

# Machine Learning Interatomic Potentials and Long-Range Physics

Dylan M. Anstine and Olexandr Isayev\*



Cite This: *J. Phys. Chem. A* 2023, 127, 2417–2431



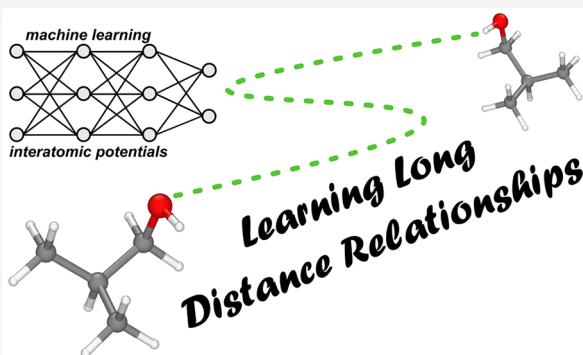
Read Online

ACCESS |

Metrics & More

Article Recommendations

**ABSTRACT:** Advances in machine learned interatomic potentials (MLIPs), such as those using neural networks, have resulted in short-range models that can infer interaction energies with near *ab initio* accuracy and orders of magnitude reduced computational cost. For many atom systems, including macromolecules, biomolecules, and condensed matter, model accuracy can become reliant on the description of short- and long-range physical interactions. The latter terms can be difficult to incorporate into an MLIP framework. Recent research has produced numerous models with considerations for nonlocal electrostatic and dispersion interactions, leading to a large range of applications that can be addressed using MLIPs. In light of this, we present a Perspective focused on key methodologies and models being used where the presence of nonlocal physics and chemistry are crucial for describing system properties. The strategies covered include MLIPs augmented with dispersion corrections, electrostatics calculated with charges predicted from atomic environment descriptors, the use of self-consistency and message passing iterations to propagate nonlocal system information, and charges obtained via equilibration schemes. We aim to provide a pointed discussion to support the development of machine learning-based interatomic potentials for systems where contributions from only nearsighted terms are deficient.



## 1. INTRODUCTION

For the last several decades *ab initio* molecular simulations have been instrumental in overcoming challenges faced by the chemical and materials sciences.<sup>1</sup> These methods have a foundation in the theories of quantum mechanics (QM) and offer computational scientists a means of understanding the atomistic- and electronic-level details governing material and molecule behavior. The value of *ab initio* molecular simulations is proven, and increasing computational power is furthering their widespread use, e.g., the development of exascale computing.<sup>2</sup> Despite this progress, many systems remain too large, and many topics require a number of simulations too great to be investigated solely by QM calculations. The rise of data-driven techniques, particularly simulations performed with machine learned interatomic potentials, has demonstrated the possibility to explore these otherwise computationally demanding areas without sacrificing *ab initio* accuracy.<sup>3</sup>

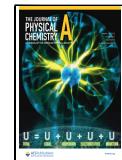
The utility of particle-based molecular and materials simulations is connected with generating an accurate potential energy surface (PES) representation, which is a landscape that underpins reactivity, phase stability, and other observable properties. Arguably, a “holy grail” for computational scientists is to efficiently sample large numbers of readily available and exceptionally accurate PESs. High-quality PESs are typically obtained using QM methods, but there is an appreciable cost

associated with the level of theory used to approximate solutions to the Schrödinger equation.<sup>4</sup> Considering an example of a multielectron ( $N_{elec}$ ) system, the number of arithmetic operations for a QM calculation could scale from order  $O(N_{elec}^3)$  to order  $O(N_{elec}^7)$  depending on the method needed to achieve the desired accuracy.<sup>5</sup> In cases where a system has more than  $10^2$  atoms, the computational requirement of QM methods is dissuading, and computational chemists and materials scientists are relegated to using lower-dimensional representations that trade PES details for efficiency. Fortunately, this accuracy–efficiency trade-off, classically thought of as pervasive and unavoidable, is beginning to be overcome by the rapid expansion of machine learned interatomic potentials (MLIPs).<sup>6–8</sup> MLIPs can be constructed using numerous techniques, such as neural networks, reinforcement learning, or kernel methods. We choose to leave discussion on the subtleties of different approaches to available reviews<sup>9–12</sup> and, instead, highlight

Received: September 28, 2022

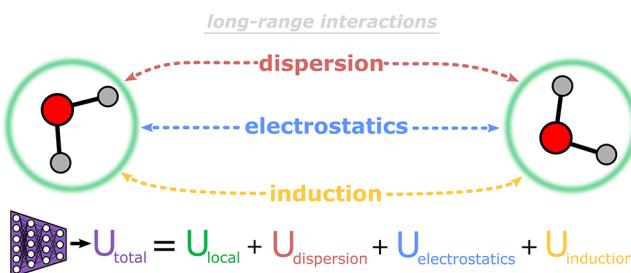
Revised: February 3, 2023

Published: February 21, 2023



herein the specifics regarding MLIPs with long-range interaction strategies.

The overarching concept of MLIPs is for a model to *learn* the relationship between a set of atomic/molecular features and accurate training data. Reference energy and forces can be obtained using high-throughput QM calculations, where the goal of the resulting MLIP model is to achieve near ab initio accuracy with orders of magnitude lower computational cost. Following successful training and validation, an MLIP can infer the interaction energies of systems with feature space representations near the training data. The inference process can be referred to as interpolation and extrapolation; however, the distinction between these terms may lack clarity in high-dimensional model spaces.<sup>13</sup> Many MLIPs reported thus far choose descriptors of local atomic environments for model training. Relying on local features as MLIP input vectors can introduce complications for simulating extended systems with relevant long-range interactions,<sup>14</sup> see Figure 1. Heterogenous



**Figure 1.** Summary of the general energetic contributions composing the total potential energy ( $U_{\text{total}}$ ) of a system.  $U_{\text{local}}$  refers to the short-range system energies and is typically inferred using a machine learning model trained on local features. Dispersion corrections, electrostatics, and induction are collectively referred to as the long-range interaction energy contributions.

bulk phases,<sup>15</sup> dispersion,<sup>16</sup> hydrogen bonding,<sup>17</sup> and extended charge transfer<sup>18</sup> are a few types of phenomena that require careful consideration when constructing an MLIP model. Therefore, while the idea of MLIPs may present itself as straightforward, complexities quickly arise depending on the physicochemical behavior governing the systems of interest.

In our experience, one objection to adopting MLIPs is a lack of demonstrated ability to capture physics and chemistry beyond short-range cutoffs (e.g., 5 Å). Recent progress has produced various MLIP models with simple to sophisticated treatments of long-range interactions,<sup>19</sup> and such an objection is beginning to lose its basis. Despite such developments, demonstrations that these MLIPs can be used to address chemical and materials science challenges are lacking. Future efforts are needed to validate that interaction mechanisms used by modern MLIPs can successfully reproduce relevant long-range physics and chemistry in an application or experimental setting. Thus, this Perspective provides a timely and focused discussion on the current state of developing MLIPs that include long-range interactions. It should be noted that our use of the term *long-range* covers the energetic contributions and structural features exhibited on length scales beyond truncated local atomic environments (see Section 2), which we will refer to as the short-range or nearsighted potentials. A brief background on short-range potentials is given to frame the challenge of including long-range interactions with MLIPs. Intricacies of machine learning practices and an extensive

description of the evolving machine learning landscape in the chemical sciences is outside the topic of this Perspective, and readers are directed to a number of reviews covering these details.<sup>6,19–22</sup> Overall, we discuss models and methods that can be applied to the two major long-range interaction components: dispersions and electrostatics, where the latter has experienced significant methodological development.

## 2. BACKGROUND

Prior to expanding on long-range physical interactions, the key concepts used in developing intrinsically short-range MLIPs are summarized. Our focus is on neural network potentials (NNPs) that utilize an atomic environment vector (AEV) input representation that shares Behler-Parrinello and ANI-like functional forms.<sup>23,24</sup> This input representation is chosen for the general illustration of the issues that nearsighted potentials face. Several other input representations exist,<sup>25–27</sup> and they can also suffer from similar deficiencies. These include spectral neighbor analysis,<sup>28</sup> smooth overlap of atomic positions,<sup>29</sup> and invariant polynomials,<sup>30</sup> to name a few. Regardless, the local interaction energy ( $U_{\text{local,MLIP}}$ ) of an N-particle system can be described as a sum of each particle's potential energy contribution ( $U_{i,\text{MLIP}}$ ).

$$U_{\text{local,MLIP}} = \sum_i^N U_{i,\text{MLIP}} \quad (1)$$

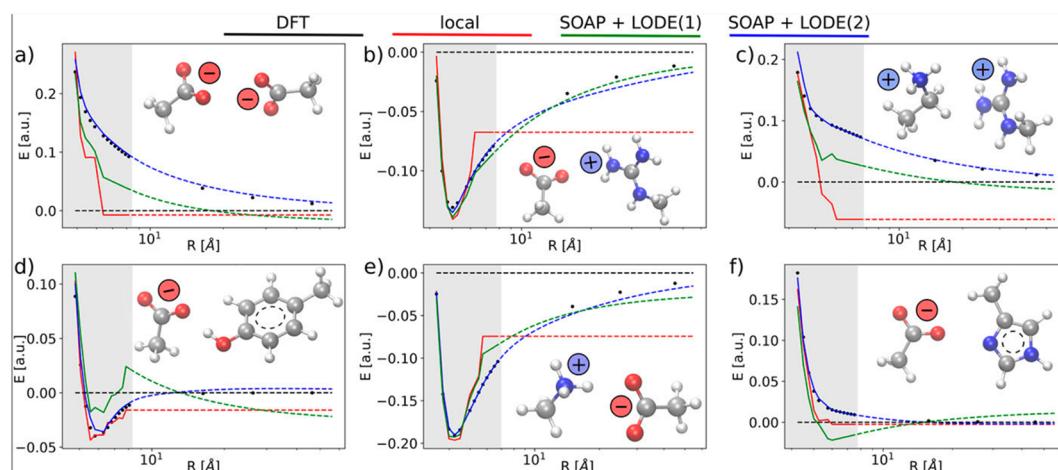
The magnitude of  $U_{i,\text{MLIP}}$  is a function of the environment that particle  $i$  is present in; i.e., the energy is defined by the collection of neighboring particles and their relative positions. NNPs utilize eq 1 by computing particle interaction energies with machine learned predictions based on atomic environment vectors as inputs. An AEV is constructed by defining a suitable cutoff function for pairwise neighbor distances ( $r_{ij}$ ) that smoothly approaches a value of zero at a defined cutoff ( $R_c$ ). ANI potentials utilize a piecewise cosine cutoff function ( $f_c$ ) with the following form.

$$f_c(r_{ij}) = \begin{cases} 0.5 \cdot \cos\left(\frac{\pi r_{ij}}{R_c}\right) + 0.5 & \text{for } r_{ij} \leq R_c \\ 0.0 & \text{for } r_{ij} > R_c \end{cases} \quad (2)$$

The cutoff function is a main component to calculate atom-centered symmetry functions ( $G$ ), which serve as invariant input vectors for machine learned interatomic potentials. Imposing a cutoff is a key topic because it enforces that AEVs are short-range by design, and therefore, the resultant neural network potential is short-range without additional considerations. The AEV, defining an atom's local environment, can be composed of several two-body radial ( $g_{m,i}^R$ ) and three-body angular symmetry ( $g_{m,i}^A$ ) functions.

$$g_{m,i}^R = \sum_{i \neq j}^N e^{-\eta(r_{ij}-R_s)^2} f_c(r_{ij}) \quad (3)$$

$$g_{m,i}^A = 2^{1-\zeta} \sum_{j,k \neq i}^N (1 + \cos(\theta_{ijk} - \theta_s)^\zeta) e^{-\eta\left(\frac{R_{ij}+R_{ik}}{2}-R_s\right)^2} f_c(r_{ij}) f_c(r_{ik}) \quad (4)$$



**Figure 2.** Example of long-range interaction relevance using charged dimer dissociation curves. Non-negligible interaction energies are captured with separation distances up to tens of nanometers. Black dots are DFT reference calculations, red lines are MLIPs with only local descriptors, and green/blue lines are nonlocal MLIP models presented by Grisafi and Ceriotti.<sup>41</sup> Reprinted with permission from ref 41. Copyright 2019, American Chemical Society.

