AARHUS UNIVERSITY

Coversheet

This is the publisher's PDF (Version of Record) of the article.
This is the final published version of the article.

How to cite this publication:

S. Chiriki, MP. V. Christiansen, and B. Hammer (2019) 'Constructing convex energy landscapes for atomistic structure optimization'. Phys. Rev. B 100, 235436, DOI: https://doi.org/10.1103/PhysRevB.100.235436

Publication metadata

Constructing convex energy landscapes for atomistic structure optimization

Siva Chiriki [ORCID], Mads-Peter V. Christiansen [ORCID], and B. Hammer [ORCID][*]

Department of Physics and Astronomy, Aarhus University, DK-8000 Aarhus C, Denmark

We propose a global optimization strategy for atomistic structure determination based on two new concepts: a few-atom complementary energy landscape and atomic role models. Global optimization of costly energy expressions may be aided by performing some of the optimization on model energy landscapes. These are often based on a sum-of-atomic-contributions form that accurately reproduces every local energy minimum of the true energy expression. However, we propose that, by not including all atomic contributions, the resulting energy landscapes may become more convex, making the search for the global optimum more facile. A role model is someone we aspire to be more like; in the same vein we define the role model of an atom to be another atom whose local environment the first atom seeks to obtain itself. Basing a complementary energy landscape on the distance of some atoms from their role models in a feature space, we arrive at a useful few-atom complementary energy landscape. We show that relaxation in this landscape is an effective mutation when employed in an evolutionary algorithm used to identify the bulk cristobalite structure of SiO_2 and the (1×4) surface reconstruction of anatase $TiO_2(001)$.

I. INTRODUCTION

Computational screening and property prediction have become important methods for novel materials design [1–4]. For some problems, dynamic and entropic effects may be of importance for making reliable predictions [5], but often, knowing the minimum energy configuration of the material is sufficient for making reliable predictions. This typically means that the lowest minimum in a high-dimensional potential energy surface must be identified.

The prevailing strategy for solving such global optimization problems consists of an algorithm that creates new possible candidate configurations and a method to evaluate whether these new configurations are more stable. There are many successful algorithms such as simulated annealing [6], basin and minima hopping [7,8], evolutionary algorithms [9–16], particle-swarm algorithms [17,18], and random structure searching [19]. The method of choice for evaluating the energy has been first-principles density functional theory (DFT), as it is adequately accurate while being computationally feasible for many systems of interest. However, replacing the DFT calculations with something less expensive is an appealing prospect and has recently become feasible with the adoption of machine learning methods (for reviews, see Refs. [20,21]). This has made it possible to replace DFT calculations with database-trained machine learning models, based on, for example, kernel methods [22–24] or neural networks [25–33]. A database of calculations is expensive to generate and therefore methods have been developed to generate the required data for surrogate models on the fly, minimizing the amount of data required, for applications in potential fitting [34–36], molecular dynamics [37–39], nudged elastic

band calculations [40–42], and indeed local [43–45] and global optimization [46–51].

Machine learning methods rely on descriptors that transform the Cartesian coordinates defining the atomic structure into vectors that can be fed into these models. Such feature vectors should encode the symmetries (translational, rotational, and permutational) that the potential energy obeys—symmetries that are not respected by the Cartesian coordinates. A range of such descriptors has been introduced in two main categories: local descriptors [25,52–56] that describe the local environment (e.g., density of atoms in a neighborhood) and global descriptors [23,57–60] that describe an entire structure.

Global optimization algorithms, such as basin-hopping and evolutionary algorithms, rely on updates to the atomic coordinates to escape local minima and discover other more stable local minima, eventually finding the global minimum. This could be a rattling of all the atomic coordinates, where the direction and magnitude of the individual displacements are random. The stochastic nature of the coordinate updates ensures exploration of the configuration space [61–63]. The performance of these global optimization techniques can thus be increased by using updates that, by incorporating the available information, maximize the chance of finding new more stable minima. This is what is done by evolutionary algorithms that utilize sophisticated crossover operations between members of a population of locally optimal structures.

Local relaxation uses the atomic forces to direct structures towards a local minimum. If the potential energy surface (PES) being optimized were convex, there would be only one minimum, the global one, and local optimization would lead to it. However, as PESs have in general many local minima, local optimization techniques usually only lead to the global minimum if the searches are started from within the hypervolume of the global minimum. In a recent

paper, Pickard proposed to add extra dimensions that become increasingly unfavorable to occupy, allowing atoms to pass through each other during relaxation, effectively increasing the hypervolume of the global minimum [64]. This can be viewed as performing the relaxation in a different landscape where the extra dimensions allow barriers to be circumpassed.

