

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/348888499>

General-Purpose Machine Learning Potentials Capturing Nonlocal Charge Transfer

Article in Accounts of Chemical Research · January 2021

DOI: 10.1021/acs.accounts.0c00689

CITATIONS

23

READS

265

4 authors, including:



Ko Tsz Wai

Georg-August-Universität Göttingen

5 PUBLICATIONS 93 CITATIONS

[SEE PROFILE](#)

General-Purpose Machine Learning Potentials Capturing Nonlocal Charge Transfer

Published as part of the Accounts of Chemical Research special issue "Data Science Meets Chemistry".

Tsz Wai Ko, Jonas A. Finkler, Stefan Goedecker,* and Jörg Behler*



Cite This: <https://dx.doi.org/10.1021/acs.accounts.0c00689>



Read Online

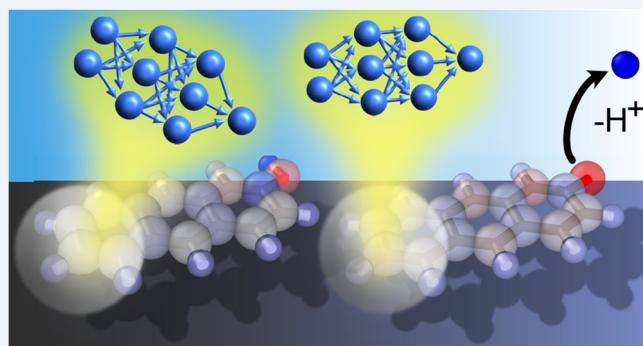
ACCESS |

Metrics & More

Article Recommendations

CONSPECTUS: The development of first-principles-quality machine learning potentials (MLP) has seen tremendous progress, now enabling computer simulations of complex systems for which sufficiently accurate interatomic potentials have not been available. These advances and the increasing use of MLPs for more and more diverse systems gave rise to new questions regarding their applicability and limitations, which has constantly driven new developments. The resulting MLPs can be classified into several generations depending on the types of systems they are able to describe. First-generation MLPs, as introduced 25 years ago, have been applicable to low-dimensional systems such as small molecules. MLPs became a practical tool for complex systems in chemistry and materials science with the introduction of high-dimensional neural network potentials (HDNNP) in 2007, which represented the first MLP of the second generation. Second-generation MLPs are based on the concept of locality and express the total energy as a sum of environment-dependent atomic energies, which allows applications to very large systems containing thousands of atoms with linearly scaling computational costs. Since second-generation MLPs do not consider interactions beyond the local chemical environments, a natural extension has been the inclusion of long-range interactions without truncation, mainly electrostatics, employing environment-dependent charges establishing the third MLP generation. A variety of second- and, to some extent, also third-generation MLPs are currently the standard methods in ML-based atomistic simulations.

In spite of countless successful applications, in recent years it has been recognized that the accuracy of MLPs relying on local atomic energies and charges is still insufficient for systems with long-ranged dependencies in the electronic structure. These can, for instance, result from nonlocal charge transfer or ionization and are omnipresent in many important types of systems and chemical processes such as the protonation and deprotonation of organic and biomolecules, redox reactions, and defects and doping in materials. In all of these situations, small local modifications can change the system globally, resulting in different equilibrium structures, charge distributions, and reactivity. These phenomena cannot be captured by second- and third-generation MLPs. Consequently, the inclusion of nonlocal phenomena has been identified as a next key step in the development of a new fourth generation of MLPs. While a first fourth-generation MLP, the charge equilibration neural network technique (CENT), was introduced in 2015, only very recently have a range of new general-purpose methods applicable to a broad range of physical scenarios emerged. In this Account, we show how fourth-generation HDNNPs can be obtained by combining the concepts of CENT and second-generation HDNNPs. These new MLPs allow for a highly accurate description of systems where nonlocal charge transfer is important.



KEY REFERENCES

- Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.¹ *The first machine learning potential applicable to a large number of atoms is introduced by expressing the energy as a sum of environment-dependent atomic energies provided by neural networks. Symmetry functions ensure the exact translational, rotational, and permutational invariance.*

Received: October 25, 2020

- 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: Condens. Matter Mater. Phys.* **2015**, *92*, 045131.² A charge equilibration neural network technique (CENT) is used to derive a machine learning potential for systems including long-range charge transfer and ionization. The charge-density-based method is well-suited for ionic materials.
- 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*, 398.³ A new generation of high-dimensional neural network potentials combining the advantages of the Behler–Parrinello approach and CENT is introduced. The method is very accurate and generally applicable to systems with nonlocal charge transfer and varying global charges.

1. INTRODUCTION

The ultimate goal of computer simulations in chemistry, molecular biology, physics, and materials science is to reach predictive power. Whether this goal can be achieved depends on several decisions that have to be made regarding the accuracy of the description of the atomic interactions, the structural complexity of the employed model systems, and the completeness of configurational sampling. Any possible setup necessarily represents a compromise because accurate electronic structure calculations are computationally too demanding for extended simulations of large systems, while realistic structural models can be studied only by more approximate potentials. Consequently, for decades the development of atomistic potentials, which are both efficient and reliable, has been a central topic of research with the goal of finding improved representations of the potential energy surface (PES) providing access to the energy and forces for given atomic configurations.

In recent years, a paradigm change has taken place in that the main approach to the construction of PESs shifted from physically inspired potentials to very general functional forms aiming to represent reference data sets obtained in electronic structure calculations as accurately as possible. Nowadays, machine learning potentials (MLP)⁴ have become a major approach for the construction of first-principles based potentials and many different methods have been proposed to date.^{1,5–13}

In general, most MLPs can be classified into different generations depending on the types of systems they can be applied to (s. Figure 1). Starting with the work of Doren and co-workers about 25 years ago,¹⁴ MLPs of the first generation have been introduced. These MLPs make use of neural networks and have in common that they are applicable to low-dimensional systems explicitly depending only on the degrees of freedom of a few atoms. While groundbreaking contributions were made in the early years, for about a decade MLPs remained a niche methodology because of the restrictions on the size of the systems.

The breakthrough in modern MLPs has been achieved by the introduction of second-generation MLPs, which enable the study of large systems containing thousands of atoms by expressing the total energy as a sum of local environment-dependent atomic energies. Since the introduction of high-dimensional neural network potentials (HDNNPs) by Behler and Parrinello in 2007, which represented the first example of a

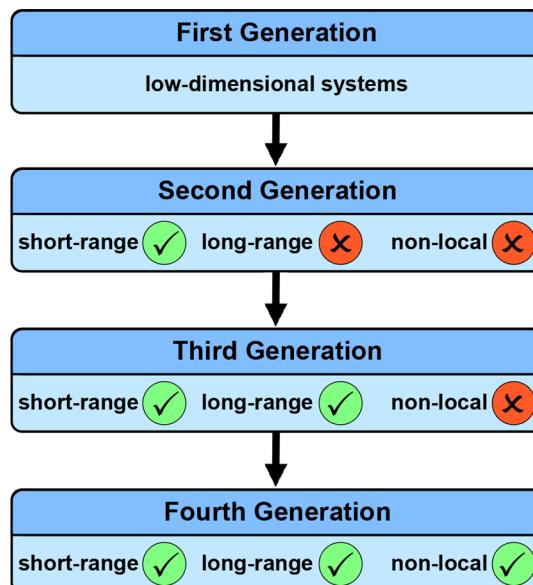


Figure 1. Four generations of machine learning potentials.

second-generation MLP, many powerful methods have been proposed^{5–7} and applied with great success.