In these equations, the values  $R_S$  and  $\theta_S$  are shifting factors used to probe specific regions of particle  $i$ 's local radial and angular environments, respectively.  $\eta$  and  $\zeta$  control the amount of the atomic environment observed by each probe. A set of shifting and width parameters is denoted by the subscript  $m$ , whose systematic variation allows one to include chemical features across the local atomic environment up to the cutoff. The collection of  $G$  values formed across sets  $m$  compose the AEV input representation used by machine learning models. Overall,  $R_S$ ,  $\theta_S$ ,  $\eta$ , and  $\zeta$  values are selected to yield an AEV resolution that distinguishes between diverse systems without unnecessary  $G$  calculations. The development of eqs 3 and 4 from the cosine cutoff function introduces a challenge for accurately representing a PES, namely, AEVs built from these functions produce machine learned (ML) models blind to long-range structure and physical phenomena.<sup>31</sup>

Depending on the construction of the AEV input representations, an ML potential model can be trained to infer atomic interaction energies and forces. To provide a concise overview, we restrict our following discussion mainly to deep NN models; however, similar considerations apply for alternative ML approaches, for example, Gaussian-approximation potentials (GAPs).<sup>32</sup> In multilayer NN models, the input representation and output energy are the physically relevant quantities. The intermediate layers consist of parameters that are optimized during training to provide the nonlinear transformation of the input representation needed for an accurate inference of energy and forces.

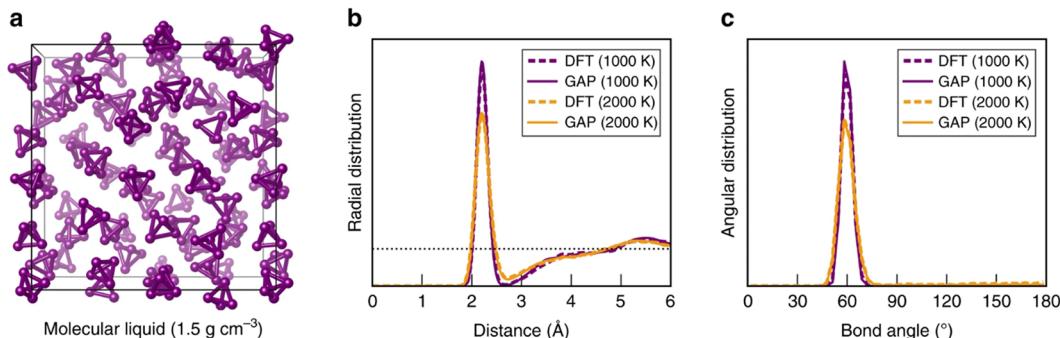
### 3. MLIPS WITH LONG-RANGE INTERACTIONS

Long-range interactions in extended chemical systems contribute to numerous physical phenomena: e.g., thermodynamic phase behavior,<sup>31</sup> variations in conformer geometry,<sup>33</sup> permeation rates of molecules through porous materials,<sup>34</sup> protein structure and dynamics,<sup>35</sup> self-assembly or directed assembly of macromolecular complexes,<sup>36</sup> and interfacial properties.<sup>37</sup> It could be suggested that interactions that occur over longer length scales can be directly predicted with an MLIP model by increasing the cutoff radius used by the atom-centered spherical symmetry functions (or equivalent local atomic descriptor). To a certain extent, it is possible to

include longer-range contributions with model components specifically trained to reproduce intermolecular interactions; for instance, AP-NET utilizes 8 Å cutoff atom-pair symmetry functions for evaluating monomer–monomer interaction energies.<sup>38</sup> There is an eventual pitfall for increasing the AEV cutoff: namely, the number of descriptor calculations increases, which yields larger compute requirements to capture physical interactions that slowly decay with separation distance. The learning task can also become more challenging due to the fact that the chemical configuration space grows with larger cutoff radii, thus, the issue of sufficient sampling compounds.

Distance scaling relationships of long-range interactions vary depending on the systems studied and the physical phenomena considered. Most often applied in classical empirical force fields are Coulomb and Lennard-Jones potentials, displaying  $r_{ij}^{-1}$  and  $r_{ij}^{-6}$ . Beyond these common functional forms, polarization contributions display  $r_{ij}^{-4}$  distance scaling for the leading interaction term (monopole-dipole) of the induction energy.<sup>39</sup> Even further, work by Ambrosetti demonstrated that the power-law scaling of van der Waals interactions can range between ca. -2 and -5 depending on the type of nanostructure and separation distance, which is a result of wavelike fluctuations in charge density.<sup>40</sup> This list of slowly decaying distance relationships is not exhaustive, but it emphasizes that MLIP developers need to carefully consider the length scales of chemical and physical interactions underpinning the target application space. As a specific example, the dissociation curves of charged dimers are cases where non-negligible long-range energy contributions are observed up to tens of angströms, see Figure 2.

Most MLIPs can be classified according to three main strategies when considering their treatment of long-range interactions and nonlocal phenomena: (1) neglect their contribution and only include short-range features, (2) augment MLIPs with standard long-range functional forms, e.g., Coulomb's law, where parameters have local environment dependency, and (3) MLIPs with long-range interactions that are sensitive to global system characteristics. The choice to adopt one of these three strategies should be physically motivated. As an example, systems with a narrow range of



**Figure 3.** Molecular dynamics simulation results using a Gaussian approximation potential (GAP) to model the interactions of condensed-phase phosphorus. (a) Visualization of the liquid-phase phosphorus simulation cell. (b) 2-Body distance correlation function, where DFT (purple) is the reference calculation and GAP (orange) is the Gaussian approximation potential ML model. (c) 3-Body angular correlation function. Reprinted with permission from ref 60. Copyright 2020, The Authors.

elements and short screening distances can oftentimes be simulated using strategy 1 with minimal loss of accuracy, which has been demonstrated across many studies.<sup>23,42–45</sup> In these cases, the adoption of strategy 2 or 3 results in additional computational expense and training complications without appreciable accuracy gain for the target application. On the contrary, employing strategy 1 for systems displaying phenomena such as long-range polarization and electrostatics will likely lead to poor predictions. The concept of choosing an appropriate model design for a given system is well-known in traditional molecular simulations and extends to those being performed with MLIPs.

**3.1. Dispersion Corrections.** van der Waals forces are a ubiquitous type of atomic interaction that originate from fluctuations in electron density distributions.<sup>46</sup> They can be partitioned into short-range repulsion terms and long-range attraction; the latter being referred to as the dispersion interaction. It is not necessary to expand on the inclusion of short-range repulsions in our MLIP discussion because they can be learned by a nearsighted model or explicitly added using a collection of empirical terms, such as the Ziegler-Biersack-Littmark (ZBL) potential.<sup>47,48</sup> For small isolated systems, which define the application space of many reported MLIPs, the contribution of dispersion interactions can be minimal, and as a result, their neglect or inclusion may not affect inference accuracy. However, the collective strength of dispersion forces in larger systems is often non-negligible, and they can be a significant factor in defining properties such as polymer cohesive energy density<sup>49</sup> and organic molecular complex stability,<sup>50</sup> to name a few. Therefore, it is worthwhile to expand on the treatment of long-range dispersion interactions in the design of an MLIP.

One level of consideration for dispersion interactions is in the construction of the reference training data. MLIP model developers are often guided by the goal of obtaining density functional theory (DFT)-quality accuracy with near empirical potential computational cost. This often equates to applying dispersion correction strategies<sup>51</sup> to DFT calculations that use, for instance, generalized gradient approximation density functionals,<sup>52</sup> or one can choose to neglect dispersion correction in the reference data and include them explicitly as part of the MLIP model. As an example, the ANI family of MLIPs is based on locally defined symmetry functions (see Section 2), yet dispersion corrections are applied to the entire system *ad hoc*, regardless of AEV cutoff, to expand the interaction range of the otherwise inherently nearsighted

models. As a counterexample, Morawietz et al.<sup>53</sup> trained a set of models, also based on locally defined symmetry functions, directly to dispersion-corrected reference data to obtain good agreement with the density maximum of bulk water studied as a function of temperature. The key distinction between these two examples is whether dispersion correction is applied “on-the-fly” during inference or learned implicitly by the MLIP. The latter strategy can be more computationally efficient; however, it will ultimately lead to incorrect predictions for systems where dispersion contributions beyond the local AEVs cannot be neglected. The MLIP training data of Morawietz et al. consisted of DFT calculations augmented with the D3 dispersion correction scheme from the 2010 landmark work of Grimme and co-worker.<sup>54</sup> D3-corrected DFT calculations are a frequently used tool for construction of MLIP reference training data for organic systems; see eq 5 for the 2-body dispersion correction functional form

$$U_{\text{Disp}}^{\text{2-body}} = \sum_{ij} \sum_{n=6,8,10,\dots} s_n \frac{C_n^{ij}}{r_{ij}^n} f_{d,n}(r_{ij}) \quad (5)$$

where  $s_n$ ,  $C_n^{ij}$ , and  $f_{d,n}$  represent the scaling factors, dispersion coefficients, and damping functions, respectively. Despite the popularity, it is worth commenting that dispersion-corrected DFT is an active field of research and defaulting to D3 corrections could incur error depending on the chemical diversity and/or system size reflected in the target application space.<sup>55,56</sup> Moreover, the importance of many-body dispersion effects is another consideration, which can result in non-negligible interaction energy contributions with increasing system size. As an example, Tkatchenko et al. used crystalline benzene as a model system to demonstrate that pairwise Tkatchenko-Scheffler dispersion corrections overestimated the crystal cohesive energies and that a many-body dispersion (MBD) correction based on coupled and uncoupled quantum harmonic oscillators substantially corrected this error.<sup>57</sup> In cases where the D3 method of Grimme is insufficient, one could seek to improve the interaction energies by including a 3-body dispersion term ( $C_9$ ) with the Axilrod-Teller form.<sup>58</sup> These two examples are select cases among many, and those looking to develop MLIPs that can be applied to systems with long-range and many-body dispersion effects are directed to the in-depth discussion provided by Hermann, DiStasio, and Tkatchenko<sup>46</sup> or Grimme and co-workers.<sup>59</sup>

For bulk systems with long screening distances, an MLIP built strictly using  $U_{\text{local}}$  trained to DFT-D reference data can

be insufficient. This is related to the previously discussed limitation of nearsighted potentials, namely, they lack knowledge of structure or interactions beyond the local atomic environment. Instances where nonlocal structure impacts local electron density fluctuations can require dedicated terms for accurate modeling. Deringer et al. successfully implemented one solution, where an explicit two-body  $r_{ij}^{-6}$ -dependent dispersion interaction term is used to reasonably reproduce the structure of condensed-phase phosphorus (see Figure 3) by training on DFT data from calculations using the PBE functional and many-body dispersion correction.<sup>60</sup> Condensed-phase phosphorus is challenging to simulate because of its mixture of covalent and noncovalent features; however, it could be successfully modeled using only  $U_{\text{local,MLIP}}$  and a two-body  $r_{ij}^{-6}$ -dependent dispersion term because the systems were chemically homogeneous and neutral. The concept of including an  $r_{ij}^{-6}$  dispersion interaction term alongside an ML-based  $U_{\text{local,MLIP}}$  calculation was also used by Wen and Tadmor for multilayer graphene, which was similarly trained with PBE and many-body dispersion-corrected DFT calculations.<sup>61</sup> The work by Muhli et al.<sup>62</sup> utilized a different strategy, where short-range descriptors were used to predict the effective atomic Hirshfield volumes needed to employ the dispersion correction scheme of Tkatchenko and Scheffler,<sup>63</sup> which enabled them to investigate the C<sub>60</sub> phase diagram and associated driving forces up to 5000 K and 1000 GPa. Their model, trained on PBE reference data from DFT calculations, qualitatively reproduced a number of the carbon structure observations obtained experimentally. The successes of these two carbon MLIPs have a similar basis as the model by Deringer et al., namely, contributions from long-range electrostatics and polarization can be neglected because the systems are chemically homogeneous and treated as closed-shell. Incorporating explicit dispersion energies with parameters (or their dependencies) obtained from ML-based predictions using local structure is an approach that conceivably encourages NNP generalizability and is worth further exploration.

