

CENT2: Improved charge equilibration via neural network techniqueEhsan Rahmatizad Khajehpasha *Physics Department, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, 45137-66731 Zanjan, I.R. Iran*Jonas A. Finkler *Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*Thomas D. Kühne  and S. Alireza Ghasemi**Dynamics of Condensed Matter and Center for Sustainable Systems Design,
Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany* (Received 25 February 2022; revised 12 April 2022; accepted 12 April 2022; published 21 April 2022)

The high computational cost of first-principles electronic structure methods together with the successful applications of machine learning (ML) techniques in atomistic simulations resulted in a surge of interest in ML-based interatomic potentials. Despite great progress in the field, there remain some challenges to be solved such as the best way of incorporating long-range interactions, as well as nonlocal charge transfer. The first generation of the charge equilibration via neural network technique (CENT) was a major step forward in concurrently taking into account both aforementioned points. Within structure prediction methods, it turned out to be a powerful tool in discovering novel polymorphs of ionic systems. On the other hand, the method is not expected to be appropriate for multicomponent systems with reference data sets in which some or all elements are subject to varying oxidation states. Here, we present the second generation of CENT, with multiple improvements to the original variant that lead to a more accurate treatment of electrostatic interactions. To do this, it aims at reproducing the electric potential function, which is directly related to the charge distribution, rather than only considering total energies. In addition, a charge-free term is added to correct for the difference between the reference energies and those obtained with the energy functional of CENT. Moreover, the Green's function within the Hartree energy is modified to substantially shield interactions from charges in the neighborhood of each point. Also, the charge density is split into ionic and electronic parts, which allows for a better approximation of the electron density. The utility of this method is examined for magnesium oxide clusters, and multiple comparisons with the first generation are made, demonstrating that much more physical electrostatic interactions can be expected from the second generation of CENT.

DOI: [10.1103/PhysRevB.105.144106](https://doi.org/10.1103/PhysRevB.105.144106)**I. INTRODUCTION**

Nowadays, atomistic modeling is routinely applied to materials of various types in order to understand microscopic processes or to calculate specific physical and chemical quantities [1]. Atomistic simulations of systems containing many particles or necessitating many calls to energy and/or force evaluations are not feasible with *ab initio* methods [2,3]. This applies even for the most popular approximation method among quantum mechanical approaches, i.e., Kohn-Sham density functional theory (KS-DFT) [4], which is currently the best trade-off between accuracy and efficiency among first-principles methods. Empirical potentials [5,6], tight-binding schemes [7], and force fields (FFs) for organic molecules and ions [8] have been successfully employed in specific applications. More sophisticated FFs have also been developed such as reactive FFs, which not only have been helpful in modeling a broader variety of compounds, but also can capture nonlocal

charge transfer in the system [9–11]. While fixed-charge FFs [5] are more efficient [8,12], reactive FFs such as ReaxFF [10] and charge-optimized potentials [11] provide higher accuracy, though at the price of higher computational cost. Despite successes with the aforementioned methods, highly accurate alternatives to achieve chemical accuracy at low cost have been desired for decades.

Recently, there has been a surge of interest in employing machine learning (ML) techniques in atomistic simulations. In particular, the use of ML techniques in constructing interatomic potentials has become a prominent alternative to standard FFs in computational chemistry and materials science [13,14]. Due to the availability of advanced and powerful computers, generating large data sets has become feasible and ML-based methods have already made a profound impact on different areas of materials science [15–23]. In recent years, methods and computer programs have been introduced that enable automatic generation of reference data points to construct an ML interatomic potential (MLIP) with minimal human intervention [24]. The key difference between the original MLIPs and FFs is the fact that the former do not

*alireza.ghasemi@uni-paderborn.de

make any prior assumption about the shape of the interaction of constituent particles in the system [25]. The MLIPs became applicable to systems of arbitrary size through the development of high-dimensional MLIPs [26,27]. An essential ingredient of MLIPs is the way they provide details of atomic environments to the machine learning engine, such that this information remains invariant under permutation, rotation, and translation. This is typically accomplished by local environment descriptors [28–30]. Even though MLIPs are relatively new, most of the aforementioned ingredients are included in a number of software packages, or libraries such as RuNNer by J. Behler, QUIP [31], aenet [32], Amp [33], DeePMD-kit [34], library-based LAMMPS [35], and FLAME [36].

The calculation of the total energy in standard high-dimensional MLIPs involves certain short-range (SR) assumptions. A different strategy in employing ML techniques for the construction of interatomic potentials was introduced in 2015 by Ghasemi and Goedecker [37]. They proposed the use of an artificial neural network (ANN) to obtain an intermediate quantity, namely the atomic electronegativity, instead of considering the output of a ML method as some sort of atomic [26] or bond energy [38]. They considered a form of functional for the total energy that has in common certain features with reactive FFs [9–11]. The method employs the charge equilibration process (CEP) to obtain atomic charges based on which the electrostatic interaction of the system is calculated. An ANN was employed for each element of the system to obtain the environment-dependent atomic electronegativities required as an input of the CEP. The method was later named charge equilibration via neural network technique (CENT) [39]. Indeed, the method not only includes the description of long-range electrostatic interactions, but also allows for a nonlocal redistribution of charge density according to environment-dependent atomic electronegativities. The method has been successfully applied to several ionic materials with a broad range of ionicity [40–47]. Inclusion of electrostatic interactions in an MLIP, e.g., via the use of the Hartree energy functional, makes the model more compliant with basic physical laws, but can also limit its range of application. In the case of CENT, one can expect that the model is best suited for ionic systems, since the charge density of the method in the Hartree term is a rough approximation of its counterpart within an electronic structure calculation.

The first attempt to incorporate long-range (LR) interactions in MLIPs was made by Artrith and Behler [48], where they proposed to split the total energy into Coulombic and SR terms, each having an independent ANN, for each component in the system. The ANN of the Coulombic term was supposed to reproduce atomic charges obtained by the Hirshfeld method. The ANN of the SR term was applied to the energy difference between the reference value and that of the Coulombic term in the same way as the standard high-dimensional ML potentials. Even though the method is a progressive step towards incorporating LR interactions into MLIPs, it retains two shortcomings: (i) atomic charges are determined via a SR process that implies only local charge transfer and (ii) the net charge is not preserved unless rescaling of the atomic charges is performed. Recently, the method was enhanced by replacing the Coulombic term by the Hartree

energy of CENT, while still determining the atomic charges via CEP. The method is known as a 4th-generation high-dimensional neural network potential (4G-HDNNP) [49]. The reference atomic charges were obtained in the same way as in the previous scheme, i.e., from the Hirshfeld method, and they remained the target quantities in the training process. Since the atomic charges obtained from the CEP are used as inputs to the NNs expressing the short-range energy, the calculation of forces is not straightforward [49,50]. This problem can be more severe if one attempts to improve the method by employing more complex charge density approximations.

