# Atomic Structure Optimization with Machine-Learning Enabled Interpolation between Chemical Elements

Sami Kaappa, Casper Larsen, and Karsten Wedel Jacobsen[*]

*Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark*

We introduce a computational method for global optimization of structure and ordering in atomic systems. The method relies on interpolation between chemical elements, which is incorporated in a machine-learning structural fingerprint. The method is based on Bayesian optimization with Gaussian processes and is applied to the global optimization of Au-Cu bulk systems, Cu-Ni surfaces with CO adsorption, and Cu-Ni clusters. The method consistently identifies low-energy structures, which are likely to be the global minima of the energy. For the investigated systems with 23–66 atoms, the number of required energy and force calculations is in the range 3–75.

The atomic-level structure of a material is often of crucial importance for its mechanical, electronic, magnetic, or chemical properties. At low temperatures, the structure can, in principle, be determined computationally by minimizing the energy of the system. However, the space of different atomic configurations of a material is huge, and structure determination is therefore a long-standing challenge in computational physics.

A number of methods to address this challenge have been introduced [1], including basin hopping [2], particle swarm optimization [3], evolutionary algorithms, [4–7], and random searches [8]. However, all of these methods require a large number of energy and force calculations, which can be time consuming if performed with, for example, density-functional theory (DFT) or higher-level quantum chemistry methods.

During the last decade, machine-learning techniques have gained impact on computational material physics [9–25]. The field of global optimization of atomic structures took a major step forward recently with the GOFEE approach [26]. In this approach, Gaussian processes are used to generate a surrogate potential energy surface (PES), which is then explored by random searching and Bayesian optimization. Thereby, a speed-up of several orders of magnitude in determining the optimal structure is in some cases achieved [26]. This method was expanded with training to forces and several other modifications in the BEACON code [27].

Global optimization is particularly challenging if many local minima exist in the PES. This would typically be the case for large systems, but it also appears for smaller systems when several chemical elements are present. In many alloys, the interchange of two atoms introduces another local minimum in the potential energy surface, but with a different energy. The number of local minima

thus becomes the number of different ordering permutations of the atoms in the lattice, which can be huge. For example, in a 64-atom unit cell with equal amounts of two different types of atoms, the number of possible permutations is on the order of $10^{18}$, and it is, of course, not possible for any algorithm to explore all these configurations.

Here we address the challenge of many local minima by interpolation between chemical elements ("ICE"). With this approach, atoms of different chemical elements can switch place not by moving in real space but by gradually changing their chemical identities. This leads to a much reduced or even vanishing barrier for the process. The idea of interpolating in chemical space has also been used in other contexts, including catalyst design [28], in combination with perturbation theory [29–32], to treat disordered alloys with the coherent potential approximation [33–35] and to incorporate chemical information in the descriptors of machine-learning models to explore the chemical space [12,36,37]. We integrate the idea of interpolation between chemical elements with BEACON into a new method ("ICE-BEACON"), which allows for global structure optimization through simultaneous optimization of chemical identities and atomic coordinates.

At the heart of the method is a fingerprint, which is a vector representation of the system configuration for a given set of atomic coordinates and fractional chemical identities of the atoms (termed "element fractions" in the following). For all element fractions equal to 0 or 1, the fingerprint is the same as the one used in BEACON [27], and it is described in detail in the Appendix. Given a database of energy and force calculations obtained, for example, with DFT, the fingerprint allows the construction of a surrogate PES using a Gaussian process. It is possible to analytically predict energies and the derivatives with respect to both atomic coordinates and element fractions on the surrogate

PES, allowing the use of gradient-based optimization algorithms.

The global optimization uses Bayesian optimization and is performed in the following way. (1) An initial database of typically two DFT calculations based on random configurations is constructed. (2) The associated surrogate PES is explored by 40 energy minimizations starting from random configurations in the space of element fractions and atomic coordinates or from previously visited energy minima with rattling of the atomic positions and with random element fractions. The element fractions are constrained to keep the right overall stoichiometry. (3) The element fractions of the resulting 40 configurations are rounded to the nearest integer (if necessary), and the configurations are evaluated with a lower-confidence-bound acquisition function (see Appendix for details). A DFT calculation for the configuration with the lowest acquisition function value is included in the database. The procedure is continued by returning to step (2).