It is well known that for many systems long-range interactions beyond the local atomic chemical environments are also important, mainly electrostatics but also dispersion interactions, and neither is considered outside the local atomic environments in second-generation MLPs. This limitation gave rise to the development of third-generation HDNNPs, in which long-range electrostatic interactions based on environment-dependent charges are explicitly included using Coulomb's law without truncation.^{15,16} Still, because the charges are treated as local properties, third-generation MLPs are unable to describe systems with strong nonlocal dependencies in the electronic structure. These can arise from changes in the total charge of the system but also from local chemical or conformational changes, which modify the electronic structure over large distances or even globally. These phenomena are very common in chemistry, molecular biology, and materials science, and in Figure 2 we give two typical examples from different fields.

Figure 2a shows the Hirshfeld charges¹⁷ of the aromatic organic molecule 2-hydroxy-anthracene obtained from density-functional theory (DFT). If this molecule is deprotonated, then the atomic charges notably change throughout the whole molecule, as can be seen for carbon atoms 6 and 7, whose partial charges change strongly although they are about 8.5 Å away from the deprotonated oxygen atom. Figure 2b shows the optimized adsorption geometries of a gold atom on a periodic MgO(001) surface. For the ideal surface, the gold atom is adsorbed on top of an oxygen atom and its Hirshfeld charge is very close to zero. If a Mg vacancy is introduced into the bottom layer of the slab outside the local environment, then the preferred adsorption site does not change but the gold atom becomes positively charged. If, on the other hand, an oxygen atom in the bottom layer is replaced by a fluorine atom, then the adsorbed Au atom has a very large negative charge and the adsorption site changes to a Mg atom. Such control of the electronic structure by doping and defects is very important for applications such as heterogeneous catalysis and semiconductor devices. All of these examples have in common that modifications in one part of the system strongly change the properties in other parts, with separations between

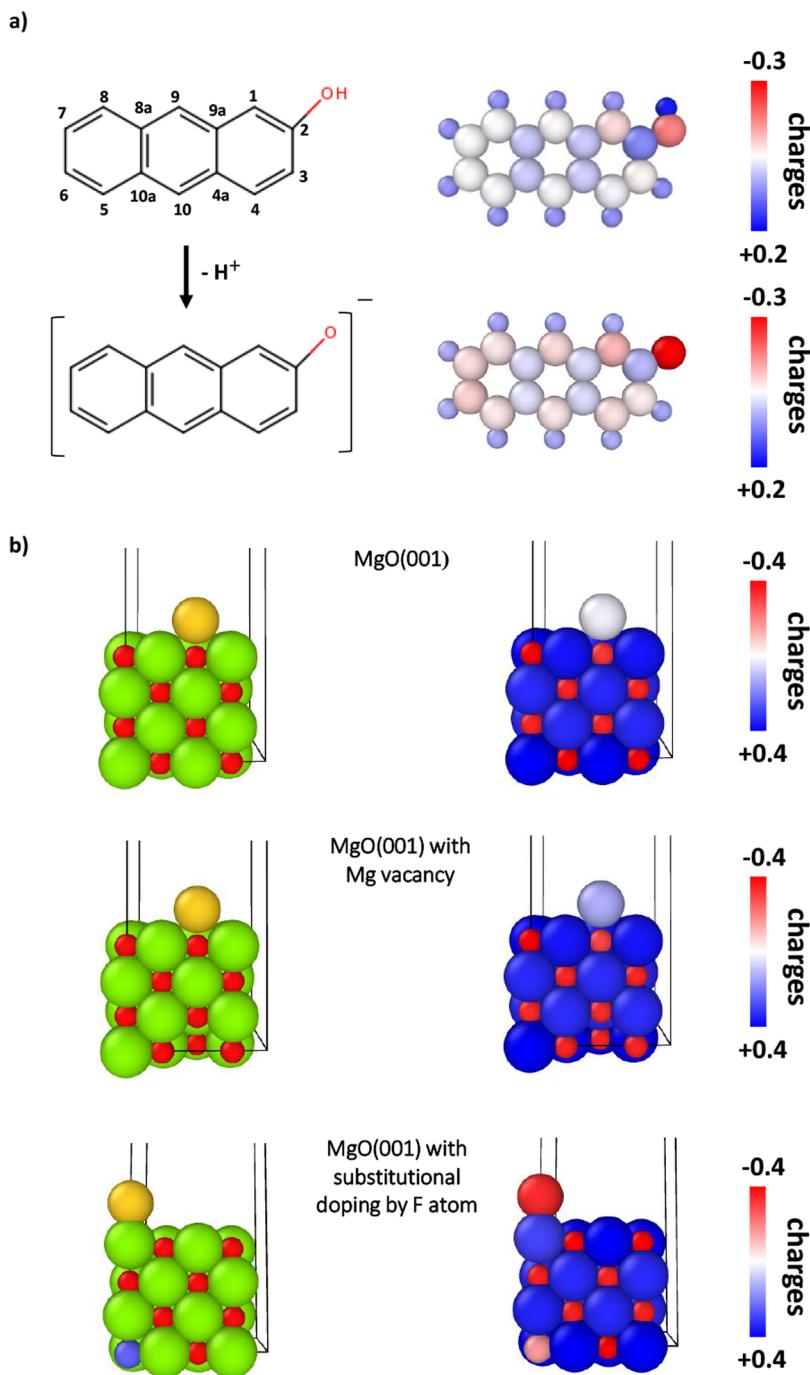


Figure 2. Examples for systems exhibiting long-range charge transfer. The structures and the corresponding atomic charges are displayed on the left- and right-hand sides, respectively. Panel (a) shows the organic molecule 2-hydroxy-anthracene together with its deprotonated form. Panel (b) shows a gold atom adsorbed on the ideal MgO(001) surface, on a surface with a Mg vacancy in a subsurface layer, and on a surface with substitutional doping by a fluorine atom. Oxygen, fluorine, magnesium, and gold atoms are colored red, blue, green, and yellow, respectively.

these interacting parts being larger than typical radii of local atomic environments considered in second- and third-generation MLPs, which are thus unable to describe these effects.

MLPs including these nonlocal effects belong to the just emerging fourth generation, which is the topic of this Account. The first fourth-generation MLP has been the charge equilibration neural network technique (CENT)¹⁸ proposed by Goedecker and Ghasemi, which employs a charge-density-based energy expression intended for ionic materials. Very recently, the development of fourth-generation MLPs has been

pursued by several groups employing different methodologies, an example being fourth-generation (4G) HDNNPs,³ which combine the advantages of the CENT approach with the numerical accuracy of HDNNPs.

