

MACHINE-LEARNING-BASED MULTI-SCALE MODELING FOR COMPLEX FLUIDS

By

Pei Ge

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Computational Mathematics, Science and Engineering – Doctor of Philosophy

2025

## **ABSTRACT**

Multi-scale modeling presents a long-standing challenge in computational mathematics and is pertinent to a wide range of applications in materials science, fluid physics, and chemical engineering. Predicting collective behaviors typically necessitates the integration of modeling dynamics across micro-scale (atomistic), meso-scale (kinetic), and macro-scale (continuum) levels, with the vast range of spatiotemporal scales posing a fundamental obstacle. Existing methods often rely on certain empirical constitutive closures or micro-macro coupling approaches. Despite their broad applications, modeling accuracy and efficiency are often challenged in real applications.

This dissertation aims to develop data-driven approaches for constructing accurate and reliable reduced models of multi-scale systems based on first-principle descriptions. The first part, including Chapters 2 and 3, focuses on constructing meso-scale reduced models of polymer kinetics directly from the full molecular dynamics. Chapter 2 discussed the many-body effect on conservative force, which is important to accurately reproduce both the probability density function of the void formation in bulk and the spectrum of the capillary wave across the fluid interface. Chapter 3 discussed the state-dependence on memory kernel and demonstrated the essential role of the broadly overlooked state-dependency nature in predicting molecule kinetics related to conformation relaxation and transition. The second part, Chapter 4, focuses on building accurate macroscale models from microscale polymer kinetics through meso-scale Langevin dynamics. A non-Newtonian hydrodynamic model is given as an example, which shows some success in systematically passing the micro-scale heterogeneous polymer structural mechanics to the macro-scale hydrodynamics without human intervention.

Copyright by  
PEI GE  
2025

## **ACKNOWLEDGEMENTS**

Completing this dissertation has been a journey of intellectual growth, perseverance, and support from many individuals. I am deeply grateful to all those who have guided and encouraged me throughout this process.

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Huan Lei, for his unwavering support, insightful guidance, and patience throughout my academic journey. His mentorship has been invaluable in shaping my research skills and fostering my academic growth. His dedication and encouragement have inspired me to persevere through challenges and strive for excellence. I am truly fortunate to have had the opportunity to learn from him. I am also sincerely thankful to my dissertation committee members, Dr. Daniel Appelö, Dr. Michael Murillo, and Dr. Yimin Xiao, for their valuable feedback and encouragement. Their expertise and perspectives have greatly enriched this work.

I am especially grateful to my parents for their steadfast belief in me. Their endless love and encouragement have been my greatest source of strength, providing the foundation for my resilience.

I would also like to thank my friends Liyao Lyu, Zhiyuan She, Shijun Liang, and many others, including Lidong Fang, Haishen Dai, Yue Zhao, and Siyu Guo, for their support and camaraderie throughout this journey.

This dissertation is as much a product of my efforts as it is of the collective support and encouragement of all these individuals. I am truly grateful.



## TABLE OF CONTENTS

CHAPTER 1	OVERVIEW . . . . .	1
1.1	Background . . . . .	1
1.2	Dissertation Contributions . . . . .	1
1.3	Dissertation Structure . . . . .	3
CHAPTER 2	COARSE-GRAINED MOLECULAR DYNAMICS MODELING . . . . .	5
2.1	Introduction . . . . .	5
2.2	Methods and Models . . . . .	9
2.3	Numerical Results . . . . .	13
2.4	Summary . . . . .	23
CHAPTER 3	DATA-DRIVEN LEARNING OF THE STATE-DEPENDENT MEMORY KERNEL . . . . .	25
3.1	Introduction . . . . .	25
3.2	Model Derivation . . . . .	27
3.3	Numerical Results . . . . .	35
3.4	Summary . . . . .	39
3.5	Other Details . . . . .	41
CHAPTER 4	DEEP LEARNING-BASED NON-NEWTONIAN FLUID MODEL . . . . .	46
4.1	Introduction . . . . .	46
4.2	Methods . . . . .	49
4.3	Numerical results . . . . .	60
4.4	Discussion . . . . .	71
CHAPTER 5	CONCLUSION AND OUTLOOK . . . . .	73
5.1	High Dimensional Stochastic Reduced Model . . . . .	74
5.2	Variational-informed Structure-preserving Macro-scale Reduced Model . . . . .	74
BIBLIOGRAPHY . . . . .		76

# CHAPTER 1

## OVERVIEW

### 1.1 Background

Multi-scale modeling is pertinent to a wide range of applications in materials science, fluid physics, and chemical engineering. Predicting collective behaviors typically necessitates the integration of modeling dynamics across micro-scale (atomistic), meso-scale (kinetic), and macro-scale (continuum) levels, with the vast range of spatiotemporal scales posing a fundamental obstacle. Empirical models are often based on oversimplified micro-scale information but do not have effective algorithms to extract the relevant information from micro-scale (E et al., 2023).

Machine learning present as been successful in dealing with high dimensional problems in scientific computing, such as physics-informed neural networks (Raissi et al., 2019a), deep operator networks (Lu et al., 2021), and Fourier neural operator (Li et al., 2021). For multi-scale modeling, a tremendous amount of progress has been made, such as deep learning-based coarse-grained molecular dynamics (Zhang et al., 2018b), machine learning-based moment closure models for the kinetic equation (Han et al., 2019a), etc. Despite the overwhelming success during the past years, some challenges as the collection of data, the generalization of the neural network, and preservation of physical constraints remain less illustrating. By addressing these challenges, this dissertation aims to utilize machine learning to construct reliable and structure-preserved reduced models with interpretable micro-scale to macro-scale mapping.

### 1.2 Dissertation Contributions

This dissertation focuses on both micro-to-meso and micro-to-macro modeling.

#### 1.2.1 Micro to Meso: Data-driven Stochastic Reduced Model

Predicting the collective behavior of complex multiscale systems is often centered around projecting the full-dimensional dynamics onto a set of resolved variables. However, an accurate construction of such a reduced model remains a practical challenge for real applications such as molecular modeling. While model reduction frameworks such as the Koopman operator (Koopman, 1931) and the Mori-Zwanzig projection formalism (Mori, 1965; Zwanzig, 1961) enable us to write

down the dynamic equations in terms of the resolved variables, the reduced model generally becomes non-Markovian with a memory term that may further depend on the resolved variables; the direct numerical evaluation involves solving the expensive full-dimensional orthogonal dynamics.

Consider the systems whose microscopic state is determined by the instantaneous positions  $\mathbf{Z}_q$  and momenta  $\mathbf{Z}_p$  of the  $N$  atoms (3D system). Denote the collection of these variables by  $\mathbf{Z}(t) = (\mathbf{Z}_p, \mathbf{Z}_q)$ , which is a vector of  $6N$  components. The Hamiltonian dynamics can be written as:

$$\frac{d\mathbf{Z}(t)}{dt} = \mathbf{J} \frac{\partial H(\mathbf{Z}(t))}{\partial \mathbf{z}}, \quad \mathbf{Z}(0) = \mathbf{z}, \quad (1.1)$$

where  $H$  is the Hamiltonian and  $\mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}$

Let  $(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2m}$  represent the resolved variables of a high-dimensional Hamiltonian system, where  $\mathbf{q}$  denotes the coarse-grained (CG) coordinates as a function of  $\mathbf{Z}_q$ , and  $\mathbf{p}$  denotes the CG momenta. Following the Zwanzig's formalism (Zwanzig, 2001; Hijón et al., 2010), the reduced dynamics takes the form

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{M}^{-1} \mathbf{p}, \\ \dot{\mathbf{p}} &= -\nabla U(\mathbf{q}) - \int_0^t \mathbf{K}(\mathbf{q}(\tau), t - \tau) \mathbf{v}(\tau) d\tau + \mathcal{R}_t, \end{aligned} \quad (1.2)$$

where  $\mathbf{M}$  is the mass matrix,  $U(\mathbf{q})$  is the free energy,  $\mathbf{v} := \dot{\mathbf{q}}$  is the velocity,  $\mathbf{K}(\mathbf{q}, t)$  is the memory, and  $\mathcal{R}_t$  is the noise whose covariance function is related to the memory following the second FDT (Vroylandt and Monmarché, 2022). The construction of the free energy will be discussed in Chapter 2 and the construction of the memory will be discussed in Chapter 3.

### 1.2.2 Micro to Macro: A Deep Learning-Based Non-Newtonian Hydrodynamic Model

A long-standing problem in the modeling of non-Newtonian hydrodynamics of polymeric flows is the availability of reliable and interpretable hydrodynamic models that faithfully encode the underlying micro-scale polymer dynamics. The main complication arises from the long polymer relaxation time, the complex molecular structure, and the heterogeneous interaction.

The empirical continuum hydrodynamic model of incompressible non-Newtonian fluids is given

as follows:

$$\begin{aligned}\nabla \cdot \mathbf{u} &= 0 \\ \rho \frac{d\mathbf{u}}{dt} &= -\nabla p + \nabla \cdot (\tau_s + \tau_p) + \mathbf{f}_{\text{ext}}\end{aligned}\tag{1.3}$$

where  $\mathbf{u}$  is the velocity,  $\rho$  is the density,  $p$  is the pressure,  $\tau_s \propto (\nabla \mathbf{u} + \nabla \mathbf{u}^T)$  is the solvent stress,  $\mathbf{f}_{\text{ext}}$  is the external force,  $\tau_p$  is the polymer stress which is generally unknown. The construction of  $\tau_p$  will be discussed in Chapter 4.

### 1.3 Dissertation Structure

The dissertation is organized as follows.

Chapter 2 constructs the free energy in CG models of both single- and two-component of polymeric fluid systems based on the recently developed deep coarse-grained potential (DeePCG) Zhang et al. (2018b) scheme, where each polymer molecule is modeled as a CG particle. In section 2.2, the free energy of the CG models is constructed by only using the training samples of the instantaneous force under the thermal equilibrium state. In section 2.3, we show that the constructed CG models can accurately reproduce both the probability density function of the void formation in bulk and the spectrum of the capillary wave across the fluid interface. More importantly, the CG models accurately predict the volume-to-area scaling transition for the apolar solvation energy, illustrating the effectiveness to probe the meso-scale collective behaviors encoded with molecular-level fidelity.

Chapter 3 focuses on constructing the memory kernel. We present a data-driven method to learn stochastic reduced models of complex systems that retain a state-dependent memory beyond the standard generalized Langevin equation with a homogeneous kernel. In section 3.2, we show that the constructed model naturally encodes the heterogeneous energy dissipation by jointly learning a set of state features and the non-Markovian coupling among the features. Section 3.3 demonstrate the limitation of the standard GLE and the essential role of the broadly overlooked state-dependency nature in predicting molecule kinetics related to conformation relaxation and transition.

Chapter 4 focuses on a micro-to-macro model named DeePN<sup>2</sup>. In section 4.2 presents that

the model retains a multi-scaled nature by mapping the polymer configurations into a set of symmetry-preserving macro-scale features. The extended constitutive laws for these macro-scale features can be directly learned from the kinetics of their micro-scale counterparts. Section 4.3 shows that DeePN<sup>2</sup> can faithfully capture the broadly overlooked viscoelastic differences arising from the specific molecular structural mechanics without human intervention.

Finally, Chapter 5 summarizes the discussed modeling and discusses the potential future work.

## CHAPTER 2

### COARSE-GRAINED MOLECULAR DYNAMICS MODELING

In this chapter, we focus on how to construct the conservative force. Conservative forces are essential in constructing CG models, as they describe the interactions between CG particles. Accurate conservative forces are critical for obtaining accurate memory and noise terms in CG models. Therefore, it is crucial to have a good approximation of the conservative forces.

In Ge et al. (2023), we construct the CG models of both single- and two-component of polymeric fluid systems based on the recently developed deep coarse-grained potential (DeePCG) (Zhang et al., 2018b) scheme, where each polymer molecule is modeled as a CG particle. By only using the training samples of the instantaneous force under the thermal equilibrium state, the constructed CG models can accurately reproduce both the probability density function of the void formation in bulk and the spectrum of the capillary wave across the fluid interface. More importantly, the CG models accurately predict the volume-to-area scaling transition for the apolar solvation energy, illustrating the effectiveness to probe the meso-scale collective behaviors encoded with molecular-level fidelity.

#### 2.1 Introduction

Molecular dynamics (MD) simulations provide a promising avenue to establish the atomistic-level understanding of many complex systems relevant to biological and materials science. Despite the overwhelming success during the past decades, a remaining bottleneck roots in the limitation of the achievable spatio-temporal scales; the gap between the micro-scale atomistic motions and many meso-scale emerging phenomena remains large. One important problem is the nano-scale interfacial fluids, which play a crucial role in the hydration and the assembly of the biomolecules and functional nano-materials (Chandler, 2005; Berne et al., 2009). However, it is well-known that such fluid systems generally exhibit complex and multifaceted nature on different scales. On the small scale (i.e., the fluid molecule correlation length), the solvation energy is determined by the molecular reorganization and scales with the volume of the void space. On the large scale, the solvation energy is determined by the free energy for maintaining a fluid-void interface and scales with the surface area. The scale-dependent behavior indicates an cross-over regime of the entropy-enthalpy transition.

While theoretical understandings (Lum et al., 1999; Rein ten Wolde et al., 2001; Hummer et al., 1996, 1998) of this ubiquitous phenomenon have been developed, computational modeling often relies on full micro-scale MD simulations to retain the multifaceted properties, which, however, remain too expensive to achieve the resolved scale for applications such as nano-scale assembly.

To accelerate the full MD simulations, many coarse-grained (CG) models have been developed. By modeling the dynamics in terms of a set of CG variables with reduced dimensionality, the coarse-grained molecular dynamics (CGMD) simulations, in principle, enable us to probe the collective behaviors on a broader scale. However, in practice, the construction of truly reliable CG models can be highly non-trivial, especially for the meso-scale interfacial fluids. There are two major challenges. The first challenge arises from the many-body nature of CG interactions. Specifically, the equilibrium density distribution of the CG model needs to match the marginal density distribution of the CG variables of the full model. Due to the unresolved atomistic degrees of freedom, the CG potential generally encodes the many-body interactions even if the full MD force field is governed by two-body interactions (Noid et al., 2008). Existing approaches often rely on various physical intuitions as well as empirical approximations (Izvekov and Voth, 2005; Noid, 2013; Lei et al., 2010; Hijón et al., 2010; Rudd and Broughton, 1998; Pagonabarraga and Frenkel, 2001; Nielsen et al., 2004; Shinoda et al., 2008; Molinero and Moore, 2009; Larini et al., 2010; Das and Andersen, 2012; Dinpajoooh and Guenza, 2017; Sanyal and Shell, 2016; Moore et al., 2016) that reproduce certain target thermodynamic quantities and/or structural distributions. For example, the pairwise additive decomposition based on direct ensemble averaging (Lei et al., 2010; Hijón et al., 2010) can recover the thermodynamic pressure but often fail to recover the pair distribution function. Conversely, the Monte Carlo and Boltzmann inverse approaches (Lyubartsev and Laaksonen, 1995; Soper, 1996; Reith et al., 2003) can reproduce the pairwise distribution function, which, however, lead to the biased predictions of the equation of state. Several studies account for the many-body effects by introducing the configuration-independent volume potential (Das and Andersen, 2010; Dunn and Noid, 2015, 2016) and the local density (Allen and Rutledge, 2008, 2009; Izvekov et al., 2010; Moore et al., 2016; Sanyal and Shell, 2016; Shahidi et al., 2020) into the pairwise interactions.

On the other hand, the accuracy of the high-order structural correlations as well as the direct applications to interfacial systems remains under-explored.

Besides the many-body effect, the fluid molecules also exhibit heterogeneous density at the interfacial vicinity. What further complicates the problem is the fact that the interfacial fluid density distribution is scale-dependent. On the small scale, the molecular reorganization generally leads to a wet interface with larger density than the bulk value. On the large scale, the fluid-void phase separation generally leads to a dry interface with lower density. The crossover implies complex molecular correlations near the interface. To capture this multi-faceted property, the constructed CG potential needs to properly embody the local particle distribution other than the homogeneous bulk distribution. Conventional structural-based CG potential functions generally show limitations to incorporate such information. Similar to the many-body dissipative particle dynamics (Pagonabarraga and Frenkel, 2001), recent studies employed the local density (Wagner et al., 2017; DeLyser and Noid, 2017; Sanyal and Shell, 2018; Jin and Voth, 2018; DeLyser and Noid, 2019, 2020; Berressem et al., 2021) as well as the density gradient (DeLyser and Noid, 2022) as the auxiliary field variables to construct the CG potential functions. While the CG models show significant improvement to reproduce the interfacial density profile, the scale-dependent interfacial energy and fluctuations have not been systematically investigated. In Ref. (Lei et al., 2015), interfacial energy is integrated into the continuum fluctuation hydrodynamic equation (Landau and Lifshitz, 1987) from the top-down perspective. Fluid particles essentially represent the Lagrangian discretization points based on the smoothed dissipative particle hydrodynamics (Serrano and Español, 2001) instead of the CG molecules; the meso-scale fluid structural properties can not be retained. Currently, the construction of reliable bottom-up CGMD models that faithfully encode the multifaceted molecular interactions remains largely open.

