Abstract

The development of machine learning sheds new light on the problem of statistical thermodynamics in multicomponent alloys. However, a data-driven approach to construct the effective Hamiltonian requires sufficiently large data sets, which is expensive to calculate with conventional density functional theory (DFT). To solve this problem, we propose to use the atomic local energy as the target variable, and harness the power of the linear-scaling DFT to accelerate the data generating process. Using the large amounts of DFT data sets, various complex models are devised and applied to learn the effective Hamiltonians of a range of refractory high entropy alloys (HEAs). The testing $R^2$ scores of the effective pair interaction model are higher than 0.99, demonstrating that the pair interactions within the 6-th coordination shell provide an excellent description of the atomic local energies for all the four HEAs. This model is further improved by including nonlinear and multi-site interactions. In particular, the deep neural networks (DNNs) built directly in the local configuration space (therefore no hand-crafted features) are employed to model the effective Hamiltonian. The results demonstrate that neural networks are promising for the modeling of effective Hamiltonian due to its excellent representation power.
1. Introduction

Density functional theory (DFT) is a powerful method to calculate the properties of materials from first principles. Combined with Monte Carlo methods, DFT not only can be applied to the ground states, but also be used to study finite temperature systems [1, 2, 3]. However, a direct combination of the two is severely hindered by the expensive computational cost. For example, using a supercell of 250-atom [4], the “brute-force” statistical simulation of the order-disorder transition in CuZn alloy calculated 600,000 configurational energies with DFT, which approximately took $10^8$ CPU hours. One less ambitious but more practical approach is to establish a surrogate model of the effective Hamiltonian from the DFT data, and use this computationally cheap model in the Monte Carlo simulation. An example of such a strategy is the cluster expansion method [5, 6, 7, 8, 9], in which the total energy is expressed in terms of the effective cluster interactions (ECIs) [10]. The EPI parameters are typically calculated with the Connolly and Williams approach (also known as the structure inversion method) [11] from the DFT energies of a set of selected structures. While cluster expansion provides a systematic approach to the construction of the effective Hamiltonian, its application to complex systems remains challenging due to rapid increase of the number of ECIs with respect to the chemical components [12, 13, 1]. This is particularly true for a class of novel materials known as high entropy alloys (HEAs) [14, 15, 16], which contain more than four principal elements, and demonstrate some exceptional mechanical properties [17, 18, 19, 20]. Other than cluster expansion, linear response methods based on the coherent potential approximation (CPA), such as the $S^{(2)}$ theory [21], the generalized perturbation method (GPM) [22], and the embedded-cluster method (ECM) [23], are also widely used to calculate the ECIs. The advantage of these methods is that they directly relate the ECIs to the underlying electronic structure, while the shortcoming is that they are limited by the single-site approximation.

A different strategy to construct the effective Hamiltonian is through supervised machine learning [24, 25]. Compared to the traditional structure inversion method, machine learning adopts a data-driven approach that generally requires large data sets. Moreover, machine learning models are generally more complex, with large numbers of parameters, therefore in principle possess better representation capability than simple models. In supervised machine learning, the training of model is achieved through adjusting the parameters so as to minimize the loss function of the training data. Machine learning has demonstrated great success in a wide range of fields, such as computer vision and natural language processing [26]. In materials science, machine learning has been applied to model physical quantities such as atomic forces [27, 28], interatomic potentials [29, 30, 25, 31], and formation energies [24, 32].

While increased model complexity generally improves the representation capability, it also makes the model sensitive to random noise. This phenomenon is known as the bias-variance trade-off [33]. Generally speaking, sufficiently large training data sets are required for the complex model to demonstrate better performance (smaller out-of-sample error) than the simple ones. As a result, for machine learning the Hamiltonian, it is critical to generate data efficiently so that more accurate models can be employed. In fact, the availability of
large data sets, and the capability to handle them, is one of the most important reasons for the resurgence of deep neural networks (DNN) in the last 10 years [34]. However, as previously mentioned, DFT methods are computationally expensive, and this is the exact reason that we resort to the surrogate model instead of directly using DFT. At first sight, the analysis seems to indicate a dilemma that huge amounts of DFT calculations are inevitable even for the surrogate model approach.