2. LOCAL MACHINE LEARNING POTENTIALS

2.1. Methodology

Most MLPs currently used in atomistic simulations can be classified as local in that the total energy depends on the atomic energies or charges, which are functions of the local chemical

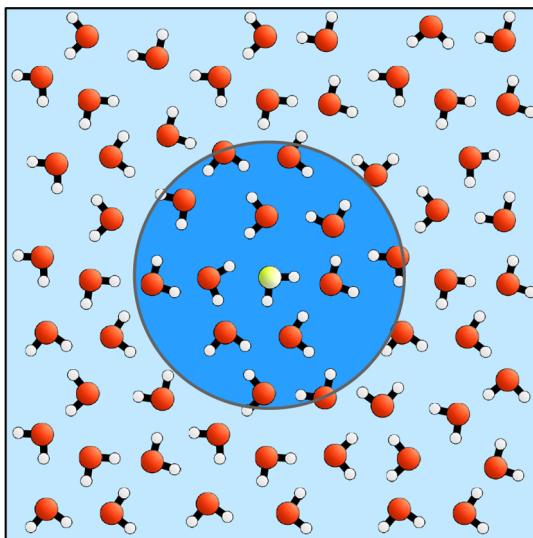


Figure 3. Local atomic environment of an oxygen atom (yellow) in water.

environments of the atoms up to a cutoff radius R_c (Figure 3). Compared to conventional empirical potentials, this cutoff is rather large, with typical values on the order of 6 Å. Consequently, the atomic energies and charges depend on the positions of a large number of neighboring atoms, which is the reason for the high accuracy that can be achieved for many systems. We note that local MLPs can also contain long-range electrostatic interactions without any truncation, but these long-range interactions do not take global dependencies in the charge distribution into account, as we will see below.

Second-generation MLPs are the most frequently used type of local MLPs, in which the total energy E_{tot} is constructed as the sum of environment-dependent atomic energies E_i ,

$$E_{\text{tot}} = \sum_{i=1}^{N_{\text{atom}}} E_i \quad (1)$$

while all interactions with atoms outside the cutoff spheres are neglected. The parameters of the MLPs are determined by minimizing the errors in reference energies and forces obtained from electronic structure calculations.

Nowadays, many second-generation MLPs are available, such as the high-dimensional neural network potentials (HDNNP) of the form proposed in 2007,¹ Gaussian approximation potentials,⁵ moment tensor potentials,⁶ and spectral neighbor analysis potentials.⁷ These methods differ in the machine learning algorithms and in the functions employed to describe the positions of the neighboring atoms inside the cutoff sphere while the underlying ansatz is essentially the same. Second-generation MLPs are very popular because accurate PESs can be obtained even for complex bonding environments, which are a substantial challenge for conventional atomistic potentials.

The structure of a second-generation HDNNP is shown schematically in Figure 4. Starting from the Cartesian coordinate vectors, \mathbf{R}_i , of the atoms, first a transformation to a suitable descriptor vector is performed, which is a local structural fingerprint of the atomic geometric environment and possesses the mandatory translational, rotational, and permutation invariance. In the case of HDNNPs, atom-centered symmetry functions (ACSFs)¹⁹ are most frequently used for this purpose. For each atom i , ACSF vector \mathbf{G}_i^A is then used as input for an

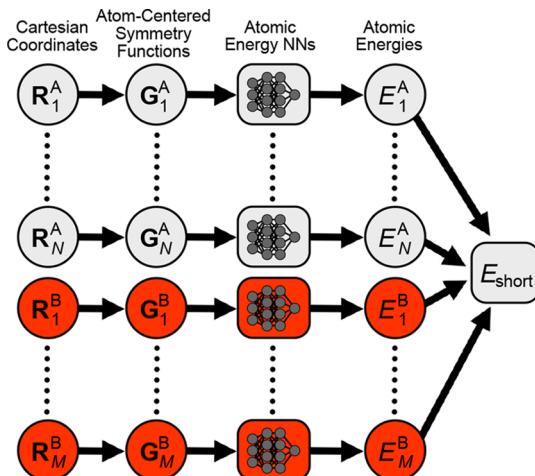


Figure 4. Structure of a second-generation HDNNP for binary system $A_N B_M$.¹

atomic neural network with element-specific architecture and weight parameters yielding the respective E_i . The total energy is then the sum of the outputs of all atomic neural networks.

In recent years, long-range interactions such as electrostatics and dispersion, which are standard components of conventional classical force fields, have attracted attention in the field of MLPs since an accurate description of long-range interactions is very important for atomistic simulations of many systems.^{20–22} For this reason, many MLPs have been extended to include long-range interactions, sometimes employing fixed element-specific charges²³ but mainly using more flexible environment-dependent charges.^{9,15,24–27} Such MLPs including long-range electrostatic interactions based on atomic charges Q_i expressed as a function of the local chemical environment define the third generation of MLPs. Here, the total energy is the sum of a short-range part consisting of environment-dependent atomic energies and a long-range part based on environment-dependent charges and the resulting electrostatic energy,

$$E_{\text{tot}} = E_{\text{short}} + E_{\text{elec}} = \sum_{i=1}^{N_{\text{atom}}} E_i + E_{\text{elec}}(\{Q_i\}) \quad (2)$$

Also, dispersion interactions depending on environment-dependent parameters can be included in third-generation MLPs. Since atomic partial charges are not quantum mechanical observables, there are different recipes for how to split up the total charge distribution obtained in electronic structure calculations into atomic charges, such as the Hirshfeld¹⁷ and Becke²⁸ schemes, which can be used as references for training the charges in third-generation MLPs.

Because in principle the short-range part can also describe electrostatic interactions within the cutoff radii, double counting of energy terms can be avoided by first training the charges. On the basis of these charges, the electrostatic energies and forces are removed from the reference values and the remaining contributions are expressed by the short-range part. As an example, the structure of a third-generation HDNNP¹⁵ is shown in Figure 5. As in second-generation HDNNPs, the short-range energy consists of a sum of environment-dependent energies provided by individual neural networks, while a second set of atomic neural networks is employed to derive environment-dependent charges used to compute the electrostatic energy. For more details about the construction and properties of second-

