

Automated characterization of spatial and dynamical heterogeneity in supercooled liquids via implementation of machine learning

Viet Nguyen  and Xueyu Song* 

Ames Laboratory and Department of Chemistry, Iowa State University, Ames, IA, United States of America

E-mail: xsong@iastate.edu

Received 7 April 2023, revised 26 June 2023

Accepted for publication 2 August 2023

Published 23 August 2023



CrossMark

Abstract

A computational approach by an implementation of the principle component analysis (PCA) with K -means and Gaussian mixture (GM) clustering methods from machine learning algorithms to identify structural and dynamical heterogeneities of supercooled liquids is developed. In this method, a collection of the average weighted coordination numbers (\overline{WCNs}) of particles calculated from particles' positions are used as an order parameter to build a low-dimensional representation of feature (structural) space for K -means clustering to sort the particles in the system into few meso-states using PCA. Nano-domains or aggregated clusters are also formed in configurational (real) space from a direct mapping using associated meso-states' particle identities with some misclassified interfacial particles. These classification uncertainties can be improved by a co-learning strategy which utilizes the probabilistic GM clustering and the information transfer between the structural space and configurational space iteratively until convergence. A final classification of meso-states in structural space and domains in configurational space are stable over long times and measured to have dynamical heterogeneities. Armed with such a classification protocol, various studies over the thermodynamic and dynamical properties of these domains indicate that the observed heterogeneity is the result of liquid–liquid phase separation after quenching to a supercooled state.

Keywords: supercooled liquids, liquid–liquid phase transition, machine learning

(Some figures may appear in colour only in the online journal)

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction

Glass plays a central role in nature and our daily lives. It is essential in food processing, preservation of wildlife animals under extreme cold [1]. Ordinary window glass, mostly made of sand (SiO_2), lime (CaCO_3) and soda (Na_2CO_3) is a best known manufactured amorphous solid product [2]. Optical wave guides use pure amorphous silica while silicon in photovoltaic cell is amorphous. In principle, glassy state is attained by supercooling a liquid below its melting temperature fast enough to avoid crystallization. Under such rapid cooling, the supercooled liquid attains mesoscopic structural disorder with ‘complex dynamics’ such as non-exponential relaxation, breakdown of Stokes–Einstein relation. Although these heterogeneities are well-known for decades [3–12], there is no direct evidence to consistently classify and correlate these heterogeneities both structurally and dynamically. These following questions remain a puzzle: What cause these heterogeneities to arise? What is the spatial order of magnitude of the domains? How much do dynamics vary among these domains? Answers to those questions could significantly impact our practical applications of glass-forming materials.

Observation of heterogeneous dynamics is directly linked to the onset of cage effect [13] where particles become trapped in local cages by their neighboring particles to prevent them from moving around as a normal liquid. The cage effect is manifested as a plateau in the self intermediate scattering function $F(k, t)$ or the mean squared displacement of particles and could be explained as following: If we take an instant snapshot of the system, we see no impressive structure change close to T_g . Let’s consider two different snapshots taken at two instants of time separated by a time interval t . We can now capture how particles move during this interval t . If the interval t is too short, the system is still in ballistic regime, there is not a significant variations of particles mobility because interaction has not kicked in to make things interesting. Meanwhile, if t is too long, larger than the relaxation time τ_r (the longest relaxation process), time average is equivalent to ensemble average, hence all particles are statistically the same and each particle will have the same mobility. t is selected such that it is long enough to capture particles interaction but short enough to avoid statistical homogeneity to observe the difference of high or low mobility of particles. Hence, such intermediate time t value is closely related to the plateau of the β -relaxation regime where particles become transiently trapped in cages and $F(k, t)$ remains constant. Only at sufficiently long times will particles break free and full relaxation takes place (α -relaxation). Particles mobility can vary several orders of magnitude [14–16]. In addition, particles with one mobility tends to form a cluster or a domain such that the system are filled with different domains of particles. In other words, particles move in cooperatively manner as a dynamically correlated mesoscopic domains with long relaxation time scales [17–21]. A dynamical length-scale ξ can be associated with the increasing dynamic heterogeneities because it measures the size of mesoscopic domains as equivalently to size of growing cooperative motion of particles [22–30].

Several theories of glass transition have been developed to seek a fundamental understanding of these spatial domains: such as the energy landscape picture [31, 32], Adam–Gibbs theory [20, 33, 34], and random first-order transition theories (RFOTs) [23], to name a few. These theories present various pictures of domains thermodynamically. Although these thermodynamic descriptions provide a simple and intuitive framework related to dynamics and spatial structures of supercooled liquids, it lacks a consistent classification protocol to characterize the structure of these mesoscopic domains. The lack of a clear characterization of these domain structures in supercooled liquids has hindered the formulation of a general theory for glass transition. Unlike crystalline solids whose structures can be easily detected due to its periodicity, no general classification scheme has been formulated for supercooled liquids to the best of our knowledge.

Meanwhile, several classification schemes are developed to identify structures in amorphous systems. The first kind of approaches include Voronoi polyhedra [35–39], bond-orientational order parameters [40, 41], the common-neighbor analysis [42–44] and topological cluster classification [45, 46] which are based on identification of a bond network among particles. However, these methods require some specific structural information *a priori* which is unknown in general except for few systems under certain situations. Other general ‘order-agnostic’ approaches [17, 47] which rely not on a specific structure but on some general properties have been developed. One of them using mutual information based on Shannon entropy [48], to determine structural length-scale. Structure in one part of the system can influence structure in another via mutual information, hence mutual information between two regions can be computed as a function of distance [49], which does not require *a priori* knowledge of the structure. Another method is to seek networks among domains. Each domain is considered as a non-interacting isolated community. By minimizing the length-scale of these communities, it minimizes the interaction among communities [50], hence leads to identification of clusters which are not specified beforehand. Another type of methods is to introduce an external, static perturbation in the form of an affine deformation of coordinate data. A drawback of these approaches is that the nature of structures identified is not as clear as the first kind of approaches because it lacks microscopic details of particles like bond network and coordination number (CN).

Given the significance of structural classification in supercooled liquids, we developed a new strategy to classify a supercooled liquid into nano-domains using some algorithms from machine learning (ML) such as the principle component analysis (PCA), K -means and Gaussian mixture (GM) clustering [51–54] both in structural and configurational spaces. This classification protocol shows improvement over discussed methods in previous paragraph because it is similar to ‘order-agnostic’ where the emergence of domains requires no prior knowledge in one hand and at the same time contains information of microscopic details as the first kind of approaches (Voronoi polyhedra, bond-orientational order parameters, etc).

The nano-domains from our approach agree with the picture in the Adam–Gibbs and RFOT theories. Based upon our classification, nature of these spatially distinct domains are clearly characterized and each of these domains is correlated with different diffusion constant distributions within a domain, hence the spatially heterogeneous dynamics naturally falls into two categories: the diffusion within various domains and the domain rearrangement dynamics which reflect the slow relaxation of the system. Structural evolution of these nano-domains is identified as coarsening kinetics from the liquid–liquid phase separation after rapid cooling or quenching. Furthermore, temperature dependence and other properties of nano-domains are also studied to support this picture. A well studied binary Lennard–Jones (LJ) model system, the Kob–Andersen model [55–57], is used to demonstrate the capability of our classification scheme for supercooled liquids since it is known that the model system does not crystallize when it is supercooled well below the melting temperature.

The paper is organized as follows. Section 2 presents a detailed presentation of the proposed method. This is followed by an extensive result presentation with discussions. Some concluding remarks are given in the final section.

2. Classification scheme

2.1. Simulation details

In this work, simulations are done with NPT ensemble (where N is the number of particles, P is pressure and T is temperature) using the molecular dynamics (MD) simulation package, LAMMPS [58]. Noose–Hoover thermostats are employed to control both external pressure (pressure is set to 0) and temperature. The atomic interaction potential used in our work is the well-known Kob–Andersen binary LJ model [55–57]. The standard form of the LJ potential can be expressed as:

$$V(r) = \begin{cases} 4\epsilon_{A,B} \left[\left(\frac{\sigma_{A,B}}{r} \right)^{12} - \left(\frac{\sigma_{A,B}}{r} \right)^6 \right] & \text{for } (r \leq r_c) \\ 0 & \text{for } (r > r_c), \end{cases} \quad (1)$$

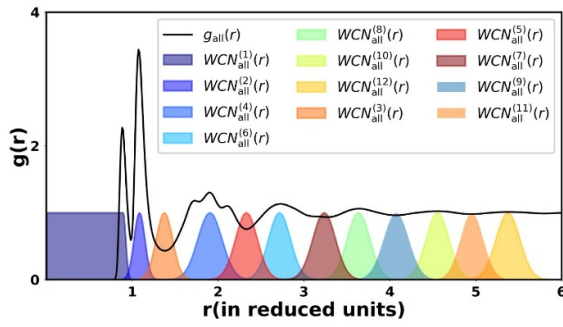
where the parameter ϵ is the potential well depth, σ is the characteristic atomic diameter and the cutting distance r_c is set to $2.5\sigma_{A,B}$. The parameters for solid Ar are adopted [59, 60]: $\sigma = 0.3405 \text{ \AA}$ and $\frac{\epsilon}{k_B} = 119.8 \text{ K}$ where k_B is the Boltzmann constant and particle mass $m = 6.69 \times 10^{-26} \text{ kg}$. The conventional reduced unit for LJ system is used: the mass unit is set to the weight of one Ar atom while the length unit in σ , energy unit in ϵ , the time unit in term of $\tau = t\sqrt{\frac{m\sigma^2}{\epsilon}}$ and reduced temperature is defined by $T^* = T\left(\frac{\epsilon}{k_B}\right)$. The system consists of 80% of A and 20% of B particles with $\epsilon_{AA} = 1$, $\sigma_{AA} = 1$, $\epsilon_{AB} = 1$, $\sigma_{AB} = 0.8$, $\epsilon_{BB} = 0.5$ and $\sigma_{BB} = 0.88$ while $m_A = m_B = 1$. Periodic boundary conditions are applied to all directions. The time step is set to 0.005τ which is about 10 femtoseconds. Number of particles of the systems studied are 5000, 16 000 and 50 000. To prepare the liquid at supercooled conditions, we first heated up the system to a high temperature to obtain a liquid state. After a short period of equilibration

and relaxation, the system is quenched to three different target temperatures $T^* = \{0.37, 0.3, 0.2\}$. The cooling process has been done by linearly decreasing temperature via re-scaling atomic velocity: $T = T_0 - \gamma n$, where γ is the cooling rate ($3.3 \times 10^{10} \text{ K s}^{-1}$ if taking Ar parameters) and n is the number of MD steps. These temperatures are reasonably selected because: $T^* = \{0.37, 0.3\}$ are below the mode-coupling temperature $T_c^* \approx 0.435$ predicted by mode-coupling theory [5, 61] but above glass transition temperature ($T_g^* = 0.25$) [4] to observe any change of dynamics [62] while $T^* = 0.2$ is below the T_g^* to study the trend of structural heterogeneity for temperature dependence. After two million time steps equilibration, the system is run for another 3 million time steps, saving configurations every 100 steps or 0.05τ . The average number density $\rho^* = \{1.14, 1.17, 1.19\}$.

