General Model for d-Center Prediction in Multi-Principal-Element Alloys

Lin Wang, Yizhan Zhang, Kayla James, and Bin Ouyang*

Cite This: https://doi.org/10.1021/jacs.5c05649



ACCESS	III Metrics & More	E Article Recommendations	s Supporting Information

ABSTRACT: The center of the d-band or d-orbital of a transition metal site serves as a key electronic structure descriptor in electrocatalysis, yet its prediction across diverse material systems remains challenging, particularly for disordered multi-principalelement alloys (MPEAs). In this study, we present a general, physically interpretable model for predicting d-center values across a wide range of surfaces and compositions, including 10,680 density functional theory (DFT)-relaxed slabs and over 1.2 million d-center values. Inspired by cluster expansion theory, the model captures local coordination environments to accurately estimate d-



Article

centers, achieving a mean absolute error (MAE) of ~0.09 eV, even when considering only the first nearest-neighbor interactions. We further demonstrate the influence of structural orientation of surface and featurization schemes for surface slabs, as well as regression methods, revealing potential for further improvements with a trade-off between model generalizability and complexity with accuracy. The resulting model offers rapid and reliable d-center estimation, enabling high-throughput screening and mechanistic interpretation in catalysis design. Additionally, the model coefficients provide a direct and accessible tool for both experimentalists and theorists to gain valuable insights into the MPEA surface electronic behavior.

1. INTRODUCTION

The computational discovery of novel electrocatalysis has been guided by the identification of electronic structure descriptors. The d-band theory has been proposed since 1995^1 and has been broadly adopted to develop electronic structure descriptors.²⁻⁴ While the d-center is frequently reported to correlate directly with the overpotentials of various catalytic processes,²⁻⁴ efficient prediction of d-center values across a broad range of materials is still absent due to the absence of both systematic data sets and universal models. The emergence of high entropy catalysis has placed new challenges in calculating and predicting the d-center values, as the disordering nature of such materials creates a diverse amount of local structures.⁵⁻¹⁰

In this work, we propose that the prediction of the d-center value can be accomplished by a universal model capable of generalizing across different bulk structures, Miller indices, elemental compositions, and stoichiometries. The model is trained on a data set comprising 10,680 DFT-relaxed surface slabs and 1,220,008 density functional theory (DFT) calculated d-center values. With a broad energy distribution ranging from -14 to 3 eV relative to the Fermi level, our model shows an impressive fitting mean absolute error (MAE) around 0.09 eV, even in its simplest form. The model is fully physically interpretable, leveraging features derived from the atomic fraction in the bulk structure and local coordination environment of surface atoms, based on the principles of cluster expansion theory.^{11–16} Notably, even when considering only the nearest-neighbor interactions, the model demonstrates high

accuracy. Further improvements can be achieved by refining the feature vectors to include longer-range interactions, distinguishing between surface and bulk atoms, and using more advanced regularization methods. This research provides a valuable reference for future catalysis studies based on the d-band theory. The model coefficients can be conveniently used by both experimentalists and theorists for the rapid estimation and analysis of various catalytic performances.

2. METHODOLOGY

First-principles DFT calculations were performed to obtain an accurate description of the structural energies and magnetic moments of the different cathode materials. All calculations were performed using the projector-augmented wave (PAW) method¹⁷ as implemented in the Vienna Ab initio Simulation Package (VASP).¹⁸ For all calculations, a reciprocal space discretization of 25 k-points per Å⁻¹ was applied, and the convergence criteria were set as 10^{-6} eV for electronic loops and 0.02 eV Å⁻¹ for ionic loops. The bulk compositions and structures adopted for DFT calculations were obtained from a

Received:	April 2, 2025
Revised:	May 19, 2025
Accepted:	June 12, 2025

previously reported high entropy data set that contains up to four constituent elements. $^{19,20}\,$

