

Physical information-enhanced graph neural network for predicting phase separation

Yaqiang Zhang(张亚强), Xuwen Wang(王煦文), Yanan Wang(王雅楠), and Wen Zheng(郑文)

Citation: Chin. Phys. B, 2024, 33 (**7**): 070702. DOI: 10.1088/1674-1056/ad4328
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

TCAS-PINN: Physics-informed neural networks with a novel temporal causality-based adaptive sampling method

Jia Guo(郭嘉), Haifeng Wang(王海峰), Shilin Gu(古仕林), and Chenping Hou(侯臣平)
Chin. Phys. B, 2024, 33 (**5**): 050701. DOI: 10.1088/1674-1056/ad21f3

Advances in neuromorphic computing: Expanding horizons for AI development through novel artificial neurons and in-sensor computing

Yubo Yang(杨玉波), Jizhe Zhao(赵吉哲), Yinjie Liu(刘胤洁), Xiayang Hua(华夏扬), Tianrui Wang(王天睿), Jiyuan Zheng(郑纪元), Zhibiao Hao(郝智彪), Bing Xiong(熊兵), Changzheng Sun(孙长征), Yanjun Han(韩彦军), Jian Wang(王健), Hongtao Li(李洪涛), Lai Wang(汪莱), and Yi Luo(罗毅)
Chin. Phys. B, 2024, 33 (**3**): 030702. DOI: 10.1088/1674-1056/ad1c58

Recent advances in protein conformation sampling by combining machine learning with molecular simulation

Yiming Tang(唐一鸣), Zhongyuan Yang(杨中元), Yifei Yao(姚逸飞), Yun Zhou(周运), Yuan Tan(谈圆),Zichao Wang(王子超), Tong Pan(潘瞳), Rui Xiong(熊瑞), Junli Sun(孙俊力), and Guanghong Wei(韦广红)
Chin. Phys. B, 2024, 33 (**3**): 030701. DOI: 10.1088/1674-1056/ad1a92

Crysformer: An attention-based graph neural network for properties prediction of crystals

Tian Wang(王田), Jiahui Chen(陈家辉), Jing Teng(滕婧), Jingang Shi(史金钢),Xinhua Zeng(曾新华), and Hichem Snoussi
Chin. Phys. B, 2023, 32 (**9**): 090703. DOI: 10.1088/1674-1056/ace247

Differentiable programming and density matrix based Hartree-Fock method

Hong-Bin Ren(任宏斌), Lei Wang(王磊), and Xi Dai(戴希)
Chin. Phys. B, 2021, 30 (**6**): 060701. DOI: 10.1088/1674-1056/abeeed

---

Physical information-enhanced graph neural network for predicting phase separation

Yaqiang Zhang(张亚强)[1], Xuwen Wang(王煦文)[1], Yanan Wang(王雅楠)[1], and Wen Zheng(郑文)[1,2,3,†]

[1] *Institute of Public-Safety and Big Data, College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Jinzhong* 030600, *China*

[2] *Shanxi Engineering Research Centre for Intelligent Data Assisted Treatment, Changzhi Medical College, Changzhi* 046000, *China*

[3] *Innovation Academy for Microsatellites of Chinese Academy of Sciences, Shanghai* 200050, *China*

(Received 3 March 2024; revised manuscript received 8 April 2024; accepted manuscript online 25 April 2024)

Although phase separation is a ubiquitous phenomenon, the interactions between multiple components make it difficult to accurately model and predict. In recent years, machine learning has been widely used in physics simulations. Here, we present a physical information-enhanced graph neural network (PIENet) to simulate and predict the evolution of phase separation. The accuracy of our model in predicting particle positions is improved by 40.3% and 51.77% compared with CNN and SVM respectively. Moreover, we design an order parameter based on local density to measure the evolution of phase separation and analyze the systematic changes with different repulsion coefficients and different Schmidt numbers. The results demonstrate that our model can achieve long-term accurate predictions of order parameters without requiring complex handcrafted features. These results prove that graph neural networks can become new tools and methods for predicting the structure and properties of complex physical systems.

