








## Finite-temperature interplay of structural stability, chemical complexity, and elastic properties of bcc multicomponent alloys from *ab initio* trained machine-learning potentials

Konstantin Gubaev <sup>1,2,\*</sup> Yuji Ikeda <sup>2</sup> Ferenc Tasnádi <sup>3</sup> Jörg Neugebauer <sup>4</sup> Alexander V. Shapeev <sup>5</sup>  
Blazej Grabowski <sup>2</sup> and Fritz Körmann <sup>1,4,†</sup>

<sup>1</sup>Materials Science and Engineering, Delft University of Technology, 2628 CD, Delft, The Netherlands

<sup>2</sup>Institute for Materials Science, University of Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart, Germany

<sup>3</sup>Department of Physics, Chemistry, and Biology (IFM), Linköping University, SE-581 Linköping, Sweden

<sup>4</sup>Computational Materials Design, Max-Planck-Institut für Eisenforschung GmbH, D-40237 Düsseldorf, Germany

<sup>5</sup>Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel St. 3, Moscow 143026, Russia



(Received 18 March 2021; accepted 14 June 2021; published 8 July 2021)

An active learning approach to train machine-learning interatomic potentials (moment tensor potentials) for multicomponent alloys to *ab initio* data is presented. Employing this approach, the disordered body-centered cubic (bcc) TiZrHfTa<sub>x</sub> system with varying Ta concentration is investigated via molecular dynamics simulations. Our results show a strong interplay between elastic properties and the structural  $\omega$  phase stability, strongly affecting the mechanical properties. Based on these insights we systematically screen composition space for regimes where elastic constants show little or no temperature dependence (elinvar effect).

DOI: [10.1103/PhysRevMaterials.5.073801](https://doi.org/10.1103/PhysRevMaterials.5.073801)

### I. INTRODUCTION

Parameter-free simulation techniques to access structural stability and mechanical properties of materials are a vital element in computer-aided materials design [1–6]. In particular, the vast composition phase space of multicomponent alloys with multiple principal elements, which are often referred to as chemically complex alloys or high entropy alloys (HEAs) [7–9], features tremendous opportunities to explore and design next-generation materials by means of computer simulations [10–12].

Many materials properties, e.g., elastic constants, feature a qualitatively very different behavior if finite-temperature contributions are taken into account [13–16]. For example, thermal expansion typically causes a strong decrease of elastic constants. However, for some materials these properties can also be remarkably insensitive to temperature. A temperature-invariant elastic behavior of materials, referred to as elinvar effect, is, e.g., known for multicomponent Ni-based alloys [17–19]. Recently, the elinvar effect has also been reported for elemental body-centered cubic (bcc) Ti [20]. The inclusion of finite-temperature effects is thus crucial for a reliable prediction of the mechanical properties of materials at elevated temperatures.

Besides the principal challenge to include finite-temperature contributions, the alloys' phase stability can further impact materials properties [16,21–23]. In Ti, for example, the high-temperature bcc phase is unstable at low temperatures (see, e.g., Ref. [24]). The inclusion of finite-temperature effects is thus unavoidable to stabilize

bcc Ti and determine its properties, in particular the elastic properties. For multicomponent alloys, the interplay between the chemical environment and finite temperature effects brings about further challenges for atomistic simulations as compared to the treatment of pure metals. For example, in chemically complex bcc alloys, being in the focus of the present work, the distinction between a *distorted* bcc lattice and an  $\omega$ -type structure is not straightforward anymore. This issue has recently been discussed and analyzed by a projected displacement scheme that allows one to disentangle these delicate structural features [25]. However, this detection scheme has been mainly limited to ground state structural analysis and did not involve any analysis on its relation to elastic properties.

Density functional theory calculations (DFT) have been successfully utilized to compute phase stability as well as elastic properties of multicomponent alloys [10,26–28]. Most of these studies were limited to 0 K. Alternative approaches that allow one to include finite-temperature effects, such as *ab initio* molecular dynamics (MD) simulations, are often too computationally expensive to study multicomponent alloys.

The advent of machine-learning interatomic potentials (MLIPs) [29,30] has made it possible to significantly reduce the computational costs of pure DFT calculations and to access materials properties with near DFT accuracy while preserving the computational efficiency [31] of classical interatomic potentials [20,32,33]. MLIPs have rapidly advanced in the past decade, resulting in a number of approaches [34–38]. They offer a particularly efficient way to tackle the challenges associated with HEAs [29]. A fundamental requirement of MLIP-based approaches is a well-defined training set which contains information about all relevant and often *a priori* unknown phases and configurations (see, e.g., Ref. [39]). One way to tackle this issue are active learning (AL)

\*kostiantyn.hubaiev@imw.uni-stuttgart.de

†koermann@mpie.de

schemes [40–44], which allow on-the-fly accumulation of the required training data. For compositionally complex alloys, particularly if they show structural instabilities, a nonautomatized and *a priori* manual selection of relevant configurations is challenging and may dramatically limit the application range of these potentials.

In the present work we propose an AL-based framework that employs DFT calculations, MLIP training, and MD simulations. Utilizing this framework we resolve the delicate interplay of structural stability ( $\omega$  vs bcc), elastic properties (elastic constant tensor), and chemical composition for refractory bcc high entropy alloys. Based on these results we derive composition-temperature dependent  $\omega$ -bcc phase diagrams and show how the structural transition is strongly reflected in the elastic properties.

## II. METHODS

### A. MLIP software

Most calculations for the present work are performed with the software named “MLIP,” which was developed at the Skolkovo Institute of Science and Technology and can be obtained from Ref. [45]. The corresponding Python interface, MLIPPY, is available from Ref. [46]. For a recent MLIP tutorial paper including several examples and applications we refer to Ref. [47].

### B. Moment tensor potentials

For studying finite-temperature properties we employ MD simulations in combination with machine-learning interatomic potentials, namely the moment tensor potentials (MTPs) proposed for single-component systems in Ref. [48] and extended to multicomponent systems in Refs. [49,50]. MTPs are local potentials, i.e., the energy  $E(\mathbf{x})$  of an atomistic configuration  $\mathbf{x}$  of size  $N_a$  is assumed to be composed of contributions from individual atoms with site energies  $E_i$ , defined by atomic coordinates  $\mathbf{r}$  and chemical types  $\mathbf{z}$  in the vicinity (neighborhood) of an atom  $i$  as

$$E(\mathbf{x}) = \sum_{i=1}^{N_a} E_i(\mathbf{r}, \mathbf{z}). \quad (1)$$

The neighborhood of the atom  $i$  includes all atoms within a given cutoff radius from the central atom. Therefore, the calculational expenses scale as for classical interatomic potentials, linearly with the number of atoms  $N_a$ . For the present work a cutoff value of 5 Å has been chosen. This includes up to the third nearest-neighbor shell for the investigated TiZrHfTa<sub>x</sub> alloys. We verified that further increasing the cutoff does not improve the quality of the interatomic potential.

The site energies from Eq. (1) are parametrized in terms of the internal MTP parameters  $\mathbf{C}_\alpha$ :

$$E(\mathbf{x}) = \sum_{i=1}^{N_a} E_i(\mathbf{r}, \mathbf{z}) = \sum_{i=1}^{N_a} V_i(\mathbf{C}_\alpha), \quad (2)$$

where  $V_i(\mathbf{C}_\alpha)$  defines the interatomic potential and form of the potential energy surface (PES) provided by MTP. The functional form of  $V_i(\mathbf{C}_\alpha)$  is given in Ref. [50]. The parameters  $\mathbf{C}_\alpha$  are determined during the training procedure. This means

essentially to solve a minimization problem with respect to the coefficients  $\mathbf{C}_\alpha$ , where the functional to be minimized is the sum of squared differences between MTP-predicted and DFT-based energies (or, optionally, forces and stresses as well) [47,50].