3. RESULTS

3.1. High-Throughput Screening of MPEA Surfaces. The high-throughput screening process begins with our database of bulk MPEA compositions, which includes a comprehensive set of ternary and quaternary equimolar compositions.^{19,20} From this collection, a subset of compositions predicted to be entropy-stabilized below 1273 K was selected. The potential for entropy stabilization below 1273 K is quantified by the $E_{hull} - 1273S_{Ideal}$ value, where E_{hull} represents the energy above the convex hull of all possible competing phases within the corresponding MPEA compositional space.²¹ Such methods have been widely used to quantify stability of multicomponent materials and have been verified by many studies, demonstrating promising performance in predicting synthesizability.^{8,15,22–24} Additional filters were applied based on material cost and aqueous stability, as illustrated in Figure 1(a). Specifically, a price cutoff of £2,000 per kilogram was used



Figure 1. (a) The workflow for high-throughput surface state calculations; (b) the elemental coverage of our database; (c) the distribution of all computed d-center values, with the Fermi energy set as zero.

to screen for low-cost compositions, based on element prices provided by CompoundChem.²⁵ In parallel, MPEA compositions containing noble metals were considered with a focus on aqueous stability. To assess this, the decomposition driving force in aqueous environments was calculated, and a decomposition energy threshold of 0.2 eV/atom was applied to identify noblemetal-containing MPEAs that exhibit relative stability according to Pourbaix diagram^{3,26,27} calculations. The above-mentioned screening workflow ultimately led to the identification of 605 bulk MPEA compositions. Their corresponding low-index surfaces were then screened and constructed, leading to a final data set of 10,680 surface slabs, each computed using DFT. For each Miller index, all possible surface terminations are enumerated to capture a diverse set of surface orientations for a given bulk composition.

The elemental coverage of this study can be illustrated in Figure 1(b), while the energy distribution of the 1,220,008 d-center (denoted as d_c) is shown in Figure 1(c), where all values are referenced to the Fermi energy. As seen in Figure 1(c), the d-center values span a broad energy range and can generally be divided into four distinct domains. The d-center values in the range of -15 eV to -10 eV are primarily associated with Sn sites.

This is due to the fully occupied 4d states and partially filled 5p states of Sn, which shift the d-center to lower energies compared to other elements. Similarly, Zn, as a d¹⁰ transition metal, contributes d-center values largely within the -10 eV to -5 eVrange, with the detailed distribution on different Miller indices shown in Figure S1. In contrast, the d-states of most other elements fall within the energy window of approximately -5 to 2 eV. A small subset of d-center values exceeding 2 eV, from Ni, Co, and Zn sites, are represented as outliers and originate exclusively from a single structure: the {111} surface of ZnCoNi with a BCC structure. This structure is shown in Figure S2, with all the corresponding outlier sites and d-center values detailed in Table S1. It should be noted that early transition metals often exhibit d-center values above the Fermi level due to their largely unoccupied d orbitals; however, this does not imply thermal instability, as many reported thermally stable metals, intermetallics, and alloys based on early transition metals also show d-centers above the Fermi level. 3,28

3.2. Linear Correlation between d-Center and Local Coordination. Inspired by the cluster expansion theory, interactions within a disordered system can be modeled as a linear combination of contributions from interacting atomic clusters. In general, this approach follows the following form:

$$X(\sigma) = X_0 + \sum_{\beta} m_{\beta} J_{\beta} \Phi_{\alpha}(\sigma)_{\alpha \in \beta}$$
(1)

In this framework, $\Phi_{\alpha}(\sigma)$ represents the cluster basis function, which can commonly take the form of an indicator basis, polynomial basis, or sinusoidal basis.¹¹ The cluster basis function $\Phi_{\alpha}(\sigma)$ is grouped and averaged over orbits β of symmetrically equivalent clusters to generate the correlation functions $\langle \Phi_{\alpha}(\sigma) \rangle_{\alpha \in \beta}$. m_{β} is the multiplicity of the orbit β , and J_{β} stands for the effective cluster interactions, which are essentially fitting parameters indicating the contribution from each cluster to the final property $X(\sigma)$. X_0 is the constant term of a cluster expansion fit. Here we propose that a simplified version of cluster expansion based on atomic fraction of each element and first nearest neighbor (1NN) pairs will work on fitting the dcenter; thus, eq 1 can be simplified as