Keywords: graph neural network, phase separation, machine learning, dissipative particle dynamics

PACS: 07.05.Mh, 64.75.Gh, 83.10.Rs, 87.64.Aa **DOI:** 10.1088/1674-1056/ad4328

1. Introduction

Phase separation, the spontaneous separation of components of a binary mixture, is a complex nonlinear process. In the last decades, there have been a large number of studies demonstrating the importance of phase separation in oil-water systems,[1] self-propelled colloids,[2–4] cellular components separation,[5,6] and liquid crystal systems.[7–9] Although phase separation is a ubiquitous phenomenon, the interaction and kinetic changes between complex environments are still difficult in the study of phase separation. Therefore, it is necessary to comprehensively use various methods such as experiments, simulations, and theories to promote the in-depth development of phase separation research.

Advances in science and technology have facilitated the study of phase separation at the microscopic level.[10–15] Researchers use advanced experimental and simulation techniques to investigate the phase separation process, including molecular dynamics (MD),[16–18] dissipative particle dynamics (DPD),[19] and Navier–Stokes (NS) equations. However, traditional simulation methods are not only expensive but also usually have to be set for short time scales to obtain high accuracy, which makes it more difficult to extend the study. In recent decades, the rapid development of machine learning has provided a new way of thinking for the realistic modeling of phase separation processes.[20–25]

Since the past decade, machine learning has been used to solve classic problems in complex physical systems.[26–34] In particular, the graph network-based simulators show extraordinary predictive capabilities for physical systems such as fluids and rigid solids, thanks to the ability of graph networks to effectively capture relational information and structural properties in the system.[35] At the same time, researchers have successfully used the features of GNNs to predict the long-term dynamics of glass systems.[36,37] Among other things, the graph neural network accurately predicted where force chains would arise upon the deformation of jammed disordered solids.[24] These results demonstrate that graph neural networks show very high predictive power in millions of molecular dynamics simulation steps, which makes the combination of machine learning and phase separation simulation a new research hotspot.

In this work, we use graph neural networks to simulate and predict phase separation processes. Specifically, we design a graph neural network model enhanced with physical information. By integrating physical features, we achieve precise predictions of particle positions. Our method exhibits significant advantages compared to classical neural network models. In addition, we also define an order parameter based on local density to describe the spatial distribution of phase separation. Experimental results show that this order parameter can effectively describe the phase separation phenomenon under different repulsion coefficients and Schmidt numbers. We applied data obtained from phase separation simulations to PIENet to achieve long-term predictions with different re-

[†]Corresponding author. E-mail: zhengwen@tyut.edu.cn

© 2024 Chinese Physical Society and IOP Publishing Ltd

pulsion coefficients. These results show that our GNN model can deeply understand the distance changes between particles, conservation forces, and the dynamics of the phase separation process from the particle positions and types at the initial moment.

2. Simulation methods and details

2.1. The PIENet model

Graph neural networks combined with relational inference models are a powerful way to discover hidden structures and physical laws from the evolution of physical locations,[35,36] with atoms corresponding to nodes and bonds corresponding to edges. Here we improve the graph network (PIENet) using physical information.

We adapt an encoder–decoder network, which has been successfully used to predict the properties of glass. Figure 1 illustrates the framework of the PIENet, which takes into account the type and position of particles as input. The model assigns a node to each particle and adds edges between particles within a "connectivity radius" to build the graph structure G. Since the computation time scales linearly with the connectivity radius, we chose a cutoff radius for particle interactions. The graph G is represented by a triple $G=(V,E,u)$, V is the set of nodes and E is the set of edges. The graph-level feature, u, can represent global properties such as temperature. For particle i and particle j, we calculate the Euclidean distance e_{ij} between particles. Only particle pairs whose Euclidean distance is smaller than the cutoff radius are selected as characteristic particles. Thus, the node features in this graph are the particle types and particle positions of all particles at the initial moment. The distance between all feature particle pairs and the conservative force calculated based on Euclidean distance as the features of edges.

