Acta Materialia 232 (2022) 117924

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat

Full length article

Efficient machine-learning model for fast assessment of elastic properties of high-entropy alloys

Guillermo Vazquez[a,*], Prashant Singh[a,b,*], Daniel Sauceda[a], Richard Couperthwaite[a], Nicholas Britt[b], Khaled Youssef[c], Duane D. Johnson[b,d], Raymundo Arróyave[a,e,f,*]

[a] Department of Materials Science and Engineering, Texas A&M University, College Station TX 77843, USA
[b] Materials Science and Technology Graduate Program, College of Arts and Sciences, Qatar University, Doha 2713, Qatar.
[c] Materials Science & Technology, Qatar University, Doha, Qatar
[d] Department of Materials Science & Engineering, Iowa State University, Ames, Iowa 50011, USA
[e] Department of Mechanical Engineering, Texas A&M University, College Station TX 77843, USA
[f] Department of Industrial and Systems Engineering, Texas A&M University, College Station TX 77843, USA

ARTICLE INFO

Article history:
Received 10 December 2021
Revised 9 March 2022
Accepted 4 April 2022
Available online 14 April 2022

Keywords:
Refractory high entropy alloys
Elastic properties
Machine learning
Descriptors
SISSO
Density-functional theory

ABSTRACT

We combined descriptor-based analytical models for stiffness-matrix and elastic-moduli with mean-field methods to accelerate assessment of technologically useful properties of high-entropy alloys, such as strength and ductility. Model training for elastic properties uses Sure-Independence Screening (SIS) and Sparsifying Operator (SO) method yielding an optimal analytical model, constructed with meaningful atomic features to predict target properties. Computationally inexpensive analytical descriptors were trained using a database of elastic properties determined from density functional theory for binary and ternary subsets of Nb-Mo-Ta-W-V refractory alloys. The optimal Elastic-SISSO models, extracted from an exponentially large feature space, give an extremely accurate prediction of target properties, similar to or better than other models, with some verified from existing experiments. We also show that electronegativity variance and elastic-moduli can directly predict trends in ductility and yield strength of refractory HEAs, and reveals promising alloy concentration regions.

1. Introduction

High Entropy Alloys (HEAs) are a novel class of materials with intriguing electronic properties [1,2] and often superior mechanical behavior [3,4] attributed to near-equiatomic (5–35 at.%) mixing of multi-principle elements [5–7]. The unique characteristic of single-phase solid-solution HEAs reveals a new and larger design space for complex solid-solution alloys [1,8–10]. Refractory-based HEAs (RHEAs), a special class of HEAs, show immense potential [11] that brings these alloys under increased focus for their superior electronic [12,13] and mechanical properties (like high-strength [14,15]), as well as very high melting temperature [11,16–18], fracture resistance [19], lower kinetic rates and increased long-term high-temperature stability [20]. RHEAs has expanded the search space needed for the discovery of viable candidates for a variety of current and anticipated applications requiring high ductility and fracture toughness, specific strength, and better mechanical performance at elevated temperatures.

Computational approaches (*e.g.*, density-functional theory (DFT), molecular dynamics (MD) and/or CALPHAD [Calculation of Phase Diagrams]) or even empirical rules are used in combination to establish relationships between atomic (*e.g.*, electronegativity, atomic mass, atomic radii, etc.) and alloy features (*e.g.*, configurational entropy and mixing enthalpy) to design or tune the application relevant properties of RHEAs [21–24]. However, excessive computational cost and uncertainty of first-principle methods limit the use of conventional approaches to explore exponentially large combinatorial design spaces (5×10^8 compositions considering 25 different elements) [25,26]. More recently, data-driven mehtods interfaced with machine-learning (ML) algorithms have enabled rapid filtering to reduce the vast alloying (Gibbs' compositional) space through fast-acting predictive models [26–28]. While successful, ML models oftentimes tend to be limited by the sparcity of the datasets available to train them [29,30]. On other occasions, ML-based approaches have resulted in models that are difficult to interpret and are

* Corresponding authors.
E-mail addresses: guillermo.vazquez@tamu.edu (G. Vazquez), psingh84@ameslab.gov (P. Singh), raymundo.arroyave@tamu.edu (R. Arróyave).

https://doi.org/10.1016/j.actamat.2022.117924
1359-6454/

thus mostly 'black boxes' establishing connections between inputs and outputs through highly non-linear and convoluted mappings.

Here, we construct predictive models of the intrinsic mechanical behavior of RHEAs to determine critical trends that are key ML training information. Notably, if we consider the space of atomic features and an arbitrary set of operators used to integrate them, the dimensionality of such a problem can easily run into $10^8 - 10^9$ range. The discovery of the most informative feature subspace thus becomes an intractable problem using conventional dimensional-reduction approaches. Based on notions of Compressed Sensing (CS) [31], which solves for the sparse solution to an undetermined systems of equations, SISSO [32–34] uses Sure-Independent Screening [35] (SIS), which selects at each iteration a new subset of features. Further reduction in the feature space and the development of a predictive model is carried out through Sparsifying Operators (SO), such as Least Absolute Shrinkage and Selection Operator (LASSO) [36] or l_0-norm regularized minimization in order to find the most sparse solution to a linear problem. The SISSO-based modeling is a form of symbolic regression, which uses trained analytical descriptors for property prediction. SISSO has already been shown to be a powerful automated-feature engineering (AFE) framework that enables the construction of accurate, interpretable predictive models for materials behavior, such as energy stability of inorganic solids [37,38] as well as oxidation behavior in ceramics [39], among others.

Knowledge of mechanical behavior is critical for the design of technologically useful alloys, however, the main focus of ML models has been on thermodynamic stability with scant attention on developing fast and inexpensive models for mechanical properties [40] due to unavailability of reliable databases. Elastic constants (*e.g.*, stiffness matrix (C_{ij})) and other engineering quantities (such as yield-strength (σ_{0y})) are extremely useful for materials design, yet, RHEA optimization has taken other design paths involving the use of simple (empirical) relations of atomic and thermodynamic properties [29,41–44]. Mechanical properties derived from elastic parameters can provide valuable insight into a RHEA's brittleness, stiffness, anisotropy, ductility, bonding character, and strength. However, our knowledge about features affecting important engineering quantities, such as strength, is very limited due to lack of efficient computational techniques. Therefore, methods or models capable of accurately predicting fundamental engineering quantities would greatly benefit the alloy design.

Here, we present descriptor-based analytical models, i.e., "Elastic SISSO", for the fast exploration of mechanical properties over the vast HEA space, which will also be useful as didactic tool for showcasing simple correlations between target properties—features identified from this method can in turn be used to train more complex 'black box' models, if so desired. We establish that dominant factors, such as high yield strength, derived from ML-predicted elastic parameters of body-centered cubic (bcc) RHEAs, consisting of 3*d* (V), 4*d* (Nb, Mo), and 5*d* (Ta,W) transition-metal elements, can be directly related to fundamental elemental quantities, *e.g.*, electronegativity variance (χ_{var}), atomic-size differences (δ), and formation enthalpy (E_{form}). Specifically, for Ta-W-Nb-Mo-V RHEA systems, we also discuss the relationship between phase stability, local environment correlation, and mechanical properties of alloying components. While there is no consensus on a threshold that describes the ductility in HEAs using Pugh's ratio [45], Poisson's ratio, or Cauchy parameter [46], to some extent all of these quantities relate to the ductility of the alloy and, with a clear trend established (and experimentally validated in subset of systems), they can be employed for design. We trained, tested, and developed descriptor-based analytical models using SISSO-based ML to predict mechanical properties of solid-solutions. This systematic study for various mechanical properties will benefit material science community in accelerating search for technological useful HEAs.

Different models and designs had been proposed to predict or induce ductility in RHEAs, for example, whether intrinsic or extrinsic, design approaches have been taken to overcome brittleness [47]. Notably, the Valence Electron Concentration (VEC) has also been linked to ductility, for example, Qi et al. studied intrinsic ductility in BCC alloys by comparing shear instability and crack initiation using DFT [48] by optimizing average VEC [49]. Chen et al. also provided distinctions between different ranges of VECs [50].

2. Methods

2.1. DFT elastic database

The stiffness matrix of binary and ternary Nb-Mo-Ta-W-V based alloys were calculated using DFT-based stress-strain approach [51,52], as implemented in the plane-wave pseudo-potential Vienna Ab-initio Simulation Package (VASP) [53,54]. Both special quasirandom structure (SQS) [55] and supercell random approximates (SCRAPs) [56] methods were used to mimic homogeneously random disordered alloys. The Perdew, Burke, and Ernzerhof (PBE) generalized gradient approximation [57] to DFT was used with a energy cut-off of 520 eV. Full (volume and ionic) optimization and charge self-consistency were done on ($2 \times 2 \times 2$) and ($4 \times 4 \times 4$) Monkhorst-Pack [58] $k-$mesh for Brillouin zone integration, respectively. A high level of convergence criteria was set for elastic property calculation in both energy and forces, i.e., 10^{-6} eV and 10^{-6} eV/Å.

As supercells used have differing structure along x, y, and z axes, when elastic distortions are applied in DFT calculations the elastic constants C_{11}, C_{22}, and C_{33}, for example, will be slightly different, rather than exactly equal due to average cubic symmetry for a homogeneous solid solution. Hence, we average appropriately (see below). Here, we applied 12 independent deformations to each investigated structure, one each for the normal and shear strains and two strains with opposing directions. Combination of binary and a ternary sets were chosen to map the stiffness data on the 5-dimensional HEA space. For binaries, ten subsystems were sampled using a 60-atom supercell in steps of 10% (with a single sub-step of 5% at composition (0.25,0.75)), while ternaries were sampled using a 64-atom supercell with base composition of (0.125,0.125,0.75) and (0.25,0.375,0.375). The equivalent compositions were generated by shuffling the atom positions for both binary and ternaries making it 10 different cases per binary and ternary, which makes up to a total of 170 combinations (110 binary and 60 ternary). The macroscopic value of the elastic stiffness coefficients was then approximated by averaging the three independent elastic constants C_{ii} and C_{ij} (with i≠j) in a cubic crystalline system [59,60], i.e., $C_{11}^{avg}=(C_{11}+C_{22}+C_{33})/3$, $C_{44}^{avg}=(C_{44}+C_{55}+C_{66})/3$, and $C_{12}^{avg}=(C_{12}+C_{13}+C_{23})/3$. The bulk (K)-, shear (G)-, and Young's (E)-moduli, Poisson's ratio (ν) were approximated using the Voight-Reuss-Hill approximation[61] as detailed by Wu et al. for the cubic system [62].

