

# Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network

S. Alireza Ghasemi,<sup>1,\*</sup> Albert Hofstetter,<sup>2</sup> Santanu Saha,<sup>2</sup> and Stefan Goedecker<sup>2</sup>

<sup>1</sup>*Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran*

<sup>2</sup>*Department of Physics, Universität Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

(Received 16 January 2015; revised manuscript received 22 June 2015; published 30 July 2015)

Based on an analysis of the short-range chemical environment of each atom in a system, standard machine-learning-based approaches to the construction of interatomic potentials aim at determining directly the central quantity, which is the total energy. This prevents, for instance, an accurate description of the energetics of systems in which long-range charge transfer or ionization is important. We propose therefore not to target directly with machine-learning methods the total energy but an intermediate physical quantity, namely, the charge density, which then in turn allows us to determine the total energy. By allowing the electronic charge to distribute itself in an optimal way over the system, we can describe not only neutral but also ionized systems with unprecedented accuracy. We demonstrate the power of our approach for both neutral and ionized NaCl clusters where charge redistribution plays a decisive role for the energetics. We are able to obtain chemical accuracy, i.e., errors of less than a millihartree per atom compared to the reference density functional results for a huge data set of configurations with large structural variety. The introduction of physically motivated quantities which are determined by the short-range atomic environment via a neural network also leads to an increased stability of the machine-learning process and transferability of the potential.

DOI: [10.1103/PhysRevB.92.045131](https://doi.org/10.1103/PhysRevB.92.045131)

PACS number(s): 61.46.Bc, 34.20.Cf, 46.15.-x

## I. INTRODUCTION

Atomistic simulations for materials are nowadays widely applied to understand and design materials. A wide spectrum of simulation methods exist, ranging from quasi-exact many-electron wave-function methods [1], density functional theory (DFT) calculations [2], and semiempirical quantum-mechanical methods [3,4] to interatomic potentials [5]. For the different methods there are well-known trade-offs between the computational costs and the accuracy of the calculations. In contrast to classical force fields, density functional calculations give sufficient accuracy for a wide range of properties and are therefore the method of choice in a huge number of studies. Due to their computational cost, DFT simulations are, however, limited in practice to systems containing less than a few thousand atoms. Since there is a great need to do highly accurate simulations for larger systems, there have been numerous efforts to improve the accuracy of force fields without increasing their cost too much. It has been widely recognized that fixed charges limit the accuracy of the established standard force fields. In the chemistry and biology communities polarizable force fields have therefore been developed [6]. However, such polarizable force fields allow only for the displacement of charges and do not allow for a true charge transfer over large distances. Charge equilibration methods [7–9] offer this possibility. In force fields such as ReaxFF [10] and the charge-optimized many-body (COMB) potential [11] it has been demonstrated that the addition of charge-equilibration terms to the conventional terms can improve the accuracy of the scheme. COMB potentials have been successfully applied to several materials such as silica [12] and titania [9].

The basic obstacle for accurate standard force fields is the parametrization problem. Certain forms of parameterized analytical functions have to be chosen to model the physical interactions, but it is not known which analytical functions are optimal. In machine-learning-based methods, on the other hand, it is not necessary to choose analytical functions. The machine-learning process will automatically find the optimal form, which is not represented in terms of classical analytic functions but, for instance, via a neural network. As a consequence, the accuracies obtained with force fields based on machine-learning processes are much higher than for standard force fields. This has been demonstrated for covalent bulk materials such as carbon and silicon, for which machine-learning-based total-energy schemes [13,14] turned out to give density functional accuracy at greatly reduced numerical cost. For finite systems such as clusters, the construction of highly accurate machine-learning-based force fields is more difficult since atoms at a surface behave quite differently from atoms in the bulk. It is not surprising that an analysis of the charge distribution of NaCl clusters obtained from density functional calculations clearly reveals a charge transfer between atoms at the surface and in the core of the cluster. The fixed charges of  $\pm 1$  electron used in established force fields such as the Tosi-Fumi force field [5] are therefore clearly inadequate to describe such systems with high accuracy. Since in the standard machine-learning-based interatomic potentials the energy of the whole system is written as a sum of the energies of individual atoms [15] and since the energy of an individual atom is determined by the short-range chemical environment, long-range charge transfers as well as ionization cannot be well described.