During the update process of the graph, the function of the processor consists of two parts: the update function and the aggregation function. The update function updates edge, node, and global features based on connected edges and nodes. The aggregation function aggregates the learned features and passes them through the decoder to obtain a representation of the underlying graph. The update functions are defined as

$$e' = F^{\mathrm{e}}\left(e,v,u\right), \tag{1}$$
$$v' = F^{\mathrm{v}}\left(e',v,u\right), \tag{2}$$
$$u' = F^{\mathrm{u}}\left(e',v',u\right). \tag{3}$$

Aggregation function:

$$\bar{e}' = \rho^{\mathrm{v}\to\mathrm{e}}\left(V'\right), \tag{4}$$
$$\bar{v}' = \rho^{\mathrm{e}\to\mathrm{v}}\left(E'\right), \tag{5}$$
$$\bar{u}' = \rho^{\mathrm{e}\to\mathrm{u}}\left(E',V'\right). \tag{6}$$

Here F^{e} is a function that updates edges based on edge, node, and global features, F^{v} and F^{u} are node and global update functions. $\rho^{\mathrm{v}\to\mathrm{e}}$ and $\rho^{\mathrm{e}\to\mathrm{v}}$ are the aggregation function of all edges and the aggregation function of all nodes, and based on this, a global feature representation is generated. As shown in Fig. 1, during the training process, the features of the nodes of the particles will be encoded directly, and the network encodes and updates edge features and merges them using two dynamic parameters (α and β). These parameters are automatically optimized by the network as it undergoes training. The graph G computes interactions among nodes via k steps (a key hyperparameter) of learned interaction of forces, to generate a sequence of updated latent graphs. The conserved forces between the pairs of feature particles as the features of the edges will be updated n times at the same time. Finally, the decoder will generate the output features at the nodes and edges to predict the positions of the particles.

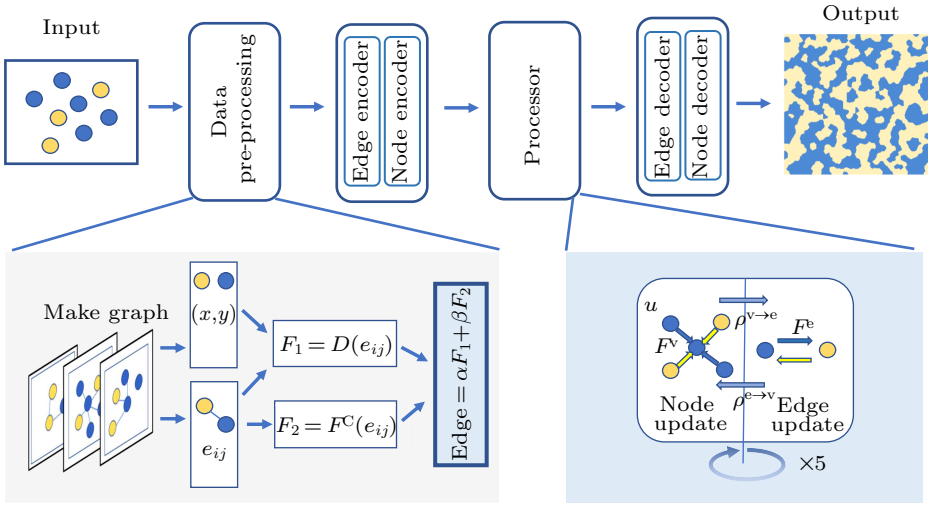


Fig. 1. Neural network framework of PIENet. The model input consists of particle positions, and node features and edge features are processed in the data pre-processing module. In the processor, node features, edge features, and global features are updated respectively. Finally, the model outputs the predicted particle positions or global features through the decoder.

Considering that the density of the simulated system is 3, we set the number of loop updates K to 5 so that the selected particle can traverse all the surrounding particles. In addition to this, the encoder and decoder are implemented by three hidden layers with 64 linear units each and are updated nonlinearly by means of an activation function (ReLU). The hidden layer enhances PIENet's nonlinear processing capabilities, making it more capable of predicting particle properties. For the error between the predicted value and the true value, we set it to the average difference between the coordinates of each particle.

$$\text{loss} = \frac{1}{N}\sum_{i=1}^{N} |X_{\text{pre}} - X_{\text{truth}}| + |Y_{\text{pre}} - Y_{\text{truth}}|. \tag{7}$$

