Geometry-enhanced graph neural network for glassy dynamics prediction

Xiao Jiang,¹ Zean Tian,¹ and Kenli Li¹

¹College of Computer Science and Electronic Engineering, Hunan University, Changsha, China

(Dated: November 24, 2022)

Abstract

Understanding the dynamic processes of the glassy system continues to be challenging. Recent advances have shown the power of graph neural networks (GNNs) for determining the correlation between structure and dynamics in the glassy system. These methods treat the glassy system as a topological graph. However, the inherent "smoothness" property of the dynamics on the graph (variation of dynamics over the graph) is ignored, resulting in a deteriorated performance of GNN for dynamic predictions over various time scales. In this paper, we first present an experimental investigation assessing the smoothness patterns of particle dynamics from the graph perspective. The results indicate that the long-time dynamics exhibit a smoother property on the graph, while the short-time dynamics reveal a distinctly non-smooth pattern. To establish the relationship between the static structure and dynamics with different smoothness patterns, we propose a novel geometry-enhanced graph neural network (Geo-GNN) model. The Geo-GNN architecture consists of a geometry feature encoder that incorporates rotation-invariant distances and angles to enhance geometric feature learning and a geometry-enhanced aggregation block that can adaptively learn the underlying smoothness patterns during message passing. Experimental results demonstrate that our method is able to capture the inherent smoothness pattern of particles and outperforms state-of-the-art baselines in predicting the dynamics across all time scales. A subsequent study has revealed that geometric features are critical for the model to accurately capture smoothness patterns in dynamics. Our research not only refines the method for predicting glassy dynamics but also provides a new lens through which to investigate the issue of causality regarding dynamical heterogeneity in glasses.
INTRODUCTION

In the physical sciences, glass is a prototypical example of non-equilibrium matter because it forms when a fluid is quenched or densified until it ceases to flow [1–6]. In contrast to more ordinary condensed matter, this solidification shown in the glass transition occurs without any discernible structural order, as seen in conventional correlation functions [5]. However, with just a minor temperature variation, the viscosity and structure relaxation time may vary by many orders of magnitude. Such glass translation phenomena can be observed in various systems with dynamic correlation sizes ranging from microscopic to macroscopic [7], but its mechanism remains a mystery in condensed matter physics.

One of the hallmarks of the glass transition is dynamic heterogeneity, where particles in glassy supercooled liquids are clustered together in the form of cooperatively rearranging regions that move much faster or slower than the average [8]. The concept of dynamic heterogeneity has been at the forefront of research on glass and non-equilibrium systems for the past two decades. Numerous attempts have been made to establish a structural signature of dynamics in supercooled liquids. Due to the lack of universal structural order, traditional approaches [9–15] provide correlations between structure and dynamics that are inadequate for constructing a theory. Recent advances in applying machine learning (ML) to scientific research open a new avenue for identifying the structural predictors associated with dynamic heterogeneity [16]. The first application of support vector machines (SVMs) to the dynamics prediction of glassy systems discovered that machine-learned softness could differentiate between slow and fast regions in liquid glass[17–19]. After that, various techniques have been proposed, from supervised approaches like linear regression and graph neural network (GNN) [20–23] to unsupervised methods like community detection and autoencoder [24–26] multiple glassy systems. In particular, GNN has achieved unprecedented accuracy in predicting the dynamic propensity by only using the static graph structure of a glassy system as input [20].

Although machine learning promises to reveal the relationship between structure and dynamics, several questions remain unresolved, such as why the association between predictions and ground-truth dynamic propensity is limited on a short timescale, particularly in the caging regime, and what missing information could help us predict the dynamic propensity at both long and short time scales? Recent works attempt to clarify the above-mentioned
concerns by modifying either the dynamic prediction objective [23] or the model architecture [21, 27]. Work in Ref. [23] has found that adding an edge decoding layer into the GNN model for predicting relative motion can improve dynamics prediction in a short-time scale. However, incorporating relative motion does not enhance dynamics prediction on longer timescales. Moreover, the work [21] proposed simplifying the GNN model with handcrafted descriptors based on message-passing rules. Such a basic model with modest complexity can produce comparable prediction results to GNN [27]. In addition, the authors highlighted that the prediction performance is less dependent on the model but relies strongly on the structural information supplied into the model. Therefore, effectively extracting geometric characteristics is essential for fitting dynamics. Ref. [20] considered the three-dimensional relative positions and assigned them as the input graph’s edge features. However, the original angular geometry of the particle in Euclidean space may be lost when it is projected onto a high-dimensional space. Moreover, the sensitivity of the coordination renders GNN susceptible to rotation. As a result, the model necessitates additional online data augmentation to enhance generalization. Unfortunately, augmentation produces a vast number of dispersed 3D positions, which hinders the learning of local structures from geometric features. This prompts us to question the validity of the GNN extraction of structural elements. Moreover, it is not inappropriate to extract structural data while ignoring the dynamics pattern of the target since the spatial correlations of dynamic targets at various times vary. Thus, extracting distinct structural features for dynamic targets at various times is desirable.

To address the above challenges, we first present a novel graph-based method for evaluating the global and local smoothness of dynamics over time. The results of the spectral graph analysis reveal that the dynamics shifts concerning time. The translation is primarily controlled by the spectral features of the dynamic propensity on the graph. This requires the model to be capable of predicting multiple levels of local smoothness from a single static structure, which GNN models presently overlook. Especially in the short-time scale, the dominant high-frequency pattern of the particle dynamics makes the GNN’s prediction accuracy declines more sharply than in the slow $\alpha$ relaxation. To alleviate this, we develop a novel geometry-enhanced Graph Neural Network (Geo-GNN), which enhances the generalization performance of dynamics prediction across a wide range of times. To provide a more accurate depiction of underlying smooth patterns while establishing the link between structure and dynamics, we meticulously developed the Geo-GNN encoding and aggrega-
tion processes. Experimental results on the Kob-Andersen system show that our model can autonomously understand the underlying dynamic patterns and surpass the existing approaches in predicting short- and long-time dynamics. To the best of our knowledge, this is the first report on using graph spectral theory in studying structural heterogeneity.