In this paper we propose that for energy landscapes defined as a sum of local contributions an efficient global minimization strategy is to locally optimize—not according to the total energy, but rather according to the sum over a subset of the local energies. This removes some local minima, effectively creating a more convex landscape that is easier to globally optimize. Quantum mechanical (QM) methods only supply the total energy and no unique way exists to decompose the total energy into a sum of atomic contributions [65]. However, many machine learning methods for reproducing quantum mechanical energies have been proposed in the form of a sum of atomic contributions [22,25,26], which opens up for the use of the method we propose here. We demonstrate the usefulness of the method when applied in combination with a recently proposed method, *cluster regularization* [66]. The cluster regularization method builds on the observation that as more and more stable configurations are identified during an evolutionary structure search, more and more clustering is observed of the atoms when represented in a feature space. Sørensen *et al.* [66] showed that this may be reversed into: by enforcing clustering of atoms in feature space, stable structures emerge faster during the evolutionary search. A cluster-distance expression was introduced that measures the sum of distances between the feature representation of all local atomic environments and their corresponding cluster centers. It was shown that this expression correlates with the energy causing its minimization to lead to effective coordinate updates. During these updates, the atoms are moved towards their local cluster centers that thus act as *attractors*. In the present work, we generalize this expression to what we call a *complementary energy* (CE). Rather than using cluster centers as the points in feature space that distances are measured from, we randomly pick attractors among the atoms in any given structure. We coin those atoms *role models*.

The paper is outlined as follows. First, the complementary energy landscape is explored in Sec. II along with the definition of role models and an example of how to apply the idea of only including certain terms in the energy expression. In Sec. III an evolutionary algorithm (EA) that incorporates the complementary energy landscape with few-atom minimizations as a mutation is presented. This is followed by demonstration runs of the EA on one SiO_2 system and three TiO_2 systems of increasing difficulty in Sec. IV. The paper ends with a concluding Sec. V.

II. COMPLEMENTARY ENERGY

A common approach when modeling potential energy surfaces is to express the total energy as a sum of local energy contributions, e.g., as a sum of atomic energies. In such a setup the total energy can be written as

$$E = \sum_i E_i. \tag{1}$$

A wide variety of models can be written in this way, such as pair potentials but also highly accurate machine learning models, be it neural networks [25] or kernel methods [22]. We propose that for the purpose of global optimization it can be beneficial to not necessarily include all terms in such an expression when evaluating the gradient. In doing so, the model potential changes character from being a surrogate energy landscape mimicking every aspect of the true energy landscape to being a landscape that hopefully shares a global minimum with the true energy landscape. We call this kind of landscape a complementary energy landscape.

The construction of potentials is facilitated by descriptors, as a way to mathematically represent atomic structures, that encode symmetries present in the Hamiltonian governing the system. The outputs of such descriptors are typically feature vectors that describe the local environment around an atom. In order to easily illustrate our findings in this work, we choose an exceedingly simple feature representation. The feature vectors we use to describe the local environment of atoms in a two-component system, with atomic types A and B, are simply

$$\mathbf{f}_i = \left[\rho_i^A, \rho_i^B, Z_i \right], \tag{2}$$

where Z_i is the atomic number of atom i and where ρ_i^Z describes the atomic density of species Z in the neighborhood of atom i given as

$$\rho_i^Z = \sum_{j \neq i, Z_j = Z} e^{-r_{ij}/\lambda} g_c(r_{ij}), \tag{3}$$

where r_{ij} is the distance between atoms i and j, λ is a length scale, and $g_c(r)$ is a cutoff function that smoothly goes to zero at the cutoff radius r_c given by

$$g_c(r) = \begin{cases} \frac{1}{2}\cos(\pi \frac{r}{r_c}) + \frac{1}{2}, & r \leqslant r_c \\ 0, & r > r_c. \end{cases} \tag{4}$$

For all figures and results presented in this work we have chosen $\lambda = 1$ Å and $r_c = 11.9$ Å.

Having chosen an expression for the feature vectors, a CE landscape can be formulated as

$$CE_{all} = \sum_i |\mathbf{f}_i - \mathbf{A}_i|, \tag{5}$$

where $\mathbf{f}_i$ is the feature vector of atom i and $\mathbf{A}_i$ is the attractor assigned to atom i. The key to the usefulness of a complementary energy landscape obviously lies in the choice of attractors. Assuming a set, J, of attractors has been decided on, the assignment may be done so that every atom has its nearest neighbor assigned to it:

