Small basis set density functional theory method for cost-efficient, large-scale condensed matter simulations

View Online   Export Citation   CrossMark

Elisabeth Keller,[1,2] (ID) Jack Morgenstein,[3] (ID) Karsten Reuter,[1] (ID) and Johannes T. Margraf[1,2,a] (ID)

AFFILIATIONS

[1] Fritz Haber Institute of the Max Planck Society, Berlin, Germany
[2] Bavarian Center for Battery Technology (BayBatt) and Chair of Physical Chemistry V, University of Bayreuth, Bayreuth, Germany
[3] Duke University, Durham, North Carolina 27708, USA

[a] Author to whom correspondence should be addressed: johannes.margraf@uni-bayreuth.de

ABSTRACT

We present an efficient first-principles based method geared toward reliably predicting the structures of solid materials across the Periodic Table. To this end, we use a density functional theory baseline with a compact, near-minimal *min+s* basis set, yielding low computational costs and memory demands. Since the use of such a small basis set leads to systematic errors in chemical bond lengths, we develop a linear pair-wise correction, available for elements $Z = 1-86$ (excluding the lanthanide series), parameterized for use with the Perdew–Burke–Ernzerhof exchange–correlation functional. We demonstrate the reliability of this corrected approach for equilibrium volumes across the Periodic Table and the transferability to differently coordinated environments and multi-elemental crystals. We examine relative energies, forces, and stresses in geometry optimizations and molecular dynamics simulations.

I. INTRODUCTION

In materials science, first-principles simulations are the state-of-the-art approach for obtaining detailed atomistic insights into the structure and properties of bulk materials, surface–adsorbate systems, interfaces, and nanoparticles. To this end, Kohn–Sham density functional theory (KS-DFT) methods employing generalized gradient approximation (GGA) functionals, such as the Perdew–Burke-Ernzerhof (PBE) functional,[1] are extremely popular. PBE reliably describes equilibrium structures, vibrational spectra, binding and cohesive energies for a broad range of materials.[2] However, to capture the behavior of structurally complex systems (e.g., defects, interfaces, and amorphous phases), large simulation cells are required, with concomitantly large demands of CPU time and memory. In addition, to describe dynamic properties or finite-temperature effects, molecular dynamics (MD) simulations with millions of simulation steps are required. Such simulations are hindered by the computational cost of typical DFT calculations.

In terms of computational scaling, the bottleneck of KS-DFT calculations lies in the solution of the KS eigenvalue problem. Commonly used direct eigensolvers, such as ELPA,[3] scale cubically ($\mathcal{O}(N^3)$) with system size. This cubic scaling KS solution step therefore dominates the total cost of the self-consistent field (SCF) cycle in the limit of large simulation cells. Considerable research is being directed toward more cost-efficient ways to solve the eigenvalue problem. These efforts include iterative eigensolvers (e.g., SLEPc[4]) with better than cubic scaling or density matrix solvers (e.g., NTPoly[5]) that bypass the diagonalization step to reach linear scaling. Both iterative eigensolvers and density matrix solvers only reach their full potential when applied to sufficiently large sparse matrices, though (e.g., for lower dimensional systems and/or insulators), whereas the computational cost is unfavorable for small- to medium-sized systems and dense bulk systems with small bandgaps (metals and semiconductors). Here, they are still outperformed by direct eigensolvers. Consequently, despite significant progress in

11 February 2025 06:27:12

hardware and algorithms, KS-DFT calculations for systems with thousands of atoms are in general far from routine.

Semiempirical electronic structure (SQM) methods are low-cost alternatives to DFT, which overcome its computational limitations by using minimal basis sets and empirical integral approximations. Popular SQM methods in quantum chemistry include the PMn methods based on the Neglect of Differential Diatomic Overlap (NDDO, e.g., PM6[6]), the extended tight-binding methods GFNn-xTB (GFN-xTB[7] and GFN2-xTB[8]), and the density functional tight-binding (DFTB) approach.[9–11]

Clearly, methods that incorporate integral approximations (e.g., within the DFTB or NDDO frameworks) have computational advantages over methods that do not. On the flipside, these approximations inevitably include empirical components, and their applicability is therefore restricted to the original domain of interest. As a consequence, SQM methods can display an unpredictable behavior and large errors when going beyond this scope. This is extensively documented in Ref. 12, where methods that display an excellent performance for molecules were found to yield poor results for metals and other solids. SQM methods thus offer low computational costs but lack the robustness and transferability of first-principles methods.

Improved robustness and accuracy can be obtained when semiempirical methods are built on top of a first-principles baseline. This is for example done in the HF-3c method,[13] which is presently the most cost-efficient method of a set of "3c" methods developed by Grimme and co-workers.[14] Specifically, the HF-3c method uses Hartree–Fock (HF) in combination with a near-minimal basis set.[13] Clearly, the lack of electron correlation and basis-set incompleteness introduces significant errors in energies and geometries. HF-3c corrects these with three atom-pairwise empirical corrections. These include corrections for the dispersion interaction, the basis set superposition error (BSSE), and a short-range basis correction targeting the basis set incompleteness error (BSIE).[13] Notably, the "3c" approach has also been extended to DFT, e.g., with the PBEh-3c[15] and r2SCAN-3c[16] methods. However, all 3c methods are tailored to obtaining geometries and thermodynamic properties of molecular systems. Yet, there is a similar demand for cost-efficient methods for obtaining reliable geometries of inorganic bulk systems at the DFT level, e.g., for initial screenings in materials discovery or for generating training data for machine learning (ML) potentials. The current work therefore introduces such an approach for bulk materials within the FHI-aims DFT code.[17]

II. METHOD

As a basis for the proposed method, we employ a cost-efficient, first-principles model using the semilocal PBE functional with a near-minimal basis set termed *min+s*. We focus on the PBE functional as a target, without additional dispersion corrections. While such corrections are important for some classes of materials, they are not used by default in general simulations of inorganic materials, as they can be detrimental (e.g., in the description of metallic systems) and parameterizations are not always broadly tested across the Periodic Table. As a consequence, all major materials databases (e.g., the Materials Project,[18] AFLOW,[19] and OQMD[20]) use plain PBE in their computational setup. Since we observed systematic underbinding

at this level of theory (and consequently overestimated lattice constants), we propose a simple empirical correction term. Together, the *min+s* basis set and the proposed linear pairwise correction (LPC) represent a robust and cost efficient method for structural relaxations of materials across the Periodic Table. This method is described in detail below.