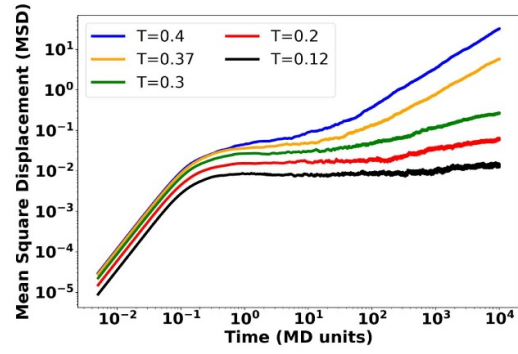
2.2. Radial distribution function (rdf) and weighted CNs (WCNs)

To investigate structural heterogeneities of a disordered system, rdf $g(r)$ is commonly employed to describe spatial local environments by means of collecting averaged CNs which describe the relative number of neighboring particles in a particular surrounding spherical shell of a particle, which is the same for all particles. However, this highly averaged CNs representation of the system lacks the details to provide realistic features of the spatial heterogeneity of a supercooled liquid. Meanwhile, for a particular configuration of the system either by a snapshot from a molecular simulation or an experimental image of supercooled colloidal system from confocal microscopy, local structures for each of an \mathbf{M} particles system can be characterized with its local coordination shell structure. Naturally, a middle ground is an order parameter that can classify these local structures of the system into a few meso-states which is useful to describe the heterogeneous structure of the system. In addition, aggregated clusters or domains, whose particles from the same meso-states should be formed in the configurational space, together tile up the whole system to make classification scheme work both structurally and configurationally. Furthermore, meso-states in the structural space and domains in the configurational space should live long enough to afford further analysis. For example these meso-states and domains can directly relate to the onset of caging effect which is attributed to plateau region of mean-square displacement trajectories in figure 1(b)). In this study, the timescale for this analysis is from 5×10^1 to 2×10^4 MD units or converted to 0.1–40 ns which associates with the plateau region at different temperatures.

Using MD simulations of this model system, the CN of a particle can be calculated. In this study, the A/B identity of the particles is ignored, which can be thought as the supercooled liquid state being generated from an effective one-component system. However, CN-based features suffer a strict cut-off value to determine whether a neighbor particle is counted as in or out of the shell. To avoid this hard assignment, WCNs [63], which utilize the normalized Gaussian distribution based on the shell structure of the system $g(r)$ (figure 1(a)) to weight



(a) WCNs based on rdf



(b) MSD trajectories

Figure 1. (a) Radial distribution function (rdf) employed by the weighted Gaussian functions for the Kob–Andersen binary LJ system at $T^* = 0.2$, regardless of particles identities. The color and superscript indicate the WCN’s positions and numbers along the rdf. (b) All-particle MSDs as a function of time at various temperatures regardless of their identities.

the contribution of each neighboring particle based on the particle’s distance to the center one. Using the relevant solvation shell features as identified maxima and minima along the rdf, the normalized Gaussian distribution functions are placed at the center of these shell features as shown in figure 1(a). The width of the Gaussian functions depends on the area that the solvation feature covers and neighboring Gaussians such that the value of the intersection is assigned to 0 or roughly 0.25 depending on whether or not the two shells are largely overlapping. However, width and the size of the overlapping areas of the Gaussians do not change the consistency of the final results. The WCNs smooth out transitions between solvation shells by counting the particles sitting at the center of the features as one while the one further away from the center feature is counted as a fraction based on the Gaussian distribution function. For each configuration, employing this WCN implementation, each component of a particle’s WCNs vector is determined by summing the weight from all surrounding particles within that shell and the dimension of the WCN vector is determined by the number of shells reasonably covering the main features of the $g(r)$, $N = 12$ in figure 1(a); other numbers of shells tested yield consistent results.

For a single configuration of the simulation, WCNs of all particles are collected from the particles’ CNs smoothed using the $g(r)$, hence the features data for the entire system is represented by a matrix $\tilde{\mathbf{X}}$ of $\mathbf{M} \times \mathbf{N}$ which is obtained from \mathbf{N} WCNs for each of \mathbf{M} particles system. WCNs are noisy and complicated in a disordered system, hence require a further step to remove some of these noises. Instead of WCNs, averaged WCNs is used which has a form: $\overline{WCN}_i = \frac{1}{N_b} \sum_j^{N_b} WCN_j$, where N_b is the number of neighboring particles in each shell plus the particle i itself.

2.3. Dimensionality reduction and clustering

For each particle, each of N features in the \overline{WCN}_s matrix is constructed separately to describe its own local solvation shell environment with respect to its surrounding particles, it is disconnected from each other to form a proper feature space. To

resolve this issue, PCA [64] is used for dimension reduction, namely to linearly transform original \overline{WCN}_s matrix into a new feature space that reduce N particles’ features to a few correlated ones. Mathematically, PCA can be done through the following three steps:

- Obtaining the mean-free data $\mathbf{X} = \tilde{\mathbf{X}} - \langle \tilde{\mathbf{X}} \rangle$ where the average is over M particles for each component of WCNs.
- Forming the correlation matrix $\mathbf{C} = \mathbf{X}^T \mathbf{X}$, which is $N \times N$.
- The principle components \mathbf{u}_i are obtained after solving the eigenvalue problem: $\mathbf{C} \mathbf{u}_i = \sigma_i^2 \mathbf{u}_i$. The eigenvalue σ_i^2 measures the variance of the data along each principle component (PC) i . PCA is optimal in term of seeking small numbers of PCs but maximizing cumulative proportion of variance explained (PVE) σ_i^2 by each principle component. In other words, the numbers of retained PCs depend on their total PVE such that the total PVE is $\geq 95\%$ of total variances presented in $\tilde{\mathbf{X}}$.

The new complete basis composes of all PCs: $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N]$ where each \mathbf{u}_i is a collective coordinate with N components corresponding to the number of features in the data input. In our study, the first three PCs retains about 85%–90%, so 6–7 PCs are sufficient enough to form PC’s basis whose PVE could be $\geq 95\%$ of total variances presented in $\tilde{\mathbf{X}}$. The new coordinates (PC representation) are generated from an inner product of original \overline{WCN}_s matrix with the PC’s basis (PC-space), mathematically, $\mathbf{Y} = \mathbf{U}^T \tilde{\mathbf{X}}$. We then use K -means clustering method to decipher hidden structures of the PC representation by classifying particles into distinct clusters called meso-states. K -means clustering is chosen because it is an unsupervised standard technique that geometrically separate particles into clusters that aggregated together because of certain similarities.

However, the K -means requires prior knowledge of the number of existing clusters K in the data structure to work effectively, which is generally unknown in most cases. An

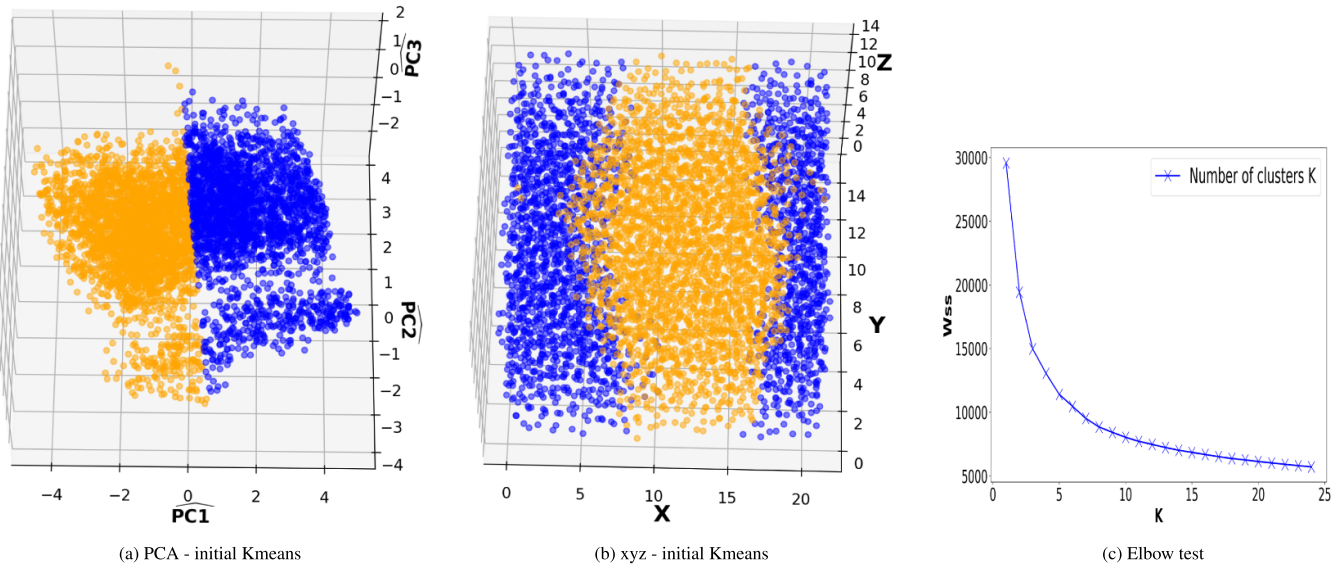


Figure 2. A 3D snapshot of meso-states of 5000 particles system at $T^* = 0.2$ from direct mapping and the Elbow test. Panels (a) and (b) are for the PC-space and configurational space while panel (c) displays the Elbow convergence test. The hat symbol for labeling axes of the PC space represents the inner product of a particle’s $WCNs$ with the PC basis, in this case for the first three PC components. Blue and orange particles belong to meso-state 1 and 2, respectively. X, Y, Z are atomic coordinates in reduced units.

implementation of the Elbow convergent test could provides a reasonable prediction of the K values. The Elbow test permits the number of clusters K being varied freely and computes the Within-Cluster sum of square distance (Wss) which is the sum of square distances between each data point and the centroid within a cluster. As the number of clusters K increase, the Wss will start to decrease and eventually become roughly constant regardless of further increasing K . The Elbow plot of the Wss against K looks like an Elbow shape where the Elbow point normally corresponds to an initial guess of K used in K -means clustering. In many cases, the Elbow plot has a clear Elbow point which indicates a good guess for K -means. In our case, initial K remains uncertain because the Elbow shape is poor to single out an Elbow point, thus we can only narrow down a possible range of K values ($K = 2-5$) (figure 2(c)). After a careful trial-and-error process with help of the co-learning strategy, $K = 2$ is selected; details of the process is discussed in the appendix A. Given $K = 2$, particles in the PC-space are classified into 2 distinct meso-states (figure 2(a)), then a direct mapping using the identities of particles in each meso-state in the PC-space also forms aggregated clusters in the configurational space as shown in figure 2(b); different projected angles of figures 2(a) and (b) to confirm the clustering structures both in PC and real space are in appendix B. Naturally, each meso-state have mixing A and B particles. On the other hand, each type of (A or B) particles itself appears as two distinct aggregated domains in the configurational space. This is clearly demonstrated in the appendix C.

