JOM, Vol. 76, No. 10, 2024
https://doi.org/10.1007/s11837-024-06753-6

Data-Driven Insights into the Structural Essence of Plasticity in High-Entropy Alloys

CHI-HUAN TUNG,[1] SHOU-YI CHANG,[1] ZHITONG BAI,[2] YUE FAN ,[2,5] SIDNEY YIP,[3] CHANGWOO DO,[4] and WEI-REN CHEN[4,6]

1.—Department of Materials Science and Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan. 2.—Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. 3.—Department of Nuclear Engineering, Massachusetts Institue of Technology, Cambridge, MA 02139, USA. 4.—Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. 5.—e-mail: fanyue@umich.edu. 6.—e-mail: chenw@ornl.gov

The heterogeneous mechanical response of a crystalline alloy with multiple principal elements was investigated using molecular dynamics simulations. The local configuration of the alloy in its quiescent state was characterized by the variables derived from the gyration tensor and the atomic electronegativity. A multivariate analysis identified the geometric and chemical factors that influenced the atomic packing variations. Upon straining, the non-affine displacement exhibited spatial heterogeneity. A statistical correlation was established between the local yield events and the specific features of the local configuration. Our findings, validated by the performance metrics analysis, provided a structural criterion for the instability mechanisms in high-entropy alloys (HEAs) and enhanced the understanding of their plasticity.

INTRODUCTION

The single-phase crystalline solid solutions, AlCoCrCuFeNi[1] and CoCrFeMnNi,[2] discovered in 2004, initiated a new era in metallurgy. The concept of multi-principal-elements with equiatomic ratios opened a new avenue for designing alloys. These so-called high entropy alloys (HEAs)[3–5] have exhibited a wide range of mechanical properties and have become a prominent research topic in modern metallurgy.[6–8]

The relationship between the mechanical properties and the microstructure of HEAs is a topic of great interest.[9–18] A key challenge is to identify the spatial regions where strain localization is likely to occur during mechanical processing. Conventional crystalline solids can be analyzed for crystallographic defects, such as dislocations, stacking faults, free surfaces, and grain boundaries, by looking for disruptions in spatial symmetry.[19] Studies of amorphous solids have explored whether specific features of packing patterns can promote localized mechanical instability.[20] One approach has been to treat the disordered structure as a packing of polyhedra[21,22] and investigate the relationship between local instability thresholds and topological features.[23–25]

A schematic two-dimensional projection of the generic structural features of relaxed HEAs is illustrated in Fig. 1. The particle arrangement resembles the configuration of crystalline solids at elevated temperatures. As in amorphous solids, the translational invariance and rotational symmetries are not strictly preserved in HEAs. Despite positional deviations, the topological features inherent to the face-centered cubic (FCC) structure remain intact. This means that topological criteria used to identify elastic anomalies in crystalline and amorphous solids cannot be applied to HEAs. Therefore, a new descriptive framework is needed to understand the structure–property relationships in HEAs. This framework should be based on variables most relevant to their deformation behavior. The challenge of developing such a framework motivates this study. In essence, we aim to develop a new way of thinking about defects in HEAs that is tailored to their unique structure.

(Received April 11, 2024; accepted June 24, 2024; published online July 11, 2024)

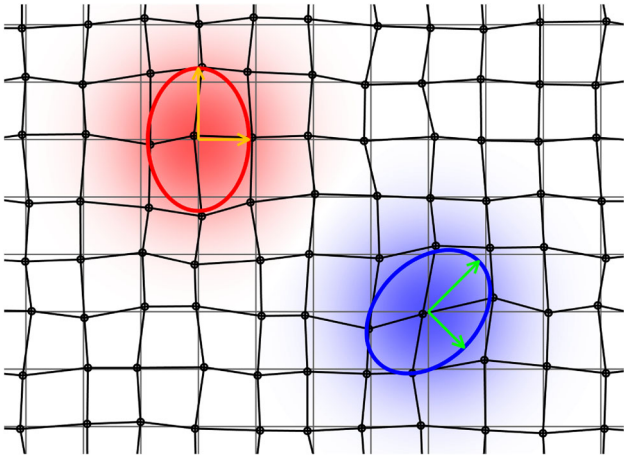

Fig. 1. Two-dimensional configurational projection of a (001) slice from a general face-centered cubic (FCC) high-entropy alloy (HEA) system, such as CoCrFeMnNi alloy. The displacement of each atom from its lattice point indicates structural deviations from crystalline perfection. Arrows indicate the orientations of local configurational units (LCUs), and the opacity of the color represents the spatial region used to calculate local physical quantities (Color figure online).

This could help us to better understand why HEAs have such good mechanical properties and to design even better HEAs in the future.

It is known in HEAs that, in addition to the intrinsic lattice distortions, local chemical environment such as electronegativity also largely dictate the system's mechanical behaviors.[26,27] Therefore, in the present study, we have combined both local structural and chemical information based on gyration tensor and electronegativity to describe the structure of HEAs from molecular dynamics (MD) simulation trajectories. We have used an exploratory data analysis approach to identify the specific configurational features of HEAs that are associated with defect nucleation and early stage plasticity in this new reference frame. Our results provide new insights into the deformation behavior of HEAs from a statistical correlation perspective.

MATERIALS AND METHODS

Atomistic Simulation and Non-affine Displacements Characterization

The initial configuration of the CoCrFeMnNi HEA was prepared by randomly assigning element types to each lattice point in the FCC perfect crystal simulation cell with the X, Y, and Z principal axes aligned with [100], [010] , and [001] crystal orientations, respectively. Periodic boundary conditions have been applied on all directions. The simulation cell contains $30 \times 30 \times 30$ FCC unit cells with overall 108,000 constituent atoms. The simulation procedures were performed using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package.[28] The force field developed by Lee and coworkers[29] was used to model the interaction among the constituent atoms of the CoCrFeMnNi HEA.[30] After being equilibrated under NPT ensemble at 77 K with zero external pressure, energy minimization was performed on the initial configuration based on the conjugate gradient algorithm to obtain the quiescent state structure. To investigate the connection between the isolated plastic activity in a strained HEA and the local geometric features at its quiescent state, simulations of HEA subjected to constant-volume tensile and shear deformation were carried out under NVT ensemble at 77 K. In Cartesian coordinates, the deformation gradient tensors $\boldsymbol{J}$ for tension applied in the z axis and shear on the xz plane with a stretch γ can be expressed as $\begin{bmatrix} \frac{1}{\sqrt{1+\gamma}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{1+\gamma}} & 0 \\ 0 & 0 & 1+\gamma \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ respectively. The strain rate used in this simulation was $5 \times 10^8\ s^{-1}$. Note that the tension along the x and y axes, as well as shear on the xy and yz planes were also performed. The simulation timestep was 0.001 ps and the trajectory was collected every 1 ps until γ reaching 0.4. The yield point was identified by common neighbor analysis from the transition point of the stress–strain curve where the von Mises stress was seen to drop. The method proposed by Falk and Langer was used to provide the non-affine measure of particle displacements from the trajectories corresponding to the yield point.[31]

Defining the Local Configurational Unit (LCU) in the Multivariate Reference Frame

To describe the LCU, we use a second-order gyration tensor R[32] defined by:

$$\mathcal{R}_{\alpha\beta} = \frac{1}{n}\sum_{I=1}^{n} r_I^{\alpha} r_I^{\beta}. \tag{1}$$

where r_I^{α} is the αth component of the position difference vector $\mathbf{r_i}$ and n is the number of particles within the nearest-neighbor cluster. The coordinate system was selected to ensure that $\sum_{I=1}^{n} \mathbf{r_i} = 0$. Therefore, $\mathbf{r_i}$ can be considered as the positional difference between the I-th nearest-neighbor particle and the center of mass of this cluster. From Eq. 1, it is clearly seen that $\mathcal{R}$ is the covariance matrix which gives the spatial correlations among the constituent particles of this system. Therefore, it provides a convenient measure of the geometric features around a tagged particle. At the initial stage of plastic deformation of the HEA, the length scale of a local plastic event was found to be around one mean inter-particle distance. Therefore, for a tagged particle, we calculated $\mathcal{R}$ from the spatial range defined by its first coordinate shell of pair distribution function $g(r)$ and used it to quantify the configurational characteristics of its local environment, such as size, shape, and orientation, which reflect its interaction with neighboring

particles. To bypass the mathematical complications caused by the abrupt length cutoff, a Gaussian-based weight function, as schematically illustrated by the color gradient in Fig. 1, has been used in calculating $\mathcal{R}$ from the trajectories. By summing the weighted contribution over all the N particles in our simulation, $\mathcal{R}_{\alpha\beta}$ can be expressed as:

$$\mathcal{R}_{\alpha\beta} = \frac{\sum_{I=1}^{N} w_i r_I^{\alpha} r_I^{\beta}}{\sum_{I=1}^{N} w_I}. \quad (2)$$