In this work, we propose to solve the above problem by a combination of two ideas: linear-scaling DFT [35] and atomic local energies. Linear-scaling DFT utilizes the nearsightedness principle [36] to reduce the computational complexity of conventional DFT from $O(N^3)$ to $O(N)$, where $N$ is the number of atoms in the supercell. Compared to conventional DFT, the speed advantage of linear-scaling is particularly significant for large supercells. Different linear-scaling DFT methods have been implemented. In this work, we focus on the locally self-consistent multiple scattering (LSMS) method [37]. The LSMS method achieves linear scaling by restricting the quantum scattering of electrons within the so-called local interaction zone, but it still evaluates the electrostatic interactions everywhere. In the LSMS method, the concept of atomic local energy is actually very natural since it always divides the space into Voronoi polyhedra corresponding to each atom. The atomic local energy can thus be obtained by integrating the energy density within the polyhedron, at a negligible computational cost. By using the local chemical environment as the input and atomic local energy as the output, as illustrated in Fig. 1, a single DFT calculation with an $N$-atom supercell generates $N$ data sets. Therefore, the combination of linear-scaling DFT and atomic local energies technically reduces the time scaling of generating one DFT data from $O(N^3)$ to $O(0)$, which renders the use of complex models with thousands of parameters feasible.

2. Results and discussion

In general, the effective Hamiltonian can be obtained from two different approaches. One is to first project the local chemical environment to the space spanned by a series of physical features, such as ECIs, then establish a mapping from the features to the atomic local energy. Note that unlike the traditional cluster expansion, the mapping between ECIs and energy is not necessarily linear. A more interesting approach is to seek a mapping directly from the local environment to the atomic local energy. Without the intermediate dimension reduction, this method is more accurate in principle, and can be easily generalized to other materials. Of course, in both cases a cutoff has to be set on the interaction range, otherwise the feature space will be of infinite dimension. In the first approach, the local energy of the $i$-th lattice site can be written as

$$E_i = H(\bar{\Sigma}_i) + \epsilon,$$

where $\epsilon$ denotes the random error due to the approximations in the model, and $\bar{\Sigma} = (\Sigma^1, \Sigma^2, \cdots, \Sigma^N)$ are the $N_f$ physical features calculated from the local chemical environment. In the second approach, the local energy is given by

$$E_i = H(\tilde{\sigma}_i) + \epsilon,$$
\[ \vec{\sigma} = (\sigma^0, \sigma^1, \ldots, \sigma^{N_L-1}) \] is a configuration of \( N_L \) atoms in the local interaction region (LIR), with \( \sigma^0 \) representing the local atom. After obtaining the local energies, the total energy is simply given by summing them over all lattice sites, i.e.,

\[ E_t = \sum_i E_i. \]  

(3)

Figure 1: (Color online) A schematic to illustrate the dependence of the local energy \( E_i \) on the local chemical environment. The grey circle demonstrates the local interaction region, with the local atom in the center. In practice, the local interaction zone for each atom includes multiple neighboring shells. The different colors of the atoms signify different chemical species. Pair interactions of the nearest neighbor and next nearest neighbor are also shown.