$$CE_{all} = \sum_i \min_{j \in J} |\mathbf{f}_i - \mathbf{A}_j|. \tag{6}$$

The gradient of CE_{all} with respect to the Cartesian coordinates of the atoms can be evaluated and CE_{all} may thus be minimized using gradient descent methods. During these minimization steps, we keep the assignments of attractors fixed and hence avoid cusps in the complementary energy landscape that would otherwise potentially develop if some atoms were to obtain new nearest attractors in the course of the minimization.
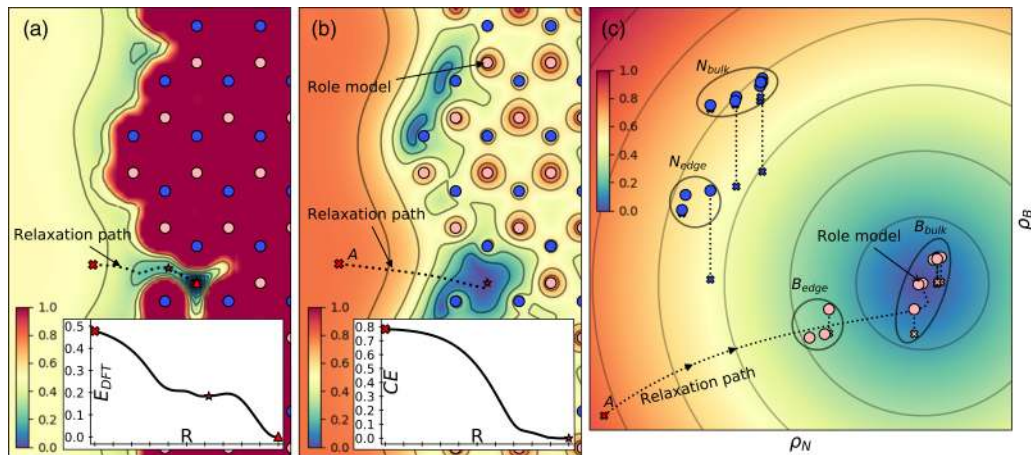
FIG. 1. (a, b) A hexagonal boron nitride nanoribbon where a boron atom has been removed from a bulk position. Coloring indicates the energy calculated from DFT and CE, respectively, the DFT energy has been cropped at the global minimum energy plus 30 eV, and both DFT and CE energies have been rescaled to be between zero and 1. Insets show the energy profiles along the relaxation paths in these landscapes. As the DFT relaxation gets stuck in the local minimum energy configuration indicated by the star, the energy beyond this point is shown along a straight line from the star to the global minimum configuration indicated by the triangle. (c) The feature-space representation of the CE relaxation procedure is shown; here ρ_N and ρ_B are feature coordinates that measure the density of nitrogen and boron atoms. The movement of a single atom in real space results in changes to the feature-space representation of several atoms. The same rescaling as in (b) has been applied.

Given Eq. (6), it still remains to decide on the set of attractors. The feature vectors of atoms appearing in the sought-after global minimum energy structure would be the optimum choice as they could lead to a zero right-hand side in Eq. (6). Obviously, these feature vectors are not available until the search has completed, and the attractors must be chosen from what is at hand during the search. In Ref. [66] the attractors were chosen as a small number of cluster centers emerging from an unsupervised clustering of all atoms in three parent structures present in the population of an evolutionary search. In this work, we rather propose to establish the set of attractors from the actually occurring atomic features in a single parent structure. We denote such atoms *role models* and the attractors are the features of the role models. An atom aspires to become more like its role model (in terms of the similarity of their local environments), and Eq. (6) gives a formal way of realizing that aspiration. An advantage of the method proposed in this work is that the attractors are points in feature space that are indeed realizable (they already exist in some structure), which is not guaranteed if cluster centers are used as in Ref. [66]. An obvious example of this is clustering points distributed on the surface of a sphere that yields a cluster center close to the center of the sphere, but far away from all the points.

Since the complementary energy is of the form of Eq. (1), the idea of only including certain terms in the summation can further be applied. To do so, one would only sum over atoms in some set I, and the *few-atom* complementary energy landscape takes the form

$$CE_{few} = \sum_{i \in I} \min_{j \in J} |\mathbf{f}_i - \mathbf{A}_j|. \tag{7}$$

We discuss below how the set of atoms may be chosen, but we start with an exceedingly simple example where only one atom is misplaced.