Although domains are generated in the real space by a simple mapping of particles’ identities in the PC-space after K -means clustering, there are two issues needed to be addressed. Firstly, the principle of K -means clustering relies on assigning

a particle to a cluster where its Euclidean distance (E-dist) to the centroid of that cluster is the closest among others. In other words, assignment of a particle depends on the E-dist measure sensitively which becomes robust for core particles of each meso-state because the difference of their distances from one state to another is well-defined. However, the E-dist criterion becomes an issue to assign interfacial particles due to the small differences in their distances to either states, so it could lead to misclassification. Secondly, even though identities of clusters are preserved from the PC-space to the configurational space the inverse transfer of the knowledge is not clear, but physically the transfer of knowledge should be bi-directional. Thus, a co-learning strategy is developed as the following:

1. Perform K -means clustering in the PC-space.
2. Use the initial knowledge of the clustering from the PC-space to perform a GM classification in the configurational space to soften the hard assignment from the K -means:
 - (a) do a direct mapping of particles identities in the PC space to identify distinct nano-domains in the real space.
 - (b) build a mixture model of multivariate Gaussian distributions of domains, then assignment of a particle belonging to a domain is determined by maximizing Gaussian probability among different domains.
3. Similar to step 2, perform GM in the PC-space from the clustering knowledge in the configurational space.
4. Iteratively perform GM classification in both spaces until convergence.

Classification of interfacial particles by the co-learning strategy converges quickly in both PC-space (figures 3(a) and

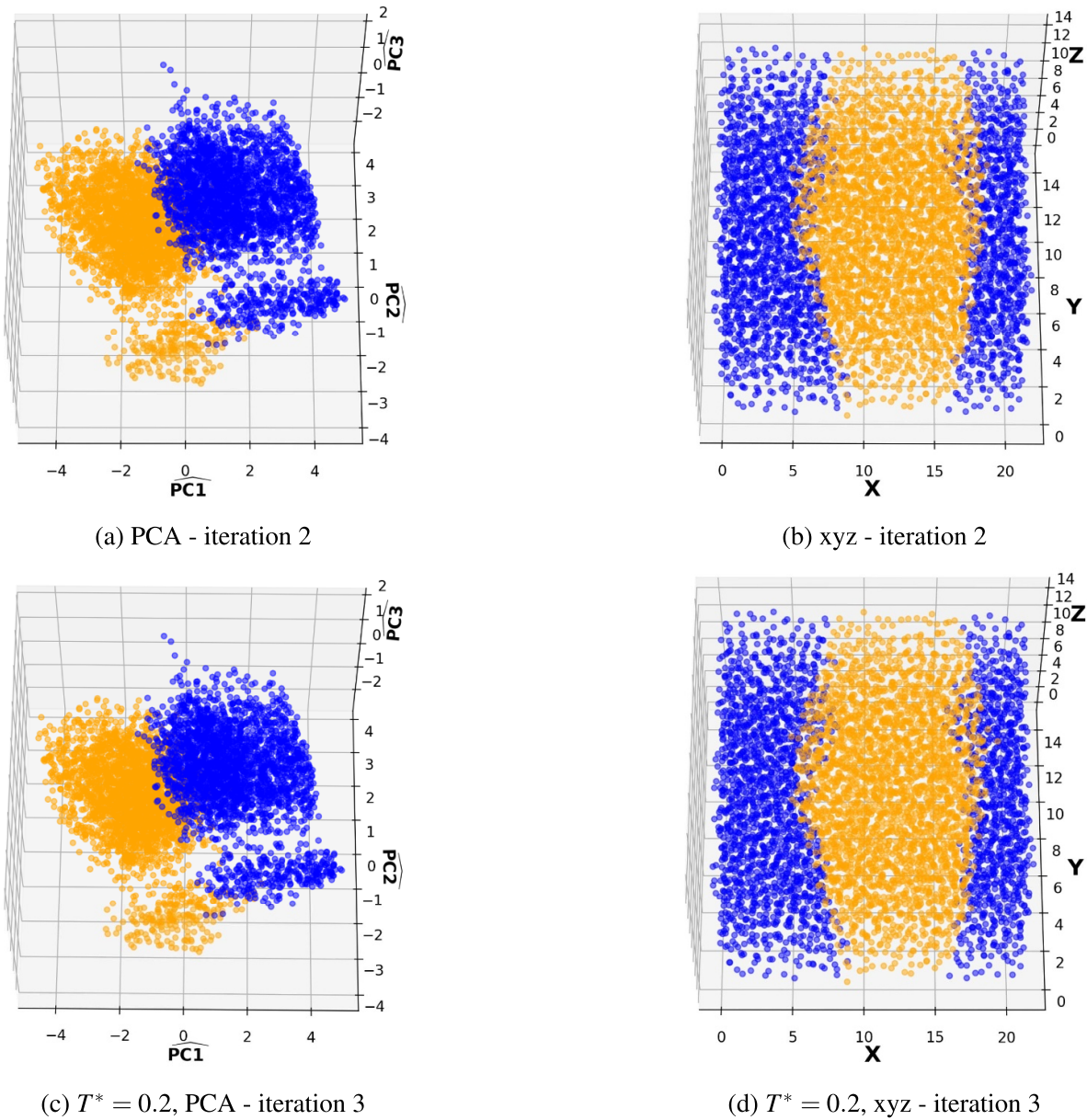


Figure 3. A 3D snapshot of 2 meso-states of 5000 particles system at $T^* = 0.2$ from our co-learning method. Panels (a) and (c) are for the PC-space and configurational space for iteration 2 while panels (b) and (d) represent clusters formed in the PC-space and configurational space for iteration 3.

(c) and configurational space (figures 3(b) and (d)) after few iterations (3–5 runs on average) as shown in the figure 3. The co-learning strategy shows improvement over K -means as it generalizes and fills the missing information from a direct information transfer from the PC-space to the configurational space. In other words, it allows a bi-directional information transfer. Firstly, correct classification of interfacial particles comes from using probabilistic clustering like GM to avoid sensitivities of E-dist criterion of the K -means. This GM clustering allows assignment of interfacial particles to two states and the decision is made by the maximum-likelihood of the Gaussian probability, which creates a boundary region of meso-states in the PC-space as shown in figure 3(c). In the configurational space, we also find that the core particles

in both domains are still the same, only interfacial particles are properly re-assigned to make the final results consistent with the direct mapping (figure 2(b)) and co-learning strategy (figure 3(d)). Secondly, the classification scheme utilizes the information from both spaces in a self-consistent manner.

3. Results and discussion

3.1. Nature of nano-domains: statics

In the previous section, a picture of structural and configurational heterogeneity is revealed by the classification of the system into meso-states (in PC-space) or nano-domains

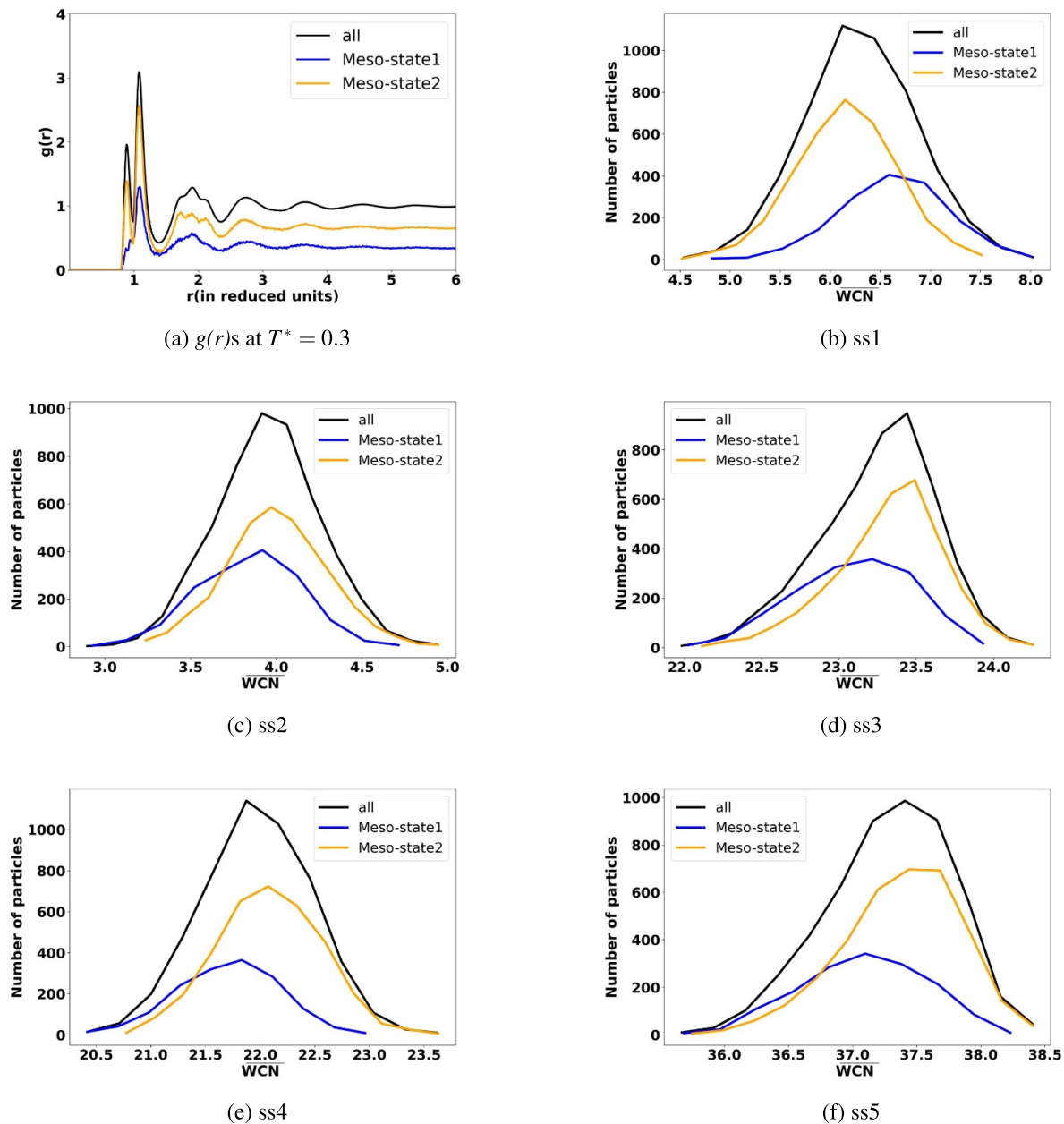


Figure 4. (a) Weighted (according to the number of particles in each meso-state) partial $g(r)$ decomposition of each meso-state at $T^* = 0.3$. Panels (b)–(f) are the weighted \overline{WCN} s distribution for the first five solvation shells of each meso-state and the whole system, respectively (ss stands for solvation shell). Given the classification of particles identities, the average \overline{WCN} distributions along each solvation shell for each meso-state is obtained by averaging over 5 ns. The partial $g(r)$ s in the figure (a) is constructed based on the particle’s identities of each state and scaled by the weight obtained from \overline{WCN} s bimodality along each solvation shell by averaging over 5 ns.