RESULTS

**Molecular dynamic simulation data** In this research, we adopt the molecular dynamic simulation data for the 80:20 Kob-Andersen model, which is a mixture of $N = 4,096$ particles: $N_A = 3277$ (large) $A$ Lennard-Jones (LJ) particles and $N_B = 819$ (small) $B$ LJ particles. The same model system has been analyzed in [20] and [21]. The readers can find more detailed information used for simulations and datasets in Ref. [20]. The focus of the state point in this research is placed on the coolest temperature, $T = 0.44$, which is close to the transition temperature predicted by Mode Coupling Theory, $T_{MCT} = 0.43$. The built dataset is composed of 800 independent configurations, of which half are used for training and the other half for testing. For each equilibrated configuration, the dynamic propensities of large particles are computed from 30 trajectories with identical beginning coordinates but random velocities selected from a Maxwell-Boltzmann distribution [28, 29]. Formally, the dynamic propensity of particle $i$ at time $t$ is defined as: $\Delta_i(t) = \langle |r_i(t) - r_i(0)| \rangle$, where $\langle .. \rangle$ represents the ensemble average and $r_i(t)$ represents the position of particle $i$ after time interval $t$. Throughout the text, the real part of the self-intermediate scattering function $s(q, t)$, where $q$ is the first diffraction peak in the static structure factor, is used to represent the corresponding time $t$.

Dynamic propensity serves as the predicted target for machine learning to identify relationships between local structure and dynamics. With the static structure of the equilibrated configurations, our objective is to predict the propensities at various times for particles of type $A$ from the embedded geometry structure.

**Smoothness of the dynamic heterogeneity.** We have shown in Fig. 1 three examples of the dynamics at three different times. One can observe that the spatial correlations of the mobility field contrast sharply between short- and long-time dynamics. In the longer timescale, particles with comparable slow or fast dynamic propensities tend to cluster together, generating the appearance of a more smooth system in which the propensities of
FIG. 1. Three examples of dynamic propensity at $S(q, t) = 0.8$, $1/e$ and 0.1, respectively. There is a striking difference between time steps in the spatial distribution of dynamic propensities. For the dynamics in the short-time scales, particles perform vibratory movements in a temporary cage produced by their neighbors and exhibit a low spatial correlation of the motion. In the long-time scales, the spatial correlation of their motion grows as the particles leave the cage and start a cooperative movement.

neighboring particles are similar [30, 31]. Nevertheless, on a shorter timescale, the system manifests itself as more non-smooth as the dynamics of spatially adjacent particles tend to behave differently. Next, we quantify the spatial correlation of particle dynamics from the graph perspective. This will help us understand how spatial correlations of propensities affect the prediction accuracy of GNN.

Given a three-dimensional input, an undirected graph for larger particles based on cutoff distance $R_c = 2\sigma_A A$ is first constructed. Inspired by the emerging field of graph signal processing [32], the dynamic propensities of particles can be considered as signals on the graph. We can now examine how the dynamics vary among the graph nodes—the smoothness of the graph—to get a quantitative measure of the spatial correlations in particle motion. The slower the propensities vary on the graph, the smoother the dynamics on the graph. In this study, two quantification techniques are proposed to explore the smoothness pattern of particle dynamics from both global and local viewpoints respectively (see Methods).

Fig. 2(a) shows the global smoothness expressed by the Rayleigh quotient of all training and test graphs. As observed, the dynamic propensities at the short-time scales where $s(q, t) \approx 0.8$, 0.7, and 0.6 are the most non-smooth. In other words, dynamic propensities at these times fluctuate dramatically between graph nodes and show a high-frequency dynamics pattern on the graph. On a longer timescale, the Rayleigh quotient gradually decreases over
FIG. 2. **Global smoothness and local smoothness of dynamic propensities.** (a) Rayleigh quotient as a function of time for all configurations at temperature $T = 0.44$. The inset displays the spatial correlation between the central and surrounding particles. A lower Rayleigh quotient relates to a system with a smoother dynamic propensity on the graph and a higher degree of spatial correlation. (b) The probability distribution function (PDF) of the local smoothness $\lambda$. With increasing time, the distribution becomes to be centered around zero.

time, offering a low-frequency pattern of dynamics. This is consistent with the spatial correlations of particle dynamics shown in Fig. 1. Specifically, for a short period, particles perform vibratory movements in a temporary cage produced by their neighbors, resulting in a low spatial correlation of the motion. Therefore, the propensities between neighboring nodes on the graph tend to differ on shorter timescales (represented by a high Rayleigh quotient). After, similar dynamics lead to the formation of clusters when particles leave the cages. The spatial correlation of their motion grows, and the dynamics at the neighboring nodes tend to be similar on longer timescales, which is shown by a low Rayleigh quotient. Moreover, we can further investigate the frequency pattern of dynamics at the particle level to comprehend the dynamics between centered particles and their first nearest neighbors. Fig. 2(b) describes the distribution of the local smoothness for all nodes. A higher absolution of $\lambda_i$ indicates that the dynamic propensity of particle $i$ differs significantly from its neighbors on the graph (see Methods). We can observe that the local smoothness for shorter-time dynamics is more diffuse and becomes centralized for longer ones, indicating that dynamic propensities become similar to their neighbors as time progresses. In the spectral graph
perspective, the high-frequency regions of the frequency spectrum of dynamic propensities diminish with time.

This raises the issue of whether high-frequency patterns originate from fast- or slow-dynamic particles. To answer this question, we show the Pearson correlation between the propensities and local smoothness for all configurations in Fig. 3(a). The local smoothness displays a strong correlation with the dynamics propensities. In the short-time scale where $s(q,t) = 0.8$, the relationship between local smoothness and propensities is nearly linear dependent. The correlations weaken with the longer relaxation time but remain significant. It can be demonstrated from the scatter plot of propensities and local smoothness in Fig. 3(b)-(d). We can see that the physical principles of how the fast- and slow-dynamic particles contribute to the high-frequency are distant. When the particle $i$ has a negative $\lambda_i$, the centered particle that is surrounded by neighboring particles with relatively high dynamic propensities tends to be stable in the system. In contrast, the particles with higher activity levels are more mobile than their neighbors, resulting in local smoothness that is positive. When the central particle and its neighbors have similar mobility, the local smoothness for
the central particle is nearly zero.