2.2. Simulation details and dataset settings

Dissipative particle dynamics (DPD) proved to be a powerful approach to simulate the dynamic behavior. In the DPD approach, a single particle is characterized by a molecule or a collection of particles. Thus, DPD allows for the simulation of evolving complex fluids for studying their behavior. The force acting on each particle, consists of three pairwise additive forces[38]

$$\boldsymbol{F}_i = \sum_{j\neq i} \left(\boldsymbol{F}_{ij}^{\text{C}} + \boldsymbol{F}_{ij}^{\text{D}} + \boldsymbol{F}_{ij}^{\text{R}}\right), \tag{8}$$

where $\boldsymbol{F}_{ij}^{\text{C}}$ is the conservative force, $\boldsymbol{F}_{ij}^{\text{D}}$ is the dissipative force, $\boldsymbol{F}_{ij}^{\text{R}}$ is the random force. The parameter a_{ij} is the repulsion coefficient between particle i and particle j, r_{c} is the cutoff radius, and T represents the temperature of the system simulation.

Here, we choose to simulate binary mixtures in two dimensions for different types of fluids by varying the repulsion coefficients between two different species of particles. The size of our simulation box is $L \times L = 20 \times 20$ with periodic boundary conditions. The fluids are set at an average number density, $\rho = 3$, with a total of 1200 particles. We use the dimensionless value of the cutoff radius $(r_{\text{c}}) = 1$, the temperature of the system is set to 1, mass of the particle $m = 1$, these settings are reasonable choices for DPD simulated fluids.

Table 1. The Schmidt number Sc corresponding to different friction coefficients γ.

γ	Sc
4.5	1
22.5	2.5
45	5
90	13

The friction coefficient is an important parameter used to control the viscosity of DPD fluids. It characterizes the strength of the drag force between interacting particles, thereby influencing the overall fluid viscosity.[39] We achieve a change in the viscosity of the system by varying the friction coefficient (γ), and Table 1 indicates the values of the friction coefficients considered in this work and the corresponding Schmidt numbers (Sc). The friction coefficients between the different species of particles are the harmonic average of their respective friction coefficients.

Our goal is to train PIENet to predict long-term dynamical features during phase separation, so we constitute different datasets. Data set 1 shows different repulsion coefficients, data set 2 shows different viscosities, and the number of particles in the system is $N = 1200$. Through two data sets, the order parameters and particle positions are predicted respectively. The above datasets were simulated in Lammps, and each configuration was run 30 times independently.

There are four hyperparameters in our PTENet: connectivity radius r, update message steps k, learning rate lr, and hidden size h. For a given dataset, we label 80% of the data as training set, 10% validation set and 10% test set. Iterations consist of 1000 training epochs using the training and validation sets to obtain optimal hyperparameters on the test set.

Table 2. Hyperparameter optimization.

Hyperparameter	Value
Connectivity radius (r)	1
Update message steps (k)	5
Learning rate (lr)	0.002
Hidden size (h)	64

3. Experimental results

3.1. Prediction of particle position

Here, we propose a new GNN model that is suitable for extracting complex relationships from simulation data calculated with high precision. It can require only the particle positions at the initial moment, form point, and edge features by connecting neighboring particles, and continuously learn the changing forces during training. The whole interaction network consists of an encoder–process–decoder architecture, in which the features of nodes and edges in the graph can be computed mutually by exchanging messages between a pair of MLP assigned for nodes and edges. Pearson correlation coefficients are often used to describe the correlation between data, ranging from -1 and 1, with closer to 1 indicating a higher correlation. We use the Pearson correlation coefficient to measure the accuracy between the predicted particle positions and the actual positions.

$$p = \sum_{i=1}^{N} \frac{\text{corrcoef}(x_{\text{pre}}, x_i) + \text{corrcoef}(y_{\text{pre}}, y_i)}{2}. \tag{9}$$

Figure 2 shows the prediction accuracy of our model for particle positions during phase separation. It can be seen that

our model has a very high accuracy (Pearson correlation coefficients are all above 0.5) in the short term, especially in the time period when the phase separation evolves faster. Experiments show that PIENet reproduces the spatial distribution of local density changes well. In particular, for short-term predictions with $t < 1000$, we find that the model can distinguish the specific force and distance changes produced by different species of particles. For periods $t > 1000$, the predictions become more diffuse but still capture the contrast between regions of variation. We compare with the classic neural network method: CNN extracts local features of data through convolution operations, and has powerful feature extraction capabilities; SVM effectively processes nonlinear high-dimensional data through kernel functions. The PIENet can process graph structure data, combine node and edge information, and help capture local and global structural features in graph data. For the prediction of particle positions, our method outperforms other methods at all time scales, with average accuracy improvements of 40.3% and 51.77% respectively.