(in real space). In order to clarify the physical interpretation of these nano-domains, it is observed that in the PC-space bimodality of \overline{WCN} s distribution along each solvation shell. Figures 4(b)–(f) provide clear evidence of two meso-states in the PC space, for example the total \overline{WCN} s distributions along first five solvation shells of the system are decomposed into distributions of each individual meso-state as there is a co-existence of two meso-states with different unique local structures. Furthermore, the bimodal distributions of the \overline{WCN} s along all shells in the PC-space can be transformed into a

construction of partial $g(r)$ s in the configurational space as shown in figure 4(a). The total $g(r)$ of the whole system is the summation of the weighted partial $g(r)$ s (blue and orange curves) representing two meso-states. In other words, the classification scheme provides a method to decompose the total $g(r)$ of the system into two partial $g(r)$ s, which represent two different meso-structures whose particles form various domains that tile up the whole configurational space.

Another quantitative measure of these distinct meso-states is to compute density and pressure profiles of the domains.

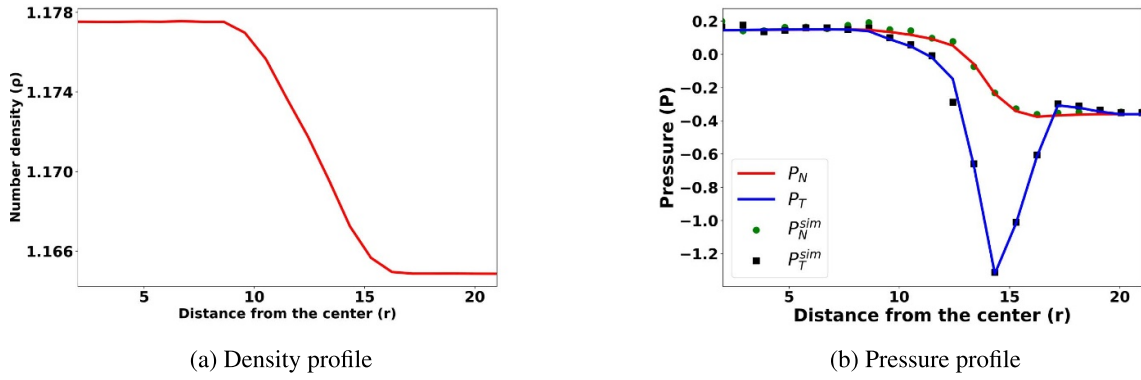


Figure 5. (a) Density profile varying with distance from the center of a domain of meso-state 1 at $T^* = 0.3$. (b) Normal and tangential components of the pressure tensor varying with the distance from the center of a domain of meso-state 1. The green spheres and black squares are normal (P_N^{sim}) and tangential (P_T^{sim}) pressures obtained in the simulation, respectively. The red line is the fit to the normal pressure (P_N) and the blue line is the tangential pressure (P_T) obtained to verify mechanical equilibrium.

Because the shapes of the domains are irregular, the thermodynamic properties of the two meso-states were calculated using a spherical region inside a domain of meso-state 1 and a thin shell in the outermost meso-state 2 region. Figure 5(a) shows the radial distribution of the atomic number density from the center of meso-state 1. Meanwhile, the six components of the pressure tensor ($p_{xx}, p_{yy}, p_{zz}, p_{xy}, p_{yx}, p_{xz}$ and p_{zx}) for each atom are computed in the Cartesian coordinate. The pressure tensor is then transformed into polar coordinate representation whose corresponding components will be ($p_{rr}, p_{\theta\theta}, p_{\phi\phi}, p_{r\theta}, p_{\theta\phi}$ and $p_{\phi r}$). It is noted that the magnitude of the off-diagonal terms is negligible compared to diagonal terms, thus the pressure tensor can be expressed as [65, 66]:

$$P(r) = P_N(r)\mathbf{e}_r\mathbf{e}_r + P_T(r)(\mathbf{e}_\theta\mathbf{e}_\theta + \mathbf{e}_\phi\mathbf{e}_\phi), \quad (2)$$

where \mathbf{e}_r , \mathbf{e}_θ and \mathbf{e}_ϕ are unit vectors, P_N and P_T are the radial or normal and transverse components of the pressure tensor, respectively. The radial profiles of the components $P_N(r)$ and $P_T(r)$ are obtained by integrating out the angular degrees of freedom over thin spherical shells extending outwards from the origin. Figure 5(b) shows the normal (P_N) and tangential (P_T) pressure profiles approximated as a spherical interface within the solid angle of the calculation. It is verified that the normal and tangential profiles satisfy the mechanical equilibrium, $\nabla \cdot \mathbf{P} = 0$, which in spherical coordinates is given by [66, 67]:

$$P_T(r) = P_N(r) + \frac{r}{2} \frac{dP_N(r)}{dr}, \quad (3)$$

where the second term is the derivative of the normal pressure with respect to distance from the center of meso-state 1.

The formation of local spherical interfaces from density and pressure profiles in figure 5 signifies a strong indication for the co-existence of two local distinct meso-states in supercooled states.

3.2. Nature of the nano-domains: dynamics

With our classification scheme, the bimodal decomposition of the $g(r)$, the density and pressure profiles seem to indicate an coexistence of two phases with domain structures after quenching, where similar liquid–liquid phase separation is also observed in a model 2D system with such classification scheme [68]. To further check the validity of such a picture, some dynamical signatures of a liquid–liquid phase separation are evaluated.

First of all, there will be two well separated relaxation time scales in such a scenario. The particles within the domains that belong to the same meso-states should have the same diffusion behavior as they are in the same thermodynamic state. After finding the nano-domains in the configurational space, core particles, the particles stay in that domain during the whole simulation time, in each domain are sorted. Core particles are colored as red and grey for meso-state 1 (blue) and black and green for meso-state 2 (orange) as shown in figures 6(a) and (b). 2D cross section of core particles is taken for the purpose of visualization. Collected core particles from each domain are then used to compute mean-square displacements to get diffusion constant by Einstein relation. Figures 6(c) and (d) show different diffusion constant distribution of different domains at different temperatures, hence supports the picture that the domains that belong to the same meso-state have the same diffusion behavior.

In the section 2.2, the stability of nano-domains is associated with the onset of cage-breaking processes which is reflected as a plateau in the self intermediate coherent function $F(k, t)$ or in diffusion dynamics via MSD (figure 1(b)). The timescale of the cage processes depends on temperature, a quantitative study will interesting, but some qualitative observations can still be made. Given the glass transition temperature being $T_g^* = 0.25$ for this system, different configurational snapshots can be used to qualitatively examine timescale of nano-domains. Figure 7 shows three different 10 ns lag

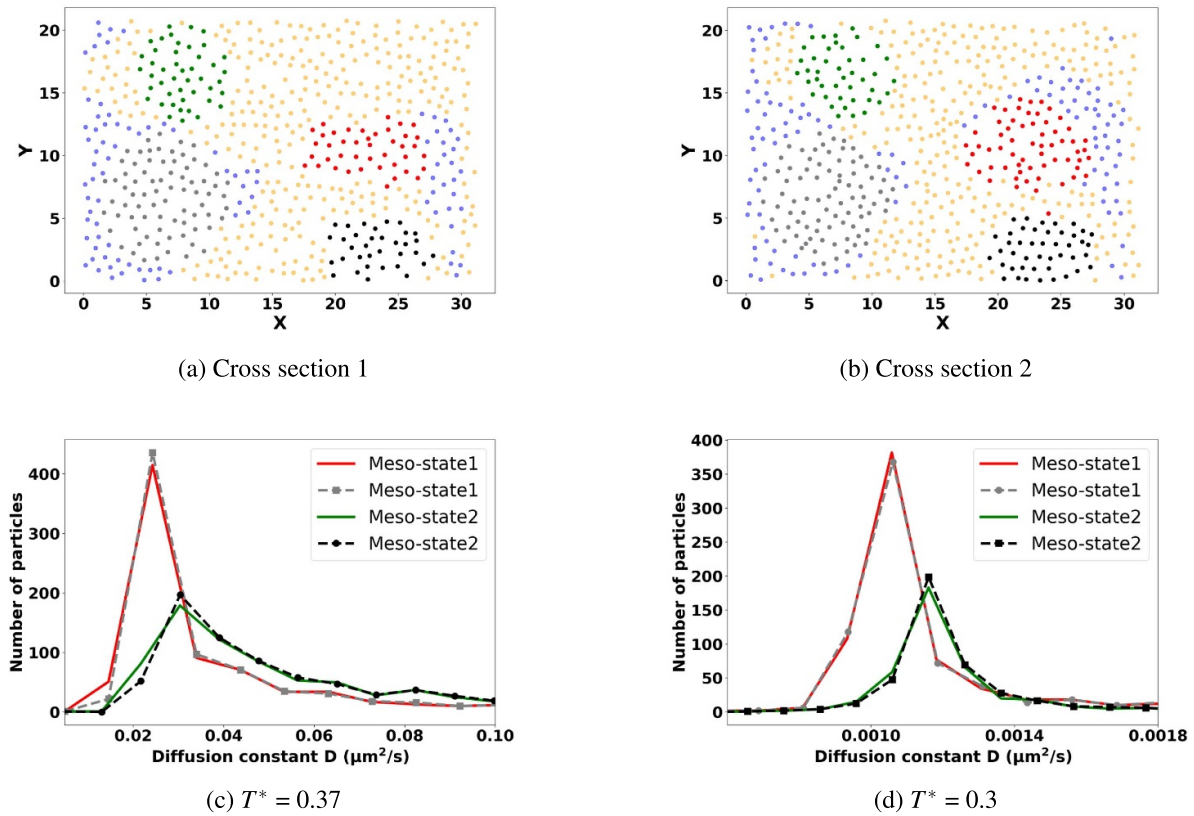


Figure 6. (a), (b) 2D representation of sorted cores particles of meso-states in bifurcated domains of a 16 000 particles system at $T^* = 0.3$. (c), (d) Distribution of diffusion constants of core particles at different temperatures.

time snapshots which are taken at three different temperatures. At $T^* = 0.2$ which is below T_g^* , almost all of particles are immobile and freeze at their local domains, so lifetime of nano-domains are indefinitely long. Meanwhile, as temperature goes up, more particles are able to escape out the cage as illustrated from $T^* = 0.3$ to $T^* = 0.37$ in figure 7, hence nano-domain shapes are changing relatively quickly and become less static. These phenomena are confirmed by quantitatively evaluating particles fluctuation of the domains as shown in figures 8(a)–(c). The magnitude of particles fluctuations increases as the temperature increases because of higher number of mobile particles.