2.1. Linear EPI

Different models of the effective Hamiltonian are investigated in this work. We first start with the linear effective pair interaction (EPI) model, in which the local energy is given by a summation of all the pair interactions within the local interaction region,

\[ E_i = \sum_f V^f \Pi^f(\vec{\sigma}_i) + V^0 + \epsilon, \]  

(4)

where \( V^f \) are the EPI parameters and \( \Pi^f \) are the number of pair interactions of type \( f \). The feature index \( f \) is actually made up of three parts \( (p, p', m) \), representing the element of the local atom, the element of the neighboring atoms, and the coordination shell, respectively. It is obvious that the EPI parameters in Eq. 4 can be obtained with linear regression. The number of data sets is \( N \times N_c \), where \( N_c \) is the number of configurations. In practice, the data sets of an \( n \)-component system are divided into \( n \) parts according to the elements, with one model fitted for each chemical component. To improve the representativeness of the data, the training sets need to contain various order and disorder structures. One simple approach is to carry out the DFT calculation using supercells of different sizes, as illustrated in Fig. 2. Within each supercell, the atoms are still randomly distributed, but due to the
periodic boundary conditions, configurations from small supercells naturally contains a range of ordered structures, while configurations from large supercells are close to random states.

The EPI model with the number of coordination shells $m_{\text{max}} = 6$ is applied to the MoNbTaW refractory HEA and the linear regression results are shown in Fig. 3. It is easy to see that this model gives a very accurate prediction of the atomic local energies, with testing $R^2$ scores higher than 0.996 for all the elements. The fitted EPI parameters are also shown in Fig. 3, and are arranged according to the $p'$ and $m$ indices. Note that due to the constraint that the total number of atoms are fixed for each coordination shell, only $n - 1$ $p'$ indices are independent for an $n$-element system, therefore we adopt the notation that $p' \neq p$ and there are a total of $(n - 1) \times m_{\text{max}}$ parameters for each element. For instance, the 18 EPI parameters in Fig. 3(a) are arranged according to the feature indices (Mo, Nb, 1), (Mo, Ta, 1), (Mo, W, 1), · · · , (Mo, W, 6). It is also easy to see that the nearest neighbor MoTa pair are the strongest pair interaction, which is in agreement with the results in literature [38, 39]. The coordination shell cutoff $m_{\text{max}}$ determines the size of the local interaction region. The impact of different $m_{\text{max}}$ on the root mean square errors (RMSEs) are shown in Fig. 5. It can be seen that the pair interactions within the first 5 shells contribute significantly to the local energies of MoNbTaW, and the effects from longer pair interactions are small. Other than MoNbTaW, we also apply the EPI model to the other three refractory HEAs and the results are listed in Tab. 1. For all the four materials, the testing $R^2$ scores of the local energies are higher than 0.99, demonstrating that the EPI model is generally a good description of the
configuration energies. The $R^2$ score of MoNbTaTiW and AlMoNbTaW are a little lower than MoNbTaW and MoNbTaVW, indicating that the addition of Ti and Al introduces stronger high-order interactions into the system.

\[
E_i = \sum_f V_f \Pi_f(\vec{\sigma}_i) + \sum_{f,f' \leq f} V_{f,f'} \Pi_f(\vec{\sigma}_i) \Pi_{f'}(\vec{\sigma}_i) + V^0 + \epsilon. \tag{5}
\]

The addition of quadratic terms increases the dimension of the feature space from $k = (n-1) \times m_{\text{max}}$ to $k \times (k+3)/2$. Using the MoNbTaTiW HEA as an example, the 6-shell quadratic EPI model thus has a total of 324 feature parameters. To avoid overfitting, these parameters are determined by ridge regression, with the L2 regularization parameter taken as $\alpha = 1.0$. The impact of the data set size on the RMSE of MoNbTaTiW is shown in Fig. 5, from which we can see that thousands of data sets are needed for the quadratic EPI model to converge. The results for all the four refractory HEAs are also shown in Tab. 1.

It is easy to see that the quadratic models indeed demonstrate better performance than the linear models.
Figure 4: (Color online) The dependence of the testing RMSE on the maximum number of coordination shells, for the linear EPI model of MoNbTaW.

