npj Computational Materials

https://doi.org/10.1038/s41524-026-01997-z

Article in Press

Dual machine learning pinpoints the Radius of Informative Structural Environments in metallic glasses

Received: 21 October 2025

Accepted: 4 February 2026

Cite this article as: Wang, M., Wang, Y., Islam, M. *et al.* Dual machine learning pinpoints the Radius of Informative Structural Environments in metallic glasses. *npj Comput Mater* (2026). https://doi.org/10.1038/s41524-026-01997-z

Muchen Wang, Yuchu Wang, Minhazul Islam, Yuchi Wang, Yunzhi Wang, Jinwoo Hwang & Yue Fan

Dual Machine Learning Pinpoints the Radius of Informative Structural Environments in Metallic Glasses

Muchen Wang[1, †], Yuchu Wang[1, †], Minhazul Islam[2], Yuchi Wang[2], Yunzhi Wang[2], Jinwoo Hwang[2], and Yue Fan[1]

[1]*Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI*

[2]*Department of Materials Sciences and Engineering, Ohio State University, Columbus, OH*

([†] *Equally contributed first authors*)

Abstract: The disordered nature of amorphous materials like metallic glasses has long hindered the establishment of well-defined structure-property relationships. Although it is widely recognized that short-range orders (SROs) within the first nearest-neighbor shell do not sufficiently characterize these materials, identifying the optimal characteristic length scale for capturing richer structural information remains elusive. Here, we resolve this ambiguity using a dual machine learning (ML) approach, which identifies the Radius of Informative Structural Environments (RISE) in a prototypical Zr-Cu metallic glass system. A top-down, reductionist approach, integrating SOAP descriptor with XGBoost model, demonstrates that the atomic environments within 5 Å radius entail maximal structural diversity and information density, leading to the optimal performance of the model on predicting given samples' configurational energies. Concurrently, a bottom-up, emergentist Vision Transformer (ViT) architecture, designed to autonomously learn structural patterns from voxelized atomic configurations, shows that its predictive performance saturates when the effective communication length between its input patches reaches an equivalent spherical radius of ~5 Å. The striking convergence of these independent ML strategies provides compelling, data-driven evidence for the existence of an intrinsic, structurally informative length scale in metallic glasses. Additional robustness checks across multiple glassy materials with various elements numbers and bonding types confirm such RISE is not an artifact of encoding parameters or system size and aligns with existing experimental and computational insights.

[1] Corresponding author, E-mail: fanyue@umich.edu

Introduction

Establishing robust and interpretable structure–property relationships stands as a central objective in materials science, enabling the prediction and design of materials with tailored properties. In crystalline solids, this endeavor has achieved remarkable success by leveraging the pivotal topological defects such as dislocations and grain boundaries. For instance, the Taylor equation links dislocation density to the material's flow stress[1], while the Hall–Petch relationship connects grain boundary concentration (*i.e.*, grain size) to the samples' yield strength[2,3]. These successes are primarily attributed to the inherent lattice periodicity and translational symmetry in crystalline systems. Conversely, in amorphous materials such as metallic glasses, the absence of long-range lattice orders and unambiguously identifiable defects presents formidable challenges[4–13], greatly complicating efforts to establish analogous structure-property relationships.

Over recent decades, substantial efforts have been directed toward understanding the atomic-level structures of metallic glasses and their implications for the samples' properties. Voronoi tessellation-based analysis of short-range orders (SROs) has been one of the instrumental efforts in characterizing the local atomic packing environments[14–18]. For example, it is widely believed that the SROs with high five-fold symmetries (*e.g.*, the icosahedron with Voronoi index <0,0,12,0> and other Kasper clusters) could enhance the sample's overall stability. However, more recent studies suggest that SROs alone are insufficient to fully describe metallic glasses' behaviors[19–27]. In particular, it was found that the same SROs, when embedded into various surrounding environments, may impose significantly different or even opposite impacts on the sample's global stability[19]. In other words, the recognition of structural correlations extending beyond nearest neighbors, *e.g.* medium-range orders, has emerged as a critical factor in building a more faithful structure-property relationship in metallic glasses. Despite the growing consensus, substantial ambiguity persists regarding the optimal characteristic length scale that is needed for capturing richer structural information. Various hypotheses have been proposed, from interpenetrating connections of icosahedra (ICOI) network[14,21,28–31], to the consequences of species-specific efficient clusters packing[32,33], and density fluctuations[34,35], *etc*. And a wide range of estimated length scales – ranging from a few angstroms to several nanometers or larger – have been assessed, even for the same alloys, based on different measurements and assumptions[22,27,36–38]. In other words, whether or not there exists an objective structurally informative characteristic length scale has remained a contentious issue. And the lack of consensus on this issue in turn prohibits the quest to build an effective model on the structure-property relationship for metallic glasses.

Data-driven techniques, such as machine learning (ML) methods, have recently emerged as powerful tools to decipher the disordered materials. Most existing ML-based studies, however, focus predominantly on linking localized atomic structural motifs to equally localized properties, such as local softness, propensity for plastic deformation, local thermal activation, *etc*[5,18,20,24,25,39]. Although these localized descriptors offer valuable insights, translating them into predictive and experimentally verifiable global material properties remains challenging. Therefore, the scope of our present study is set to directly correlate a given sample's global property with its structural features. Here, we choose the configurational energy of the entire sample as our global property of interest. This choice is motivated by several considerations: firstly, it is an unambiguously defined and computationally accessible metric with good reproducibility, thereby simplifying initial model development; secondly, the enthalpies of metallic glasses have a well-documented

experimental correlation with the critical global mechanical properties[40–44]. In particular, under the same loading condition, a sample with higher enthalpy is in general more ductile than that of a sample at a lower enthalpy state. It hence underscores the significance of configurational energy in dictating metallic glasses' macroscopic responses.

More specifically, to tackle the above-identified critical knowledge gap, namely whether or not there exists an intrinsic characteristic length scale that dictates the metallic glasses' global property, we employ two complementary machine learning (ML) architectures to primarily investigate a widely adopted prototypical $Zr_{44}Cu_{56}$ metallic glass model, aiming to provide an unbiased, data-driven perspective on this matter. The first approach embodies a top-down, reductionist strategy combining the smooth overlap atomic position[45] (SOAP) encoding and a subsequent Extreme Gradient Boosting[46] (XGBoost) model. More specifically, by decomposing the input samples' configurations into atomic motifs with controlled sizes and then systematically varying the predefined sizes, we identify a distinct optimal unit size at approximately a radius of 5 Å, under which the ML model's performance peaks, suggesting that this is a focal domain encapsulating the most effective structural information. The excess configurational entropy also presents the broadest distribution within this length scale, providing a physical rationale for such an optimal performance. Complementarily, the second approach adopts a bottom-up, emergentist perspective, employing a Vision Transformer[47] (ViT) architecture that autonomously identifies structural patterns without pre-imposed size constraints. By systematically adjusting the effective communication length between the model's smallest input patches, we observe a saturation in the ML model's predictive performance beyond the effective communication length of 8 Å (equivalent to a spherical radius of 5 Å), further reinforcing the characteristic length scale of such Radius of Informative Structural Environments (RISE). The remarkable convergence of results from these independent, unbiased approaches provides compelling evidence supporting the existence of a distinct and structurally informative length scale in metallic glasses. To further validate the generality of our findings, we examined a diverse set of glassy models – ranging from ternary metallic alloys, to binary glasses with intermetallic bonding, and even to single-element amorphous solids with covalent bonding – and observed consistent behavior across all cases. Consequently, these findings may lay a foundational basis for developing more accurate, robust, and interpretable structure-property models for amorphous materials in future research.

