JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

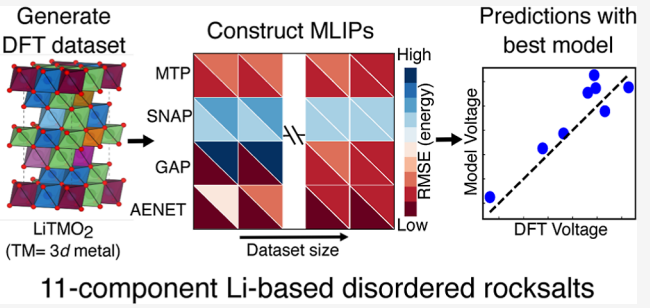
Constructing and Evaluating Machine-Learned Interatomic Potentials for Li-Based Disordered Rocksalts

Published as part of Journal of Chemical Theory and Computation virtual special issue "Machine Learning and Statistical Mechanics: Shared Synergies for Next Generation of Chemical Theory and Computation".

Vijay Choyal, Nidhish Sagar, and Gopalakrishnan Sai Gautam*

Cite This: https://doi.org/10.1021/acs.jctc.4c00039 Read Online

ABSTRACT: Lithium-based disordered rocksalts (LDRs), which are an important class of positive electrode materials that can increase the energy density of current Li-ion batteries, represent a significantly complex chemical and configurational space for conventional density functional theory (DFT)-based high-throughput screening approaches. Notably, atom-centered machine-learned interatomic potentials (MLIPs) are a promising pathway to accurately model the potential energy surface of highly disordered chemical spaces, such as LDRs, where the performance of such MLIPs has not been rigorously explored yet. Here, we represent a comprehensive evaluation of the accuracy, transferability, and ease of training of five atom-centered MLIPs, including the artificial neural network potentials developed by the atomic energy network (AENET), the Gaussian approximation potential (GAP), the spectral neighbor analysis potential (SNAP) and its quadratic extension (qSNAP), and the moment tensor potential (MTP), in modeling a 11-component LDR chemical space. Specifically, we generate a DFT-calculated data set of 10,842 configurations of disordered $LiTMO_2$ and TMO_2 compositions, where TM = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, and/or Cu. To provide a point-of-comparison on the performance of atom-centered MLIPs, we also trained the neural equivariant interatomic potential (NequIP) on a subset of our data. Importantly, we find AENET to be the best potential in terms of accuracy and transferability for energy predictions, while MTP is the best for atomic forces. While AENET is the fastest to train among the MLIPs considered at low number of epochs (300), the training time increases significantly as epochs increase (3300), with a corresponding reduction in training errors (~60%). Note that AENET and GAP tend to overfit in small data sets, with the extent of overfitting reducing with larger data sets. Finally, we observe AENET to provide reasonable predictions of average Li-intercalation voltages in layered, single-TM $LiTMO_2$ frameworks, compared to DFT (~10% error on average). Our study should pave the way both for discovering novel disordered rocksalt electrodes and for modeling other configurationally complex systems, such as high-entropy ceramics and alloys.



INTRODUCTION

Lithium-ion batteries (LIBs) are the workhorse energy storage technology that is powering modern portable electronics and electric vehicles, and remains a crucial ingredient in our transition to a carbon-free society.[1−3] Among the various components that constitute LIBs, the cathode or the positive electrode, which is usually made of lithium, combinations of few transition metals (TMs), and oxygen resulting in a $LiTMO_2$ composition,[4,5] is most critical in determining the energy density and cost of the overall electrochemical cell.[1,6,7] Thus, increasing the performance of cathode materials is an ongoing and active area of research. One pathway to increase the energy density of LIB cathodes is to utilize anionic redox (i.e., oxygen redox), which can be reversibly accessed in cathode materials with a Li-excess composition, i.e., $Li_{1+x}TM_{1-y}O_2$, where x and y are positive numbers less than 1.[8,9] Importantly, cathode frameworks that adopt a Li-excess composition along with a fluorinated disordered rocksalt structure (i.e., DRX cathodes) have displayed significant promise, with reversible and facile anionic redox, high voltages with minimal volume changes during charge/discharge, and macroscopic Li-ion transport via percolating networks.[10−21] Thus, designing and optimizing DRX cathodes can significantly improve the energy density and (potentially) cost of LIBs.[22]

Received: January 10, 2024
Revised: April 29, 2024
Accepted: May 15, 2024

The structure of a conventional $LiTMO_2$ cathode, such as $LiCoO_2$, is a layered framework, with Li-ions (green polyhedra) and TM-ions (blue polyhedra) occupying distinct layers along the *c*-axis, as shown in Figure 1a. Such layered

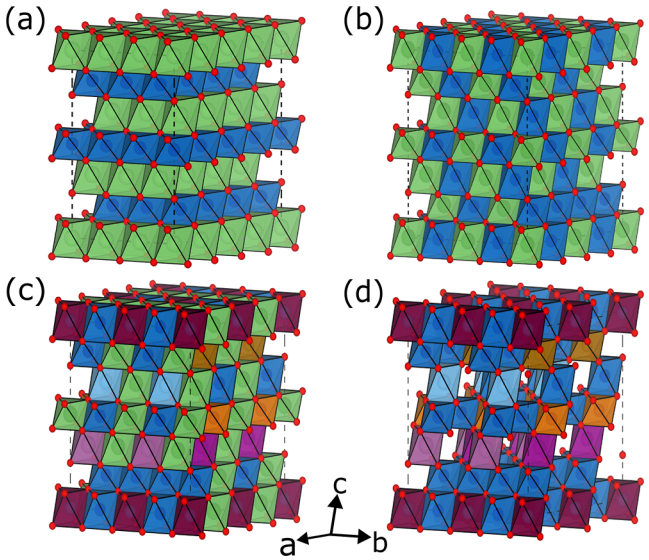


Figure 1. Structural schematic of (a) layered $LiTMO_2$, (b) disordered, single-TM $LiTMO_2$, (c) disordered multi-TM $LiTMO_2$, and (d) disordered, delithiated, multi-TM TMO_2. Green spheres/polyhedra indicate Li, red spheres indicate O, and TMs are indicated by other colors of spheres/polyhedra.

frameworks consist of a face centered close-packed anionic lattice (red spheres indicate anions in Figure 1), with Li and TM cations occupying octahedral voids of adjacent close-packed {111} planes, which results in distinct Li and TM layers.[5] Thus, the layered $LiTMO_2$ cathode is a specific ordering of the rocksalt structure, created by the anionic (O) and cationic (Li + TM) sublattices. Note that the Li and TM ordering within a layered cathode can undergo disordering, i.e., large fractions of Li sites can be occupied by TM and vice versa, resulting in the breaking of long-range order, as displayed in Figure 1b.[8] Note that disordering can occur in cathodes with multiple TMs also (see Figure 1c), and subsequent topotactic Li deintercalation can result in a disordered TMO_2 configuration as well (Figure 1d). However, disordered cathodes with a stoichiometric Li content (i.e., $LiTMO_2$) do not show appreciable electrochemical activity,[23−25] since there are not enough connected Li-conducting pathways available within the structure for macroscopic Li transport,[26] prompting the use of Li-excess compositions.[10] Nevertheless, $LiTMO_2$ remains the base or stoichiometric composition around which the Li, TM, and anionic content (e.g., fluorination) is changed for designing energy-dense DRX.

Off late, materials screening approaches using high-throughput density functional theory (DFT[27,28]) calculations with/without machine learning (ML), have resulted in a significant number of theory-predicted candidates in a variety of applications, including batteries.[10,29−42] Indeed, some of the theoretical predictions have also been validated by subsequent experiments.[43−49] However, modeling disordered rocksalt compositions, and subsequently performing a computational screening across various compositions is nontrivial, owing to the configurational complexity and length-scale of the system. Specifically, disordered rocksalts do not have significant long-range order, necessitating large supercells, which in turn results in multiple symmetrically distinct Li-TM arrangements to consider. The computational complexity becomes particularly severe in case of disordered structures containing multiple TMs, which is usually the case for several of the Li-excess cathodes that show good performance.[8] Thus, a conventional high-throughput screening approach is not practical for new DRX discovery, and this also applies for the general field of high entropy ceramics.[50,51]

An alternative method to model disordered multicomponent systems such as disordered rocksalt cathodes is to employ the so-called machine-learned interatomic potentials (MLIPs), which can provide "quick" and "accurate" estimates of energies and atomic forces within a given disordered configuration.[52−54] Specifically, MLIPs are usually trained on a (smaller) DFT-based training set and act as a mathematical approximation of the underlying potential energy surface (PES)[54] of the chemical system under consideration. Conventionally, MLIPs have been constructed on an atom-centered basis, i.e., the MLIP fingerprints the local environment around an atom-of-interest, and are designed to ensure invariance with rotation, reflection, translation, and permutations of the underlying atoms. Once constructed and validated, an MLIP can be used for larger-size and longer-time-scale simulations.

Examples of atom-centered MLIPs include the artificial neural network (ANN) potential[55,56] developed as part of the atomic energy network (AENET) package,[33,57,58] the Gaussian approximation potential (GAP),[54,59,60] the moment tensor potential (MTP),[61−63] the spectral neighbor analysis potential (SNAP), and its quadratic version (qSNAP).[64,65] Training such potentials typically involves the generation of a sufficiently diverse data set (usually with DFT), minimizing the root mean squared errors (RMSEs) and/or the mean absolute errors (MAEs) against target metrics (such as total energies and atomic forces), and optimization of relevant hyperparameters. These MLIPs are a significant improvement over classical force fields, in terms of displaying near-DFT accuracy and being sufficiently quick in predicting energies, forces, and stresses.[33,66−69] While the performance of these potentials has been compared for single-component systems,[52] a rigorous benchmarking and testing of these potentials on a disordered system with a large number of components is missing so far.

