

# ABSTRACT

DAIGLE, SAMUEL EVAN. Interaction of Defects and Disorder in High-Entropy Ceramics: Simulation of Random Solid Solutions from First Principles. (Under the direction of Donald Brenner).

Atom pair interaction models are introduced for the local atomic environment dependence of defect energies in early-transition metal high-entropy carbides and diborides. Separate models are presented for vacancy formation energies, bulk energies, and stacking fault energies fit to density functional theory (DFT) simulations, with an emphasis on the interpretability of model parameters. Defect properties in high-entropy materials are presented relative to the defect energy distribution tails, requiring modeling techniques which are extensible away from the ideal disordered state. For vacancy formation energies in the (Hf,Nb,Ta,Ti,Zr) $B_2$  high-entropy diboride, a simplified physics-informed reduction of the commonly used cluster expansion is introduced for a robust and efficient parameterization of the near-neighbor shells surrounding a vacancy. The predicted total vacancy concentration is found to be higher than the average of the constituent binary compositions with a temperature dependence that deviates from traditional Maxwell-Boltzmann statistics, converging to the mean vacancy concentration at infinite temperature. This relationship is produced by a shift in the relative contribution of each atomic configuration as a function of temperature, highlighting the importance of the lower tail of the vacancy formation energy distribution at temperatures below 2000 K. The parameterization is further shown to give insight into ordering tendencies in the bulk high-entropy diboride, as the atom pairs with a bidirectional or “bonded” interaction in the near-neighbor shells can be directly compared with respect to the average interaction to produce relative pair affinities, with sensitivity analysis performed on the model parameters to obtain uncertainty estimates for all pair affinities. The parameterization of pair energy models is analyzed with respect to the physical interpretation of sparse parameter sets produced by lasso regression, automatic relevance detection, and stepwise parameter subset selection. As a result of coupling between the remaining interaction parameters, rejected parameters can

be retained in the effective interactions such that the parameter values are equal to zero only with respect to an arbitrary reference state. These hidden parameters prevent metrics such as the Bayesian information criterion from capturing the uncertainty in the parameter set and reduce the interpretability of the model parameters by virtue of the shifting reference energy. It is shown that any objective function for parameter subset selection must consider this reference energy in order to make physics-informed decisions on the parameter space. Finally, a proximity weighted rule of mixtures model is introduced for the local environment dependence of stacking fault energies in the high-entropy transition metal carbides. The intrinsic stacking fault is found to be predictable from a rule of mixtures of the constituent binary carbides in the vicinity of the defect, and the unstable stacking fault is predicted by a new application of the Evans-Polanyi-Semenov principle for chemical reaction path energies. Large concentrations of negative stacking fault energies in the stacking fault energy distributions are considered with respect to Shockley partial dislocation separation distances and dislocation pinning at high energy lattice sites. Setting the upper bound of partial dislocation separation as the distance associated with the mean stacking fault energy plus one standard deviation of the distribution, all group IV, V, VI high-entropy transition metal carbides are predicted to exhibit finite partial separations in agreement with experimentally demonstrated phase stability of compositions with valence electron concentration as high as 9.6.

© Copyright 2024 by Samuel Evan Daigle

All Rights Reserved

Interaction of Defects and Disorder in High-Entropy Ceramics:  
Simulation of Random Solid Solutions from First Principles

by  
Samuel Evan Daigle

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Materials Science and Engineering

Raleigh, North Carolina  
2024

APPROVED BY:

---

Ralph Smith

---

Douglas Irving

---

Jacob Jones

---

Donald Brenner  
Chair of Advisory Committee

# DEDICATION

To my family.

## **BIOGRAPHY**

The author was born in Cleveland, OH. He attended Rensselaer Polytechnic Institute where he earned his B.S. in Materials Engineering in 2012. After working for several years as a Materials Engineer, he moved to Raleigh, NC to pursue his Ph.D. at North Carolina State University where he studied the behavior of disordered crystals using first principles methods. During his graduate studies, he was a fellow of the Science and Engineering of the Atomic Scale NSF Traineeship, working to solve interdisciplinary problems at the interface of materials science, statistics, and machine learning.

## ACKNOWLEDGEMENTS

Thank you to all of the colleagues and mentors through my time at NC State who have contributed to the work presented in this dissertation. Specifically, I would like to extend my deepest gratitude to my advisor Dr. Don Brenner for his support and mentorship in fostering my growth as a researcher as well as to the rest of my committee, Dr. Doug Irving, Dr. Ralph Smith, and Dr. Jacob Jones. I would also like to thank Ashleigh Wright and all those involved with the SEAS program for promoting a collaborative environment between disciplines and encouraging new approaches to solving mathematical problems in materials science.

To my group members and office mates, past and present, it has been an honor to work with all of you. To Zsolt Rák and Josh Harris, thank you for your advice and willingness to offer guidance as I was first getting my feet wet in this research area. Leah Granger, Mina Lim, Kelsey Mirrielees, Jonathon Baker, and all the others, I have enjoyed every day working with you. To Skyler Kauffman, Jalen Pryor, and Belicia Castillo, thank you for your contributions and your excitement in tackling new challenges.

Most importantly, thank you to my family and friends who have helped me get to this point. To my early teachers who saw in me an untapped academic potential, and everyone along the way who pushed me in the right direction. To my parents, who have supported my artistic and academic endeavors in equal measure, thank you for everything big and small that made me who I am. To my wife Erin Bennett, thank you for always keeping me afloat and pushing me to bigger and better things.

Lastly, this work would not have been possible without the financial support from the U.S. Office of Naval Research through award No. N00014-15-1-2863, No. N00014-21-1-2515 (SPICES), and No. N00014-23-1-2758, and from the National Science Foundation through grant No. DGE-1633587 (Science and Engineering of the Atomic Structure at North Carolina State University). Additionally, this work benefited from the computing resources provided by North Carolina State University High Performance Computing Services (RRID:SCR.022168).

# TABLE OF CONTENTS

List of Tables . . . . .	vii
List of Figures . . . . .	viii
Chapter 1 Introduction . . . . .	1
Chapter 2 Background and Motivation . . . . .	4
2.1 High-Entropy Alloys: Structure and Properties . . . . .	5
2.1.1 Entropy and the Crystal Lattice . . . . .	6
2.1.2 Lattice Distortion . . . . .	10
2.1.3 Short-Range Order . . . . .	12
2.2 High-Entropy Ceramics . . . . .	14
2.3 Defects . . . . .	18
Chapter 3 Vacancy Formation Energies in High Entropy Crystals: Application to a High-Entropy Diboride . . . . .	22
3.1 Abstract . . . . .	23
3.2 Introduction . . . . .	23
3.3 Methods . . . . .	25
3.3.1 Pair approximation model . . . . .	26
3.4 Results and Discussion . . . . .	29
3.4.1 Vacancy concentration . . . . .	35
3.4.2 Configurational entropy . . . . .	39
3.4.3 Pair affinities . . . . .	39
3.5 Conclusions . . . . .	42
3.6 Appendix: Configurational entropy derivation . . . . .	43
3.7 Acknowledgments . . . . .	44
Chapter 4 Interfacial Defect Properties of High-Entropy Carbides: a New Evans-Polanyi-Semenov Relation . . . . .	45
4.1 Abstract . . . . .	46
4.2 Background and Motivation . . . . .	46
4.3 Calculation Details . . . . .	50
4.4 Results and Discussion . . . . .	51
4.4.1 Stacking Fault Energies . . . . .	52
4.4.2 Evans-Polanyi-Semenov Relations . . . . .	57
4.4.3 Shockley Partial Dislocation Separations . . . . .	63
4.5 Conclusion . . . . .	69
4.6 Acknowledgments . . . . .	70
Chapter 5 Ordering Phenomena in High-Entropy Ceramics . . . . .	71
5.1 Parameter Selection . . . . .	72
5.1.1 Stochastic Stepwise Subset Optimization . . . . .	73

5.2	Bulk Energy Pair Interaction Models . . . . .	76
5.2.1	Pair Model Regressors . . . . .	78
5.2.2	Mean-Field Normalization . . . . .	84
5.2.3	Ordering and Phase Stability . . . . .	88
<b>Chapter 6 Conclusions . . . . .</b>		<b>92</b>
<b>References . . . . .</b>		<b>95</b>

## LIST OF TABLES

Table 3.1	Root mean squared error (in meV) from least squares fitting as well as cross-validation errors from LOOCV and 5x 4-fold CV for unconstrained and RPA models. SRO is the prediction error for the set of pseudo-ordered simulations. Sets $B$ and $R$ denote the indices of shells modeled as bonded and relaxation interactions, respectively. . . . .	33
Table 3.2	Atom pair affinities (meV/bond) and standard deviations ( $\pm 2\sigma$ ) for cation near neighbors in the first four shells of the RPA model with $B = f1, 2, 3, 4g$ and $R = f5, 6, 7g$ . . . . .	40
Table 4.1	High-entropy carbide compositions studied for each calculation. $h11\bar{2}i f111g$ slip path calculations consisted of 12 simulations at evenly spaced points along the gamma surface. USF calculations were performed at the predicted peak location as presented in Section 4.4.1. Compositions have been experimentally validated as single-phase [1, 2] except as marked for non-validated ( $y$ ) and multi-phase ( $z$ ) [2] compositions. . . . .	49
Table 4.2	Binary carbide stacking fault energies $\gamma^{isf}$ in J/m <sup>2</sup> for the single metal rocksalt carbides used in Eq. 4.2. . . . .	52
Table 4.3	Root mean squared fitting errors for the weighted RoM stacking fault models $\gamma_{\omega}^{isf}$ and $\gamma_c^{isf}$ (J/m <sup>2</sup> ). . . . .	56
Table 4.4	Linear coefficients $\alpha$ (slope) and $\beta$ (intercept) for the first three moments of the predicted ISF distribution along with fitting error (J/m <sup>2</sup> ) for 94 HETMC compositions in Fig. 4.9. . . . .	68
Table 5.1	Feature selection for the RPA vacancy pair model for (Hf,Nb,Ta,Ti,Zr)B <sub>2</sub> .	76
Table 5.2	Supercell configurations for the bulk energy pair model training set. . . . .	77
Table 5.3	Fitting and cross validation scores for the tested linear regression methods.	82
Table 5.4	5-fold cross validation results for lasso regression and automatic relevance determination. . . . .	83
Table 5.5	Pair interaction energies calculated from the zero mean-field convergence of OLS regression. . . . .	88
Table 5.6	(Hf,Mo,Ti,W,Zr)C pair model fitting and cross validation scores. . . . .	89
Table 5.7	Pair interaction energies calculated from the zero mean-field convergence of OLS regression. . . . .	90

## LIST OF FIGURES

Figure 2.1	Schematic representation of ordered and disordered structures of an arbitrary binary composition. . . . .	6
Figure 2.2	Entropy of an $N$ -component system. . . . .	8
Figure 2.3	Entropy contribution of adding an $N^{\text{th}}$ component to an equiatomic composition. . . . .	9
Figure 2.4	Schematic representation of the substitutional size effect on the disordered lattice. . . . .	10
Figure 2.5	Schematic representation of dislocation pinning as Schockley partial dislocations are arrested at lattice sites with large Peierls friction barriers. . . . .	11
Figure 2.6	Publication history of high-entropy alloys and ceramics. . . . .	14
Figure 2.7	Schematic representation of the selective oxidation observed experimentally in (Hf,Nb,Ta,Ti,Zr)C. . . . .	15
Figure 2.8	Gaussian process prediction (a) and uncertainty (b) of formation enthalpy for (Hf,Nb,Ta,Ti,Zr)C as a ternary system with (Hf,Zr)C, (Nb,Ta)C, and TiC end members. Cross-section for a fixed (Hf,Zr)C composition (c) is provided to illustrate the shape of the Gaussian process function. . . . .	17
Figure 2.9	Predicted solubility limit of TiC in the high-entropy (Hf,Nb,Ta,Ti,Zr)C after dealloying, represented as a ternary system of (Hf,Zr)C, (Nb,Ta)C, and TiC. . . . .	18
Figure 2.10	Comparison of radial distribution functions for a high-entropy oxide, carbide, and nitride separated into metal-metal and anion-anion distances. . . . .	19
Figure 2.11	Comparison of radial distribution functions for two high-entropy carbides (Hf,Nb,Ta,Ti,Zr)C and (Hf,Mo,Ta,W,Zr)C . . . . .	21
Figure 3.1	Selection frequency of each parameter set in the top 50 models by BIC. . . . .	31
Figure 3.2	(a) Convergence of the reduced pair approximation model ( $B = \overline{f}1, 2, 3, 4g$ , $R = \overline{f}5, 6, 7g$ , and $U = \overline{f}$ ) and the unconstrained model ( $U = \overline{f}1, \dots, 7g$ ) as a function of training sample size for 5x k-fold cross validation up to and including LOOCV; and (b) cross-validation error as a function of largest shell $K$ in the model for 4-fold and LOOCV (training sample size of 169 and 224 respectively). Cross-validation errors for the 50 top models by BIC shown in grey. . . . .	32
Figure 3.3	A sample construction of the SRO structure in which a near-neighbor shell (red) is replaced with a single element. . . . .	34
Figure 3.4	DFT simulated vacancy formation energies ( ) plotted against LOOCV predictions from the RPA model with $B = \overline{f}1, 2, 3, 4g$ and $R = \overline{f}5, 6, 7g$ . SRO predictions (+) included to illustrate an extrapolation of the model. SRO vacancies are not included in any model training sets. . . . .	35

Figure 3.5	Vacancy formation energies of all simulations and the predicted total distribution using the RPA model. Here, the parameter set of $B = \bar{f}1, 2, 3g$ is used for illustrative purposes. The shaded region represents the vacancy configurations which account for 99% of the total concentration at 2000 K. Vertical lines indicate the binary diboride vacancy formation energies of each constituent element. . . . .	36
Figure 3.6	Cation vacancy concentration of the high-entropy diboride structure as a function of inverse temperature, calculated by individual site contributions as compared to the vacancy concentration of the uniform mean formation energy and the concentration in each of the binary transition metal diborides. . . . .	38
Figure 4.1	DFT calculated ISF energies as a function of RoM values using layer-dependent weighting for five compositions . . . . .	53
Figure 4.2	Calculated gamma surface energies along the $\frac{1}{6}h11\bar{2}/\bar{f}111g$ slip direction for four atomic configurations of (a) (Hf,Nb,Ta,Ti,Zr)C and (b) (Mo,Nb,Ta,V,W)C. ISF and midpoint are indicated by the dashed vertical lines, while (c) shows the relationship between ISF and the interpolated peak position and a fitted third order polynomial for initial N-R estimate. . . . .	55
Figure 4.3	USF energy plotted as a function of the ISF energy at (a) the midpoint along the slip path, (b) and (c) at the maximum along the path as estimated by N-R. (a) and (b) are broken down by composition. A:(Hf,Nb,Ta,Ti,Zr)C, B:(Mo,Nb,Ta,V,W)C, C: (Hf,Nb,Ta,Ti,V)C, D:(Nb,Ta,Ti,V,W)C. Lines are linear fits to the data. . . . .	58
Figure 4.4	USF peak shift from Newton-Raphson search along $h11\bar{2}/\bar{f}111g$ gamma surface. Fitting function obtained from Eq. 4.5 using data from the four compositions specified in Table 4.1. . . . .	60
Figure 4.5	Shifted energy curve model using Eq. 4.6 plotted along the $\frac{1}{6}h11\bar{2}/\bar{f}111g$ slip direction for (top) positive, (middle) zero, and (bottom) negative ISF. . . . .	62
Figure 4.6	Maximum USF energies (blue circles) and deviations from the center point along the $h11\bar{2}/\bar{f}111g$ gamma surface where these energies are found (red squares) for data plotted in Fig. 4.3. The solid lines are two parameter fits for an energy curve overlap model using Eq. 4.6. . . . .	63
Figure 4.7	Depiction of two Shockley partial dislocations separated by a stacking fault along the horizontal plane. Brightly colored atoms at the edge dislocations correspond to the local atomic environment that contributes to the Peierls stress, with blue regions representing negative stacking fault regions and orange peaks on the energy landscape representing the high energy pinning sites preventing partial dislocation separation. . . . .	65
Figure 4.8	ISF distributions for three compositions. The dashed vertical lines represent the mean and the mean plus one standard deviation for each distribution. . . . .	66

Figure 4.9	For 94 HETMC compositions, (a) mean, standard deviation, and skew of the predicted ISF energy distributions as a function of the same quantities calculated from the five corresponding single-metal carbide constituents and (b) the sum of first two moments as a function of VEC.	67
Figure 5.1	Schematic of the stochastic stepwise search algorithm for optimized parameter subset selection.	75
Figure 5.2	Comparison of fitting errors for grouped and individually optimized vacancy-neighbor pair interaction models.	76
Figure 5.3	Tuning of the hyperparameter $\lambda$ for lasso regression using the Akaike and Bayesian information criteria.	79
Figure 5.4	(Hf,Nb,Ta,Ti,Zr)C parameter subset comparison between models fit with ordinary least squares, Bayesian ridge regression, lasso, and automatic relevance detection.	81
Figure 5.5	Number of parameters included in the BIC optimized lasso regression for each training split of leave-one-out cross validation.	83
Figure 5.6	(Hf,Nb,Ta,Ti,Zr)C pair interactions in the first two neighbor shells for leave-one-out cross validation splits of the BIC optimized lasso regression with 25 parameters.	85
Figure 5.7	(Hf,Nb,Ta,Ti,Zr)C pair model parameter subset comparison between zero mean-field models fit with ordinary least squares, Bayesian ridge regression, lasso, and automatic relevance detection.	87
Figure 5.8	(Hf,Mo,Ti,W,Zr)C pair interactions in the first two neighbor shells for zero mean-field adjusted linear regressions.	89
Figure 5.9	Comparison of (Hf,Nb,Ta,Ta,Zr)C and (Hf,Mo,Ti,W,Zr)C bulk pair interaction distributions using parameters from zero mean-field adjusted ordinary least squares regression.	90

# CHAPTER

## 1

# INTRODUCTION

The first high-entropy ceramics were presented by Rost et al. in 2015 [3]. By leveraging the disorder phenomena of multi-principal component materials in combination with the thermomechanical properties of ceramics, a new class of materials was proposed that could open the door to a host of new applications. In theory, one could now introduce disorder onto any crystalline lattice or sublattice through careful selection of alloying elements such that the new composition would retain the base crystal structure. This expanded the field of high-entropy materials to encapsulate the diverse range of mechanical, thermal, and functional properties of ceramics, with new opportunities to tune these properties across unexplored phase spaces.

Since the publication of the first high-entropy oxide, many high-entropy ceramics have been developed on the premise that the disorder inherent to a solid solution can be isolated

to the cation sublattice of an ionic material. Examples of high-entropy ceramics in literature include oxides, carbides, diborides, nitrides, carbo-nitrides, silicides, perovskites, fluorite oxides and other related materials [4–18]. In all cases, the disorder on a lattice site introduces lattice distortions and electronic effects which can combine to create unique properties.

In general, the properties of a high-entropy material can be described in one of two ways. In the first case, a material property will reflect a rule of mixtures such that the high-entropy composition behaves as an average of its constituents. This behavior is typical for properties such as lattice constant, density, and bulk modulus [19–21]. In the second (and perhaps more interesting) case, the combination of constituent elements in a disordered lattice will result in some wholly new property. In the first reported metallic high-entropy alloys (HEAs), this was the case for the simultaneous increase in fracture toughness and strength, as well as a newfound high-temperature mechanical stability [22]. For some high-entropy ceramics, the disordered lattice has resulted in exceptionally low electronic and thermal conductivities [19] as well as hardness values exceeding the rule of mixtures for certain diboride compositions [23].

Alongside the exciting possibilities of this new class of materials, there are challenges that must be overcome. From a simulation perspective, many-component materials carry a large computational burden as large supercells must approximate the disorder in a system while avoiding tiling artifacts across periodic boundaries. Further, characterizing the degree of disorder and modeling the configuration dependence of materials properties across many-dimensional phase space is not a trivial issue. Approaches to overcoming the dimensionality problem in high-entropy material simulation include the special quasi-random supercell method [24], which can struggle to capture local effects of a high-entropy lattice diverging from ideal mixing, and high-throughput small cell modeling [25], which introduce unavoidable short-range order artifacts and may not be suitable for modeling defects with long-range interactions. Understanding the physical implications of any assumptions that are made in the simulation of a high-entropy material is critical for developing modeling techniques that capture the effects of disorder on the crystal lattice.

Research into high-entropy materials has sought to find new combinations of properties which might enable new technologies. Some of the central questions guiding the research in this work relate to the existence of new candidate compositions for ultra-high temperature ceramics (UHTCs) and ultra-hard materials, specifically within the class of early-transition metal carbides and diborides. To better understand the properties and potential limitations of these materials, this work seeks to characterize defects and disorder in high-entropy lattices through the use of first principles atomic simulation and statistical modeling of the relationships in the local atomic environment.

This work will address the following motivating questions as they relate to high-entropy materials:

1. How do defects behave on a disordered lattice?
2. What are the effects of defect energy distributions on material properties?
3. How can simplified models be used to efficiently capture the configuration dependence of defect properties?

## CHAPTER

# 2

# BACKGROUND AND MOTIVATION

Since their introduction, HEAs have been found to have remarkable combinations of mechanical properties. As a nascent field relative to traditional metallurgical alloying, the rapid discovery of novel properties has driven immense interest in this area of research. The equimolar CrMnFeCoNi “Cantor alloy” [26], one of the seminal HEA compositions, exhibits uniquely high fracture toughness while maintaining a yield strength equivalent to some of the strongest steel alloys [22]. More recently, refractory HEAs containing early transition metals such as the equiatomic VNbMoTaW HEA have demonstrated high-temperature mechanical stability comparable to or better than commercial Ni superalloys [27]. These mechanical properties are largely driven by disorder and distortions on the crystal lattice which can lead to changes in deformation mechanisms and kinetics. By superimposing these disordered crystal structures onto the cation sublattice of refractory ceramics, it was proposed that

the existing high melting point and high hardness of UHTCs such as HfC could be improved with the addition of entropy. Guided by this idea, the group IVB, VB, and VIB early-transition metal carbides and diborides were chosen as candidate materials to explore the effects of the high-entropy lattice in ceramics.

This chapter will discuss the motivating literature for the investigation of high-entropy ceramics, including a brief history of HEAs, an overview of the early-transition metal carbides and diborides, and the lattice defects studied in this work.

## 2.1 High-Entropy Alloys: Structure and Properties

The first high-entropy materials, reported separately by Yeh and Cantor in 2004 [26, 28], represented a significant departure from the traditional alloying techniques that had been used and refined over thousands of years dating back to the bronze age. Instead of adding small amounts of an alloying element to a single principal metal, the approach was introduced wherein many alloying elements existed in equal proportion such that none (or all) of them could be considered the principal element. These materials were called equiatomic multicomponent alloys [26], multi-principle component materials, or high-entropy alloys [28].

After observing the combination of strength and toughness in the Cantor alloy, studies by Otto [29] and Gludovatz [30] explored the temperature dependent mechanical behavior. Otto showed high-temperature yield and tensile strengths approaching those of commercial Ni superalloys, while Gludovatz demonstrated an increase in both strength and ductility at low temperatures, continuing down to cryogenic temperatures. At 77 K, the yield strength and ultimate tensile strength of CrMnFeCoNi were found to increase by 85% and 70% respectively as compared to the room temperature values, from approximately 400 and 750 MPa to over 750 and 1250 MPa respectively. Further, the strain at failure increased from approximately 0.5 to 0.7. Expanding to other alloys, Senkov et al. [27] reported that medium-entropy and high-entropy refractory alloys NbMoTaW and VNbMoTaW retained yield stresses above

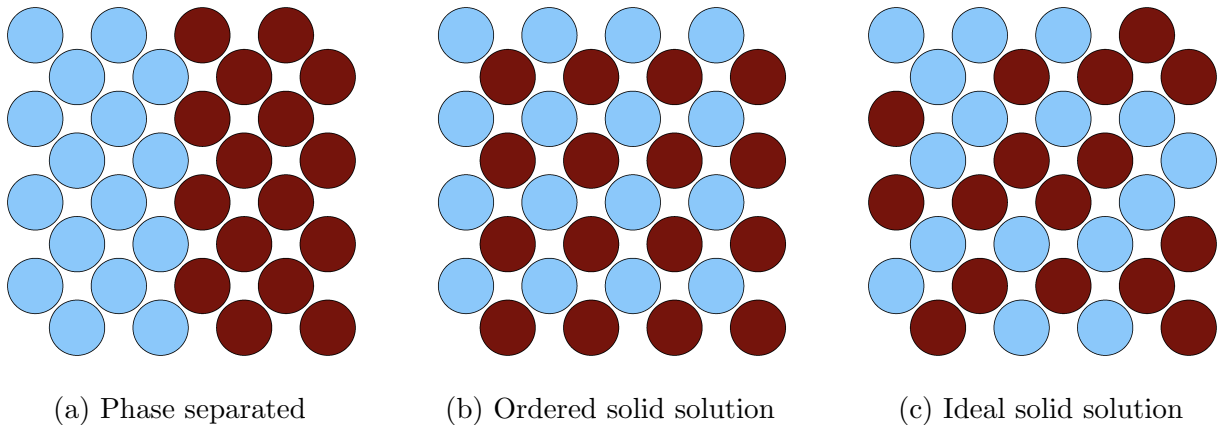


Figure 2.1: Schematic representation of ordered and disordered structures of an arbitrary binary composition.

500 and 800 MPa respectively at 1000 °C, and exhibited a gradual decrease beyond that up to 1600 °C where both alloys remained above 400 MPa. These properties have brought structural HEAs to the forefront of materials research, with potential applications as new state-of-the-art aerospace alloys [31].

The early HEA discoveries have provided a motivation to explore the centers of four-, five-, and six-dimensional phase diagrams in search of novel combinations of material properties, with early compositions including as many as nine metallic components [28]. The early successes of the field have driven widespread interest in the effect of entropy (and the resulting disorder) on phase stability, solubility limits, and properties across a wide range of materials.

### 2.1.1 Entropy and the Crystal Lattice

Many of the most studied high-entropy materials exist as a single-phase random solid solution on a crystal lattice [32]. As opposed to ordered phases where certain elements might occupy their own lattice site, the elements see each lattice site as equivalent, as illustrated in Fig. 2.1. Phase stability and lattice disorder in crystalline materials are dictated by the thermodynamic quantities of the atomic interactions between constituent elements.

It has long been understood that there are two thermodynamic driving forces for phase

stability. The Gibbs free energy is presented as

$$G = H - TS. \quad (2.1)$$

That is, the free energy  $G$  of a system is equal to the enthalpy of formation  $H$  minus the temperature-dependent contribution from the entropy  $S$ , with systems preferring the lowest free energy state available.

In conventional alloy design, the change in enthalpy is the primary consideration. With small entropies, the solubility of one element into the crystal lattice of another is dictated by the availability of a lower enthalpic state. Even in well-known random solid solutions such as AuCu, the solubility exists primarily due to a favorable enthalpy of mixing, while the order-disorder phenomena within the lattice are driven by entropic effects [33].

HEAs were developed, in part, on the premise of increasing the entropic contribution to the largest extent possible, maximizing entropy as a driving force for phase stability [28]. This can be accompanied by a negative enthalpy of formation  $\Delta H$ , or, in the case of an entropy-stabilized material, the entropic contribution to the free energy can overcome a positive  $\Delta H$ . Given that the entropic contribution is temperature dependent, entropy-stabilized phases are metastable, exhibiting a reversible phase transformation with sufficient anneal times at temperatures below

$$T = \frac{\Delta H}{\Delta S}. \quad (2.2)$$

From a statistical thermodynamics viewpoint, the configurational entropy of a system can be derived from the available microstates as

$$S = -k_B \sum_i^N x_i \ln x_i. \quad (2.3)$$

Here,  $x_i$  represent the atomic fraction of each atom type  $i$  and  $N$  is the number of atomic

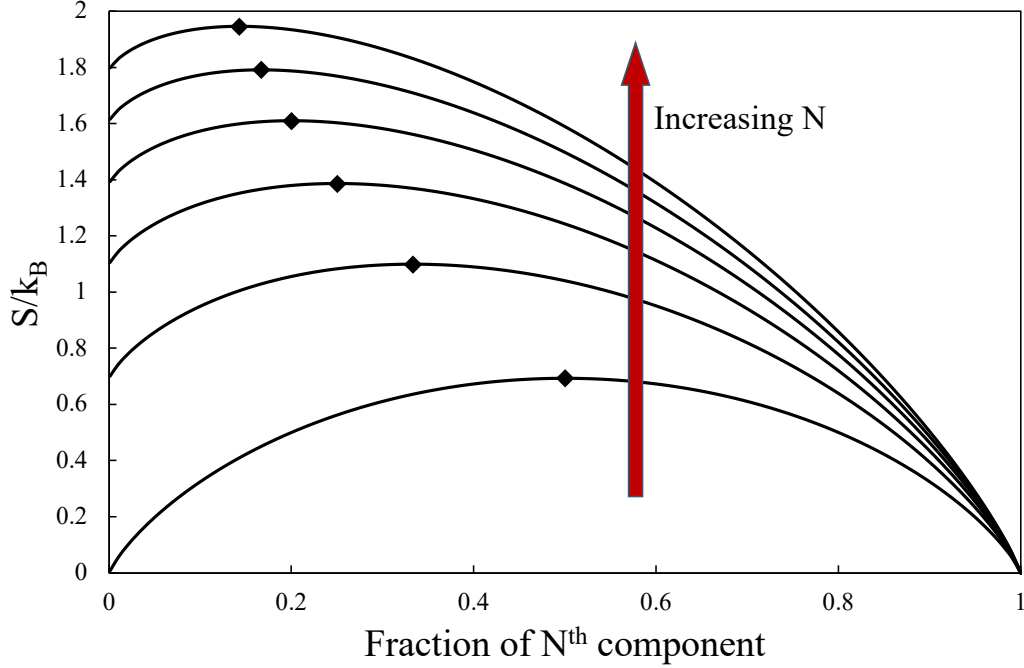


Figure 2.2: Entropy of an  $N$ -component system with  $(N - 1)$  equiatomic species, varying the concentration of the  $N^{\text{th}}$  component. Black diamonds represent the equiatomic  $N$  component system.

species in the system. This equation always finds a maximum at the equimolar composition

$$x_i = \frac{1}{N} \quad (2.4)$$

as demonstrated by Rost et al. [3] by reducing the atomic fraction of any single component such that the total entropy is given as

$$S = k_B \left[ x \ln x + (N - 1) \frac{1-x}{N-1} \ln \frac{1-x}{N-1} \right]$$

$$S = k_B \left[ x \ln x + (1-x) \ln \frac{1-x}{N-1} \right]. \quad (2.5)$$

The composition dependence for the  $N^{\text{th}}$  component is illustrated in Fig. 2.2, where the equimolar composition for each  $N$  is marked by a black diamond.

The configurational entropy for an ideal disordered multi-component system always

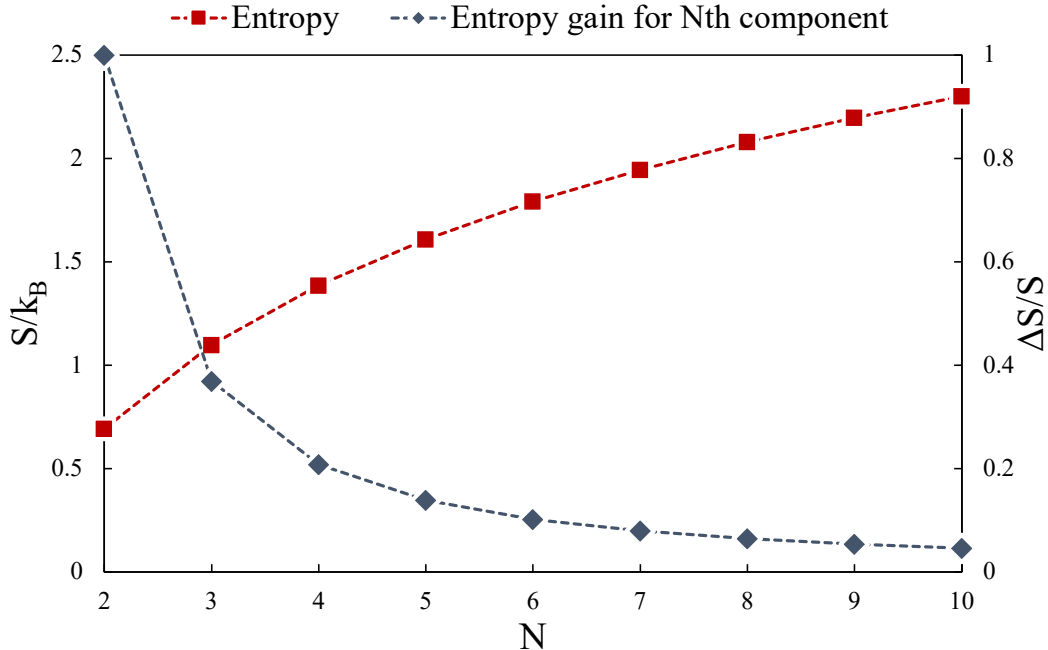
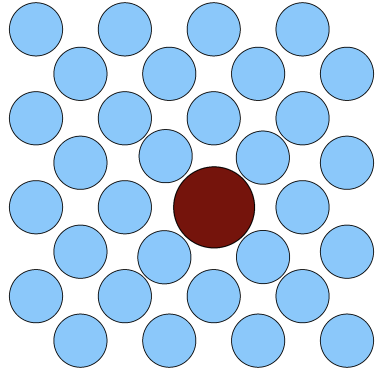


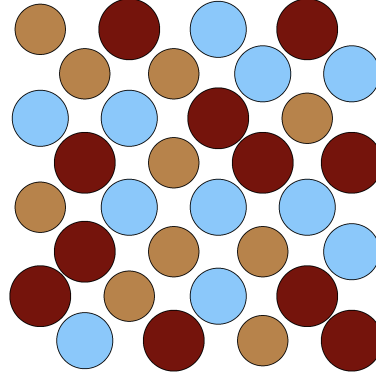
Figure 2.3: Entropy contribution of adding an  $N^{\text{th}}$  component to an equiatomic composition.

increases with the addition of an  $N + 1$  species. This increase, however, lessens with increasing  $N$ , resulting in marginal entropy gains as the number of species in a composition grows larger. Viewed in terms of the balance between entropy and enthalpy, the entropy gain beyond a certain  $N$  may no longer be able to overcome a positive enthalpy of formation for the  $N + 1$  species. When considering the extent to which the entropy gain can be exploited by increasing  $N$ , we must consider the decay in entropy contribution from each additional component, plotted in Fig. 2.3 as entropy  $S$  and the fractional entropy contribution  $\Delta S/S$  for the  $N^{\text{th}}$  component as a function of  $N$ .