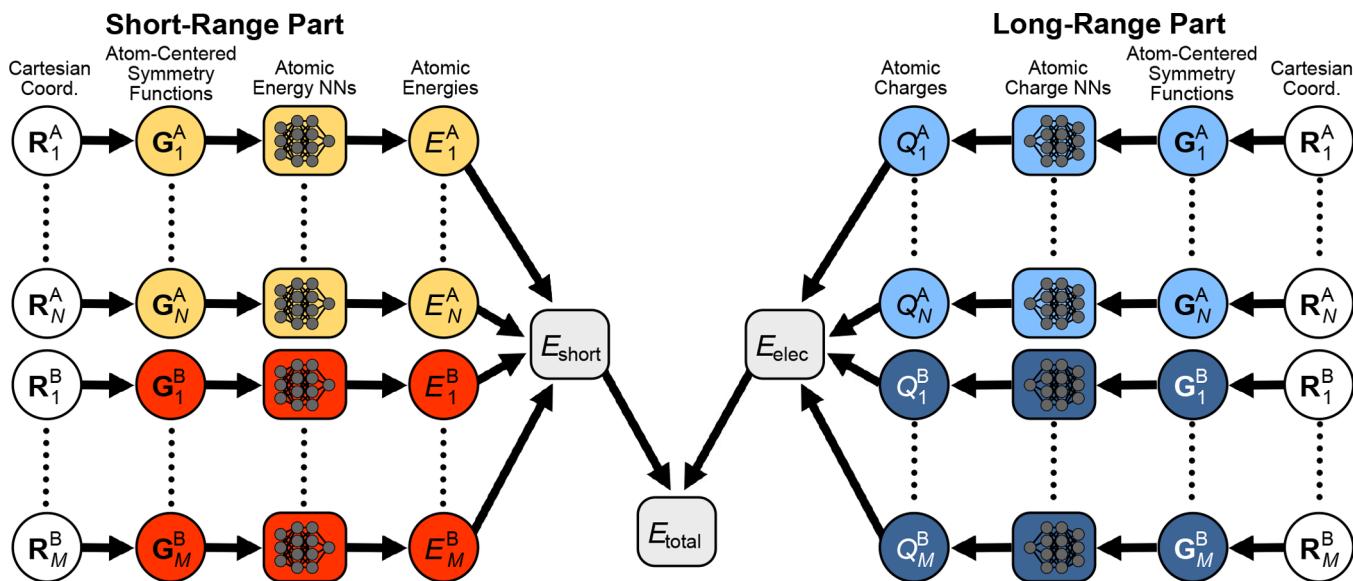


Figure 5. Structure of a third-generation HDNNP for binary system A_NB_M .¹⁶

and third-generation HDNNPs, the interested reader is referred to several reviews.^{29–31}

We finally note that recently it has been shown that long-range electrostatic effects can be also captured by introducing a nonlocal long-distance equivariant (LODE) representation,³² which is dependent on a Coulomb-type atomic density potential.³³

Potentials of the second and sometimes third generation are the current workhorse methods in atomistic simulations employing MLPs, and their accuracy has been demonstrated for numerous problems in chemistry, physics, and materials science, showing that the locality assumption holds well in many systems, such as in solids including metals, semiconductors, insulators, and liquids with an efficient screening of long-range interactions.

2.2. Limitations of Local Machine Learning Potentials

In spite of the success of local MLPs, there are numerous systems exhibiting strong nonlocal effects. Recently, Deringer and Csányi have suggested a statistical analysis to determine good cutoff radii based on the convergence of the forces with respect to this radius.³⁴ They found that with the standard cutoff radii it is not possible to reach very high accuracy for sp^2 -hybridized carbon systems. Furthermore, Parsaeifard, Finkler, and Goedecker³⁵ have demonstrated on a large data set of C_{60} structures that local quantities, such as atomic charges, energies, and atom projected densities of states, can differ significantly for atoms that share the same local structure but have different far-away environments. This indicates nonlocal effects.

We will now demonstrate the limitations of current local MLPs using the example of a third-generation HDNNP and the molecule shown in Figure 6. This molecule, which is discussed in more detail in ref 3, consists of a linear chain of 10 sp^2 -hybridized carbon atoms saturated by two hydrogen atoms. The electronic charge density of this system is strongly modified by protonating one of the terminal carbon atoms, as can be seen in the changes of the atomic Hirshfeld charges obtained by DFT. In these calculations, the molecular structure has been kept fixed apart from the positions of the additional proton and its neighboring hydrogen atom, which have been relaxed. Thus, the atomic environment of the terminal carbon atom at the opposite end

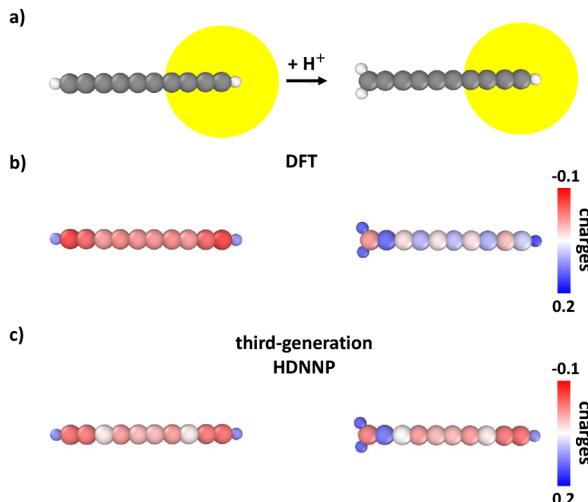


Figure 6. Illustration of the limitations of third-generation HDNNPs in the description of atomic charges:³ (a) DFT-optimized geometry of $C_{10}H_2$ and structure of the protonated molecule in which only the hydrogen positions at the left end have been relaxed. Hydrogen and carbon atoms are colored gray and white, respectively. The yellow circle shows the cutoff sphere of the right carbon atom defining its atomic environment, which is identical in both molecules. Panel (b) shows the Hirshfeld charges obtained from DFT, which are very different for the highlighted carbon atom. The atomic charges predicted by the third-generation HDNNP displayed in (c) are incorrectly the same.

remains completely identical. As a consequence, the charge of this atom predicted by a third-generation HDNNP is the same in both molecules and corresponds roughly to the average of the DFT charges in both systems because in the training process the HDNNP tries to find the best compromise. The atomic energies and forces of second-generation HDNNPs exhibit the same behavior and are the same for both molecules, although the electronic structures are different.

In addition to the inability to describe long-range charge transfer, this example shows another limitation of local MLPs, which is the restriction to a single global charge state because no information about the total charge of the system is included. It is important to realize that although in the present case we have

used HDNNPs as an example method, these limitations are general and equally apply to all types of second- and third-generation MLPs.

3. NONLOCAL MACHINE LEARNING POTENTIALS

The limitations of local MLPs for systems with significant nonlocal charge transfer are now well understood, and the focus has shifted to the development of a fourth generation of potentials applicable to these systems. Starting with the charge equilibration neural network technique¹⁸ introduced in 2015, an increasing number of new methods are just emerging. Fourth-generation (4G) HDNNPs³ combine the strengths of CENT and second-generation HDNNPs in describing long-range and short-range interactions, respectively, and are therefore applicable to a broad range of systems. Both CENT and 4G-HDNNPs are built on charge equilibration techniques, which have also been used by other types of potentials. In addition, a series of other fourth-generation methods with similar capabilities are now available.

3.1. Charge Equilibration Schemes

The empirical description of charge transfer is based on the concept of electronegativity. Computational methods based on this principle have been developed in the form of charge equilibration schemes.³⁶ In these schemes, one first considers how the energy of an isolated atom changes as a function of its charge. This function is then approximated by a second-order Taylor expansion. The coefficients of the linear term can be interpreted as generalized electronegativities whereas the coefficients of the quadratic terms represent the so-called hardness. The equilibrated charge distribution is the one that minimizes the total energy of the system. Rappe and Goddard³⁶ have provided values for the electronegativities and hardness terms based on the experimental ionization potentials, electron affinities, and covalent radii. For molecules containing light elements, they could show that the resulting charge equilibration scheme correctly reproduces experimental trends.

In combination with classical empirical terms such as bond stretching and bending, charge equilibration schemes form an integral part in a series of modern force fields such as COMB³⁷ and ReaxFF.³⁸ Here, the electronegativities do not depend only on the chemical elements but also on the atomic environments via simple analytical functions.