2.2. Phase stability analysis

High-throughput formation enthalpy calculation for Nb-Ta-Mo-W-V HEAs were performed using DFT-based Green's function Korringa-Kohn-Rostoker electronic-structure methods [63,64], in which the coherent-potential approximation accounts properly for averaging over all chemically disorder configurations in structures [65]. The gradient-corrected Perdew, Burke, and Ernzerhof (PBE) exchange-correlation functional [57] was used within each atomic site for charge distributions and the total energy. A semi-circular

Table 1
ML-predicted RHEAs with higher strength (σ_{0y}; GPa) compared to equiatomic MoNbTaW and MoNbTaVW. Results are arranged in increasing density.

MPEAs	Mo	Nb	Ta [at.%]	V	W	ρ [g/cc]	C_{11} [GPa]	E [GPa]	K [GPa]	σ_{0y} [GPa]	σ_{1300C} [GPa]	G/K
MoNbTaW	25	25	25	-	25	13.68	344	224	235	1.23	0.22	0.33
MoNbTaVW	20	20	20	20	20	12.17	320	197	223	1.76	0.35	0.31
RHEA0	19.99	0.70	30.02	44.91	4.38	10.69	285	160	205	1.842	0.29	0.268
RHEA1	19.57	8.83	13.76	40.82	17.02	10.83	311	182	218	1.807	0.32	0.295
RHEA2	17.68	9.08	10.84	41.31	21.09	10.99	317	187	220	1.781	0.32	0.302
RHEA3	4.67	15.32	6.17	42.79	31.05	11.43	312	182	220	1.811	0.32	0.293
RHEA4	8.64	1.94	14.54	46.28	28.60	11.82	318	192	223	1.767	0.31	0.302
RHEA5	16.84	2.09	34.60	34.75	11.72	12.05	298	178	213	1.913	0.35	0.285
RHEA6	3.17	23.98	5.89	28.22	38.74	12.56	328	200	229	1.864	0.38	0.317
RHEA7	10.31	8.63	12.87	31.20	36.99	12.98	343	219	235	1.870	0.40	0.334
RHEA8	23.31	1.31	32.20	22.91	20.27	13.17	335	218	230	1.897	0.41	0.324
RHEA9	1.07	0.78	40.59	26.80	30.76	14.51	313	202	226	1.942	0.41	0.300

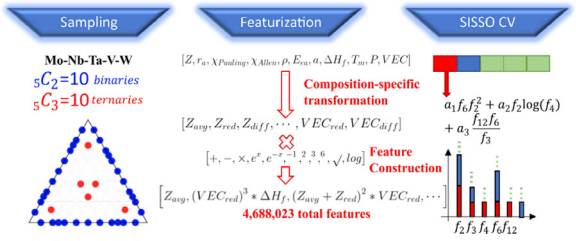

Fig. 1. Schematic diagram for sampling, feature construction, and feature analysis processes.

contour and Gauss-Laguerre quadratures using 24 complex energies was used for integration. A 24× 24 × 24 Monkhort-Pack [58] *k*-mesh was used for Brillouin-zone integration of bcc Nb-Ta-Mo-W-V HEAs.

2.3. Misfit volume calculation for strength prediction

The misfit volume of the alloys can be calculated as the derivatives of atomic volume of elements in HEAs with respect to composition as $\Delta V_n = \frac{\partial V_{\text{HEA}}}{\partial c_n} - \sum_{m=1}^{N} c_m \frac{\partial V_{\text{HEA}}}{\partial c_m}$. The expression can be derived from the expression $\sum_n c_n = 1$ using conditions $V_{\text{HEA}} = V_{\text{HEA}}(c_1, c_2, \ldots, c_{N-1})$ and $\partial V_{\text{HEA}}/\partial c_N = 0$ [66].

3. Results

3.1. Elastic SISSO - featurization

Featurization of the alloy sampling is carried out via a stoichiometric approach, extrapolating elemental properties. Bartel et al. proposed these alloy-specific transformations in their Gibb's energy model whereby three different features are retrieved for each elemental property, using the stoichiometrically-weighted average (avg), the stoichiometrically-weighted harmonic mean (red), and the stoichiometrically mean difference (diff) of the atomic feature in question [37], which are showcased in supplemental Table S3.

We use atomic properties: atomic number, period, atomic radius, Pauling (molecular) and Allen (solid-state) electronegativity, density, heat of formation, BCC lattice constant, melting point, electron affinity and valence electron count (*i.e.*, $Z, P, r_a, \chi_{Pauling}, \chi_{Allen}, \rho, \Delta H, a, T_m, E_{ea}$, and VEC). These 11 elemental features – once transformed – make up a 33-dimensional primary alloy feature space are fed into the SISSO based AFE framework. The feature construction was carried out by applying the operator set $\left[+, -, \times, e^x, e^{-x}, ^{-1}, ^2, ^3, ^6, \sqrt{}, log\right]$ to the 33 alloy-specific features. As shown in schematic Fig. 1, all features were created by combination of the primary features and operators, which were recursively added to the feature set two times.

3.2. Training, model-generation, and cross-validation

During each SISSO iteration, the SIS method selects 1000 features, which are added to a subset to which the l_0 norm regularized minimization Sparsifying Operator is applied so as to obtain the best linear regressor with a number of terms equal to the current iteration out of this subset by doing this every iteration (1*D*, 2*D*, and 3*D* descriptors). Consequently, for the last descriptors $\binom{3000}{3} \sim 4.5 \times 10^9$ combinations were tested. Out of the 3 descriptors (1D, 2D, and 3D), the 3D descriptor shows the best accuracy, both in training and cross-validation (CV) estimation, which indicates that over-fitting, a recurring caveat in sparse regressor, may start occurring at dimensions higher than 3.

The SISSO models used in this work were put to test in a couple of different ways: (i) a 5-fold/10-fold CV to understand the over-fitting and frequency analysis of features found in CV, and (ii) the descriptors trained on full data. Thus, we only discuss the most complete descriptors and apply them here for further sampling of mechanical properties over complex composition space. As the systems have cubic symmetry, the SISSO model were retrained for two independent elastic constants [C_{11}; C_{12}] along with (K; G; E; ν; (inverse) Pugh's ratio, G/K; C_{44}) that can be derived from Voigt-Reuss-Hill approximations in terms of C_{11} and C_{12} [61].

3.3. Elastic parameter C_{11} descriptor - cross validation and error analysis

$$C_{11} = 11.6 + 1.14\left[VEC_{avg}^3\sqrt{ln(P_{avg})}\right] + 47.7\left[VEC_{diff}^{36}\right] - 0.029a_{avg}^6\left[VEC_{red} + VEC_{diff}\right]$$

The parity plot in Fig. 2a shows a good predictability even for a small number of folds. The 5-fold CV results for the stiffness matrix parameter C_{11}, for which the coefficient of determination indicates that an independent SISSO descriptor trained in 80% of the data is able to determine the rest 20% consistently. The inset graph shows the error distribution for the model and an approximation for the probability density, which follows a normal distribution. While for the evaluation of these results five different analytical models were used, Eq. (1) shows a final model using all data available for the extrapolation to a higher dimension space in the next section.

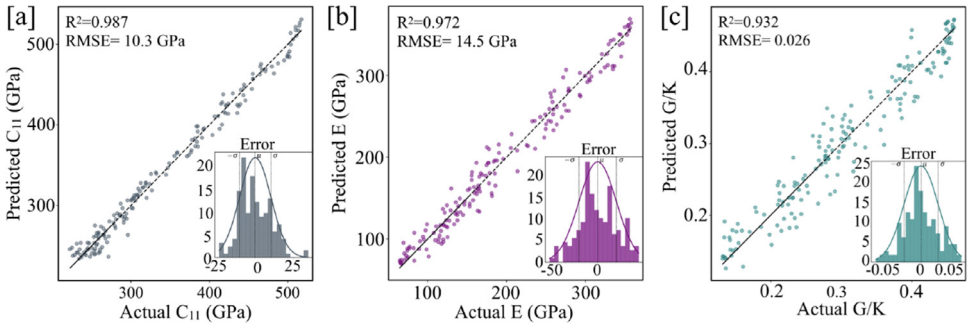

Fig. 2. Actual (DFT) vs predicted (ML) (a) stiffness parameter (C_{11}; GPa), (b) Young's moduli (E; GPa), and (c) (inverse) Pugh's ratio (PR=G/K) from five-fold cross-validation test. Inset in a-c shows the error distribution around the mean, which is virtually zero for all the three cases.

3.4. Young's moduli (E) descriptor - cross validation and error analysis

$$E = -469 + 1.31\frac{VEC_{avg}T_{m,avg}}{\sqrt{\Delta H_{f,avg}}} - 0.335\frac{VEC_{red}\Delta H_{f,avg}}{VEC_{avg}-VEC_{diff}} + 1.37\frac{VEC_{diff}}{E_{ea,diff}\chi^2_{Pauling,diff}} \quad (2)$$

Fig. 2 b corresponds to the results for the Young's moduli (E), the coefficients of determination for E is overall lower than for C_{11}. In the case shown here, the error distribution for the 5-fold CV once again has nearly a zero mean error. The final model is shown in Eq. (4), which in comparison appears to be more connected to the composition-specific features of the electronegativities.