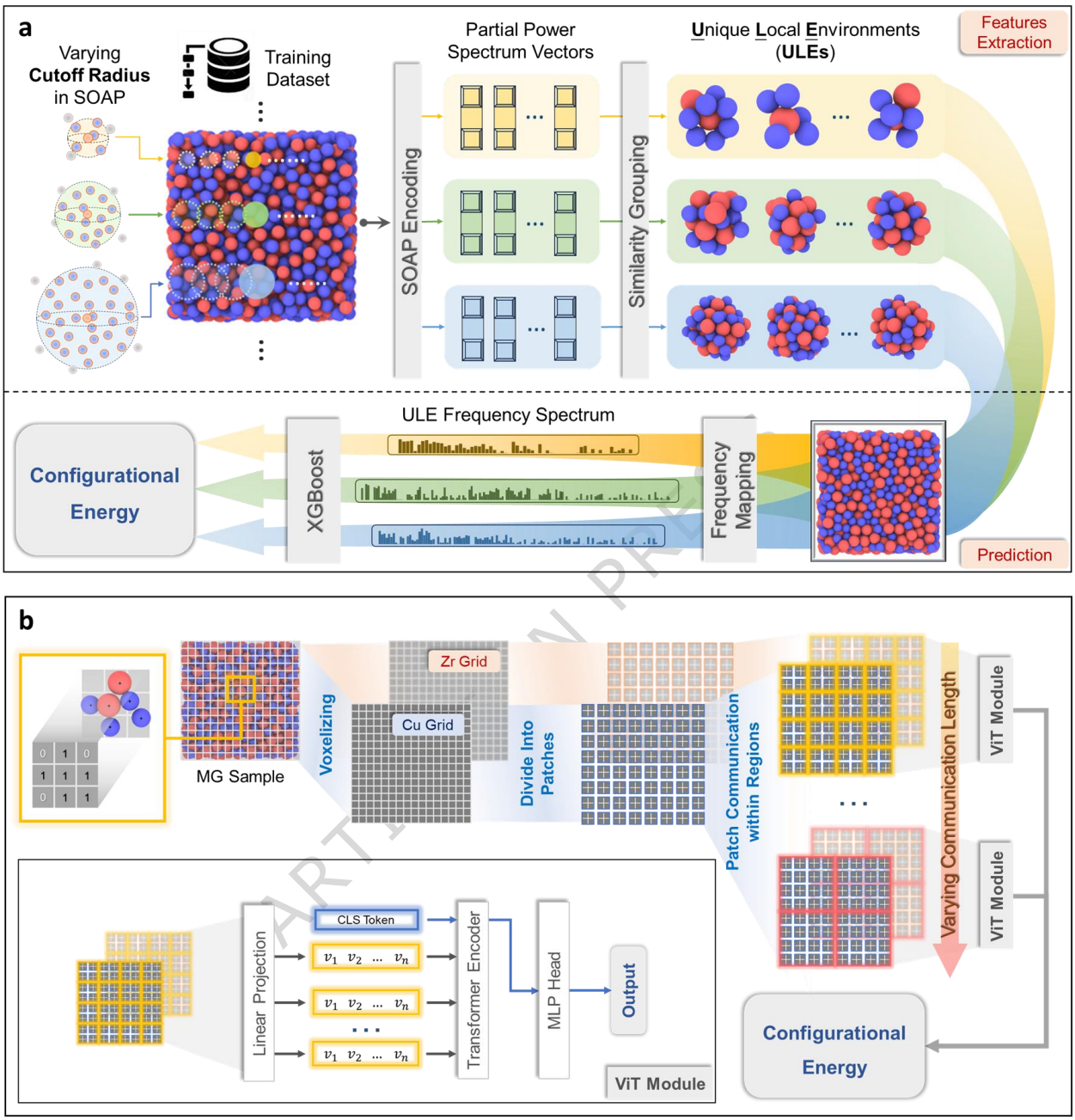

Fig. 1 | Architectures of the models. (a) Reductionist model architecture. Each sample is first processed using the SOAP descriptor, which encodes local atomic environments as partial power spectrum vectors. By systematically varying the SOAP cutoff radius (3.0–9.0 Å), the model captures structural features across multiple length scales. Similarity grouping is then performed to group local environments into ULEs, reducing redundancy while preserving diversity. These ULEs are used to construct ULE frequency spectrum that represent each sample's composition. An XGBoost model, trained on these frequency vectors, predicts the configurational energy of each sample. (b) Emergentist model architecture (visualized in 2D for clarity). Each sample is voxelized into a pair of binary 3D grids, with separate channels representing Cu and Zr atomic species. The simulation box is discretized into voxels, which are then partitioned into non-overlapping patches to enhance spatial abstraction and computational efficiency. To explore multiscale structural correlations, we define regions of patch communication based on a tunable communication length, ranging from 2 to 32 Å. Within each defined region, patch embeddings are generated through 3D convolutions and processed by a ViT module. The ViT

module integrates patch information via multi-head self-attention calculations, with a CLS token encoding the global representation of the structure. The output from the transformer is passed through a multi-layer perceptron to predict the sample's configurational energy.

Results

In the following Results, we analyze two complementary ML frameworks – a reductionist SOAP-based XGBoost model and an emergentist ViT model – to identify the characteristic length scale in metallic glasses. Both models are systematically evaluated across different cutoff or communication lengths and alloy systems, allowing direct comparison of their predictive performance and length-scale behavior. Together, these analyses establish the robustness and physical relevance of the identified optimal length scale.

Reductionist Model Architecture

Our top-down reductionist model adopts a systematic approach rooted in the core principle of reductionism, which involves decomposing complex systems into simpler, discrete components to elucidate their roles and contributions to the whole. The overall workflow of the architecture is illustrated in **Fig. 1a**.

In accordance with this principle, we first decompose the metallic glass samples into distinct local atomic environments, each represented by the partial power spectrum vectors (or SOAP spectrum vectors) via SOAP encoding. A key strength of the SOAP descriptor lies in its inherent rotational and translational invariance, which is guaranteed by its spherical harmonics expansion algorithm. It hence ensures that structural comparisons remain consistent regardless of the orientation or positional differences of atomic environments. We systematically vary the encoding cutoff radius of the SOAP descriptor while holding all other parameters fixed. Specifically, the cutoff radius is adjusted incrementally from 3.0 to 9.0 Å, covering spatial scales that range from the immediate 1st nearest-neighbor shell to distances extending beyond the 3rd nearest-neighbor shell. This carefully chosen range encompasses the spatial scales where additional information beyond SROs is expected to emerge[21,28–31].

Following this encoding step, local atomic environments are categorized into Unique Local Environments (ULEs) through clustering based on their SOAP spectrum vectors and a similarity threshold. This clustering step significantly reduces redundancy by grouping structurally similar atomic environments, effectively simplifying the dataset. Such data reduction is essential, as it enhances computational efficiency and minimizes the risk of overfitting – a particularly pressing concern when analyzing extensive datasets comprising millions of atomic environments, each represented by high-dimensional vectors containing thousands of features. Subsequently, each local atomic environment within a given sample is assigned to its closest matching ULE, thereby constructing a frequency vector that quantitatively captures the distribution of ULEs within each sample.

It is worth emphasizing that, to focus our analysis on the intrinsic length scale of RISE, all metallic glass samples in the present study were thermally prepared and stress-free. Because this way the

influences from extrinsic factors, such as mechanical loading-induced shear bands and structural heterogeneities, can be avoided. More detailed procedures concerning sample preparation, ULE identification, and the subsequent construction of these frequency vectors are elaborated in the Methods section. Ultimately, we employ the XGBoost algorithm – optimized using Bayesian hyperparameter tuning techniques – to establish a surrogate model to correlate the ULE frequency spectrums with the configurational energies of the metallic glass samples.

Model's Non-monotonic Performance & Maximum Information Density

The model training of such reductionist architecture has been highly successful. For example, **Fig. 2a** presents a typical model's performance with the SOAP cutoff radius at 5 Å, yielding a test-set root mean square error (RMSE) of 0.90 eV (*i.e.*, 0.45 meV/atom) and a coefficient of determination (R^2) of 0.9749. We then compared the model's performance under various cutoff radii, and remarkably, as shown in **Fig. 2b** (blue data points), the RMSE of the model exhibits a non-monotonic variation and minimizes at an intermediate cutoff radius around 5 Å, roughly around the 2nd nearest neighbor distance. After scrutinization, the optimal ML performance at such a length scale is attributed to the two factors.

The first factor is the effective number of ULEs under various cutoff radii. It is worth noting that, after the very first step of SOAP encoding, there is always the same number of raw SOAP spectrum vectors, regardless of the specific cutoff radius value. This is because the SOAP encoding was applied across all individual atoms in each sample of the same dataset. However, the effective ULEs count diverges after the similarity grouping stage. Even though the exact same similarity threshold values have been applied (details in the Discussion and Methods sections below), the number of ULEs under various cutoff radii presents a non-monotonic profile (red data points in **Fig. 2b**), echoing the model's performance variation. This is because, at a very small cutoff radius, the SOAP descriptor essentially captures only the SROs, so that the local environments are less diverse, and fewer ULEs are identified. On the other hand, at a very large cutoff radius, the local environment becomes increasingly similar for many close-by atoms, causing a reduced distinctiveness of atomic environments and, subsequently, a smaller number of ULEs. In contrast, at the intermediate cutoff, there is the highest number of ULEs, suggesting the greatest structural richness and diversity. To assess the robustness of the ULE construction and model performance, we performed five independent trials with different random initialization seeds. Despite starting from distinct atomic configurations, the resulting number of ULEs converged to nearly identical values (red error bars in **Fig. 2b**, more details in Supplemental Material Note 1). The corresponding model accuracy also remained stable (blue error bars in **Fig. 2b**). These results confirm that the model is insensitive to initialization and consistently yields an equally informative set of ULEs, supporting the reliability of the reductionist framework.

To further validate these findings and better understand the physical origin of the non-monotonic structural richness, we examined the mean value of excess configurational entropy[48] (S_2) of each resolved individual ULE. More specifically, we first calculated the S_2 value for every atom according to its radius distribution function (more details in Methods), then computed the average S_2 within each ULE by summing the individual atomic S_2 values and dividing by the number of atoms assigned to that ULE. **Fig. 2c** shows the comparison of the distributions of S_2 under various cutoff radii. It can be seen that the configurational entropy covers the broadest range at the 5 Å

cutoff radius, reflecting that the maximum information density is entailed within such a characteristic length scale.