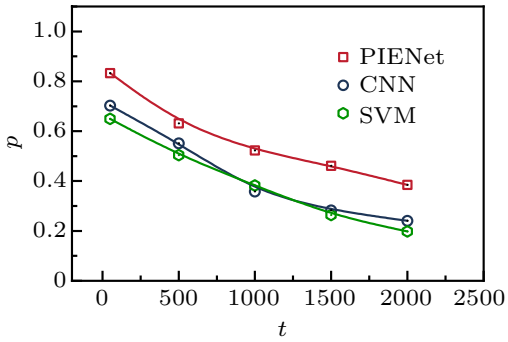


Fig. 2. Pearson correlation coefficient of particle position relative to simulation time for different neural network methods.

In fact, our model first learns the short-time dynamics on the edges and then gains the ability to predict the particle interaction forces on the nodes, which can be understood from the Pearson coefficients and the training duration. Overall, GNN uses the information learned from the interaction forces and distance variations of paired particles to fix prediction errors over long time periods with high accuracy within the model.

3.2. Simulation of phase separation with different repulsion coefficients

To characterize the evolution of phase separation, we designed an order parameter to describe the distribution of the local density of the two components. The length of the simulation box is L, so we divide the box into $L \times L$ grids and count the number of A particles and B particles in each grid. The numbers of A particles and B particles in the i-th grid are NA_i and NB_i. When particles of the same species are gathered in a lattice, and the order parameter reaches its maximum value. This approach allows us to analyze the spatial distribution of particles and their evolution. The order parameter φ is defined as follows:[33]

$$\varphi = \left\langle \sum_{i=1}^{L^2} \frac{|NA_i - NB_i|}{(NA_i + NB_i)L^2} \right\rangle. \tag{10}$$

The angular brackets indicate the min–max normalization of the order parameters. In DPD simulations, the repulsion coefficients between different particles are the key factors driving the system into the phase separation state. The repulsion coefficient determines the segregation strength of the phase separation process. Weak segregation has been referred to as a shallow quench, and strong segregation is referred to as a deep quench.[15,40] In our simulation, for the interaction between different species, we chose $\alpha_{ij} = 25$ as the interaction repulsion coefficient of A–A and B–B. For the A–B interaction, we chose two typical α_{ij} values, 36 and 50 respectively, corresponding to shallow quenching and deep quenching.

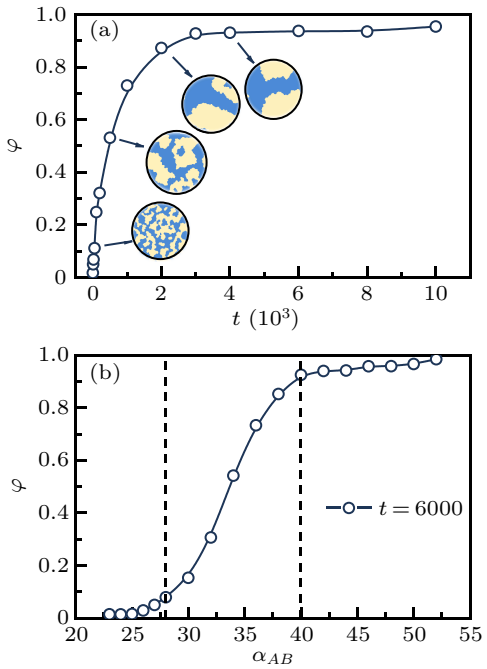


Fig. 3. (a) The plot of the variation of the order parameter φ concerning time t. We took the data of 14 points in the simulation step. (b) $t = 6000$, the variation of the parameter φ concerning the repulsion coefficient, the two dashed lines correspond to $\alpha_{AB} = 28$ and $\alpha_{AB} = 40$.