In this work, we aim to address the above challenges by constructing CG models of meso-scale interfacial fluids based on the deep molecular dynamics (DeePMD) scheme (Zhang et al., 2018a,c). DeePMD is initially developed for learning the many-body interactions from the ab initio molecular dynamics, and has been applied to construct the deep coarse-grained (DeePCG) model



(Zhang et al., 2018b) of liquid water in bulk. Unlike the conventional forms of the inter-molecular potential function, the DeePMD represents each particle as an agent and the relative positions of its neighboring particles as the local environment. Rather than approximating the total potential of the full system by an unified parametric function, the DeePMD directly maps the local environment of each agent to the potential energy of that particle through a neural network that strictly preserves the spatial symmetries and the particle permutation invariance. Accordingly, the construction does not rely on the empirical decomposition (e.g., pairwise, three-body) of the high-dimensional particle configuration space. This unique feature is particularly suited for modeling the many-body potential of CGMD models, where the ensemble-averaged interaction between two CG particles further depends on the other neighboring CG particles and can not be represented by a pairwise additive function. Moreover, the heterogeneous particle density distribution across the fluid interface can be naturally incorporated into the CG potential function as the local environment of each particle. Accordingly, the constructed CG models can accurately model the multifaceted, scale-dependent interfacial fluctuations and apolar solvation without additional human intervention.

We demonstrate the effectiveness of the CG models by considering both the single- and two-component fluids in presence of thermal interfacial fluctuations. As discussed in Ref. (Chandler, 2005), the scale-dependent hydrophobic effects can be general for solvent molecules with attractive interactions; polymeric liquids are therefore used as the benchmark problem. We compare the numerical results from the full MD simulations and the CG description that represents each molecule as a single particle located at the center of mass. By merely using training samples under equilibrium thermal fluctuations, the constructed CG models accurately predict the high-order correlations, the local compressibility and the interfacial capillary wave for both single- and two-component fluids. In contrast, the empirical CG potential constructed based on the pairwise approximation shows apparent deviations. Furthermore, we conduct the rare-event sampling simulations to estimate the probability of the void formulation in bulk. The predictions of CG model show good agreement with the full MD results. More importantly, the CG models accurately predict the volume-to-area scaling transition for the solvation energy, and therefore, pave the way for modeling the nanoscale

assembly in aqueous environment.

Before wrapping up this section, we note that the present work focuses on the collective, quasi-equilibrium properties determined by the conservative potential function of a set of extensive CG variables; see Refs. (John and Csányi, 2017; Chan et al., 2019) for relevant work. For the conformational free energy of non-extensive CG variables, several machine-learning based approaches (Stecher et al., 2014; Mones et al., 2016; Lemke and Peter, 2017; Galvelis and Sugita, 2017; Schneider et al., 2017; Zhang et al., 2018d; Zavadlav et al., 2018; Wang et al., 2019) have been developed; see also a recent review (Noé et al., 2020) and the references therein. Furthermore, to accurately predict the dynamic properties, memory and coherent noise terms (Mori, 1965; Zwanzig, 1973) arising from the unresolved variables need to be properly introduced into the CG model (Lei et al., 2010; Hijón et al., 2010; Lei et al., 2016; Lei and Li, 2021), which are left to future investigations.

## 2.2 Methods and Models

### 2.2.1 Full Model of the Polymeric Fluids

We consider the micro-scale models of the star polymer melt similar to Ref. (Hijón et al., 2010). The full system consists of  $M$  molecules with a total number of  $N$  atoms. Each polymer molecule consists of a “center” atom connected by  $N_a$  arms with  $N_b$  atoms per arm. The positions of the atoms are denoted by  $\mathbf{q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N]$ , where  $\mathbf{q}_i$  represents the position of the  $i$ -th atom. The potential function is governed by the pairwise and bond interactions, i.e.,

$$V(\mathbf{q}) = \sum_{i \neq j} V_p(q_{ij}) + \sum_k V_b(l_k), \quad (2.1)$$

where  $V_p$  is the pairwise interaction between both the intra- and inter-molecular atoms except the bonded pairs.  $q_{ij} = \|\mathbf{q}_i - \mathbf{q}_j\|$  is the distance between the  $i$ -th and  $j$ -th atoms.  $V_b$  is the bond interaction between the neighboring particles of each polymer arm and  $l_k$  is the length of the  $k$ -th bond. The bond potential  $V_b$  is chosen to be the harmonic potential, i.e.,

$$V_b(l) = \frac{1}{2} k_s (l - l_0)^2, \quad (2.2)$$

where  $k_s$  and  $l_0$  represent the elastic coefficient and the equilibrium length  $l_0$ , respectively. In this study, all the physical quantities take the reduced unit. The atom mass is chosen to be unity.

We investigate three fluid systems with micro-scale potential governed by Eq. (2.1). In Sec. 2.3.1, we consider the polymeric fluids in bulk and examine if the CG models can retain the many-body interactions and the local compressibility. In particular, we choose  $N_a = 12$ ,  $N_b = 6$ ,  $\sigma = 2.415$ ,  $\epsilon = 1.0$ ,  $k_s = 1.714$ ,  $l_0 = 2.77$  similar to Ref. (Hijón et al., 2010).  $V_p$  takes the form of the Lennard–Jones potential with cut-off  $r_c$ , i.e.,

$$V_p(r) = \begin{cases} V_{\text{LJ}}(r) - V_{\text{LJ}}(r_c), & r < r_c \\ 0, & r \geq r_c \end{cases} \quad V_{\text{LJ}}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right], \quad (2.3)$$

where  $\epsilon = 1.0$  is the dispersion energy and  $\sigma = 2.415$  is the hardcore distance. Also we choose  $r_c = 2^{1/6}\sigma$  so that  $V_p$  recovers the Weeks-Chandler-Andersen potential. The full system consists of  $N = 2120$  polymer molecules in a cubic domain  $180 \times 180 \times 180$  (in reduced unit) with periodic boundary condition imposed along each direction. The Nosé-Hoover thermostat is employed to conduct the canonical ensemble simulation with  $k_B T = 3.96$ .

In Sec. 2.3.2, we consider the polymeric fluid in presence of fluid-void interface. Micro-scale model parameters are similar to Sec. 2.3.1 except that  $r_c = 2.5\sigma$  and  $k_B T = 1.7$ . Simulations are conducted in a domain  $180 \times 180 \times 200$  with periodic boundary condition imposed along the  $x$ - and  $y$ -direction. At the equilibrium, the fluid shows a clear fluid-void interface near  $z = 20$  and  $z = 180$ , respectively.

In Sec. 2.3.3, we consider a two-component polymeric fluid. Micro-scale model of the polymer molecule is similar to the single-component fluid system with  $N_a = 15$ ,  $N_b = 12$ ,  $k_s = 20.0$ ,  $l_0 = 1.5$ ,  $k_B T = 0.5$ . The full system consists of 3488 molecules in a domain  $200 \times 200 \times 120$  with periodic boundary condition imposed along each direction. The pairwise interaction  $V_p$  is chosen to be quadratic, i.e.,

$$V_p(r) = \begin{cases} \frac{a}{2r_c} (r - r_c)^2, & r < r_c \\ 0, & r \geq r_c \end{cases}. \quad (2.4)$$

Specifically, we consider two sets of the pairwise interaction: (I)  $a^{11} = 6.0$ ,  $a^{12} = 3.0$ ,  $a^{22} = 6.0$ ,  $r_c = 1.5$ , where  $a^{12}$  represents the pairwise interaction between the component-1 and component-2 atoms. (II)  $a^{11} = 3.0$ ,  $a^{12} = 60.0$ ,  $a^{22} = 3.0$ ,  $r_c^{11} = 1.5$ ,  $r_c^{12} = 2.5$ ,  $r_c^{22} = 1.5$ . The fluid shows a full mixture and interfacial separated state for the two cases respectively.

### 2.2.2 Coarse-grained Models

For all of the three systems, we construct the CG models by representing each molecule as an individual particle. The positions of the CG particles are denoted by  $\mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_M]$ , where  $\mathbf{Q}_i = \mathbf{Q}_i(\mathbf{q})$  represents the center of mass (COM) of the  $i$ -th molecule. The conservative potential  $U(\mathbf{Q})$  is determined by the marginal density function of  $\mathbf{Q}$  with respect to the equilibrium density function of the full model, i.e.,

$$\rho(\mathbf{Q}) = \int e^{-V(\mathbf{q})/k_B T} \prod_{i=1}^M \delta(\mathbf{Q}_i(\mathbf{q}) - \mathbf{Q}_i) d\mathbf{q} / \int e^{-V(\mathbf{q})/k_B T} d\mathbf{q}, \quad (2.5)$$

$$U(\mathbf{Q}) = -k_B T \ln \rho(\mathbf{Q}).$$

In DeePCG, a neural network  $\tilde{U}(\mathbf{Q}; \Theta)$  is used to represent the CG potential  $U(\mathbf{Q})$ , where  $\Theta$  represents the neural network parameters. To keep the extensive property, the total energy is decomposing into local contributions of the individual CG particles:

$$\tilde{U}(\mathbf{Q}; \Theta) = \sum_{i=1}^M \tilde{U}_{nn}(\mathbf{D}(\tilde{\mathbf{Q}}^i); \Theta), \quad (2.6)$$

where  $\tilde{U}_{nn}$  is the local potential of an individual particle,  $\tilde{\mathbf{Q}}^i \in \mathbb{R}^{N_i \times 4}$  is the generalized coordinates of the  $i$ -th particle. It represents the local environment of the  $i$ -th particle relative to its  $N_i$  neighboring particles within cutoff  $R_c$ . In particular, the  $j$ -th row is defined as  $\tilde{\mathbf{Q}}_j^i = (s(r_j), s(r_j)x_j/r_j^2, s(r_j)y_j/r_j^2, s(r_j)z_j/r_j^2)$ , where  $\mathbf{r}_j = (x_j, y_j, z_j)$  denotes the relative position between the  $i$ -th particle and its  $j$ -th local neighbor.  $s(r)$  is a smooth differentiable function that decays to 0 at  $r = R_c$ , which ensures the force also smoothly decays to zero at the cut-off.  $\mathbf{D} \in \mathbb{R}^{M_1 \times M_2}$  is the symmetry preserving features of each particle. The entry of  $D$  can be written as:

$$D_{j,l}(\tilde{\mathbf{Q}}^i) = \left( \sum_{k=1}^{N_i} g_{1,j}(s(r_k); \Theta) \tilde{\mathbf{Q}}_k^i \right) \left( \sum_{k=1}^{N_i} g_{2,l}(s(r_k); \Theta) \tilde{\mathbf{Q}}_k^i \right)^T, \quad (2.7)$$

where  $\{g_{1,j}(s(r); \Theta)\}_{j=1}^{M_1}$  and  $\{g_{2,l}(s(r); \Theta)\}_{l=1}^{M_2}$  are neural networks mapping from the scalar  $r$  to multiple features, and  $M_1$  and  $M_2$  are the number of customized features.  $D_{j,l}$  preserves the translational and rotational invariance; the summation over index  $k$  ensures the permutational symmetry. In this study,  $s(r)$  is chosen as

$$s(r) = \begin{cases} \frac{1}{r}, & r \leq R_{cs} \\ \frac{1}{r} \left[ \frac{1}{2} \cos \left( \pi \frac{r-R_{cs}}{R_c-R_{cs}} \right) + \frac{1}{2} \right], & R_{cs} < r \leq R_c \\ 0, & r > R_c \end{cases} \quad (2.8)$$

where  $R_{cs} = 0.97R_c$  is a smooth cut-off parameter.

In principle,  $\tilde{U}(\mathbf{Q}; \Theta)$  can be trained by minimizing the difference of the predicted force terms between the full micro-scale and the CG models, i.e.,  $\langle \|\nabla \tilde{U}(\mathbf{Q}; \Theta) - \nabla U(\mathbf{Q})\|^2 \rangle_{\mathbf{Q}}$ , where  $\langle \cdot \rangle_{\mathbf{Q}}$  represents the conditional expectation with respect to the constraints of  $\mathbf{Q}$ , i.e.,  $\prod_{i=1}^M \delta(\mathbf{Q}_i(\mathbf{q}) - \mathbf{Q}_i)$ . However, the evaluation of the force term  $-\nabla U(\mathbf{Q})$  relies on the constraint sampling with respect to  $\delta(\mathbf{Q}(\mathbf{q}) - \mathbf{Q})$ , which can be computational expensive. On the other hand, we note that the instantaneous force  $\mathcal{F}(\mathbf{Q})$  follows  $\mathcal{F}(\mathbf{Q}) = -\nabla U(\mathbf{Q}) + \mathcal{R}(\mathbf{Q})$ , where  $\mathcal{R}(\mathbf{Q})$  is the zero-mean fluctuation force. Therefore, we have  $\langle \|\nabla \tilde{U}(\mathbf{Q}; \Theta) - \nabla U(\mathbf{Q})\|^2 \rangle_{\mathbf{Q}} = \langle \|\nabla \tilde{U}(\mathbf{Q}; \Theta) + \mathcal{F}(\mathbf{Q})\|^2 \rangle_{\mathbf{Q}} + \langle \|\mathcal{R}(\mathbf{Q})\|^2 \rangle_{\mathbf{Q}}$ , where the last term does not involve in the training. Accordingly, we can transform the training by minimizing the empirical loss

$$L = \sum_{i=1}^S \sum_{j=1}^M \left\| \nabla \tilde{U}(\mathbf{Q}^{(i)}; \Theta) + \mathcal{F}_j(\mathbf{Q}^{(i)}) \right\|^2, \quad (2.9)$$

where the superscript represents the index of  $S$  configurations. For the three micro-scale models specified in Sec. 2.2.1, we collect training samples from 50-, 200-, 250-long (in reduced unit) trajectories from the full MD simulations. 5000 snapshots are used to train the CG potential function for each case. The networks are trained by the Adam stochastic gradient descent method (Kingma and Ba, 2015). In particular, we emphasize that all the training samples are collected from thermal equilibrium states. As shown in Sec. 2.3, the constructed CG potentials naturally encode the

many-body and heterogeneous interfacial interactions, which enable us to accurately predict rare events such as the probability of the void formation and scale-dependent apolar solvation energy.

## 2.3 Numerical Results

### 2.3.1 Bulk Fluids

Let us start with the CG model of fluids in bulk. Due to the constraint terms in Eq. (2.5), the marginal probability density function  $\rho(\mathbf{Q})$  generally can not be represented in form of the simple two-point correlation  $\rho^{(2)}(\mathbf{Q}_i, \mathbf{Q}_j)$ . Accordingly, the CG potential function  $U(\mathbf{Q})$  generally exhibits the many-body nature and can not be exactly constructed in form of the pairwise interaction. This limitation was verified in earlier studies on the CG modeling of polymeric fluids (Lei et al., 2010; Hijón et al., 2010), where the CG interactions are constructed based on the pairwise decomposition, i.e.,

$$U(\mathbf{Q}) \approx \sum_{i \neq j} U^{(2)}(Q_{ij}) \quad (2.10)$$

$$\frac{dU^{(2)}(r)}{dr} = - \langle \mathcal{F}_{ij}(\mathbf{Q}_{ij}) \cdot \mathbf{e}_{ij} \rangle_{Q_{ij}=r},$$

where  $\mathbf{e}_{ij} = \mathbf{Q}_{ij}/Q_{ij}$  represents the unit vector between the  $i$ -th and  $j$ -th particle.

To examine the model accuracy, we simulate the CG models with  $U(\mathbf{Q})$  constructed in form of both Eq. (2.6) and Eq. (2.10). Fig. 2.1 shows the obtained radial distribution functions (RDFs). Predictions from the full MD and the reduced model based on the DeePCG potential (2.6) show good agreement. In contrast, the pairwise CG potential (2.10) yields pronounced over-estimations of the peak value near  $r = 16$  due to the over-simplification of the many-body CG potential using the two-body interaction; see also Refs. (Lei et al., 2010; Hijón et al., 2010).

The many-body nature of  $U(\mathbf{Q})$  is also manifested in the angular distribution functions (ADFs)  $p(\theta)$ , where  $\theta$  is the angle determined by relative positions of three molecules.

$$P(\theta; A_{rc}) = \frac{1}{W} \left\langle \sum_i \sum_{j \neq i} \sum_{k > j} \delta(\theta - \theta_{jik}) \right\rangle \quad (2.11)$$

where  $\theta_{jik}$  is the angle between  $\mathbf{Q}_{ji}$  and  $\mathbf{Q}_{ki}$ , and  $W$  is a normalization factor. The summation is over all the triplet  $i, j, k$ , such that  $\|\mathbf{Q}_i - \mathbf{Q}_j\| \leq A_{rc}$  and  $\|\mathbf{Q}_i - \mathbf{Q}_k\| \leq A_{rc}$ . Fig. 2.2 shows the

ADFs within four different cut-off regimes. Similar to the RDF, predictions of the DeePCG model agree well with the full MD model while the pairwise approximation yields apparent deviations.