In practice, the enthalpic and entropic contributions to the total free energy can be difficult to isolate, in particular with respect to the influence of entropy on phase stability. Take, for example, the Cantor alloy, which was one of the first reported single-phase solid solution HEAs in 2004 [26] and quickly became ubiquitous in HEA research. It took until 2016—more than a decade later—that it was discovered to be metastable at moderate temperatures, when Otto et al. demonstrated a reversible decomposition forming bcc-Cr, L1<sub>0</sub>-NiMn, and B2-FeCo



(a) Dilute substitutional solid solution



(b) Ternary equimolar solid solution

Figure 2.4: Schematic representation of the substitutional size effect on the disordered lattice.

at temperatures below about 800 °C [34]. Due to the uncertain nature of phase stability and ordering in many-component systems, there is a degree of controversy over the nomenclature used to discuss them, with some researchers preferring terms such as multi-principal element alloys or compositionally complex alloys to avoid any implication of the role of entropy in phase stability.

In this work, we define a high-entropy phase as any single-phase crystalline material with at least five disordered components sharing a lattice such that there exists a large configurational entropy. As subset of high-entropy materials, entropy-stabilized materials are defined as metastable materials which exhibit a reversible phase transformation to a single phase driven by entropy contributions at high temperature.

### 2.1.2 Lattice Distortion

Many of the unique properties of HEAs have been attributed to the large distortions of the disordered crystal lattice. Because the lattice constant for HEAs typically follows closely with a rule of mixtures, the size mismatch between components of an HEA, shown schematically in Fig. 2.4, forces the individual atoms to be displaced from their ideal lattice sites. The effects of this distortion can be thought of in terms of substitutional solid solution strengthening,

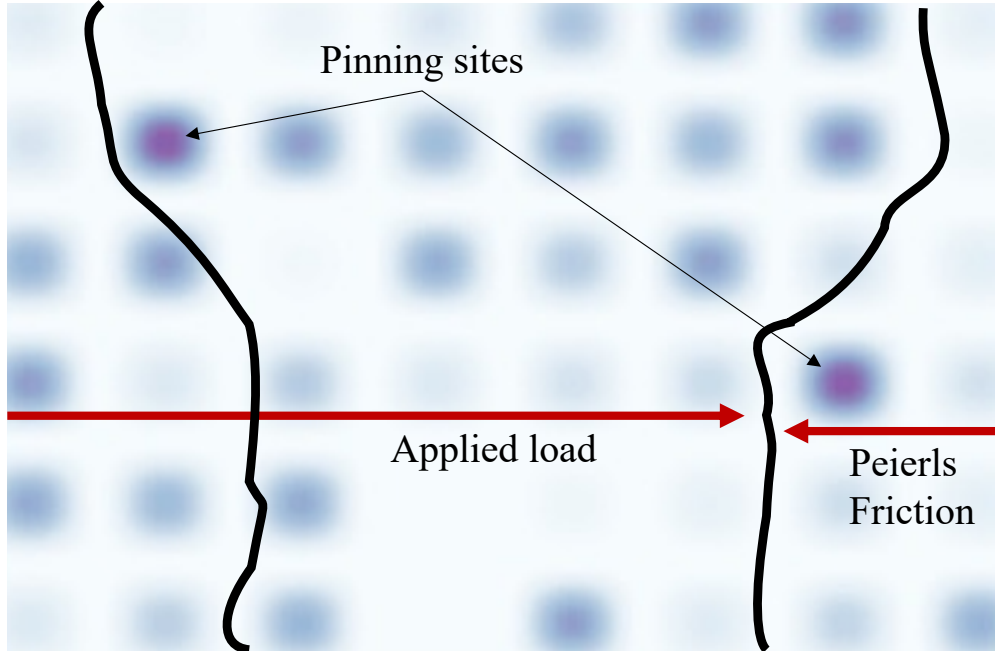


Figure 2.5: Schematic representation of dislocation pinning as Schockley partial dislocations (black lines) are arrested at lattice sites with large Peierls friction barriers. Darker lattice sites correspond to larger energy barriers for dislocation motion.

wherein a solute atom with large size mismatch creates a localized stress field in the crystal lattice. These stress fields interact with those created by dislocations and prevent slip from occurring in the vicinity of the substitution site. In the case of HEAs, each lattice site sees a different local environment, with neighbors shifted from their ideal positions. The varied stress field in the lattice interferes with dislocation motion and provides a strengthening mechanism that can be modelled as a sum of the interaction energies of a dislocation with the lattice [35]. This interpretation is supported experimentally by results from Okamoto et al. [36] that measured the average atomic displacements in the Cantor alloy lattice by X-ray diffraction as a function of temperature, and found the fluctuations in distortion due to thermal vibrations produced the expected increases in solid solution strengthening.

An interesting aspect of dislocation motion in HEAs is the so called “jerky dislocation motion” phenomenon observed by Utt et al. [37] in single crystal CoCrFeMnNi via in situ transition electron microscopy (TEM) under uniaxial load. The TEM images show a bowed

Shockley partial dislocation moving in the crystal through a kink-pair mechanism with adjacent portions of the dislocation gliding separately as they overcome the energy barrier for dislocation motion. The results are recreated through molecular dynamics simulation and interpreted in terms of a varied Peierls friction across the landscape of the distorted lattice. This force balance is illustrated in Fig. 2.5 for an arbitrary cubic lattice.

Given the difficulty of accurately measuring the local lattice strain experimentally, the exact mechanism of strengthening in HEAs is not settled [38]. DFT simulations can help provide insights into the local stresses, distortions, and energy barriers, but cannot easily capture the time- and length-scales necessary for deformation phenomena.

### 2.1.3 Short-Range Order

When studying HEAs, it is often convenient to assume that the lattice is an ideally mixed solid solution. This assumption is not always true, however, and represents another barrier to describing both entropic and enthalpic contributions to the free energy. Any given configuration of atoms on a high-entropy lattice will have a distinct enthalpy of mixing. As such, it is more accurate to describe the enthalpy as a distribution with a composition-dependent distribution width from the difference between high- and low-energy configurations.

Within a given enthalpy distribution, there can exist certain configurations (atomic pairs, triplets, chains, etc.) with lower energies that are preferred at short length-scales without disrupting the long-range disorder of the crystal lattice. One can define this short-range order (SRO) by assessing the near-neighbor environment of each element in a composition with respect to the relative concentration of each other component. The Warren-Cowley SRO parameter  $\alpha_{ij}^r$  is given as

$$\alpha_{ij}^r = 1 - \frac{P_{ij}^r}{2c_i c_j}, \quad (2.6)$$

for an  $ij$  pair of atoms at distance  $r$ , where  $P_{ij}^r$  is the probability of finding an  $ij$  pair, and  $c_i$  is the concentration of  $i$  in the composition [39, 40].

The presence of SRO can have a significant impact on the mechanical and thermal properties of a material. Yin et al. have analyzed the effects of SRO on dislocation mobility in MoNbTaW, finding that ordering increases the mobility of edge dislocations while reducing the mobility of screw dislocations [41], while Antillon et al. demonstrated SRO strengthening due to an increase in the critical resolved shear stress for edge and screw dislocations at lower annealing temperatures in a CoFeNiTi alloy [42].

A study by Chen et al. found that a large degree of SRO could lead to a “pseudo-composite” microstructure with different sets of ordered clusters behaving as soft and hard filler particles to increase both the strength and ductility of a simulated CoCuFeNiPd alloy [43], replicating experimental results for high strength and ductility in CoCuFeNiPd with clustering at short and medium length scales [44]. Separate work by Chen et al. identified a melting temperature dependence in AlCoFeNiCu with varying Cu content, attributing a decrease in melting temperature to increased SRO and chemical segregation driven by larger concentrations of Cu [45]. In both of these results, however, the ordering was present at length scales longer than what is typically considered to be SRO. In these cases it can be difficult to differentiate between SRO and chemical segregation, with the two often coexisting across temperature and composition ranges.

SRO in many-component materials can be difficult to quantify, both experimentally and by simulation, and remains an open topic with respect to nominally disordered phases [38]. Molecular dynamics and Monte Carlo simulations provide avenues for understanding the ordering tendencies of different compositions as well as possible effects of ordering on mechanical properties, but are dependent on well-trained interatomic potentials. The choice of interatomic potential is particularly important when secondary phases may be present, adding significant complexity in the case of HEAs.

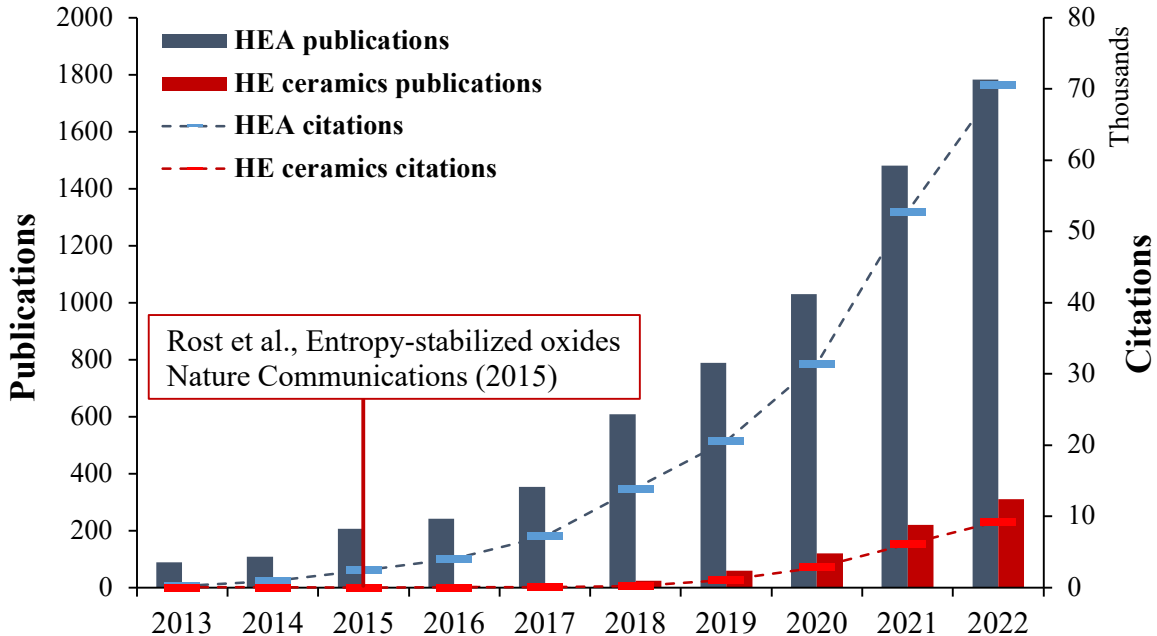


Figure 2.6: Publication history of journal articles on HEAs and high-entropy ceramics since 2013. Data collected from Web of Science (webofscience.com) using boolean topic searches: [(“high-entropy”/“entropy-stabilized”)&alloy] for HEAs and [(“high-entropy”/“entropy-stabilized”)&(carbide/oxide/nitride/ceramic)& alloy] for high-entropy ceramics.

## 2.2 High-Entropy Ceramics

Since the publication of the first entropy-stabilized oxide, many high entropy ceramics have been developed on the premise that the disorder inherent to a solid solution can be isolated to the cation sublattice of an ionic material. Examples of high entropy ceramics in literature include oxides, carbides, diborides, nitrides, sulfides, spinels, and perovskites [4]. In addition to the structural and thermal barrier applications driving this research, (Mg,Co,Ni,Cu,Zn)O has been studied for use as an oxidation catalyst and energy storage material [46, 47], while the (Mg,Co,Ni,Fe)O<sub>1.2</sub> mixed-phase rocksalt-spinel has been proposed for applications in water-splitting [48]. The ability to introduce the lattice disorder of high-entropy materials onto any given ionic lattice has paved the way for studies into a wide variety of structural, thermal, and functional properties. As a result, the field of high-entropy ceramics research is following in the footsteps of HEAs, with publication and citation rates matching or exceeding

the growth of the HEA literature in the early 2010s, as shown in in Fig. 2.6.

Searching for materials which took advantage of the extent properties of the metallic HEAs, our research focused on candidate materials for structural and thermal barrier coating materials. For cutting-edge aerospace applications, leading edge components on hypersonic-flight and re-entry vehicles can be exposed to temperatures as high has 2400 °C. The early-transition metal carbides, nitrides, and diborides are a class of high hardness, high melting temperature materials which are frequently studied as candidate materials for these types of extreme environment applications [49, 50]. Most of the group IV and V carbides, nitrides, and diborides are classified as UHTCs, defined as a melting temperature above 3000 °C [51]. In particular,  $\text{Hf}_x\text{Ta}_{1-x}\text{C}$  alloys are considered to be the high melting materials known, with melting temperatures exceeding 4000 K while exhibiting high hardness and good ablation resistance [52]. Based on the properties of the constituent binary ceramics, it is expected that the high-entropy combinations will make good candidate materials for use in extreme environments.

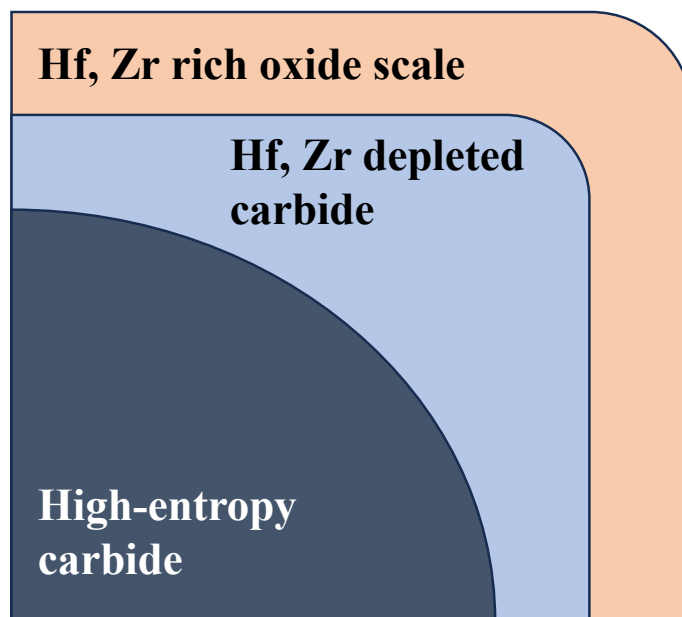


Figure 2.7: Schematic representation of the selective oxidation observed experimentally in (Hf,Nb,Ta,Ti,Zr)C.

While the high-entropy (Hf,Nb,Ta,Ti,Zr)C is predicted to have a melting temperature on the order of its constituent binary carbides ( $\approx 3000$  °C) [53], experimental oxidation studies by Backman et al. [54, 55] showed that it exhibited relatively poor oxidation resistance, with selective oxidation behavior driving dealloying in the oxidation affected region near the surface. Oxidation experiments at 1700 °C in a 1 mol% oxygen environment led to the formation of a TiO<sub>2</sub> oxide phase along with a Hf and Zr rich oxide scale, with a noted depletion of Hf and Zr around the edge of the sample cross-section, as shown schematically in Fig. 2.7.

Given that the entropy of a many-component system is dependent on the stoichiometry of the composition, dealloying as a result of selective oxidation could disrupt the free energy balance, perhaps changing the solubility limit of the components in a metastable structure. A brief simulation was performed to test this premise, wherein a pseudo-ternary carbide corresponding to the selective oxidation in (Hf,Nb,Ta,Ti,Zr)C was considered, with the end-members specified as (Hf,Zr)C, (Ta,Nb)C, and TiC. The enthalpy for the ternary phase diagram was calculated using Gaussian process regression for a small set of disordered 80-atom DFT cells in the vicinity of the equiatomic composition. The Gaussian process produces a continuous function for the enthalpy surface with respect to the concentration of the end-members, and is accompanied by a model uncertainty for each point on the surface. The prediction and uncertainty of the Gaussian process enthalpy surface are plotted in Fig. 2.8. Notably, the uncertainty increases dramatically outside of the region of interest as the model was not trained on these configurations.

The ideal solubility limit of each end-member can be found by taking the derivative of the total free energy (i.e. the sum of the Gaussian process enthalpy of formation and the ideal entropy as given in Eq. 2.5) with respect to the concentration and setting it equal to zero. Evaluating the sign of the first derivative indicates whether an increase in the concentration of the specified end member will increase or decrease the total free energy. Demonstrated for the model system at 2000 K in Fig. 2.9, we see that the solubility limit of TiC reduces in (Hf,Nb,Ta,Ti,Zr)C relative to the remaining species after a 35% reduction of (Hf,Zr)C

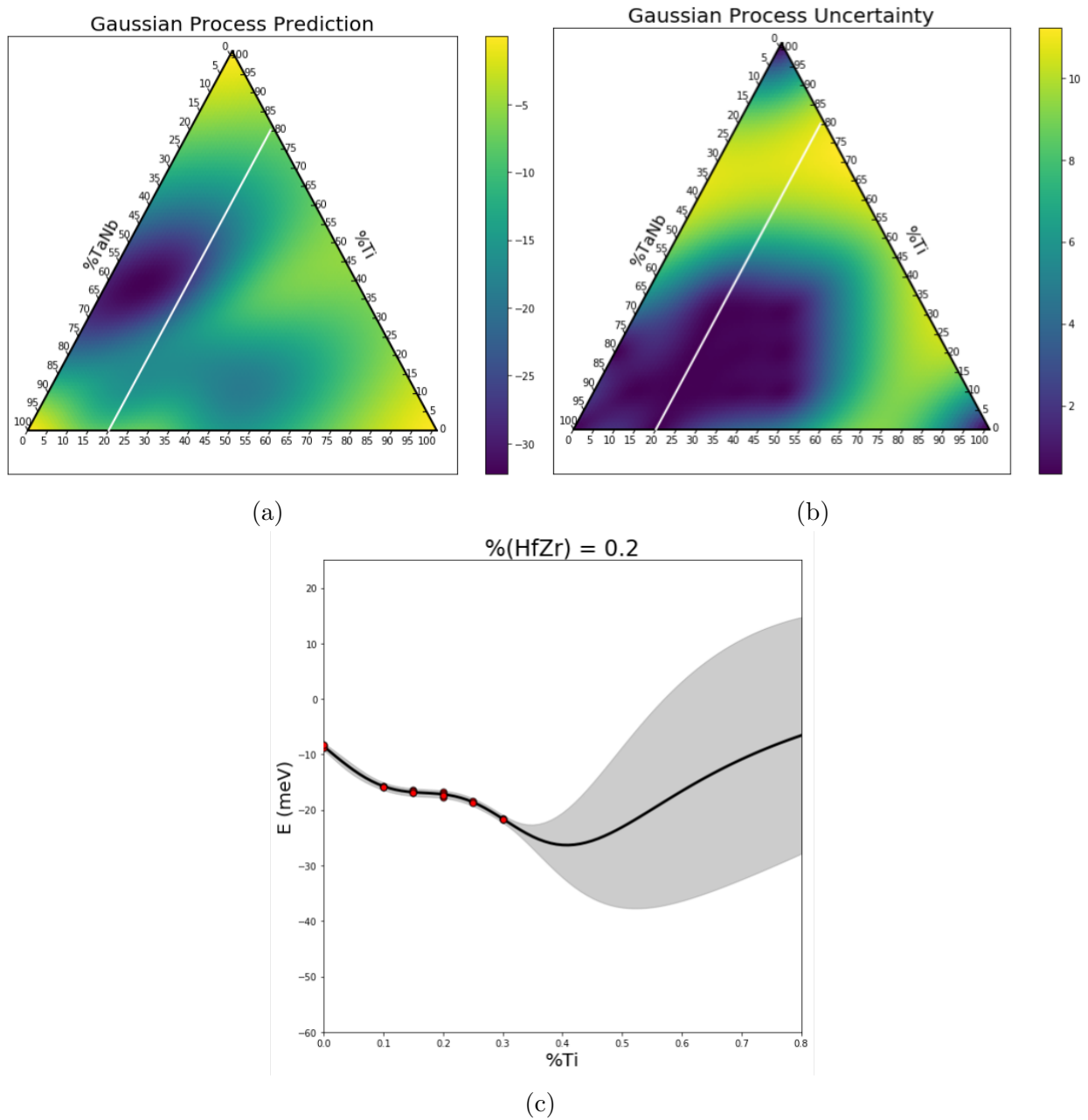


Figure 2.8: Gaussian process prediction (a) and uncertainty (b) of formation enthalpy for (Hf,Nb,Ta,Ti,Zr)C as a ternary system with (Hf,Zr)C, (Nb,Ta)C, and TiC end members. Cross-section for a fixed (Hf,Zr)C composition (c) is provided to illustrate the shape of the Gaussian process function.

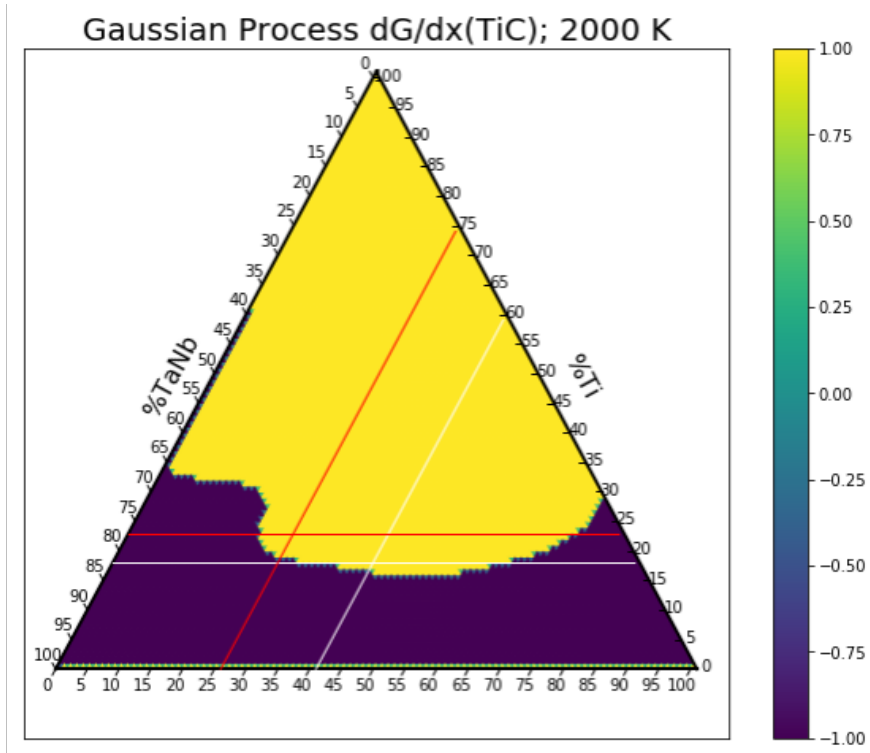


Figure 2.9: Predicted solubility limit of TiC in the high-entropy (Hf,Nb,Ta,Ti,Zr)C after dealloying, represented as a ternary system of (Hf,Zr)C, (Nb,Ta)C, and TiC. Composition indicated for the equimolar phase (white) and the composition after 35% depletion of (Hf,Zr)C e.g. into an oxide film (red). Positive values of the free energy derivative indicate a supersaturation of the component in the medium-entropy depleted system.

corresponding with the selective oxidation seen experimentally.

## 2.3 Defects

Defects in high-entropy ceramics are the primary focus of this research. While the energies and concentrations for defects such as vacancies are well understood through the statistical thermodynamics lens, the diverse local interactions in high-entropy materials complicate the issue and require new frameworks to understand and model defect properties.

In the same way that dislocation motion in HEAs is influenced by the fluctuations in local composition and strain, defect energies are defined by the local atomic environment. To better understand the variations in the local environment, it is instructive to analyze lattice

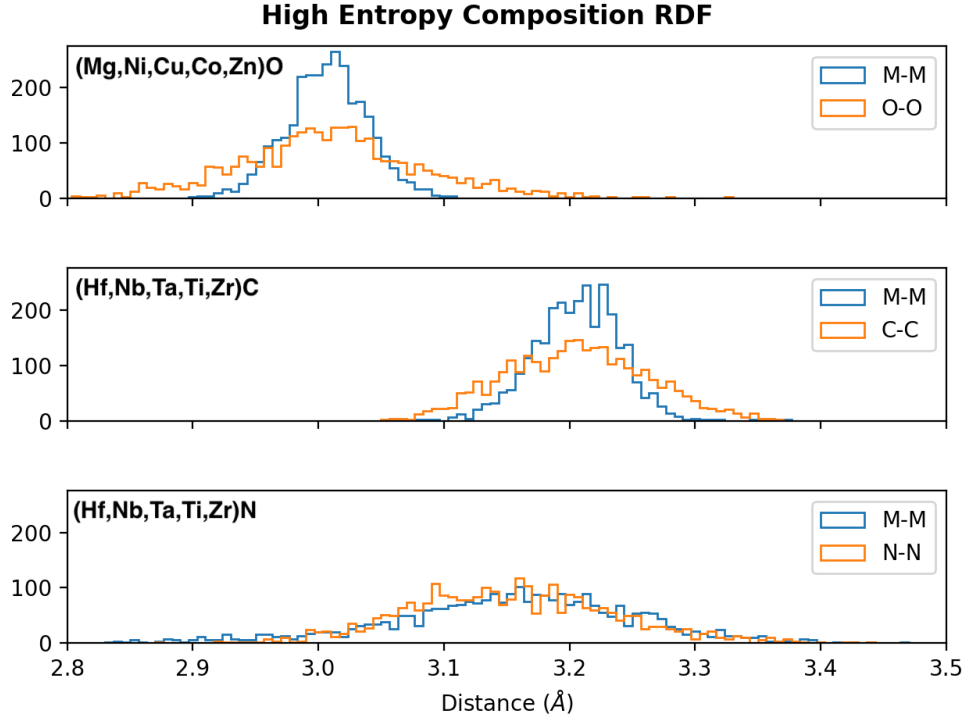


Figure 2.10: Comparison of the RDFs for a high-entropy oxide, carbide, and nitride separated into metal-metal and anion-anion distances. Separating the sublattice partial RDFs demonstrates the localization of the disorder on the oxygen sublattice, while the nitride exhibits uniform distortion across sublattices.

distortions as they exist in the high entropy ceramics. In Fig. 2.10, the first peak of the partial radial distribution function (RDF), as generated from DFT simulation, is plotted for metal-metal and anion-anion neighbor pairs in three different high-entropy rocksalt ceramics:  $(\text{Mg,Ni,Cu,Co,Zn})\text{O}$ ,  $(\text{Hf,Nb,Ta,Ti,Zr})\text{C}$ , and  $(\text{Hf,Nb,Ta,Ti,Zr})\text{N}$ . Between these three systems, there is a distinct difference in the relative behavior of the two sublattices. For the oxide and carbide compositions, the distortion on the lattice is concentrated on the anion sublattice, with the disordered metal sublattice exhibiting a narrower distribution of pair distances. In the oxide, this localization of the distortion on the oxygen sublattice is consistent with the experimental results of Rost [56]. In the case of the nitride, however, the bond character changes such that the sublattices are distorted uniformly.

Expanding this analysis to a second carbide composition, we see in the radial distributions

plotted in Fig. 2.11 that the group IV-V-VI carbide (Hf,Mo,Ta,W,Zr)C mirrors the uniform distortion of the nitride sublattice rather than the localized distortions of the group IV-V carbide. Breaking the interactions down into the RDF for each atom type and counting all metal and carbon neighbors, as plotted in Fig. 2.11(b), we find that the group VI transition metals account for the majority of the large distortions on the cation sublattice, with Zr-X and Ta-X interactions exhibiting narrower RDF peaks. For the analysis of high-entropy defects in this work, it is important to consider this element-dependent distortion behavior, as large deviations from the ideal lattice site can contribute excess strain energies to a defect energy calculation.

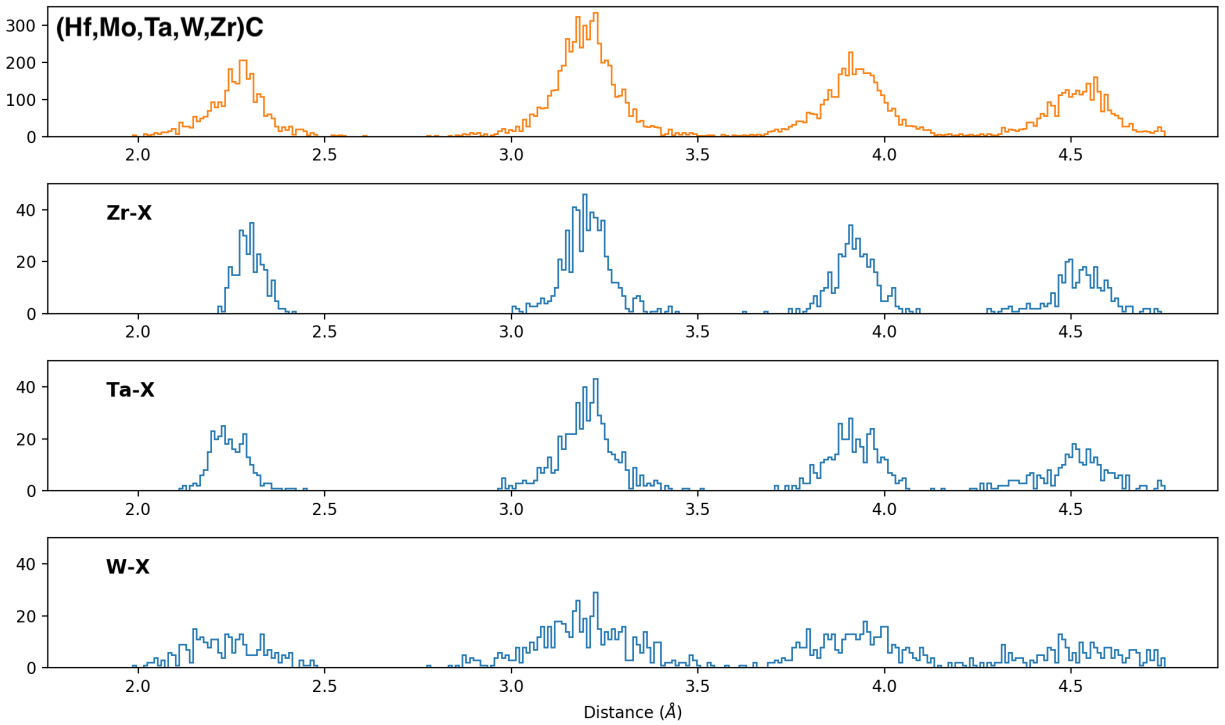
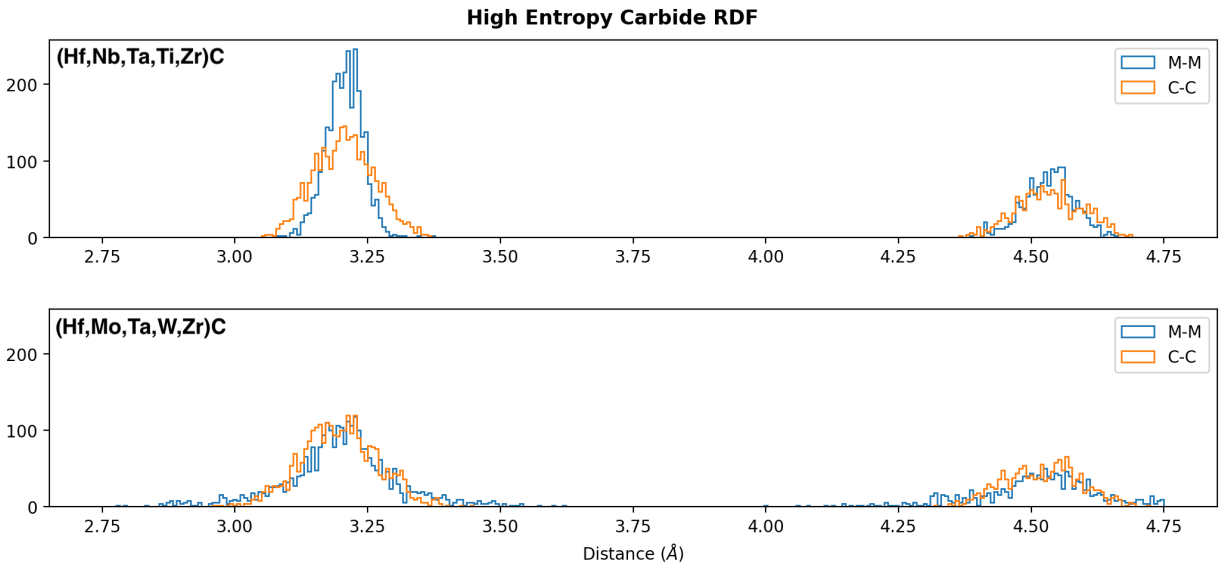


Figure 2.11: Comparison of radial distribution functions for two high-entropy carbides (Hf,Nb,Ta,Ti,Zr)C and (Hf,Mo,Ta,W,Zr)C separated into (a) metal-metal and carbon-carbon distances, and (b) isolated  $i$ -X pairs in the composition containing Mo and W, representing the distance from a specific atom type to its near-neighbor environment.