Example with one misplaced atom

To illustrate how the method works in a case where the optimum attractor can be inferred, we present in Fig. 1 an
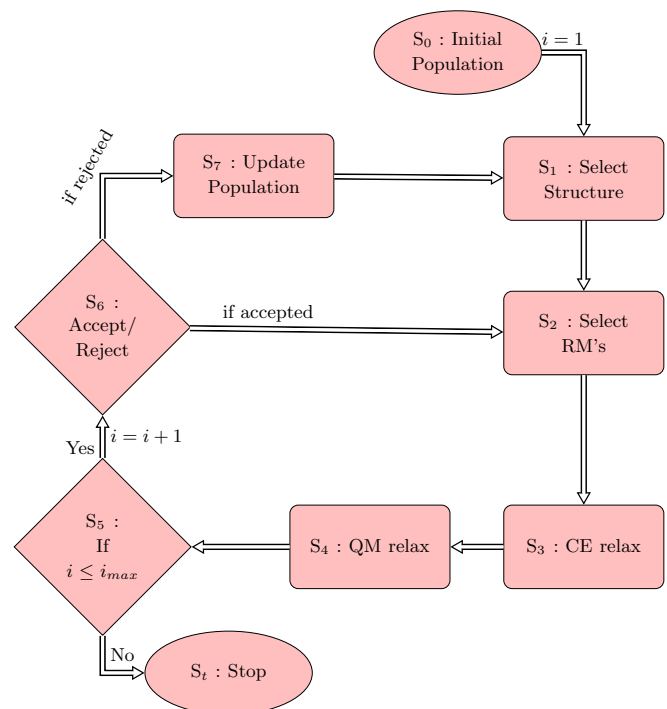
FIG. 2. Flow chart of the complementary energy assisted evolutionary algorithm. Algorithm starts with step S_0, proceeds in a loop (steps S_1–S_7), and stops if the number of iterations reaches a maximum (i_{max}). After selecting a structure, a set of role models is chosen and a CE relaxation is performed followed by a relaxation with the quantum mechanical energy, which is repeated until the energy no longer improves.
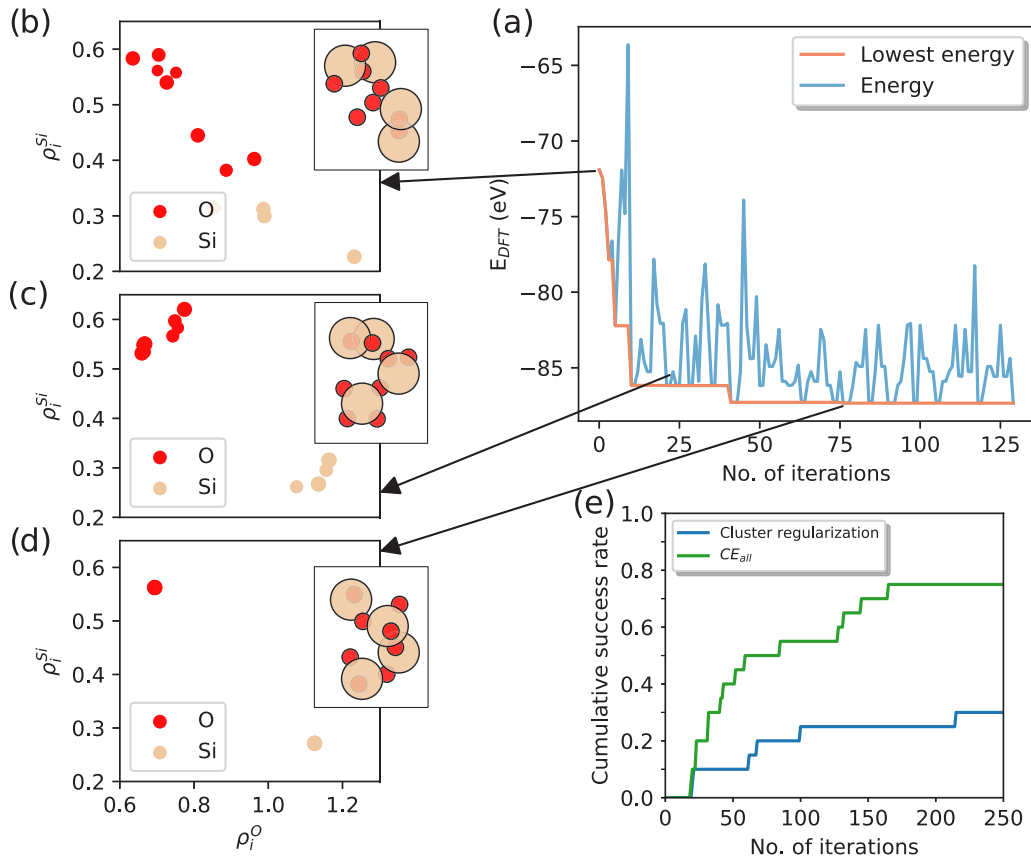
FIG. 3. (a) Potential energy as a function of iteration, along with best energy found. Feature space representation of structures at (b) 0, (c) 23, and (d) 76 iterations along with corresponding geometries in the insets. (e) Success rate for cluster regularization and CE_{all} applied on a SiO_2 cristobalite-type structure.

example where a boron atom has been removed from the bulk of a hBN nanoribbon and placed outside of it. In Fig. 1(a) the true DFT energy is used to relax the atom, but there is an energy barrier blocking the path to the global minimum. With a single-atom CE relaxation method, Eq. (7), the global minimum structure is reachable as shown in Fig. 1(b), where the role model acting as the attractor is a boron atom in a bulk position. Finally Fig. 1(c) shows the feature-space representation of the CE relaxation. Note that only one boron atom is allowed to move in real space but this causes other atoms to move in feature space. The local minimum that the DFT relaxation finds is likely due to obtaining bonds to two nitrogen atoms, a similar bonding configuration as that of an edge atom; the CE relaxation also moves through this position but does not recognize it as a local minimum due to the convex property of the CE as seen from the feature-space view in Fig. 1(c). The CE is able to move the boron atom to its optimal position effectively, whereas relaxations with the real energy encounters barriers and standard perturbations, such as a rattle mutation, are unlikely to move the boron atom to a position where a normal relaxation would lead to the optimal bulk position.

In Fig. 1, the role model was decided on manually to make the best possible case for the illustration. In practice, role models will not be known before a search is initialized, but may, as illustrated in the next section, be chosen stochastically during the search.