Besides the equilibrium correlations, we further examine the fluid local compressibility. While this property plays an important role in the nano-scale hydrophobicity, canonical solvation theories generally refer to the fluids at the proximity of the vapor-liquid coexistent phase. Here we examine this property of bulk fluids for the validation of the constructed many-body CG potential  $\tilde{U}(\mathbf{Q})$ ; the discussion of the apolar solvation energy is postponed to Sec. 2.3.2. Specifically, we examine the rare event of the void formation in bulk. Following Ref. (Patel et al., 2010), we define the smoothed molecule number within a probing spherical volume centered at  $\mathbf{Q}_c$  by

$$\hat{n}(\{\mathbf{Q}_i\}_{i=1}^M) = \sum_{i=1}^M \frac{1}{2} \left( 1 + 2 \tanh \left( \frac{R - Q_i}{h} \right) \right), \quad (2.12)$$

where  $R$  is the radius of the probing sphere,  $Q_i = \|\mathbf{Q}_i - \mathbf{Q}_c\|$  is the distance between the COM of molecule  $i$  (or equivalently, the CG particle) and the spherical center, and  $h = 1.0$  represents the smooth length.

By Eq. (2.12), particle number  $\hat{n}$  is differentiable with respect to the individual molecule position  $\mathbf{Q}_i$ . Similar to Ref. (Patel et al., 2010), we can probe the probability of the void formation by establishing a replica of umbrella sampling by imposing the bias potential

$$U_{\text{bias}}(\hat{n}; n_j) = \frac{k_n}{2} (\hat{n} - n_j)^2, \quad (2.13)$$

where  $k_n$  is the magnitude of the bias potential and  $n_j$  is the target value of the particle number inside the domain, as shown in Fig. 2.3(a). We set  $k_n = 21.9$  and establish 40 independent simulations with  $n_j$  evenly distributed between 0 and 7.5. For each replica, we collect  $8 \times 10^5$  samples of  $\hat{n}$  from a 1600-long trajectory. By using the weighted histogram analysis method (Kumar et al., 1992b), we can stitch the joint probability density  $\rho(\hat{n}, n_j)$  to construct  $\rho(\hat{n})$ . Fig. 2.3(b) shows the probability density  $\rho(\hat{n})$  obtained from the full MD and the reduced model. The predictions of the DeePCG model agree well with the full MD model over the full regime of  $\hat{n}$ .

Finally, we examine the normalized density fluctuation  $\delta n / \langle n \rangle$  within a spherical volume of various sizes, where  $\langle n \rangle$  is the average particle number and  $\delta n = \sqrt{\langle (\hat{n} - \langle n \rangle)^2 \rangle}$  is the standard

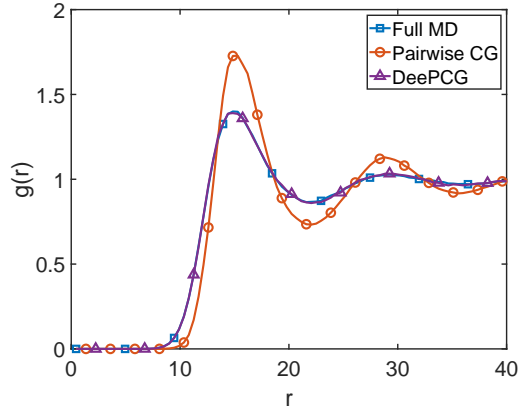


Figure 2.1 Radial distribution function  $g(r)$  of the molecule COM obtained from the full MD simulation, the CG model using the pairwise force approximation by Eq. (2.10), and the DeePCG model.

deviation. Specifically, we define the particle number by Eq. (2.12) with two different smooth length  $h = 1.0$  and  $h = 0.1$ , respectively. The latter case essentially represents each molecule as a simple point and counts the particle number as integers, and therefore, yields larger density fluctuations. As shown in Fig. 2.3(c), the full MD and CG model show good agreement for both cases, indicating that the CG model can faithfully capture the high-order correlations and the local compressibility beyond the continuum thermodynamic limit.

### 2.3.2 Single-component Interfacial Fluids

Besides the many-body interactions, another hallmark of interfacial fluids is the heterogeneous molecular distribution across the fluid interface, which leads to scale-dependent interfacial interactions and fluctuations. On the macro-scale level, the interfacial interactions can be generally described by continuum models such as the Young-Laplace equation (Rowlinson and Widom, 2002); the apolar solvation energy is proportional to the interfacial area and characterized by the surface tensor. However, on the length scale comparable to the correlation length of the fluid molecules, the interfacial energy often exhibits a cross-over regime representing the volume-dependent to area-dependent scaling transition. Therefore, the meso-scale interfacial energy provides a crucial metric to validate the accuracy of the CG model.

First, we examine the interfacial thermal fluctuations. With the micro-scale model specified in Sec. 2.2.1, the fluid molecule interaction consists of both the short-range repulsion and long-range



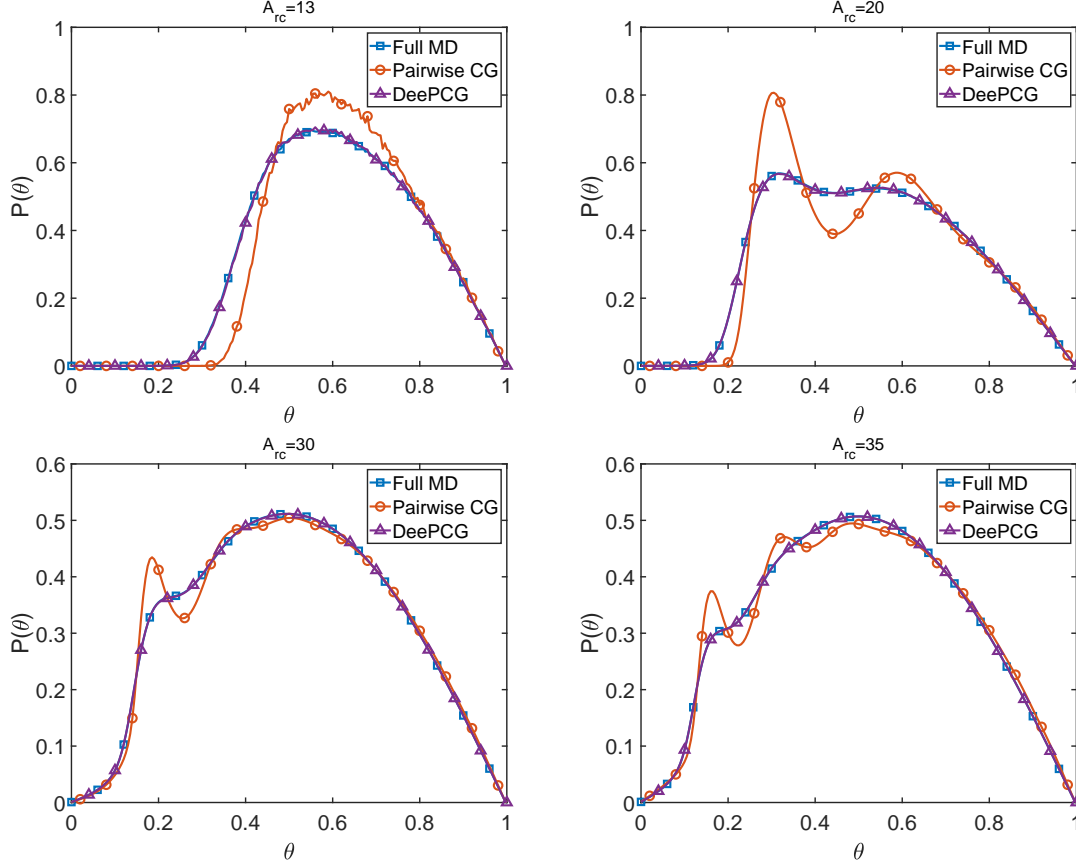


Figure 2.2 Angular distribution function  $p(\theta)$  of the molecule COM obtained from the full MD simulation, the pairwise CG model and the DeePCG model with different cut-off regimes  $A_{rc}$ .

attraction. Under the thermal equilibrium states, the fluid system exhibits the fluid-void interfaces near  $z = 20$  and  $z = 180$ . The periodic boundary condition is imposed along the  $x$ - and  $y$ -direction.

To quantify the molecule distribution near the interface at  $z = z_0$ , we define the smoothed density field  $\rho_s(\mathbf{R})$  by

$$\rho_s(\mathbf{R}) = \sum_{i=1}^M W(\|\mathbf{R} - \mathbf{Q}_i\|, h) \quad (2.14)$$

on the  $N_x \times N_y \times N_z$  lattice grids. Specifically,  $\mathbf{R}^{(i,j,k)} := (x^i, y^j, z^k)$ , where  $(x^i, y^j) = (i, j) \times dl$ ,  $dl = L/N_x$  and  $z^k = z_0 - h + k \times dz$ ,  $dz = 2h/N_z$ .  $\mathbf{Q}_i$  represents the COMs of the neighboring molecules for each grid point.  $W(r, h)$  represents the quintic spline kernel function (Morris et al., 1997) with finite support  $h$ . In this study, we set  $h = 30.0$ ,  $dl = 1.8$  and  $dz = 0.2$ .

The smoothed density field  $\rho_s(\mathbf{R})$  enables us to define the instantaneous surface (IS) height

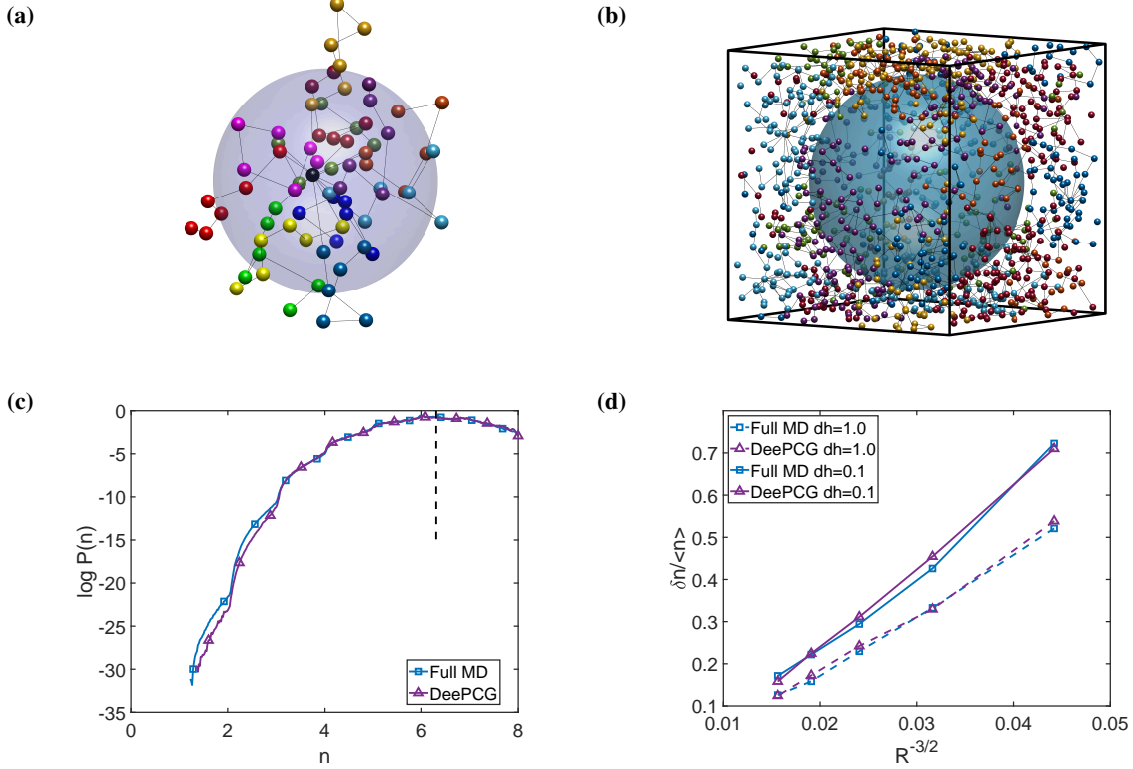


Figure 2.3 The density fluctuation and the molecule number distribution within a spherical probing volume. (a) A sketch of the star polymer with the black atom as the center. Atoms in the same arm have the same color. The transparent particle represents the coarse-grained molecule. (b) A sketch of the instantaneous molecule position under bias potential (2.13). The iso-surface in blue color represents the interface of the void space. (c) The probability density function of the molecule number within a spherical volume of radius  $R = 16.0$ . The vertical dashed line represents the average molecule number under equilibrium. (d) The normalized density fluctuations within a spherical volume of radius  $R$  between 8.0 and 16.0. The particle number is defined by Eq. (2.12) with the resolution length  $h$  set to be 0.1 (solid lines) and 1.0 (dashed lines).

$\tilde{h}(x, y)$  as the iso-surface of the fluid density (Willard and Chandler, 2010), i.e.,

$$\rho_s(x, y, \tilde{h}(x, y)) = \rho_0/2, \quad (2.15)$$

where  $\rho_0$  is the bulk fluid density, as shown in Fig. 2.4(a). Accordingly, we can compute the IS density distribution  $\tilde{\rho}(z)$  along the  $z$ -direction, where the reference position is chosen to be  $\tilde{h}(x, y)$  for each grid point  $(x, y)$ . As shown in Fig. 2.4(c),  $\tilde{\rho}(z)$  exhibits apparent oscillations across the instantaneous surface. The peaks near  $z = 6$  and  $z = 16$  represent the first and the second layer of the fluid molecule near the interface. Alternatively, we can compute the density distribution  $\rho(z)$  with respect to the plane at the average of the instantaneous height  $\langle \tilde{h}(x, y) \rangle$ , i.e., the Gibbs dividing

surface. Different from  $\tilde{\rho}(z)$ ,  $\rho(z)$  shows a smooth transition from 0 to the bulk value across the interface. For both definitions, the predictions from the CG model agree well with the full MD simulations. We emphasize that the learning of the DeePCG potential does not involve any human intervention such as the definitions of the density field and the interface height. The consistent predictions between the MD and CG models validate that constructed DeePCG potential  $\tilde{U}(\mathbf{Q}; \Theta)$  faithfully captures the intrinsic fluid structure near the interface.

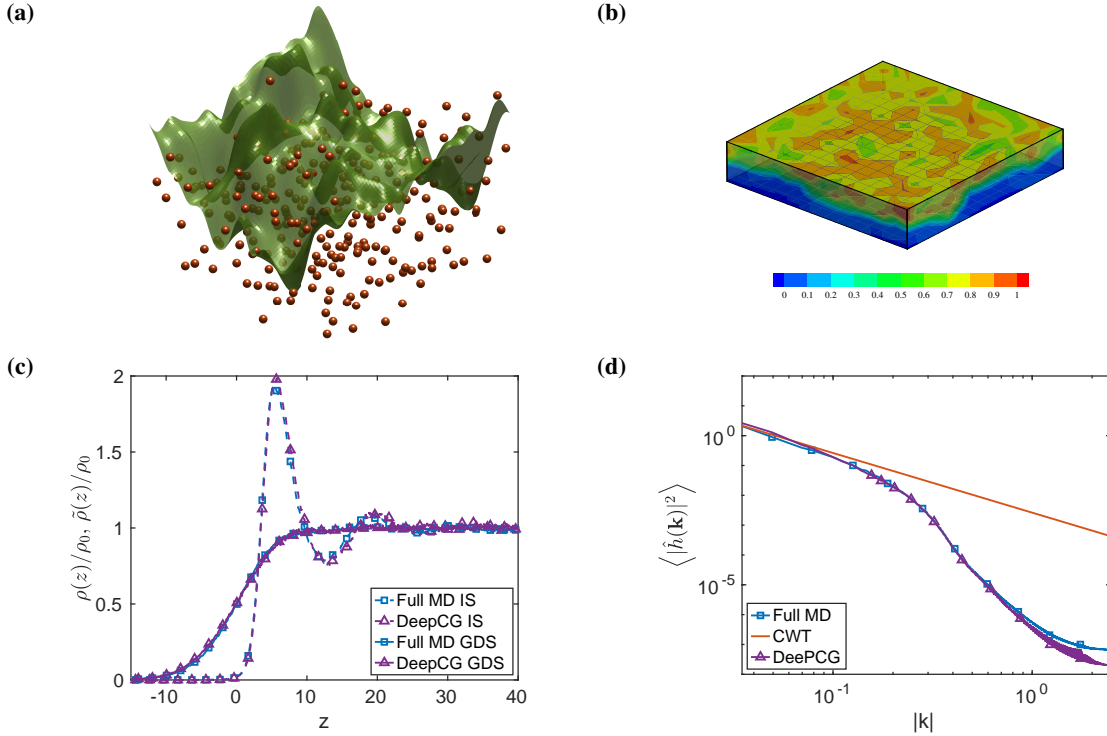


Figure 2.4 The fluid density and the fluctuating interface of the single-component interfacial fluid system. (a) The interface defined by Eq. (2.15) with molecules (red) and interface (green). (b) A sketch of the instantaneous density field defined by Eq. (2.14). (c) The average density profile across the Gibbs dividing surface (GDS) and the instantaneous interface (IS) defined by Eq. (2.15). (d) The ensemble average of the capillary wave spectrum of the fluctuating interface. The solid line in red represents the CWT fitting using Eq. (2.17) at the low wave number.