Here, we perform a comprehensive analysis of the accuracy, transferability, and ease of training of five MLIPs, including MTP, SNAP, qSNAP, GAP, and AENET, on a highly diverse, 11-component data set involving disordered $LiTMO_2$ compositions (TM = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, and/or Cu). Thus, the primary objective of this work is to benchmark the performance of the aforementioned MLIPs on a 11-component configurationally complex data set. To train and test the potentials, we generate a DFT-calculated data set of 10,842 configurations involving different compositions of $LiTMO_2$, which was divided into training and testing data subsets of varied sizes. Subsequently, we quantify the accuracy via errors on energies and forces within training data sets, transferability via the similarity of errors across the training and corresponding testing sets, and ease of training through the computational training time required for each potential. We also train a neural equivariant interatomic potential (NequIP) on a subset of our total data set to compare the performance of a graph-network-based potential versus the atom-centered MLIPs considered. Importantly, we find that AENET provides

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

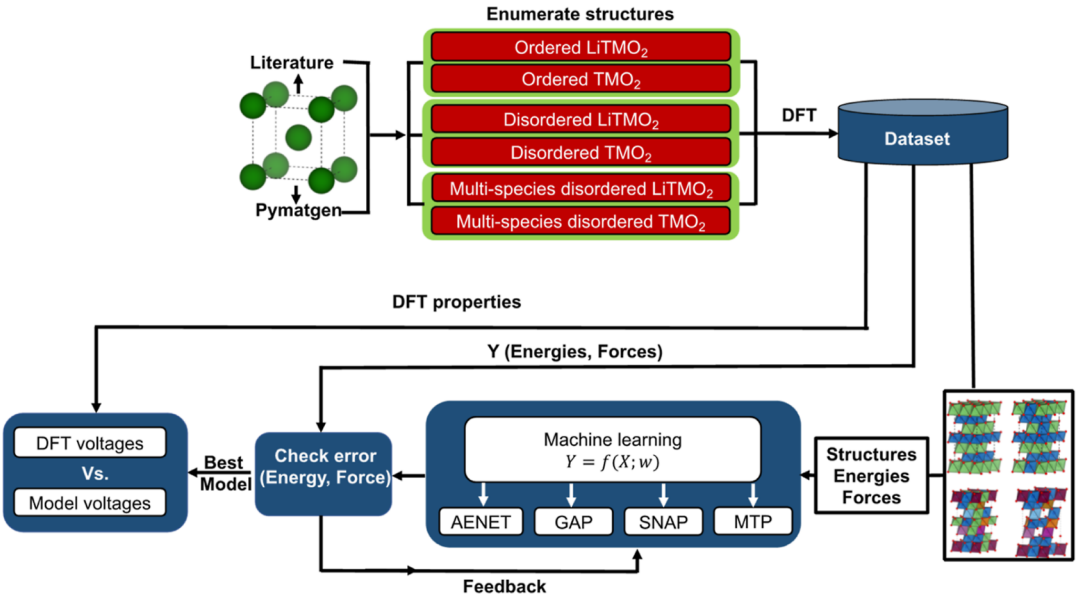


Figure 2. Workflow of structure enumeration, DFT calculations, MLIP training, and voltage predictions employed in this work. For all MLIPs, the inputs are the DFT-calculated structures, energies, and forces, and the outputs are energies and forces.

Table 1. Summary of Different MLIP Frameworks Considered in This Work[a]

MLIP	MTP	SNAP	GAP	AENET
idea	many-body interactions within a cutoff radius represented via moment tensors	local atomic density (weighted delta functions) projected on a 4D hypersphere	local atomic density modeled via a SOAP kernel	feed forward neural network with local bonding environment as the input layer
descriptor	moment tensors consisting of radial distribution function and outer products of position vectors of neighboring atoms	hyperspherical projection of atomic density expanded in terms of bispectrum components	atomic density as weighted sum of Gaussians	radial and angular distribution functions
training algorithm	BFGS	linear regression	Gaussian process regression	limited memory BFGS
basis functions	Chebyshev polynomials (for radial basis) and contracted moment tensors	hyperspherical harmonics	equispaced Gaussians (for radial basis) and spherical harmonics (for angular basis)	Chebyshev polynomials

[a]SOAP and BFGS represent smooth overlap of atomic positions[59,77] and Broyden−Fletcher−Goldfarb−Shanno,[78] respectively.

the best accuracy and transferability for total energies and is the easiest to train for a low number of epochs, while MTP is the best for atomic forces, exhibits robust transferability for total energies, and is difficult to train, highlighting that different MLIPs may be suited for different target metrics/applications. To further probe the accuracy of AENET on derived properties, we benchmark the potential-calculated average Li intercalation voltages versus DFT-estimates in ordered, layered, single-TM $LiTMO_2$ compositions (TM = Ti, V, Cr, Mn, Fe, Co, Ni, or Cu), with AENET showing a reasonable MAE ∼ 0.34 V (∼10%), across all systems. We hope that our study and the potentials we have constructed provide new avenues for discovering novel DRX cathodes for LIBs, and a framework for screening through computationally complex, disordered, multicomponent systems.

METHODS

Workflow. An overview of the data generation, calculations employed, training of MLIPs, and quantifying predictive errors is shown in Figure 2. DFT total energies and atomic forces of six categories of training structures, i.e., structures involving single or multiple TMs, with or without Li, and ordered or disordered configurations, were calculated to create the training (and test) data sets (see Data Set Generation subsection for additional details). A 90:10 random split of the data set was used as training and test data sets, respectively, for the construction of all MLIPs. Note that the same random split(s) of structures was used for training/testing of all MLIPs. Subsequently, atomic descriptors (see Table 1 and Constructing MLIPs subsection) were generated for the structures in the training set, which were used to train the MLIPs with the DFT total energies and/or atomic forces being the target properties. The hyperparameters for each MLIP framework were optimized during the training process to provide the minimum RMSE and MAE with respect to DFT-calculated energies and/or forces and are provided in Table S1 of the Supporting Information. Among the trained MLIPs, we chose the best MLIP (i.e., the AENET potential), which gave the lowest RMSE in energies and performed a comparison of the model-predicted Li-intercalation voltages versus DFT calculations in ordered, layered, single-TM $LiTMO_2$-TMO_2 compositions.

Data Set Generation. The $LiTMO_2$ disordered rocksalt data set was generated by enumerating Li and TM arrangements within the cation sublattice of the rocksalt structure. Specifically, we used the structures enumerated by Artrith et al.[33] with various cation arrangements up to a total of 18 cation sites per cell using the enumeration approach of Hart et al.[70−72] Thus, Artrith et al.[33] generated a total of 10,046 structures based on nine TMs (i.e., TM = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, and Cu). Subsequently, we generated specific

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

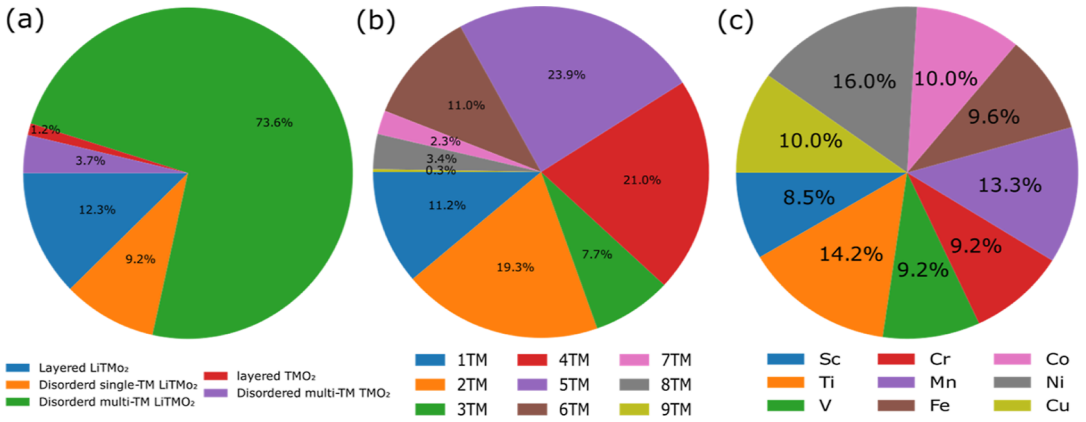


Figure 3. Pie charts indicating the various splits within the data set used. (a) Split across various layered, and disordered $LiTMO_2$ and TMO_2, (b) the percentage of multi-TM combinations among disordered-$LiTMO_2$, and (c) the percentage of each TM present in disordered $LiTMO_2$ structures. The notation "1TM" in panel (b) indicates a single TM (out of Sc, Ti, V, Cr, Mn, Fe, Co, Ni, or Cu) being present in a given structure. "2TM" (3TM and so on) indicate possible combinations of two TMs (three TMs and so on) being present in a given structure, such as Ti + V, Co + Ni, etc.

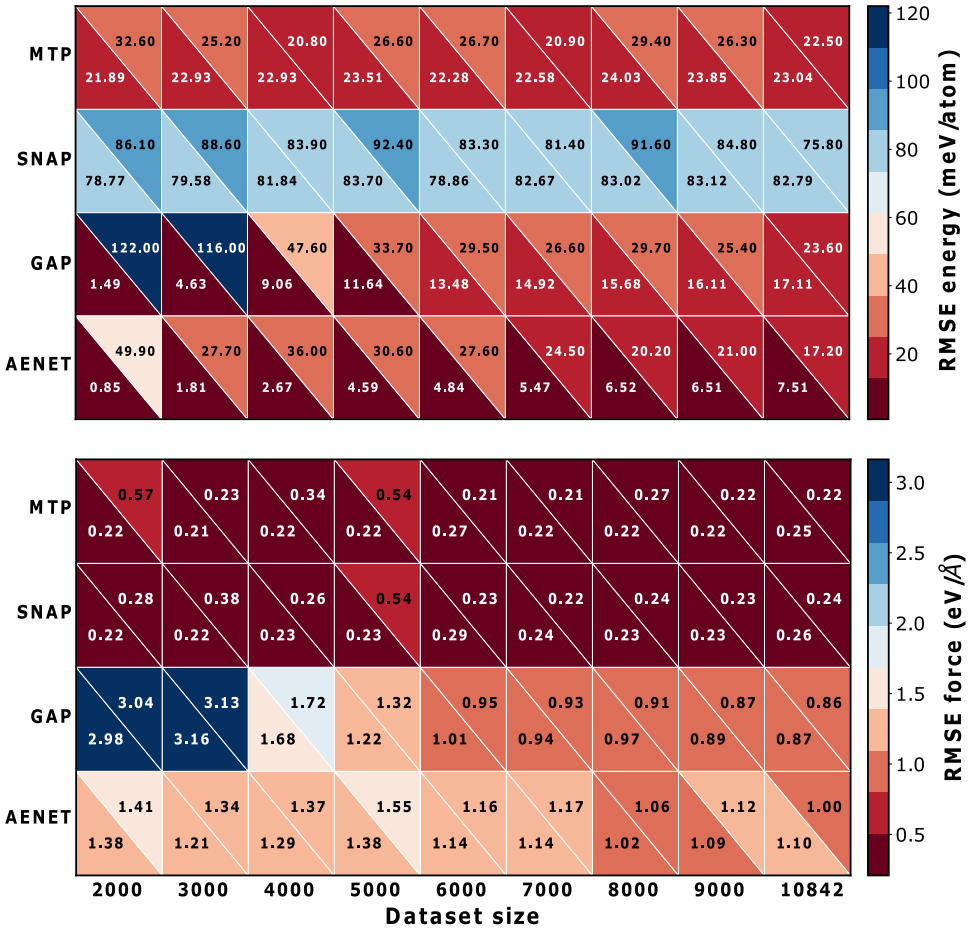


Figure 4. RMSE in energies (top panel) and forces (bottom panel) for the MLIPs considered in this work, as a function of increasing total data set size. Lower left and upper right triangles within each square represent training and test errors, respectively.