$$d_{c}^{M} = \sum_{i} c_{i} J_{i} + \sum_{i} N_{i}^{M} J_{i}^{M} + d_{0}^{M}$$
(2)

 c_i represents the atomic fraction of each metal in the calculated structure, and J_i shows the contribution of each element to the d-center. N_i^M denotes the coordination number of a given species *i* that appears as a nearest neighbor (NN) to the target species M, while J_i^M represents the contribution of species *i* to the shift in the d-center of species *M* relative to the baseline value d_0^M . The term d_0^M is the constant carrying all of the unconsidered interactions. With this formulation, the d-center model can be parametrized by doing linear regression on the complete data set of calculated d-center values.

The cluster expansion approach has been widely used to predict various scalar properties of alloy-based systems. The successful application of cluster expansion to predict configurational energy,²⁹ thermal conductivity,³⁰ volume change,³¹ and band gap³² supports its feasibility for further adaptation to predict d-center values in MPEA systems. Particularly, as shown in eq 2, the singlet term describes the reference energy level of the d-electron at a certain composition, in analogy to the elemental chemical potential as being interpreted in conventional cluster expansion for total energy.^{11,33} The singlet

coefficient can be interpreted as the "reference chemical potential" of d electrons for a given metal. The pairwise term and potential high-order term can then be used to interpret the influence of local bonding configurations. The 1NN and 2NN coefficients describe how interactions with first- and second-nearest neighbors shift the d-center. A negative coefficient indicates that the interaction stabilizes the d-center, whereas a positive coefficient indicates destabilization. Therefore, the values of the fitting coefficients can serve as a guideline for further optimizing the chemical short-range order on the surfaces of synthesized catalysts.^{13–15,20,34,35}

The schematic of the featurization process is illustrated in Figure 2(a). Fivefold cross-validation was applied to the



Figure 2. (a) Schematic of how the feature vector is constructed, where the atomic fraction of each element and coordination numbers of 1NN metal elements surrounding a central metal site are used as input features; (b) the fitting result of all d-center values, with the orange and brown vertical dashed lines indicating the d-center value of FCC Pt (111) and FCC Ir (111), respectively. The data are scaled within the range of [-4 eV, 0 eV] to zoom in on the region of particular interest for finding sites close to Pt/Ir. Full range is shown in Figure S6. The black dashed lines illustrate a 0.5 eV shift to a perfect fit, which is used to guide the eye for the spread of the data.

regression model to evaluate various values of the regularization parameter in the L2 penalty term. The resulting learning curve is shown in Figure S3, and the validation MAE for each fold across different alpha values is summarized in Table S2. These results help demonstrate the robustness and generalizability of the regression model. The regression results of the final model that has all 1,220,008 d-centers are presented in Figure 2(b), while the results for the d-centers in the validation data set are shown in Figure S4. The elemental dependent fitting can be found in Figure S5, while the fitting with the full energy range can be found in Figure S6. Impressively, all data fit well within this linear model, yielding a mean absolute error (MAE) of 0.09 eV. We also trained the model with L1 regularization. As shown in Table S3 and Figure S7, L1 regularization does not surpass L2 regularization performance, except at very small penalty strengths, where it essentially reverts to ordinary least-squares fitting. As the penalty strength increases, L1 regularization drives more coefficients to zero (see Table S4), undermining the model's physical completeness; since the feature sets are mutually orthogonal, none should be eliminated. These results underscore the model's robustness and high predictive accuracy without the need for more complex regularization schemes.

Other effects, including lattice distortion, electron transfer, and orbital hybridization, are implicitly captured through the cluster expansion¹¹ of the d-center values. By expanding up to pair clusters, the model inherently includes all physical interactions within atomic pairs. The low MAE observed indicates that contributions from higher-order, many-body interactions are relatively minor, confirming that pairwise

interactions dominate the physics governing the d-center in this system. Furthermore, by using the same feature vector, we can also fit regression models on predicting other electronic descriptors such as d-width and achieve reasonable accuracy, as shown in Figure S8.