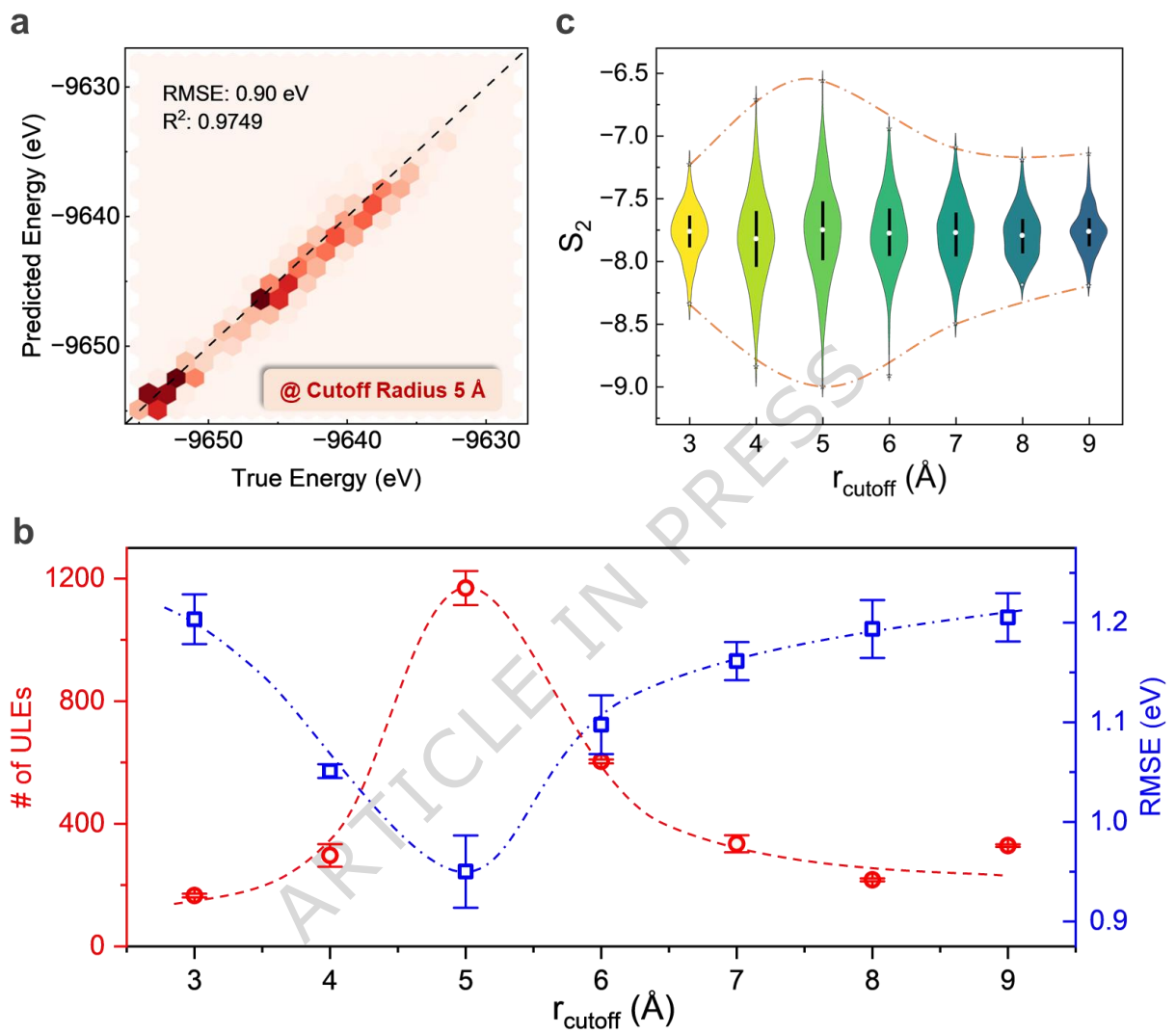

Fig. 2 | Reductionist model performance and peak of information density. (a) Prediction heatmap of the reductionist model trained using SOAP descriptors at 5 Å cutoff radius. (b) Relationship between the cutoff radius, number of ULEs, and model performance. Error bars represent the variation obtained from five independently seeded ULE searches. More details in Supplemental Material Note 1. (c) The distribution of average S_2 for ULEs at each cutoff radius, and the stars mark the upper and lower bounds of S_2 at each cutoff.

Comparison of Information Quality

We further evaluated the information quality captured under different SOAP cutoff radii. To ensure a fair comparison across radii, we ranked the retrieved ULEs based on their SHAP values at each prescribed cutoff, obtained via SHapley Additive exPlanations[49] (see Supplemental Material Note

2). Then we compared the model's predictive performance under various cutoff radii by using the same fraction of the most important ULEs. **Fig. 3** shows the performance landscape with varying fractions at 1%, 2%, 4%, 8%, 16%, 32%, 64%, and 100%, respectively. It can be seen that the RMSE depends jointly on the retained fraction of ULEs and the cutoff radius, and not surprisingly, a higher fraction of ULEs leads to better performance. Remarkably, however, the best performance consistently occurs when the cutoff radius is 5 Å, regardless of the retained fraction of ULEs. As an example, we present the detailed RMSE distributions under two scenarios in **Fig. 3**, using 8% and 100% of retrieved ULEs, respectively. Evidently, the optimal cutoff radius that yields the narrowest and most symmetric error distributions is 5 Å. In contrast, the performance at other suboptimal radii is poorer, as characterized by their broader error distributions and higher variance. These results highlight that the local environments within a 5 Å cutoff radius provide the highest-quality information for the ML model. Since no prior assumptions about structural significance were imposed during training, these findings reveal an intrinsic characteristic length scale – the RISE – that maximizes information density and quality, driving the model's optimal performance.

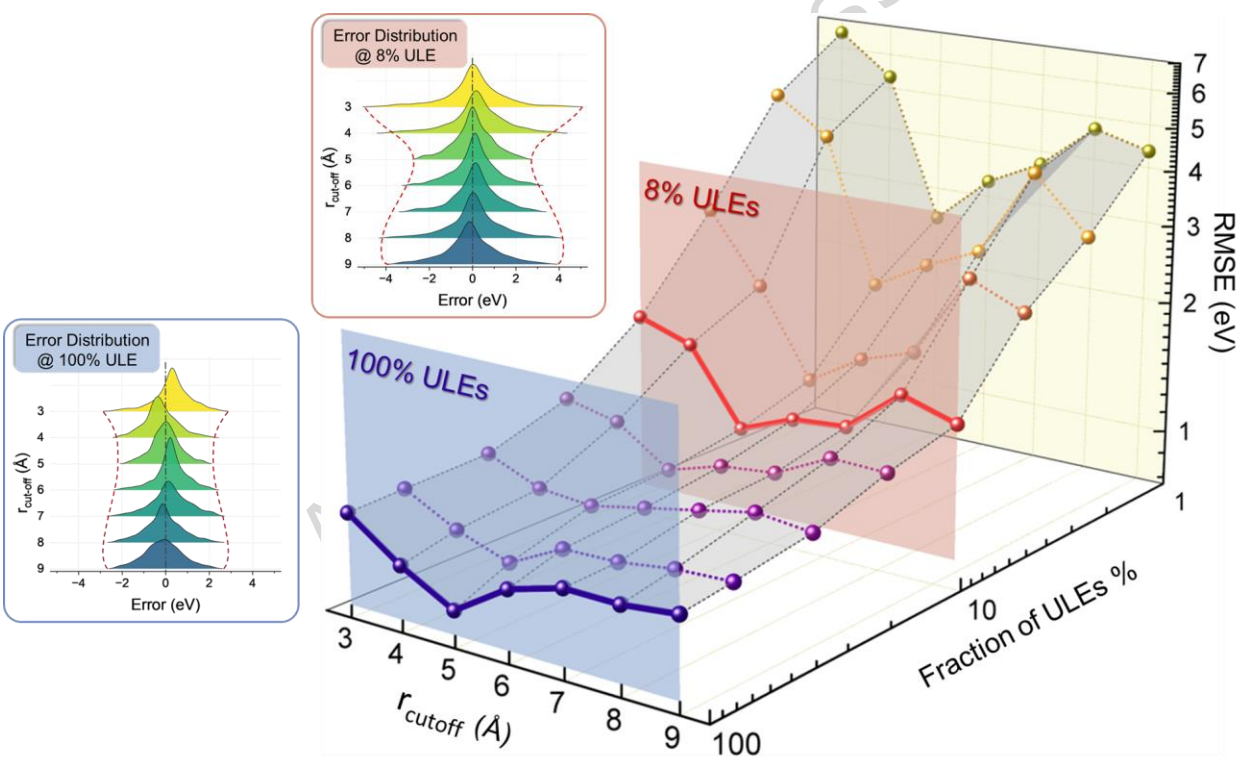

Fig. 3 | Performance landscape of the reductionist model. RMSE as a function of SOAP cutoff radius and the fraction of top-ranked ULEs retained for prediction. The subpanels display the prediction error distributions for models at different cutoffs trained using the top 8% and the full 100% of ULEs.

Emergentist Model Architecture

To validate the findings by the top-down model above, we further employ a completely different bottom-up emergentist model architecture depicted in **Fig. 1b**, which is founded on the principle

that global behaviors of a given system arise from collective interactions among its simple constituents. More specifically, inspired by recent advancements in ViTs, which excel at learning hierarchical representations from grid-based data without predefined local features, our approach begins by voxelizing each atomic configuration into a two-channel 3D grid. Each channel corresponds to one atomic species (e.g., Cu and Zr). In particular, we discretize the simulation box (with a side length of 32.43 Å) into a 32 × 32 × 32 voxel grid, yielding voxels with a side length of approximately 1.01 Å, smaller than typical interatomic spacing, to ensure that each voxel contains at most one atom. This results in a binary 3D grid per atomic species channel, with voxel values indicating the presence (1) or absence (0) of an atom type. Consequently, each sample is represented by two overlaid binary 3D grids, preserving atomic identities and spatial arrangements. Instead of directly feeding individual voxels into the model, we partition these grids into non-overlapping 2×2×2 patches, yielding 16×16×16 patches per sample. Aggregating voxel data into patches enhances computational efficiency and maintains essential spatial relationships. To systematically probe spatial correlations at different scales, we restrict patch communication within cubic regions defined by the communication length, ranging from 2 to 32 Å. In practice, this communication length is implemented as a block-wise partition of the patch grid, such that self-attention is computed only among patches within a fixed cubic region, rather than via a strict distance-based (radial) cutoff between patch centers. As a result, the physically relevant interaction range is more naturally interpreted as an effective receptive field, which is jointly determined by the communication length and the finite patch size. This formulation allows the model to capture a spectrum of structural information, from local atomic arrangements to broader system-wide correlations. To avoid boundary artifacts that could bias spatial encoding, periodic padding is applied to all voxel grids prior to patch extraction, ensuring translational continuity consistent with the physical periodic boundary conditions (PBC) of the simulation box.

These patch-based representations are subsequently processed by the ViT module (detailed in Method) to predict the configurational energy. Within the ViT module, patches from both atomic species channels undergo linear projection via 3D convolutional kernels, transforming raw voxel values into embedding vectors suitable for transformer input. A *classification* (CLS) token, serving as the global structural representation, is prepended, and positional embeddings are added to preserve spatial context within the 3D grid. This sequence is input into a transformer encoder, where multi-head self-attention calculation selectively integrates structural information from patches within each cubic region, corresponding to the chosen communication length. The encoded representation of the CLS token is then passed through a multi-layer perceptron (MLP), producing the final configurational energy prediction. Thus, the ViT module generates an energy estimate specifically informed by structural features at its designated communication length.

It is worth noting that the use of cubic patches is primarily driven by computational and geometric constraints. Implementing spherical receptive patches, although conceptually appealing for direct comparison with the spherical cutoff radius in the reductionist model, is technically impractical because non-overlapping spheres cannot tessellate flat three-dimensional space, while overlapping spheres would result in substantial data redundancy and conflicts in attention mapping. Moreover, the cubic tessellation naturally aligns with the orthorhombic simulation box and PBCs typically employed in atomistic simulations, providing an efficient and physically consistent discretization scheme. These considerations collectively justify our adoption of cubic patch partitioning in the emergentist model.