To further examine the interfacial fluctuations, we evaluate the Fourier spectrum of the instantaneous height  $\tilde{h}(x, y)$ , i.e.,

$$\hat{h}(\mathbf{k}) = \frac{1}{L^2} \int_0^L \int_0^L \tilde{h}(x, y) e^{-ik_x x - ik_y y} dx dy, \quad (2.16)$$

where  $\mathbf{k} = (k_x, k_y)$  is the 2D wave number. Fig. 2.4 shows the ensemble average of the spectrum  $\langle |\hat{h}(\mathbf{k})|^2 \rangle$ . On the low wave number limit, the interfacial energy is governed by the surface tensor

with equi-partition distribution among the individual Fourier modes following the capillary wave theory (CWT) (Buff et al., 1965; Evans, 1979), i.e.,

$$\langle |\hat{h}(\mathbf{k})|^2 \rangle = \frac{k_B T L^2}{\gamma |\mathbf{k}|^2}, \quad (2.17)$$

where  $\gamma$  is the surface tension. At low wave number,  $\langle |\hat{h}(\mathbf{k})|^2 \rangle$  obtained from numerical simulations shows good agreement with the CWT theory. As the wave number increases, the spectrum gradually deviates from the CWT prediction, indicating that there exists strong correlations between the height fluctuations of neighboring sites on the molecular scales. Nevertheless, the predictions from the CG model agree well with the MD results over the entire wave number regime. In particular, the good agreement in the high wave number regime shows that the CG model can accurately capture the local roughness of the interface, which is extremely sensitive to the molecule spatial correlations and the many-body interactions.

Next, we examine the meso-scale, size-dependent apolar solvation energy. Similar to the bulk system considered in Sec. 2.3.1, we examine the probability density function of the number of molecule  $P(\hat{n})$  within a spherical volume of radius  $R = 25.0$ . As shown in Fig. 2.5(a), the predictions from the full MD and the CG model agree well over the full regime of  $\hat{n}$ . In particular, at the quasi-equilibrium regime, the interfacial energy is mainly determined by the fluid compressibility;  $P(\hat{n})$  and  $\hat{n}$  follow the quadratic relationship, i.e.,  $P(\hat{n}) \propto (\hat{n} - \langle n \rangle)^2 / \delta n^2$ . Since both  $\hat{n}$  and  $\delta n^2$  scale with the volume, the free energy  $-k_B T \ln P(\hat{n})$  scale with the volume near  $\langle n \rangle$ . In contrast,  $P(\hat{n})$  deviates from the quadratic relationship as  $\hat{n}$  decreases and yields a larger value of  $P(0)$ . The fat tail arises from the formation of a clear void-fluid interface. In particular, on the scale beyond the correlation length of fluid molecules, the local molecular reorganization is insufficient to accommodate the phase separation. Accordingly, the interfacial energy scales with the surface area of the void space.

The multi-faceted nature of the interface energy can be further examined by computing the apolar solvation free energy  $\Delta G = -k_B T \ln P(0)$  for the different sizes of the void space. By the theory of Pratt and his co-worker (Hummer et al., 1996), for the small void space,  $\Delta G$  is governed

by the molecule number fluctuations with the Gaussian distribution, i.e.,

$$\Delta G \approx \frac{1}{2}k_B T \hat{n}^2 / \delta n^2 + \frac{1}{2}k_B T \ln 2\pi \delta n^2, \quad (2.18)$$

where  $\hat{n}^2 / \delta n^2$  scales with the space volume  $4/3\pi R^3$ . On the large scale,  $\Delta G$  is determined by the macro-scale surface tensor  $\gamma$ , i.e.,  $\Delta G \approx 4\pi R^2 \gamma$ .

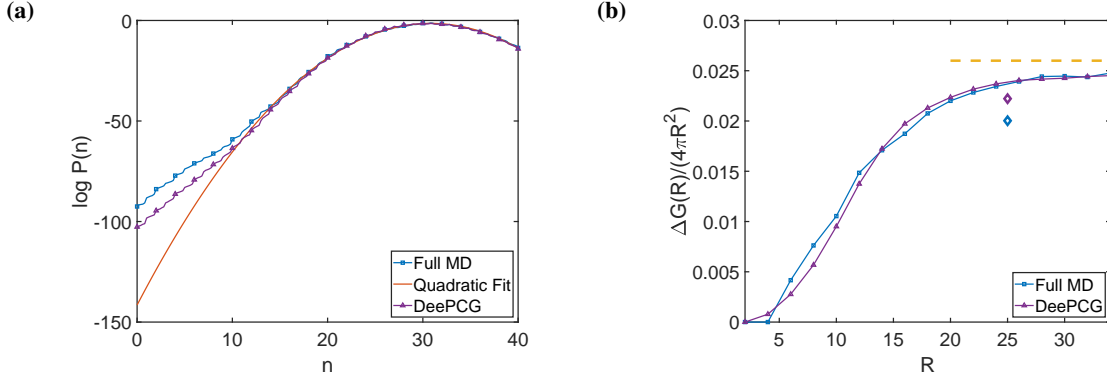


Figure 2.5 (a) The probability density function of the molecule number within a spherical volume of radius  $R = 25.0$ . The red line represents the quadratic fitting; the deviation near  $n = 0$  arises from the formation of a clear fluid-void interface, where free energy approximately scales with the area of the interface. (b) Normalized solvation free energy  $\Delta G(R)/4\pi R^2$  obtained from the thermal integration sampling by Eq. (2.19). The transition from the volume- to area-scaling occurs between  $R = 15$  and  $25$ . The two symbols represent the predictions from the probability of the void space  $-k_B T \ln P(0)$  for  $R = 25$  in (a). The dashed horizontal line represents the macro-scale limit with the surface tensor  $\gamma$  obtained from the fluctuating interface using CWT (Eq. (2.17)) presented in Fig. 2.4.

To quantify the cross-over regime, we conduct the thermal integration sampling of  $\Delta G(R)$  with  $R$  between 0 and 34. The integration force  $\frac{d\Delta G}{dR}$  is estimated by imposing the biased potential, i.e.,

$$\frac{d\Delta G(R)}{dR} = \left\langle \sum_{i=1}^M \nabla_{\mathbf{Q}_i} U_{bias} \cdot \frac{\mathbf{Q}_i - \mathbf{Q}_c}{\|\mathbf{Q}_i - \mathbf{Q}_c\|} \right\rangle, \quad (2.19)$$

where  $U_{bias}(\hat{n}; 0)$  is defined by Eq. (2.13) with  $k_n = 29.20$  and  $h = 0.4$ . Fig. 2.5(b) shows the obtained solvation energy  $\Delta G(R)$  normalized by the surface area. The predictions of the CG and the full MD models show good agreement. In particular, at small value of  $R$ ,  $\Delta G(R)/4\pi R^2$  grows with  $R$  and implies the volume-scaling regime. The transition from the volume- to the area-scaling occurs between  $R = 15$  and  $R = 25$ . For  $R > 30$ ,  $\Delta G(R)/4\pi R^2$  approaches the value of the macro-scale surface tensor  $\gamma$  estimated from the interfacial fluctuations by the CWT theory (2.17) shown in Fig. 2.4.

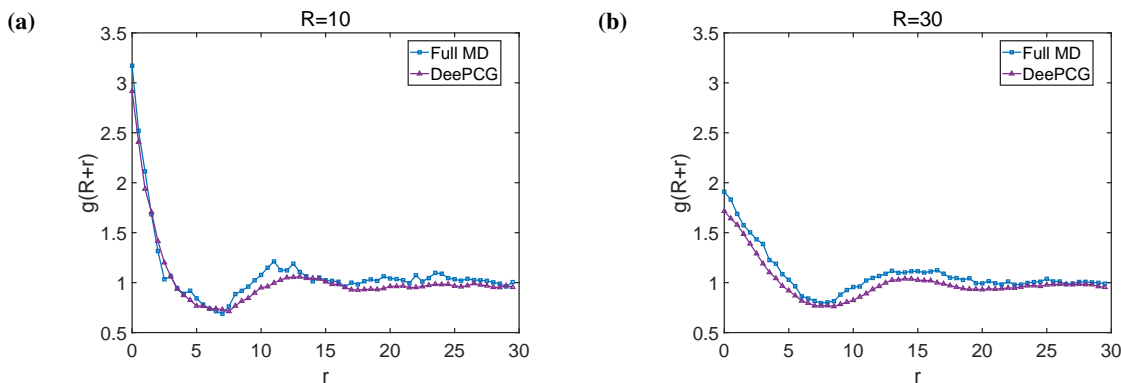


Figure 2.6 The average equilibrium fluid density with a distance  $r + R$ , where  $R$  is the radius of the spherical void space with  $R = 10$  (left) and  $R = 30$  (right).

The scale-dependent interfacial energy is also manifested in the solvent density distribution near the vicinity of the void space. Fig. 2.6 shows the normalized radial distribution function  $g(r + R)$  adjacent to the interface. For  $R = 10$ , solvation is governed by the local compressibility and molecule re-organization, leading to the high fluid density adjacent to the interface. For  $R = 30$ , solvation leads to the clear fluid-void interface and fluid density is closer to the bulk value. The CG model accurately captures the transition and agrees well with the full MD results for both cases.

### 2.3.3 Two-component Fluids

We first consider a two-component fluid system that takes the parameter set (I) specified in Sec. 2.2.1. Therefore, the full MD system can maintain a full mixture state. The reduced model is represented by the CG particles of two different types. The equilibrium state reaches a full mixture state as well. Fig. 2.7 shows the radial distribution functions of the COM of the molecules. Due to the “hydrophilic” interactions between type-1 and -2 molecules, the pair distribution between type 1-2 shows more a pronounced peak at  $R = 11.5$  as compared with the distribution between type 1-1 at  $R = 12.5$ . Similar to Sec. 2.3.1, we compute the angular distribution functions among the molecules of both types. For all of the correlation functions, the CG and full MD models show good agreement.

Next, we consider the parameter set (II) specified in Sec. 2.2.1. Due to the “hydrophobic” interaction between the two molecule types, the system develops into an immiscible state with a clear interface between the two components, as shown in Fig. 2.8(a). To examine the heterogeneous

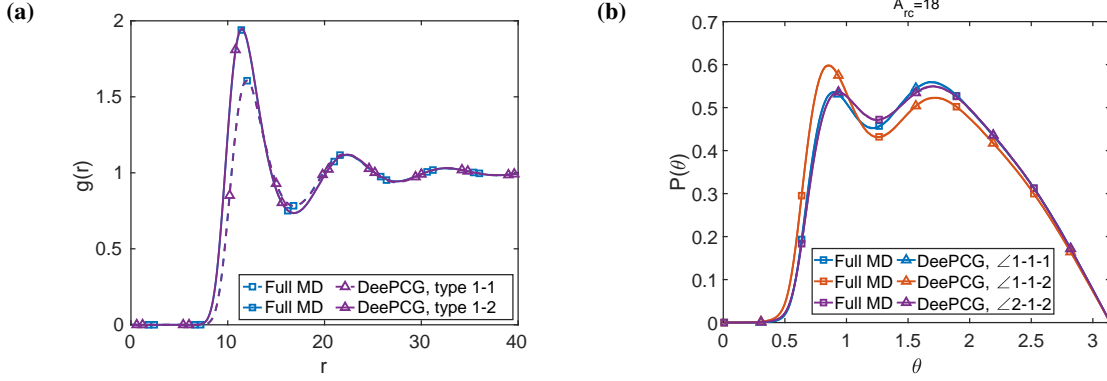


Figure 2.7 (a) Radial distribution function  $g(r)$  of the two-component, miscible polymer fluid system among the molecule COM of type 1-1, type 1-2. (b) Angular distribution function  $P(\theta)$  of the same system among the molecule COM of type 1.

fluid particle distribution, we analyze the radial distribution functions of the fluid particle on the  $x$ - $y$  plane at different regimes. Fig. 2.8(b) shows the planar RDFs sampled at  $z = 60$  (interface) and  $z = 30$  (bulk). In particular, the planar RDF near the interfacial regime shows more pronounced peaks and structural oscillations compared with the RDF in the bulk regime. For both cases, the predictions from the CG model show good agreement with the full MD simulations.

To further quantify the fluid density across the interface, we define the density field  $\rho_s(\mathbf{R})$  by Eq. (2.14) on the lattice grids across the average interface of the two components (i.e., GDS) and the instantaneous height  $\tilde{h}(x, y)$  as the iso-surface of the fluid density of a single component (i.e., IS). For this system, we set  $h = 40$ ,  $dl = 2.0$  and  $dz = 1.0$ . Fig. 2.8(c) shows the density profiles  $\tilde{\rho}(z)$  and  $\rho(z)$  across the interface based on the definition of IS and GDS, respectively. Similar to the single-component fluid system,  $\tilde{\rho}(z)$  shows pronounced oscillations that represent the intrinsic multi-layer fluid structure across the interface. In contrast,  $\rho(z)$  shows a smooth transition across the interface due to the ensemble-averaged definition of the interface plane. The consistent predictions between the MD and CG models validate the accuracy of the constructed DeePCG potential.

Finally, we examine the thermal fluctuations across the interface. Fig. 2.8(d) shows the ensemble average of the Fourier spectrum density  $\langle |\hat{h}(\mathbf{k})|^2 \rangle$  of the instantaneous height  $\tilde{h}(x, y)$  defined by Eq. (2.16). Similar to the single-component interfacial fluid system,  $\langle |\hat{h}(\mathbf{k})|^2 \rangle$  agrees well with the CWT theory at the low wave number and deviates from the  $1/|\mathbf{k}|^2$  scaling at high wave number

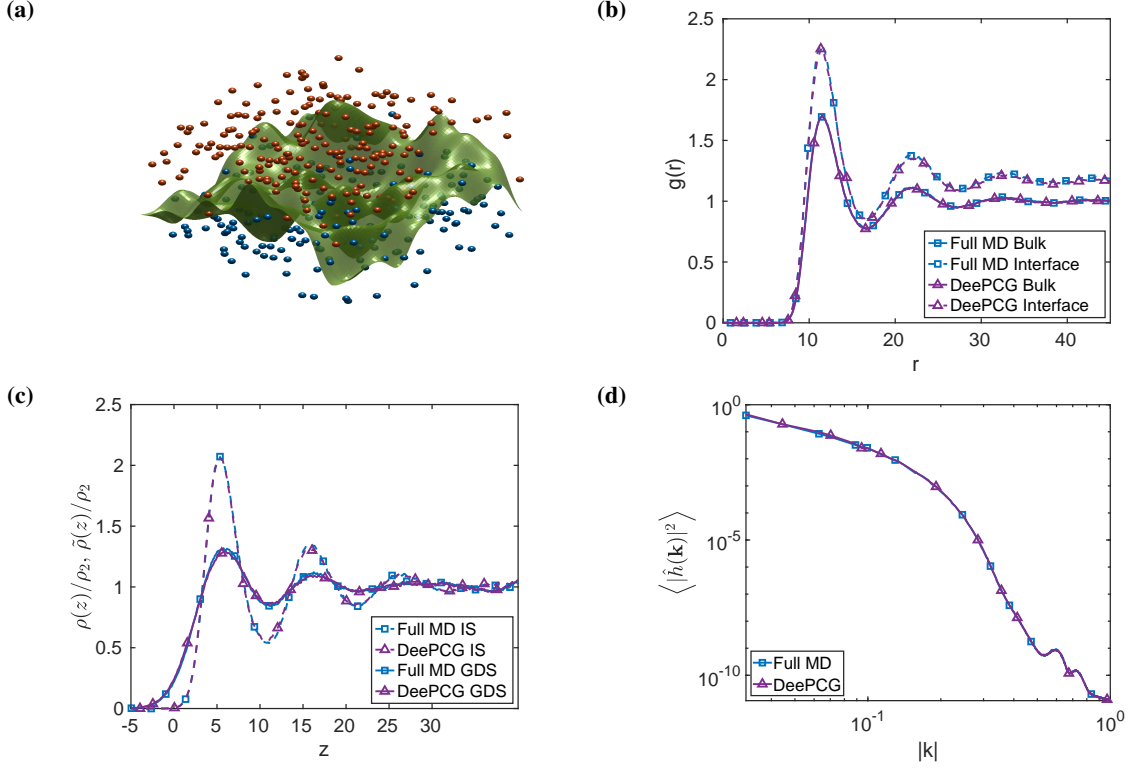


Figure 2.8 The fluid density and the fluctuating interface of the two-component, immiscible fluid system. (a) The interface defined by Eq. (2.16) with type-1 (blue) and type-2 molecules (red). (b) Radial distribution function  $g(r)$  of type-2 molecules on the  $x$ - $y$  plane near the bulk ( $z = 30$ ) and the interface ( $z = 60$ ). (c) The average density profile across the Gibbs dividing surface and the instantaneous surface defined by Eq. (2.15). (d) The capillary wave spectrum of the fluctuating interface.

due to the local spatial correlations between the molecules. The predictions from CG and full MD models show good agreement over the full regime.