In addition to achieving numerical accuracy, the fitting procedure also provides valuable insights into the underlying factors that drive d-center shifts when a given metal is paired with different species. The fitted interaction coefficients, denoted as J_i^M in eq 2, are visualized in Figure 3. Most of these



Figure 3. Heatmap showing the fitting coefficient for contributions to d-center values for all metals evaluated. Mg and Al do not have d-electrons, so the corresponding rows are colored gray.

coefficients are negative, indicating that compositional disorder generally lowers the d-center values. It is important to note that the d-center value is typically defined as the first moment of the complete density of states (DOS) of d orbitals. Therefore, it is related only to the shape rather than the occupancy of d orbitals. The predominance of negative interaction coefficients suggests that the DOS of the d orbital is up-skewed, while the centroid is shifted downward. This phenomenon can be interpreted as a result of favorable chemical bonding between metals in the selected MPEAs. The formation of metal–metal bonds in these alloys tends to lower the electronic energy levels, reflecting increased bonding interactions. This observation aligns with the generally negative formation energies observed in most MPEAs, indicating that these mixed-metal systems are energetically more stable than their pure metal counterparts.

4. DISCUSSIONS

Even though the model presented in Figure 2 shows reasonable accuracy in capturing the d-center values, it remains uncertain whether further improvements are possible. In this section, we analyze in detail how surface structure and regression setup influence the final fitting accuracy.

4.1. Structural Dependency for the General Model. The structural orientation of surface atoms is determined by two physical factors: the bulk structure and the Miller index. It is therefore reasonable to hypothesize that the performance of the model described in eq 2 may vary depending on these structural characteristics. Figure 4(a) illustrates the three typical MPEA bulk structures considered in this study. When the linear model

pubs.acs.org/JACS

Article



Figure 4. (a) The three bulk atomic structures of our MPEA data set; (b) the parity plot by fitting the d-center independently for surfaces generated from three bulk structures; (c) demonstration of all Miller indices presented for FCC and BCC; (d) the parity plot by fitting the d-center independently for different surface configurations. All of the dashed diagonal lines in (b) and (d) illustrate a 0.5 eV shift to a perfect fit.

is fitted separately for surfaces generated for each bulk structure, only marginal improvements in MAE are observed: 0.09 eV for all BCC, 0.08 eV for all FCC, and 0.06 eV for all HCP, as shown in Figure 4(b), and performance on validation data is shown in Figure S9(a). Similarly, for FCC and BCC surface slabs, multiple Miller indices, including (110), (111), and (112), were sampled during the high-throughput DFT calculations, as shown in Figure 4(c). However, as shown in Figure 4(d), fitting the model based on a specific Miller index does not yield significant improvements in accuracy, with all MAEs remaining close to 0.08 eV. Figure S9(b) displays the performance solely on the basis of the validation data. These results suggest that the trend observed in Figure 2(b) is broadly applicable and largely independent of both the bulk structure and surface orientation. In contrast with prior reported attempts to fit various d-centers with different physical intuitions and formulisms,³⁶ our model shows the first time that a universal adaptable model can be generalized into arbitrary MPEA compositions.

4.2. Impact of Regression Method for the General Model. In addition to surface structure, another two key factors influencing model accuracy are the featurization and regression setup. As shown in Figure 2(a), the model illustrated in Figure 2(b) adopts the simplest configuration by considering only 1NN. While this approach already yields reasonable results, cluster expansion theory suggests that incorporating extended feature vectors, such as including interactions beyond 1NN, can generally enhance model performance. Furthermore, the model shown in Figure 2(b) does not differentiate between 1NN atoms located at the surface and those deeper in the bulk. As a result, secondary effects arising from variations in the local coordination environment are neglected. Finally, regularization has been shown to play a crucial role in ensuring the robustness of cluster expansion models,¹¹ especially when the available data are limited. All three of these factors-extended interactions, surface/bulk distinction, and regularization-are explored, with corresponding results summarized in Figure 5.