All the DFT calculations presented in this Letter are performed with GPAW [38] using the Atomic Simulation Environment [39,40] and the Perdew-Burke-Ernzerhof (PBE) functional [41] unless otherwise stated. The local relaxations in the surrogate potential energy surface are carried out by sequential least squares programming as implemented in the SCIPY package [42,43], which enables the use of both equality and inequality constraints with efficient gradient-driven optimization. In ICE-BEACON, we need equality constraints to fix the number of atoms of different elements in the system, while inequality constraints are used in order to limit the element fractions to be between 0 and 1.

We first consider a AuCu bulk system in a fixed fcc structure with a 64-atom supercell, with 32 atoms of both gold and copper, and a lattice parameter of 3.767 Å. In Fig. 1, we illustrate a single local optimization of the atomic ordering by relaxing the element fractions of the atoms within the surrogate potential energy surface. The training set consists of four DFT calculations with different orderings of the atoms.

Starting from random fractions (step index 0), the run ends up with the correct layered structure of Au and Cu [44,45]. The energy evolution is smooth and the descent optimization is able to guide the system into the low-energy minimum without trapping in states with noninteger element fractions.

We also see that some of the parameters stay at 0 or 1 during multiple steps, but then change from 0 to 1 (or from 1 to 0) later in the run. This shows that the method not only pushes all the fractions to the closest zeros and ones (which would be close to a local minimum in coordinate space), but it is able to circumvent energy barriers by interpolating in chemical element space.

We now consider global ordering optimization of the AuCu, $Au_3Cu$, and $AuCu_3$ bulk systems. The total number of atoms for each supercell is 64 (see Fig. 2). Even though all atoms are fixed at lattice positions, we train the surrogate PES also on forces and thereby include additional information about the local environment around the atoms. Figure 2 shows the success curves for each system. The
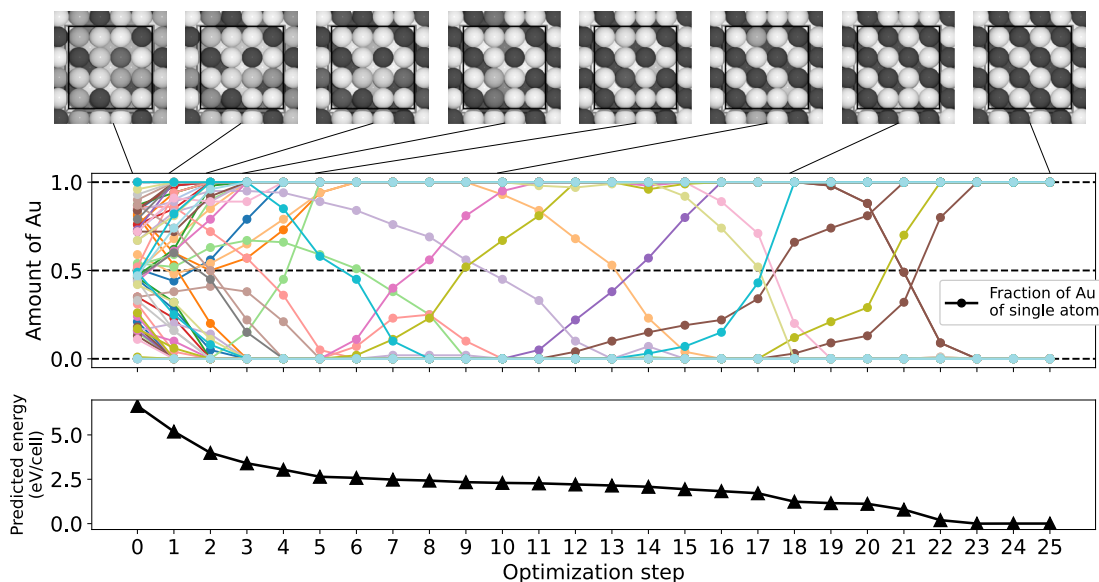