3.5. Pugh's ratio (PR) descriptor - cross validation and error analysis

Lastly, the inverse Pugh's ratio model in Fig. 2c is performing with less accuracy than the C_{11} and E. Evaluating a single model, rather than two, adds uncertainty instead of performing at the lowest accuracy of the two models G and K (see Table. S1). The final analytical equation for Pugh's ratio is shown in Eq. (3).

$$PR = 3.82 - 0.293\left[\frac{\chi_{Allen,avg}a_{avg}}{\sqrt{VEC_{avg}}}\right] - 0.0191\left[\frac{exp(VEC_{diff})}{\chi_{Pauling,avg}E_{ea,red}}\right] - 0.105\left[\frac{\rho_{red}}{E_{ea,diff}\chi_{Pauling,diff}\Delta H_{f,avg}}\right]$$

While Bartel et al. trained a SISSO descriptor for the thermodynamic properties of crystalline compounds by supplying compound-specific properties, the final descriptor for Gibbs' energy only chose the reduced mass from these transformed features, besides calculated atomic volume and temperature [37]. Descriptors shown here are dependent solely on compound-specific features, we hypothesize the complexity in elastic properties of binary and ternary compounds can be captured by non-linear features arranged in a linear combination. The complexity of the models is then assumed to translate to higher dimensions. Elastic stiffness constants naturally arise from tensor analysis of the strain-stress response of materials and the factors governing this response get as complex as the number of constituents grows. The SISSO-chosen descriptors shown here deviate enough from the linearity of the common rule-of-mixtures estimation to deliver a better approximation as the complexity of the alloy grows. The compromise for accuracy when using this descriptor is then compared to results of DFT.

3.6. Target property descriptor analysis

If a feature is selected in each term of five three-dimensional (3D) SISSO equations in 5-fold cross-validation test, then the highest frequency a feature can get is 15. To better understand these results, we present a descriptor analysis in Fig. 3, which consists of counting features by appearance inside a term for every equation and displaying 10 features with highest frequency. The most frequent features are the one those consistently appear in the analytical 3D models. The final equations may resemble the alloy-specific features chosen by the CV method, however, this selection of features is not guaranteed in other cases.

Final models (Eqs. (1)–(3)) and CV analysis within SISSO framework show some common alloy-specific features. For example, majority of descriptors include the VEC and Pauling electronegativity as a shared feature. The feature analysis in Fig. 3 is provided to understand feature presence and different mathematical operations in SISSO models. Moreover, the ranking of the atomic properties by appearance in the cross-validation method is not analogous to feature importance since they are not the operator coupled features used in SISSO, but it gives a qualitative idea about the basic building blocks of the descriptors, i.e., the composition-weighted features.

While it appears that the accuracy is correlated to the confidence of the CV in selecting the same alloy-specific features as in analytical SISSO models, this is not necessarily true for each case. For example, C_{11} and E have higher confidence in their top features in contrast to the Pugh's ratio model. Yet, the diverse features selected in each run are inherently correlated in that they create multiple top performing descriptors once coupled with the sparsyfying operators in SISSO method. Fig. 3 reveals the elemental features such as VEC selected in each model, which was found most significant for predicting elastic properties and consistently appears in the cross-validation.

Analytical models such as those available in Eq. (1)–(3) and Table S1 are easy to use, which will expand new non-data-intensive approaches for rapid space exploration of complex materials including high-entropy alloys. For engineering purposes where the ductility is an important performance factor, an inexpensive analytical model can be useful for composition optimization from large alloy design space.

4. Application to the quinary alloy

4.1. Interpretation of target properties and its interdependence on physical features

Unlike other 'black-box' ML methods, it is possible to create a direct relation with a specific feature from the get-go by analyzing the analytical function. We can go one step forward and evaluate a

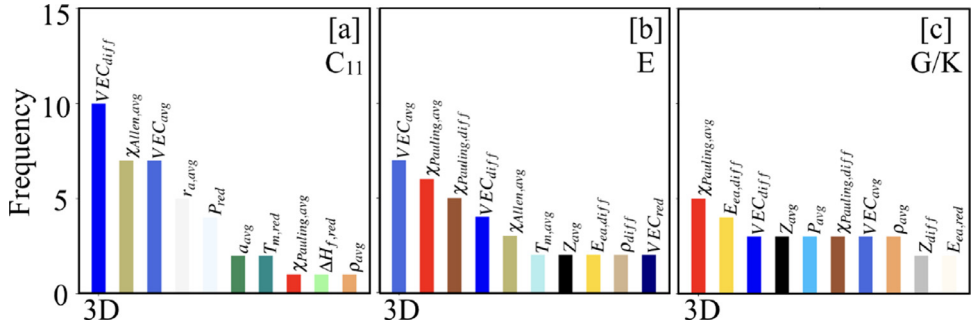

Fig. 3. (a-c) Descriptor analysis ranks feature appearance in C_{11}, E, and Pugh's ratio in the CV method.

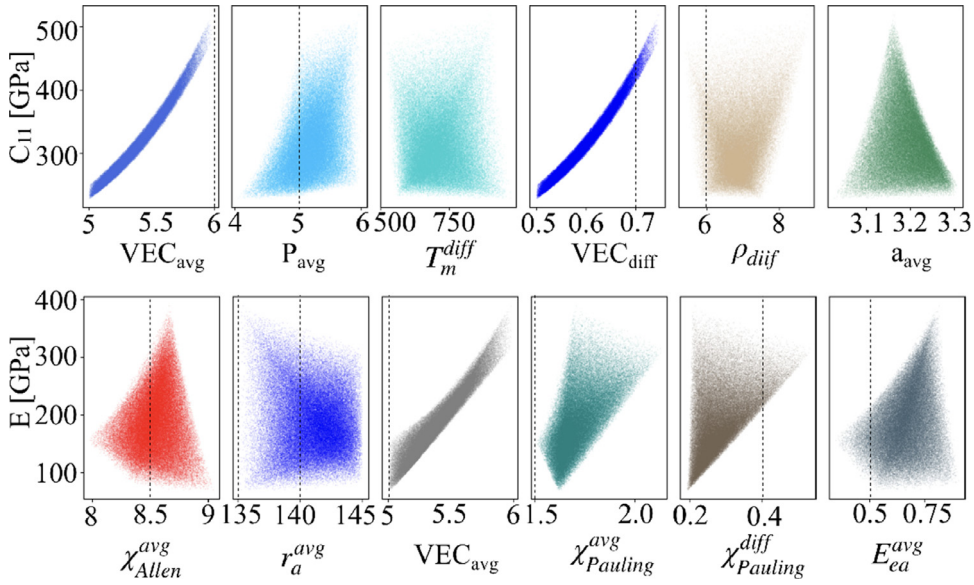

Fig. 4. Chosen features (top-panel) stiffness matrix element (C_{11}), and (bottom-panel) Young's moduli (E) descriptor.

quinary random sampling and reveal how a subset of features (for each model, e.g., C_{11}, and E) contributes to the ML models given in Fig. 4. These analytical models for elastic parameters were evaluated over 100,000 points in the Mo-Nb-Ta-V-W RHEA space.

Stiffness constant C_{11}: Analytical model for the C_{11} in Eq. (1) strongly depends on the VEC, as obvious from Fig. 4 (top-panel). The symmetry of the system constrains the values of all the weighted transformations for VEC to be completely correlated, i.e., the relation of C_{11} with the average, harmonic average and difference-average is always the same. Therefore, the average of the periods (in periodic table) are correlated to some extent with the number of electrons, especially, when only looking at five different elements distributed over 3 different periods. This feature is present in the final descriptor, but only harmonic average of periods appeared during cross-validation. The period in this system is constrained to integer values, which indicates an even stronger correlation between different averaging methods. Here, the average melting point difference between lowest and highest points indicate a higher C_{11}, so not only high differences between constituents yields a high C_{11}, but also small difference stabilizes the high stiffness. For example, a maximum in C_{11} was obtained for the optimal range of average lattice parameter (via Vegard's rule), i.e., 3.15Å and 3.20Å . In supplemental Fig. S3, we provide a short discussion on the usefulness of descriptor-based analytical approaches for property prediction in complex alloys over that of simple rules. The observations that the difference on electronegativity correlates well to mechanical properties of HEAs in this work agrees well with existing literature [67].

Young's modulus E: Descriptor for E in Eq. (2) strongly depends on VEC and χ, as shown in Fig. 4 (bottom-panel). As expected, average electronegativities have different correlation to E due to the difference in their physical origin. Allen average electronegativity has an optimum value for E, while Pauling has an overall positive correlation. The $\chi^{diff}_{Pauling}$ shows good correlation with $\chi^{avg}_{Pauling}$ [see Fig. 4 (top-panel)] except for the onset where $\chi^{diff}_{Pauling}$ converges at single point while $\chi^{avg}_{Pauling}$ shows wide distribution. Moreover, we need to be cautious here because former represents the averages of absolute χ and later is average of difference in χ. On the other hand, the electron affinity shows a higher range of stiffness at elevated values, and VEC shows a similar correlation with C_{11}. This last statement is true for all features, and E and C_{11} are so closely correlated that the shape of feature-target correlations are similar.

The average of lattice constant has proven to deviate by a small margin from the experimental lattice parameter, which follows a negative slope after a maximum value of 3.17Å for both C_{11} and E. Gorban et al. [68] also observed the same trend, which was attributed to size mismatch and weakening atomic interactions at a larger lattice parameter. A positive trend in modulus of elasticity at a higher electron count has been observed, *i.e.*, $6e^-$ [68,69], SISSO models presented in this work show similar behavior.

