

# Accelerating high-throughput searches for new alloys with active learning of interatomic potentials

Konstantin Gubaev<sup>a</sup>, Evgeny V. Podryabinkin<sup>a</sup>, Gus L.W. Hart<sup>b</sup>, Alexander V. Shapeev<sup>a,\*</sup>

<sup>a</sup> Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel str. 3, Moscow 143026, Russia

<sup>b</sup> Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602, USA

## ARTICLE INFO

### Keywords:

Alloy phase prediction  
Machine learning  
Active learning  
Interatomic potentials  
Cluster expansion  
Moment Tensor Potentials

## ABSTRACT

We propose an approach to materials prediction that uses a machine-learning interatomic potential to approximate quantum-mechanical energies and an active learning algorithm for the automatic selection of an optimal training dataset. Our approach significantly reduces the amount of density functional theory (DFT) calculations needed, resorting to DFT only to produce the training data, while structural optimization is performed using the interatomic potentials. Our approach is not limited to one (or a small number of) lattice types (as is the case for cluster expansion, for example) and can predict structures with lattice types not present in the training dataset. We demonstrate the effectiveness of our algorithm by predicting the convex hull for the following three systems: Cu-Pd, Co-Nb-V, and Al-Ni-Ti. Our method is three to four orders of magnitude faster than conventional high-throughput DFT calculations and explores a wider range of materials space. In all three systems, we found unreported stable structures compared to the AFLOW database. Because our method is much cheaper and explores much more of materials space than high-throughput methods or cluster expansion, and because our interatomic potentials have a systematically improvable accuracy compared to empirical potentials such as embedded atom model, it will have a significant impact in the discovery of new alloy phases, particularly those with three or more components.

## 1. Introduction

Advances in computer power, improvements in first-principles methods, and the generation of large materials databases like AFLOWLIB [1], OQMD [2], CMR [3], NOMAD [4], and Materials Project [5] have enabled modern data analysis tools to be applied in the field of materials discovery [6–8]. There have been growing efforts in computational search for materials with superior properties, including metallic alloys [9–11], semiconductor materials [12], and magnetic materials [13]. In this work we consider the problem of predicting stable phases in multicomponent alloys. A typical prediction algorithm consists of sampling structures across the configurational space and evaluating their energies. The sampling is done by searching through structures that are either selected from some carefully assembled pool of possible structures, often called crystal prototypes [14], or are generated by some sampling algorithm, see, e.g., Refs. [15,16].

The evaluation of the energy of the structures in the pool is often done with density functional theory (DFT). Even despite its favorable accuracy/efficiency trade-off as compared to other quantum-mechanical algorithms, the DFT calculations remain the bottleneck in

materials prediction workflows, making an exhaustive search impractical. Machine learning (ML) for materials prediction has the potential to dramatically reduce the number of quantum-mechanical calculations performed and thus reduce the computational expense of predicting new materials via computation. The reduction of the computational time is achieved by constructing a surrogate model that “interpolates” the quantum-mechanical training data and makes subsequent energy evaluations much faster (by orders of magnitude). This is similar in spirit to the cluster expansion method which has been broadly used in different materials discovery applications [17–19,12]. Cluster expansion is quite successful when the stable structures are derivatives of a particular structure (fcc, bcc, etc.) but is not useful when this is not the case. Its accuracy also converges slowly when atomic size mismatch is not negligible [20]. Additionally, more classical machine-learning algorithms such as decision trees [21], support vector machines [22], and other ML algorithms [23,24] have been tried. Surrogate models such as the cluster expansion and standard machine learning approaches do not have the broad applicability and exceptional accuracy of the moment tensor potentials-based [25] approach we demonstrate here.

\* Corresponding author.

E-mail address: [alexander@shapeev.com](mailto:alexander@shapeev.com) (A.V. Shapeev).

The two important features of our approach are a completely general form for the interatomic potentials and an active learning algorithm for generating and refining the training set. In this work we extend the approach [26] for predicting the structure of a single-component material. In our approach, a ML model reproduces DFT for off-equilibrium structures that are not restricted to any lattice. Furthermore, the model learns the DFT interaction actively (on-the-fly) while equilibrating the candidate structures, completely automating the construction of the training set. Thus, structural optimization of the training structures can be performed via the interatomic potentials, rather than via DFT, further accelerating the construction of the training set.

Our method is based on moment tensor potentials (MTPs [25]) and the active learning algorithm [27]. Namely, we solve the following problem: given a set of elements, find the most stable structures (in the sense of lying on the convex hull of formation enthalpies) consisting of these elements, each characterized by composition, unit cell geometry and atomic positions within the unit cell. In this work we extend the interatomic potential [25] and active learning algorithm [27] to handle atomistic configurations with multiple types of atoms, similarly to the approach used in cheminformatics [28]. The differences between the algorithms from Ref. [28] and this work include that (1) we need derivatives of the energy, whereas in Ref. [28] we needed only the energy (or other predicted properties); and (2) that in Ref. [28] we were concerned with the selection from a finite set of predefined structures, whereas in this work we need to solve the problem predicting the energy with a fitted potential and assembling the training set used for the fitting at the same time (in other words, exploring the potential energy landscape and constructing the training set at the same time).

The idea of applying neural networks, as a broad class of machine-learning algorithms, to constructing interatomic potentials was pioneered in Ref. [29]. Application of Gaussian process regression, another class of machine-learning algorithms, was then proposed in Ref. [30]. The promising results obtained in these works have motivated many research groups to pursue this research direction [31–50,25,51]. However, the application of such algorithms to the problem of materials prediction has proven difficult since following such a methodology requires one to collect all the representative structures in the training set which is as hard as predicting materials structure itself. In our view, it is the active learning [27,26,33,52,53] that paves the way for machine-learning interatomic potentials to accelerate computational materials discovery.