## CHAPTER

### 3

# VACANCY FORMATION ENERGIES IN HIGH ENTROPY CRYSTALS: APPLICATION TO A HIGH-ENTROPY DIBORIDE

Reproduced from S. E. Daigle and D. W. Brenner, “Statistical Approach to Obtaining Vacancy Formation Energies in High-Entropy Crystals from First Principles Calculations: Application to a High-Entropy Diboride,” *Phys. Rev. Materials*, 4(12), 123602, Dec. 2020, doi: 10.1103/PhysRevMaterials.4.123602 [57].

## 3.1 Abstract

A reduced pair approximation model for vacancy formation energy in multicomponent materials is proposed as an alternative to the commonly used cluster expansion method. By imposing physical constraints to the interaction coefficients, lower rank models are obtained with improved accuracy as measured by Bayesian information criterion and cross-validation. Additionally, reduced models can outperform the full parameterization in high-entropy compounds with as much as 50% less training data. The results are presented for cation vacancies in the high-entropy diboride  $\text{Hf}_{0.2}\text{Zr}_{0.2}\text{Ti}_{0.2}\text{Ta}_{0.2}\text{Nb}_{0.2}\text{B}_2$  calculated by density functional theory simulations of large cell special quasirandom structures. Further, the calculation of vacancy concentrations from a distribution of energies is considered, wherein the chemical disorder on lattice sites gives rise to non-Arrhenius temperature dependence. Preferential clustering and the possibility of short-range order in the high-entropy lattice are explored through pair affinities derived from model coefficients.

## 3.2 Introduction

High-entropy ceramics, a relatively recent variation on high-entropy alloys, have emerged as a new class of material with unique and potentially important functional and structural properties. Possible applications for these materials include catalysis, barriers for thermal and environmental protection, substrates for water splitting, and as energy storage materials [58]. The high-entropy ceramics reported to date have two sublattices: an ordered sublattice with one or two species, and a second sublattice with multiple, randomly-arranged species. High-entropy rocksalt oxides [59] and carbides [60], for example, consist of an fcc sublattice occupied by oxygen or carbon, respectively, and a chemically disordered interpenetrating fcc sublattice containing multiple metallic species. Other examples of high-entropy ceramics include spinel and perovskite oxides, metal diborides, nitrides, and silicides [18, 23, 59–77].

While the presence of multiple species is important for thermochemical stability as well as

structural and functional properties, their presence also creates special challenges to predicting phase stability and defect properties. In general, as the number of species in a given phase increases, the availability of reliable thermodynamic phase data tends to decrease. This has led to the development of quantities such as the Entropy Forming Ability [78] introduced by Curtarolo and co-workers that can be used to efficiently screen for potentially stable high-entropy compounds from first principles data. Similarly, the calculation of vacancy formation energies becomes complex due to the range of compositions and configurations available to a vacancy site on the chemically disordered lattice. Unlike in chemically ordered crystals, vacancies in crystals with random mixing cannot be represented by a single defect energy. Each vacancy must instead be treated individually with respect to not only the species removed, but also the corresponding occupations of the local environment [79, 80]. The number of possible occupations can be exceedingly large for these systems, particularly when multiple neighbor shells need to be considered to accurately represent the vacancy formation energies. Furthermore, the way in which diverse vacancies influence the bulk vacancy concentration must be carefully considered where a single energy would have otherwise been used in the relevant thermodynamic equations.

Much of the literature on vacancies in disordered systems has centered on the cluster expansion method and its application to the simplest case of binary systems [79–82]. This method takes the linear sum of interactions for every atom cluster in a material to a given distance and accuracy (e.g., pair, triplet, or n-tuple) to calculate the relative energy of ordering in the local atomic structure. When applied to vacancies, the vacancy can be treated as an  $N + 1$  species in the cluster expansion. While the cluster expansion method can predict accurate energies across the configurational and compositional space, the number of coefficients generally scales as roughly  $N^x$  with  $N$  as the number of components and  $x$  as the number of atoms in the highest order cluster [83]. For a high-entropy compound with five or more principal components, this rapidly becomes intractable, requiring a tremendous number of simulations to properly fit the interaction coefficients.

This work proposes a simplified neighbor interaction model derived from generalizations of the physical interactions in the local environment and guided by parameter space statistics to enable efficient prediction of cation vacancy formation energies, vacancy concentrations, and cluster affinities in many-component materials. The model is applied to the high-entropy diboride  $\text{Hf}_{0.2}\text{Zr}_{0.2}\text{Ti}_{0.2}\text{Ta}_{0.2}\text{Nb}_{0.2}\text{B}_2$ , which has a crystalline planar AlB<sub>2</sub> structure with one plane containing boron and the other plane containing the randomly arranged metal cations. Parameter sets within this model were determined using the Bayesian information criterion (BIC) [84] through a goodness of fit to first principles calculations combined with a penalty imposed on the size of the parameter space. The predictive performance of the model was tested via regression across a range of training sample sizes through repeated k-fold and leave-one-out (LOOCV) cross-validation. Using the model, a dense sampling of the discrete distribution of cation vacancy formation energies was generated and used to estimate vacancy concentrations as a function of temperature. The method, which includes uncertainty estimates for the interaction parameters, is general and therefore can be applied to other high-entropy compounds.

### 3.3 Methods

DFT energies of bulk and vacancy structures were calculated using plane-wave projector-augmented wave pseudopotential methods [85] implemented in the Vienna Ab initio Simulation Package [86–88]. The generalized gradient approximation as parameterized by Perdew *et al.* [89] was used for the exchange-correlation potential. The cut-off energy for the plane wave basis was set to 415 eV with a  $3 \times 3 \times 3$   $\Gamma$ -centered k-point mesh. Five distinct configurations of a 135 atom (45 cation) special quasirandom structure (SQS) [24] were constructed on a skewed hexagonal lattice such that vacancy-vacancy distances were greater than 10 Å and artifacts from periodic lattice strains on tiled vectors were minimized. Vacancy simulations were relaxed with respect to their structures while fixing the supercell volume to that of the

bulk in order to capture the relevant strains in the local lattice.

Vacancy formation energies  $E^{vf}$  were calculated for each cation site in the five SQS configurations and for each constituent binary diboride using the expression

$$E^{vf} = E^v - E^{bulk} + \mu . \quad (3.1)$$

Here,  $E^v$  is the energy of the vacancy-containing supercell,  $E^{bulk}$  is the energy of the bulk supercell, and  $\mu$  is the chemical potential of the vacating species taken as the DFT energy per atom of the elemental solid phase. Interaction parameters were fit to 225 vacancies simulated at every cation site in the SQS supercells (45 vacancies for each cation type).

### 3.3.1 Pair approximation model

Unlike the cluster expansion method, which references the summation to the total energy, solving explicitly for the vacancy formation energy allows one to consider instead the contribution of each neighbor interaction to the vacancy formation directly. Following the pair approximation model in Ref. [90], the vacancy formation energy can be represented as

$$E^{vf} = nE^b + E^r - \mu , \quad (3.2)$$

where  $E^b$  is the energy required to break each of  $n$  bonds between the atom and its neighbors,  $E^r$  is the electronic and structural relaxation of the local environment surrounding the vacancy, and  $\mu$  is the chemical potential of the vacating atom (alternatively in this framework, the atom can be placed on the surface of an infinite bulk with half of its bonds restored). In the case of a multicomponent system, a linear contribution from each neighbor is assumed and the equation is rewritten as

$$E^{vf} = \sum_{j=1}^N n_j (E_{ij}^b + E_j^r) + E_i^{r0} - \mu_i \quad (3.3)$$

or, for multiple neighbor shells,

$$E^{vf} = \sum_{j=1}^N \sum_{k=1}^K n_{jk} (E_{ijk}^b + E_{jk}^r) + E_i^{r0} - \mu_i . \quad (3.4)$$

Here,  $i$  represents the atomic species of the removed atom,  $j$  is the species of a given neighboring atom, and  $k$  is the index of the neighbor shell. Each bond  $E_{ijk}^b$  is represented by an  $ij$  pair in shell  $k$ , whereas the relaxation component  $E_{jk}^r$  is independent of the vacating atom  $i$ , and  $E_i^{r0}$  is independent of any bonds.

In the case of the diboride (or for any other ionic and covalent materials), the terms in the equation also have embedded in them the neighbor dependent contributions to the  $i$ -type bonding with the anion sublattice, as this cannot be isolated in the DFT energies. Because the boron sites are compositionally invariant, they do not need to be considered explicitly. Rather, the bonding and relaxation terms can be considered as the effective interactions of the variants in the composition.

With Eq. (3.4) as a starting point, physical constraints can be applied to reduce the parameter space. First, the bond energy between two atoms in the compound must be non-directional, i.e.,

$$E_{ijk}^b = E_{jik}^b . \quad (3.5)$$

Additionally, while bonded interactions are unlikely to extend beyond the first few neighbor shells, structural and electronic relaxation contributions can be longer ranged. From this, the vacancy interactions of distant neighbors can be reduced to only their relaxation components by imposing

$$E_{ijk}^b = 0 \quad (3.6)$$

for any shell  $k$  that is determined to have a non-influential bonded interaction.

A final reduction of the coefficients is made with the introduction of semi-arbitrary parameters for each interaction term. Specifically, it can be shown that for any vector  $U$ :

$$U + U^T = M , \quad (3.7)$$

where  $U^T$  is the transpose of  $U$  and  $M$  is a symmetric matrix  $M_{ij} = M_{ji}$ . Thus,

$$A + U + U^T = B$$

and

$$A + U = B - U^T \quad (3.8)$$

for some arbitrary symmetric matrices  $A$  and  $B$ . Applied to the bond matrices for a given shell, this yields

$$E^b + E^r = E^b - [E^r]^T , \quad (3.9)$$

where  $E^b - [E^r]^T$  is an equivalent solution to the pair approximation model with the relaxation terms indexed over vacating atom  $i$ . Defining the set of neighbor shells  $B$  such that bonding parameters  $E_{ijk}^b$  are symmetric and nonzero for all  $k \in B$ , the relaxation terms can now be summed to a single  $i$ -wise vector with elements  $E_i^r$  such that

$$E_i^r = E_i^{r0} - \mu_i \sum_j^N \sum_{k \in B} n_{jk} E_{ik}^r . \quad (3.10)$$

Because  $\sum_j n_{jk}$  is constant for each shell  $k$  in a given crystal structure,  $E_i^r$  is independent of the neighbor shell compositions.

Applying these assumptions to the pair approximation in Eq. (3.4), the total parameteri-

zation for the five-cation case ( $N = 5$ ) is reduced to

$$E^{vf} = \sum_{j=1}^5 \left( \sum_{k \in B} n_{jk} E_{ijk}^b + \sum_{k \in R} n_{jk} E_{jk}^r \right) + E_i^r \quad (3.11)$$

for neighbor shells  $B$  that involve energy contributions from broken bonds, and for shells  $R$  that only contribute to the relaxation energy. The determination of shells included in  $B$  and  $R$  is made through statistical evaluation of the model with different parameter sets. In the form presented in Eq. (3.11), there remains an additional degree of freedom for each contributing neighbor shell. These can be set as constant for a fixed reference point in the results, e.g., by setting each  $E_{11k}^b$  and  $E_{1k}^r$  equal to zero.

This reduced pair approximation (RPA) model considers only the composition of neighbor shells, representing a simplification of other common methods that introduce additional coefficients to account for the configurations of neighbor pairs, triplets and n-tuples. In the case of the five-cation high-entropy diboride, this model can represent the  $3.45 \cdot 10^{16}$  possible compositions of the first seven neighbor shells of a vacancy (cutoff radius of 7 Å) with as few as 53 linear degrees of freedom. As demonstrated in the following section, imposing the generalized physical constraints allows for an accurate representation of vacancy formation energies in high-entropy systems without the need to consider the many possible decorations of the local environment.

### 3.4 Results and Discussion

For the selection of bonded and relaxation sets  $B$  and  $R$  in Eq. (3.11), the BIC is compared for each disjoint combination of  $k \in B$ ,  $k \in R$ , and  $k \in reject$  including all shell indices  $k = 8$ . The BIC for a given set of parameters is calculated from the goodness of fit of the DFT results with a penalty imposed on the size of the parameter space, specifically

$$BIC = k_M \ln n - 2 \ln \hat{L}, \quad (3.12)$$

for sample size  $n$ , model dimension  $k_M$ , and maximum likelihood  $\hat{L}$ . For BIC calculation in this work, errors are assumed to be independent and identically distributed. For the Gaussian case, the log likelihood is calculated from the variance estimate  $\hat{\sigma}^2$  of multiple linear regression as

$$\hat{L} = \frac{1}{2}n[\ln 2\pi + \ln \hat{\sigma}^2 + 1] . \quad (3.13)$$

Further, the formulation in Eq. (3.11) is expanded to include  $k \geq U$  for unconstrained parameterization of neighbor shells wherein the assumptions in Section 3.3 are ignored in favor of the formulation in Eq. (3.4). This allows for rejection of the proposed framework for any shell  $k$  during model selection.

Within the 50 best models (as ranked by BIC) in the 65,536 parameter set combinations up to eight neighbor shells, the selection frequency of each shell classification is calculated and plotted in Fig. 3.1. In a given shell, a high rate of unconstrained parameters might suggest that the proposed model assumptions are flawed, while a low but non-zero rate points to the information contained in that shell being valuable enough to outweigh an inefficient parameterization. The selection frequencies show that the reduced parameterizations dominate in the models with the best expected performance, as well as a trend toward parameter sets without a bond energy term (and, eventually, rejected parameters) with increasing distance from the vacancy site.

To quantify the predictive performance of the selected models for  $\text{Hf}_{0.2}\text{Zr}_{0.2}\text{Ti}_{0.2}\text{Ta}_{0.2}\text{Nb}_{0.2}\text{B}_2$ , cross-validation scores were calculated on models fit to the DFT results using multiple regression across a range of training sample sizes through repeated k-fold cross-validation and LOOCV. Plotted in Fig. 3.2 is the performance and convergence behavior of an optimized parameterization ( $B = f1, 2, 3, 4g$ ,  $R = f5, 6, 7g$ ,  $U = ;$ , and  $reject = f8g$ ) compared against the rest of the top 50 models and the fully unconstrained model for  $K = 7$ . Additionally, cross-validation error is plotted against the

largest shell  $K$  for the optimized and unconstrained models. This shows improved predictive performance with the inclusion of each of the first seven neighbor shells, in agreement with the parameter selection frequency in Fig. 3.1.

For all sample sizes and cutoff radii, the RPA outperforms the unconstrained model. In particular, the RPA model is shown to be more resilient to overfitting when including distant pairs, even with smaller sample sizes. The resilience of the model is, in part, gained from the symmetry of the parameters linking the coefficients from different vacancy types. Whereas the unconstrained parameterization acts as separate, independent models for each  $i$ -type vacancy, the RPA coefficients are symmetric across  $ij$  terms, effectively increasing the size of the training set available to each parameter while removing unnecessary degrees of freedom.

Of note for models with  $K = 2$  and relatively few parameters, the unconstrained model is only able to approach the performance of the optimized RPA, not surpass it, suggesting that

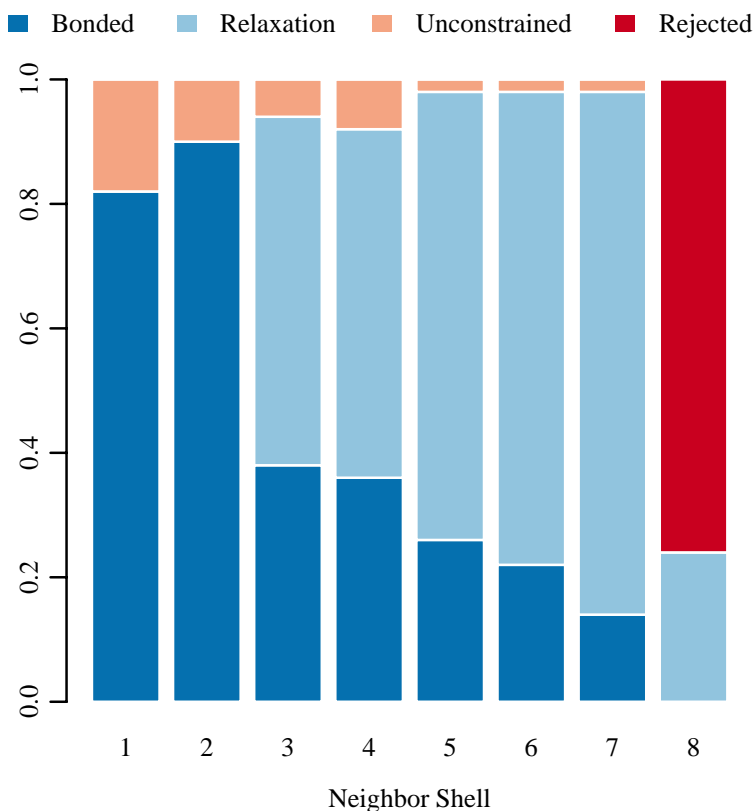


Figure 3.1: Selection frequency of each parameter set in the top 50 models by BIC.

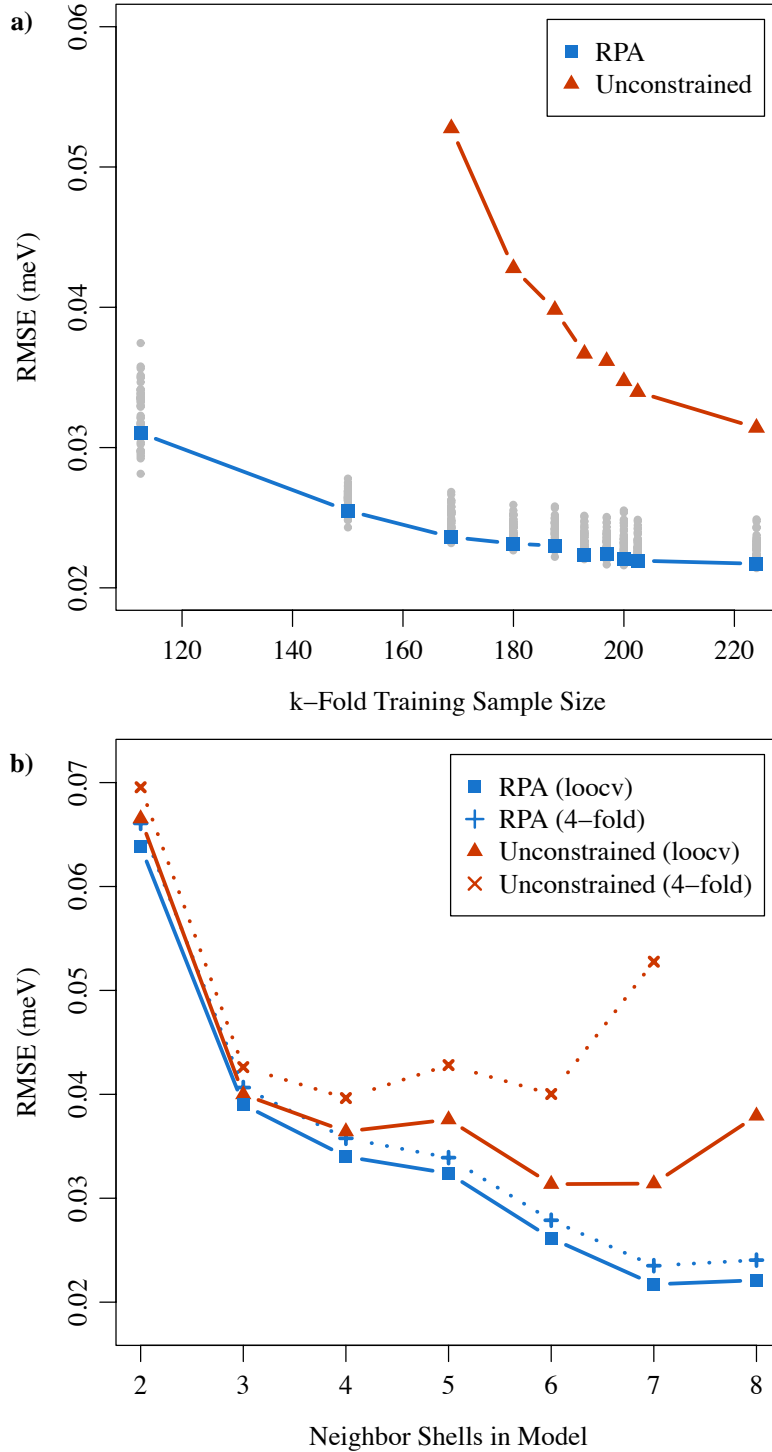


Figure 3.2: (a) Convergence of the reduced pair approximation model ( $B = f1, 2, 3, 4g$ ,  $R = f5, 6, 7g$ , and  $U = ;$ ) and the unconstrained model ( $U = f1, \dots, 7g$ ) as a function of training sample size for 5x k-fold cross validation up to and including LOOCV; and (b) cross-validation error as a function of largest shell  $K$  in the model for 4-fold and LOOCV (training sample size of 169 and 224 respectively). Cross-validation errors for the 50 top models by BIC shown in grey.

Table 3.1: Root mean squared error (in meV) from least squares fitting as well as cross-validation errors from LOOCV and 5x 4-fold CV for unconstrained and RPA models. SRO is the prediction error for the set of pseudo-ordered simulations. Sets  $B$  and  $R$  denote the indices of shells modeled as bonded and relaxation interactions, respectively.

Assumption	Parameters	BIC	RMSE	LOOCV	4-Fold CV	SRO
$U = \bar{f}1, \dots, 7g$	145	-601.7	11.0	31.4	52.8	101.4
$B = \bar{f}1, 2, 4g; R = \bar{f}3, 5, 6, 7g$	63	-878.6	15.9	22.3	24.4	77.0
$B = \bar{f}1, 2, 5g; R = \bar{f}3, 4, 6, 7g$	63	-876.6	16.0	22.4	24.2	74.3
$B = \bar{f}1, 2, 3, 6g; R = \bar{f}4, 5, 7g$	73	-869.3	14.4	21.4	23.2	90.6
$B = \bar{f}1, 2g; R = \bar{f}3, \dots, 7g$	53	-864.9	18.5	24.3	25.8	79.4
$B = \bar{f}1, 2, 3, 4g; R = \bar{f}5, 6, 7g$	73	-863.1	14.6	21.7	23.6	81.2

the RPA assumptions do not remove information from the model.

The optimizations of the RPA model converge quickly to root mean square prediction errors on the order of 20 meV across a 2 eV range of calculated formation energies. RPA models yield 30% lower LOOCV errors than those of the unconstrained model for training size  $n = 224$ , and 50% lower 4-fold cross-validation errors ( $n = 169$ ). Comparing the fitting errors and LOOCV prediction errors within each model, the RPA values are overfit by less than 50%, compared to 185% overfitting in the unconstrained case. Model parameters and fitness metrics are presented in Table 3.1.

The RPA model is shown to predict the DFT results for the stoichiometric and maximally disordered SQS; however, that result says little about the applicability of the model to non-ideally mixed local arrangements. By its formulation, every site in the SQS should closely approximate ideal mixing in a random compound. To investigate if the model can be extrapolated to atomic clusters that do not follow the assumptions imposed by the SQS formulation, a set of pseudo-ordered structures is constructed in which the nearest neighbor shells are selectively replaced, in turn, with each constituent metal species to approximate certain short-range order (SRO) clusters as illustrated in Fig. 3.3.

Due to supercell size constraints, each shell is calculated separately, and the surrounding bulk is left unchanged from the SQS, that is, we do not impose a deficient region surrounding

the SRO shell to maintain the equimolar stoichiometry of the high-entropy compound. As a result, the supercells are no longer fully consistent with the assumptions of our model, most importantly that the chemical potential reference assumes the compound to be compositionally invariant. It might also be expected that any ignored cluster interactions would play a larger role when present in such large concentrations, in addition to any volume or strain effects of the composition changes. These measurements are, as such, only useful as a qualitative guide to assess the applicability of the model as configurations deviate from ideal mixing. Despite the deficiencies of both the model and the simulations for representing these SRO cases, the predictions are largely in agreement with the simulated results. The modeled vacancy formation energies and the SRO predictions are plotted against the DFT energies in Fig. 3.4. RMSE around 75-80 meV for the SRO structures across 3 eV of predictions suggests reasonable extensibility of the RPA model beyond the explicitly formulated case.

Based on this result, the remaining analysis is conducted across the entire range of local

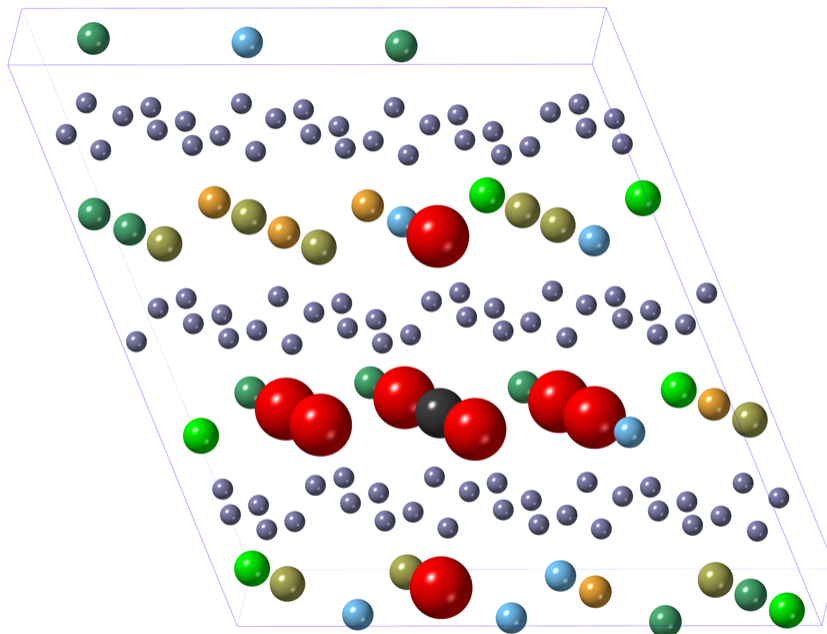


Figure 3.3: A sample construction of the SRO structure in which a near-neighbor shell (red) is replaced with a single element.

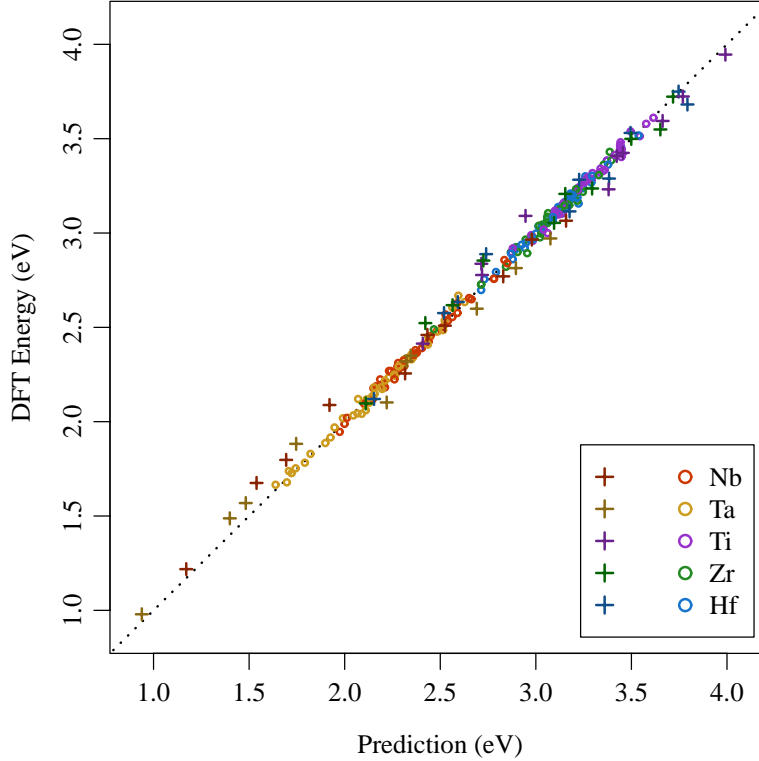


Figure 3.4: DFT simulated vacancy formation energies ( ) plotted against LOOCV predictions from the RPA model with  $B = f1, 2, 3, 4g$  and  $R = f5, 6, 7g$ . SRO predictions (+) included to illustrate an extrapolation of the model. SRO vacancies are not included in any model training sets.

compositions. Any errors that might emerge as the local order deviates from the ideal random SQS structures are omitted for the purposes of this work.

### 3.4.1 Vacancy concentration

Because the vacancy formation energy in disordered multicomponent compounds does not take on a single value, the traditional formalisms for calculating the vacancy concentration in a material are no longer straightforward. In the case of the proposed neighbor model for a five-component diboride, there are over  $3.45 \cdot 10^{16}$  possible compositions of the first seven neighbor shells. Each of these neighbor environments should be considered in the derivation of the vacancy concentration.

In Fig. 3.5 the histogram of vacancy formation energies from the 225 DFT calculations,

representing a set of ideal random arrangements, is compared to the predicted distribution of formation energies across all combinatorial compositions of the local environment. The binary diboride cation vacancy formation energies of each constituent metal are included for reference. Using only the explicitly calculated SQS structures results in a clear bias of the formation energy distribution toward the peaks. The peaks of the distribution represent the most abundant configurations in an ideally mixed compound; however, these are not the most energetically favorable configurations and therefore may not be the largest contributors to the total vacancy concentration. By applying the model across the entire composition space, a complete picture of the vacancy formation energy distribution can be used to calculate the precise contributions to the global vacancy concentration.

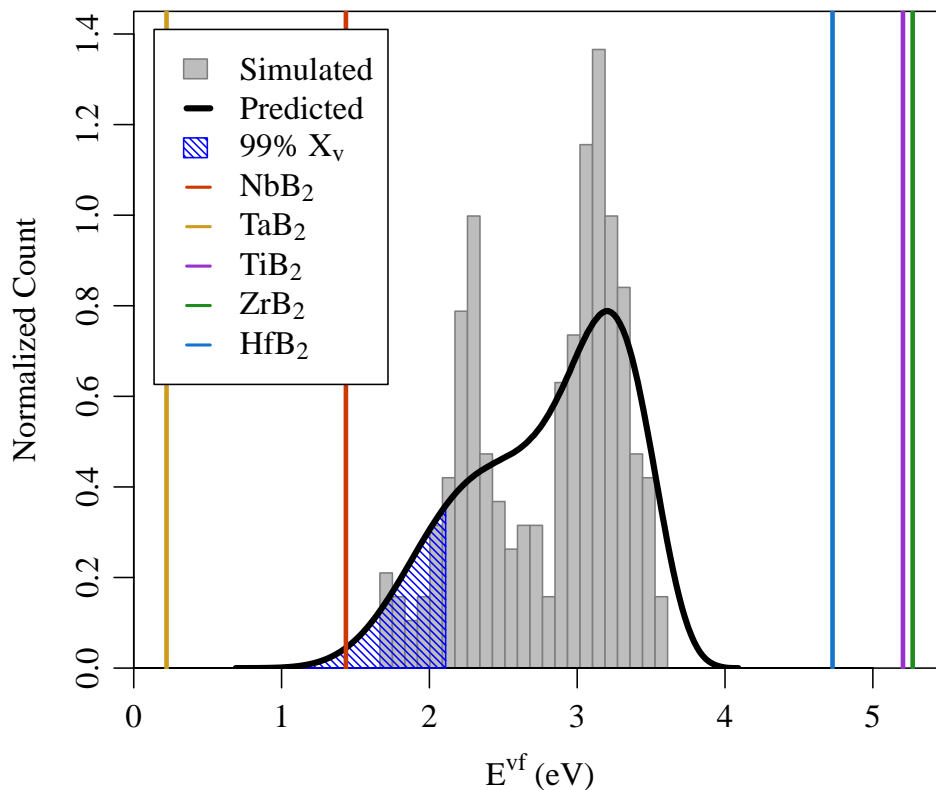


Figure 3.5: Vacancy formation energies of all simulations and the predicted total distribution using the RPA model. Here, the parameter set of  $B = f1, 2, 3g$  is used for illustrative purposes. The shaded region represents the vacancy configurations which account for 99% of the total concentration at 2000 K. Vertical lines indicate the binary diboride vacancy formation energies of each constituent element.