where $w_I = e^{-\frac{|\mathbf{r_i}|^2}{2\sigma^2}}$ is the weight function. The variance σ has been chosen to be the position of the first peak of $g(r)$ to emphasize the contribution from the nearest-neighbor particles. Because $\mathcal{R}$ is symmetric positive semidefinite, according to the spectral theorem[33] $\mathcal{R}$ can be diagonalized as:

$$\mathcal{R} = \boldsymbol{Q}\Lambda_{\mathcal{R}}\boldsymbol{Q}^T. \quad (3)$$

In Eq. 3 $\Lambda_{\mathcal{R}}$ is a diagonal matrix with three entries λ_1, λ_2 and λ_3 in ascending order which are the eigenvalues of $\mathcal{R}$ and often called principal moments. $\boldsymbol{Q}$ is an orthogonal matrix whose three columns are the orthonormal eigenvectors of $\mathcal{R}$ and $\boldsymbol{Q}^T$ is the transpose of $\boldsymbol{Q}$. Additionally, the sign of each column in $\boldsymbol{Q}$ were selected to make $Q_{33} > 0$ and $\det(\boldsymbol{Q}) = 1$ to ensure the uniqueness and avoid improper rotations. Taking $\boldsymbol{Q}$ as a rotation matrix, following the *zxz*-convention it can be expressed as:

$$\boldsymbol{Q} = R_1 R_2 R_3. \quad (4)$$

$$R_1 = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix},$$

$$R_3 = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where ψ, θ , and ϕ are the Euler angles.

By decomposing $\mathcal{R}$ into its eigenvalues and eigenvectors to eliminate the off-diagonal entries, the configurational features of $\mathcal{R}$ can be intuitively identified in this eigenspace. The size and shape of the local environment around a tagged particle, which is termed as the LCU, can be directly linked to the principal moments and its orientation can be described by the Euler angles, as displayed by the ellipsoids and arrows in Fig. 1. For a defect-free FCC perfect crystal consisting of only one element, $\mathcal{R}$ for any constituent atom is simply $\frac{1}{3}\mathrm{tr}(\mathcal{R}) \cdot \mathrm{I}_3$ because of the inherent octahedral symmetry. The breakdown of structural symmetry in HEAs is the result of complicated inter-particle interaction originating from the chemical difference between the constituent atoms. In the context of excess local electronegativity $\langle\chi\rangle$, which is defined as the difference between the averaged electronegativity of atoms in an LCU and the global average,[26] we have demonstrated a connection between the distributions of different types of atoms and the non-affinity of local deformation. Therefore, in this study, $\langle\chi\rangle$ is labeled to the tagged particle of each LCU to reflect its compositional characteristics. Along with other six structural variables of $\mathcal{R}$, a vector space $\mathbb{R}^7$ is set up to facilitate the interpretation of configurational features of HEAs. The thus-constructed $\mathbb{R}^7$ space also lays down the foundation of further principal component analysis (PCA; see Appendix 1) to extract insights into how those seven variables are collectively determining the incipient of local plasticity in HEAs.

RESULTS

Independent Variables Characterizing the Local Configuration of HEA Identified by PCA

Following the LCU characterization in section "Defining the Local Configurational Unit (LCU) in the Multivariate Reference Frame," each of the 108,000 particles within our simulated HEA system has been individually labeled based on the corresponding eigenvectors and eigenvalues of $\mathcal{R}$, as well

Table I. The principal axes identified by PCA analysis

Principal components (PCs)		1	2	3	4	5	6	7
$\Lambda_{\mathcal{R}}$	λ_1	− 0.36	0.00	0.00	0.00	0.48	− 0.70	− 0.39
	λ_2	− 0.39	− 0.02	0.00	0.00	0.43	0.04	0.81
	λ_3	− 0.43	0.00	0.00	0.00	0.36	0.71	− 0.43
Q	ϕ	0.00	0.00	− 0.60	0.80	0.01	0.00	0.00
	θ	− 0.02	1.00	− 0.01	− 0.01	0.00	0.00	0.02
	ψ	− 0.01	0.02	0.80	0.60	0.00	0.00	0.00
	$\langle\chi\rangle$	0.73	0.01	0.01	0.00	0.68	0.09	− 0.01
Percentage of the total variance		28%	24%	18%	16%	8%	4%	2%

The percentages of total variance projected onto the principal axes are presented in descending order.

as the excess local electronegativity denoted as $\langle\chi\rangle$. This data is structured into a $7 \times 108,000$ feature matrix, denoted as $\boldsymbol{F}$. To eliminate potential intercorrelations among input variables, we conducted PCA.[33,34] This analytical process re-expresses the data as a set of new orthogonal variables, facilitating the extraction of crucial configurational information from the original dataset.

Table I presents the singular vectors and variances from PCA. In the first and fifth principal components, denoted as PC1 and PC5, an inverse relationship is observed between $\langle\chi\rangle$ and the eigenvalues λ_1, λ_2 , and λ_3. Specifically, in PC1, the rounded values of all lambda entries are -0.4. Considering that the trace of $\Lambda_{\mathcal{R}}$ corresponds to the radius of gyration R_G of the associated LCU, this finding suggests that alterations in $\langle\chi\rangle$ predominantly impact the size of the LCU. This observation underscores the influence of chemical disparities among HEA atoms on configurational features, primarily through changes in the average interparticle distance.

Furthermore, in PC1, the negative correlation between $\langle\chi\rangle$ and the entries of $\Lambda_{\mathcal{R}}$ indicates that an LCU characterized by a more negative $\langle\chi\rangle$ tends to be larger than those with less negative $\langle\chi\rangle$. Despite PC1 being the most significant principal component, explaining 28% of the original data's variance, it is essential to note that this correlation does not imply causation. This is evident in the positive correlation revealed by PC5, which accounts for 8% of the variance.

One can define the principal component score as $\boldsymbol{S} \equiv U^T\boldsymbol{F} = \Sigma V^T$ where U and V are the left and right singular vectors of $\boldsymbol{F}$ (see A). The I^{th} row of $\boldsymbol{S}$, which is denoted as S_I, is the projection of the centralized data on the I^{th} principal axes U_I. The fact that the data distributes within an ellipse area in the $\mathbb{R}^2$ subspace subtended by U_1 and U_5 with an aspect ratio of $\sqrt{3.5}$ in Fig. 2a demonstrates that $\langle\chi\rangle$ is only correlated with, instead of proportional to, the trace of $\Lambda_{\mathcal{R}}$.

Effect of Compositional Difference on Local Geometric Features of HEA

Using trajectories obtained within proximity to yield points, the non-affine parameter D^2 for each LCU is computed following the methodology proposed by Falk and Langer.[31] Subsequently, we identify LCUs characterized by the 1st and 99th percentiles of D^2 in descending order. The geometric attributes of these LCUs at the quiescent state are then classified to establish their correlation with local plasticity events.

For this analysis, we introduce ternary discrete variables $d_j \in \{0, 1, 2\}$, where $d_j = 1$ signifies the j^{th} LCUs within the 1st percentile, $d_j = 2$ designates LCUs within the 99th percentile, and $d_j = 0$ otherwise. Consequently, the extreme groups can be denoted as $D_1^2 \equiv \{j \mid d_j = 1\}$ and $D_{99}^2 \equiv \{j \mid d_j = 2\}$, respectively.

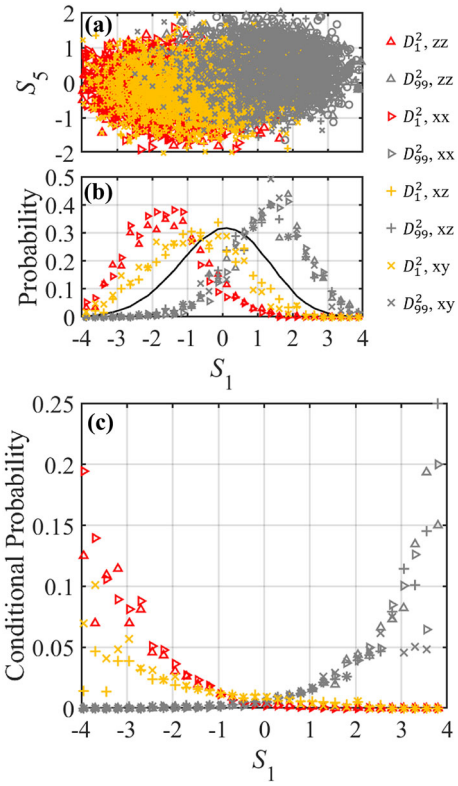

Fig. 2. (a) The joint distribution of (S_1, S_5) for LCUs characterized by the 1st percentile of D^2 (red and yellow symbols) and 99th percentile of D^2 (gray symbols). The different loading conditions are represented by the scheme of symbol given in Fig. 3a. A clustering phenomenon of data belonging to the same percentile group is clearly revealed. (b) The probability density functions for all LCUs $p(S_1)$ (black curve), and the probability density for the two extreme groups, $p(S_1|j \in D_1^2)$ and $p(S_1|j \in D_{99}^2)$. (c) The conditional probabilities $P(j \in D_1^2|S_1)$ and $P(j \in D_{99}^2|S_1)$ calculated from the results given in (b). It is clearly seen that the spatial regions characterized by negative anomalies of S_1 are more likely to undergo non-affine deformation and those characterized by higher S_1 tend to deform elastically (Color figure online).