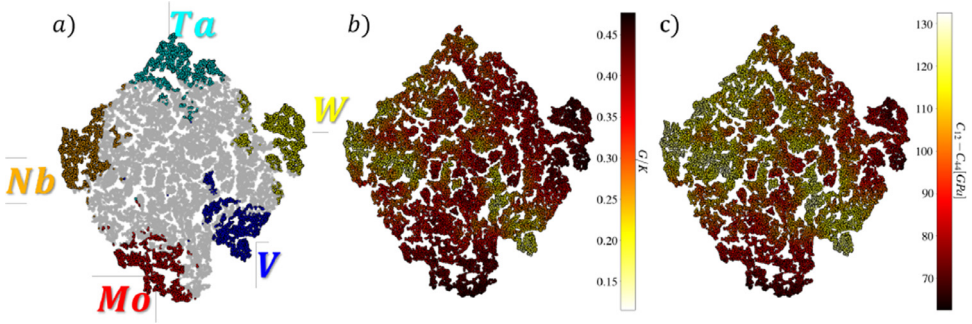

Fig. 5. (a) t-SNE dimensional reduction of the quinary space with highlighted unary-enriched zones, the color scheme indicates that this point contains more than 50% molar of this respective element, (b) Pugh's ratio, and (c) Cauchy pressure $C_{12}-C_{44}$ mapped in top of t-SNE, as predicted by the SISSO models.

4.2. Pugh's ratio – evaluating derived parameters from stiffness matrix

The inverse Pugh's ratio (G/K) and Cauchy pressure ($C_{12}-C_{44}$) are empirically defined quantities, which are commonly used to predict the ductility of complex alloys. Originally, Pugh [45] discussed the ratio of (G/K $<$ 0.57 or K/G $>$1.75) and found it correlated with the ductility of elemental metals (fcc, bcc, and hcp), where bulk moduli (K) assesses the resistance to fracture while shear modulus (G) the tendency for increased fracture resistance after the onset of plastic deformation. Indeed, Gschneidner Jr. et al. used existing experimental data from large number of materials to categorize brittle-to-ductile transition at G/K $<$ 0.57, [70] where the lower (higher) ratio indicates ductility (brittleness). Here, we use the inverse Pugh's ratio (PR=G/K) simply to span a $0 \leq PR \leq 1$ range, and ductility is indicated by G/K $<$ 0.57. For HEAs, the randomly sampled 100,000 compositions in the quinary alloy space (composition vector) were mapped on a two-dimensional space using a t-distributed Stochastic Neighbor Embedding (t-SNE) scheme in Fig. 5a, where the clustering algorithm separates the unary-rich composition. The percentage of composition is more evenly distributed away from t-SNE corners in Fig. 5a. The (inverse) Pugh's ratio (G/K) in Fig. 5b is below the anisotropy-specific ductile-to-brittle critical point, as evaluated using the descriptor in Eq. (3). The HEAs have a positive Cauchy pressure ($C_{12}-C_{44} < 0$), as shown in Fig. 5c.

Recently, Senkov et al. [71] expanded the correlation between these two quantities by taking the ratio of Cauchy pressure and bulk moduli. The anisotropy-specific critical value for the Pugh's ratio in Fig. 5b was found in good agreement with Pettifor's condition for incipient brittleness, i.e., $C_{12}-C_{44} < 0$. Although this would be considered ductile for elements and simple compounds, the ductility of HEAs is more complex, and such predictions must be seen as qualitative. Physically, the Pugh's ratio predicted for RHEAs in Fig. 5b shows the ability of an alloy to deform, where smaller value (G/K $<$ 0.57) shows higher deformability (ductile) while higher value (G/K $>$ 0.57) signifies lower deformability (brittle). Models for Pugh's ratio and Cauchy pressure show similar trends, i.e., the minimum for one in the quinary space is the maximum for the other. Notably, both models predict high ductility for HEAs in the same composition region. The model also shows a soft gradient between consecutive alloys in the t-SNE. Similarly, Pugh's ratio follows expected positive correlation with *VEC*, meaning that the alloy system is expected to show higher ductility as the VEC_{avg} decreases. Both the bulk moduli (K) and shear moduli (G) have a positive slope with VEC_{avg}, whereas *G* has a steeper increment. This suggests that the alloy tends to get harder for compression or shear. Yet, the shear mechanism will be more affected as the bond energy of the crystal becomes stronger. The average and the difference of the lattice parameter show different correlations to the Pugh's ratio.

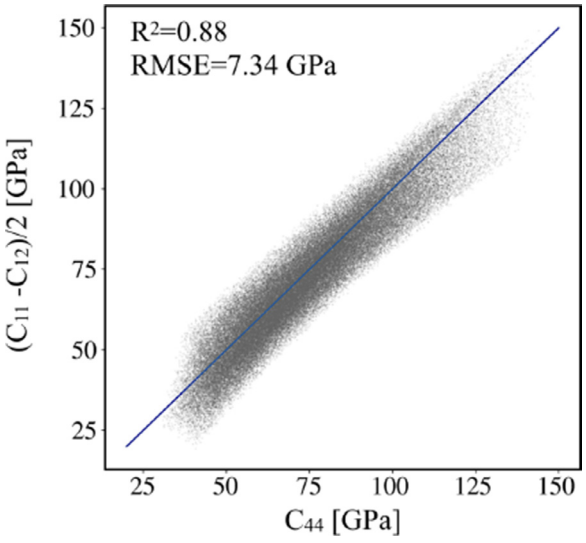

Fig. 6. The descriptor predicted C_{44} was compared with C_{44} estimated using independent elastic parameters, i.e., $C_{44} \approx (C_{11}-C_{12})/2$, assuming zero anisotropy of disorder bcc Bravais lattice.

4.3. SISSO prediction of C_{44} and isotropy dilemma

DFT methods often do not provide accurate estimate of shear constant (C_{44}) in isotropic disordered solids such bcc and fcc alloys. However, the C_{44} can be directly estimated using the independent stiffness constants (C_{11} and C_{12}) in a system with cubic symmetry, assuming isotropic behavior. To exemplify this, we show in Fig. 6 a comparison between the ML predicted C_{44} and others derived from relation $[C_{11}-C_{12}]/2$. Notably, the C_{44} derived from the isotropic assumption correlates well to ML predictions. The small prediction error of 7.34 GPa (RMSE) in C_{44} corresponds to an average deviation of 8.6% from the model.

4.4. High-throughput prediction of high strength HEAs and its origin

HEAs are expected to not only match the current state-of-the-art structural materials properties of conventional as well complex alloys, but they are expected to have lower density, optimal stiffness, i.e., low brittleness or higher ductility. However, the large composition space of HEAs restricts the exploration of physical space beyond equiatomic compositions. We show in Fig. 7 that our computationally inexpensive analytical ML models fa-

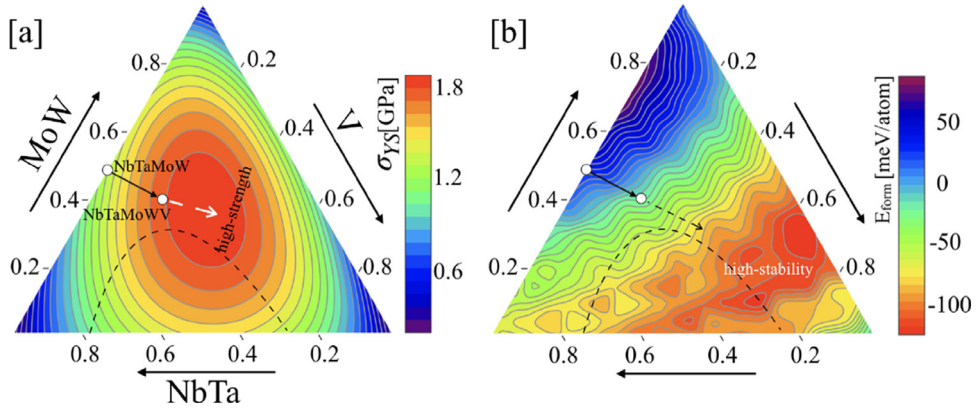

Fig. 7. (a) Yield-strength (σ_{0y}) and (b) formation enthalpy (ΔH_{form} at 0 K) vs composition (c_i, i=Nb/Ta, Mo/W, V). The parabolic-dashed line shows the high electronegativity variance ($\Delta\chi_{var}$), while arrow-head from NbTaMoW to NbTaMoWV RHEA (solid-white circle) shows experimentally observed increasing trend. Our observation suggests that its not the single parameter but the interplay of different quantities, such as chemical entropy, $\Delta\chi_{var}$, and ΔH_{form}, that are responsible for higher strength.

cilitates quick search of technologically useful alloys over multi-dimensional spaces.

The SISSO predicted analytical model of stiffness matrix (C_{ij}) (Eq. (1) and see supplemental Table S1) were used to estimate the zero-temperature flow stress (τ_{y0}), or critically resolved shear stress for edge dislocations in a BCC matrix, using the reduced-order model developed by Maresca and Curtin [66]. The main parameters ruling the strength model are the elastic constants and the misfit volume, the latter is calculated through Vegard's 'law' using experimental values of atomic volumes. Our DFT results for solid-solution and associated elemental volumes could be used but the critically resolved shear stress is related more to the yield-strength of the material through the Taylor factor. Therefore, for the current work, we used calculated σ_{y0} for the yield-strength at 0 Kas well as the approximation for finite-temperature ($1300^{\circ}C$) and finite strain-rate (0.001), which builds up on the 0K approximation. Our estimates of σ_{0y} in Fig. 7a suggests towards a region in (Nb-Ta)-(Mo-W)-V compositions space that shows higher strength, which is different from high mixing entropy (ΔS_{mix}) region (marked by solid white dot) doesn't necessarily correlate with mechanical or electronic behavior. In contrast to the usual expectations, however, we found that the region of high yield-strength in Fig. 7a directly correlates with higher thermodynamic stability (ΔH_{mix}) as shown in Fig. 7b. The increasing trends (ML predicted) in yield-strength from quaternary NbTaMoW (1.12 GPa) to quinary NbTaMoWV (2.00 GPa) matches with experiments [72,73], where adding V was found to increase the strength of the alloy.