A. PBE/*min+s* baseline

1. Basis set specification

In FHI-aims, KS orbitals are expressed in terms of numeric atom-centered orbital (NAO) basis functions ϕ_i,[17]

$$\phi_i(\boldsymbol{r}) = \frac{u_i(r)}{r} Y_{lm}(\Omega) \tag{1}$$

with localized numerical tabulated radial functions $u_i(r)$ and spherical harmonics $Y_{lm}(\Omega)$. For each element, the NAO basis set is hierarchically constructed from a minimal free-atom basis set by iteratively adding additional radial functions until a required level of energy convergence is reached. FHI-aims provides predefined numerical settings for different levels of convergence. These settings define a set of NAO basis functions and correspondingly adjusted integration grids, multipole expansions for the Hartree potential, etc. The most commonly used settings in FHI-aims are termed *light* and *tight*. For GGA functionals, such as PBE, the *tight* settings are essentially fully converged and recommended for highly precise energy calculations, whereas the *light* settings are much more cost-efficient and often used for structural relaxations and *ab initio* molecular dynamics.[17] For large systems, where the solution of the KS equations dominates the computational cost, even *light* settings can become computationally prohibitive in terms of CPU time and memory consumption, however.

In this work, we employ a cost-efficient near-minimal NAO basis set, which we denote as *min+s*. The *min+s* basis consists of a minimal set of basis functions (including the full valence shell of the corresponding element) and one additional s-type function, which grants some amount of radial flexibility to the basis set at a negligible computational cost. Specifically, it substantially improves equilibrium volumes V_0 compared to the pure minimal basis set. To illustrate this point, we benchmarked the relative equilibrium volume error using the PBE functional for 71 ground-state elemental crystals[21] against multiple small basis set choices (*min*, *min+s*, *min+sp*, *min+spd*) compared to the converged tight basis set; see Fig. S1. This reveals that *min+s* systematically improves the description of all elements by reducing the error in V_0 approximately by a factor of 2 compared to using the min basis set alone. Including higher angular momentum p and d functions further decreases the error in V_0, but also lowers the computational benefit of using small basis sets compared to the converged tight settings. In contrast, the cost of *min+s* differs only negligibly from the min basis set.

All basis functions are localized within a basis set cutoff inherited from the *light* settings, where the cutoff is chosen to be as small as possible without significantly affecting the accuracy of computations. For completeness, the *min+s* basis set cutoffs for each element are provided in the supplementary material. To further reduce the computational cost, only the chemically relevant valence and shallow-core electrons are considered by using the frozen core approximation in the solution of the KS equations. Here,

we are using the implementation by Yu *et al.* (FC99+C+V) with an energy cutoff of −100 eV (if not otherwise stated).[22,23] Note that despite improving the description of equilibrium volumes, the *min+s* basis is inadequate for accurately describing the variationally optimal electron density and KS wavefunction. This can be seen from density difference plots between converged and small basis electron densities for the Cu dimer (see Fig. S2), revealing that the small basis sets fail to capture the redistribution of charge in a chemical bond. Clearly, a spherically symmetrical s-function cannot describe such anisotropic effects. Nonetheless, the density difference plots do reveal small improvements in the overall density going from *min* to *min+s*.

As an alternative to increasing the basis set size, there are also recent approaches for designing adaptive small basis sets, such as the NOTCH model of Wang and Neese[24] or the adaptive q-vSZP basis set reported by Müller *et al.*.[25] In general, this increases the flexibility of the basis but also the empiricism of the underlying models. For example, NOTCH defines the contraction of a Gaussian Type Orbital (GTO) basis set via bond length-dependent contraction coefficients. Similarly, Ref. [25] uses an atom-in-molecule adaptively contracted GTO minimal basis with effective atomic charge and coordination number dependent contraction coefficients. The current work avoids the empirical parameterization of such contractions at the cost of lower flexibility. Alternatively, the ONETEP code use similar ideas to optimize atom centered basis functions in a completely non-empirical variational approach.[26] However, this significantly increases the computational effort of the calculations,

compared to a static (near) minimal basis set. Note that unlike GTO methods, the NAO approach used herein does not require deep contractions of primitive basis functions, since the NAOs already represent atomic orbitals accurately. Consequently, NAOs are well suited for small basis methods, such as the one used herein. Meanwhile, the adaptive contraction approach cannot be directly transferred to NAOs.

2. Computational cost

In Fig. 1(a), the computational cost of the *min+s* basis set is compared to the *light* and *tight* settings for a single SCF iteration on CsPbBr$_3$ supercells with up to 2560 atoms. PBE/*min+s* reduces the total wall-clock time on average to 52% and 21% compared to PBE/*light* and PBE/*tight*. Beyond the runtime, large-scale simulations are also often limited by their memory demand. Here, PBE/*min+s* reduces the required memory usage on average to 15% and 28% compared to PBE/*light* and PBE/*tight*, respectively. The computational time and memory savings due to *min+s* are thereby mainly due to the reduced number of basis functions and lighter integration grids (compared to *tight*). However, the overall reduction varies according to the contained elements, as well as the size and density of the examined system. Generally, the savings are largest for elements with low atomic numbers and decrease for heavy elements with many core orbitals. In this respect, CsPbBr$_3$ is far from the best case scenario. Nevertheless, the *min+s* settings clearly lead to substantial savings compared to the *light* and especially *tight* settings.
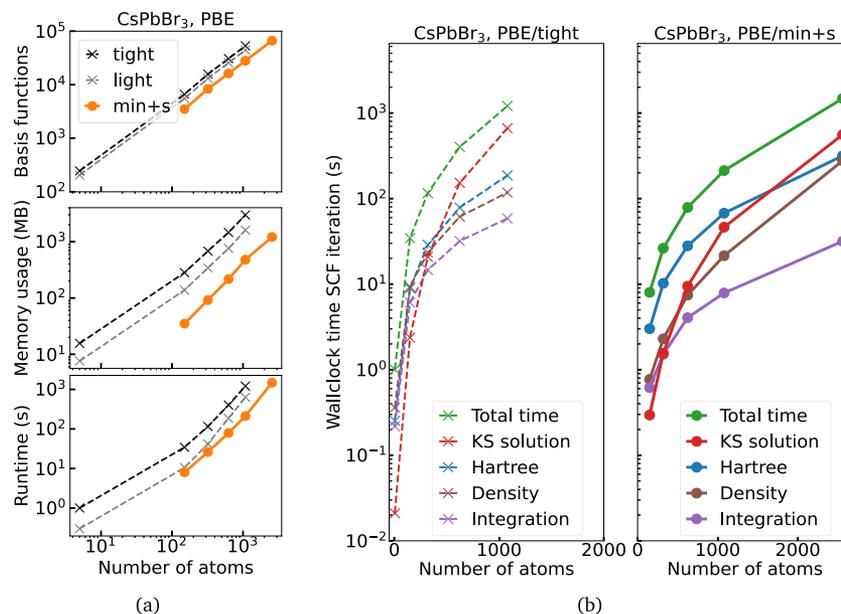

FIG. 1. Scaling behavior of the computational cost for self-consistent field (SCF) iterations employing the *min+s*, *light*, and *tight* settings as a function of system size for CsPbBr$_3$ supercells with 5, 150, 320, 625, 1080, and 2560 atoms. (a) Comparison of the number of basis functions, highest tracked memory usage, and total runtime. (b) Subtimings for the first SCF iteration with the *min+s* and *tight* basis sets. The 5-, 150-, 320-, 625-, 1080-, and 2560-atom supercells have been computed on two nodes with 40 CPUs each with large memory nodes 192 GB on the HPC cluster Cobra (processor type: Intel Skylake 6148, processor clock: 2.4 GHz). For *min+s* settings, five-atom computations were excluded due to ELPA errors from too small matrices for the chosen number of CPUs. For *light* and *tight* settings, the 2560-atom supercell could not be computed due to memory limitations.

