neural networks consist of two fully connected hidden layers with 64 nodes each. A similar architecture is also used for Flibe, except with 32 nodes in each layer. The Gaussian activation function was used in hidden layer nodes and linear activation in the output node. In each salt system, 80% of the total data is used for training and tuning, and the remaining 20% is withheld and used as the test set. The training is conducted using the PANNA package [14]. Molecular dynamics simulations with the neural network potentials are then performed in LAMMPS [16].

3   Results

3.1 LiF Binary Salt Potential

3.1.1   Pair Interactions, Bulk EOS and Surfaces

A variety of atomic configurations are used to train and test the neural networks as described in section 2.1. In addition, various high-index surfaces that were not seen by the network were also calculated and compared to the DFT values to assess the ability of the network to extrapolate. First, dimer interactions predicted by the neural network are shown in Figure 1a) for Li-F, F-F and Li-Li, respectively. In all cases, the neural network predictions and DFT calculations agree very well. The flattening of the tails on the potential energy curve is accurately predicted by the neural network, which means that the medium-range contributions to force are accurately represented. The high-energy repulsion at short interatomic distance is also accurately reproduced. At short interatomic distances, accuracy is limited by the DFT data themselves due to the use of pseudopotentials to model the nucleus and core electrons. These high-energy interactions are important when simulating liquids at high temperatures where a large number of configurations are sampled, and high energy collisions can occur. Further, accurate calculation of these

energies reduces energy drift due to time integration for molecular dynamics simulations in the microcanonical ensemble (*NVE*). In all species pairs, the equilibrium distance (minimum energy) is also reproduced. Across the range of atomic distances, this accuracy is found sufficient for molecular dynamics.
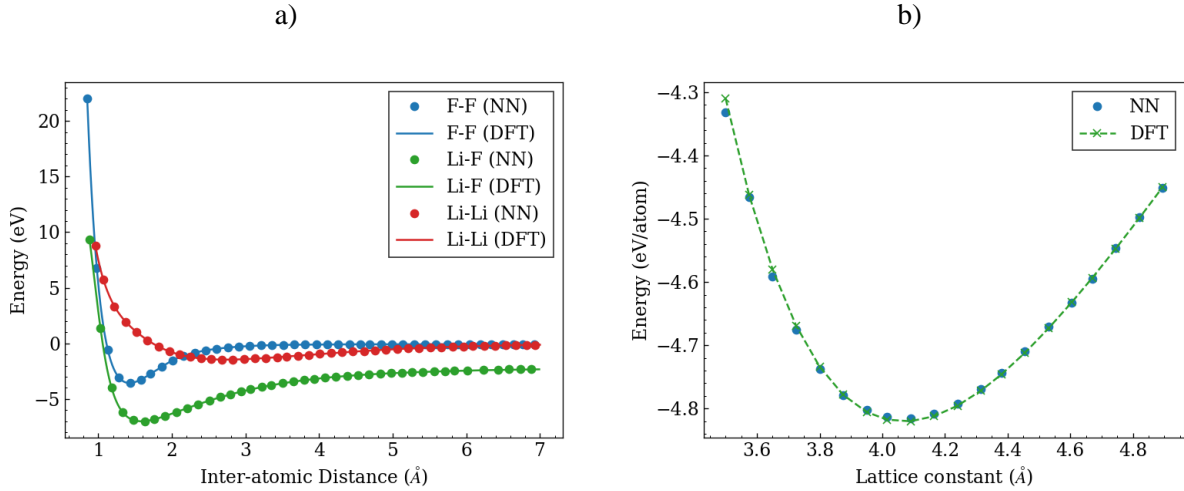
a)   b)

Figure 1: DFT versus NN predictions for a) pair interactions for Li-F, Li-Li and F-F and b) Equation of state for bulk B1 phase LiF at 0K [17].

Using the same neural network, the equation of state for bulk solid phase B1 is also fit as shown in Figure 1b). Energy data was collected for different cell volumes, and the equilibrium lattice constant is calculated as 4.07 Å at the minimum energy, which is consistent with previous calculations within $\pm$ 0.01 Å [18]. The neural network prediction is in good agreement with DFT calculations with a mean average error (MAE) of < 8 meV/atom. At large compression, the error is higher due to poorer sampling in that region. This accuracy can be improved by adding more high-energy configurations into the training data. For the calculations here, the accuracy is sufficient since most simulations are performed either at or near equilibrium density. Moreover, the energy $E(\mathbf{r})$ for pair and EOS interactions is smooth and continuous, which allows for the accurate calculation of forces and stresses.

Surface energy calculations were also performed with neural networks and compared to DFT results under different local environments (coordination, species density, symmetry). The slabs were constructed with 11 − 13 layers. The DFT calculations were performed using a $3 \times 3 \times 1$ **k**-point mesh.

Graphical representations of some of the calculated surfaces are shown in Figure 2. The surface energy γ is calculated as follows:

$$\gamma = \frac{E_{\text{slab}} - NE_{\text{bulk}}}{2A} \qquad (5)$$

where E_{slab} is the calculated energy of the slab, N is the number of formula units of LiF in the slab, E_{bulk} is the energy of a bulk crystal per unit, and A is the area of the slab's surface. In each cell, a 20 Å thick vacuum is added on top of the slabs to limit interactions with periodic images. In this work, we are interested in neural network prediction of DFT calculations in low coordination environments rather than providing detailed study of LiF surfaces, which has been the work of previous studies [19]. Thus, optimization and testing of various minimum energy structures was not performed. We therefore note that the surface energies calculated here should not be taken as precise determination of values that should be compared to experimental values. As such, calculations were performed with only ionic relaxation of the two outer layers. For the (111) surfaces, ½ of the atoms were removed from the top and bottom layer to preserve a 1:1 Li:F stoichiometry. Here, there are two possible surfaces: one with Li-termination $(111)_{\text{Li}}$ and one with F-termination $(111)_{\text{F}}$ as shown in Figure 2b) and 2c). We note that this differs from other proposed (111) surfaces, such as the octopolar reconstructed surface where ¼ of the atoms in the outermost layer and ¾ of the atoms on the subsequent layer are removed [19].