Constant charges are also present in the majority of force fields for biomolecules, and finding good charges for use in such force fields is highly nontrivial. Even though the electrostatic part of the interaction energy is not the dominating one in these force fields, it is to be expected that variable charges

\*[aghasemi@iasbs.ac.ir](mailto:aghasemi@iasbs.ac.ir)

could also lead in this context to considerable improvements of the accuracy. Environment-dependent charges obtained from neural networks (NN) were recently introduced [16], but this approach does not give the correct total charge of the system and can, for instance, lead to fluctuations of the total charge in molecular dynamics simulations. The fact that the total charge cannot be fixed also prevents the treatment of ionized systems. In addition, atomic reference charges must be provided in this approach. The extraction of such charges from *ab initio* calculations is always ambiguous.

A machine-learning-based approach which, in the spirit of our paper, does not directly aim at the total energy but at an intermediate physical quantity is the work of Snyder and coworkers [17,18], in which they propose to construct machine-learning-based kinetic energy functionals. Whereas in this work the charge density can, in principle, adopt any form, we restrict our charge density to an approximate superposition of atomic charge densities. In this way we exploit the well-known fact that the charge density in molecules, clusters, and solids is given within a first approximation by a superposition of atomlike charge densities. The distribution of the electronic charge density is determined by atomic environment-dependent electronegativities. These electronegativities are in turn determined by the short-range environment of the associated atom and can easily be predicted by a neural-network process. To demonstrate the power of our approach we choose ionic clusters in which a correct description of the charge density is essential since bonding is mediated through charge transfer. The fact that the charge distribution can redistribute itself over the whole cluster will allow us to treat both surface and bulk atoms with very high accuracy, and we will demonstrate that density functional accuracy can be obtained with such an approach for clusters. In contrast to conventional force fields, our approach also allows us to describe ionized systems without the need of any reparametrization.

## II. METHOD

We consider a system consisting of  $\frac{N}{2}$  sodium and  $\frac{N}{2}$  chloride atoms. We postulate the following simple form for the total-energy expression:

$$U_{\text{tot}}(\{q_i\}) = \sum_{i=1}^N \left( E_i^0 + \chi_i q_i + \frac{1}{2} J_{ii} q_i^2 \right) + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (1)$$

where  $E_i^0$  are some fixed reference energies which in our implementation are set to energies of isolated atoms,  $q_i$  are atomic charges, and  $\chi_i$  is the environment-dependent atomic electronegativity of atom  $i$  whose functional dependence is determined by a neural-network approach. To describe the hardness [19] of atom  $i$  we introduce element-dependent atomic hardness terms  $J_{ii}$ . The charge density of the system  $\rho(\mathbf{r})$  is, in our approach, assumed to be a superposition of spherically symmetric Gaussian functions centered at atomic positions, each normalized to the corresponding atomic charge

$q_i$  given by

$$\rho_i(\mathbf{r}) = \frac{q_i}{\alpha_i^3 \pi^{\frac{3}{2}}} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_i|^2}{\alpha_i^2}\right).$$

With this choice for the atomic charge densities, the total energy of Eq. (1) can be calculated analytically.

