High accuracy and transferability of a neural network potential through charge equilibration for calcium fluoride

Somayeh Faraji,[1] S. Alireza Ghasemi,[1,*] Samare Rostami,[1] Robabe Rasoulkhani,[1] Bastian Schaefer,[2] Stefan Goedecker,[2] and Maximilian Amsler[3]

[1]*Institute for Advanced Studies in Basic Sciences, P.O. Box 45195-1159, Zanjan, Iran*

[2]*Department of Physics, Klingelbergstrasse 82, 4056 Basel, Switzerland*

[3]*Department of Materials Science and Engineering, Northwestern University, Evanston, Illinois 60208, USA*

(Received 14 November 2016; revised manuscript received 15 January 2017; published 16 March 2017)

We investigate the accuracy and transferability of a recently developed high-dimensional neural network (NN) method for calcium fluoride, fitted to a database of *ab initio* density functional theory (DFT) calculations based on the Perdew-Burke-Ernzerhof (PBE) exchange correlation functional. We call the method charge equilibration via neural network technique (CENT). Although the fitting database contains only clusters (i.e., nonperiodic structures), the NN scheme accurately describes a variety of bulk properties. In contrast to other available empirical methods the CENT potential has a much simpler functional form, nevertheless it correctly reproduces the PBE energetics of various crystalline phases both at ambient and high pressure. Surface energies and structures as well as dynamical properties derived from phonon calculations are also in good agreement with PBE results. Overall, the difference between the values obtained by the CENT potential and the PBE reference values is less than or equal to the difference between the values of local density approximation (LDA) and Born-Mayer-Huggins (BMH) with those calculated by the PBE exchange correlation functional.

DOI: 10.1103/PhysRevB.95.104105

I. INTRODUCTION

Atomistic simulations have become a widespread and integral part of materials research in physics and chemistry, providing a powerful tool to predict materials properties prior to their synthesis. The accuracy of quantities obtained from such calculations however strongly depends on the model employed to describe the atomic interactions. A wide spectrum of methods have been developed, ranging from quasiexact many-electron wave function approaches [1], density functional theory (DFT) calculations [2], and semiempirical quantum mechanical methods [3,4] to empirical, classical potentials [5]. Methods based on quantum mechanics in general allow an accurate prediction of physical and chemical properties of materials. However, the associated high computational cost may prevent their use for target applications with systems containing a large number of atoms. DFT has proven to provide a convenient compromise between accuracy and efficiency for many applications and has thus become the quasistandard in materials science. Nevertheless, its range of application still remains limited to few hundreds of atoms whenever a large number of energy and force evaluations are required, such as in molecular dynamics simulations or in structure prediction. Numerous efforts have been therefore made to develop empirical potentials which preserve the accuracy of *ab initio* methods at reduced computational cost.

A variety of force fields (FF) have been proposed to model ionic materials, e.g., the Buckingham potential [6], which are several orders of magnitude faster than *ab initio* methods. In a classical FF, the total energy is a sum of atomic energies arising from physically motivated analytic functions to model the atomic interactions which contain parameters often fitted to experimental or (more accurate) computational reference data. Most force fields take into account classical electrostatic interactions, but the charges giving rise to these interactions are often fixed, limiting the accuracy of such FFs. Therefore, polarizable FFs [7] and charge equilibration methods [8] have been developed to overcome these limitations. In some of the most advanced charge equilibration schemes, such as the charge optimized many body (COMB) potentials [9] and the ReaxFF [10] force field, the atomic charges depend on the environment via some analytic functions. These potentials have been successfully applied to study materials and other condensed matter systems [11–13]. Nevertheless, the basic shortcoming of all standard FFs, namely that the optimal analytical functional form is unknown, still remains in these FFs.

In contrast to conventional FFs, machine learning based potentials do not require a predefined, physically motivated functional form to model the atomic interactions. Such schemes have recently attracted considerable attention since they are capable of reproducing a large number of *ab initio* reference data with a very high accuracy. High dimensional artificial NN [14] (ANNs) and Gaussian approximation potentials [15] (GAP) are two of the most promising approaches and have been successfully used in many studies involving atomistic simulations [14–21]. In these methods, the total energy is again a sum of atomic or pairwise energies, which are however directly obtained from a machine learning process. In fact, the majority of these methods are based on purely mathematical models and do not account for any sort of physical or chemical principles. However, there are several attempts to combine machine learning techniques with physically motivated energy functionals such as the Coulombic interaction for zinc oxide bulk [22] and the bond order potential for small silicon clusters [23]. Although such potentials give excellent results for systems with chemical environments comparable to those used during the training process, they often fail to describe structures dissimilar to all training data: if a potential is

*aghasemi@iasbs.ac.ir

trained with molecular structures, it cannot be used for bulk environments with periodic boundary conditions.

A new approach was recently proposed by Ghasemi *et al.* [24], in which a short range intermediate physical quantity, namely an environment dependent atomic electronegativity, is obtained from an ANN process while the total energy is computed from a charge equilibration process. The method is thus called charge equilibration via neural network technique (CENT). By allowing the electronic charge to distribute itself in an optimal way in the system, both neutral and charged systems can be equally described with a high accuracy [24].

In this work we train the CENT potential to calcium fluoride (CaF_2) clusters and investigate its transferability to systems with different boundary conditions by comparing various bulk and surface properties to *ab initio* results. CaF_2 is representative of the fluoride-structured halides, an important class of ionic materials with numerous technological applications [25–28]. As a chemically inert, wide band gap semiconductor, it can be grown in high-grade single crystals with excellent optical transmission properties over a wide energy range and exhibits a very high elastic compressibility [29–32]. Furthermore, CaF_2 is also well known for being a fast-ion conductor [33,34]. Here we evaluate the accuracy of the CENT potential for bulk properties such as enthalpy differences between polymorphs at high pressure, surface geometries, vacancy formation energies, and lattice dynamics. Our results show that the CENT potential trained on clusters can successfully predict crystalline properties with accuracies close to DFT results. In fact, the difference between the CENT results and the DFT reference methods used for the fitting process is less than or comparable to the differences between both popular DFT exchange correlation functionals and the empirical BMH potential.

II. METHODS

A. Feed-forward artificial neural network

ANN techniques are inspired by biological neural structures and provide a framework that can be used to approximate mathematical functions. In an ANN there are many nodes which are connected through so-called weights. These nodes are commonly distributed in layers, organized as an input, (multiple) hidden and an output layer. In atomistic simulations, ANNs are used to calculate total energies, where some representation of the atomic coordinates is passed into the input layer, and the output layer, which only contains a single node, provides the desired scalar quantity (energy). Such ANNs usually work in a feed-forward fashion where signals are only transmitted in one direction along the layers. In high dimensional ANN [14,35] potentials, the total energy is the sum of atomic energies, each obtained from an ANN process. A major disadvantage of currently available ANN potentials is that they can only reproduce bulk properties for structures that have been used in the fitting data base, as demonstrated by Eshet *et al.* [36].