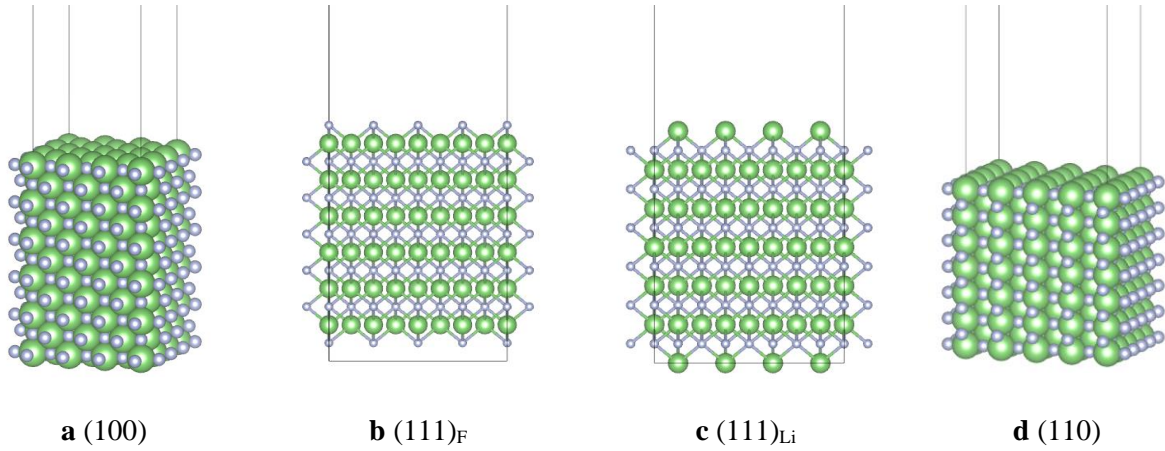| **a** (100) | **b** $(111)_{\text{F}}$ | **c** $(111)_{\text{Li}}$ | **d** (110) |

Figure 2: LiF surfaces with different surface terminations and 1:1 Li:F stoichiometry.

The surface energies calculated from DFT are compared to neural network in Table 1 for the (100), (110) , $(111)_{Li}$ and $(111)_F$ surfaces. The DFT calculated surface energies ranged from 34.6 to 67.2 meV/ $Å^2$, representing a broad range of values. We note that the surface energy for (100) is in excellent agreement with previous calculations despite various approximations made in the system setup and structural optimization [20]. In all cases, the predicted surface energies are relatively close to the DFT energies, with the greatest error in $(111)_{Li}$, where the neural network under-predicts the surface energy by 4.1 meV/$Å^2$. Considering that that local chemical coordinations for these surfaces are significantly different (in species and number) compared to bulk crystal or liquid configurations which comprised of the vast majority of training data, these calculations suggest that the neural network potential has good transferability and performs well over a range of environments.

Table 1: Calculated surface energies from DFT versus neural networks.

Surface	DFT (meV/ $Å^2$)	NN (meV/ $Å^2$)
(100)	34.6	36.8
(110)	47.0	49.9
$(111)_{Li}$	63.5	59.4
$(111)_F$	67.2	64.0

3.1.2 Solid-to-Liquid Phases of LiF

The neural network interatomic potential was used to calculate energies and forces for atomic configurations, including solid crystalline LiF, semi-crystalline LiF near the melting point and liquid LiF. A single NNIP that accurately simulates all of these configurations is desirable since it would allow for phase transitions over the range of potential operating conditions. Specifically, the prediction of salt properties near the phase boundaries is useful in understanding melting and freezing transitions that are critical to the safety of molten salt reactors [7]. Snapshots of the atomic configurations are shown in Figure

3. As shown, semi-crystalline contains large localized atomic displacements while maintaining a partially ordered structure, while liquid configurations are highly disordered. The disorder of the liquid state is confirmed by calculation of radial distribution functions in previous work [9].
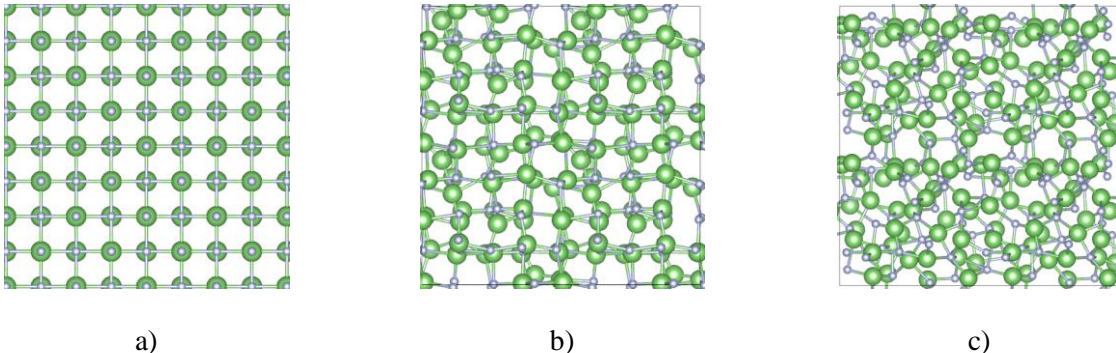

a)   b)   c)

Figure 3: Graphical representation of LiF as a) solid crystalline bulk B1 phase, b) disordered semi-crystal near the melting point of 1120 K, and c) molten LiF at 1200 K.

The test errors of NNIP vs. DFT results for all three phases are shown in Figure 4. Both the training and test errors in energy and force predictions are shown in Table 2. Figure 4a) shows the ab initio vs. neural network energy calculation for the test data, where data is colored based on configuration type: solid, semi-crystalline, and liquid salt. As shown, energy prediction is very accurate across the range of configurations and coordination environments. The error histograms fit to Gaussians are shown in Figure 4b). As temperature increases, the error distributions widen with standard deviations of 0.75, 1.0 and 1.8 eV/atom for solid, semi-crystalline, and liquid salt, respectively. This is caused by the wider distribution of energy values at higher temperatures creating a distribution in configuration-space sampling, which drives neural network prediction error. Lastly, the ab initio forces are compared with neural network calculations in Figure 4c, which again shows excellent agreement with a MAE of 0.06 eV/ Å.