$$U_{\text{tot}}(\{q_i\}, \{\mathbf{r}_i\}) = \sum_{i=1}^N \left[ E_i^0 + \chi_i q_i + \frac{1}{2} \left( J_{ii} + \frac{2\gamma_{ii}}{\sqrt{\pi}} \right) q_i^2 \right] + \sum_{i>j}^N q_i q_j \frac{\text{erf}(\gamma_{ij} r_{ij})}{r_{ij}}, \quad (2)$$

where  $\gamma_{ij} = \frac{1}{\sqrt{\alpha_i^2 + \alpha_j^2}}$  and  $r_{ij}$  is the distance between atoms  $i$  and  $j$ . The atomic charges  $q_i$  are implicitly environment dependent through  $\chi_i$ , as will be explained below, and are therefore implicit functions of the atomic positions. The implicit dependence of  $q_i$  is obtained by minimizing the total energy of Eq. (2) with respect to  $q_i$ . This leads to a linear system of equations where all matrix elements depend on the atomic positions,

$$\frac{\partial U_{\text{tot}}}{\partial q_i} = 0, \quad \forall i = 1, \dots, N \implies \sum_{j=1}^N A_{ij} q_j + \chi_i = 0, \quad (3)$$

where

$$A_{ij} = \begin{cases} J_{ii} + \frac{2\gamma_{ii}}{\sqrt{\pi}} & \text{for } i = j, \\ \frac{\text{erf}(\gamma_{ij} r_{ij})}{r_{ij}} & \text{otherwise.} \end{cases}$$

The fact that the electrostatic interaction energy of any continuous charge density is always positive and that  $J_{ii}$  are positive constants implies that the matrix  $A$  is positive definite and that the linear system of equations therefore always has a unique solution which gives the minimal electrostatic energy. The linear system equation (3) has to be solved under the constraint that the atomic charges sum up to the correct overall charge  $q_{\text{tot}} = \sum_{i=1}^N q_i$ . Adding this constraint via the Lagrange multipliers leads to the modified linear system of equations

$$\tilde{A} \mathbf{Q} = \mathbf{X}, \quad (4)$$

where  $\mathbf{Q}$  and  $\mathbf{X}$  are vectors of the dimension  $(N+1)$  and  $\tilde{A}$  is a  $(N+1) \times (N+1)$  matrix. In expanded form Eq. (4) is given as

$$\left( \begin{array}{c|c} & 1 \\ \hline A_{i,j} & \vdots \\ & 1 \\ \hline 1 & \dots & 1 & 0 \end{array} \right) \begin{pmatrix} q_1 \\ \vdots \\ q_N \\ \lambda \end{pmatrix} = \begin{pmatrix} -\chi_1 \\ \vdots \\ -\chi_N \\ -q_{\text{tot}} \end{pmatrix}. \quad (5)$$

In this way, we can allow for charge transfer over long distances without the need to find the presumably extremely complicated explicit long-range environment dependence of  $q_i$ . All that is needed to get the implicit long-range dependence of  $q_i$  on the atomic positions is an explicit short-range dependence of  $\chi_i$ , which is fixed once and for all by our neural network together with the solution of a simple linear system of equations. In addition, the total charge of the system is conserved, unlike in the method given by Ref. [16]. This

approach is physically motivated since in a Kohn-Sham density functional calculation the total energy is also minimized with respect to the charge-density distribution. So the approach can be considered some kind of constrained minimization of the total energy, where the form of the charge density is restricted to a sum over Gaussian functions with amplitudes  $q_i$ . In Kohn-Sham density functional theory, the total energy consists, of course, not only of the electrostatic term but also the kinetic and exchange-correlation terms which oppose a charge distribution that would merely minimize the electrostatic part of the total energy. In our scheme this opposing force is represented by the constraints on the form of the charge density and the environment dependence of  $\chi_i$ .

Even though our energy expression is missing the kinetic and exchange-correlation terms, the physically important energy differences can be calculated in a way similar to that in density functional theory based on the Hellmann-Feynman theorem. Energy differences are, by definition, given by the integral over the atomic forces times the displacements. The long-range force acting on the atoms is the classical electrostatic force arising from the charge distribution. In this way it is actually also guaranteed that the charges obtained by solving the linear system of equations are an accurate representation of the physical charge density within the limitations imposed by our adopted form of the charge density. Unreasonable charge densities would lead to the wrong forces and hence to the wrong energy differences between different structures.