combinations of TMs that were missing from the data set of Artrith[33] et al.[26] [e.g., $Li_9(Sc,Mn,V,Cr,Fe,Co,Cu,Ni)_9O_{18}$ and structures with all nine TMs] using pymatgen's advanced transformations module,[73] resulting in an additional ∼400 structures. Also, we generated the structures without Li-ion to include Li-deficient conditions encountered at the top of charge. Specifically, we randomly selected a total of 500 structures from each of the n-TM combinations (n = 1 to 9) and removed all Li atoms from such structures, bringing our final data set to a total of 10,842 structures. A pie-chart visualizing the various splits in the total data set generated is given in Figure 3.

Constructing MLIPs. All MLIPs investigated in this work are atom centered, i.e., the MLIPs express the PES as a sum of

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

atomic energies (ϵ_i), where each atomic energy is a function of the local environment of each atom that obeys the underlying symmetry of the environment, such as rotational and translational invariance. The range of the local environment is determined by the cutoff radius (R_c or R_{cut}) hyperparameter, with a corresponding switching function (f_c, a cosine function for example) that ensures that the contribution of neighboring atoms to the atomic energy decay smoothly to zero as the distance from the atom-of-interest approaches R_c. However, the MLIPs differ in the way the local environments are described mathematically and the functional basis/expression used to map the descriptors to the PES. The key concepts and model parameters used in all MLIPs considered are described in the Supporting Information and tabulated in Table 1. We also refer the readers to the subset of the literature here.[64,74−76]

DFT Calculations. We performed Hubbard U corrected[79] DFT calculations using the Perdew−Burke−Ernzerhof functionalization of the generalized gradient approximation,[80] with projector-augmented wave potentials,[81,82] as implemented in the Vienna ab initio simulation package (VASP, version 6.1.2).[83,84] The calculations were performed, without preserving any symmetry, for all the 10,842 structures until total energies converged to within 0.01 meV. Note that we performed a single self-consistent field calculation for the converged structures obtained from Artrith et al.,[33] while we performed a full structure relaxation (i.e., relax ionic positions, cell volume, and cell shape) for all the structures we enumerated and for Li-intercalation voltage calculations. For all structures that were fully relaxed, we converged both the total energies and atomic forces to within 0.01 meV and ±30 meV/Å, respectively.

Γ-centered k-point meshes with a density of 1000 divided by the number of atoms were used, in accordance with the work of Artrith et al.[33] The plane wave kinetic energy cutoff was set to 520 eV and VASP input parameters were generated using pymatgen,[73] where the parameters were compatible with the Materials Project.[85] The reference atomic energies in the AENET package were calculated by placing isolated Li, TM, and O atom, at the origin of a 18 × 19 × 20 Å^3 cell (i.e., each reference energy was calculated for an isolated atom). Table S2 summarizes Hubbard U values and the corresponding calculated atomic energies. We used the materials ML (MAML[52]) python package as an interface to construct the GAP, SNAP, qSNAP, and MTP potentials. For voltage calculations, we considered layered, single-TM $LiTMO_2$ structures obtained from the inorganic crystal structure database (ICSD[86]) and the corresponding delithiated versions, except $LiTiO_2$, where we obtained the structure via ionic substitution of layered-$LiVO_2$. While DFT-calculated voltages were obtained from a full structure relaxation of layered-$LiTMO_2$ and layered-TMO_2 structures, we obtained AENET energies (and voltages) using the corresponding initial structures used for DFT calculations.

RESULTS

Training and Test Errors. The RMSE in energies (in units of meV/atom) and forces (in eV/Å) errors for all the four MLIPs are summarized in the top and bottom panels of Figure 4, respectively, as a function of increasing total data set size, with the parity plots between MLIP- and DFT-calculated values compiled in Figures S2−S11. Specifically, we display the progression of errors for total data set sizes from 2000 until 9000 in steps of 1000, and for the full data set of 10,842 configurations, where each data set is split into a 90:10 training/test set. Note that each smaller data set is a randomly chosen subset from the 10,842 structures, thus including the diversity and complexity of the local atomic environments to an extent. Each row of Figure 4 represents an MLIP, while the top and bottom triangles within each box represent the errors associated with the training and test sets, respectively, for each total data set size. The errors displayed in Figure 4 correspond to the best set of hyperparameters that we identified for each MLIP (see Table S1). We do not include qSNAP error data in Figure 4 since we observed qSNAP to provide similar metrics as SNAP when trained on small data sets (i.e., 2000 and 3000, see Figure S1), indicating marginal improvement over SNAP. Moreover, we encountered numerical and convergence difficulties with qSNAP while attempting to train it on larger data set sizes. Hence, we do not include qSNAP in comparison to the other MLIPs for the rest of this study.

In general, the MLIPs considered show higher training errors in energy with an increase in training data set size (lower triangles in Figure 4), although the increase is not always monotonous and the extent of increase is different for different potentials. For example, the training set energy errors increased from 21.89 to 23.04 meV/atom in MTP, 78.77 to 82.79 meV/atom in SNAP, 1.49 to 17.11 meV/atom in GAP, and 0.85 to 7.51 meV/atom in AENET. While AENET and GAP display largely monotonic increase in error with data set size, indicating that these frameworks become better fitted with more data, MTP and SNAP are nonmonotonic and show relatively low increase (∼5%) in error as the total data set size is increased from 2000 to 10,842. Thus, the choice of the data set itself may play a role in the accuracy of MTP and SNAP fits and more data may not necessarily improve the fit. Importantly, AENET exhibits the lowest error with our full training data set, indicating high accuracy during training. Indeed, AENET's RMSE of ∼7.5 meV/atom is not far away from the typical error expected within DFT calculations (∼1 meV/atom). In contrast, SNAP shows the highest error across all data set sizes, which may be due to the high number of components in our data set.

Different from training errors on energies, test errors (upper triangles in Figure 4) monotonically decrease with increasing data set size for all MLIPs, except MTP, which indicates an improvement in the transferability of all models with increasing training data set size. Additionally, the MLIPs do display higher test errors compared to training errors, which is expected, except for MTP and SNAP, which show lower test than training errors for larger data sets. Notably, GAP and AENET display significantly high test errors compared to training, especially in small data sets (<6000 data points), which is indicative of a high degree of overfitting. For example, GAP's test error is 122 meV/atom compared to a training error of 1.49 meV/atom for a total data set size of 2000, while AENET's test and training errors for the same data set size are 49.9 and 0.85 meV/atom, respectively. Thus, both AENET and GAP are susceptible to overfit to the available data, especially when the data set size is of the order of few thousand points or below. However, both AENET and GAP models become more transferable with increase in the overall data set size, as exhibited by similar training and test errors as the data set size approaches 10,842. The improvement in the accuracy and transferability of AENET is consistent with prior observations in literature using ANNPs as well.[87,88] Interestingly, both MTP and SNAP show similar training and test

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

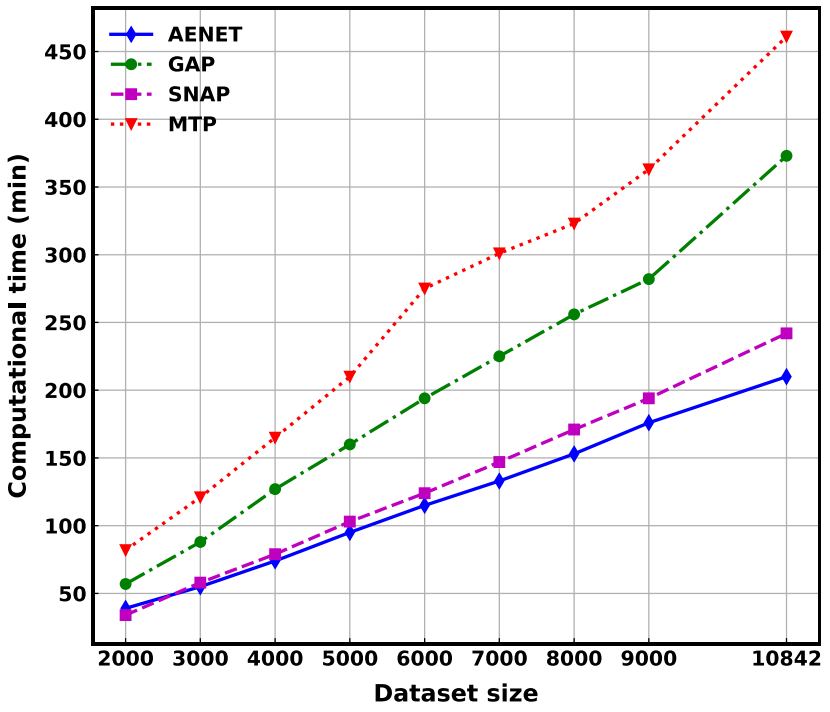


Figure 5. Computational time (in minutes) for training vs total data set size for all MLIPs considered in this work.