The prediction performance of GNN will be heavily impacted by the smoothness attribute associated with particle dynamics. For instance, neighboring particles often have distinct dynamic propensities on short timescales. Suppose we use a low-pass filter to force the representation of linked particles to be identical. In that case, the performance will be severely degraded since the low-frequency information is inadequate to infer a non-smooth pattern of dynamics. However, information related to high frequency that captures the difference between nodes may be preferable. Many existing works agree that the GNN works like a low-pass filter and might smooth out the difference between neighboring nodes by removing high-frequency signals [33, 34]. This could cause the learned representation of central particles to converge to the averages of the local structure. This is trivial for glassy dynamical predictions when the spatial correlation scale of dynamic propensities is large since the smoothness is dominant and the fast- or slow-moving particles gather in patches. But when the spatial correlation length scale decreases, it gets non-trivial due to the non-smooth pattern of dynamics. Therefore, GNN is easily affected by whether these dynamics are smooth or not. As a result, GNN predicts long-term dynamics better with smoother targets but is limited to short-time scale with more non-smooth targets, as shown in [20]. In addition, the essence of averaging the local structure is still a low-pass filter [21], thus performing poorly with non-smooth dynamics. This reminds us that the existing GNN low-pass filter is far from optimum to produce reliable predictions for glassy dynamics. For the glassy dynamic prediction to be more accurate, the latent features that a GNN learns must be able to show the underlying smoothness pattern. In this paper, we strengthen the capability of GNN to learn smoothness patterns of dynamics in two ways: by increasing the feature extraction ability and making feature aggregation more adaptable across a wide range of smoothness patterns.

**The Geo-GNN model.**

The details of the proposed architecture are shown in Fig. 4. It primarily consists of two components: 1) a feature encoder for describing locally geometric environments and 2) an effective feature aggregation framework for preventing aggregated particle features from being smooth when the actual local dynamics are more non-smooth. The 3D edge type, distance, and angle are added for each constructed graph as initial data. The embedding layer further processes these details to yield representations for edges and triplets. Next,
the stack of triplet-to-edge interaction layers propagates edge and triplet information, giving
the learned local structural features at each edge. Then, a series of edge-to-node and node-
to-node interaction operations are performed to aggregate the features of each particle. The
output features of the layer are then sent to the multi-layer perceptron (MLP) for the final
predictions of particle propensity. In the following, we discuss the different components of
the model.

**Graph input layer** Following [20], we firstly establish a basis graph consisting of the node
set $\mathcal{V}$ and edge set $\mathcal{E}$ from the input 3D particle glass system. Each particle is regarded
as a node $i \in \mathcal{V}$. Two directional edges $(i, j)$ and $(j, i)$ connect the pairs of nodes $i$ and $j$ with opposite directions when the interparticle distance $d_{ij} = \|r_i - r_j\|$ is smaller than the
cutoff threshold. Since the primary relative positions are changeable and inconsistent across
various coordinate systems, we consider the rotational-invariant distance as the edge feature
providing the two-body geometry information. Then, we have two directed edges with the
same feature as $d_{ij} = d_{ji}$ but are in opposing directions. We associated the type labels with
inges based on the particle types of paired nodes to discriminate the edges. For the binary
system with particle types $A$ and $B$, there will be four kinds of edges labeled 0, 1, 2, and 3,
which correspond to $A \rightarrow A$, $A \rightarrow B$, $B \rightarrow A$, and $B \rightarrow B$, respectively.

In contrast to the present GNN, which only includes two-body connections, we start
with the basic graph and include three-body connections $\mathcal{A}$ that correspond to triplets of
particles. Given pair of edges $(i, j)$ and $(j, k)$ on the basis graph, the three-body connections
$(i, j, k) \in \mathcal{A}$ with $j \neq i$, are formed when the target node $i \in \mathcal{N}(j)$ interacts with the source
$k \in \mathcal{N}(j)$ by the passing node $j$. Then, angles $\theta_{ijk} = \arccos\left(\frac{r_{ij} \cdot r_{jk}}{\|r_{ij}\| \|r_{jk}\|}\right)$ between two edges
can be explicitly incorporated into the graph as the initialize triplet features. Instead of
using relative locations and distances, our model may leverage higher-order geometric data
to describe the local structure of particles in more detail.

**Graph embedding layer** Before the start of the geometric encoding, the scalar properties
(i.e., edge type, distance, and angle) are first embedded to get initial scalar embeddings
for all edges and triplets. In contrast to the GNN, we do not use any initial features
tied to individual nodes. The embedding layer first converts edge types to their one-hot
representations and further embeds them into dense vectors $z_{ij}$:

\[
z_{ij} = 1_{\text{Hot}}(Z_{i,j})W_z, \tag{1}
\]
where $\mathbf{1}_{\text{HoT}}()$ is the one-hot encoder of $Z_{i,j}$ and $W^z$ is the trainable embedding matrix which is initialized randomly and optimized in the training process. Next, we adopt the radial basis functions layer to encode the interparticle distances: $\mathbf{e}_{ij} = [R_1(r_{ij})]|...|R_N(r_{ij})]$, where $|$ represents concatenation over the scalars to get $K$-dimensional embedding, $R_k(r_{ij}) = e^{-\beta(r_{ij} - \mu_k)}$ with $\mu_k$ being uniformly selected between zero and the distance cutoff $R_c$, while $\beta$ is a learnable scaling parameter. The edge type and distance can be further combined to get the initial latent features for edges:

$$h_{ij}^0 = \text{MLP}(z_{ij}\|\mathbf{e}_{ij}).$$  \hfill (2)

For the initial triplet features, we use the projection of spherical harmonics on $\theta_{ijk}$ to encode the three-body correlations, which are given by:

$$t_{ijk} = [Y_0^0(\theta_{ijk})]|...|Y_{N_0}^0(\theta_{ijk})]|W_\theta,$$  \hfill (3)

where $Y_1^0(\theta_{ijk})$ is the $l$-th order real spherical harmonic and $W_\theta$ is the trainable parameters. Empirically, these representations can provide detailed geometry information to the network and help our model converge fast during training.

**Geometric feature encoder** After the initial embeddings are established, Geo-GNN conducts angle-aware convolutions in the triplet-to-edge layer to comprehend the geometric environment. The rules of the convolution that propagate the features from the triplet to the edge can be referred to as the message passing on a line graph [35–37]. As shown in Fig. 4(d), the embedded environment for particle pairs is updated by employing the following angle-informed convolution:

$$m_{ij}^L = \sum_{k \in N_j \setminus \{i\}} W_0^L t_{ijk} \odot W_e^L h_{jk}^{L-1},$$  \hfill (4)

$$h_{ij}^L = \sigma(W_e^L[\sigma(W_e^Lh_{ij}^{L-1})\|m_{ij}^L]).$$  \hfill (5)

where $\sigma$ is the activation function, $\odot$ denotes the element-wise multiplication and $W_0^L$, $W_e^L$, $W_e'\ell$ and $W_e'\ell'$ are trainable weight matrix. After performing $L$ triplet-to-edge layers, information from multi-hops away can be incorporated into the learned geometric descriptors associated with pairs of neighboring particles. To allow the model easily propagate the features from earlier layers forward, we use the residual update as in Eq. 5 by concatenating...
the newly updated geometric information with the features from the previous $L-1$ layer to yield the final geometric environment embeddings.