B. Charge equilibration via neural network techniques

In 2015, Ghasemi *et al.* proposed a new method to generate interatomic potentials for ionic systems based on NN by interpolating the electronegativity rather than the total energy [24]. Here we provide a concise summary of its essential concepts. In this scheme, the total energy of a system containing N atoms is given by

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

where E_i^0 are the energies of individual, isolated atom i, χ_i is the environment dependent atomic electronegativity of atom i whose functional dependence is determined by the ANN, q_i are the atomic charges, J_{ii} is the element-dependent atomic hardness [37] of atom i, and $\rho(\mathbf{r})$ is the charge density of the system, which in our implementation is given by a superposition of spherical Gaussian functions centered at the atomic positions $\mathbf{r}_i$, each normalized to the corresponding atomic charge q_i. Thus, Eq. (1) can be rewritten as follows:

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

where $\gamma_{ij} = \frac{1}{\sqrt{\alpha_i^2+\alpha_j^2}}$ and α_i are the widths of the Gaussian atomic charge densities, and r_{ij} is the distance between the atoms i and j. The atomic charges q_i are implicitly environment dependent through the atomic electronegativity χ_i. As a consequence, long distance charge transfers for a given system are possible while the total charge of the system is conserved. In the charge equilibration process the atomic charges are distributed such that the electrostatic energy is minimized. This is done by differentiating U_{tot} with respect to the atomic charges using Lagrange multipliers, which constrain the total charge to a constant value, leading to a system of linear equations of the form $\tilde{A}\mathbf{Q} = -\chi$, where $\tilde{A}$ is a $(N+1)\times(N+1)$ matrix, and $\mathbf{Q}$ and χ are $(N+1)$-dimensional vectors. In the case of free boundary conditions and a small number of atoms the system of linear equations is solved directly, whereas for large systems it is solved iteratively. In contrast, for periodic systems it is always solved iteratively irrespective of the system size. Since the system of linear equations is well-conditioned the total number of iterations to reach sufficient convergence never exceeds 20 based on extensive tests. When performing molecular dynamics simulations or local geometry relaxations with relatively small atomic displacements in each time step, the initial guess for the first iteration can be taken from the atomic charges obtained in the previous step, hence significantly reducing the number of iterations to reach convergence.

The input layer of a high dimensional NN is a set of symmetry functions $\{\mathbf{G}_i\}$ to describe the chemical environment of each atom. We use the symmetry functions proposed in Ref. [35], which have the advantage that the array representing the chemical environments is invariant under translation and rotation of the structure. The adopted functional form of the radial and angular symmetry functions used in this work is

described in Ref. [24]. We used in total 70 symmetry functions, 16 and 54 for the radial and angular part, respectively, parametrized according to Ref. [38]. The cutoff radius for the symmetry functions were well tested and a value of 6.35 Å was sufficient to generate accurate potentials. The appropriate values for Gaussian widths are 0.53 Å for both calcium and fluorine atoms. The atomic ANNs contain two hidden layers with three nodes each. Similar to Ref. [24], we found that such a small number of nodes in the hidden layers is sufficient to reduce the root mean square error to less than 0.33 mHa per atom upon validation tests (see Sec. II C). Furthermore, the optimal values found for the atomic hardnesses of calcium and fluorine are 0.01 and 0.02 in atomic units, respectively.

Due to the well-defined functional form of the individual terms in the ANNs, an analytic expression of the atomic forces can be derived as follows:

$$\mathbf{F}_j = -\sum_{i=1}^{N}\left(q_i \frac{\partial \chi_i}{\partial \mathbf{r}_i}\right) + \sum_{i>j}^{N} q_i q_j \frac{\partial V_{ij}}{\partial \mathbf{r}_j}, \tag{3}$$

where $V_{ij} = \frac{erf(\gamma_{ij})}{r_{ij}}$. The first term in Eq. (3) is calculated during the neural network process based on the chain rule for differentiation. In the case of fully periodic systems the Hartree energy in Eq. (1) and its derivative, i.e., the second term in Eq. (3), cannot be directly computed and is evaluated in Fourier space. To calculate the Hartree energy for slab systems, e.g., to evaluate surface energies, the P^3D method [39] is employed which is both efficient and accurate while preserving the original boundary condition. In the P^3D method, the Poisson equation is solved by expanding the charge density and potential in terms of plane waves in the periodic directions and by using finite elements in the nonperiodic dimension.

C. Preparation of the training data

The reference calculations to construct the ANN potential were performed using DFT with the Perdew-Burke-Ernzerhof (PBE) [40] exchange-correlation functional as implemented in the FHI-aims code [41]. A set of reference structures consisting of 2800 charge-neutral CaF_2 clusters with sizes ranging from 24 to 99 atoms was used to train the ANN potential. These structures were not generated all at once. Instead, the final training data set was built up iteratively since it is of great importance to train a NN potential on a set of diverse structures. Therefore, we first generated a smaller set of reference data points with random structures to train a preliminary ANN potential. Then, using this ANN potential together with the minima hopping global geometry optimization method [42,43] (MHM), the low energy regions of the (preliminary) energy landscape were explored. Local geometry optimizations within the MHM were performed with the recently developed and highly efficient stabilized quasinewton minimizer [44] to reduce computational cost. The resulting low energy structures from these MHM runs were carefully filtered to avoid duplicate structures and to ensure a large diversity in structural motifs within the augmented data set, a task performed by comparing structures with the fingerprint method proposed in Ref. [45]. Thereby, a measure of similarity is given by the configurational distance $d(p,q)$ between two structures p and q according to

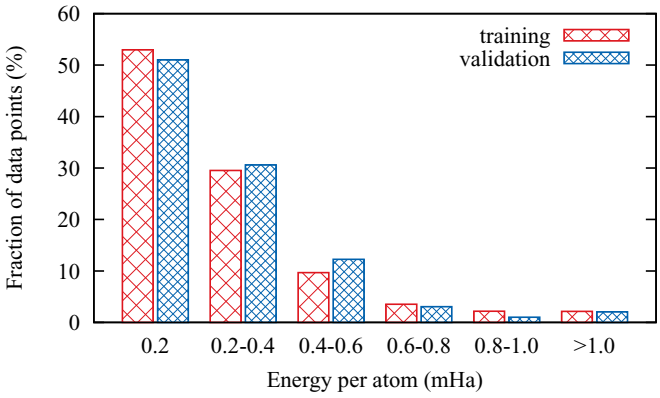

FIG. 1. Distribution of the errors in the total energies per atom with respect to reference DFT calculations for all data points in the training and validation sets. The number of data points are normalized such that the total number of data points gives 100%. In total, the training and validation sets contain 2379 and 419 structures, respectively.

$$d(p,q) = min_P\left(\sum_{k}^{N}\left|\mathbf{G}_k^p - \mathbf{G}_{P(k)}^q\right|^2\right)^{1/2}, \tag{4}$$

where P is a permutation function to match an atom k in structure p with atom $P(k)$ in structure q. The optimal permutation P which minimizes the configurational distance

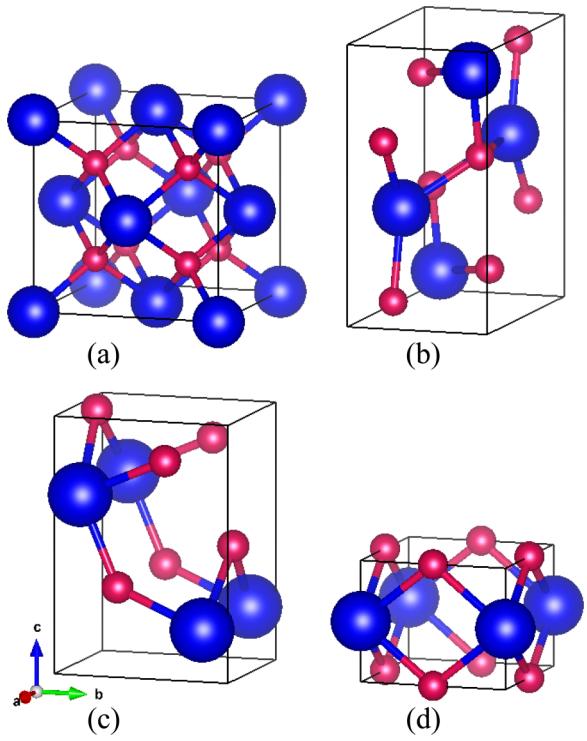

FIG. 2. Unit cell of the four different CaF_2 polymorphs: (a) cubic fluorite ($Fm\bar{3}m$), (b) $P4/mmm$, (c) $Pmc2_1$, and (d) $Pnma$. The blue (large) and purple (small) spheres denote Ca and F atoms, respectively.