The training process of MTPs is parallelized and employs an efficient quasi-Newton BFGS [51] minimization method which does not require second derivatives of the minimized functional with respect to the fitting coefficients (only first derivatives).

In the whole process, the computational time for generating the training set (i.e., performing the required DFT calculations) consumes the vast majority of the computational expenses.

The special functional form of MTPs takes into account multibody interactions while the total number of parameters to fit, and thus the amount of required DFT data, scales as  $m^2$  with the number of species  $m$ . The quadratic scaling makes MTPs extremely efficient in particular for multicomponent alloys.

### C. Active learning: Overview

The main goal of the AL procedure is to simultaneously sample the training data (i.e., atomistic configurations) and to train the MTP on the most relevant sampling data. This significantly reduces any active intervention by the user which is particularly important when dealing with complex alloys in high-dimensional chemical composition space as is the case for HEAs, where otherwise very elaborate and sophisticated manual selection of training data would be required [39].

For MTPs, the AL relies on a *measure of extrapolation* for a given atomistic configuration. This quantity  $\gamma$  is called *extrapolation grade* and it provides an uncertainty measure of the MTP-predicted quantities for a given configuration. Mathematical details as well as different interpretations of the extrapolation grade for MTPs can be found in Ref. [52].

The actual determination of  $\gamma$  is based on the so-called MaxVol criterion [53]. This criterion provides a measure of how much the current training set should be expanded in terms of volume in the highly multidimensional space of the MTP internal parameters [52]. The goal of the AL procedure is to maximize the volume of the training domain, thus ensuring that no point (configuration) remains outside this domain and thus no extrapolation happens.

Typical values for  $\gamma$  classifying the status of a given configuration are as follows.

$\gamma < 1$ : interpolation case.

$1 < \gamma < 2-5$ : “safe” extrapolation case.

$5-10 < \gamma$ : “severe” extrapolation case.

During AL simulations the extrapolation grade is simultaneously computed along with the application of the MTP, for example, in MD. If  $\gamma$  exceeds some predefined threshold value, the potential can be retrained with additional data and the MD simulation can be restarted. If the given threshold value is not exceeded, the potential interpolates (or, depending on the actual value of  $\gamma$ , extrapolates) during the entire simulation run and no retraining is performed.

The AL procedure starts either with an empty training set (i.e., randomly initialized potential parameters) or with an

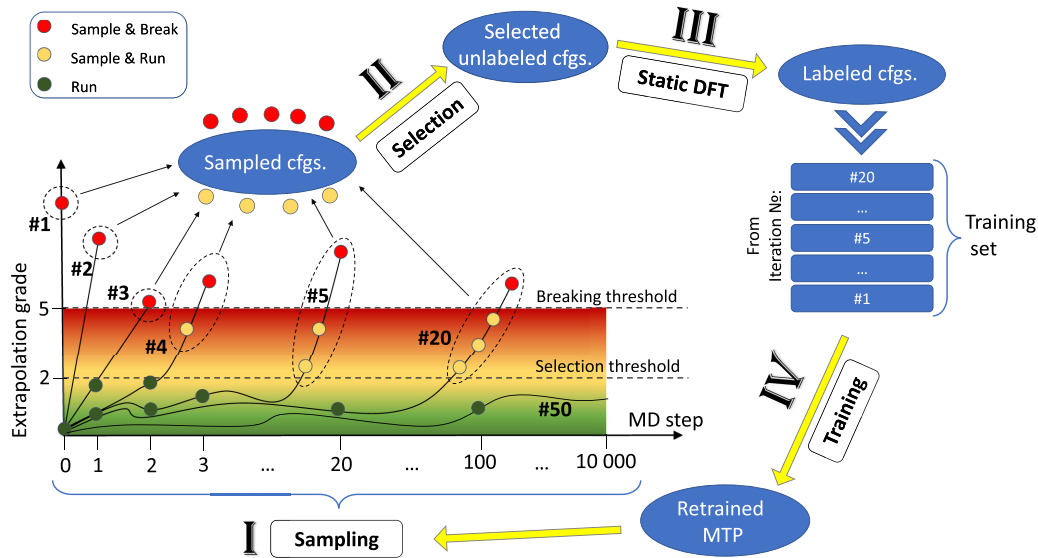


FIG. 1. Graphical illustration of the active learning scheme. (I) Configurations are sampled from MD trajectories (# $N$  is the trajectory of a certain AL iteration). (II) Sampled configurations are shortlisted based on the MaxVol criterion. (III) After assigning DFT energies/forces/stresses they are appended to the training set. (IV) Training of a new MTP able to “work” in larger configurational space. This procedure is repeated until no new configurations are sampled from the MD trajectory (AL iteration #50). All numerical values are approximate, but they provide a correct overview of the typical AL procedure and are valid for the present research as well.

initial training set, e.g., of configurations with presumably high relevance based on *a priori* knowledge of the material under investigation. For a high entropy alloy, for example, this could be a set of randomly disordered configurations. In our case we initiated the training set with DFT relaxation trajectories of random configurations with different Ta content. During the simulation and first MD steps, the initial point in configurational space is extended and the AL algorithm complements the training set with the relevant configurations from the so far unknown regions of the configurational space. Once enough relevant configurations are gathered into the training set, the retrained potential can be used to continue the simulation without further retraining.

#### D. Active learning: Scheme

The AL scheme used in the present work is based on the approach proposed for single-component systems in Ref. [52] and extended to multicomponent cases in Refs. [49,50]. For the MD simulations we use the LAMMPS [54] package ([55]) in combination with the MTPs from the MLIP package as force fields. In the present work, the AL approach with MTPs is applied to a multicomponent problem in combination with MD. The illustration of the algorithm is provided in Fig. 1 and consists of the following steps.

(I) Sampling. A LAMMPS simulation starts and continues until it encounters a case of significant extrapolation (trajectories #1 to #20 in Fig. 1). The simulation is then terminated and all configurations with  $\gamma$  exceeding the predefined *selection threshold* are sampled from the current MD trajectory for the next AL step.

This process can be parallelized to accelerate the convergence and gathering of relevant training data by conducting several simulations in parallel (differing, e.g., by initial ve-

locity distributions, element compositions, starting structures, etc.).

If the AL algorithm does not detect further configurations beyond the threshold at this step, the potential is sufficiently trained and ready for performing the actual simulations of interest.

(II) Selection of samples. After each stage of sampling, the accumulated list of configurations to be added to the training list could, in principle, contain similar configurations, e.g., with very similar atomic positions. Since the DFT calculations consume more than 90% of the whole AL procedure, only the most relevant samples providing the most unique information to the potential should be considered. To that end, we again resort to the MaxVol criterion to shortlist not more than 50–100 configurations to be sent to DFT in each sampling step. Note that for the MaxVol criterion and hence for the selection procedure the atomic geometric arrangement is sufficient (i.e., unlabeled data), while energies or forces are not required.

(III) DFT processing of selected samples. At this stage we perform static DFT calculations for the selected samples assigning DFT energies, forces, and stresses to them. With this the data becomes “labeled” and suitable for (re)training.

(IV) (Re)training. The new set of labeled data is added to the training set and the potential is retrained. After retraining, during subsequent MD runs the potential will not consider these previously sampled configurations as extrapolations anymore and thus not break the MD calculation as sketched in Fig. 1.