In this paper, we present the second generation of CENT, which splits the total energy into two terms, but differently from that of a 4G-HDNNP. The new variant has three major improvements in comparison to the original CENT: (i) an SR term in the form of a high-dimensional ANN is added to the energy functional of the original CENT, which can be thought of as allowing the constant atomic energies in the original CENT to become environment-dependent quantities through the use of an ANN for each one, (ii) the total charge density is split into ionic and electronic distributions, and (iii) the CENT part is trained to reproduce a modified electric potential in which SR electrostatic interactions are shielded by means of modifying the Green's function. In the next section, we describe the method in detail. In the following, we will refer to the first generation of CENT as CENT1 and the new variant as CENT2.

II. METHODS

In this new scheme, the CENT energy is not intended to account for the entire total energy, but only part of it. Rather, we demand that the form of the CENT functional be able to accurately reproduce LR contributions of the electrostatic interactions. In pursuit of this objective, we split the total energy as follows:

$$U = U_{\text{SR}} + U_{\text{CENT}}. \quad (1)$$

Therein, U_{SR} is a solely SR and charge-free contribution to the total energy. The second term includes charge-dependent components, and it is expected to provide an approximate treatment of the electrostatic interactions. For a system of N atoms, this is given by

$$U_{\text{CENT}} = \sum_{i=1}^N \left[(q_i + z_i)\chi_i + \frac{1}{2}J_i(q_i + z_i)^2 \right] + U_{\text{SLR}}, \quad (2)$$

where J_i is the hardness and q_i and z_i are the electronic and ionic charges of atom i , respectively. Moreover, U_{SLR} is the shielded long-range (SLR) electrostatic energy that does not contain short-range interactions. Similar to CENT1, χ_i is the atomic environment-dependent electronegativity of atom i . Atomic charges are obtained by the minimization of the total energy, i.e., by the CEP. The role of U_{CENT} in CENT2 differs from that in CENT1 conceptually. In CENT1, χ_i values are predicted by an ANN such that the CENT energy reproduces the reference total energies. In the new variant, χ_i values are supposed to result in a charge density, which reproduces, accurately, only the LR portion of the electric potential. This

is done by introducing a U_{trial} quantity, defined as

$$U_{\text{trial}} = \int \rho_{\text{trial}}(\mathbf{r}; \mathbf{R}_p) V_{\text{SLR}}(\mathbf{r}) d^3r, \quad (3)$$

where ρ_{trial} is a trial charge density with which the potential is probed in the vicinity of a probing point \mathbf{R}_p . In our implementation, the trial charge density is considered to be a normalized Gaussian function with exponent length equal to one Bohr. We consider probing points on a uniform hexahedral mesh covering the interior points of the structure and partially its exterior points, i.e., points close to the surface of the structure. Lastly, V_{SLR} is a modified electric potential in which the electrostatic interaction of points arising from charges in its neighborhood is substantially shielded. A systematic way of shielding can be done by modifying the kernel (Green's function) in the Coulomb integral, i.e.,

$$V_{\text{SLR}} = \int \kappa(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r}') d^3r', \quad (4)$$

where κ is a shielded Coulombic Green's function, and it is given by

$$\kappa(r) = \frac{1}{r} - \underbrace{\frac{e^{-\lambda^n r^n}}{r}}_{\kappa_s(\mathbf{r})}. \quad (5)$$

Therein, λ is a shielding factor, and n is an integer greater than or equal to one. All numerical results presented in this study are computed with $n = 4$. The Green's function in Eq. (5) shields the SR contributions of electrostatic interactions, while keeping all long-range contributions intact. Once the method is implemented for a given n , the degree of the shielding can be controlled systematically by adjusting only a single parameter, namely λ . The smaller the value we consider for λ , the stronger the shielding of SR contributions will be. Currently, we do not have a very efficient implementation of Eq. (4) that obtains V_{SLR} for free boundary conditions. Instead, we treat the Coulombic and shielding terms separately and then add them together. The former term can be calculated using the Poisson solvers available in almost every DFT package, and some are very well modularized as a library, like the BigDFT PSolver [51]. The latter term can be computed in Fourier space on an expanded cell, since it decays rapidly, especially if $n \geq 2$ is employed, as we have done in our study. A detailed derivation is presented in Appendix. Finally, Eq. (4) can be very efficiently calculated with a similar implementation to that in Ref. [52].

Similar to CENT1, we consider the total charge density as a superposition of atom-centered functions with finite range. In contrast to CENT1, we separate the charge density into ionic and electronic parts as is customary in *ab initio* methods, i.e., $\rho(\mathbf{r}) = \rho^n(\mathbf{r}) + \rho^e(\mathbf{r})$. This allows us to search for more flexible forms of the electron density that can expectantly resemble the reference DFT counterpart to a better degree without changing ionic atomic charge densities. We consider $\rho^n(\mathbf{r})$ as the superposition of atom centered spherically symmetric Gaussian functions with narrow width β . The electron density is expanded in terms of spherically symmetric Gaussian functions, that is to say multiple functions with different widths α_i , on each atom. In this study, we consider three and two Gaussian functions for oxygen and magnesium atoms,

respectively. The splitting of charge density into ionic and electronic terms is not only a step towards grasping electrostatic interaction in atomistic systems, it is also an avenue by which better approximations for the electron density that can account for atomic-scale polarizations might be developed in the future. Therefore, we have

$$\begin{aligned} \rho_i^n(\mathbf{r}) &= z_i G_i(\mathbf{r}; \beta), \\ \rho_i^e(\mathbf{r}) &= \begin{cases} q'_i \sum_{j=1}^2 c_j G_j(\mathbf{r}; \alpha_j) & \text{for Mg} \\ q'_i \sum_{j=1}^2 c_j G_j(\mathbf{r}; \alpha_j) + 2G_i(\mathbf{r}; \beta) & \text{for O,} \end{cases} \\ \sum_{j=1}^2 c_j &= 1 \quad \text{and} \quad i = 1, \dots, N, \end{aligned} \quad (6)$$

where $q'_i = q_i$ for Mg atoms and $q'_i = q_i - 2$ for O atoms. Furthermore, $G_i(w)$ is a Gaussian function with width w centered on atom i at \mathbf{r}_i given by

$$G_i(\mathbf{r}; w) = \frac{1}{w^3 \pi^{\frac{3}{2}}} e^{-\frac{|\mathbf{r}-\mathbf{r}_i|^2}{w^2}}. \quad (7)$$