III. EVOLUTIONARY ALGORITHM

The complementary energy relaxation has been employed as a mutation in an evolutionary algorithm; the flow chart for this algorithm is shown in Fig. 2. The algorithm is initialized by generating ten locally relaxed structures that are employed as the initial population (S_0 in the flow chart). The main loop of the algorithm consists of picking a structure from the population and applying CE and QM relaxations sequentially. The CE relaxation steps require the selection of a number of role models. These are taken from the feature vectors of a number of atoms appearing in the single structure picked from the population. In order to avoid choosing essentially equivalent role models, a filter is applied requiring a 0.01 minimum distance in feature space. After filtering, R_1 role models are picked randomly from the remaining features as attractors for the next CE step, where R_1 is a random number between a predefined lower and upper bound that should be decided upon based on the system studied. The change in number of role models in each iteration allows the algorithm to escape from local minima. In the present work, both few-atom CE and all-atom CE have been employed; in the case of few-atom relaxation, a random number R_2 between 1 and the total number of atoms is chosen and the R_2 atoms with the highest local CE are included in the relaxation with the remaining constrained to their current positions. This results in a new candidate structure that is relaxed with the QM
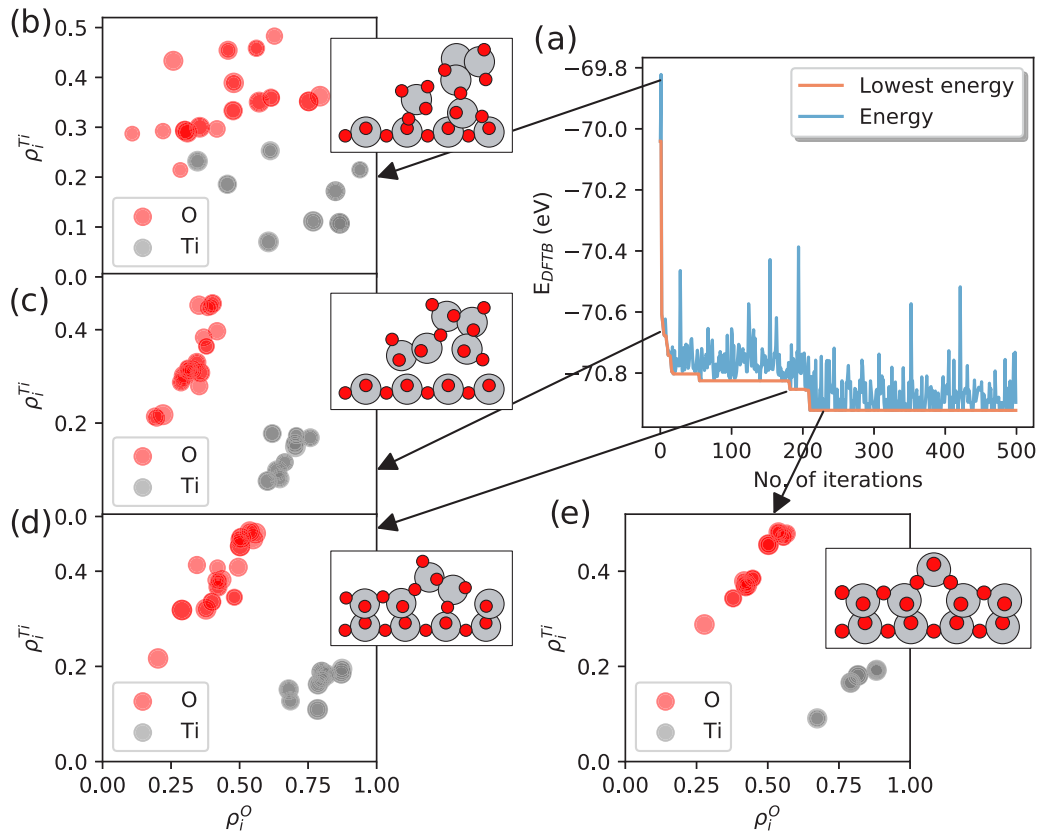
FIG. 4. (a) Potential energy as a function of iteration, along with best energy found. Feature space representation of structures after (b) 0, (c) 4, (d) 181, and (e) 230 iterations along with corresponding geometries in the insets.

method of choice. If the energy of the resulting structure is lower than before the CE relaxation another CE-QM step is performed; if not, the most recently accepted structure (if any) is added to the population and the most energetic member of the population is discarded. Finally, the process may be repeated, starting the CE-QM relaxations for a new parent structure chosen at random. The algorithm continues to update the population members for a predetermined number of steps, and may be started a great number of times to establish reliable statistics on its performance.

IV. RESULTS

The evolutionary algorithm described in the previous section has been applied on two kinds of systems: (i) a bulk system, namely, cristobalite-type SiO_2, and (ii) a TiO_2-anatase surface reconstruction with increasing number of layers. For both systems our proposed CE method is compared to the cluster regularization method. In the case of a SiO_2 cristobalite-type structure, the experimental cell parameters ($a = b = 4.52$ Å, $c = 7.10$ Å, and $\alpha = \beta = \gamma = 90°$) were optimized with linear combination of atomic orbitals DFT with the Perdew-Burke-Ernzerhof functional as implemented in GPAW, which is also the level of theory used for QM relaxations [67,68]. The optimized cell parameters are $a = b = 4.75$ Å, $c = 6.85$ Å, and $\alpha = \beta = \gamma = 90°$ with a $2 \times 2 \times 2$ **k**-point grid for Brillouin zone sampling. The number of role models is chosen randomly between 2 and 4. The success

curves are shown in Fig. 3(e) for both the methods. Along with success curves, the energy vs iteration plots for one of the restarts with feature representation at 0, 23, and 76 iterations are provided in Figs. 3(a)–3(d). For the initial structure, the feature space is scattered and, after a few iterations, it converts to the fewer number of features as shown in Figs. 3(b) and 3(c). In Fig. 3(a), the algorithm starts finding the global minimum after 76 iterations, containing only two distinct features shown in Fig. 3(d). The success curves shown in Fig. 3(e) indicate that CE_{all} achieves 75% success around 180 iterations whereas cluster regularization [66] only achieved 30% success around 220 iterations. The minimum number of iterations required before finding the global minimum in one restart is the same for both algorithms. The lower success rate for cluster regularization may be due to slow convergence of cluster centers in feature space.