TABLE I. Lattice constants (Å), bulk moduli B_0 (GPa), cohesive energies E^f (eV/atom), and the energy differences E'^f (eV/atom) with respect to cubic CaF_2 of four CaF_2 phases at ambient condition. The entries denoted with BMH, CENT, PBE, and LDA contain results from the present work, whereas the other results were taken from literature, either from calculations (Theory) or experiments (Expt.). The values in parentheses next to the BMH, CENT, and LDA values indicate the deviation from PBE results.

Phase	Method	a_0	b_0	c_0	B_0	E^f	E'^f
$Fm\bar{3}m$	BMH	5.460 (−0.055)			88.30 (10.34)	−9.048 (−3.246)	
	CENT	5.483 (−0.032)			85.40 (7.44)	−5.835 (−0.033)	
	PBE	5.515			77.96	−5.802	
	LDA	5.332 (−0.183)			101.23 (23.27)	−6.379 (−0.577)	
	Expt.	5.466 [49]			87 [49]		
	Expt.	5.463 [50]			82 [51]		
	Theory	5.5252 [52]			77 [52]	−4.290 [53]	
	Theory	5.4872 [54]			79.54 [54]		
	Theory	5.509 [55]			79 [55]		
	Theory	5.579 [56]			77.06 [56]		
$Pnma$	BMH	5.979 (−0.033)	3.607 (−0.040)	6.963 (−0.122)	96.52 (12.93)	−8.978 (−3.229)	0.070
	CENT	5.998 (−0.014)	3.608 (−0.039)	7.062 (−0.023)	104.28 (20.69)	−5.799 (−0.050)	0.036
	PBE	6.012	3.647	7.085	83.59	−5.749	0.053
	LDA	5.797 (−0.215)	3.482 (−0.165)	6.846 (−0.239)	111.34 (27.75)	−6.347 (−0.598)	0.032
	Expt.	5.63 [49]	3.49 [49]	7.15 [49]	75 [52]		
	Expt.	5.96 [57]	3.58 [57]	7.03 [57]	70.92 [54]		
	Expt.	6.018 [58]	3.614 [58]	7.023 [58]	66 [55]		
	Theory	6.0274 [52]	3.6682 [52]	7.0413 [52]			
	Theory	5.9826 [54]	3.6303 [54]	7.0499 [54]			
	Theory	6.020 [55]	3.660 [55]	7.108 [55]			
	Theory	5.9909 [56]	3.6025 [56]	7.0717 [56]	86.78 [56]		
$Pmc2_1$	BMH	3.511 (−0.030)	4.069 (−0.096)	5.999 (0.019)	79.347 (7.41)	−8.778 (−3.160)	0.270
	CENT	3.516 (−0.025)	4.035 (−0.130)	5.925 (−0.055)	82.998 (11.06)	−5.659 (−0.041)	0.176
	PBE	3.541	4.165	5.98	71.933	−5.618	0.184
	LDA	3.395 (−0.146)	4.048 (−0.117)	5.785 (−0.195)	93.552 (21.62)	−6.170 (−0.552)	0.208
$P4/mmm$	BMH	3.028 (0.008)	3.648 (−0.099)	3.728 (−0.019)	85.513 (9.99)	−8.724 (−3.164)	0.324
	CENT	2.926 (−0.094)	3.687 (−0.060)	3.687 (−0.060)	95.338 (19.81)	−5.637 (−0.077)	0.198
	PBE	3.020	3.747	3.747	75.528	−5.560	0.243
	LDA	3.621 (0.601)	3.621 (−0.126)	2.927 (−0.820)	98.724 (23.20)	−6.133 (−0.573)	0.246

is obtained through the Hungarian or Munkres algorithm [46]. The complete process of performing MHM runs, filtering the structures, and then retraining the ANN potential was repeated several times. The structures are thus reasonably diverse and contain various morphologies. They are disordered structures similar to amorphous solids rather than containing ordered crystalline motifs. In fact, they do not resemble any well-known CaF_2 crystalline phases such as those illustrated in Fig. 2. Also, they do not contain vacancylike structures or structures containing void regions. The surface of the structures are curved and do not resemble surfaces which are cut out of bulk crystalline CaF_2 structures along different crystallographic planes. The final dataset contained 2800 clusters, and the following approach was used to validate the training of CENT: 85% of the total data were randomly selected as training points to obtain a CENT potential, which was then used to predict the energies of the remaining 15% of the structures for validation, a common procedure to detect potential overfitting. This process was repeated ten times, based on which a RMS error of 0.33 mHa/atom was achieved with a standard deviation of ±0.04 mHa/atom. Figure 1 shows a representative error distribution for one of these training processes. The error in the validation data is only slightly larger compared to the errors in the training data, and only a fraction of less than 3% of the total data results in an error larger than 1.0 mHa/atom. In machine learning techniques, one way of avoiding overfitting is early stopping, which we have also employed in our trainings. In addition, since we do not use the information of atomic forces as training data, the predicted forces are closely monitored to analyze the accuracy: in the case of overfitting the error in the forces would increase while the error in the energies continues to decrease.

III. NUMERICAL RESULTS

In this section we examine the transferability of the CENT potential. Trained exclusively on a dataset consisting of cluster reference structures as described in Sec. II C, CENT was used to compute various bulk and surface properties. The results are compared to DFT results to assess the accuracy of the CENT potential. Furthermore, to provide a comparison between the results of the CENT potential and a FF, the BMH form of

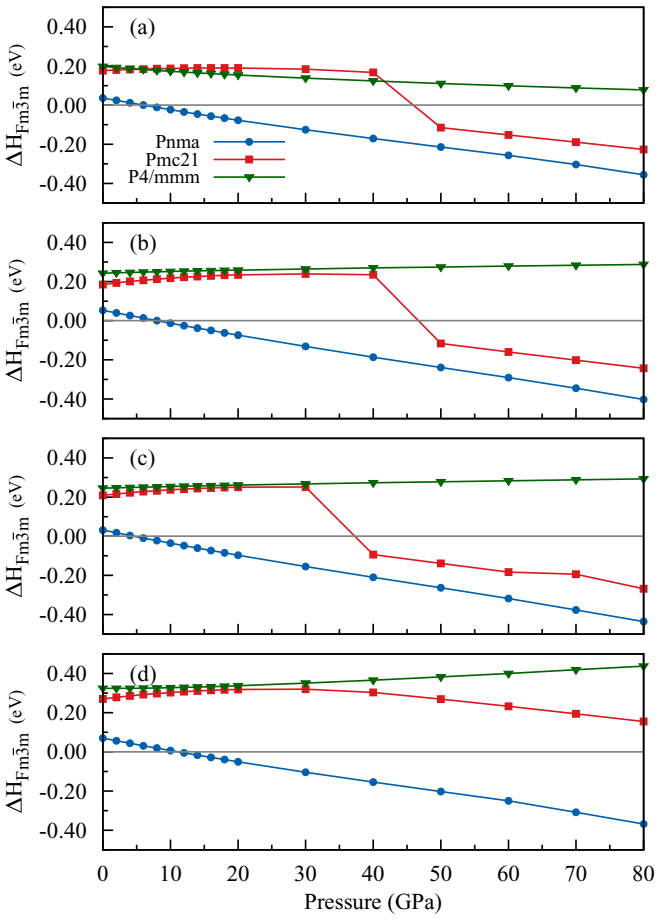

FIG. 3. Enthalpy differences per atom with respect to the $Fm\bar{3}m$ phase as a function of pressure for three different polymorphs, using CENT (a), PBE (b), LDA (c), and BMH (d).