The dashed parabolic region marked in Fig. 7a,b shows the region of high electronegativity variance ($\Delta\chi_{var}$) (also see supplemental Fig. S1a). The $\Delta\chi_{var}$ increases with increase in V and Nb/Ta more compared to Mo/W because of higher V electronegativity compared to others. We hypothesize that the large $\Delta\chi_{var}$ may lead to strong charge imbalance in the neighboring environment of V as large χ pulls more charge, which adds local lattice distortion and hence provide local solid-solution strengthening. The strong correlation between σ_{0y} in Fig. 7a and ΔH_{mix} in Fig. 7b possibly arises from V addition that enhances the local-lattice distortion and the alloy stability, as discussed by Song et al. [12] and Singh et al. [56] Our assessment suggests that optimal combination of entropy (as a proxy for alloy complexity), phase stability, and $\Delta\chi_{var}$ results into higher yield-strength.

The Young's moduli (E) is another important design characteristics that shows the tensile stiffness of materials such as HEAs. General convention suggests that maximizing E should also maximize the strength, e.g., high σ_{0y}. To understand this, we plot E vs σ_{0y} [in GPa] in Fig. 8a that shows a optimal E-region (150<E<250) in which σ_{0y} is higher (marked by dashed lines), however, beyond the marked region, σ_{0y} falls off quickly. We found an inverse relation between E in Fig. 8b and σ_{0y} in Fig. 8c with $\Delta\chi_{var}$, i.e., E (σ_{0y}) decreases (increases) with increasing (decreasing) $\Delta\chi_{var}$. This confirms our idea of finding optimal tensile stiffness range rather than maximizing it while designing new RHEAs.

From our analysis, we predict new RHEA compositions as shown in Table 1, where C_{11}, E, K, τ_{0y}, and σ_{YS} are tabulated along with chemical composition of each alloying element. The RHEAs are arranged in order of increasing density (ρ in g/cc), which becomes critical depending on area of application. Notably, predicted RHEAs with higher σ_{YS} directly correlates with higher V (20+ at.%) concentration, this further emphasizes our hypothesis of optimal concentration of V in RHEAs improves both phase stability and mechanical behavior. This can be attributed to the fact that higher Allen-scale electronegativity of V pulls charges from neighboring sites and creates strong solid-solution strengthening through local lattice distortions. The size effect (δ) plot in supplemental Fig. S2b also shows the higher effect in the optimal region with higher σ_{YS} marked in Fig. 7d. Moreover, alloys with higher V concentration in predicted RHEAs show increase in energy stability, see supplemental Table S2.

Discussion on ductility: The criteria to evaluate the ductile behavior based on elastic constants in BCC HEA has been reported at several instances [45,74,75]. However, it is impossible that all material properties or all type of alloys can be fit using same kind of model or approach. Nonetheless, the Pugh ratio offers good correlation with ductility of materials [45], and, as noted above, supported by existing experimental data from large number of materials [70]. Indeed, this is reflected in our data, see Figs. 8a and 9a. Poisson's ratio ν, considering the relationship between ν and (G/K), has the potential to assess the ductile behavior of MPEAs (Fig. 9b), where the brittle-to-ductile transition limit corresponds to $\nu > 0.26$ [76]. Similarly, the large positive Cauchy pressure in Fig. 9c represents an increased degree of the metallic bonding that correlates with ductility [76], whereas the negative value is indicative of brittleness (or more covalent bonding) [77]. Recently, Lee et al. [76] have reported ductility in NbTaTiV RHEA, which further supports our idea that state-of-the-art, ab-initio calculations and machine-learning approaches may be helpful for refractory-based alloys. This also establishes the quality of prediction of high-

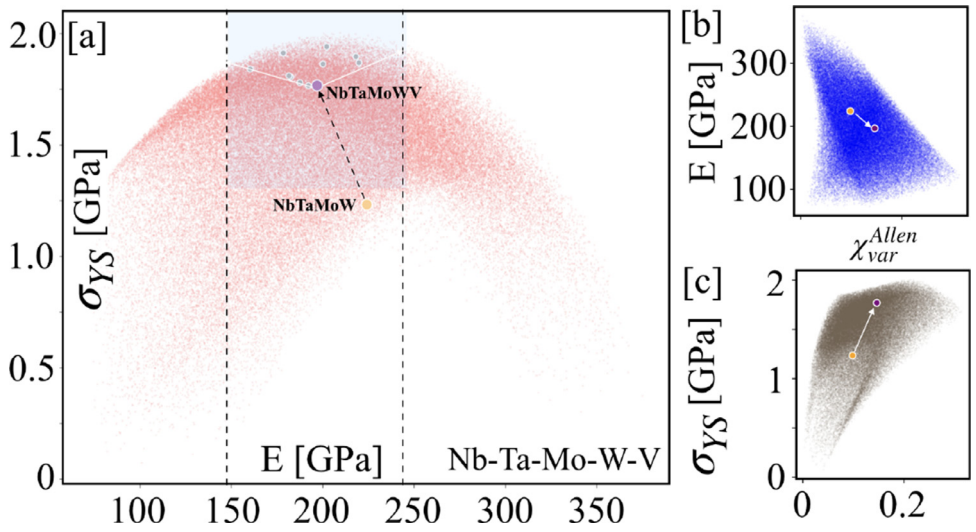

Fig. 8. (a) Compositions with higher yield-strength (σ_{0y}; 1.5< σ_{0y} <2.25 GPa) away from high-entropy region (see Table 1) are highlighted vs optimal Young's moduli range (E; 150 < E < 250 GPa). (b,c) Electronegativity variance (χ_{var}) was found to show strong correlation with (E, σ_{0y}).

Table 2
The C_{11}, C_{12}, C_{44}, bulk-moduli, and elastic-moduli of diverse HEAs (some outside the training chemistry) predicted from the SISSO model show good agreement with DFT.

System	C_{11}		C_{12}		C_{44}		K		E	
	ML	DFT	ML	DFT	ML	DFT	ML	DFT	ML	DFT
MoNbTaVW	321	328	172	169	74	79	223	222	197	213
CrMoNbTaV	315	330	160	159	77	85	213	216	206	226
CrMoNbTaW	370	378	172	168	99	105	241	238	244	275
CrMoNbVW	363	372	169	170	97	101	236	237	235	265
CrMoTaVW	372	393	171	170	100	111	242	244	248	290
CrNbTaVW	322	360	166	156	78	102	222	224	204	266
NbTaTiVZr	171	190	150	104	10	43	146	132	116	116

Table 3
The misfit-volume and strength calculation with different elemental misfits results into expected deviations.

	ΔV_{Mo}	ΔV_{Nb}	ΔV_{Ta}	ΔV_{V}	ΔV_{W}	$\sum_n c_n V_n^2$	$\tau_{y.0}$ (GPa)	C_{11}	C_{12}	C_{44}
Mo-Nb-Ta-V-W										
Maresca and Curtin's reduced model	-0.628	1.713	1.877	-2.484	-0.478	2.650	0.610	346.8	157.7	90.5
Vegard's law (misfit factor)/SISSO (elastic constants)	-0.681	1.760	1.810	-2.424	-0.465	2.586	0.576	320.8	172.0	73.0
Mo-Nb-Ta-W										
Maresca and Curtin's reduced model	-1.293	1.135	1.168	-	-1.010	1.336	0.450	375.5	167.3	101.6
Vegard's law (misfit factor)/SISSO (elastic constants)	-1.287	1.154	1.204	-	-1.071	1.396	0.402	344.2	182.0	86.2

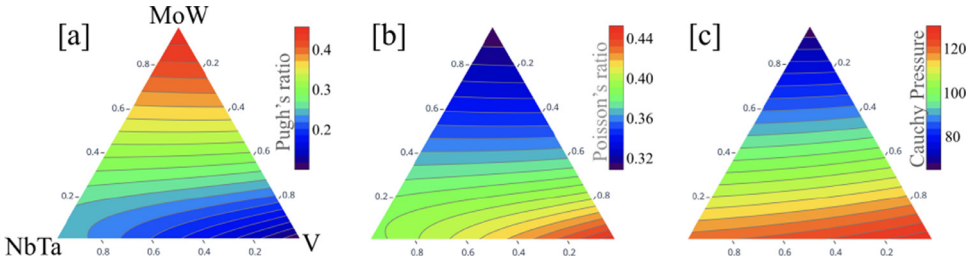

Fig. 9. We plot the (a) Pugh's ratio, (b) Poisson's ratio, and (c) Cauchy pressure (C_{12}-C_{44} in GPa) that highlights the optimal region of strength lies slightly away from high-entropy region.

strength regions in Fig. 7a.