An important issue when dealing with a multicomponent system is how to reconcile the identity of a vacancy. When approaching the issue from a bulk reference, as in the case of the cluster expansion method, it can be stated that vacancies do not have an identity. While this is ostensibly true in all cases, when using a chemical potential reference it becomes critical to consider what previously occupied the vacancy site.

Further, Zhang and Sluiter have suggested that the composition can be constrained such that the vacancy concentration can be solved with the use of Lagrange multipliers [80]. While it might be mathematically convenient to impose such constraints (and in the case of the cluster expansion method, perhaps necessary), it is not necessarily an accurate representation of a system that interacts with its environment.

Alternatively, a framework can be constructed in which each local environment (as defined by the RPA model) is considered to be a distinct lattice site, and the occupation of each lattice site is calculated independently. Applying traditional Boltzmann statistics to each site configuration yields an equation for the total vacancy concentration:

$$X_v = \sum_{\sigma} p(\sigma) e^{-\frac{E(\sigma)}{kT}} , \quad (3.14)$$

summed over all configurations  $\sigma$ , where  $p(\sigma)$  and  $E(\sigma)$  represent the probability and vacancy formation energy, respectively, of a given site. Here,  $p(\sigma)$  is calculated using the purely combinatorial approach, however it would be trivial to add another layer of Boltzmann statistics to this term to reflect the relative probabilities of certain clusters in the case where SRO is a significant factor. This approach is similar to the method employed by Ruban [79], differing only in the use of a discrete distribution over a continuous approximation.

The vacancy concentrations derived from the predicted distribution were converged by random sampling of the available configuration space for two trillion arrangements. The temperature dependence of  $X_v$  is plotted in Fig. 3.6. Due to the unequal contributions of the varied energies, the total vacancy concentration deviates from Arrhenius behavior. The

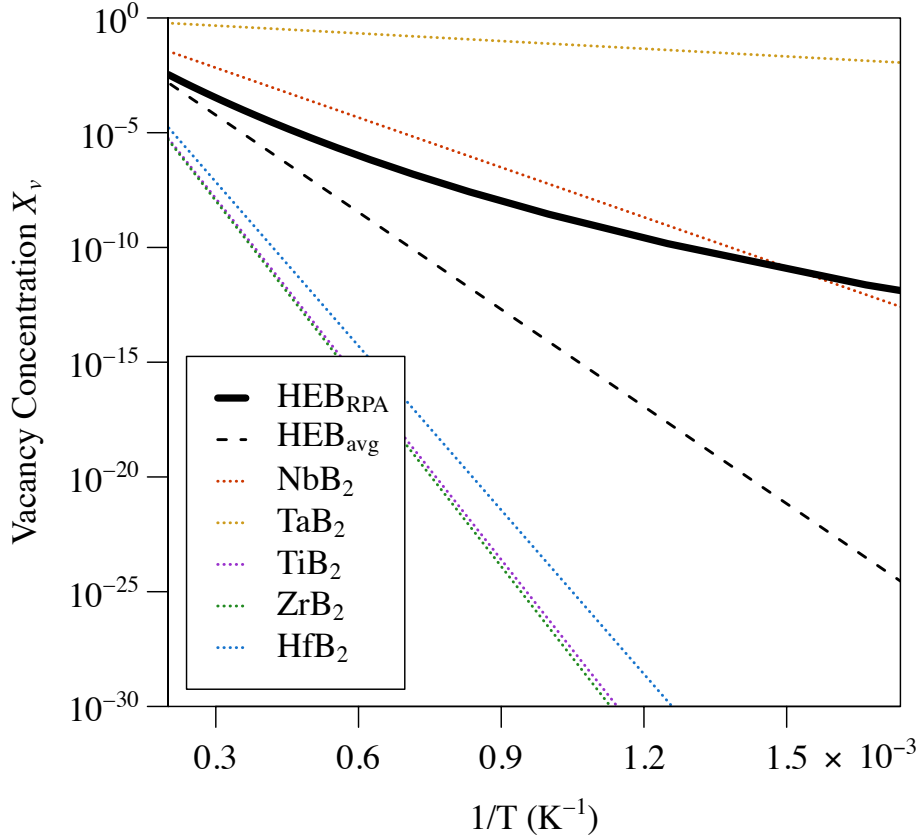


Figure 3.6: Cation vacancy concentration of the high-entropy diboride structure as a function of inverse temperature, calculated by individual site contributions as compared to the vacancy concentration of the uniform mean formation energy and the concentration in each of the binary transition metal diborides.

logarithm of  $X_v$  now takes on a nonlinear dependence on inverse temperature wherein the lower energy vacancies dominate at low temperature and the concentration converges to that of the mean formation energy at ultra-high temperatures. From this behavior, the calculation of the entire energy distribution becomes critical, as the energies in the left tail of the distribution have the largest contribution to the vacancy characteristics of the bulk. To further illustrate this uneven contribution of the different local environments, Fig. 3.5 plots the vacancy configurations that contribute 99% of the total vacancies in the bulk at 2000 K, all of which occur below the first peak of the distribution (primarily Nb and Ta vacancies).

Generally, whether by this isolated site method or by an alternative approach explored in the literature [79, 80, 91, 92], the vacancy concentration in a high-entropy system is

expected to be within the range of the constituent compounds, but above the average. Early experimental results from a recent positron annihilation study of a high-entropy alloy support this finding [93].

### 3.4.2 Configurational entropy

An additional factor in calculating the concentration of vacancies in a disordered material is the effect of a disordered multicomponent lattice on the entropic contribution to vacancy formation. This contribution was incorrectly derived in Ref. [94] as scaling with a factor of  $e^{(N-1)}/N$  for an  $N$  component system. Rather, the entropic contribution scales with  $1/N$  for small concentrations as derived in the appendix 3.6.

This solution is intuitive if the vacancy is considered to be an  $N + 1$  species, as suggested in Refs. [80, 81], due to the reduced configurational entropy contribution of the  $n^{\text{th}}$  species as illustrated for high-entropy oxides in the supplemental materials of Ref. [59]. Specifically, for each additional component added to a compound, there is a smaller gain in configurational entropy for the equivalent increase in concentration, always reaching a maximum entropy at the equiatomic composition. The decrease in the entropic contribution to the vacancy formation energy, in turn, results in a lower vacancy concentration, calculated as

$$X_v = \frac{1}{N} \sum_{\sigma} p(\sigma) e^{-\frac{E(\sigma)}{kT}} . \quad (3.15)$$

This decrease is significant, though not as large as the differences in concentration between frameworks for addressing the enthalpy distribution.

### 3.4.3 Pair affinities

Because the pair approximation method implemented in this work is compositionally invariant and fixed to a chemical potential reference, it does not provide the same insight as the cluster expansion method into the bulk energies across the composition space of high-

entropy materials. It does, however, retain the ability to approximate pair affinities from the interaction coefficients within the constraints of the formulation. The pseudo-ordered structure calculations in this work demonstrate that while the prediction accuracy may suffer to some extent with varying composition and cluster order, the constraints on disorder and composition are not overly restrictive for application of the model to configurations that deviate from the ideally mixed SQS structure.

To find any pair affinities that might lead to short-range order in the compound, the relative energy of two configurations is taken as the bonds gained minus the bonds lost. To form an arbitrary additional  $ij$  bond, an  $iX$  and a  $jX$  bond must be broken, restoring an  $XX$

Table 3.2: Atom pair affinities (meV/bond) and standard deviations ( $\approx 2\sigma$ ) for cation near neighbors in the first four shells of the RPA model with  $B = f1, 2, 3, 4g$  and  $R = f5, 6, 7g$ .

$(i,j,1)$	Ti		Zr		Hf		Nb		Ta	
Ti	4.6	5.3	20.1	3.9	8.2	4.4	13.3	3.8	19.6	3.7
Zr			5.2	5.8	0.5	4.1	5.2	4.0	10.1	4.5
Hf					7.5	4.5	7.5	3.9	8.8	3.5
Nb							12.2	6.6	13.9	3.9
Ta									24.5	5.0
$(i,j,2)$	Ti		Zr		Hf		Nb		Ta	
Ti	5.7	7.6	14.5	6.6	9.9	6.2	11.1	5.7	19.1	5.5
Zr			0.9	9.4	0.0	6.9	2.1	6.0	13.2	5.8
Hf					11.6	5.8	13.3	6.1	8.2	5.9
Nb							14.7	9.1	11.9	5.9
Ta									28.6	8.3
$(i,j,3)$	Ti		Zr		Hf		Nb		Ta	
Ti	4.5	5.2	0.2	3.3	2.8	2.9	1.8	3.0	0.3	2.9
Zr			4.1	3.7	2.8	2.7	1.5	3.0	5.3	2.9
Hf					0.8	2.9	2.0	2.9	2.8	2.9
Nb							0.2	4.7	1.5	3.4
Ta									6.3	4.0
$(i,j,4)$	Ti		Zr		Hf		Nb		Ta	
Ti	4.2	7.1	9.0	4.4	2.0	4.4	2.6	3.9	4.2	4.1
Zr			4.8	7.0	6.3	4.1	1.2	4.1	0.9	4.7
Hf					6.1	4.7	5.3	4.2	5.1	4.0
Nb							1.5	5.0	5.4	3.7
Ta									3.6	5.4

bond in the process (where X matches the composition of the bulk). Recalling from Eq. (3.9) that for each fitted bonding parameter

$$E_{ijk}^b = E_{ijk}^b + E_{ik}^r + E_{jk}^r , \quad (3.16)$$

it is apparent that the relaxation terms will cancel out in any redistribution of bonds, so long as composition is conserved. As such, the arbitrary parameters  $E_{ijk}^b$  can be used directly for this analysis.

Furthermore, variances of the pair affinities can be computed using approximated local sensitivities. The covariance matrix of the parameters is found using

$$V = \sigma_0^2 [S^T S]^{-1} \quad (3.17)$$

for model variance  $\sigma_0^2$  from the sample size  $n$ , number of parameters  $p$ , and residuals matrix  $R$  as

$$\sigma_0^2 = \frac{1}{n-p} R R^T \quad (3.18)$$

and local sensitivity matrix  $S$ , calculated as the partial derivatives of the  $n$  samples with respect to each parameter. The terms of the local sensitivity matrix are defined as

$$S_{np} = \frac{dE_n^{vf}}{d\theta_p} \quad (3.19)$$

for each simulated vacancy formation energy  $E_n^{vf}$  and each interaction parameter  $\theta_p$ . Applied to the linear RPA model, the terms reduce to the parameter coefficients for each configuration, i.e., the number of each neighbor interaction  $n_{ijk}$  and  $n_{jk}$  in Eq. (3.11). The variances are extended to pair affinities using the expression

$$\sigma_{ij}^2 = c^T V c , \quad (3.20)$$

where  $c$  is a vector representing the linear combination of the total bonds added and removed. The derived pair affinities and uncertainties are presented in Table 3.2.

The largest positive affinities are noted for Ti-Zr, Ta-Nb, and Ta-Ta pairs in both the first and second shells, as well as negative affinities for the first and second Ti-Nb, Ti-Ta, and Zr-Nb pairs. While the magnitude of pair affinities generally decreases after the first and second neighbors, there is not a simple distance dependence of the interactions. Given the similar bond radii of the in-plane and out-of-plane neighbors that comprise the first two neighbor shells (3.1 Å and 3.4 Å respectively), orientational and boron sublattice bonding effects likely play a role.

A separate analysis would be necessary to determine the degree of short-range order that could result from these relative energies. Above a certain magnitude, one might expect to see clustering among high affinity pairs and segregation of negative affinity pairs; however, the trade-off between entropic stabilization and short-range order is beyond the scope of this work.

### 3.5 Conclusions

An efficient method for prediction of diverse vacancy formation energies across the vast composition space of high-entropy materials is applied to the high-entropy diboride  $\text{Hf}_{0.2}\text{Zr}_{0.2}\text{Ti}_{0.2}\text{Ta}_{0.2}\text{Nb}_{0.2}\text{B}_2$ . A reduced set of bonding and relaxation coefficients with 73 linear degrees of freedom is shown to outperform the unconstrained parameterization, achieving an RMSE for k-fold cross-validation of less than 25 meV on training sets as small as 150 vacancy simulations. This model framework enables rapid evaluation of vacancies in compounds across the high-entropy composition space.

Accurate evaluation of the vacancy formation energy distribution, particularly with respect to the left tail, is presented as critical for the calculation of vacancy concentrations in disordered materials. The bulk vacancy concentration is estimated as the combined total of

the isolated contributions across a dense sampling of the available vacancy configurations.

Pair affinities derived from pair interaction coefficients in the RPA model are presented as a means to investigate possible short-range order in high-entropy materials and likely cluster pairs are highlighted. Further derivation of the interplay between entropy and clustering interactions is left for future analyses.

### 3.6 Appendix: Configurational entropy derivation

The derivation for vacancy concentration in a multicomponent material mirrors the classical thermodynamic derivation, starting from the configurational entropy  $S_c$  for an equiatomic compound as

$$S_c = k[\ln N! - m \ln \left(\frac{N-n}{m}\right)! - n \ln n!] \quad (3.21)$$

for number of atoms  $N$ , number of species  $m$ , and number of vacancies  $n$  where composition is maintained. Applying Sterling's approximation and taking the derivative with respect to  $n$ :

$$S_c = k[N \ln N - (N-n) \ln \left(\frac{N-n}{m}\right) - n \ln n] \quad (3.22)$$

$$\Delta S_c = k \left[ \ln \left(\frac{N-n}{m}\right) - \ln n \right], \quad (3.23)$$

and for small  $n$ :

$$\Delta S_c = -k \ln \left(\frac{mn}{N-n}\right) \approx -k \ln \left(\frac{mn}{N}\right). \quad (3.24)$$

Setting the free energy equal to zero and rearranging to solve for vacancy concentration  $X_v$ ,

$$\Delta H + kT \ln \left(\frac{mn}{N}\right) = 0 \quad (3.25)$$

$$X_v = \frac{1}{m} \exp \left(-\frac{\Delta H}{kT}\right). \quad (3.26)$$

## 3.7 Acknowledgments

This work was supported by a Multi-Disciplinary University Research Initiative through Award No. N00014-15-1-2863 from the U.S. Office of Naval Research and by the National Science Foundation under Grant No. DGE-1633587 (Science and Engineering of the Atomic Structure at North Carolina State University).

## CHAPTER

# 4

# INTERFACIAL DEFECT PROPERTIES OF HIGH-ENTROPY CARBIDES: A NEW EVANS-POLANYI-SEMENOV RELATION

Reproduced from S. E. Daigle, S. Curtarolo, W. G. Fahrenholtz, J. P. Maria, D. E. Wolfe, E. Zurek, and D. W. Brenner, “Interfacial Defect Properties of High-Entropy Carbides: Stacking Faults, Shockley Partial Dislocations, and a New Evans-Polanyi-Semenov Relation,” in review. ArXiv preprint, Nov. 2023, doi: 10.48550/arXiv.2311.00834 [95].

## 4.1 Abstract

Using first principles calculations,  $\bar{1}11g$  intrinsic stacking fault (ISF) energies in Group IVB, VB, and VIB high-entropy transition metal carbides are shown to be predictable from an optimized rule of mixtures based on the properties of the single metal carbide constituents present near the stacking fault. A composition-independent linear relationship is demonstrated between the ISF energies and the unstable stacking fault (USF) energies along the  $h11\bar{2}i/\bar{1}11g$  gamma surface slip path. Treating the ISF and USF energies as analogous to the heat of reaction and transition state barrier in chemical reactions, this linear relationship represents a new application of the Evans-Polanyi-Semenov principle. Further, a full defect energy distribution can be obtained from the predicted ISF energies with only the composition as an input for the mixed early-transition metal carbides. Applying a model that balances the elastic repulsion between partial dislocations with the distribution of ISF energies, we show that Shockley partial edge dislocations should remain bound for all valence electron concentration values up to about 9.6, even when the average stacking fault energy is negative.

## 4.2 Background and Motivation

The first high-entropy ceramic, a five cation rocksalt oxide, was reported in 2015 [3]. Since then, the number of high-entropy ceramics has rapidly expanded and now includes a range of different oxides, carbides, diborides, nitrides, carbo-nitrides, perovskites and other related materials [4–16]. One of the central questions associated with these materials is the degree to which the structural, thermal, mechanical, electronic and chemical properties can be predicted by a rule of mixtures (RoM) based on their constituent compounds, and which properties are unique to a high-entropy ceramic. For example, the rocksalt oxide with cations  $\text{Mg}^{+2}$ ,  $\text{Ni}^{+2}$ ,  $\text{Co}^{+2}$ ,  $\text{Cu}^{+2}$ , and  $\text{Zn}^{+2}$  was found to have a lattice constant that is well described by a RoM and Bader charges that are transferable between the binary, ternary, and high-entropy compositions [96]. However, adding aliovalent cations such as  $\text{Li}^{+1}$  and  $\text{Sr}^{+3}$  results in a

range of cation and oxygen charges, multiple valence states, and a reduction in thermal conductivity due to enhanced phonon scattering from the different charge states [19, 97]. For stoichiometric high-entropy carbides with transition metals (HETMCs) from Groups IVB, VB, and VIB, theory predicts that lattice constants, bulk moduli, and cohesive energies are all well described by a RoM from the constitutive rocksalt binaries [21]. The carbon vacancy formation energies, on the other hand, cannot be accurately estimated from a simple RoM, but they can effectively be estimated using a neural network trained on computational energies [98].

A convenient computational tool to help to understand slip in crystals is the gamma surface. This is generated by moving two adjacent regions of a crystal along a given interfacial plane and calculating the potential energy after relaxing the system in the direction normal to the interface. The potential energy is given as a function of the relative in-plane displacement of the slabs to produce a potential energy landscape equivalent to a topological map. The gamma surface was introduced by Vitek to find energetically stable stacking faults in bcc materials [99], and has since become an invaluable tool for understanding slip and related mechanical deformations. For example, intrinsic stacking fault (ISF) energies (energy valleys) are used to predict partial dislocation separations, while unstable stacking fault (USF) energies (peaks in the gamma surface) can be used to predict energetically favorable slip systems. USF and ISF energies are also used in analytic expressions to predict quantities such as fracture and twinning tendencies. Tadmor and Bernstein, for example, developed an expression for predicting twinning that includes the ratio ISF/USF [100, 101].

For the research reported below, it is important to distinguish between quasi-static shear calculations [102], where energies may include a component of shear stress along the slip plane, and the gamma surface. The former attempts to better describe the physics of the shear process, while the latter is intended as a convenient tool with features such as the ISF and USF energies that yield insight into mechanical deformation. Hence, stress along a given slip path from a gamma surface is not necessarily a relevant physical quantity with respect

to quantities such as dislocation mobility [103].

Experimental and computational studies have established that the rocksalt Group IVB carbides TiC, ZrC, and HfC and the Group VB carbides VC, NbC, and TaC prefer to slip along  $h110i$  directions [104]. However, the Group IVB carbides slip along  $f110g$  planes, while the Group VB carbides slip along  $f111g$  planes. Studies of mixed IVB-VB carbides have reported either slip on both planes, or on  $f111g$  planes exclusively [105, 106]. By examining the gamma surface with energies calculated from Density Functional Theory (DFT), Thompson, Weinberger and co-workers have shown that the transition from slip on  $f110g$  to  $f111g$  planes can be attributed to a lowering of the barrier along the  $h112i/f111g$  slip direction due to a decrease in the high symmetry ISF energy as composition goes across the columns [104, 107]. Further, slip along the  $[11\bar{2}]$  direction to the low energy ISF followed by slip along  $[\bar{1}2\bar{1}]$  on  $f111g$  planes leads to a net slip along the  $[01\bar{1}]$  direction. The decomposition of  $h110i/f111g$  slip is illustrated for fcc Ni by Shang et al. using an alias shear deformation model [108].

Ma et al. have used DFT to calculate the energies along the slip and twinning paths on  $f111g$  planes for rocksalt HETMCs containing Zr, Nb, Ta, Hf, Ti, and V [109]. They report that USF energies for the single metal carbides are higher than those for compositions with mixed cations, suggesting that dislocation nucleation would be more prevalent in the latter compared to the binaries. Our calculations described below do not reproduce this result. Instead, we find that the binaries and the high-entropy compositions studied all fall along well-behaved relations regarding ISF and USF energies. In addition, the Ma et al. study did not characterize the shifts in the position of the USF away from the point halfway between the ideal system and the ISF configuration along the slip direction. As shown below, accounting for this shift is critical to accurately calculating the USF energy, and leads to a previously unreported relationship between the ISF and USF.

The intent of the research reported in this paper was to use DFT calculations to explore the degree to which ISF and USF energies along the  $h112i/f111g$  gamma surface path in stoichiometric Group IVB-VB-VIB HETMCs can be estimated from a RoM from the

Table 4.1: High-entropy carbide compositions studied for each calculation.  $h11\bar{2}/f111g$  slip path calculations consisted of 12 simulations at evenly spaced points along the gamma surface. USF calculations were performed at the predicted peak location as presented in Section 4.4.1. Compositions have been experimentally validated as single-phase [1, 2] except as marked for non-validated ( $y$ ) and multi-phase ( $z$ ) [2] compositions.

Calculation	Compositions	
ISF	(Cr,Mo,Nb,Ta,W)C	(Cr,Mo,Nb,V,W)C
	(Hf,Mo,Nb,Ta,W)C <sup>y</sup>	(Hf,Mo,Ta,W,Zr)C <sup>z</sup>
	(Hf,Mo,Ti,W,Zr)C	(Hf,Mo,V,W,Zr)C <sup>z</sup>
	(Hf,Nb,Ta,Ti,V)C	(Hf,Nb,Ta,Ti,W)C
	(Hf,Nb,Ta,Ti,Zr)C	(Hf,Ta,Ti,W,Zr)C
	(Mo,Nb,Ta,Ti,V)C <sup>y</sup>	(Mo,Nb,Ta,V,W)C
	(Mo,Nb,Ta,W,Zr)C <sup>y</sup>	(Nb,Ta,Ti,V,W)C
$h11\bar{2}/f111g$ slip path	(Hf,Nb,Ta,Ti,Zr)C	(Mo,Nb,Ta,V,W)C
USF	(Hf,Nb,Ta,Ti,V)C	(Hf,Nb,Ta,Ti,Zr)C
	(Mo,Nb,Ta,V,W)C	(Nb,Ta,Ti,V,W)C

corresponding values for the single metal rocksalt carbides. The HETMCs compositions in this study are listed in Table 4.1.

As discussed in more detail below, three general conclusions come from the research. First, based on the DFT results, a RoM with an optimized weighting scheme based on atomic layers near the stacking fault provides energies within  $\pm 0.2$  J/m<sup>2</sup> for both the average and local energies of the ISF. Second, by including Group VIB elements, where rocksalt is not the most energetically stable structure, the average ISF energy of a high-entropy composition can be negative. This is despite these HETMC compositions producing single phase materials, as predicted by entropy descriptors and confirmed by experiment [1]. A random cation arrangement leads to regions of positive and negative energies, which we evaluate statistically. Similar to other work on the effective roughness for slip in high-entropy materials [37, 110, 111], it is proposed that these statistical distributions contribute to the Peierls stress through a “mean+1” model discussed below that prevents partial dislocations from separating into extended stacking faults. Finally, it was found that the USF position is not always at the high-symmetry point halfway between the ideal system and the ISF,

but rather can exist earlier or later along the slip path. This peak shift is apparent from literature gamma surfaces for transition metal carbides, such as those presented by Yu et al. [104], but has not previously been quantified. We find that the peak shift can be defined as a monotonic function of the ISF. When the correct peak position is used, there is a linear relationship between the USF and the ISF energies that spans both net negative and net positive ISF values.

As a result of these relationships, the ISF energy can be found via a RoM, while the associated USF and the peak shift can be estimated from monotonic functions. The linear relation between the USF and ISF is analogous to the well-established Evans-Polanyi-Semenov (EPS) relation [112], where for a class of similar chemical reactions, the energy barrier is linearly proportional to the heat of reaction.

### 4.3 Calculation Details

The energies were calculated from DFT using the plane-wave projector-augmented-wave pseudopotential methods [85] in the Vienna Ab initio Simulation Package [86–88]. The exchange-correlation potential used was the Perdew et al. [89] parameterization of the generalized gradient approximation. The plane-wave basis cutoff energy was 520 eV with a  $6 \times 6 \times 1$   $\Gamma$ -centered k-point mesh. Spin polarization was enabled for all simulations to account for possible magnetism in the group VIB transition metals.

Gamma surface configurations for the rocksalt carbides were constructed as translations of an 80-atom slab with two (111) free surfaces separated by at least 12 Å of vacuum. The x and y supercell vectors were defined as the  $\frac{1}{2}[11\bar{2}]$  and  $[\bar{1}10]$  directions, respectively. For each composition, 50 random arrangements of the cation sites were generated. Relaxations were performed with respect to the atomic positions in the [111] direction, while keeping the supercell dimensions fixed along with the x and y coordinates. Lattice parameters for the high-entropy carbide compositions were determined by the average result of energy

minimization performed on 10 randomly arranged 80-atom bulk rocksalt structures with supercell vectors of  $[210]$ ,  $[\bar{1}20]$ , and  $[002]$ .

All calculations in these studies were performed with full carbon stoichiometry. Transition metal carbides are often carbon deficient, especially at elevated temperatures, and carbon vacancies can affect stacking fault energies [113]. However, vacancy effects in small supercells require additional considerations and are not captured in this work.

Gamma surface energies  $\gamma$  were calculated for each configuration and the constituent binaries using the expression

$$\gamma = \frac{E_{\gamma} - E_{slab}}{A} \tag{4.1}$$

where  $E_{\gamma}$  and  $E_{slab}$  are the energy of the translated and vacuum slab configurations, respectively, and  $A$  is the cross-sectional area of the supercell in the  $(111)$  plane. The local environment of any point on the  $(111)$  gamma surface can be defined by the translation vector between the upper and lower portions of the slab and the local atomic environment in the nearby  $(111)$  cation planes.

## 4.4 Results and Discussion

The compositions in Table 4.1 were studied to determine the relationship between ISF energy, USF energy, and the USF position along the  $h11\bar{2}i\bar{f}111g$  slip path. Because the shape of the gamma surface varies with composition at the stacking fault, a larger set of calculations was required to find the USF peak position for a given configuration. Compositions were divided into three separate simulation regimes to efficiently characterize the critical points of the slip path.

Table 4.2: Binary carbide stacking fault energies  $\gamma^{isf}$  in J/m<sup>2</sup> for the single metal rocksalt carbides used in Eq. 4.2.

	IVB		VB		VIB
Ti	2.757	V	0.453	Cr	-0.777
Zr	2.645	Nb	0.667	Mo	-1.002
Hf	2.691	Ta	0.572	W	-1.353

#### 4.4.1 Stacking Fault Energies

Stacking faults in high-entropy compounds differ from those of single-component materials in that the energy is dependent on the arrangement of nearby atoms in the structure. Similar to vacancies in high-entropy alloys and ceramics, the defect energy can be defined as a distribution of energies based on the local atomic environment [21, 57, 98], in this case defined as the composition along the interface. The energy dependence on the local arrangement of atoms requires a sample of different atomic configurations from each tested composition. The sample can then be used to fit a model and predict the total energy distribution across the configurational space, enabling analysis of the impact of defects on bulk properties and phenomena [57]. Here, DFT energies were calculated for 50 random ISF configurations of 14 HEC compositions (see Table 4.1).

ISF energies in HETMCs can be approximated by a weighted RoM of the single metal carbide stacking fault energies for atoms within 0.5 nm of the defect plane as

$$\gamma_{\omega}^{isf} = \sum_i \sum_k \frac{\omega_k n_{ik} \gamma_i^{isf}}{n_k}. \quad (4.2)$$

Here,  $n_{ik}$  are the number of atoms of each alloying element  $i$  in each layer  $k$ ,  $\gamma_i^{isf}$  are the  $f111g$  stacking fault energies in the binaries, and  $\omega_k$  are the fitted layer weights. For the 80 atom supercells in this work,  $n_k = 4$ , and the normalized weights  $\omega_{1-4}$  are found to be 0.589, 0.321, 0.045, and 0.045, respectively. While the layer weights are calculated as fitted parameters, they decrease with increasing distance from the defect, corresponding with a

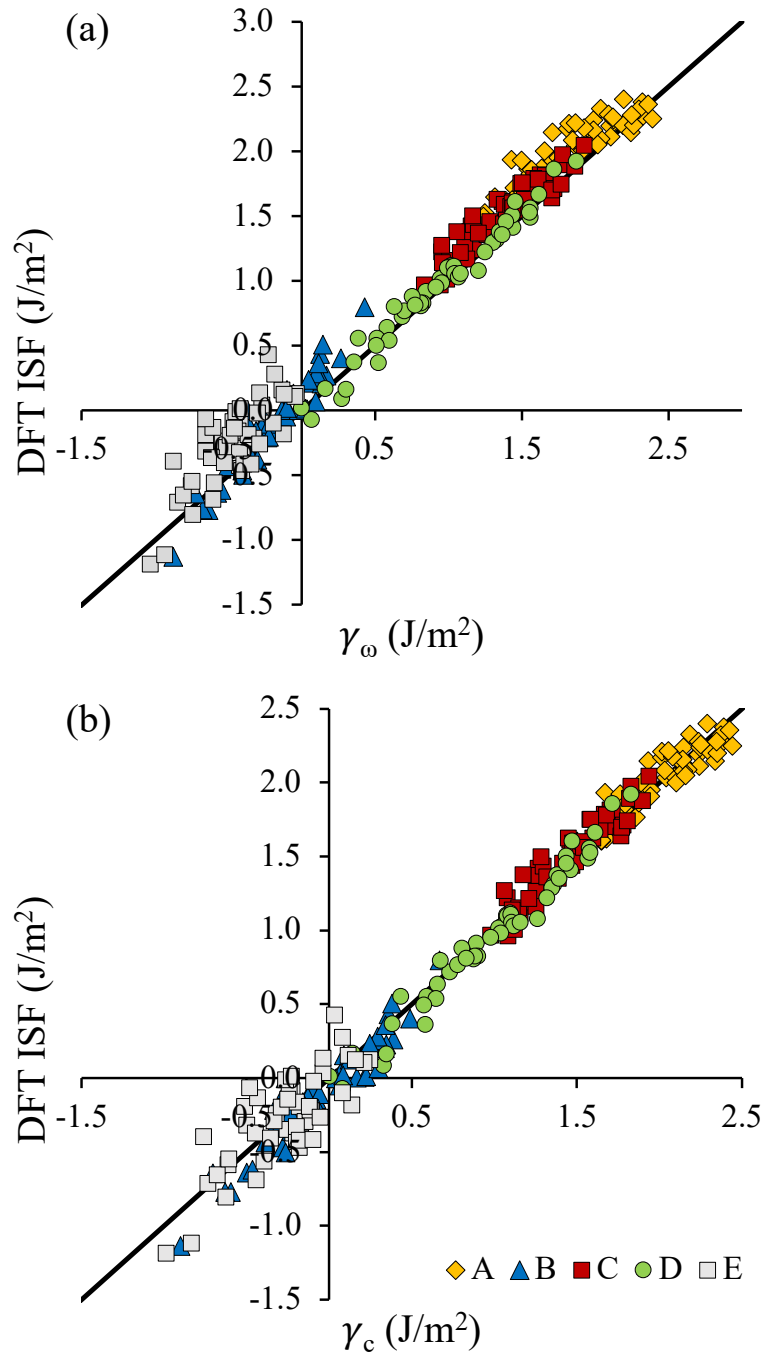


Figure 4.1: DFT calculated ISF energies as a function of RoM values using layer-dependent weighting for five compositions A:(Hf,Nb,Ta,Ti,Zr)C, B:(Mo,Nb,Ta,V,W)C, C:(Hf,Nb,Ta,Ti,V)C, D:(Hf,Nb,Ta,Ti,W)C, and E:(Cr,Mo,Nb,V,W)C. RoM fit from (a) layer weights only using Eq. 4.2 and (b) corrected RoM using Eq. 4.3.

qualitative decay function representing some unknown physical interaction. The energies of the single metal carbide stacking faults are given in Table 4.2.

Here we note that the ISF energies for group VIB metal carbides are all negative, as these compositions prefer the hexagonal WC or Mo<sub>2</sub>C structure (P $\bar{6}$ m2) to the rocksalt structure (Fm $\bar{3}$ m). In HETMCs, these elements can be present as equimolar components [1, 2], resulting in negative stacking faults for certain stacking fault configurations within the stable structure. The consequences of negative stacking faults in HETMCs are considered in Section 4.4.3.

Using the single metal carbide stacking fault energies in Table 4.2 and the given weights  $\omega$ , the optimized RoM gives a reasonable approximation of the high-entropy ISF energies, as plotted for five compositions in Fig. 4.1a. However, comparing the RoM prediction to the DFT data in Fig. 4.1, it is apparent that compositions with high and low ISF energies are underestimated by a RoM and that the distribution of energies within each composition is itself biased. This bias is accounted for by fitting two separate quadratic functions, first to the bulk RoM ISF energies  $\bar{\gamma}$  for each composition and second to the local weighted RoM  $\gamma_\omega$  as expressed in Eq. 4.2. The results are plotted for the same five compositions in Fig. 4.1b. These empirical corrections reduce the root mean squared error by 26% from 182 mJ/m<sup>2</sup> to 135 mJ/m<sup>2</sup>. The error for high [(Hf,Nb,Ta,Ti,Zr)C] and low [(Cr,Mo,Nb,V,W)C] valence electron concentration (VEC) compositions are reduced by 55% and 27% respectively. The fitted model is given as

$$\gamma_c^{isf} = (0.456\bar{\gamma} - 0.335)^2 + (0.303\gamma_\omega - 1.835)^2 + 3.393, \tag{4.3}$$

with model errors for each composition presented in Table 4.3. In all cases, the ISF energies for stacking fault configurations in HETMCs are well described as an average of the binary single metal carbide constituents near the interface.