However, $2G_i(\mathbf{r}; \beta)$ applies only for O atoms, where the $2s$ orbital has been considered to be as localized as the ionic charges. This choice is made for the simplicity of implementation and does not influence the results, since the strong shielding causes the ionic and electronic contributions of the $2s$ orbital to the electric potential to cancel out at points for which the interaction is not shielded. We use the same value of $\beta = 0.5 a_0$ both for oxygen and for magnesium atoms where a_0 is the Bohr radius. If the shielding of the interaction is strong enough, as we suggest in this paper, widths well below one bohr, which are suitable for ion densities, lead to nearly identical shielded electric potentials. In principle, one can let c_j values vary independently for each atom, which would require employing a different χ for each c_j and an independent ANN accordingly. To do this, one would need to modify the first term of Eq. (2). In this study, however, we keep this fixed and allow only q_i to vary. As a result, we consider only one environment-dependent atomic electronegativity for each atom. The values of α_j and c_j are different for different elements in the system. We obtained them by fitting to the reference electron density of the magnesium oxide (MgO) dimer. We used the coefficients $c_2 = 7.011$ and $c_3 = -6.011$ both for Mg and O atoms. Two Gaussian functions of width $\alpha_1 = 1.93 a_0$ and $\alpha_2 = 1.8335 a_0$ for Mg atoms and $\alpha_1 = 1.05 a_0$ and $\alpha_2 = 0.9975 a_0$ for O atoms, respectively, lead to an optimal fit for the electron density of the dimer. It is worth mentioning that the reference electron density based on the aforementioned pseudopotentials calculations has no peak on atoms and our choice of approximate atomic charges imitates this behavior. Once q_i values are obtained via the CEP, i.e., by minimization of the energy with respect to the charges so that $\frac{\partial U}{\partial q_i} = 0$, the modified electrostatic energy is calculated via

$$U_{\text{SLR}} = \frac{1}{2} \int \rho(\mathbf{r}) V_{\text{SLR}}(\mathbf{r}) d^3r. \quad (8)$$

It is worth mentioning that in our implementation U_{SLR} is not the target quantity during the training process.

A. Training

In order to train the ANN of atomic electronegativities, we generate reference trial energies based on the ground-state electron density of reference structures. Since we have not implemented our procedure in a DFT code, we save the ground-state electron density of each reference structure into a file, which is subsequently read by the FLAME code [36]. The program generates a large number of trial energies at probing points on a regular mesh with grid spacing of $1.5 a_0$ that extends beyond the structure on each side by $2 a_0$. The reference trial energies are obtained using Eq. (3), where V_{SLR} is based on the ground-state electron density of the reference structure plus the ion density, as described in Sec. II.

We define the cost function as follows:

$$\text{cost function} = \sum_i^{N_s} (U_{i,\text{trial}} - U_{i,\text{trial}}^{\text{ref}})^2, \quad (9)$$

where N_s is the number of reference training data points, namely the total number of trial energies of reference structures. Direct training of the ANN associated with the cost function above is computationally impractical due to the huge number of trial energies. Instead, we split the cost function into groups, each one consisting of the trial energies of one structure. In this way, not only can we perform the calculations in parallel, we can also obtain q_i for a given structure and then perform the inverse CEP to obtain the reference χ_i . The latter allows us to bypass multiple execution of the CEP for each structure during the training process. Indeed, obtaining q_i from trial energies is similar to obtaining atomic charges from the electrostatic potential, which is known to be ill conditioned. In our method, this is not a serious problem for two reasons: (i) we have a better approximation of the charge density than that from the typical choice of point particles, and (ii) we employ the Green's function given in Eq. (5), which allows us to probe the interior points of the structure, whereas typically only exterior points sufficiently far from the structure are probed in the standard approaches so that the Coulomb potential is a good approximation. Nevertheless, we employed a regularization scheme with a control parameter whose value is adjusted according to the second derivative matrix and is equal to one thousandth of the largest eigenvalue. Once q_i are obtained for each structure and the reverse CEP is applied, the resulting environment-dependent atomic electronegativities can be used to train an ANN for the U_{CENT} term.

Then, for each structure, U_{CENT} is calculated and subtracted from the DFT total energies to obtain reference U_{SR} values. In the second step, the standard high-dimensional ANN [26] was employed to treat the U_{SR} term. For this step, the ANN used in the first step, i.e., ANN in U_{CENT} , was kept fixed. The components of the training process in CENT2 are depicted in Fig. 1.

B. Reference Training Data

The reference total energies and forces of all structures were calculated using the BigDFT code [53] with the PBE

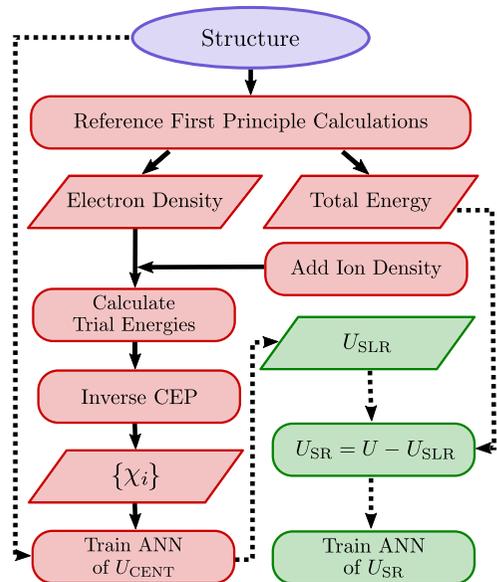


FIG. 1. Workflow of the CENT2 training procedure.

exchange-correlation functional [54]. The Mg $3s$ and O $2s$ and $2p$ orbitals were treated as valence within the framework of the norm-conserving pseudopotentials [55]. Grid spacing of $0.4 a_0$ was used in all BigDFT calculations, where the electron density was on a finer mesh with a grid spacing of $0.2 a_0$. Reference structures were generated using the minima hopping algorithm [56] based on the CENT1 potential, whose performance was illustrated in Refs. [36] and [46], respectively. The reference set consisting of 8038 charge-neutral MgO clusters with sizes ranging from 26 to 120 atoms was used to train the ANN potential. The reference data set was split into the training and validation data sets, containing 6430 and 1608 structures, respectively. The same architecture was used for all four ANNs, two per element, one for U_{SR} and the other for atomic electronegativities in U_{CENT} . Each ANN had two hidden layers of 15 nodes, and the output layer had one node. We employed atom-centered symmetry functions [28], both radial and angular, which were 70 in total and determined the number of nodes in the input layer.

Very small values for the shielding factor lead to large shielding of electrostatic interaction, which is in principle undesirable. In the next section, we will examine three values and compare the impact of this parameter. However, it is worth having an estimate of the contributions of the two terms to the total energy, namely U_{SR} and U_{CENT} , and their distributions. Figure 2 illustrates the histogram of the total energy, as well as its two component terms per atom for the structures in the reference data set. The calculations were performed with $\lambda = 0.16 a_0^{-1}$ using the reference electron densities added to the ion charge densities. The variation of the total energies is comparable to those of its two constituents.

III. RESULTS AND DISCUSSION

A. Error estimates for all reference data

In this study, we present only results for $n = 4$ [see Eq. (5)], which leads to nearly perfect shielding of small distances.

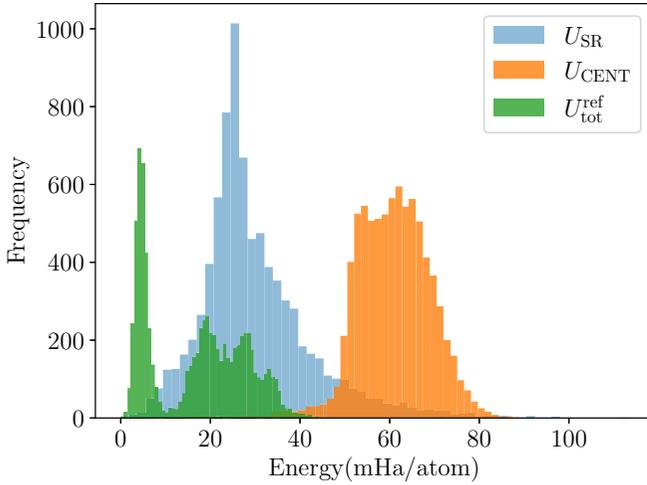


FIG. 2. Histogram of the total energy U_{SR} and U_{CENT} in the reference dataset. Within each contribution, the minimum value is subtracted from all other values.