## 2.4 Summary

In this study, we constructed coarse-grained models of meso-scale interfacial polymeric fluids based on the DeePCG scheme (Zhang et al., 2018b). In particular, the constructed CG potential can accurately encode the many-body interactions arising from the unresolved atomistic interactions, as well as the heterogeneous molecule distributions near the interface. This unique feature ensures that the constructed CG models can retain the consistent invariant distribution with the full MD model and faithfully capture the multi-facted, scale-dependent interfacial energy without additional human intervention. The training process only requires the MD samples of the instantaneous force field without further ad hoc assumptions and approximations of the CG potential functions.



While we focus on the polymeric fluids in this study, the present CG models can be generalized for complex fluids and soft matter systems where the many-body and heterogeneous effects are often pronounced. In particular, the constructed CG potential functions accurately reproduce the pairwise and high-order correlation functions while the empirical approximations show limitations. Moreover, the accurate predictions of the local compressibility and the full-range spectrum of the interfacial fluctuations demonstrate the validity of the CG models to probe the collective behaviors across the molecular and continuum scales. More importantly, the CG models successfully predict the probability of the void formation as a rare event and the transition of the volume- to area-scaling of solvation energy. The accurate predictions on such properties show the promise of the present models to study the challenging problems relevant to nanoscale assembly processes (Miller et al., 2007), where the full MD simulations often show limitation to achieve the resolved spatio-temporal scale.

Finally, we note that the present study focuses on the quasi-equilibrium properties of the reduced model. The zero-rate shear viscosity predicted by the DeePCG model is 55.56% less than the value of the full MD model. The predictive modeling of the dynamic properties further relies on the accurate construction of the memory and fluctuation terms that represent the unresolved energy-dissipation processes (Hijón et al., 2010; Lei et al., 2016; Lei and Li, 2021; She et al., 2023). Also, it is worth exploring the construction of CG potential function with certain generalization abilities that account for the different temperature (Zhang et al., 2020) and model resolution (Empereur-mot et al., 2022). We will pursue these problems in future studies.

## CHAPTER 3

### DATA-DRIVEN LEARNING OF THE STATE-DEPENDENT MEMORY KERNEL

In this chapter, we focus on how to construct the memory kernel. In Ge et al. (2024), we present a data-driven method to learn stochastic reduced models of complex systems that retain a state-dependent memory beyond the standard generalized Langevin equation (GLE) with a homogeneous kernel. The constructed model naturally encodes the heterogeneous energy dissipation by jointly learning a set of state features and the non-Markovian coupling among the features. Numerical results demonstrate the limitation of the standard GLE and the essential role of the broadly overlooked state-dependency nature in predicting molecule kinetics related to conformation relaxation and transition.

#### 3.1 Introduction

Predicting the collective behavior of complex multiscale systems is often centered around projecting the full-dimensional dynamics onto a set of resolved variables. However, an accurate construction of such a reduced model remains a practical challenge for real applications such as molecular modeling. While model reduction frameworks such as the Koopman operator (Koopman, 1931) and the Mori-Zwanzig projection formalism (Mori, 1965; Zwanzig, 1961) enable us to write down the dynamic equations in terms of the resolved variables, the reduced model generally becomes non-Markovian with a memory term that may further depend on the resolved variables; the direct numerical evaluation involves solving the expensive full-dimensional orthogonal dynamics. In practice, one common approximation is to ignore such state-dependency; the reduced model is simplified as the standard generalized Langevin equation (GLE) (Zwanzig, 2001) with a memory kernel that only depends on time. Several approaches (Lange and Grubmüller, 2006; Darve et al., 2009; Ceriotti et al., 2009; Baczewski and Bond, 2013; Davtyan et al., 2015; Lei et al., 2016; Russo et al., 2019; Jung et al., 2017; Lee et al., 2019; Ma et al., 2019; Wang et al., 2020; Zhu and Venturi, 2020; Klippenstein and van der Vegt, 2021; Vroylandt et al., 2022; She et al., 2023; Xie et al., 2022) have been developed to construct the memory kernel such that certain dynamic properties (e.g., the two-point correlations) can be properly reproduced. Despite its broad application, the validity of

the standard GLE for real multiscale systems remains less understood (Hänggi, 1997; Klippenstein et al., 2021).

Intuitively, the above model reduction problem is somewhat analogous to hiking on a mountain where the landscape map and the path roughness represent the free energy and the memory term, respectively. In general, we should not expect homogeneous path roughness at the different locations (e.g., the valleys and the ridges), which, conversely, needs to be inferred from the hiking records. Indeed, studies based on full molecular dynamics (MD) simulations (Posch et al., 1984; Straub et al., 1987, 1990; Plotkin and Wolynes, 1998; Luo et al., 2006; Best and Hummer, 2006, 2010; Hinczewski et al., 2010; Satija et al., 2017; Morrone et al., 2012; Daldrop et al., 2017) and sophisticated projection operator construction (Deutch and Oppenheim, 1971; Zwanzig, 1973, 1992; Berezhkovskii and Szabo, 2011; Glatzel and Schilling, 2022; Vroylandt, 2022; Vroylandt and Monmarché, 2022; Ayaz et al., 2022a; Jung and Jung, 2023) show that the extracted memory term can exhibit a pronounced state-dependent nature, where the implications for the collective behaviors remain under-explored. For extensive MD systems, a recent study (Lyu and Lei, 2023) on reduced modeling of polymer melt shows that the heterogeneous inter-molecular energy dissipation (i.e., the memory) can be crucial for transport on the hydrodynamic scale. However, for canonical non-extensive problems such as biomolecule systems, a quantitative understanding of the state-dependent memory effect on the reduced dynamics remains an open problem. Several recent works (Lei et al., 2016; Lee et al., 2019; Satija and Makarov, 2019; Grogan et al., 2020; Singh et al., 2021; Ayaz et al., 2021; Vroylandt et al., 2022; Dalton et al., 2023) model the non-Markovian effect for transition dynamics based on the standard GLE. While elegant semi-analytical studies (Straub et al., 1988; Singh et al., 1990; Carmeli and Nitzan, 1983; Tarjus and Kivelson, 1991; Krishnan et al., 1992; Voth, 1992; Straus et al., 1993; Haynes et al., 1993, 1994; Cossio et al., 2015) on idealized 1D double-well potential provide theoretical insights into the state-dependent nature, quantitative modeling that retains the reduced dynamics consistent with the full MD model, including collective properties such as transition and conformation relaxation, relies on accurate construction and efficient simulation of a reduced model beyond the standard GLE.

This work presents a data-driven approach for learning a new stochastic reduced model that retains a state-dependent memory for non-extensive systems. Instead of dealing with the orthogonal dynamics (Darve et al., 2009; Vroylandt and Monmarché, 2022; Lyu and Lei, 2023), the training only relies on the trajectory samples and does not directly solve the Mori-Zwanzig projection formalism. The main idea is to seek a generalized representation of the memory as the composition of a set of state-dependent features, which encodes the coupling between the resolved and unresolved variables and will be learned using three-point correlation functions. Efficient training is achieved by constructing the encoders using a set of sparse bases, whose correlations can be efficiently pre-computed. The time-dependent component is directly learned in the Fourier space which enables the efficient evaluation of the convolution term via the FFT and meanwhile ensures non-negative energy dissipation (i.e., model stability). To simulate the model, coherent noise can be introduced that strictly satisfies the second fluctuation-dissipation theorem (FDT) and retains a consistent invariant distribution. The present model, with a new memory form, essentially reveals a caveat in model reduction of multiscale systems and provides a reliable approach for simulating the stochastic reduced dynamics beyond empirical models. It enables us to probe open problems such as the effect of state-dependent memory on molecular kinetics. Numerical results show that the broadly overlooked state-dependency can play a crucial role. In particular, the standard GLE is insufficient to capture the collective properties such as conformation relaxation and transition rate distribution, which, fortunately, can be reproduced by the present model.

### 3.2 Model Derivation

Let  $(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2m}$  represent the resolved variables of a high-dimensional Hamiltonian system, where  $\mathbf{q}$  denotes the coarse-grained (CG) coordinates as a function of the position variables of the full model, and  $\mathbf{p}$  denotes the CG momenta. Following the Zwanzig’s formalism (Zwanzig, 2001; Hijón et al., 2010), the reduced dynamics takes the form

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{M}^{-1} \mathbf{p}, \\ \dot{\mathbf{p}} &= -\nabla U(\mathbf{q}) - \int_0^t \mathbf{K}(\mathbf{q}(\tau), t - \tau) \mathbf{v}(\tau) d\tau + \mathcal{R}_t,\end{aligned}\tag{3.1}$$

where  $\mathbf{M}$  is the mass matrix,  $U(\mathbf{q})$  is the free energy,  $\mathbf{v} := \dot{\mathbf{q}}$  is the velocity,  $\mathbf{K}(\mathbf{q}, t)$  is the memory, and  $\mathcal{R}_t$  is the noise whose covariance function is related to the memory following the second FDT (Vroylandt and Monmarché, 2022). Before proceeding to the construction of  $\mathbf{K}(\mathbf{q}, t)$ , we note that the rigorous form based on Zwanzig’s formalism depends on both  $\mathbf{q}$  and  $\mathbf{p}$ . Here we focus on the state-dependence on  $\mathbf{q}$  and assume it is independent of  $\mathbf{p}$  (Hijón et al., 2010). Furthermore,  $\mathbf{M}$  generally depends on  $\mathbf{q}$ ; the current choice of  $\mathbf{q}$  leads to a constant mass matrix [see Refs. (Lee et al., 2019; Ayaz et al., 2022a) and Section 3.5.1]. Also, the construction of the free energy  $U(\mathbf{q})$  can be nontrivial; several canonical methods based on enhanced sampling (Torrie and Valleau, 1977; Kumar et al., 1992a; Darve and Pohorille, 2001; Laio and Parrinello, 2002) and temperature acceleration (Rosso et al., 2002; Maragliano and Vanden-Eijnden, 2006; Abrams and Tuckerman, 2008; Maragliano and Vanden-Eijnden, 2008) have been developed to facilitate the phase space exploration. We assume the phase space can be effectively explored and  $U(\mathbf{q})$  is known a priori.

Instead of rigorously constructing  $\mathbf{K}(\mathbf{q}, t)$  from the full model, we ask the question of which forms of  $\mathbf{K}$  can generate a memory effect. One common approach is to embed the memory in a larger Markovian dynamics with a set of auxiliary variables. An essential observation is that the memory term can be generally written as

$$\mathbf{K}(\mathbf{q}(\tau), t - \tau) \approx C^+ \circ \exp((t - \tau)\mathcal{L}_{\text{aux}}) \circ C^-, \quad (3.2)$$

where  $\mathcal{L}_{\text{aux}}$  is the Liouville operator corresponding to the auxiliary dynamics and  $C^\pm$  are channels representing the coupling of the resolved and auxiliary variables. As a special case, if the coupling and the auxiliary dynamics take a linear form, the embedded memory recovers the standard GLE kernel, i.e.,  $\mathbf{K}(\mathbf{q}, t) = \mathbf{K}(t)$  (e.g., see Refs. (Lei and Li, 2021; She et al., 2023)). Therefore, to construct the reduced model beyond the standard GLE, the coupling channels need to properly retain certain kinds of state-dependency nature. This motivates us to represent  $C^\pm$  by seeking a set of state-dependent features  $\phi(\mathbf{q}) = [\phi_1(\mathbf{q}), \dots, \phi_n(\mathbf{q})]$ , where  $\phi : \mathbb{R}^m \rightarrow \mathbb{R}^{n \times m}$  essentially encode the nonlinear coupling between the resolved and unresolved variables and the detailed form will be specified later.  $\exp(t\mathcal{L}_{\text{aux}})$  induces the non-Markovian interactions among the features with a time lag  $t$  characterized by a kernel function, i.e.,  $C^+ \circ \exp((t - \tau)\mathcal{L}_{\text{aux}}) \circ C^- = \phi(\mathbf{q}(t))^T \Theta(t - \tau) \phi(\mathbf{q}(\tau))$ ,

where  $\Theta : \mathbb{R}^+ \rightarrow \mathbb{R}^{n \times n}$  and component  $\Theta_{ij}(t - \tau)$  represents the dissipation between features  $\phi_i(\mathbf{q}(t))$  and  $\phi_j(\mathbf{q}(\tau))$ . In the remainder of this work, we use  $\phi_t$  to denote  $\phi(\mathbf{q}(t))$ .

With the above observation, we propose the following form to model the reduced dynamics (3.1), i.e.,

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{M}^{-1} \mathbf{p}, \\ \dot{\mathbf{p}} &= -\nabla U(\mathbf{q}) - \int_0^t \phi_t^T \Theta(t - \tau) \phi_\tau \mathbf{v}(\tau) d\tau + \mathcal{R}_t,\end{aligned}\tag{3.3}$$

where encoders  $\{\phi_i(\mathbf{q})\}_{i=1}^n$  and kernel  $\Theta(t)$  need to be determined. As a special case, at the Markovian limit  $\Theta(t) \propto \delta(t)$ , Eq. (3.3) recovers the Langevin dynamics and the quadratic form  $\phi^T \phi$  ensures positive energy dissipation. Also, by choosing  $\Theta(t)$  to be diagonal with individual components corresponding to certain frequency modes, Eq. (3.3) reduces to the heat bath model (Zwanzig, 1973) with a nonlinear coupling of bath coordinates. On the other hand, the present model enables an adaptive choice of the number of spatial features and a more general form of  $\Theta(t)$  with the off-diagonal components capturing the non-Markovian coupling among the features, which turns out to be crucial for reproducing the collective dynamics.

### 3.2.1 Coherent Noise and Invariant Density of the Reduced Model

We emphasize that Eq. (3.3) should not be viewed as a direct approximation of Zwanzig's projection formalism. Rather, it serves as a reduced model that faithfully retains the state-dependent memory effect. To construct the model, we represent encoders  $\{\phi_i(\mathbf{q})\}_{i=1}^n$  and kernel  $\Theta(t)$  in form of

$$\begin{aligned}\phi_i(\mathbf{q}) &= \mathbf{H}_i^T \psi(\mathbf{q}), \\ \Theta(t) &= e^{-\alpha t} \sum_{k=0}^{N_\omega} \hat{\Theta}_k \cos(\omega_k t),\end{aligned}\tag{3.4}$$

where  $\psi(\mathbf{q}) = [\psi_1(\mathbf{q}), \dots, \psi_{N_b}(\mathbf{q})]$  is a set of sparse bases and  $\mathbf{H} = [\mathbf{H}_1^T, \dots, \mathbf{H}_n^T]$  are trainable coefficients,  $\omega_k = \frac{2\pi}{T_c} k$  and  $T_c$  is the time domain cut-off of the kernel. We note that  $e^{-\alpha t}$  should not be viewed as the bases to approximate  $\Theta(t)$  (e.g.,  $\{e^{-\alpha_i t} \cos(\beta_i t), e^{-\alpha_i t} \sin(\beta_i t)\}_{i=1}^{N_\alpha}$ ; see Refs. (Lei et al., 2016; Lee et al., 2019)). Rather,  $\Theta(t)$  is mainly characterized by the Fourier series expansion on  $[0, T]$ , and the exponential term  $e^{-\alpha t}$  is essentially a regularization term to eliminate

the periodicity while maintaining the semi-positive definiteness condition.  $\Theta(t)$  needs to preserve positive semi-definiteness. Hence, we represent Fourier modes  $\hat{\Theta}_k = \Gamma_k \Gamma_k^T$ , where  $\Gamma_k \in \mathbb{R}^{n \times n}$  is a low-triangular matrix to be determined along with  $\alpha \geq 0$ .

For the fluctuation term  $\mathcal{R}_t$ , we represent it as a noise in the form of  $\mathcal{R}_t = \phi_t^T \tilde{\mathbf{R}}(t)$ , where  $\tilde{\mathbf{R}}(t)$  is a Gaussian random process whose covariance function determined by  $\Theta(t)$ , i.e.,  $\langle \tilde{\mathbf{R}}(t) \tilde{\mathbf{R}}(\tau)^T \rangle = k_B T \Theta(t - \tau)$ . This choice avoids dealing with the orthogonal dynamics to calculate the fluctuation term. Furthermore, we can show that this choice enables the reduced model to retain a consistent invariant density function.