#### E. Active learning: Usage

Here we specify the application of the AL approach to the problem addressed in the present work, i.e., the prediction of

elastic constants for a given set of alloys and temperatures by means of MD simulations.

Before carrying out the elastic constants calculations one needs to determine the temperature-dependent equilibrium volume for each alloy. This is done via standard *NPT* simulations. Already during these initial *NPT* simulations, the sampling scheme provides relevant configurations for training the potential. Eventually, the goal is to predict elastic constants via stress-strain relations. In practice the applied strains are about  $\approx 1\%$ . The respective displacements of individual atoms for such strains are much smaller as compared to typical displacements due to thermal vibrations at elevated temperatures. Given that MTPs are local potentials that decompose a given configuration into a set of neighborhoods, the structures entering the training set from the *NPT* calculations provide already a subset of the configurations required for the strain simulations. This way, starting the AL with *NPT* simulations at different temperatures does not only provide the equilibrium volumes versus temperature but allows us already to gather relevant samples for training of the potential for the subsequent elastic constant (strain) simulations.

For the Ti-Zr-Hf-Ta alloys in the focus of our investigations further retraining is, however, needed. The main reason for that is a phase transformation in the considered composition range. At temperatures below 900 K some alloys tend to transform into  $\omega$ -like structures (Sec. II F). If MTPs are trained only on bcc structures, encountering  $\omega$  structures results in high extrapolation grades (see Fig. 1), which indicates the necessity of further retraining. Nevertheless, it is important to stress that, even though the MTP is initially trained on bcc configurations only, the AL approach successfully reveals regimes of bcc lattice instability and enables the prediction of the transition to  $\omega$ -like structures [25].

In Sec. IV we study several alloy compositions in a broad temperature range (300–2500 K). A single MTP has been constructed with the above AL scheme and used for all compositions and temperatures. The potential has been trained on 96 chemical configurations of  $\text{TiZrHfTa}_x$  with varying Ta concentrations between 0 and 33 at. %. For each concentration, 12 random configurations have been considered. Apart from greatly simplifying the process of training, using a single potential provides coherent results among different temperatures/concentrations.

### F. Structural descriptor for $\omega$ vs bcc

For unary metals with ideal bcc lattice we can distinguish the low-temperature  $\omega$  phase and the high-temperature bcc phase based on their crystallographic symmetry differences (note though that the defective  $\omega$  structures with line chain defects have to be carefully considered [24]). In contrast, for chemically disordered alloys like the here investigated  $\text{TiZrHfTa}_x$  alloys (Sec. IV), this is not straightforwardly possible, because the ideal crystallographic symmetry is broken due to lattice distortions caused by the chemical disorder. To overcome this difficulty, an approach utilizing projected atomic displacements was recently proposed [25]. This approach utilizes the fact that the  $\omega$  structure can be obtained by a martensitic transformation from the bcc structure via atomic displacements along one of the  $\langle 111 \rangle_{\text{bcc}}$  directions. In

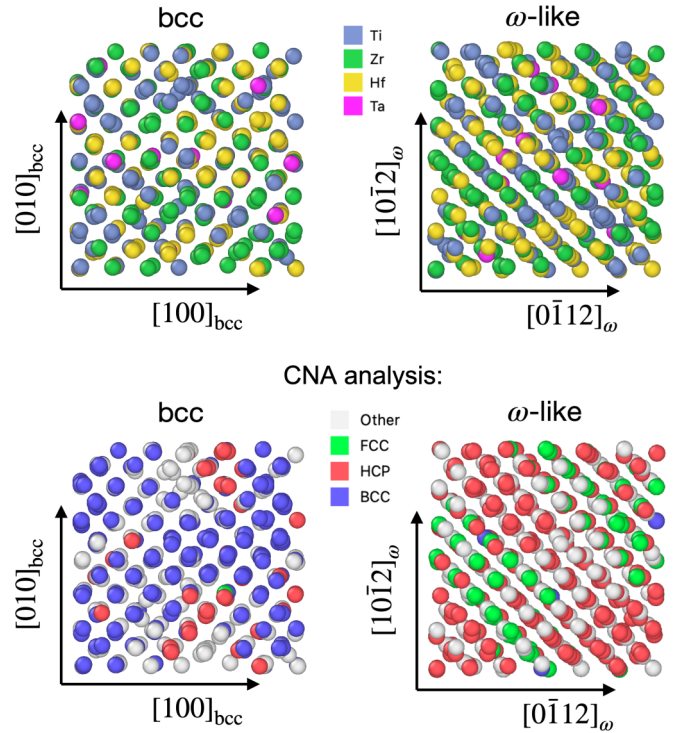


FIG. 2. Planar projections of bcc and  $\omega$ -like structures obtained by averaging over the MD trajectories of two SQS samples, along with the results of the adaptive CNA analysis (bottom panels) [56].

particular, the magnitudes of atomic displacements along the possible directions of such a transformation are analyzed. The relevant details of this approach are given in the following.

In a first step, the average atomic positions  $\langle \mathbf{r}_i \rangle$  of an MD trajectory are computed, with  $i$  labeling the atoms. The displacements with respect to the ideal bcc lattice sites  $\mathbf{r}_i^{\text{bcc,ideal}}$  are then given by

$$\langle \mathbf{d}_i \rangle = \langle \mathbf{r}_i \rangle - \mathbf{r}_i^{\text{bcc,ideal}}. \quad (3)$$

Note that the such obtained  $\langle \mathbf{d}_i \rangle$  do *not* contain any time dependency of atomic vibrations due to the averaging over MD steps in  $\langle \mathbf{r}_i \rangle$  and also that the reference  $\mathbf{r}_i^{\text{bcc,ideal}}$  are adequately shifted so that the sum of the displacements over atoms is zero. The averaged atomic positions do include, however, the implicit temperature-dependent impact of vibrations, e.g., the stabilization of the bcc phase at higher temperatures. Examples of averaged atomic geometries are given in Fig. 2. The such obtained  $\langle \mathbf{d}_i \rangle$  are then projected onto the four possible directions of the  $\omega$ -bcc transformation  $[111]_{\text{bcc}}$ ,  $[1\bar{1}\bar{1}]_{\text{bcc}}$ ,  $[\bar{1}\bar{1}1]_{\text{bcc}}$ , and  $[\bar{1}1\bar{1}]_{\text{bcc}}$  (Fig. 3) with further averaging over all MD snapshots and over all atoms providing

$$p_{[111]_{\text{bcc}}} \equiv \overline{|\langle \mathbf{d} \rangle \cdot \mathbf{e}_{[111]_{\text{bcc}}}|} = \frac{1}{n} \sum_{i=1}^n |\langle \mathbf{d}_i \rangle \cdot \mathbf{e}_{[111]_{\text{bcc}}}|, \quad (4)$$

and three similar quantities corresponding to the other directions:

$$p_{[1\bar{1}\bar{1}]_{\text{bcc}}}, p_{[\bar{1}\bar{1}1]_{\text{bcc}}}, p_{[\bar{1}1\bar{1}]_{\text{bcc}}}. \quad (5)$$



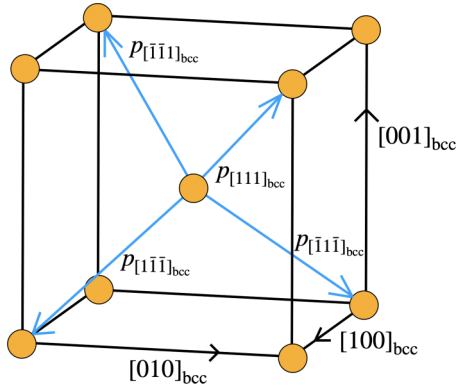


FIG. 3. Illustration of the  $p_{(111)_{bcc}}$  directions onto which the atomic displacements are projected. These directions correspond to the possible martensitic variants of the transformation from bcc to  $\omega$ , which, for pure elemental bcc systems, would be symmetrically equivalent.