Overall, the agreement between neural network and DFT calculation is excellent with all energy test errors < 2.3 meV/atom and force errors < 0.08 eV/Å in all cases. This is near the precision of the DFT calculations used to generate the reference data, where **k**-point density and plane-wave energy cutoffs were converged up to a precision of 2 meV/atom. The low error in the test set combined with the good force

prediction suggest that the NNIP is well trained to interpolate between configurations and is robust across

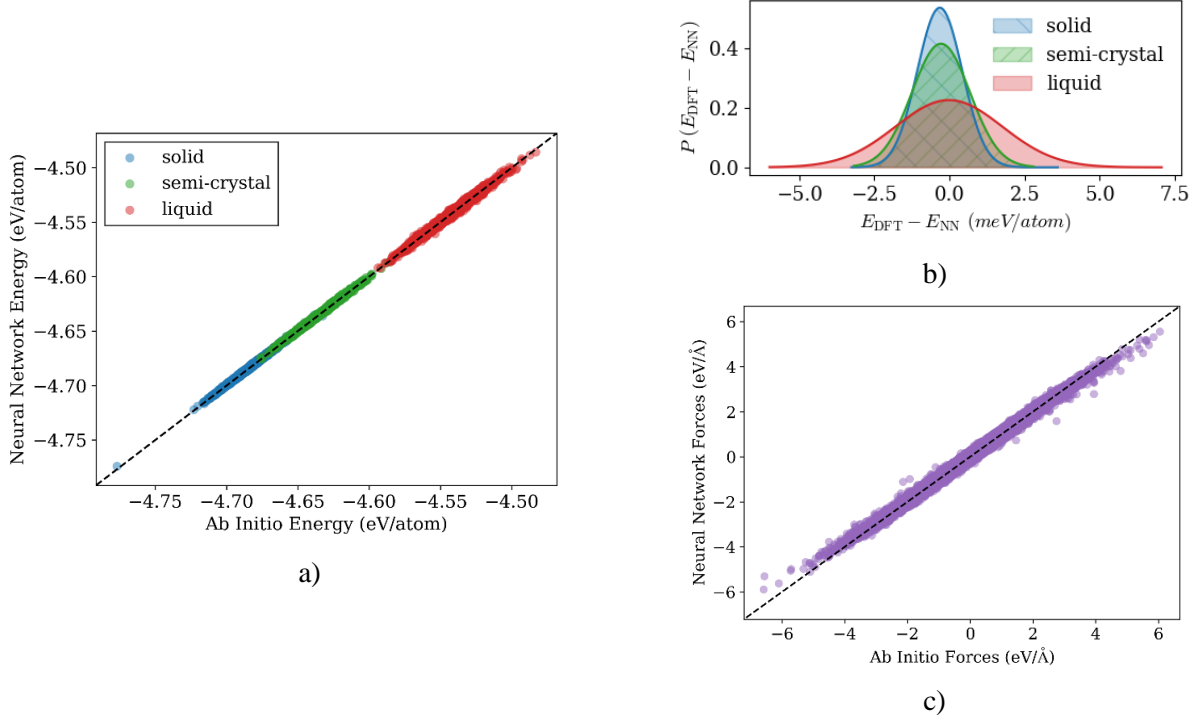a range of chemical environments.

Figure 4: LiF neural network test errors. a) Parity plot for semi-crystalline, solid and liquids calculated from neural networks and AIMD, b) normalized probability density function of energy errors, and c) force parity plot for all calculations.

Table 2: Energy and forcers training error for different phases of LiF.

	Error (MAE)	Solid	Semi-Crystalline	Liquid
Energy	Train	1.0	1.0	1.3
(meV/atom)	Test	1.6	2.3	2.1
Forces	Train	0.04	0.04	0.04
(eV/ Å)	Test	0.08	0.05	0.06

3.2 LiF-BeF₂ Potential

3.2.1 Energy and Force Prediction

For Flibe, a potential is trained by sampling Flibe configurations from both normal operating and extreme pressure and temperature conditions (973K, 20% compression at 973K, 5000 K) shown in Figure 5c). The neural network and ab initio energies and forces are compared in Figure 5, while the training and test errors are shown in Table 3. For both energies and forces, NNIP shows excellent agreement with the DFT data regardless of different atomic species or configurations. Specifically, the training MAE in energies and forces are 1.75 meV/atom and 0.081 eV/A, respectively. The test errors are slightly higher at an energy MAE of 2.10 meV/atom and force MAE of 0.084 eV/A. These energies are well within chemical precision, on the same order of errors inherent in DFT calculation due to **k**-point sampling, and approximations in the selection of a finite plane-wave basis set. The good agreement of forces in the test set indicates that the potential energy surface is not overfit since the gradients are accurately computed in the test set, which was not used to train the network. Moreover, the NNIP shows a good approximation in different structural environments as shown in Figure 5c). The formation of BeF₄ tetrahedra can be seen at 973 K, which was also found in previous studies [9][21]. At 5000 K, tetrahedra start to dissociate as result of high energy collisions. In contrast, more BeF₄ tetrahedral and tetrahedral chains emerge when Flibe is compressed due to increased density and reduced ionic mobility.

To further test the versatility and transferability of the NNIP, we calculated the formation energy of purely crystalline Flibe, which is depicted in Figure 7c) [22][23], which was not used in training. The Flibe crystal structure is shown in Figure 5c). The formation energy E_f per atom is calculated:

$$E_f = E_{\text{Flibe}} - \frac{n_{\text{F}}}{N} E_{\text{F}} - \frac{n_{\text{Li}}}{N} E_{\text{Li}} - \frac{n_{\text{Be}}}{N} E_{\text{Be}} \qquad (6)$$

where E_{flibe} is the energy per atom of the crystalline Flibe configuration, E_{F}, E_{Li} and E_{Be} are the per-atom energies of the ground-state structures of each species [24][25][26], and n_{F}, n_{Li}, n_{Be} and N are the number

of atoms of each element (24, 12 and 6 respectively) and total number of atoms in the computed Flibe cell. The DFT calculation is consistent with methods described in Section 2.2. Here, we find that the formation energies from DFT and NN are in agreement at -3.356 and -3.333 eV/atom, respectively. Since the crystal structure was never seen by the neural network during training, this suggests that the Flibe NNIP is well fit and able to extrapolate to different structures and chemical environments.

a)

b)