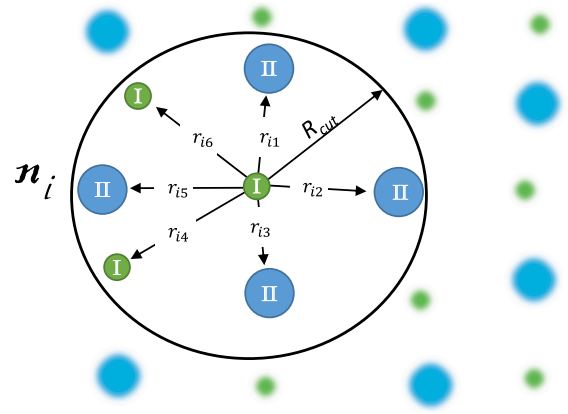
This paper is organized as follows: in Section 2 we introduce the algorithms we use, including the moment tensor potentials (Section 2.1), active learning (Section 2.2), and the “relaxation while learning on-the-fly” algorithm (Section 2.3). In Section 3 we test the proposed algorithm on predicting the stable structures of the Cu-Pd, Co-Nb-V, and Al-Ni-Ti systems and discuss the performance of our algorithm. In particular, we compare our results to those obtained by high-throughput DFT calculations as reported in the AFLOW database [10,1]. In all three systems we have discovered new structures below the reported convex hull of ground-state structures. Finally, in Section 4 we make concluding remarks.

## 2. Methodology

### 2.1. Machine-learning potentials

We use the moment tensor potentials (MTPs) for approximating a quantum-mechanical energy. The potential is parametrized by a set of parameters  $\theta$  that are found from minimizing the loss functional expressing that the predicted energy  $E$  is close to the reference quantum-mechanical energy  $E^{\text{qm}}$ :

$$L(\theta) = \sum_j (E(\theta, x^{(j)}) - E^{\text{qm}}(x^{(j)}))^2 \rightarrow \min, \quad (1)$$



**Fig. 1.** Partitioning scheme: energy  $E$  is composed from contributions  $V_i$  of individual neighborhoods  $n_i$ . The neighborhood  $n_i$  of the  $i$ th atom is described by the relative position of neighboring atoms,  $r_{ij}$ , and the types of atoms  $z_j$  (I or II in this illustration).

where  $x^{(j)}$  are the configurations in the training set and  $E^{\text{qm}}(x^{(j)})$  are their reference energies.

Our model is local, which we enforce by partitioning the energy,  $E$ , into the contributions,  $V$ , of individual atomic neighborhoods. To define a neighborhood of the  $i$ th atom, we let  $r_{ij}$  be the position of  $j$ th atom relative to the  $i$ th atom (thus,  $r_{ij}$  is a vectorial quantity) and  $z_j$  be the type of the  $j$ th atom. Then  $n_i$  is the collection of  $r_{ij}$  and  $z_j$ , and  $E(x) = \sum_i V(n_i)$ . The locality of the model is expressed by the requirement that  $V$  does not depend on atoms that are farther from  $i$  than some cutoff distance  $R_{\text{cut}}$ , which is usually around 5 Å. An illustration of an atomic neighborhood is sketched in Fig. 1. Mathematically, each atom in the neighborhood introduces four degrees of freedom, on which  $n_i$  depends: these are three coordinates in Euclidean space, and a discrete variable representing the chemical type. Typically, neighborhoods include a few dozen atoms, which means that the function  $V(n_i)$  depends on the order of hundred scalar variables. To somewhat reduce the dimensionality, we embed all physical symmetries into  $V(n)$  so they will not have to be learned by the model. These symmetries arise from the isotropy and translational symmetry of the physical space, and from the fact that the interaction between atoms does not depend on their ordering (see Fig. 2).

As in the work [25] devoted to the single-component moment tensor potentials,  $V(n)$  is linearly expanded through a set of *basis functions*  $B_\alpha$ :

$$V(n) = \sum_\alpha \xi_\alpha B_\alpha(n). \quad (2)$$

The basis functions, in turn, depend on the set of *moment tensor descriptors*

$$M_{\mu,\nu}(n_i) = \sum_j f_\mu(|r_{ij}|, z_i, z_j) \underbrace{r_{ij} \otimes \dots \otimes r_{ij}}_{\nu \text{ times}}, \quad (3)$$

where the index  $j$  enumerates all the atoms in the neighborhood  $n_i$ . The functions  $f_\mu(|r_{ij}|, z_i, z_j)$  depend only on the interatomic distances and atomic types, therefore we call them *radial functions*. The terms  $r_{ij} \otimes \dots \otimes r_{ij}$  contain the angular information about the neighborhood  $n_i$  and are tensors of rank  $\nu$ . We next explain how to construct the basis functions from the moment tensor descriptors, following which we present a simple illustration of the structure of the descriptors and basis functions.

The functions  $B_\alpha(n_i)$  enumerate all possible contractions of any number of  $M_{\mu,\nu}(n_i)$  yielding a scalar. Note that  $M_{\mu,\nu}(n_i)$  are invariant, by construction, with respect to translations of the system and permutations of equivalent atoms. Their scalar contractions are invariant with respect to rotations of the neighborhood. Thus the resulting function

Download English Version:

<https://daneshyari.com/en/article/11026648>

Download Persian Version:

<https://daneshyari.com/article/11026648>

[Daneshyari.com](https://daneshyari.com)