Another signature of the liquid–liquid phase separation that follows the quenching from a high temperature equilibrium (normal liquid) state to a super-cooled state is the scaling law of the domain size growth [69–71]. In this case, the two meso-states are the equilibrium thermodynamic states with domains formed either via spinodal decomposition or nucleation such as shown in figure 3(d) [68, 72]. It is well-established that the growth of characteristic domain size follows an algebraic growth law in time [69–71] $L(t) \sim t^{1/3}$ for conserved scalar order parameters (even though the A/B particles are treated as the same in our classification, but the growth dynamics still follows the conserved order parameter scaling law as the real dynamics is still constrained by the swapping of A/B identity) [73–77] which can be tested

by the calculation of equal-time correlation function $C(r, t)$ from our classification. Considering a scalar order parameter ψ , the equal-time correlation is: $C(r, t) = N^{-1} \sum_i \psi_i(t) \psi_{i+r}(t)$ where N is the number of particles, $i + r$ indicates a neighboring particle displaced by a distance r relative to the reference particle i with $\psi_i = +1$ for particles identities of state 1 and $\psi_i = -1$ for particles identities of state 2, hence the product of $\psi_i(t) \psi_{i+r}(t)$ will be $+1$ between pair of particles from the same state and will be -1 otherwise. Since the domain identities for each particle are known from the classification scheme, the result of $C(r, t)$ indeed confirms the scaling law of domains growth $L(t) \sim t^{1/3}$ as shown in figure 8(d).

Finally, similar to the 2D case [68], the number of domains belonging to the same meso-state will increase to tile up the whole system as the system size increases. However, 3D domains are hard to visualize, so their spatial structures are illustrated by taking different 2D cross sections at different configurations. Figures 9(a) and (c) shows two different cross sections of two domains that belong to the same meso-state 1. These two domains are also shown by taking a second snapshot at 2 ns later in figures 9(b) and (c). As the system size continues to increase, the meso-state1 will split into more domains as shown in figure 10 for 50 000 particles system. It should be emphasized that the structure of a domain might disappear in some regions of the space along different cross sections

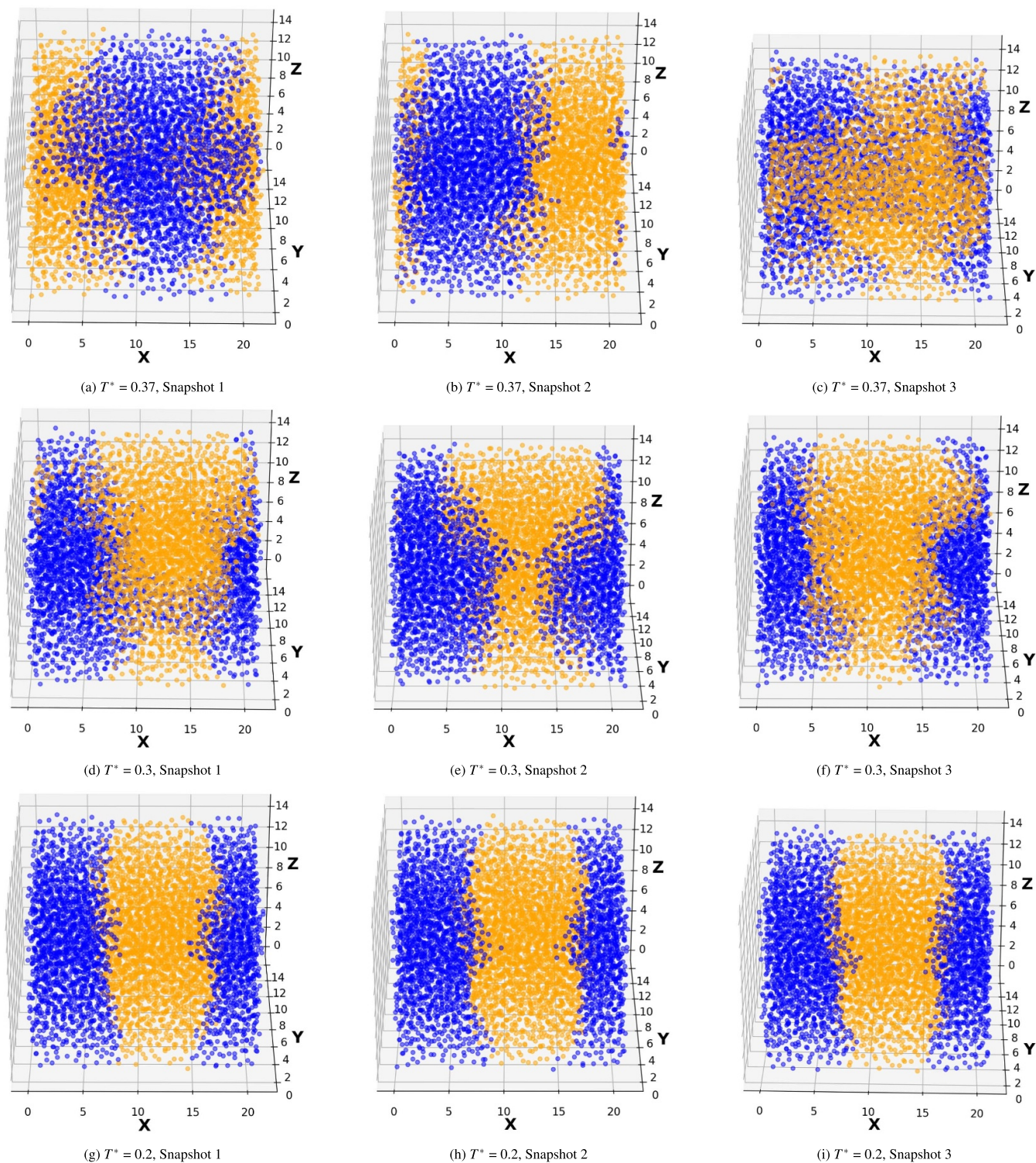


Figure 7. 3D snapshots of meso-states at different temperatures. The snapshots are at every 10 ns lag time.

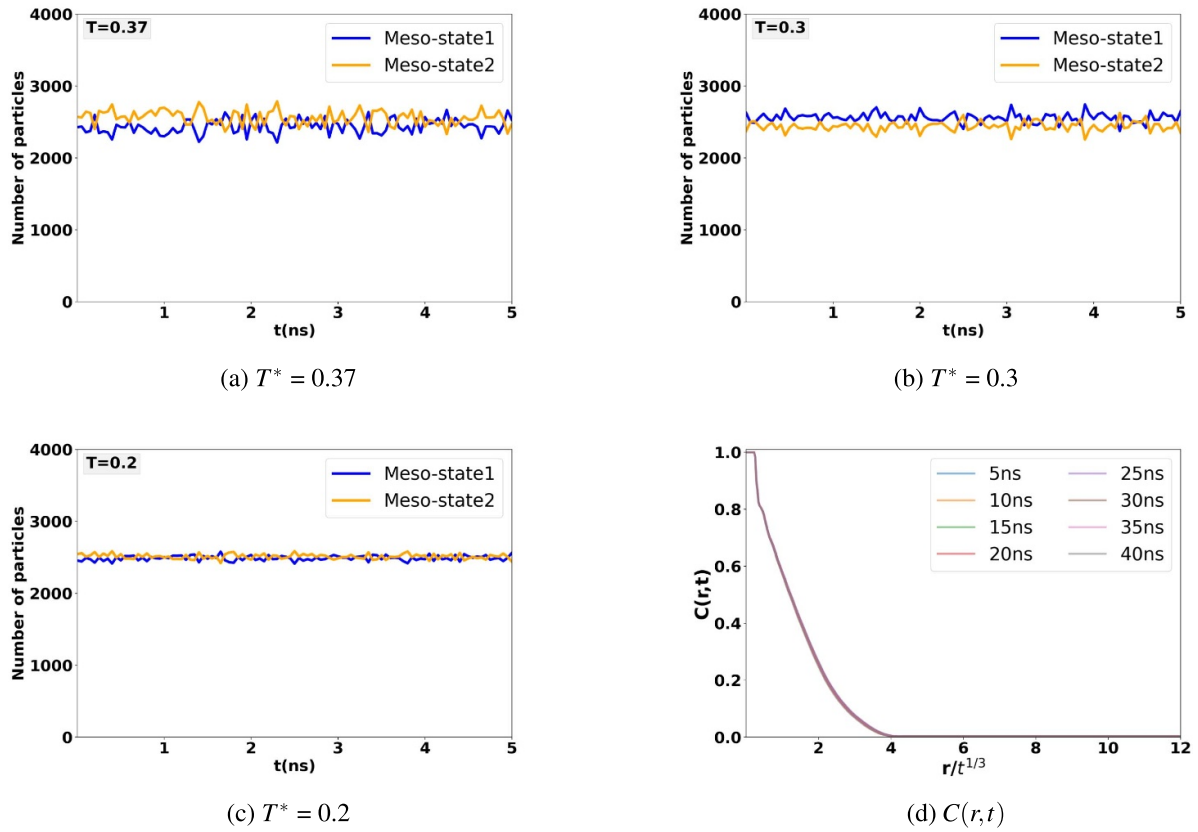


Figure 8. (a)–(c) Particle number fluctuations of meso-states at different temperatures over 5 ns. (d) Equal-time correlation $C(r,t)$ at $T^* = 0.2$.