**Proposition 3.2.1.** For reduced model (3.3) with  $\Theta(t) = e^{-\alpha t} \sum_{k=0}^{N_\omega} \hat{\Theta}_k \cos(\omega_k t)$ , by choosing the fluctuation term  $\mathcal{R}_t = \phi_t^T \tilde{\mathbf{R}}(t)$ , where  $\tilde{\mathbf{R}}(t)$  is a Gaussian random process satisfying

$$\langle \tilde{\mathbf{R}}(t) \tilde{\mathbf{R}}(\tau)^T \rangle = k_B T \Theta(t - \tau), \quad (3.5)$$

the reduced model has an invariant distribution

$$\rho_{\text{eq}}(\mathbf{q}, \mathbf{p}) \propto \exp \left\{ - \left[ U(\mathbf{q}) + \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} / 2 \right] / k_B T \right\}. \quad (3.6)$$

**Proof.** Let us introduce auxiliary variables

$$\begin{aligned} \mathbf{z}_{k,1} &= - \int_0^t e^{-\alpha(t-\tau)} \Gamma_k \cos(\omega_k(t-\tau)) \phi_\tau^T \mathbf{v}(\tau) d\tau + \mathbf{R}_{k,1}(t), \\ \mathbf{z}_{k,2} &= - \int_0^t e^{-\alpha(t-\tau)} \Gamma_k \sin(\omega_k(t-\tau)) \phi_\tau^T \mathbf{v}(\tau) d\tau + \mathbf{R}_{k,2}(t), \end{aligned} \quad (3.7)$$

where  $\Gamma_k^T \Gamma_k = \hat{\Theta}_k$  and  $\mathbf{R}_{k,1}(t)$  is a Gaussian random process satisfying

$$\langle \mathbf{R}_{j,1}(t) \mathbf{R}_{k,1}(\tau)^T \rangle = k_B T \delta_{jk} e^{-\alpha(t-\tau)} \cos(\omega_k(t-\tau)), \quad (3.8)$$

where  $\delta_{jk}$  is the Kronecker delta. Accordingly, the second equation of Eq. (3.3) can be written as

$$\dot{\mathbf{p}} = -\nabla U(\mathbf{q}) + \phi(\mathbf{q})^T \sum_k \Gamma_k^T \mathbf{z}_{k,1}, \quad (3.9)$$

and  $\mathbf{R}_{j,2}(t)$  will be specified later.

Let  $\mathbf{z}_k = [\mathbf{z}_{k,1}, \mathbf{z}_{k,2}]$  and  $\mathbf{R}_k = [\mathbf{R}_{k,1}, \mathbf{R}_{k,2}]$ , we can rewrite Eq. (3.7) by

$$\begin{aligned}\mathbf{z}_k &= - \int_0^t e^{-\alpha(t-\tau)} \begin{pmatrix} \cos(\omega_k(t-\tau))I & \sin(\omega_k(t-\tau))I \\ -\sin(\omega_k(t-\tau))I & \cos(\omega_k(t-\tau))I \end{pmatrix} \begin{pmatrix} \Gamma_k \phi_\tau \mathbf{v}(\tau) \\ 0 \end{pmatrix} d\tau + \mathbf{R}_k(t) \\ &= - \int_0^t \exp \left[ \begin{pmatrix} -\alpha I & \omega_k I \\ -\omega_k I & -\alpha I \end{pmatrix} (t-\tau) \right] \begin{pmatrix} \Gamma_k \phi_\tau \mathbf{v}(\tau) \\ 0 \end{pmatrix} d\tau + \mathbf{R}_k(t).\end{aligned}\quad (3.10)$$

By taking the time derivative of Eq. (3.10) with respect to  $t$ , we have

$$\frac{d\mathbf{z}_k}{dt} = \underbrace{\begin{pmatrix} -\alpha I & \omega_k I \\ -\omega_k I & -\alpha I \end{pmatrix}}_{\triangleq \mathbf{J}} \mathbf{z}_k - \begin{pmatrix} \Gamma_k \phi_\tau \mathbf{v}(t) \\ 0 \end{pmatrix} + \frac{d\mathbf{R}_k}{dt} - \mathbf{J}\mathbf{R}_k(t). \quad (3.11)$$

Furthermore, we note that  $\mathbf{R}_k(t)$  can be modeled as a generalized Ornstein–Uhlenbeck process and

$\frac{d\mathbf{R}_k}{dt} - \mathbf{J}\mathbf{R}_k(t)$  can be represented by

$$\frac{d\mathbf{R}_k}{dt} - \mathbf{J}\mathbf{R}_k(t) = \Lambda_k \dot{\mathbf{W}}_{k,t}, \quad (3.12)$$

where  $\dot{\mathbf{W}}_{k,t}$  is the standard white noise and  $\Lambda_k \Lambda_k^T = -k_B T (\mathbf{J} + \mathbf{J}^T)$ . With this choice, the covariance of  $\mathbf{R}_k(t) = [\mathbf{R}_{k,1}, \mathbf{R}_{k,2}]$  is given by

$$\langle \mathbf{R}_k(t) \mathbf{R}_k(\tau)^T \rangle = k_B T e^{-\alpha(t-\tau)} \begin{pmatrix} \cos(\omega_k(t-\tau))I & \sin(\omega_k(t-\tau))I \\ -\sin(\omega_k(t-\tau))I & \cos(\omega_k(t-\tau))I \end{pmatrix}$$

such that Eq. (3.8) remains valid. Using Eqs. (3.7)(3.9)(3.11), we can write the reduced model (3.3) in the form of

$$\begin{aligned}\frac{d}{dt} \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \\ \cdots \\ \mathbf{z}_{k,1} \\ \mathbf{z}_{k,2} \\ \cdots \end{pmatrix} &= \begin{pmatrix} 0 & I & \cdots & 0 & 0 & \cdots \\ -I & 0 & \cdots & \phi(\mathbf{q})^T \Gamma_k^T & 0 & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & -\Gamma_k \phi(\mathbf{q}) & \cdots & -\alpha I & -\omega_k I & \cdots \\ 0 & 0 & \cdots & \omega_k I & -\alpha I & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \nabla U(\mathbf{q}) \\ \mathbf{v} \\ \cdots \\ \mathbf{z}_{k,1} \\ \mathbf{z}_{k,2} \\ \cdots \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \cdots \\ \Lambda_k \dot{\mathbf{W}}_{k,t} \\ \cdots \end{pmatrix} \\ &\triangleq \mathbf{K} \nabla F(\mathbf{q}, \mathbf{p}, \cdots, \mathbf{z}_{k,1}, \mathbf{z}_{k,2}, \cdots) + \Lambda \dot{\mathbf{W}}_t,\end{aligned}\quad (3.13)$$



where  $\mathbf{K}$  is the first matrix of the right-hand-side of Eq. (3.13),  $F(\mathbf{q}, \mathbf{p}, \dots, \mathbf{z}_{k,1}, \mathbf{z}_{k,2}, \dots) = U(\mathbf{q}) + \frac{1}{2}\mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} + \frac{1}{2} \sum_{k=1}^{N_\omega} (\mathbf{z}_{k,1}^T \mathbf{z}_{k,1} + \mathbf{z}_{k,2}^T \mathbf{z}_{k,2})$  is the total free energy of the extended system, and  $\Lambda = \text{diag}(0, 0, \dots, \Lambda_k, \dots)$ . Using (3.12), it is easy to show  $\Lambda \Lambda^T = -k_B T (\mathbf{K} + \mathbf{K}^T)$ . Therefore, the gradient system (3.13) (i.e., the reduced model (3.3)) has the invariant density function

$$\rho_{\text{eq}}(\mathbf{q}, \mathbf{p}, \mathbf{z}) = \exp[-F(\mathbf{q}, \mathbf{p}, \mathbf{z})/k_B T].$$

### 3.2.2 Training Method of the Reduced Model

To learn the reduced model (3.3), we need to choose appropriate metrics such that the state-dependent non-Markovian nature can be manifested. While auto-correlations such as  $c_{vv}(t) = \langle \mathbf{v}(t) \mathbf{v}(0)^T \rangle$  merely characterize the overall memory effect, a crucial observation is that the correlation conditional with different initial state  $\mathbf{q}_0$  further depends on the local energy dissipation and therefore naturally encode the signatures of the heterogeneous memory effect. Accordingly, we right-multiply the second equation of (3.3) by  $\mathbf{v}(0)$  and take the conditional expectation on  $\mathbf{q}_0 = \mathbf{q}^*$ , i.e.,

$$\begin{aligned} \mathbf{g}(t; \mathbf{q}^*) &= \int_0^t \langle \phi_t^T \Theta_{t-\tau} \phi_\tau \mathbf{v}_\tau \mathbf{v}_0^T | \mathbf{q}_0 = \mathbf{q}^* \rangle d\tau \\ &= \int_0^t \langle \text{Tr} [\Theta_{t-\tau} \mathbf{H} \psi_\tau \mathbf{v}_\tau \mathbf{v}_0^T \psi_t^T \mathbf{H}^T] | \mathbf{q}_0 = \mathbf{q}^* \rangle d\tau \\ &= \int_0^t \text{Tr} [\Theta_{t-\tau} \mathbf{H} \mathbf{C}_{\psi, \psi}(t, \tau; \mathbf{q}^*) \mathbf{H}^T] d\tau, \end{aligned}$$

where  $\mathbf{g}(t; \mathbf{q}^*) := \langle [\dot{\mathbf{p}}_t + \nabla U(\mathbf{q}_t)] \mathbf{v}_0^T | \mathbf{q}_0 = \mathbf{q}^* \rangle$  and  $\mathbf{C}_{\psi, \psi}(t, \tau; \mathbf{q}^*) := \langle \psi_\tau \mathbf{v}_\tau \mathbf{v}_0^T \psi_t^T | \mathbf{q}_0 = \mathbf{q}^* \rangle$  is a three-point correlation characterizing the coupling among the bases. Since  $\psi(\mathbf{q})$  is sparse,  $\psi_\tau \psi_t^T$  can be evaluated with  $O(1)$  complexity and hence  $\mathbf{C}_{\psi, \psi}(t, \tau; \mathbf{q}^*)$  can be efficiently pre-computed. Accordingly, we can train the reduced model in terms of coefficients  $\mathbf{H}$  for encoders  $\phi(\mathbf{q})$  as well as matrices  $\{\Gamma_k\}_{k=1}^{N_\omega}$  and  $\alpha$  for kernel  $\Theta(t)$  by minimizing the empirical loss

$$\begin{aligned} L &= \sum_{l=1}^{N_q} \sum_{k=1}^{N_t} \left\| \tilde{\mathbf{g}}(t_k; \mathbf{q}^{(l)}) - \mathbf{g}(t_k; \mathbf{q}^{(l)}) \right\|^2, \\ \tilde{\mathbf{g}}(t_k; \mathbf{q}^{(l)}) &= \sum_{j=1}^k \text{Tr} [\Theta(t_k - t_j) \mathbf{H} \mathbf{C}_{\psi, \psi}(t_k, t_j; \mathbf{q}^{(l)}) \mathbf{H}^T] \delta t, \end{aligned} \tag{3.14}$$

where  $\{\mathbf{q}^{(l)}\}_{l=1}^{N_q}$  represent configuration samples within the phase space. For systems with pronounced free energy barriers,  $\mathbf{q}^{(l)}$  can be collected along with free energy construction (Maragliano and Vanden-Eijnden, 2008), whereas the conditional correlations for each  $\mathbf{q}^{(l)}$  need to be sampled from unbiased equilibrium trajectories.  $\widetilde{\mathbf{g}}(\cdot)$  represents the prediction by the reduced model which depends on the trainable variables and the pre-computed correlation  $\mathbf{C}_{\psi,\psi}$ .  $\delta t$  is the time step.

Besides the conditional correlation functions, we can also introduce the loss function with respect to the overall correlation function, i.e.,

$$L_2 = \sum_{k=1}^{N_t} \|\widetilde{\mathbf{g}}_2(t_k) - \mathbf{g}_2(t_k)\|^2$$

$$\widetilde{\mathbf{g}}_2(t_k) = \sum_{j=1}^k \text{Tr} \left[ \Theta(t_k - t_j) \mathbf{H} \overline{\mathbf{C}}_{\psi,\psi}(t_k, t_j) \mathbf{H}^T \right] \delta t,$$

where  $\mathbf{g}_2(t) = \langle [\dot{\mathbf{p}}_t + \nabla U(\mathbf{q}_t)] \mathbf{v}_0^T \rangle$  is the overall correlation and  $\overline{\mathbf{C}}_{\psi,\psi}(t, \tau) = \langle \psi_\tau \mathbf{v}_\tau \mathbf{v}_0^T \psi_t^T \rangle$ . In particular, if there is scale separation between  $c_{vv}(t)$  and  $c_{qq}(t)$  (e.g.,  $c_{vv}(t)$  decays much faster than  $c_{qq}(t)$ ; see Fig. 3.4 and Fig. 3.5), we may approximate  $\overline{\mathbf{C}}_{\psi,\psi}(t, \tau)$  by two-point correlations, i.e.,  $\overline{\mathbf{C}}_{\psi,\psi}(t, \tau) \approx \langle \psi_\tau \otimes \psi_t^T \rangle : \langle \mathbf{v}_\tau \mathbf{v}_0^T \rangle$ .

Efficient training is achieved by using the following numerical methods to evaluate  $\widetilde{\mathbf{g}}$  and  $\widetilde{\mathbf{g}}_2$ . Specifically,  $\psi_\tau \psi_t^T$  can be efficiently pre-computed with  $O(1)$  complexity by using the sparse piecewise linear basis functions. Furthermore, we can use the low-rank representation (e.g., based on the singular value decomposition) of  $\mathbf{C}_{\psi,\psi}$  and  $\overline{\mathbf{C}}_{\psi,\psi}$  to accelerate the matrix production  $\mathbf{H} \mathbf{C}_{\psi,\psi} \mathbf{H}^T$ . In addition, the convolution on index  $j$  can be efficiently evaluated by the Fast Fourier Transform algorithm (Cooley and Tukey, 1965).

For the present study,  $\psi$  is chosen as the uniform piecewise linear basis function defined on  $[2.8, 4.1]$  with  $N_b = 66$ ,  $T_c = 200$  and  $N_\omega = 2000$ . We use  $2 \times 10^6$  short trajectories extracted from the full MD simulation (see Sec. 3.3.1 for details) where each one consists of 300 time-series samples to sample the correlation for  $\mathbf{q}^*$  at the saddle point. We use  $2 \times 10^6 \sim 8 \times 10^6$  short trajectories for other individual points.

While  $L_2$  alone is insufficient to characterize the emergence of the state-dependent memory, it serves as a necessary condition and can facilitate the learning of the reduced model. In practice,

we can use both loss functions to train the reduced model with  $N_q = 65$ ,  $N_t = 300$  for  $L$ , and  $N_t = 30000$  for  $L_2$ . Specifically, the training is conducted by the Adam (Kingma and Ba, 2015) optimization method in three stages with 2000, 6000, and 6000 steps respectively. For the first stage, we only use  $L_2$  to train the model with a constant learning rate of 0.04. For each step of the following two stages, 16 initial states (i.e.,  $q^{(l)}$ ) are randomly selected as one training batch to evaluate the total loss  $L_t = L + L_2$ . For both stages, the initial learning rate is  $1 \times 10^{-2}$  and the exponential decay rate is 0.9 per 150 steps. In practice, the number of  $N_b$  and  $N_q$  can be further reduced. As shown in the GitHub repository, with  $N_b = 8$ ,  $N_q = 26$ , the constructed reduced model can accurately recover the MD results.

### 3.2.3 Simulation of the Reduced Model

We simulate the reduced model (3.3) on  $t \in [0, T]$  by generating the noise term  $\mathcal{R}_t = \phi_t^T \tilde{\mathbf{R}}(t)$ , where  $\tilde{\mathbf{R}} : \mathbb{R}^+ \rightarrow \mathbb{R}^n$  is a Gaussian random process, using the constructed Fourier modes rather than Markovian embedding (e.g., see Ref. (Ayaz et al., 2022b)). Specifically, we have proved that by choosing  $\langle \tilde{\mathbf{R}}(t) \tilde{\mathbf{R}}(\tau)^T \rangle = k_B T \Theta(t - \tau)$ , the reduced model retains a consistent equilibrium density, i.e.,  $\rho_{\text{eq}}(\mathbf{q}, \mathbf{p}) \propto \exp \left\{ -\beta \left[ U(\mathbf{q}) + \frac{1}{2} \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p} \right] \right\}$  (Prop. 3.2.1). Accordingly, we can generate  $\{\tilde{\mathbf{R}}(t_i)\}_{i=0}^N$  similar to Refs. (Berkowitz et al., 1983; Ogorodnikov and Prigarin, 1996) by

$$\tilde{\mathbf{R}}(t_i) = \beta^{-1/2} \sum_{k=0}^{2N} \tilde{\Theta}_k^{1/2} [\cos(\omega_k t_i) \xi_k + \sin(\omega_k t_i) \eta_k], \quad (3.15)$$

where  $\tilde{\Theta}_k$  are the Fourier (essentially cosine) modes of  $\Theta(|t|)$  on  $[-T, T]$  (Berkowitz et al., 1983; Ogorodnikov and Prigarin, 1996);  $\xi_k$  and  $\eta_k$  are independent Gaussian random vectors. Specifically, for large simulation time  $T$ , the Fourier modes  $\tilde{\Theta}_k$  is given by

$$\begin{aligned} \tilde{\Theta}_k &= \sum_{j=1}^{N_\omega} \int_{-T}^T e^{-\alpha|t|} \hat{\Theta}_j \cos(\omega_j t) \cos(\omega_k t) dt \\ &= \sum_{j=1}^{N_\omega} \int_0^T e^{-\alpha t} \hat{\Theta}_j (\cos((\omega_j - \omega_k)t) + \cos((\omega_j + \omega_k)t)) dt \\ &\approx \sum_{j=1}^{N_\omega} \left( \frac{\alpha \hat{\Theta}_j}{\alpha^2 + (\omega_j - \omega_k)^2} + \frac{\alpha \hat{\Theta}_j}{\alpha^2 + (\omega_j + \omega_k)^2} \right). \end{aligned} \quad (3.16)$$

Therefore  $\tilde{\mathbf{R}}(t)$  can be generated using the Fast Fourier Transform algorithm (Cooley and Tukey, 1965) using  $O(N \log N)$  complexity. Also, the convolution term  $\int_0^t \phi_i^T \Theta(t - \tau) \phi_\tau \mathbf{v}(\tau) d\tau$  in Eq. (3.3) can be computed using the fast convolution method developed in Ref. (Schädle et al., 2006) with  $O(N \log N)$  complexity.