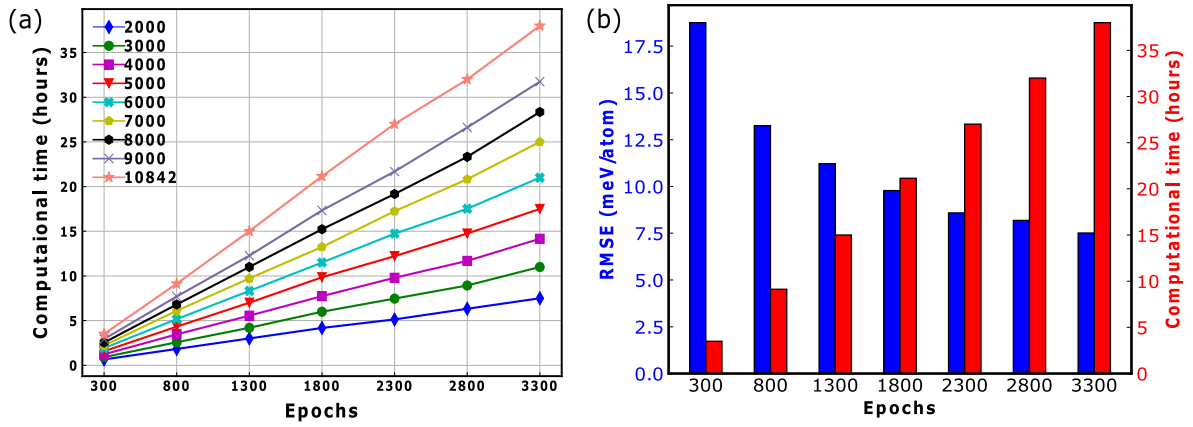


Figure 6. (a) Computational training time (in hours) vs number of epochs for different total data set sizes using AENET. (b) Training energy RMSE (left *y*-axis and blue bars) using AENET for the total 10,842 data set with increasing number of epochs, with the corresponding computational training time (in hours) also indicated (right *y*-axis and red bars).

errors across all data set sizes (<11 meV/atom deviation) indicating good transferability of MTP and SNAP, though the accuracy of these MLIPs may not be as good as AENET or GAP. Thus, for the full data set (10,842 points), AENET gives both best accuracy and transferability for total energy evaluations, while MTP provides best transferability for smaller data sets (<7000 points).

Compared to training and predicting total energies, all MLIPs do display significantly larger errors in training or predicting atomic forces, as compared to DFT (lower panel of Figure 4), which is expected given that force data is usually noisier compared to energy data. Indeed, the lowest force training error that we encountered is 0.21 eV/Å by MTP for the 3000 data set, which is at least an order of magnitude higher than the errors typically encountered in DFT calculations (~0.03−0.05 eV/Å). In terms of training the MLIPs, AENET does not yet have a provision to weight forces during training, while we did not explicitly include forces during the training of GAP to lower computational time and limit memory usage. Using multiple computing cores is one possible strategy with GAP to include forces during training. Consequently, AENET and GAP show significantly high training errors on atomic forces, compared to MTP and SNAP. For example, AENET's RMSE on forces in the 10,842 training set is 1.10 eV/Å, while GAP displays an RMSE of 0.87 eV/Å, significantly higher than MTP or SNAP (0.25−0.26 eV/Å). While the force training errors on GAP can be certainly brought lower by a careful optimization of the "default_sigma" hyperparameter that controls the relative weights between forces and energies apart from the regularity of the fit,[75] AENET's training algorithm requires an extension if force errors need to be brought down systematically for any data set.

All MLIPs, except AENET, display good transferability in terms of force predictions, given the similarity of training and

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

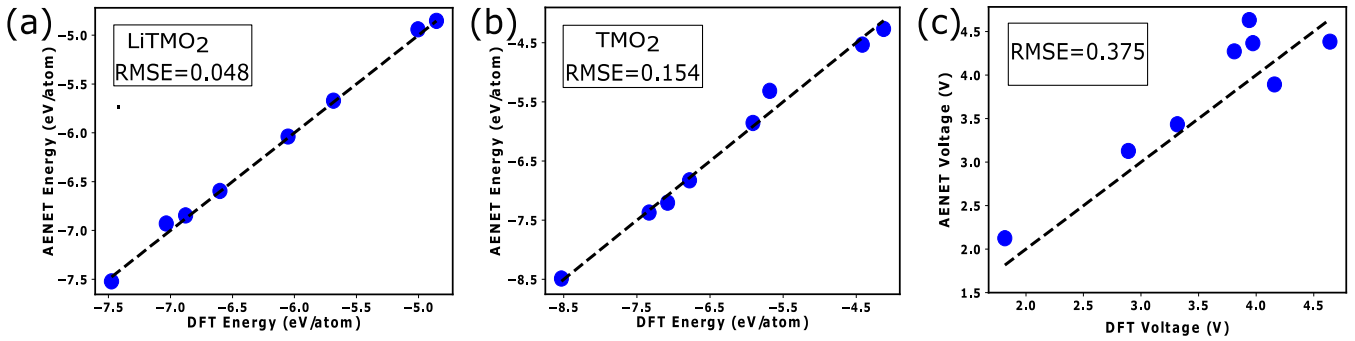


Figure 7. Comparison of AENET and DFT calculated energies of $LiTMO_2$ [panel (a)] and TMO_2 [panel (b)] compounds, where TM = Ti, V, Cr, Mn, Fe, Co, Ni, or Cu. Panel (c) represents a similar AENET-DFT comparison in terms of average Li intercalation voltage, where each voltage is calculated vs Li metal.

test force errors, particularly at large data set sizes. For example, MTP, SNAP, and GAP display <0.03 eV/Å deviation within training and test errors for the 10,842 data set. Thus, given both energy and force errors during training and their transferability to the test data set, we find MTP to be the best performer, with marginal improvement in accuracy with increasing data set size, and most suited for any dynamics simulations among the MLIPs considered. However, in terms of energy training and transferability, we find AENET to be the best performer, with a systematic improvement in accuracy and transferability with increasing data, and the most suited for any quick evaluation of total energies for a given structure, i.e., for static calculations. Note that the overall utility of a given MLIP is also determined by the computational time taken for training, which is discussed in the following section.

Ease of Training. The computational time taken for training the MLIPs as a function of increasing total data set size (from 2000 to 10,842), for the set of optimal hyperparameters (Table S1), is displayed in Figure 5. We used a single core of an Intel Xeon Gold 6271 central processing unit (CPU), with a maximum random access memory of 128 gigabytes, without any hyperthreading for training all MLIPs. For AENET, the computational training time plotted in Figure 5 is for 300 epochs (a hyperparameter for AENET training), while we have performed an extended analysis on the dependence of the training time on number of epochs later this section (see Figure 6). As expected, all MLIPs require larger training times with increasing data set size. Importantly, we find AENET (blue diamonds in Figure 5) to be the swiftest to train for all data set sizes at 300 epochs, except 2000 data set size, where SNAP (purple squares) is the fastest. Indeed, AENET is more than twice as fast (~210 min) compared to MTP (~460 min). Also, we find MTP (red triangles) to be the slowest to train for all data set sizes, followed by GAP (green circles).

Moreover, MTP and GAP both become progressively harder to train with increasing data set sizes. For example, MTP and GAP require an additional ~37 and ~28 min per additional 1000 total data points included, respectively, while AENET and SNAP are significantly easier to train, requiring additional times of ~15 and ~20 min per additional 1000 data points. Thus, the marginal improvement in accuracy on energies and forces with increasing data points exhibited by MTP is counterweighed, to an extent, by the significantly higher computational times required to train the potential. Note that we did not explicitly include forces during the training of GAP and we expect the computational training time of GAP to significantly increase, if forces were to be included as well. Therefore, we find that AENET is the easiest to train at 300 epochs, compared to the other MLIPs considered in this work.

Note that one of the key hyperparameters that influences (training) errors and computational time in AENET is the number of epochs used during training. The AENET results displayed in Figure 4 correspond to 3300 epochs of training. To further probe the role of this hyperparameter, we examine the variation of the computational training time as a function of increasing total data set size and number of epochs (Figure 6a) and the change in training RMSE and computational training time for the full 10,842 data set with increasing epochs (Figure 6b).

For all data set sizes, we observe an increase in computational time with increasing epochs (Figure 6a). More importantly, the rate of increase in computational time is more rapid with increasing epochs for larger data sets, indicating that the training time can quickly escalate as the complexity of the training set increases. For example, the training time required for 3300 epochs over the full 10,842 data set is ~38 h (or ~2280 min, Figure 6a), which is ~5× the training time required by MTP for the full data set (Figure 5). However, we do note that the energy RMSEs for the training set decreases steadily with increasing epochs (Figure 6b), from ~18.7 meV/atom at 300 epochs to ~7.5 meV/atom at 3300 epochs (i.e., a ~ 60% drop in errors). Nevertheless, the rapid increase in computational time for the full 10,842 data set may not be worth the marginal gain in accuracy, especially from ~27 h, ~ 8.5 meV/atom at 2300 epochs to ~38 h, ~ 7.5 meV/atom at 3300 epochs, or a ~ 41% gain in training time for an increase in accuracy of ~12% (Figure 6b). Therefore, AENET can be a quick MLIP to train compared to MTP, GAP, or SNAP, only if a low number of epochs gives reasonable accuracy and transferability for total energies.

Voltage Predictions. The accuracy in predicting material properties is critical for evaluating the performance and determining the utility of MLIPs. Hence, we choose the prediction of average intercalation voltages, in ordered, layered, single-TM $LiTMO_2$ compounds (TM = Ti, V, Cr, Mn, Fe, Co, Ni, or Cu) as a test of accuracy, versus DFT calculations. Specifically, we examine the accuracy of the AENET potential in determining voltages since it provided the best accuracy and transferability in prediction energies within disordered systems (see Figure 4). Note that we refer to AENET being the better potential, than MTP, strictly for energy predictions. For voltage calculations, we utilize the AENET potential trained at 2300 epochs (since training errors declined by only ~1.0

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

meV/atom from 2300 to 3300 epochs, see Figure 6b). Note that the average Li intercalation voltage, versus Li metal, can be approximately determined using the following relation

$$V = -\left(\frac{E_{\mathrm{LiTMO_2}} - E_{\mathrm{TMO_2}} - E_{\mathrm{Li}}}{F}\right) \quad (1)$$

where $E_{\mathrm{LiTMO_2}}$, $E_{\mathrm{TMO_2}}$, and E_{Li} are the total energies per formula unit of the lithiated structure, delithiated compound, and pure Li, respectively, from DFT or AENET, with the $p - V$ and entropic contributions ignored. F is the Faraday's constant. Note that we used the DFT calculated total energy for Li in case of AENET predictions also since the energy scales of our fitted AENET potential were quite different from the energy scale of a pure Li in its body centered cubic ground state configuration. We performed either structure relaxation using DFT or a static calculation using AENET, from the ICSD starting structures, to obtain the energies listed in eq 1. All DFT and AENET calculated voltage values are presented in Table S3 of the Supporting Information.