These quantities are sorted according to their values in descending order and denoted as

$$p_1, p_2, p_3, p_4. \quad (6)$$

The measure of proximity of an arbitrary structure with respect to the  $\omega$  structure, i.e., the structural descriptor, is then characterized either by  $p_1$  or by  $p_1 - p_2$  [25].

In our case we use the  $p_1 - p_2$  measure as it provides a zero value for a purely bcc-like structure, when no displacement direction is preferred over another. In contrast to this, for  $\omega$ -like structures (Fig. 7) an anisotropy is observed in one of the four possible displacement directions due to the preferred transformation from bcc towards  $\omega$ -like structures. The unit for the  $p_i$  values is  $c_\omega$ , the lattice parameter of the  $\omega$ -structure coherent to the given bcc structure with lattice parameter  $a_{bcc}$  (taken for each composition/temperature correspondingly). It is given as

$$c_\omega = \frac{\sqrt{3}}{2} a_{bcc}. \quad (7)$$

We provide the results of the application of this scheme to the series of TiZrHfTa<sub>x</sub> alloys in Sec. IV.

### III. TECHNICAL DETAILS

#### A. DFT calculations

DFT calculations were carried out with VASP [57–60] using the projector augmented wave (PAW) method [61] and the Perdew-Burke-Ernzerhof generalized gradient approximation (PBE-GGA) [62] for the exchange correlation energy.

The plane wave cutoff energy was chosen to be 250 eV and  $k$ -point meshes of  $3 \times 3 \times 3$  for 54- and 128-atom supercells were employed for the different alloys. The Methfessel-Paxton method (order 1) with a smearing parameter of 0.1 eV was used. The following numbers of electrons were treated as valence in the PAW potentials: 10 for Ti, 12 for Zr, 10 for Hf, and 11 for Ta. We used 54-atom supercells for Ta concentrations of 0 at. %, 5.6 at. %, 11.1 at. %, 16.6 at. %, and 33.3 at. % and 128-atom supercells for 2.1 at. %, 3.5 at. %, and 8.3 at. % to improve the resolution in the composition

TABLE I. Accuracy of the MTP developed for TiZrHfTa<sub>x</sub>. Note that a single MTP is utilized to describe the TiZrHfTa<sub>x</sub> system with varying Ta content from 0 at. % to 33 at. % including chemically disordered bcc and  $\omega$ -like structures in a wide temperature range as given in the table. Errors are measured on a validation set sampled from MD runs at different temperatures which were not part of the training set.

Number of fitting parameters	932
Temperature range (K)	100–2500
Size of the final training set	3853
Energy RMSE (meV/atom)	8.6
Force componentwise RMSE (eV/Å)	0.148

space. In total 3432 and 421 configurations of 54-atom and 128-atom supercells, respectively, were eventually used to train the potential.

#### B. MTP fitting

An overview of the MTP fitting data and quality of fitting is provided in Table I. The MTP energy accuracy for the TiZrHfTa<sub>x</sub> alloys is 8.6 meV/atom. We attribute this comparably large value to the fact that a single MTP was used to fit a broad range of temperatures including the (pre)melting region, a range of different Ta concentrations, and also two crystallographically different phases, namely  $\omega$ -like and bcc. To verify this we inspect the mean average energy difference,

$$\langle \Delta E \rangle = \langle \Delta E_{\text{DFT}} - E_{\text{MTP}} \rangle := \Delta E(x, T), \quad (8)$$

where the average is performed for each considered composition and temperature over six different configurations. The results are shown in Fig. 4, revealing that the largest contribution to the overall energy RMSE comes from the high-

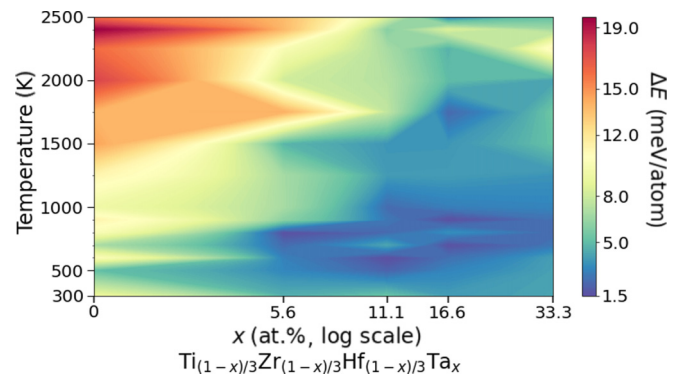


FIG. 4. Distribution of the energy per atom errors vs temperature and Ta content for the TiZrHfTa<sub>x</sub> alloys. The largest contribution to the RMSE of the energies,  $\Delta E(x, T)$ , in Table I is stemming from the region where the simulations reveal overheating (premelting). Also, energies of  $\omega$ -like structures (low Ta content and low temperatures) are predicted less accurately than those for the bcc structures (the structural distinction is discussed in Sec. IV). Note that fit preference was given to forces rather than energies to improve the reproduction of the dynamic and elastic properties. The Ta concentration on the  $x$  axis is plotted in logarithmic scale to improve the visualization (“0” represents here zero Ta concentration though).

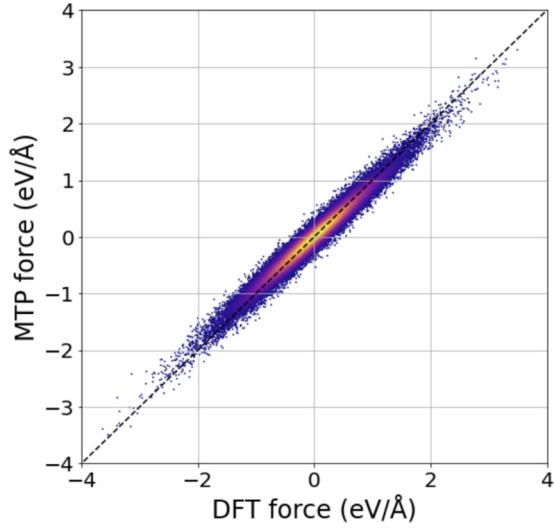


FIG. 5. Correlation between DFT and MTP forces from  $NVT$  MD trajectories for  $T \leq 1500$  K.

temperature region, which, as will be discussed in Sec. IV below, can be attributed to an overheating regime in the underlying MD simulations [63]. To also illustrate the quality of the MTP forces specifically for the crystalline structures ( $\omega$  and bcc), we provide a correlation plot in Fig. 5 for the DFT and MTP forces for the respective MD trajectories from the validation set.

### C. MD simulations

For the MD simulations we used LAMMPS with the fitted MTP as an interatomic potential. For each temperature  $T$  the equilibrium volumes  $V = V_{\text{eq}}(T)$  were obtained from the initial  $NPT$  runs. To compute the elastic constants as well as the displacement projections to detect the structural  $\omega$ -bcc stability (Sec. IV), simulations were carried out utilizing  $NVT$  ensembles with a time step of 1 fs for in total 60 000 MD steps. The first 10 000 MD steps from these were neglected for equilibration and the remaining 50 000 MD steps were used to perform the average for computing the stresses and atomic displacements.