The USF energy is another useful tool for understanding dislocation slip in fcc materials. However, because it is not necessarily at a high symmetry position, it can be expensive to

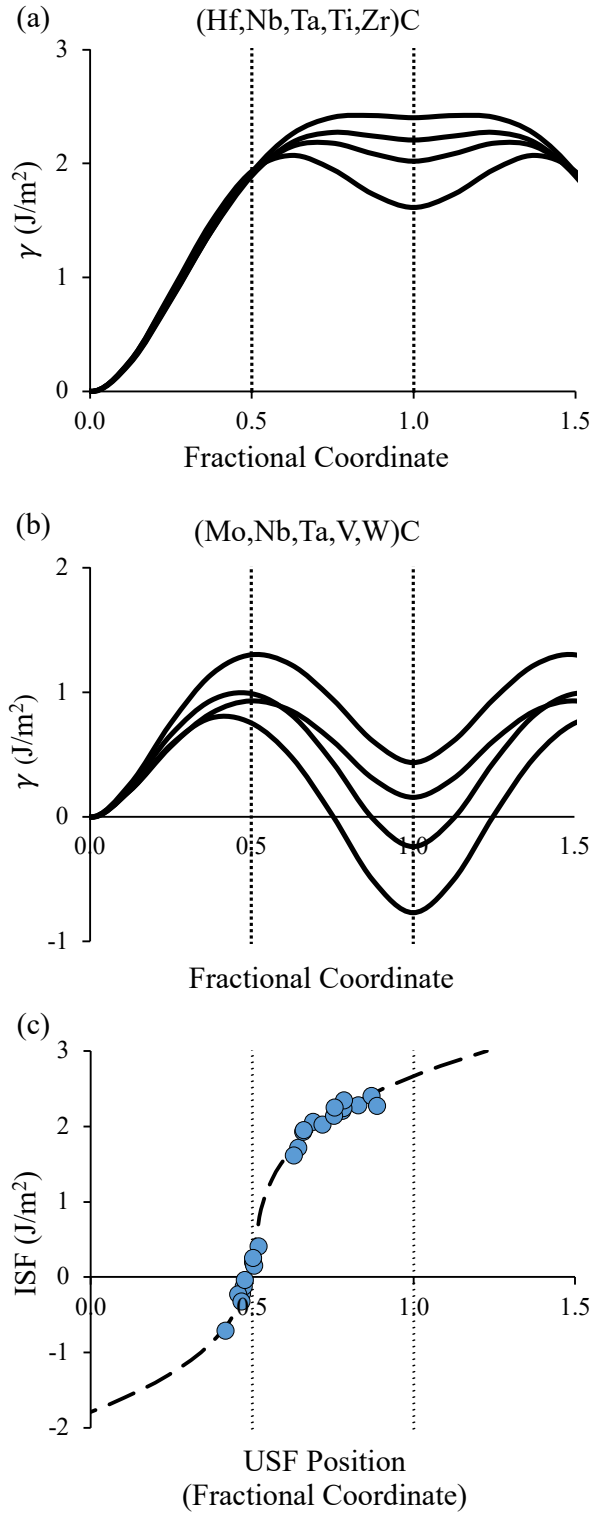


Figure 4.2: Calculated gamma surface energies along the  $\frac{1}{6}\langle h11\bar{2}i \rangle / \langle 111g \rangle$  slip direction for four atomic configurations of (a) (Hf,Nb,Ta,Ti,Zr)C and (b) (Mo,Nb,Ta,V,W)C. ISF and midpoint are indicated by the dashed vertical lines, while (c) shows the relationship between ISF and the interpolated peak position and a fitted third order polynomial for initial N-R estimate.

Table 4.3: Root mean squared fitting errors for the weighted RoM stacking fault models  $\gamma_\omega^{isf}$  and  $\gamma_c^{isf}$  (J/m<sup>2</sup>).

Composition	RMSE ( $\gamma_\omega^{isf}$ )	RMSE ( $\gamma_c^{isf}$ )
(Cr,Mo,Nb,Ta,W)C	0.294	0.220
(Cr,Mo,Nb,V,W)C	0.263	0.192
(Hf,Mo,Nb,Ta,W)C	0.164	0.121
(Hf,Mo,Ta,W,Zr)C	0.163	0.135
(Hf,Mo,Ti,W,Zr)C	0.111	0.124
(Hf,Mo,V,W,Zr)C	0.146	0.126
(Hf,Nb,Ta,Ti,V)C	0.171	0.103
(Hf,Nb,Ta,Ti,W)C	0.082	0.088
(Hf,Nb,Ta,Ti,Zr)C	0.227	0.102
(Hf,Ta,Ti,W,Zr)C	0.145	0.117
(Mo,Nb,Ta,Ti,V)C	0.146	0.151
(Mo,Nb,Ta,V,W)C	0.163	0.123
(Mo,Nb,Ta,W,Zr)C	0.170	0.113
(Nb,Ta,Ti,V,W)C	0.137	0.136

calculate in DFT compared to the ISF. This is illustrated by the energies along the  $\frac{1}{6}\langle 11\bar{2}i \rangle$  slip path for two high-entropy compositions and four arrangements per composition plotted in Fig. 4.2.

While the intrinsic stacking fault is always at the high symmetry position, the USF can vary significantly from the center of the slip path between compositions, and between arrangements within the same composition. The traditional approach is to scan the gamma surface with a grid along the dislocation vector. Because the location of the USF peak is not fixed along this vector, it is often required to use a fine grid to calculate an accurate maximum value for the energy barrier. This is especially expensive for high-entropy materials, where multiple configurations are necessary.

Here, an estimate of the USF peak location is combined with a Newton-Raphson (N-R) root-finding algorithm to predict the coordinates of the barrier on the gamma surface and minimize computation time required to simulate the relevant configurations. For the N-R method, the root of a differentiable function can be found with increasing accuracy by calculating the x-intercept of the tangent line at an initial guess and iterating until sufficient

accuracy has been obtained. In the case of the USF, we must consider the second derivative of the gamma surface to locate the local maximum of the function.

Using a coarse grid for several random configurations of two compositions (Hf,Nb,Ta,Ti,Zr)C and (Mo,Nb,Ta,V,W)C representing the range of ISF energies for the HETMCs, the shape of the gamma surface can be characterized along the  $h11\bar{2}i$  vector to identify the critical points (Fig. 4.2). The USF is near the high symmetry position at  $\frac{1}{12}[11\bar{2}]$  for (Mo,Nb,Ta,V,W)C, but is shifted toward the ISF for (Hf,Nb,Ta,Ti,Zr)C. From these two compositions, a preliminary function for peak shift is obtained as a third order polynomial with respect to the ISF and used to initiate the N-R search. Starting from this function, plotted in Fig. 4.2c, the DFT energy at the predicted peak and at one point on either side are used to find the finite differences second derivative. Extrapolating this slope to the x-intercept gives the improved estimate for the USF peak location, where the final DFT energy can be calculated. Given knowledge of the shape of the gamma surface, the first N-R iteration is sufficient despite beginning from an arbitrarily chosen polynomial fit.

Plotted in Fig. 4.3 are the USF energies calculated for the  $h11\bar{2}i/f111g$  gamma surface as a function of ISF energy. For the data plotted in Fig. 4.3a, the USF energy is calculated at the midpoint along the slip direction ( $\frac{1}{12}[11\bar{2}]$ ), while the USF energy plotted in Fig. 4.3b is the maximum along the entire slip path. There are clear linear relations in both that are discussed in the following section in terms of the EPS principle.

#### 4.4.2 Evans-Polanyi-Semenov Relations

The EPS relation (also referred to as the Bell–Evans–Polanyi or Brønsted–Evans–Polanyi principle) states that for a given set of related chemical reactions, the activation energy  $E_A$  for a chemical process is proportional to the heat of reaction  $\Delta H$  as

$$E_A = E_A^0 + \alpha\Delta H \tag{4.4}$$

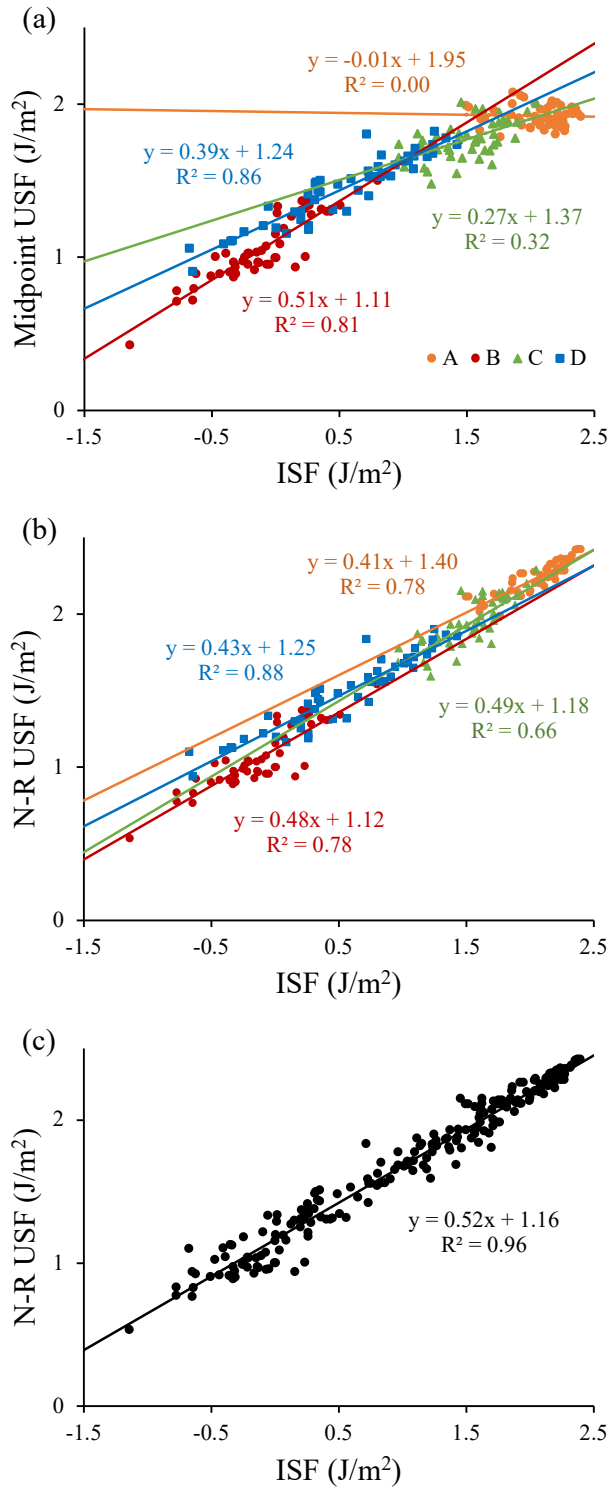


Figure 4.3: USF energy plotted as a function of the ISF energy at (a) the midpoint along the slip path, (b) and (c) at the maximum along the path as estimated by N-R. (a) and (b) are broken down by composition. A:(Hf,Nb,Ta,Ti,Zr)C, B:(Mo,Nb,Ta,V,W)C, C: (Hf,Nb,Ta,Ti,V)C, D:(Nb,Ta,Ti,V,W)C. Lines are linear fits to the data.

where  $E_A^0$  and  $\alpha$  are constants [112]. This simple relation has impacted multiple areas of chemistry, including combustion [114], organic synthesis [115, 116], photo-chemistry [117, 118], polymerization [119], interstellar chemistry [120], heterogeneous catalysis [121–124], gel degradation [125], surface diffusion [126], energetic materials sensitivity [127–129], materials discovery [130], and fundamental chemical kinetics [131–133]. Adding to this body of work, we propose the ISF and USF energies as an analogous set of quantities to the heat of reaction and activation energy, respectively.

One central premise of HETMCs has been that their properties are continuously tunable by a RoM through the IVB-VIB elements, for example as discussed in Section 4.4.1 in the context of stacking faults. VEC is another quantity calculated as a RoM from constituent elements, used both as a theoretical construct within the electronic band structure [134] and as a way of accounting for composition, including cations as well as carbon and nitrogen fraction [135]. In the context of the EPS relation, it is relevant to explore (1) whether the EPS relation is maintained for the continuous range of ISF and USF values on a gamma surface, and if so (2) whether the  $E_A^0$  and  $\alpha$  values are constant or composition dependent.

The data plotted In Fig. 4.3a, where the USF energies are calculated at the the high symmetry position, is grouped by composition with a separate linear fit for each. For compositions A and B, the  $R^2$  values indicate that the data is relatively well described by an EPS relation, but with different intercepts and slopes ( $E_A^0$  and  $\alpha$  values). The linear fit for composition C is less accurate (smaller  $R^2$  value), while there appears to be little correlation for composition D. Hence, if using the midpoint energy, the EPS relation seems to be composition dependent. However, when using the maximum energy along the slip path for the USF (Fig. 4.3b), all four data sets show a linear correlation and similar slopes and intercepts. Plotted in Fig. 4.3c is the data combined for all compositions, again using the maximum energy along the slip path. It is clear that an EPS relation exists between the ISF and USF energies, and that this relation holds across compositions when the accurate USF energy is used.

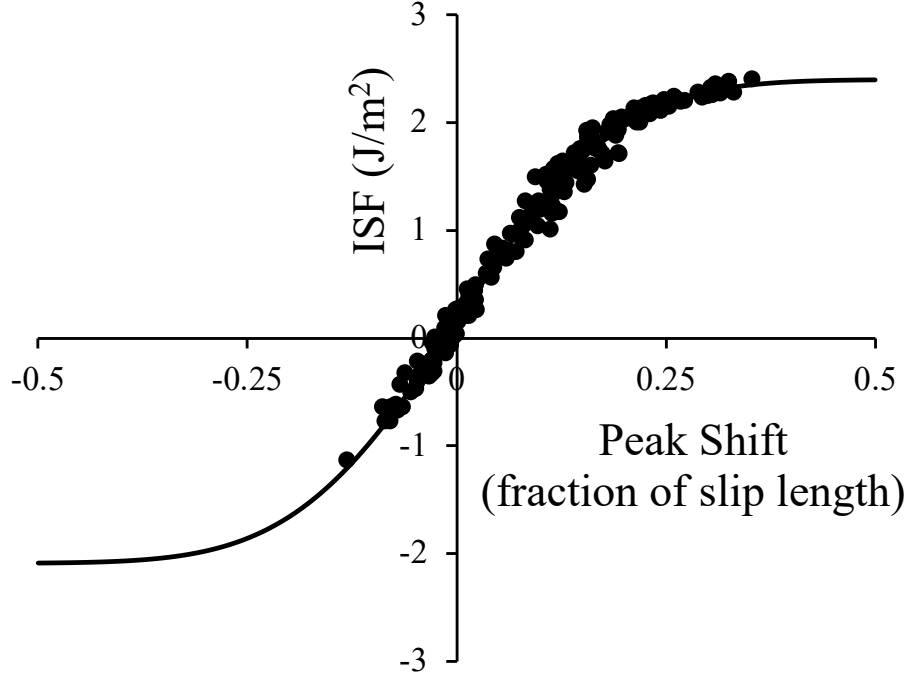


Figure 4.4: USF peak shift from Newton-Raphson search along  $h11\bar{2}i\bar{f}111g$  gamma surface. Fitting function obtained from Eq. 4.5 using data from the four compositions specified in Table 4.1.

Plotted in Fig. 4.4 are deviations of the USF position from the center point along the  $h11\bar{2}i\bar{f}111g$  gamma surface for the data in Fig. 4.3c. There is a monotonic, continuous relationship between the position of the USF (the transition state) and the ISF energy (the reaction product). The larger the ISF, the later the USF occurs along the slip direction.

Along with the data for ISF as a function of USF peak shift in Fig. 4.4, the error function

$$I(x) = B \operatorname{erf}(\alpha x) + \Delta \quad (4.5)$$

is plotted along the entire length of the slip path using fitting parameters  $B = 2.245 \text{ J/m}^2$ ,  $a = 4.605$ , and  $\Delta = 0.559 \text{ J/m}^2$ . Although the physical interpretation of an error function is unclear, it provides a good quantitative fit to the data (a logistic function provides a similar fit). In particular, the asymptotic behavior of the error function is consistent with the range of possible positions of the USF along the slip path, existing only on the region between the

ideal structure and the ISF.

This function enables the evaluation of critical points on the gamma surface while minimizing the number of necessary calculations along the slip path. Given an ISF energy, which is well approximated by a weighted RoM, one can identify the position of the USF and perform a single DFT calculation without the need for an iterative search algorithm along the slip path.

The relationships in Fig. 4.3 and Fig. 4.4 can be understood using an energy curve overlap model [112, 136], illustrated in Fig. 4.5. In this model there are two curves that correspond to local energy minima on the gamma surface along the  $[11\bar{2}]$  sliding direction. One curve has a minimum at the ideal structure, while the other has a minimum at the ISF, shifted uniformly in energy according to the ISF energy. The energy along the entire path within this model is the minimum of the two curves at any given point. Fig. 4.5 illustrates the model for the three cases of  $ISF > 0$ ,  $ISF = 0$ , and  $ISF < 0$ , with the barrier corresponding to the energy where the two curves intersect (the maximum in the solid curve). In the case of  $ISF > 0$ , the position of the barrier along the sliding path is later than then center position. Consistent with the simulation data, as the ISF energy decreases the value and position of the USF decreases along the slip direction.

The energy curves in Fig. 4.6 are a two-parameter fit assuming Gaussian functions of the form

$$E(x) = \frac{A}{\sigma} \frac{1}{2\pi} \left[ 1 - \exp \left[ \frac{1}{2} \left( \frac{x - x_0}{\sigma} \right)^2 \right] \right] \quad (4.6)$$

for the two overlapping energy curves. Here  $x_0 = 0$  for the curve starting at the ideal structure,  $x_0 = 1$  for the curve starting at the ISF,  $\sigma = 0.7$  and  $A = 9.15 \text{ J/m}^2$ . The simple model is able to qualitatively describe both the relations between USF and ISF, and between ISF and the shift in USF position along the  $[11\bar{2}]$  direction.

Similar fits to the data can be achieved with other potential energy approximations

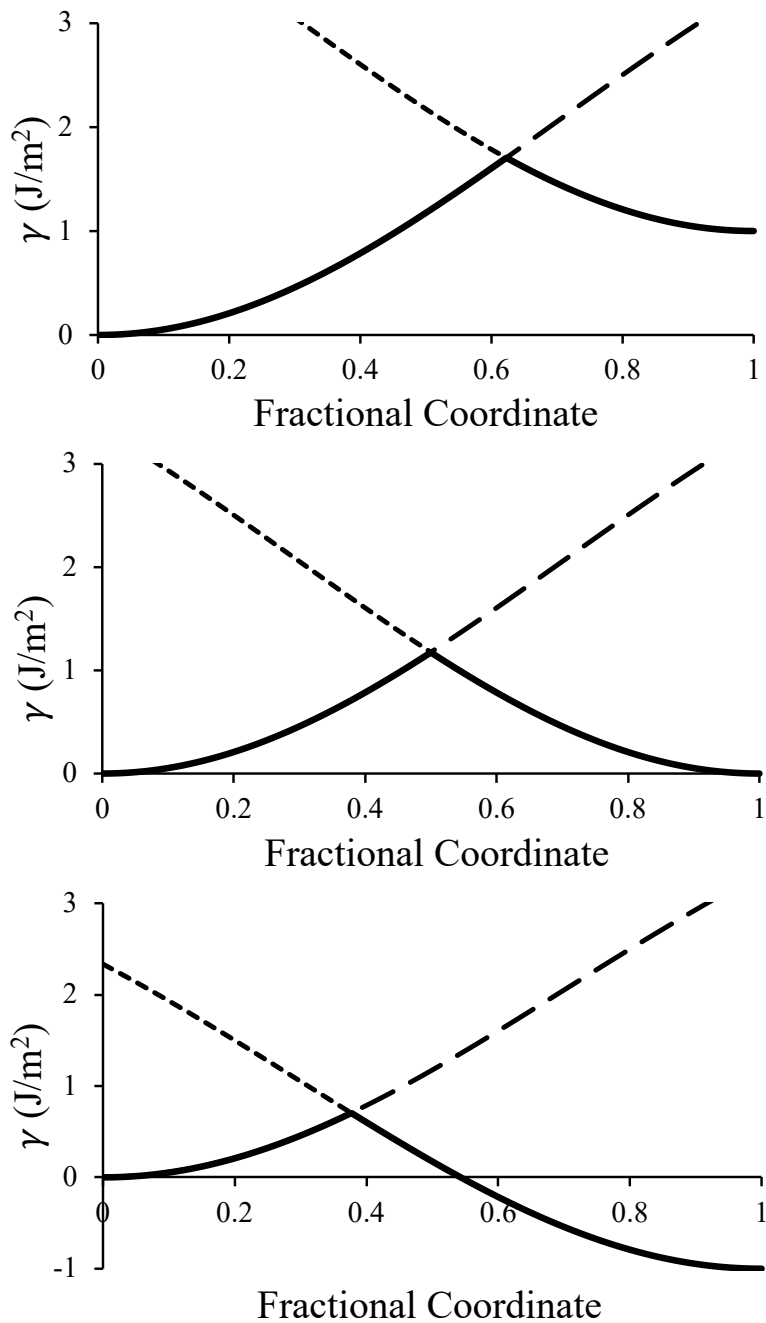


Figure 4.5: Shifted energy curve model using Eq. 4.6 plotted along the  $\frac{1}{6}\langle 11\bar{2} \rangle / \langle 111 \rangle$  slip direction for (top) positive, (middle) zero, and (bottom) negative ISF.

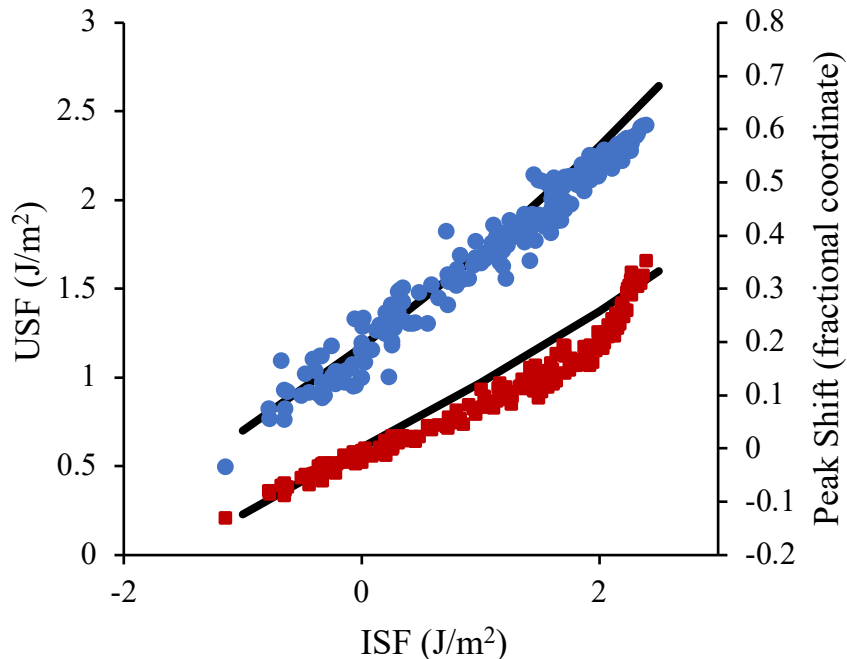


Figure 4.6: Maximum USF energies (blue circles) and deviations from the center point along the  $h11\bar{2}i/f111g$  gamma surface where these energies are found (red squares) for data plotted in Fig. 4.3. The solid lines are two parameter fits for an energy curve overlap model using Eq. 4.6.

such as quadratic energy wells, sine fitting, or a Lennard-Jones potential. None of the approximations tested, however, have a clear physical interpretation or other advantages over the two-parameter Gaussian fit.

### 4.4.3 Shockley Partial Dislocation Separations

Pairs of Shockley partial dislocations in an fcc crystal are connected by a stacking fault such that the equilibrium distance between partials is determined by a balance of the elastic repulsion and the stacking fault energy. One of the interesting issues in high-entropy materials is that they can have negative stacking fault energies while still being phase stable [111, 137]. In non-random crystals, negative stacking fault energies suggest that any Shockley partial dislocations that form should move as far apart as possible, because both the elastic repulsion and the stacking fault formation energy would favor infinite distances. However, large numbers

of stacking faults are not always observed experimentally. As an example in a HETMC, the simulations in Section 4.4.1 indicate that (Mo,Nb,Ta,V,W)C has an average stacking fault energy of  $-0.09 \text{ J/m}^2$  such that the elastic repulsion between any Shockley partials should create a significant stacking fault area. Experimentally, however, this composition exhibits sharp x-ray diffraction peaks that suggest a low defect density [2].

One suggestion to account for the apparent lack of stacking faults associated with Shockley partials is that “roughness” from the atomic disorder in high-entropy materials inhibits the ability of the partials to fully separate [37, 110]. While the relation between Peierls stress and stacking fault energy is unclear [137], it is reasonable to assume that this roughness is qualitatively related to the variations in stacking fault energies from a random composition along the interface as the partial dislocations attempt to separate.

This concept is illustrated in Fig. 4.7, which shows a schematic of two Shockley partial edge dislocations separated by a stacking fault along the blue plane. The peaks on the energy landscape represent a measure of the stacking fault energy contributions to the Peierls stress. A negative stacking fault energy should lower the Peierls stress, while a positive stacking fault energy (denoted by the orange peaks) should increase the Peierls stress. Because the extra half plane exists along the length of an edge dislocation, the positive regions will pin the plane at different points despite the negative stacking fault regions exhibiting a lower local Peierls barrier. Taken to an extreme, even when the net stacking fault energy is negative, positive regions would prevent the Shockley dislocations from moving.

In a set of atomic simulations by Shih et al. [111] that modeled random NiCo alloys with dissociated Shockley partial edge dislocations, the partials remained at finite separations that were less than the distance at which the mean plus one standard deviation ( $\mu + \sigma$ ) of the stacking fault energy distribution equaled the repulsive elastic stress. For this mean + 1 model, we generalize the trend for partial edge dislocation separation length  $s$  with the

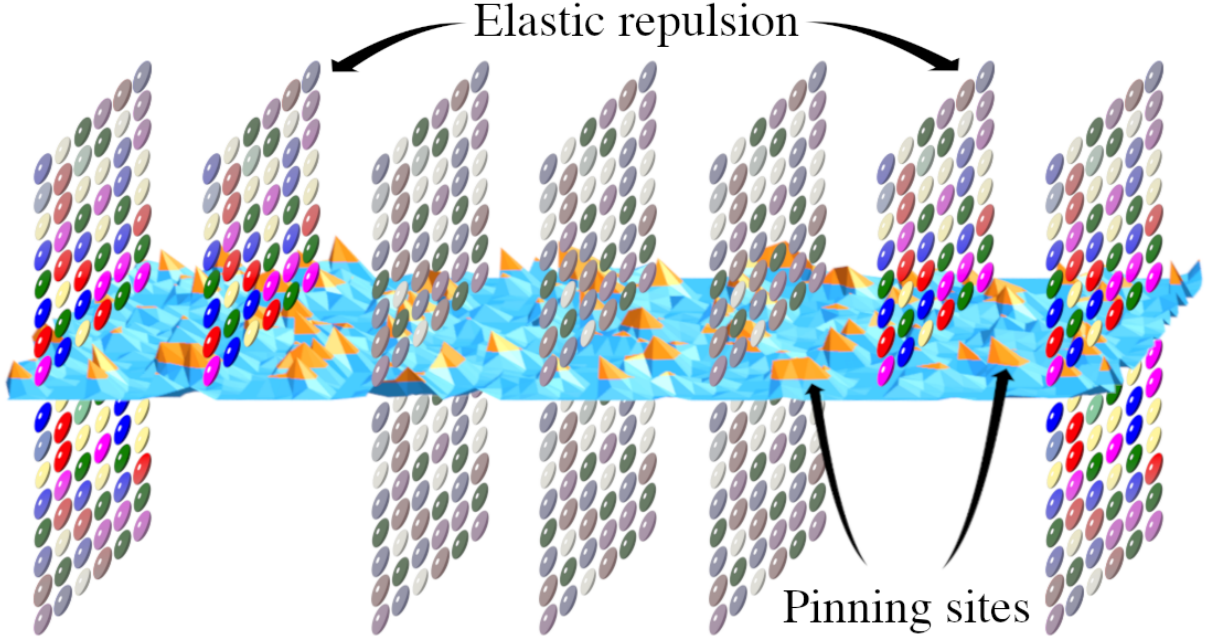


Figure 4.7: Depiction of two Shockley partial dislocations separated by a stacking fault along the horizontal plane. Brightly colored atoms at the edge dislocations correspond to the local atomic environment that contributes to the Peierls stress, with blue regions representing negative stacking fault regions and orange peaks on the energy landscape representing the high energy pinning sites preventing partial dislocation separation.

expression

$$s < \frac{Gb^2}{2\pi(\mu + \sigma)} \quad (4.7)$$

where  $G$  is the shear modulus,  $b$  is the Burger's vector, and  $\mu + \sigma$  is again the sum of mean and standard deviation of the stacking fault energy distribution.

Sampling from a random distribution of interface cation arrangements, a predicted distribution of ISF energies can be generated for each composition. Critically, the distribution is dependent on the volume that is considered to be part of the local environment. Taken to both extremes for the five cation system, the distribution of energies for a single-atom volume will consist of five peaks corresponding to the five constituent elements, each accounting for 20% of the total distribution. Similarly, as the volume of the local environment increases to

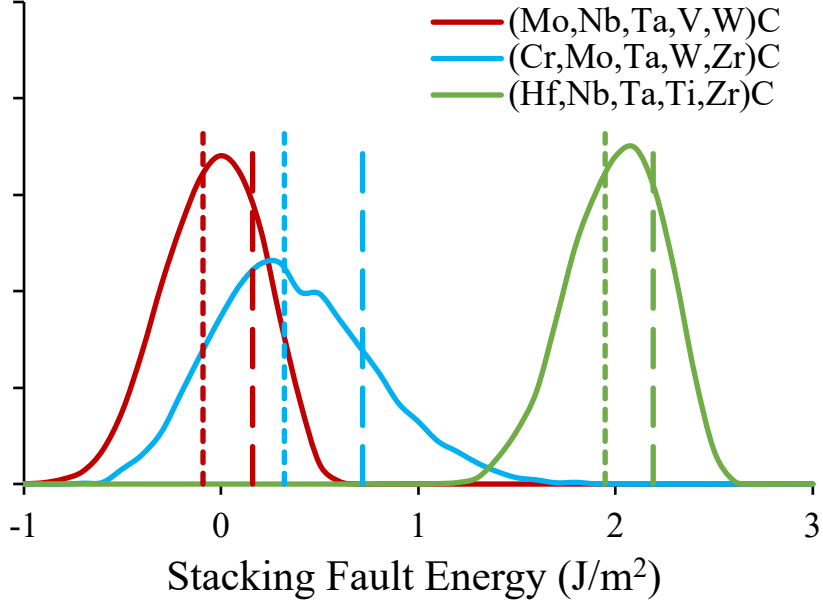


Figure 4.8: ISF distributions for three compositions. The dashed vertical lines represent the mean and the mean plus one standard deviation for each distribution.

infinity, the distribution will approach the delta function centered on the distribution mean.

The volume of the local area can be defined by sampling for each element from a gamma distribution such that

$$X_i \sim \Gamma(\alpha = \frac{n}{N}, \beta = 1) \quad (4.8)$$

where  $\alpha$  and  $\beta$  are the  $\Gamma$  shape and rate parameters,  $n$  is the number of atoms in the local environment,  $N$  is the number of distinct elements on the lattice site.

In this work, the depth of the local environment was determined from the fitting in Section 4.4.1. Sampling each (111) plane separately, the distribution of interface cation arrangements is built from  $X_i \sim \Gamma(1, 1)$ , equating to a naive assumption of  $n = N$ . Without knowing the shape or size of the critical area, this assumption corresponds with a cutoff distance on the order of  $2b$ .