3.2. CENT Scheme

As we have seen that atomic charges often show a nonlocal dependence, the central quantity of the CENT potential^{2,39} is the element- and environment-dependent electronegativity χ_i , which is assumed to be a local property. Since it is very difficult to derive analytical functions for its environment dependence empirically, in the CENT method a neural network is used to find the best functional form employing ACSFs as structural fingerprints (Figure 7).

CENT was initially developed for ionic materials, and it turned out that no additional terms are needed for an accurate prediction of the total energy of these systems, which is given by

$$E_{\text{tot}}(\{Q_i\}) = \sum_{i=1}^{N_{\text{atom}}} \left(E_i^0 + \chi_i Q_i + \frac{1}{2} J_{ii} Q_i^2 \right) + \frac{1}{2} \iint \frac{\rho(\mathbf{R}) \rho(\mathbf{R}')}{|\mathbf{R} - \mathbf{R}'|} d\mathbf{R} d\mathbf{R}' \quad (3)$$

The first two terms describe the energetics of the individual atoms with atomic charge Q_i . J_{ii} is the element-dependent

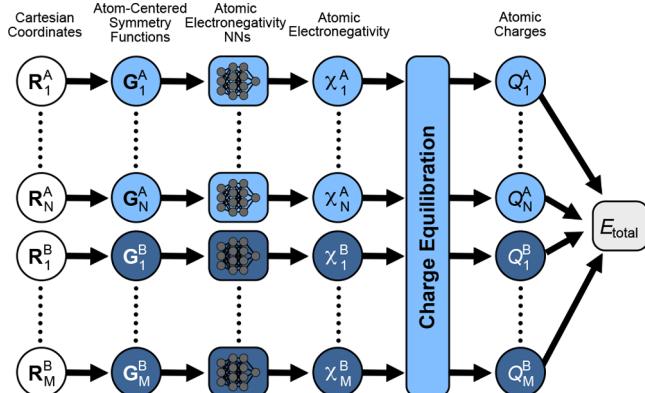


Figure 7. Charge equilibration neural network technique (CENT) for binary system $A_N B_M$.

hardness term, and the last term in the energy expression describes the long-range electrostatic interactions. In the case of periodic boundary conditions, the pairwise Coulomb interaction is replaced by an Ewald summation.⁴⁰ The charge density $\rho(\mathbf{R})$ is taken to be a superposition of spherically symmetric atomic Gaussian densities ρ_i centered at the atomic positions,

$$\rho_i(\mathbf{R}) = \frac{Q_i}{\alpha_i^3 \pi^{3/2}} \exp\left(-\frac{|\mathbf{R} - \mathbf{R}_i|^2}{\alpha_i^2}\right)$$

With this choice, the total energy of eq 3 can be calculated analytically

$$E_{\text{tot}}(\{Q_i\}, \{\mathbf{R}_i\}) = \sum_{i=1}^{N_{\text{atom}}} \left(E_i^0 + \chi_i Q_i + \frac{1}{2} \left(J_{ii} + \frac{2\gamma_{ii}}{\sqrt{\pi}} \right) Q_i^2 \right) + \sum_{i>j}^{N_{\text{atom}}} Q_i Q_j \frac{\text{erf}(\gamma_{ij} R_{ij})}{R_{ij}} \quad (4)$$

where $\gamma_{ij} = \frac{1}{\sqrt{\alpha_i^2 + \alpha_j^2}}$ and R_{ij} is the distance between atoms i and j .

Since the energy is a quadratic function of atomic charges Q_i , its minimization with respect to Q_i leads to a linear system of equations. The solution provides the equilibrated charges and E_{tot} . Even though CENT depends only on local atomic electronegativities optimized in the training, a global redistribution of charges over the entire system is obtained.

The minimization of the energy expression in eq 3 has to be done under the linear constraint of conservation of the global charge. Varying this global charge allows us to treat not only neutral but also ionized systems. The constraint gives rise to a system of equations with an additional row and column.² It can be written as

$$\begin{pmatrix} \mathbf{A} & | & 1 \\ \vdots & | & \vdots \\ 1 & | & 0 \end{pmatrix} \begin{pmatrix} Q_1 \\ \vdots \\ Q_{N_{\text{atom}}} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\chi_1 \\ \vdots \\ -\chi_{N_{\text{atom}}} \\ Q_{\text{tot}} \end{pmatrix} \quad (5)$$