### D. Chemical disorder

To mimic chemical disorder, we used 432-atom bcc supercells and populated the lattice sites with the constitutive elements according to the special quasirandom structure (SQS) approach [64]. For generating the SQS we used the ATAT package [65,66] and the MCSQS procedure [67], utilizing a Monte Carlo algorithm to minimize the pair correlations for the first two shells. The resulting correlation mismatch was on average less than 0.01. The following Ta atomic concentrations,  $x$ , were considered for  $\text{TiZrHfTa}_x$ , namely 0 at. %, 2.1 at. %, 3.5 at. %, 5.6 at. %, 8.3 at. %, 11.1 at. %, 16.6 at. %, and 33.3 at. %.

For detecting the  $\omega$ -bcc stability, 24 SQS structures per composition and per temperature were considered and four different SQS structures were chosen to compute the elastic constants.

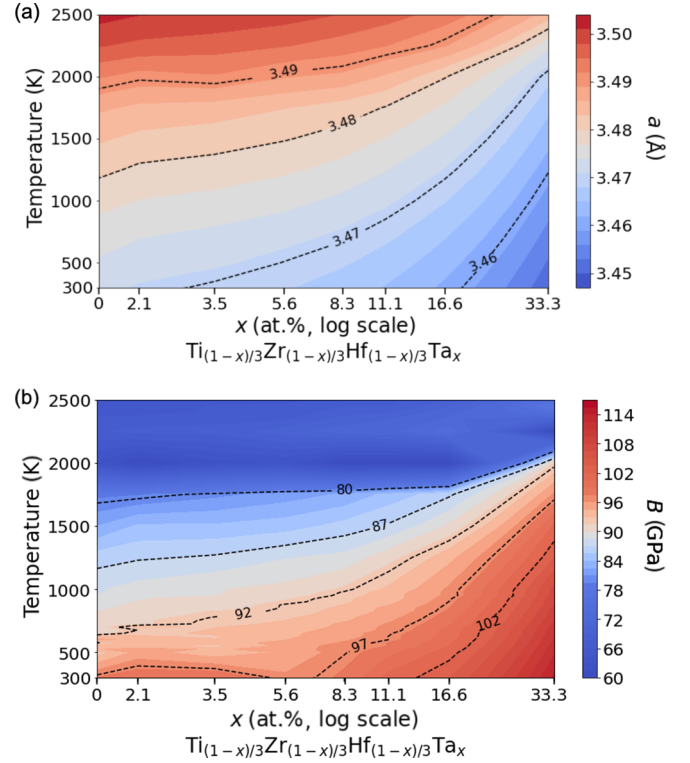


FIG. 6. Composition- and temperature-dependent equilibrium lattice constant  $a(x, T)$  and bulk modulus  $B(x, T)$  of  $\text{TiZrHfTa}_x$ . The bulk modulus reveals a fingerprint of the  $\omega$ -bcc transition at low temperatures and small Ta concentrations, which is not observed in the lattice parameter.

### E. Elastic constants

We computed the elastic constants via stress-strain relations with deformations of  $\pm 1\%$  applied in the uniaxial regime, separately for each  $C_{ij}$  (Voigt notation). We then averaged the elastic constants over the three crystallographically equivalent directions assuming cubic symmetry as

$$\tilde{C}_{11} := \frac{1}{3}(C_{11} + C_{22} + C_{33}), \quad (9)$$

$$\tilde{C}_{12} := \frac{1}{3}(C_{12} + C_{13} + C_{23}), \quad (10)$$

$$\tilde{C}_{44} := \frac{1}{3}(C_{44} + C_{55} + C_{66}), \quad (11)$$

with the bulk modulus given as

$$B = \frac{\tilde{C}_{11} + 2\tilde{C}_{12}}{3}. \quad (12)$$

Note that the above formulas assume cubic symmetry as anticipated for the random solid solution. This symmetry may be broken in the presence of  $\omega$ -like structures.

## IV. RESULTS AND DISCUSSION

We first discuss the composition and temperature dependence of the equilibrium lattice constant,  $a(x, T)$ , of the  $\text{TiZrHfTa}_x$  alloys as derived from the  $NPT$  simulations. As shown in Fig. 6(a) we do not observe any remarkable peculiarities at low temperatures. The lattice constant decreases smoothly with increasing Ta concentration, which is a direct

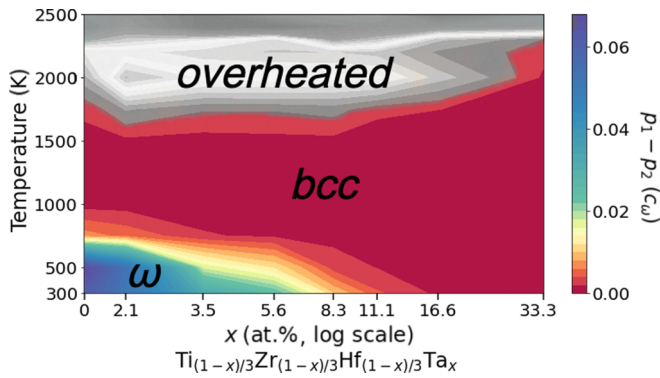


FIG. 7. Structural descriptor, i.e., the averaged difference between the two largest displacements,  $(p_1 - p_2)(x, T)$ , for  $\text{TiZrHfTa}_x$  depending on Ta concentration  $x$  and temperature. Different structural phases can be clearly classified according to the different value regimes of the structural descriptor. The high-temperature region provides high values of  $(p_1 - p_2)$  due to the loss of crystalline structure; therefore, our  $\omega$ -bcc descriptor is not meaningful here.

consequence of the smaller size of Ta as compared to the larger sizes of Zr and Hf and also to the element-averaged size of  $\text{TiZrHf}$ . The negligible impact of the  $\omega$  to bcc transformation on the lattice constant and thus volume is consistent with the results found for pure Ti [24]. It is a consequence of the displacive nature of the transformation. But, as we will discuss further below, the transformation can have a significant impact on computed elastic properties. A fingerprint of the transformation can be already found in the bulk modulus dependency shown in Fig. 6(b). At room temperature an anomaly is observed between 0 and about 7 at. % Ta concentration.

In order to detect the structural  $\omega$ -bcc phase transition we utilize the recently proposed projected displacement method [25] as also described in Sec. II F above. As a descriptor for this purpose we utilize the difference between the two largest projected displacement values  $(p_1 - p_2)(x, T)$ , depending on Ta concentration and temperature. This is shown in Fig. 7, where the values have been determined by averaging over 24 SQS structures. We observe three very distinct regimes. At low temperatures and low Ta content, the values for  $(p_1 - p_2)$  are clearly nonzero, indicating the presence of  $\omega$ -like structures. For larger Ta content and higher temperatures  $(p_1 - p_2)$  approaches zero (no preferential displacement direction), revealing the stability of the bcc phase. At high temperatures we observe a strong anisotropic behavior of  $(p_1 - p_2)$ . We trace back this feature to an artifact due to overheating of the material. Regarding the  $\omega$ -bcc stability there is a clear and well-pronounced gradient in  $(p_1 - p_2)$  allowing us to distinguish both phases.