Figure 7 displays AENET calculated values versus DFT estimates, in terms of total energies (in eV/atom) of the fully lithiated compositions (panel a), fully delithiated compositions (panel b), the average voltages (panel c, in V vs Li metal). The black lines in each panel of Figure 7 indicate parity lines, with the degree of agreement between AENET and DFT values quantified in terms of RMSEs in each panel. Importantly, AENET estimates the total energies of both the lithiated and delithiated compositions quite accurately, with RMSEs of 0.048 and 0.154 eV/atom (with corresponding MAEs of 0.035 and 0.117 eV/atom), respectively, where the errors for lithiated compositions are similar to the test errors observed in Figure 4. These RMSEs represent percentage errors of 0.55% for $LiTMO_2$ and 2.12% for TMO_2 compositions. The higher errors within the TMO_2 compositions are expected since our data set has significantly more $LiTMO_2$ configurations than TMO_2.

In terms of voltage calculations, AENET does suffer from a compounding of error (instead of a cancellation of error) versus DFT, with the RMSE at ∼0.375 V (and MAE of ∼0.342 V) compared to DFT for all systems. The voltage errors represent a percentage error of 10.05% against DFT, which is higher than the percentage errors on the energy estimates. Notably, AENET's largest voltage errors are in the $LiFeO_2$ and $LiTiO_2$ systems (∼17%). For comparison, we have also included MTP calculated voltage values in Table S6, with MTP's overall performance (mean percentage error of 12.18%) expectedly slightly worse than AENET. Nevertheless, we expect our AENET potential to be useful for obtaining "quick" voltage predictions for several disordered lithium TM oxide compositions, with reasonable accuracy, provided that the potential is not utilized for a full structure relaxation.

DISCUSSION

We have performed a comprehensive evaluation on the accuracy, transferability, and ease of training of four different MLIPs, namely, AENET, GAP, SNAP, and MTP, on a 11-component lithium-containing disordered rocksalt chemical space (Figure 1), which is quite relevant for designing energy-dense cathodes in Li-ion batteries. Specifically, we have quantified the training and test errors (Figure 4) and the computational training time (Figure 5) of the above-mentioned MLIPs on a DFT-calculated data set of 10,842 configurations (Figure 3) based on a computational workflow (Figure 2). While we found MTP to be accurate and transferable in terms of force predictions, we found AENET to be significantly more accurate and transferable for energy predictions. In terms of ease of training, we found AENET to be fastest to train at 300 epochs, with training time increasing ∼10× at 3300 epochs and training errors decreasing by ∼60% (Figure 6). While MTP is ∼2× slower than AENET at 300 epochs, it is ∼5× faster than AENET at 3300 epochs (Figures 5 and 6). We also observed AENET's voltage predictions at 2300 epochs (Figure 7) to be reasonable compared to DFT for layered $LiTMO_2$. Thus, we hope that our work propels the use of MLIPs to model highly disordered ceramic and/or metallic systems.

For the training of any MLIP on any chemical system, the following factors are important: (i) accuracy in terms of low errors on energy and force predictions (and any other relevant properties) on the training data set, (ii) transferability, as quantified by the similarity of errors between the training and test data sets, and (iii) quality of the data set, in terms of sampling a sufficient number of relevant local chemical environments, and (iv) computational cost, be it for generating the data set or for training the MLIP itself. While these factors have a degree of dependence on one another, we found that an increase in data set size generally lead to higher accuracy and better transferability, especially for force predictions, although the rate of reduction in errors begin diminishing at different rates for different MLIPs. For example, both MTP and SNAP do not show significant changes in their train/test force errors beyond a 7000 data set size, while AENET and GAP do not display a similar behavior. Thus, increasing the data set size is not always a solution for improving certain property predictions, while diversity of the local environments sampled within the data set and the choice of a given MLIP itself may play a bigger role in improving predictions.

One of the biggest challenges associated with the DRX system is predicting the right local ordering, which is influenced by both the TM and the extent of excess-Li,[89] where sampling the excess-Li is one of the challenges posed by the data set we have used in this work. Other limitations of our DFT-calculated data set include the lack of TM layer stacking arrangements, slab gliding, lattice parameter variations, Li-vacancy long-range orderings, and other structural variations that may be observed in rocksalt-based systems, especially at intermediate Li concentrations. Our calculated data set also samples more lithiated configurations than delithiated compositions, indicating imbalance in sampling the two compositions. Thus, the MLIPs constructed here will have limited utility for predicting voltage profiles. However, average voltages between fully lithiated ($LiTMO_2$) and delithiated (TMO_2) compositions are not dependent on specific voltage profiles since these are thermodynamically averaged quantities. Thus, our trained MLIPs, especially AENET, should have reasonable utility in predicting average voltages across the entire range of Li content, which can still be used as a screening criterion. Also, note that the "ground truth" for the trained MLIPs in this work is the DFT-calculated data set. Hence, errors made by DFT in predicting average voltages, in specific chemical systems, will translate to our MLIPs as well. Such translation of DFT errors into MLIPs is purely mathematical and is less dependent on underlying chemical or structural factors. Another component to note while comparing experimental voltage data to DFT or MLIP

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

predictions is the depth of charge observed experimentally, which can depend on several factors such as the actual composition of the electrode (including dopants), electrolyte stability, size of the electrode particles, and any binder or electronically conducting medium (e.g., carbon coating) that are added to the electrode.

One of the overarching concerns while constructing MLIPs is the tendency for the frameworks to be overfitted, which is generally indicated by a significant difference in errors in the test versus training data sets. Importantly, we find that both AENET and GAP tend to overfit at small data set sizes (<3000 data points, see Figure 4), with the tendency to overfit decreasing significantly at larger data sets. Note that the definition of a small data set size that results in overfitting of AENET or GAP will be dependent on the underlying chemical system, especially on the number of components present. Indeed, previous studies on fewer component systems have shown remarkably good fits with AENET and GAP.[57,90−92] Nevertheless, we advocate a careful evaluation of the train and test errors with increasing data set size to be done while using AENET or GAP to ensure appropriate fits. Surprisingly, we found MTP and SNAP to be quite resilient to overfitting and show remarkably similar training and test errors across all data set sizes. However, the accuracy on energies for both MTP and SNAP may have been limited by the high number of components (11) considered in this study.

The significant difference in force predictions between MTP and AENET can be attributed to the corresponding definitions of the loss functions used in training the respective models. For instance, MTP accounts for forces (and stresses) in its training loss function, which allows MTP to achieve a better balance between predicting energies and forces. On the other hand, the version of AENET that we used trained only on energies, resulting in higher errors on force predictions. Another factor that enables MTP to train well on both energies and forces is that MTP efficiently accounts for many-body interactions via the construction of moment tensors, up to a maximum "level". In the case of SNAP and GAP, both employ different mathematical frameworks to describe the local atomic density around a reference atom, rendering limited scope for such models to describe many-body and angular interactions. Nevertheless, the framework of AENET, given its neural network architecture and the various weights and biases involved, provides it greater flexibility compared to MTP, SNAP, and GAP, allowing it to train the best on large data sets of total energies.

In terms of ease of training, we found AENET and MTP to be fastest and slowest potentials to train, across all data set sizes (except 2000 data points, Figure 5). However, AENET is fastest only at a low number of epochs (300), which results in higher errors. Indeed, increasing the number of epochs does significantly slow AENET down (Figure 6), with marginal reductions in training/test errors. Thus, if AENET were to be used for other systems, the number of epochs is an important hyperparameter to be optimized to get the best trade-off of computational time and accuracy. Another point to note is that not all of the MLIP frameworks that are available for users are optimized to train on multiple CPUs, which can significantly reduce training time. This is crucial since not all users have access to cutting-edge graphic processing units (GPUs).

For performing predictions that can be practically utilized for the design of battery systems, we examined the accuracy of our AENET potential for average voltage predictions in layered, single-TM $LiTMO_2$, as compared to DFT. While we found high accuracy in energy predictions of $LiTMO_2$ and TMO_2 ordered configurations using AENET, the accuracy does drop for voltage predictions due to compounding of error, especially in Fe and Ti containing systems. Another factor that may contribute to larger errors in Fe and Ti systems are errors associated with DFT in modeling the Fe^{4+} and Ti^{3+} oxidation states. To further probe the quality of AENET's voltage predictions on disordered lithium containing rocksalt systems, more DFT calculations are required on chemical combinations and/or configurations that we have not sampled so far. While this is an immediate follow-up exercise that we intend to take, we do expect AENET to provide quick, qualitative voltage predictions reasonably well for disordered rocksalt configurations as well. Thus, high-throughput screening approaches to unearth novel, high-voltage disordered rocksalt cathode compositions may be feasible with AENET potentials. Nevertheless, given our preliminary voltage predictions (Table S3), we do expect Fe-, Co-, and Ni-based disordered rocksalt compositions to be promising.

Apart from data set generation that was constrained by computational costs, another limitation of our study is that we have explored MLIPs that are atom-centered. For example, AENET and MTP are parametrizations of radial and angular distributions of atoms around a central atom of interest, while GAP and SNAP employ descriptors to quantify the local density around an atom. There are several graph-network-based (neural net) potentials that have been developed recently, such as SchNet,[93] NequIP,[94] MEGNet,[95] and CHGNet[96] with the work by Reiser et al.[97] providing a well-compiled summary of available graph-based MLIPs. Most graph network potentials can be combined with deep learning to provide high accuracy for multicomponent systems, as illustrated by MEGNet[98] and CHGNet.[96] Long range interactions, such as electrostatics and dispersions, have also been included with such graph-based potentials.[99] Note that incorporating deep learning often results in better performance with "large" data sets, such as the Materials Project,[85] and, in turn, requires large computational training time.