Three shielding factors 0.20, 0.18, and 0.16 a_0 were examined, where $\frac{1}{\lambda}$ for these three values approximately corresponds to 1.35, 1.50, and 1.68 times the average nearest-neighbor distance in the core of our large MgO clusters, respectively. An estimate of the shielding strength can be found as the value of $e^{-\lambda+r^4}$ at the nominal nearest-neighbor distance. For the aforementioned λ values, it is about 0.74, 0.82, and 0.88, respectively. These values imply that the electrostatic interaction between nearest-neighbor atoms is predominantly suppressed.

Energy and force root mean square errors (RMSEs) of training and validation data sets are listed for the three aforementioned λ values in Table I. To avoid overfitting, we consider early stopping, thus the RMSE values of the 25th epoch are presented in the table. The results indicate that one can obtain a smaller value of energy RMSE by considering a stronger shielding, e.g., $\lambda = 0.18 a_0^{-1}$ rather than $\lambda = 0.20 a_0^{-1}$. But, $\lambda = 0.18 a_0^{-1}$ appears to be nearly optimal. Nevertheless, we also consider $\lambda = 0.16 a_0^{-1}$, since it leads to a non-negligible decrease in force RMSE. Fig. 3 depicts the learning curves for $\lambda = 0.16 a_0^{-1}$. The energy and force RMSE evolutions both indicate that the training process is saturated at the 25th epoch, beyond which errors decrease marginally.

Reference U_{SLR} values can in principle be included in the training process. However, we instead considered the quantity as a test to assess the reliability and transferability of the method. It is rather astonishing that reference U_{SLR} values

TABLE I. Energy and force RMSE values at the 25th epoch of CENT2 training.

λ (a_0^{-1})	Energy RMSE (mHa/atom)		Force RMSE (mHa a_0^{-1})	
	Training	Validation	Training	Validation
0.20	0.32	0.39	10.6	10.4
0.18	0.26	0.31	9.1	9.5
0.16	0.27	0.32	8.1	8.0

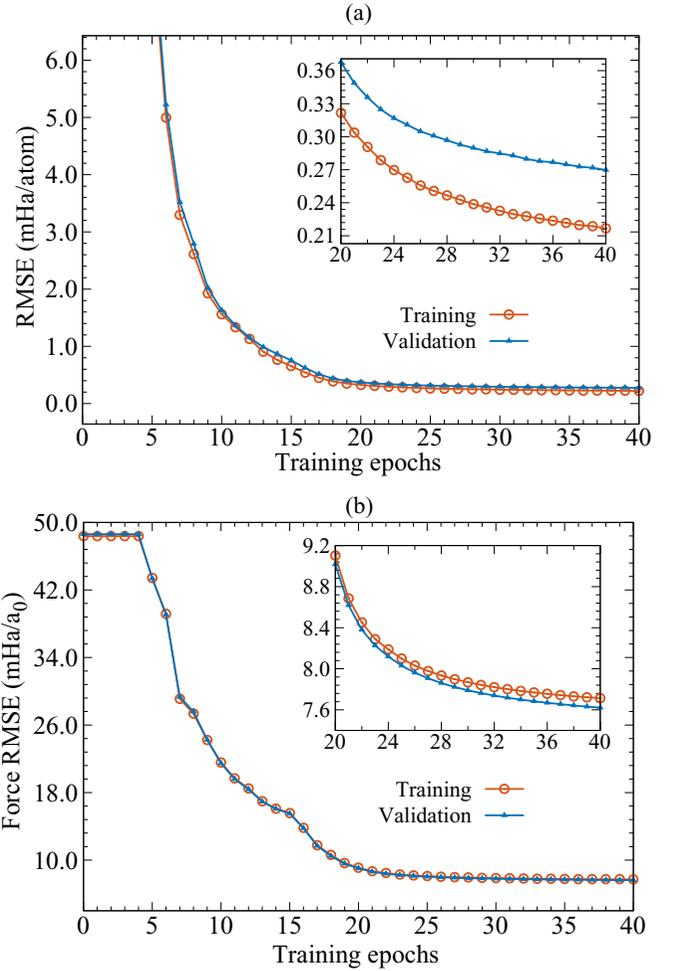


FIG. 3. The convergence of the energy and force RMSE values with respect to the training iterations for training and validation datasets.

were reproduced within a small error when excluded in the training process. Figure 4 illustrates a remarkably adequate correlation between the reference U_{SLR} values and those predicted by CENT2, as obtained in our calculations for $\lambda =$

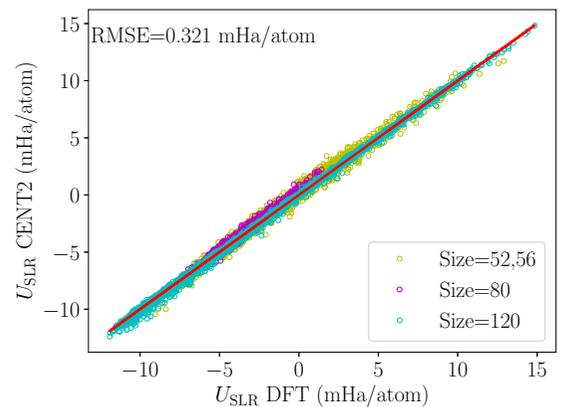


FIG. 4. Comparison of reference U_{SLR} energies vs CENT2 U_{SLR} values. There is a decent correlation although U_{SLR} is not a target quantity of the training procedure.

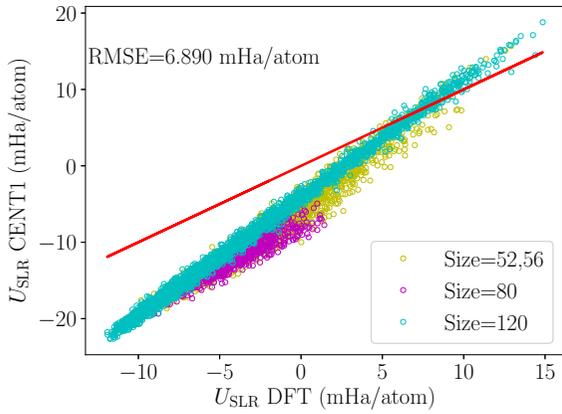


FIG. 5. Comparison of reference U_{SLR} energies versus CENT1 U_{SLR} values.