Traditional methods typically rely on hand-crafted descriptors such as symmetry functions [38] to describe the particle environment in a rotationally invariant way. The radial symmetry function is used to characterize the radial density environment, while the angular symmetry function to characterize the bond orientations. These methods first project the distance and angle between neighbor particles onto a combination of basis functions. The sum of the expanded features of nearby particles is then used to describe the local environment around the central particle. Naturally, this method results in coarse-grained descriptors, which fail to capture the essential details of dynamic smoothness. Here, we characterize the local environment for all edges instead of nodes to achieve a more fine-grained description. As shown in Eq. 4, angle-aware convolution can be viewed as a weighted sum of the bond orientations in the local environment. But unlike in the hand-crafted descriptors, we can learn the weights of the sum from the neural network. Then the embedded environment descriptor $m_{ij}^L$ can contain more expressive three-body information.

Feature aggregation When the geometric environment features for two nearby particles are extracted, we aggregate the structural features related to dynamics at the node level by using two edge-to-node layers and one node-to-node layer. According to the previous assumption, it is reasonable to suggest that structural features should likewise exhibit smoothness patterns that correlate to dynamic heterogeneity. Thus, the aggregated node feature must contain the non-smooth parts that define the local smoothness pattern between nearby nodes without compromising the smooth feature shared by neighboring nodes. Therefore, beyond averaging the local structure, we design a self-attention mechanism to assign different attention weights between nodes that can adaptively handle the propagation of high-frequency and low-frequency structural features.

In the first step of edge-to-node aggregation, we initialize the message with a skip connection from the edge embeddings $h_{ij}^0$ to this layer for keeping more original two-body geometric information. Next, we augment the edge message from node $i$ to $j$ with the learned high-order geometric details by concatenating the newly updated node features $h_{ij}^0$ and $h_{ij}^0$ with environment embeddings $h_{ij}^L$ learned by the geometric feature encoder. Specifically, two
layers of edge-to-node message passing are given:

$$\alpha_{ij} = \tanh(W_\alpha e_{ij})$$ (6)

$$h_i = \sum_{j \in \mathcal{N}_i} \alpha_{ij} \odot W_0 h_{ij}^0$$ (7)

$$\alpha'_{ij} = \tanh(W_{\alpha'}[h_i \parallel h_{ij}^L \parallel h_j])$$ (8)

$$h_i = \sigma(\sum_{j \in \mathcal{N}_i} \alpha'_{ij} \odot \sigma(W^1_i[h_i \parallel h_{ij}^L \parallel h_j]))$$, (9)

where \(\tanh\) is the hyperbolic tangent activation function, whose output is ranged in \((-1, 1)\).

Here, we propose a geometric-enhanced self-attention strategy where the importance of the neighbor nodes is calculated from the learned geometric feature. Naturally, the neighbor nodes that received a highly important score contribute more to the aggregation of the target node feature. Traditional attention schemes, such as GAT [39], typically use the softmax to force the attention coefficients to be positive. But the aggregation based on non-negative attention scores is still equivalent to averaging the local structure, which remains unfavorable for learning high-frequency structural features. By relaxing the attention weights to negative values through the \(\tanh\) activation function, the target node feature can be augmented not only by the similarity with the non-negative attentions but also by the differences of neighboring nodes with negative attention. Thus, we can use the neighboring difference to learn more high-frequency dynamic patterns beyond averaging the local structure. Similar attention strategies have also been unitized in other domains [40, 41].

To provide more accurate predictions, we further fine-tune the node feature to enhance the high-frequency components from its neighboring environment rather than conducting the node regression directly. The following node-to-node aggregation is performed:

$$h_i = \sum_{j \in \mathcal{N}_i} MLP(h_i \parallel h_i - h_j).$$ (10)

This layer explores the dissimilarity between a node’s self-feature and its local environment. Intuitively, the aggregation of the difference between nodes works similarly to a high-pass filter, enabling it to learn more high-frequency information and preventing the node from being homogeneous. Combining the self-feature with the neighbor difference allows the model to adaptively understand the crucial component from the supervised dynamic propensity.

**Geo-GNN predicting propensity**
We first assess how well the proposed Geo-GNN does in predicting dynamical propensi-
ities. In this study, performance is evaluated by the Pearson correlation value between the
predicted and actual propensities. A higher correlation coefficient indicates better perfor-
mance. Fig. 5(a) shows the comparison results at the temperatures $T = 0.44$ and $T = 0.56$.
Our model gives the best performance across all times at the two-state points. It demon-
strated that our approach is superior at capturing the underlying structure and dynamics.
When compared with the state-of-the-art GNN, it is noticeable that our approach results in
significantly different prediction behaviors. Considering the period from the cage dynamics
to the relaxation time at the lowest temperature $T = 0.44$, our model shows overwhelming
superiority to GNN in predicting the short-time dynamic propensities. Especially, Geo-CNN
achieves the best performance in the caging regime where $S(q, t) = 0.8$, while GNN has the
worst prediction performance at this time. This implies that Geo-GNN understands more
structural elements for dynamic predictions than GNN can in these time scales. After step-
ing into the diffusion region, GNN starts to close the gap, but our model is still much
superior. Similarly, after the glass plateau ends, the performance of 3D convolutional net-
works approaches that of the baseline GNN. In addition to better prediction accuracy, we
also discover that our rotation-invariant model, which uses distance and angle as inputs, has
more stability since it does not rely on an extra augmentation step. This differs from the
rotation-sensitive GNN, which uses relative position as the geometric prior.