To briefly explore the performance of graph-network-based potentials, we trained NequIP on the data set size of 2000 using the hyperparameters listed in Table S4. The performance of NequIP on energies and forces, on the training and test data subsets are listed in Table S5. Importantly, we find the RMSE on energies to be lower than AENET for the test data set, indicating that NequIP is more transferable and is less prone to overfitting than AENET for small data set sizes. In terms of forces, we observe NequIP to be significantly better than MTP both on the training and test RMSE values, signifying that NequIP may outperform MTP on the full data set as well for force predictions. However, the computational training time on the 2000-sized data set for NequIP was 279 h on the same hardware resources that we had used for training the other MLIPs in this work, highlighting the disparity in the ease of training between atom-centered and graph-network-based potentials. Moreover, we were unable to get a converged NequIP for larger data set sizes, nor access to high-performance GPUs on which NequIP is designed to run significantly faster than CPUs. Thus, we believe that graph-network-based potentials can have significant utility for modeling PESs of diverse chemical spaces, if the training routines are better optimized and parallelized. For now, NequIP (and SchNet) have not been used for high-component

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(>5) chemical spaces besides our brief attempt here, as far as we know; thus, their performance remains unexplored in high-component systems.

On the other hand, deep learning can be combined with atom centered basis to generate highly accurate MLIPs for low-component systems, as characterized by the DeePMD[100] framework. Also, symbolic regression can be used to construct highly transferable potentials with minimal parameters.[101] Nevertheless, we believe that the potentials explored in this study are highly relevant for describing low-to-medium component systems where data is scarce, and are may be useful for molecular dynamics simulations given their ease of training and computational speed during predictions. Going forward, developing interatomic potential frameworks that can deal with data scarcity in the broad field of materials science and are resilient against overfitting[102] will be of high utility in studying disordered and/or large scale systems over long times.

CONCLUSIONS

In this work, we have quantified the accuracy, transferability, and ease of training of five atom-centered MLIPs, namely, MTP, SNAP, qSNAP, GAP, and AENET, in their ability to model the PES of disordered, 11-component, $LiTMO_2$ compositions (TM = Sc, Ti, V, Cr, Mn, Fe, Co, Ni, and/or Cu), which form the base of designing energy-dense DRX cathodes for advanced LIBs. We also trained the graph-network-based NequIP on our 2000 sized data subset to provide a point-of-comparison to the performance of the atom-centered MLIPs. To train these potentials, we generated a DFT-calculated data set of 10,842 configurations and created a 90:10 random split of the set for training and testing. Using magnitude of training errors and similarity between training and test errors as the metrics of accuracy and transferability, respectively, we found AENET to be the best potential for predicting total energies, while MTP was the best performer for atomic forces. While both AENET and GAP tended to overfit in small data sets, the extent of overfitting and the transferability of these potentials improve considerably with increase in data set size. In terms of ease of training, we observed AENET (MTP) to exhibit the smallest (largest) computational training time, where the number of epochs is low (~300) for AENET. However, note that AENET's training time does increase significantly with increasing the number of epochs during training, which may hamper swift training of AENET potentials in specific systems. Finally, we found AENET to also provide reasonable predictions of derived quantities, such as average Li-intercalation voltages, in layered, ordered, single-TM $LiTMO_2$ compositions, versus DFT estimates. We hope that our work inspires more studies on using atom-centered MLIPs to model configurationally complex systems, resulting in the discovery of new DRX cathodes and other high-entropy ceramic/metallic systems.

ASSOCIATED CONTENT

Data Availability Statement

All computed data, relevant scripts, and the best MLIPs constructed in this work are available freely for all via our GitHub repository: https://github.com/sai-mat-group/mlip-disordered-rocksalts.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.4c00039.

Description of all MLIP frameworks and the corresponding set of hyperparameters used, Hubbard U values used in DFT calculations, DFT-, MTP-, and AENET-calculated voltages, NequIP hyperparameters, training and test scores, RMSE comparisons between SNAP and qSNAP, and compilation of energy and force parity plots across all MLIPs and data set sizes (PDF)

AUTHOR INFORMATION

Corresponding Author

Gopalakrishnan Sai Gautam − *Department of Materials Engineering, Indian Institute of Science, Bengaluru 560012 Karnataka, India*; orcid.org/0000-0002-1303-0976; Email: saigautamg@iisc.ac.in

Authors

Vijay Choyal − *Department of Materials Engineering, Indian Institute of Science, Bengaluru 560012 Karnataka, India*

Nidhish Sagar − *Department of Materials Engineering, Indian Institute of Science, Bengaluru 560012 Karnataka, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.4c00039

Author Contributions

V.C. and N.S. performed all DFT calculations and construction of MLIPs, under the supervision of G.S.G. V.C. wrote the initial draft of the paper. G.S.G. edited and refined the draft. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

G.S.G. acknowledges financial support from the Indian Institute of Science (IISc) Seed grant, SG/MHRD/20/0020 and SR/MHRD/20/0013, the IISc-Axis Bank Centre for Mathematics and Computing for travel support, and the Science and Engineering Research Board (SERB) of the Department of Science and Technology, Government of India, under sanction numbers SRG/2021/000201 and IPA/2021/000007. V.C. thanks the Institute of Eminence Postdoctoral fellowship awarded by the Indian Institute of Science for financial assistance. G.S.G. and V.C. thank Aqshat Seth, a project associate at IISc, for computational help during manuscript revision. A portion of the DFT calculations showcased in this work were performed with the computational resources provided by the Supercomputer Education and Research Center, Indian Institute of Science. We acknowledge National Supercomputing Mission (NSM) for providing computing resources of "PARAM Siddhi-AI", under National PARAM Supercomputing Facility (NPSF), C-DAC, Pune and supported by the Ministry of Electronics and Information Technology (MeitY) and Department of Science and Technology (DST), Government of India.

REFERENCES

(1) Kang, K.; Meng, Y. S.; Bréger, J.; Grey, C. P.; Ceder, G. Electrodes with High Power and High Capacity for Rechargeable Lithium Batteries. *Science* **2006**, *311* (5763), 977−980.

(2) Tarascon, J.-M. Is Lithium the New Gold? *Nat. Chem.* **2010**, *2*, 510.

(3) Whittingham, M. S. Ultimate Limits to Intercalation Reactions for Lithium Batteries. *Chem. Rev.* **2014**, *114*, 11414−11443.