As indicated in Table I, the principal component, S_1, demonstrates its significance in the variations within the feature matrix. It will be meaningful to investigate how the local configurations will affect the initiation of non-affine deformations by observing the distribution of D_1^2 and D_{99}^2 in the space of S_1 and S_5, which also represent the different combination of LCU size and $\langle\chi\rangle$. The collection of 2-tuples $(S_1, S_5) = \{(S_{1j}, S_{5j}) \mid 1 \le j \le N \cap j \in \mathbb{N}\}$ for all LCUs belonging to the two extreme groups were extracted and are presented in Fig. 2a in the $\mathbb{R}^2$ space spanned by U_1 and U_5. The two perpendicular coordinates are labeled S_1 and S_5. The up-pointing triangles, right-pointing triangles, plus signs, and x marks represent the data collected from LCUs that underwent plastic deformation under uniaxial tensile stress along the z direction, uniaxial tensile stress along the x direction, shear stress along the xz direction, and shear stress along the xy direction,

respectively, as schematically illustrated in Fig. 3a. The red and yellow symbols represent the LCUs in the 1st percentile of D^2, and the gray symbols represent the LCUs in the 99th percentile of D^2. Figure 2a shows that, regardless of the deformation conditions, the results from the same percentile group tend to cluster together.

To analyze the data in the U_1 vector space, we project them onto the S_1 axis and compute the probability density functions (PDFs) $p(S_1)$ for all LCUs from the quiescent state trajectories of HEA. The PDFs of S_1 conditioned on $j \in D_1^2$ and $j \in D_{99}^2$ are denoted as $p(S_1|j \in D_1^2)$ and $p(S_1|j \in D_{99}^2)$, respectively. As shown in Fig. 2b, the PDF of S_1 (black curve) can be well approximated by a Gaussian function with the form $p(S_1) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{S_1^2}{2\sigma^2}\right]$ where $\sigma^2 = 1.52$. The bell-shaped curves of all $p(S_1|j \in D_1^2)$ and $p(S_1|j \in D_{99}^2)$ indicate that S_1 follows a normal distribution for both the two extreme groups in the U_1 vector space. The mean of $p(S_1|j \in D_1^2)$ is around -2 for the LCUs that tend to undergo plastic deformation under uniaxial tensile stress, and around -1 for those that tend to yield under shear stress. All the $p(S_1|j \in D_{99}^2)$ is around $S_1 = 2$ for the LCUs that are more resistant to deformation.

The likelihood of an LCU undergoing yield based on its configurational features described by S_1 is not immediately apparent from the various probability distributions presented in Fig. 2b. Instead, Bayes' theorem allows us to deduce the conditional probabilities $P(j \in D_1^2|S_1)$ and $P(j \in D_{99}^2|S_1)$, representing the susceptibility of LCUs when S_1 is given. This can be expressed as $P(A|B) \equiv \frac{p(B|A)}{p(B)}P(A)$, where the uppercase P denotes discrete probability. Analyzing the yellow and red symbols in Fig. 2c, a clear monotonic decrease in $P(j \in D_1^2|S_1)$ is observed with an increase in S_1 across all deformation conditions. Conversely, the gray symbols exhibit an opposite trend in $P(j \in D_{99}^2|S_1)$. This leads us to the conclusion that spatial regions characterized by negative anomalies of S_1 are more prone to undergoing non-affine deformation, while those characterized by higher S_1 tend to deform elastically. Importantly, this observation differs intrinsically from the stability criterion of amorphous solids, where both positive and negative anomalies of atomic level volumetric strain could cause mechanical instability.[35,36] We attribute this discrepancy to the anharmonicity of the interaction potential employed in our simulation, where the energy cost and restoring force for compressive deformation are significantly larger than those for tensile deformation in regions characterized by lower S_1, given the same amount of strain.

Figure 3b presents an example of the spatial distributions of S_1. The non-affine deformation within strained HEA is prominently localized

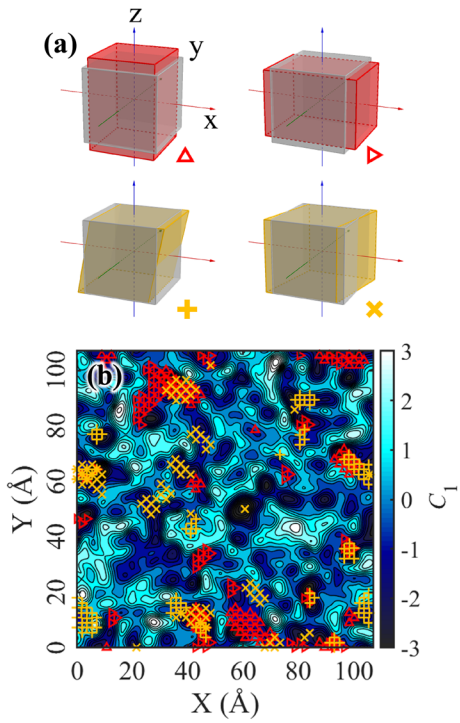

Fig. 3. (a) Schematic of four loading conditions: uniaxial tensile stress along the z direction (up-pointing triangle), uniaxial tensile stress along the *x* direction (right-pointing triangle), shear stress along the xz direction (plus sign), and shear stress along the xy direction (x marks). (b) One example of the contour plot of S_1. The heterogeneity of its spatial distribution can be clearly revealed by the blue gradient color scheme. The yellow and red symbols indicate the locations of non-affiine particle rearrangements in CoCrFeMnNi subjected to the mechanical loadings illustrated in (a) (Color figure online).

within the dark blue spatial regions where S_1 falls below its mean value of 0. Nevertheless, under diverse deformation conditions, the precise locations of irreversible particle rearrangement exhibit significant variation and minimal overlap. The insights derived from PCA, as outlined in Table I, offer a preliminary understanding of the physical basis of this phenomenon: nearly half of the original data's variance is captured by the second, third, and fourth principal axes (PC2, PC3, and PC4). These axes represent distinct linear combinations of ψ, θ, and ϕ (refer to "Defining the Local Configurational Unit (LCU) in the Multivariate Reference Frame"). Thus, achieving a comprehensive understanding of HEA plasticity requires a fundamental exploration of the influence of LCU orientation on local plastic activity.

Influence of Orientation of LCU on the Plasticity of Strained HEA

In this investigation, the geometric configuration of a LCU is elucidated through an ellipsoid characterized by three principal axes corresponding to the eigenvectors of $\Lambda_{\mathcal{R}}$, converging at the center of

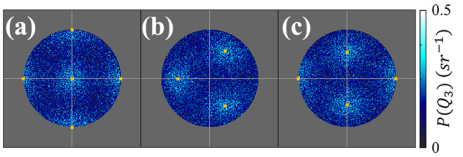

Fig. 4. The probability distribution $p(Q_3)$ viewed along the (a) ⟨001⟩, (b) ⟨111⟩, and (c) ⟨110⟩ directions. The results show that the preferred orientation of LCUs is along the ⟨001⟩ direction of FCC crystal structure. The data are coarse-gained based on the approach given in Appendix B.

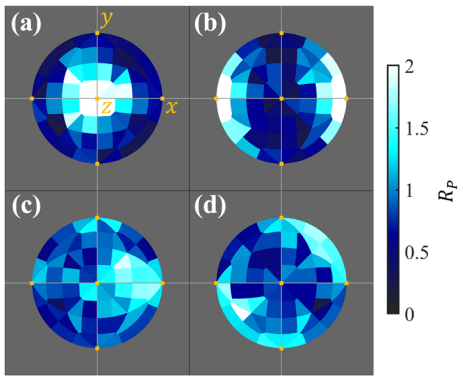

Fig. 5. The spatial distributions of R_P on the unit sphere for (a) uniaxial tensile stress along the z direction, (b) uniaxial tensile stress along the *x* direction, (c) shear stress along the xz direction, and (d) shear stress along the xy direction.

symmetry. Adhering to the framework proposed by Theodorou and Suter to quantitatively define the conformation of an unperturbed linear polymer chain,[37] we designate the axes of this configurational ellipsoid as Q_1, Q_2 , and Q_3, aligning with the directions of minor, median, and major axes, respectively. Consequently, the orientation of an LCU can be precisely determined based on the alignment of its Q_3. Based on the methodology outlined in Appendix B, we compute the coarse-grained probability distribution $P_c(Q_3)$. The azimuthal equidistant projections, observed along the 4-fold axes (C_4) of ⟨001⟩, 3-fold axes (C_3) of ⟨110⟩ , and 2-fold axes (C_2) of ⟨111⟩ crystallographic directions, are presented in Fig. 4a–c, respectively. Upon scrutinizing the distribution of the projected data points, it is clearly seen that the preferred orientation of Q_3 for all LCUs is aligned with the ⟨001⟩ direction of the FCC crystal structure.