The predicted ISF distributions derived from this  $\Gamma$  sampling and the modified RoM model in Section 4.4.1 are plotted in Fig. 4.8 for three different compositions. The mean

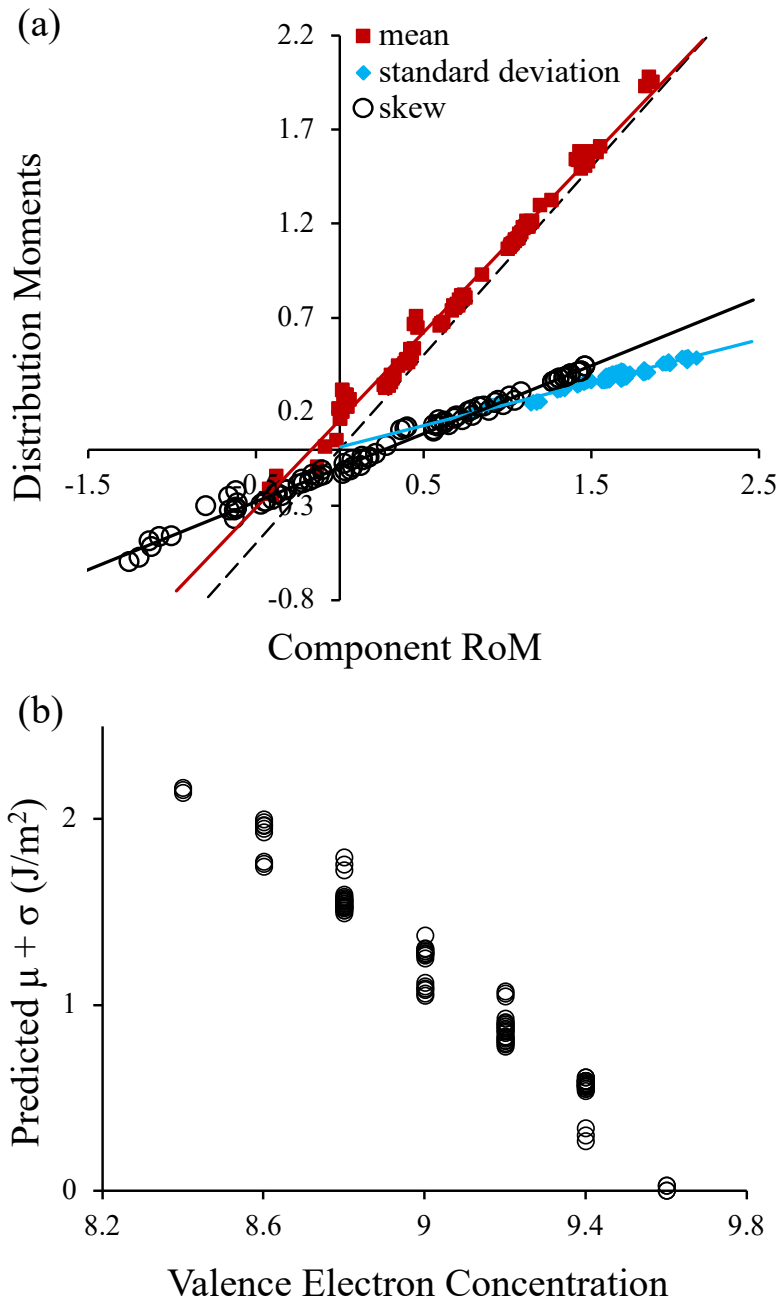


Figure 4.9: For 94 HETMC compositions, (a) mean, standard deviation, and skew of the predicted ISF energy distributions as a function of the same quantities calculated from the five corresponding single-metal carbide constituents and (b) the sum of first two moments as a function of VEC.

Table 4.4: Linear coefficients  $\alpha$  (slope) and  $\beta$  (intercept) for the first three moments of the predicted ISF distribution along with fitting error ( $\text{J}/\text{m}^2$ ) for 94 HETMC compositions in Fig. 4.9.

	$\alpha$	$\beta$	RMSE
Mean	0.932	0.156	0.058
Standard deviation	0.228	0.015	0.015
Skew	0.361	0.096	0.026

and mean plus one standard deviation for each distribution are indicated by the dashed vertical lines. For the (Mo,Nb,Ta,V,W)C composition, the mean stacking fault energy and standard deviation are  $-0.09 \text{ J}/\text{m}^2$  and  $0.25 \text{ J}/\text{m}^2$ , respectively. Using  $G = 193 \text{ GPa}$  and  $b = 0.437 \text{ nm}$  [138] in Eq. 4.7 yields a Shockley partial separation  $s$  of less than 18 nm. Hence, although 64% of the stacking fault distribution is negative, the model predicts a finite partial separation.

Plotted in Fig. 4.9 are the mean, standard deviation and skew of the stacking fault energy distributions for each HETMC composition calculated from a random sample of configurations as a function of the same quantities calculated from just the five constituent single metal carbide ISFs. There are strong linear relations such that the first three moments of the full ISF distributions for a given composition can be estimated from only the binaries, adding another facet to the RoM. The linear coefficients are given in Table 4.4.

The mean + 1 stacking fault energies, as calculated from the predicted distribution moments, are plotted as a function of their VEC in Fig. 4.9b for 94 HETMC compositions consisting of five cation combinations from Cr, Hf, Mo, Nb, Ta, Ti, V, W, and Zr. There is a monotonic decrease in energy, suggesting that the separation between Shockley partial dislocations increases with increasing VEC. However, because the mean + 1 model estimates a lower bound, partial dislocation separation distances are expected to be finite for all VEC values 9.6.

## 4.5 Conclusion

First principles calculations were used to establish new RoMs for interfacial defects in group IVB, VB, and VIB stoichiometric transition metal carbides. Energies of  $\bar{f}111g$  ISFs were shown to be predictable from an optimized RoM based on values from the respective single metal carbide compositions. A linear relationship was discovered between these stacking fault energies and the USF energies along the  $h11\bar{2}/\bar{f}111g$  slip path on the gamma surface. Provided that the positional shift of the USF from the high symmetry position is accounted for, the linear coefficient is independent of composition. The ISF and USF energies are analogous to the heat of reaction and transition state barrier in chemical reactions, and this linear relationship is similarly analogous to the EPS relationship that is ubiquitous across chemical systems, but not applied to plasticity in solids.

It was further found that the mean, standard deviation, and skew for the distribution of stacking fault energies that result from disorder of cations around the boundary are predictable from linear relations that connect the stacking fault energies of the binaries to the corresponding high-entropy compositions. Using a model where the upper tail of the ISF distribution balances the elastic repulsion between partial dislocations, we demonstrate why Shockley partial edge dislocations remain bound for HETMC compositions with VEC

9.6 despite certain compositions exhibiting negative mean ISF energies. The mean plus one standard deviation is proposed as an upper bound for determining the partial separation. This is consistent with experiment, where compositions predicted to have negative stacking fault energies such as (Mo,Nb,Ta,V,W)C are phase stable with relatively low defect densities [2].

The RoMs and linear relations discovered in this study are important for understanding and predicting the defect structure and mechanical properties of HETMCs. In addition, the EPS relation has been an invaluable tool for estimating reaction kinetics across diverse areas of chemistry, and based on our results we anticipate that it will find other applications in predicting mechanical properties of high-entropy materials. Future work on this topic should include vacancy effects on gamma surface energies, and may require more sophisticated

machine learning techniques for quantitative predictability [98, 139].

## 4.6 Acknowledgments

We would like to gratefully acknowledge the DoD SPICES MURI sponsored by the Office of Naval Research (Naval Research contract N00014-21-1-2515) and for their financial support of this work, as well as the Office of Naval Research for funding through grant N00014-23-1-2758. We also acknowledge the computing resources provided by North Carolina State University High Performance Computing Services Core Facility (RRID:SCR\_022168).

## CHAPTER

# 5

# ORDERING PHENOMENA IN HIGH-ENTROPY CERAMICS

The previous chapters have demonstrated that pair interaction models are an invaluable tool for predicting defect energies and characterizing the local atomic environment in high-entropy materials. This usefulness goes beyond defect energies, as was briefly discussed for pair affinities in the bulk high-entropy diboride in Section 3.4.3. In that case, a model trained on and optimized for vacancy configurations yielded insights into the relative enthalpic favorability of neighbor pairs for all shells with a bonded interaction term. The model for vacancies was robust to overfitting and required a minimal training set as compared to the more traditional cluster expansion method. As such, further analysis of this parameterization was performed to explore the bulk interaction energies in high-entropy ceramics.

## 5.1 Parameter Selection

There are several methods commonly employed in modeling order-disorder phenomena in solid solutions. The cluster expansion represents an exhaustive set of parameters to describe the pairs, triplets, and n-tuples that exist in a disordered material [81]. This method works across composition gradients and has been proven to capture accurate configurational enthalpy landscapes in binary systems, but scales poorly beyond two-component solid solutions. When applied to high-entropy systems, the parameter space can include as many as 1730 features for a six-component alloy. Work by Nataraj et al. [140] demonstrated that this vast parameter space can be reduced by careful selection of cluster interactions in the model, yielding parameter sets of 292, 230, and 562 features for LOOCV optimized models of NbTiVZr, HfNbTaTiZr, and AlHfNbTaTiZr respectively. The resultant models describe the dataset well, but require large training sets and notably omit highly strained configurations of the high-entropy lattice, accounting for nearly half of the configurations generated for the six-component AlHfNbTaTiZr.

Other efforts to model ordering in high- or medium-entropy alloys have used the embedded atom method (EAM) to fit empirical molecular dynamics potentials. A study by Li et al. [141] trained an EAM potential for the ternary NiCoCr system to model ordering and its effects on dislocation motion, using over 3,000 atomic configurations for the training dataset. Separately, an EAM potential fit with a machine learning protocol was used by Xu et al. [142] to investigate ordering and vacancy mediated diffusion in the ternary NiFeCr alloy, requiring 72,000 samples for training.

In both of these methods, large training sets are necessary for fitting the model and the sample configuration supercells must also be large enough to capture order and disorder on the lattice at various length scales. The consequence of minimum size constraints is compounded in the case of high-entropy ceramics as the anion sublattice doubles the number of atoms in the supercell for rocksalt carbides or nitrides and triples it for the hexagonal transition-metal diborides.

A less computationally expensive model has been proposed in several recent publications by Zhang, Liu, and their collaborators [143, 144], consisting only of pair interactions and fit to large supercell local density approximations of the electronic structure using the locally self-consistent multiple scattering method [145]. The model was used to explore short-range order in the four-component NbMoTaW alloy as well as two HEAs, NbMoTaWV and NbMoTaTiW, achieving a good model fit with as few as 100 configurations in the training dataset. This effective pair interaction model is similar to the reduced pair approximation developed in Chapter 3 for vacancies in (Hf,Nb,Ta,Ti,Zr)B<sub>2</sub>, and the work demonstrates that the linear combination of pairwise interactions can also be sufficient to describe the bulk enthalpies of different lattice decorations in a many-component disordered solid solution.

Driven by the insights from the high-entropy diboride system and the early successes of bulk energy pair models in literature, the following sections will explore the parameterization of pair models and possible applications in the simulation of ordering phenomena.

### 5.1.1 Stochastic Stepwise Subset Optimization

The reduced pair approximation model is derived in Chapter 3 and given for a five component system in Eq. 3.11 as

$$E^{vf} = \sum_{j=1}^5 \left( \sum_{k \in B} n_{jk} E_{ijk}^b + \sum_{k \in R} n_{jk} E_{ijk}^r \right) + E_i^r \quad (5.1)$$

for bidirectional  $ij$  pair interactions  $E_{ijk}^b$  in neighbor shells  $k \in B$  and unidirectional vacancy interactions  $E_{ijk}^r$  in neighbor shells  $k \in R$ . The assumptions in the derivation reduce the number of features included in a pair model from 145 to 53-73 (depending on the chosen training optimization) while maintaining a physical interpretation of the interaction parameters. The paradigm takes advantage of grouped sets of parameters in each neighbor shell to establish a tractable parameter space which can be mapped in its entirety to find the optimal parameter subset, whether by BIC or any other optimization criteria.

Despite its convenience, the constraint of grouped interaction parameters is imposed arbitrarily, because a bidirectional interaction of an A-B pair in shell  $k$  does not imply an A-C interaction. Each parameter in the model quantifies the contribution to the total energy from a given pair and while it is reasonable to expect a radial decay for such interactions, the rate of decay may be element dependent. Eliminating the grouping constraint results in a parameter space with possible subsets  $\theta$  scaling with the total available interaction parameters as  $2^N - 1$ . For  $k = 7$  in a five-component system, the 145 available unconstrained features yield a parameter space on the order of  $10^{43}$ . This can be reduced by, for example, imposing all first shell neighbors as bidirectional, but this would not significantly reduce the cost of an exhaustive parameter space search. The removal of this constraint increases the complexity of the parameter space, but allows for new avenues of parameter selection to further optimize the model.

One strategy for finding the optimal parameter subset is to use a stepwise parameter search [146], wherein parameters are chosen one at a time based on which available parameter optimizes the search criteria, e.g., BIC. This can be done forward (starting with zero parameters and adding one at each step), backward (starting with all available parameters and removing one at each step), or both (a parameter may be either added or removed at each step). Two problems arise from this approach. First, while the stepwise search is typically powerful for identifying influential features, it can be susceptible to outliers forcing the selecting of non-influential features as well [147]. Excess parameters introduced in this way may reduce the interpretability of the model by erroneously implying the existence of physical interactions. Second, due to the relationship of each parameter with the mean-field, the parameter selected at each step will be highly dependent on the parameters currently included in the model. This results in a final parameter set that is dependent on the starting point of the search, again reducing the interpretability of the model.

To solve the second problem, the stepwise parameter search is performed inside of a stochastic loop such that after optimization of  $\theta$ , a fraction of the selected parameters are

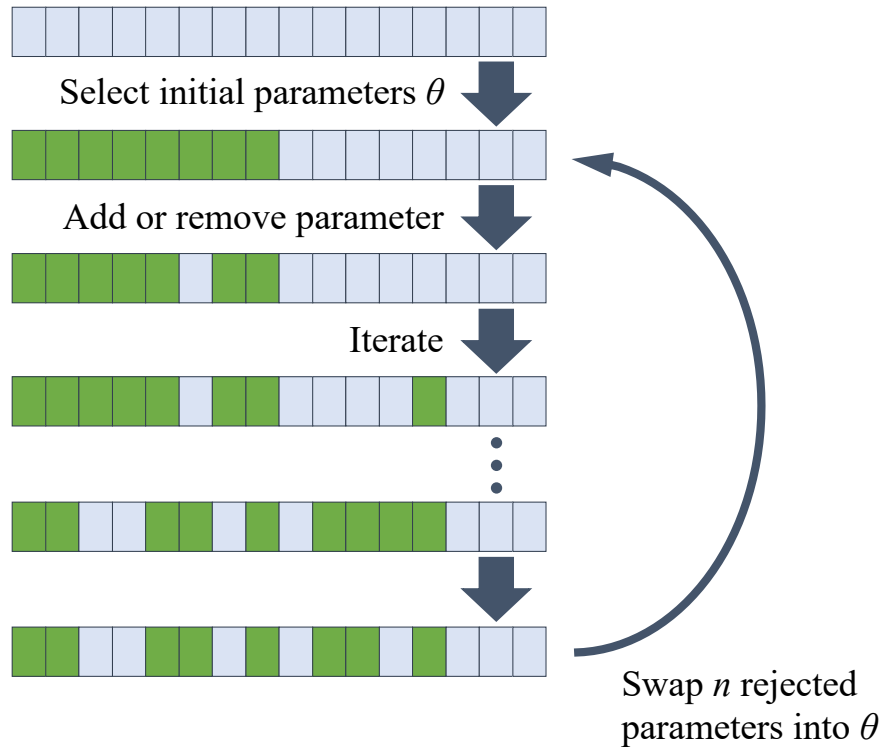


Figure 5.1: Schematic of the stochastic stepwise search algorithm for optimized parameter subset selection.

discarded and replaced with an equal sized set of rejected parameters, followed by a re-initialization of the stepwise search. The stochastic stepwise parameter search is illustrated in Fig. 5.1. This algorithm reduces the likelihood of choosing a poorly optimized parameterization as a result of the chosen starting point and consistently identifies models which improve on the predictive performance of the RPA model. Table 5.1 and Fig. 5.2 compare a 47 parameter model produced by stochastic stepwise optimization to the 63 parameter model ( $B = 1, 2, 3, 4; R = 5, 6, 7$ ), and a short range unconstrained pair model ( $U = 1, 2, 3$ ). The stochastic stepwise search is able to reduce the parameter subset of the previously optimized 63 parameter model to 40-50 features and reduce the LOOCV score by 20%. While this method does not produce a unique solution, it indicates that further reductions to the parameter subset can improve the predictive power of a pair interaction model.

Table 5.1: Feature selection for the RPA vacancy pair model for (Hf,Nb,Ta,Ti,Zr)B<sub>2</sub>.

Parameters	CV score (meV)	BIC	Assumptions
75	40.0	-609.8	all bonds (k=3)
63	22.3	-878.6	grouped (k=7)
47	17.9	-1016.2	stepwise optimized (k=7)

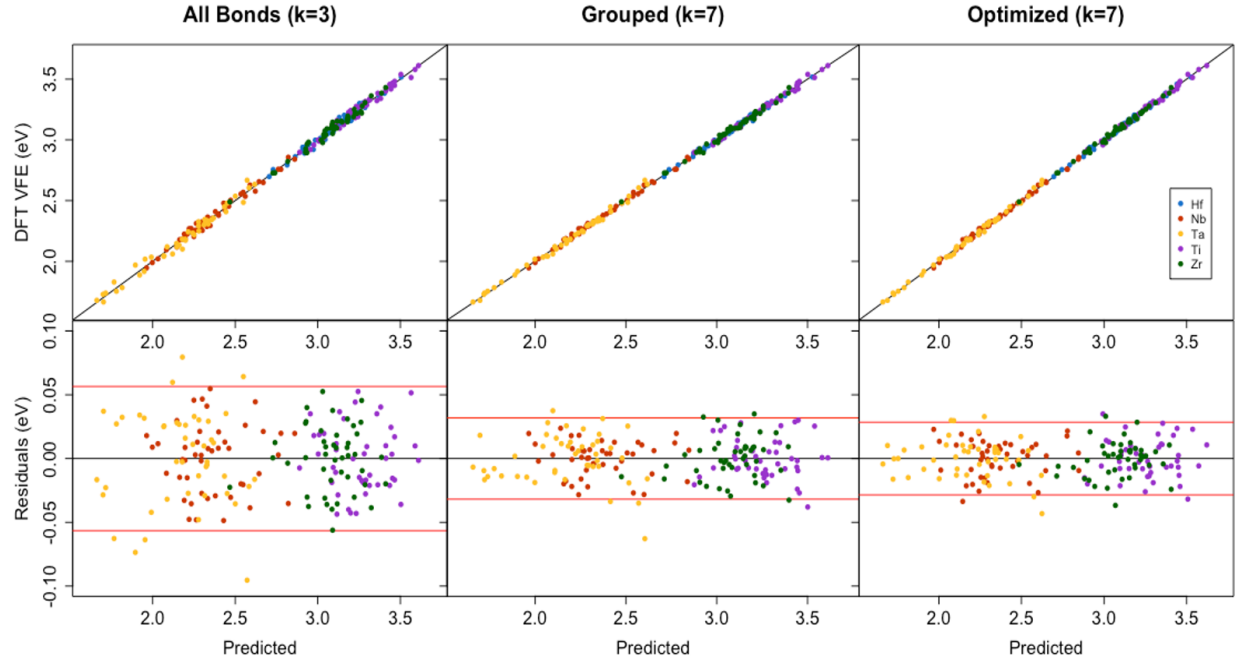


Figure 5.2: DFT vacancy energies and residual errors plotted against the predicted energy for pair models parameterized as all bonds for three neighbor shells (left), grouped and symmetry constrained bonds as in Ch. 3 for seven neighbor shells (middle), and stochastic stepwise BIC optimization (right).

## 5.2 Bulk Energy Pair Interaction Models

To evaluate pair interaction models for bulk energy predictions, the high-entropy carbide (Hf,Nb,Ta,Ti,Zr)C was simulated by DFT in a varied set of small- to medium-sized supercells. A study by Liu et al. [148] showed that ordering artifacts across periodic boundaries in small cell simulations of high-entropy alloys can bias a pair model and lead to inaccurate predictions of the total energy. It was further demonstrated that ordering across periodic boundaries is easily quantifiable in the context of SRO using the Warren-Cowley SRO parameters, and that

a composite training set using varied supercell sizes and geometries eliminates the prediction biases.

The set of disordered supercells in the initial training set for this work was comprised of 17 supercell geometries with between 20 and 100 atoms. The small cell calculations provide an inexpensive set of disordered structures for an initial fit while capturing a degree of SRO. Additional 240-atom large cell configurations were generated and added to the training set through an iterative Monte Carlo process following the approach used by Liu [144]. Generating supercell configurations in this way allows for the inclusion of structures that diverge from ideal random arrangements and capture the lower tail of the configurational energy distribution. The complete set of supercells used for training the bulk energy pair models is presented in Table 5.2.

The pair model for the bulk energy with fixed composition is simply given as the sum of all atomic pairs

$$E = \sum_{i,j,k} n_{ijk} E_{ijk}^{\text{pair}} \quad (5.2)$$

for all  $ij$  pairs in neighbor shell  $k$  surrounding a lattice site. Noting that  $ij$  pairs are symmetric, there exist 15 distinct pairs in each neighbor shell for a 5-component system, scaling as  $N(N + 1)/2$  for  $N$  components. However, constraining the system to a fixed composition reduces the linear degrees of freedom of the parameters by  $N - 1$ . Specifically, one can sum

Table 5.2: Supercell configurations for the bulk energy pair model training set.

Cell size (atoms)	Geometries	Total Configurations
20	4	200
25	1	50
30	2	100
40	4	200
60	3	90
80	2	100
100	1	50
240	1	140

all  $ijk$  self-interaction terms to a single composition-dependent constant, and consider only with the pair fractions  $c_{ijk}$  of the remaining species. Stated as the energy per atom,

$$E/\text{atom} = \sum_{j>i,k} c_{ijk} E_{ijk}^{\text{pair}}. \quad (5.3)$$

For the rocksalt carbide, an 8 Å cutoff radius was imposed for pair terms, yielding 6 neighbor interaction shells and 61 parameters for the simple linear regression case.

Looking to reduce the number of parameters further, three regression techniques are implemented for the dataset: lasso regression, Bayesian ridge regression, and automatic relevance detection (ARD). Each technique attempts to reduce overfitting by penalizing large parameters and, in the case of lasso and ARD, eliminating parameters that do not meaningfully contribute to the model fit.

### 5.2.1 Pair Model Regressors

Lasso regression (least absolute shrinkage and selection operator) [149] is a parameter optimization technique that directly penalizes each parameter in a linear regression proportionally with its magnitude. The lasso penalty utilizes the  $L1$  norm (that is, the absolute value of the parameters) in the optimization function as

$$\min_{\beta} \left[ (Y - X\beta)^t (Y - X\beta) + \lambda \sum_i j\beta_j \right], \quad (5.4)$$

where  $Y$  are the observed values,  $X$  are the input variables, and  $\beta$  are the model parameters.  $\lambda$  is a tunable hyperparameter specifying the degree of regularization in the regression. The primary advantage of lasso regression over a similar technique in ridge regression (penalizing the loss function with the  $L2$  norm  $\beta^2$ ) is that the linear response of the  $L1$  norm  $j\beta_j$  with parameter magnitude acts as a form of parameter subset selection and can lead to sparse parameterizations. For this work, the hyperparameter  $\lambda$  is tuned using the Bayesian

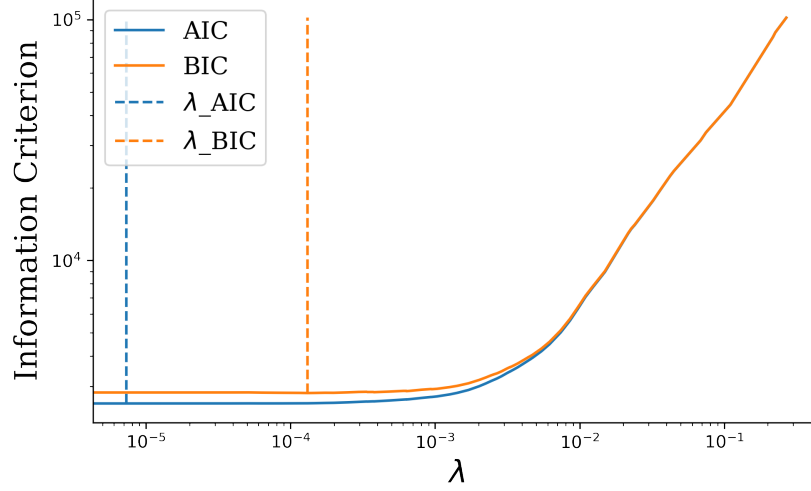


Figure 5.3: Tuning of the hyperparameter  $\lambda$  for lasso regression using the Akaike and Bayesian information criteria.

information criterion (BIC), as defined in Eq. 3.12, to further reward the elimination of parameters and find the smallest parameter set  $\beta$ . The BIC along with another similar subset selection metric in the Akaike information criterion

$$AIC = 2k_M - 2 \ln \hat{L} \quad (5.5)$$

are shown as a function of  $\lambda$  in Fig. 5.3.

Bayesian ridge regression [150] is another variation of linear regression that seeks to limit overfitting in a model by imposing a penalty for the magnitude of each parameter. Deriving the  $L_2$  regularization from a Bayesian framework allows the regression parameters to be treated as random variables, yielding parameter uncertainties in addition to constraining the parameter values. A similar form of Bayesian regression for a cluster expansion with a physics based prior was proposed by Mueller and Ceder [151], and has seen use in HEA ordering simulations [143, 144]. Starting from Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5.6)$$

where  $P(A|B)$  is the probability of  $A$  given  $B$ , the maximum likelihood estimate is derived for a pair model [151] as

$$\hat{V} = (X_w^\top X_w + \Lambda)^{-1} X_w^\top y_w. \quad (5.7)$$

Here  $X_w$  and  $y_w$  are the weighted input and output vectors for the model, and  $\Lambda$  is the regularization matrix which is inversely related to the covariance matrix for the pair interactions in the model. The regularization serves to decrease non-influential parameters, but does not eliminate them. Instead, the minimization of interaction parameters serves to center the parameters around a mean-field interaction such that each interaction is quantified as either more or less favorable than the average pair interaction on the lattice.

ARD regression, also called sparse Bayesian learning or relevance vector machine, is similar to Bayesian ridge regression, but replaces the spherical Gaussian prior with an elliptic Gaussian distribution such that each model parameter has its own hyperparameter within the prior [150]. This formulation is more flexible in its treatment of the parameter space and can more readily produce sparse parameterizations.

These three regularized linear regression algorithms, as implemented in the Scikit-learn Python package [150, 152], are compared with ordinary least squares linear regression for the pair interaction terms for (Hf,Nb,Ta,Ti,Zr)C in Fig. 5.4 and Table 5.3. ARD and lasso regression result in sparse parameter sets and the lowest BIC scores, however lasso regression sacrifices accuracy as a result of the reduced parameterization. Bayesian ridge and ordinary least squares perform very similarly, with both exhibiting the lowest training error, LOOCV, and 5-fold CV scores. Despite the discrepancy in number of parameters, the ARD model suffers only a 15% reduction in fitting error and cross-validation scores with over 50% fewer independent parameters than the models produced by Bayesian ridge and linear regression.

The parameters chosen by the sparse models are largely consistent with the higher magnitude parameters in the unreduced parameter subsets, and can be compared to give

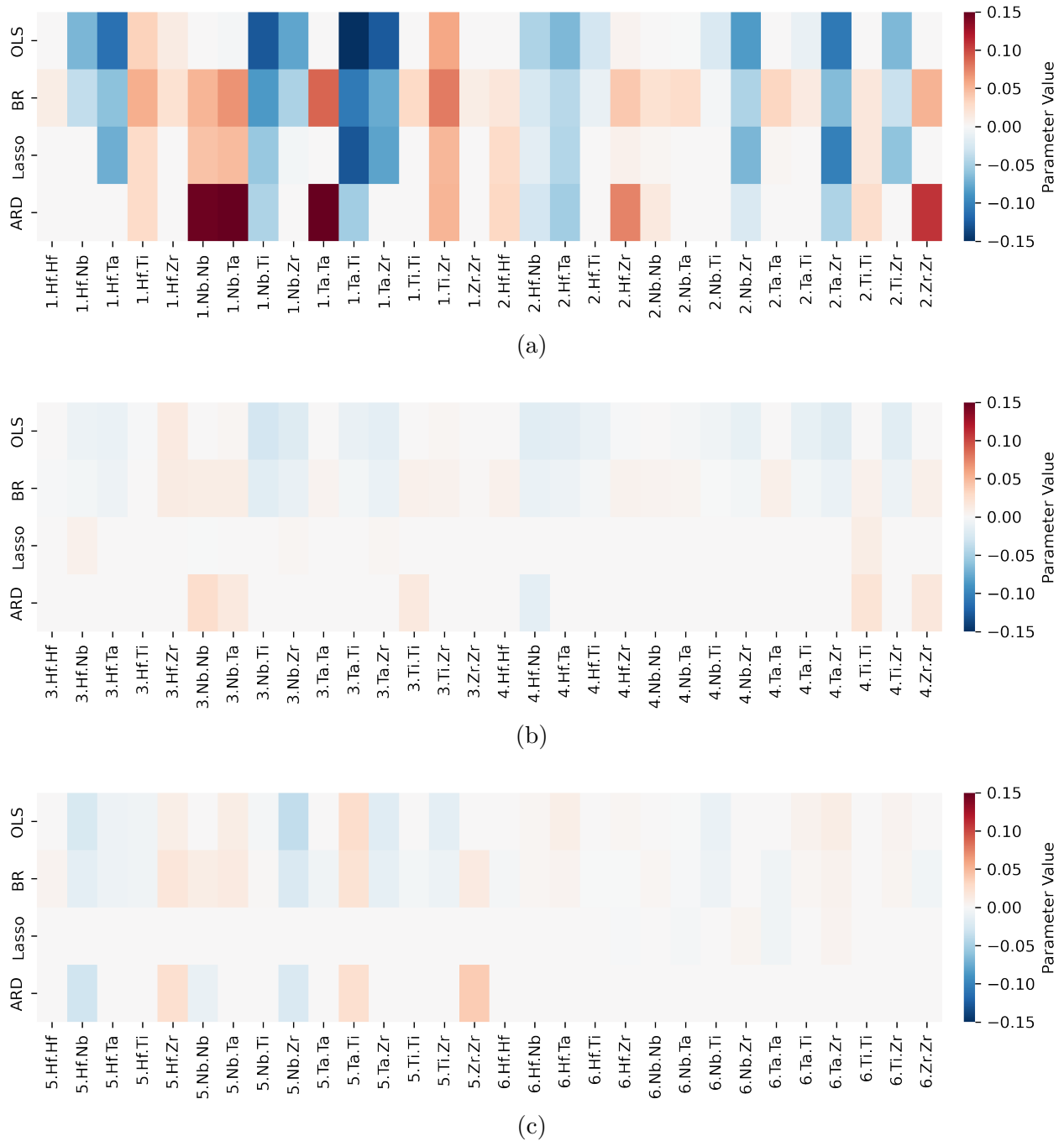


Figure 5.4: Parameter comparison between models fit with ordinary least squares (OLS), Bayesian ridge regression (BR), lasso, and automatic relevance detection (ARD). Parameter plots are separated in (a), (b), and (c) by interaction distance with parameters indicated as  $k.i.j$  for an  $ij$  pair in shell  $k$ .

Table 5.3: Fitting and cross validation scores for the tested linear regression methods. For each parameterization, the maximum number of independent parameters are given to account for coupled parameters and fitted model intercepts. Errors given in meV/atom.

Model	Parameters (ind.)	RMSE	BIC	LOOCV	5-fold CV
OLS	90 (61)	0.971	592.5	1.064	1.082
Bayesian Ridge	90 (61)	0.971	592.5	1.060	1.068
Lasso	30 (31)	1.819	186.4	1.696	1.523
ARD	28 (29)	1.113	172.7	1.250	1.220

insight into potential ordering in the bulk. Nb-Nb, Nb-Ta, and Ta-Ta pairs in the first shell along with Hf-Zr and Zr-Zr pairs in the second shell exhibit strong positive interaction terms corresponding with unfavorable mixing, while Nb-Ti, Ta-Ti, and Ti-Zr first-shell pairs and Hf-Ta, Nb-Zr, and Ta-Zr second-shell pairs exhibit the largest negative interaction terms (favorable mixing). The models also indicate relatively large interactions in the fifth shell for Hf-Nb (negative), Nb-Zr (negative), Ta-Ti (positive), and Zr-Zr (positive) pairs. The relationship between the largest interaction parameters of each model supports a physical interpretation of the models that relies on strong interactions exhibiting large interaction coefficients.

Contrasting with the early results for vacancy interactions, this regressor comparison fails to provide evidence that a sparse parameter set is necessarily better for modeling bulk pair interactions. Given the excess degrees of freedom in the parameterizations of the pair-interaction models, it is unclear if a reduction in parameters provides the benefits that one might expect. Metrics such as the BIC, which heavily penalizes larger parameterizations, rely on an assumption of independence and do not capture the interactions of mathematically coupled parameters. As such, the BIC may not always be a reliable tool for finding the optimal parameter subset in these systems.

To further explore the strength of the sparse parameterizations, the individual 5-fold cross validation splits are presented in Table 5.4. With the lasso algorithm optimizing the parameterization of each training split with respect to the BIC, it is apparent that the cross

Table 5.4: 5-fold cross validation results for lasso regression and automatic relevance determination. Fitting and cross validation root mean squared errors for individual splits of cross validation given in meV/atom.