FIG. 1.   A single local optimization run of element fractions in AuCu bulk with 64 atoms in a unit cell. The training set consists of four DFT calculations. The structures at different steps are shown above with the color saturation denoting the corresponding element fractions (black, Au; white, Cu). The black lines indicate the computational supercell boundaries. Middle: the evolution of the element fractions (where Au corresponds to 1 and Cu to 0) is shown for all atoms along the optimization. Lower: the energy predicted by the surrogate model is shown. The optimization is seen to proceed to the ground state without getting trapped in metastable states.
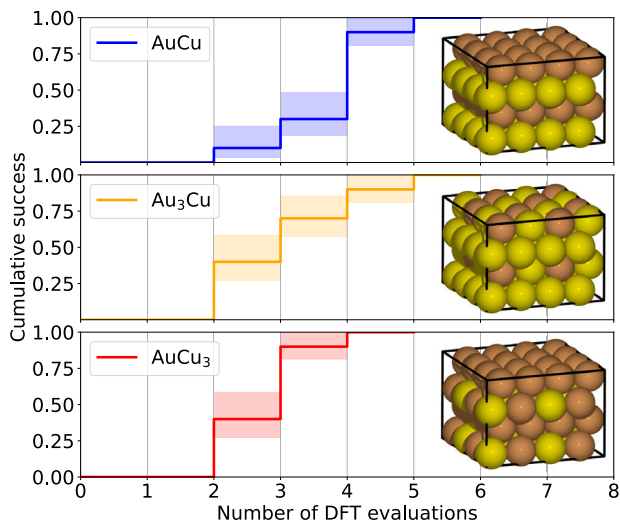
FIG. 2. Success curves for optimizing the ordering of $Au_xCu_y$ in a bulk fcc lattice using ICE-BEACON. For each system, ten individual optimization runs are performed. The number of DFT evaluations in the training set is shown on the $x$ axis, while the fraction of successful runs, where the correct structure is identified, is shown on the $y$ axis. The indicated uncertainties are calculated with bootstrapping.
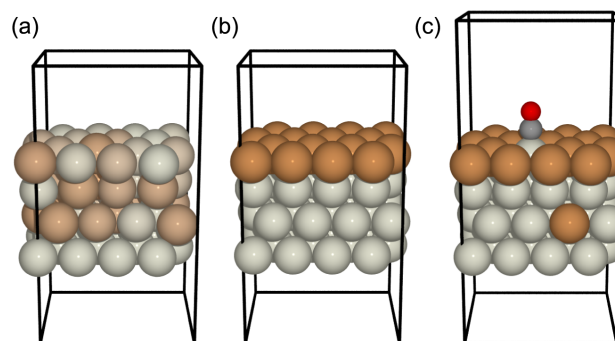


FIG. 3. Initial and optimal CuNi surface structures with 16 Cu atoms out of a total of 64. (a) Typical initial structure before optimization of the clean surface. The colors are linearly scaled between those of Ni and Cu corresponding to the element fractions of the atoms. (b) Optimal structure for the clean surface system as found with ICE-BEACON with simultaneous optimization of coordinates and fractions. The structure is found after three DFT calculations in all of four runs. (c) Optimal structure after introducing a CO molecule at the surface. We see that the CO molecule pulls a Ni atom to the surface due to the stronger chemisorption. During the optimizations, the bottommost layer is fixed to be Ni in fixed positions. Colors: Cu, brown; Ni, light gray; C, dark gray; O, red.

essential features of the PES are learned very quickly, as we only need at maximum five DFT evaluations to obtain the correct ordering in all of the ten runs per system. The identified structures match the experimentally determined ones at low temperatures [44]. Experimentally, AuCu exhibits an interlayer expansion between the Au and Cu layers breaking the cubic symmetry [44,45], but including the distortion does not change the result, as discussed further in the Supplemental Material [46].

We note that the algorithm does not know when the global minimum ordering is found, and it thus continues the search after the eight DFT evaluations shown in the figure. It should be emphasized that the algorithm itself determines which configurations to evaluate with DFT based on the acquisition function. This ensures a balance between relevant low-energy structures and exploration of new configurations with large uncertainties.

To further investigate the applicability of ICE-BEACON, we study a CuNi alloy with an fcc(111) surface. The system is setup as a four-layer slab with a $4 \times 4$ surface supercell and a fixed in-plane lattice constant of 3.583 Å (64 atoms in total). We consider the situation with 16 Cu atoms. The coordinates and chemical identities of the bottommost Ni layer are fixed during the optimizations. The GPAW calculations are run in spin-paired mode, but we have verified that spin polarization has no effect on the order of the energies for the final structures. The RPBE functional [48] is used for the CuNi surface systems.

The initial structures for the optimizations in the surrogate potential energy surfaces are obtained by rattling the atomic coordinates by 0.05 Å around the lattice sites and assigning random element fractions to the atoms in the three top layers keeping the total number of Cu atoms at 16. One example of an initial structure is shown in Fig. 3(a).