$0.16 a_0^{-1}$. Unfortunately, CENT1 U_{SLR} values do not correlate with those of the reference U_{SLR} , as shown in Fig. 5. This is due to the poor charge redistribution in that scheme. As shown in Fig. 6, U_{SLR} values obtained from Hirshfeld charges do not have a decent correlation with the reference values as well. It is to be noted that the RMSE of 0.321 (mHa/atom) for the U_{SLR} values is slightly larger than the total energy RMSE, indicating that further improvements on electron density approximation are expected to lower the error of the total energy as well as the long-range part. For example, adding p -type orbitals for the treatment of atomic-scale polarization may lead to a more accurate electrostatic interaction and consequently a smaller error within U_{SLR} .

B. Analyses on V_{SLR} for an in-plane distorted rocksalt-like 64-atom structure

CENT1 was successfully applied to multiple ionic systems over a broad range of ionicity. Indeed, the method exhibited the great ability of using ML to interpolate an intermediate physical quantity with an unforeseeable degree of transferability [39] and predictive power for novel crystalline structures [40,41]. Nevertheless, it has a few drawbacks and limitations such as the applicability exclusively to ionic systems. More-

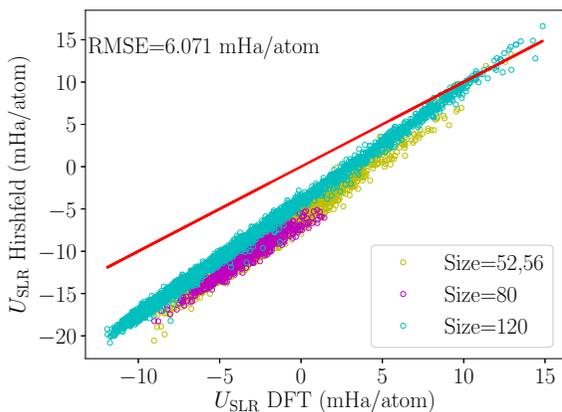


FIG. 6. Comparison of reference U_{SLR} energies versus Hirshfeld U_{SLR} values.

over, the large variation of atomic charges is unavoidable if the training process is carried on far enough to achieve a small error in energy. This is attributed to the approximate charge density used in CENT1 together with the presumably inappropriate distribution of charges imposed by CEP. Apparently, the latter cannot be resolved as long as one aims at reproducing the total energies. The former can be relieved to some extent if a more genuine form of the charge density is adopted. This cannot be done without separating the charge density into ionic and electronic parts. As detailed in Sec. II, the aforementioned problems are abundantly addressed through the changes made in CENT2. The target training quantity of the U_{CENT} term is no longer the total energy, but the reference U_{trial} values. Also, the charge density is improved by splitting it into ionic and electronic densities, as well as employing multiple Gaussian functions for each atomic energy. Here we investigate the electric potential function within the CENT1 and CENT2 methods, and also that arising from Hirshfeld atomic charges, in detail for a planar cluster structure. The accurate electric potential is the key element to obtain both electrostatic energy and forces.

The configuration for which the following tests were carried out was generated by making random in-plane displacements of atoms from equilibrium for an in-plane distorted rocksalt-like 64-atom structure (see Fig. 7). We first compare a vector quantity that is a term in the equation for calculating atomic electrostatic forces. In fact, the quantity probes the long-range electrostatic force exerted on ions arising from the total charge, i.e., the sum of ionic and electronic charge densities. The quantity for the atom at position \mathbf{r}_i is calculated with the following equation:

$$\tilde{\mathbf{f}}_i = \int V_{\text{SLR}}(\mathbf{r}) \frac{\partial \rho_i^n(\mathbf{r})}{\partial \mathbf{r}_i} d^3 r. \quad (10)$$

Three methods are considered, namely via CENT1, CENT2, and Hirshfeld charges. In each method, the resulting charge density is supplied to Eq. (4), with which the atomic electrostatic forces are calculated. However, in the case of Hirshfeld and CENT1 charges, the electron density is formed by the superposition of Gaussian charges with $q_i = Q - q_i^{(\text{Hirshfeld})}$, where Q is 2 for Mg and 6 for O atoms. Indeed, Q indicates the number of valence electrons considered in the reference DFT calculations. For an unbiased comparison, the ionic charges are considered to be identical in all three methods. When using atomic charges, one can envision that only long-range contributions of electrostatic interactions are treated fairly accurately. Therefore, we calculated forces in all of the aforementioned methods with the same formula, namely Eq. (4). The reference DFT atomic electrostatic forces, obtained by BigDFT from the ground state density and the same formula as in the other three methods, are presented in Fig. 7. DFT forces are superimposed with those from each of the three methods in the panels of Fig. 7 for a better comparison. Apparently, CENT1 LR atomic forces [Fig. 7(a)] have a poor agreement with those of DFT, whereas CENT2 and the results from Hirshfeld are far more accurate. However, CENT2 LR atomic forces [Fig. 7(b)] have lower error than those of Hirshfeld [Fig. 7(c)] at the periphery of the structure. The RMSE of atomic forces for CENT1, CENT2, and Hirshfeld are 0.049, 0.013, and 0.022 ($\text{Ha } a_0^{-1}$), respectively.

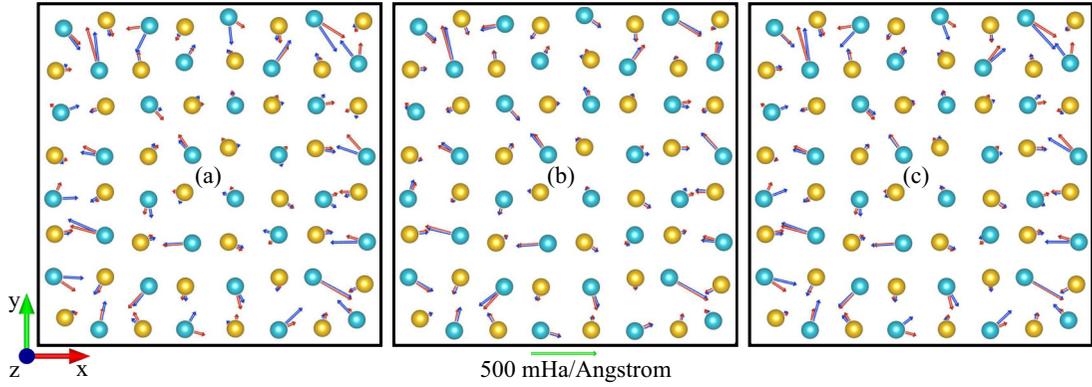


FIG. 7. Comparison of \vec{f} defined in Eq. (10) obtained with charge densities of different methods: (a) DFT (red arrows) vs CENT1 (blue arrows), (b) DFT (red arrows) vs CENT2 (blue arrows), and (c) DFT (red arrows) vs Hirshfeld charges (blue arrows). Oxygen and magnesium atoms are shown as blue and gold spheres, respectively. RMSE values of the CENT1, CENT2, and Hirshfeld \vec{f} vectors with respect to the DFT vectors are 0.029, 0.008, and 0.022 (Ha a_0^{-1}), respectively.