where \mathbf{A}_{ij} can be expressed as

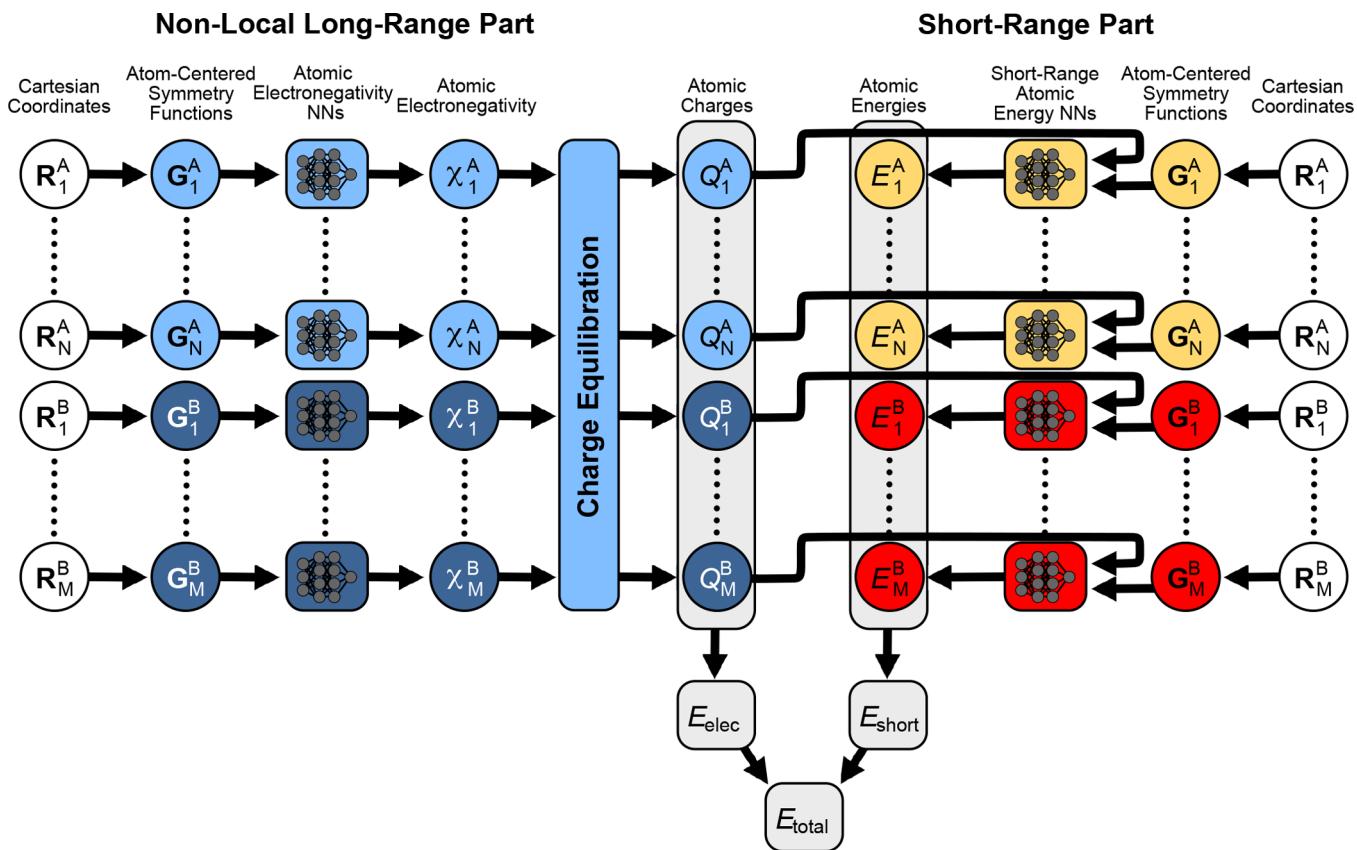


Figure 8. Structure of a 4G-HDNNP for a binary system A_NB_M .

$$A_{ij} = \begin{cases} I_i + \frac{2\gamma_{ii}}{\sqrt{\pi}} & \text{if } i = j \\ \frac{\text{erf}(\gamma_{ij}R_{ij})}{R_{ij}} & \text{otherwise} \end{cases} \quad (6)$$

A unique solution to this equation system always exists. This scheme is able to describe both neutral and charged ionic systems with the same accuracy² without the need for reparameterization.

For materials such as oxides, it was used in structure predictions,⁴¹ showing that the method not only is able to interpolate between structures contained in the training set but also is able to extrapolate to entirely new structures.

3.3. 4G-HDNNPs

While CENT represented the first MLP taking the global charge distribution into account, its numerical accuracy is typically lower compared to that of local MLPs and its applicability is restricted to predominantly ionic materials. Accurate local MLPs, on the other hand, can be constructed only for systems without substantial long-range charge transfer. For these reasons, recently the advantages of CENT and HDNNPs have been combined, resulting in fourth-generation high-dimensional neural network potentials (4G-HDNNP),³ which are nonlocal and able to capture long-range charge transfer in a wide range of molecules and materials.

As in CENT, in 4G-HDNNPs the global charge distribution is determined by a charge equilibration scheme relying on local atomic electronegativities. The resulting atomic charges are then used to compute the electrostatic energy, which is combined

with short-range atomic energies that accurately describe bonding in complex local environments.

The general structure of a 4G-HDNNP is shown in Figure 8. First, atomic charges are predicted by a charge equilibration process using environment-dependent atomic electronegativities expressed by neural networks. The atomic environments are described by ACSFs, and the same set of linear equations (eq 5) as in CENT is used to predict atomic charges. However, there is an important difference in the training process. In CENT, the neural networks predicting the atomic electronegativities are trained to minimize the error in the CENT total energy with respect to the reference total energies. The atomic charges Q_i of CENT serve merely as an intermediate quantity in this process and are not necessarily physical. In 4G-HDNNPs, the neural networks predicting the electronegativities are trained to minimize the difference between the predicted charges and reference charges from electronic structure calculations. In the same way as in CENT, a constraint to conserve the total charge of the system is included. On the basis of this training, the 4G-HDNNP charges are physically as meaningful as the reference charges and depend on the global system.

The Coulomb interaction of the determined charges is one of the two energy contributions of 4G-HDNNPs, which use the total energy expression

$$E_{\text{tot}} = E_{\text{short}} + E_{\text{elec}} \quad (7)$$

To account for the remaining short-range energy, a second set of atomic neural networks is trained to predict the energy difference E_{short} , which is obtained by subtracting the electrostatic energies E_{elec} from the electronic structure reference energies E_{ref} .

$$E_{\text{short}} = E_{\text{ref}} - E_{\text{elec}} = \sum_{i=1}^{N_{\text{atom}}} E_i(\mathbf{G}_i, Q_i) \quad (8)$$

In the same way, short-range forces are also extracted for the training process. Similar to second-generation HDNNPs, this second set of neural networks addresses primarily the influence of the local chemical environment. However, there is one important difference in that each of the atomic neural networks has one additional input neuron providing the atomic charge of the respective atom, as obtained from the charge equilibration process. In this way, nonlocal information enters the short-range part, allowing a response to global changes in the electronic structure. Like the CENT approach, 4G-HDNNPs can also be simultaneously trained for systems with different global charges enabling the construction of a single MLP for systems in different charge states.

To illustrate the capabilities of the 4G-HDNNP method, we now investigate its performance for the example of a gold atom adsorbed on a doped and undoped MgO surface as shown in Figure 2b. As can be seen in Table 1, the charge transfer caused

Table 1. Total Energy Difference ($E_{\text{Mg}} - E_{\text{O}}$) (in meV) between the Gold Atom Placed on Top of Mg or O on a Doped and an Undoped Substrate, Respectively, Calculated Using DFT and Second (2G)-, Third (3G)-, and Fourth (4G)-Generation HDNNP

	doped	undoped
DFT	-39.3	362.4
2G-HDNNP	118.4	118.4
3G-HDNNP	119.2	113.8
4G-HDNNP	-21.7	348.9

by the substitutional doping induces a preference of the gold atom to reside on top of a magnesium atom in DFT, which is well described by 4G-HDNNP, while the second- and third-generation HDNNPs fail to distinguish these surfaces.

Figure 9 shows the adsorption energy for the gold atom as a function of its vertical distance from a surface oxygen atom. For both the doped and the undoped substrate, the 4G-HDNNP adsorption energy is much closer to DFT than the energy obtained with the second-generation HDNNP. Furthermore,

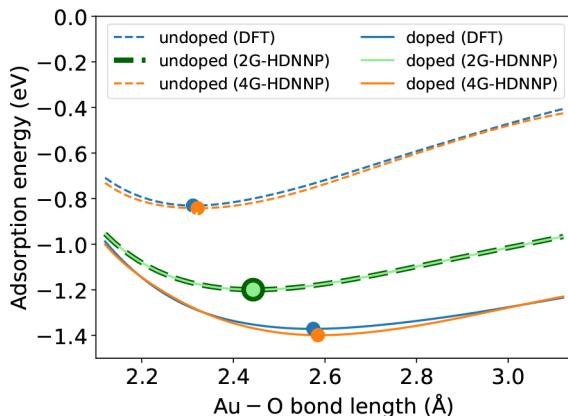


Figure 9. Comparison of DFT, second- (2G), and fourth-generation (4G) HDNNP adsorption energies for gold placed on top of oxygen on a doped and undoped MgO substrate. Equilibrium bond lengths are marked by a dot.

only the 4G-HDNNP is able to resolve the different equilibrium bond lengths, in agreement with DFT.

3.4. Further Nonlocal MLPs

The development of fourth-generation MLPs is a very active field of research, and an increasing number of methods have been published recently or are currently under development. The defining property of fourth-generation MLPs is their ability to describe long-range effects due to charge transfer as well as electrostatic interactions without truncation using an explicit Coulomb term for very large systems. Apart from CENT^{2,39} and 4G-HDNNP,³ the Becke population neural network (BpopNN)⁴² method is also based on directly including information about the atomic charges, which after being redistributed over the entire system are used for the calculation of the Coulomb energy.