To obtain better insights into the time savings for different system sizes, we examine the major contributors to the total wall-clock time of the first SCF iteration in Fig. 1(b). For small system sizes, the major contributors to the total time are linear scaling grid-based computational steps (Hartree potential, density update, and integration step). For large system sizes, the cubic scaling KS solution dominates. By using the *min+s* basis set [Fig. 1(b), right] instead of the *tight* basis set [Fig. 1(b) left], the computational cost for both grid-based computational steps and the KS solution is decreased, yielding lower computational cost for small-, medium-, and large-scale systems. Even better, the relative savings increase with increasing system size (see Table 1 of the supplementary material) due to the increased sparsity of matrices obtained with the *min+s* basis set. Crucial for enabling large-scale computations is the crossover point, for which the cost of the cubic scaling KS solution exceeds the linear scaling grid-based computational steps. PBE/*min+s* shifts this crossover point to significantly larger system sizes (around 1000 atoms). In turn, this pushes the KS bottleneck to larger system sizes and enables cost-efficient large-scale simulations.

3. Basis set incompleteness errors for crystals

Using the compact *min+s* basis clearly leads to computational advantages, but this inevitably has negative effects on the predictive accuracy of the calculations, due to basis set incompleteness errors (BSIEs). For example, it has been observed that small basis sets lead to systematically overestimated bond lengths in organic molecules.[13,28] To examine whether similar systematic trends can be observed in solids, we computed energy–volume curves and derived equilibrium properties (such as equilibrium volumes V_0) of bulk crystals across the Periodic Table. Note that for simplicity, we will use the term BSIE for the discrepancy between *min+s* and *tight* settings in the following, although there are also (smaller) contributions from grid densities and other factors to this difference.

For illustration, the PBE/*min+s* and PBE/*tight* energy–volume curves of bulk silicon in the diamond crystal structure are shown in Fig. 2(a). The corresponding equilibrium volumes are obtained from a Birch–Murnaghan equation-of-state fit,[29,30] performed using the atomic simulation environment (ASE).[31] This reveals that

PBE/*min+s* overestimates the equilibrium volume V_0 of silicon by 6%, compared to PBE/*tight*. Beyond this, the BSIE also impacts the shape of the energy–volume curve. This discrepancy in shape and equilibrium volume between two energy–volume curves can be assessed with the Δ-value introduced by Lejaeghere *et al.* as follows:[21]

$$\Delta = \sqrt{\frac{\int \Delta E^2(V)\, dV}{\Delta V}}. \tag{2}$$

These Δ-values will be used as an optimization target further below.

Importantly, the overestimation of equilibrium volumes due to BSIE is not just observed for silicon, but for a wide range of elements and bonding types, including ionic, covalent, metallic, and molecular systems. This is shown for a dataset of 128 crystals in Fig. 3 (see Table I). The volumes are commonly overestimated by 5%–10%, and in some cases by more than 40%. Analyzing the BSIE for different bonding types, we find that the error is smallest for systems that are reasonably similar to free atoms, such as noble gases or ionic materials. In contrast, the overestimation increases for more complex bonding situations, e.g., in covalent, metallic, and molecular systems. This can be rationalized by considering that the minimal basis set is obtained from isolated atom calculations in the NAO scheme. A near minimal basis is thus well suited for free atoms (the limit of infinite volume in an energy–volume curve), and the BSIE will be more pronounced for smaller volumes. Overall, this results in underbinding and an overestimation of equilibrium volumes.

B. Linear pairwise correction (LPC) for BSIE
1. Method definition

Having observed the systematic BSIE effect on energy–volume curves, we now aim to correct the potential energy surface (PES) in such a way that overestimated bond lengths are shortened without adversely affecting the PES and the related thermodynamic ensemble. To this end, we draw on the literature of minimally invasive biases that have been developed for large-scale biomolecular
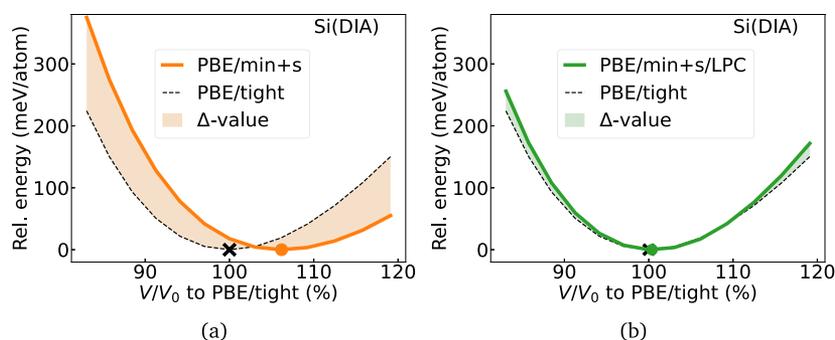
FIG. 2. Energy–volume curves for bulk silicon obtained with (a) PBE/*min+s* and (b) PBE/*min+s*/LPC, both relative to PBE/*tight*. The Δ-values show the dissimilarity of the E(V) curves obtained from PBE/*min+s* and PBE/*min+s*/LPC compared to PBE/*tight*. The equilibrium volumes obtained with PBE/*min+s*, PBE/*min+s*/LPC, and PBE/*tight* are marked by dots and crosses.
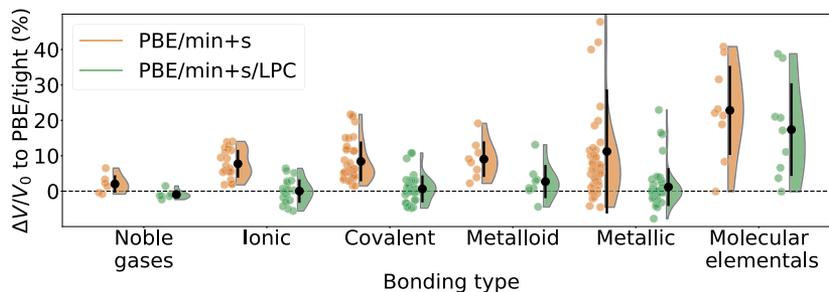
FIG. 3. Performance of PBE/*min+s* and PBE/*min+s*/LPC for equilibrium volumes of bulk systems with different bonding types, relative to PBE/*tight* (see Table I). The mean errors and standard deviations are indicated in black. The original data are plotted on the left side of each distribution.