Buckingham potential was used as implemented in the GULP code [47] and parametrized according to Ref. [48]. Note that the BMH potential was originally fitted to bulk properties and is thus expected to perform well for periodic systems.

A. Bulk properties

As a first benchmark we computed the bulk properties of four well-known CaF_2 crystalline phases: the cubic fluorite ($Fm\bar{3}m$) and three orthorhombic cotunnite $PbCl_2$-type structures with space groups $Pnma$, $Pmc2_1$, and $P4/mmm$ (see Fig. 2). Table I contains the lattice constants, bulk moduli, formation energies, and the energy differences with respect to the most stable $Fm\bar{3}m$ phase from CENT, BMH, and DFT calculations together with experimental data whenever available. The formation energies obtained with the CENT potential are close to the reference training method, i.e., PBE. Based on the average errors, which are around 1%, 11%, and 56% for CENT, LDA, and BMH, respectively, we conclude that the error of CENT results with respect to PBE is much smaller than the differences between LDA or BMH and PBE. Furthermore, the standard deviations between the predicted CENT, LDA, and BMH energy differences and the PBE values are 0.028, 0.018, and 0.069 (eV/atom), respectively, indicating that the energies from the CENT potential and LDA correlate

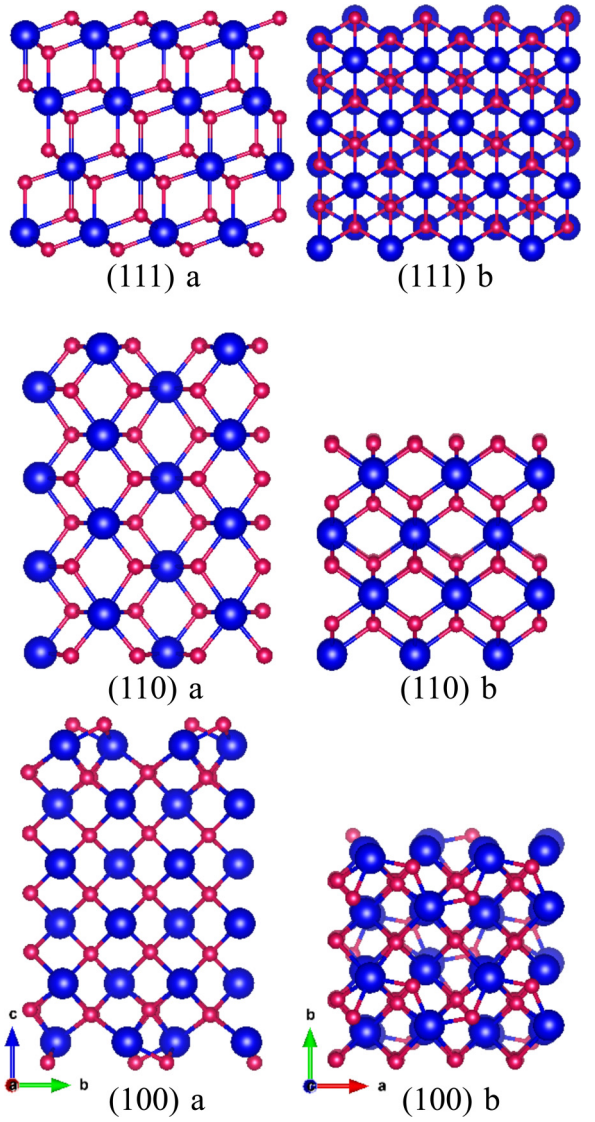

FIG. 4. Three different low-index surfaces of CaF_2 shown from the side along the a axis (denoted by a) and the top along the c axis (denoted by b) after a local relaxation. The blue (large) and purple (small) spheres denote Ca and F atoms, respectively.

well with the PBE results. The lattice constants and the bulk moduli obtained with CENT and BMH are considerably closer to the PBE results than those obtained with LDA.

Pressure induced phase transitions in CaF_2 have been reported and are well studied both experimentally and theoretically (see Refs. [49,51,52,55,57,59–62]). Cubic $Fm\bar{3}m$ is known to transform into the orthorhombic $Pnma$ phase at a pressure between 8 and 10 GPa, accompanied by an increase of the coordination number of Ca from 8 to 9. Xray diffraction and Raman spectroscopy have shown that this high pressure phase is stable up to 49 GPa at room temperature [49,51]. In Fig. 3, the enthalpy of the $Pnma$ phase together with the $Pmc2_1$ and $P4/mmm$ structures are plotted with respect to the $Fm\bar{3}m$ phase as a function of pressure. Our calculations show that the phase transition from $Fm\bar{3}m$ to $Pnma$ occurs at 4.6, 8, 6, and 10 GPa, in LDA, PBE, CENT, and BMH, respectively, all values close to experimental measurements. The transition is

TABLE II. Thickness, atomic relaxation, and rumpling (all in Å) of the (111), (110), and (100) surfaces. The quantities in parentheses indicate differences with respect to PBE results.

Surface	Method	Thickness	Relaxation	Rumpling
(111)	BMH	11.100(−0.065)	−0.005(−0.024)	−0.002(−0.054)
	CENT	11.213(0.048)	0.033(0.014)	0.052(0.049)
	PBE	11.165	0.019	0.003
	LDA	10.778(−0.387)	0.010(−0.009)	0.004(0.001)
(110)	BMH	19.435(−0.091)	−0.176(−0.089)	−0.123(−0.174)
	CENT	19.368(−0.158)	−0.130(−0.043)	0.079(0.028)
	PBE	19.526	−0.087	0.051
	LDA	18.879(−0.647)	−0.102(−0.015)	0.036(−0.015)
(100)	BMH	15.251(−0.436)	−0.490(0.020)	−0.019(−0.136)
	CENT	15.543(−0.144)	−0.425(0.085)	0.121(0.004)
	PBE	15.687	−0.510	0.117
	LDA	15.068(−0.619)	−0.429(0.081)	0.041(−0.076)

accompanied by a collapse of the unit cell volumes by 13.6%, 8.16%, 5.7%, and 8.81% for LDA, PBE, CENT, and BMH, respectively, comparable to the experimental value of 8.3% given in Ref. [62].

B. Surfaces

Next, we investigated the accuracy of the CENT potential for surface properties, which have been the subject of extensive experimental and theoretical studies in the past [63–69]. Puchin *et al.* [63] for example performed a systematic theoretical investigation of various surfaces with respect to their relative stabilities and the effect of surface relaxation/reconstruction on band gaps. Sicher *et al.* performed extensive structure prediction calculations with the MHM [42,43] to compare the stability of hydrated (100) and (111) surfaces, giving insight into different experimentally observed morphologies in CaF_2 nanocrystals (not yet published).