As in Ref. [15], the input layers of the neural networks are fed with the so-called symmetry functions. A symmetry function is a transformation of the Cartesian atomic coordinates onto a set of numbers describing the chemical environments of an atom. The symmetry functions depend explicitly on the positions of all atoms in the local chemical environment embedded in a sphere with a given cutoff radius. We adopted the functional form of the symmetry functions from Ref. [20] with a modification of the cutoff function as described below. Our notation is consistent with Ref. [20].

$$G_i^2 = \sum_j e^{-\eta(R_{ij}-R_s)^2}, \quad (6)$$

where  $\eta$  is the width of the Gaussian and  $R_s$  is its center. The symmetry function of type  $G^5$  for atom  $i$  is

$$G_i^5 = 2^{1-\zeta} \sum_j (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(R_{ij}^2+R_{ik}^2)} f_c(R_{ij}) f_c(R_{ik}), \quad (7)$$

where  $\eta$  defines the Gaussian widths and the parameter  $\zeta$  controls the angular resolution. The parameters  $\lambda$  can have values  $+1$  and  $-1$ , which shift the maxima of the function in the parentheses. We used 8 radial functions of type  $G^2$  and 43 angular functions of type  $G^5$ . Both types of symmetry functions,  $G^2$  and  $G^5$ , have parameters by which information of different spatial regions of atomic environment is extracted; for example, all eight radial functions have the same functional form of type  $G^2$  but with different values for  $\eta$ . We employed a cutoff function  $f_c$ , which is different from what was done in Ref. [20]. We used the following polynomial as the cutoff

function:

$$f_c(R) = \left(1 - \frac{R^2}{R_c^2}\right)^3, \quad (8)$$

where  $R_c$  is the cutoff radius. This type of cutoff function has the advantage that its first and second derivatives are continuous at the cutoff radius.

In summary, the algorithm consists of the following steps. For each atom in the system, the characteristic array of symmetry functions values is calculated and given as an input to the NN. The output of the NN is then the environment-dependent electronegativity of the atom. Next, the subsequent solution of the linear equations gives the atomic charges. In the last step, the total energy and atomic forces are calculated using Eq. (2) and the equations given in Appendix B.

### III. APPLICATION AND RESULTS

In order to construct and examine the method, we generated a large number of reference data points using the BIGDFT [21] code. An accurate evaluation of the electrostatic term in density functional calculations is of great importance, particularly for ionized clusters. In the BIGDFT code, the Hartree potential is calculated using the method given in Ref. [22], which enables us to have accurate energetics for the ionized reference data points. The DFT calculations are performed with the local-density approximation (LDA). A large number of data points consisting of NaCl neutral and ionized ( $+1$  and  $+2$ ) clusters ranging from 8 to 80 atoms with step sizes of 4 are generated. Data points are split into three sets: training, validation, and test. Only clusters with fewer than 44 atoms are contained in the training set. The 64-atom clusters are reserved for testing, and the bigger ones are used as the validation data set. Training data points are used for updating the neural-network weights during a training process. Validation data points are used to select among different training runs. Since a large diversity of geometries in the reference data structures is crucial to good training for potentials based on machine-learning schemes, we have used great care in generating the reference data structures. The large diversity of structures in our data sets is discussed in Appendix A.