(1) 5000 K   **(2)** 973 K   **(3)** Compressed 973 K   **(4)** Crystalline
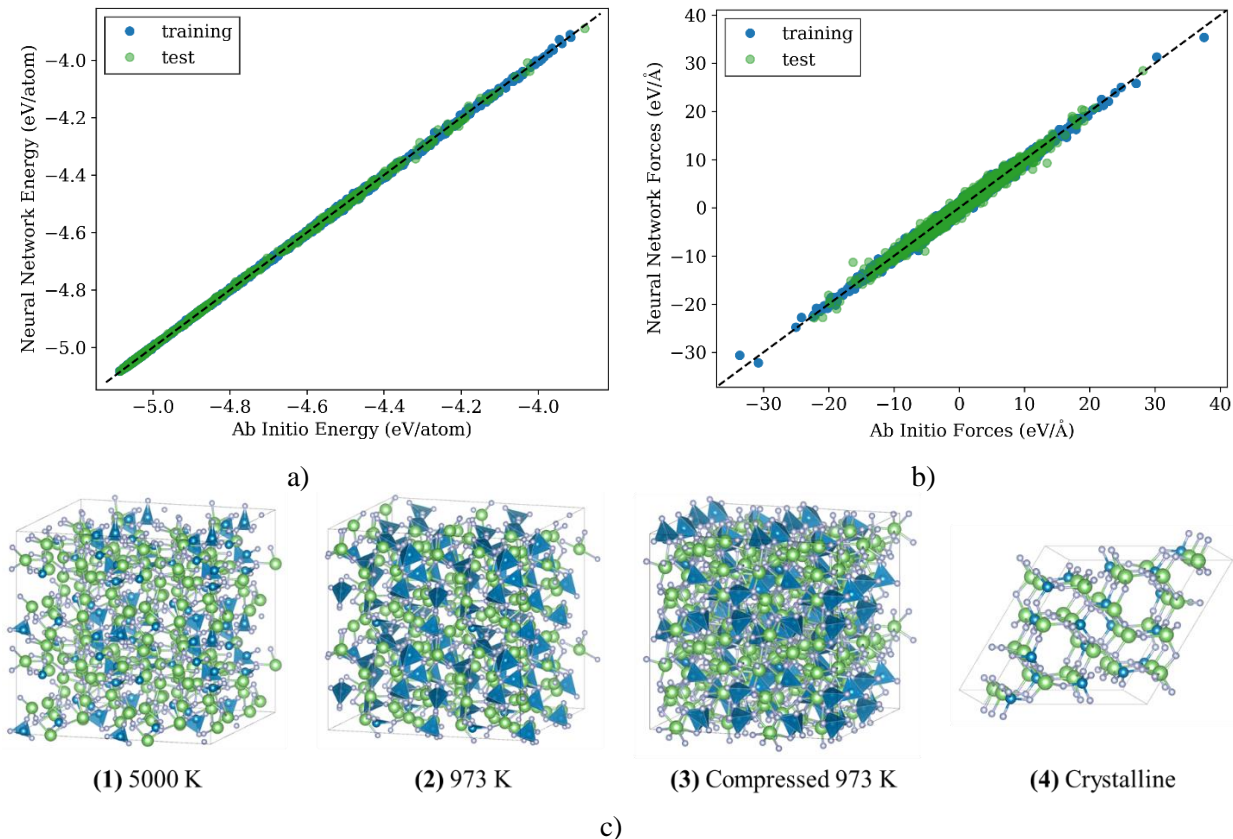
c)

Figure 5: a) NNIP vs. AIMD prediction of potential energies, b) NNP vs. AIMD prediction of forces, and c) Structures of Flibe at (1) 5000 K, (2) 973 K, and (3) 973K under 15% used to train the NNIP. Crystalline Flibe shown in (4) used in to test the trained network. Blue and green represent Be and Li, respectively.

Table 3: Neural network potential energy and force mean average error (MAE) for training, validation and test data sets.

	Train (80%)	**Test (10%)**
Energy (meV/atom)	1.75	2.10
Forces (eV/Ang)	0.081	0.084

13

The ability of the neural network potential to extrapolate over longer time scale is also tested. An MD simulation was performed in LAMMPS using the NNIP [16]. First, energy minimization was performed on the atomic configuration. Then, the system was heated and equilibrated from 100 to 973 K over 15 ps. Finally, NVT simulation was performed at 973K for 130 ps using a time step of 1 fs, which exceeds the timescale of simulations performed to generate the training data. Then snapshots along the NNIP-MD trajectory were uniformly sampled to perform static DFT calculations on each of them. The calculated energies from DFT were compared to those calculated by the NNIP, as shown in Figure 6. Compared to DFT, the neural network predictions for energy were very accurate with less than 2 meV/atom error. Forces were also accurately calculated with an error of less than 0.08 eV/A. These results are consistent with the errors found during training and testing. This shows that the NNIP is capable of interpolating ab initio potential energy surface, thus extending accurate structure sampling over a longer time scale (far beyond the time scales involved in AIMD in generating training data). This greatly increases the confidence that the neural network can be used in large length scale and long timescale molecular dynamics simulations that are used for property prediction.
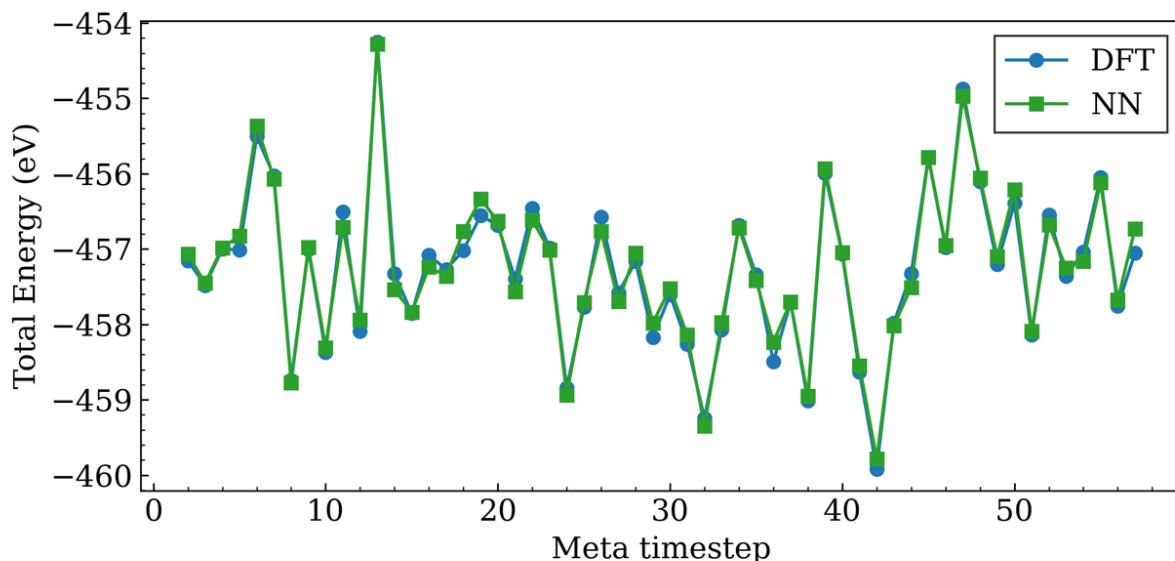

Figure 6: Neural network trajectory with each meta time step representing 2 ps. Configurations are recalculated in DFT and the ab initio energies are compared.