**3.2. Inference of Ab Initio Point Charges.** Augmenting a short-range MLIP with an explicit electrostatic energy term ( $U_{\text{es}}$ ), see eq 6, is a straightforward approach to include long-range interactions between charge sites into an ML-based model. From a quantum mechanical view of atoms,  $U_{\text{es}}$  is a complicated function of the electron density, and it is oftentimes computationally convenient to condense this distribution onto localized point charges. One difficulty that arises when using this strategy to account for long-range electrostatic interactions is that mapping electron density/electrostatic potentials onto point charges is an ambiguous task that lacks a unique solution. This has led to numerous electron density partitioning schemes<sup>64–66</sup> that all give distinct atom-centered point charges.<sup>67</sup> The consequence is that NNPs can be constructed using a variety of charge assignment strategies that vary in transferability, yielding an important consideration for interpreting simulation results.

$$U_{\text{system}} = U_{\text{local,MLIP}} + U_{\text{es}} \quad (6)$$

Some developers opt to augment their ML-based short-range potentials with Coulomb's law and Ewald summation using static formal charges. This strategy is useful when integer charges can be rationally assigned using fundamental chemical principles; for example, this has been demonstrated in simulations of crystalline nitride materials.<sup>68</sup> Electrostatic

interactions based on formal charges are simple but narrow in the types of systems they can be applied to, and such a strategy is not applicable to most organic systems. Atom-centered charges of organic species are typically noninteger and exhibit variation with the atomic local environment. This originates from chemically identical molecules with unique conformations displaying differences in electron density distribution. One can address conformational-dependent partial charges by training an ML model to predict them from the AEV input representations. The system energy can be reformulated as a local interaction energy with monopole electrostatic contributions as

$$U_{\text{system}} = U_{\text{local,MLIP}} + \sum_{i=1}^{N-1} \sum_{j>i}^n \frac{q_i(\{g_{m,i}^{\text{R}}, g_{m,i}^{\text{A}}\}) q_j(\{g_{m,j}^{\text{R}}, g_{m,j}^{\text{A}}\})}{4\pi\epsilon r_{ij}} \quad (7)$$

where  $q_i$  and  $q_j$  are the charges on atoms  $i$  and  $j$  separated by a distance  $r_{ij}$ , and  $\epsilon$  is the screened dielectric constant. We are omitting dispersion and induction terms from eq 7 for clarity. It is important to note that this equation is only valid when phenomena related to electron density (re)distribution—for example, induction or multipole electrostatics—are dominated by features on a length scale less than the  $U_{\text{local,MLIP}}$  cutoff. Instances where this is not the case are discussed further in Section 4. Moreover, eq 7 can only be appropriately applied for systems where the dielectric screening is homogeneous. Regardless, learning conformation-dependent charges is used by several MLIPs, including PhysNet,<sup>69</sup> HIP-NN,<sup>70</sup> and models by Behler and co-workers.<sup>71,72</sup> It is worth highlighting that charges predicted from ML models may need to be adjusted to ensure they sum to the net charge of the system. Choosing to use simple charge distribution normalization or more complex species-specific weighting schemes can affect model transferability and accuracy.

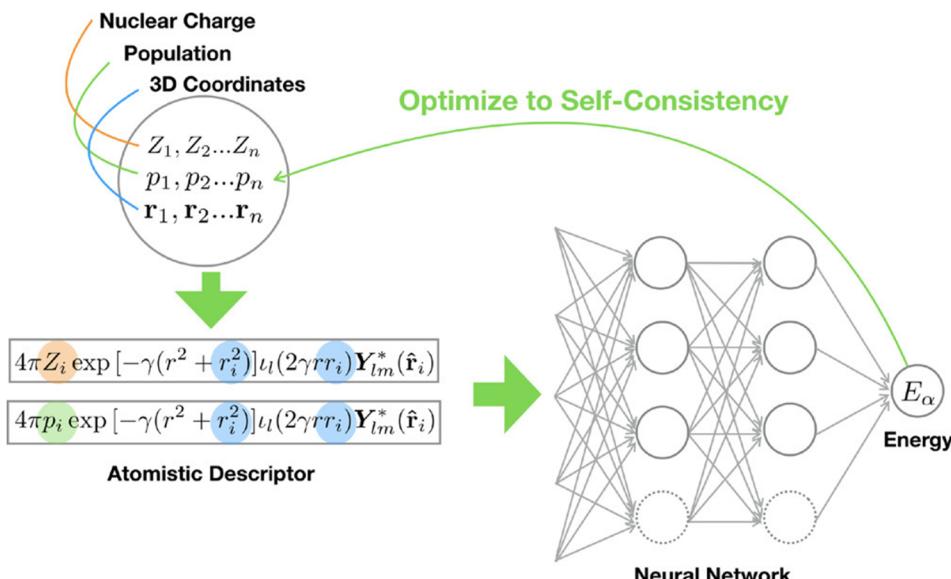
An alternative to predicting atom-centered point charges has recently been used in two studies reported by Zhang et al.,<sup>73,74</sup> where electrostatics are calculated using the concept of maximally localized Wannier functions.<sup>75</sup> Their approach maintains the high-level form of eq 6 but expands the  $U_{\text{es}}$  term as a function of Wannier centers (WCs) with a charge value of 2e- and ions having a charge value of atomic nuclei + core electrons. As an example, a water molecule has four WCs each carrying 2e- and three nuclear positions with 1e, 1e, and 6e for the hydrogens and oxygen, respectively. The total system energy is calculated by representing the potential energy of WC and nuclei-centered charges with Gaussian charge distributions ( $U_{G_i}$ ).

$$U_{\text{system}} = U_{\text{local,NN}} + U_{G_i} \quad (8)$$

The calculation of  $U_{G_i}$  can be carried out in Fourier space as

$$U_{G_i} = \frac{1}{2\pi V} \sum_{m \neq 0, |m| \leq L} \frac{e^{(-\frac{\pi^2 m^2}{\rho^2})}}{m^2} S^2(m) \quad (9)$$

where  $L$  is the Fourier space cutoff,  $V$  is the volume of the simulation cell, and  $S(m)$  is the charge weighted structure factor of the nuclei and WC positions. Model implementation requires two ML components to be trained using input from the local atomic environments: (1) to calculate  $U_{\text{local,NN}}$  and (2) to provide the WC positions. This framework has not yet been applied to a diversity of systems with nonlocal effects, but



**Figure 4.** Overview of the BpopNN architecture. A self-consistent charge optimization scheme is used to iteratively refine atom-centered partial point charges and minimize the total predicted energy with respect to the overall charge distribution. Reprinted with permission from ref 77. Copyright 2020, American Chemical Society.

the current design effectively captures long-range electrostatic interactions, which, for example, have been demonstrated in different types of water dissociation curves. Regardless, it should be emphasized that the concept of using WCs as the basis for electrostatics is attractive because they are not restricted to the spherical simplifications of atom-centered point charges. This may lead to better MLIP predictions because of a flexible representation of the electron density distribution.

**3.3. Models Based on Iterative Refinement.** In contrast to single-step predictions of electrostatic contributions, a number of MLIP models have devised schemes based on iterative convergence. We choose to divide these models into self-consistency and message-passing methods. The defining feature of self-consistent approaches is the iterative refinement toward a converged prediction from an initial estimate, where the objective value is related to an output which itself is used to derive new inputs (for example, charges). Differing from the cyclic convergence of a self-consistent procedure, message-passing methods use consecutive calculation steps that are meant to propagate chemical information across a system through communication between neighboring atoms. These propagation steps, i.e., message passes, can be thought of as increasing the amount of nonlocal details available at each atomic site that can be used to predict chemical properties such as atomic charges. Additional message-passing steps alleviate the difficulties associated with methods only using truncated local environments, e.g., those discussed in Section 2, because input representations are refined using information about distant chemical structures that is acquired through communication with intermediate neighboring atoms. This iterative propagation of information is related to the idea of molecular graph-like representations, where atoms (nodes) are connected to their neighbors (edges) to form a “message-passing network”. One of the greatest advantages of MLIP models that use message passing is their ability to learn their own flexible chemical representations that are not bound by nearsighted features, which conceptually encourages extensi-

bility and generalizability. Several diverse MLIP models applying these strategies are highlighted herein.

The work by Gao and Remsing reports a long-range MLIP approach called self-consistent field neural network (SCFNN),<sup>76</sup> which combines an iterative refinement approach with maximally localized Wannier centers (WCs) for calculating electrostatics. The strategy of SCFNN is to execute two MLIP components sequentially: (1) a set of neural networks for predicting WC positions and the change in those positions due to an effective electric field, and (2) a set of networks for predicting local configurational and effective field forces on the atoms. A self-consistency procedure is implemented to converge the WC positions. Following initial WC estimates, via a neural network, a loop is carried out where perturbations to the WC positions are calculated based on the effective electric field, which itself is a function of the WC positions. After applying this perturbation, a new effective field is calculated, and the procedure repeats until the inferred perturbation to the WC position is below a chosen threshold ( $\sim 10^{-4}$  Å). Eventually a converged electric field is obtained, and the authors use it as input alongside AEV symmetry functions to calculate atomic forces with the second set of neural networks. SCFNN has been applied to a single chemical system, water, where structural features, long-range polarization, and electronic properties were reproduced with appreciable accuracy. This procedure by Gao and Remsing has some resemblance with the works of Zhang et al.,<sup>73,74</sup> however, the separation of short- and long-range electrostatics, use of neural networks instead of explicit functional forms, and assumption of linear response are notable distinguishing features. Both studies report quality results for water systems, and the extent to which the approaches are generalizable to diverse systems is worth further investigation.