In this study, we introduce a probability ratio R_P defined as $\frac{P(j\in D_1^2|Q_3)}{P(j\in D_1^2)}$ for the extreme group associated with the 1st percentile of D^2. This ratio serves to assess the impact of LCU orientation on local plastic events. The correlation between these statistical events becomes apparent through the deviation of R_P from 1. Figure 5a–d illustrates the coarse-grained distribution of R_P on the unit sphere under different loading conditions: uniaxial tensile stress along the z direction, uniaxial tensile stress along the x direction, shear stress along the xz direction, and shear stress along the xy direction, respectively. The gradient of the blue color indicates the magnitude of R_P. Upon examining the distribution of light blue domains, characterized by $R_P > 1$, in relation to the corresponding deformation gradient, a clear pattern emerges. It reveals that non-affine deformations near each yield point are promoted by the compatibility between the orientation of LCU and the direction of applied loading. In all cases, the Q_3 of LCUs undergoing non-affine deformation tends to align with the major principal axis of applied strain Q_{E3}, defined by $\boldsymbol{E} = Q_E \Lambda_E Q_E^T$, where $\boldsymbol{E} = \frac{1}{2}(\boldsymbol{J}^{\mathbf{T}}\boldsymbol{J} - \boldsymbol{I}_3)$, representing the green strain tensor of the deformation gradient $\boldsymbol{J}$ (refer to "Atomistic Simulation and Non-affine Displacements Characterization"). For a HEA subjected to uniaxial tensile stress, the corresponding Q_{E3} aligns with the tensile axes which is [001] in Fig. 5a and [100] in 5(b). For a sheared HEA, in the zero-strain limit ($\gamma \to 0$), Q_{E3} aligns with the angle bisector between shear direction and normal vectors of shear plane which is [101] in Fig. 5c and [110] in Fig. 5d. Furthermore, we propose an approach to quantitatively measure the specific alignment between the overall shape of LCUs and the principal strain orientation using spherical harmonics expansions as outlined in Appendix C.[32] Leveraging the symmetric property of Y_2^0, it becomes evident that the degree of alignment between the LCU orientation and the direction of external strain can be quantitatively assessed by the positive deviation of its second moment G_2^0 in the mass distribution function in the principal strain coordinate.

Shape Distributions of LCU in the Quiescent State of HEA

Having explored the static dimensions and orientations of LCUs and their associations with the plasticity of a strained HEA, we can now investigate the impact of LCU shape on local particle arrangements. Various methods exist to characterize the shape of an object, and, in this study, the shape distribution of LCUs is determined through the probability distribution of the 6th and 7th principal components (PC6 and PC7) of $\boldsymbol{C}$, as detailed in Table I. However, it is noteworthy that the variance retained by PC6 and PC7 in the original data is only 4% and 2%, respectively. This observation leads to the conclusion that the configurational differences in local environments within HEA are too subtle to be effectively discerned based on the shapes of LCUs characterized by the geometric quantities associated with PC6 and PC7. In essence, the shapes of various LCUs in a quiescent HEA, as depicted by these principal components, are fundamentally similar due to the limited variance captured by PC6 and PC7.

The geometric configuration of LCU can be quantified through linear combinations of λ_1, λ_2 and λ_3 in PC6 and PC7 provided by PCA, as outlined in Table I. Denoting these linear combinations as the dot product $S_I = U_I \cdot \lambda$, where $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ forms a vector in $\mathbb{R}^3$ space following the proposal by Šolc and Stockmayer.[38] To facilitate the discussion, the 0.09 contribution of $\langle\chi\rangle$ to PC6 is considered a statistical fluctuation due to its marginal influence. Consequently, the unit vector of PC6, U_6, is determined as $\frac{1}{\sqrt{2}}(-1,0,1)$ and $S_6 = \frac{1}{\sqrt{2}}(\lambda_3 - \lambda_1)$. Geometrically, this quantity measures the deviation of the shape of an LCU from an elliptical form toward that of a perfect sphere. Similarly, the unit vector of PC7, U_7, is established as $\frac{1}{\sqrt{6}}(-1,2,-1)$, yielding $S_7 = \frac{1}{\sqrt{6}}[2\lambda_2 - (\lambda_1 + \lambda_3)]$. Given the circular symmetry, U_7 serves as an indicator to assess whether the quadric surface of an LCU can be characterized as oblate $S_7 > 0$ or prolate $S_7 < 0$.

Figure 6a illustrates the $\mathbb{R}^2$ subspace formed by U_6 and U_7. It is noteworthy to mention that Theodorou and Suter introduced two principal axes of gyration, namely, asphericity b and acylindricity c, to describe the instantaneous shape of a linear polymer chain in a solution.[37] Expressing b and c in terms of dot product, $b = U_b \cdot \lambda$ and $c = U_c \cdot \lambda$ where $U_b = (-\frac{1}{2}, -\frac{1}{2}, 1)$ and $U_c = (-1,1,0)$. The outcomes for all possible shapes are confined within the light blue-marked right triangular region. The three vertices correspond to three asymptotic geometric shapes: a perfect sphere, a zero-thickness disc, and a volumeless line. Given the ascending order of λ_1, λ_2 , and λ_3, the collection of all possible 3D elliptical shapes should form a right triangle. Here, the hypotenuse, long leg, and short leg represent the limits of $\lambda_1 = 0$, $\lambda_2 = \lambda_1$ and $\lambda_3 = \lambda_2$, respectively.

In Fig. 6b, $p(S_6, S_7)$ for all LCUs at the quiescent state, as well as the extreme groups $p(S_6, S_7 | j \in D_{99}^2)$ and $p(S_6, S_7 | j \in D_1^2)$ for two deformation conditions, are presented. All data points cluster near the origin, indicating that the shapes of all LCUs deviate only slightly from a perfect sphere. Additionally, the data exhibits reflection symmetry concerning the axis of $S_7 = 0$, and, from the distribution, it can be concluded that no discernible correlation exists between the shape of LCU and the local plasticity of strained HEA.

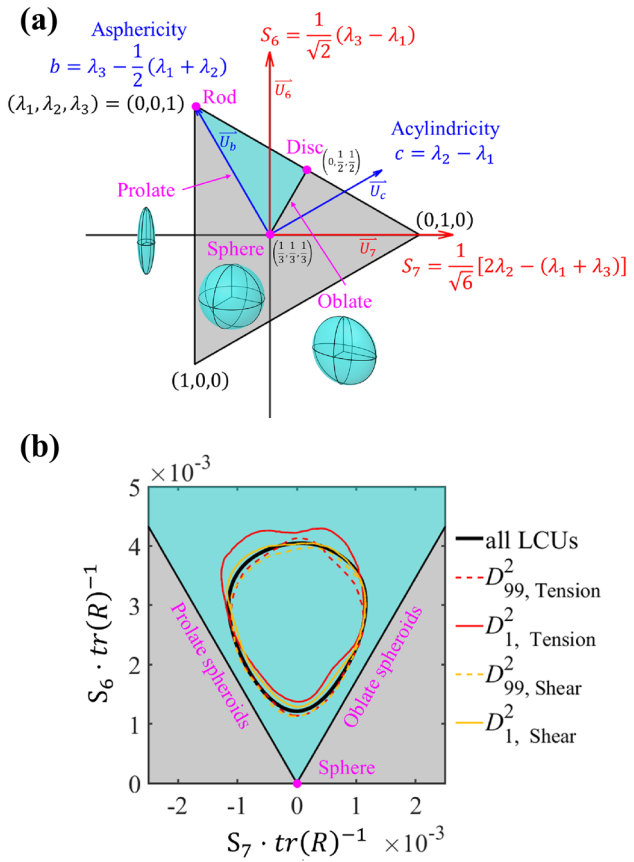

Fig. 6. (a) The R^2 subspace spanned by U_6 and U_7 contains all possible ellipsoids with tr(R) = 1. The red vectors are U_6 and U_7 from PCA given in Table I and the blue vectors are the axes of asphericity and acylindricity defined by Theodorou and Suter.[37] (b) The joint probability distribution of $p(S_6, S_7)$ normalized by tr(R) for the extreme groups D_{99}^2 and D_1^2 when subjected to tension and shear deformation, respectively. Contours specify the 68% confident level (Color figure online).

Classification of Configurational Features Using Linear Discriminant Analysis

Utilizing the independent configurational variables derived from PCA, we can further categorize the LCUs based on their susceptibility to non-affine displacement. Building upon the insights provided in Figs. 2, 3, 4, 5, and 6, the information pertaining to local configurations can be distilled and represented in a reference coordinate system comprising the principal axes of S_1, G_2^0, and S_6. This representation accounts for the influences of chemical composition, as well as size, orientation, and shape on the LCUs. The dataset is structured as a $3 \times 108,000$ feature matrix $\boldsymbol{f} = (S_1,\ G_2^0,\ S_6) = \left\{\left(S_{1j},\ G_{2j}^0,\ S_{6j}\right) \mid 1 \le j \le N \cap j \in \mathbb{N}\right\}$ within the $\mathbb{R}^3$ Cartesian space, where each row is individually converted to its corresponding standard scores. Examining the outcomes presented in Fig. 7, it is observed that the data from the D_1^2 and D_{99}^2 groups exhibit distinct clustering. $\boldsymbol{m}_1$ ad $\boldsymbol{m}_{99}$, along with the variances Σ_1 and Σ_{99}, contribute to the separation of these groups.