For the  $\text{TiZrHf}$  alloy (i.e., no Ta at all) we predict a transition temperature of about 700 K. This value is slightly lower compared to the value of 900 K found using explicit AIMD simulations [25]. The origin is the larger supercell with 432 atoms utilized in the present work, compared to the 54-atom supercells used in Ref. [25]. As discussed in Ref. [24] for the case of pure Ti, a 54-atom supercell size may overestimate the  $\omega$ -bcc transition temperature by about 200 K. A careful supercell size convergence test revealed that the 432-atom

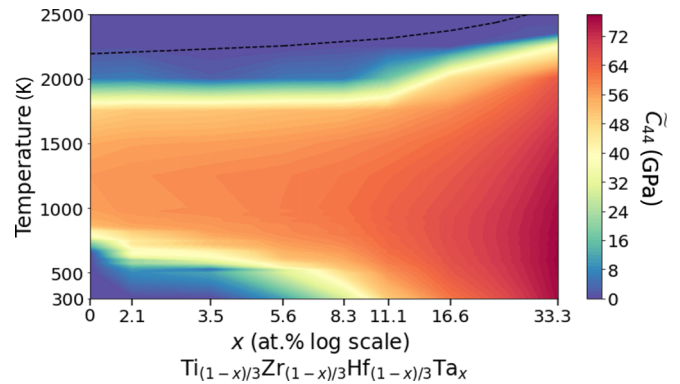


FIG. 8. Composition- and temperature-dependent elastic constant  $\tilde{C}_{44}(x, T)$  for  $\text{TiZrHfTa}_x$  assuming cubic symmetry. Negative values are cropped to zero and correspond to the dark-blue area. The dashed black line at the top of the plot corresponds to the melting temperature. The melting temperature has been approximated by linearly mixing the melting temperatures of the individual constituents.

cells used in this study provide well converged transition temperatures [24].

We now discuss how the structural stability impacts the elastic properties. For this purpose we consider  $NVT$  ensembles (see Sec. II E above) with fixed cubic shape. The volume at each temperature is fixed to the corresponding equilibrium one. The results for  $\tilde{C}_{44}$  are presented in Fig. 8. Due to the fixed cubic shape, the elastic constants are negative (cropped to zero for visualization purposes) in the regime where the  $\omega$  phase becomes stable. This is consistent with the pronounced anisotropic character of the displacive projections shown in Fig. 7. An important outcome of comparing Figs. 7 and 8 is the clear correlation of the  $\omega$ -bcc transition with the magnitude of the elastic constant  $\tilde{C}_{44}$  (other elastic constants are provided in the Appendix). The observed negative elastic constants for high temperatures indicate that the bcc phase becomes unstable against the liquid phase [68]. Since the actual computation of the melting temperature is beyond the present scope, we do not analyze this further. We note, however, that the composition trend of the derived negative elastic constants at high temperatures is consistent with the melting temperatures derived by linear mixing from the experimental values of the individual constituents.

In order to further highlight the correlation and interplay of the  $\omega$  descriptors and elastic constants we show in Fig. 9 the temperature dependencies of the  $\tilde{C}_{44}(T)$  elastic constant along with the  $(p_1 - p_2)(T)$  structural descriptors for several Ta concentrations. A striking observation is how sensitively the temperature dependence of the elastic constants depends on the Ta concentration. A near-zero  $\tilde{C}_{44}$  elastic constant indicates that the bcc phase is dynamically unstable. It is also consistent with the fact that the structural descriptors  $(p_1 - p_2)$  become nonzero and thus indicate  $\omega$  formation. This correlation between  $\tilde{C}_{44}$  and  $(p_1 - p_2)$  is not only observed for severely unstable cases, Figs. 9(c) and 9(d), but is also present for alloys which are bcc stable in the whole temperature range, Fig. 9(a). At very high temperatures, overheating results in negative  $\tilde{C}_{44}$  elastic constants and finite values of  $(p_1 - p_2)$ . Both results are related to the loss of crystal structure, indi-

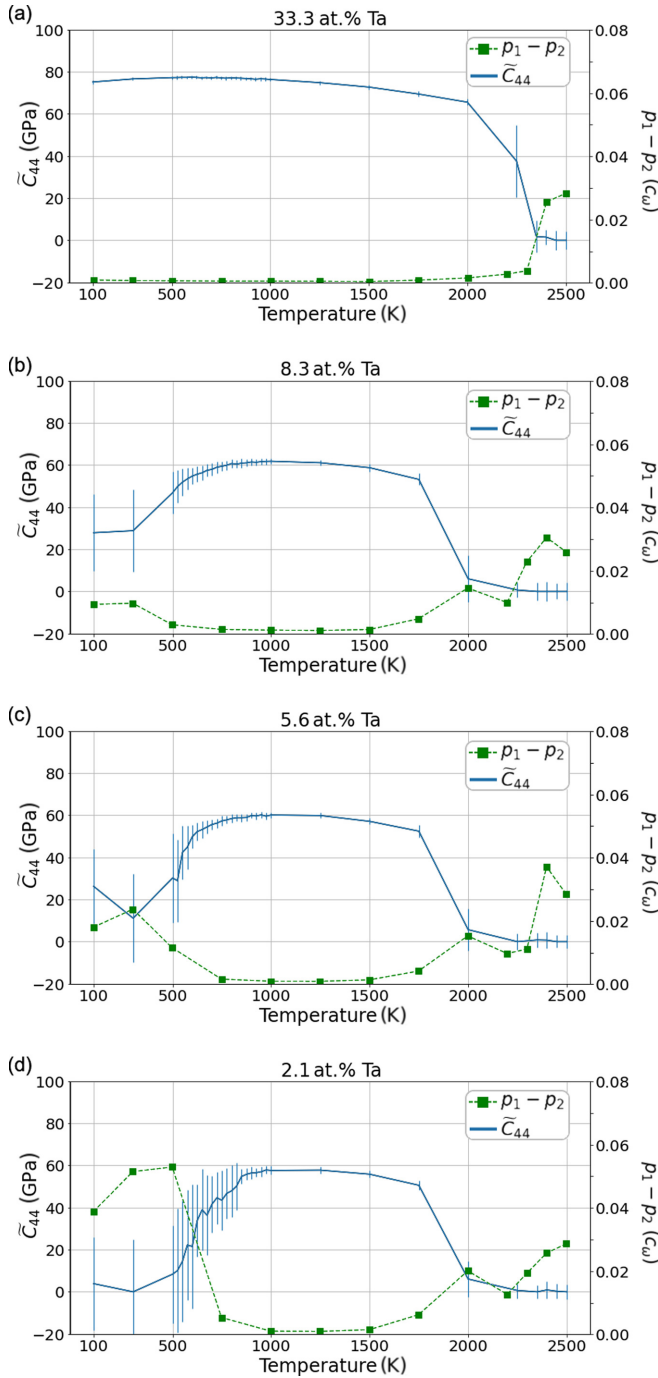


FIG. 9. Comparison between the  $\tilde{C}_{44}$  elastic constant and the order parameter of the  $\omega$ -bcc transition. Error bars correspond to the two- $\sigma$  confidence interval. The large errors at low temperatures are caused by the instability of the bcc phase.

cating that the bcc phase becomes unstable against the liquid phase. Since this phase is not in the focus of the present work we do not analyze it further.

The temperature dependence for many of the considered alloys shows an anomalous behavior: The elastic constants increase with temperature until they reach a maximum after which they decrease. The origin of this behavior is a competition between two factors: The stabilization of the bcc phase

with temperature results in an increase of elastic constants, while the volume expansion with increasing temperature decreases them. This competition can be used to design alloys with elastic constants that are constant over large temperature ranges, as can be seen, e.g., in Fig. 9(a). Similar effects have been recently reported for bcc Ti [20]. With our ML-based scheme multicomponent alloys can be effectively screened to achieve desired temperature-dependent behavior of elastic properties.