TABLE I. Test set of 128 materials, including noble gases,[21] ionic binary compounds,[27] covalent semiconductor binary compounds,[27] metalloids,[21] metals,[21] and molecular elemental crystals.[21]

Bonding type	Number of materials	Materials
Noble gases	6	He, Ne, Ar, Kr, Xe, Rn[21]
Ionic	21	Alkali halides AB with A = Li, Na, K, Rb, Cs and B = F, Cl, Br, I[27]
Covalent	37	AlAs(ZB), AlN(WUR), AlN(ZB), AlP(ZB),AlSb(ZB), BAs(ZB), BP(ZB), CdS(WUR), CdS(ZB), CdSe(WUR), CdSe(ZB), CdTe(ZB), GaAs(ZB), GaN(WUR), GaN(ZB), GaP(ZB), GaSb(ZB), HgS(ZB), HgSe(ZB), HgTe(ZB), InAs(ZB), InN(WUR), InP(ZB), InSb(ZB), MgO(RS), MgS(RS), MgSe(RS), PbS(RS), PbSe(RS), PbTe(RS), SiC(ZB), ZnO(WUR), ZnS(WUR), ZnS(ZB), ZnSe(ZB), ZnTe(ZB), C(DIA)[27]
Metalloid	8	B, Si, Ge, As, Se, Sb, Po, Te[21]
Metallic	47	Li, Na, K, Rb, Cs, Be, Mg, Ca, Sr, Ba, Sc, Y, Lu, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, W, Mn, Tc, Re, Fe, Ru, Os, Co, Rh, Ir, Ni, Pd, Pt, Cu, Ag, Au, Zn, Cd, Hg, Al, Ga, In, Sn, Tl, Pb, Bi[21]
Molecular elemental crystals	9	H_2, O_2, N_2, F_2, P, S, Cl_2, Br_2, I_2[21]

simulations. In particular, Pitera and Chodera introduced a linear form of bias based on a maximum entropy argument, which distorts the unbiased statistical ensemble the least.[32,33] With this goal in mind, we define a simple linear pairwise correction (LPC) $e_{LPC,AB}$, which is fast to evaluate and easy to parameterize,

$$e_{LPC,AB} = c_{Z_A Z_B} \cdot (r_{AB} - r_{cut,Z_A Z_B}) \qquad (3)$$

with $r_{AB} = |\mathbf{r}_{AB}|$ being the absolute distance between atoms A and B, $c_{Z_A Z_B}$ denoting the element-pair dependent correction strength (with $c_{Z_A Z_B} \geq 0$), and $r_{cut,Z_A Z_B}$ denoting an element-pairwise cutoff radius. The latter provides a measure for the onset of the correction, which should be short-ranged and act mainly on directly bonded atoms, while interactions between next-nearest neighbors and beyond are removed by a switching function (see below). The effect of the LPC on the silicon energy–volume curve is shown in Fig. 2(b). This confirms that the underbinding is corrected, without otherwise distorting the potential energy surface.

To avoid the need for parameterizing all element-pairs in the Periodic Table, the pairwise parameters $c_{Z_A Z_B}$ and $r_{cut,Z_A Z_B}$ are determined via arithmetic means of the contributing species,

$$c_{Z_A Z_B} = \frac{c_{Z_A} + c_{Z_B}}{2}, \qquad (4)$$

$$r_{cut,Z_A Z_B} = \frac{r_{cut,Z_A} + r_{cut,Z_B}}{2}. \qquad (5)$$

The use of the arithmetic mean is a common choice for cutoff radii. It is, e.g., also used for Lennard-Jones potentials. For the correction strength $c_{Z_A Z_B}$, use of the geometric mean was also explored. However, this proved problematic in cases where the parameterization yielded values of c_{Z_A} close to zero (see below). With the geometric mean, all pairwise corrections involving these elements would be zero, while the arithmetic mean is more well behaved in this case.

The full correction term E_{LPC} is obtained by summing up the pairwise corrections $e_{LPC,AB}$ for each atom pair AB, multiplied by the aforementioned switching function,

$$E_{LPC} = \frac{1}{2} \sum_{A}^{N_{unit}} \sum_{B \neq A}^{N_{super}} e_{LPC,AB} \cdot f_{switch}(r_{AB}, r_{cut,Z_A Z_B}). \qquad (6)$$

The switching function f_{switch} of width w = 0.5 Å is given by[34]

$$f_{\text{switch}}\left(r_{AB}, r_{\text{cut},Z_A Z_B}\right) = \begin{cases} 1 & \text{if } r_{AB} < r_{\text{cut},Z_A Z_B} - w, \\ \frac{1}{2}\left(\cos\left(\frac{\pi}{w} \cdot \left(r_{AB} - r_{\text{cut},Z_A Z_B} + w\right)\right) + 1\right) & \text{if } r_{\text{cut},Z_A Z_B} - w \le r_{AB} \le r_{\text{cut},Z_A Z_B}, \\ 0 & \text{if } r_{AB} > r_{\text{cut},Z_A Z_B}. \end{cases} \tag{7}$$

f_{switch} ensures a smooth transition to zero as $r_{\text{cut},Z_A Z_B}$ is approached. This in turn leads to continuous derivatives, which enable force and stress evaluations. Finally, the total corrected energy $E_{min+s/\text{LPC}}$ consists of the first-principles energy $E_{\text{PBE}/min+s}$ obtained with the PBE/$min+s$ baseline and the LPC correction term E_{LPC},

$$E_{min+s/\text{LPC}} = E_{\text{PBE}/min+s} + E_{\text{LPC}}. \tag{8}$$

The expressions for LPC forces and stresses are given in the supplementary material.

2. Parameterization

Clearly, the accuracy of the LPC ultimately depends on an appropriate parameterization of the correction strength c_{Z_A} and the cutoff r_{cut,Z_A}. The latter mainly serves to ensure that the correction is applied to all relevant short-range interactions, while leaving long-range interactions unaltered. For most elements, this can be achieved by setting r_{cut,Z_A} to 2.5 times the corresponding elemental single-bond covalent radius ($r_{\text{cut},Z_A} = 2.5 r_{\text{cov},Z_A}$), with radii taken from Refs. 35 and 36. Exceptions are made for some elements that require larger cutoffs to cover the most common bonding situations. These exceptions include noble gases for which van-der-Waals radii are used ($r_{\text{cut},Z_A} = 2.5 r_{\text{vdW},Z_A}$ with r_{vdW,Z_A} for He, Ne, Ar, Kr, Xe from Refs. 37 and 38). Furthermore, slightly larger cutoffs are used for S, Hg, Pb, Se, Be ($r_{\text{cut},Z_A} = 3 r_{\text{cov},Z_A}$), of which all display diverse bonding patterns in different crystal polymorphs.