In Fisher's linear discriminant analysis,[33,39] the solution to classification problems involving two classes entails identifying the optimal decision plane defined by its normal vector $\boldsymbol{v}$ and a threshold value c_{th}. If $\boldsymbol{v}^T\boldsymbol{f} > c_{\text{th}}$, the sample is more likely to belong to group D_1; conversely, if $\boldsymbol{v}^T\boldsymbol{f} < c_{\text{th}}$, the sample is assigned to group D_{99}. The separation

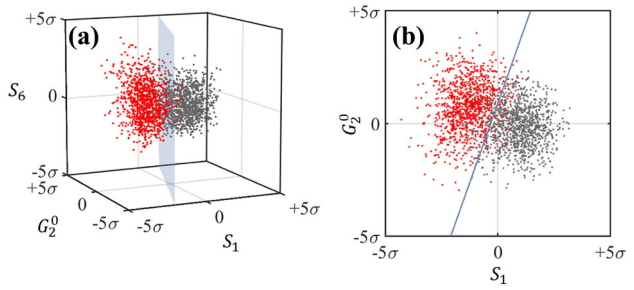

Fig. 7. (a) Three-dimensional and (b) two-dimensional scatter plots of independent configurational variables used to characterize LCU morphologies. Extreme groups D_1^2 (red) and D_{99}^2 (gray) can be separated by a properly chosen decision surface (light blue) (Color figure online).

Table II. The normal vector $\boldsymbol{v}$, threshold value c_{th} of the decision plane and the separation ratio R_S for four loading conditions

	J_{ZZ}	J_{XX}	J_{XZ}	J_{XY}
v_1	− 1.5741	− 1.5784	− 0.8240	− 0.8292
v_2	0.5569	0.4727	0.2866	0.3002
v_3	− 0.0088	− 0.0317	0.0086	0.0017
R_S	3.9618	3.9427	1.2751	1.4050
c_{th}	0.45	0.45	0.00	0.05

ratio R_S, quantifying the separation distance between the two groups, is given by $R_S = \frac{\left(\boldsymbol{v}^T\boldsymbol{m}_1 - \boldsymbol{v}^T\boldsymbol{m}_{99}\right)^2}{\boldsymbol{v}^T(\Sigma_1+\Sigma_{99})\boldsymbol{v}}$. To maximize R_S, the optimal decision plane's normal vector $\boldsymbol{v}$ is obtained by solving the generalized eigenvalue problem, yielding $\boldsymbol{v} = (\Sigma_1 + \Sigma_{99})^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_2)$.

Table II presents the values of $\boldsymbol{v}$ for four distinct deformation conditions. Notably, the magnitude of the first component of $\boldsymbol{v}$, associated with S_1, is approximately three times that of the second component related to G_2^0. The third component corresponding to S_6 is found to be negligible. The table further underscores the compelling evidence that the chemical composition, size, and orientation of LCU are pivotal factors in understanding the specific atomic packing features relevant to local yielding events in a strained HEA.

To determine the optimal classification threshold c_{th}, we used the evaluation metrics described in Appendix D to assess the efficacy of the classification. Figure 8 shows the receiver operating characteristic (ROC) curves[39] for four different deformation conditions. The optimal c_{th} was chosen to maximize accuracy (ACC). For the uniaxial condition, the optimal c_{th} produced an (FPR, TPR) of approximately (0.1, 0.9), while the shear condition produced an (FPR, TPR) of approximately (0.2, 0.8). As defined in Appendix D, the area under the ROC curve (AUC) is the probability that the $\boldsymbol{v}^T\boldsymbol{f}$ of a

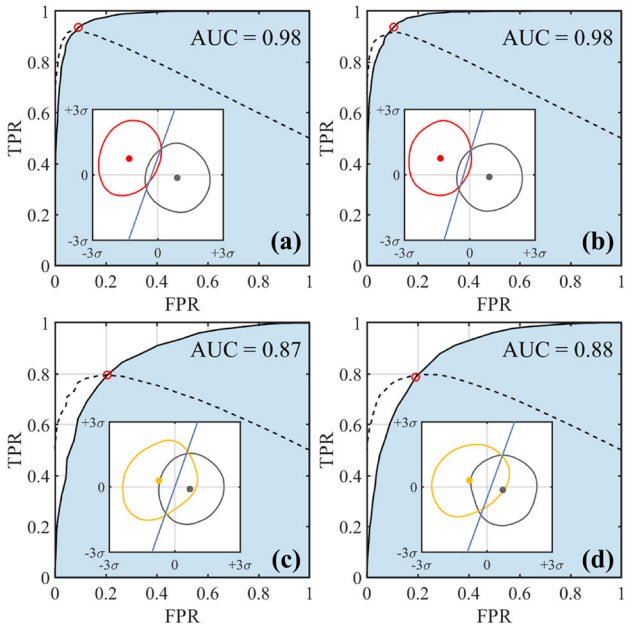

Fig. 8. Receiver operating characteristic (ROC) curves for uniaxial tensile stress along the z direction (a), uniaxial tensile stress along the *x* direction (b), shear stress along the xz direction (c), and shear stress along the xy direction (d). Solid lines show true positive rate (TPR) versus false positive rate (FPR) at different classification thresholds c_{th}, dashed lines show classification accuracy, red circles indicate the conditions with maximum accuracy, and insets show decision boundaries and 2D distributions of extreme group D_1^2 (red and yellow) and D_{99}^2 (gray), where contours specify the 68% confidence level and dots represent group means (Color figure online).

randomly chosen conditionally positive sample (D_1^2) is higher than that of a randomly chosen conditionally negative sample (D_{99}^2). Therefore, the AUC is commonly used to measure the quality of a classification model. For uniaxial tension conditions, the AUCs were around 0.98, and, for shear conditions, they were around 0.87. The corresponding decision boundaries are shown in the insets of Fig. 8. The results demonstrate that our method can successfully discriminate the plasticity of LCUs based on the extracted features.

DISCUSSION

This study's consideration of structure–property relationships in HEAs raises several aspects that warrant further discussion. To further elucidate the physical insights presented in Figs. 5, 7, and 8, we employ the calculation of the potential energy landscape (PEL) for HEA as an independent approach. The PEL offers an alternative perspective, allowing the interpretation of elementary topological fluctuations, such as collective rearrangements of particles. These fluctuations involve the hopping between neighboring local minima in the PEL.[40–44] In this investigation, we employ the activation relaxation technique (ART)[45–47] to

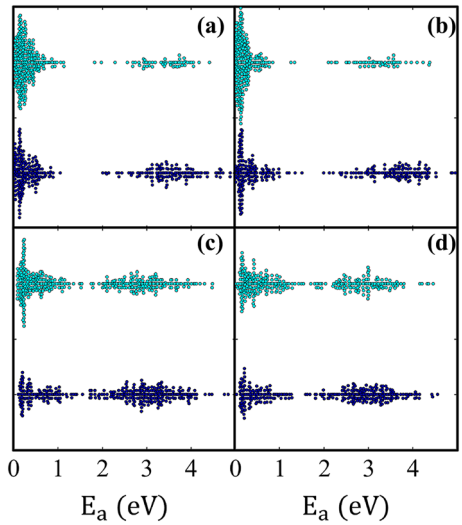

Fig. 9. Distributions of activation barriers for atomic rearrangements in the 50 LCUs with the highest G_2^0 (light blue dots) and with the lowest G_2^0 (dark blue dots) correspond to the loading conditions in Fig. 5a–d. The profile of activation barriers is characterized by a low-energy collective motion mode with E_a<1 eV, and a high-energy mode related to formation of point defects such as Frenkel pair. In the low-energy regime, the probability of observing the local particle rearrangement in the LCUs aligned with the applied load is seen to be statistically higher than that in the LCUs misaligned with applied load. The results of this independent thermodynamic calculation support the physical picture presented in Fig. 5 obtained from configurational analysis (Color figure online).

explore the PEL structures of HEA. This method establishes key connections between neighboring metastable states in the PEL, devoid of any presumptions about predetermined mechanisms. The LCUs are categorized based on G_2^0, and the distributions of activation barriers, as identified by the ART algorithms, are calculated. The results, depicted in Fig. 9, encapsulate approximately 350 activation barriers from groups containing 50 LCUs with the highest G_2^0 (light blue dots) and those with the lowest G_2^0 (dark blue dots). For the four explored loading conditions, the activation barriers associated with particle rearrangement can be partitioned into a low-energy collective motion mode $E_a < 1$ eV and a high-energy mode linked to the formation of point defects from a perfect crystal, such as Frenkel pairs. Notably, the PEL calculations suggest that the alignment between LCUs and Q_{E3} of the applied load promotes the low-barrier collective motion mode. This finding provides crucial validation, underscoring the pivotal role of the inherent orientation of LCUs in influencing the mechanical properties of highly distorted crystalline systems within HEAs. We would like to note that here we only report the probability distribution of the scalar-valued energy barriers. It would also be meaningful to examine the spatial correlation between LCU orientations and the transformation events identified in ART in the future studies.