To make our predictions more interpretable, we show the Rayleigh quotient of Geo-GNN
estimated propensities in Fig. 5(b). The Geo-GNN predicted global smoothness for all test
data set shifts similar to the ground truth as the time increased. This thoroughly explains
the predicted dynamics that exhibit the high-frequency patterns on the short timescale and
the low-frequency patterns on the long timescale. As clarified in the previous section, the
short-time dynamic propensity shows the non-smooth property. Gaining access to these
non-smooth patterns of dynamics allows our model to improve its accuracy on shorter-time
predictions. However, the low-pass property of GNN and CNN makes them hard to extract
the non-smooth features corresponding to dynamics, thus causing performance degradation
at these times. The smoothness properties of diffusive dynamics require predictions that
depend more on the shared characteristics of neighboring particles. So, it is easy to figure out
the relationship between structures with the smooth pattern of dynamics using methods like
Geo-GNN, GNN, and CNN that build on the aggregated features of the local environment.
We show the target and predicted dynamic propensity in Fig. 5(c). For better visualization, we project the three-dimensional positions of each particle into two-dimensional space under periodic boundary conditions. We can see that the predicted propensities of our model agree well with the targets. The predicted dynamics between neighbors tend to be dissimilar at $S(q, t) = 0.8$ and become smooth at relaxation time where $S(q, t) = 1/e$ and the diffusive regime where $S(q, t) = 0.1$, which are consistent with our previous analysis.

It should be noted that the Pearson correlation coefficients across short-time scales at high-temperature appear to be the highest, and the predicted smoothness at these points fit the ground truth better than the predictions in the lowest temperature. Thus, we guess that the models’ inability to achieve the upper bound of the correlation values is due to the complexity of smoothness in the dynamics. That is, when the inherent structure-dynamic correlates well in the system, the reason why the model performs poorly is that it has not adequately learned the smoothness pattern of dynamics from the structure. The subsequent experiments confirm this perspective.

**Contribution of geometric feature encoder** Our approach has been proven to be successful in identifying structure-dynamic relationships. It offers a new state-of-the-art method to help us further strengthen the prediction of dynamics in glassy systems. It is still unclear what part of structural information hidden inside a glassy system affects the smoothness patterns of dynamic heterogeneity and, by extension, plays a significant role in dynamics prediction. Next, we try to figure out what’s inside the black box by conducting detailed experiments. These experiments will help us find the physical insight behind the proposed model in glassy dynamics prediction.

We begin our exploration with the message-passing routines for structure feature extraction. GNN designs its feature vectors using several recurrent layers whose weights are shared across all stages. For each layer, updated nodes and edges are reassigned to the graph for the subsequent iteration. In contrast, we separate the edge update from the node aggregation and use a forward-learning technique to learn the features hierarchically. Thus, it can help us examine the role of each layer in the prediction. Our next step is to investigate how the depth of the geometric feature encoder influences dynamic tendency prediction. As shown in Fig 6(a), when we drop all triplet-to-edge layers and maintain the graph embedding layer and node aggregation layers in the Geo-GNN, such a simplified model (Geo-GNN(0)) behaves quite similarly to the GNN ones. There is only a slight difference in the short-time where
$S(q, t) = 0.8$. This inevitably makes us doubt the ability of GNN to extract structure features. Note that GNN implicitly learns the distance and angle features by taking the relative positions as the geometry inputs. In contrast, we explicitly incorporate the distance and angle information at the geometric feature encoder. As observed, it is possible to improve performance by using only one triplet-to-edge layer (Geo-GNN$^{(1)}$). The correlations could improve by adding a second triplet to the edge layer. However, the inclusion of a third layer only slightly improves the performance. What is more intriguing is that, at the timescale of the cage dynamic, the model gains more by increasing the number of layers.

We may wonder which geometry factors affect the prediction performance of the model. To answer this question, we retain the distance information at the input graph and validate how much the angular feature in the geometric input affects the model’s performance. Intriguingly, all revised models yield the same performance at interesting times without angle features corresponding to triplet atoms, as shown in Fig. 6(c). The increase in message-passing shells does not help improve accuracy. Thus, the three-body connections accompanied by distance information can not provide sufficient structural details. It is also observed that knowing angle information for GNN only slightly improves the performance in short-time scales and achieves the same performance in predicting long-time dynamics by using distance as input. However, Geo-GNN came to a different conclusion. By including the angle features, the model can gain more performance improvement from the increase in the number of layers, particularly in predicting the short-time scale dynamics. The impact of increasing the encoder layer is weakened in the large timescale but still plays an important role. Thus, we believe that adding angle information is critical for improving dynamics prediction across all timescales. This observation is consistent with the conclusions in the previous research$^{[21]}$. But more important here is that we can further understand this phenomenon from the correlation between dynamic smoothness.

Can the model figure out how smooth the dynamics are by making the triplet-to-edge layer deeper? To clarify this, in Fig 6(b), we report the Pearson correlation between the actual and predicted local smoothness. Consistent with the performance of dynamics predictions, the correlations about local smoothness also grow gradually as the model accesses longer-range orientation order. Even adding a single layer may significantly enhance the ability to learn the smoothness of dynamic heterogeneity and improve the accuracy of the predictions. Especially at the short-time scale, the depth of the geometric feature encoder is crucial
for capturing spatial heterogeneity at these times. We can interpret this phenomenon by exploring the smoothness poetry of dynamics. While the studied area for the feature encoder is small, the model is unable to capture the relevant features since the nearby nodes tend to have distinct dynamics, particularly on a short-time scale. Informative nodes with the same dynamics are usually located outside the local multi-hop neighborhood. Thus, taking the long-range dependencies allows the model to learn more informative features with the help of the distant nodes. Moreover, we also find that the model without the feature encoder layer performs poorly in capturing the smoothness of dynamic heterogeneity. Especially in the short time, the smoothness of the prediction results deviates significantly from the initial Geo-GNN predictions, resulting in less accuracy of dynamic propensity prediction. It is reasonable to conclude the model can better reveal the underlying smoothness patterns if combined with longer-range geometric feature information.

Overall, it is clear from the previous experiments that understanding the smoothness pattern of particle dynamics is crucial to enhancing the precision of dynamic prediction. Furthermore, the key to understanding this heterogeneity is explaining the long-range angular order in the local structure.

**Contribution of feature aggregation layer**

After completing the geometric feature encode, the features on the edges must be aggregated to the nodes. Using two edge-to-node layers and one node-to-node layer, we have well-designed the feature aggregation process to prevent homogeneity of aggregated node features during message passing. Next, the proposed strategy is validated by replacing the whole aggregation process with the MLP, as in GNN. Specifically, the updated feature for each node $i$ is given by:

$$h_i = MLP\left(\sum_{j \in N_i} h_{ij}^{0} \parallel \sum_{j \in N_i} h_{ij}^{L}\right),$$

which are then regressed to the propensity of node $i$.