Xie et al. reported the so-called Becke Population Neural Network (BpopNN),<sup>77</sup> which focuses on refining  $q$  predictions obtained from modified smooth overlap of atomic position descriptors. An innovative feature of BpopNN is the use of a self-consistent charge update scheme (SCF-q). The method-

ology seeks to train a ML model to learn a potential energy functional form ( $U_{\text{BpopNN}}$ ) that depends on the nuclear charges ( $Z$ ), Becke populations ( $P$ ), and the atomic positions ( $r$ ).

$$\begin{aligned} U_{\text{BpopNN}} = & U_{\text{local,NN}}(Z, P, r) + U_{\text{es}}(Z, P, r) \\ & + U_{\text{intra}}(Z, P, r) \end{aligned} \quad (10)$$

The potential energy contributions having a dependency on  $P$  enable the SCF-q scheme to be used. Following model training on constrained DFT reference data,<sup>78</sup> inference can be carried out by minimizing the total energy with respect to the Becke populations  $\left(\frac{\nabla U_{\text{BpopNN}}}{\nabla P} \approx 0\right)$ , which can be performed using optimizers implemented in automatic differentiation packages, e.g., stochastic gradient descent or the Adaptive moment solver (Adam).<sup>79</sup> In practice, initial guess charges are supplied, and iterative optimization steps are used to arrive at an accurate distribution of atom-centered charges. The selection of initial guess charges is a point worth emphasizing because they can influence the optimized charge distribution identified by the self-consistency procedure. An overview of the BpopNN model workflow and architecture is provided in Figure 4. To the best of our knowledge, BpopNN has only been applied to proof-of-concept examples of interaction energies, structural transitions, and charge distributions of lithium hydride systems.

Message-passing neural networks (MPNNs) are a versatile class of machine learning models that use iterative updates as part of their predictions.<sup>80</sup> Each iteration, referred to as a message-passing step ( $t$ ), increases the amount of nonlocal information available to every atom in a system. While a clear practicality issue exists with performing large numbers of message passes, for many organic systems (those without large nonlocal effects)  $t = 3$  can be sufficient to reach accurate predictions.<sup>81</sup> It is worth expanding upon the general message-passing framework to present the mechanism that MPNN-based MLIPs can use to include nonlocal interactions. The central idea of an MLIP with message passing is for atoms in a system to maintain an abstract hidden state representation ( $h_i^t$ ) that is updated through communication with nearby neighbors. The details of  $h_i^t$  are not informed by the model developer, and the MLIP learns them during training. In most cases, this necessitates the training of two neural networks: a message-passing function and a hidden state update function. Neural networks are not strictly required, and any sufficiently flexible and learnable model can be used. Regardless, the updated hidden state maintained by atom  $i$  ( $h_i^{t+1}$ ) is determined using the update function ( $U_t$ ), which transforms the previous hidden state ( $h_i^t$ ) based on the cumulative message ( $m_i^{t+1}$ ) received.

$$h_i^{t+1} = U_t(h_i^t, m_i^{t+1}) \quad (11)$$

The message-passing function ( $M_t$ ) is responsible for defining the contribution that each neighbor of atom  $i$ , denoted as  $N(i)$ , makes to the cumulative message. The atomic neighbors are defined up to a cutoff distance, e.g., atomic separations  $\leq 5.0 \text{ \AA}$ .

$$m_i^{t+1} = \sum_{j \in N(i)} M_t(h_i^t, h_j^t, e_{ij}) \quad (12)$$

$M_t$  is learned through training—similar to  $U_t$ —and is a function of each atom's hidden state and any additional

features ( $e_{ij}$ ) between atoms  $i$  and  $j$ . The fact that eq 12 has a dependency on both  $h_i^t$  and  $h_j^t$  is crucial because it supports that every message-passing iteration propagates system information from further away atoms. Thus, increasingly nonlocal features can be incorporated via more message-passing steps to refine each atom's hidden state. After a number of message-passing iterations, the hidden states can be utilized for the inference of atomic properties, including energies, charges, electronegativity, and so on.

Several models have been based on iterative update/message-passing schemes, including Deep Tensor Neural Networks (DTNNs),<sup>82</sup> SchNet,<sup>83</sup> SpookyNet,<sup>84</sup> and AIMNet.<sup>81</sup> To demonstrate the transferability of using iterative schemes, we expand on the AIMNet model, which was reported by Zubatyuk et al. in 2019 to have applicability to neutral organic molecules composed of HCNOFSCI.<sup>81</sup> In AIMNet the hidden state representation takes the form of a 16-dimensional embedding vector ( $a_z$ ) that is used for defining atomic features. As an aside, the atomic embedding strategy is used because it is a solution to the unfavorable scaling problem that species-specific networks face when the number of elements in an application space grows. Regardless, the radial and angular atomic environments, as defined in Section 2, are multiplied with the atomic embeddings according to

$$G_{m,i}^R = \sum_j g_{m,i,j}^R \otimes a_j \quad (13)$$

and

$$G_{m,i}^A = \sum_{j,k} g_{m,i,j,k}^A \otimes F_{\text{NN}_i}([a_j \cdot a_k, a_j + a_k]) \quad (14)$$

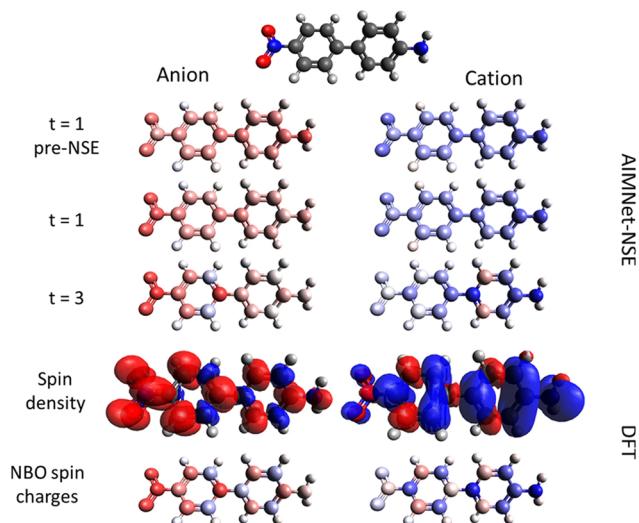
where  $g_{m,i,j}^R$  and  $g_{m,i,j,k}^A$  are vectors of the  $j$  and  $j,k$  neighboring atom components of eqs 3 and 4, respectively. In eq 14,  $F_{\text{NN}}$  is a neural network trained to construct an effective angular embedding vector, and subscripts  $i$ ,  $j$ , and  $k$  are the atoms composing the molecular geometry component. The  $a_z$  values are updated according to the scheme described above, where neighboring messages are directly accounted for up to  $\sim 5 \text{ \AA}$  with each iterative step. The final atomic feature vector ( $f_i$ ) used for molecular property prediction is a concatenation of the terms from eqs 13 and 14 that are flattened and undergo a nonlinear transformation via another neural network.

$$f_i = F_{\text{NN}_2}(G_{m,i}^R, G_{m,i}^A) \quad (15)$$

AIMNet uses  $f_i$  to predict system features, such as atomic partial charges, which accurately reproduce those derived from DFT calculations. Furthermore, additional nonlocal details are beginning to be included in message-passing models. For example, SpookyNet<sup>84</sup> and AIMNet-NSE<sup>85</sup> can both model spin states, which are valuable for simulations including reactive events and open-shell species. These models are motivated by the fact that MLIPs that rely only on nuclear degrees of freedom lack an ability to accurately infer interaction energies of species other than neutral singlets. There are a number of cases such as bond-breaking or transition-state chemistry, where failing to account for a proper electronic spin state results in a drastic misrepresentation of the potential energy surface; for example, see Figure 4b of ref 84. To highlight one working mechanism, AIMNet-NSE introduced a neutral spin equilibration procedure that is applied as

$$q_i^s = \tilde{q}_i^s + \frac{f_i^s}{\sum_{j=1}^N f_j^s} \left( Q^s - \sum_{j=1}^N \tilde{q}_j^s \right) \quad (16)$$

where  $Q^s$  is the total molecular spin charge state, and  $\tilde{q}_i^s$  and  $f_i^s$  are partial spin-polarized atomic charges and weight factors predicted by the neural network, respectively. Subsequent message-passing steps further improve the AIMNet-NSE model prediction of localized spin charge states. As an example, Figure 5 shows spin charge state predictions of the



**Figure 5.** Example of using AIMNet with neutral spin equilibration and message passing to accurately assign atom-centered partial spin charges of anionic and cationic 4-amino-4'-nitrobiphenyl. Each message pass ( $t$ ) refines the spin-charge density distribution, resulting in accurate predictions in comparison with NBO analysis. Reprinted with permission from ref 85. Copyright 2021, The Authors.

anionic and cationic forms of 4-amino-4'-nitrobiphenyl as a function of message passes. It can be seen that, for  $t = 1$ , i.e., without neutral spin equilibration, the spin charge-densities are approximately equivalent. However, after  $t = 3$  more accurate wave-like spin charge density distributions are predicted, showing improved agreement with DFT natural bond orbital (NBO) analysis. Overall, the ability of AIMNet-NSE to cover an application space of neutral, cationic, and anionic species with accurate spin charge density distributions is a significant step toward MLIPs with broader generalizability.

#### 3.4. Potentials Incorporating Charge Equilibration.

The empirical interatomic potential simulation community has been interested in calculating geometry-dependent atomic charges for several decades. For instance, force fields such as ReaxFF<sup>86</sup> have accomplished this task by using charge equilibration (Qeq) schemes. For the purpose of clearly distinguishing MLIP methodologies in this Perspective, we chose to narrowly define Qeq to refer to methods that determine the charge distribution of a given configuration by solving a system of linear equations that include the interaction energies between charge densities, the atomic partial charges, and the atomic electronegativities. An important feature of this formulation is that the electrostatic interactions between all pairs of atoms are typically used to construct the charge–charge interaction energy tensor, and thus, solving the Qeq linear equations results in a fully global charge redistribution. Among the most notable Qeq methodologies proposed is the

one by Rappe and Goddard,<sup>87</sup> which has found recent use in MLIP formulations. The fact that Qeq methods are well-established molecular simulation techniques and are not directly restricted by nearsighted approximations makes them an attractive option to be utilized by MLIP practitioners. It should be noted that many of the MLIPs discussed below are methodologically consistent with that of Rappe and Goddard, but several charge equilibration variants exist, such as the electronegativity-equalization method (EEM),<sup>88</sup> split-charge equilibration (SQE),<sup>89</sup> and atom-condensed Kohn–Sham DFT approximated to second order (ACKS2),<sup>90</sup> to name a few.

The Qeq approach is described with the following functional form

$$U_{\text{es}, \text{Qeq}} = \sum_{i=1}^N (q_i \chi_i + 0.5 J_i q_i^2) + \sum_{j>i} \frac{q_i q_j}{r_{ij}} \quad (17)$$