In contrast to the cutoffs, the correction strengths c_{Z_A} need to be more carefully tuned for each element, in order to minimize the systematic underbinding caused by the BSIE. To this end, a training set including a range of common homoelemental bonding situations for each element (namely the dimer, graphite, diamond, β-tin, bcc, and fcc prototypes) was used, as first reported in Ref. 39. This consistent set of structures covers coordination numbers from 1 to 12 and is thus representative of the diverse bonding situations encountered in solids. However, this diversity also has a downside, in that less important high energy configurations (e.g., fcc oxygen) can dominate the error when optimizing the parameters, leading to unbalanced parameterizations.

To ensure that the LPC is robust for the important low energy configurations of each element, the structures are therefore weighted according to a Boltzmann distribution centered on the energetically most stable structure (at the PBE/$tight$ level), with the Boltzmann factor p_i of structure $i \in I = \{\text{dimer}, \text{graphite}, \text{diamond}, \beta-\text{tin}, \text{bcc}, \text{and fcc}\}$ given as

$$p_i = \frac{1}{Z} \exp\left(\frac{-\Delta E_i}{kT}\right), \tag{9}$$

with

$$\Delta E_i = E_{0,i} - \min\left(\{E_{0,j}| j \in I\}\right) \tag{10}$$

and the normalization factor $Z = \sum_{j \in I} \exp\left(\frac{-\Delta E_j}{kT}\right)$. Here, kT is the product of the Boltzmann constant k and the temperature T. In the fitting procedure, kT is fixed to 0.25 eV, which provides a good balance between the correct description of low energy configurations and a qualitative description of higher energy configurations. To further reduce the influence of outliers (e.g., in many cases the dimers), structures with Boltzmann factors smaller than ten percent are excluded from the fitting procedure.

Based on these structures, the c_{Z_A} parameters for all elements were optimized with the Nelder–Mead method,[40,41] for fixed $r_{\text{cut},A}$, minimizing the loss function $L(c_{Z_A})$ by summing over the Boltzmann-factor p_i [Eq. (9)] weighted Δ_i-values [Eq. (2)] for each structure i in the training set,

$$L(c_{Z_A}) = \sum_{i \in I} \Delta_i(E_{min+s/\text{LPC}}, E_{\text{PBE}/tight}) \cdot p_i. \tag{11}$$

While this approach for reference data generation worked well for most elements, exceptions were made for H, O, N, F, Cl, Br, and I. This was necessary due to the fact that these elements form molecular dimers, which dominate the loss function when following the procedure described above. As the chemistry of molecular dimers is very different from important classes of solids (i.e., hydrides, oxides, nitrides, and halides), the corresponding parameters were reoptimized for representative binary compounds, keeping the parameters of the other elements fixed. For further details regarding the dataset, please refer to the supplementary material.

Overall, we thus obtained LPC parameters for elements with $Z = 1$–86 (excluding the lanthanide series) for use with the $min+s$ basis set and the PBE functional. The $min+s$ basis set and the LPC correction are accessible through the corresponding species defaults in FHI-aims starting from version 231 212 in `species_defaults/defaults_2020/minimal+s`, which automatically include the keyword to enable the basis set error corrections (currently parameterized for use with the $min+s$ basis set and the PBE functional) for energies, forces, and stresses.

III. RESULTS AND DISCUSSION

Having defined the LPC, we first benchmark PBE/$min+s$/LPC for equilibrium volumes of crystals against a PBE/$tight$ reference. Subsequently, we examine the performance of PBE/$min+s$/LPC for geometry optimizations and MD simulations.

A. Equilibrium volumes of crystalline solids

The performance of PBE/*min+s*/LPC is examined for a test set of monoelemental and binary solid state systems categorized into predominantly ionic, covalent, metalloid, and metallic materials and molecular elemental solids (see Table I). For this test set, the equilibrium volumes obtained with PBE/*min+s*/LPC and PBE/*min+s* are compared to the PBE/*tight* reference. As mentioned above, PBE/*min+s* overestimates most equilibrium volumes, whereas PBE/*min+s*/LPC reduces the BSIE significantly and obtains reliable equilibrium geometries for most materials; see Fig. 3.

Even though the LPC was mostly fitted on monoelemental reference structures, the method is transferable to multi-elemental materials. Indeed, basis set errors are generally larger for monoelemental systems (mainly represented in the metallic, covalent, and molecular elemental classes) compared to poly-elemental (e.g., ionic) materials. Overall, PBE/*min+s*/LPC shows the largest residual errors for molecular elemental dimer structures such as O_2 or N_2, where the performance is only slightly better than PBE/*min+s*. This can be attributed to the fact that the LPC was fitted to binary compounds (such as oxides and nitrides) for these elements and is thus not well suited for the corresponding elemental molecular systems. Indeed, the results indicate that fitting on elemental dimers would likely lead to an over-correction for solid binaries, vindicating the selected fitting strategy.

It should be noted that the energy–volume curves are computed by applying uniform strain to the cell. Under these conditions, the curves for molecular dimers mainly reflect a stretching or compression of the covalent bonds in the dimers. In principle, one could also compute energy–volume curves under the condition of fixed dimer geometries. In this case, the curve would likely be dominated by the basis set superposition error (BSSE), which leads to overbinding in non-covalent interactions. BSSE effects are not addressed by the LPC correction, since they occur on a different length scale (i.e., on the order of van-der-Waals radii) and have the opposite sign. When using a dispersion correction, BSSE effects can be compensated to a certain degree by adjusting the damping factor and onset of damping of the switching function. Furthermore, a method analogous to the semiempirical geometrical counterpoise correction (gCP) of Kruse and Grimme[42] could be developed for solids. These developments are beyond the scope of the current paper, however. At this stage, the PBE/*min+s*/LPC method is not recommended for systems that are dominated by van-der-Waals interactions.

To test the transferability of the PBE/*min+s*/LPC method further, we examined its performance for a test set of 100 organo-metal halide perovskite materials ABX_3 (structures from Ref. 43) and 100 quaternary kesterite materials A_2BCX_4 (the first 100 structures from Ref. 44); see Fig. 4. Because these materials consist of three to four different elements and may include small organic molecules, this represents a challenging test for the PBE/*min+s*/LPC method.