Saturation of Model's Prediction beyond Critical Communication Length

The model training of such emergentist architecture has been highly successful, and as an example, **Fig. 4a** shows the training performance with an 8 Å communication length. Unlike the top-down reductionist approach that relies on explicitly defined local atomic environments, such a bottom-up emergentist model learns structural correlations directly from the voxelized atomic grids. With that being said, due to the cubic geometry inherent in the model input, the communication length used in attention computations cannot be directly compared to the spherical cutoff radius employed in SOAP-based encoding. To facilitate comparison, we define a spherical equivalent radius by matching the volume of a cube to that of a sphere, yielding $r \cong 0.62l$. This transformation allows us to interpret the cubic communication length within a meaningful and comparable structural length scale. **Fig. 4b** shows the predictive performance as a function of the communication length. A clear saturation emerges beyond a communication length of 8 Å, corresponding to an equivalent radius of approximately 5 Å, remarkably consistent with the RISE identified by the reductionist model discussed above. This convergence strongly suggests that the 5 Å radius represents a critical structural length scale, beyond which additional structural details only marginally enhance model accuracy. The consistently low and stable RMSE values observed beyond this length confirm that the essential structural features required to predict a sample's global configurational energy are effectively captured within this scale, and that extending the range further offers no additional meaningful information.

Model Interpretation and Important Region Identification

To probe how the model internally identifies and utilizes structural information across different scales, we employed the Integrated Gradients[50] (IG) method to highlight the spatial regions that contribute most significantly to energy prediction. Interestingly, we observed that for the majority of samples, the most influential regions identified by IG begin to spatially overlap once the communication length reaches 8 Å or larger, as shown in the example in the left panel of **Fig. 4c**. This suggests that the model's attention converges toward a consistent structural region beyond this scale. To quantitatively evaluate this behavior, we constructed a correlation map (**Fig. 4d**), in which each grid value $D(i,j)$ represents the average normalized distance between the centers of the most important regions for any two given communication lengths l_i and l_j, ranging from 2 Å to 16 Å. Formally, we define:

$$D(i,j) = \frac{1}{n}\sum_{k=1}^{n} \frac{\left\| r_{l_i}^{(k)} - r_{l_j}^{(k)} \right\|_{PBC}}{L \cdot \sqrt{3}/2} \tag{1}$$

where $r_{l_i}^{(k)}$ and $r_{l_j}^{(k)}$ denote the centers of the most important regions (as identified by IG) in sample k for communication length l_i and l_j, respectively. The operator $\|\cdot\|_{PBC}$ denotes the Euclidean distance under PBC, and L is the side length of the cubic simulation box. The

normalization factor $L \cdot \sqrt{3}/2$ corresponds to the maximum possible point-to-point distance under PBC, ensuring comparability across all samples. The variable n denotes the total number of samples used to compute the average. Additional methodological details are provided in the Methods section.

The resulting correlation map reveals a distinct trend: while the distances between influential regions are relatively large and vary randomly at smaller communication lengths, a notable shift occurs in the upper-right quadrant of the map, corresponding to communication length ≥ 8 Å. This trend provides quantitative evidence that when the communication length exceeds 8 Å, the spatial locations of the most important regions become increasingly converging.

Together, these observations reinforce that a communication length of 8 Å, equivalent to a spherical radius of approximately 5 Å, adequately captures critical structural correlations necessary for accurate configurational energy prediction. These results echo and strengthen conclusions from the reductionist model, affirming that a 5 Å radius represents a meaningful and physically relevant scale, where the model's attention transitions from highly scattered regions whose locations are sensitive to the chosen communication length, to spatially coherent domains that more reliably contribute to the sample's global stability.

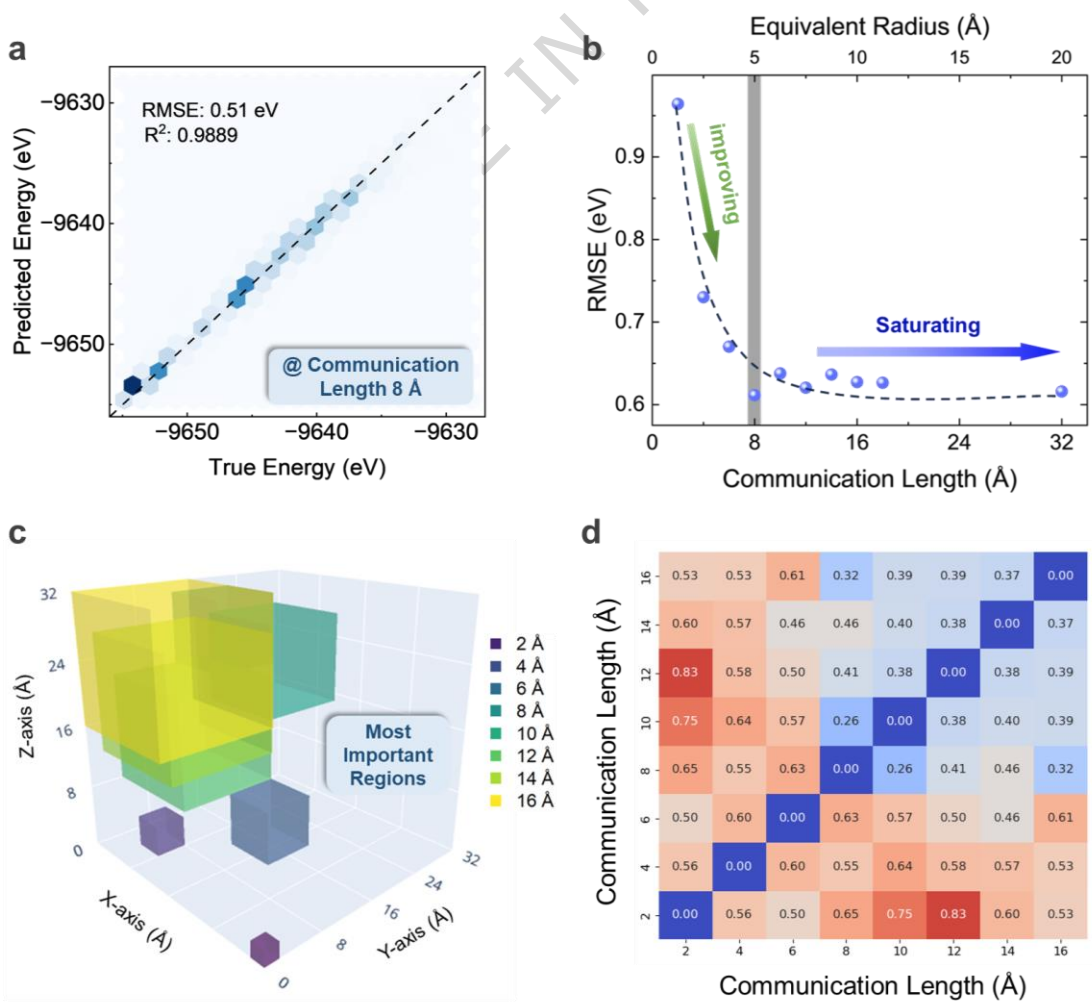

Fig. 4 | Results for the emergentist model. (a) Prediction heatmap of the emergentist model trained at communication length 8 Å. (b) Model performance as a function of the communication length, with the equivalent radius scale shown on the upper axis. (c) A representative example of the most important region for various separation sizes in the 3D simulation box. (d) Correlation map showing the normalized distances between the most important regions identified by IG across different communication lengths. Note that the correlation map is constructed by evaluating all pairwise distances between communication lengths ranging from 2 Å to 16 Å, and averaging the results across the full dataset of 10,000 samples

Discussion

A central finding of this work is that two conceptually distinct ML paradigms, namely a reductionist model based on allocating similar medium-range structural motifs and an emergentist model that learns correlations from fine-grained patch-level representations, converge on the same characteristic length scale governing informative structural environments in metallic glasses. This agreement is nontrivial, given the fundamentally different ways in which structural information and correlations are encoded in the two frameworks, and suggests that the identified RISE reflects an intrinsic physical property of the glassy structure rather than an artifact of a particular model or representation.

While it is encouraging that both the reductionist model and the emergentist model pinpoint the same RISE, the trends in their performance differ markedly and are worth discussing. Specifically, the reductionist model exhibits a clear performance peak at a 5 Å cutoff radius, whereas the emergentist model shows a saturation behavior (rather than a non-monotonic variation) beginning at the equivalently same length. Understanding such contrast provides insight into how each model processes structural information and what it reveals about the underlying physics of metallic glasses.

As aforementioned, the performance peak in the reductionist model synchronizes with the peak in ULE count, indicating that the local environments at a 5 Å cutoff offer the greatest combination of information quantity and quality. At this scale, the atomic neighborhoods are large enough to capture relevant structural diversity, yet not so extended that information becomes redundant.

Nevertheless, one might question whether this peak could arise from limitations of the SOAP encoding itself. Specifically, in our implementation, the SOAP spectrum vectors have a fixed dimensionality (see Methods for details). As a result, increasing the cutoff radius causes more atomic information to be packed into a descriptor of the same length, potentially compressing the feature space and blurring distinctions between structurally different environments. This interplay, where small cutoffs lack sufficient structural context and large cutoffs suffer from over-compression, could *in principle* produce a non-monotonic trend in both ULE count and model performance, with an artificial peak at some intermediate cutoff.

To examine such a possibility, we intentionally reduced the dimensionality of the SOAP spectrum vector by ~43% (see Methods), further increasing the level of information compression. Should the 5 Å peak were merely an artifact of descriptor resolution, one would then expect the ULE count and model performance to shift toward a smaller optimal cutoff under the more compressed encoding. However, as evidently seen in **Fig. 5a**, both the ULE count and the model's predictive

accuracy still peak robustly at 5 Å, even with the shortened descriptor. This invariance strongly supports the interpretation that the 5 Å length scale *reflects a genuine, intrinsic* structural feature of the metallic glass, a regime where atomic environments are maximally diverse yet not overwhelmed by redundant or overlapping information. In other words, the RISE emerges as a robust physical length scale, rather than a byproduct of descriptor design.