Our second test system is the anatase $TiO_2(001)$-(1×4) surface reconstruction with one, two, and three atomic layers. For this system density-functional tight binding (DFTB) as implemented in DFTB+ is used [69]. DFTB has been used to study TiO_2 systems previously, in cases where either the system size or the number of calculations has been too large for a full DFT treatment [70,71]. We use DFTB parameters from Dolgonos *et al.* and lattice parameters $a = 3.94$ Å and $c = 9.47$ Å together with a 2×1 **k**-point grid for Brillouin zone sampling [72]. We refrain, however, from using the self-consistent-charge correction. The number of role models is chosen randomly between 4 and 6. To understand the
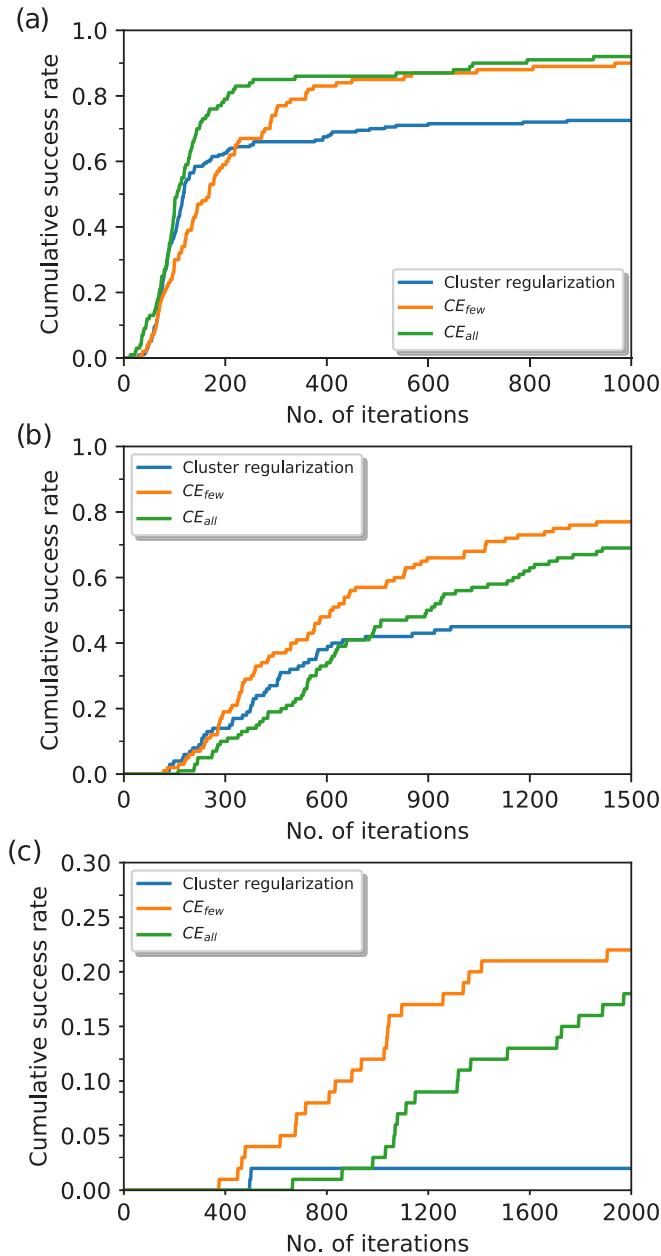
FIG. 5. Cumulative success rate for (a) two-layer, (b) three-layer, and (c) four-layer anatase TiO$_2$ systems. The CE role model method outperforms the cluster regularization method especially for the most difficult four-layer system.

mechanism of the algorithm we first analyze a trajectory for one run of the algorithm on the two-layer anatase-TiO$_2$-(001) surface in Fig. 4. The energy as a function of the number of iterations is shown in Fig. 4(a) along with the lowest energy found thus far. Initially the energy is very high, and the feature representation of the structure is scattered. The algorithm quickly enforces a less scattered feature space by moving atoms towards environments equivalent to that of their role models, as seen from the structure in Fig. 4(c), which has a much less scattered feature space, even though the found structure is separated from the fixed bottom layer. As the algorithm progresses, the feature representation becomes less scattered while the best energy found decreases. The

global minimum structure is found after 230 iterations. This structure has a well-ordered feature representation with six distinct clusters of features, as shown in Fig. 4(e), which the CE relaxation efficiently enforces.

In order to validate the performance of the complementary energy assisted evolutionary algorithm we benchmarked it against the previously used cluster regularization method [66] utilizing either the all-atom CE or the few-atom CE approach. Both methods were used in predicting the structure of a reconstructed two-, three-, and four-layer anatase TiO$_2$(001) surface and the resulting success plots are shown in Fig. 5. The success plots are obtained from 100 restarts for each TiO$_2$ system for all three algorithms. The number of iterations is varied depending on the system size; thus 1000, 1500, and 2000 iterations are carried out for two-, three-, and four-layer TiO$_2$ systems, respectively.