The comparison results with different node aggregation strategies are shown in Fig. 7. It can be demonstrated that the two-step aggregation approach we suggested predicts outcomes substantially better than MLP with the identical encoding feature. This further indicates that our approach can adaptively learn the underlying dynamic heterogeneity at various times. As seen in Fig. 7(b), utilizing MLP to aggregate node features decreases the correlation between the predicted and actual local smoothness. Our proposed two-step aggregation
can help us learn the local smoothness pattern of dynamics more effectively than MLP, and thus it can prevent the nodes from being too similar. However, directly averaging the messages on the edges may obscure some critical structural information required to predict the underlying smoothness patterns.

**DISCUSSION**

In this study, we discovered that the non-smoothness of glassy dynamics on the graph is an intrinsic characteristic. We provide supporting evidence for this finding and demonstrate that capturing the smoothness of dynamic heterogeneity is vital for dynamics prediction. Therefore, we introduce a geometry-enhanced graph neural network model (Geo-GNN) with enhanced dynamic prediction accuracy and robustness. The results of ablation experiments demonstrated that the angle order is critical in predicting the smoothness patterns of dynamic propensities. This discovery paves the way for future expansions and enhancements in addressing the structure dynamics relationship. Future research will focus on the smoothness of the structural features in the glassy system and how it affects the smoothness of short- and long-term dynamics.

**METHODS**

**Graph signal and smoothness** An undirected graph $G = (V, E)$ can be used to represent a single static snapshot of the glassy system, where $V$ is the set of nodes that correspond to particles and $E$ is the edge set that reveals the connections between the nodes. An undirected edge links two nodes if their distance is smaller than a defined threshold. Here, we only consider the particles of type $A$ when constructing the graph. The graph can be further described by an adjacency matrix $A \in \{0, 1\}^{N_A \times N_A}$, where $A_{ij} = 1$ indicates that a connection exists from $i$ to $j$ in the graph (i.e., $(i, j) \in E$) and $A_{ij} = 0$ otherwise. Since $A$ is unweighted, we have $A_{ij} = A_{ji}$. The representation of the graph Laplacian can be obtained by $L = D - A$, where $D$ is the degree matrix with the diagonal elements $d_{ii} = \sum_{j=1}^{N_A} A_{ij}$ and $d_{ij} = 0$ for $i \neq j$.

Each node $i$ in a built graph can be assigned with its dynamic propensity $\Delta_i(t)$ at time $t$. A vector $\Delta(t)$ stacking these scalar values is called a graph signal. The Rayleigh quotient
of a graph signal $\Delta(t)$ supported on the graph is defined as

$$R(L, \Delta(t)) = \frac{\Delta(t)^T L \Delta(t)}{\Delta(t)^T \Delta(t)} = \frac{\sum_{i \in V} \sum_{j \in \mathcal{N}_i} (\Delta_i(t) - \Delta_j(t))^2}{\sum_{i \in V} \Delta_i^2(t)},$$  \hspace{1cm} (12)$$

where $\mathcal{N}_i$ is the neighbour set of particle $i$. The Rayleigh quotient can be thought of as a way to measure how smooth a signal is on a graph. A higher relay quotient value indicates the graph signal fluctuates strongly across the nodes, which corresponds to the high-frequency graph signals. The Rayleigh quotient has also been used to understand graph neural networks from the graph filtering perspective [43].

To quantify the dynamical heterogeneity at the particle level, we also defined the local smoothness of propensity for each node as follows:

$$\lambda_i(t) = \sum_{j \in \mathcal{N}(i)} \frac{\Delta_i(t) - \Delta_j(t)}{\Delta_j(t)}.$$  \hspace{1cm} (13)$$

The local smoothness measures how much the signal values at central node $i$ differ from neighboring nodes. The larger the absolute value of local smoothness, the more significant the difference between the central particle and its surrounding neighbors.

**Training details** The loss function is calculated by minimizing the L2-norm of the difference between the predicted and actual probabilities of type A particles. The architecture and hyper-parameters were optimized at the temperature $T = 0.44$ and $S(q, t) = 1/e$. The Geo-GNN is implemented by three layers of triplet-to-edge for the geometric feature encoder, two layers of edge-to-node with 32 hidden units, and one layer of node-to-node, each having 32 hidden units. The cutoff and number of the radial basis function as $5.0\sigma_{AA}$ and $N_r = 64$, respectively. We use $N_{\theta} = 16$ for the representation of the angle. In all our tests, we train Geo-GNN using the Adam optimizer [44] with an initial learning rate of 0.001.

**Data availability.** The MD simulation trajectories of the Kob-Anderson system are available in Ref [20].

**Code availability.** Geo-GNN is implemented using PyTorch Geometric [45]. The code is available from https://github.com/xjiang-hnu/Geo-gnn.

[1] Angell, C. A. Formation of glasses from liquids and biopolymers. *Science* **267**, 1924–1935 (1995).
[2] Ediger, M. D., Angell, C. A. & Nagel, S. R. Supercooled liquids and glasses. *The journal of physical chemistry* **100**, 13200–13212 (1996).

[3] Debenedetti, P. G. & Stillinger, F. H. Supercooled liquids and the glass transition. *Nature* **410**, 259–267 (2001).

[4] Lubchenko, V. & Wolynes, P. G. Theory of structural glasses and supercooled liquids. *Annual Review of Physical Chemistry* **58**, 235 (2007).

[5] Berthier, L. & Biroli, G. Theoretical perspective on the glass transition and amorphous materials. *Rev. Mod. Phys.* **83**, 587 (2011).

[6] Keys, A. S., Garrahan, J. P. & Chandler, D. Calorimetric glass transition explained by hierarchical dynamic facilitation. *Proceedings of the National Academy of Sciences* **110**, 4482–4487 (2013).

[7] Angelini, T. E. *et al.* Glass-like dynamics of collective cell migration. *Proceedings of the National Academy of Sciences* **108**, 4714–4719 (2011).

[8] Ediger, M. D. Spatially heterogeneous dynamics in supercooled liquids. *Annu. Rev. Phys. Chem.* **51**, 99–128 (2000).

[9] Leocmach, M. & Tanaka, H. Roles of icosahedral and crystal-like order in the hard spheres glass transition. *Nat. Commun.* **3**, 974 (2012).

[10] Malins, A., Eggers, J., Royall, C. P., Williams, S. R. & Tanaka, H. Identification of long-lived clusters and their link to slow dynamics in a model glass former. *J. Chem. Phys.* **138**, 12A535 (2013).

[11] Tong, H. & Tanaka, H. Revealing hidden structural order controlling both fast and slow glassy dynamics in supercooled liquids. *Phys. Rev. X* **8**, 011041 (2018).

[12] Tong, H. & Tanaka, H. Structural order as a genuine control parameter of dynamics in simple glass formers. *Nat. Commun.* **10**, 1–10 (2019).