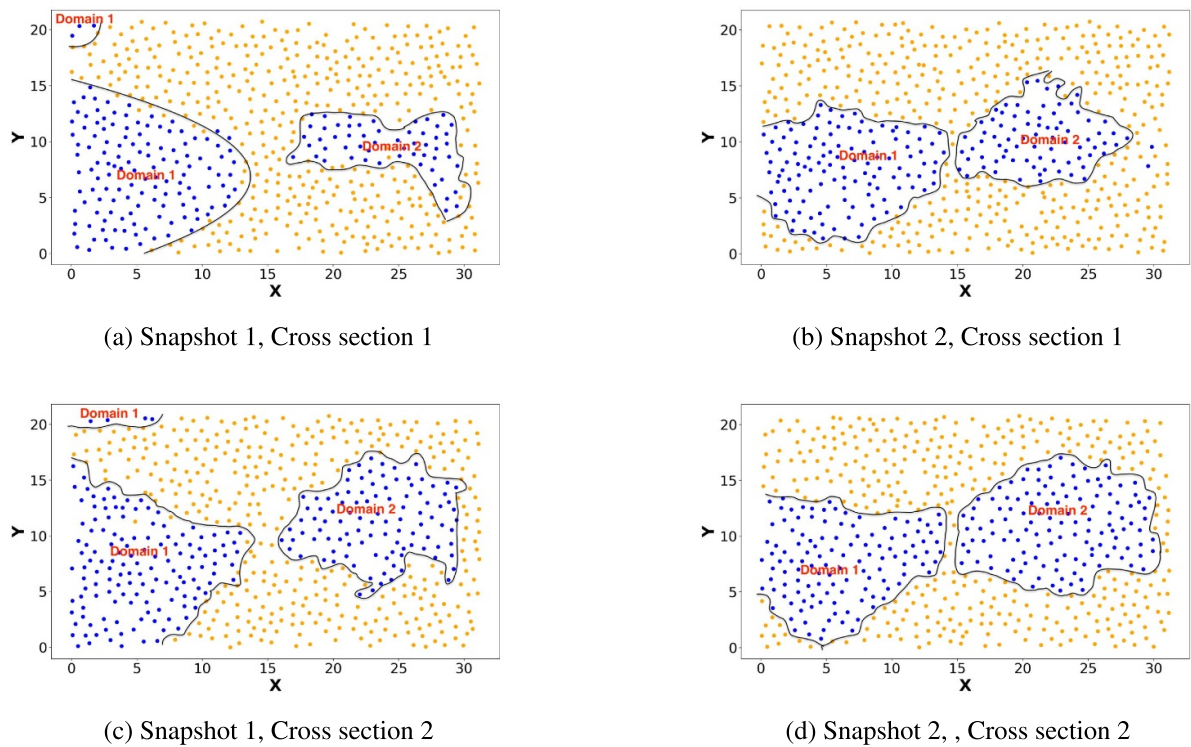


Figure 9. 2D representation of bifurcated domains for meso-states of 16 000 particles system at $T^* = 0.2$, Panels (a), (c) and (b), (d) are two different cross section along Z-direction of two different snapshots. The two snapshots are at 2 ns lag time. The green lines are an guide to the eye for the boundaries among bifurcated domains and the numbers represent the various domains of a specific meso-state.

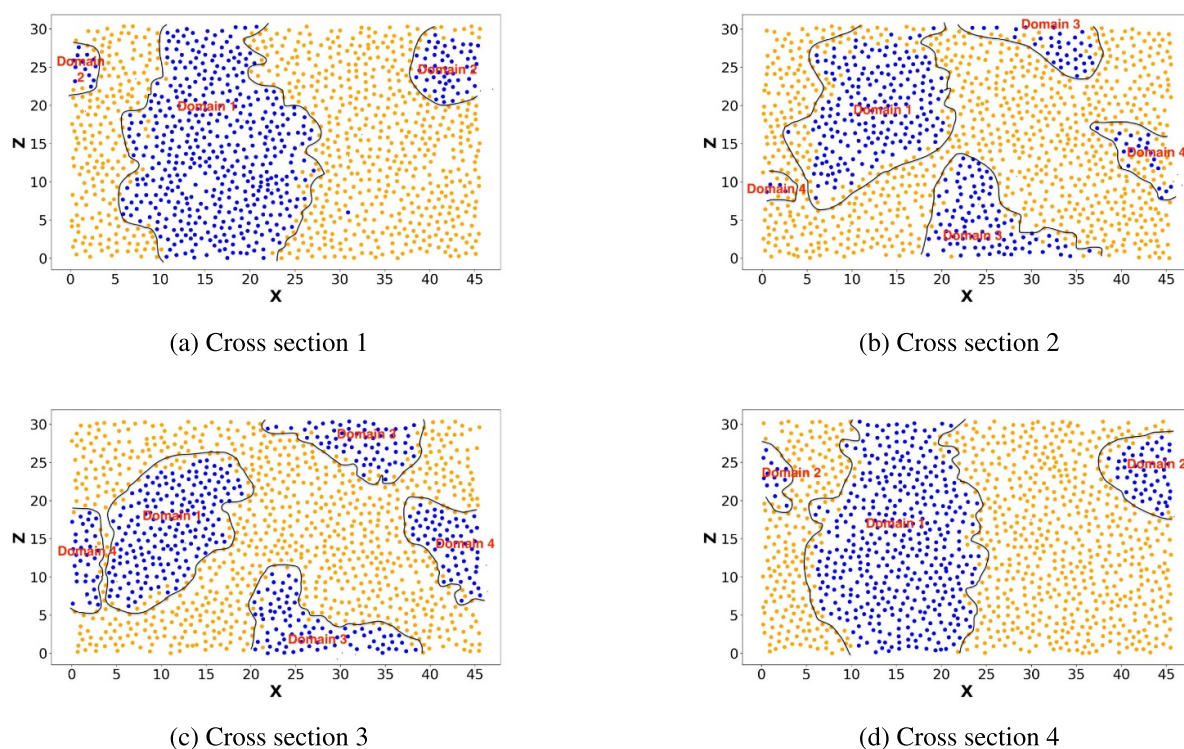


Figure 10. 2D cross sections along Y -axis to represent bifurcated domains of meso-states of 50 000 particles system at $T^* = 0.2$. The green lines are an guide to the eye for the boundaries among bifurcated domains and the numbers represent the various domains of a specific meso-state.

as happened to the bifurcated domain 2,3,4 to illustrate the finiteness of each domain in 3D configurational space.

4. Concluding remarks

ML methods are used to develop a scheme to identify spatially co-existing meso-states or nano-domains both structurally and configurationally. The physical interpretation of these meso-states are explicitly demonstrated by the observation of bimodality of \overline{WCN} s distribution along each solvation shell, the corresponding construction of weighted partial $g(r)$; the formation of interfaces from calculations of the pressure and density profiles. Given the classification, heterogeneous dynamics of these nano-domains are captured by the difference in the collective distribution of diffusion constants; spatial characterization of these nano-domains is used to evaluate their lifetimes to understand of cage effect for longer relaxation dynamics. Furthermore, kinetic domain growth scaling law calculation presents a direct evidence to indicate that such domains are the result of liquid-liquid phase separation when the system is at supercooled condition from quenching.

Using the classification scheme developed in this report, the L-L phase separation behaviors can be studied in details. The observed domain structures provide a natural molecular realization of the Adam-Gibbs' Cooperative Rearranging Regions or the mosaic picture of ROFT. These domain structures naturally lead to two types of relaxation dynamics, the intra-domain relaxation is largely due to diffusion inside a domain

and the inter-domain relaxation which is related to the coarsening kinetics of the first-order phase transitions. Therefore, the classification scheme provides a platform for further extensive statistical mechanics analysis of supercooled liquids.

Data availability statement

The data that supports the findings of this study are openly available in DataShare, Iowa State University's Open Data Repository, at <https://doi.org/10.25380/iastate.23925933>.

Acknowledgments

This work is supported by the Division of Chemical and Biological Sciences, Office of Basic Energy Sciences, U.S. Department of Energy, under Contact No. DE-AC02-07CH11358 with Iowa State University.

Appendix A. Selection of $K = 2$

In the main text, the Elbow test cannot determine an effective K value for initial K -means clustering but a possible range of K clusters. For a selection of $K = 2$ as described in the main text, we first constructed K -means clustering models with various numbers of K (from 2 to 4) in the PC-space as shown in figures A1(a), (c), (e) and in the configurational space by direct mapping in figures A2(a), (c) and (e). For all models of

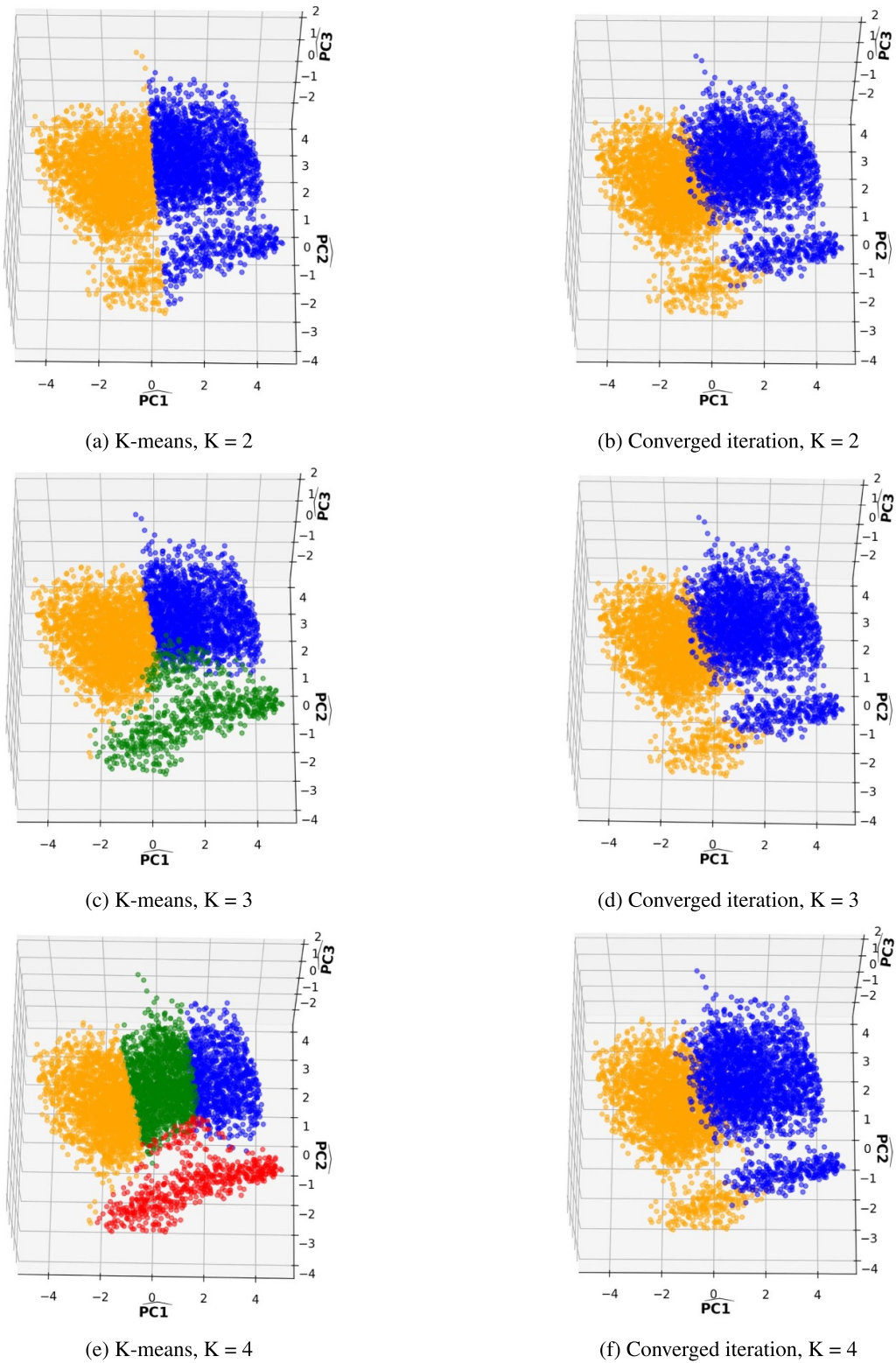


Figure A1. Comparison for different K values for initial K -means and after converged iteration of 5000 particles in PC-space at $T^* = 0.2$.