As shown in Fig. 3(a), it can be seen that the evolution of phase separation becomes more and more stable as time changes φ gradually from 0 to 1, while the A and B particles also keep gathering. And φ remains relatively stable after $t = 4000$, which indicates that the order parameter φ can describe the process of phase separation well. Figures 3(b) show the effect of the repulsion coefficient on the order parameter at $t = 6000$. Based on the values of φ, we identify three different regimes.[15] For α_{AB} values below 28, φ remains around 0.02. Beyond this critical threshold, the system enters a weak separation state, with φ gradually increasing as α_{AB} rises. How-

ever, when α_{AB} exceeds 40, φ becomes relatively stable, nearing a value of 0.95. Therefore, we hope that the change of φ with α_{AB} will become one of the new indicators for determining the boundary between shallow and deep quenching of the degree of phase separation.

3.3. Simulation of phase separation for symmetric and asymmetric Schmidt numbers

In phase separation processes with binary fluids, the viscosity differences of different fluids greatly affect the dynamic behavior of phase separation. The fluids in a multiphase system do not only interact with themselves but also with one another. It is shown that the Schmidt number determines the relative importance of momentum and mass diffusion, and it plays an important role in phase separation dynamics.

As shown in Fig. 4(a), φ varies changes with time for symmetric Schmidt number, and as the Schmidt number increases, the viscosity between the particles increases, which greatly affects the rate of phase separation. It can be seen that the asymmetric Schmidt number accelerates the rate of phase separation, but still maintains the normal trend. Figures 4(b) is the relationship between φ and t under the asymmetric Schmidt number. When the viscosity difference between different components is large, the rate of phase separation will be affected, which also leads to a deeper degree of phase separation. This is mainly due to the large viscosity gap between different components. The reduction in viscosity increases the diffusion rate of the fluid, and the fluid mechanics due to the asymmetry of viscosity at the fluid interface greatly changes the coarsening mechanism.

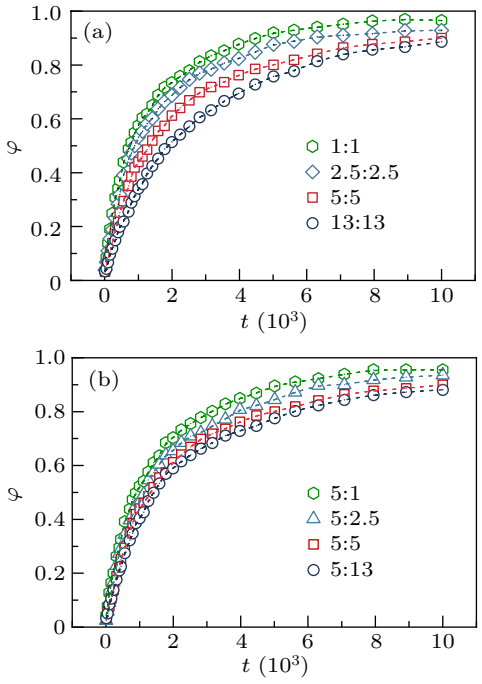


Fig. 4. (a) The variation of the order parameter φ with time at symmetric Schmidt number. (b) The variation of the order parameter φ with time at asymmetric Schmidt number.

3.4. Prediction of order parameters

Figures 5(a) show the effect of the repulsion coefficient α_{AB} on the phase separation behavior. When the repulsion coefficient of phase separation $\alpha_{AB} = 25$, the value of φ is stable at about 0.02, and no phase separation is observed. As the repulsion coefficient increases, the mixture starts separating, but this separation is relatively weak, forming an A-rich and B-rich phases. When α_{AB} exceeds a certain critical value, it can be observed that the phase separation behavior changes from a weak separation state to a strong separation state. We perform the predictions of the local density order parameters for the phase separation processes with different repulsion coefficients.