Figure 5(a) illustrates the possible variations in the regression setup. First, we compared the model performance by including or excluding the coordination number of second-nearest-



Figure 5. (a) Illustration of the top view of the 1NN (green circles) and 2NN (pink circles) of targeted atomic site (red circle) as the top panel; the atomic structure below illustrates the side view of a surface model with the atoms from the top row to bottom row being surface, subsurface, and bulk; illustrations of how the feature vector is concatenated for including different NN, and atomic sites are placed in the lower panels. (b) Donut histogram comparing the impact of 1NN and 2NN, as well as the surface/bulk distinction for each element. (c) Donut histogram comparing the impact of 1NN and 2NN, as well as the regression algorithm (least squares fitting/ L2 regularization). For both (b) and (c) the inner shell of the histograms indicates the model considering both 1NN and 2NN.

neighbor (2NN) atoms, as illustrated by the top panel. Since the feature vector used in Figure 2 did not distinguish the atomic environment at the surface or in the bulk, we developed an extended feature vector that treats the same species differently depending on its position within the surface slab (e.g., surface site, subsurface site, or bulk site). Lastly, the comparison between least-squares fitting and L2 regularization has also been made. Figure 5(b) demonstrates the elemental dependency fitting error when considering different NN setups as well as simple or extended feature vectors. In the extended feature vector, all atoms are categorized as surface, subsurface, or bulk atoms according to the schematic shown in Figure 5(a), while the simple feature vector considers atoms at all three positions equivalently. Particularly, the bars going inward toward the inner dashed circle show the fitting MAE when considering only 1NN, while the bars going outward to the outer dashed circle show the fitting MAE when considering both 1NN and 2NN. The unmeshed bars indicate the MAE from a simple feature vector, while the meshed bars indicate the MAE from extended feature vectors. The inner and outer dashed circles are the maximum fitting errors across all elements, which are set as 0.12 eV for both 1NN and 2NN. Detailed MAEs are listed in Table S5. Therefore, it can be inferred from Figure 5(b) and Table S5 that the inclusion of 2NN slightly improves the fitting, while the distinguishing of surface/bulk atoms does not have a significant impact on fitting MAEs.

A similar comparison is also made in Figure 5(c), where the inner and outer dashed circles also indicate the maximum error from fitting across all elements. Additionally, the circle hatch bars indicate the results from direct least-squares fitting, while the meshed bars represent results using linear regression based on a ridge, a type of L2 regularization. Detailed MAEs are listed in Table S6. Similar conclusions can be drawn from Figure 5(c), as the inclusion of 2NN has divergent roles in influencing fitting results, while the level of regularization seems to have little influence on the final MAE result. Based on the comparisons made in Figure 4 and Figure 5, it is thus clearer to conclude that even though there is still space for further improving the fitting, the improvement may not be very substantial.

The demonstrated model shows remarkable performance in capturing the d-center values of MPEA surfaces; it can also be further adapted into predicting low-entropy systems, particularly if more specialized training data are implemented. As shown in Table S7, the model shows reasonable capability when predicting the pure Pt. As discussed earlier, the shape of the DOS for the d orbital is largely influenced by the skewness of the d orbitals within the same metal, while the stability of chemical bonds plays a significant role. The stability can be indirectly linked to the chemical potentials of the elements, a property in which MPEAs differ considerably from pure metals and intermetallics. Despite the large and systematic training data set used for MPEAs, it does not include any data from the lowentropy regime. As a result, the model's applicability to pure metals and intermetallics is limited, even if they share the same crystal structure as MPEAs.

Another important aspect that needs attention is that the GGA level of theory is used in this work to calculate all of the dcenter values for fitting the presented model. The choice of GGA considers the reasonable trade-off between reliability and computational affordability. Caution is advised when transferring our fitted parameters to niche systems where GGA is known to be less reliable, e.g., some magnetic systems with Fe, Co, and Ni. In such cases, it is recommended to refit the model with high-fidelity calculations or even experimental data to capture the correct d-center values. This work provides comprehensive support of the validity of using cluster expansion to predict d-center values without any preference or bias on actual quantum chemistry method. Therefore, only a modest, system-specific data set is needed to update the coefficients, substantially reducing the overall computational effort.