In order to avoid overfitting, all training runs are performed with only 15 epochs. As in Refs. [14,20], a hyperbolic tangent function is used as the activation function of hidden layers. Also the activation function of the output layer (values of electronegativities) is taken to be the hyperbolic tangent function rather than the commonly used linear function. This allows us to avoid a strong variation of atomic electronegativities from one atom to another within a structure, which in turn may result in too large variations in atomic charges. In contrast to the standard NN potentials (e.g., Refs. [14,16]), we can achieve in our approach small errors with a small number of nodes in NN hidden layers. Among several fits using various NN architectures and several training runs with different initial random numbers for NN weights, we obtained the best compromise between small root-mean-square error (RMSE) and transferability of the potential for training and validation data sets with the architecture 51-3-3-1, i.e., 51 symmetry functions, two hidden layers each containing 3 nodes, and the single-node output layer. In this way, the number of degrees

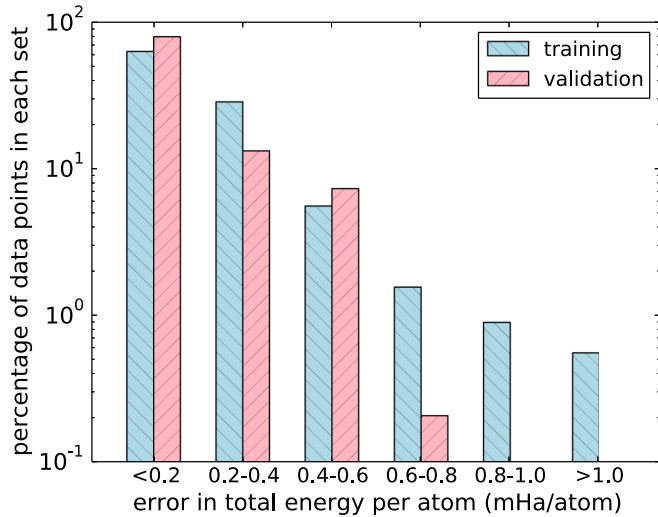


FIG. 1. (Color online) Distribution of error in total energy per atom with respect to reference DFT calculations for all data points in the training and validation sets. The number of data points is normalized to the total number of data points in each data set times 100, indicating the percentage. In total the training and validation sets contain 15 191 and 7658 structures, respectively.

of freedom has been reduced by one order of magnitude compared to that in the method of Ref. [16], which requires several thousand degrees of freedom. The reason why our method requires many fewer degrees of freedom comes from the fact that it is easier to fit with a short-range scheme (i.e., neural network with short-range symmetry functions) a quantity which itself, namely, electronegativity, is intrinsically short range. Appropriate values for Gaussian widths and atomic hardnesses  $J_{ii}$  can be obtained by a hand-tuning process. Gaussian widths of 1.0 and 2.0 bohrs are found to be optimal for sodium and chlorine atoms, respectively. Furthermore, values of 0.2 and 0.1 a.u. are found to be suitable for  $J_{ii}$  for sodium and chlorine atoms, respectively. The hand-tuning procedure is done such that errors in total energy decrease during the training process within a few epochs while avoiding large variation of atomic electronegativities from one atom to another within a structure which in turn prevents unreasonable charge transfer among atoms.

The obtained RMSE of the training and validation sets is 0.26 mHa per atom, the RMSE of the neutral test set is 0.13 mHa per atom, and the RMSE of the ionized test set (including  $q_{\text{tot}} = +1$  and  $q_{\text{tot}} = +2$ ) is 0.44 mHa per atom. The error distribution of the training and validation sets is illustrated in Fig. 1. For a few structures, the error rises to about 1 mHa per atom, while for the majority of structures it is well below this value and thus far below the so-called chemical accuracy of 1.6 mHa per particle. Since large structures are less challenging for our potential due to a smaller surface-to-volume ratio and our choice of selecting larger structures in the validation data set, the error of the validation data points is in some energy intervals less than that of the training data points. Figure 2 shows LDA, the Perdew-Burke-Ernzerhof (PBE) functional [23], and NN energies for different compression and expansion ratios relative to the equilibrium geometry of the respective method