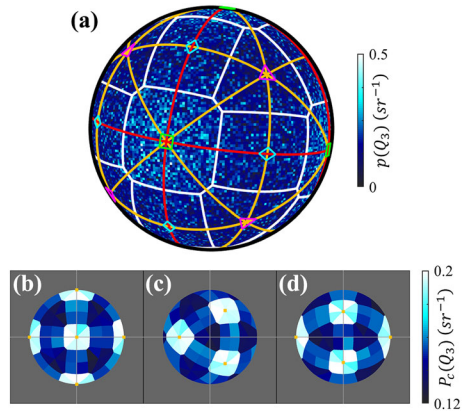

Fig. 10. (a) The partition of the unit sphere based on the mirror planes in an octahedron perpendicular to the 4-fold axes (C_4) of ⟨001⟩ (red curves) and 2-fold axes (C_2) of ⟨110⟩ (yellow curves). The 4-fold, 3-fold and 2-fold axes are specified by the green square symbols, magenta triangle symbols, and cyan diamond symbols, respectively, while the Voronoi boundaries are displayed as white curves. The coarse-grained probability distribution $p_c(Q_3)$ viewed along the (b) ⟨001⟩, (c) ⟨111⟩, and (d) ⟨110⟩ directions (Color figure online).

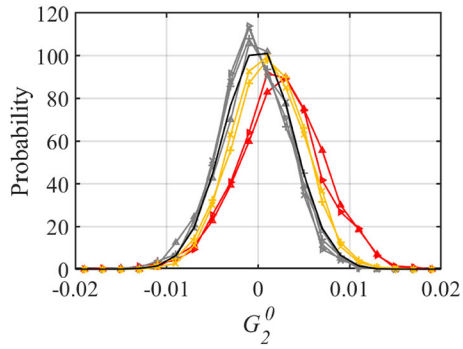

Fig. 11. The probability density functions for all LCUs $p(G_2^0)$ (black curve), and the probability density for the two extreme groups, $p(G_2^0|j \in D_1^2)$ (red and yellow curves) and $p(G_2^0|j \in D_{99}^2)$ (gray curves) (Color figure online).

It is well acknowledged that the movement of pre-existing crystal defects, like dislocations, significantly impacts the mechanical properties of crystalline alloys. Computational studies indicate that common defects, including pre-existing dislocations, grain boundaries, and free surfaces, exhibit activation barriers within the range of 0.1–0.5 eV.[48–50] On the other hand, the barriers associated with sustainable plasticity under laboratory strain rates, were estimated to be around $25k_BT$,[51] approximately 10^{-1} eV. When considering the number density of pre-existing crystalline defects, typically around 10^{-4} per atom, our calculations predict a detectable population of low-energy collective motions (around 10^{-1} eV) at approximately 10^{-1} per atom. These findings from the PEL calculations suggest that the nucleation of defects linked to non-affine collective motions in the D_1^2 group are indeed non-negligible.

It is also worth noting that, at the current stage, we are focusing solely on cryogenetic conditions, where the static atomic structure predominantly influences the overall mechanical behavior, and thermal vibrations have minimal impact. At elevated temperatures, the method of evaluating the gyration tensor for LCUs can be adjusted by considering the time-averaged positions of particles.

CONCLUSION

Using atomistic molecular dynamics simulations, we investigated the local packing of the CoCrFeMnNi high-entropy alloy (HEA). We have quantitatively described the local configuration using the local coordination unit (LCU), which is defined by the entries of the local gyration tensor and local excess electronegativity $\langle\chi\rangle$ to emphasize the interparticle collectivity. Based on this descriptive framework, we extracted the configurational features of a pre-strained HEA and statistically correlated them with the local plasticity in a strained HEA using eigenvector-based multivariate analyses. Our method was able to identify that the local non-affine displacement in a strained HEA is essentially determined by three key factors: the fluctuation of $\langle\chi\rangle$, the size of the LCU, and the orientational compatibility between the LCU and the deformation gradients of the applied strains. We provided evidence that our approach is probabilistically capable of forecasting the defect nucleation sites from the compositional heterogeneous matrix, based on the extracted configurational features. While the present study only focuses on the Cantor alloy model, we believe our findings could also be extrapolated to other HEA systems. Considering that the local composition of an N-element alloy can be expressed in an N–1 dimensional space, the currently proposed approach of evaluating averaged electronegativity effectively projects the LCU onto one of the directions in this space. Therefore, the varying affinity or bond strength between specific types of elements can also be represented as the projection of LCU composition in different directions. Additionally, data mining approaches such as PCA can help to identify the most significant combinations of element types that impact mechanical strength in future studies.

One particularly intriguing perspective is to investigate the influence of missing or extra atoms on the plasticity of HEAs. We hypothesize that the local configuration near point defects will be significantly altered by the change in atomic coordination. The same approach could also be expanded to label the heterogeneous local environment surrounding extended defects in multicomponent solid solutions and predict their properties. To pursue this ambitious study, we would need to identify new principal configuration variables and classify the local atomic structure accordingly. Our methodology provides a foundation for such inference studies.

While our simulation force field prevents a comprehensive discussion of the effect of chemical short-range ordering (CSRO)[17,52–54] on particle arrangement, we believe that a predictive configurational criterion can still be established to a priori identify potential plastic yielding sites. We anticipate that the salient configurational features most relevant to the mechanical failure of CSRO-characterized HEAs will be successfully identified through the multivariate statistical analysis proposed in this study.

Finally, our findings delineating the effect of local cooperativity on plasticity are highly relevant to the design of numerous emerging engineering alloys, such as compositionally complex alloys (CCAs),[55,56] based on the multi-element principle. Our work also identifies a crucial aspect for future first-principle theories of such complex multi-element alloys.

ACKNOWLEDGEMENTS

This research was performed at the Spallation Neutron Source, which is a DOE Office of Science User Facilities operated by Oak Ridge National Laboratory. C.-H. T. and S.-Y.C. acknowledge the support by the Ministry of Science and Technology, Taiwan, under Grant No. MOST 108-2221-E007-054-MY3. Z. B. and Y. F. acknowledge the support by the National Science Foundation under Grant No. DMR-1944879.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

APPENDIX A: PRINCIPAL COMPONENT ANALYSIS (PCA) FOR FEATURE EXTRACTION

With this descriptive framework given in "Defining the Local Configurational Unit (LCU) in the Multivariate Reference Frame," a methodology to extract the configurational features from the trajectories of the quiescent state can be established. What is not known from our earlier computational study is the influence of compositional difference on the configurational features of LCU. One way to provide a definite answer of this question is to delineate the correlation of $\langle\chi\rangle$ with other input variables from $\mathcal{R}$ in the $\mathbb{R}^7$ space via PCA: first, the original data obtained from the trajectories of N particles is partitioned into three groups: eigenvalues of $\Lambda_{\mathcal{R}}$, Euler angles, and $\langle\chi\rangle$. The average value of each group is respectively calculated then subtracted from the original data so that the each resulting variable will have a zero mean. This processed data are further normalized by the variance of original variables of each group and arranged into a $7 \times N$ matrix $\boldsymbol{F}$ where three groups

of processed variables are presented as columns. Using singular value decomposition,[33] $\boldsymbol{F}$ can be decomposed into

$$\boldsymbol{F} = U\Sigma V^T. \tag{5}$$

where U and V are unitary matrices whose columns are the eigenvectors of $\boldsymbol{F}\boldsymbol{F}^T$ and $\boldsymbol{F}^T\boldsymbol{F}$ respectively. Σ^2 is a diagonal matrix whose entries are the eigenvalues of $\boldsymbol{F}\boldsymbol{F}^T$. Because $\boldsymbol{F}$ is mean-centered, by definition $\frac{\boldsymbol{F}\boldsymbol{F}^T}{N-1}$ is the covariance matrix $\boldsymbol{C}$ with dimensions of 7×7 which conceals the information of correlation between different input variables. $\boldsymbol{C}$ can be expressed by the following block matrix form:

$$\boldsymbol{C} = \frac{\boldsymbol{F}\boldsymbol{F}^T}{N-1} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \boldsymbol{C}_{13} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \boldsymbol{C}_{23} \\ \boldsymbol{C}_{31} & \boldsymbol{C}_{32} & \boldsymbol{C}_{33} \end{bmatrix}. \tag{6}$$

where $\boldsymbol{C}_{\mathrm{ij}} \equiv \frac{\boldsymbol{F}_i\boldsymbol{F}_j^T}{N-1}$ and the sub-matrix $\boldsymbol{F}_i$ are the processed variables of each group. The dimensions of $\boldsymbol{F}_i$ are therefore $N \times M_i$ where M_i is the number of variables in each group. Because the data in $\boldsymbol{F}$ are normalized, the trace of each sub-block $\boldsymbol{C}_{\mathrm{ii}}$ is equal to M_i. This standardization process is required to avoid the undesired complication, caused by the difference in the dimensionality and unit of input variables and ensure the comparability among the correlations of different variables. From Eqs. 5 and 6, it is found that $\boldsymbol{C}$ can be expressed as $\boldsymbol{C} = \frac{U\Sigma^2U^T}{N-1}$. Given the physical meaning of Eq, 3, this eigendecoposition of $\boldsymbol{C}$ can be equivalently considered as spinning the data around onto a new basis which has desired properties that facilitates the feature extraction. In this regard, the column vectors in U are the principal axes which form an orthonormal basis in $\mathbb{R}^7$. The eigenvalues of $\frac{\Sigma^2}{N-1}$ are the percentages of the variance of the original data projected onto each corresponding principal axis. In this reference framework established by PCA, the configurational features of HEA can be extracted from the expressions of the basic vectors in terms of the input variables and the variance projected onto them.