K -means clustering in the previous step, we then performed the co-learning strategy for each $K = 2, 3, 4$ to sort out the one K that all models of K -means converge in both the PC and real space. Final results presented in both PC (figures A1(b),

(d) and (f) and configurational space (figures A2(b), (d) and (f)) show convergence for $K = 2$ for all K -means models, thus support our $K = 2$ choice. Physically for an one-component system the Gibbs phase rule will lead to $K = 2$ as well.

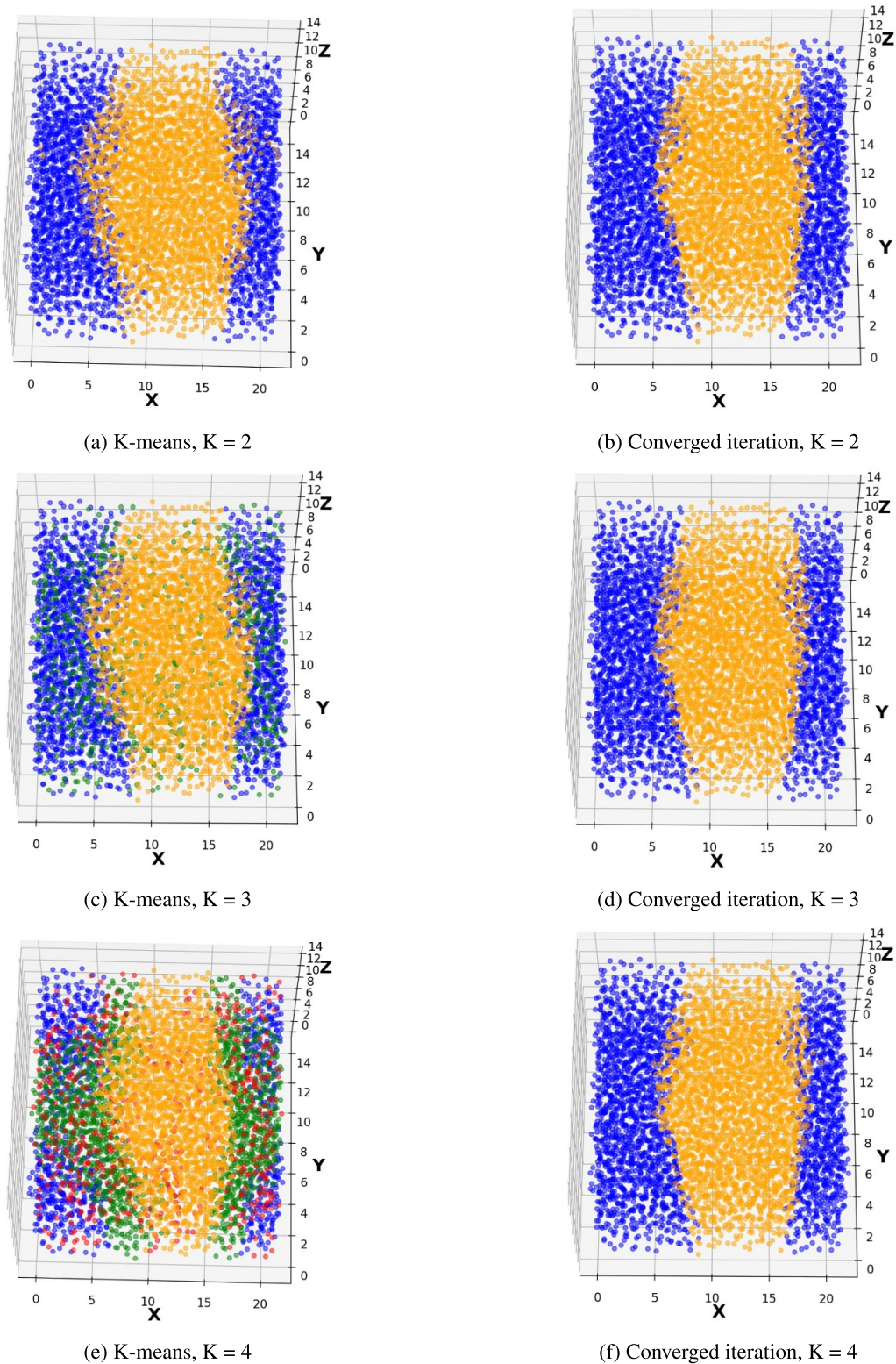


Figure A2. Comparison for different K values for initial K -means and after converged iteration of 5000 particles in real space at $T^* = 0.2$.

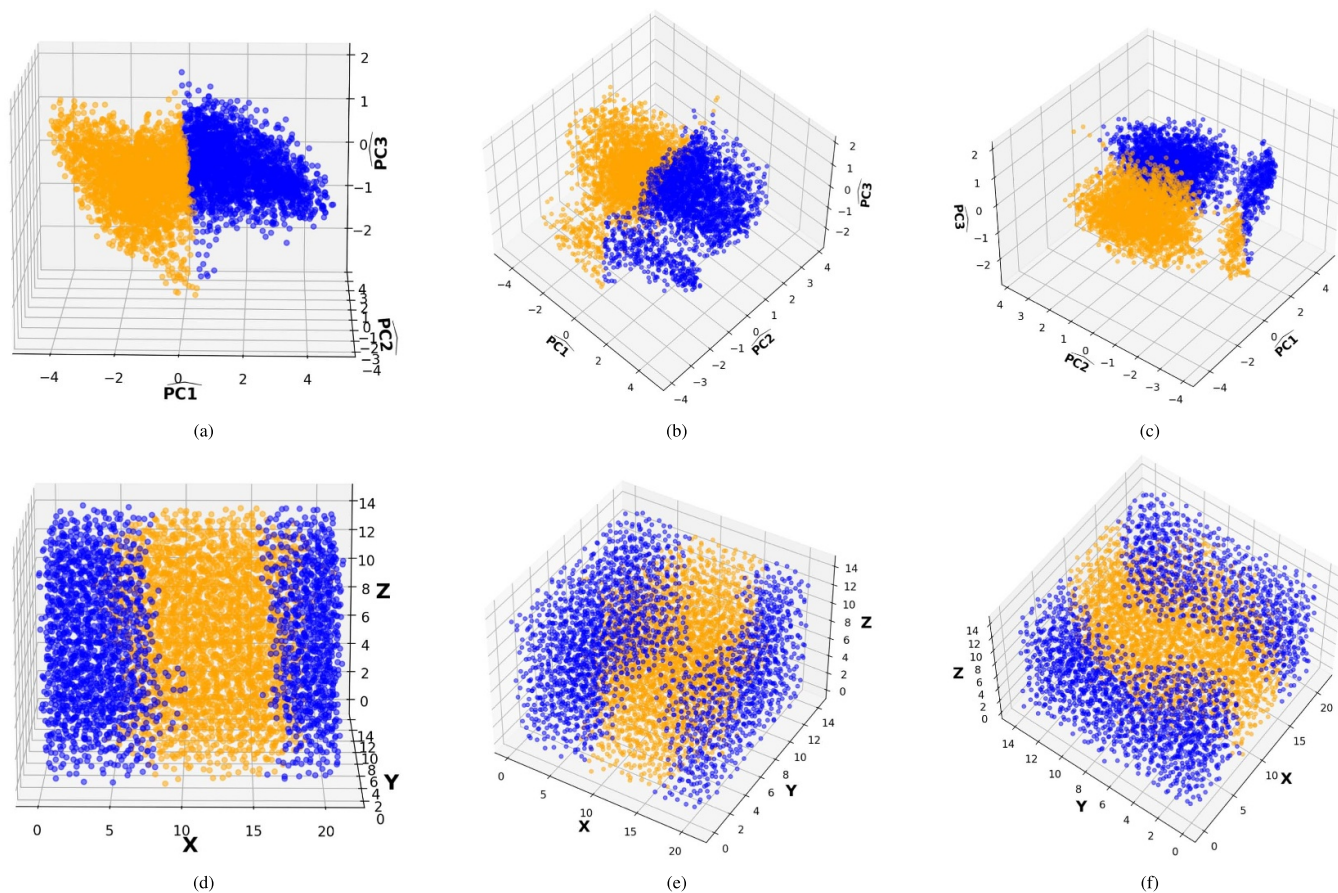


Figure B1. Different angles projection to display the local meso-states structure in PC and configurational spaces from the figures 2(a) and (b) in the main text. Panels (a)–(c) are for PC-space and (d)–(f) are for the configurational space.

Appendix B. Angle projection to visualize the meso-state structure

In addition to the figures in the main text, in figure B1 different angle projections of figures 2(a) and (b) are presented here to provide different view for the domain structure.

Appendix C. Classification of A/B particles type

In the main text, the identity of particles’ type (A/B) is ignored when collecting \overline{WCNs} for the classification of particles into meso-states. Indeed each meso-state consists of a mixture of A and B particles shown in figure C1.