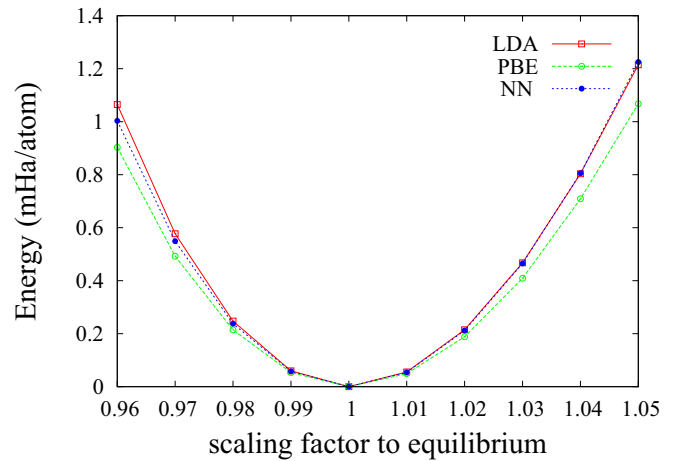


FIG. 2. (Color online) Energy vs compression/expansion ratios (scaling factor) relative to the equilibrium geometry of the respective method for a 64-atom NaCl cubic structure.

for a 64-atom NaCl cubic structure. The result shown in Fig. 2 demonstrates clearly the transferability of our method since such compressed and expanded configurations were included in neither the training nor validation set.

Figure 3 shows the drastic improvement in accuracy of our force field compared to a standard Tosi-Fumi force field with constant charges. All energies in Fig. 3 are relative to the average energy per atom of all structures for each method.

In order to check whether the entire low-energy configurational space is well described by our interatomic potential we sampled the entire low-energy configurational space for different cluster sizes with the minima hopping method [24] (MH), checking for local minima on the NN potential-energy surface that correspond to unphysical configurations. No physically unreasonable configurations were found in all these

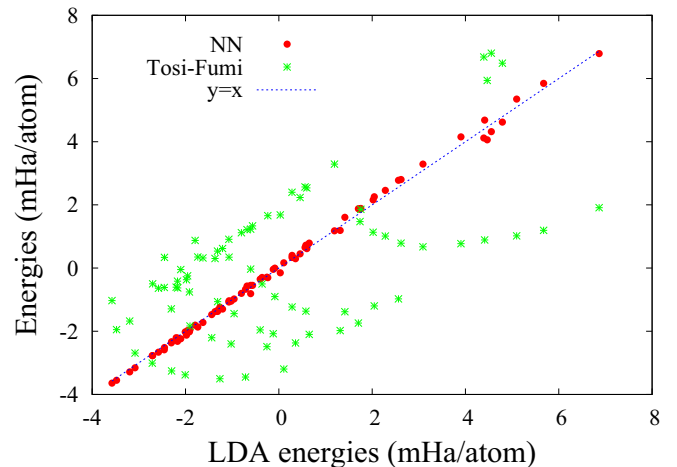


FIG. 3. (Color online) Comparison of the correlation in energy between the LDA reference values and the NN values as well as between the LDA values and the values obtained by the Tosi-Fumi force field. Only neutral systems are considered since the Tosi-Fumi force field is not applicable to ionized systems.

runs, indicating that the interatomic potential describes well the entire low-energy configurational space.

#### IV. CONCLUSION

Based on a machine-learning algorithm, a scheme was presented to reproduce with high accuracy potential-energy surfaces of quantum-mechanical origin. Instead of predicting atomic energies directly, we use NN methods to predict environment-dependent atomic electronegativities from which atomic charges are obtained by a charge equilibration process [7,8]. Once the charges are available, the total energy can be calculated easily. Applying the method to neutral and ionized sodium chloride clusters shows that unprecedented accuracy can be obtained for a particularly difficult system, namely, clusters in which the atomic environment differs drastically between surface and core atoms. The potential is highly transferable and does not give rise to any unphysical structures.