Given that the presented model already accounts for the presence of various surface roughness (examples shown in Figure S10), Figure 4 demonstrates that the model's performance is not highly sensitive to either surface orientation or bulk structure but more directly connected to the actual nearest neighbor sites. We thus conclude that the model can be further extended to systems with various structural imperfections and defects, as well as nanostructures. The development of an efficient model for predicting d-center values will also facilitate further extensions to properties more directly related to catalytic performance, such as overpotentials in various catalytic reactions.

5. CONCLUSIONS

We present a general model for predicting d-center values that applies to any MPEA surface, regardless of elemental composition, bulk structure type, and Miller indices. With a comprehensive data set containing 1,220,008 d-centers from 10,680 surfaces, we demonstrate that predicting d-center values can be achieved with a simpler model than previously expected. This study will help accelerate the rapid screening of effective electrocatalysis.

ASSOCIATED CONTENT

Data Availability Statement

The data supporting the findings are available within the main text and SI. More detailed data can be made available upon reasonable request.

1 Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.5c05649.

Structural and surface-dependent d-center of Zn; crossvalidation by applying L2 regularization; the performance of validation data in regression fitting; elemental dependent fitting; fitting with full energy range; cross-validation by applying L1 regularization; linear fitting to other electronic descriptors (d-width); the validation MAE of the general models with various structures; the validation MAE when comparing various feature vector constructions for the model; the testing MAE when predicting pure Pt with various surfaces; the example of surface steps in the structures (PDF)

AUTHOR INFORMATION

Corresponding Author

Bin Ouyang – Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32304, United States; orcid.org/0000-0002-8181-6815; Email: bouyang@fsu.edu

Authors

- Lin Wang Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32304, United States
- Yizhan Zhang Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32304, United States
- Kayla James Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32304, United States; Orcid.org/0009-0005-0187-4732

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.5c05649

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Zhengda He for his valuable contributions to this research. The authors acknowledge support from the startup funding from Florida State University. This work is also supported by ACS-PRF. The computational resources are provided by the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science and the U.S. Department of Energy under contract no. DE-AC02-05CH11231, and Research Computing Center (RCC) at Florida State University. The computation and data processing are also supported by the supercomputing resources from the Department of Energy's Office of Energy Efficiency and Renewable Energy at the National Renewable Energy Laboratory.

REFERENCES

(1) Hammer, B.; Nørskov, J. K. Electronic factors determining the reactivity of metal surfaces. *Surf. Sci.* **1995**, 343 (3), 211–220.

(2) He, Z.; Wang, J.; Ouyang, B. Design Principle of Carbon-Supported Single-Atom Catalysts - Interplay between d-Orbital Periodicity and Local Hybridization. *Chem. Mater.* **2024**, *36* (3), 1405–1412.

(3) He, Z.; Ouyang, B. Charting the electronic structure for discovering low-cost intermetallic catalysts. *Journal of Materials Chemistry A* **2024**, *12* (33), 21987–21996.

(4) Resasco, J.; Abild-Pedersen, F.; Hahn, C.; Bao, Z.; Koper, M. T. M.; Jaramillo, T. F. Enhancing the connection between computation and experiments in electrocatalysis. *Nature Catalysis* **2022**, *5* (5), 374–381.

(5) Batchelor, T. A. A.; Pedersen, J. K.; Winther, S. H.; Castelli, I. E.; Jacobsen, K. W.; Rossmeisl, J. High-Entropy Alloys as a Discovery Platform for Electrocatalysis. *Joule* **2019**, *3* (3), 834–845.

(6) Zeng, Y.; Ouyang, B.; Liu, J.; Byeon, Y. W.; Cai, Z.; Miara, L. J.; Wang, Y.; Ceder, G. High-entropy mechanism to boost ionic conductivity. *Science* **2022**, *378* (6626), 1320–1324.