(4) Amatucci, G. G.; Tarascon, J. M.; Klein, L. C. CoO2, The End Member of the Li x CoO2 Solid Solution. *J. Electrochem. Soc.* **1996**, *143* (3), 1114−1123.
(5) Delmas, C.; Carlier, D.; Guignard, M. The Layered Oxides in Lithium and Sodium-Ion Batteries: A Solid-State Chemistry Approach. *Adv. Energy Mater.* **2021**, *11* (2), 2001201.
(6) Goodenough, J. B.; Park, K.-S. The Li-Ion Rechargeable Battery: A Perspective. *J. Am. Chem. Soc.* **2013**, *135* (4), 1167−1176.
(7) Whittingham, M. S. Lithium Batteries and Cathode Materials. *Chem. Rev.* **2004**, *104* (10), 4271−4302.
(8) Clément, R. J.; Lun, Z.; Ceder, G. Cation-Disordered Rocksalt Transition Metal Oxides and Oxyfluorides for High Energy Lithium-Ion Cathodes. *Energy Environ. Sci.* **2020**, *13*, 345−373.
(9) Ben Yahia, M.; Vergnet, J.; Saubanère, M.; Doublet, M.-L. Unified Picture of Anionic Redox in Li/Na-Ion Batteries. *Nat. Mater.* **2019**, *18* (5), 496−502.
(10) Lee, J.; Urban, A.; Li, X.; Su, D.; Hautier, G.; Ceder, G. Unlocking the Potential of Cation-Disordered Oxides for Rechargeable Lithium Batteries. *Science* **2014**, *343* (6170), 519−522.
(11) Yabuuchi, N.; Takeuchi, M.; Nakayama, M.; Shiiba, H.; Ogawa, M.; Nakayama, K.; Ohta, T.; Endo, D.; Ozaki, T.; Inamasu, T.; Sato, K.; Komaba, S. High-Capacity Electrode Materials for Rechargeable Lithium Batteries: Li3NbO4-Based System with Cation-Disordered Rocksalt Structure. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (25), 7650−7655.
(12) Yabuuchi, N.; Tahara, Y.; Komaba, S.; Kitada, S.; Kajiya, Y. Synthesis and Electrochemical Properties of Li4MoO5-NiO Binary System as Positive Electrode Materials for Rechargeable Lithium Batteries. *Chem. Mater.* **2016**, *28* (2), 416−419.
(13) Kan, W. H.; Chen, D.; Papp, J. K.; Shukla, A. K.; Huq, A.; Brown, C. M.; McCloskey, B. D.; Chen, G. Unravelling Solid-State Redox Chemistry in Li1.3Nb0.3Mn0.4O2 Single-Crystal Cathode Material. *Chem. Mater.* **2018**, *30* (5), 1655−1666.
(14) Chen, R.; Ren, S.; Knapp, M.; Wang, D.; Witter, R.; Fichtner, M.; Hahn, H.; Knapp, M.; Fichtner, M.; Hahn, H.; Chen, R.; Ren, S.; Wang, D.; Witter, R. Disordered Lithium-Rich Oxyfl Uoride as a Stable Host for Enhanced Li + Intercalation Storage. *Adv. Energy Mater.* **2015**, *5*, 1401814.
(15) Ren, S.; Chen, R.; Maawad, E.; Dolotko, O.; Guda, A. A.; Shapovalov, V.; Wang, D.; Hahn, H.; Fichtner, M. Improved Voltage and Cycling for Li+ Intercalation in High-Capacity Disordered Oxyfluoride Cathodes. *Adv. Sci.* **2015**, *2* (10), 2198−3844.
(16) Ji, H.; Urban, A.; Kitchaev, D. A.; Kwon, D.-H.; Artrith, N.; Ophus, C.; Huang, W.; Cai, Z.; Shi, T.; Kim, J. C.; Kim, H.; Ceder, G. Hidden Structural and Chemical Order Controls Lithium Transport in Cation-Disordered Oxides for Rechargeable Batteries. *Nat. Commun.* **2019**, *10*, 592.
(17) Sathiya, M.; Rousse, G.; Ramesha, K.; Laisa, C. P.; Vezin, H.; Sougrati, M. T.; Doublet, M.-L.; Foix, D.; Gonbeau, D.; Walker, W.; Prakash, A. S.; Ben Hassine, M.; Dupont, L.; Tarascon, J.-M. Reversible Anionic Redox Chemistry in High-Capacity Layered-Oxide Electrodes. *Nat. Mater.* **2013**, *12* (9), 827−835.
(18) Radin, M. D.; Vinckeviciute, J.; Seshadri, R.; Van der Ven, A. Manganese Oxidation as the Origin of the Anomalous Capacity of Mn-Containing Li-Excess Cathode Materials. *Nat. Energy* **2019**, *4* (8), 639−646.
(19) Ji, H.; Wu, J.; Cai, Z.; Liu, J.; Kwon, D.-H.; Kim, H.; Urban, A.; Papp, J. K.; Foley, E.; Tian, Y.; Balasubramanian, M.; Kim, H.; Clément, R. J.; McCloskey, B. D.; Yang, W.; Ceder, G. Ultrahigh Power and Energy Density in Partially Ordered Lithium-Ion Cathode Materials. *Nat. Energy* **2020**, *5* (3), 213−221.
(20) Cao, X.; Li, H.; Qiao, Y.; Jia, M.; He, P.; Cabana, J.; Zhou, H. Achieving Stable Anionic Redox Chemistry in Li-Excess O2-Type Layered Oxide Cathode via Chemical Ion-Exchange Strategy. *Energy Storage Mater.* **2021**, *38*, 1−8.
(21) Yabuuchi, N. Rational Material Design of Li-Excess Metal Oxides with Disordered Rock Salt Structure. *Curr. Opin. Electrochem.* **2022**, *34*, 100978.
(22) Kitchaev, D. A.; Lun, Z.; Richards, W. D.; Ji, H.; Clément, R. J.; Balasubramanian, M.; Kwon, D. H.; Dai, K.; Papp, J. K.; Lei, T.; McCloskey, B. D.; Yang, W.; Lee, J.; Ceder, G. Design Principles for High Transition Metal Capacity in Disordered Rocksalt Li-Ion Cathodes †. *Energy Environ. Sci.* **2018**, *11*, 2159−2171.
(23) Sakurai, Y.; Arai, H.; Yamaki, J.-I. Preparation of Electrochemically Active α-LiFeO2 at Low Temperature. *Solid State Ionics* **1998**, *113*−*115*, 29−34.
(24) Obrovac, M. N.; Mao, O.; Dahn, J. R. Structure and Electrochemistry of LiMO2 (M = Ti, Mn, Fe, Co, Ni) Prepared by Mechanochemical Synthesis. *Solid State Ionics* **1998**, *112*, 9−19.
(25) Pralong, V.; Gopal, V.; Caignaert, V.; Duffort, V.; Raveau, B. Lithium-Rich Rock-Salt-Type Vanadate as Energy Storage Cathode: $Li_{2-x}VO_3$. *Chem. Mater.* **2012**, *24* (1), 12−14.
(26) Urban, A.; Lee, J.; Ceder, G. The Configurational Space of Rocksalt-Type Oxides for High-Capacity Lithium Battery Electrodes. *Adv. Energy Mater.* **2014**, *4* (13), 1400478.
(27) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, 864−871.
(28) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133−A1138.
(29) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65* (11), 1501−1509.
(30) Kapoor, V.; Singh, B.; Sai Gautam, G.; Cheetham, A. K.; Canepa, P. Rational Design of Mixed Polyanion Electrodes $Na_xV_2P_{3-i}(Si/S)_iO_{12}$ for Sodium Batteries. *Chem. Mater.* **2022**, *34* (7), 3373−3382.
(31) Tekliye, D. B.; Kumar, A.; Weihang, X.; Mercy, T. D.; Canepa, P.; Sai Gautam, G. Exploration of NaSICON Frameworks as Calcium-Ion Battery Electrodes. *Chem. Mater.* **2022**, *34* (22), 10133−10143.
(32) Allam, O.; Kuramshin, R.; Stoichev, Z.; Cho, B. W.; Lee, S. W.; Jang, S. S. Molecular Structure-Redox Potential Relationship for Organic Electrode Materials: Density Functional Theory-Machine Learning Approach. *Mater. Today Energy* **2020**, *17*, 100482.
(33) Artrith, N.; Urban, A.; Ceder, G. Efficient and Accurate Machine-Learning Interpolation of Atomic Energies in Compositions with Many Species. *Phys. Rev. B* **2017**, *96*, 014112.
(34) Wang, C.; Aoyagi, K.; Wisesa, P.; Mueller, T. Lithium Ion Conduction in Cathode Coating Materials from On-the-Fly Machine Learning. *Chem. Mater.* **2020**, *32* (9), 3741−3752.
(35) Houchins, G.; Viswanathan, V. An Accurate Machine-Learning Calculator for Optimization of Li-Ion Battery Cathodes. *J. Chem. Phys.* **2020**, *153* (5), 054124.
(36) Sendek, A. D.; Cubuk, E. D.; Ransom, B.; Nanda, J.; Reed, E. J. Machine-Learning and Data-Intensive Methods for Accelerating the Development of Rechargeable Battery Chemistries: A Review. *Transition Metal Oxides for Electrochemical Energy Storage*; Wiley, 2022; pp 393−409.
(37) Hart, G. L. W.; Mueller, T.; Toher, C.; Curtarolo, S. Machine Learning for Alloys. *Nat. Rev. Mater.* **2021**, *6* (8), 730−755.
(38) Varley, J. B.; Miglio, A.; Ha, V.-A.; van Setten, M. J.; Rignanese, G.-M.; Hautier, G. High-Throughput Design of Non-Oxide p-Type Transparent Conducting Materials: Data Mining, Search Strategy, and Identification of Boron Phosphide. *Chem. Mater.* **2017**, *29* (6), 2568−2573.
(39) Aykol, M.; Hummelshøj, J. S.; Anapolsky, A.; Aoyagi, K.; Bazant, M. Z.; Bligaard, T.; Braatz, R. D.; Broderick, S.; Cogswell, D.; Dagdelen, J.; Drisdell, W.; Garcia, E.; Garikipati, K.; Gavini, V.; Gent, W. E.; Giordano, L.; Gomes, C. P.; Gomez-Bombarelli, R.; Balaji Gopal, C.; Gregoire, J. M.; Grossman, J. C.; Herring, P.; Hung, L.; Jaramillo, T. F.; King, L.; Kwon, H.-K.; Maekawa, R.; Minor, A. M.; Montoya, J. H.; Mueller, T.; Ophus, C.; Rajan, K.; Ramprasad, R.; Rohr, B.; Schweigert, D.; Shao-Horn, Y.; Suga, Y.; Suram, S. K.; Viswanathan, V.; Whitacre, J. F.; Willard, A. P.; Wodo, O.; Wolverton, C.; Storey, B. D. The Materials Research Platform: Defining the Requirements from User Stories. *Matter* **2019**, *1* (6), 1433−1438.