Model	Parameters	Training Error	CV Error
Lasso	21	2.925	2.759
	31	1.396	1.356
	45	1.106	1.114
	50	0.997	1.051
	53	0.962	1.312
ARD	26	1.157	1.129
	26	1.120	1.162
	27	1.130	1.112
	26	1.133	1.230
	34	0.996	1.302

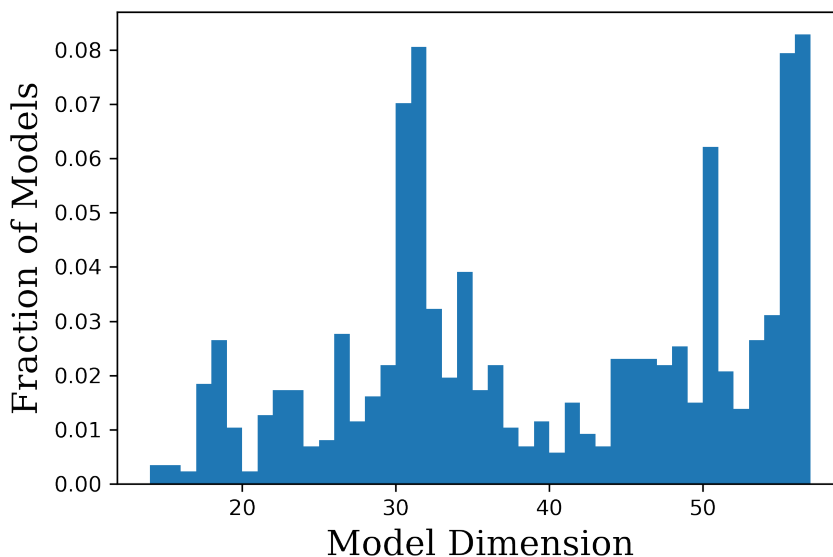


Figure 5.5: Number of parameters included in the BIC optimized lasso regression for each training split of leave-one-out cross validation.

validation scores of lasso regression are not evidence of overfitting, but rather the result of a correlation between model dimension and model precision. The size of the parameterization is discovered to be highly dependent on the sample split and the cross validation errors are found to be consistent with the fitting errors. ARD regression does not improve on the accuracy of ordinary least squares, but is found to reliably produce sparse parameterizations

with only minimal information loss. Neither lasso nor ARD produces a unique solution for the parameterization, introducing ambiguity into the physical interpretation of the interaction terms.

### 5.2.2 Mean-Field Normalization

The complete distribution of LOOCV parameterizations is presented in Fig. 5.5 and shows that the the model dimension varies from 14 to 58 parameters with only a single sample omitted from the training set in each split. The multi-modal nature of the distribution could suggest a grouping behavior of the parameters where the inclusion of certain interaction terms must coincide with a set of coupled terms to maintain a balance in the mean-field interaction. In total, 111 of the 873 LOOCV lasso models contain 25 parameters. These 111 models consist almost exclusively of first- and second-shell interaction terms and have an average fitting error of 2.792 meV/atom. The parameters in the first two shells for each model are shown in Fig. 5.6.

Two sets of interaction terms stand out in the analysis of these parameterizations. In the first shell, the Ta-Ta interaction is either zero or large and, by its inclusion, results in an increase in the Nb-Ta, Hf-Ta, Ta-Ti, and Ta-Zr interaction terms. This same effect is present in the second shell for the Zr-Zr interaction acting on Hf-Zr, Nb-Zr, Ta-Zr, and Ti-Zr in the same way. The highly unstable nature of BIC optimized lasso regression makes it unreliable for producing a consistent physically interpretable model. Further, it is clear that any reduction of the parameter space interferes with a direct interpretation of individual parameter values.

If we consider the pair interaction terms as the difference in the enthalpy contribution of atoms in a given pair compared with the mean-field interaction for each, it becomes apparent that the reduction of a parameter space only serves to place the dropped terms at a fixed

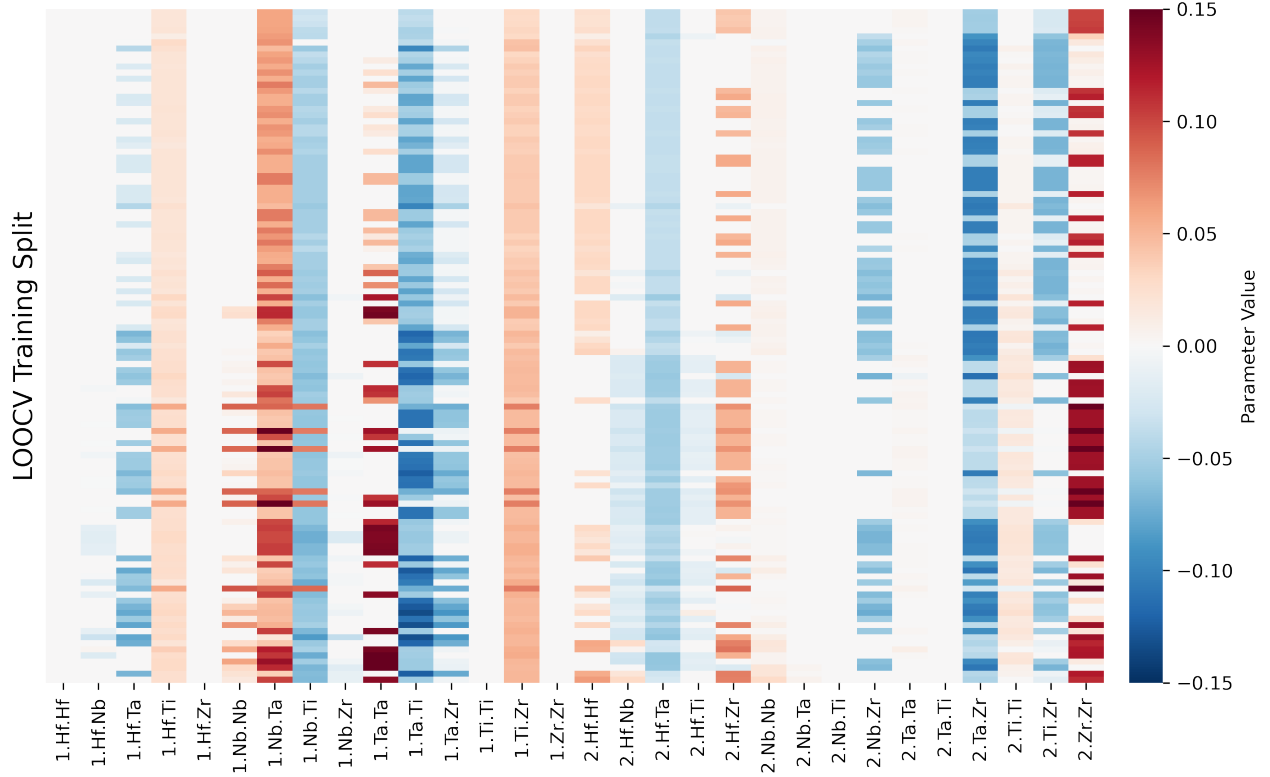


Figure 5.6: (Hf,Nb,Ta,Ti,Zr)C pair interactions in the first two neighbor shells for leave-one-out cross validation splits of the BIC optimized lasso regression with 25 parameters.

distance from the mean-field energy. Assigning a mean-field energy

$$E_{ik}^{\text{MF}} = \frac{1}{m} \sum_j E_{ijk} \quad (5.8)$$

for the average interaction in an  $m$ -component system of each atom  $i$  with the atoms in shell  $k$ , we can define pair interactions  $E_{ijk}$  relative to the mean-field as

$$E_{ijk} = E_{ijk} - (E_{ik}^{\text{MF}} + E_{jk}^{\text{MF}} - \sum_x E_{xk}^{\text{MF}}), \quad (5.9)$$

representing the difference in energy between and  $ij$  pair and the average  $ix$  and  $jx$  pairs on a disordered lattice, with  $x$  as the global composition. The equivalent interaction parameters

$E_{ijk}$  are defined with a fixed physical reference such that

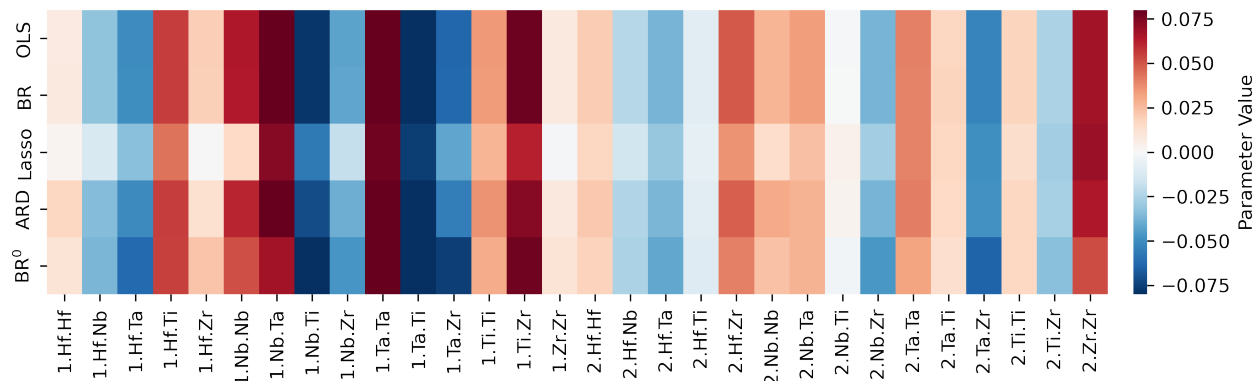
$$E_{ik}^{\text{MF}} = 0 \quad (5.10)$$

for all  $i$ . The shifts in mean-field energy are captured by a corresponding shift in the intercept of the model, given as

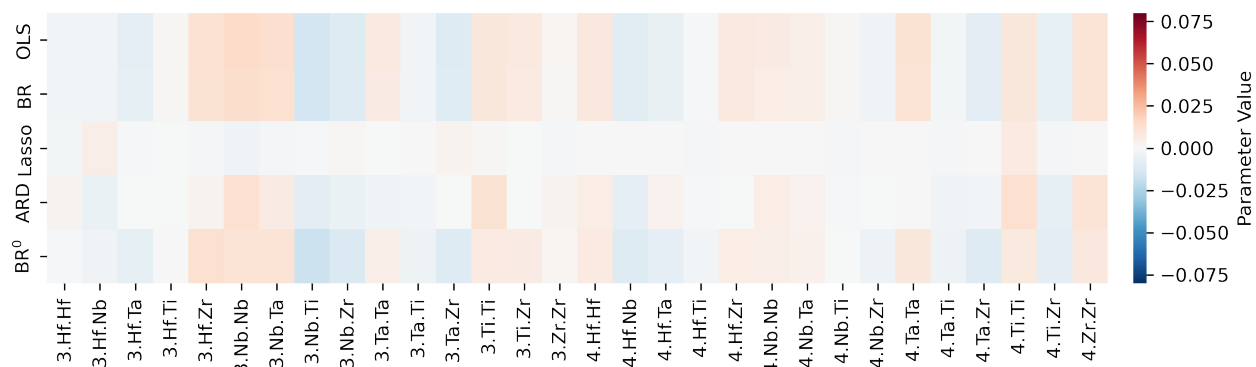
$$E_0 = E_0 + \frac{1}{mk} \sum_{i,k} E_{ik}^{\text{MF}}. \quad (5.11)$$

We can now compare the regression model coefficients at the converged zero mean-field solution to isolate the energy contribution of each pair, finding that all regression algorithms yield nearly identical parameterizations, as shown in Fig. 5.7. BIC optimized lasso regression and ARD omit certain long-range interaction shells entirely (fifth- and sixth-shell interactions, respectively), which is retained in the corresponding zero mean-field parameterizations. Within the remaining shells, however, the interaction coefficients for ARD, Bayesian ridge regression, and ordinary least squares all converge to approximately the same values. Interestingly, the initial solution to Bayesian ridge regression produces a similar result to the zero mean-field converged result, but still underestimates the largest positive and negative interactions. This can be explained intuitively by the  $L2$  regularization penalty driving the parameters to a solution where the mean-field is close to zero, but this diverges when interaction energies at either extreme incur too large a penalty.

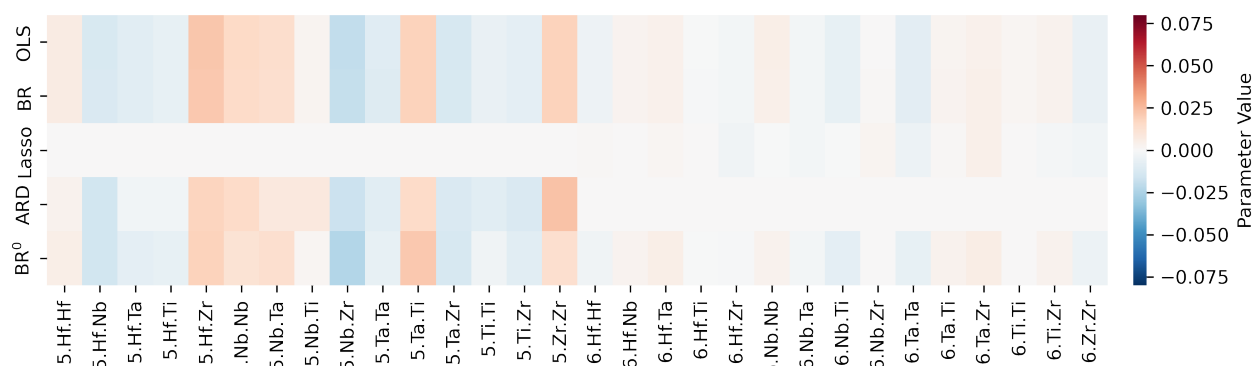
Based on these findings, mean-field normalization of any pair-interaction model is presented as a necessary step before applying any regularized regression technique, as the regularization penalty must be tied to a physically meaningful reference state. The parameter space optimizations from Bayesian ridge, lasso, and ARD regression do not meaningfully improve the model fit and distort the pair energies based on objective functions tied to arbitrary interaction parameter values rather than direct fitting of pair interactions. Further optimization of a physics-informed optimization function that incorporates this mean-field normalization is left for future research.



(a)



(b)



(c)

Figure 5.7: (Hf,Nb,Ta,Ti,Zr)C pair model parameter subset comparison between models fit with ordinary least squares (OLS), Bayesian ridge regression (BR), lasso, and automatic relevance detection (ARD) followed by mean-field convergence. Parameter plots are separated in (a), (b), and (c) by interaction distance with parameters indicated as  $k.i.j$  for an  $ij$  pair in shell  $k$ . The initial Bayesian ridge regression model  $BR^0$  is included for comparison.

Table 5.5: Pair interaction energies calculated from the zero mean-field convergence of OLS regression. Pairs indicated as  $k.i.j$  represent an  $ij$  pair in shell  $k$

Pair Interaction	$E_{ijk}^0$ (meV)	Pair Interaction	$E_{ijk}^0$ (meV)
1.Ta.Ti	-90.4	1.Ta.Ta	116.2
1.Nb.Ti	-78.0	1.Nb.Ta	88.0
1.Ta.Zr	-63.6	1.Ti.Zr	78.7
2.Ta.Zr	-53.7	2.Zr.Zr	67.0
1.Hf.Ta	-50.2	2.Hf.Zr	65.0

For the high-entropy carbide composition, the zero mean-field interaction energies can now be used directly as pair affinities in a given shell to give insight into potential ordering behavior. The five highest and lowest interaction parameters are given in Table 5.5. The large negative interaction energies in Ti-Ta, Nb-Ti, Ta-Zr, and Hf-Ta pairs indicate enthalpic mixing between group IVB and VB carbides, while the self-interactions between Ta-Ta and Nb-Nb pairs carry the largest energy penalty. Interestingly, the parameterization still features relatively large interactions between Nb-Zr, Ta-Ti, Hf-Zr, and Zr-Zr pairs in the fifth shell. It is unclear if this is a physically significant interaction, or an artifact of the small supercells used for the training configurations.

### 5.2.3 Ordering and Phase Stability

We next apply the same regression techniques to the group IV-VI high-entropy carbide (Hf,Mo,Ti,W,Zr)C that has been experimentally observed as a multi-phase composition [78]. Comparing the pair model coefficients between stable and unstable high-entropy compositions, we expect to find a difference in the distribution of pair energies signifying an abundance of enthalpically favorable and unfavorable configurations.

The model errors for bulk energy predictions in (Hf,Mo,Ti,W,Zr)C are found to be significantly higher than for (Hf,Nb,Ta,Ta,Zr)C, with training errors on the order of 4 meV/atom as shown in Table 5.6. Some of this uncertainty is likely introduced by the lattice distortions of the group VIB elements, with large distortions conflicting with the assumptions

Table 5.6: Fitting and cross validation scores for (Hf,Mo,Ti,W,Zr)C using the tested linear regression methods. As before, the maximum number of independent parameters are provided. Errors given in meV/atom.

Model	Parameters (ind.)	RMSE	BIC	LOOCV	5-fold CV
OLS	90 (61)	4.326	595.8	4.741	4.836
Bayesian Ridge	90 (61)	595.8		4.709	4.777
Lasso	36 (37)	4.636	228.5	4.984	4.846
ARD	39 (40)	4.369	148.9	4.663	4.803

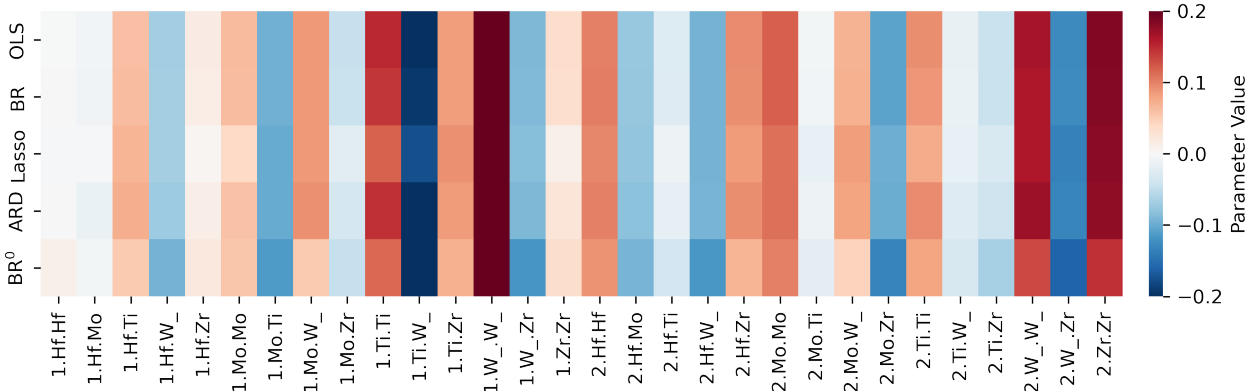


Figure 5.8: (Hf,Mo,Ti,W,Zr)C pair interactions in the first two neighbor shells for zero mean-field adjusted linear regressions.

of a simple pair interaction model [140]. While still reduced by 35-40%, ARD and lasso produce less sparse models than for the (Nb,Ta) composition, but the errors in this case are consistent with ordinary least squares and Bayesian ridge regression. The zero mean-field approach from Section 5.2.1 yields the same result here, with all four models converging to nearly identical parameterizations. For (Hf,Mo,Ti,W,Zr)C, none of the regression methods reject any set of interactions in the first six neighbor shells. Comparison of the first two shells is presented in Fig. 5.8.

The largest pair interactions in (Hf,Mo,Ti,W,Zr)C are presented in Table 5.7, with the full distribution of coefficients for each composition plotted against their percentiles in Fig. 5.9. Similar to the (Nb,Ta) composition, the preferred atomic interactions are between unlike atomic species, namely the interactions between group IVB and group VIB atoms. The Ti-W pair exhibits a particularly strong affinity and, when taken alongside the strong positive

Table 5.7: Pair interaction energies calculated from the zero mean-field convergence of OLS regression. Pairs indicated as  $k.i.j$  represent an  $ij$  pair in shell  $k$

Pair Interaction	$E_{ijk}^0$ (meV)	Pair Interaction	$E_{ijk}^0$ (meV)
1.Ti.W	-198.7	1.W.W	272.0
2.W.Zr	-127.8	2.Zr.Zr	184.8
2.Mo.Zr	-107.4	2.W.W	166.1
1.Mo.Ti	-96.0	1.Ti.Ti	151.4
2.Hf.W	-93.9	2.Mo.Mo	119.2

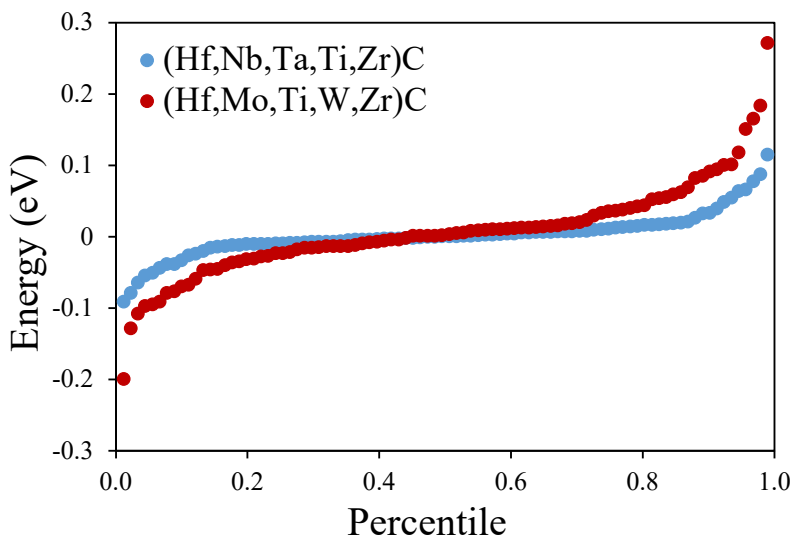


Figure 5.9: Comparison of (Hf,Nb,Ta,Ta,Zr)C and (Hf,Mo,Ti,W,Zr)C bulk pair interaction distributions using parameters from zero mean-field adjusted ordinary least squares regression.

energy associated with W-W pairs in both the first and second shells, may correspond with a tendency to form a secondary phase. The (Ti,W)C ternary system has been shown to exhibit complete miscibility at high temperatures [153, 154], and could suggest the availability of (Ti,W)-rich secondary phases to facilitate the segregation of W from the rocksalt structure. Comparing to the stable (Hf,Nb,Ta,Ti,Zr)C composition, we see that the largest positive and negative interaction parameters here are more than two times larger. Further, the distribution is broader throughout, resulting in fewer pair interactions near the mean-field energy that can be considered degenerate.

This surface level analysis of the bulk pair interactions is generally in agreement with the experimental result of a multi-phase composition, but does not provide specific information or

predictions on the degree of ordering, phase transition temperatures, or the possible availability of metastable structures. To explore the temperature-dependent ordering tendencies of these atomic pairs, techniques such as Metropolis Monte Carlo could be implemented to enable the configurational evolution of a large supercell structure toward an equilibrium ordering.

## CHAPTER

# 6

## CONCLUSIONS

High-entropy ceramics are a promising new class of materials with the potential for myriad applications of their tunable structural and functional properties. In this research, it is shown that the interactions in the local atomic environment are critical for understanding the atomic-scale phenomena that drive materials properties, and that robust and efficient parameterizations of the environment dependence are necessary for establishing property distributions that span the vast range of compositional and configurational degrees of freedom inherent to high-entropy systems.

For vacancy simulations, it was shown that a set of physical constraints applied to the simplest form of pair interactions in the high-entropy diboride  $(\text{Hf,Nb,Ta,Ti,Zr})\text{B}_2$  would reduce the parameter space by over 55%. The reduced pair model proved to be robust to overfitting and could produce reliable vacancy energy predictions with a training sample of

fewer than 200 DFT simulations. Further the parameters from the fitting provide insight into the pair interactions in the material, and act as a first step in understanding the ordering tendencies of high-entropy materials.

The defect energies were shown to be a distribution based on the local environment, allowing for a combinatorial solution to the total defect distribution and an accurate evaluation of the lower tail. Because the most energetically favorable configurations will have a larger contribution to the bulk vacancy concentration, it was shown through analysis of Boltzmann statistics that the interactions in the lower third of the distribution contribute nearly all of the vacancies in the bulk, even at temperatures as high as 2000 K. The vacancy energy distribution in high-entropy materials is also shown to cause a deviation in the Arrhenius relationship between vacancy concentration and temperature. Instead, the vacancy concentration behaves as the mean of the distribution only at infinite temperature while diverging to larger concentrations as temperature decreases.

Stacking faults in the set of group IV, V, and VI early-transition metal rocksalt carbides can be similarly parameterized by the interactions in the environment surrounding the defect plane. In this case, a composition-independent and configuration-agnostic parameterization of the local composition yields a rule-of-mixtures fit with only minor empirical corrections. The unstable stacking fault and barrier position along the  $\{11\bar{2}\}/\{111\}g$  gamma surface slip path were found to be well-defined with respect to the intrinsic stacking fault energy by the Evans-Polanyi-Semenov principle. This relation is ubiquitous in studies of chemical reaction kinetics, but had not previously been applied to the deformation mechanisms in solids.

As with vacancies, the full distribution of stacking fault energies is obtainable for all compositions studied, with the first three distribution moments defined by empirical linear fits to the moments of the distribution of binary energies in the composition. The relative width of the distribution is tied to the sampling of the local environment, with samples taken from a gamma distribution defined by the number of atoms contributing to the effective local interaction. The full stacking fault energy distribution can then be used to estimate a lower

bound for the equilibrium stacking fault separation, predicting positive effective stacking fault energies and finite separations for all early-transition metal high-entropy carbide compositions up to a valence electron concentration of 9.6.

A third pair interaction parameterization was implemented for bulk energy calculations from the linear sum of atom pairs in one stable and one unstable high-entropy carbides. Comparing different parameterizations of the pair terms, it was found that regularized regression techniques did little to change the parameter space, with a “hidden” mean-field energy retaining information from the omitted parameters. Reducing this term to zero yields a parameter subset with a well-defined physical interpretation such that the relative pair affinities are easily compared. Mean-field normalization of the pair-interaction model is presented as a necessary step in order to implement regularization techniques with physically interpretable constraints on the parameter space.

The parameterizations of the local environment for defects and bulk energies give insight into the complexities of the high-entropy phase space. Continuing to find properties in high-entropy materials that can be defined as a function of their local environments will enhance understanding of the mechanisms that enable composition tuning with respect to macroscopic properties. The stacking fault energies distributions provide insight into barriers for dislocation motion, and can be used to inform experiment to establish meaningful relationships between the atomic pair interactions and bulk mechanical properties. Similarly, the bulk energy pair interactions can inform predictions of phase stability and ordering, with the parameters themselves valuable as inputs to a Monte Carlo type annealing simulation. Other defects or defect complexes including anion vacancies, oxygen interstitials, twins, and grain boundaries provide additional opportunities for the implementation of local environment models to enhance our understanding of disorder phenomena in high-entropy materials.

## REFERENCES

- [1] Kevin Kaufmann, Daniel Maryanovsky, William M. Mellor, Chaoyi Zhu, Alexander S. Rosengarten, Tyler J. Harrington, Corey Oses, Cormac Toher, Stefano Curtarolo, and Kenneth S. Vecchio. Discovery of high-entropy ceramics via machine learning. *npj Comput. Mater.*, 6(1):42, May 2020.
- [2] Pranab Sarker, Tyler Harrington, Cormac Toher, Corey Oses, Mojtaba Samiee, Jon-Paul Maria, Donald W. Brenner, Kenneth S. Vecchio, and Stefano Curtarolo. High-entropy high-hardness metal carbides discovered by entropy descriptors. *Nat. Commun.*, 9(1):4980, November 2018.
- [3] Christina M. Rost, Edward Sachet, Trent Borman, Ali Moballegh, Elizabeth C. Dickey, Dong Hou, Jacob L. Jones, Stefano Curtarolo, and Jon-Paul Maria. Entropy-stabilized oxides. *Nat. Commun.*, 6(1):8485, September 2015.
- [4] Corey Oses, Cormac Toher, and Stefano Curtarolo. High-entropy ceramics. *Nat. Rev. Mater.*, 5(4):295–309, February 2020.
- [5] Lun Feng, William G. Fahrenholtz, and Donald W. Brenner. High-Entropy Ultra-High-Temperature Borides and Carbides: A New Class of Materials for Extreme Environments. *Annu. Rev. Mater. Res.*, 51(1):165–185, July 2021.
- [6] Sicong Jiang, Tao Hu, Joshua Gild, Naixie Zhou, Jiuyuan Nie, Mingde Qin, Tyler Harrington, Kenneth Vecchio, and Jian Luo. A new class of high-entropy perovskite oxides. *Scripta Mat.*, 142:116–120, January 2018.
- [7] Mingde Qin, Qizhang Yan, Yi Liu, and Jian Luo. A new class of high-entropy M3B4 borides. *J. Adv. Ceram.*, 10(1):166–172, February 2021.
- [8] George N. Kotsonis, Saeed S. I. Almishal, Francisco Marques Dos Santos Vieira, Vincent H. Crespi, Ismaila Dabo, Christina M. Rost, and Jon-Paul Maria. High-entropy oxides: Harnessing crystalline disorder for emergent functionality. *J. Am. Ceram. Soc.*, 106(10):5587–5611, October 2023.
- [9] Nilesh P. Gurao and Krishanu Biswas. High-Entropy Materials: Critical Review and Way Forward. *Curr. Sci. India*, 118(10):1520, May 2020.
- [10] Rui-Zhi Zhang and Michael J. Reece. Review of high entropy ceramics: design, synthesis, structure and properties. *J. Mater. Chem. A*, 7(39):22148–22162, 2019.
- [11] Yichen Wang. Processing and properties of high entropy carbides. *Adv. Appl. Ceram.*, 121(2):57–78, February 2022.
- [12] Abhishek Sarkar, Ruzica Djenadic, Di Wang, Christina Hein, Ralf Kautenburger, Oliver Clemens, and Horst Hahn. Rare earth and transition metal based entropy stabilised perovskite type oxides. *J. Eur. Ceram. Soc.*, 38(5):2318–2327, May 2018.

- [13] Mingde Qin, Qizhang Yan, Haoren Wang, Kenneth S. Vecchio, and Jian Luo. High-entropy rare earth tetraborides. *J. Eur. Ceram. Soc.*, 41(4):2968–2973, April 2021.
- [14] Keith Bryce, Yueh-Ting Shih, Liping Huang, and Jie Lian. Calcium-Magnesium-Aluminosilicate (CMAS) corrosion resistance of high entropy rare-earth phosphate (Lu<sub>0.2</sub>Yb<sub>0.2</sub>Er<sub>0.2</sub>Y<sub>0.2</sub>Gd<sub>0.2</sub>)PO<sub>4</sub>: A novel environmental barrier coating candidate. *J. Eur. Ceram. Soc.*, 43(14):6461–6472, November 2023.
- [15] Weiwei Sang, Wenbo Xie, Ruiyi Hou, Siqu Li, Hongsong Zhang, Shuaixia Liu, and Xiaoge Chen. Crystal structure and thermophysical properties of (Gd<sub>0.25</sub>Sm<sub>0.25</sub>Yb<sub>0.25</sub>Y<sub>0.25</sub>)<sub>3</sub>TaO<sub>7</sub> high-entropy oxide. *Ceram. Int.*, 49(17):29358–29363, September 2023.
- [16] Xinyu Ping, Bin Meng, Xiaohua Yu, Zhiyuan Ma, Xiaoyu Pan, and Wu Lin. Structural, mechanical and thermal properties of cubic bixbyite-structured high-entropy oxides. *Chem. Eng. J.*, 464:142649, May 2023.
- [17] Joshua Gild, Mojtaba Samiee, Jeffrey L. Braun, Tyler Harrington, Heidy Vega, Patrick E. Hopkins, Kenneth Vecchio, and Jian Luo. High-entropy fluorite oxides. *Journal of the European Ceramic Society*, 38(10):3578–3584, August 2018.
- [18] Joshua Gild, Jeffrey Braun, Kevin Kaufmann, Eduardo Marin, Tyler Harrington, Patrick Hopkins, Kenneth Vecchio, and Jian Luo. A high-entropy silicide:(mo<sub>0.2</sub>nb<sub>0.2</sub>ta<sub>0.2</sub>ti<sub>0.2</sub>w<sub>0.2</sub>)si<sub>2</sub>. *J. Materiomics*, 5(3):337–343, 2019.
- [19] M. Lim, Zs. Rak, J. L. Braun, C. M. Rost, G. N. Kotsonis, P. E. Hopkins, J.-P. Maria, and D. W. Brenner. Influence of mass and charge disorder on the phonon thermal conductivity of entropy stabilized oxides determined by molecular dynamics simulations. *J. Appl. Phys.*, 125(5):055105, February 2019.
- [20] Sufyan M. Shaikh, V. S. Hariharan, Satyesh K. Yadav, and B. S. Murty. CALPHAD and rule-of-mixtures: A comparative study for refractory high entropy alloys. *Intermetallics*, 127:106926, December 2020.
- [21] Mina Lim and Donald W. Brenner. Predicting properties of high entropy carbides from their respective binaries. *Comp. Mater. Sci.*, 226:112255, June 2023.
- [22] Zezhou Li, Shiteng Zhao, Robert O. Ritchie, and Marc A. Meyers. Mechanical properties of high-entropy alloys with emphasis on face-centered cubic alloys. *Progress in Materials Science*, 102:296–345, May 2019.
- [23] Joshua Gild, Yuanyao Zhang, Tyler Harrington, Sicong Jiang, Tao Hu, Matthew C Quinn, William M Mellor, Naixie Zhou, Kenneth Vecchio, and Jian Luo. High-entropy metal diborides: a new class of high-entropy materials and a new type of ultrahigh temperature ceramics. *Sci. Rep.*, 6:37946, 2016.
- [24] A Van de Walle, P Tiwary, M De Jong, D L Olmsted, M Asta, A Dick, D Shin, Yi Wang, L Q Chen, and Z K Liu. Efficient stochastic generation of special quasirandom structures. *Calphad*, 42:13–18, 2013.