We find that PBE/*min+s*/LPC remains a consistent improvement over PBE/*min+s* here, although the volumes are still overestimated by ~5% on average. The improvement is particularly significant for the kesterites, where volumes are overestimated by up to 25% at the PBE/*min+s* level. Notably, the residual error of PBE/*min+s*/LPC is fairly systematic so that most systems are still underbound. This points to limitations of the simple functional form of the LPC, which cannot distinguish different crystal environments.
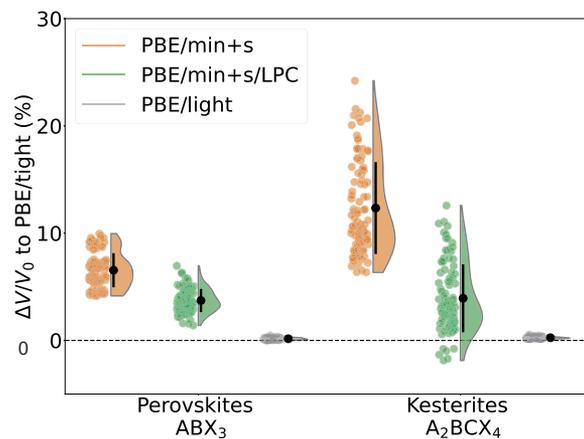

FIG. 4. Performance of PBE/*min+s*, PBE/*min+s*/LPC, and PBE/*light* for equilibrium volumes of 100 organo-metal halide perovskites ABX_3 (structures from Ref. 43) and 100 kesterites A_2BCX_4 (structures from Ref. 44), relative to PBE/*tight*. The mean errors and standard deviations are indicated in black. The raw data are plotted on the left side of each distribution.

Nonetheless, the current approach represents an improvement over the baseline, even in this extrapolative setting.

B. Applications

One of the main use-cases for methods such as PBE/*min+s*/LPC is the (pre-)relaxation of medium to large simulation cells, e.g., in the context of *ab initio* thermodynamics studies or materials screening.[45] While we have established that the proposed method will yield improved equilibrium volumes, the task of geometry optimization can potentially start from structures that are far from equilibrium.

To demonstrate this, the relative volume deviations for full unit-cell relaxations of a compressed and rattled 512-atom silicon supercell are shown in Fig. 5. The initial lattice constant was 4.95 Å (corresponding to a volume compression of 26% in the 512-atom supercell), and the atomic positions were randomly displaced from the diamond structure with a standard deviation of 0.21 Å. This cell was subsequently relaxed using the trust-radius enhanced Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm,[17,46,47] with a convergence criterion of $|F_{max}| \leq 0.01$ eV/Å. Note that this medium-sized system was deliberately chosen to allow comparison with fully converged PBE/*tight* calculations. Even here, the PBE/*min+s* based models show significant computational benefits, with 3–4 times lower memory demands and calculation times. Full timings and memory usage statistics of each method are provided in the supplementary material.

To gauge how the LPC affects the shape of the PES far from equilibrium, we can examine the optimization path PBE/*min+s*/LPC compared to PBE/*min+s* and PBE/*tight*; see Fig. 5. We observe that the initial optimization steps are nearly identical for PBE/*min+s*/LPC and PBE/*min+s*, displaying a fast expansion of the cell volume. In contrast, the PBE/*tight* benchmark approaches the equilibrium volume at a gradual pace. After the first five steps, the LPC correction steers the optimization toward the PBE/*tight*
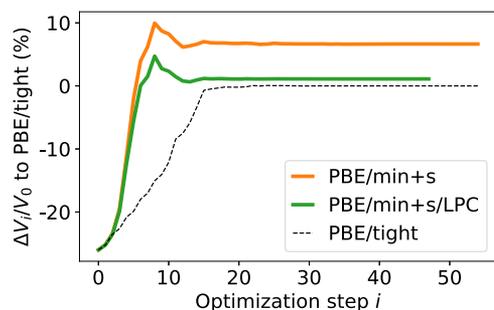
FIG. 5. Geometry optimization path obtained with LPC corrected PBE/"minimal+s" (energy cutoff −200 eV), PBE/*min+s*, and PBE/*tight* starting from a stretched and rattled Si(DIA) 512-atom supercell. For each optimization step i, the relative volume deviation compared to the relaxed reference geometry (PBE/*tight* $|F_{max}| \leq 0.01$ eV/Å) is plotted.

(a)

(b)

FIG. 6. (a) The relaxed Si(111)-DAS 7×7 surface reconstruction relaxed at the PBE/*min+s*/LPC level, in top and side views. The lower eight sub-surface layers are not shown for clarity. The Si atoms are color-coded analogously to Ref. 50: dumbbell atoms (orange), rest atoms (pink), dimer atoms (red), adatoms (green), bulk-like atoms in the top layer (dark blue), and second-layer atoms (light blue). The initial geometry was taken from Ref. 50. (b) Relaxed rest atom configurations from the Si(111)-DAS 7×7 surface overlaid with the PBE/*tight* reference (dark green).

equilibrium volume, diverging from the PBE/*min+s* path, which overestimates the final relaxed volume by 7%, compared to a 1% overestimation with PBE/*min+s*/LPC. This behavior illustrates the conservative nature of the LPC, in that it yields the correct results around equilibrium and otherwise leaves the baseline method mostly unaffected. This is an important property, because it avoids the generation of spurious minima on the PES, which can result from non-linear corrections.[48,49]

Moving beyond bulk systems, it is also important to establish the transferability of the LPC to surface systems. Surface slab calculations often require large supercells, in order to accommodate complex reconstructions and obtain sufficiently bulk-like properties for central atoms. An example for such a system is the Si(111)-dimer-adatom stacking fault (DAS) 7×7 reconstruction.[51] We use an initial structure with 1033 atoms obtained from Ref. 50. Therein, Shen *et al.* employed an ML force field to relax the structure, due to the high computational cost of first-principles methods for such systems.

We performed local relaxations of this surface structure, at both the PBE/*min+s*/LPC and PBE/*tight* levels, with a convergence criterion of $|F_{max}| \leq 0.01$ eV/Å. Importantly, the PBE/*min+s*/LPC structure displays all the characteristic features of the DAS 7×7 reconstruction: two triangular faulted and unfaulted half unit cells with adatoms, rest atoms, and dimers; see Fig. 6. Beyond these qualitative features, the structure is also quantitatively in good agreement with the PBE/*tight* benchmark: The distances between the adatoms shown deviate on average by 0.37%, the dimer bond lengths deviate on average by 1.7%, and the distances of the rest atoms to the atoms in the top layer deviate on average by 2.1%. Absolute distance errors can be found in the supplementary material.