From Fig. 5(b), it can be seen that the GNN model predicts φ with higher accuracy in the shallow quench ($\alpha_{AB} = 30$) than in the deep quench ($\alpha_{AB} = 50$). Further, our model shows excellent performance as the simulation time varies, with an overall average error rate within 5%. The GNN model is capable of accurately predicting the local dynamic behavior under different repulsion parameters, even without the description parameters of the particle environment during the prediction process. Despite the forces between particles are different, the model only requires basic geometric information to learn the features of the particles and edges during training, resulting in excellent predictive results. These results show that GNNs are very sensitive to changes in local density, especially distance changes due to interactions between different particles.

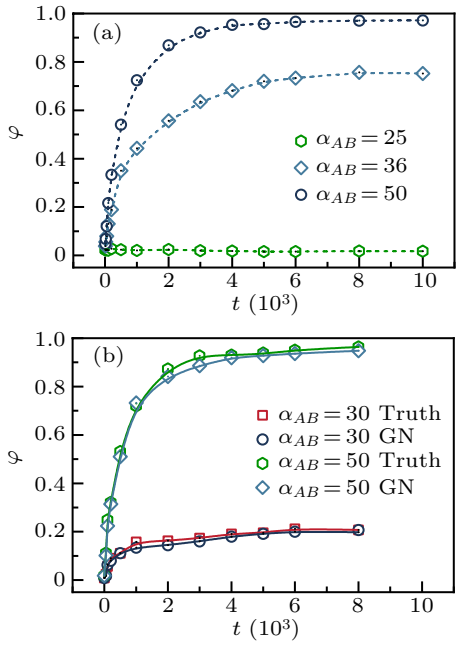


Fig. 5. (a) Variation of the local density order parameter φ with simulation time t for different repulsion coefficients. (b) The true and predicted values of the order parameter φ for different repulsion coefficients.

4. Conclusions

In this work, we design a data-driven GNN model based on the particle features of the initial position and learn the

conserved forces and relative distances between the characteristic particle pairs. We successfully predict the order parameter changes and make long-term predictions for the particle positions. Furthermore, we simulated the phase separation of binary mixtures under different conditions and measured the evolution of the phase separation by introducing an order parameter of local density. The results show that the order parameter can effectively describe the process of two-phase separation and is validated for different systems. Experiments demonstrate that our model has strong robustness and can be suitable for phase separation models in a variety of regimes.

The machine learning model of our proposed model can be used in a variety of complex systems because it can gain the ability to predict long-term particle positions based on the dynamic properties of the structure, which is particularly interesting for physical simulations. Moving forward, we believe that graph networks enhanced with physical information can be applied to more difficult classification and prediction tasks. Specifically, we are interested in applying PIENet to investigate phase separation processes in three dimensions. The method is to add the z coordinate of the particle as a feature of the node and calculate the Euclidean distance between the three-dimensional particle pairs as a feature of the edge. At the same time, how to effectively deal with complex molecular structures and dynamic interaction forms in high-dimensional space is a key issue in specific molecular simulation processes. In the future, graph neural networks are expected to play an important role in improving the accuracy of molecular simulations and increasing computational efficiency, pushing the field of molecular simulations to a new stage of development.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant No. 11702289) and the Key Core Technology and Generic Technology Research and Development Project of Shanxi Province, China (Grant No. 2020XXX013).