3.2.2 Local Structure of Flibe

To test the ability of the neural network potential to calculate disordered liquid structures in molten salts, the radial distribution function (RDF) and angular distributions were computed from NN-based MD was compared to AIMD calculation. The AIMD simulation of a system with 91 atoms (26 Li, 52 F, 26 Be) at 973K over 40 ps. For NNIP MD simulations, the simulation cell was replicated in a 3×3×3 arrangement, which resulted in a much larger simulation cell of 2457 atoms. This yielded better statistics in the RDF calculation and allowed for examination of accuracy at larger length scales that are inaccessible by AIMD. The system was thermally equilibrated at 973K over 15 ps before data was collected. Then canonical ensemble simulation with the neural network potential was performed at 973K for 50 ps. The raw radial distribution functions and angular distribution for NN and DFT are compared in Figure 7.

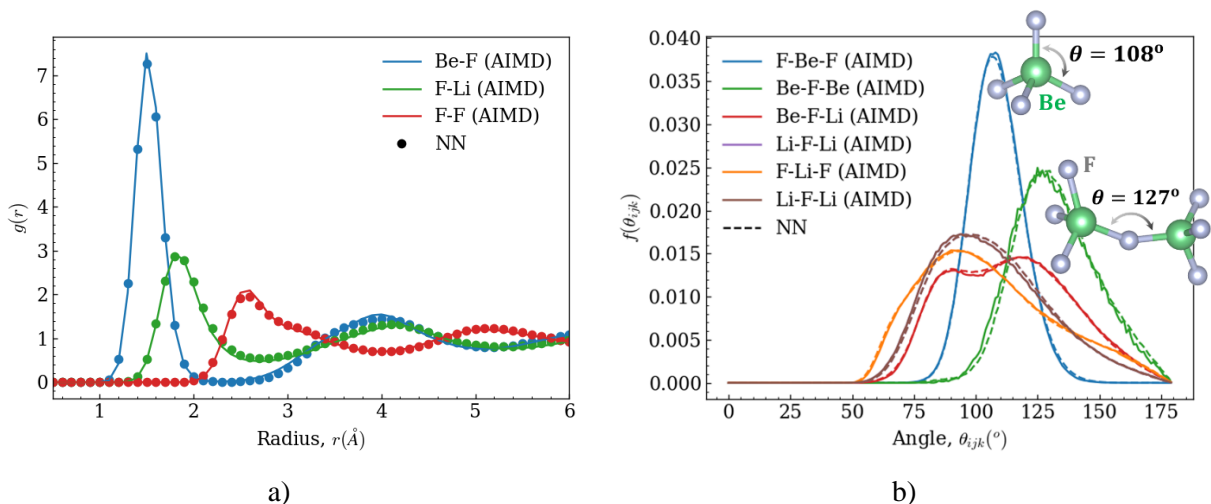a)   b)

Figure 7: a) Neural network potential and DFT radial distribution functions obtained for Be-F, Li-F and F-F pairs in Flibe. b) Angular distribution functions of Flibe.

For the RDFs in Figure 7a), NN (●) correctly predicts the DFT values (—) for all species pairs (Be-F, F-Li, F-F). NN-based RDFs follow the curvature of the DFT-based RDFs closely over the range from 0 – 6 Å, precisely matching the location and magnitude of the minima and maxima, and the peak widths and heights. This indicates that the DFT coordination distances and the coordination numbers between atomic species is well reproduced by the NNIP. Furthermore, the calculated first peak distances for Be-F, Li-F and F-F are 1.5, 1.8, and 2.5 Å, which are close to their experimental values of 1.58, 1.85, and 2.56 Å, respectively

[27]. In addition, the calculated fluorine coordination numbers (integral of first peak in the RDF) of Li and Be are 4.0 and 4.0, compared to experimental values of 4.0, and 3.7 respectively. This indicates that DFT, and consequently NNIPs can accurately capture the local environment of ions in Flibe.

To calculate the angular distribution, the angle formed by triplets were tallied based specified bond cut off distances, and the distributions were normalized $\left(\int_0^{2\pi} f(\theta)d\theta = 1\right)$ to give $f(\theta_{ijk})$, where i, j, k represent the atomic species of each atom in the triplet. The cutoff distances were specified using the first minimum of each RDF, which are 2.7 Å for Li-F, and 2.3 Å for Be-F. The angular distributions are found in Figure 7b), where the two prominent peaks are observed for F-Be-F and Be-F-Be. The peak value of F-Be-F is located at 108°, which represents a BeF_2 tetrahedral configuration, and the Be-F-Be represents corner-sharing tetrahedra with a median angle of 127°. These observations agree with our previous study [9], where the detailed analysis was performed on the local and extended structures in Flibe, and Be-F_4 tetrahedral chains were observed with up to 4 Be atom centers. As shown in Figure 7b), the three-body interactions are accurately modeled by the NN (--), with the angular distributions almost exactly overlapping the AIMD (—) results in every species combination. In classical MD potentials, this was an issue in initial rigid ion models (RIM) that were used for molten salts, which were not able to predict the induced dipole in the fluorine atom of Be-F-Be [21]. While this can be corrected by adding contributions that account for ionic polarizability, the classical approach of testing and fitting physics-based models necessarily requires significant time for model design and domain knowledge to determine the relevant effects. This is particularly difficult in cases where there are many possible species and interactions are not yet well-understood. In contrast, NN architectures provide arbitrarily rich parameter spaces, which should scale well to chemically complex systems.

Overall, these results show that the NNIP can accurately predict local structures and chemistry that are calculated by electronic structure methods like DFT, further demonstrating the robustness of the NNIP. Since a much smaller simulation cell was used to train the neural network, these tests show that the potential extrapolates well to larger systems where many more local configurations are sampled. Furthermore, due to long-range Coulomb effects and the medium-range cut off distance of NNIP (up to 7 Å), neural networks

alone might not be expected to achieve high accuracy in ionic systems. Therefore, other studies have proposed explicit treatment on charge information together with the neural network prediction [28]. Yet, this was not found to be necessary in molten salts as shown in this work, and previous work on NaCl molten salt [29]. The authors believe that this is likely due to the charge screening from local ionic solvation shells that limit long-range interactions.