Structural relaxations yield the ground state geometry (or some meta-stable state) of a system at $T = 0$ K. In practice, we are often also interested in finite temperature properties, however. These can be accessed via molecular dynamics (MD) simulations, which by definition also explore non-equilibrium regions of the PES. To explore the performance of PBE/*min+s*/LPC in this setting, we performed MD simulations for bulk copper at $T = 100, 300, 500,$ and 2000 K. For each method, a 108-atom fcc-Cu supercell was
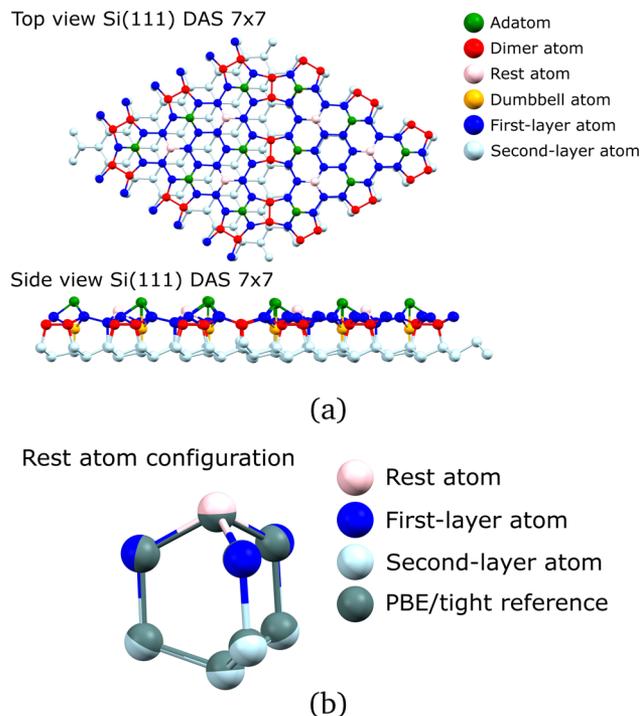
first relaxed and subsequently equilibrated for 3 ps in the NVT ensemble (using the Nosé–Hoover thermostat), followed by 3 ps production runs in the NVE ensemble. A 3 fs time step was used throughout.

The corresponding radial distribution functions (RDFs) are shown in Fig. 7. Due to its underbinding tendencies (and consequently too large unit cell), the PBE/*min+s* model significantly overestimates interatomic distances at low to moderate temperatures (100, 300, and 500 K). In contrast, the corrected PBE/*min+s*/LPC model faithfully reproduces the PBE/*light* reference, albeit with slightly broadened features at 100 and 300 K. This trend reverses somewhat as the RDF features broaden with higher temperatures so that agreement between PBE/*min+s*/LPC and the reference is essentially perfect at 500 K, and the RDF is slightly overstructured at 2000 K. Overall, these simulations show that the LPC is also beneficial in molecular dynamics simulations, despite the focus on equilibrium volumes in the parameterization. Furthermore, these simulations indicate that the LPC has no adverse effects on the numerical stability of dynamics trajectory, as energy conservation during NVE simulations is unaffected; see Fig. 7 of the supplementary material.
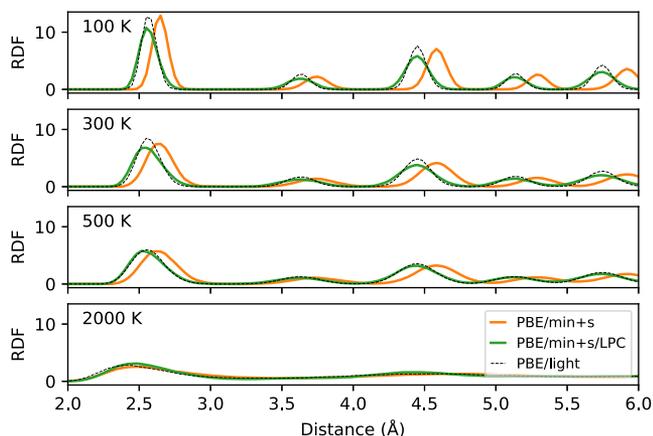
FIG. 7. Performance of PBE/*min+s*/LPC for MD simulations of bulk copper compared to PBE/*min+s* and PBE/*light*. Radial distribution functions for 108 atom supercells in the NVE ensemble after equilibrating at T = 100, 300, 500, and 2000 K averaged over 600 snapshots after equilibration.

IV. CONCLUSION

In this paper, we proposed a semiempirical small basis set density functional method for cost-efficient, large-scale material simulations denoted PBE/*min+s*/LPC. This method is parameterized for elements up to radon (Z = 1–86, excluding the lanthanide series). The method employs a well-balanced, near-minimal *min+s* NAO basis set, which leads to significant savings in terms of computational cost and memory demand, compared to fully converged calculations. In order to address the systematic overestimation of bond lengths caused by basis set incompleteness, a minimally invasive pairwise correction is used. The resulting method reliably provides accurate equilibrium volumes for mono- and poly-elemental crystals in diverse bonding situations. Despite focusing on equilibrium structures for the parameterization, PBE/*min+s*/LPC does not deteriorate the quality of the baseline method when out of equilibrium (e.g., for distorted structures or in MD simulations). While the proposed method is geared toward use in the FHI-aims code, the underlying concepts could easily be transferred to other codes using atom centered basis functions. The LPC approach is also applicable to other functionals, since the underlying cause (underbinding due to basis set incompleteness) is not functional dependent. However, it would be advisable to adjust the parameters when using different functionals or basis sets.

It should be noted that SQM methods are commonly benchmarked (and often also fitted) with respect to experimental structures or higher-level reference methods. The current method uses a different approach by fitting on converged PBE data instead. This is because one of the main use cases of the method is to pre-relax structures for subsequent calculations with a converged basis. In this context, being close to the target method is more important than being close to experiment. PBE itself has of course been benchmarked extensively to higher level computational methods as well as experiment for solid-state and molecular properties.[2,21] Extending PBE/*min+s*/LPC with a dispersion correction is also a worthy goal. Here, a careful balance between BSSE effects (which mimic dispersion to a certain extent) and the dispersion correction must be achieved. This is beyond the scope of the current work.

We envision that methods such as PBE/*min+s*/LPC will be useful in the space between fully converged first-principles methods (which offer high accuracy at a high computational cost) and efficient ML potentials (which are computationally efficient but not always reliable, depending on the availability of adequate training data). For example, they can be used for initial relaxations or MD trajectories to generate realistic atomistic configurations for training an ML potential. In this case, the systematic volume errors of pure PBE/*min+s* would be problematic, because they would bias the configurations away from the target region. PBE/*min+s*/LPC can also be useful as a pre-relaxation method, when fully converged DFT structures are required.