- [25] Yoav Lederer, Cormac Toher, Kenneth S. Vecchio, and Stefano Curtarolo. The search for high entropy alloys: A high-throughput ab-initio approach. *Acta Materialia*, 159:364–383, October 2018.
- [26] B. Cantor, I.T.H. Chang, P. Knight, and A.J.B. Vincent. Microstructural development in equiatomic multicomponent alloys. *Materials Science and Engineering: A*, 375-377:213–218, July 2004.
- [27] O. N. Senkov, G. B. Wilks, J. M. Scott, and D. B. Miracle. Mechanical properties of Nb<sub>25</sub>Mo<sub>25</sub>Ta<sub>25</sub>W<sub>25</sub> and V<sub>20</sub>Nb<sub>20</sub>Mo<sub>20</sub>Ta<sub>20</sub>W<sub>20</sub> refractory high entropy alloys. *Intermetallics*, 19(5):698–706, May 2011.
- [28] J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang. Nanostructured High-Entropy Alloys with Multiple Principal Elements: Novel Alloy Design Concepts and Outcomes. *Advanced Engineering Materials*, 6(5):299–303, 2004.
- [29] F. Otto, A. Dlouhý, Ch. Somsen, H. Bei, G. Eggeler, and E. P. George. The influences of temperature and microstructure on the tensile properties of a CoCrFeMnNi high-entropy alloy. *Acta Materialia*, 61(15):5743–5755, September 2013.
- [30] Bernd Gludovatz, Anton Hohenwarter, Dhiraj Catoor, Edwin H. Chang, Easo P. George, and Robert O. Ritchie. A fracture-resistant high-entropy alloy for cryogenic applications. *Science*, 345(6201):1153–1158, September 2014.
- [31] Daniel B. Miracle, Jonathan D. Miller, Oleg N. Senkov, Christopher Woodward, Michael D. Uchic, and Jaimie Tiley. Exploration and Development of High Entropy Alloys for Structural Applications. *Entropy*, 16(1):494–525, 2014.
- [32] Easo P. George, Dierk Raabe, and Robert O. Ritchie. High-entropy alloys. *Nat Rev Mater*, 4(8):515–534, August 2019.
- [33] H. Okamoto, D. J. Chakrabarti, D. E. Laughlin, and T. B. Massalski. The Au-Cu (Gold-Copper) system. *JPE*, 8(5):454–474, October 1987.
- [34] F. Otto, A. Dlouhý, K.G. Pradeep, M. Kuběnová, D. Raabe, G. Eggeler, and E.P. George. Decomposition of the single-phase high-entropy alloy CrMnFeCoNi after prolonged anneals at intermediate temperatures. *Acta Materialia*, 112:40–52, June 2016.
- [35] Céline Varvenne, Aitor Luque, and William A. Curtin. Theory of strengthening in fcc high entropy alloys. *Acta Materialia*, 118:164–176, October 2016.
- [36] Norihiko L. Okamoto, Koretaka Yuge, Katsushi Tanaka, Haruyuki Inui, and Easo P. George. Atomic displacement in the CrMnFeCoNi high-entropy alloy – A scaling factor to predict solid solution strengthening. *AIP Advances*, 6(12):125008, December 2016.

- [37] Daniel Utt, Subin Lee, Yaolong Xing, Hyejin Jeong, Alexander Stukowski, Sang Ho Oh, Gerhard Dehm, and Karsten Albe. The origin of jerky dislocation motion in high-entropy alloys. *Nat. Commun.*, 13(1):4777, August 2022.
- [38] D. B. Miracle and O. N. Senkov. A critical review of high entropy alloys and related concepts. *Acta Materialia*, 122:448–511, January 2017.
- [39] J. M. Cowley. Short-Range Order and Long-Range Order Parameters. *Phys. Rev.*, 138(5A):A1384–A1389, May 1965.
- [40] Y. Rao and W.A. Curtin. Analytical models of short-range order in FCC and BCC alloys. *Acta Materialia*, 226:117621, March 2022.
- [41] Sheng Yin, Yunxing Zuo, Anas Abu-Odeh, Hui Zheng, Xiang-Guo Li, Jun Ding, Shyue Ping Ong, Mark Asta, and Robert O. Ritchie. Atomistic simulations of dislocation mobility in refractory high-entropy alloys and the effect of chemical short-range order. *Nat Commun*, 12(1):4873, August 2021.
- [42] E. Antillon, C. Woodward, S. I. Rao, B. Akdim, and T. A. Parthasarathy. Chemical short range order strengthening in a model FCC high entropy alloy. *Acta Materialia*, 190:29–42, May 2020.
- [43] Shuai Chen, Zachary H. Aitken, Subrahmanyam Pattamatta, Zhaoxuan Wu, Zhi Gen Yu, David J. Srolovitz, Peter K. Liaw, and Yong-Wei Zhang. Simultaneously enhancing the ultimate strength and ductility of high-entropy alloys via short-range ordering. *Nat Commun*, 12(1):4953, August 2021.
- [44] Qingqing Ding, Yin Zhang, Xiao Chen, Xiaoqian Fu, Dengke Chen, Sijing Chen, Lin Gu, Fei Wei, Hongbin Bei, Yanfei Gao, Minru Wen, Jixue Li, Ze Zhang, Ting Zhu, Robert O. Ritchie, and Qian Yu. Tuning element distribution, structure and properties by composition in high-entropy alloys. *Nature*, 574(7777):223–227, October 2019.
- [45] Shuai Chen, Tian Wang, Xiaoyan Li, Yuan Cheng, Gang Zhang, and Huajian Gao. Short-range ordering and its impact on thermodynamic property of high-entropy alloys. *Acta Materialia*, 238:118201, October 2022.
- [46] Hao Chen, Jie Fu, Pengfei Zhang, Honggen Peng, Carter W. Abney, Kecheng Jie, Xiaoming Liu, Miaofang Chi, and Sheng Dai. Entropy-stabilized metal oxide solid solutions as CO oxidation catalysts with high-temperature stability. *J. Mater. Chem. A*, 6(24):11129–11133, June 2018.
- [47] Qingsong Wang, Abhishek Sarkar, Di Wang, Leonardo Velasco, Raheleh Azmi, Subramshu S. Bhattacharya, Thomas Bergfeldt, Andre Düvel, Paul Heitjans, Torsten Brezesinski, Horst Hahn, and Ben Breitung. Multi-anionic and -cationic compounds: new high entropy materials for advanced Li-ion batteries. *Energy Environ. Sci.*, 12(8):2433–2442, August 2019.

- [48] Shang Zhai, Jimmy Rojas, Nadia Ahlborg, Kipil Lim, Michael F. Toney, Hyungyu Jin, William C. Chueh, and Arun Majumdar. The use of poly-cation oxides to lower the temperature of two-step thermochemical water splitting. *Energy Environ. Sci.*, 11(8):2172–2178, August 2018.
- [49] Xiaochao Jin, Xueling Fan, Chunsheng Lu, and Tiejun Wang. Advances in oxidation and ablation resistance of high and ultra-high temperature ceramics modified or coated carbon/carbon composites. *Journal of the European Ceramic Society*, 38(1):1–28, January 2018.
- [50] Ambreen Nisar, Rubia Hassan, Arvind Agarwal, and Kantesh Balani. Ultra-high temperature ceramics: Aspiration to overcome challenges in thermal protection systems. *Ceramics International*, 48(7):8852–8881, April 2022.
- [51] Eric Wuchina, Elizabeth Opila, Mark Opeka, Bill Fahrenholtz, and Inna Talmy. UHTCs: Ultra-High Temperature Ceramic Materials for Extreme Environment Applications. *Electrochem. Soc. Interface*, 16(4):30, December 2007.
- [52] Omar Cedillos-Barraza, Dario Manara, K. Boboridis, Tyson Watkins, Salvatore Grasso, Daniel D. Jayaseelan, Rudy J. M. Konings, Michael J. Reece, and William E. Lee. Investigating the highest melting temperature materials: A laser melting study of the TaC-HfC system. *Sci Rep*, 6:37962, December 2016.
- [53] Tyler J. Harrington, Joshua Gild, Pranab Sarker, Cormac Toher, Christina M. Rost, Olivia F. Dipolo, Cameron McElfresh, Kevin Kaufmann, Eduardo Marin, Lucas Borowski, Patrick E. Hopkins, Jian Luo, Stefano Curtarolo, Donald W. Brenner, and Kenneth S. Vecchio. Phase stability and mechanical properties of novel high entropy transition metal carbides. *Acta Materialia*, 166:271–280, March 2019.
- [54] Lavina Backman, Joshua Gild, Jian Luo, and Elizabeth J. Opila. Part I: Theoretical predictions of preferential oxidation in refractory high entropy materials. *Acta Materialia*, 197:20–27, September 2020.
- [55] Lavina Backman, Joshua Gild, Jian Luo, and Elizabeth J. Opila. Part II: Experimental verification of computationally predicted preferential oxidation of refractory high entropy ultra-high temperature ceramics. *Acta Materialia*, 197:81–90, September 2020.
- [56] Christina M. Rost, Zsolt Rak, Donald W. Brenner, and Jon-Paul Maria. Local structure of the  $\text{Mg}_x\text{Ni}_x\text{Co}_x\text{Cu}_x\text{Zn}_x\text{O}(x=0.2)$  entropy-stabilized oxide: An EXAFS study. *Journal of the American Ceramic Society*, 100(6):2732–2738, 2017.
- [57] S. E. Daigle and D. W. Brenner. Statistical approach to obtaining vacancy formation energies in high-entropy crystals from first principles calculations: Application to a high-entropy diboride. *Phys. Rev. Mater.*, 4(12):123602, December 2020.
- [58] Corey Oses, Cormac Toher, and Stefano Curtarolo. High-entropy ceramics. *Nat. Rev. Mater.*, 5:295–309, 2020.

- [59] Christina M Rost, Edward Sacht, Trent Borman, Ali Moballegh, Elizabeth C Dickey, Dong Hou, Jacob L Jones, Stefano Curtarolo, and Jon Paul Maria. Entropy-stabilized oxides. *Nat. Commun.*, 6:8485, 2015.
- [60] Tyler J Harrington, Joshua Gild, Pranab Sarker, Cormac Toher, Christina M Rost, Olivia F Dippo, Cameron McElfresh, Kevin Kaufmann, Eduardo Marin, Lucas Borowski, et al. Phase stability and mechanical properties of novel high entropy transition metal carbides. *Acta Mater.*, 166:271–280, 2019.
- [61] Sicong Jiang, Tao Hu, Joshua Gild, Naixie Zhou, Jiuyuan Nie, Mingde Qin, Tyler Harrington, Kenneth Vecchio, and Jian Luo. A new class of high-entropy perovskite oxides. *Scr. Mater.*, 142:116–120, 2018.
- [62] Abhishek Sarkar, Qingsong Wang, Alexander Schiele, Mohammed Reda Chellali, Subramshu S Bhattacharya, Di Wang, Torsten Brezesinski, Horst Hahn, Leonardo Velasco, and Ben Breitung. High-entropy oxides: Fundamental aspects and electrochemical properties. *Adv. Mater.*, 31(26):1970189, 2019.
- [63] Christina M Rost, Zsolt Rak, Donald W Brenner, and Jon Paul Maria. Local structure of the  $\text{mgxnixcoxcuxznxo}$  ( $x= 0.2$ ) entropy-stabilized oxide: An exafs study. *J. Am. Ceram. Soc.*, 100(6):2732–2738, 2017.
- [64] D Bérardan, S Franger, A K Meena, and N Dragoe. Room temperature lithium superionic conductivity in high entropy oxides. *J. Mater. Chem. A*, 4(24):9536–9541, 2016.
- [65] Abhishek Sarkar, Ruzica Djenadic, Nandhini J Usharani, Kevin P Sanghvi, Venkata S K Chakravadhanula, Ashutosh S Gandhi, Horst Hahn, and Subramshu S Bhattacharya. Nanocrystalline multicomponent entropy stabilised transition metal oxides. *J. Eur. Ceram. Soc.*, 37(2):747–754, 2017.
- [66] Abhishek Sarkar, Christoph Loho, Leonardo Velasco, Tiju Thomas, Subramshu S Bhattacharya, Horst Hahn, and Ruzica Djenadic. Multicomponent equiatomic rare earth oxides with a narrow band gap and associated praseodymium multivalency. *Dalton Trans.*, 46(36):12167–12176, 2017.
- [67] Abhishek Sarkar, Leonardo Velasco, Di Wang, Qingsong Wang, Gopichand Talasila, Lea de Biasi, Christian Kübel, Torsten Brezesinski, Subramshu S Bhattacharya, Horst Hahn, et al. High entropy oxides for reversible energy storage. *Nat. Commun.*, 9(1):1–9, 2018.
- [68] Qingsong Wang, Abhishek Sarkar, Zhenyou Li, Yang Lu, Leonardo Velasco, Subramshu S Bhattacharya, Torsten Brezesinski, Horst Hahn, and Ben Breitung. High entropy oxides as anode material for li-ion battery applications: A practical approach. *Electrochem. Commun.*, 100:121–125, 2019.
- [69] Ralf Witte, Abhishek Sarkar, Robert Kruk, Benedikt Eggert, Richard A Brand, Heiko Wende, and Horst Hahn. High-entropy oxides: An emerging prospect for magnetic rare-earth transition metal perovskites. *Phys. Rev. Mater.*, 3(3):034406, 2019.

- [70] Giovanna Tallarita, Roberta Licheri, Sebastiano Garroni, Roberto Orru, and Giacomo Cao. Novel processing route for the fabrication of bulk high-entropy metal diborides. *Scr. Mater.*, 158:100–104, 2019.
- [71] Da Liu, Tongqi Wen, Beilin Ye, and Yanhui Chu. Synthesis of superfine high-entropy metal diboride powders. *Scr. Mater.*, 167:110–114, 2019.
- [72] Paul H Mayrhofer, Alexander Kirnbauer, Ph Ertelthaler, and Christian M Koller. High-entropy ceramic thin films; a case study on transition metal diborides. *Scr. Mater.*, 149:93–97, 2018.
- [73] Elinor Castle, Tamás Csanádi, Salvatore Grasso, Ján Dusza, and Michael Reece. Processing and properties of high-entropy ultra-high temperature carbides. *Sci. Rep.*, 8(1):1–12, 2018.
- [74] Xueliang Yan, Loic Constantin, Yongfeng Lu, Jean François Silvain, Michael Nastasi, and Bai Cui. (hf0.2zr0.2ta0.2nb0.2ti0.2)c high-entropy ceramics with low thermal conductivity. *J. Am. Ceram. Soc.*, 101(10):4486–4491, 2018.
- [75] Jieyang Zhou, Jinyong Zhang, Fan Zhang, Bo Niu, Liwen Lei, and Weimin Wang. High-entropy carbide: A novel class of multicomponent ceramics. *Ceram. Int.*, 44(17):22014–22018, 2018.
- [76] Yan Yang, Wei Wang, Guo Yong Gan, Xue Feng Shi, and Bi Yu Tang. Structural, mechanical and electronic properties of (tanbhtizr)c high entropy carbide under pressure: Ab initio investigation. *Phys. B*, 550:163–170, 2018.
- [77] Ján Dusza, Peter Švec, Vladimír Girman, Richard Sedlák, Elinor G Castle, Tamás Csanádi, Alexandra Kovalčíková, and Michael J Reece. Microstructure of (hf-ta-zr-nb)c high-entropy carbide at micro and nano/atomic level. *J. Eur. Ceram. Soc.*, 38(12):4303–4307, 2018.
- [78] Pranab Sarker, Tyler Harrington, Cormac Toher, Corey Oses, Mojtaba Samiee, Jon Paul Maria, Donald W Brenner, Kenneth S Vecchio, and Stefano Curtarolo. High-entropy high-hardness metal carbides discovered by entropy descriptors. *Nat. Commun.*, 9(1):1–10, 2018.
- [79] Andrei V Ruban. Thermal vacancies in random alloys in the single-site mean-field approximation. *Phys. Rev. B*, 93(13):134115, 2016.
- [80] Xi Zhang and Marcel H F Sluiter. Ab initio prediction of vacancy properties in concentrated alloys: The case of fcc cu-ni. *Phys. Rev. B*, 91(17):174107, 2015.
- [81] A Van der Ven and G Ceder. Vacancies in ordered and disordered binary alloys treated with the cluster expansion. *Phys. Rev. B*, 71(5):054102, 2005.
- [82] Erna K Delczeg-Czirjak, Lorand Delczeg, Levente Vitos, and Olle Eriksson. Monovacancy formation energies and fermi surface topological transitions in pd-ag alloys. *Phys. Rev. B*, 92(22):224107, 2015.

- [83] Nongnuch Artrith, Alexander Urban, and Gerbrand Ceder. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. *Phys. Rev. B*, 96(1):014112, 2017.
- [84] Gideon Schwarz et al. Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464, 1978.
- [85] Peter E Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50(24):17953, 1994.
- [86] Georg Kresse and Jürgen Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47(1):558, 1993.
- [87] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54(16):11169, 1996.
- [88] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.*, 6(1):15–50, 1996.
- [89] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865, 1996.
- [90] B T M Loh and C T Liu. Reply to comment on some calculations of vacancy characteristics in substitutional alloys. *Scr. Metall.*, 6(7):565–566, 1972.
- [91] Weiliang Chen, Xueyong Ding, Yuchao Feng, Xiongjun Liu, Kui Liu, Z P Lu, Dianzhong Li, Yiyi Li, C T Liu, and Xing Qiu Chen. Vacancy formation enthalpies of high-entropy fecocrni alloy via first-principles calculations and possible implications to its superior radiation tolerance. *J. Mater. Sci. Technol.*, 34(2):355–364, 2018.
- [92] A Esfandiarpour and M N Nasrabadi. Vacancy formation energy in cunico equimolar alloy and cunico high entropy alloy: ab initio based study. *Calphad*, 66:101634, 2019.
- [93] Kazuki Sugita, Naoki Matsuoka, Masataka Mizuno, and Hideki Araki. Vacancy formation enthalpy in cocrfemni high-entropy alloy. *Scr. Mater.*, 176:32–35, 2020.
- [94] Zhijun Wang, C T Liu, and Peng Dou. Thermodynamics of vacancies and clusters in high-entropy alloys. *Phys. Rev. Mater.*, 1(4):043601, 2017.
- [95] Samuel E. Daigle, Stefano Curtarolo, William G. Fahrenholtz, Jon-Paul Maria, Douglas E. Wolfe, Eva Zurek, and Donald W. Brenner. Interfacial defect properties of high-entropy carbides: Stacking faults, Shockley partial dislocations, and a new Evans-Polanyi-Semenov relation, November 2023. arXiv:2311.00834 [cond-mat].
- [96] Zs. Rak, C. M. Rost, M. Lim, P. Sarker, C. Toher, S. Curtarolo, J.-P. Maria, and D. W. Brenner. Charge compensation and electrostatic transferability in three entropy-stabilized oxides: Results from density functional theory calculations. *J. Appl. Phys.*, 120(9):095105, September 2016.

- [97] Jeffrey L. Braun, Christina M. Rost, Mina Lim, Ashutosh Giri, David H. Olson, George N. Kotsonis, Gheorghe Stan, Donald W. Brenner, Jon-Paul Maria, and Patrick E. Hopkins. Charge-Induced Disorder Controls the Thermal Conductivity of Entropy-Stabilized Oxides. *Adv. Mater.*, 30(51):1805004, 2018.
- [98] Xi Zhao, Sen Yu, Jiming Zheng, Michael J. Reece, and Rui-Zhi Zhang. Machine learning of carbon vacancy formation energy in high-entropy carbides. *J. Eur. Ceram. Soc.*, 43(4):1315–1321, April 2023.
- [99] V. Vitek. Intrinsic stacking faults in body-centred cubic crystals. *Philos. Mag.*, 18(154):773–786, October 1968.
- [100] N. Bernstein and E. Tadmor. Tight-binding calculations of stacking energies and twinnability in fcc metals. *Phys. Rev. B*, 69(9):094116, March 2004.
- [101] E Tadmor. A first-principles measure for the twinnability of FCC metals. *J. Mech. Phys. Solids*, 52(11):2507–2519, November 2004.
- [102] Michal Jahnátek, Jürgen Hafner, and Marián Krajčí. Shear deformation, ideal strength, and stacking fault formation of fcc metals: A density-functional study of Al and Cu. *Phys. Rev. B*, 79(22):224103, June 2009.
- [103] Xue-Chun Zhang, Shuo Cao, Lian-Ji Zhang, Rui Yang, and Qing-Miao Hu. Unstable stacking fault energy and peierls stress for evaluating slip system competition in body-centered cubic metals. *J. Mater. Res. Technol.*, 22:3413–3422, January 2023.
- [104] Hang Yu, Mohammadreza Bahadori, Gregory B. Thompson, and Christopher R. Weinberger. Understanding dislocation slip in stoichiometric rocksalt transition metal carbides and nitrides. *J. Mater. Sci.*, 52(11):6235–6248, June 2017.
- [105] Chase J. Smith, Xiao-Xiang Yu, Qianying Guo, Christopher R. Weinberger, and Gregory B. Thompson. Phase, hardness, and deformation slip behavior in mixed Hf<sub>x</sub>Ta<sub>1-x</sub>C. *Acta Mater.*, 145:142–153, February 2018.
- [106] Tamás Csanádi, Vladimír Girman, Łukasz Maj, Jerzy Morgiel, Michael J. Reece, and Ján Dusza. Hardness anisotropy and active slip systems in a (Hf-Ta-Zr-Nb)C high-entropy carbide during nanoindentation. *Int. J. Refract. Met. Hard. Mater.*, 100:105646, November 2021.
- [107] Nicholas De Leon, Xiao-xiang Yu, Hang Yu, Christopher R. Weinberger, and Gregory B. Thompson. Bonding Effects on the Slip Differences in the B 1 Monocarbides. *Phys. Rev. Lett.*, 114(16):165502, April 2015.
- [108] S. L. Shang, W. Y. Wang, Y. Wang, Y. Du, J. X. Zhang, A. D. Patel, and Z. K. Liu. Temperature-dependent ideal strength and stacking fault energy of fcc Ni: a first-principles study of shear deformation. *J. Phys.: Condens. Matter*, 24(15):155402, March 2012.

- [109] Li Ma, Zhipeng Wang, Jinli Huang, Guohua Huang, Min Xue, Pingying Tang, and Touwen Fan. Effect of alloying elements on stacking fault energies and twinnabilities in high-entropy transition-metal carbides. *J. Am. Ceram. Soc.*, 104(12):6521–6532, December 2021.
- [110] Yichen Wang, Tamás Csanádi, Hangfeng Zhang, Ján Dusza, Michael J. Reece, and Rui-Zhi Zhang. Enhanced Hardness in High-Entropy Carbides through Atomic Randomness. *Adv. Theory Simul.*, 3(9):2000111, September 2020.
- [111] Mulaine Shih, Jiashi Miao, Michael Mills, and Maryam Ghazisaeidi. Stacking fault energy in concentrated alloys. *Nat. Commun.*, 12(1):3590, June 2021.
- [112] M. G. Evans and M. Polanyi. Inertia and driving force of chemical reactions. *T. Faraday Soc.*, 34(0):11–24, January 1938.
- [113] Haimin Ding, Xiaoliang Fan, Kaiyu Chu, Xinchun Zhang, and Xiangfa Liu. The influence of carbon vacancies on the stacking fault energy of TiC. *J. Eur. Ceram. Soc.*, 34(7):1893–1897, July 2014.
- [114] Qinghui Meng, Xuan Lin, Yitong Zhai, Lidong Zhang, Peng Zhang, and Liusi Sheng. A theoretical investigation on Bell-Evans-Polanyi correlations for hydrogen abstraction reactions of large biodiesel molecules by H and OH radicals. *Combust. Flame*, 214:394–406, April 2020.
- [115] Paula Pla, Yang Wang, and Manuel Alcamí. When is the Bell–Evans–Polanyi principle fulfilled in Diels–Alder reactions of fullerenes? *Phys. Chem. Chem. Phys.*, 22(16):8846–8852, 2020.
- [116] Gene G. Wubbels. The Bell–Evans–Polanyi Principle and the regioselectivity of electrophilic aromatic substitution reactions. *Tetrahedron Lett.*, 56(13):1716–1719, March 2015.
- [117] Yi Chen, Kai-Hsin Chang, Fan-Yi Meng, Sheng-Ming Tseng, and Pi-Tai Chou. Broadening the Horizon of the Bell–Evans–Polanyi Principle towards Optically Triggered Structure Planarization. *Angew. Chem. Int. Ed.*, 60(13):7205–7212, March 2021.
- [118] Zong-Ying Liu, Yu-Chen Wei, and Pi-Tai Chou. Correlation between Kinetics and Thermodynamics for Excited-State Intramolecular Proton Transfer Reactions. *J. Phys. Chem. A*, 125(30):6611–6620, August 2021.
- [119] Elena F. Sheka. Virtual Free-Radical Polymerization of Vinyl Monomers in View of Digital Twins. *Polymers*, 15(14):2999, July 2023.
- [120] Robin T. Garrod, Mihwa Jin, Kayla A. Matis, Dylan Jones, Eric R. Willis, and Eric Herbst. Formation of Complex Organic Molecules in Hot Molecular Cores through Non-diffusive Grain-surface and Ice-mantle Chemistry. *Astrophys. J. Suppl. Ser.*, 259(1):1, March 2022.

- [121] Jun Cheng, P. Hu, Peter Ellis, Sam French, Gordon Kelly, and C. Martin Lok. Brønsted-Evans-Polanyi Relation of Multistep Reactions and Volcano Curve in Heterogeneous Catalysis. *J. Phys. Chem. C*, 112(5):1308–1311, February 2008.
- [122] Rishi G. Agarwal and James M. Mayer. Coverage-Dependent Rate-Driving Force Relationships: Hydrogen Transfer from Cerium Oxide Nanoparticle Colloids. *J. Am. Chem. Soc.*, 144(45):20699–20709, November 2022.
- [123] B.M. Pascucci, G.S. Otero, P.G. Belevi, and M.M. Branda. Understanding the effects of metal particle size on the NO<sub>2</sub> reduction from a DFT study. *Appl. Surf. Sci.*, 489:1019–1029, September 2019.
- [124] Thomas Kropp and Manos Mavrikakis. Brønsted–Evans–Polanyi relation for CO oxidation on metal oxides following the Mars–van Krevelen mechanism. *J. Catal.*, 377:577–581, September 2019.
- [125] Hu Jia, Zheng Kang, Sanxi Li, Yanfei Li, Junrui Ge, and Dalong Feng. Thermal degradation behavior of seawater based temporary plugging gel crosslinked by polyethyleneimine for fluid loss control in gas well: Kinetics study and degradation prediction. *J. Disper. Sci. Technol.*, 42(9):1299–1310, July 2021.
- [126] Mgcini Keith Phuthi, Archie Mingze Yao, Simon Batzner, Albert Musaelian, Boris Kozinsky, Ekin Dogus Cubuk, and Venkatasubramanian Viswanathan. Accurate Surface and Finite Temperature Bulk Properties of Lithium Metal at Large Scales using Machine Learning Interaction Potentials, preprint; 2023.
- [127] Svatopluk Zeman. Modified Evans–Polanyi–Semenov relationship in the study of chemical micromechanism governing detonation initiation of individual energetic materials. *Thermochim. Acta*, 384(1-2):137–154, February 2002.
- [128] Guangfu Luo, Xizheng Wang, Michael R. Zachariah, and Rohan Mishra. Ignition Threshold of Perovskite-Based Oxides for Solid Fuel Oxidation from First-Principles Calculations. *J. Phys. Chem. C*, 123(29):17644–17649, July 2019.
- [129] Marc J. Cawkwell, Jack Davis, Nicholas Lease, Frank W. Marrs, Alexandra Burch, Suyana Ferreira, and Virginia W. Manner. Understanding Explosive Sensitivity with Effective Trigger Linkage Kinetics. *ACS Phys. Chem. Au*, 2(5):448–458, September 2022.
- [130] Eva Zurek. Discovering New Materials via A Priori Crystal Structure Prediction. In *Reviews in Computational Chemistry*, pages 274–326. John Wiley & Sons, Ltd, 2016.
- [131] César Barrales-Martínez and Pablo Jaque. A deeper analysis of the role of synchronicity on the Bell–Evans–Polanyi plot in multibond chemical reactions: a path-dependent reaction force constant. *Phys. Chem. Chem. Phys.*, 24(24):14772–14779, 2022.
- [132] Shantanu Roy, Stefan Goedecker, and Vladimir Hellmann. Bell-Evans-Polanyi principle for molecular dynamics trajectories and its implications for global optimization. *Phys. Rev. E*, 77(5):056707, May 2008.

- [133] A. F. Shestakov. Quantum-Chemical Verification of the Polanyi–Semenov Generalized Relationship. *Dokl. Phys. Chem.*, 393(4-6):339–342, December 2003.
- [134] Seung-Hoon Jhi, Jisoon Ihm, Steven G. Louie, and Marvin L. Cohen. Electronic mechanism of hardness enhancement in transition-metal carbonitrides. *Nature*, 399(6732):132–134, May 1999.
- [135] Karthik Balasubramanian, Sanjay V. Khare, and Daniel Gall. Valence electron concentration as an indicator for mechanical properties in rocksalt structure nitrides, carbides and carbonitrides. *Acta Mater.*, 152:175–185, June 2018.
- [136] Evgenii T Denisov. New empirical models of radical abstraction reactions. *Russ. Chem. Rev.*, 66(10):859–876, October 1997.
- [137] Zongrui Pei, Siyuan Zhang, Yinkai Lei, Fan Zhang, and Mingwei Chen. Decoupling between Shockley partials and stacking faults strengthens multiprincipal element alloys. *Proc. Natl. Acad. Sci. U.S.A.*, 118(51):e2114167118, December 2021.
- [138] Shi-Yu Liu, Shuoxin Zhang, Shiyang Liu, De-Jun Li, Zhiqiang Niu, Yaping Li, and Sanwu Wang. Stability and mechanical properties of single-phase quinary high-entropy metal carbides: First-principles theory and thermodynamics. *J. Eur. Ceram. Soc.*, 42(7):3089–3098, July 2022.
- [139] Chunghee Nam. Prediction of mechanical properties of high-entropy ceramics by deep learning with compositional descriptors. *Mater. Today Commun.*, 35:105949, June 2023.
- [140] Chiraag Nataraj, Edgar Josué Landinez Borda, Axel van de Walle, and Amit Samanta. A systematic analysis of phase stability in refractory high entropy alloys utilizing linear and non-linear cluster expansion models. *Acta Materialia*, 220:117269, November 2021.
- [141] Qing-Jie Li, Howard Sheng, and Evan Ma. Strengthening in multi-principal element alloys with local-chemical-order roughened dislocation pathways. *Nat Commun*, 10(1):3563, August 2019.
- [142] Biao Xu, Jun Zhang, Shihua Ma, Yaoxu Xiong, Shasha Huang, J. J. Kai, and Shijun Zhao. Revealing the crucial role of rough energy landscape on self-diffusion in high-entropy alloys based on machine learning and kinetic Monte Carlo. *Acta Materialia*, 234:118051, August 2022.
- [143] Jiaxin Zhang, Xianglin Liu, Sirui Bi, Junqi Yin, Guannan Zhang, and Markus Eisenbach. Robust data-driven approach for predicting the configurational energy of high entropy alloys. *Materials & Design*, 185:108247, January 2020.
- [144] Xianglin Liu, Jiaxin Zhang, Junqi Yin, Sirui Bi, Markus Eisenbach, and Yang Wang. Monte Carlo simulation of order-disorder transition in refractory high entropy alloys: A data-driven approach. *Computational Materials Science*, 187:110135, February 2021.

- [145] Yang Wang, G. M. Stocks, W. A. Shelton, D. M. C. Nicholson, Z. Szotek, and W. M. Temmerman. Order-N Multiple Scattering Approach to Electronic Structure Calculations. *Phys. Rev. Lett.*, 75(15):2867–2870, October 1995.
- [146] Arthur S. Goldberger. Stepwise Least Squares: Residual Analysis and Specification Error. *Journal of the American Statistical Association*, 56(296):998–1000, 1961.
- [147] C. Agostinelli. Robust stepwise regression. *Journal of Applied Statistics*, 29(6):825–840, August 2002.
- [148] Xianglin Liu, Jiaxin Zhang, Markus Eisenbach, and Yang Wang. Machine learning modeling of high entropy alloy: the role of short-range order, June 2019. arXiv:1906.02889 [cond-mat, physics:physics].
- [149] Robert Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.
- [150] Michael E. Tipping. Sparse bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.*, 1:211–244, September 2001.
- [151] Tim Mueller and Gerbrand Ceder. Bayesian approach to cluster expansions. *Phys. Rev. B*, 80(2):024103, July 2009.
- [152] 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, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [153] B. Haldar, D. Bandyopadhyay, R. C. Sharma, and N. Chakraborti. The Ti-W-C (Titanium-Tungsten-Carbon) System. *JPE*, 20(3):337–343, May 1999.
- [154] Erwin Rudy. Constitution of ternary titanium-tungsten-carbon alloys. *Journal of the Less Common Metals*, 33(2):245–273, November 1973.