## V. CONCLUSION

An active learning approach to construct moment tensor potentials based on DFT data for multicomponent systems has been provided. It facilitates the determination of finite-temperature elastic properties using fast MLIP molecular dynamics simulations with near DFT accuracy. The approach has been applied to study the structural-elastic interplay of TiZrHfTa<sub>x</sub> high entropy alloys. The proposed approach is generic and allows one to screen various alloy compositions over a wide range of temperatures. Its numerical efficiency allows one to utilize it for large system sizes while still preserving near-DFT accuracy. The computed temperature-dependent elastic properties of TiZrHfTa<sub>x</sub> high entropy alloys feature composition-dependent peculiarities. These peculiarities are connected to the stabilization of the bcc phase with increasing Ta concentration and temperature. We showed that based on this insight alloy compositions can be designed to have nearly temperature-independent elastic properties. The presented method therefore allows one to advance high-throughput tools for assessing elastic properties from zero kelvin [69] to finite-temperature calculations. This will open new avenues to design elastically temperature-invariant multicomponent alloys.

## ACKNOWLEDGMENTS

We gratefully acknowledge fruitful discussions with I. A. Abrikosov at Linköping University. Funding from the collaborative DFG-RFBR Grant (Grants No. DFG KO 5080/3-1, No. DFG GR 3716/6-1, and No. RFBR 20-53-12012), from the European Research Council (ERC) under the EU's

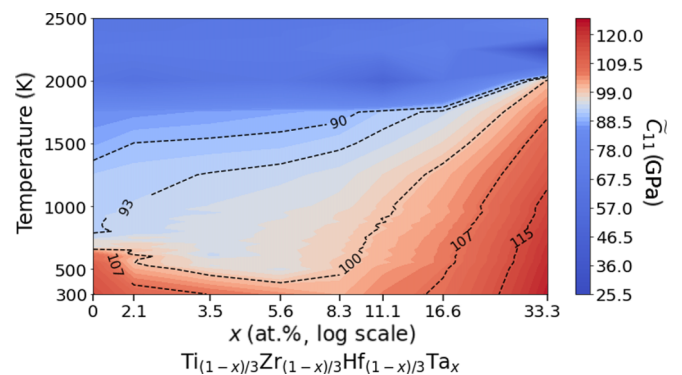


FIG. 10. Composition- and temperature-dependent elastic constant  $\tilde{C}_{11}(x, T)$  for TiZrHfTa<sub>x</sub> assuming cubic symmetry. The heat map contains a fingerprint of the  $\omega$ -bcc transition at low temperatures and small Ta concentrations.



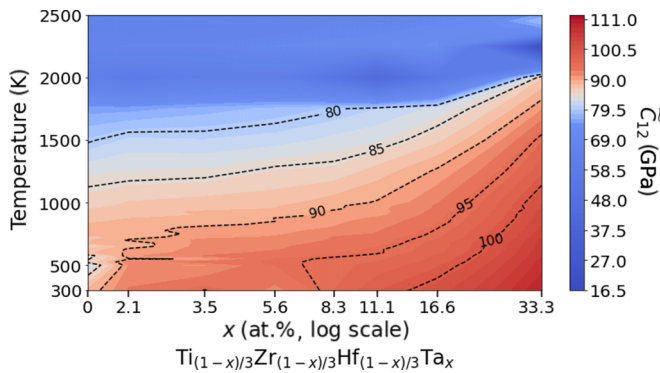


FIG. 11. Composition- and temperature-dependent elastic constant  $\tilde{C}_{12}(x, T)$  for  $\text{TiZrHfTa}_x$  assuming cubic symmetry.

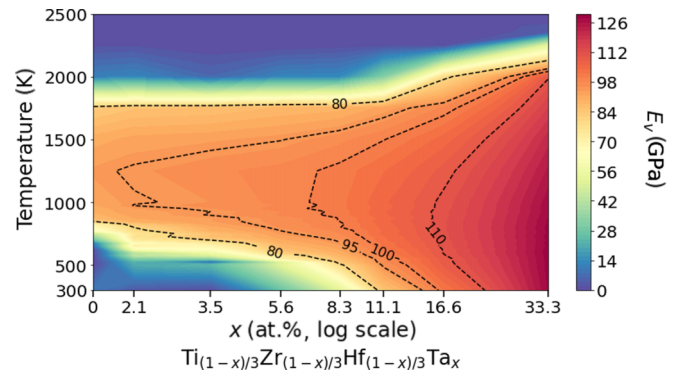


FIG. 12. Composition- and temperature-dependent homogenized Voigt elastic modulus  $E_V(x, T)$  for  $\text{TiZrHfTa}_x$ . Negative values are cropped to zero and correspond to the dark-blue area.

Horizon 2020 Research and Innovation Programme (Grant No. 865855), DFG priority program (SPP 2006), DFG Grant No. 405621160, and from NWO/STW (VIDI Grant No. 15707) are gratefully acknowledged. B.G. acknowledges the support by the Stuttgart Center for Simulation Science (SimTech).

## APPENDIX

For completeness we provide here the results for the computed elastic constants  $\tilde{C}_{11}$  (Fig. 10) and  $\tilde{C}_{12}$  (Fig. 11), as well

as the homogenized elastic modulus (Fig. 12) derived from the Voigt ratio:

$$E_V = \frac{(\tilde{C}_{11} - \tilde{C}_{12} + 3\tilde{C}_{44})(\tilde{C}_{11} + 2\tilde{C}_{12})}{(2\tilde{C}_{11} + 3\tilde{C}_{12} + \tilde{C}_{44})}.$$

- [1] M. Babanli, F. Prima, P. Vermaut, L. Demchenko, A. Titenko, S. Huseynov, R. Hajiyev, and V. Huseynov, *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing* (Springer, New York, 2018), pp. 937–944.
- [2] T. Zhou, Z. Song, and K. Sundmacher, *Engineering* **5**, 1017 (2019).
- [3] Y. Cho, W. J. Cho, I. S. Youn, G. Lee, N. J. Singh, and K. S. Kim, *Acc. Chem. Res.* **47**, 3321 (2014).
- [4] W. Yu and A. D. MacKerell, *Antibiotics* (Springer, New York, 2017), pp. 85–106.
- [5] Y. Noda, M. Otake, and M. Nakayama, *Sci. Technol. Adv. Mater.* **21**, 92 (2020).
- [6] G. B. Goh, N. O. Hodas, and A. Vishnu, *J. Comput. Chem.* **38**, 1291 (2017).
- [7] E. P. George, D. Raabe, and R. O. Ritchie, *Nat. Rev. Mater.* **4**, 515 (2019).
- [8] D. B. Miracle and O. N. Senkov, *Acta Mater.* **122**, 448 (2017).
- [9] F. Otto, Y. Yang, H. Bei, and E. P. George, *Acta Mater.* **61**, 2628 (2013).
- [10] Y. Ikeda, B. Grabowski, and F. Körmann, *Mater. Charact.* **147**, 464 (2019).
- [11] J. Rickman, G. Balasubramanian, C. Marvel, H. Chan, and M.-T. Burton, *J. Appl. Phys.* **128**, 221101 (2020).
- [12] K. Kaufmann, D. Maryanovsky, W. M. Mellor, C. Zhu, A. S. Rosengarten, T. J. Harrington, C. Oses, C. Toher, S. Curtarolo, and K. S. Vecchio, *npj Comput. Mater.* **6**, 1 (2020).
- [13] D. Ma, M. Yao, K. G. Pradeep, C. C. Tasan, H. Springer, and D. Raabe, *Acta Mater.* **98**, 288 (2015).
- [14] X. Wang, L. Liu, M. Wang, X. Shi, G. Huang, and L. Zhang, *Calphad* **48**, 89 (2015).
- [15] N. Shulumba, O. Hellman, L. Rogström, Z. Raza, F. Tasnádi, I. A. Abrikosov, and M. Odén, *Appl. Phys. Lett.* **107**, 231901 (2015).
- [16] X. Zhang, W. Huang, J. Chen, C. Liu, H. Yu, L. Zhao, and W. Jiang, *Vacuum* **157**, 312 (2018).
- [17] Z. Deng, K. Chu, Q. Li, Y. Onuki, and Q. Sun, *Scr. Mater.* **187**, 197 (2020).
- [18] A. Ahadi, R. Khaledialidusti, T. Kawasaki, S. Harjo, A. Barnoush, and K. Tsuchiya, *Acta Mater.* **173**, 281 (2019).
- [19] J. Cui and X. Ren, *Appl. Phys. Lett.* **105**, 061904 (2014).
- [20] A. V. Shapeev, E. V. Podryabinkin, K. Gubaev, F. Tasnádi, and I. A. Abrikosov, *New J. Phys.* **22**, 113005 (2020).
- [21] Y. Qiang, S. Jian, X. Hui, and W.-Y. Guo, *Trans. Nonferrous Met. Soc. China* **17**, 1417 (2007).
- [22] S. Wang, M. Wu, D. Shu, G. Zhu, D. Wang, and B. Sun, *Acta Mater.* **201**, 517 (2020).
- [23] N. Wei, T. Jia, X. Zhang, T. Liu, Z. Zeng, and X. Yang, *AIP Adv.* **4**, 057103 (2014).
- [24] D. Korbmayer, A. Glensk, A. I. Duff, M. W. Finnis, B. Grabowski, and J. Neugebauer, *Phys. Rev. B* **100**, 104110 (2019).
- [25] Y. Ikeda, K. Gubaev, J. Neugebauer, B. Grabowski, and F. Körmann, *npj Comput. Mater.* **7**, 1 (2021).
- [26] D. Ma, B. Grabowski, F. Körmann, J. Neugebauer, and D. Raabe, *Acta Mater.* **100**, 90 (2015).