Table 1: Summary of the testing $R^2$ scores, root mean square errors (RMSE), and maximum absolute deviations (MAD) for four different refractory HEAs. Both linear regression (LR) and quadratic regression (QR) of the EPI model are employed.

| Material     | Element | LR-$R^2$ | QR-$R^2$ | LR-RMSE | QR-RMSE | LR-MAD | QR-MAD |
|--------------|---------|----------|----------|----------|----------|--------|--------|
| MoNbTaW      | Mo      | 0.99624  | 0.99668  | 0.00131  | 0.00123  | 0.00684| 0.00566|
|              | Nb      | 0.99647  | 0.99675  | 0.00142  | 0.00136  | 0.00592| 0.00539|
|              | Ta      | 0.99640  | 0.99675  | 0.00153  | 0.00145  | 0.00707| 0.00642|
|              | W       | 0.99632  | 0.99673  | 0.00140  | 0.00132  | 0.00683| 0.00609|
| MoNbTaVW     | Mo      | 0.99852  | 0.99882  | 0.00208  | 0.00186  | 0.00935| 0.00762|
|              | Nb      | 0.99823  | 0.99874  | 0.00233  | 0.00197  | 0.01401| 0.00839|
|              | Ta      | 0.99818  | 0.99876  | 0.00249  | 0.00206  | 0.01363| 0.00911|
|              | V       | 0.99819  | 0.99882  | 0.00212  | 0.00171  | 0.01445| 0.01251|
|              | W       | 0.99852  | 0.99885  | 0.00225  | 0.00198  | 0.01072| 0.00874|
| MoNbTaTiW    | Mo      | 0.99480  | 0.99575  | 0.00294  | 0.00266  | 0.01848| 0.01312|
|              | Nb      | 0.99203  | 0.99389  | 0.00356  | 0.00311  | 0.02696| 0.02017|
|              | Ta      | 0.99141  | 0.99331  | 0.00385  | 0.00340  | 0.02273| 0.01988|
|              | Ti      | 0.98915  | 0.99113  | 0.00382  | 0.00345  | 0.02310| 0.02293|
|              | W       | 0.99475  | 0.99561  | 0.00315  | 0.00288  | 0.01975| 0.01372|
| AlMoNbTaW    | Al      | 0.99535  | 0.99645  | 0.00393  | 0.00343  | 0.02231| 0.01729|
|              | Mo      | 0.99075  | 0.99299  | 0.00583  | 0.00508  | 0.02472| 0.02684|
|              | Nb      | 0.99043  | 0.99187  | 0.00573  | 0.00529  | 0.02724| 0.02584|
|              | Ta      | 0.99054  | 0.99203  | 0.00592  | 0.00543  | 0.02908| 0.03731|
|              | W       | 0.99107  | 0.99317  | 0.00604  | 0.00528  | 0.03114| 0.03290|

2.3. Configuration space

In contrast to the EPI models that make use of the lattice symmetry for dimension reduction, the following models work directly on the chemical configuration space. In the simplest case, one can express the atomic local energies as a linear summation of contributions
Figure 5: (Color online) The dependence of the testing RMSE on the number of data sets for each element. The results are from the quadratic EPI model of MoNbTaW. The data points correspond to 248, 744, 1240, 1736, 2232, 2728, 3224, 3720, and 4216 data sets, respectively.

from each lattice sites in the local interaction region, i.e.,

\[ E_i = \sum_{j \in \text{LIR}} V^j \sigma^j_i + V^0 + \epsilon. \]  

(6)

Note that \( \sigma^j_i \) are categorical variables and an embedding procedure is needed to represent them in real space. Since each lattice site has to be occupied by one of the \( n \) elements, we choose to use \( n - 1 \) dummy variables to encode \( \sigma^j_i \). For example, the representations of Mo, Nb, Ta, Ti, W in the MoNbTaTiW HEA are \((0, 0, 0, 0)\), \((1, 0, 0, 0)\), \((0, 1, 0, 0)\), \((0, 0, 1, 0)\), and \((0, 0, 0, 1)\), respectively. As a result, for \( m_{\text{max}} = 6 \) (64 atoms within the LIR), the number of features for one element increases from 24 in the linear EPI model to 256. Due to lattice symmetry, \( V^j \) of the same bonding type should have the same value, so this configuration space linear model is essentially the same as the linear EPI model. However, the benefits of the configuration space representation are that the complete information about the local chemical environment is preserved, and the inclusion of higher order interactions is straightforward. For example, adding quadratic interactions to the nearest neighbors, the local energies can be written as

\[ E_i = \sum_{j \in \text{LIR}} V^j \sigma^j_i + \sum_{j,k>j \in \text{LIR}} V^{jk} \sigma^j_i \sigma^k_i + V^0 + \epsilon, \]  