[13] Marín-Aguilar, S., Wensink, H. H., Foffi, G. & Smallenburg, F. Tetrahedrality dictates dynamics in hard sphere mixtures. *Phys. Rev. Lett.* **124**, 208005 (2020).

[14] Widmer-Cooper, A. & Harrowell, P. Predicting the long-time dynamic heterogeneity in a supercooled liquid on the basis of short-time heterogeneities. *Phys. Rev. Lett.* **96**, 185701 (2006).

[15] Widmer-Cooper, A., Perry, H., Harrowell, P. & Reichman, D. R. Irreversible reorganization in a supercooled liquid originates from localized soft modes. *Nat. Phys.* **4**, 711–715 (2008).
[16] Clegg, P. S. Characterising soft matter using machine learning. *Soft Matter* **17**, 3991–4005 (2021).

[17] Cubuk, E. D. *et al.* Identifying structural flow defects in disordered solids using machine-learning methods. *Phys. Rev. Lett.* **114**, 108001 (2015).

[18] Ma, X. *et al.* Heterogeneous activation, local structure, and softness in supercooled colloidal liquids. *Phys. Rev. Lett.* **122**, 028001 (2019).

[19] Schoenholz, S. S., Cubuk, E. D., Sussman, D. M., Kaxiras, E. & Liu, A. J. A structural approach to relaxation in glassy liquids. *Nat. Phys.* **12**, 469–471 (2016).

[20] Bapst, V. *et al.* Unveiling the predictive power of static structure in glassy systems. *Nature Physics* **16**, 448 (2020).

[21] Boattini, E., Smallenburg, F. & Filion, L. Averaging local structure to predict the dynamic propensity in supercooled liquids. *Physical Review Letters* **127**, 088007 (2021).

[22] Wang, Q. & Zhang, L. Inverse design of glass structure with deep graph neural networks. *Nature communications* **12**, 1–11 (2021).

[23] Shiba, H., Hanai, M., Suzumura, T. & Shimokawabe, T. Unraveling intricate processes of glassy dynamics from static structure by machine learning relative motion. *arXiv preprint arXiv: Arxiv-2206.14024* (2022).

[24] Paret, J., Jack, R. L. & Coslovich, D. Assessing the structural heterogeneity of supercooled liquids through community inference. *The Journal of Chemical Physics* **152**, 144502 (2020).

[25] Boattini, E. *et al.* Autonomously revealing hidden local structures in supercooled liquids. *Nat. Commun.* **11**, 1–9 (2020).

[26] Bihani, V., Manchanda, S., Ranu, S. & Krishnan, N. M. A. Unsupervised graph neural network reveals the structure-dynamics correlation in disordered systems. *arXiv preprint arXiv: Arxiv-2206.12575* (2022).

[27] Alkemade, R. M., Boattini, E., Filion, L. & Smallenburg, F. Comparing machine learning techniques for predicting glassy dynamics. *The Journal of Chemical Physics* **156**, 204503 (2022).

[28] Widmer-Cooper, A., Harrowell, P. & Fynewever, H. How reproducible are dynamic heterogeneities in a supercooled liquid? *Phys. Rev. Lett.* **93**, 135701 (2004).

[29] Widmer-Cooper, A. & Harrowell, P. On the study of collective dynamics in supercooled liquids through the statistics of the isoconfigurational ensemble. *The Journal of Chemical Physics*
126, 154503 (2007).

[30] Karmakar, S., Dasgupta, C. & Sastry, S. Short-time beta relaxation in glass-forming liquids is cooperative in nature. Phys. Rev. Lett. 116, 085701 (2016). URL https://link.aps.org/doi/10.1103/PhysRevLett.116.085701.

[31] Tah, I. & Karmakar, S. Signature of dynamical heterogeneity in spatial correlations of particle displacement and its temporal evolution in supercooled liquids. Physical Review Research 2, 022067 (2020).

[32] Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A. & Vandergheynst, P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE signal processing magazine 30, 83–98 (2013).

[33] Nt, H. & Maehara, T. Revisiting graph neural networks: All we have is low-pass filters. arXiv preprint arXiv:1905.09550 (2019).

[34] Zhou, J. et al. Graph neural networks: A review of methods and applications. AI Open 1, 57–81 (2020).

[35] Gasteiger, J., Groß, J. & Günemann, S. Directional message passing for molecular graphs. In International Conference on Learning Representations (2020). URL https://openreview.net/forum?id=B1eWbxStPH.

[36] Choudhary, K. & DeCost, B. Atomistic line graph neural network for improved materials property predictions. npj Computational Materials 7, 1–8 (2021).

[37] Fang, X. et al. Geometry-enhanced molecular representation learning for property prediction. Nature Machine Intelligence 4, 127–134 (2022).

[38] Behler, J. & Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. Phys. Rev. Lett. 98, 146401 (2007).

[39] Veličković, P. et al. Graph attention networks. arXiv preprint arXiv:1710.10903 (2017).

[40] Bo, D., Wang, X., Shi, C. & Shen, H. Beyond low-frequency information in graph convolutional networks. In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, 3950–3957 (2021).

[41] Yang, L. et al. Diverse message passing for attribute with heterophily. Advances in Neural Information Processing Systems 34, 4751–4763 (2021).

[42] Van Der Maaten, L. Accelerating t-sne using tree-based algorithms. The Journal of Machine Learning Research 15, 3221–3245 (2014).
[43] Hoang, N., Maehara, T. & Murata, T. Revisiting graph neural networks: Graph filtering perspective. In 2020 25th International Conference on Pattern Recognition (ICPR), 8376–8383 (IEEE, 2021).

[44] Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).

[45] Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. In ICLR Workshop on Representation Learning on Graphs and Manifolds (2019).

ADDITIONAL INFORMATION

Acknowledgements This work was supported by

Author contributions. X.J. developed the software and performed the analysis. X.J., Z.T., and K.L. contributed to the interpretation of the results. X.J. and Z.T. conceived the idea and approach presented in this work. All authors contributed to the writing of the paper.

Competing interests. The authors declare no competing interests.