As for the emergentist model, it does not impose any fixed local definition or encoding constraint. Instead, it adaptively learns spatial correlations across scales via attention-based mechanisms. As the communication length increases, the model accesses broader structural context, leading to continuous performance gains until saturation. The observed saturation at 8 Å (or effectively a spherical radius of 5 Å) indicates that the model has captured essentially all relevant structural information needed to predict the global configurational energy, and further expansion adds only marginal or redundant context.

We also examined the consistency of the identified RISE with respect to the system size. In particular, we repeated the full analysis using a larger set of metallic glass samples consisting of 4,000 atoms, while maintaining the same atomic density and processing conditions. Both the reductionist and emergentist models were retrained on this new dataset, and their performances were assessed using the same protocols as before. As seen in **Fig. 5b-c**, despite small variations in quantitative precision, the consistent models' performance and characteristic RISE length are retained. Interestingly, while the reductionist model's RMSE per atom remains consistent across system sizes, the emergentist model exhibits a prominent improvement in RMSE per atom after saturation when trained on the larger 4000-atom system. This can be attributed to the more continuous and dense spatial representation of the atomic structure within the larger simulation box. With more atoms uniformly distributed across the same atomic density, the voxelized input contains richer spatial context and denser patch coverage, allowing the model to learn smoother and more stable patterns. Additionally, the increased sample size effectively reduces the relative contribution of boundary artifacts and sparsity in patch representations, particularly at larger communication lengths. As a result, the model benefits from enhanced structural coherence and improved learning dynamics, leading to a lower predictive error per atom. These findings hence further confirm that the observed optimal length scale is not dependent on system size or model parameters but instead reflects the true physics feature of the Zr-Cu metallic glasses.

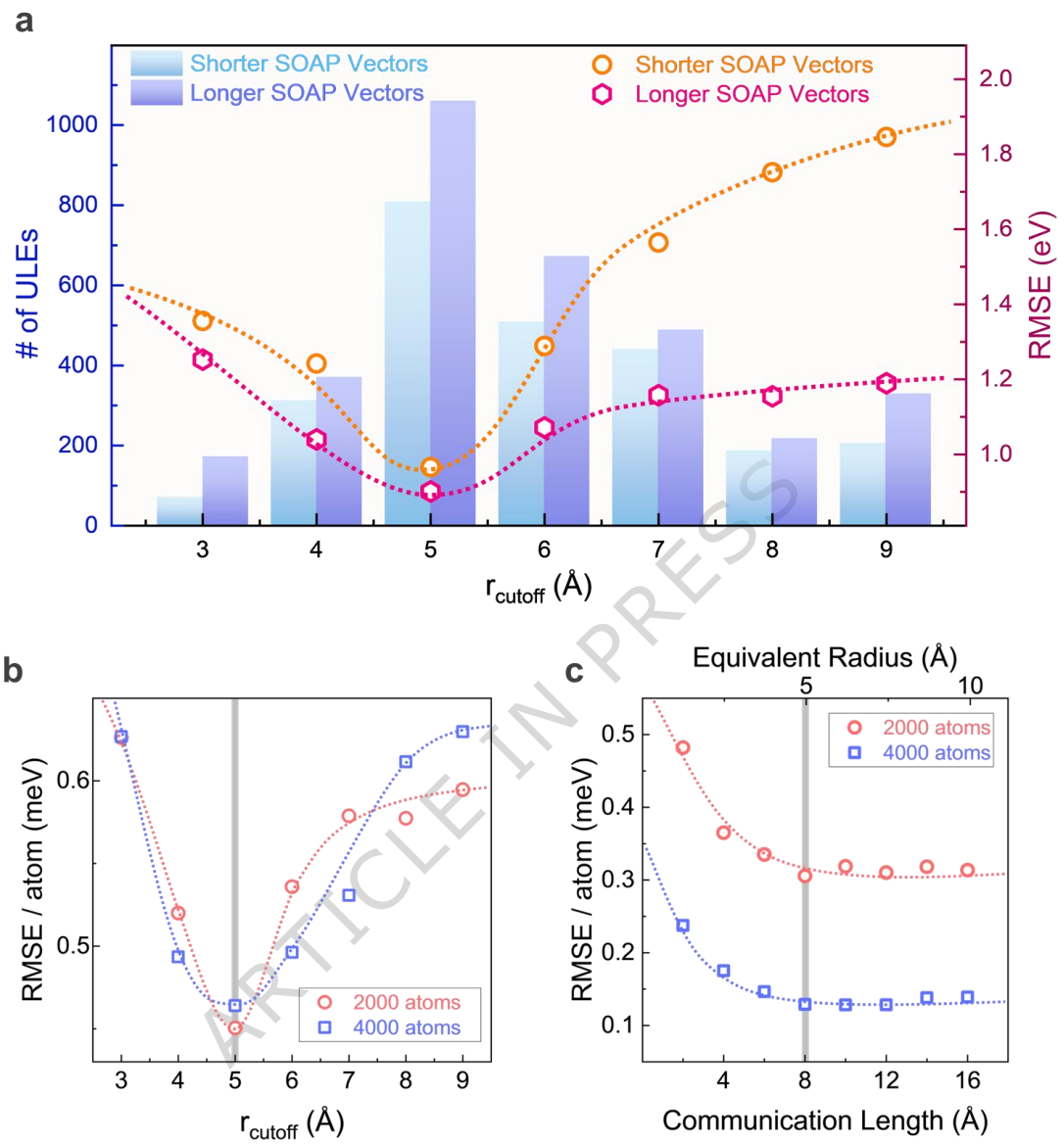

Fig. 5 | Result consistency across encoding resolution and system sizes. (a) Number of ULEs (left axis) and reductionist model accuracy (right axis) as functions of SOAP cutoff radius, using SOAP spectrum vectors with different lengths. (b) Reductionist model performance (RMSE per atom) on 2000- and 4000-atom systems as a function of SOAP cutoff radius. (c) Emergentist model performance (RMSE per atom) on 2000- and 4000-atom systems as a function of communication length.

To further examine the universality of the identified RISE across different alloy chemistries, we first extended our analysis to two ternary Zr-Cu-Al metallic glass systems, $Zr_{40}Cu_{50}Al_{10}$, and $Zr_{65}Cu_{25}Al_{10}$, following the same dual machine-learning procedures used for the binary Zr–Cu system. As shown in **Fig. 6a-b**, for both compositions, the reductionist model achieves optimal predictive performance at a 5 Å cutoff radius, while the emergentist model saturates beyond an 8 Å communication length (or an equivalent radius of 5 Å). This consistent convergence demonstrates that the identified RISE continues to exist in more compositionally complex alloys. In addition, we further examined a chemically very distinct binary glass model – $Pd_{80}Si_{20}$ with

intermetallic bonding. Markedly, both ML frameworks again yield consistent results (demonstrated in **Fig. 6c-d**), with the reductionist model peaking at a 4 Å cutoff and the emergentist model saturating beyond a 6 Å communication length (or an equivalent radius of 3.7 Å). This smaller RISE than that in Zr-Cu based metallic glasses is not surprising and consistent with the well-documented structural features of Pd–Si glasses, which exhibit pronounced short-range and chemical ordering characterized by dense Pd-centered coordination polyhedra and less obvious medium-range network among Si atoms[51,52]. As confirmed by the radial distribution function (RDF) comparison between $Pd_{80}Si_{20}$ and Zr-Cu-Al (**Fig. 6e**), $Pd_{80}Si_{20}$ displays a sharper first peak and a faster decay of correlations beyond 4 Å, indicating a more localized structural correlation length. This reflects the highly efficient local packing and reduced topological frustration in Pd-Si, such that the structural information relevant to configurational energy prediction saturates at a shorter range than in Zr-Cu-based glasses. The consistency between the two ML approaches across these systems reinforces that the RISE represents a genuine, physically meaningful structural length scale in metallic glasses. Although beyond the main scope of the present study, we have extended a similar analysis to covalent glassy systems, including amorphous carbon and amorphous silicon[53,54] (see Supplementary Material Note 3). In these cases, the dual ML approaches once again converge on a single, consistent characteristic length scale. These findings are consistent with recent 4D-STEM experimental observations, which show much lower degree of MRO in covalent-bonded materials[55,56]. Such alignment is encouraging, as it lends credence to the notion that our dual ML framework not only can identify a consistent intrinsic RISE length in amorphous materials, but also can reveal its dependence on the atomic bonding features.