In the present work we considered three low-index surfaces, namely (111), (110), and (100), which are illustrated in Fig. 4. Equilibrium lattice parameters of 5.332, 5.515, 5.483, and 5.460 Å were used for LDA, PBE, CENT, and BMH calculations, respectively. The BMH, LDA, and PBE calculations were performed using slabs in a supercell including a vacuum region of more than 14 Å to avoid interactions between periodic images. No vacuum region was required for the CENT calculations since the P^3D method was used to solve the Poisson equation with surface boundary conditions. Polarized surfaces require additional care due to nonvanishing dipole interactions. The (111) and (110) surfaces are dipole free, whereas (001) is a polar surface (so-called Tasker-type-III structure) consisting of oppositely charged planes. Therefore, 50% of the fluorine atoms are moved from the top of the slab to the bottom to suppress the dipole moment.

After preparing the slabs from the ideal bulk structure, all atoms were fully relaxed to release the strain induced by the broken bonds at the surfaces. The amount of this relaxation can be defined as the mean difference between the first and second surface layers before and after relaxation. The same quantity, but with respect to the change of the distance between the second and third sublayers, is called rumpling. The effect of the relaxation on the structure is shown in Fig. 4, and all numerical quantities are summarized in Table II. Our calculations are qualitatively in good agreement with previous reports. A detailed discussion of each of the three surfaces is given below.

(111) Surface: We modeled the (111) surface with 12 atomic layers, which can also be interpreted as four layers of combined F-Ca-F triple layers. The perfectly flat (111) surface after cleaving from the bulk is F^- terminated, forming a hexagonal array on the surface. The symmetry elements of the surface are inversion and mirror planes perpendicular to the surface, as well as the translational in-plane symmetries of the slab. The ions Ca^{2+} and F^- in the second and third atomic layers have a threefold symmetry, which is preserved even after relaxation. The thickness of the complete 12-layer slab, defined as the distance between the outmost atoms of the whole slab, is listed in Table II after a local geometry relaxation, indicating that the value obtained by the CENT potential is very close to the PBE result. All the results show that the outmost sublayers

TABLE III. Surface energies (in J/m^2) for the (111), (110), and (100) surfaces.

	Present work				Others	
	BMH	CENT	PBE	LDA	Expt.	Theory
(111)	0.538(0.161)	0.539(0.162)	0.377	0.563(0.186)	0.45 [71] 0.45–0.5 [73]	0.39 [72], 0.437 [32], 0.52 [65], 0.467 [63] 0.39 [74], 0.497 [75], 0.476 [64]
(110)	0.870(0.276)	0.763(0.169)	0.594	0.815(0.221)		0.61 [72], 0.717 [32], 0.82 [65], 0.819 [63] 0.969 [75], 0.760 [64]
(100)	1.964(1.070)	1.032(0.138)	0.894	1.195(0.301)	0.9–0.95 [73]	0.84 [72], 0.957 [32], 1.189 [63]

move only very little along the nonperiodic z direction upon relaxation, which is in good agreement with the results of Jockisch *et al.* [70]. While in the unrelaxed structure the outmost F and Ca layers are separated by 0.792, 0.796, 0.770 Å in CENT, PBE, and LDA, respectively, their distances increase to 0.824, 0.815, and 0.779 Å upon relaxation. However, in the BMH results, this distance decreases from 0.766 Å to 0.761 Å. In the layers further away from the surface, i.e., deeper in the slab, this relaxation effect vanishes rapidly. The relative distances between units of the F-Ca-F triple layer however increases when moving away from the surface towards the center of the slab, in agreement with earlier theoretical results.

(110) Surface: For the (110) surface we considered a slab with eight atomic layers. In contrast to the (111) surface where atoms only relax within the c direction, a reconstruction within the ab plane is observed for the (110) surface. The outmost sublayer of the slab contains both Ca and F atoms: While the F atoms move outwards away from the slab upon relaxation, the Ca atoms move inwards. Together, the F and Ca in the outer layer undergo a net inward relaxation of 0.176, 0.130, 0.087, and 0.102 Å for BMH, CENT, PBE, and LDA, respectively.

(100) Surface: We used a model with 12 atomic layer for the (100) surface. Overall, the amount of relaxation within this surface is considerably larger compared to the former two surfaces, accompanied by a strong reconstruction as illustrated in Fig. 4. The outmost layer of F atoms is pushed inward towards the center of the slab by 0.490, 0.425, 0.510, and 0.429 Å in BMH, CENT, PBE, and LDA, respectively.

The surface energy $E_f^{(hkl)}$ is defined as the work required to separate a crystal into two parts along a certain plane (hkl). In order to assess the accuracy of the CENT potential in predicting the stability of the three surfaces, their values were calculated according to

$$E_f^{(hkl)} = \frac{1}{2A}\left(E_s^{(hkl)} - nE_b\right), \tag{5}$$

where A is the surface area, $E_s^{(hkl)}$ is the total energy of the equilibrium structure of the slab, E_b is the bulk total energy per f.u., and n is the number of f.u. in the slab. The factor $\frac{1}{2}$ accounts for the presence of two surfaces at either side of the slab.

According to our DFT and CENT results, the surface energy of the (111) surface is roughly half of the (100) surface. The energy of the (110) surface is consistently higher than the (111) and lower than the (100) surface. All potentials are qualitatively in agreement with other theoretical and experimental studies as shown in Table III. However, there are significant quantitative discrepancies depending on which potentials or DFT functionals are used. In particular, BMH significantly overestimates the (100) surface energy, which can be readily attributed to the fixed charges used in the potential and the polar character of the surface slab, whereas CENT on the other hand performs much better due to the charge equilibration process. Overall, the values obtained with CENT are within the error bars spanned by the various theoretical results and reasonably close to the PBE values.

TABLE IV. Defect formation energies E_f^A (eV) of F and Ca centers calculated by BMH, CENT, PBE, and LDA.

Atom	BMH	CENT	PBE	LDA
Ca	17.942	9.671	15.526	17.065
F	3.377	6.001	8.735	9.432

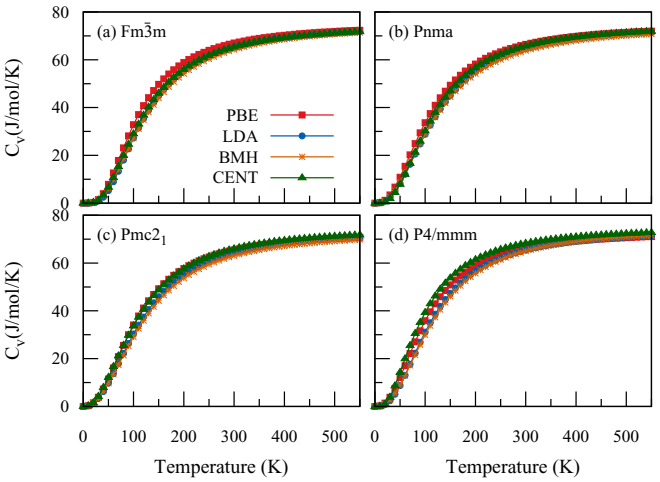

FIG. 5. Heat capacity of CaF_2 in four different phases: (a) cubic fluorite ($Fm\bar{3}m$), (b) ($Pnma$), (c) $Pmc2_1$, and (d) $P4/mmm$.