LR atomic forces are related to V_{SLR} , which in turn is associated with the charge density. We illustrate V_{SLR} in Figs. 8 and 9. The latter shows V_{SLR} along three lines, which are two diagonals and one edge of the structure. Clearly, the CENT1 potential function and that of Hirshfeld charges correlate weakly with that of DFT. It is to be noted that both methods result in electric potential functions that do not correlate with that of DFT, if the remaining valence electronic and ionic densities are not added. Among all three methods, CENT2 has the best correlation along all three lines. The error of CENT1 V_{SLR} is as large as that of the reference method, i.e., V_{SLR} of DFT. This is as can be seen both in the color map (Fig. 8) and the potential along the three lines (Fig. 9). The poor performance of the CENT1 method in reproducing V_{SLR} indeed pertains to the fact that in CENT1 one trains the ANN such that U_{CENT} gives accurate total energies, while in the case of CENT2 one employs the term to reproduce a physical quantity that the scheme is capable of reproducing. In addition, the charge density in CENT2 is split into ionic and electronic parts, which allows the employment of a better approximation for the electronic term. In fact, CENT2 offers a route to approximate the electron density more accurately, as discussed earlier. In fact, it can even be improved in the future with local charges, which can capture atomic-scale polarization. For this, one must also add a term in Eq. (2) to equilibrate atomic polarizations, in the same way as atomic charges, through a new environment-dependent quantity.

Having a better approximation of the electron density would result in a more accurate electrostatic interaction, which may also ease the task of training for the short-range term, i.e., U_{SR} .

IV. CONCLUSION

In this paper, we presented the second generation of the CENT scheme that, similar to the original variant, accounts for nonlocal charge transfer. There has been several changes to the method, which result in a more accurate treatment of long-range interaction due to a more reliable charge redistribution. The method can also ease further improvements that might be possible by adding extra terms to the energy functional and by employing more accurate approximations to the electronic charge density. The second generation of CENT has three major differences in comparison to the original CENT: (i) an SR term in the form of high-dimensional ANN is added to the original CENT energy functional, (ii) the total charge density is split into ionic and electronic distributions, and (iii) the CENT is trained to reproduce a modified electric potential in which SR electrostatic interactions are shielded by modifying the Green's function. We applied the method to MgO clusters of various size and diverse structural motifs. It is worth noting that the calculation of forces does not require any sophisticated trick and the same standard approach of the first

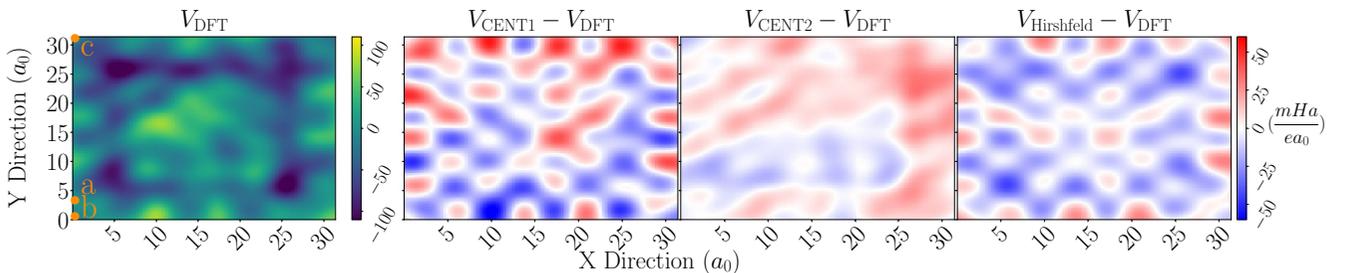


FIG. 8. Color map of V_{SLR} within the plane formed by atoms for the in-plane distorted rocksalt-like 64-atom structure. Left panel depicts V_{SLR} , as computed by DFT, and the other three panels (from left, CENT1, CENT2, and Hirshfeld) show the difference with respect to the DFT values.

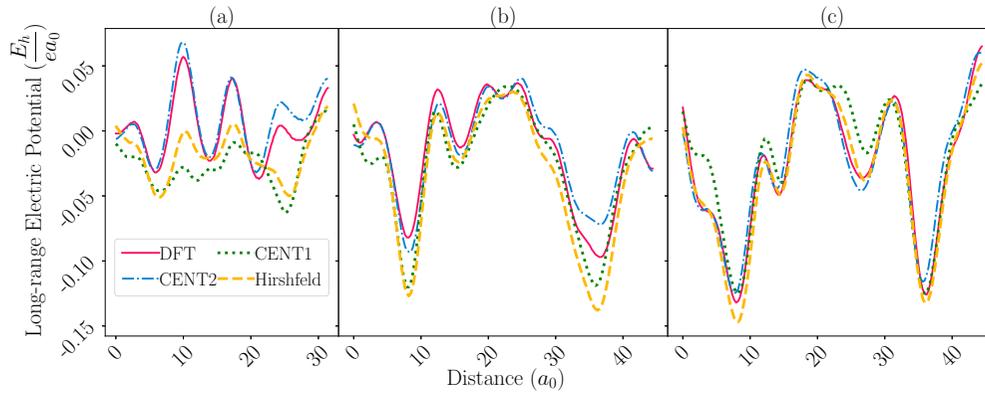


FIG. 9. Comparison of V_{SLR} within the plane formed by atoms as a function of distance as calculated with the charge densities of DFT, CENT1, CENT2, and Hirshfeld along three lines, (a) x direction, (b) $y = x$ direction, and (c) $y = -x$ direction. The distance in each panel is with respect to the corresponding point of a, b, and c shown in the left panel of Fig. 8.

generation is employed. We illustrated that the new variant is superior to the first generation when electrostatic interactions are concerned. We envisage that the new variant has the capability of being utilized for nonionic systems, as well as ionic materials of varying stoichiometry, where an element has a variety of oxidation states. In addition, the method can be improved upon in a more systematic and straightforward way than is possible for its rivals. The method is expected to have even better performance for extended systems. Implementation of the method for boundary conditions suitable for surface problems and bulk materials will be presented in the future.

ACKNOWLEDGMENTS

S.A.G. and E.R.K. thank S. Goedecker for fruitful discussions. Furthermore, S.A.G. appreciates financial support from the Alexander von Humboldt Foundation by means of a Georg Forster research fellowship. The authors acknowledge generous grants of computing time on OCuLUS and super-computer NOCTUA from the Paderborn Center for Parallel Computing (PC²). Also, the authors gratefully acknowledge support from the scientific computing core (sciCORE) facility at the University of Basel.