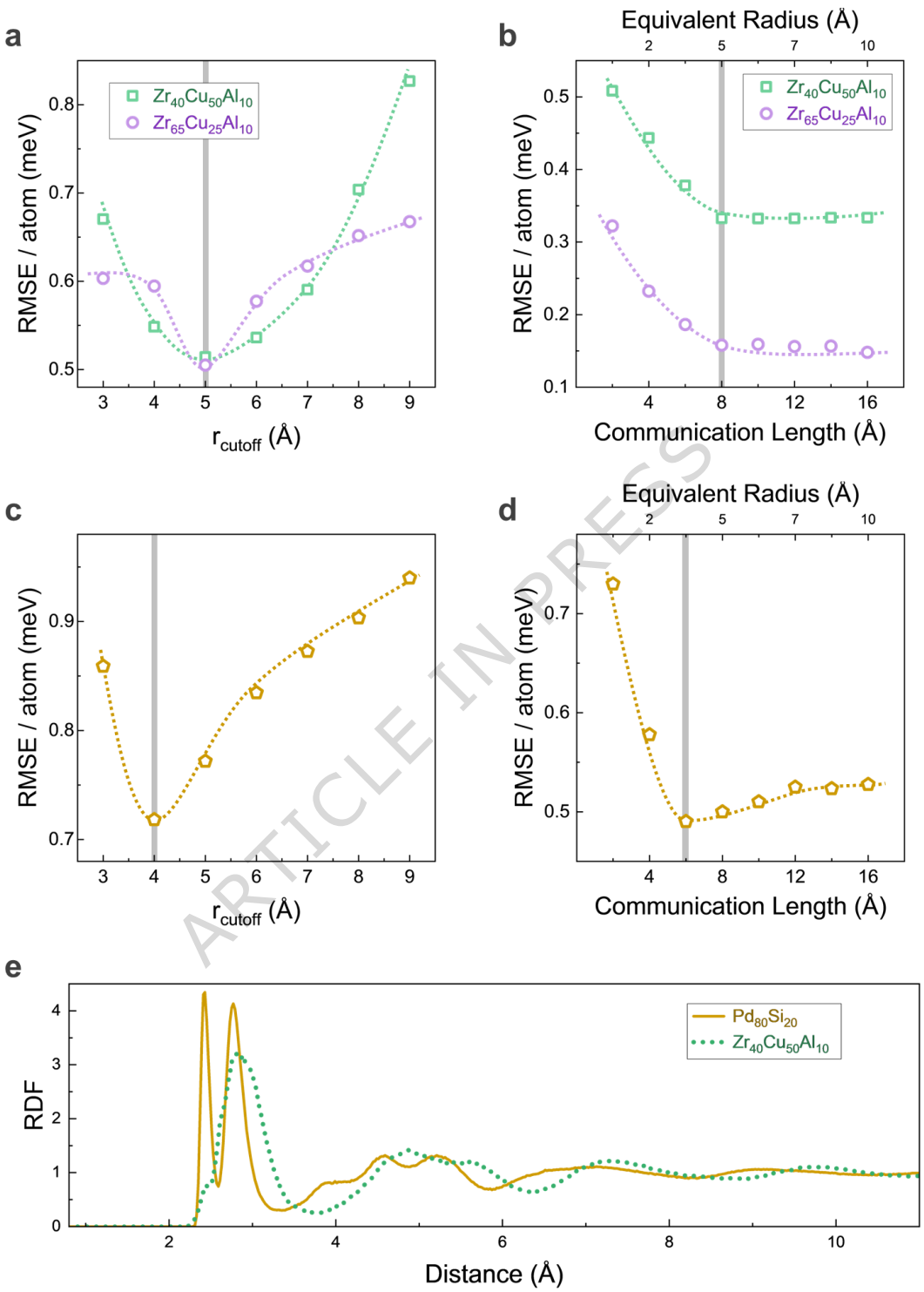

Fig. 6 | Model performance across different metallic glass systems. (a) Prediction accuracy (RMSE) of the reductionist model for Zr-Cu-Al systems as a function of the SOAP cutoff radius. (b) Prediction accuracy (RMSE) of the emergentist model for Zr-Cu-Al systems as a function of the communication length. (c) Prediction accuracy (RMSE) of the reductionist model for $Pd_{80}Si_{20}$ as a function of the SOAP cutoff radius. (d) Prediction accuracy (RMSE) of the emergentist model for $Pd_{80}Si_{20}$ as a function of the communication length. (e) RDF comparison between $Pd_{80}Si_{20}$ and $Zr_{40}Cu_{50}Al_{10}$.

We further note that the robustness of the identified characteristic length scale is not limited by the size effect of the simulation box. In all systems studied, the extracted RISE is significantly smaller than the simulation box dimensions, ensuring that it is not artificially constrained by PBC. More importantly, the reliability of the RISE identification relies on sufficient statistical sampling of disordered configurations. The datasets used in this work contain 10,000 distinct samples, providing a large and diverse sampling pool of local environments. This extensive sampling ensures convergence of model performance trends and supports the statistical stability of the extracted length scale across different metallic glass systems.

Markedly, the particular length scale in Zr-Cu based metallic glass (*i.e.* roughly 2nd nearest neighbor distance) aligns with multiple prior studies across experimental and computational domains. For instance, structural comparisons between atoms with contrasting local flexibility have further revealed pronounced differences in atomic environments at this shell[18], suggesting it as a key regime for property-relevant variation. In addition, spatially resolved analyses of RDFs have shown the emergence of ordered motifs, such as Bergman-type clusters, centered around this length scale[31]. Also, independent measurements of structural coherence suggest a rapid decay beyond this shell, positioning this length scale as the upper bound of correlated local order[57]. As another example, the spatial connectivity of interpenetrating icosahedral clusters - associated with dynamic slowdown and energetic stabilization near the glass transition - has been observed to emerge within this radial regime[58]. Very recently, particle-level dynamical defect units have also been identified based on quasi-localized vibrational modes, revealing mechanically active structures characterized by a comparable spatial extent[59]. Although these defect units are defined from a dynamical perspective rather than from static structural correlations, their characteristic size further supports the notion that this radial regime represents a fundamental length scale governing medium-range organization and physical response in metallic glasses. Collectively, these observations substantiate our conclusion that the 2nd nearest neighbor distance defines the onset of medium-range structural organization in metallic glasses and encodes the critical motifs governing their behaviors.

Identifying the intrinsic characteristic length scale that can offer more effective information beyond SROs provides a physically grounded framework, not only for interpreting the structural origin of key properties in metallic glasses, but also for building capable meso-scale models with the most effective unit/voxel size. With this length scale now clarified, a natural next step is to resolve the more detailed topological characteristics of the motifs that populate this regime. In addition, the question of whether these motifs exhibit universal symmetries, alloy-specific variations, or dynamic heterogeneity will be essential, since it not only could deepen our understanding of disorder in metallic glasses but may also inform the predictive design of glass-forming alloys with tailored mechanical and thermal properties. Admittedly, probing specific key atomic motifs within this length scale is beyond the scope of the present study. Nevertheless, such information could be accessed by analyzing patch-wise correlations within attention regions of the emergentist model, or by employing complementary advanced topological analysis techniques, such as persistent homology[60–64], which would warrant future investigation.

Before concluding, it is worth acknowledging that our current work primarily focuses on predicting the global configurational energy, justified by its clear definition, computational accessibility, and experimentally validated correlations with several key mechanical properties

such as strength and ductility. With that being said, admittedly, some other critical behaviors concerning spatial heterogeneities imposed by extrinsic factors (*e.g.*, mechanical loading-induced shear banding) cannot be directly predicted by our present model in its current form. Those complex behaviors involve hierarchies spanning multiple length scales and capturing them would likely require more intricate and extended atomic environments. Nevertheless, the methodologies and architectures developed here provide a versatile framework that could be extended to investigate location-specific phenomena. For example, the emergentist model framework can be adapted to examine local structural correlations by incorporating a modified mask transformer decoder[65], enabling the direct prediction of localized properties at the patch level. Such an approach would facilitate systematic exploration of the characteristic length scales at which structural information most effectively predicts localized behaviors, such as shear band formation. It warrants further investigation in the future, which would expectedly shed light on better bridging the gap between atomic-level structures and spatially heterogeneous mechanical behaviors in metallic glasses.

Methods

Metallic Glass sample preparation. The 3D $Zr_{44}Cu_{56}$ metallic glasses are simulated using molecular dynamics (MD) with the LAMMPS package and a set of embedded-atom method[66] (EAM) potentials, which employs a cutoff distance of 6.5 Å, such that atomic interactions beyond 5 Å are explicitly included in the simulations. Each simulation box contained 2000 atoms. The system is first equilibrated at 2000 K, followed by a quench to 0 K at varying cooling rates ranging from 10^{10} to 10^{13} K/s in the NVT ensemble using a Nose-Hoover thermostat. To capture a broader range of structural stability, an additional annealing step is applied: samples are heated to 500-750 K, below the glass transition temperature, and held at each target temperature for sufficient time to allow structural relaxation. Such rate-controlled cooling and sub-T_g annealing treatment can together generate glassy samples spanning a wide variety of energy scales (see Supplemental Material Note 6), equivalent to around 8 orders of magnitude effective cooling rates[67]. PBC is applied in all three directions. The final structures are relaxed to mechanical equilibrium using a conjugate gradient algorithm. A total of 10,000 metallic glass samples is generated, covering a wide spectrum of configurational energies. These samples are then randomly divided into training, validation, and test sets in a 3:1:1 ratio for subsequent ML tasks. Specifically, the training and validation sets are used for model fitting and hyperparameter tuning, while all RMSE values reported throughout the manuscript are computed on the independent test set.

SOAP descriptor encoding. The SOAP descriptor used in the reductionist model encodes local atomic environments by expanding a Gaussian-smeared atomic density onto a set of orthonormal basis functions composed of spherical harmonics and radial basis functions. Specifically, for each atomic species Z, the local atomic density is defined as:

$$\rho^{z}(r) = \sum_{i \in Z} \exp\left(-\frac{|r - R_i|^2}{2\sigma^2}\right) \tag{2}$$

where R_i denotes the position of atom i and σ is the Gaussian smearing width. This density is expanded as:

$$c^{Z}_{nlm} = \int g_n(r) Y_{lm}(\theta, \phi) \rho^{Z}(r) dr \tag{3}$$

where $g_n(r)$ are radial basis functions and $Y_{lm}(\theta, \phi)$ are real spherical harmonics. The rotationally invariant SOAP power spectrum is then constructed as:

$$p^{Z_1 Z_2}_{nn'l} = \pi \sqrt{\frac{8}{2l+1}} \sum_{m} c^{Z_1}_{nlm} c^{Z_2}_{n'lm} \tag{4}$$

yielding a partial power spectrum vector that characterizes the local environment of each atom.

For a sample containing 2000 atoms, the encoding procedure yields 2000 individual SOAP spectrum vectors. In the standard model, the descriptor is configured with 11 (n_{max}) radial basis functions and a maximum spherical harmonics degree (l_{max}) of 12, using a Gaussian smearing width of 0.5. This results in a SOAP spectrum vector of 3289 dimensions. The reduced model employs 9 (n_{max}) radial basis functions and a spherical harmonics degree (l_{max}) of 10, yielding a shortened vector of 1881 dimensions, approximately 57% of the original length.