C. Defects

Understanding and controlling the nature and concentration of defects in materials, also referred to as defect engineering, is an important task to tune materials properties, and atomistic simulations are often used to provide a better theoretical insight [76,77]. The so-called F-centered defect is often encountered in CaF_2, where an electron is trapped in an anion vacancy. The defect formation energies can be computed using the following equation

$$E_f^A = E(A) + E(V) - E(P), \tag{6}$$

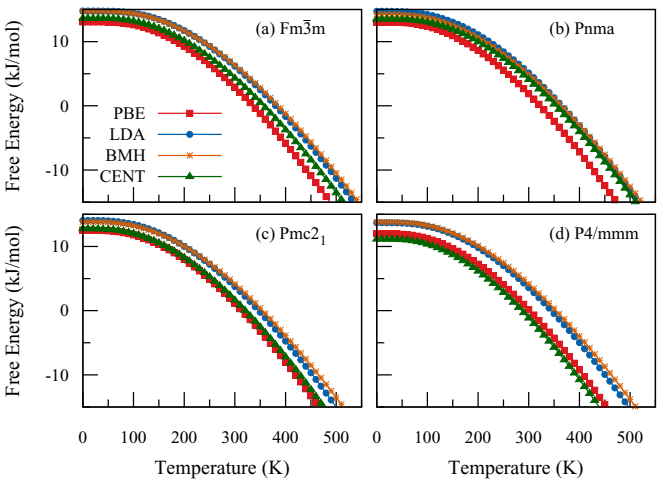

FIG. 6. Free energy of CaF_2 in four different phases: (a) cubic fluorite ($Fm\bar{3}m$), (b) ($Pnma$), (c) $Pmc2_1$, and (d) $P4/mmm$.

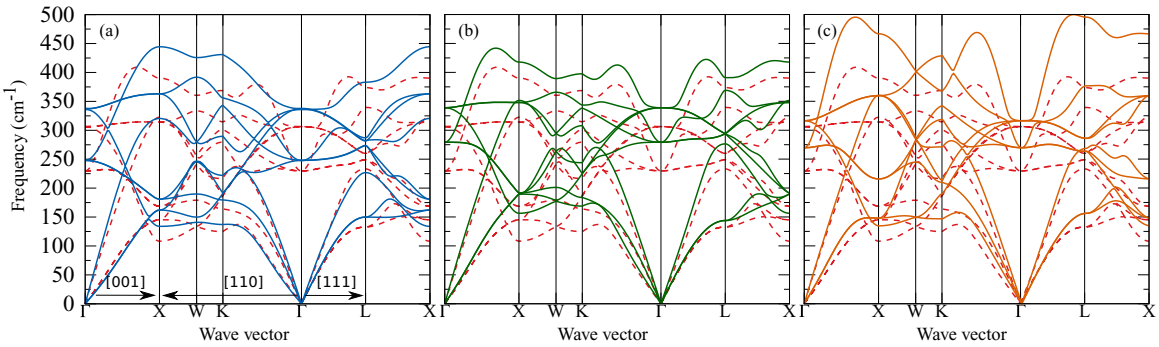

FIG. 7. Phonon dispersion of CaF_2 in the cubic fluorite ($Fm\bar{3}m$) structure. The results from the CENT, LDA, and BMH are shown in the panels (a), (b), and (c), respectively. The red (dashed) lines correspond to the reference data obtained with PBE.

where $E(A)$ is the energy for an isolated Ca or F atom and $E(V)$ and $E(P)$ are the total energies of the bulk with and without vacancy, respectively.

We performed defect calculations of bulk CaF_2 with 3×3 supercells containing 324 atoms with the BMH and CENT potentials and DFT with both the LDA and PBE exchange correlation functionals. A comparison of the defect formation energies listed in Table IV shows that the formation of a F vacancy is energetically more favorable than a Ca vacancy in all four potentials. However, unlike all other quantities where the CENT results were rather accurate, the values of E_f^A using CENT differ significantly from LDA and PBE. Although this outcome might be somewhat discouraging, predicting defect energies can be considered one of the most difficult tests to pass. Still, the CENT potential outperforms other force fields with respect to the accuracy of these values [78,79]. CENT could be further improved by specifically including such defect data points in the training set.

D. Lattice dynamics

Up to now we had mainly assessed the accuracy of properties that can be derived from computing total energies. Here we analyze how the CENT potential performs when considering higher order derivatives of the energy landscapes in lattice dynamics calculations. Phonon calculations were carried out with the Phonopy package [80] to obtain the phonon dispersion, free energies, and heat capacities. All structures were relaxed with a tight convergence criterion of 1×10^{-4} eV Å^{-1} before atomic displacements with an amplitude of 0.01 Å were created in $3 \times 3 \times 3$ supercells. k-point meshes of $4 \times 4 \times 4$ were used in all DFT calculations, and q meshes of $45 \times 45 \times 45$ were used for the Brillouin zone integration.

The heat capacities and free energies at constant volume (C_V) as a function of temperature are shown in Figs. 5 and 6, respectively. Overall, the results of the CENT potential and PBE calculations agree very well. In the case of $Fm\bar{3}m$, the LDA, PBE, BMH, and CENT calculations results in C_v values of 65.107, 66.999, 64.266, and 65.346 (J/mol/K) at room temperature, respectively, which are very close to the experimental result of 67.11 (Jmol^{-1}K^{-1}) [81].

Figure 7 shows the phonon band structure of CaF_2 in the $Fm\bar{3}m$ phase computed from the forces of the BMH and CENT potentials and DFT with the LDA and PBE functionals. The acoustic modes are qualitatively similar in all methods, however, quantitatively the results from the CENT potential agree better with PBE than with LDA and BMH. Furthermore, due to the crystal symmetry and the degeneracy of the transversal branches, there are only six phonon branches along the [100] and [111] directions, whereas along the [110]

TABLE V. The frequencies of the maximum longitudinal optical mode ($LO_{\max}$), minimum transversal acoustic modes ($TA_{\min}$), and the separation of longitudinal and transversal acoustic modes at the K and L points ($LA - TA$) calculated by BMH, CENT, PBE, and LDA. All values are in cm^{-1} and the numbers in parentheses are the differences with respect to PBE.

	method	Γ	X	W	K	L
	BMH	0	135 (26)	149 (16)	147 (23)	156 (24)
	CENT	0	134 (25)	141 (8)	138 (14)	150 (18)
$TA_{\min}$	PBE	0	109	133	124	132
	LDA	0	157 (48)	177 (44)	169 (45)	143 (11)
	BMH	317 (11)	466 (76)	403 (43)	429 (60)	495 (122)
	CENT	336 (30)	445 (55)	426 (66)	430 (61)	383 (10)
$LO_{\max}$	PBE	306	390	360	369	373
	LDA	339 (33)	417 (27)	389 (29)	398 (29)	391 (18)
	BMH	0			43 (4)	107 (6)
	CENT	0			42 (3)	77 (24)
$LA - TA$	PBE	0			39	101
	LDA	0			13 (26)	133 (32)

TABLE VI. Elastic constants C_{ij}, Young modulus E, shear modulus G, and Poisson ratio ν for CaF_2 of the $Fm\bar{3}m$ phase. All values, except ν, are in units of GPa.

	Present work				Previous work	
	BMH	CENT	PBE	LDA	Expt.	Theory
C_{11}	168.8	174.7	153.5	183	174 [83], 164 [84]	168 [85], 159 [86], 191 [86]
C_{12}	48.2	39.2	40.6	60.3	56 [83], 44 [84]	48 [85], 41 [86], 61 [86]
C_{44}	45.6	22.7	27.3	35.8	35.93 [83], 34 [84]	40 [85], 30 [86], 40 [86]
E	147.4	94.2	95.2	116.4	107.5 [87]	119.85 [85]
G	51.0	35.8	36.7	44.5	40.7 [87], 41.1 [87]	47.08 [85]
ν	0.2	0.3	0.3	0.3	0.28 [51]	

direction the degeneracy is resolved and all nine phonon branches are split. For a quantitative comparison of the three methods we list in Table V the maximum and minimum values of frequencies which are related to the longitudinal optical modes ($LO_{\max}$) and transversal acoustic modes ($TA_{\min}$) as well as the separation of the longitudinal and transversal acoustic modes at the K and L points at the zone boundaries. Note that the LO/TO splitting was neglected.