### 3.3 Numerical Results

The present model enables us to simulate the reduced dynamics beyond the standard GLE and systematically investigate open problems like the state-dependent memory effect on the collective dynamics of complex systems such as molecule kinetics.

#### 3.3.1 Full Atomistic Model

In this work, we consider the full micro-scale model of benzyl bromide (see Fig. 3.1 for a sketch of the molecule structure) in an aqueous environment. The general AMBER (Wang et al., 2004) force field is used for the benzyl bromide molecule and the partial charges of molecule atoms were set by the restrained electrostatic potential (RESP) approach (Bayly et al., 1993). The rigid TIP3P water model (Jorgensen et al., 1983) is used for the water molecules and the bond lengths and angles were held constant through the SHAKE algorithm (Ryckaert et al., 1977; Miyamoto and Kollman Peter, 2004). Long-range electrostatic interactions were calculated using a Particle Mesh Ewald summation with a relative error set to be  $10^{-4}$ .

The full system consists of one benzyl bromide molecule and 2400 water molecules with the periodic boundary condition imposed along each direction. The isothermal-isobaric thermostat (Martyna et al., 1994) is used to equilibrate the system for 16 ns at 298K and 1 bar using a time step of 1 fs. Following the equilibration, the box size is scaled to be near  $41.5 \times 41.5 \times 41.5 \text{ \AA}^3$ . The simulation was run for a production period of  $1.5 \mu\text{s}$  in a canonical ensemble with a Nosé-Hoover thermostat (Nosé, 1984; Hoover, 1985). The numerical results of this work are presented in Å for length, picosecond for time, and gram per mole for mass.

The resolved variable  $q$  is defined as the distance between the bromine atom and the ipso-carbon atom. The free energy is obtained from the probability density function  $\rho(q)$  (see the inset plot of Fig. 3.3(a)), i.e.,  $U(q) = -k_B T \ln \rho(q)$ , where  $\rho(q)$  is directly obtained from the full MD samples

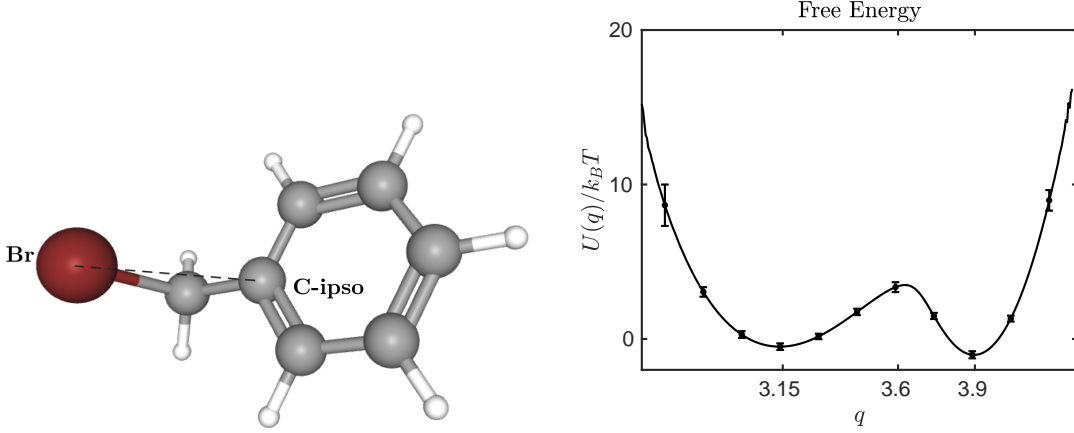


Figure 3.1 Left: A sketch of the molecule benzyl bromide. The resolved variable is defined as the distance between the bromine atom and the ipso-carbon atom. Right: The free energy of the resolved variable  $q$ . The error bar represent the 95% confidence interval.

using the kernel density estimation. To verify the accuracy of the constructed  $U(q)$ , we calculate the expectation of  $q\nabla U(q)$  on the sample. The numerical result gives  $0.996k_B T$  and is close to the theoretical prediction  $\langle q\nabla U(q) \rangle = \int q\nabla U(q)e^{-U(q)/k_B T} dq \equiv k_B T$ .

### 3.3.2 Limitation of the Standard GLE

Let us start with the standard GLE by setting features  $\phi(\mathbf{q}) \equiv \mathbf{I}$  in Eq. (3.3), which capture the dynamics on the resolved scale considered in Refs. (Ayaz et al., 2021; Dalton et al., 2023). We right-multiply  $\mathbf{q}(0)$  to Eq. (3.3) and compute the correlation functions, i.e.,  $h(t) = \int_0^t \Theta(t - \tau) c_{vq}(\tau) d\tau$ , where  $h(t) = \langle [\dot{\mathbf{p}}_t + \nabla U(\mathbf{q}_t)] \mathbf{q}_0^T \rangle$ . The standard GLE kernel  $\Theta(t)$  (i.e.,  $\mathbf{K}(t)$  in Eq. (3.1)) can be obtained using the Fourier transform of the integral equation. If the reduced dynamics (3.1) can be simplified as the standard GLE, then  $c_{vq}(t)$  should be accurately reproduced. Fig. 3.2 shows the prediction of  $c_{vq}(t)$  from the standard GLE and the full MD model. The apparent deviations imply non-negligible state-dependency. To further probe this effect, we compute  $h''(t; \mathbf{q}^*) = g'(t; \mathbf{q}^*)$  conditional with different initial states  $\mathbf{q}^*$ . Unlike a unified short-time correlation (i.e.,  $g'(0; q^*) = -k_B T \Theta(0)/m$ ) predicted by the standard GLE, the large dispersion reveals the heterogeneous nature of the energy dissipation process.

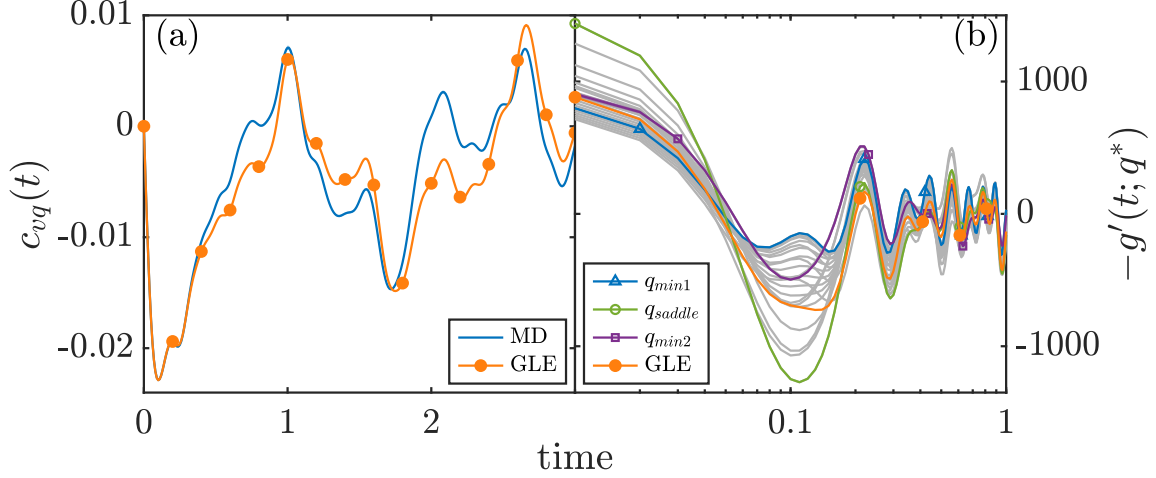


Figure 3.2 Correlation functions predicted by the standard GLE and the full MD: (a) Overall  $c_{vq}(t)$  and (b)  $-g'(t; \mathbf{q}^*)$  conditional with  $\mathbf{q}^*$  representing various initial states (gray lines), including two local minima and the saddle point (see inset of Fig. 3.3(a)). The large dispersion implies the limitation of the standard GLE, which predicts a single curve in short time.

### 3.3.3 GLE with State-dependent Memory

To capture the state-dependent memory, we train the present model (3.3) with a different number of features. Fig. 3.3(a-b) shows the obtained encoder  $\phi(\cdot)$  using one feature and  $\Theta(t)$  is scaled with  $\Theta(0) = 1$ .  $\phi$  exhibits apparent deviation from a uniform distribution. In particular, it shows a peak value near the saddle point  $q = 3.65$ , implying a larger effective friction near the regime. This result supports a similar assumption in semi-analytical studies (Straus et al., 1993) on improving Kramers' theory (Kramers, 1940). Also, it explains the short-time dispersion shown in Fig. 3.2, where  $g'(t; \mathbf{q}^*)$  at the saddle point is significantly larger than the local minima. Fig. 3.3(c-d) shows the obtained encoders  $\{\phi_i(\cdot)\}_{i=1}^n$  with  $n = 4$  features and the diagonal components of  $\Theta(t)$ . Compared with the case of  $n = 1$ , the larger variation of  $\phi_i$  enables a better representation of the state-dependent memory.

Next, we examine the conditional correlations  $c_{vq}(t; \mathbf{q}^*)$  and  $c_{vv}(t; \mathbf{q}^*)$ . As shown in Fig. 3.4, for both local minima and the saddle point, the predictions of the present model using four features agree well with the MD results. In contrast, the predictions of the standard GLE show apparent deviations for  $\mathbf{q}^*$  at the saddle point. Also, the present model using four features with a diagonal  $\Theta(t)$  (see Section 3.5.3) shows improved short-time predictions but remains insufficient for long-time

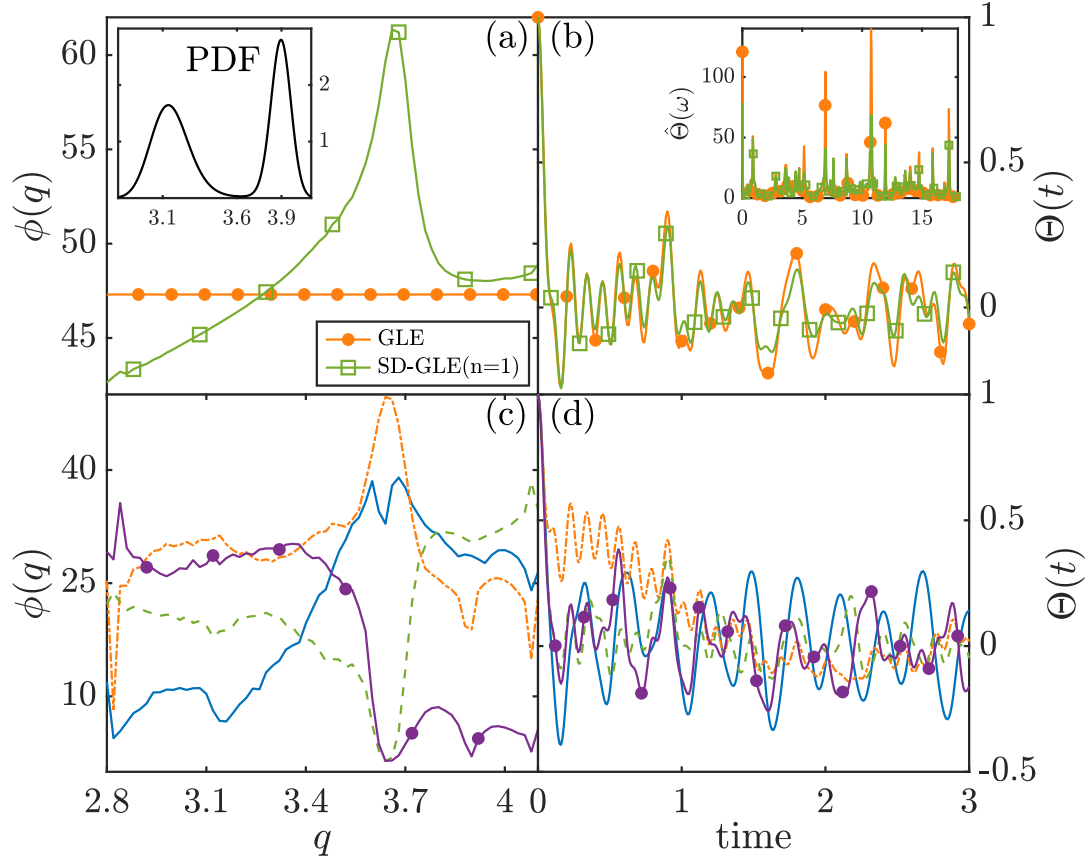


Figure 3.3 The state features  $\phi$  and diagonal components of the matrix-valued kernel  $\Theta(t)$  for the present model with state-dependent memory (SD-GLE) trained using (a-b) one feature and (c-d) four features. Inset plots: (a) probability density function (PDF) of  $q$ , where  $\phi(q)$  near the saddle point shows a peak; (b) Fourier modes of  $\Theta(t)$ .

correlations. This indicates the complex global variation of the memory term, which can not be represented by a simple state-dependent re-scaling of a kernel function; the non-Markovian coupling among multiple features is crucial to capture the heterogeneous energy dissipation over the full space.

Finally, we examine the collective behavior related to molecule kinetics. Fig. 3.5(a) shows the position correlation  $c_{qq}(t)$  characterizing the molecule conformation relaxation. Compared with the MD results, the standard GLE shows a significant underestimation of the relaxation time. This discrepancy is possibly due to the larger effective friction near the saddle point (see Fig. 3.3(a)), which essentially dampens the transition between the two local minima. The standard GLE overlooks such state-dependency and therefore yields a faster relaxation. This limitation is

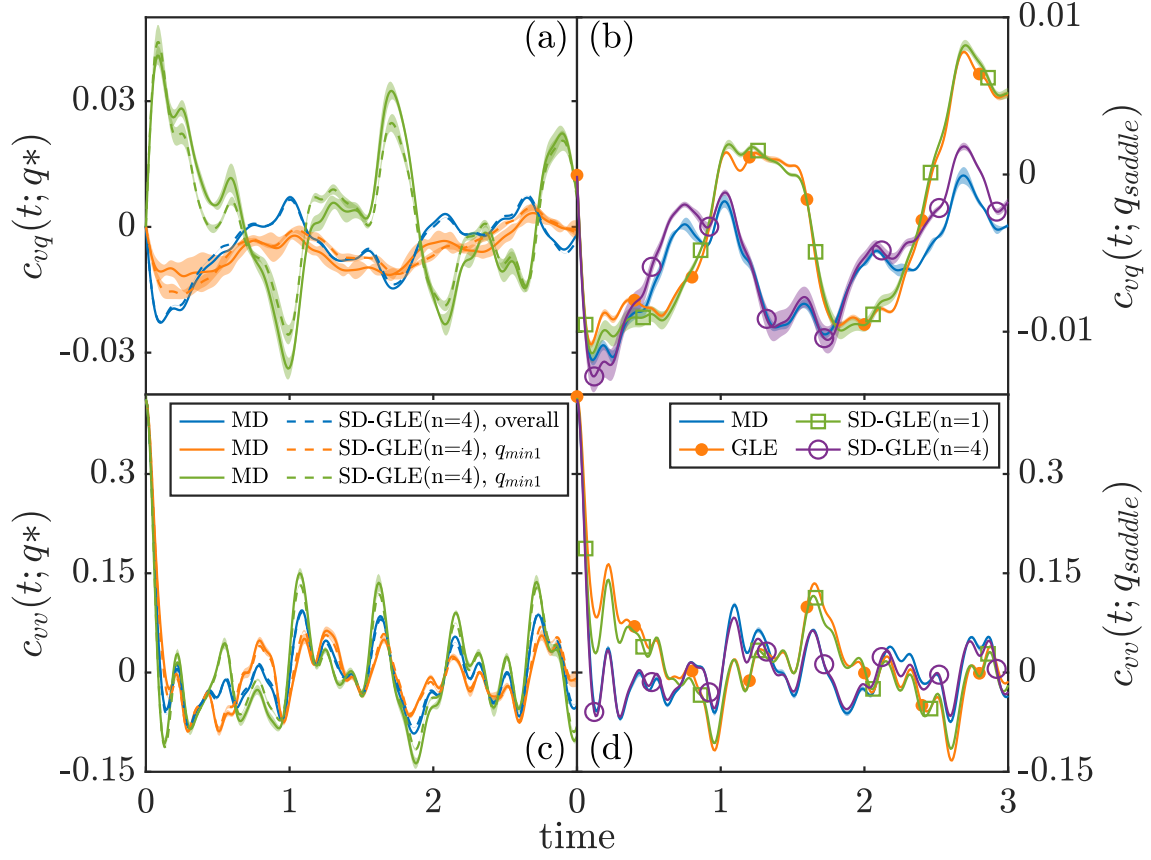


Figure 3.4 Overall and conditional correlation functions predicted by the full MD and various reduced models for two local minima and the saddle point: (a-b)  $c_{vq}$  and (c-d)  $c_{vv}$ . Shaded regimes represent the 95% confidence interval; same for Fig. 3.5.

consistently reflected in the distribution of the transition time. For this system, the free energy barrier is approximately  $3.5k_B T$  (see Fig. 3.1 right); the transition time is obtained from the simulation trajectories of the MD and various reduced models. As shown in Fig. 3.5(b), the standard GLE predicts a larger probability for the short transition time, indicating a smaller overall friction than the local (i.e., saddle point) value. Fortunately, the heterogeneous non-Markovianity can be faithfully retained in the present model. In particular, the constructed model using one feature yields a better prediction than the standard GLE. As we increase to four features, the predictions recover the MD results.