Materials & Correspondence. Correspondence and requests for materials should be addressed to Z.T. (zeantian@hnu.edu.cn).
FIG. 4. Illustration of the geometry-enhanced graph neural network (Geo-GNN). (a) The three-dimensional glass system is represented by a graph where particles are taken as nodes, two nodes at a distance less than 2 LJ-units are connected by an edge. The type and distance of pairs of nodes are assigned as the edge feature. Angles corresponding to triplet particles are assigned for two connected edges. (b) Workflow of Geo-GNN: graph embedding layer, geometric feature encoder, feature aggregation layer, and final regression layer. (c) Message passing scheme for triplet-to-edge layer, edge-to-node layer and node-to-node layer. (d) The triplet-to-edge layer generates the environment embedding by combing the angle representation and embeddings from the last layer.
FIG. 5. **Experiments results on KA system.** (a) Pearson correlation coefficient between predicted and actual propensities for A particles in the KA system at temperature $T = 0.44$ and $T = 0.56$. The results of support vector machine (SVM), convolution neural network (CNN), and graph neural network (GNN) are taken from Ref. [20]. (b) Relay quotient of the ground-truth propensities and the predicted propensities for A particles. (c)-(e) A two-dimensional visualization of the actual propensities (upper row) and Geo-GNN predicted (lower row) for A particles at temperature $T = 0.44$. The three-dimensional positions under periodic boundary conditions are mapped to two-dimensional space by the t-SNE algorithm [42] for visualization.
FIG. 6. Experiments results on different numbers of the triplet-to-edge layer. (a) Pearson coefficient between the actual and predicted propensity at temperature $T = 0.44$ with varying number of triplet-to-edge layers. (b) Pearson coefficient between the actual and predicted local smoothness with a varying number of triplet-to-edge layers. Geo-GNN$(n)$ refers to Geo-GNN using $n$ layers of triplet-to-edge. (c) Person correlation differences with or without angle feature incorporating in the triplet-to-edge layer.
FIG. 7. Experiments results without geometric feature encoder. (a) Pearson coefficient between the actual and predicted propensity at temperature $T = 0.44$. MLP refers to the Geo-GNN using the MLP to aggregate the learned geometric feature from edge to node. (b) Pearson coefficient between the actual and predicted local smoothness.
Supplementary Information: Geometry-enhanced graph neural network for glassy dynamics prediction

Jiang et al.

Supplementary Note 1: impact of edge threshold. In this note, we first explore how the performance of our proposed Geo-GNN depends on the edge threshold employed in the construction of the graph. To this end, we train Geo-GNN at \( s(q,t) = 0.8, 1/e, \) and 0.1 using the graphs with different edge thresholds as the inputs. As shown in Supplementary Fig. 1, when the edge threshold is greater than 1.25\( \sigma_{AA} \), the performance of our model is weakly affected by the edge threshold, especially for the long-time dynamics. As for the short timescale, the accuracy slightly improves when the threshold is increased. Even though we achieve the greatest performance with an edge threshold of 2.0\( \sigma_{AA} \), the resulting model complexity is rather high due to the large number of triplets corresponding to the graph (see Supplementary Table. I). However, the graph with an edge threshold of 1.5\( \sigma_{AA} \) may reduce the number of triplets by over an order of magnitude, resulting in significant computation time savings. Therefore, in the following experiments of this supplemental material, we train the Geo-GNN on graphs with an edge threshold of 1.5\( \sigma_{AA} \) to balance performance and efficiency.

![Supplementary Figure 1](image)

Supplementary Figure 1. Pearson correlation between the dynamic propensity as predicted by Geo-GNN and ground-truth as a function of edge threshold at temperature \( T = 0.44 \).
Supplementary Table I. The number of edges and triplets with different edge thresholds.

| Edge threshold | 1.0   | 1.25  | 1.5   | 2.0   |
|----------------|-------|-------|-------|-------|
| number of edges| \( \simeq 11800 \) | \( \simeq 47000 \) | \( \simeq 61600 \) | \( \simeq 167000 \) |
| number of triplets| \( \simeq 43000 \) | \( \simeq 51000 \) | \( \simeq 875000 \) | \( \simeq 6677000 \) |
Supplementary Note 2: parameter details. We optimized the parameters of Geo-GNN at the state point where $T = 0.44$ and $t$ is chosen at $S(q, t) = 1/e$. The parameter search space and final parameters for the Geo-GNN architecture are specified in Supplementary Table. II. The final parameters are marked in bold. The detailed comparison results of different dimensions of the hidden layers and the length of the radial basis function are shown in Supplementary Fig. 2 and Supplementary Fig. 3. We can find that the results have little dependence on these parameters. It is possible to train the networks to get comparable results with a hidden dimension of just 16. Moreover, we also replace the Gaussian expanded representations of distance with an MLP encoder (labeled as “1” in Supplementary Fig. 3). It can be seen that there is no difference in the results.

Instead of using spherical harmonics to represent the angle, we also utilize the MLP to encode the angle into a vector representation for dynamic prediction. As shown in Supplementary Fig. 4, the spherical harmonics representation of the angle improves the performance for all times, especially for the short-time scales. Therefore, such a representation of angle may offer more reliable information on structural order than just utilizing angle alone.

Supplementary Table II. The parameter search space for Geo-GNN at the temperature $T = 0.44$ and $S(q, t) = 1/e$.

| Parameter                          | Values                  |
|-----------------------------------|-------------------------|
| hidden dimensions                 | [16, 32, 64]            |
| length of radial basis function   | [16, 64, 128]           |
| length of spherical harmonics     | 16                      |
| number of triplet-to-edge layer   | [1, 2, 3]               |
| number of edge-to-node layer      | 2                       |
| number of node-to-node layer      | 1                       |
| dropout                           | [0, 0.1, 0.2]           |
| batch size                        | 1                       |
Supplementary Figure 2. Determining the dimensions of the hidden layer.

Supplementary Figure 3. Determining the length of radial basis function.

Supplementary Figure 4. Determining the representation strategy of angle. Angle-MLP refers to Geo-GNN using MLP instead of spherical harmonics for representing the angle.
**Supplementary Note 3: the detailed comparison between models.** Supplementary Table. III illustrates the critical distinction between Geo-GNN and GNN.

Supplementary Table III. The detailed comparison between the proposed Geo-GNN and GNN.

| Model                | Geo-GNN | GNN  |
|----------------------|---------|------|
| node size            | 3277    | 3277 |
| node feature size    | -       | 1 (type) |
| edge size            | \(\simeq 167000\) | \(\simeq 167000\) |
| edge feature size    | 2 (type, distance) | 3 (relative position) |
| triplet size         | \(\simeq 6677000\) | - |
| triplet feature size | 1 (angle) | - |
| data augmentation    | without | with |
| rotation-invariant   | True    | False |
| recurrent message passing | False | True |
| number of parameters | 30274   | 70721 |