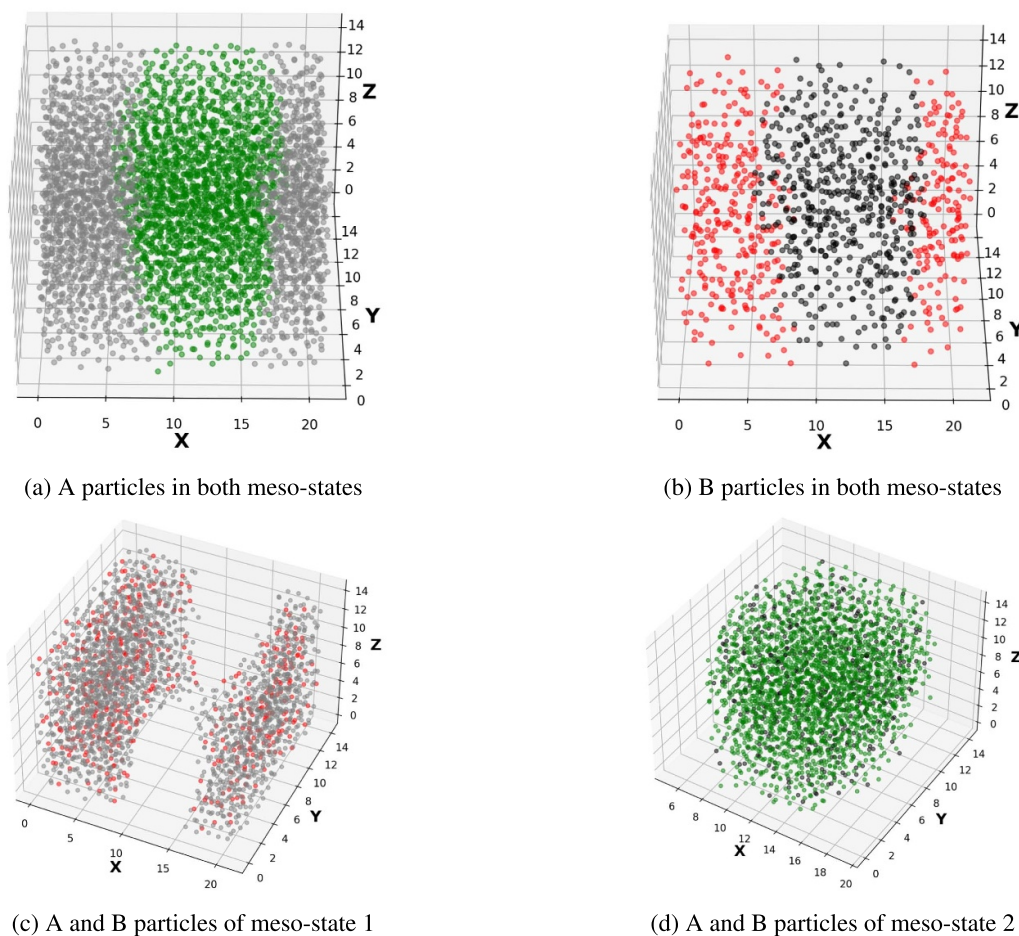


Figure C1. A 3D projection of classified A/B particles in configurational space of 5000 particles system at $T^* = 0.2$. Panels (a) and (b) are classified A or B particles in both meso-states while panels (c) and (d) are A and B particles of a particular meso-state. Grey and red are A and B particles belonging to meso-state 1, green and black are A and B particles of meso-state 2. Meso-state 1 has 2052 A particles and 459 B particles while the meso-state 2 has 1948 A particles and 541 B particles, so the ratio of A/B particles in each meso-state is roughly 4.47 and 3.60, respectively, which is different from the bulk one.

ORCID iDs

Viet Nguyen  <https://orcid.org/0000-0003-2961-3217>

Xueyu Song  <https://orcid.org/0000-0001-5142-4223>

References

- [1] Crowe J H, Carpenter J F and Crowe L M 1998 *Annu. Rev. Physiol.* **60** 73
- [2] Debenedetti P G, Truskett T M, Lewis C P and Stillinger F H 2001 *Advances in Chemical Engineering* (Academic) pp 21–79
- [3] Sastry S, Debenedetti P G and Stillinger F H 1998 *Nature* **393** 554
- [4] Andersen H C 2005 *Proc. Natl Acad. Sci.* **102** 6686
- [5] Kob W, Donati C, Plimpton S J, Poole P H and Glotzer S C 1997 *Phys. Rev. Lett.* **79** 2827
- [6] Gotze W and Sjogren L 1992 *Rep. Prog. Phys.* **55** 241
- [7] Cubuk E D, Schoenholz S S, Rieser J M, Malone B D, Rottler J, Durian D J, Kaxiras E and Liu A J 2015 *Phys. Rev. Lett.* **114** 108001
- [8] Stillinger F H 1995 *Science* **267** 1935
- [9] Smessaert A and Rottler J 2013 *Phys. Rev. E* **88** 022314
- [10] Candelier R, Widmer-Cooper A, Kummerfeld J K, Dauchot O, Biroli G, Harrowell P and Reichman D R 2010 *Phys. Rev. Lett.* **105** 135702
- [11] Kawasaki T and Tanaka H 2014 *Phys. Rev. E* **89** 062315
- [12] Yang X, Liu R, Yang M and Chen K 2016 *Phys. Rev. Lett.* **116** 238003
- [13] Doliwa B and Heuer A 1998 *Phys. Rev. Lett.* **80** 4915
- [14] Glotzer S C 2000 *J. Non-Cryst. Solids* **274** 342
- [15] Sillescu H 1999 *J. Non-Cryst. Solids* **243** 81
- [16] Ediger M D 2000 *Annu. Rev. Phys. Chem.* **51** 99
- [17] Royall C P and Williams S R 2015 *Phys. Rep.* **560** 1
- [18] Cavagna A 2009 *Phys. Rep.* **476** 51
- [19] Donati C, Douglas J F, Kob W, Plimpton S J, Poole P H and Glotzer S C 1998 *Phys. Rev. Lett.* **80** 2338
- [20] Adam G and Gibbs J H 1965 *J. Chem. Phys.* **43** 139
- [21] Vidal Russell E and Israeloff N E 2000 *Nature* **408** 695
- [22] Berthier L, Biroli G, Bouchaud J-P, Cipelletti L, Masri D E, L'Hôte D, Ladieu F and Pierno M 2006 *Science* **310** 1797
- [23] Kirkpatrick T R, Thirumalai D and Wolynes P G 1989 *Phys. Rev. A* **40** 1045
- [24] Viot P, Tarjus G and Kivelson D 2000 *J. Chem. Phys.* **112** 10368
- [25] Garrahan J P and Chandler D 2002 *Phys. Rev. Lett.* **89** 035704
- [26] Hurley M M and Harrowell P 1995 *Phys. Rev. E* **52** 1694
- [27] Bennemann C, Donati C, Baschnagel J and Glotzer S C 1999 *Nature* **399** 246

- [28] Donati C, Franz S, Glotzer S C and Parisi G 2002 *J. Non-Cryst. Solids* **307–310** 215
- [29] Whitelam S, Berthier L and Garrahan J P 2004 *Phys. Rev. Lett.* **92** 185705
- [30] Berthier L 2004 *Phys. Rev. E* **69** 020201
- [31] Goldstein M 1969 *J. Chem. Phys.* **51** 3728
- [32] Berthier L and Biroli G 2011 *Rev. Mod. Phys.* **83** 587
- [33] Gibbs J H and DiMarzio E A 1958 *J. Chem. Phys.* **28** 373
- [34] Bouchaud J-P and Biroli G 2004 *J. Chem. Phys.* **121** 7347
- [35] Bernal J D 1959 *Nature* **183** 141
- [36] Bernal J D 1960 *Nature* **185** 68
- [37] Finney J 1970 *Proc. R. Soc. A* **319** 479–93
- [38] Anikeenko A V and Medvedev N N 2007 *Phys. Rev. Lett.* **98** 235504
- [39] Anikeenko A V, Medvedev N N and Aste T 2008 *Phys. Rev. E* **77** 031101
- [40] Steinhardt P J, Nelson D R and Ronchetti M 1983 *Phys. Rev. B* **28** 784
- [41] Lechner W and Dellago C 2008 *J. Chem. Phys.* **129** 114707
- [42] Tsuzuki H, Brancio P S and Rino J P 2007 *Comput. Phys. Commun.* **177** 518
- [43] Faken D and Jónsson H 1994 *Comput. Mater. Sci.* **2** 279
- [44] Honeycutt J D and Andersen H C 1987 *J. Phys. Chem.* **91** 4950
- [45] Williams S R 2007 arXiv:0705.0203
- [46] Malins A, Williams S R, Eggers J and Royall C P 2013 *J. Chem. Phys.* **139** 234506
- [47] Dunleavy A J, Wiesner K, Yamamoto R and Royall C P 2015 *Nat. Commun.* **6** 6089
- [48] Shannon C E 1948 *Bell Syst. Tech. J.* **27** 379
- [49] Dunleavy A J, Wiesner K and Royall C P 2012 *Phys. Rev. E* **86** 041505
- [50] Ronhovde P, Chakrabarty S, Hu D, Sahu M, Sahu K K, Kelton K F, Mauro N A and Nussinov Z 2012 *Sci. Rep.* **2** 329
- [51] James G, Witten D, Hastie T and Tibshirani R 2014 *An Introduction to Statistical Learning with Application in R (Springer Texts in Statistics)* (Springer)
- [52] Scherer M K, Trendelkamp-Schroer B, Paul F, Pérez-Hernández G, Hoffmann M, Plattner N, Wehmeyer C, Prinz J-H and Noé F 2015 *J. Chem. Theory Comput.* **11** 5525
- [53] Bishop C 2006 *Pattern Recognition and Machine Learning* (Springer)
- [54] Murphy K P 2012 *Machine Learning: A Probabilistic Perspective* (The MIT Press)
- [55] Kob W and Andersen H C 1995 *Phys. Rev. E* **51** 4626
- [56] Kob W and Andersen H C 1994 *Phys. Rev. Lett.* **73** 1376
- [57] Middleton T F and Wales D J 2001 *Phys. Rev. B* **64** 024205
- [58] Thompson A P *et al* 2022 *Comput. Phys. Commun.* **271** 108171
- [59] Montero de Hijes P, Espinosa J R, Bianco V, Sanz E and Vega C 2020 *J. Phys. Chem. C* **124** 8795
- [60] Bai X-M and Li M 2006 *J. Chem. Phys.* **124** 124707
- [61] Janssen L M C 2018 *Front. Phys.* **6** 97
- [62] Schröder T B and Dyre J C 2020 *J. Chem. Phys.* **152** 141101
- [63] Rudzinski J F, Radu M and Bereau T 2019 *J. Chem. Phys.* **150** 024102
- [64] Shlens J 2014 *Educational* **51** (arXiv:1404.1100)
- [65] Gunawardana K G S H and Song X 2018 *J. Chem. Phys.* **148** 204506
- [66] Rowlinson J S and Widom B 2013 *Molecular Theory of Capillarity* (Courier Corporation)
- [67] Ballal D, Lu Q, Raju M and Song X 2019 *J. Chem. Phys.* **151** 134108
- [68] Nguyen V and Song X 2023 arXiv:2301.05990
- [69] Aranson I 2011 *J. Stat. Phys.* **142** 220
- [70] Binder K 1975 Dynamics of first order phase transitions *Fluctuations, Instabilities and Phase Transitions* ed T Riste (Springer) pp 53–86
- [71] Humayun K and Bray A J 1991 *J. Phys. A: Math. Gen.* **24** 1915
- [72] Brickley J, Nguyen V and Song X 2023 Unpublished results
- [73] Bray A J, Humayun K and Newman T J 1991 *Phys. Rev. B* **43** 3699
- [74] Mazenko G F 1990 *Phys. Rev. B* **42** 4487
- [75] Mazenko G F, Valls O T and Zannetti M 1988 *Phys. Rev. B* **38** 520
- [76] Liu F and Mazenko G F 1991 *Phys. Rev. B* **44** 9185
- [77] Bray A J 1990 *Phys. Rev. B* **41** 6724