4.5. Application of the proposed SISSO model to systems outside the training chemistry

In Table 2, we demonstrate the application of the SISSO models to several stoichiometric HEAs with chemistries outside the chemical space used in the model training. As shown in Table 2, we found good agreement between descriptor predicted elastic constants and DFT. If we focus on ML vs DFT results in Table 2, error in prediction for most cases is within 1–5%. Notably, the C_{ij} for NbTaTiVZr is the only exception that shows large error, i.e., C_{11} (171 vs 190 GPa, i.e., ∼10% error), C_{12} (150 vs 104 GPa, i.e., ∼44% error), C_{44} (10 vs 43 GPa, i.e., ∼77% error). In our limited understanding to this aspect, we believe the possible reason in under predicting C_{ij}'s can be two-fold - (i) Ti/Zr not included in the train-

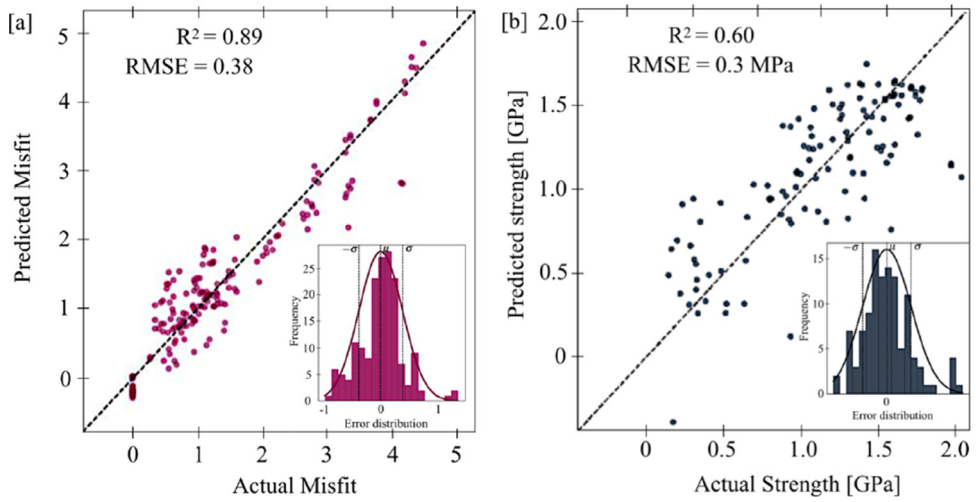

Fig. 10. Actual vs predicted (a) misfit volume, and (b) strength (σ; GPa) from ten-fold cross-validation test. Inset in a,b shows the error distribution around the mean, which is virtually zero for both the cases.

ing datasets, and (ii) crystal anisotropy of Ti/Zr (hexagonal parent phase). Here, we mainly attribute this fact to crystal anisotropy of Ti/Zr as C_{ij} are strongly direction dependent, which is not included in our training dataset. Despite minor disagreements, these results are promising, as they tentatively suggest that the model is capable of making safe extrapolations. We want to emphasize that more insights are needed into the extrapolation of the SISSO based descriptors and the chosen featurization for the complete generalization of the proposed descriptor. A more thorough investigation, beyond the scope of this work, is necessary to ascertain the ability of this model to extrapolate, with reasonable accuracy, beyond the training chemistry.

4.6. Analytical model for misfit-volume and yield-strength

While Maresca and Curtin [66] estimated the misfit volume using experimental values, all the numbers and details are not clearly stated. We emphasize that the misfit volume calculated with different atomic volumes are expected to show deviations when used with the reduced strength model, as shown in Table 3. Another source of uncertainty comes from elastic constants determined from the rule of mixing and the SISSO descriptors. However, descriptors-based analytical models are more close to reality as all possible alloying effects are included, while the rule of mixtures do not.

We developed the analytical SISSO model to predict alloy misfit volume factor (see method section) trained on same binary and ternary, which is similar to analytical descriptors in Eqs. (1)–(3). The misfit volume descriptor can directly be used with Maresca and Curtin's reduced strength model [66]. In Fig. 10a, the 10-fold cross-validation test shows the accuracy of 0.89 (R^2) with an RMSE of 0.38. While this is a good value for predictions within the training data, the uncertainty transmitted to the strength model may grow larger. The model presented in Eq. (4) shows that accurate analytical models for misfit volumes can be developed, however, we need more accurate high-fidelity data from experiments or ab-initio methods.

$$V^{misfit} = -0.0654 - 93800\frac{a_{diff}}{\chi^{avg}_{Pauling}\sqrt{a_{avg}}} + 376000\frac{a_{diff}}{\chi^{avg}_{Pauling}(a_{avg}+a_{red})} - 132000\frac{a_diff}{\chi^{avg}_{Pauling}a_{avg}a_{red}}$$

We further emphasize that the ML-trained strength models for MPEAs are not so readily available in literature due to the lack of the training dataset. However, we could develop a sparse strength database for HEAs consisting of 114 experimental HEAs (fcc, bcc, and fcc+bcc). The featurization used was also alloy specific but simpler, using only averages and thermodynamical parameters. A 10-fold cross-validation test was used to showcase the overall confidence of the SISSO model in Fig. 10b. An error using the 3D descriptor is prone to overfitting as cross-validation accuracy dropped to 0.60 from 0.74 (trained on all whole dataset). We believe that the possibility of large prediction error comes from limited experimental data with unknown experimental conditions, sample type and quality, and other unknown critical quantities. This indicates towards the need of more reliable dataset creation from well-defined experimental sources and conditions for robust machine-learning strength model. Our work also points out the fact that for intelligent design in materials discovery, i.e., in absence of reliable data, mean-filed approaches can be useful for accelerated filtering of multi-dimensional hyperspace such as high-entropy alloys.

Caveats in DFT vs ML-predicted shear moduli: In refractory-based alloys, the shear or trigonal elastic parameter (C_{44}) of individual elements, especially, Nb is underestimated in most electronic-structure methods (DFT: 17 GPa [78]; Expt: 31 GPa [79]), independent of DFT methods or exchange-correlation functional [80]. This issue has been discussed at length in literature, *e.g.*, by Koci et al. [78], who attribute it to transverse-phonon mode-driven anomalous band dispersion; more recently, as DFT typically freezes core states, it was associated with warping of the low-lying core levels during shearing, which was not addressed previously. Nonetheless, this issue limits the reliable data generation for RHEAs. And, interestingly, our SISSO-trained analytical model shows better accuracy for Nb, i.e., 24 GPa compared to experiments (31 GPa) [79].

5. Summary and discussion

In this work, descriptor-based machine-learning (ML) framework models were developed to efficiently scan and predict HEA elastic properties. The power of descriptor-based analytical models for fast exploration of the alloy space was exemplified for refractory based Nb-Ta-Mo-W-V HEAs. The reliable, optimal, and interpretable analytical descriptors were trained with SISSO based ML method on an elastic property database obtained from density-functional theory calculations. A detailed analysis of target properties was also carried out to correlate common elemental/alloy

features for optimized descriptor to better interpret proposed analytical models, distinctly different from black-box ML models.

The descriptor-predicted stiffness matrix (C_{ij}) of Nb-Ta-Mo-W-V HEAs were used to assess technologically useful quantities, such as yield-strength, that identify high-strength regions that correlate more with optimal combination of entropy (not high-entropy region), regions with large size effect (δ), large electronegativity variance (χ_{var}), and regions of high phase stability (lower formation energy). Our predicted trends match limited existing experiments, further establishing inexpensive descriptor-based methods can accelerate design of technologically useful alloys.

Elastic relations, like trends in Cauchy stability (C_{12}-$C_{44} < 0$), show similar behavior as yield-strength in the quinary space suggesting a more direct approach to estimate ductility using analytical descriptors. To further emphasize this point, we note that the alloy compositions with excellent mechanical properties is not necessarily those with highest chemical entropy or valence-electron count. For example, compositions with high-strength and high phase-stability (lower formation enthalpy) were found in lower-entropy regions, again questioning the focus on maximizing entropy to achieve better mechanical behavior.

Our results emphasize that computationally inexpensive models are important for thorough and accurate search of the vast HEA composition space to identify regions with desirable target properties. For that, an unconstrained search of the alloy space permits a further possibility to optimize HEA compositions. The application of the model to chemistries (beyond those used to train it) are promising, although additional work is needed to ensure the safety of such extrapolations. Regardless, the model is considered to be useful for further design of the Nb-Ta-Mo-W-V based systems with superior mechanical proeprties.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and supplement. Also, the data that support the plots within this paper and other findings of this study are available from the corresponding authors upon reasonable request. Supporting data and code for all figures generated using descriptors in Eq. (1)–(3) are available at *Github* [81].

Declaration of Competing Interest

The manuscript entitled “Efficient machine-learning model for fast assessment of elastic properties of high-entropy alloys” by G. Vazquez et al has not been published previously and is not under consideration for publication elsewhere. This manuscript is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

The authors declare no competing interests.

Acknowledgements

ARPA-E ULTIMATE project *Batch-wise Improvement in Reduced Design Space using a Holistic Optimization Technique* (BIRDSHOT) under primary contract No. DE-AR0001427 is acknowledged. GV, KY and RA acknowledge the support of QNRF under Project No. NPRP11S-1203-170056. RC was supported by NSF-CMMI Grant No. 1663130. DS was supported by NSF Grant No. 1545403. NB was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI). Methods developed at Ames Laboratory (DDJ and PS) and used herein were supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Science & Engineering Division. Ames Laboratory operated by Iowa State University for the U.S. DOE under contract DE-AC02-07CH11358. Calculations were conducted using the advanced computing resources provided by Texas A&M High-Performance Research Computing.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2022.117924.