APPENDIX

1. Calculation of kernel in Fourier space with e^{-r^n} shielding

In order to calculate the modified electrostatic energies of Eq. (4) with the kernel in Eq. (5), we decided on a two-step implementation in which the Coulombic contribution is assumed to be computed using the standard methods available in DFT packages. Therefore, we provide details for the implementation of the second part, i.e., the screening term with the Green's function $\kappa_s(\mathbf{r}) = \frac{e^{-\lambda^n r^n}}{r}$. The screening term is a short-range one, whose extent of impact is determined by n and λ . If these two parameters are chosen to make the range of the screened potentials sufficiently small, then one can expand the cell beyond the charge density to the extent where the potential is nonzero. This allows one to calculate the potential function arising from the screening part of the kernel in the

Fourier space as

$$V_s(\mathbf{r}) = \frac{1}{(2\pi)^3} \int \tilde{\kappa}_s(\mathbf{k}) \tilde{\rho}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} d^3k, \quad (\text{A1})$$

where $\tilde{\kappa}_s$ and $\tilde{\rho}$ are the Fourier transform of the κ_s and the charge density, respectively. To calculate $\tilde{\kappa}_s$, we do as follows:

$$\begin{aligned} \tilde{\kappa}_s(\mathbf{k}) &= \int \kappa_s(\mathbf{r}') e^{-i\mathbf{k}\cdot\mathbf{r}'} d^3r' \\ &= \frac{4\pi}{k} \int_0^\infty e^{-\lambda^n r'^n} \sin(kr') dr'. \end{aligned} \quad (\text{A2})$$

The $\tilde{\kappa}_s(\mathbf{k})$ function in Eq. (A2) has a compact form for $n = 1$, given by

$$\tilde{\kappa}_s(\mathbf{k}) = \frac{4\pi}{k^2 + \lambda^2}. \quad (\text{A3})$$

Indeed, such a Green's function results in the Yukawa potential. It neither decays sufficiently fast at mid-range and long distances, unless for large shielding factors, nor does it shield strongly the impact of charges in the vicinity of an observation point. The method of expanding the cell is impractical, or at least inefficient for $n = 1$, unless the screening factor is large. Therefore, the method given in Ref. [52] should be used. However, we do not recommend the use of $n = 1$ in CENT2, since such screening does not enforce the exclusion of the SR interaction to the extent that is required. For $n = 2$, the kernel in Fourier space is given by

$$\tilde{\kappa}_s(\mathbf{k}) = \frac{4\pi F(\frac{k}{2\lambda})}{k\lambda}, \quad (\text{A4})$$

where F is the Dawson function, also denoted as D_+ . For $n > 2$, we take a numerical approach, because the kernel in Fourier space is given by hypergeometric functions, whose calculations are numerically inefficient. With change of variable, Eq. (A2) yields

$$\tilde{\kappa}_s(\mathbf{k}) = \frac{4\pi}{k\lambda} \int_0^\infty e^{-r^n} \sin(tr) dr, \quad (\text{A5})$$

where $t = \frac{k}{\lambda}$. Expanding the sine function, we have

$$\tilde{G}(\mathbf{k}) = \frac{4\pi}{k\lambda} \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} \left(\frac{k}{\lambda}\right)^{2m+1} \int_0^{\infty} e^{-r^n} r^{2m+1} dr. \quad (\text{A6})$$

Using the identity

$$\int_0^{\infty} e^{-r^n} r^m dr = \frac{1}{n} \Gamma\left(\frac{m+1}{n}\right) \quad (\text{A7})$$

yields

$$\tilde{\kappa}(\mathbf{k}) = \frac{4\pi}{nk\lambda} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{k}{\lambda}\right)^{2m+1} \Gamma\left(\frac{2m+2}{n}\right)}{(2m+1)!}. \quad (\text{A8})$$

This series converges very rapidly if $\frac{k}{\lambda}$ is not large. Therefore, we consider this solution only for $\frac{k}{\lambda} < 1$, for which the series certainly converge rapidly, and we need the first few terms. For $\frac{k}{\lambda} > 1$, we use the discrete Fourier transform. From integration by parts, we have

$$\int_0^{\infty} e^{-r^n} \sin(tr) dr = \frac{1}{t} - \frac{n}{t} \int_0^{\infty} r^{n-1} e^{-r^n} \cos(tr) dr. \quad (\text{A9})$$

Therefore, we get

$$\tilde{\kappa}(\mathbf{k}) = \frac{4\pi}{k^2} - \frac{4\pi n}{k^2} \int_0^{\infty} r^{n-1} e^{-r^n} \cos(tr) dr. \quad (\text{A10})$$

We define

$$g(t) := \int_0^{\infty} r^{n-1} e^{-r^n} \cos(tr) dr. \quad (\text{A11})$$

The Fourier transform of $g(t)$ is

$$\tilde{g}(\omega) = \begin{cases} 0 & \omega = 0 \\ \pi \omega^{n-1} e^{-\omega^n} & \omega > 0 \\ \pi (-\omega)^{n-1} e^{-\omega^n} & \omega < 0. \end{cases} \quad (\text{A12})$$

Using the fast Fourier transform, we calculate $\tilde{\kappa}(k)$ on a one-dimensional grid, and using cubic splines, we can thus evaluate the function at the points in the reciprocal space.

2. Solution to a Gaussian charge density

Since the system of linear equations must be solved during the training process, building and factorizing the matrix would help make this efficient. To do this, we employ the solution of the potential for the Gaussian charge density. For the charge density, given in Eq. (7), normalized to q , we have

$$\tilde{\rho}(\mathbf{k}) = q e^{-\frac{w^2 k^2}{4}}. \quad (\text{A13})$$

Therefore, the potential can be written as

$$V_s(\mathbf{r}) = \frac{q}{2\pi^2} \int_0^{\infty} e^{-\lambda^n r^n} dr' \int \frac{1}{k} e^{-\frac{w^2 k^2}{4}} \sin(kr') e^{i\mathbf{k}\cdot\mathbf{r}} d^3k.$$

It simplifies to

$$V_s(\mathbf{r}) = \frac{q}{\pi^{\frac{1}{2}} w r} \int_0^{\infty} e^{-\lambda^n r^n} \left[e^{-\frac{(r-r')^2}{w^2}} - e^{-\frac{(r+r')^2}{w^2}} \right] dr',$$

and at the origin to

$$V_s(0) = \frac{4q}{\pi^{\frac{1}{2}} w^3} \int_0^{\infty} e^{-\lambda^n r^n} e^{-\frac{r^2}{w^2}} r' dr'.$$

For $n = 1$, it has a compact form of

$$V_s(\mathbf{r}) = \frac{q e^{-\frac{w^2 \lambda^2}{4} - \lambda r}}{2r} \left[\operatorname{erf}\left(\frac{r}{w} - \frac{w\lambda}{2}\right) - e^{2\lambda r} \operatorname{erfc}\left(\frac{w\lambda}{2} + \frac{r}{w}\right) + 1 \right]. \quad (\text{A14})$$

For $n = 2$, it has also a compact form, i.e.,

$$V_s(\mathbf{r}) = \frac{q}{r\sqrt{w^2\lambda^2 + 1}} e^{-\frac{\lambda^2 r^2}{w^2\lambda^2 + 1}} \operatorname{erf}\left(\frac{r}{w\sqrt{w^2\lambda^2 + 1}}\right). \quad (\text{A15})$$