3.2.3 Ionic Diffusion

Modeling atomic transport behavior is of practical relevance in molten salt applications, as it enables the prediction of important properties like viscosity, heat conduction and species diffusion. Previous efforts have found that AIMD can accurately predict ionic self-diffusivity for a variety of binary and ternary fluoride salts [9]. This provides motivation for using a neural network potential to model the more expensive ab initio PES to calculate diffusivity and perform molecular dynamics for understanding transport behavior. The self-diffusivity is calculated using the block-averaging method, from the slope of the mean squared displacement (MSD) as a function of time [30]. Here, the Einstein relationship is used:

$$D = \frac{1}{6} \lim_{t \to \infty} \frac{d}{dt} (\text{MSD}) = \frac{1}{6} \lim_{t \to \infty} \frac{d}{dt} \left(\frac{1}{N} \frac{1}{n_t} \sum_{j=1}^{n_t} \sum_{i=1}^{N} \left(r_i(t_j + dt) - r_i(t_j) \right)^2 \right) \tag{7}$$

where MSD is determined for and averaged over subsets n_t of the total trajectory, and over all the atoms N of the same elemental type, $r_i(t)$ are the coordinates of atomic i at time t, and dt is the length of time spanned by each subset in time. For this calculation, a 91 atom simulation is used to allow for direct comparison between DFT and NN. The system is equilibrated at 973K for 10 ps, and the simulation is performed over 70 ps using both NN potentials and AIMD. The calculated diffusivities for both NN and AIMD are shown in Table 4, where the uncertainty values represent the 95% confidence intervals the predictions.

Table 4: Neural network potential versus AIMD-calculated self-diffusivity of ions in Flibe at 973K.

	$D_{NN} \times 10^{-5}(\text{cm}^2\text{s}^{-1})$	$D_{AIMD} \times 10^{-5}(\text{cm}^2\text{s}^{-1})$
Be	0.87 ± 0.13	0.83 ± 0.10
F	1.93 ± 0.19	1.73 ± 0.17
Li	5.12 ± 0.50	5.67 ± 0.52

In all cases (Be, F, and Li), the relative magnitudes of ionic diffusivities are reproduced and agree qualitatively with expectations based on the known structure and dynamics of Flibe [31][32]. Beryllium diffusivity is lowest at $0.83 - 0.87 \times 10^{-5}$ cm^2/s due to the presence of BeF$_4$ tetrahedral polymer-like chains shown in Figure 5c) and 7b). In contrast, the LiF in Flibe (66.6%LiF-33.3%BeF$_2$) is found to dissociate, producing free fluorine and lithium ions. As a result, the fluorine diffusivity is higher (some fluorine still bound to Be) at $1.73 - 1.93 \times 10^{-5}$ cm^2/s, and the lithium diffusivity is the highest at $5.12 - 5.67 \times 10^{-5}$ cm^2/s. Further, the neural network prediction falls within the uncertainty ($\pm 10 - 15\%$) for all cases, which is within the uncertainty of typical experiments. This provides a good indication that NNIP dynamics are accurate and bound only by the accuracy of the ab initio functionals used to train the model. Given reliable DFT functionals, NNIPs should reasonably be capable of calculating more demanding dynamics and transport properties such as viscosity or heat transfer coefficient.

3.2.4 Computational Performance of Neural Network Potential

The computational demand for DFT is compared against the neural network potential for Flibe in this section. Simulations for this test were performed on the Engaging cluster in the Massachusetts Green High Performance Computing Center (MGHPCC) facility and the Frontera Cluster in the Texas Advanced Computing Center (TACC). The Engaging cluster consists of nodes with 2×16 Intel Xeon 2.1 GHz cores, connected with an Infiniband interconnect with bandwidth of 4×14 Gb/s node-to-node transfer and 0.7 ms of latency. On Frontera, nodes consisted of 28×2 Intel Xeon Platinum 8280 2.7 GHz processors connected via a HDR-100 Mellanox InfiniBand. Parallelization in DFT is performed using VASP recommendations

based on machine architecture, with 32 cores (one node) operating on each orbital [12]. MD with neural network potentials are performed in a parallel installation of LAMMPS [33].

The computational demand of neural networks is compared to DFT as shown in Figure 8. The required core-hours to run 1000 time steps of molecular dynamics simulation is shown for Flibe systems of different sizes. As expected, DFT scales roughly as $\sim O(N^3)$ where N is the number of atoms. In contrast, the neural network potentials scale as $\sim O(N)$. Thus, the computational gain from using a neural network potential grows with increasing system size. With a system with 238 atoms in a 14.2×14.2×14.2 Å^3, periodic cell effects are very small and a variety of calculated thermokinetic properties are generally expected to be well converged [34]. For a system of this size, the required calculation time is 3379.0 CPU-hours for 1000 DFT time steps. In comparison, the same calculations can be done using the neural network potential with 5.3 CPU-hours, which is 630 times faster.
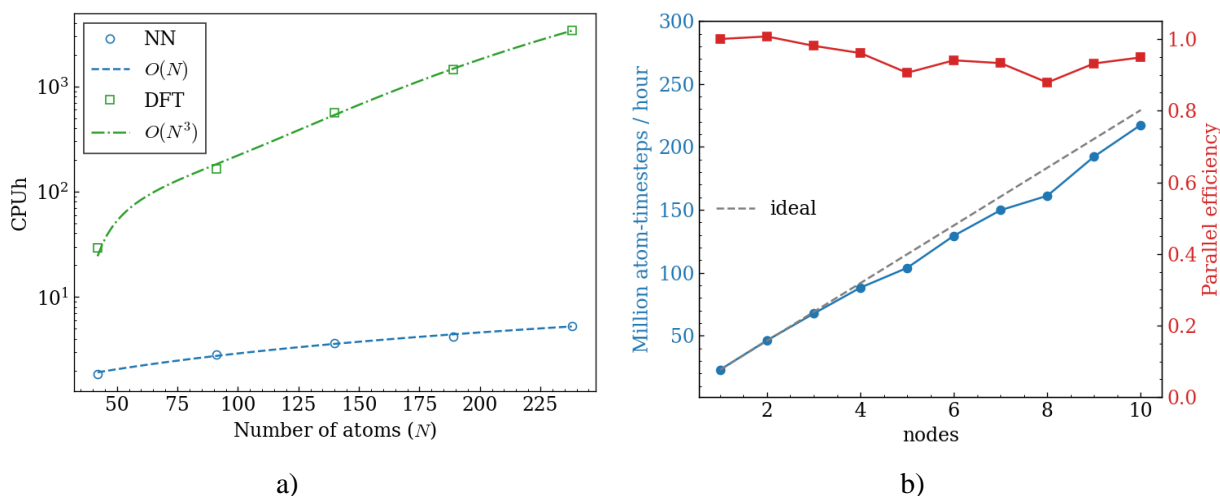

Figure 8: a) CPU-hours for 1000 MD time steps versus number of atoms, b) computational speedup for Flibe system with 91,000 atoms. Each node represents 56 cores on the TACC Frontera system. The black dashed line represents ideal efficiency.