APPENDIX B: THE COARSE GRAIN SCHEME FOR PARTITIONING THE UNIT SPHERE BY THE OCTAHEDRAL SYMMETRY ELEMENTS

In order to keep the true relative sizes and the orientation intact, we project the 3D orientation distribution function (ODF) of Q_3 onto a 2D space based the method of azimuthal equal-area projection proposed by Lambert.[57,58] The coordinates of this 2D space are defined as $(u, v) = \left(\frac{x\sqrt{\frac{1-z}{2}}}{x^2+y^2}, \frac{y\sqrt{\frac{1-z}{2}}}{x^2+y^2}\right)$ where $\frac{(x,y,z)}{\sqrt{x^2+y^2+z^2}}$ defines the orientation of Q_3 in the original 3D space. One can calculate the probability density $p(Q_3)$ from the ODF of Q_3. From the data presented in Fig. 2, the correlation between the orientation of LCU at the quiescent state in terms of Q_3 and the local plastic activity can be best revealed by the conditional probabilities. To avoid the undesired singularity issue caused the presence of marginal distribution of Q_3, the results presented in Fig. 10a need to be properly coarse-grained. Specifically, as illustrated in Fig. 10a, the unit sphere can be partitioned into eight spherical triangles by the mirror planes perpendicular to the C_4 axes based on the octahedral symmetry of the FCC structure.[59] Each of these spherical triangles, characterized by three right angles defined by three geodesics, can be further partitioned into six scalene triangles by the mirror planes perpendicular to the C_2 axes. Overall, the unit sphere is broken up into 48 triangular facets with the edges forming a Delaunay network[60] of the collection of the intersections between the unit sphere and all C_4, C_3, and C_2 axes. Each triangular facet contains three vertices corresponding to the C_4, C_3, and C_2 axes, respectively. Each facet is further divided into three domains based on the scheme of Voronoi diagram.[60] Each domain is used as the coarse-grained unit. The total projections of Q_3 for each domain associated with the vertices corresponding to C_4, C_3, and C_2 axes are recorded and accordingly the corresponding coarse-gained probability distribution $P_c(Q_3)$ has been calculated and is illustrated in Fig. 10b–d.

APPENDIX C: ANGULAR MASS DISTRIBUTION FUNCTION OF A DEFORMED OBJECT EXPANDED BY REAL SPHERICAL HARMONICS EXPANSIONS (RSHE)

In terms of the principal coordinate Q_E, the gyration tensor $\mathcal{R}$ defined in "Defining the Local Configurational Unit (LCU) in the Multivariate Reference Frame" can be transformed into $\mathcal{R}_E = Q_E^T\mathcal{R}Q_E$. In this vector space, the angular mass distribution function of an ellipsoid described by $\mathcal{R}_E$ can accordingly be expressed as:

$$\rho(\boldsymbol{r}) = (2\pi)^{-\frac{2}{3}}\exp\left[-\frac{1}{2}\boldsymbol{r}^T(\mathcal{R}_E)^{-1}\boldsymbol{r}\right]. \tag{7}$$

In terms of real spherical harmonics, $\rho(\boldsymbol{r})$ can be expanded as:

$$\rho(\boldsymbol{r}) = \sum_{l,m}\rho_l^m(r)Y_l^m(\boldsymbol{r}). \tag{8}$$

where Y_l^m is the real spherical harmonics of order l and degree m. The second moment of $\rho_l^m(r)$ can be defined as:[32]

$$G_l^m = \frac{\int drr^2 \rho_l^m(r)}{\int drr^2}. \tag{9}$$

In terms of G_l^m, the nine entries of $\mathcal{R}_E$ can be expressed as:

$$\mathcal{R}_E = \begin{bmatrix} -\frac{\sqrt{5}}{15}G_2^0 + \frac{1}{\sqrt{15}}G_2^2 & \frac{1}{\sqrt{15}}G_2^{-2} & \frac{1}{\sqrt{15}}G_2^1 \\ \frac{1}{\sqrt{15}}G_2^{-2} & -\frac{\sqrt{5}}{15}G_2^0 - \frac{1}{\sqrt{15}}G_2^2 & \frac{1}{\sqrt{15}}G_2^{-1} \\ \frac{1}{\sqrt{15}}G_2^1 & \frac{1}{\sqrt{15}}G_2^{-1} & \frac{2\sqrt{5}}{15}G_2^0 \end{bmatrix} + \frac{\mathrm{tr}(\mathcal{R}_\mathrm{E})}{3} \cdot \boldsymbol{I}_3. \tag{10}$$

From Eqs. 9 and 10, G_2^0 is found to take the following expression:

$$G_2^0 = \frac{\sqrt{5}}{2}\left(2\mathcal{R}_{E,\,33} - \mathcal{R}_{E,\,11} - \mathcal{R}_{E,\,22}\right). \tag{11}$$

In Fig. 11, we give the probability distribution of G_2^0 for the two extreme groups associated with 1st percentile and 99th percentile of D^2 which are specified as $p\left(G_2^0 | j \in D_1^2\right)$ and $p\left(G_2^0 | j \in D_1^2\right)$ respectively. The curves of $p\left(G_2^0 | j \in D_1^2\right)$ are found to shift positively from that of $p\left(G_2^0 | j \in D_{99}^2\right)$ in both tension and shear conditions.

APPENDIX D: CALCULATION OF RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES AND AREA UNDER THE ROC CURVE (AUC)

In a general binary classification problem, the performance of classification model is affected by the selection of decision threshold c_th. One can define the true positive rate (TPR) to represent the sensitivity of a model:[39]

$$\mathrm{TPR} = \frac{N(\text{true positive})}{N(\text{condition positive})}. \tag{12}$$

Likewise, to represent the probability of false alarm, the FPR is defined as

$$\mathrm{FPR} = \frac{N(\text{false negative})}{N(\text{condition negative})}. \tag{13}$$

where $N(\text{true positive})$, $N(\text{condition positive})$, $N(\text{false negative})$ and $N(\text{condition negative})$ represent the instances of each event specified in the parentheses. One can examine the tradeoff between TPR and FPR as a function of $c_\mathrm{th}$from the ROC curve where TPR are plotted against FPR. In our study, if $c_\mathrm{th} = \max(\boldsymbol{v}^T\boldsymbol{f})$, all the data will be classified negatively by our model classifies and thus $(\mathrm{FPR}, \mathrm{TPR}) = (0, 0)$. Likewise, if $c_\mathrm{th} = min(\boldsymbol{v}^T\boldsymbol{f})$, all the data will be classified positively and $(\mathrm{FPR}, \mathrm{TPR}) = (1, 1)$. A perfect classification model is characterized by the maximum sensitivity with giving no false alarm. Therefore $(\mathrm{FPR}, \mathrm{TPR}) = (0, 1)$ for this ideal condition. In our case, $N(\text{condition negative}) = N(\text{condition positive}) = 0.5$. The accuracy (ACC) can be defined as[39]

$$\begin{aligned} ACC &\equiv \frac{N(\text{true positive}) + N(\text{true negative})}{N(\text{condition positive}) + N(\text{condition negative})} \\ &= \frac{1}{2}(TPR - FPR + 1). \end{aligned} \tag{14}$$

APPENDIX E: SYMBOLS AND NOTATIONS USED IN THIS PAPER

See Table III.