E. Elastic properties

The final materials properties that we examined are the mechanical quantities of CaF_2 in the bulk $Fm\bar{3}m$ phase. In order to characterize the elastic behavior at equilibrium we calculated the single crystal second order elastic stiffness constants C_{ij}, the shear modulus G, the Young's modulus E, and the Poisson's ratio ν using the ELASTIC software package [82]. The results obtained with BMH, CENT, LDA, and PBE are listed in Table VI. Although the values of these properties are pressure and temperature dependent we list the experimental and theoretical values in Table VI at ambient conditions. Due to symmetry there are only three independent elastic constants for cubic crystals: C_{11}, C_{12}, and C_{44}. Comparing the elastic constants in the principal direction C_{11} shows that it is reproduced reasonably well by BMH and CENT, and the differences between CENT and BMH potentials and PBE are less than that between LDA and PBE. The CENT shear constant in the main diagonal of the elasticity matrix C_{44} is also in good agreement with PBE. Finally, the off-diagonal element C_{12} obtained by CENT has a marginal deviation and it differs from that of PBE by merely 3.4%, significantly less than the difference between both LDA and BMH with PBE. Overall, the elastic constants obtained with the LDA functional are larger than those of PBE, a direct consequence of the well-known overbinding behavior of the LDA functional, also leading to smaller lattice constants, a larger bulk modulus, and larger elastic constants. Similarly, the BMH potential also significantly overestimates the elastic constants with respect to PBE. The calculations of G, E, and ν were performed within the Hill approximation since the arithmetic average of the Reuss and Voigt limits have been shown to give good estimates of the actual values. Based on the results presented in Table VI we conclude that, similar to the elastic constants, the CENT potential describes the elastic properties of CaF_2 with an acceptable level of accuracy. Furthermore, the values of the Poisson's ratio obtained by BMH, CENT, PBE, and LDA are in the range of $\frac{1}{5} \leqslant \nu \leqslant \frac{1}{2}$, which is consistent with experimental observations.

IV. DISCUSSION AND CONCLUSIONS

Accurate force fields are urgently called for to efficiently model large systems with thousands or tens of thousands of atoms. The recent development of a new generation of force fields based on machine learning techniques is an important step in this direction. Since such ANN potentials are not based on physically motivated functional forms, the common perception up to now was that they have limited transferability, i.e. that they cannot reproduce highly accurate results for structural motifs which are not contained in the training data set. The recently developed CENT method however differs significantly from previously available ANN potentials. By aiming to predict a physically motivated quantity that is highly local using ANN, namely the environment dependent atomic electronegativity, a good transferability from clusters to extended bulk systems can be readily achieved by fitting the ANN to a data set of small molecules only.

In this paper we have examined the transferability of the CENT method by generating a potential for CaF_2 based on clusters and evaluating its performance on various bulk properties. For each property that we investigated we compared CENT to DFT results obtained with the LDA and PBE functionals as well as to the empirical BMH potential. With the exception of vacancy formation energies, all results show that CENT can reproduce the PBE results close to density functional accuracy. This is more impressive considering that the largest clusters used in our training data set consist of only 99 atoms with no structural motifs predominant in crystalline structures. The BMH potential on the other hand gives good results for bulk structural properties of cubic CaF_2, which is not surprising given its parameters were originally fit to this phase. However, it performs poorly for most other polymorphs and properties investigated in this study [e.g., the surface energy of the polar (100) surface], and CENT consistently produces results in better agreement with PBE. In fact, this favorable behavior of CENT is not only limited to CaF_2 but extends to other ionic materials, demonstrating that an ANN force field can be generated with a high predictive power of materials properties at conditions which significantly differ from those of the reference data.

ACKNOWLEDGMENTS

M.A. acknowledges support from the Novartis Universität Basel Excellence Scholarship for Life Sciences and the Swiss National Science Foundation (P300P2-158407). The Swiss

National Supercomputing Center in Lugano (projects s700 and s707), the Extreme Science and Engineering Discovery Environment (XSEDE) (which is supported by National Science Foundation Grant No. OCI-1053575), the Bridges system at the Pittsburgh Supercomputing Center (PSC) (which is supported by NSF Award No. ACI-1445606), the Quest high performance computing facility at Northwestern University, and the National Energy Research Scientific Computing Center (DOE: DE-AC02-05CH11231) are gratefully acknowledged.

[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] R. A. Buckinghum, Proc. R. Soc. London, Ser. A **168**, 264 (1938).
[7] B. G. Dick and A. W. Overhauser, Phys. Rev. **112**, 90 (1958).
[8] A. K. Rappe and I. W. A. Goddard, J. Phys. Chem. **95**, 3358 (1991).
[9] J. Yu, S. B. Sinnott, and S. R. Phillpot, Phys. Rev. B **75**, 085311 (2007).
[10] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, J. Phys. Chem. A **105**, 9396 (2001).
[11] 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).
[12] Y.-T. Cheng, T. Liang, J. A. Martinez, S. R. Phillpot, and S. B. Sinnott, J. Phys.: Condens. Matter **26**, 265004 (2014).
[13] J. E. Muller, A. C. T. van Duin, and W. A. Goddard III, J. Phys. Chem. C **114**, 4939 (2010).
[14] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[15] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).
[16] C. M. Handley and J. Behler, Eur. Phys. J. B **87**, 1 (2014).
[17] T. Morawietz, V. Sharma, and J. Behler, J. Chem. Phys. **136**, 064103 (2012).
[18] W. J. Szlachta, A. P. Bartók, and G. Csányi, Phys. Rev. B **90**, 104108 (2014).
[19] F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, Int. J. Quantum Chem. **115**, 1094 (2015).
[20] B. Meredig, A. Agrawal, S. Kirklin, J. E. Saal, J. W. Doak, A. Thompson, K. Zhang, A. Choudhary, and C. Wolverton, Phys. Rev. B **89**, 094104 (2014).
[21] M. Rupp, R. Ramakrishnan, and O. A. von Lilienfeld, J. Phys. Chem. Lett. **6**, 3309 (2015).
[22] N. Artrith, T. Morawietz, and J. Behler, Phys. Rev. B **83**, 153101 (2011).
[23] M. Malshe, R. Narulkar, L. M. Raff, M. Hagan, S. Bukkapatnam, and R. Komanduri, J. Chem. Phys. **129**, 044111 (2008).
[24] S. A. Ghasemi, A. Hofstetter, S. Saha, and S. Goedecker, Phys. Rev. B **92**, 045131 (2015).
[25] W. Liang, V. Ilchenko, D. Eliyahu, A. Savchenkov, A. Matsko, D. Seidel, and L. Maleki, Nat. Commun. **6**, 7371 (2015).
[26] C. Cazorla and D. Errandonea, Phys. Rev. Lett. **113**, 235902 (2014).
[27] A. Lyberis, A. J. Stevenson, A. Suganuma, S. Ricaud, F. Druon, F. Herbst, D. Vivien, P. Gredin, and M. Mortier, Opt. Mater. (N.Y.) **34**, 965 (2012).
[28] L. Sang, M. Liao, Y. Koide, and M. Sumiya, Appl. Phys. Lett. **98**, 103502 (2011).
[29] J. Barth, R. L. Johnson, M. Cardona, D. Fuchs, and A. M. Bradshaw, Phys. Rev. B **41**, 3291 (1990).
[30] M. Verstraete and X. Gonze, Phys. Rev. B **68**, 195123 (2003).
[31] N. Sata, G. Y. Shen, M. L. Rivers, and S. R. Sutton, Phys. Rev. B **65**, 104114 (2002).
[32] H. Shi, R. I. Eglitis, and G. Borstel, Phys. Rev. B **72**, 045109 (2005).
[33] W. Hayes, Contemp. Phys. **19**, 469 (1978).
[34] W. Hayes and A. M. Stoneham, *Defects and Defect Processes in Nonmetallic Solids* (Dover Publications, Mineola, NY, 2004).
[35] J. Behler, J. Chem. Phys. **134**, 074106 (2011).
[36] H. Eshet, R. Z. Khaliullin, T. D. Kuhne, J. Behler, and M. Parrinello, Phys. Rev. B **81**, 184107 (2010).
[37] W. J. Mortier, K. V. Genechten, and J. Gasteiger, J. Am. Chem. Soc. **107**, 829 (1985).
[38] N. Artrith and A. Urban, Comput. Mater. Sci. **114**, 135 (2016).
[39] S. A. Ghasemi, A. Neelov, and S. Goedecker, J. Chem. Phys. **127**, 224102 (2007).
[40] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[41] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun **180**, 2175 (2009).
[42] S. Goedecker, J. Chem. Phys. **120**, 9911 (2004).
[43] M. Amsler and S. Goedecker, J. Chem. Phys. **133**, 224104 (2010).
[44] B. Schaefer, S. A. Ghasemi, S. Roy, and S. Goedecker, J. Chem. Phys. **142**, 034112 (2015).
[45] 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).
[46] H. W. Kuhn, Nav. Res. Log. **2**, 83 (1955).
[47] J. D. Gale, J. Chem. Soc., Faraday Trans. **93**, 629 (1997).
[48] C. Cazorla and D. Errandonea, J. Phys. Chem. C **117**, 11292 (2013).
[49] L. Gerward, J. S. Olsen, S. Steenstrup, M. Malinowski, S. Åsbrink, and A. Waskowska, J. Appl. Crystallogr. **25**, 578 (1992).
[50] J. M. Leger, J. Haines, A. Atouf, O. Schulte, and S. Hull, Phys. Rev. B **52**, 13247 (1995).
[51] S. Speziale and T. S. Duffy, Phys. Chem. Miner. **29**, 465 (2002).
[52] X. Wu, S. Qin, and Z. Y. Wu, Phys. Rev. B **73**, 134103 (2006).
[53] F. Hund and K. Lieck, Z. Anorg. Allg. Chem. **271**, 17 (1952).
[54] S. X. Cui, W. X. Feng, H. Q. Hu, Z. B. Feng, and Y. W. Wang, Comp. Mater. Sci **47**, 71 (2009).
[55] H. Shi, W. Luo, B. Johansson, and R. Ahujia, J. Phys. Condens. Matter **21**, 415501 (2009).