References

[1] R. Zhang, S. Zhao, J. Ding, Y. Chong, T. Jia, C. Ophus, M. Asta, R.O. Ritchie, A.M. Minor, Short-range order and its impact on the CrCoNi medium-entropy alloy, Nature 581 (7808) (2020) 283–287, doi:10.1038/s41586-020-2275-z.
[2] P. Singh, A.V. Smirnov, D.D. Johnson, Ta-Nb-Mo-W refractory high-entropy alloys: anomalous ordering behavior and its intriguing electronic origin, Phys. Rev. Mater. 2 (2018) 055004, doi:10.1103/PhysRevMaterials.2.055004.
[3] Z. Li, K.G. Pradeep, Y. Deng, D. Raabe, C.C. Tasan, Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off, Nature 534 (7606) (2016) 227–230, doi:10.1038/nature17981.
[4] S.S. Sohn, A.K.d. Silva, Y. Ikeda, F. Körmann, W. Lu, W.S. Choi, B. Gault, D. Ponge, J. Neugebauer, D. Raabe, Ultrastrong medium-entropy single-phase alloys designed via severe lattice distortion, Adv. Mater. 31 (8) (2019) 1807142, doi:10.1002/adma.201807142.
[5] J.W. Yeh, S.K. Chen, J.Y. Gan, S.J. Lin, T.S. Chin, T.T. Shun, C.H. Tsau, S.Y. Chang, Formation of simple crystal structures in Cu-Co-Ni-Cr-Al-Fe-Ti-V alloys with multiprincipal metallic elements, Metall. Mater. Trans. A 35 (2004) 2533–2536, doi:10.1007/s11661-006-0234-4.
[6] J.W. Yeh, S.K. Chen, S.J. Lin, J.Y. Gan, T.S. Chin, T.T. Shun, C.H. Tsau, S.Y. Chang, Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes, Adv. Eng. Mater. 6 (5) (2004) 299–303, doi:10.1002/adem.200300567.
[7] B. Cantor, I.T. Chang, P. Knight, A.J. Vincent, Microstructural development in equiatomic multicomponent alloys, Mater. Sci. Eng. A (2004) 213–218, doi:10.1016/j.msea.2003.10.257. 375-377 (1-2 SPEC. ISS.)
[8] N. Ishizu, J. Kitagawa, New high-entropy alloy superconductor Hf21Nb25Ti15V15Zr24, Results Phys. 13 (2019) 102275. https://www.sciencedirect.com/science/article/pii/S2211379719307764.
[9] E.P. George, D. Raabe, R.O. Ritchie, High-entropy alloys, Nat. Rev. Mater. 4 (8) (2019) 515–534, doi:10.1038/s41578-019-0121-4.
[10] Y. Ikeda, B. Grabowski, F. Körmann, Ab initio phase stabilities and mechanical properties of multicomponent alloys: a comprehensive review for high entropy alloys and compositionally complex alloys, Mater. Charact. 147 (2019) 464–511. https://www.sciencedirect.com/science/article/pii/S1044580318306636.
[11] D. Miracle, O. Senkov, A critical review of high entropy alloys and related concepts, Acta Mater. 122 (2017) 448–511. https://www.sciencedirect.com/science/article/pii/S1359645416306759.
[12] H. Song, F. Tian, Q.-M. Hu, L. Vitos, Y. Wang, J. Shen, N. Chen, Local lattice distortion in high-entropy alloys, Phys. Rev. Mater. 1 (2017) 023404, doi:10.1103/PhysRevMaterials.1.023404.
[13] P. Singh, S. Gupta, S. Thimmaiah, B. Thoeny, P.K. Ray, A. Smirnov, D.D. Johnson, M.J. Kramer, Vacancy-mediated complex phase selection in high entropy alloys, Acta Mater. 194 (2020) 540–546. https://www.sciencedirect.com/science/article/pii/S1359645420303396.
[14] Y. Zhang, T.T. Zuo, Z. Tang, M.C. Gao, K.A. Dahmen, P.K. Liaw, Z.P. Lu, Microstructures and properties of high-entropy alloys, Prog. Mater. Sci. 61 (2014) 1–93. https://www.sciencedirect.com/science/article/pii/S0079642513000789.
[15] S. Maiti, W. Steurer, Structural-disorder and its effect on mechanical properties in single-phase TaNbHfZr high-entropy alloy, Acta Mater. 106 (2016) 87–97. https://www.sciencedirect.com/science/article/pii/S1359645416300179.
[16] O.N. Senkov, J.M. Scott, S.V. Senkova, F. Meisenkothen, D.B. Miracle, C.F. Woodward, Microstructure and elevated temperature properties of a refractory TaNbHfZrTi alloy, J. Mater. Sci. 47 (9) (2012) 4062–4074, doi:10.1007/s10853-012-6260-2.
[17] O.N. Senkov, D.B. Miracle, K.J. Chaput, J.P. Couzinie, Development and exploration of refractory high entropy alloys - a review, J. Mater. Res. 33 (19) (2018) 3092–3128, doi:10.1557/jmr.2018.153.
[18] P. Singh, A. Sharma, A.V. Smirnov, M.S. Diallo, P.K. Ray, G. Balasubramanian, D.D. Johnson, Design of high-strength refractory complex solid-solution alloys, npj Comput. Mater. 4 (1) (2018) 16, doi:10.1038/s41524-018-0072-0.
[19] B. Gludovatz, A. Hohenwarter, D. Catoor, E.H. Chang, E.P. George, R.O. Ritchie, A fracture-resistant high-entropy alloy for cryogenic applications, Science 345 (6201) (2014) 1153–1158. https://science.sciencemag.org/content/345/6201/1153.
[20] S.-Y. Chang, C.-E. Li, Y.-C. Huang, H.-F. Hsu, J.-W. Yeh, S.J. Lin, Structural and thermodynamic factors of suppressed interdiffusion kinetics in multi-component high-entropy materials, Sci. Rep. 4 (1) (2014) 4162, doi:10.1038/srep04162.

[21] Y. Zhang, Y. Zhou, J. Lin, G. Chen, P. Liaw, Solid-solution phase formation rules for multi-component alloys, Adv. Eng. Mater. 10 (6) (2008) 534–538, doi:10.1002/adem.200700240.

[22] S. Guo, C. Ng, J. Lu, C.T. Liu, Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys, J. Appl. Phys. 109 (10) (2011) 103505, doi:10.1063/1.3587228.

[23] B. Steingrimsson, X. Fan, X. Yang, M.C. Gao, Y. Zhang, P.K. Liaw, Predicting temperature-dependent ultimate strengths of body-centered-cubic (bcc) high-entropy alloys, npj Comput. Mater. 7 (1) (2021) 152, doi:10.1038/s41524-021-00623-4.

[24] Z. Pei, J. Yin, J.A. Hawk, D.E. Alman, M.C. Gao, Machine-learning informed prediction of high-entropy solid solution formation: beyond the Hume-Rothery rules, npj Comput. Mater. 6 (1) (2020) 50, doi:10.1038/s41524-020-0308-7.

[25] R. Feng, P.K. Liaw, M.C. Gao, M. Widom, First-principles prediction of high-entropy-alloy stability, npj Comput. Mater. 3 (1) (2017) 50, doi:10.1038/s41524-017-0049-4.

[26] S.Y. Lee, S. Byeon, H.S. Kim, H. Jin, S. Lee, Deep learning-based phase prediction of high-entropy alloys: optimization, generation, and explanation, Mater. Des. 197 (2021) 109260, doi:10.1016/j.matdes.2020.109260.

[27] J.M. Rickman, G. Balasubramanian, C.J. Marvel, H.M. Chan, M.T. Burton, Machine learning strategies for high-entropy alloys, J. Appl. Phys. 128 (22) (2020) 221101, doi:10.1063/5.0030367.

[28] Z. Zhou, Y. Zhou, Q. He, Z. Ding, F. Li, Y. Yang, Machine learning guided appraisal and exploration of phase design for high entropy alloys, npj Comput. Mater. 5 (1) (2019) 128, doi:10.1038/s41524-019-0265-1.

[29] Y. Zhang, C. Ling, A strategy to apply machine learning to small datasets in materials science, npj Comput. Mater. 4 (1) (2018) 25, doi:10.1038/s41524-018-0081-z.

[30] M. de Jong, W. Chen, R. Notestine, K. Persson, G. Ceder, A. Jain, M. Asta, A. Gamst, A statistical learning framework for materials science: application to elastic moduli of k-nary inorganic polycrystalline compounds, Sci. Rep. 6 (1) (2016) 34256, doi:10.1038/srep34256.

[31] L.M. Ghiringhelli, J. Vybiral, E. Ahmetcik, R. Ouyang, S.V. Levchenko, C. Draxl, M. Scheffler, Learning physical descriptors for materials science by compressed sensing, New J. Phys. 19 (2) (2017) 023017.

[32] R. Ouyang, S. Curtarolo, E. Ahmetcik, M. Scheffler, L.M. Ghiringhelli, SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates, Phys. Rev. Mater. 2 (8) (2018) 83802, doi:10.1103/PhysRevMaterials.2.083802.

[33] R. Ouyang, E. Ahmetcik, C. Carbogno, M. Scheffler, L.M. Ghiringhelli, Simultaneous learning of several materials properties from incomplete databases with multi-task SISSO, J. Phys. Mater. 2 (2) (2019) 24002, doi:10.1088/2515-7639/ab077b.

[34] R. Ouyang, A data-driven method combining symbolic regression and compressed sensing toward accurate & interpretable models, 2017, https://github.com/rouyang2017/SISSO.

[35] J. Fan, J. Lv, Sure independence screening for ultrahigh dimensional feature space, J. R. Stat. Soc. Ser. B 70 (5) (2008) 849–911.

[36] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B 58 (1) (1996) 267–288.

[37] C.J. Bartel, S.L. Millican, A.M. Deml, J.R. Rumptz, W. Tumas, A.W. Weimer, S. Lany, V. Stevanović, C.B. Musgrave, A.M. Holder, Physical descriptor for the Gibbs energy of inorganic crystalline solids and temperature-dependent materials chemistry, Nat. Commun. 9 (1) (2018), doi:10.1038/s41467-018-06682-4. arXiv:1805.08155

[38] C.J. Bartel, C. Sutton, B.R. Goldsmith, R. Ouyang, C.B. Musgrave, L.M. Ghiringhelli, M. Scheffler, New tolerance factor to predict the stability of perovskite oxides and halides, Sci. Adv. 5 (2) (2019) arXiv:1801.07700, doi:10.1126/sciadv.aav0693.

[39] D. Sauceda, P. Singh, A.R. Falkowski, Y. Chen, T. Doung, G. Vazquez, M. Radovic, R. Arroyave, High-throughput reaction engineering to assess the oxidation stability of max phases, npj Comput. Mater. 7 (1) (2021) 6, doi:10.1038/s41524-020-00464-7.

[40] B. Yin, W.A. Curtin, First-principles-based prediction of yield strength in the RhIrPdPtNiCu high-entropy alloy, npj Comput. Mater. 5 (1) (2019) 14, doi:10.1038/s41524-019-0151-x.

[41] X. Yang, Y. Zhang, Prediction of high-entropy stabilized solid-solution in multi-component alloys, Mater. Chem. Phys. 132 (2–3) (2012) 233–238, doi:10.1016/j.matchemphys.2011.11.021.

[42] A. Takeuchi, A. Inoue, Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element, Mater. Trans. 46 (2005) 2817–2829, doi:10.2320/matertrans.46.2817.