-
- [1] J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopoulos, H. Reis, and M. K. Shukla, *Handbook of Computational Chemistry* (Springer, Cham, Switzerland, 2017).
- [2] L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella, and L. Genovese, *WIREs Comput. Mol. Sci.* **7**, e1290 (2017).
- [3] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann *et al.*, and *J. Chem. Phys.* **152**, 194103 (2020).
- [4] W. Kohn and L. Sham, *Phys. Rev.* **137**, A1697 (1965).
- [5] F. G. Fumi and M. P. Tosi, *J. Phys. Chem. Solids* **25**, 31 (1964).
- [6] J. Tersoff, *Phys. Rev. B* **30**, 4874 (1984).
- [7] Th. Frauenheim, F. Weich, T. Kohler, S. Uhlmann, D. Porezag, and G. Seifert, *Phys. Rev. B* **52**, 11492 (1995).
- [8] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996).
- [9] A. K. Rappe and W. A. Goddard, *J. Phys. Chem.* **95**, 3358 (1991).
- [10] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, *J. Phys. Chem. A* **105**, 9396 (2001).
- [11] J. Yu, S. B. Sinnott, and S. R. Phillpot, *Phys. Rev. B* **75**, 085311 (2007).
- [12] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [13] W. Andreoni and S. Yip, *Handbook of Materials Modeling* (Springer, Cham, Switzerland, 2020).
- [14] L. Fiedler, K. Shah, M. Bussmann, and A. Cangi, *Phys. Rev. Materials* **6**, 040301 (2022).
- [15] F. Brockherde, L. Vogt, L. Li, M. E. Tuckerman, K. Burke, and K.-R. Müller, *Nat. Commun.* **8**, 872 (2017).
- [16] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, and O. A. von Lilienfeld, *J. Chem. Theory Comput.* **13**, 5255 (2017).
- [17] O. Isayev, C. Oses, C. Toher, E. Gossett, S. Curtarolo, and A. Tropsha, *Nat. Commun.* **8**, 15679 (2017).
- [18] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J.

- Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik, *ACS Cent. Sci.* **4**, 268 (2018).
- [19] A. Mansouri Tehrani, A. O. Oliynyk, M. Parry, Z. Rizvi, S. Couper, F. Lin, L. Miyagi, T. D. Sparks, and J. Brgoch, *J. Am. Chem. Soc.* **140**, 9844 (2018).
- [20] P. S. Gromski, A. B. Henson, J. M. Granda, and L. Cronin, *Nat. Rev. Chem.* **3**, 119 (2019).
- [21] J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev, and A. E. Roitberg, *Nat. Commun.* **10**, 2903 (2019).
- [22] J. Kessler, F. Calcavecchia, and T. D. Kühne, *Adv. Theor. Simul.* **4**, 2000269 (2021).
- [23] S. A. Ghasemi and T. D. Kühne, *J. Chem. Phys.* **154**, 074107 (2021).
- [24] H. Mirhosseini, H. Tahmasbi, S. R. Kuchana, S. A. Ghasemi, and T. D. Kühne, *Comput. Mater. Sci.* **197**, 110567 (2021).
- [25] S. Lorenz, M. Scheffler, and A. Gross, *Phys. Rev. B* **73**, 115431 (2006).
- [26] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [27] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [28] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- [29] A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B* **87**, 184115 (2013).
- [30] L. Zhu, M. Amsler, T. Fuhrer, B. Schaefer, S. Faraji, S. Rostami, S. A. Ghasemi, A. Sadeghi, M. Grauzinyte, C. Wolverton, and S. Goedecker, *J. Chem. Phys.* **144**, 034203 (2016).
- [31] G. Csányi, S. Winfield, J. Kermode, M. C. Payne, A. Comisso, A. De Vita, and N. Bernstein, *IoP Comput. Phys. Newsletter Spring* 1 (2007).
- [32] N. Artrith and A. Urban, *Comput. Mater. Sci.* **114**, 135 (2016).
- [33] A. Khorshidi and A. A. Peterson, *Comput. Phys. Commun.* **207**, 310 (2016).
- [34] H. Wang, L. Zhang, J. Han, and W. E, *Comput. Phys. Commun.* **228**, 178 (2018).
- [35] A. Singraber, J. Behler, and C. Dellago, *J. Chem. Theory Comput.* **15**, 1827 (2019).
- [36] M. Amsler, S. Rostami, H. Tahmasbi, E. R. Khajehpasha, S. Faraji, R. Rasoulkhani, and S. A. Ghasemi, *Comput. Phys. Commun.* **256**, 107415 (2020).
- [37] S. A. Ghasemi, A. Hofstetter, S. Saha, and S. Goedecker, *Phys. Rev. B* **92**, 045131 (2015).
- [38] K. V. J. Jose, N. Artrith, and J. Behler, *J. Chem. Phys.* **136**, 194111 (2012).
- [39] S. Faraji, S. A. Ghasemi, S. Rostami, R. Rasoulkhani, B. Schaefer, S. Goedecker, and M. Amsler, *Phys. Rev. B* **95**, 104105 (2017).
- [40] H. A. Eivari, S. A. Ghasemi, H. Tahmasbi, S. Rostami, S. Faraji, R. Rasoulkhani, S. Goedecker, and M. Amsler, *Chem. Mater.* **29**, 8594 (2017).
- [41] R. Rasoulkhani, H. Tahmasbi, S. A. Ghasemi, S. Faraji, S. Rostami, and M. Amsler, *Phys. Rev. B* **96**, 064108 (2017).
- [42] S. Rostami, M. Amsler, and S. A. Ghasemi, *J. Chem. Phys.* **149**, 124106 (2018).
- [43] S. Faraji, S. A. Ghasemi, B. Parsaeifard, and S. Goedecker, *Phys. Chem. Chem. Phys.* **21**, 16270 (2019).
- [44] E. R. Khajehpasha, S. Goedecker, and S. A. Ghasemi, *J. Comput. Chem.* **42**, 699 (2021).
- [45] S. Rostami, N. Seriani, S. A. Ghasemi, and R. Gebauer, *Phys. Rev. Materials* **5**, 063605 (2021).
- [46] H. Tahmasbi, S. Goedecker, and S. A. Ghasemi, *Phys. Rev. Materials* **5**, 083806 (2021).
- [47] S. Rostami, S. A. Ghasemi, and S. Goedecker, *Phys. Rev. Mater.* **5**, 123603 (2021).
- [48] N. Artrith, T. Morawietz, and J. Behler, *Phys. Rev. B* **83**, 153101 (2011).
- [49] T. W. Ko, J. A. Finkler, S. Goedecker, and J. Behler, *Nat. Commun.* **12**, 398 (2021).
- [50] P. P. Poier, L. Lagardère, J.-P. Piquemal, and F. Jensen, *J. Chem. Theory Comput.* **15**, 6213 (2019).
- [51] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Beylkin, *J. Chem. Phys.* **125**, 074105 (2006).
- [52] A. Cerioni, L. Genovese, A. Mirone, and V. A. Sole, *J. Chem. Phys.* **137**, 134108 (2012).
- [53] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, A. Bergman, and R. Schneider, *J. Chem. Phys.* **129**, 014109 (2008).
- [54] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [55] C. Hartwigsen, S. Goedecker, and J. Hutter, *Phys. Rev. B* **58**, 3641 (1998).
- [56] S. Goedecker, *J. Chem. Phys.* **120**, 9911 (2004).