References

[1] Fletcher P D, Howe A M and Robinson B H 1987 *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* **83** 985
[2] Buttinoni I, Bialké J, Kümmel F, Löwen H, Bechinger C and Speck T 2013 *Phys. Rev. Lett.* **110** 238301
[3] Tung C, Harder J, Valeriani C and Cacciuto A 2016 *Soft Matter* **12** 555
[4] Richard D, Löwen H and Speck T 2016 *Soft Matter* **12** 5257
[5] Baumgart T, Hammond A T, Sengupta P, Hess S T, Holowka D A, Baird B A and Webb W W 2007 *Proc. Nat. Acad. Sci. USA* **104** 3165
[6] Chong P A and Forman-Kay J D 2016 *Current Opinion in Structural Biology* **41** 180
[7] Zola R S, Evangelista L, Yang Y C and Yang D K 2013 *Phys. Rev. Lett.* **110** 057801
[8] Bažec M and Žumer S 2006 *Phys. Rev. E* **73** 021703
[9] Jeon Y, Jamil M, Lee H D and Rhee J 2008 *Pramana* **71** 559
[10] Connell S D and Smith D A 2006 *Molecular Membrane Biology* **23** 17
[11] Girelli A, Rahmann H, Begam N, Ragulskaya A, Reiser M, Chandran S, Westermeier F, Sprung M, Zhang F, Gutt C, *et al.* 2021 *Phys. Rev. Lett.* **126** 138004
[12] Novik K E and Coveney P V 2000 *Phys. Rev. E* **61** 435
[13] Shimizu R and Tanaka H 2015 *Nat. Commun.* **6** 7407
[14] Das K and Das S K 2023 *Phys. Rev. E* **107** 044116
[15] Gidituri H, Akella V, Vedantam S and Panchagnula MV 2019 *J. Chem. Phys.* **150** 234903
[16] Laradji M, Mouritsen O G, Toxvaerd S and Zuckermann M J 1994 *Phys. Rev. E* **50** 1243
[17] Ahmad S, Das S K and Puri S 2012 *Phys. Rev. E* **85** 031140
[18] Perlmutter J D and Sachs J N 2011 *Biophys. J.* **100** 491a
[19] Warren P B 1998 *Current Opinion in Colloid & Interface Science* **3** 620
[20] Cai H, Vernon R M and Forman-Kay J D 2022 *Biomolecules* **12** 1131
[21] Zhang P and Chern G W 2021 *Phys. Rev. Lett.* **127** 146401
[22] Lopez C A, Vesselinov V V, Gnanakaran S and Alexandrov B S 2019 *J. Chem. Theory Comput.* **15** 6343
[23] Chang Y C, Chen Y J, Chen P Y, Chen Y C, Maqbool F, Ho T Y and Chiang Y Y 2023 *ACS Appl. Mater. Interfaces* **15** 12473
[24] Mandal R, Casert C and Sollich P 2022 *Nat. Commun.* **13** 4424
[25] Chu X, Sun T, Li Q, Xu Y, Zhang Z, Lai L and Pei J 2022 *BMC Bioinformatics* **23** 72
[26] Noé F, Tkatchenko A, Müller K R and Clementi C 2020 *Annu. Rev. Phys. Chem.* **71** 361
[27] Zhang K, Li X, Jin Y and Jiang Y 2022 *Soft Matter* **18** 6270
[28] Boattini E, Smallenburg F and Filion L 2021 *Phys. Rev. Lett.* **127** 088007
[29] Mokhtar F, Kansal R, Diaz D, Duarte J M, Pata J, Pierini M and Vli mant J R 2021 arXiv:2111.12840v1 [physics.data-an]
[30] Shen Z A, Luo T, Zhou Y K, Yu H and Du P F 2021 *Briefings in Bioinformatics* **22** bbab051
[31] Ha S and Jeong H 2021 *Sci. Rep.* **11** 12804
[32] Wang R, Fang F, Cui J and Zheng W 2022 *Sci. Rep.* **12** 500
[33] Zhao H, Wang R, Zhao C and Zheng W 2023 *Chin. Phys. B* **32** 056402
[34] Liu Y, Wang R, Zhao C and Zheng W 2022 *Chin. Phys. B* **31** 116401
[35] Sanchez-Gonzalez A, Godwin J, Pfaff T, Ying R, Leskovec J and Battaglia P 2020 *Proceedings of the* 37*th International Conference on Machine Learning*, July 13–18, 2020, Virtual, p. 8459
[36] Bapst V, Keck T, Grabska-Barwińska A, Donner C, Cubuk E D, Schoenholz S S, Obika A, Nelson A W, Back T, Hassabis D, *et al.* 2020 *Nat. Phys.* **16** 448
[37] Shiba H, Hanai M, Suzumura T and Shimokawabe T 2023 *J. Chem. Phys.* **158** 084503
[38] Groot R D and Warren P B 1997 *J. Chem. Phys.* **107** 4423
[39] Visser D, Hoefsloot H C and Iedema P D 2006 *J. Comput. Phys.* **214** 491
[40] Zhao Y, Liu H, Lu Z Y and Sun C C 2008 *Chin. J. Chem. Phys.* **21** 451