SUPPLEMENTARY MATERIAL

The supplementary material contains additional figures on performance and timings and also contains parameters for elements with Z = 1–86.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Elisabeth Keller: Formal analysis (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal). **Jack Morgenstein**: Methodology (supporting); Software (supporting). **Karsten Reuter**: Funding acquisition (equal); Supervision (equal). **Johannes T. Margraf**: Conceptualization (equal); Formal analysis (equal); Methodology (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

REFERENCES

[1] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865–3868 (1996).
[2] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán, Phys. Rev. B **79**, 155107 (2009).

[3] A. Marek, V. Blum, R. Johanni, V. Havu, B. Lang, T. Auckenthaler, A. Heinecke, H.-J. Bungartz, and H. Lederer, J. Phys.: Condens. Matter **26**, 213201 (2014).

[4] V. Hernandez, J. E. Roman, and V. Vidal, ACM Trans. Math. Softw. **31**, 351–362 (2005).

[5] W. Dawson and T. Nakajima, Comput. Phys. Commun. **225**, 154–165 (2018).

[6] J. J. P. Stewart, J. Mol. Model. **13**, 1173–1213 (2007).

[7] S. Grimme, C. Bannwarth, and P. J. Shushkov, J. Chem. Theory Comput. **13**, 1989–2009 (2017).

[8] C. Bannwarth, S. Ehlert, and S. Grimme, J. Chem. Theory Comput. **15**, 1652–1671 (2019).

[9] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner, Phys. Rev. B **51**, 12947–12957 (1995).

[10] G. Seifert, D. Porezag, and T. Frauenheim, Int. J. Quantum Chem. **58**, 185–192 (1996).

[11] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Phys. Rev. B **58**, 7260–7268 (1998).

[12] M. Cui, K. Reuter, and J. T. Margraf, J. Chem. Theory Comput. **20**, 5276–5290 (2024).

[13] R. Sure and S. Grimme, J. Comput. Chem. **34**, 1672–1685 (2013).

[14] E. Caldeweyher and J. G. Brandenburg, J. Phys.: Condens. Matter **30**, 213001 (2018).

[15] S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen, J. Chem. Phys. **143**, 054107 (2015).

[16] S. Grimme, A. Hansen, S. Ehlert, and J.-M. Mewes, J. Chem. Phys. **154**, 064103 (2021).

[17] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175–2196 (2009).

[18] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, APL Mater. **1**, 011002 (2013).

[19] M. Esters, C. Oses, S. Divilov, H. Eckert, R. Friedrich, D. Hicks, M. J. Mehl, F. Rose, A. Smolyanyuk, A. Calzolari, X. Campilongo, C. Toher, and S. Curtarolo, Comput. Mater. Sci. **216**, 111808 (2023).

[20] J. Shen, S. D. Griesemer, A. Gopakumar, B. Baldassarri, J. E. Saal, M. Aykol, V. I. Hegde, and C. Wolverton, J. Phys.: Mater. **5**, 031001 (2022).

[21] K. Lejaeghere, V. Van Speybroeck, G. Van Oost, and S. Cottenier, Crit. Rev. Solid State Mater. Sci. **39**, 1–24 (2014).

[22] K. Koepernik and H. Eschrig, Phys. Rev. B **59**, 1743–1757 (1999).

[23] V. W.-z. Yu, J. Moussa, and V. J. Blum, J. Chem. Phys. **154**, 224107 (2021).

[24] Z. Wang and F. Neese, J. Chem. Phys. **158**, 184102 (2023).

[25] M. Müller, A. Hansen, and S. Grimme, J. Chem. Phys. **159**, 164108 (2023).

[26] J. C. A. Prentice et al., J. Chem. Phys. **152**, 174111 (2020).

[27] W. P. Huhn and V. Blum, Phys. Rev. Mater. **1**, 033803 (2017).

[28] S. Vuckovic and K. J. Burke, J. Phys. Chem. Lett. **11**, 9957–9964 (2020).

[29] F. D. Murnaghan, Proc. Natl. Acad. Sci. U. S. A. **30**, 244–247 (1944).

[30] F. Birch, Phys. Rev. **71**, 809–824 (1947).

[31] A. H. Larsen et al., J. Phys.: Condens. Matter **29**, 273002 (2017).

[32] J. W. Pitera and J. D. Chodera, J. Chem. Theory Comput. **8**, 3445–3451 (2012).

[33] A. D. White and G. A. Voth, J. Chem. Theory Comput. **10**, 3023–3030 (2014).

[34] A. P. Bartók and G. Csányi, Int. J. Quantum Chem. **115**, 1051–1057 (2015).

[35] P. Pyykkö and M. Atsumi, Chem.-Eur. J. **15**, 186–197 (2009).

[36] B. Cordero, V. Gómez, A. E. Platero-Prats, M. Revés, J. Echeverría, E. Cremades, F. Barragán, and S. Alvarez, Dalton Trans. **2008**, 2832–2838.

[37] A. Bondi, J. Phys. Chem. **68**, 441–451 (1964).

[38] M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. A **113**, 5806–5812 (2009).

[39] B. Cheng, R.-R. Griffiths, S. Wengert, C. Kunkel, T. Stenczel, B. Zhu, V. L. Deringer, N. Bernstein, J. T. Margraf, K. Reuter, and G. Csanyi, Acc. Chem. Res. **53**, 1981–1991 (2020).

[40] F. Gao and L. Han, Comput. Optim. Appl. **51**, 259–277 (2012).

[41] P. Virtanen et al., Nat. Methods **17**, 261–272 (2020).

[42] H. Kruse and S. Grimme, J. Chem. Phys. **136**, 154101 (2012).

[43] I. E. Castelli, J. M. García-Lastra, K. S. Thygesen, and K. W. Jacobsen, APL Mater. **2**, 081514 (2014).

[44] M. Pandey and K. W. Jacobsen, Phys. Rev. Mater. **2**, 105402 (2018).

[45] A. Bruix, J. T. Margraf, M. Andersen, and K. Reuter, Nat. Catal. **2**, 659–670 (2019).

[46] J. Nocedal and S. J. Wright, *Numerical Optimization*, *Springer Series in Operation Research and Financial Engineering*, 2nd ed. (Springer, New York, 2006).

[47] B. G. Pfrommer, M. Côté, S. G. Louie, and M. L. Cohen, J. Comput. Phys. **131**, 233–240 (1997).

[48] J. J. P. Stewart, J. Mol. Model. **10**, 155–164 (2004).

[49] R. W. Janes and R. A. Palmer, J. Mol. Struct.: THEOCHEM **339**, 95–101 (1995).

[50] Y. Shen, S. I. Morozov, K. Luo, Q. An, and W. A. Goddard III, J. Am. Chem. Soc. **145**, 20511–20520 (2023).

[51] K. Takayanagi, Y. Tanishiro, S. Takahashi, and M. Takahashi, Surf. Sci. **164**, 367–392 (1985).