(7)

which amounts to include all the triplet interactions within the first coordination shell. Using MoNbTaTiW as an example, the calculation results of this model (1ST) are shown in Tab. 2. Compared to the data in Tab. 1, it is easy to see that this model performs better than the linear EPI model, but the improvement is very small, indicating that these triplet interactions are not important for the MoNbTaTiW HEA.

2.4. Neural Network Model

Deep neural networks (DNNs) provide a general solution to the problem of constructing effective Hamiltonian in the configuration space. This is a very attractive approach because
Table 2: Summary of the testing $R^2$ score, root mean square error (RMSE), and maximum absolute deviation (MAD) for the configuration space models of MoNbTaTiW. 1ST stands for the first-shell-triplet model and NN stands for the neural network model.

| Material Element | 1ST-$R^2$ | NN-$R^2$ | 1ST-RMSE | NN-RMSE | 1ST-MAD | NN-MAD |
|------------------|-----------|----------|-----------|----------|---------|--------|
| Mo               | 0.99505   | 0.99306  | 0.00287   | 0.00340  | 0.01341 | 0.01706|
| Nb               | 0.99292   | 0.98823  | 0.00335   | 0.00432  | 0.02200 | 0.02547|
| Ta               | 0.99263   | 0.98977  | 0.00357   | 0.00420  | 0.01863 | 0.02237|
| Ti               | 0.98980   | 0.98703  | 0.00370   | 0.00417  | 0.02432 | 0.03418|
| W                | 0.99485   | 0.99305  | 0.00312   | 0.00363  | 0.01454 | 0.02421|

Figure 6: A schematic of the neural networks used to model the effective Hamiltonian. The input layer is made up of dummy variables representing the local chemical environment.

it involves no “handcrafted” features and is exact in principle, according to the universal approximation theorem [40]. A schematic of the neural network architecture employed in this work is shown in Fig. 6 and applied to the MoNbTaTiW HEA. In practice, 20 neurons are adopted for each of the 3 hidden layers, where the rectified linear (ReLU) activation functions are used. The input layer is made up of 256 dummy variables ($m_{max} = 6$) and the calculation details are described in the method section. The total number of epochs is 150, and the convergence of RMSE with respect to the epoch number are shown in Fig. 7(a). The $R^2$ scores, RMSEs, and MADs are listed in Tab. 2 and the comparisons between the predicted energy and the DFT data are shown in Fig. 7(b)-(f). Compared to the linear EPI results, it is easy to note that the neural networks produce very high training scores, which demonstrates the superior representation capability of DNN. The testing scores are reasonably good, but generally worse than the other methods. This can also be seen from Fig. 8 where a comparison of the RMSEs from the four different models are shown. The
relatively large difference between the training and testing errors of the neural network model indicates that the results can be improved by adding more data sets, giving prediction errors closer to the other models.

![Figure 7](image)

Figure 7: (Color online) (a) Convergence of MoNbTaTiW RMSE with respect to the number of epochs. (b)-(f) Comparison of the local energy predicted by the neural network model with the ones from DFT. The blue circles represent the training data and the filled red circles represent the testing data. (b) Mo, (c) Nb, (d) Ta, (e) Ti, and (f) W.

Finally, we would like to point out that, while the neural network model did not show better prediction accuracy for the materials studied, the results are still very encouraging because they demonstrated that with sufficiently large data sets, a direct mapping from the local chemical environment to the atomic local energy is feasible. The versatile DNN models are of great potential to describe materials where non-linear atomic interactions are important. These interactions are difficult to include in the traditional methods, but
Figure 8: (Color online) Comparison of the training (blue) and testing (orange) RMSEs from different models. LR stands for the linear EPI model. EPI\_QR stands for the quadratic EPI model. 1ST represents the first shell triplet model and NN denotes the neural network model.

Straightforward for neural networks since they work in the configuration space and uses nonlinear activation functions.

In conclusion, we developed an approach to obtain large amounts of DFT data sets by employing the $O(N)$ LSMS method to calculate the atomic local energies. Such a method reduces the time scaling of generating a single data point to $O(0)$, which not only substantially speeds up the construction of effective Hamiltonian, but also allows for the use of complex models to describe non-linear atomic interactions. Using the large DFT data sets, a range of refractory HEAs were studied with four different models, among them the quadratic EPI model demonstrated the best overall performance. Neural networks were successfully applied to establish a mapping directly from the local chemical environment to the atomic local energy, which provides a promising tool for the future study of complex alloys.
3. Method

The DFT data are calculated with the locally self-consistent multiple scattering (LSMS) method [37], which is a real space implementation of the Korringa-Kohn-Rostoker (KKR) method [41, 42] and scales linearly with respect to the number of atoms. The angular momentum cutoff in the LSMS method is set as 3. The local interaction zone (LIZ) with 59 atoms is found to be accurate enough and adopted after testing with both KKR method and LSMS method with LIZ = 283. Note that the LIZ in the LSMS only restricts the electron quantum scattering, while the LIR in the effective Hamiltonian is a cutoff on all the atomic interactions. The scalar-relativistic equations are solved to properly treat the heavier elements in the system. For simplicity, the atoms are assumed to be on perfect BCC lattice sites. For MoNbTaW this was found to only affect the low temperature ground states [31].

The MoNbTaW data sets for the linear EPI model are calculated with supercells of 64, 128, and 256 atoms. 10 chemical configurations are generated randomly for each supercell, therefore the total number of training data is 4,480, with 1,120 data for each element. For the 5-element HEAs, the data sets for the linear EPI model are calculated with supercells of 20, 40, 80, and 160 atoms. 10 different configurations are generated randomly for each supercell, so the total number of training data sets is 3,000, with 600 data for each element. The same number of testing data are calculated using configurations different from the training ones.

More data sets are used for the other three models. For MoNbTaW, supercells of 64, 128, 256, and 512 atoms are used, giving a total of 38,400 training data and the same number of testing data. For MoNbTaVW, supercells of 20, 40, 80, 160, and 320 atoms are used, giving a total of 50,800 training data and the same number of training data. For MoNbTaTiW, supercells of 20, 40, 80, and 160, and 320 atoms are used, giving a total of 62,000 training data and 29,840 testing data. For AlMoNbTaW, supercells of 20, 40, 80, and 160 atoms are used, giving a total of 30,000 training data and the same number of testing data.

The Scikit-learn [43] package is used for the linear and quadratic regression models. PyTorch is used for the neural network models, where mini-batch gradient descent, adaptive moment estimation (Adam) [44], and back-propagation algorithms are employed to update the neural network parameters. The mini-batch size is chosen as 16 and the learning rate is set as 0.001. The loss function is simply the energy mean square error (MSE), with an additional L2 regularization penalty of $\alpha = 1 \times 10^{-8}$.

4. Data availability

The data that support the findings of this study are available from the corresponding author upon request.

5. Code availability

The data that support the findings of this study are available from the corresponding author upon request.
6. Acknowledgements

The work of X. L. and M. E. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Science and Engineering Division. J. Z. was supported by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory. This research used resources of the Oak Ridge Leadership Computing Facility, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

7. Contributions

X. L. conceived the project, calculated the data, and wrote the first draft. X. L. and J. Z. analyzed the results. Y. W. and M. E. developed the LSMS code. Y. W. modified the LSMS code to calculate the atomic local energy. All authors contributed to the final manuscript.

8. Competing interests

The authors declare no competing interests.

References

[1] M. Widom, Modeling the structure and thermodynamics of high-entropy alloys, Journal of Materials Research 33 (2018) 2881–2898.

[2] M. Eisenbach, Z. Pei, X. Liu, First-principles study of order-disorder transitions in multicomponent solid-solution alloys, Journal of Physics: Condensed Matter 31 (2019) 273002.

[3] A. Fernández-Caballero, J. S. Wróbel, P. M. Mummery, D. Nguyen-Manh, Short-range order in high entropy alloys: Theoretical formulation and application to Mo-Nb-Ta-V-W system, Journal of Phase Equilibria and Diffusion 38 (2017) 391–403.

[4] S. N. Khan, M. Eisenbach, Density-functional Monte-Carlo simulation of CuZn order-disorder transition, Phys. Rev. B 93 (2016) 024203.

[5] R. Kikuchi, A theory of cooperative phenomena, Phys. Rev. 81 (1951) 988–1003.

[6] J. Sanchez, F. Ducastelle, D. Gratias, Generalized cluster description of multicomponent systems, Physica A: Statistical Mechanics and its Applications 128 (1984) 334 – 350.

[7] R. Feng, P. K. Liaw, M. C. Gao, M. Widom, First-principles prediction of high-entropy-alloy stability, npj Computational Materials 3 (2017) 50.

[8] V. Blum, G. L. W. Hart, M. J. Walorski, A. Zunger, Using genetic algorithms to map first-principles results to model hamiltonians: Application to the generalized ising model for alloys, Phys. Rev. B 72 (2005) 165113.
[9] O. Levy, G. L. W. Hart, S. Curtarolo, Uncovering compounds by synergy of cluster expansion and high-throughput methods, Journal of the American Chemical Society 132 (2010) 4830–4833. PMID: 20218599.

[10] A. V. Ruban, I. A. Abrikosov, Configurational thermodynamics of alloys from first principles: effective cluster interactions, Rep. Prog. Phys. 71 (2008) 046501.

[11] J. W. D. Connolly, A. R. Williams, Density-functional theory applied to phase transformations in transition-metal alloys, Phys. Rev. B 27 (1983) 5169–5172.

[12] C. Jiang, B. P. Uberuaga, Efficient ab initio modeling of random multicomponent alloys, Phys. Rev. Lett. 116 (2016) 105501.

[13] A. Seko, Y. Koyama, I. Tanaka, Cluster expansion method for multicomponent systems based on optimal selection of structures for density-functional theory calculations, Physical Review B 80 (2009) 165122.

[14] J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, S.-Y. Chang, Nanostructured high-entropy alloys with multiple principal elements: Novel alloy design concepts and outcomes, Advanced Engineering Materials 6 (2004) 299–303.

[15] B. Cantor, I. Chang, P. Knight, A. Vincent, Microstructural development in equiatomic multicomponent alloys, Materials Science and Engineering: A 375-377 (2004) 213 – 218.

[16] E. P. George, D. Raabe, R. O. Ritchie, High-entropy alloys, Nature Reviews Materials 4 (2019) 515–534.

[17] B. Gludovatz, A. Hohenwarter, K. V. S. Thurston, H. Bei, Z. Wu, E. P. George, R. O. Ritchie, Exceptional damage-tolerance of a medium-entropy alloy CrCoNi at cryogenic temperatures, Nature Communications 7 (2016) 10602.

[18] B. Gludovatz, A. Hohenwarter, D. Catoor, E. H. Chang, E. P. George, R. O. Ritchie, A fracture-resistant high-entropy alloy for cryogenic applications, Science 345 (2014) 1153–1158.

[19] O. Senkov, G. Wilks, J. Scott, D. Miracle, Mechanical properties of Nb25Mo25Ta25W25 and V20Nb20Mo20Ta20W20 refractory high entropy alloys, Intermetallics 19 (2011) 698 – 706.

[20] D. Miracle, O. Senkov, A critical review of high entropy alloys and related concepts, Acta Materialia 122 (2017) 448 – 511.

[21] B. L. Gyorffy, G. M. Stocks, Concentration waves and fermi surfaces in random metallic alloys, Phys. Rev. Lett. 50 (1983) 374–377.

[22] F. Ducastelle, F. Gautier, Generalized perturbation theory in disordered transitional alloys: Applications to the calculation of ordering energies, Journal of Physics F: Metal Physics 6 (1976) 2039.
[23] A. Gonis, X. G. Zhang, A. J. Freeman, P. Turchi, G. M. Stocks, D. M. Nicholson, Configurational energies and effective cluster interactions in substitutionally disordered binary alloys, Phys. Rev. B 36 (1987) 4630–4646.

[24] A. R. Natarajan, A. Van der Ven, Machine-learning the configurational energy of multicomponent crystalline solids, npj Computational Materials 4 (2018) 56.

[25] Q. Gao, S. Yao, J. Schneider, M. Widom, Machine Learning methods for interatomic potentials: application to boron carbide, arXiv e-prints (2015) arXiv:1512.09110.

[26] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436.

[27] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, K.-R. Müller, Machine learning of accurate energy-conserving molecular force fields, Science Advances 3 (2017).

[28] Z. Li, J. R. Kermode, A. De Vita, Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces, Phys. Rev. Lett. 114 (2015) 096405.

[29] V. L. Deringer, G. Csányi, Machine learning based interatomic potential for amorphous carbon, Phys. Rev. B 95 (2017) 094203.

[30] A. P. Bartók, J. Kermode, N. Bernstein, G. Csányi, Machine learning a general-purpose interatomic potential for silicon, Phys. Rev. X 8 (2018) 041048.

[31] T. Kostiuchenko, F. Körmann, J. Neugebauer, A. Shapeev, Impact of lattice relaxations on phase transitions in a high-entropy alloy studied by machine-learning potentials, npj Computational Materials 5 (2019) 55.

[32] W. Ye, C. Chen, Z. Wang, I.-H. Chu, S. P. Ong, Deep neural networks for accurate predictions of crystal stability, Nature Communications 9 (2018) 3800.

[33] A high-bias, low-variance introduction to machine learning for physicists, Physics Reports 810 (2019) 1 – 124.

[34] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.

[35] S. Goedecker, Linear scaling electronic structure methods, Rev. Mod. Phys. 71 (1999) 1085–1123.

[36] E. Prodan, W. Kohn, Nearsightedness of electronic matter, Proceedings of the National Academy of Sciences 102 (2005) 11635–11638.

[37] Y. Wang, G. M. Stocks, W. A. Shelton, D. M. C. Nicholson, Z. Szotek, W. M. Temmerman, Order-N multiple scattering approach to electronic structure calculations, Phys. Rev. Lett. 75 (1995) 2867–2870.

[38] W. P. Huhn, M. Widom, Prediction of A2 to B2 phase transition in the high-entropy alloy Mo-Nb-Ta-W, JOM 65 (2013) 1772–1779.
[39] J. Zhang, X. Liu, S. Bi, J. Yin, G. Zhang, M. Eisenbach, Robust data-driven approach for predicting the configurational energy of high entropy alloys, Materials & Design 185 (2020) 108247.

[40] M. Leshno, V. Y. Lin, A. Pinkus, S. Schocken, Multilayer feedforward networks with a nonpolynomial activation function can approximate any function, Neural Networks 6 (1993) 861 – 867.

[41] J. Korringa, On the calculation of the energy of a Bloch wave in a metal, Physica 13 (1947) 392 – 400.

[42] W. Kohn, N. Rostoker, Solution of the Schrödinger equation in periodic lattices with an application to metallic lithium, Phys. Rev. 94 (1954) 1111–1120.

[43] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[44] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, arXiv e-prints (2014) arXiv:1412.6980.