The result of an ICE-BEACON optimization with simultaneous optimization of the atomic positions and the element fractions is shown in Fig. 3(b). This structure was proposed as the global minimum by all of four individual optimization runs after three DFT calculations. The copper atoms are seen to migrate to the surface, which is also what is found experimentally [49].

In Fig. 3(c) we show the result of an ICE-BEACON calculation where a CO molecule is placed on top of the CuNi(111) surface. Again both atomic coordinates and element fractions are optimized. A nickel atom is pulled to the surface with the CO in an on-top position. This is in agreement with the fact that the adsorption energy is larger on Ni(111) than on Cu(111) [50], and this overcomes the segregation energy of copper. Experimentally CO is observed to occupy top sites in Cu-Ni(111) alloy surfaces with mixing of the metal atoms [51], even though CO prefers hollow sites on pure Ni(111) and Cu(111) [50]. The effect of Ni atoms backsegregating to the surface exposed to CO has also been observed experimentally [52].

The on-top site of CO is observed in, on the average, 20 DFT calculations in five out of a total of eight optimization runs. One of the other runs suggests a bridge site, and one run suggests a hollow hcp adsorption site. The last run does not find a low-energy structure. In all cases, we find that the CO molecule lifts Ni atoms to the neighboring sites. See
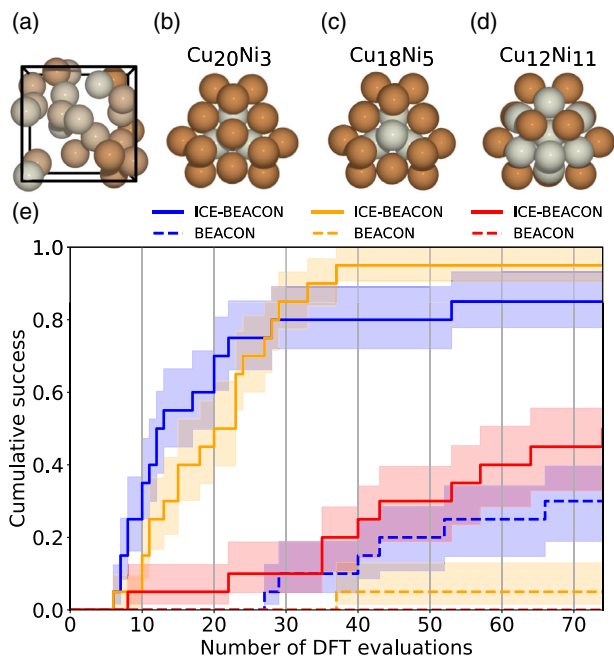
FIG. 4. (a) Example of a random CuNi initial structure of a 23-atom cluster generated in a cubic box (black lines) before optimization in ICE-BEACON. The colors are linearly scaled between those of Ni and Cu, corresponding to the element fractions of the atoms. (b)–(d) The identified global minimum structures for $Cu_{20}Ni_3$, $Cu_{18}Ni_5$, and $Cu_{12}Ni_{11}$, respectively, as determined with ICE-BEACON. (e) Success curves for optimizing the clusters with BEACON and ICE-BEACON up to 75 DFT calculations based on 20 runs each. ICE-BEACON requires considerably less DFT calculations than BEACON, and in the case of 11 nickel atoms, where the number of possible orderings exceeds $10^6$, only ICE-BEACON finds the optimal structure.

Fig. S2 in the Supplemental Material [46] for the structures and their energies.

As the most challenging example, we consider CuNi clusters, where the metal atoms are not ascribed to lattice sites, and both element fractions and atomic coordinates are simultaneously optimized. We consider CuNi clusters with a total of 23 atoms and varying content of nickel (3, 5, and 11 atoms). In Fig. 4, we compare the results of ICE-BEACON with BEACON and find a considerable reduction in the number of DFT calculations necessary to find the optimal structure when using ICE-BEACON.

For each cluster stoichiometry, 20 ICE-BEACON optimization runs are performed and success is declared at the first encounter of an energy within 0.05 eV of the minimum energy encountered across all BEACON and ICE-BEACON runs for the respective cluster stoichiometry. Subsequently, it is checked that the structures within this energy range are, in fact, identical. The initial structures for the random searches in the surrogate PES are chosen with random initial atomic coordinates within a box of size $8 \times 8 \times 8$ Å$^3$ and with random element fractions [see Fig. 4(a) for an example]. The computational unit cell is a cube with side length 16 Å.