Table III. Summary of the key variables used in this paper along with their physical meanings

Symbol	Meaning
R, $\mathcal{R}_{\alpha\beta}$	Gyration tensor and its entries
$\boldsymbol{Q}$	Eigenvectors of R
$\boldsymbol{F}$	$7 \times N$ feature matrix
$\boldsymbol{S}$	Singular matrix of $\boldsymbol{F}$
G_l^m	Second moment of LCU associated with spherical harmonics of order l and degree m
D^2	Non-affine parameter
D_1^2, D_{99}^2	Set of particle with the 1st and 99th D^2
$\boldsymbol{J}$	Deformation gradient tensor
U_i	Coefficients associated with LCU shapes

REFERENCES

1. J. Yeh, S. Chen, S. Lin, J. Gan, T. Chin, T. Shun, C. Tsau, and S. Chang, *Adv. Eng. Mater.* 6(5), 299 (2004).
2. B. Cantor, I. Chang, P. Knight, and A. Vincent, *Mater. Sci. Eng. A* 375, 213 (2004).
3. Y. Zhang, T.T. Zuo, Z. Tang, M.C. Gao, K.A. Dahmen, P.K. Liaw, and Z.P. Lu, *Prog. Mater Sci.* 61, 1 (2014).
4. Z. Li, K.G. Pradeep, Y. Deng, D. Raabe, and C.C. Tasan, *Nature* 534(7606), 227 (2016).
5. D.B. Miracle and O.N. Senkov, *Acta Mater.* 122, 448 (2017).
6. E.P. George, W. Curtin, and C.C. Tasan, *Acta Mater.* 188, 435 (2020).
7. E.P. George, D. Raabe, and R.O. Ritchie, *Nat. Rev. Mater.* 4(8), 515 (2019).
8. H. Diao, R. Feng, K. Dahmen, and P. Liaw, *Curr. Opin. Solid State Mater. Sci.* 21(5), 252 (2017).
9. F. Otto, A. Dlouhý, C. Somsen, H. Bei, G. Eggeler, and E.P. George, *Acta Mater.* 61(15), 5743 (2013).
10. Z. Zhang, M. Mao, J. Wang, B. Gludovatz, Z. Zhang, S.X. Mao, E.P. George, Q. Yu, and R.O. Ritchie, *Nat. Commun.* 6(1), 1 (2015).
11. C. Varvenne, A. Luque, and W.A. Curtin, *Acta Mater.* 118, 164 (2016).
12. C. Varvenne, G.P.M. Leyson, M. Ghazisaeidi, and W.A. Curtin, *Acta Mater.* 124, 660 (2017).
13. S.I. Rao, C. Woodward, T.A. Parthasarathy, and O. Senkov, *Acta Mater.* 134, 188 (2017).
14. J. Li, Q. Fang, B. Liu, and Y. Liu, *Acta Mater.* 147, 35 (2018).
15. L. Zhang, Y. Xiang, J. Han, and D.J. Srolovitz, *Acta Mater.* 166, 424 (2019).
16. F.G. Coury, M. Kaufman, and A.J. Clarke, *Acta Mater.* 175, 66 (2019).
17. Q. Ding, Y. Zhang, X. Chen, X. Fu, D. Chen, S. Chen, L. Gu, F. Wei, H. Bei, and Y. Gao, *Nature* 574(7777), 223 (2019).
18. L. Li, Q. Fang, J. Li, B. Liu, Y. Liu, and P.K. Liaw, *Mater. Sci. Eng.* A, 139323 (2020).
19. J.M. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems* (Cambridge University Express, Cambridge, 1979).
20. Z. Wang and W.-H. Wang, *Natl. Sci. Rev.* 6(2), 304 (2019).
21. P.J. Steinhardt, D.R. Nelson, and M. Ronchetti, *Phys. Rev. Lett.* 47(18), 1297 (1981).
22. T. Ninomiya, Topological disorder in condensed matter, in *Topological Disorder in Condensed Matter*, vol 46. ed. by F. Yonezawa, and T. Ninomiya (Springer, Berlin, 1983), p. 159.
23. H. Peng, M. Li, and W. Wang, *Phys. Rev. Lett.* 106(13), 135503 (2011).
24. J. Ding, S. Patinet, M.L. Falk, Y. Cheng, and E. Ma, *Proc. Natl. Acad. Sci. U.S.A.* 111(39), 14052 (2014).
25. X. Yang, R. Liu, M. Yang, W.-H. Wang, and K. Chen, *Phys. Rev. Lett.* 116(23), 238003 (2016).
26. C.-H. Tung, G.-R. Huang, Z. Bai, Y. Fan, W.-R. Chen, S.-Y. Chang, Structural origin of plasticity in strained high-entropy alloy. arXiv:2005.07088 (2020).
27. H.S. Oh, S.J. Kim, K. Odbadrakh, W.H. Ryu, K.N. Yoon, S. Mu, F. Körmann, Y. Ikeda, C.C. Tasan, and D. Raabe, *Nat. Commun.* 10(1), 1 (2019).
28. S. Plimpton, *J. Comput. Phys.* 117(1), 1 (1995).
29. B.-J. Lee, J.-H. Shim, and M.I. Baskes, *Phys. Rev. B* 68, 144112 https://doi.org/10.1103/PhysRevB.68.144112 (2003).
30. W.-M. Choi, Y.H. Jo, S.S. Sohn, S. Lee, and B.-J. Lee, *NPJ Comput. Mater.* 4(1), 1 (2018).
31. M.L. Falk and J.S. Langer, *Phys. Rev. E* 57(6), 7192 (1998).
32. G.-R. Huang, Y. Wang, C. Do, L. Porcar, Y. Shinohara, T. Egami, and W.-R. Chen, Determining gyration tensor of orienting macromolecules through their scattering signature. *J. Phys. Chem. Lett.* 10(14), 3978 (2019).
33. G. Strang, *Linear Algebra and Learning from Data* (Wellesley-Cambridge Press, Wellesley, 2019).
34. I.T. Jolliffe, *Principal Components in Regression Analysis* (Springer, New York, 1986).
35. T. Egami, *Prog. Mater. Sci.* 56(6), 637 (2011).
36. J. Ding and Y. Cheng, *Appl. Phys. Lett.* 104(5), 051903 (2014).
37. D.N. Theodorou and U.W. Suter, *Macromolecules* 18(6), 1206 (1985).
38. K. Šolc and W.H. Stockmayer, *J. Chem. Phys.* 54(1), 2756 (1971).
39. K.P. Murphy, *Machine Learning: A Probabilistic Perspective* (MIT press, Cambridge, 2012).
40. Y. Fan, T. Iwashita, and T. Egami, *Nat. Commun.* 5, 5083 (2014).
41. H.J. Hwang, R.A. Riggleman, and J.C. Crocker, *Nat. Mater.* 15(9), 1031 (2016).
42. Y. Fan, T. Iwashita, and T. Egami, *Nat. Commun.* 8, 15417 (2017).
43. C. Liu, P. Guan, and Y. Fan, *Acta Mater.* 161, 295 (2018).
44. P. Cao, M.P. Short, and S. Yip, *Proc. Natl. Acad. Sci. U.S.A.* 116(38), 18790 (2019).
45. G.T. Barkema and N. Mousseau, *Phys. Rev. Lett.* 77(21), 4358 (1996).
46. E. Machado-Charry, L.K. Béland, D. Caliste, L. Genovese, T. Deutsch, N. Mousseau, and P. Pochet, *J. Chem. Phys.* 135(3), 034102 (2011).
47. E. Cances, F. Legoll, M.C. Marinica, K. Minoukadeh, and F. Willaime, *J. Chem. Phys.* 130(11), 114711 (2009).
48. T. Zhu, J. Li, and S. Yip, *Phys. Rev. Lett.* 93(2), 025503 (2004).
49. T. Zhu, J. Li, A. Samanta, A. Leach, and K. Gall, *Phys. Rev. Lett.* 100(2), 025502 (2008).
50. M. McPhie, S. Berbenni, and M. Cherkaoui, *Comput. Mater. Sci.* 62, 169 (2012).
51. A. Cottrell, *Philos. Mag.* 86(25–26), 3811 (2006).
52. J. Ding, Q. Yu, M. Asta, and R.O. Ritchie, *Proc. Natl. Acad. Sci. U.S.A.* 115(36), 8919 (2018).
53. E. Antillon, C. Woodward, S. Rao, B. Akdim, and T. Parthasarathy, *Acta Mater.* 190, 29 (2020).
54. F. Zhang, S. Zhao, K. Jin, H. Xue, G. Velisa, H. Bei, R. Huang, J. Ko, D. Pagan, and J. Neuefeind, *Phys. Rev. Lett.* 118(20), 205501 (2017).
55. A.M. Manzoni and U. Glatzel, *Mater. Charact.* 147, 512 (2019).
56. H.M. Daoud, A.M. Manzoni, N. Wanderka, and U. Glatzel, *JOM* 67(10), 2271 (2015).
57. J.P. Snyder and P.M. Voxland, *An Album of Map Projections* (US Government Printing Office, Washington, 1989).
58. G. Marsaglia, *Ann. Math. Stat.* 43(2), 645 (1972).
59. M.J. Buerger, *Elementary Crystallography: An Introduction to the Fundamental Geometric Features of Crystals* (The MIT Press, Cambridge, 1978).
60. T. Ogawa, Problems in a digital description of a configuration of atoms and some other geometrical topics in physics, in *Topological Disorder in Condensed Matter*. ed. by F. Yonezawa, and T. Ninomiya (Springer, Berlin, 1983), pp. 60–77.