(7) Ouyang, B.; Zeng, Y. The rise of high-entropy battery materials. *Nat. Commun.* **2024**, *15* (1), 973.

(8) Wang, L.; He, Z.; Ouyang, B. Data driven design of compositionally complex energy materials. *Comput. Mater. Sci.* 2023, 230, 112513.
(9) Sun, Y.; Dai, S. High-entropy materials for catalysis: A new

frontier. *Sci. Adv.* **2021**, *7* (20), No. eabg1600. (10) Ouyang, B.; Chakraborty, T.; Kim, N.; Perry, N. H.; Mueller, T.;

Aluru, N. R.; Ertekin, E. Cluster Expansion Framework for the Sr(Ti1-xFex)O3-x/2 (0 < x < 1) Mixed Ionic Electronic Conductor: Properties Based on Realistic Configurations. *Chem. Mater.* **2019**, *31* (9), 3144–3153.

(11) Barroso-Luque, L.; Zhong, P.; Yang, J. H.; Xie, F.; Chen, T.; Ouyang, B.; Ceder, G. Cluster expansions of multicomponent ionic materials: Formalism and methodology. *Phys. Rev. B* **2022**, *106* (14), 144202.

(12) Sanchez, J. M.; de Fontaine, D. The fee Ising model in the cluster variation approximation. *Phys. Rev. B* **1978**, *17* (7), 2926–2936.

(13) Ouyang, B.; Artrith, N.; Lun, Z.; Jadidi, Z.; Kitchaev, D. A.; Ji, H.; Urban, A.; Ceder, G. Effect of Fluorination on Lithium Transport and Short-Range Order in Disordered-Rocksalt-Type Lithium-Ion Battery Cathodes. *Adv. Energy Mater.* **2020**, *10* (10), 1903240.

(14) Lun, Z.; Ouyang, B.; Cai, Z.; Clément, R. J.; Kwon, D.-H.; Huang, J.; Papp, J. K.; Balasubramanian, M.; Tian, Y.; McCloskey, B. D.; et al.

Design Principles for High-Capacity Mn-Based Cation-Disordered Rocksalt Cathodes. *Chem.* **2020**, *6* (1), 153–168.

(15) Lun, Z.; Ouyang, B.; Kwon, D. H.; Ha, Y.; Foley, E. E.; Huang, T. Y.; Cai, Z.; Kim, H.; Balasubramanian, M.; Sun, Y.; et al. Cationdisordered rocksalt-type high-entropy cathodes for Li-ion batteries. *Nat. Mater.* **2021**, *20* (2), 214–221.

(16) Huang, J.; Ouyang, B.; Zhang, Y.; Yin, L.; Kwon, D. H.; Cai, Z.; Lun, Z.; Zeng, G.; Balasubramanian, M.; Ceder, G. Inhibiting collective cation migration in Li-rich cathode materials as a strategy to mitigate voltage hysteresis. *Nat. Mater.* **2023**, *22* (3), 353–361.

(17) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59* (3), 1758–1775.

(18) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6* (1), 15–50.

(19) Wang, L.; Ouyang, B. Phase Selection Rules of Multi-Principal Element Alloys. *Adv. Mater.* **2024**, *36* (16), No. e2307860.

(20) Wang, L.; H, T.; Ouyang, B.. The Impact of Domain Knowledge on Universal Machine Learning Models. *ChemRxiv* **2024**.

(21) Ong, S. P.; Wang, L.; Kang, B.; Ceder, G. Li-Fe-P-O2 Phase Diagram from First Principles Calculations. *Chem. Mater.* **2008**, 20 (5), 1798–1807.

(22) Ouyang, B.; Wang, J.; He, T.; Bartel, C. J.; Huo, H.; Wang, Y.; Lacivita, V.; Kim, H.; Ceder, G. Synthetic accessibility and stability rules of NASICONs. *Nat. Commun.* **2021**, *12* (1), 5752.