Figures 4(b)–4(d) show the identified structures. For three and five nickel atoms, the structure is a triple icosahedron, where six atoms are shared by two of the icosahedra and five atoms by all three. The latter are exactly the five nickel atoms in $Cu_{18}Ni_5$. The cluster with 11 nickel atoms exhibits an icosahedron of 11 nickel atoms and 2 copper atoms, with the icosahedron covered by additional copper atoms (in the back in the figure). The structure with three nickel atoms agrees with previous calculations based on an interatomic potential [53].

In the cases with three and five nickel atoms, ICE-BEACON finds the optimal structures with typically 20 DFT calculations, while BEACON requires considerably more and does often not find the correct structure within 75 DFT calculations [see Fig. 4(e)]. With 11 nickel atoms, the number of possible different Cu/Li decorations for given atomic positions becomes larger than $10^6$ and BEACON does not identify the optimal structure in any of the 20 runs. ICE-BEACON finds the structure in half of the runs with 75 DFT calculations.

To summarize, we have introduced a method to simultaneously optimize atomic positions and chemical ordering in atomic-scale systems. The method was illustrated with applications to bulk, surface, and cluster systems. The introduction of element fractions enables a significant improvement for multielement systems compared to a method like BEACON, which only works in coordinate space.

A striking feature of the method, and a key to its performance, is that the interpolation between chemical elements allows for the switching of atoms of different types without an energy barrier. As a result, the element fractions are not trapped in many metastable states during optimization. Another remarkable feature is that, in most cases, the final values for the element fractions are close to 0 or 1 as for real atomic materials. These features are present for all applications shown here, but to which extent this holds for other applications further research will show.

The approach can easily be extended to systems with several different chemical elements and simultaneous optimization of the unit cell, and further efficiency may be obtained by training the model to smaller subsystems as well. For systems where several different concentrations are investigated, the surrogate model may benefit from a common database with all the DFT calculations.

The code for ICE-BEACON is included in the GPATOM package [54]. The current limitation with respect to speed is due to the calculation of the fingerprint and its gradients at each surrogate relaxation step, which, however, only scales as $\mathcal{O}(N_{atoms})$. For large systems, the inversion of the $C$ matrix (see Appendix) may be limiting because it scales cubically with the number of atoms and the number of DFT calculations included in the training set. Memory issues related to the matrix inversion is currently limiting the system size to around 100 atoms, but larger systems could be treated with a more memory efficient parallelization.

*Appendix: Surrogate model and acquisition function.*—
The fingerprint is the same that we use in BEACON with the essential difference that now the atoms are represented as fractional: each atom possesses the element fraction $q_A$ of one of the two elements and the fraction $q_B = 1 - q_A$ of the other element. We use

$$\rho_{AB}^R(r, q; \mathbf{x}) = \sum_{\substack{i \in A \\ j \in B}} q_{i,A} q_{j,B} \frac{1}{r_{ij}^2} f_c(r_{ij}; R_c^R) e^{-|r - r_{ij}|^2 / 2\delta_R^2} \quad \text{(A1)}$$

for a radial distribution function and

$$\rho_{ABC}^\alpha(\theta, q; \mathbf{x}) = \sum_{\substack{i \in A \\ j \in B \\ k \in C}} [q_{i,A} q_{j,B} q_{k,C} f_c(r_{ij}; R_c^\alpha) f_c(r_{jk}; R_c^\alpha)$$
$$\cdot e^{-|\theta - \theta_{ijk}|^2 / 2\delta_\alpha^2}] \quad \text{(A2)}$$

for an angular distribution, where $q_{i,A} \in [0, 1]$ is the element fraction describing how much atom $i$ is of element $A$. In Eqs. (A1)) and (A2), $r_{ij}$ is the distance between atoms $i$ and $j$, and $f_c$ is a smooth cutoff function with cutoff distances $R_c$. The cutoff distances have $R_c^R = 8$ and $R_c^\alpha = 4$ Å. The values for $\delta_R = 0.4$ Å and $\delta_R = 0.4$ rad are also kept fixed. The full fingerprint is obtained by concatenating the vectors created with Eqs. (A1)) and (A2) for different element pairs and triples.