The BpopNN method is a generalization of the energy minimization in CENT, which contains the electrostatic and local atomic energy terms. This local atomic energy depends not only on the short-range environment but also on the atomic population (i.e., the atomic charges). The functional dependence of the atomic energy on its population is learned from a set of constrained DFT calculations.⁴³ This energy is then minimized with respect to the populations through a self-consistent cycle called SCF-q. This method is similar to 4G-HDNNPs in that it contains an internal representation of the electronic structure in the form of global atomic charges.

Other methods such as deep tensor neural networks (DTNN),⁴⁴ the atoms-in-molecules network (AIMNet),^{10,45} and message passing neural networks (MPNN)⁴⁶ rely on an iterative exchange of information between neighboring atoms, which step by step increases the range of the atomic interactions that can be described. Thus, they cannot be unambiguously assigned to a specific generation of MLPs. Although Coulomb interactions are not explicitly included in the form of a $1/r$ term in the AIMNet method, they can be learned by the model up to the range that is defined by the number of message-passing iterations, thus approaching fourth-generation potentials in the limit of message passing through the entire system.

In DTNN,⁴⁴ the dependence on the charge is abandoned in favor of an atomic description vector that is repeatedly refined by pairwise interaction with neighboring atoms. Each of these repeated refinements increases the effective cutoff radius of the atomic interactions. Similarly, AIMNet¹⁰ uses repeated update steps, during which information is passed between nearby atoms, to refine an internal representation of each atom. In recent work,⁴⁵ AIMNet was extended to predict atomic charges and spins for systems with multiple charge states. This allows for the calculation of several quantities such as atomic electronegativity and hardness as well as condensed Fukui functions.

MPNNs⁴⁶ operate on a similar principle. A graph-based representation of the atomic system is used and information is passed along its edges. The method was successfully applied to the QM9 data set.⁴⁷ MPNNs have also been used to predict atomic charges in a method called electron-passing neural networks (EPNN).⁴⁸ Other related methods are the hierarchically interacting particle neural network (HIP-NN)⁴⁹ or SchNet,⁸ where information is also passed between nearby atoms.

4. CONCLUSIONS

The development of atomistic potentials has seen rapid progress in recent years. In particular, the introduction of machine

learning potentials nowadays allows us to achieve a level of accuracy that is comparable to electronic structure calculations, drastically extending the time and length scales of computer simulations of complex systems, for which reliable empirical potentials are not available. Most MLPs currently in use rely on a local description in the form of atomic energies with impressive applications for a variety of systems. Even though long-range electrostatic interactions have been considered on the basis of fixed or local environment-dependent charges, these approaches break down quantitatively and qualitatively if the global electronic structure of a system is modified by distant parts of the system outside the local environments or by changes in the total charge.

The resulting long-range charge transfer and nonlocal dependencies of the atomic charges can be described by a new fourth generation of MLPs, which is just emerging as a general tool for atomistic simulations. Several different methods have now been proposed with very encouraging first examples illustrating their capabilities. These “non-local” MLPs close an important gap in the applicability of MLPs for many interesting systems in chemistry, biochemistry, and materials science and are expected to allow the replacement of demanding electronic structure calculations by MLPs for more and more applications in the future. This methodical advance will further improve the general accuracy of MLPs, which in the future is expected to shift the focus of potential development more and more to the quality of the underlying reference electronic structure data, posing a natural limit for the level of accuracy that can be achieved by atomistic potentials. The possibility to calculate accurate energies and forces several orders of magnitude faster than with DFT will be a game changer in the field of atomistic simulations, allowing for reliable simulations of unprecedented size. Since by definition third- and fourth-generation MLPs cover all types of short- and long-range interactions, on the basis of local and global properties such as electrostatics and dispersion the concept of a possible future fifth generation of MLPs is yet unclear, but interesting new developments can be expected in this very dynamic field of research.

AUTHOR INFORMATION

Corresponding Authors

Stefan Goedecker – Department of Physics, Universität Basel, 4056 Basel, Switzerland; Email: stefan.goedecker@unibas.ch

Jörg Behler – Universität Göttingen, Institut für Physikalische Chemie, Theoretische Chemie, 37077 Göttingen, Germany; orcid.org/0000-0002-1220-1542; Email: joerg.behler@uni-goettingen.de

Authors

Tsz Wai Ko – Universität Göttingen, Institut für Physikalische Chemie, Theoretische Chemie, 37077 Göttingen, Germany

Jonas A. Finkler – Department of Physics, Universität Basel, 4056 Basel, Switzerland

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.accounts.0c00689>

Author Contributions

T.W.K. and J.A.F. contributed equally to this work.

Notes

The authors declare no competing financial interest.

Biographies

Tsz Wai Ko received his bachelor degree in engineering physics in 2016 and his master of philosophy in theoretical and computational physics in 2018 from Hong Kong Polytechnic University. Currently he is a Ph.D. student in theoretical chemistry at the Georg-August-Universität Göttingen under the supervision of Prof. Jörg Behler. He is working on the development of machine learning potentials with accurate long-range interactions.

Jonas A. Finkler obtained his master's degree in physics from the University of Basel, where he joined the group of Stefan Goedecker as a Ph.D. student in 2019. His main research interests are the development of new algorithms and simulation methods ranging from advanced sampling techniques to machine learning potentials.

Stefan Goedecker obtained his diploma in physics in 1987 from the Technical University of Munich and his Ph.D. in 1992 from EPFL Lausanne. He was then a postdoc at the Cornell Theory Center and at the Max-Planck Institut for solid-state research in Stuttgart. In 2000, he obtained a staff position at CEA Grenoble. Since 2003, he has been a professor of computational physics at the University of Basel. He is interested in the development and improvement of algorithms for a wide spectrum of methods for atomistic simulations.

Jörg Behler graduated in chemistry at the Universität Dortmund in 2000. He then joined the Fritz-Haber-Institut der Max-Planck-Gesellschaft in Berlin, where he obtained his Ph.D. in 2004. After a postdoctoral stay at the same place, he moved to the ETH Zurich in 2006. In 2007, he established an independent research group at the Ruhr-Universität Bochum. In 2017, he became a full professor for theoretical chemistry at the Georg-August-Universität Göttingen. His main research interest is the development of machine learning potentials with applications in chemistry and materials science.

ACKNOWLEDGMENTS

We are grateful for financial support by the Deutsche Forschungsgemeinschaft (DFG) (BE3264/13-1, project number 411538199) and the Swiss National Science Foundation (SNF) (project number 182877).