ULE identification and ULE frequency vector construction. ULEs are identified from the reductionist model by clustering normalized SOAP spectrum vectors based on a predefined similarity threshold. For each atom, a normalized SOAP spectrum vector is computed, and these vectors are sequentially compared against a growing set of representative ULEs. A new atomic environment is classified as a distinct ULE if its similarity score with all existing ULEs fell below the specified threshold. The similarity score is defined as:

$$M_S = 1 - \sqrt{\frac{\left\| p_i - p_j \right\|^2}{2}} \tag{5}$$

where p_i and p_j denote two SOAP normalized spectrum vectors. A threshold of 0.83 is used for the results presented in the main text. It is worth noting that we also examined other similarity threshold values to test the sensitivity of the model's performance (see Supplemental Material Note 4). It is confirmed that the performance peak remains at the same SOAP cutoff radius.

Once the global ULE set is established, each atomic environment in every sample is assigned to its closest matching ULE based on the same similarity metric. A ULE frequency vector is then constructed for each sample, where each element represents the fraction of atoms assigned to a given ULE. This vector serves as a compact structural fingerprint, summarizing the distribution of distinct local motifs while abstracting away individual atomic positions and identities.

Configurational entropy calculation. The configurational entropy (S_2), a widely used metric for quantifying local structural order, is computed using LAMMPS. S_2 is defined as:

$$S_2 = -2\pi\rho k_B \int_0^\infty [g(r)\ln g(r) - g(r) + 1]r^2 dr \quad (6)$$

where ρ is the system density, g(r) is the radial distribution function, and k_B is the Boltzmann constant. A Gaussian smoothing width of 0.3 Å is applied. The local density around each atom is used to normalize the radial distribution function, enabling accurate characterization of structural ordering on the atomic scale. To evaluate the entropy associated with each ULE, we first calculated the S_2 value for every atom, then computed the average S_2 within each ULE by summing the individual atomic values and dividing by the number of atoms assigned to that ULE. Distributions of these average S_2 values were then analyzed across the full set of ULEs obtained under each SOAP cutoff radius. This allows us to assess the diversity of structural order captured at different length scales. Note that to numerically compute S_2, an upper limit for the integral is needed. For the results presented in the main text, the upper limit of 7.0 Å was employed. We also examined other upper limit values (see Supplemental Material Note 5), and they do not affect the results in **Fig. 2c**.

Vit module specification. The atomic structure input is divided into non-overlapping patches, which are embedded into a lower-dimensional patch sequence using a 3D convolutional layer with a kernel size and stride equal to the patch size. A CLS token is prepended to the sequence to aggregate global structural information. Positional embeddings are initiated randomly and are added to the patch sequence to preserve spatial information. The resulting sequence is passed through a dropout layer (p = 0.1) before being input to the transformer encoder. Based on the defined communication length, the 3D patch grid (16×16×16) is partitioned into cubic subregions (e.g., 2×2×2 or 4×4×4 patches). For each region, we define a token group that includes the CLS token and all patch tokens within the subregion. Attention is computed only within each region using a standard multi-head self-attention (MHSA) mechanism in the TransformerEncoderBlock layer, followed by a feedforward network. Specifically, for a given token group with feature matrix $X \in \mathbb{R}^{N\times d}$, the MHSA operation is implemented through the following matrix formulation:

$$Attention(X) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (7)$$

where $Q = XW_Q$, $K = XW_K$, and $V = XW_V$ are the query, key, and value matrices, respectively, and W_Q, W_K, and W_V are learned projection matrices. This operation enables adaptive aggregation of patch-level features within each communication region.

Two TransformerEncoderBlock layers are applied to each region independently. Each block includes LayerNorm, MHSA, and a two-layer MLP with ReLU activation and residual connections. During evaluation, dropout is disabled. After processing all regional segments, the updated patch tokens are concatenated back into a global sequence. The CLS token is extracted, passed through a final LayerNorm layer, and fed into a fully connected regression head to predict the sample's configurational energy. The model is trained using mean squared error (MSE) loss with the Adam optimizer (learning rate 1×10^{-3}), and a StepLR scheduler (step size 20, decay factor 0.1) over 40 epochs with a batch size of 32. Model performance is evaluated across multiple communication lengths to probe structure-property relationships at varying spatial scales.

Data availability

The data supporting the findings of this study is available at https://zenodo.org/records/18066041.

Code availability

The code for the reductionist and emergentist models, including model training, is available at https://github.com/muchen1453/RISE.

Acknowledgements

The authors acknowledge the supports of NSF-DMR-2406530 (Minhazul Islam, Yuchi Wang, Yunzhi Wang, Jinwoo Hwang) and NSF-DMR-2406531 (Muchen Wang, Yuchu Wang, Yue Fan).

Author contributions

Muchen Wang: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation; Yuchu Wang: Methodology, Formal analysis, Data curation; Minhazul Islam: Writing – review & editing; Yuchi Wang: Writing – review & editing; Yunzhi Wang: Writing – review & editing, Writing – original draft, Resources, Conceptualization; Jinwoo Hwang: Writing – review & editing, Writing – original draft, Resources, Conceptualization; Yue Fan: Writing – review & editing, Writing – original draft, Supervision, Resources, Investigation, Conceptualization.

Competing Interests

The authors declare no competing financial or non-financial interests.

References

1. Taylor, G. I. The mechanism of plastic deformation of crystals. *Proceedings of the Royal Society of London A* 145, 362–387 (1934).
2. Hall, E. O. The Deformation and Ageing of Mild Steel: III Discussion of Results. 747–753 (1951).

3. Petch, N. J. The Cleavage Strength of Polycrystals. *The Journal of the Iron and Steel Institute* 174, 25–31 (1953).
4. Ding, J. *et al.* Universal structural parameter to quantitatively predict metallic glass properties. *Nat Commun* 7, 1–10 (2016).
5. Fan, Z. & Ma, E. Predicting orientation-dependent plastic susceptibility from static structure in amorphous solids via deep learning. *Nat Commun* 12, 1–13 (2021).
6. Ding, J. & Ma, E. Computational modeling sheds light on structural evolution in metallic glasses and supercooled liquids. *NPJ Comput Mater* 3, 1–12 (2017).
7. Wang, Q., Zhang, L. F., Zhou, Z. Y. & Yu, H. Bin. Predicting the pathways of string-like motions in metallic glasses via path-featurizing graph neural networks. *Sci Adv* 10, 2799 (2024).
8. Ren, F. *et al.* Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. *Sci Adv* 4, eaaq1566 (2018).
9. Teich, E. G., Galloway, K. L., Arratia, P. E. & Bassett, D. S. Crystalline shielding mitigates structural rearrangement and localizes memory in jammed systems under oscillatory shear. *Sci Adv* 7, 3392–3404 (2021).
10. Cheng, Y. Q. & Ma, E. Atomic-level structure and structure–property relationship in metallic glasses. *Prog Mater Sci* 56, 379–473 (2011).
11. Fan, Y., Iwashita, T. & Egami, T. Energy landscape-driven non-equilibrium evolution of inherent structure in disordered material. *Nat Commun* 8, 1–7 (2017).
12. Liu, C., Guan, P. & Fan, Y. Correlating defects density in metallic glasses with the distribution of inherent structures in potential energy landscape. *Acta Mater* 161, 295–301 (2018).
13. Ma, E. Tuning order in disorder. *Nat Mater* 14, 547–552 (2015).
14. Sheng, H. W., Luo, W. K., Alamgir, F. M., Bai, J. M. & Ma, E. Atomic packing and short-to-medium-range order in metallic glasses. *Nature* 439, 419–425 (2006).
15. Wu, Y. *et al.* Substantially enhanced plasticity of bulk metallic glasses by densifying local atomic packing. *Nat Commun* 12, 1–9 (2021).
16. Cao, Y. *et al.* Continuous polyamorphic transition in high-entropy metallic glass. *Nat Commun* 15, 1–9 (2024).
17. Maldonis, J. J., Banadaki, A. D., Patala, S. & Voyles, P. M. Short-range order structure motifs learned from an atomistic model of a Zr50Cu45Al5 metallic glass. *Acta Mater* 175, 35–45 (2019).
18. Fan, Z., Ding, J. & Ma, E. Machine learning bridges local static structure with multiple properties in metallic glasses. *Materials Today* 40, 48–62 (2020).
19. Liu, C. *et al.* Concurrent prediction of metallic glasses' global energy and internal structural heterogeneity by interpretable machine learning. *Acta Mater* 259, 119281 (2023).
20. Wang, Q. & Jain, A. A transferable machine-learning framework linking interstice distribution and plastic heterogeneity in metallic glasses. *Nat Commun* 10, 1–11 (2019).
21. Luo, S., Khong, J. C., Huang, S., Yang, G. & Mi, J. Revealing in situ stress-induced short- and medium-range atomic structure evolution in a multicomponent metallic glassy alloy. *Acta Mater* 272, 119917 (2024).
22. Zhao, P., Li, J., Hwang, J. & Wang, Y. Influence of nanoscale structural heterogeneity on shear banding in metallic glasses. *Acta Mater* 134, 104–115 (2017).
23. Miyazaki, N., Wakeda, M., Wang, Y. J. & Ogata, S. Prediction of pressure-promoted thermal rejuvenation in metallic glasses. *NPJ Comput Mater* 2, 1–9 (2016).
24. Wang, Q. *et al.* Predicting the propensity for thermally activated β events in metallic glasses via interpretable machine learning. *NPJ Comput Mater* 6, 1–12 (2020).
25. Wang, Q. & Zhang, L. Inverse design of glass structure with deep graph neural networks. *Nat Commun* 12, 1–11 (2021).
26. Hilke, S. *et al.* The influence of deformation on the medium-range order of a Zr-based bulk metallic glass characterized by variable resolution fluctuation electron microscopy. *Acta Mater* 171, 275–281 (2019).
27. Nomoto, K. *et al.* Medium-range order dictates local hardness in bulk metallic glasses. *Materials Today*