Energies and forces, $\mu = (E, -F)$, are calculated with the usual expression for a Gaussian process [55,56]

$$\mu(\mathbf{x}, Q) = \mu_p(\mathbf{x}) + K(\rho(\mathbf{x}, Q), P)C(P, P)^{-1}[y - \mu_p(X)], \quad \text{(A3)}$$

where $\mathbf{x}$ is the atomic coordinates, $Q$ are the element fractions, $\mu_p(\mathbf{x})$ is the prior function, $K$ and $C$ are covariance matrices, $P$ is a matrix containing the fingerprints for the training data, and $y$ is the training targets. In Eq. (A3), the covariance matrix $K$ is the only quantity that contains the information about the element fractions that are stored in the vector $Q$ for each atom. For details about the construction of the matrices $K$ and $C$ and the vector $y$, we refer the reader to the article about BEACON [27]. We use a constant prior $\mu_p(\mathbf{x}) = \mu_p$ and a squared-exponential kernel with a distance measure given by the Euclidean distance between the fingerprint vectors. Along an ICE-BEACON run, the prior constant and the kernel hyperparameters are updated by maximizing the marginal log-likelihood as also described in Ref. [27].

Within the Bayesian optimization, we use an acquisition function $f$ of the form of the lower confidence bound, which reads

$$f(\mathbf{x}) = E(\mathbf{x}) - 2\Sigma(\mathbf{x}), \quad \text{(A4)}$$

where $\Sigma(\mathbf{x})$ is the uncertainty of the predicted energy, provided by the Gaussian process [27,55,56].

*kwj@fysik.dtu.dk

[1] J. Zhang and V. A. Glezakou, Int. J. Quantum Chem. **121**, 044114 (2020).
[2] D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).
[3] Z. Chen, W. Jia, X. Jiang, S.-S. Li, and L.-W. Wang, Comput. Phys. Commun. **219**, 35 (2017).
[4] L. B. Vilhelmsen and B. Hammer, J. Chem. Phys. **141**, 044711 (2014).
[5] S. V. Lepeshkin, V. S. Baturin, Y. A. Uspenskii, and A. R. Oganov, J. Phys. Chem. Lett. **10**, 102 (2019).
[6] S. Lysgaard, D. D. Landis, T. Bligaard, and T. Vegge, Top. Catal. **57**, 33 (2014).
[7] M. Jäger, R. Schäfer, and R. L. Johnston, Nanoscale **11**, 9042 (2019).
[8] C. J. Pickard and R. J. Needs, J. Phys. Condens. Matter **23**, 053201 (2011).
[9] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, npj Comput. Mater. **5**, 83 (2019).
[10] L. Himanen, A. Geurts, A. S. Foster, and P. Rinke, Adv. Sci. **6**, 1900808 (2019).
[11] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Nature (London) **559**, 547 (2018).
[12] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, and M. Ceriotti, Sci. Adv. **3**, e1701816 (2017).
[13] J. Behler, J. Chem. Phys. **145**, 170901 (2016).
[14] A. Denzel and J. Kästner, J. Chem. Theory Comput. **16**, 5083 (2020).
[15] T. D. Huan, R. Batra, J. Chapman, S. Krishnan, L. Chen, and R. Ramprasad, npj Comput. Mater. **3**, 37 (2017).
[16] E. Garijo del Río, J. J. Mortensen, and K. W. Jacobsen, Phys. Rev. B **100**, 104103 (2019).
[17] E. Garijo del Río, S. Kaappa, J. A. Garrido Torres, T. Bligaard, and K. W. Jacobsen, J. Chem. Phys. **153**, 234116 (2020).
[18] V. L. Deringer and G. Csányi, Phys. Rev. B **95**, 094203 (2017).
[19] J. A. Garrido Torres, P. C. Jennings, M. H. Hansen, J. R. Boes, and T. Bligaard, Phys. Rev. Lett. **122**, 156001 (2019).
[20] O.-P. Koistinen, F. B. Dagbjartsdóttir, V. Ásgeirsson, A. Vehtari, and H. Jónsson, J. Chem. Phys. **147**, 152720 (2017).
[21] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Sci. Adv. **3**, e1603015 (2017).
[22] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, J. Chem. Phys. **148**, 241722 (2018).