As illustrated in Fig. 5(a) for the two-layer system, both CE methods achieve 85% success, whereas cluster regularization saturates around 65% success. For this small system the CE$_{all}$ method has a steeper success curve than the CE$_{few}$. The performance gain is due to choosing attractors from role models, as both few- and all-atom CE relaxations obtain similar success rates. However, for the case of the three-layer TiO$_2$ system, the CE$_{few}$ method reaches a higher success rate than CE$_{all}$ as well as cluster regularization as shown in Fig. 5(b). The cluster regularization method achieves a 40% success rate whereas the CE$_{few}$ method achieves an ~80% success rate and both methods start finding the global minimum after the same number of iterations. The success curve for the four-layer TiO$_2$ system is shown in Fig. 5(c); here the CE$_{few}$ method starts finding the global minimum sooner than the CE$_{all}$ method and both methods convincingly outperform the cluster regularization method. These results show that, especially for large systems, the few-atom CE strategy is beneficial.

V. CONCLUSION

In conclusion, a method of generating candidate structures for global optimization searches has been presented. The method relies on a complementary energy landscape defined through distances in a feature space. The method is aided by only including a subset of the atoms in the complementary energy expression. The concept of atomic role models has been introduced as a way to choose the attractors that parametrize the complementary energy. The role models need not be known prior to a search, but may be chosen stochastically among atoms in a population of favorable structures encountered during an ongoing search. The method has been implemented in an evolutionary algorithm and applied to systems of varying difficulty, where it leads to increased performance. The method can be used in any global optimization algorithm that relies on perturbing the atomic coordinates to generate new candidate structures. The complementary energy can be viewed as an intentionally inaccurate approximate potential energy landscape; however, the few-atom relaxation method could also be applied to more accurate surrogate potentials, e.g., those produced by neural networks or kernel methods that can be written as a sum of local contributions. This also opens the possibility of choosing which atoms to include in

the complementary energy relaxation based on "real" local energies from a surrogate model rather than based on the local *complementary* energies, possibly further increasing the performance of the method.

[1] J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff, and J. K. Nørskov, Nat. Mater. **5**, 909 (2006).

[2] C. J. Pickard and R. J. Needs, Phys. Rev. Lett. **97**, 045504 (2006).

[3] A. R. Organov, J. Chen, C. Gatti, Y. Ma, Y. Ma, C. W. Glass, Z. Liu, T. Yu, O. O. Kurakevych, and V. L. Solozhenko, Nature (London) **457**, 863 (2009).

[4] Z. A. Piazza, H.-S. Hu, W.-L. Li, Y.-F. Zhao, J. Li, and L.-S. Wang, Nat. Commun. **5**, 3113 (2014).

[5] C. Toher, C. Oses, D. Hicks, and S. Curtarolo, npj Comput. Mater. **5**, 69 (2019).

[6] S. Kirkpatrick, S. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).

[7] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A. **101**, 5111 (1997).

[8] S. Goedecker, J. Chem. Phys. **120**, 9911 (2004).

[9] R. L. Johnston, Dalton Trans. **22**, 4193 (2003).

[10] A. R. Oganov and C. W. Glass, J. Chem. Phys. **124**, 244704 (2006).

[11] S. M. Woodley and C. Catlow, Comput. Mater. Sci. **45**, 84 (2009).

[12] L. B. Vilhelmsen and B. Hammer, J. Chem. Phys. **141**, 044711 (2014).

[13] Q. Zhu, A. R. Oganov, A. O. Lyakhov, and X. Yu, Phys. Rev. B **92**, 024106 (2015).

[14] S. Lysgaard, J. S. G. Mýrdal, H. A. Hansen, and T. Vegge, Phys. Chem. Chem. Phys. **17**, 28270 (2015).

[15] B. C. Revard, W. W. Tipton, A. Yesypenko, and R. G. Hennig, Phys. Rev. B **93**, 054117 (2016).

[16] F. Curtis, X. Li, T. Rose, A. Vazquez-Mayagoitia, S. Bhattacharya, L. M. Ghiringhelli, and N. Marom, J. Chem. Theory Comput. **14**, 2246 (2018).

[17] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Phys. Rev. B **82**, 094116 (2010).

[18] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Comput. Phys. Commun. **183**, 2063 (2012).

[19] C. J. Pickard and R. Needs, J. Phys. Condens. Mat. **23**, 053201 (2011).

[20] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Nature (London) **559**, 547 (2018).

[21] M. Rupp, O. A. von Lilienfeld, and K. Burke, J. Chem. Phys. **148**, 241401 (2018).

[22] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).

[23] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Phys. Rev. Lett. **108**, 058301 (2012).

[24] G. Schmitz, I. H. Godtliebsen, and O. Christiansen, J. Chem. Phys. **150**, 244113 (2019).

[25] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

[26] J. Behler, Phys. Chem. Chem. Phys. **13**, 17930 (2011).

[27] H. Zhai and A. N. Alexandrova, J. Chem. Theory Comput. **12**, 6213 (2016).

[28] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Sci. Adv. **3**, e1603015 (2017).

[29] J. S. Smith, O. Isayev, and A. E. Roitberg, Chem. Sci. **8**, 3192 (2017).

[30] S. Jindal, S. Chiriki, and S. S. Bulusu, J. Chem. Phys. **146**, 204301 (2017).

[31] S.-D. Huang, C. Shang, P.-L. Kang, and Z.-P. Liu, Chem. Sci. **9**, 8644 (2018).

[32] C. Wang, A. Tharval, and J. R. Kitchin, Mol. Simulat. **44**, 623 (2018).

[33] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, Phys. Rev. Lett. **120**, 143001 (2018).