(23) Wang, L.; Sunariwal, N.; He, Y.; Kim, D. h.; Yeon, D. h.; Zeng, Y.; Cabana, J.; Ouyang, B. Elemental Stability Rules for High Entropy Disordered Rocksalt Type Li-Ion Battery Positive Electrodes. *Adv. Energy Mater.* **2025**, *15*, 2404982.

(24) Wang, L.; Wang, J.; Ouyang, B. Computational Investigation of MAX as Intercalation Host for Rechargeable Aluminum-Ion Battery. *Adv. Energy Mater.* **2023**, *13* (46), 2302584.

(25) Periodic Table of Element Prices. 2019, https://www. compoundchem.com/wp-content/uploads/2019/12/23-%E2%80%93-Periodic-Table-of-Element-Prices.pdf (Accessed September 21, 2023).

(26) Singh, A. K.; Zhou, L.; Shinde, A.; Suram, S. K.; Montoya, J. H.; Winston, D.; Gregoire, J. M.; Persson, K. A. Electrochemical Stability of Metastable Materials. *Chem. Mater.* **2017**, 29 (23), 10159–10167.

(27) Persson, K. A.; Waldwick, B.; Lazic, P.; Ceder, G. Prediction of solid-aqueous equilibria: Scheme to combine first-principles calculations of solids with experimental aqueous states. *Phys. Rev. B* **2012**, *85* (23), 235438.

(28) Landrum, G. A.; Dronskowski, R. The Orbital Origins of Magnetism: From Atoms to Molecules to Ferromagnetic Alloys. *Angew. Chem., Int. Ed. Engl.* **2000**, *39* (9), 1560–1585.

(29) Samuely, P.; Szabó, P.; Kačmarčík, J.; Meerschaut, A.; Cario, L.; Jansen, A. G. M.; Cren, T.; Kuzmiak, M.; Šofranko, O.; Samuely, T. Extreme in-plane upper critical magnetic fields of heavily doped quasitwo-dimensional transition metal dichalcogenides. *Phys. Rev. B* 2021, 104 (22), 224507.

(30) Togo, A.; Chaput, L.; Tanaka, I.; Hug, G. First-principles phonon calculations of thermal expansion inTi3SiC2, Ti3AlC2, and Ti3GeC2. *Phys. Rev. B* **2010**, *81* (17), 174301.

(31) Zhao, X.; Tian, Y.; Lun, Z.; Cai, Z.; Chen, T.; Ouyang, B.; Ceder, G. Design principles for zero-strain Li-ion cathodes. *Joule* **2022**, *6* (7), 1654–1671.

(32) Xu, X.; Jiang, H. Cluster expansion based configurational averaging approach to bandgaps of semiconductor alloys. *J. Chem. Phys.* **2019**, *150* (3), 034102.

(33) Fontaine, D. D. Cluster Approach to Order-Disorder Transformations in Alloys. In *Solid State Physics*, Ehrenreich, H., Turnbull, D., Eds.; Vol. 47; Academic Press, 1994; pp 33–176.

(34) Li, L.; Ouyang, B.; Lun, Z.; Huo, H.; Chen, D.; Yue, Y.; Ophus, C.; Tong, W.; Chen, G.; Ceder, G.; et al. Atomic-scale probing of short-range order and its impact on electrochemical properties in cationdisordered oxide cathodes. *Nat. Commun.* **2023**, *14* (1), 7448. (35) Cai, Z.; Zhang, Y.-Q.; Lun, Z.; Ouyang, B.; Gallington, L. C.; Sun, Y.; Hau, H.-M.; Chen, Y.; Scott, M. C.; Ceder, G. Thermodynamically Driven Synthetic Optimization for Cation-Disordered Rock Salt Cathodes. *Adv. Energy Mater.* **2022**, *12* (21), 2103923.

(36) Nanba, Y.; Koyama, M. Physical Significance of Descriptors to Predict the Band Center of High-Entropy Nanoalloys. *J. Comput. Chem.* **2025**, 46 (8), No. e70086.