- [27] J. Miao, C. Slone, T. Smith, C. Niu, H. Bei, M. Ghazisaeidi, G. Pharr, and M. J. Mills, *Acta Mater.* **132**, 35 (2017).
- [28] F. Tian, L. K. Varga, J. Shen, and L. Vitos, *Comput. Mater. Sci.* **111**, 350 (2016).
- [29] A. Ferrari, B. Dutta, K. Gubaev, Y. Ikeda, P. Srinivasan, B. Grabowski, and F. Körmann, *J. Appl. Phys.* **128**, 150901 (2020).
- [30] J. Behler, *J. Chem. Phys.* **145**, 170901 (2016).
- [31] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood *et al.*, *J. Phys. Chem. A* **124**, 731 (2020).
- [32] I. Novikov, B. Grabowski, F. Kormann, and A. Shapeev, [arXiv:2012.12763](https://arxiv.org/abs/2012.12763).
- [33] B. Grabowski, Y. Ikeda, P. Srinivasan, F. Körmann, C. Freysoldt, A. I. Duff, A. Shapeev, and J. Neugebauer, *npj Comput. Mater.* **5**, 1 (2019).
- [34] S. Chmiela, H. E. Sauceda, I. Poltavsky, K.-R. Müller, and A. Tkatchenko, *Comput. Phys. Commun.* **240**, 38 (2019).
- [35] L. Zhang, J. Han, H. Wang, R. Car, and W. E, *Phys. Rev. Lett.* **120**, 143001 (2018).
- [36] M. Gastegger and P. Marquetand, *J. Chem. Theory Comput.* **11**, 2187 (2015).
- [37] G. P. Pun, R. Batra, R. Ramprasad, and Y. Mishin, *Nat. Commun.* **10**, 1 (2019).
- [38] T. Lookman, P. V. Balachandran, D. Xue, and R. Yuan, *npj Comput. Mater.* **5**, 1 (2019).
- [39] X.-G. Li, C. Chen, H. Zheng, Y. Zuo, and S. P. Ong, *npj Comput. Mater.* **6**, 1 (2020).
- [40] J. S. Smith, B. Nebgen, N. Mathew, J. Chen, N. Lubbers, L. Burakovsky, S. Tretiak, H. A. Nam, T. Germann, S. Fensin *et al.*, *Nat. Commun.* **12**, 1 (2021).
- [41] R. Jinnouchi, F. Karsai, and G. Kresse, *Phys. Rev. B* **100**, 014105 (2019).
- [42] N. Artrith and J. Behler, *Phys. Rev. B* **85**, 045439 (2012).
- [43] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky, *npj Comput. Mater.* **6**, 1 (2020).
- [44] J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, and A. E. Roitberg, *J. Chem. Phys.* **148**, 241733 (2018).
- [45] <https://mlip.skoltech.ru/download/>.
- [46] <https://gitlab.com/ashapeev/mlip-2/-/tree/mlippy>.
- [47] I. S. Novikov, K. Gubaev, E. Podryabinkin, and A. V. Shapeev, *Mach. Learn.: Sci. Technol.* **2**, 025002 (2020).
- [48] A. V. Shapeev, *Multiscale Model. Simul.* **14**, 1153 (2016).
- [49] K. Gubaev, E. V. Podryabinkin, and A. V. Shapeev, *J. Chem. Phys.* **148**, 241727 (2018).
- [50] K. Gubaev, E. V. Podryabinkin, G. L. Hart, and A. V. Shapeev, *Comput. Mater. Sci.* **156**, 148 (2019).
- [51] R. Battiti and F. Masulli, *International Neural Network Conference* (Springer, New York, 1990), pp. 757–760.
- [52] E. V. Podryabinkin and A. V. Shapeev, *Comput. Mater. Sci.* **140**, 171 (2017).
- [53] S. Goreinov, I. V. Oseledets, D. V. Savostyanov, E. E. Tyrtshnikov, and N. Zamarashkin, *Matrix Methods: Theory Algorithms and Applications: Dedicated to the Memory of Gene Golub* (World Scientific, Singapore, 2010), pp. 247–256.
- [54] S. Plimpton, *J. Comput. Phys.* **117**, 1 (1995)
- [55] <http://lammps.sandia.gov>.
- [56] A. Stukowski, V. V. Bulatov, and A. Arsenlis, *Modell. Simul. Mater. Sci. Eng.* **20**, 085007 (2012).
- [57] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- [58] G. Kresse and J. Hafner, *Phys. Rev. B* **49**, 14251 (1994).
- [59] G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- [60] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [61] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [62] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [63] At each composition and temperature, the volume has been fixed to the equilibrium one from *NPT* simulations and for each the configuration after 5000 MD steps has been chosen (time step 1 fs). The six configurations are chosen from the set of special quasirandom structures.
- [64] A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, *Phys. Rev. Lett.* **65**, 353 (1990).
- [65] A. van de Walle, M. D. Asta, and G. Ceder, *Calphad* **26**, 539 (2002).
- [66] A. van de Walle, *Calphad* **33**, 266 (2009).
- [67] A. van de Walle, P. Tiwary, M. M. de Jong, D. L. Olmsted, M. D. Asta, A. Dick, D. Shin, Y. Wang, L.-Q. Chen, and Z.-K. Liu, *Calphad* **42**, 13 (2013).
- [68] M. Born, *J. Chem. Phys.* **7**, 591 (1939).
- [69] M. De Jong, W. Chen, T. Angsten, A. Jain, R. Notestine, A. Gamst, M. Sluiter, C. K. Ande, S. Van Der Zwaag, J. J. Plata *et al.*, *Sci. Data* **2**, 150009 (2015).