[34] J. S. Smith, B. Nebgen, N. Lubbers, and O. Isayev, J. Chem. Phys. **148**, 241733 (2018).

[35] Y.-H. Tang and W. A. de Jong, J. Chem. Phys. **150**, 044107 (2019).

[36] J. J. Kranz, M. Kubillus, R. Ramakrishnan, O. A. von Lilienfeld, and M. Elstner, J. Chem. Theory Comput. **14**, 2341 (2018).

[37] Z. Li, J. R. Kermode, and A. De Vita, Phys. Rev. Lett. **114**, 096405 (2015).

[38] M. Gastegger, J. Behler, and P. Marquetand, Chem. Sci. **8**, 6924 (2017).

[39] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam, Phys. Rev. Lett. **122**, 225701 (2019).

[40] A. A. Peterson, J. Chem. Phys. **145**, 074106 (2016).

[41] O.-P. Koistinen, F. B. Dagbjartsdóttir, V. Ásgeirsson, A. Vehtari, and H. Jónsson, J. Chem. Phys. **147**, 152720 (2017).

[42] J. A. Garrido Torres, P. C. Jennings, M. H. Hansen, J. R. Boes, and T. Bligaard, Phys. Rev. Lett. **122**, 156001 (2019).

[43] E. Garijo del Río, J. J. Mortensen, and K. W. Jacobsen, Phys. Rev. B **100**, 104103 (2019).

[44] A. Denzel and J. Kästner, J. Chem. Phys. **148**, 094114 (2018).

[45] G. Schmitz and O. Christiansen, J. Chem. Phys. **148**, 241704 (2018).

[46] S. Q. Wu, M. Ji, C. Z. Wang, M. C. Nguyen, X. Zhao, K. Umemoto, R. M. Wentzcovitch, and K. M. Ho, J. Phys. Condens. Matter **26**, 035402 (2013).

[47] T. K. Patra, V. Meenakshiundaram, J.-H. Hung, and D. S. Simmons, ACS Comb. Sci. **19**, 96 (2017).

[48] Q. Tong, L. Xue, J. Lv, Y. Wang, and Y. Ma, Faraday Discuss. **211**, 31 (2018).

[49] E. L. Kolsbjerg, A. A. Peterson, and B. Hammer, Phys. Rev. B **97**, 195424 (2018).

[50] M. Van den Bossche, H. Grönbeck, and B. Hammer, J. Chem. Theory Comput. **14**, 2797 (2018).

[51] M. Van den Bossche, J. Chem. Phys. **123**, 3038 (2019).

[52] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).

[53] B. Huang and O. A. von Lilienfeld, J. Chem. Phys. **145**, 161102 (2016).

[54] V. L. Deringer and G. Csányi, Phys. Rev. B **95**, 094203 (2017).

[55] N. Artrith, A. Urban, and G. Ceder, Phys. Rev. B **96**, 014112 (2017).

[56] F. A. Faber, A. S. Christensen, B. Huang, and O. A. von Lilienfeld, J. Chem. Phys. **148**, 241717 (2018).

[57] M. Valle and A. R. Oganov, Acta Crystallogr. A **66**, 507 (2010).

[58] K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. Von Lilienfeld, K.-R. Müller, and A. Tkatchenko, J. Phys. Chem. Lett. **6**, 2326 (2015).

[59] X. Chen, M. S. Jørgensen, J. Li, and B. Hammer, J. Chem. Theory Comput. **14**, 3933 (2018).

[60] S. A. Meldgaard, E. L. Kolsbjerg, and B. Hammer, J. Chem. Phys **149**, 134104 (2018).

[61] M. S. Jørgensen, U. F. Larsen, K. W. Jacobsen, and B. Hammer, J. Chem. Phys. A **122**, 1504 (2018).

[62] T. Yamashita, N. Sato, H. Kino, T. Miyake, K. Tsuda, and T. Oguchi, Phys. Rev. Materials **2**, 013803 (2018).

[63] F. Häse, C. Roch, Loïc M. Kreisbeck, and A. Aspuru-Guzik, ACS Cent. Sci. **4**, 1134 (2018).

[64] C. J. Pickard, Phys. Rev. B **99**, 054102 (2019).

[65] M. Yu, D. R. Trinkle, and R. M. Martin, Phys. Rev. B **83**, 115113 (2011).

[66] K. H. Sørensen, M. S. Jørgensen, A. Bruix, and B. Hammer, J. Chem. Phys. **148**, 241734 (2018).

[67] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen, Phys. Rev. B **71**, 035109 (2005).

[68] A. H. Larsen, M. Vanin, J. J. Mortensen, K. S. Thygesen, and K. W. Jacobsen, Phys. Rev. B **80**, 195112 (2009).

[69] B. Aradi, B. Hourahine, and T. Frauenheim, J. Phys. Chem. A **111**, 5678 (2007).

[70] D. Selli, G. Fazio, G. Seifert, and C. D. Valentin, J. Chem. Theory Comput. **13**, 3862 (2017).

[71] D. Selli, G. Fazio, and C. D. Valentin, J. Chem. Phys. **147**, 164701 (2017).

[72] G. Dolgonos, B. Aradi, N. H. Moreira, and T. Frauenheim, J. Chem. Theory Comput. **6**, 266 (2010).