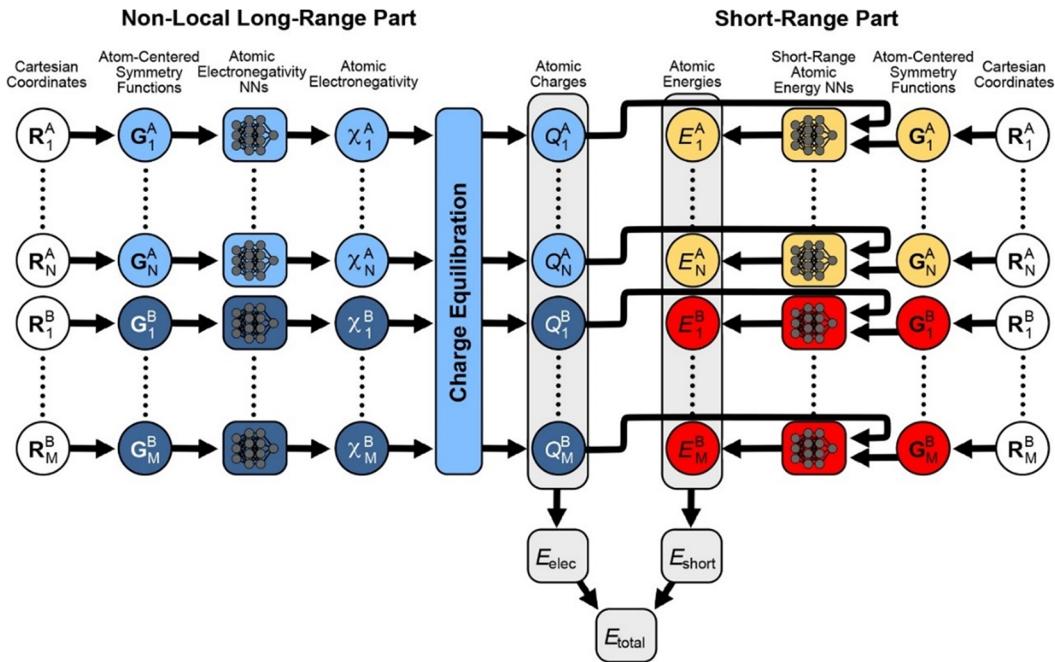
where  $\chi_i$  and  $J_i$  are the electronegativity and atomic hardness of atom  $i$ , respectively. The Coulomb calculation can be carried out using a long-range summation technique, e.g., Ewald or Wolf summation,<sup>91,92</sup> but it has been omitted from eq 17 for clarity. The objective of Qeq is to use optimization techniques to solve the charge distribution  $\{q_1, \dots, q_N\}$  that minimizes  $U_{\text{es}, \text{Qeq}}$ . This optimization is performed with the constraint of

$$Q = \sum_i q_i \quad (18)$$

which ensures that the sum of partial atomic charges is equal to the net charge of the system ( $Q$ ). The distinguishing feature between MLIPs that employ Qeq formalisms for their electrostatic calculations is their treatment of  $\chi_i$  and  $J_i$ . A straightforward strategy is to apply parameters from existing empirical data sets. This was employed by Yoo et al.<sup>93</sup> in developing a CHNO reactive MLIP, where they used  $\chi_i$  and  $J_i$  from ReaxFF to model small molecule bond dissociation and various chemical features and reaction properties of 1,3,5-trinitroperhydro-1,3,5-triazine (RDX).

Instead of adopting predetermined Qeq parameters, the work of Nokikov and Shapeev applied regression techniques to derive system-specific values for silica.<sup>94</sup> Their study used moment tensor potentials augmented with a standard Qeq approach to evaluate phonon spectra, structural properties, and the elastic tensor of  $\alpha$ -quartz. Interestingly, they found the Qeq scheme with parameters obtained via iterative optimization yield no significant increase in accuracy but contributed further to prediction uncertainty. The idea of a simple combination between short-range MLIPs and charge equilibration is compelling, but this result indicates a more intimate connection between local and long-range model components might be required for predictive accuracy.

To highlight an MLIP that utilizes charge equilibration with integrated short- and long-range components, the so-called fourth-generation Behler-Parrinello NNPs (4GNPP) of Behler, Goedecker, and co-workers<sup>95</sup> is discussed. 4GNPP is a slightly modified local Behler-Parrinello model combined with the CENT architecture, which is an ML-based charge equilibration framework originally aimed at ionic crystal applications.<sup>96</sup> The total energy of a 4GNPP uses the same high-level form of eq 6, where the calculation occurs in two parts. In the first step, the CENT scheme is applied to calculate  $U_{\text{es}}$  from the interaction of atomic charge densities. A Gaussian



**Figure 6.** Schematic overview of the 4GNNP architecture. A charge equilibration scheme (shown on the left) is employed to obtain the globally optimized partial point charges used in calculating long-range electrostatic interactions. The short-range potential (shown on the right) involves the 2nd generation BPNNP with an additional input parameter for the charge value on each atom. Reprinted with permission from ref 97. Copyright 2021, American Chemical Society.

functional form is used to distribute the charges centered on atomic positions as

$$\rho_i(r_i) = \frac{q_i}{\alpha_i^3 \pi^{3/2}} \exp\left(-\frac{|r - r_i|^2}{\alpha_i^2}\right) \quad (19)$$

with  $r_i$  and  $\alpha_i$  being the atomic position and distribution width parameter, respectively. Qeq is applied to determine the solution of  $\{q_1, \dots, q_N\}$  that minimizes  $U_{\text{es,CENT}}$ , which is expressed similarly to eq 17 as a truncated Taylor expansion in the following form

$$U_{\text{es,CENT}} = \sum_{i=1}^N \left( U_i^0 + q_i \chi_i + 0.5 \left( J_i + \frac{2}{\alpha_i \sqrt{2} \sqrt{\pi}} \right) q_i^2 \right) + \sum_{j>i}^N \frac{q_i q_j}{r_{ij}} \operatorname{erf} \left( \frac{r_{ij}}{\sqrt{\alpha_i^2 + \alpha_j^2}} \right) \quad (20)$$

where  $U_i^0$  is an atomic reference energy. A key contribution of CENT is that  $\chi_i$  is assumed to be environment-dependent and predicted with an ML model from an AEV input representation. It is worth highlighting that there is methodological nuance between the training procedure of the standalone CENT approach and 4GNNPs, and the interested reader is directed to refs 19 and 95 on this point. Regardless, the charge optimization process is a global system operation carried out with electronegativity parameters that vary depending on the local arrangement of atoms. This is the main mechanism that allows for the 4GNNP to account for events like nonlocal charge transfer.

The short-range contribution ( $U_{\text{local,NN}}$ ) is calculated following the optimization of  $\{q_1, \dots, q_N\}$ . Consistent with the description in Section 2, the calculation  $U_{\text{local,MLIP}}$  in the 4GNNP framework is a summation of per atom energies

predicted from element-specific neural networks. A change is made to the input feature vector to also include the charge value determined in the CENT step. Considering this charge value is obtained from a global process, it brings a unique contribution to what would otherwise be a neural network operating on strictly short-range features. An illustrative example of the 4GNNP architecture is shown in Figure 6. The prospect of the 4GNNP scheme is promising with demonstrated success for challenging systems, including ionic long linear alkanes, small sodium chloride clusters, and interactions with a magnesium oxide surface, to name a few.<sup>95,97</sup>

#### 4. POLARIZATION, ML/MM, AND BEYOND

Our discussion up to this point has mainly focused on MLIPs using pairwise dispersion, some comments on the importance of many-body effects, and electrostatics via Coulomb's law (predominately via point charges). Nevertheless, it is known that intermolecular interactions that extend beyond local atomic environments consist of electrostatics, dispersion, and induction;<sup>98</sup> therefore, MLIP models neglecting one or more of these terms are incomplete by design. Most MLIPs discussed in the previous sections that make such assumptions are either applied pragmatically or they do so with a justifiable basis from a chemical understanding of the target application space. Regardless, long-range interactions between molecules and materials are defined by the electric fields produced by the electron density distribution of each species, and thus, electrostatics are accurately described using a multipole expansion. Long-range interactions are often coupled, and neighboring molecules and materials produce deformations of electron densities from the isolated reference state (i.e., polarization) with an associated induction energy. Therefore, an MLIP model that aims to accurately represent an

application space where polarization effects have significant contributions beyond  $U_{\text{local},\text{MLIP}}$  must include explicit functional forms or features operating on the length scales of relevant electron density redistribution. One of the simplest approximations of polarization effects in an MLIP is an electrostatic potential calculated via partial atomic charges that vary as a function of atomic environments. As we have previously noted, reducing a complex electron density distribution to atom-centered point charges can produce non-negligible error,<sup>99</sup> and as a result, this is a poor approximation to account for the induction energy of a system with long-range polarization.

To the best of our knowledge, the only model that includes a dedicated mechanism for long-range polarization contributions is the previously discussed report by Gao and Remsing (see Section 3.3). Their approach operates through self-consistent determination of Wannier center locations through linear perturbations in response to the effective electric fields at these positions. Gao and Remsing's SCFNN model was shown to accurately predict the high-frequency dielectric constant of water, which demonstrates reliable long-range electrostatic screening effects. These results are encouraging; however, the SCFNN model relies on linear response and the partitioning of DFT data into short- and long-range contributions. It is worth commenting that these dependencies do not limit Gao and Remsing's study of chemically homogeneous water systems, but systematically extending this framework to a broad range of chemistry does not appear straightforward. We are unaware of any other MLIP that includes a dedicated explicit mechanism for incorporating induction energy beyond local features. Momentarily disregarding the issues associated with partial point charges (i.e., monopole electrostatics), it could be argued that charge equilibration or message-passing methods can incorporate longer-range charge transfer effects and are better suited for polarizable systems. While this is partially accurate, these methods also have limitations arising from local approximations. For message-passing MLIPs, such as AIMNet, intermediate particles are required to allow the update function to propagate information about nonlocal structure. Thus, a polarizing body whose nearest neighbor is more than one AEV cutoff distance away is effectively treated as an independent system from the message-passing perspective. This is less of an issue for condensed phases, but it can cause incorrectly predicted physics in low-density systems (e.g., vapors). Turning to the Qeq methods, these are performed as a global system operation that allows charge to redistribute by minimizing the electrostatic energy via the distribution of partial charges, which can conceivably capture a degree of polarization effects. For a specific example, the aforementioned 4GNPP predicts electronegativities from local features before carrying out Qeq; however, these are chemically ambiguous properties that also suffer from truncated AEVs, and a similar limitation exists for describing long-range polarization contributions. It is potentially worthwhile to consider strategies used in classical molecular simulations to provide inspiration for MLIP models designed for long-range polarizable systems. A frequently applied empirical model that is based on a multipole expansion is the AMOEBA force field, where permanent atomic multipoles (up to the quadrupole) and polarizable dipoles are used to calculate electrostatic interactions.<sup>100,101</sup> Adapting this representation to an MLIP framework could be valuable for modeling polarizable systems. Briefly, in addition to interactions between permanent

electrostatic multipoles, AMOEBA explicitly includes many-body polarization as

$$\begin{bmatrix} \alpha_1^{-1} & -T_{12}^{\text{d-d}} & \cdots & -T_{1N}^{\text{d-d}} \\ -T_{21}^{\text{d-d}} & \alpha_2^{-1} & \cdots & -T_{2N}^{\text{d-d}} \\ \vdots & \vdots & \ddots & \vdots \\ -T_{N1}^{\text{d-d}} & -T_{N2}^{\text{d-d}} & \cdots & \alpha_N^{-1} \end{bmatrix} \begin{bmatrix} \mu_1^{\text{ind}} \\ \mu_2^{\text{ind}} \\ \vdots \\ \mu_N^{\text{ind}} \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_N \end{bmatrix} \quad (21)$$

, where  $\alpha$ ,  $T^{\text{d-d}}$ ,  $\mu^{\text{ind}}$ , and  $E$  are the atomic polarizability tensors (usually isotropic), Thole damped interaction tensors,<sup>102</sup> induced dipoles, and the polarization electric field at each site, respectively.