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(40) Jain, A.; Hautier, G.; Ong, S. P.; Persson, K. New Opportunities for Materials Informatics: Resources and Data Mining Techniques for Uncovering Hidden Relationships. *J. Mater. Res.* **2016**, *31* (8), 977−994.
(41) Canepa, P.; Bo, S.-H.; Sai Gautam, G.; Key, B.; Richards, W. D.; Shi, T.; Tian, Y.; Wang, Y.; Li, J.; Ceder, G. High Magnesium Mobility in Ternary Spinel Chalcogenides. *Nat. Commun.* **2017**, *8* (1), 1759.
(42) Zhang, H.; Wang, Z.; Cai, J.; Wu, S.; Li, J. Machine-Learning-Enabled Tricks of the Trade for Rapid Host Material Discovery in Li-S Battery. *ACS Appl. Mater. Interfaces* **2021**, *13* (45), 53388−53397.
(43) Kundu, S.; Solomatin, N.; Kraytsberg, A.; Ein-Eli, Y. MgSc2Se4 Solid Electrolyte for Rechargeable Mg Batteries: An Electric Field-Assisted All-Solid-State Synthesis. *Energy Technol.* **2022**, *10* (11), 2200896.
(44) Black, A. P.; Frontera, C.; Torres, A.; Recio-Poo, M.; Rozier, P.; Forero-Saboya, J. D.; Fauth, F.; Urones-Garrote, E.; Arroyo-de Dompablo, M. E.; Palacín, M. R. Elucidation of the Redox Activity of Ca2MnO3.5 and CaV2O4 in Calcium Batteries Using Operando XRD: Charge Compensation Mechanism and Reversibility. *Energy Storage Mater.* **2022**, *47*, 354−364.
(45) Huang, J.; Ouyang, B.; Zhang, Y.; Yin, L.; Kwon, D.-H.; Cai, Z.; Lun, Z.; Zeng, G.; Balasubramanian, M.; Ceder, G. Inhibiting Collective Cation Migration in Li-Rich Cathode Materials as a Strategy to Mitigate Voltage Hysteresis. *Nat. Mater.* **2023**, *22* (3), 353−361.
(46) Zhang, Z.; Zou, Z.; Kaup, K.; Xiao, R.; Shi, S.; Avdeev, M.; Hu, Y.; Wang, D.; He, B.; Li, H.; Huang, X.; Nazar, L. F.; Chen, L. Correlated Migration Invokes Higher Na+-Ion Conductivity in NaSICON-Type Solid Electrolytes. *Adv. Energy Mater.* **2019**, *9* (42), 1902373.
(47) Blanc, L. E.; Choi, Y.; Shyamsunder, A.; Key, B.; Lapidus, S. H.; Li, C.; Yin, L.; Li, X.; Gwalani, B.; Xiao, Y.; Bartel, C. J.; Ceder, G.; Nazar, L. F. Phase Stability and Kinetics of Topotactic Dual Ca2+-Na + Ion Electrochemistry in NaSICON NaV2(PO4)3. *Chem. Mater.* **2022**, *35* (2), 468−481.
(48) Qian, X.; He, J.; Mastronardo, E.; Baldassarri, B.; Yuan, W.; Wolverton, C.; Haile, S. M. Outstanding Properties and Performance of CaTi0.5Mn0.5O3-δ for Solar-Driven Thermochemical Hydrogen Production. *Matter* **2021**, *4* (2), 688−708.
(49) Bayliss, R. D.; Key, B.; Sai Gautam, G.; Canepa, P.; Kwon, B. J.; Lapidus, S. H.; Dogan, F.; Adil, A. A.; Lipton, A. S.; Baker, P. J.; Ceder, G.; Vaughey, J. T.; Cabana, J. Probing Mg Migration in Spinel Oxides. *Chem. Mater.* **2020**, *32* (2), 663−670.
(50) Zhang, R. Z.; Reece, M. J. Review of High Entropy Ceramics: Design, Synthesis, Structure and Properties. *J. Mater. Chem. A* **2019**, *7*, 22148−22162.
(51) Oses, C.; Toher, C.; Curtarolo, S. High-Entropy Ceramics. *Nat. Rev. Mater.* **2020**, *5* (4), 295−309.
(52) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A.; Ong, S. P. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *J. Phys. Chem. A* **2020**, *124* (4), 731−745.
(53) Li, X.-G.; Hu, C.; Chen, C.; Deng, Z.; Luo, J.; Ong, S. P. Quantum-Accurate Spectral Neighbor Analysis Potential Models for Ni-Mo Binary Alloys and Fcc Metals. *Phys. Rev. B* **2018**, *98* (9), 094104.
(54) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31* (46), 1902765.
(55) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High- Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
(56) Behler, J. Neural Network Potential-Energy Surfaces in Chemistry: A Tool for Large-Scale Simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930−17955.
(57) Artrith, N.; Urban, A. An Implementation of Artificial Neural-Network Potentials for Atomistic Materials Simulations: Performance for TiO2. *Comput. Mater. Sci.* **2016**, *114*, 135−150.
(58) Cooper, A. M.; Kästner, J.; Urban, A.; Artrith, N. Efficient Training of ANN Potentials by Including Atomic Forces via Taylor Expansion and Application to Water and a Transition-Metal Oxide. *NPJ. Comput. Mater.* **2020**, *6*, 54.
(59) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 184115.
(60) Szlachta, W. J.; Bartók, A. P.; Csányi, G. Accuracy and Transferability of Gaussian Approximation Potential Models for Tungsten. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *90*, 104108.
(61) Gubaev, K.; Podryabinkin, E. V.; Hart, G. L. W.; Shapeev, A. V. Accelerating High-Throughput Searches for New Alloys with Active Learning of Interatomic Potentials. *Comput. Mater. Sci.* **2019**, *156*, 148−156.
(62) Podryabinkin, E. V.; Shapeev, A. V. Active Learning of Linearly Parametrized Interatomic Potentials. *Comput. Mater. Sci.* **2017**, *140*, 171−180.
(63) Shapeev, A. V. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Model Simul* **2016**, *14* (3), 1153−1173.
(64) Wood, M. A.; Thompson, A. P. Extending the Accuracy of the SNAP Interatomic Potential Form. *J. Chem. Phys.* **2018**, *148* (24), 241721.
(65) Thompson, A.; Swiler, L.; Trott, C.; Foiles, S.; Tucker, G. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *J. Comput. Phys.* **2015**, *285*, 316−330.
(66) Nguyen, T. T.; Székely, E.; Imbalzano, G.; Behler, J.; Csányi, G.; Ceriotti, M.; Götz, A. W.; Paesani, F. Comparison of Permutationally Invariant Polynomials, Neural Networks, and Gaussian Approximation Potentials in Representing Water Interactions through Many-Body Expansions. *J. Chem. Phys.* **2018**, *148*, 241725.
(67) Nyshadham, C.; Rupp, M.; Bekker, B.; Shapeev, A. V.; Mueller, T.; Rosenbrock, C. W.; Csányi, G.; Wingate, D. W.; Hart, G. L. W. Machine-Learned Multi-System Surrogate Models for Materials Prediction. *npj Comput. Mater.* **2019**, *5*, 51.
(68) Deringer, V. L.; Csányi, G. Machine Learning Based Interatomic Potential for Amorphous Carbon. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 094203.
(69) Wang, J.; Panchal, A. A.; Sai Gautam, G.; Canepa, P. The Resistive Nature of Decomposing Interfaces of Solid Electrolytes with Alkali Metal Electrodes. *J. Mater. Chem. A Mater.* **2022**, *10* (37), 19732−19742.
(70) Hart, G. L. W.; Forcade, R. W. Algorithm for Generating Derivative Structures. *Phys. Rev. B* **2008**, *77* (22), 224115.
(71) Hart, G. L. W.; Forcade, R. W. Generating Derivative Structures from Multilattices: Algorithm and Application to Hcp Alloys. *Phys. Rev. B* **2009**, *80* (1), 014120.
(72) Hart, G. L. W.; Nelson, L. J.; Forcade, R. W. Generating Derivative Structures at a Fixed Concentration. *Comput. Mater. Sci.* **2012**, *59*, 101−107.
(73) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314−319.
(74) Novikov, I. S.; Gubaev, K.; Podryabinkin, E. V.; Shapeev, A. V. The MLIP Package: Moment Tensor Potentials with MPI and Active Learning. *Mach Learn Sci. Technol.* **2021**, *2* (2), 025002.
(75) Bartók, A. P.; Csányi, G. Gaussian Approximation Potentials: A Brief Tutorial Introduction. *Int. J. Quantum Chem.* **2015**, *115*, 1051−1057.
(76) Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B. *ACM Trans. Math Software* **1997**, *23* (4), 550−560.
(77) 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.

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(78) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Math Program* **1989**, *45* (1−3), 503−528.
(79) Anisimov, V. I.; Zaanen, J.; Andersen, O. K. Band Theory and Mott Insulators: Hubbard U Instead of Stoner I. *Phys. Rev. B* **1991**, *44* (3), 943−954.
(80) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(81) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 17953−17979.
(82) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758−1775.
(83) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15−50.
(84) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169−11186.
(85) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, 011002.
(86) Hellenbrandt, M. The Inorganic Crystal Structure Database (ICSD)—Present and Future. *Crystallogr. Rev.* **2004**, *10* (1), 17−22.
(87) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121* (16), 10037−10072.
(88) Kocer, E.; Ko, T. W.; Behler, J. Neural Network Potentials: A Concise Overview of Methods. *Annu. Rev. Phys. Chem.* **2022**, *73* (1), 163−186.
(89) Lun, Z.; Ouyang, B.; Kwon, D.-H.; Ha, Y.; Foley, E. E.; Huang, T.-Y.; Cai, Z.; Kim, H.; Balasubramanian, M.; Sun, Y.; Huang, J.; Tian, Y.; Kim, H.; McCloskey, B. D.; Yang, W.; Clément, R. J.; Ji, H.; Ceder, G. Cation-Disordered Rocksalt-Type High-Entropy Cathodes for Li-Ion Batteries. *Nat. Mater.* **2021**, *20* (2), 214−221.
(90) Deringer, V. L.; Bernstein, N.; Bartók, A. P.; Cliffe, M. J.; Kerber, R. N.; Marbella, L. E.; Grey, C. P.; Elliott, S. R.; Csányi, G. Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics. *J. Phys. Chem. Lett.* **2018**, *9* (11), 2879−2885.
(91) Fujikake, S.; Deringer, V. L.; Lee, T. H.; Krynski, M.; Elliott, S. R.; Csányi, G. Gaussian Approximation Potential Modeling of Lithium Intercalation in Carbon Nanostructures. *J. Chem. Phys.* **2018**, *148* (24), 241714.
(92) Zhang, C.; Sun, Q. Gaussian Approximation Potential for Studying the Thermal Conductivity of Silicene. *J. Chem. Phys.* **2019**, *126*, 105103.
(93) 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* (24), 241722.
(94) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E(3)-Equivariant Graph Neural Networks for Data-Efficient and Accurate Interatomic Potentials. *Nat. Commun.* **2022**, *13* (1), 2453.
(95) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31* (9), 3564−3572.
(96) Deng, B.; Zhong, P.; Jun, K.; Riebesell, J.; Han, K.; Bartel, C. J.; Ceder, G. CHGNet: Pretrained Universal Neural Network Potential for Charge-Informed Atomistic Modeling. *arXiv* **2023**, arXiv:2302.14231.
(97) Reiser, P.; Neubert, M.; Eberhard, A.; Torresi, L.; Zhou, C.; Shao, C.; Metni, H.; van Hoesel, C.; Schopmans, H.; Sommer, T.; Friederich, P. Graph Neural Networks for Materials Science and Chemistry. *Commun. Mater.* **2022**, *3*, 93.
(98) Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Ong, S. P. Learning Properties of Ordered and Disordered Materials from Multi-Fidelity Data. *Nat. Comput. Sci.* **2021**, *1* (1), 46−53.
(99) Westermayr, J.; Chaudhuri, S.; Jeindl, A.; Hofmann, O. T.; Maurer, R. J. Long-Range Dispersion-Inclusive Machine Learning Potentials for Structure Search and Optimization of Hybrid Organic-Inorganic Interfaces. *Digital Discovery* **2022**, *1* (4), 463−475.
(100) 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* (14), 143001.
(101) Hernandez, A.; Balasubramanian, A.; Yuan, F.; Mason, S. A. M.; Mueller, T. Fast, Accurate, and Transferable Many-Body Interatomic Potentials by Symbolic Regression. *npj Comput. Mater.* **2019**, *5* (1), 112.
(102) Chang, R.; Wang, Y.-X.; Ertekin, E. Towards Overcoming Data Scarcity in Materials Science: Unifying Models and Datasets with a Mixture of Experts Framework. *npj Comput. Mater.* **2022**, *8* (1), 242.

https://doi.org/10.1021/acs.jctc.4c00039
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX