RESEARCH ARTICLE | MARCH 12 2024

Machine-learning-assisted modeling of alloy ordering phenomena at the electronic scale through electronegativity

Special Collection: Era of Entropy: Synthesis, Structure, Properties, and Applications of High-Entropy Materials

Dingqi Zhao ; Xi Jin; Junwei Qiao ; Yong Zhang ; Peter K. Liaw

Appl. Phys. Lett. 124, 111902 (2024)
https://doi.org/10.1063/5.0188516

Applied Physics Letters ARTICLE pubs.aip.org/aip/apl

Machine-learning-assisted modeling of alloy ordering phenomena at the electronic scale through electronegativity

Cite as: Appl. Phys. Lett. **124**, 111902 (2024); doi: 10.1063/5.0188516
Submitted: 21 November 2023 · Accepted: 19 February 2024 ·
Published Online: 12 March 2024

Dingqi Zhao,[1] Xi Jin,[1] Junwei Qiao,[1,2,a)] Yong Zhang,[3] and Peter K. Liaw[4]

AFFILIATIONS

[1]College of Materials Science and Engineering, Taiyuan University of Technology, Taiyuan 030024, China
[2]Key Laboratory of Interface Science and Engineering in Advanced Materials, Ministry of Education, Taiyuan University of Technology, Taiyuan 030024, China
[3]State Key Laboratory for Advanced Metals and Materials, University of Science and Technology Beijing, Beijing 100083, China
[4]Department of Materials Science and Engineering, The University of Tennessee, Knoxville, Tennessee 37996-2200, USA

Note: This paper is part of the APL Special Collection on Era of Entropy: Synthesis, Structure, Properties, and Applications of High Entropy Materials.
[a)]**Author to whom correspondence should be addressed:** qiaojunwei@gmail.com

ABSTRACT

Many studies attribute the excellent properties of high-entropy alloys to the ordering-phenomena. It can be known from density functional theory that the macroscopic properties of the system can be described by the electron density. Electronegativity is related to electron density, and models describing ordering can be established based on electronegativity scales through machine learning. In this study, a large dataset was established and predicted the ordered state corresponding to the alloy composition. The accuracy of the model on the test set was 94%. Furthermore, this study used different methods to explain the machine learning model and learned more model information.

In numerous research fields, methods to build complex system models through machine learning (ML) have attracted much attention.[1–5] Modeling and interpreting models through ML provide an interesting perspective on understanding complex systems: not all microscopic details affect the macroscopic properties of the system, and causal emergence occurs at different levels in the complex system.[6–8] High-entropy alloys (HEAs) belong to complex systems that have more principal elements added than traditional alloys.[9] The birth of the design concept of HEAs was accompanied by such a conjecture: If more types of alloying elements are added to the alloy system, the alloy will be a disordered solid solution rather than an intermetallic phase.[9] This conjecture can be viewed as an extension of the cooperation phenomenon[10] in binary alloys (also known as the order–disorder transformation[11,12] and superlattice–sublattice problem[13]). To summarize, the positions of atoms in crystal lattices are not independent, but will affect each other, resulting in ordering phenomena in the alloy.[14–16] The existence of this effect will directly affect the microscopic particle distribution and then affect the macroscopic properties of the alloy.[17–19]

In the initial research, people hoped to establish a phenomenological model to describe the statistical laws between macroscopic variables and ordered phenomena.[20–22] These phenomenological models have promoted the discovery of new HEA compositions.[23–26] Electronegativity has long been used as such a system variable, such as the Hume–Rothery criterion in binary alloys,[20] and there are also related electronegativity criteria in HEAs.[23] The concept of electronegativity was proposed in 1932.[27] It is believed that electronegativity can represent the polarizability of atoms,[28] the distribution of electrons,[29] and the properties of electron orbits.[30] It can be known from the density functional theory[31] that the macroscopic properties of matters can be regarded as a functional function of the electron density. The relevant definition of electronegativity is closely related to the electron density, so ones associate electronegativity with the atomic radius,[32] ionization energy,[33] electron affinity,[33] compound formation enthalpy,[34] and other properties. In the past 90 years, the related concepts of electronegativity have continued to develop. Not only various electronegativity scales corresponding to different experiments have been proposed[35] but the related concepts of electronegativity have also

been constantly updated. It is interesting to develop a model describing ordering in alloys starting from electronegativity. It can help us understand the relationship between the electronegativity and ordering in alloy systems and can give interesting explanations about alloy-ordering phenomena from the electronic level, providing interesting ideas for alloy designing.

The earliest electronegativity scale was proposed by Pauling.[27] It is a set of standards abstracted from thermochemical experimental data to describe the ability of atoms in a molecule to attract electrons to themselves. In 1963, Whitehead, Hinze, and Jaffe proposed the concept of orbital electronegativity,[30] defining electronegativity as the derivative of the energy of a certain valence state, corresponding to the number of electrons occupied in the orbit. They believed that electronegativity is not a fixed property of atoms, but a property of atomic orbitals. In 1994, Ghosh redefined electronegativity based on the density functional theory[36–38] and divided chemical bonds into three regions: two atomic regions and one bond region. The bonding process was explained as follows: the electronegativity before bonding is higher than the average electronegativity of the two atoms, so electrons flow from the atomic region to the bonding region. This explanation can describe the bonding process of diatomic molecules. In 2019, Rahm, Zeng, and Hoffmann made corresponding corrections to the traditional Allen electronegativity[29] through a large number of density functional calculations,[39] considering that electronegativity is the average electron binding energy of ground state atoms. Electronegativity is not a measurable value of an isolated atom, and related electronegativity scales often come from different experimental measurements or abstract methods. In total, 10 electronegativity scales were chosen in this study to establish relevant models describing the ordering in alloy systems. In this study, macroscopic variables describing the system are obtained by calculating the properties of the alloy system corresponding to different electronegativity scales. This process is also called feature engineering. The corresponding features of the first raw moment of the relevant electronegativity scale corresponding to the alloy system are suffixed by the letter a, and the characteristics of the second central moment are suffixed by δ. The first raw moment is the weighted average value of electronegativity, and the second central moment is the weighted standard deviation of electronegativity, which is employed to describe the level of electronegativity deviation from the average value in the alloy system (see the supplementary material for more information about feature engineering and different electronegativity scales).

In fact, ordering always exists in real alloys, and what often needs to be discussed is the degree and form of ordering. In many studies, the ordering phenomenon observed by instruments is often named as intermetallic phase. The ordering method in complex alloy systems is more complicated than that in dilute alloys.[40] The order in the alloy can be compared to the critical opalescence phenomenon. When the order reaches a certain degree or range, it will be identified by the instrument. It seems inappropriate to use the traditional naming method. In this study, we prefer to use different degrees of symmetry to describe alloy systems. The relevant standards used here come from Miracle and Senkov.[9] It is a good solution to describe the order of alloy systems through static concentration waves.[41] This description based on Fourier transform can relate the concentration waves to the symmetry of the alloy system. The alloy system can be divided into two states of high and low symmetry through engineering-precision resolution. An order phase found in alloys means that the alloy system has low symmetry. In fact, the distinction for this task is difficult, for example: in the work of Zhang *et al.*,[42] the accuracy of identifying solid solutions and non-solid solutions was 88.7%. In the work of Huang *et al.*,[43] the accuracy of identifying was 78.9%. In Pei's work,[44] solid solutions were identified with an accuracy of 93%.

Accuracy is not the only indicator for evaluating reality. In fact, in some early works, simple criteria can achieve good classification results,[23–25] which may be due to the small database, leading to significant biases. When the database is large enough, some simple criteria seem to lose their effectiveness. Here, the relevant effects are shown in Fig. 1. First, all data are normalized (projected to the range of 0–1 through the normal distribution according to the characteristic quantile), and the four radar charts in Fig. 1 are drawn. In the radar chart, an alloy composition will correspond to a polygon, and the vertices of the polygon represent the features corresponding to this alloy composition. A random subset in the database was selected to draw the radar diagrams in Figs. 1(a) and 1(b). A total of 200 alloy components were superimposed with the center as the origin. The darker superimposed color means that the same type of alloys has similar properties. Figure 1(a) represents the low symmetry state, and Fig. 1(b) represents the high symmetry state. As can be found from Fig. 1, there are obvious differences between the two-symmetry states. This feature means that it is easy to find specific features in small databases to distinguish macroscopic phenomena. For example, low-symmetry systems tend to have higher *E_n_G_a* features, while high-symmetry systems do not. However, after drawing more data (4000 groups in total) into a radar chart, as presented in Figs. 1(c) and 1(d), the polygons are almost filled, and the contrast between two symmetry states is not distinguishable. It does not work to use simple criteria to make distinctions in big data. Complex systems require more complex models, and ML has demonstrated excellent capabilities in building models that describe complex systems.[45–47]

The ordering phenomena in alloys are governed by the same model,[16,48] and a unified model can be established to describe the ordering in multicomponent HEAs and dilute alloys.[44] By collecting the data of HEAs and traditional alloys together, a dataset with 4000 alloy data components is constructed. In the dataset, alloy components in two symmetry states account for half each. A prediction accuracy of 94% was achieved in this work. (Accuracy is a common classifier evaluation standard, and the corresponding function is provided in the supplementary material.) The algorithm selected in this study is the decision forest[49] in supervised learning. Supervised learning uses a training set to learn a model and then uses the model to predict the test set. In this work, different random seeds are used to control the division of training and test sets. 70% of the data are employed for model learning, and the remaining 30% is used to test the model's predictive ability. This method is also known as simple cross-validation, and K-fold cross-validation with different K values also provides in the supplementary material. Features constitute the feature space, and the macroscopic phenomena of the alloy system correspond to the output space.[50] In supervised learning, it is assumed that there is a joint distribution between the feature space and the output space. The training and test data rely on this joint distribution.[50] The joint distribution can also be understood as the mapping between the feature and the output space. The purpose of ML tools is to find an optimal mapping that can transform the feature space into the output space. If the supervised learning model can achieve high prediction accuracy on never-seen

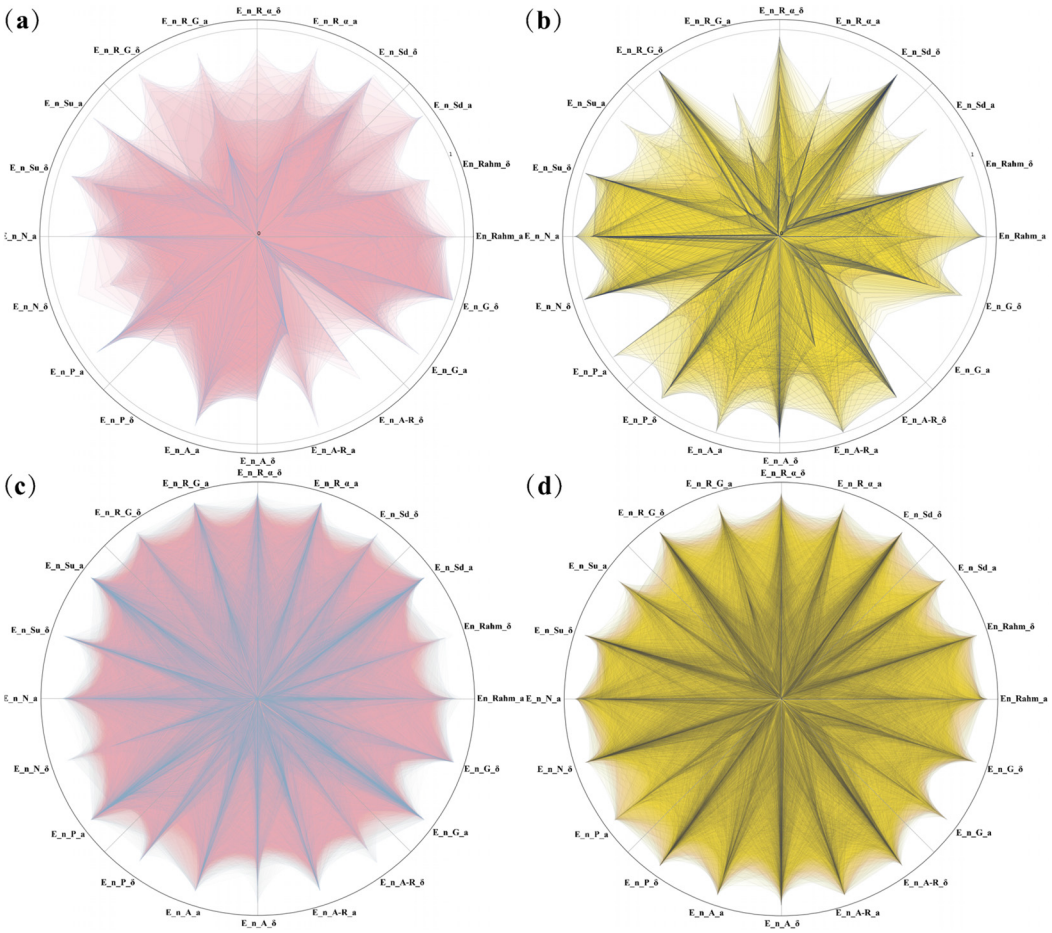

FIG. 1. Alloy feature radar chart, the vertices of the polygon correspond to the features of the alloy composition. (a) 200 random high-symmetry alloys. (b) 200 random low-symmetry alloys. (c) 2000 random high-symmetry alloys. (d) 2000 random low-symmetry alloys.

datasets, it means that the relevant model has learned the relationship between macroscopic variables and ordering phenomena. We tried several common algorithm frameworks to test the effects of different algorithms on this task (more details are provided in the supplementary material), and the results showed that the forest algorithm based on the tree structure showed greater advantages.

The decision tree model is a common classifier, while the decision forest can be regarded as an ensemble algorithm. By combining decision-tree models in a parallel manner, the decision forest becomes a strong learner and has better performance in combating overfitting. The classification-decision-tree model consists of nodes and directed edges. There are two types of nodes, namely, one is internal nodes, and the other is leaf nodes. Internal nodes will contain features, while leaf nodes represent related categories. In this work, the optimal model is found by adjusting the hyperparameters of the decision-forest model. The method of adjusting hyperparameters is the grid method, and the optimal hyperparameter combination is found by traversing limited hyperparameters (see the supplementary material for more details). The result of the model is shown in Fig. 2. The hyperparameter that has the greatest impact on the accuracy of the decision tree model is the depth of the tree. The relationship between the depth of the tree and the accuracy of the model is plotted in Fig. 2. As shown in Fig. 2(a), the forest algorithm achieves the highest accuracy when the depth of the tree in the weak classifier reaches 18. After that, as the depth of the tree increases, the model accuracy almost remains unchanged. Another important parameter is the number of sub-learners. The best hyperparameters was obtained through 77 sets of parameter combinations. When the forest model contains 30 tree models, the accuracy of the model does not increase as the number increase. In the six sets of hyperparameter comparisons about min samples leaf, the results show that when min samples leaf is equal to 2, the model is the most accurate. In the six sets of hyperparameter comparisons about min samples split, when number is equal to 3, the model is the most accurate (relevant results are shown in the supplementary material).

The basic unit of the tree model is a simple if-then judgment, which can be explained as a competition to a certain extent. When a related feature reaches a certain level, it means that the competition corresponding to the feature in the system reaches an overwhelming level. As exhibited in Fig. 2(a), the accuracy description of the relevant

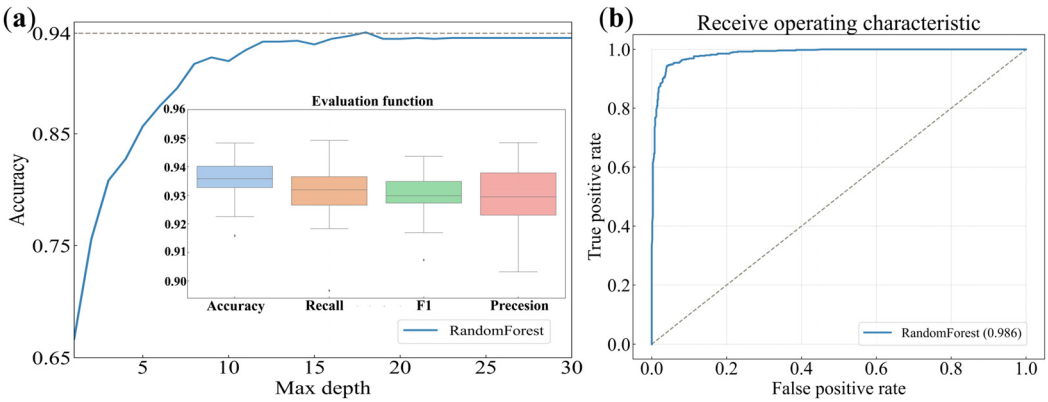

FIG. 2. The effect of the model on the test set. (a) Accuracy increases with tree depth, and the box plot shows the model evaluation effect in different test sets. (b) Receiver operating characteristic curve.

model shows that the model has a high accuracy of up to 94%. Considering that the performance of the model may be affected by database partitioning, 30 different random seeds were selected to average the performance of the model, and the relevant results were represented in a box graph. To comprehensively demonstrate the predictive performance of the model, Fig. 2(a) shows four model evaluation functions.

A receiver-operating characteristic (ROC) is a common computational method used to evaluate the effectiveness of classifiers. The horizontal axis of the ROC curve image represents the false positive rate (FPR), while the vertical axis represents the true positive rate (TPR). The shape and position of the ROC curve can reflect the performance of the classifier. The closer the ROC curve is to the upper left corner, the better the performance of the classification model. In addition to the ROC curve, the area under curve (AUC) is also one of the indicators for evaluating classifier performance. AUC is the area below the ROC curve and is usually considered a comprehensive indicator for evaluating classifier performance. The range of AUC values is between 0 and 1, and the larger the value, the better the performance of the classifier. When AUC is equal to 1, it indicates that the classifier has perfectly classified all samples; when AUC is equal to 0.5, it indicates that the performance of the classifier is equivalent to random guessing, meaning that the classifier has not learned any useful information. As shown in Fig. 2(b), the AUC of the relevant model obtained can reach 0.986.

Another advantage of the tree model is that one can interpret the model by feature information in internal nodes. This study chose three interpretation methods, one regarding the importance of Gini index[49] features method, which we will show in the supplementary material. The permutation feature importance method and the drop column feature importance method are shown in Fig. 3. The results in Fig. 3 come from the average of fivefold cross-validation under the control of 10 random seeds, which means that the data come from 50 sets of tests.

The importance ranking obtained by the permutation feature importance method[51,52] is shown in Fig. 3(a). First, the best feature combination is found to be 20 features through a recursive feature elimination method (see the supplementary material for more details), and the results are used as the baseline. Second, the corresponding features are permuted one by one and the impact of the new features on the accuracy of the model is observed. The abscissa is the improvement effect of the second model. 0.1 represents a 10% increase in model prediction accuracy. Figure 3(b) shows the relevant results obtained by the drop column feature importance method. By removing each feature in turn, the impact of relevant features on the model effect is obtained. It can be seen from Fig. 3(b) that the correlation results obtained by this method fluctuate greatly because the model has been retrained, but one thing is clear: removing any features will degrade the performance of the model, and the correlation results are consistent with those obtained by the recursive feature elimination method. Hence, the results obtained by the drop column feature importance method are not as significant as the permutation feature importance method.

The controversy surrounding what electronegativity is has not yet been resolved and is also beyond the scope of this study. Considering that most studies follow Pauling's explanation method: electronegativity is the force of atoms to attract electrons, and then the second central moment of electronegativity describes the fluctuation of electron density in the system. The top rankings in both ranking methods are second central moment features. This is because the second central moment describes the average deviation of the electron attraction ability of different particles in the system and is more closely related to the electron density. In both interpretation methods, $E\_n\_N\_\delta$ ranks high, which means that the Nagle electronegativity scale contains more relevant information that can describe the ordering phenomenon of the alloy system. Nagle electronegativity is related to the polarizability radius.[28] Nagle assumes that atoms are approximately spherical and believes that electronegativity is closely related to atomic size. The electronegativity of different particles comes from the attraction ability of the atomic nucleus to electrons. Therefore, in the formulation of the corresponding scale, electronegativity is regarded as a function of the valence electrons and the atomic radius. This idea has a long history. For example, when the ratio of polarizability to molar volume is larger than one, it is believed that the element will exhibit metallic properties. This polarizability radius can describe the relative volume of atoms, thereby determining electron delocalization or itinerancy.

In traditional phenomenological models, such as the Miedema model,[34] relevant models are established by calling the abstract information of particles. The ML method can provide another way of

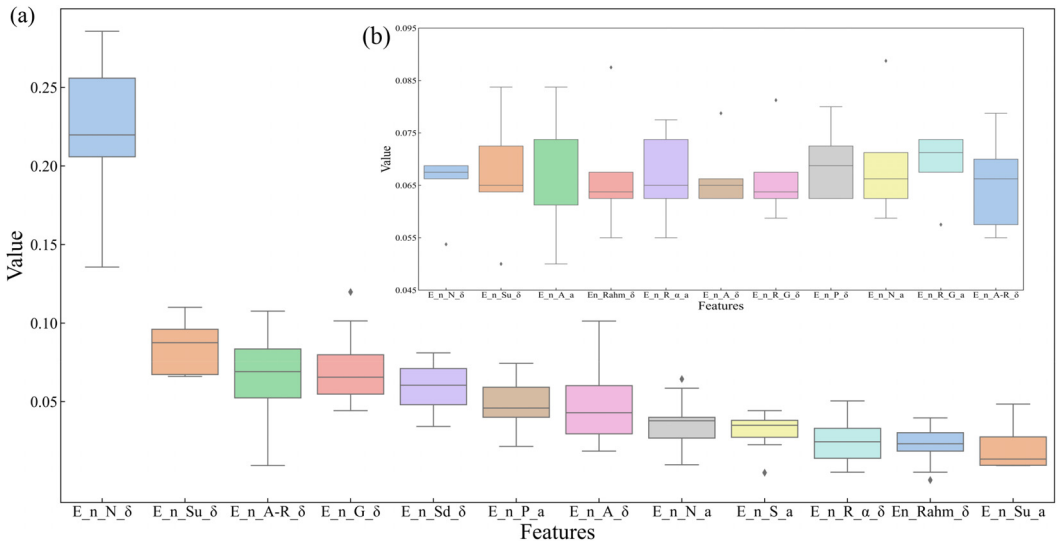

FIG. 3. Rank the importance of features in the model. (a) Permutation feature importance method. (b) Drop column feature importance method.

modeling, first establishing a sufficiently accurate model, and then analyzing the physical background of each feature. The corresponding standard of atomic radius is derived from the abstract description of experimental values. One can establish different phenomenological models based on the hypothetical meaning of "atomic size." It is believed that these phenomenological models can affect the macroscopic properties through elastic distortion or stacking effects. In fact, it is not the artificial meaning that matters, but the relevant information that these scales themselves carry to describe the large-scale phenomena of the system. The process of modeling is to discover the relationship between these scales and macroscopic phenomena or to establish the joint distribution between them.[53] The artificially defined meaning itself should not be an obstacle to the application of related concepts. Moreover, it is important to find the level of causal emergence to describe the macroscopic phenomena of the system.

See the supplementary material for the details on feature engineering, algorithm hyperparameters, and model validation.

The authors would like to acknowledge the financial support of the National Natural Science Foundation of China (No. 52271110). P.K.L. acknowledges the support from (1) the National Science Foundation (Nos. DMR-1611180, 1809640, and 2226508) and (2) the U.S. Army Research Office (Nos. W911NF-13-1-0438 and W911NF-19-2-0049).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Dingqi Zhao: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Writing – original draft (equal); Writing – review & editing (equal). **Xi Jin:** Supervision (equal); Validation (equal); Writing – review & editing (equal). **Junwei Qiao:** Resources (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal). **Yong Zhang:** Supervision (equal); Validation (equal); Writing – review & editing (equal). **Peter K. Liaw:** Validation (equal); Visualization (equal); Writing – review & editing (equal).

DATA AVAILABILITY

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

REFERENCES

[1]P. Raccuglia, K. C. Elbert, P. D. Adler *et al.*, Nature **533**, 73 (2016).
[2]C. C. Fischer, K. J. Tibbetts, D. Morgan, and G. Ceder, Nat. Mater. **5**, 641 (2006).
[3]S. Curtarolo, G. L. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, Nat. Mater. **12**, 191 (2013).
[4]R. D. King, J. Rowland, S. G. Oliver *et al.*, Science **324**, 85 (2009).
[5]H. Li, S. Xu, M. Wang, Z. Chen, F. Ji, K. Cheng, Z. Gao, Z. Ding, and W. Yang, J. Mater. Chem. A **8**, 17987 (2020).
[6]A. Nir, E. Sela, R. Beck, and Y. Bar-Sinai, Proc. Natl. Acad. Sci. U. S. A. **117**, 30234 (2020).
[7]G. L. Hart, T. Mueller, C. Toher, and S. Curtarolo, Nat. Rev. Mater. **6**, 730 (2021).
[8]W. J. Murdoch, C. Singh, K. Kumbier, R. Abbasi-Asl, and B. Yu, Proc. Natl. Acad. Sci. U. S. A. **116**, 22071 (2019).
[9]D. B. Miracle and O. N. Senkov, Acta Mater. **122**, 448 (2017).
[10]G. H. Wannier, Rev. Mod. Phys. **17**, 50 (1945).
[11]D. De Fontaine, *Solid State Physics* (Elsevier, 1994), p. 33.
[12]F. C. Nix and W. Shockley, Rev. Mod. Phys. **10**(1), 1 (1938).
[13]A. T. Santoro and A. Mighell, Acta Crystallogr. Sect. A **29**, 169 (1973).
[14]P. C. Clapp and S. C. Moss, Phys. Rev. **142**, 418 (1966).
[15]P. C. Clapp and S. C. Moss, Phys. Rev. **171**, 754 (1968).
[16]S. G. Brush, Rev. Mod. Phys. **39**, 883 (1967).
[17]R. Zhang, S. Zhao, J. Ding, Y. Chong, T. Jia, C. Ophus, M. Asta, R. O. Ritchie, and A. M. Minor, Nature **581**, 283 (2020).
[18]Z. Lei, X. Liu, Y. Wu *et al.*, Nature **563**, 546 (2018).

[19]Q. Ding, Y. Zhang, X. Chen *et al.*, Nature **574**, 223 (2019).
[20]U. Mizutani, *Hume-Rothery Rules for Structurally Complex Alloy Phases* (CRC Press, London, 2008).
[21]D. G. Pettifor, Mater. Sci. Technol. **8**, 345 (1992).
[22]D. G. Pettifor, Mater. Sci. Technol. **4**, 675 (1988).
[23]M. G. Poletti and L. Battezzati, Acta Mater. **75**, 297 (2014).
[24]Y. Zhang, Y. J. Zhou, J. P. Lin, G. L. Chen, and P. K. Liaw, Adv. Eng. Mater. **10**, 534 (2008).
[25]S. Guo and C. T. Liu, Prog. Nat. Sci.: Mater. Int. **21**, 433 (2011).
[26]M. C. Troparevsky, J. R. Morris, P. R. C. Kent, A. R. Lupini, and G. M. Stocks, Phys. Rev. X **5**, 011041 (2015).
[27]L. Pauling, J. Am. Chem. Soc. **54**, 3570 (2002).
[28]J. K. Nagle, J. Am. Chem. Soc. **112**, 4741 (2002).
[29]L. C. Allen, J. Am. Chem. Soc. **111**, 9003 (2002).
[30]J. Hinze, M. Whitehead, and H. H. Jaffe, J. Am. Chem. Soc. **85**, 148 (1963).
[31]W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).
[32]W. Gordy, Phys. Rev. **69**, 604 (1946).
[33]R. S. Mulliken, J. Chem. Phys. **2**, 782 (1934).
[34]A. R. Miedema, J. Less Common Met. **32**, 117 (1973).
[35]R. C. Artem, I. G. Vladimir, M. Z. Ekaterina, and A. C. Rafael, Russ. Chem. Rev. **67**, 375 (1998).
[36]S. K. Ghosh, Int. J. Quantum Chem. **49**, 239 (1994).
[37]T. K. Ghanty and S. K. Ghosh, J. Phys. Chem. **100**, 17429 (1996).
[38]T. K. Ghanty and S. K. Ghosh, J. Mol. Struct.: THEOCHEM **309**, 143 (1994).
[39]M. Rahm, T. Zeng, and R. Hoffmann, J. Am. Chem. Soc. **141**, 342 (2019).
[40]A. J. Ardell, "Intermetallics as precipitates and dispersoids in high-strength alloys," in *Intermetallic Compounds—Principles and Practice* (Wiley Chichester, Sussex, 1995).
[41]A. G. Khachaturyan, *Theory of Structural Transformations in Solids* (Courier Corporation, 2013).
[42]Y. Zhang, C. Wen, C. Wang, S. Antonov, D. Xue, Y. Bai, and Y. Su, Acta Mater. **185**, 528 (2020).
[43]W. Huang, P. Martin, and H. L. Zhuang, Acta Mater. **169**, 225 (2019).
[44]Z. Pei, J. Yin, J. A. Hawk, D. E. Alman, and M. C. Gao, npj Comput. Mater. **6**, 50 (2020).
[45]J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[46]G. Carleo and M. Troyer, Science **355**, 602 (2017).
[47]D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, Science **334**, 1518 (2011).
[48]K. G. Wilson, Rev. Mod. Phys. **55**, 583 (1983).
[49]L. Breiman, *Classification and Regression Trees* (Routledge, 2017).
[50]S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms* (Cambridge University Press, 2014).
[51]L. Breiman, Mach. Learn. **45**, 5–32 (2001).
[52]A. Fisher, C. Rudin, and F. Dominici, J. Mach. Learn. Res. **20**(177), 1 (2019).
[53]E. P. Hoel, L. Albantakis, and G. Tononi, Proc. Natl. Acad. Sci. U. S. A. **110**, 19790 (2013).