In addition to scaling performance with respect to N (atoms), a strong scaling test across multiple processors was also tested by performing MD simulations using 56 – 560 cores. The results are shown in Figure 8b). The system simulated is Flibe at 973 K with 91,000 atoms. At each time step, the evaluation of energies and forces requires relatively simple matrix multiplication and algebraic operations required by

the calculation of forward pass of the neural network. This is performed via a LAMMPS interface plugin provided by PANNA, which is written in C++ [14]. Thus, the parallelization shown in Figure 8b), benefits from LAMMPS optimization, which provides robust spatial decomposition algorithms using message passing interface (MPI) parallelization [35][33]. As shown, scaling is near-ideal (> 98%) if more than 30,000 atoms are used per node. The efficiency drops slightly if more nodes are used. At 9,100 atoms per node (10 nodes), the efficiency is 94%. These results are consistent with tests performed for other packages [35]. This demonstrates that the calculations generally scale well with the number of cores used for a large simulation. Further, additional speed improvements in the future can likely be gained by pruning the network to reduce the total number of floating-point operations per energy calculation or by looking at different parallelization strategies. This includes optimizing a hybrid MPI (distributed memory) and OpenMP (shared memory) parallelization scheme for calculations using many processors, or utilizing graphics-based acceleration methods for significantly larger systems [36].

From our previous study, the calculation of salt properties such as equation of state or diffusion activation energy can require tens-to-hundreds of thousands of DFT calculations [9]. In such cases, the use of robust and flexible NNIP molecular dynamics provides more than 3 orders of magnitude computational time reduction over AIMD, reducing the time to determine a single thermophysical property from several weeks to hours. Therefore, the NNIP could greatly increase the efficiency of high-throughput calculation of molten salt properties. Further, larger systems with more interesting properties like radiation damage or defect chemistry can potentially be studied.

4 Conclusions

Robust and versatile NNIPs have been developed for the common salt constituent LiF, and a prototype salt Flibe (66%LiF-33%BeF2). We showed that a diverse set of local atomic environments could be captured using radial and angular symmetry functions based on a sum of Gaussian functions and used as inputs to train versatile interatomic potentials. For LiF, we showed that a NNIP can simultaneously capture dimer interactions (F-F, Li-Li, Li-F), deformed solids at 0K, solids near the melting point, through

20

the melting transition and towards high-temperature liquids with DFT-level accuracies of < 2 meV/atom. Further, the LiF NNIP can accurately predict energies of surfaces that were not used for training, suggesting that neural networks could properly extrapolate to configurations differing from experience. Such behavior is highly desirable in NNIPs, where the physical interactions are not explicitly encoded, and the model parameters are difficult to interpret. The Flibe NNIP trained is able to simulate salt at 973 K (operating temperature of most reactor systems), compressed Flibe up to 20%, and high temperature salt exceeding 5000 K. By introducing structures under extreme conditions, configurational sampling was greatly improved. The Flibe NNIP has been integrated into the LAMMPS [16] molecular dynamics engine using the PANNA package [14] and shows highly consistent accuracy for both extended timescale and length scale. Particularly, the NNIP structural (local coordination environment) and transport property prediction of Flibe is consistent with DFT and previous experimental results.

In this work, NNIPs have been shown highly adaptable to a variety of local environments. This feature is very desirable for the applications to real systems which involve many components and span wide thermodynamic conditions. This demonstrates a significant advantage over classical interatomic potentials that are typically fit for a specific application. Due to its ability to approximate arbitrarily complex functions, neural networks enable robust modeling of interatomic interactions of arbitrarily complex systems. Moreover, these potentials are able to accelerate property computations by more than three orders of magnitude, and far beyond the accessible time and length scales of ab initio methods, while showing strong scaling efficiency with a large number of processors. As such, neural networks are demonstrated to provide a promising direction for complex multi-component ionic liquid systems such as molten salts that would be difficult to study using either ab initio or classical methods alone. In future work, these potentials will be applied to study increasingly complex systems, which could enable dramatically accelerated screening and allow unprecedented interpolation across compositional and real operating space.

5 References

[1] Y. Wang *et al.*, "Current understanding and challenges of solar-driven hydrogen generation using polymeric photocatalysts," *Nat. Energy*, vol. 4, no. 9, pp. 746–760, 2019, doi: 10.1038/s41560-019-0456-5.

[2] S. Frangini and A. Masi, "Molten carbonates for advanced and sustainable energy applications: Part II. Review of recent literature," *Int. J. Hydrogen Energy*, vol. 41, no. 42, pp. 18971–18994, 2016, doi: 10.1016/j.ijhydene.2016.08.076.

[3] T. Voisin, A. Erriguible, and C. Aymonier, "A new solvent system: Hydrothermal molten salt," *Sci. Adv.*, vol. 6, no. 17, 2020, doi: 10.1126/sciadv.aaz7770.

[4] S. Bell, T. Steinberg, and G. Will, "Corrosion mechanisms in molten salt thermal energy storage for concentrating solar power," *Renew. Sustain. Energy Rev.*, vol. 114, p. 109328, 2019, doi: https://doi.org/10.1016/j.rser.2019.109328.

[5] S. Delpech, C. Cabet, C. Slim, and G. S. Picard, "Molten fluorides for nuclear applications," *Mater. Today*, vol. 13, no. 12, pp. 34–41, 2010, doi: 10.1016/S1369-7021(10)70222-4.

[6] A. L. Rollet and M. Salanne, "Studies of the local structures of molten metal halides," *Annu. Reports Prog. Chem. - Sect. C*, vol. 107, pp. 88–123, 2011, doi: 10.1039/c1pc90003j.

[7] D. F. Williams, L. M. Toth, and K. T. T. Clarno, "Assessment of Candidate Molten Salt Coolants for the Advanced High-Temperature Reactor (AHTR)," Oak Ridge National Laboratory, Oak Ridge, Tennessee, 2006. doi: ORNL/TM-2006/69.

[8] M. Salanne, B. Rotenberg, S. Jahn, R. Vuilleumier, C. Simon, and P. A. Madden, "Including many-body effects in models for ionic liquids," *Theor. Chem. Acc.*, vol. 131, no. 3, p. 1143, 2012, doi: 10.1007/s00214-012-1143-9.