44, 48–57 (2021).
28. Lee, M., Lee, C. M., Lee, K. R., Ma, E. & Lee, J. C. Networked interpenetrating connections of icosahedra: Effects on shear transformations in metallic glass. *Acta Mater* 59, 159–170 (2011).
29. Wang, J. *et al.* Clustering-mediated enhancement of glass-forming ability and plasticity in oxygen-minor-alloyed Zr-Cu metallic glasses. *Acta Mater* 261, 119386 (2023).
30. Tang, L. *et al.* Short- and medium-range orders in Al90Tb10 glass and their relation to the structures of competing crystalline phases. *Acta Mater* 204, 116513 (2021).
31. Fang, X. W. *et al.* Spatially resolved distribution function and the medium-range order in metallic liquid and glass. *Sci Rep* 1, 1–5 (2011).
32. Laws, K. J., Miracle, D. B. & Ferry, M. A predictive structural model for bulk metallic glasses. *Nat Commun* 6, 1–10 (2015).
33. Wang, Q., Li, J. H., Liu, J. B. & Liu, B. X. Atomistic Design of Favored Compositions for Synthesizing the Al-Ni-Y Metallic Glasses. *Sci Rep* 5, 1–13 (2015).
34. Egami, T. & Ryu, C. W. Origin of medium-range atomic correlation in simple liquids: Density wave theory. *AIP Adv* 13, 85308 (2023).
35. Ryu, C. W., Dmowski, W. & Egami, T. Ideality of liquid structure: A case study for metallic alloy liquids. *Phys Rev E* 101, 030601 (2020).
36. Im, S. *et al.* Direct determination of structural heterogeneity in metallic glasses using four-dimensional scanning transmission electron microscopy. *Ultramicroscopy* 195, 189–193 (2018).
37. Deng, J. W., Du, K. & Sui, M. L. Medium range order of bulk metallic glasses determined by variable resolution fluctuation electron microscopy. *Micron* 43, 827–831 (2012).
38. Hwang, J. & Voyles, P. M. Variable Resolution Fluctuation Electron Microscopy on Cu-Zr Metallic Glass Using a Wide Range of Coherent STEM Probe Size. *Microscopy and Microanalysis* 17, 67–74 (2011).
39. Cubuk, E. D. *et al.* Identifying structural flow defects in disordered solids using machine-learning methods. *Phys Rev Lett* 114, 108001 (2015).
40. Li, B. *et al.* Superior mechanical properties of a Zr-based bulk metallic glass via laser powder bed fusion process control. *Acta Mater* 266, 119685 (2024).
41. Costa, M. B. & Greer, A. L. Enthalpy of anelasticity and rejuvenation of metallic glasses. *Acta Mater* 265, 119609 (2024).
42. Wang, W. H. The elastic properties, elastic models and elastic perspectives of metallic glasses. *Prog Mater Sci* 57, 487–656 (2012).
43. Wang, W. H., Dong, C. & Shek, C. H. Bulk metallic glasses. *Materials Science and Engineering R: Reports* 44, 45–89 (2004).
44. Greer, A. L., Cheng, Y. Q. & Ma, E. Shear bands in metallic glasses. *Materials Science and Engineering: R: Reports* 74, 71–132 (2013).
45. De, S., Bartók, A. P., Csányi, G. & Ceriotti, M. Comparing molecules and solids across structural and alchemical space. *Physical Chemistry Chemical Physics* 18, 13754–13769 (2016).
46. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 13-17-August-2016, 785–794 (2016).
47. Dosovitskiy, A. *et al.* An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *ICLR 2021 - 9th International Conference on Learning Representations* https://arxiv.org/pdf/2010.11929 (2020).
48. Piaggi, P. M. & Parrinello, M. Entropy based fingerprint for local crystalline order. *J Chem Phys* 147, 114112 (2017).
49. Lundberg, S. M. & Lee, S. I. A Unified Approach to Interpreting Model Predictions. *Adv Neural Inf Process Syst* 2017-December, 4766–4775 (2017).
50. Sundararajan, M., Taly, A. & Yan, Q. Axiomatic Attribution for Deep Networks. in *Proceedings of the 34th International Conference on Machine Learning* (2017).
51. Durandurdu, M. Ab initio modeling of metallic Pd80Si20 glass. *Comput Mater Sci* 65, 44–47 (2012).
52. Yue, G. Q. *et al.* Local structure order in Pd78Cu6Si16 liquid. *Sci Rep* 5, 1–6 (2015).

53. Deringer, V. L. & Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys Rev B* 95, 094203 (2017).
54. Deringer, V. L. *et al.* Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics. *J Phys Chem Lett* 9, 2879–2885 (2018).
55. Abbasi, M. *et al.* In situ observation of medium range ordering and crystallization of amorphous TiO2 ultrathin films grown by atomic layer deposition. *APL Mater* 11, (2023).
56. Meng, J. *et al.* Experimentally informed structure optimization of amorphous TiO2 films grown by atomic layer deposition. *Nanoscale* 15, 718–729 (2023).
57. Ma, D., Stoica, A. D. & Wang, X. L. Power-law scaling and fractal nature of medium-range order in metallic glasses. *Nat Mater* 8, 30–34 (2009).
58. Wu, Z. W. *et al.* Critical scaling of icosahedral medium-range order in CuZr metallic glass-forming liquids. *Sci Rep* 6, 1–7 (2016).
59. Hu, Y. C. & Tanaka, H. Unveiling hidden particle-level defects in glasses. *Nature Communications 2025 16:1* 16, 5321- (2025).
60. Edelsbrunner, H., Letscher, D. & Zomorodian, A. Topological persistence and simplification. *Discrete Comput Geom* 28, 511–533 (2002).
61. Hiraoka, Y. *et al.* Hierarchical structures of amorphous solids characterized by persistent homology. *Proceedings of the National Academy of Sciences* 113, 7035–7040 (2016).
62. Liu, S. *et al.* Turing pattern and chemical medium-range order of metallic glasses. *Materials Today Physics* 38, 101254 (2023).
63. Zheng, S. *et al.* Active phase discovery in heterogeneous catalysis via topology-guided sampling and machine learning. *Nature Communications* 16, 1–13 (2025).
64. Townsend, J., Micucci, C. P., Hymel, J. H., Maroulas, V. & Vogiatzis, K. D. Representation of molecular structures with persistent homology for machine learning applications in chemistry. *Nat Commun* 11, 1–9 (2020).
65. Strudel, R., Garcia, R., Laptev, I. & Schmid, C. Segmenter: Transformer for Semantic Segmentation. *Proceedings of the IEEE International Conference on Computer Vision* 7242–7252 (2021) doi:10.1109/ICCV48922.2021.00717.
66. Cheng, Y. Q., Ma, E. & Sheng, H. W. Atomic Level Structure in Multicomponent Bulk Metallic Glass. *Phys Rev Lett* 102, 245501 (2009).
67. Liu, C. & Fan, Y. Emergent Fractal Energy Landscape as the Origin of Stress-Accelerated Dynamics in Amorphous Solids. *Phys Rev Lett* 127, 215502 (2021).

Figure Titles and Legends

Fig. 1 | Architectures of the models. (a) Reductionist model architecture. Each sample is first processed using the SOAP descriptor, which encodes local atomic environments as partial power spectrum vectors. By systematically varying the SOAP cutoff radius (3.0-9.0 Å), the model captures structural features across multiple length scales. Similarity grouping is then performed to group local environments into ULEs, reducing redundancy while preserving diversity. These ULEs are used to construct ULE frequency spectrum that represent each sample's composition. An XGBoost model, trained on these frequency vectors, predicts the configurational energy of each sample. (b) Emergentist model architecture (visualized in 2D for clarity). Each sample is voxelized into a pair of binary 3D grids, with separate channels representing Cu and Zr atomic species. The simulation box is discretized into voxels, which are then partitioned into non-overlapping patches to enhance spatial abstraction and computational efficiency. To explore multiscale structural correlations, we define regions of patch communication based on a tunable communication length, ranging from 2 to 32 Å. Within each defined region, patch embeddings are generated through 3D convolutions and processed by a ViT module. The ViT module integrates patch information via multi-head self-attention calculations, with a CLS

token encoding the global representation of the structure. The output from the transformer is passed through a multi-layer perceptron to predict the sample's configurational energy.

Fig. 2 | Reductionist model performance and peak of information density. (a) Prediction heatmap of the reductionist model trained using SOAP descriptors at 5 Å cutoff radius. (b) Relationship between the cutoff radius, number of ULEs, and model performance. Error bars represent the variation obtained from five independently seeded ULE searches. More details in Supplemental Material Note 1. (c) The distribution of average S_2 for ULEs at each cutoff radius, and the stars mark the upper and lower bounds of S_2 at each cutoff.

Fig. 3 | Performance landscape of the reductionist model. RMSE as a function of SOAP cutoff radius and the fraction of top-ranked ULEs retained for prediction. The subpanels display the prediction error distributions for models at different cutoffs trained using the top 8% and the full 100% of ULEs.

Fig. 4 | Results for the emergentist model. (a) Prediction heatmap of the emergentist model trained at communication length 8 Å. (b) Model performance as a function of the communication length, with the equivalent radius scale shown on the upper axis. (c) A representative example of the most important region for various separation sizes in the 3D simulation box. (d) Correlation map showing the normalized distances between the most important regions identified by IG across different communication lengths. Note that the correlation map is constructed by evaluating all pairwise distances between communication lengths ranging from 2 Å to 16 Å, and averaging the results across the full dataset of 10,000 samples

Fig. 5 | Result consistency across encoding resolution and system sizes. (a) Number of ULEs (left axis) and reductionist model accuracy (right axis) as functions of SOAP cutoff radius, using SOAP spectrum vectors with different lengths. (b) Reductionist model performance (RMSE per atom) on 2000- and 4000-atom systems as a function of SOAP cutoff radius. (c) Emergentist model performance (RMSE per atom) on 2000- and 4000-atom systems as a function of communication length.

Fig. 6 | Model performance across different metallic glass systems. (a) Prediction accuracy (RMSE) of the reductionist model for Zr-Cu-Al systems as a function of the SOAP cutoff radius. (b) Prediction accuracy (RMSE) of the emergentist model for Zr-Cu-Al systems as a function of the communication length. (c) Prediction accuracy (RMSE) of the reductionist model for $Pd_{80}Si_{20}$ as a function of the SOAP cutoff radius. (d) Prediction accuracy (RMSE) of the emergentist model for $Pd_{80}Si_{20}$ as a function of the communication length. (e) RDF comparison between $Pd_{80}Si_{20}$ and $Zr_{40}Cu_{50}Al_{10}$.