The error in the energies of our interatomic potential compared to the reference density functional data is far less than the error in total energies arising from the use of different exchange-correlation functionals. This shows that, in principle, even accuracies higher than those obtainable from density functional theory could be achieved if in the training and validation sets energies are calculated with a more accurate method.

#### ACKNOWLEDGMENTS

Computational resources were provided by the CSCS under Project No. s499, and the support of M.-G. Giuffreda is greatly acknowledged. S.S. was supported by the SNF.

#### APPENDIX A: NATURE OF THE REFERENCE DATA STRUCTURE

The reference data contain cluster structures with a large variety of sizes and shapes. As an example, we describe the nature of the structures of the clusters containing 16 atoms. The structures can be classified as regular and irregular. The regular structures include planar shapes, rods, ladders, and cubes, as well as a mixture of all these geometries. In the irregular structures, atoms are randomly distributed with the only constraint that bond lengths are neither too short nor too long. Ten selected 16-atom NaCl cluster structures, found in the reference data, are illustrated in Fig. 4. Other cluster sizes contain similar types of structures; however, the diversity of geometries grows as the number of atoms in the cluster increases.

#### APPENDIX B: ATOMIC FORCES

The calculation of atomic forces is straightforward. The force exerted on atom  $j$  is given by

$$\mathbf{F}_j = -\frac{\partial U_{\text{tot}}}{\partial \mathbf{r}_j} - \sum_{k=1}^N \frac{\partial U_{\text{tot}}}{\partial q_k} \frac{\partial q_k}{\partial \mathbf{r}_j}, \quad (\text{B1})$$

where the second term vanishes as required by Eq. (3). So we do not have to calculate the derivative of the charges  $q_i$  with

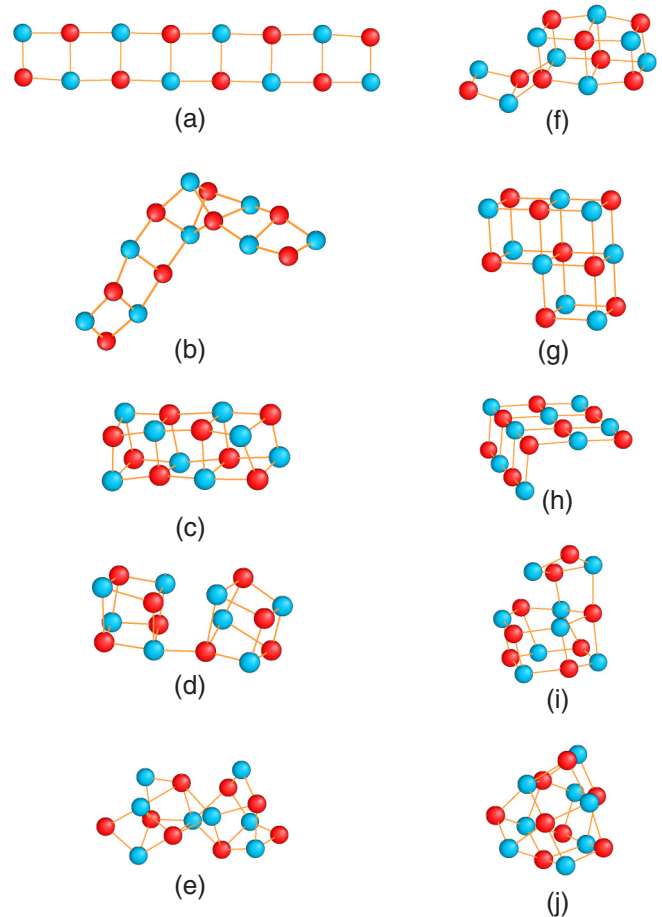


FIG. 4. (Color online) Illustration of structural diversity in the reference data sets for the case of 16-atom clusters: (a) ladder-type structure, (b) ladder-type structure which is bent, (c) cuboid structure, (d) cuboid structure in which three bonds are broken, (e) a nearly elongated random structure, (f) a ladder structure attached to a cuboid, (g) two cuboid structures which share a cube, (h) a rodlike structure attached to a plane, (i) a compact random structure, and (j) another compact random structure.