[9] S. T. Lam, Q.-J. Li, J. P. Mailoa, C. Forsberg, R. Ballinger, and J. Li, "The Impact of Hydrogen

Valence on Its Bonding and Transport in Molten Fluoride Salts," *J. Mater. Chem. A*, 2020, doi: 10.1039/d0ta10576g.

[10] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Phys. Rev.*, vol. 140, no. 4A, pp. A1133----A1138, Nov. 1965, doi: 10.1103/PhysRev.140.A1133.

[11] T. D. Kuhne, "Ab initio molecular dynamics," *Methods Mol. Biol.*, vol. 924, no. 1, pp. 29–42, 2013, doi: 10.1007/978-1-62703-17-5_2.

[12] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Phys. Rev. B - Condens. Matter Mater. Phys.*, vol. 54, no. 16, pp. 11169–11186, 1996, doi: 10.1103/PhysRevB.54.11169.

[13] J. S. Smith, O. Isayev, and A. E. Roitberg, "ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost," *Chem. Sci.*, vol. 8, no. 4, pp. 3192–3203, Apr. 2017, doi: 10.1039/c6sc05720a.

[14] R. Lot, F. Pellegrini, Y. Shaidu, and E. Küçükbenli, "PANNA: Properties from Artificial Neural Network Architectures," *Comput. Phys. Commun.*, vol. 256, p. 107402, 2020, doi: https://doi.org/10.1016/j.cpc.2020.107402.

[15] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," pp. 1–15, 2014, [Online]. Available: http://arxiv.org/abs/1412.6980.

[16] S. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," *J. Comput. Phys.*, vol. 117, no. 1, pp. 1–19, Mar. 1995, doi: 10.1006/jcph.1995.1039.

[17] A. Jain and R. C. Dixit, "B 1 to B 2 structural phase transition in LiF under pressure," *AIP Conf. Proc.*, vol. 1953, no. May 2018, 2018, doi: 10.1063/1.5032633.

[18] G. J. Finch and S. Fordham, "The effect of crystal-size on lattice-dimensions," *Proc. Phys. Soc. London*, vol. 48, pp. 85–94, 1936.

[19] M. Rubbo, M. Bruno, and M. Prencipe, "Thermodynamic study of reconstructed crystal surfaces.: The octopolar (111) face of LiF crystals," *Surf. Sci.*, vol. 632, pp. 180–184, 2015, doi: 10.1016/j.susc.2014.10.015.

[20] B. M. E. Van Der Hoff and G. C. Benson, "A quantum-mechanical calculation of the surface energy of crystalline lithium fluoride," *J. Chem. Phys.*, vol. 22, no. 3, pp. 475–480, 1954, doi:

10.1063/1.1740093.

[21]   M. Salanne, C. Simon, P. Turq, N. Ohtori, and P. A. A. Madden, "Modeling of Molten Salts," in *Molten Salts Chemistry*, 1st ed., no. i, Elsevier Inc., 2013, pp. 1–16.

[22]   D. M. Collins, M. C. Mahar, and F. W. Whitehurst, "Bonding in lithium tetrafluoroberyllate(II)," *Acta Crystallogr. Sect. B Struct. Sci.*, vol. 39, pp. 303–306, 1983.

[23]   P. Seiler, "Estimation of ionicity coefficients in Li2 Be F4 crystals by X-ray diffraction," *Acta Crystallogr. Sect. B Struct. Sci.*, vol. 49, pp. 223–235, 1993.

[24]   K. Lejaeghere, V. Van Speybroeck, G. Van Oost, and S. Cottenier, "Error estimates for solid-state density-functional theory predictions: an overview by means of the ground-state elemental crystals," *Crit. Rev. Solid State Mater. Sci.*, vol. 39, pp. 1–24, 2014.

[25]   A. W. Overhauser, "Crystal structure of lithium at 4.2 K," *Phys. Rev. Lett.*, vol. 53, no. 1, pp. 64–65, 1984, doi: 10.1103/PhysRevLett.53.64.

[26]   F. K. Larsen and N. K. Hansen, "Diffraction study of the electron density distribution in beryllium metal," *Acta Crystallogr. Sect. B Struct. Sci.*, vol. 40, pp. 169–179, 1984.

[27]   F. Vaslow and A. H. Narten, " Diffraction pattern and structure of molten BeF 2 –LiF solutions ," *J. Chem. Phys.*, vol. 59, no. 9, pp. 4949–4954, 1973, doi: 10.1063/1.1680711.

[28]   T. Morawietz, V. Sharma, and J. Behler, "A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges," *J. Chem. Phys.*, vol. 136, no. 6, 2012, doi: 10.1063/1.3682557.

[29]   J. L. Qingjie Li, Emine Kucukbenli, Stephen Lam, Boris Khaykovich, Efthimios Kaxiras, "A neural network interatomic potential for molten NaCl," *Bull. Am. Phys. Soc.*, vol. 65, 2020.

[30]   D. C. Rapaport and D. C. R. Rapaport, *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004.

[31]   R. J. Heaton, R. Brookes, P. A. Madden, M. Salanne, C. Simon, and P. Turq, "A First-Principles Description of Liquid BeF2 and Its Mixtures with LiF: 1. Potential Development and Pure BeF2," *J. Phys. Chem. B*, vol. 110, no. 23, pp. 11454–11460, 2006, doi: 10.1021/jp061000+.

[32]   M. Salanne *et al.*, "A first-principles description of liquid BeF2 and its mixtures with LiF: 2. Network formation in LiF-BeF2," *J. Phys. Chem. B*, vol. 110, no. 23, pp. 11461–11467, 2006, doi: 10.1021/jp061000+.

[33] S. Plimpton, "Short-Range Molecular Dynamics," *J. Comput. Phys.*, vol. 117, no. 6, pp. 1–42, 1997, doi: 10.1006/jcph.1995.1039.

[34] H. O. Nam, A. Bengtson, K. Vörtler, S. Saha, R. Sakidja, and D. Morgan, "First-principles molecular dynamics modeling of the molten fluoride salt with Cr solute," *J. Nucl. Mater.*, vol. 449, no. 1–3, pp. 148–157, 2014, doi: 10.1016/j.jnucmat.2014.03.014.

[35] A. Singraber, J. Behler, and C. Dellago, "Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials," *J. Chem. Theory Comput.*, vol. 15, no. 3, pp. 1827–1840, 2019, doi: 10.1021/acs.jctc.8b00770.

[36] J. Glaser *et al.*, "Strong scaling of general-purpose molecular dynamics simulations on GPUs," *Comput. Phys. Commun.*, vol. 192, pp. 97–107, 2015, doi: 10.1016/j.cpc.2015.02.028.