[56] Y. Y. Qi, Y. Cheng, M. Liu, X. R. Chen, and L. C. Cai, Physica B **426**, 13 (2013).
[57] K. F. Seifert, Ber. Bunsenges. Physik. Chem. **70**, 1041 (1966).
[58] E. Morris, T. Groy, and K. Leinenweber, J. Phys. Chem. Sol. **62**, 1117 (2001).
[59] D. P. Dandekar and J. C. Jamieson, Trans. Am. Crystallogr. Assoc. **5**, 19 (1969).
[60] S. Dorfman, F. Jiang, Z. Mao, A. Kubo, V. Prakapenka, and T. Duffy, in *AGU Fall Meeting Abstracts*, Vol. 1 (2007), p. 1522.
[61] J. H. Burnett, Z. H. Levine, and E. L. Shirley, Phys. Rev. B **64**, 241102(R) (2001).
[62] F. S. El'kin, O. B. Tsiok, L. G. Khvostantsev, and V. V. Brazhkin, J. Exp. Theor. Phys. **100**, 971 (2005).
[63] V. Puchin, A. Puchina, M. Huisinga, and M. Reichling, J. Phys.: Condens. Matter **13**, 2081 (2001).
[64] P. Tasker, J. Phys. Colloques **41**, C6 (1980).
[65] N. H. de Leeuw and T. G. Cooper, J. Mater. Chem. **13**, 93 (2003).
[66] J. Vogt, J. Henning, and H. Weiss, Surf. Sci. **578**, 57 (2005).
[67] R. Eglitis, H. Shi, and G. Borstel, Surf. Rev. Lett. **13**, 149 (2006).
[68] S. Hirth, F. Ostendorf, and M. Reichling, Nanotechnology **17**, S148 (2006).
[69] Y. Kakinuma, S. Azami, and T. Tanabe, CIRP Ann. Manuf. Techn. **64**, 117 (2015).
[70] A. Jockisch, U. Schroder, F. De Wette, and W. Kress, J. Phys.: Condens. Matter **5**, 5401 (1993).
[71] J. J. Gilman, J. Appl. Phys. **31**, 2208 (1960).
[72] P. Maldonado, J. Godinho, L. Evins, and P. M. Oppeneer, J. Phys. Chem. C **117**, 6639 (2013).
[73] A. Velichko, V. Ilyushin, N. Filimonova, and A. Katsyuba, J. Synch. Investig. **7**, 488 (2013).
[74] N. De Leeuw, J. A. Purton, S. Parker, G. Watson, and G. Kresse, Surf. Sci. **452**, 9 (2000).
[75] P. Lindan, Philos. Mag. Lett. **71**, 75 (1995).
[76] J. B. Yi, C. C. Lim, G. Z. Xing, H. M. Fan, L. H. Van, S. L. Huang, K. S. Yang, X. L. Huang, X. B. Qin, B. Y. Wang, T. Wu, L. Wang, H. T. Zhang, X. Y. Gao, T. Liu, A. T. S. Wee, Y. P. Feng, and J. Ding, Phys. Rev. Lett. **104**, 137201 (2010).
[77] M. Nowotny, L. Sheppard, T. Bak, and J. Nowotny, J. Phys. Chem. C **112**, 5275 (2008).
[78] K. Choudhary, T. Liang, A. Chernatynskiy, S. R. Phillpot, and S. B. Sinnott, J. Phys.: Condens. Matter **27**, 305004 (2015).
[79] A. Kumar, A. Chernatynskiy, T. Liang, K. Choudhary, M. J. Noordhoek, Y.-T. Cheng, S. R. Phillpot, and S. B. Sinnott, J. Phys.: Condens. Matter **27**, 336302 (2015).
[80] A. Togo and I. Tanaka, Scr. Mater. **108**, 1 (2015).
[81] S. Andersson and G. Backstrom, J. Phys. C **20**, 5951 (1987).
[82] R. Golesorkhtabar, P. Pavone, J. Spitaler, P. Puschnig, and C. Draxl, Comput. Phys. Commun. **184**, 1861 (2013).
[83] D. Huffman and M. Norwood, Phys. Rev. **117**, 709 (1960).
[84] *Non-Tetrahedrally Bonded Elements and Binary Compounds I*, Landolt-Börnstein–Group III Condensed Matter Vol. 41C, edited by O. Madelung, U. Rössler, and M. Schulz (Springer-Verlag, Heidelberg, 1998).
[85] R. Srinivasan, Proc. Phys. Soc. **72**, 566 (1958).
[86] M. Råsander and M. Moram, J. Chem. Phys. **143**, 144104 (2015).
[87] D. Chung and W. Buessem, J. Appl. Phys. **38**, 2535 (1967).