REFERENCES

- (1) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
- (2) 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: Condens. Matter Mater. Phys.* **2015**, *92*, 045131.
- (3) 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*, 398.
- (4) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **2016**, *145*, 170901.
- (5) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.
- (6) Shapeev, A. V. Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Model. Simul.* **2016**, *14*, 1153–1173.
- (7) 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.
- (8) Schütt, K.; Kindermans, P.-J.; Felix, H. E. S.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. Schnet: A continuous-filter convolutional

- neural network for modeling quantum interactions. *Adv. Neural. Inf. Process. Syst.* **2017**, 991–1001.
- (9) Unke, O. T.; Meuwly, M. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges. *J. Chem. Theory Comput.* **2019**, 15, 3678–3693.
- (10) 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, eaav6490.
- (11) Drautz, R. Atomic cluster expansion for accurate and transferable interatomic potentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2019**, 99, 014104.
- (12) Bereau, T.; DiStasio, R. A., Jr; Tkatchenko, A.; von Lilienfeld, O. A. Non-covalent interactions across organic and biological subsets of chemical space: Physics-based potentials parametrized from machine learning. *J. Chem. Phys.* **2018**, 148, 241706.
- (13) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **2017**, 8, 3192–3203.
- (14) Blank, T. B.; Brown, S. D.; Calhoun, A. W.; Doren, D. J. Neural network models of potential energy surfaces. *J. Chem. Phys.* **1995**, 103, 4129–4137.
- (15) Artrith, N.; Morawietz, T.; Behler, J. High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, 83, 153101.
- (16) Morawietz, T.; Sharma, V.; Behler, J. A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges. *J. Chem. Phys.* **2012**, 136, 064103.
- (17) Hirshfeld, F. L. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta* **1977**, 44, 129–138.
- (18) 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: Condens. Matter Mater. Phys.* **2015**, 92, 045131.
- (19) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **2011**, 134, 074106.
- (20) Barker, J. Reaction field, screening, and long-range interactions in simulations of ionic and dipolar systems. *Mol. Phys.* **1994**, 83, 1057–1064.
- (21) Sagui, C.; Darden, T. A. Molecular dynamics simulations of biomolecules: Long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, 28, 155–179.
- (22) French, R. H.; Parsegian, V. A.; Podgornik, R.; Rajter, R. F.; Jagota, A.; Luo, J.; Asthagiri, D.; Chaudhury, M. K.; Chiang, Y.-M.; Granick, S.; Kalinin, S.; Kardar, M.; Kjellander, R.; Langreth, D. C.; Lewis, J.; Lustig, S.; Wesolowski, D.; Wetlaufer, J. S.; Ching, W.-Y.; Finnis, M.; Houlihan, F.; von Lilienfeld, O. A.; van Oss, C. J.; Zemb, T. Long range interactions in nanoscale science. *Rev. Mod. Phys.* **2010**, 82, 1887.
- (23) Deng, Z.; Chen, C.; Li, X.-G.; Ong, S. P. An electrostatic spectral neighbor analysis potential for lithium nitride. *NPJ. Comput. Mater.* **2019**, 5, 75.
- (24) Darley, M. G.; Handley, C. M.; Popelier, P. L. A. Beyond point charges: Dynamic polarization from neural net predicted multipole moments. *J. Chem. Theory Comput.* **2008**, 4, 1435–1448.
- (25) Yao, K.; Herr, J. E.; Toth, D. W.; Mckintyre, R.; Parkhill, J. The TensorMol-0.1 model chemistry: A neural network augmented with long-range physics. *Chem. Sci.* **2018**, 9, 2261–2269.
- (26) Bereau, T.; Andrienko, D.; von Lilienfeld, O. A. Transferable atomic multipole machine learning models for small organic molecules. *J. Chem. Theory Comput.* **2015**, 11, 3225–3233.
- (27) Bleiziffer, P.; Schaller, K.; Riniker, S. Machine learning of partial charges derived from high-quality quantum-mechanical calculations. *J. Chem. Inf. Model.* **2018**, 58, 579–590.
- (28) Becke, A. D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, 88, 2547–2553.
- (29) Behler, J. First principles neural network potentials for reactive simulations of large molecular and condensed systems. *Angew. Chem., Int. Ed.* **2017**, 56, 12828.
- (30) Behler, J. Representing potential energy surfaces by high-dimensional neural network potentials. *J. Phys.: Condens. Matter* **2014**, 26, 183001.
- (31) Behler, J. Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem.* **2015**, 115, 1032–1050.
- (32) Grisafi, A.; Ceriotti, M. Incorporating long-range physics in atomic-scale machine learning. *J. Chem. Phys.* **2019**, 151, 204105.
- (33) Grisafi, A.; Nigam, J.; Ceriotti, M. Multi-scale approach for the prediction of atomic scale properties. *Chem. Sci.*, in press **2021**, DOI: 10.1039/DOSC04934D.
- (34) Deringer, V. L.; Csányi, G. Machine-learning based interatomic potential for amorphous carbon. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, 95, 094203.
- (35) Parsaeifard, B.; Finkler, J. A.; Goedecker, S. Detecting non-local effects in the electronic structure of a simple covalent system with machine learning methods. arXiv:2008.11277 2020.
- (36) Rappe, A. K.; Goddard, W. A., III Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **1991**, 95, 3358.
- (37) Yu, J.; Sinnott, S. B.; Phillpot, S. R. Charge optimized many-body potential for the Si/SiO₂ system. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, 75, 085311.
- (38) vanDuin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A., III ReaxFF: A reactive force field for hydrocarbons. *J. Phys. Chem. A* **2001**, 105, 9396.
- (39) Amsler, M.; Rostami, S.; Tahmasbi, H.; Rahmatizad, E.; Faraji, S.; Rasoulkhani, R.; Ghasemi, S. A. FLAME: A library of atomistic modeling environments. *Comput. Phys. Commun.* **2020**, 256, 107415.
- (40) Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, 369, 253–287.
- (41) Eivari, H. A.; Ghasemi, S. A.; Tahmasbi, H.; Rostami, S.; Faraji, S.; Rasoulkhani, R.; Goedecker, S.; Amsler, M. Two-dimensional hexagonal sheet of TiO₂. *Chem. Mater.* **2017**, 29, 8594–8603.
- (42) 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, 4256–4270.
- (43) Kaduk, B.; Kowalczyk, T.; Van Voorhis, T. Constrained density functional theory. *Chem. Rev.* **2012**, 112, 321–370.
- (44) Schütt, K. T.; Arbabzadah, F.; Chmiela, S.; Müller, K. R.; Tkatchenko, A. Quantum-chemical insights from deep tensor neural networks. *Nat. Commun.* **2017**, 8, 13890.
- (45) Zubatyuk, R.; Smith, J.; Nebgen, B. T.; Tretiak, S.; Isayev, O. Teaching a neural network to attach and detach electrons from molecules. *ChemRxiv* **2020**, DOI: 10.26434/chemrxiv.12725276.v1.
- (46) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural message passing for quantum chemistry. arXiv:1704.01212 2017.
- (47) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; Von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Sci. Data* **2014**, 1, 140022.
- (48) Metcalf, D.; Jiang, A.; Spronk, S.; Cheney, D.; Sherrill, C. D. Electron-passing neural networks for atomic charge prediction in systems with arbitrary molecular charge. *J. Chem. Inf. Model.* **2021**, 61, 115–122.
- (49) Lubbers, N.; Smith, J. S.; Barros, K. Hierarchical modeling of molecular energies using a deep neural network. *J. Chem. Phys.* **2018**, 148, 241715.