An important application where polarization effects can dictate system properties is the so-called machine learning/molecular mechanics (ML/MM) method (e.g., refs 103–105), which is related to the quantum mechanics/molecular mechanics (QM/MM) approach presented in a landmark work by Warshel and Levitt.<sup>106</sup> For an example of ML/MM, Inizan et al. recently applied ANI MLIPs to calculate chemically accurate solute–solute interactions, whereas solute–solvent and solvent–solvent interactions were simulated with AMOEBA.<sup>107</sup> In ML/MM, inference is affected by the electric field induced by the molecular mechanics region, and therefore, training beyond unperturbed gas-phase QM data is required in the absence of explicit long-range features. The challenge can compound when mutual polarization is considered; i.e., the MM components are described using a classical polarizable force field, such as AMOEBA.<sup>108</sup> The difficulty of simulating these systems may eventually be overcome by methods using only MLIPs with high-fidelity long-range electrostatics/polarization, i.e., ultimately replacing the MM region, need for embedding schemes, and complicated Hamiltonian formulations. A strategy to perform ML/MM simulations for highly polar systems with a mechanical embedding scheme is to train an MLIP model to infer atomic contributions to the multipole moments. As an example, reports by Poplier and co-workers have demonstrated the use of data-driven techniques,<sup>109–111</sup> particularly Gaussian process regression, to infer atomic electrostatic moments derived in the framework of the quantum theory of atoms-in-molecules.<sup>112</sup> The challenge of developing MLIPs with accurate polarization and electrostatic multipoles is ongoing, but a number of works are starting to address these issues.<sup>70,113–115</sup>

## 5. OUTLOOK AND CONCLUDING REMARKS

MLIPs are reaching a mature status, and their use in molecular simulations is becoming a frequented method in the computational chemist's toolbox. This Perspective highlighted specific models and methods that are expanding the application space of MLIPs beyond systems with structure/properties dominated by short-range physics and chemistry. We emphasized that the approximation of atomic interaction locality can constrain the variety of systems an MLIP can simulate. To overcome this, models that capture interactions and environment changes beyond an atom's immediate vicinity have been reported, indicating that MLIP-driven simulations for bulk, biological, and material systems are emerging areas.

Debatably, the bottleneck to increasing MLIP capabilities is the accumulation of reference data for training, instead of limitations resulting from the underlying model architecture. It is currently unclear the amount and diversity of system

sampling required to train MLIPs with reliable long-range interactions. The issue of limited data is persistent across ML fields, and active learning is one strategy that can address this challenge.<sup>116</sup> In the paradigm of active learning, redundant training practices are reduced by maximizing training data diversity and limiting data set size such that each sample meaningfully contributes to refining model parameters. For instance, an active learning strategy known as query by committee was used in producing the ANI-1x model,<sup>117</sup> where less than 25% of the original ANI-1 training data was needed to achieve equal accuracy.<sup>118</sup> Active learning strategies are key for building the next generation of MLIPs considering that complex models often require more training data. Moreover, MLIP accuracy and transferability are dependent on the composition of the reference data set. The development of methods that minimize the training size and maximize the value of each data point is essential to build MLIPs for practical use. Although not reported, there is envisioned value in systematic studies that establish “rule of thumb” design principals<sup>119</sup> between accuracy, MLIP architecture with long-range interactions, and active learning strategies.

The emergence of MLIPs with nonlocal interactions creates opportunities for insight into complex systems with exceptional accuracy. A shift in research focus from “improving accuracy on standard benchmarks” to performing simulations aimed at measurable scientific progress is timely and encouraged. The practice of judging MLIP models against each other will continue in all likelihood; however, we advocate for these future comparisons to occur on sets of standard benchmarks of experimental ground truths. Unfortunately, such experimental data sets either do not exist or are not yet widely embraced by MLIP developers, and efforts to curate these, particularly for noncovalent interactions, is a worthwhile pursuit. By our assessment, the number of different MLIP models greatly exceeds the cases in which they have been uniquely successful for understanding a chemical or materials science challenge. Most MLIPs have only been tested on a handful of systems in simple trial studies, and as a consequence, the area of applied MLIP modeling lags behind model development. Exploring new ML methods and developing complex model architectures are appreciable pursuits; however, it is worthwhile to interrogate such efforts for the value they provide beyond existing MLIPs. We emphasize that MLIPs are subject to an adage of classical simulations: a model should be no more or no less complex than what the application space demands. In the context of this Perspective, the complexity involved in treating long-range interactions should be judicious, for example, based on rational physical and chemical knowledge. MLIP models should be designed with treatments of electrostatics, polarization, dispersion, and many-body effects that coincide with relevant material or chemical phenomena that dictate the properties or structures of interest. On one hand, the treatment of long-range interactions in the design of an MLIP can be intuited by the molecular simulation practitioner: we have highlighted examples in Section 3.1 where it is possible to rationalize only using  $U_{\text{local,MLIP}}$  and explicit dispersion correction because the condensed-phase systems were neutral and chemically homogeneous. On the other hand, some tasks may require using a MLIP in a bulk phase application setting to understand its deficits and incorporate refinements. Such cases further support the importance of moving MLIP development beyond benchmarks and proof-of-concept studies.

Many reports now exist where nearsighted MLIPs have been applied to simulate large-scale systems ( $>10^7$  atoms), for example, Smith et al. for aluminum,<sup>120</sup> Guo et al. for copper,<sup>121</sup> and Lu et al. for water and copper.<sup>122</sup> With the advent of exascale computing,<sup>123</sup> it is expected that simulations for similar system sizes using MLIPs with long-range interactions will soon follow. MLIP-based simulations on this scale can yield unprecedented understanding into challenges faced across physics, chemistry, and engineering disciplines.<sup>124</sup> However, a remaining limitation to deploying MLIPs for such efforts is efficient integration with software capable of performing molecular simulations, particularly molecular dynamics and Monte Carlo methods. Examples of MLIP implementations into existing molecular simulation software are the DeePMD-kit,<sup>125</sup> AENET,<sup>126</sup> SNAP,<sup>28</sup> and TorchANI.<sup>127</sup> Despite these successes, there are a number of long-range MLIP models, such as those using message passing or Qeq, that are challenging to efficiently plug into existing software, and further development effort is needed. There is also appreciable optimization space available at the hardware and algorithm levels, for example, see the works of Guo et al.<sup>121</sup> and Galvelis et al.<sup>128</sup> Constructing MLIP interfaces with simulation packages that fully leverage the power of accelerated computing architectures is a nontrivial task. Unified effort between MLIP development, model implementation, and performance optimization is an ongoing need.

This year marks nearly a decade and a half of high-dimensional MLIP research, during which models have progressed to cover a span of short-ranged to nonlocal physics and chemistry. Even though broadly generalizable models for condensed matter with long-range interactions are in an infantile state, system/species specific MLIPs are positioned to revolutionize computational chemistry and materials science fields. Many challenging areas remain where molecular simulations using MLIPs with long-range interactions have yet to experience widespread success. Condensed-phase chemical reactions, interactions with light, and complicated catalytic events are a few areas in the chemical sciences that can benefit from MLIP model development. For materials science, magnetic materials, complex behavior of defects, and composite interfaces are modestly explored areas where MLIP models with long-range interactions can be transformative. The field of MLIPs has experienced rapid progression, and the application space continues to expand. Thus, we envision simulating systems with complex short- and long-range behavior using accurate MLIPs will be the basis of many future scientific breakthroughs.

## AUTHOR INFORMATION

### Corresponding Author

**Oleksandr Isayev** – Department of Chemistry, Mellon College of Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States;  [orcid.org/0000-0001-7581-8497](https://orcid.org/0000-0001-7581-8497); Email: [oleksandr@oleksandrisayev.com](mailto:oleksandr@oleksandrisayev.com)

### Author

**Dylan M. Anstine** – Department of Chemistry, Mellon College of Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.jpca.2c06778>

**Notes**

The authors declare no competing financial interest.

**Biographies**

**Dr. Dylan M. Anstine** is a postdoctoral research fellow in the Department of Chemistry at Carnegie Mellon University (CMU). He joined the research laboratory of Dr. Olexandr Isayev in November of 2021, where he actively works on developing machine learned interatomic potentials for crystal structure predictions, chemical reactions, and bulk systems. Prior to joining CMU, he received his Ph.D. in Materials Science and Engineering at the University of Florida under the advisement of Dr. Coray M. Colina. His Ph.D. research combined Monte Carlo and molecular dynamics simulation techniques to study adsorption-induced restructuring in polymers of intrinsic microporosity.



**Dr. Olexandr Isayev** is an Assistant Professor at the Department of Chemistry at Carnegie Mellon University. In 2008, Olexandr received his Ph.D. in computational chemistry. He was Postdoctoral Research Fellow at the Case Western Reserve University and a scientist at the government research lab. Before CMU, he was a faculty at UNC Eshelman School of Pharmacy, the University of North Carolina at Chapel Hill. Olexandr received the “Emerging Technology Award” from the American Chemical Society (ACS) and the GPU computing award from NVIDIA. The research in his lab focuses on connecting artificial intelligence (AI) and molecular modeling with chemical sciences.

**ACKNOWLEDGMENTS**

The work performed by D.A. and O.I. was made possible by the Office of Naval Research (ONR) through support provided by the Energetic Materials Program (MURI Grant No. N00014-21-1-2476). We acknowledge the Extreme Science and Engineering Discovery Environment (XSEDE) award CHE200122, which is supported by NSF Grant No. ACI-

1053575. This research is part of the Frontera computing project at the Texas Advanced Computing Center. Frontera is made possible by the National Science Foundation award OAC-1818253. The authors acknowledge colleagues and coauthors, especially Dr. Justin S. Smith, Dr. Roman Zubatyuk, Dr. Adrian Roitberg, and Dr. Segeri Tretiak, for their invaluable discussions and contributions to the development of NNPs.

**REFERENCES**

- (1) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine Learning for Molecular and Materials Science. *Nature* **2018**, *559* (7715), 547–555.
- (2) Kothe, D.; Lee, S.; Qualters, I. Exascale Computing in the United States. *Comput. Sci. Eng.* **2019**, *21* (1), 17–29.
- (3) Friederich, P.; Häse, F.; Proppe, J.; Aspuru-Guzik, A. Machine-Learned Potentials for next-Generation Matter Simulations. *Nat. Mater.* **2021**, *20* (6), 750–761.
- (4) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115* (19), 2315–2372.
- (5) Bartlett, R. J.; Musial, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79* (1), 291–352.
- (6) Gokcan, H.; Isayev, O. Learning Molecular Potentials with Neural Networks. *WIREs Comp. Mol. Sci.* **2022**, *12*, No. e1564.
- (7) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31* (46), 1902765.
- (8) Botu, V.; Batra, R.; Chapman, J.; Ramprasad, R. Machine Learning Force Fields: Construction, Validation, and Outlook. *J. Phys. Chem. C* **2017**, *121* (1), 511–522.
- (9) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121* (16), 10073–10141.
- (10) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights into Chemical Systems. *Chem. Rev.* **2021**, *121* (16), 9816–9872.
- (11) Glielmo, A.; Husic, B. E.; Rodriguez, A.; Clementi, C.; Noé, F.; Laio, A. Unsupervised Learning Methods for Molecular Simulation Data. *Chem. Rev.* **2021**, *121* (16), 9722–9758.
- (12) Manzhos, S.; Carrington, T. Neural Network Potential Energy Surfaces for Small Molecules and Reactions. *Chem. Rev.* **2021**, *121* (16), 10187–10217.
- (13) Balestrieri, R.; Pesenti, J.; LeCun, Y. Learning in High Dimension Always Amounts to Extrapolation. *arXiv preprint* 2021, DOI: [10.48550/arXiv.2110.09485](https://doi.org/10.48550/arXiv.2110.09485)
- (14) Behler, J.; Csányi, G. Machine Learning Potentials for Extended Systems: A Perspective. *Eur. Phys. J. B* **2021**, *94* (7), 142.
- (15) Hart, G. L. W.; Mueller, T.; Toher, C.; Curtarolo, S. Machine Learning for Alloys. *Nat. Rev. Mater.* **2021**, *6* (8), 730–755.
- (16) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comput. Mol. Sci.* **2011**, *1* (2), 211–228.
- (17) Režáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets 2: Hydrogen Bonding in an Extended Chemical Space. *J. Chem. Theory Comput.* **2020**, *16* (10), 6305–6316.
- (18) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-Range Charge-Transfer Excited States in Time-Dependent Density Functional Theory Require Non-Local Exchange. *J. Chem. Phys.* **2003**, *119* (6), 2943–2946.
- (19) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121* (16), 10037–10072.
- (20) Artrith, N.; Butler, K. T.; Coudert, F.-X.; Han, S.; Isayev, O.; Jain, A.; Walsh, A. Best Practices in Machine Learning for Chemistry. *Nat. Chem.* **2021**, *13* (6), 505–508.