[23] B. Kalita, L. Li, R. J. McCarty, and K. Burke, Acc. Chem. Res. **54**, 818 (2021).

[24] J. Wellendorff, K. T. Lundgaard, A. Møgelhøj, V. Petzold, D. D. Landis, J. K. Nørskov, T. Bligaard, and K. W. Jacobsen, Phys. Rev. B **85**, 235149 (2012).

[25] A. Grisafi, A. Fabrizio, B. Meyer, D. M. Wilkins, C. Corminboeuf, and M. Ceriotti, ACS Cent. Sci. **5**, 57 (2019).

[26] M. K. Bisbo and B. Hammer, Phys. Rev. Lett. **124**, 086102 (2020).

[27] S. Kaappa, E. G. del Río, and K. W. Jacobsen, Phys. Rev. B **103**, 174114 (2021).

[28] C. J. H. Jacobsen, S. Dahl, B. S. Clausen, S. Bahn, A. Logadottir, and J. K. Nørskov, J. Am. Chem. Soc. **123**, 8404 (2001).

[29] F. Weigend, J. Chem. Phys. **141**, 134103 (2014).

[30] O. A. von Lilienfeld, R. D. Lins, and U. Rothlisberger, Phys. Rev. Lett. **95**, 153002 (2005).

[31] G. F. von Rudorff and O. A. von Lilienfeld, Phys. Rev. Research **2**, 023220 (2020).

[32] B. Huang and O. A. von Lilienfeld, arXiv:2012.07502.

[33] P. Soven, Phys. Rev. **156**, 809 (1967).

[34] A. V. Ruban, I. A. Abrikosov, D. Y. Kats, D. Gorelikov, K. W. Jacobsen, and H. L. Skriver, Phys. Rev. B **49**, 11383 (1994).

[35] I. A. Abrikosov, A. V. Ruban, H. L. Skriver, and B. Johansson, Phys. Rev. B **50**, 2039 (1994).

[36] M. J. Willatt, F. Musil, and M. Ceriotti, Phys. Chem. Chem. Phys. **20**, 29661 (2018).

[37] M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzsenyi, and P. Marquetand, J. Chem. Phys. **148**, 241709 (2018).

[38] J. Enkovaara et al., J. Phys. Condens. Matter **22**, 253202 (2010).

[39] Atomic Simulation Environment (ASE), https://wiki.fysik.dtu.dk/ase/ (2020).

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

[41] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[42] D. Kraft, ACM Trans. Math. Softw. **20**, 262 (1994).

[43] P. Virtanen et al. and (SciPy 1.0 Contributors), Nat. Methods **17**, 261 (2020).

[44] H. Okamoto, D. J. Chakrabarti, D. E. Laughlin, and T. B. Massalski, J. Phase Equilib. **8**, 454 (1987).

[45] I. Uszyński, J. Janczak, and R. Kubiak, J. Alloys Compd. **206**, 211 (1994).

[46] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.127.166001 for supplemental results for AuCu bulk and CO@CuNi(111) surface systems, which includes Ref. [47].

[47] E. B. Tadmor, G. S. Smith, N. Bernstein, and E. Kaxiras, Phys. Rev. B **59**, 235 (1999).

[48] B. Hammer, L. B. Hansen, and J. K. Nørskov, Phys. Rev. B **59**, 7413 (1999).

[49] P. R. Webber, M. A. Morris, and Z. G. Zhang, J. Phys. F **16**, 413 (1986).

[50] K. M. Gameel, I. M. Sharafeldin, and N. K. Allam, Phys. Chem. Chem. Phys. **21**, 11476 (2019).

[51] X. H. Feng, M. R. Yu, S. Yang, G. Meigs, and E. Garfunkel, J. Chem. Phys. **90**, 7516 (1989).

[52] K. Wandelt and C. R. Brundle, Phys. Rev. Lett. **46**, 1529 (1981).

[53] E. Hristova, Y. Dong, V. G. Grigoryan, and M. Springborg, J. Phys. Chem. A **112**, 7905 (2008).

[54] ASE-GPAtom, https://gitlab.com/gpatom/ase-gpatom/ (2021).

[55] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006).

[56] J. Wu, M. Poloczek, A. G. Wilson, and P. Frazier, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., New York, 2017), pp. 5267–5278, https://proceedings.neurips.cc/paper/2017/hash/64a08e5f1e6c39faeb90108c430eb120-Abstract.html.