[43] S. Guo, C. Liu, Phase stability in high entropy alloys: formation of solid-solution phase or amorphous phase, Prog. Nat. Sci. 21 (6) (2011) 433–446, doi:10.1016/S1002-0071(12)60080-X.

[44] S. Guo, C. Ng, J. Lu, C.T. Liu, Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys, J. Appl. Phys. 109 (10) (2011) 103505, doi:10.1063/1.3587228.

[45] S. Pugh, XCII. Relations between the elastic moduli and the plastic properties of polycrystalline pure metals, London Edinb. Dublin Philos.Mag. J. Sci. 45 (367) (1954) 823–843, doi:10.1080/14786440808520496. https://www.tandfonline.com/action/journalInformation?journalCode=tphm20.

[46] D.G. Pettifor, Theoretical predictions of structure and related properties of intermetallics, Mater. Sci. Technol. 8 (4) (1992) 345–349, doi:10.1179/mst.1992.8.4.345. https://www.tandfonline.com/action/journalInformation?journalCode=ymst20.

[47] S.W. Wu, G. Wang, Y.D. Jia, J. Yi, Q.J. Zhai, C.T. Liu, B.A. Sun, H.J. Chu, J. Shen, P.K. Liaw, T.Y. Zhang, Enhancement of strength-ductility trade-off in a high-entropy alloy through a heterogeneous structure, Acta Mater. 165 (2019) 444–458, doi:10.1016/j.actamat.2018.12.012.

[48] L. Qi, D.C. Chrzan, Tuning ideal tensile strengths and intrinsic ductility of bcc refractory alloys, Phys. Rev. Lett. 112 (11) (2014), doi:10.1103/PhysRevLett.112.115503.

[49] S. Sheikh, S. Shafeie, Q. Hu, J. Ahlström, C. Persson, J. Veselý, J. Zýka, U. Klement, S. Guo, Alloy design for intrinsically ductile refractory high-entropy alloys, J. Appl. Phys. 120 (16) (2016) 164902, doi:10.1063/1.4966659.

[50] R. Chen, G. Qin, H. Zheng, L. Wang, Y. Su, Y.L. Chiu, H. Ding, J. Guo, H. Fu, Composition design of high entropy alloys using the valence electron concentration to balance strength and ductility, Acta Mater. 144 (2018) 129–137, doi:10.1016/j.actamat.2017.10.058.

[51] Y. Le Page, P. Saxe, Symmetry-general least-squares extraction of elastic data for strained materials from ab initio calculations of stress, Phys. Rev. B Condens. Matter Mater.Phys. 65 (10) (2002) 1–14, doi:10.1103/PhysRevB.65.104104.

[52] S. Shang, Y. Wang, Z.K. Liu, First-principles elastic constants of α- and θ-Al2O3, Appl. Phys. Lett. 90 (10) (2007) 9901, doi:10.1063/1.2711762.

[53] G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals, Phys. Rev. B 47 (1993) 558–561, doi:10.1103/PhysRevB.47.558.

[54] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B 59 (1999) 1758–1775, doi:10.1103/PhysRevB.59.1758.

[55] S.H. Wei, L.G. Ferreira, J.E. Bernard, A. Zunger, Electronic properties of random alloys: special quasirandom structures, Phys. Rev. B 42 (1990) 9622, doi:10.1103/PhysRevB.42.9622.

[56] R. Singh, A. Sharma, P. Singh, G. Balasubramanian, D.D. Johnson, Accelerating computational modeling and design of high-entropy alloys, Nat. Comput. Sci. 1 (1) (2021) 54–61, doi:10.1038/s43588-020-00006-7.

[57] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868, doi:10.1103/PhysRevLett.77.3865.

[58] H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integrations, Phys. Rev. B 13 (1976) 5188–5192, doi:10.1103/PhysRevB.13.5188.

[59] L.Y. Tian, Q.M. Hu, R. Yang, J. Zhao, B. Johansson, L. Vitos, Elastic constants of random solid solutions by SQS and CPA approaches: the case of fcc Ti-Al, J. Phys. Condens. Matter 27 (31) (2015), doi:10.1088/0953-8984/27/31/315702.

[60] F. Tasnádi, M. Odén, I.A. Abrikosov, Ab initio elastic tensor of cubic Ti 0.5 Al 0.5 N alloys: dependence of elastic constants on size and shape of the supercell model and their convergence, Phys. Rev. B Condens. Matter Mater.Phys. 85 (14) (2012) 144112, doi:10.1103/PhysRevB.85.144112.

[61] R. Hill, The elastic behaviour of a crystalline aggregate, Proc. Phys. Soc. London Sect. A 65 (5) (1952) 349–354, doi:10.1088/0370-1298/65/5/307.

[62] Z.J. Wu, E.J. Zhao, H.P. Xiang, X.F. Hao, X.J. Liu, J. Meng, Crystal structures and elastic properties of superhard Ir N2 and Ir N3 from first principles, Phys. Rev. B Condens. Matter Mater.Phys. 76 (5) (2007), doi:10.1103/PhysRevB.76.054115.

[63] D.D. Johnson, D.M. Nicholson, F.J. Pinski, B.L. Gyorffy, G.M. Stocks, Density-functional theory for random alloys: total energy within the coherent-potential approximation, Phys. Rev. Lett. 56 (1986) 2088–2091, doi:10.1103/PhysRevLett.56.2088.

[64] D.D. Johnson, F.J. Pinski, Inclusion of charge correlations in calculations of the energetics and electronic structure for random substitutional alloys, Phys. Rev. B 48 (1993) 11553–11560, doi:10.1103/PhysRevB.48.11553.

[65] A. Alam, D.D. Johnson, Optimal site-centered electronic structure basis set from a displaced-center expansion: improved results via a priori estimates of saddle points in the density, Phys. Rev. B 80 (2009) 125123, doi:10.1103/PhysRevB.80.125123.

[66] F. Maresca, W.A. Curtin, Mechanistic origin of high strength in refractory bcc high entropy alloys up to 1900k, Acta Mater. 182 (2020) 235–249.

[67] H.S. Oh, S.J. Kim, K. Odbadrakh, W.H. Ryu, K.N. Yoon, S. Mu, F. Körmann, Y. Ikeda, C.C. Tasan, D. Raabe, et al., Engineering atomic-level complexity in high-entropy and complex concentrated alloys, Nat. Commun. 10 (1) (2019) 1–8.

[68] N. Krapivka, S. Firstov, et al., High-entropy alloys: interrelations between electron concentration, phase composition, lattice parameter, and properties, Phys. Met. Metall. 118 (10) (2017) 970–981.

[69] V. Gorban, N. Krapivka, S. Firstov, D. Kurilenko, Role of various parameters in the formation of the physicomechanical properties of high-entropy alloys with bcc lattices, Phys. Met. Metall. 119 (5) (2018) 477–481.

[70] K. Gschneidner, A. Russell, A. Pecharsky, J. Morris, Z. Zhang, T. Lograsso, D. Hsu, C.C. Lo, Y. Ye, A. Slager, et al., A family of ductile intermetallic compounds, Nat. Mater. 2 (9) (2003) 587–591.

[71] O.N. Senkov, D.B. Miracle, Generalization of intrinsic ductile-to-brittle criteria by Pugh and Pettifor for materials with a cubic crystal structure, Sci. Rep. 11 (1) (2021) 4531, doi:10.1038/s41598-021-83953-z.

[72] O.N. Senkov, G.B. Wilks, J.M. Scott, D.B. Miracle, Mechanical properties of Nb25Mo25Ta 25W25 and V20Nb20Mo 20Ta20W20 refractory high entropy alloys, Intermetallics 19 (5) (2011) 698–706, doi:10.1016/j.intermet.2011.01.004.

[73] O.N. Senkov, J.M. Scott, S.V. Senkova, D.B. Miracle, C.F. Woodward, Microstructure and room temperature properties of a high-entropy TaNbHfZrTi alloy, J. Alloys Compd. 509 (20) (2011) 6043–6048, doi:10.1016/j.jallcom.2011.02.171.

[74] F. Tian, L.K. Varga, N. Chen, J. Shen, L. Vitos, Ab initio design of elastically isotropic TiZrNbMoV$_x$ high-entropy alloys, J. Alloys Compd. 599 (2014) 19–25.

[75] X. Gu, A. McDermott, S.J. Poon, G.J. Shiflet, Critical Poisson's ratio for plasticity in Fe–Mo–C–B–Ln bulk amorphous steel, Appl. Phys. Lett. 88 (21) (2006) 211905.
[76] C. Lee, G. Kim, Y. Chou, B.L. Musicó, M.C. Gao, K. An, G. Song, Y.-C. Chou, V. Keppens, W. Chen, et al., Temperature dependence of elastic and plastic deformation behavior of a refractory high-entropy alloy, Sci. Adv. 6 (37) (2020), doi:10.1126/sciadv.aaz4748.
[77] D. Pettifor, Theoretical predictions of structure and related properties of intermetallics, Mater. Sci. Technol. 8 (4) (1992) 345–349.
[78] L. Koči, Y. Ma, A.R. Oganov, P. Souvatzis, R. Ahuja, Elasticity of the superconducting metals V, Nb, Ta, Mo, and W at high pressure, Phys. Rev. B 77 (2008) 214101, doi:10.1103/PhysRevB.77.214101.
[79] P. Söderlind, O. Eriksson, J.M. Wills, A.M. Boring, Theory of elastic constants of cubic transition metals and alloys, Phys. Rev. B 48 (1993) 5844–5851, doi:10.1103/PhysRevB.48.5844.
[80] N. Nagasako, M. Jahnátek, R. Asahi, J. Hafner, Anomalies in the response of v, nb, and ta to tensile and shear loading: ab initio density functional theory calculations, Phys. Rev. B 81 (2010) 094108, doi:10.1103/PhysRevB.81.094108.
[81] G. Vazquez, Efficient machine-learning model for fast assessment of elastic properties of high-entropy alloys, 2021, https://github.com/vtguille/elasisso.