- (21) Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C. Machine Learning for Molecular Simulation. *Annu. Rev. Phys. Chem.* **2020**, *71* (1), 361–390.
- (22) Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115* (16), 1032–1050.
- (23) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401.
- (24) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203.
- (25) Pozdnyakov, S. N.; Willatt, M. J.; Bartók, A. P.; Ortner, C.; Csányi, G.; Ceriotti, M. Incompleteness of Atomic Structure Representations. *Phys. Rev. Lett.* **2020**, *125* (16), 166001.
- (26) Willatt, M. J.; Musil, F.; Ceriotti, M. Atom-Density Representations for Machine Learning. *J. Chem. Phys.* **2019**, *150* (15), 154110.
- (27) Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134* (7), 074106.
- (28) Thompson, A. P.; Swiler, L. P.; Trott, C. R.; Foiles, S. M.; Tucker, G. J. Spectral Neighbor Analysis Method for Automated Generation of Quantum-Accurate Interatomic Potentials. *J. Comput. Phys.* **2015**, *285*, 316–330.
- (29) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87* (18), 184115.
- (30) Shapeev, A. v. Moment Tensor Potentials: A Class of Systematically Improvable Interatomic Potentials. *Multiscale Model. Simul.* **2016**, *14* (3), 1153–1173.
- (31) Yue, S.; Muniz, M. C.; Calegari Andrade, M. F.; Zhang, L.; Car, R.; Panagiotopoulos, A. Z. When Do Short-Range Atomistic Machine-Learning Models Fall Short? *J. Chem. Phys.* **2021**, *154* (3), 034111.
- (32) 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* (13), 136403.
- (33) Parsaeifard, B.; De, D. S.; Finkler, J. A.; Goedecker, S. Fingerprint-Based Detection of Non-Local Effects in the Electronic Structure of a Simple Single Component Covalent System. *Condens. Matter* **2021**, *6* (1), 9.
- (34) Park, H. B.; Kamcev, J.; Robeson, L. M.; Elimlech, M.; Freeman, B. D. Maximizing the Right Stuff: The Trade-off between Membrane Permeability and Selectivity. *Science* **2017**, *356* (6343), No. eaab0530.
- (35) Zhou, H.-X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018**, *118* (4), 1691–1741.
- (36) Lutz, J.-F.; Lehn, J.-M.; Meijer, E. W.; Matyjaszewski, K. From Precision Polymers to Complex Materials and Systems. *Nat. Rev. Mater.* **2016**, *1* (5), 16024.
- (37) Niblett, S. P.; Galib, M.; Limmer, D. T. Learning Intermolecular Forces at Liquid-Vapor Interfaces. *J. Chem. Phys.* **2021**, *155* (16), 164101.
- (38) Glick, Z. L.; Metcalf, D. P.; Koutsoukas, A.; Spronk, S. A.; Cheney, D. L.; Sherrill, C. D. AP-Net: An Atomic-Pairwise Neural Network for Smooth and Transferable Interaction Potentials. *J. Chem. Phys.* **2020**, *153* (4), 044112.
- (39) Kaplan, I. G. *Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials*; John Wiley & Sons, 2006.
- (40) Ambrosetti, A.; Ferri, N.; DiStasio, R. A.; Tkatchenko, A. Wavelike Charge Density Fluctuations and van Der Waals Interactions at the Nanoscale. *Science* **2016**, *351* (6278), 1171–1176.
- (41) Grisafi, A.; Ceriotti, M. Incorporating Long-Range Physics in Atomic-Scale Machine Learning. *J. Chem. Phys.* **2019**, *151* (20), 204105.
- (42) Li, W.; Ando, Y.; Minamitani, E.; Watanabe, S. Study of Li Atom Diffusion in Amorphous Li<sub>3</sub>PO<sub>4</sub> with Neural Network Potential. *J. Chem. Phys.* **2017**, *147* (21), 214106.
- (43) Artrith, N.; Kolpak, A. M. Understanding the Composition and Activity of Electrocatalytic Nanoalloys in Aqueous Solvents: A Combination of DFT and Accurate Neural Network Potentials. *Nano Lett.* **2014**, *14* (5), 2670–2676.
- (44) Li, Q.-J.; Küçükbenli, E.; Lam, S.; Khaykovich, B.; Kaxiras, E.; Li, J. Development of Robust Neural-Network Interatomic Potential for Molten Salt. *Cell Rep. Phys. Sci.* **2021**, *2* (3), 100359.
- (45) Artrith, N.; Urban, A. An Implementation of Artificial Neural-Network Potentials for Atomistic Materials Simulations: Performance for TiO<sub>2</sub>. *Comput. Mater. Sci.* **2016**, *114*, 135–150.
- (46) Hermann, J.; DiStasio, R. A.; Tkatchenko, A. First-Principles Models for van Der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications. *Chem. Rev.* **2017**, *117* (6), 4714–4758.
- (47) Ziegler, J. F.; Biersack, J. P. The Stopping and Range of Ions in Matter. In *Treatise on Heavy-Ion Science: Vol. 6: Astrophysics, Chemistry, and Condensed Matter*; Bromley, D. A., Ed.; Springer US: Boston, MA, 1985; pp 93–129. DOI: [10.1007/978-1-4615-8103-1\\_3](https://doi.org/10.1007/978-1-4615-8103-1_3).
- (48) Byggmästar, J.; Hamedani, A.; Nordlund, K.; Djurabekova, F. Machine-Learning Interatomic Potential for Radiation Damage and Defects in Tungsten. *Phys. Rev. B* **2019**, *100* (14), 144105.
- (49) Small, P. A. Some Factors Affecting the Solubility of Polymers. *J. Appl. Chem.* **1953**, *3* (2), 71–80.
- (50) Wagner, J. P.; Schreiner, P. R. London Dispersion in Molecular Chemistry—Reconsidering Steric Effects. *Angew. Chem.* **2015**, *54* (42), 12274–12296.
- (51) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comp. Mol. Sci.* **2011**, *1* (2), 211–228.
- (52) Perdew, J. P.; Yue, W. Accurate and Simple Density Functional for the Electronic Exchange Energy: Generalized Gradient Approximation. *Phys. Rev. B* **1986**, *33* (12), 8800–8802.
- (53) Morawietz, T.; Singraber, A.; Dellago, C.; Behler, J. How van Der Waals Interactions Determine the Unique Properties of Water. *Proc. Nat. Acad. Sci.* **2016**, *113* (30), 8368.
- (54) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H–Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.
- (55) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 Dispersion Coefficient Model. *J. Chem. Phys.* **2017**, *147* (3), 034112.
- (56) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-XTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671.
- (57) Tkatchenko, A.; DiStasio, R. A.; Car, R.; Scheffler, M. Accurate and Efficient Method for Many-Body van Der Waals Interactions. *Phys. Rev. Lett.* **2012**, *108* (23), 236402.
- (58) Axilrod, B. M.; Teller, E. Interaction of the van Der Waals Type Between Three Atoms. *J. Chem. Phys.* **1943**, *11* (6), 299–300.
- (59) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116* (9), 5105–5154.
- (60) Deringer, V. L.; Caro, M. A.; Csányi, G. A General-Purpose Machine-Learning Force Field for Bulk and Nanostructured Phosphorus. *Nat. Commun.* **2020**, *11* (1), 5461.
- (61) Wen, M.; Tadmor, E. B. Hybrid Neural Network Potential for Multilayer Graphene. *Phys. Rev. B* **2019**, *100* (19), 195419.
- (62) Muhli, H.; Chen, X.; Bartók, A. P.; Hernández-León, P.; Csányi, G.; Ala-Nissila, T.; Caro, M. A. Machine Learning Force Fields Based on Local Parametrization of Dispersion Interactions: Application to the Phase Diagram of C<sub>60</sub>. *Phys. Rev. B* **2021**, *104* (5), 54106.
- (63) Tkatchenko, A.; Scheffler, M. Accurate Molecular van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **2009**, *102* (7), 73005.