respect to the atomic positions. This is analogous to the use of the Hellmann-Feynman theorem in standard DFT, which tells us that the derivative of the wave function (and consequently also of the charge density) with respect to the atomic positions is not required. So the formula simplifies to

$$\mathbf{F}_j = -\sum_{i=1}^N \left( q_i \frac{\partial \chi_i}{\partial \mathbf{r}_j} \right) + \sum_{i>j}^N q_i q_j \frac{\partial V_{ij}}{\partial \mathbf{r}_j},$$

where  $V_{ij} = \frac{\text{erf}(\gamma_{ij} r_{ij})}{r_{ij}}$ . The derivatives of  $\chi_i$  with respect to atomic positions are calculated in the neural-network process. The first term describes the force acting on atom  $j$  arising from the local charge distribution around this atom. The second term in the force expression is the contribution arising from the electrostatic energy of the global charge distribution excluding the self-energy of a Gaussian.

- [1] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, New York, 2008).
- [2] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [3] C. M. Goringe, D. R. Bowler, and E. Hernández, *Rep. Prog. Phys.* **60**, 1447 (1997).
- [4] D. A. Papaconstantopoulos and M. J. Mehl, *J. Phys. Condens. Matter* **15**, R413 (2003).
- [5] M. P. Tosi and F. G. Fumi, *J. Phys. Chem. Solids* **25**, 45 (1964).
- [6] B. G. Dick and A. W. Overhauser, *Phys. Rev.* **112**, 90 (1958).
- [7] A. K. Rappe and W. A. Goddard, III, *J. Phys. Chem.* **95**, 3358 (1991).
- [8] C. E. Wilmer, K. C. Kim, and R. Q. Snurr, *J. Phys. Chem. Lett.* **3**, 2506 (2012).
- [9] Y.-T. Cheng, T.-R. Shan, T. Liang, R. K. Behera, S. R. Phillpot, and S. B. Sinnott, *J. Phys. Condens. Matter* **26**, 315007 (2014).
- [10] A. C. T. vanDuin, S. Dasgupta, F. Lorant, and W. A. Goddard, III, *J. Phys. Chem. A* **105**, 9396 (2001).
- [11] J. Yu, S. B. Sinnott, and S. R. Phillpot, *Phys. Rev. B* **75**, 085311 (2007).
- [12] T.-R. Shan, B. D. Devine, J. M. Hawkins, A. Asthagiri, S. R. Phillpot, and S. B. Sinnott, *Phys. Rev. B* **82**, 235302 (2010).
- [13] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [14] J. Behler, R. Martoňák, D. Donadio, and M. Parrinello, *Phys. Status Solidi B* **245**, 2618 (2008).
- [15] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [16] N. Artrith, T. Morawietz, and J. Behler, *Phys. Rev. B* **83**, 153101 (2011).
- [17] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, *Phys. Rev. Lett.* **108**, 253002 (2012).
- [18] J. C. Snyder, M. Rupp, K. Hansen, L. Blooston, K.-R. Müller, and K. Burke, *J. Chem. Phys.* **139**, 224104 (2013).
- [19] W. J. Mortier, K. V. Genechten, and J. Gasteiger, *J. Am. Chem. Soc.* **107**, 829 (1985).
- [20] J. Behler, *J. Chem. Phys.* **134**, 074106 (2011).
- [21] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, O. Zilberberg, A. Bergman, M. Rayson, and R. Schneider, *J. Chem. Phys.* **129**, 014109 (2008).
- [22] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Belykin, *J. Chem. Phys.* **125**, 074105 (2006).
- [23] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [24] S. Goedecker, *J. Chem. Phys.* **120**, 9911 (2004).