- (64) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.
- (65) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Chim. Acta* **1977**, *44* (2), 129–138.
- (66) Marenich, A. v.; Jerome, S. v.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8* (2), 527–541.
- (67) Verstraelen, T.; Pauwels, E.; de Proft, F.; van Speybroeck, V.; Geerlings, P.; Waroquier, M. Assessment of Atomic Charge Models for Gas-Phase Computations on Polypeptides. *J. Chem. Theory Comput.* **2012**, *8* (2), 661–676.
- (68) Deng, Z.; Chen, C.; Li, X.-G.; Ong, S. P. An Electrostatic Spectral Neighbor Analysis Potential for Lithium Nitride. *NPJ. Comput. Mater.* **2019**, *5* (1), 75.
- (69) Unke, O. T.; Meuwly, M. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.* **2019**, *15* (6), 3678–3693.
- (70) Sifain, A. E.; Lubbers, N.; Nebgen, B. T.; Smith, J. S.; Lokhov, A. Y.; Isayev, O.; Roitberg, A. E.; Barros, K.; Tretiak, S. Discovering a Transferable Charge Assignment Model Using Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9* (16), 4495–4501.
- (71) Morawietz, T.; Behler, J. A Density-Functional Theory-Based Neural Network Potential for Water Clusters Including van Der Waals Corrections. *J. Phys. Chem. A* **2013**, *117* (32), 7356–7366.
- (72) Artrith, N.; Morawietz, T.; Behler, J. High-Dimensional Neural-Network Potentials for Multicomponent Systems: Applications to Zinc Oxide. *Phys. Rev. B* **2011**, *83* (15), 153101.
- (73) Zhang, L.; Chen, M.; Wu, X.; Wang, H.; E, W.; Car, R. Deep Neural Network for the Dielectric Response of Insulators. *Phys. Rev. B* **2020**, *102* (4), 41121.
- (74) Zhang, L.; Wang, H.; Muniz, M. C.; Panagiotopoulos, A. Z.; Car, R.; E, W. A Deep Potential Model with Long-Range Electrostatic Interactions **2022**, *156* (12), 124107.
- (75) Marzari, N.; Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. Maximally Localized Wannier Functions: Theory and Applications. *Rev. Mod. Phys.* **2012**, *84* (4), 1419–1475.
- (76) Gao, A.; Remsing, R. C. Self-Consistent Determination of Long-Range Electrostatics in Neural Network Potentials. *Nat. Commun.* **2022**, *13* (1), 1572.
- (77) Xie, X.; Persson, K. A.; Small, D. W. Incorporating Electronic Information into Machine Learning Potential Energy Surfaces via Approaching the Ground-State Electronic Energy as a Function of Atom-Based Electronic Populations. *J. Chem. Theory Comput.* **2020**, *16* (7), 4256–4270.
- (78) Kaduk, B.; Kowalczyk, T.; van Voorhis, T. Constrained Density Functional Theory. *Chem. Rev.* **2012**, *112* (1), 321–370.
- (79) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv preprint 2014 DOI: 10.48550/arXiv.1412.6980*
- (80) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural Message Passing for Quantum Chemistry. In *Proceedings of the 34th International Conference on Machine Learning*; Precup, D.; Teh, Y. W., Eds.; Proceedings of Machine Learning Research, 2017; Vol. 70, pp 1263–1272.
- (81) Zubatyuk, R.; Smith, J. S.; Leszczynski, J.; Isayev, O. Accurate and Transferable Multitask Prediction of Chemical Properties with an Atoms-in-Molecules Neural Network. *Sci. Adv.* **2019**, *5* (8), No. eaav6490.
- (82) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-Chemical Insights from Deep Tensor Neural Networks. *Nat. Commun.* **2017**, *8* (1), 13890.
- (83) 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.
- (84) Unke, O. T.; Chmiela, S.; Gastegger, M.; Schütt, K. T.; Sauceda, H. E.; Müller, K.-R. SpookyNet: Learning Force Fields with Electronic Degrees of Freedom and Nonlocal Effects. *Nat. Commun.* **2021**, *12* (1), 7273.
- (85) Zubatyuk, R.; Smith, J. S.; Nebgen, B. T.; Tretiak, S.; Isayev, O. Teaching a Neural Network to Attach and Detach Electrons from Molecules. *Nat. Commun.* **2021**, *12* (1), 4870.
- (86) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105* (41), 9396–9409.
- (87) Rappe, A. K.; Goddard, W. A. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95* (8), 3358–3363.
- (88) Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-Equalization Method for the Calculation of Atomic Charges in Molecules. *J. Am. Chem. Soc.* **1986**, *108* (15), 4315–4320.
- (89) Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. A Generalization of the Charge Equilibration Method for Nonmetallic Materials. *J. Chem. Phys.* **2006**, *125* (9), 094108.
- (90) Verstraelen, T.; Ayers, P. W.; van Speybroeck, V.; Waroquier, M. ACKS2: Atom-Condensed Kohn-Sham DFT Approximated to Second Order. *J. Chem. Phys.* **2013**, *138* (7), 074108.
- (91) Wolf, D.; Kebinski, P.; Phillpot, S. R.; Eggebrecht, J. Exact Method for the Simulation of Coulombic Systems by Spherically Truncated, Pairwise R-1 Summation. *J. Chem. Phys.* **1999**, *110* (17), 8254–8282.
- (92) Ewald, P. P. Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253–287.
- (93) Yoo, P.; Sakano, M.; Desai, S.; Islam, M. M.; Liao, P.; Strachan, A. Neural Network Reactive Force Field for C, H, N, and O Systems. *NPJ. Comput. Mater.* **2021**, *7* (1), 9.
- (94) Novikov, I. S.; Shapeev, A. v. Improving Accuracy of Interatomic Potentials: More Physics or More Data? A Case Study of Silica. *Mater. Today Commun.* **2019**, *18*, 74–80.
- (95) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A Fourth-Generation High-Dimensional Neural Network Potential with Accurate Electrostatics Including Non-Local Charge Transfer. *Nat. Commun.* **2021**, *12* (1), 398.
- (96) Ghasemi, S. A.; Hofstetter, A.; Saha, S.; Goedecker, S. Interatomic Potentials for Ionic Systems with Density Functional Accuracy Based on Charge Densities Obtained by a Neural Network. *Phys. Rev. B* **2015**, *92* (4), 45131.
- (97) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. General-Purpose Machine Learning Potentials Capturing Nonlocal Charge Transfer. *Acc. Chem. Res.* **2021**, *54* (4), 808–817.
- (98) Stone, A. *The Theory of Intermolecular Forces*; oUP oxford, 2013.
- (99) Williams, D. E. Representation of the Molecular Electrostatic Potential by Atomic Multipole and Bond Dipole Models. *J. Comput. Chem.* **1988**, *9* (7), 745–763.
- (100) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114* (8), 2549–2564.
- (101) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933–5947.
- (102) Thole, B. T. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59* (3), 341–350.
- (103) Lier, B.; Poliak, P.; Marquetand, P.; Westermayr, J.; Oostenbrink, C. BuRNN: Buffer Region Neural Network Approach for Polarizable-Embedding Neural Network/Molecular Mechanics Simulations. *J. Phys. Chem. Lett.* **2022**, *13* (17), 3812–3818.
- (104) Bösel, L.; Thürlemann, M.; Riniker, S. Machine Learning in QM/MM Molecular Dynamics Simulations of Condensed-Phase Systems. *J. Chem. Theory Comput.* **2021**, *17* (5), 2641–2658.
- (105) Zeng, J.; Giese, T. J.; Ekesan, S.; York, D. M. Development of Range-Corrected Deep Learning Potentials for Fast, Accurate Quantum Mechanical/Molecular Mechanical Simulations of Chemical Reactions in Solution. *J. Chem. Theory Comput.* **2021**, *17* (11), 6993–7009.

- (106) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249.
- (107) Inizan, T. J.; Plé, T.; Adoua, O.; Ren, P.; Gökcen, H.; Isayev, O.; Lagardère, L.; Piquemal, J.-P. Scalable Hybrid Deep Neural Networks/Polarizable Potentials Biomolecular Simulations Including Long-Range Effects. *arXiv preprint* 2022 . DOI: 10.48550/arXiv.2207.14276
- (108) Loco, D.; Lagardère, L.; Caprasecca, S.; Lippurini, F.; Mennucci, B.; Piquemal, J.-P. Hybrid QM/MM Molecular Dynamics with AMOEBA Polarizable Embedding. *J. Chem. Theory Comput.* **2017**, *13* (9), 4025–4033.
- (109) Mills, M. J. L.; Popelier, P. L. A. Intramolecular Polarisable Multipolar Electrostatics from the Machine Learning Method Kriging. *Comput. Theor. Chem.* **2011**, *975* (1), 42–51.
- (110) Fletcher, T. L.; Popelier, P. L. A. Multipolar Electrostatic Energy Prediction for All 20 Natural Amino Acids Using Kriging Machine Learning. *J. Chem. Theory Comput.* **2016**, *12* (6), 2742–2751.
- (111) Mills, M. J. L.; Popelier, P. L. A. Polarisable Multipolar Electrostatics from the Machine Learning Method Kriging: An Application to Alanine. *Theor. Chem. Acc.* **2012**, *131* (3), 1137.
- (112) Bader, R. F. W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18* (1), 9–15.
- (113) Veit, M.; Wilkins, D. M.; Yang, Y.; DiStasio, R. A.; Ceriotti, M. Predicting Molecular Dipole Moments by Combining Atomic Partial Charges and Atomic Dipoles. *J. Chem. Phys.* **2020**, *153* (2), 024113.
- (114) Schütt, K.; Unke, O.; Gastegger, M. Equivariant Message Passing for the Prediction of Tensorial Properties and Molecular Spectra. In *Proceedings of the 38th International Conference on Machine Learning*; Meila, M.; Zhang, T., Eds.; Proceedings of Machine Learning Research, 2021; Vol. 139, pp 9377–9388.
- (115) Thürlemann, M.; Bösel, L.; Riniker, S. Learning Atomic Multipoles: Prediction of the Electrostatic Potential with Equivariant Graph Neural Networks. *J. Chem. Theory Comput.* **2022**, *18* (3), 1701–1710.
- (116) Seung, H. S.; Opper, M.; Sompolinsky, H. Query by Committee. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*; COLT '92; Association for Computing Machinery: New York, NY, 1992; pp 287–294. DOI: 10.1145/130385.130417.
- (117) Smith, J. S.; Zubatyuk, R.; Nebgen, B.; Lubbers, N.; Barros, K.; Roitberg, A. E.; Isayev, O.; Tretiak, S. The ANI-1ccx and ANI-1x Data Sets, Coupled-Cluster and Density Functional Theory Properties for Molecules. *Sci. Data.* **2020**, *7* (1), 134.
- (118) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less Is More: Sampling Chemical Space with Active Learning. *J. Chem. Phys.* **2018**, *148* (24), 241733.
- (119) Artrith, N.; Butler, K. T.; Coudert, F.-X.; Han, S.; Isayev, O.; Jain, A.; Walsh, A. Best Practices in Machine Learning for Chemistry. *Nat. Chem.* **2021**, *13* (6), 505–508.
- (120) Smith, J. S.; Nebgen, B.; Mathew, N.; Chen, J.; Lubbers, N.; Burakovsky, L.; Tretiak, S.; Nam, H. A.; Germann, T.; Fensin, S.; et al. Automated Discovery of a Robust Interatomic Potential for Aluminum. *Nat. Commun.* **2021**, *12* (1), 1257.
- (121) Guo, Z.; Lu, D.; Yan, Y.; Hu, S.; Liu, R.; Tan, G.; Sun, N.; Jiang, W.; Liu, L.; Chen, Y.; et al. Extending the Limit of Molecular Dynamics with Ab Initio Accuracy to 10 Billion Atoms. *arXiv e-prints* 2022, DOI: 10.48550/arXiv.2201.01446
- (122) Lu, D.; Wang, H.; Chen, M.; Lin, L.; Car, R.; E, W.; Jia, W.; Zhang, L. 86 PFLOPS Deep Potential Molecular Dynamics Simulation of 100 Million Atoms with Ab Initio Accuracy. *Comput. Phys. Commun.* **2021**, *259*, 107624.
- (123) Mniszewski, S. M.; Belak, J.; Fattebert, J.-L.; Negre, C. F. A.; Slattery, S. R.; Adedoyin, A. A.; Bird, R. F.; Chang, C.; Chen, G.; Ethier, S.; et al. Enabling Particle Applications for Exascale Computing Platforms. *Int. J. High Perform. Comput. Appl.* **2021**, *35* (6), 572–597.
- (124) Ramprasad, R.; Batra, R.; Pilania, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine Learning in Materials Informatics: Recent Applications and Prospects. *NPJ. Comput. Mater.* **2017**, *3* (1), 54.
- (125) Wang, H.; Zhang, L.; Han, J.; E, W. DeePMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.
- (126) Chen, M. S.; Morawietz, T.; Mori, H.; Markland, T. E.; Artrith, N. AENET-LAMMPS and AENET-TINKER: Interfaces for Accurate and Efficient Molecular Dynamics Simulations with Machine Learning Potentials. *J. Chem. Phys.* **2021**, *155* (7), 074801.
- (127) Gao, X.; Ramezanghorbani, F.; Isayev, O.; Smith, J. S.; Roitberg, A. E. TorchANI: A Free and Open Source PyTorch-Based Deep Learning Implementation of the ANI Neural Network Potentials. *J. Chem. Inf. Model.* **2020**, *60* (7), 3408–3415.
- (128) Galvelis, R.; Varela-Rial, A.; Doerr, S.; Fino, R.; Eastman, P.; Markland, T. E.; Chodera, J. D.; Fabritiis, G. de. NNP/MM: Fast Molecular Dynamics Simulations with Machine Learning Potentials and Molecular Mechanics. *arXiv preprint* 2022, DOI: 10.48550/arXiv.2201.08110

## □ Recommended by ACS

### Transferable Force Fields from Experimental Scattering Data with Machine Learning Assisted Structure Refinement

Brennon L. Shanks, Michael P. Hoepfner, et al.

DECEMBER 05, 2022

THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS

READ ▶

### Decomposing Chemical Space: Applications to the Machine Learning of Atomic Energies

Frederik Ø. Kjeldal and Janus J. Eriksen

MARCH 16, 2023

JOURNAL OF CHEMICAL THEORY AND COMPUTATION

READ ▶

### Dielectric Saturation in Water from a Long-Range Machine Learning Model

Harender S. Dhatarwal, Richard C. Remsing, et al.

APRIL 14, 2023

THE JOURNAL OF PHYSICAL CHEMISTRY B

READ ▶

### Δ-Machine Learned Potential Energy Surfaces and Force Fields

Joel M. Bowman, Qi Yu, et al.

DECEMBER 17, 2022

JOURNAL OF CHEMICAL THEORY AND COMPUTATION

READ ▶

Get More Suggestions >