### 3.4 Summary

In summary, to plan an optimal hiking trail on a mountain, a landscape map is generally insufficient; the local path roughness needs to be properly considered. Similarly, to predict the



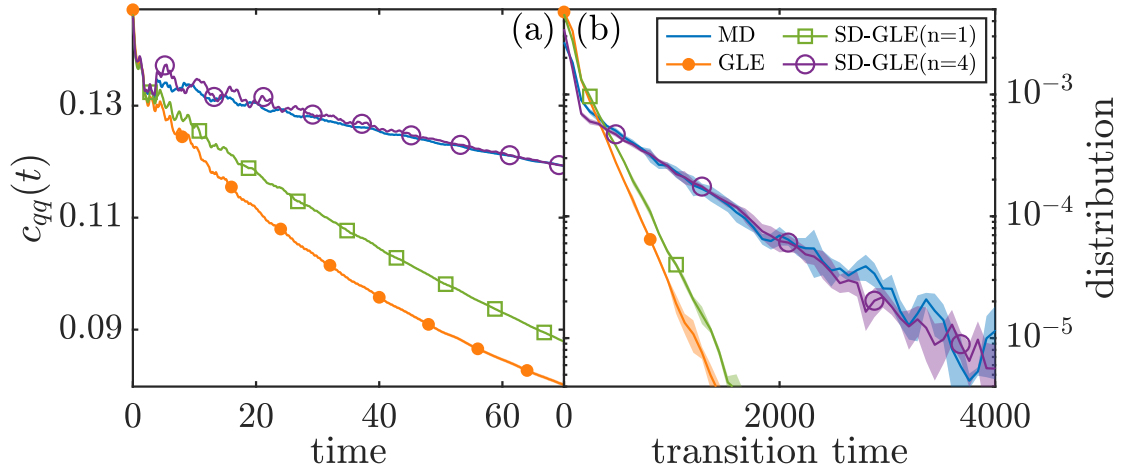


Figure 3.5 Collective molecule behaviors predicted by the full MD and the various reduced models: (a) overall conformation relaxation and (b) distribution of the transition time between the two local minima.

reduced dynamics of a multi-scale system, the state-dependent memory may need to be modeled to account for the heterogeneous energy dissipation arising from the unresolved dynamics, which, however, has been broadly overlooked. While the crucial role of the non-Markovian effect that complements the conservative free energy has been gradually recognized, the formulation of the memory term remains largely empirical (e.g., the standard GLE). The current work focuses on this caveat and presents a data-driven approach to learning such a stochastic reduced model beyond the standard GLE, where the complex state-dependent memory can be naturally encoded in the non-Markovian interactions among a set of features in terms of the resolved variables. The training does not rely on the explicit knowledge of the full model and only utilizes the trajectory samples, where the three-point correlations can be efficiently pre-computed. Numerical results of a molecule system demonstrate the crucial role of the state-dependent non-Markovianity on collective behavior, where the standard GLE shows limitations due to the over-simplified assumption of a homogeneous memory kernel. In contrast, the present model accurately predicts the molecule kinetics including the transition time distribution, and provides a reliable approach to simulate stochastic reduced dynamics of multiscale problems that faithfully retains the collective behaviors and rare event properties (E and Vanden-Eijnden, 2010) beyond empirical models.

### 3.5 Other Details

#### 3.5.1 Mass Matrix of the Reduced Model

Generally the mass matrix should depend on the resolved variables for the general cases, and we refer to Refs. (Ayaz et al., 2022a; Lee et al., 2019) for further discussions and the reduced dynamics with position-dependent mass. However, in this study we focus on the effect of the state-dependent non-Markovian memory on the collective behavior of complex systems. Therefore, we choose the coarse-grained resolved variables such that the corresponding mass matrix is a constant.

Specifically, we define  $q = \|\mathbf{Q}_1 - \mathbf{Q}_2\|$ , where  $\mathbf{Q}_1$  and  $\mathbf{Q}_2$  are the atom coordinates of the full model (see Fig. 3.1 and Sec. 3.3.1 for details). Accordingly, we have  $\dot{q} = \mathbf{Q}_{12}^T \dot{\mathbf{Q}}_{12} / q$  and its covariance follows

$$\begin{aligned} \langle \dot{q} \dot{q} \rangle &= \left\langle \frac{1}{q^2} \mathbf{Q}_{12}^T \dot{\mathbf{Q}}_{12} \dot{\mathbf{Q}}_{12}^T \mathbf{Q}_{12} \right\rangle \\ &= \left\langle \frac{1}{q^2} \text{Tr} [(\mathbf{Q}_{12} \mathbf{Q}_{12}^T)(\dot{\mathbf{Q}}_{12} \dot{\mathbf{Q}}_{12}^T)] \right\rangle \\ &= \left\langle \frac{1}{q^2} \text{Tr}(\mathbf{Q}_{12} \mathbf{Q}_{12}^T) \right\rangle (M_1^{-1} + M_2^{-1}) k_B T \\ &= (M_1^{-1} + M_2^{-1}) k_B T, \end{aligned} \tag{3.17}$$

where  $M_1$  and  $M_2$  represent the mass of two atoms and we have used the fact that the distribution of  $\mathbf{Q}_{12}$  and  $\dot{\mathbf{Q}}_{12}$  are independent. Therefore, the mass matrix of  $q$  is a constant  $M \equiv M_1 M_2 / (M_1 + M_2)$ .

#### 3.5.2 Limitations of the Standard GLE near the Local Minima

Fig. 3.6 shows the predictions of the conditional correlations  $c_{qv}(t, q^*)$  and  $c_{vv}(t, q^*)$  for  $q^*$  representing the two local minima. Similar to the results of the saddle point (see Fig. 3.5(b)), the predictions of the standard GLE show apparent deviations from the full MD results due to the ignorance of the state-dependent memory nature. In contrast, the predictions of the present model with four features can accurately recover the MD predictions.

#### 3.5.3 Other Forms of the State-dependent Memory Term

For comparison, we also consider other forms of the reduced model. In particular, we retain the encoders  $\phi$  in Eq. (3.3) but set  $\Theta(t)$  to be diagonal, i.e., we ignore the non-Markovian coupling among the different state features. The reduced model is trained using four features. Fig. 3.7 shows

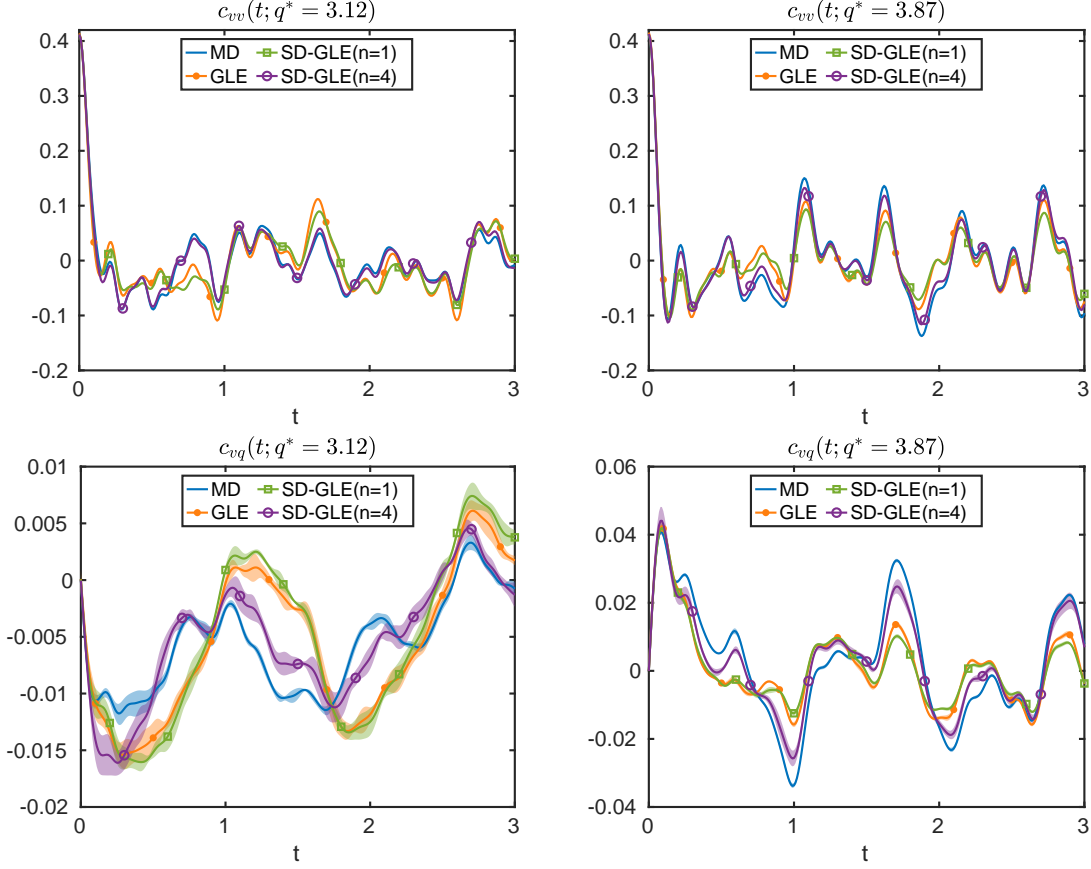


Figure 3.6 The conditional correlation functions  $c_{vq}(t, q^*)$  and  $c_{vv}(t, q^*)$  for the two local minima predicted by the full MD, the standard GLE, and the present model (SD-GLE) constructed using one and four spatio-features. Left:  $q^* = 3.07$ ; Right:  $q^* = 3.87$ . The predictions by the standard GLE show apparent discrepancies with the full MD results. Shaded regimes represent the 95% confidence interval.

the conditional correlation  $c_{vv}(t; q^*)$  obtained from the full MD and different reduced models. The prediction of the constructed model (labeled by “SD-GLE-Diag”) shows apparent deviations from the full MD result with incremental improvement over the stand GLE. The large discrepancy reveals the complex state-dependent nature; the non-Markovian effect can be neither approximated by ansatz like  $\gamma(q)\theta(t)$  as a simple generalization/re-scaling of the initial value at  $t = 0$ , nor represented by the coupling with the independent bath variables. Instead, the non-Markovian coupling among the various state-features retained in the present model plays a crucial for accurately modeling the heterogeneous energy dissipation arising from the unresolved intramolecular interactions and reproducing the collective dynamics.

Furthermore, we note that Ref. (Vroylandt and Monmarché, 2022) develops an efficient approach

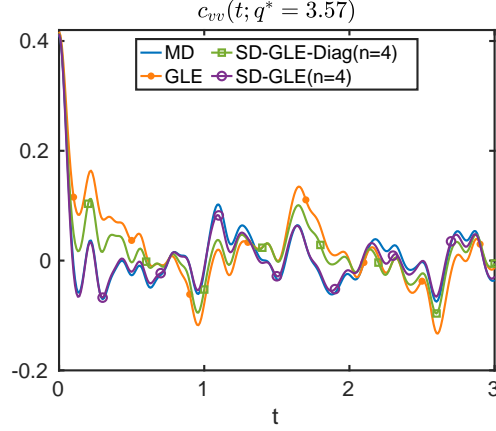


Figure 3.7 The conditional correlation functions  $c_{vv}(t, q^*)$  for the saddle point predicted by the full MD, the standard GLE, the reduced models constructed using four state-features with diagonal  $\Theta(t)$  (SD-GLE-Diag) and full  $\Theta(t)$  (SD-GLE). The large discrepancy between the SD-GLE-Diag model and the full MD results implies the complexity of the state-dependency of the memory term, which can not be well represented by the coupling of independent bath variables. The non-Markovian interactions among the state-features are essential to capture the heterogeneous energy dissipation process.

to compute the memory function based on the finite-rank approximation of the Zwanzig's projection formalism. The method can be used to efficiently extract the state-dependent memory and probe the physics insights from the trajectory samples of the full MD model. The present work focuses on training a reduced model with heterogeneous memory that enables generating coherent noise and conducting stochastic simulations.

Specifically, the memory term takes the form  $K(q(\tau), t - \tau)$  (with a simple change of variable  $s = t - \tau$  following the notation) in Ref. (Vroylandt and Monmarché, 2022) and  $\tilde{K}(q(t), q(\tau), t - \tau) = \phi(q(t))^T \Theta(t - \tau) \phi(q(\tau))$  in the present work. In particular, by setting  $q(t) = q(\tau)$ , the memory term in two forms should have similar prediction. Following this argument, we use the method in Ref. (Vroylandt and Monmarché, 2022) to calculate  $K(q(\tau), t - \tau)$  and compare it with  $\tilde{K}(q(t), q(\tau), t - \tau)$  of the present model with  $q(t) = q(\tau) = q^*$ . Fig. 3.8 shows the obtained memory functions for  $q^*$  taking the saddle point and two local minima. The prediction by the two approaches show good agreement.

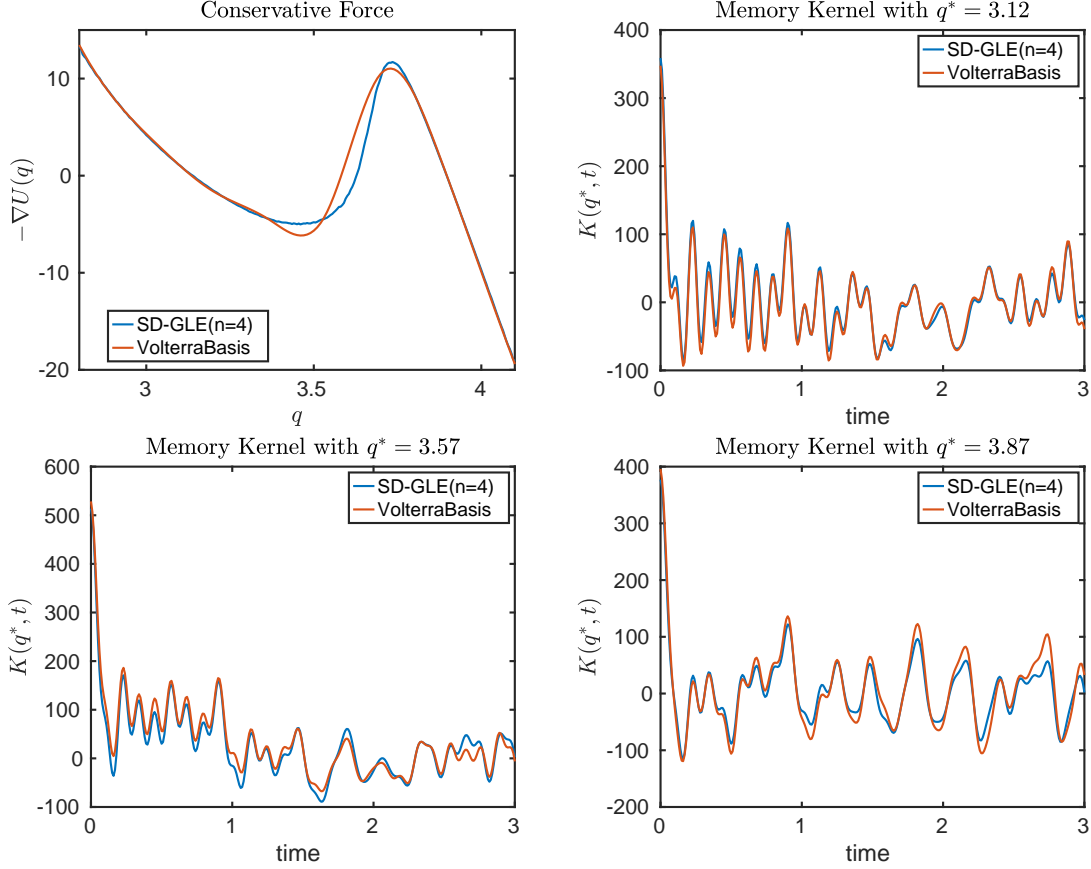


Figure 3.8 The conservative force and memory kernel obtained from Ref. (Vroylandt and Monmarché, 2022) (labeled as Volterra Basis) and the present model. The memory term takes the form  $K(q(\tau), t - \tau)$  and  $\tilde{K}(q(t), q(\tau), t - \tau) = \phi(q(t))^T \Theta(t - \tau) \phi(q(\tau))$ , respectively. Specifically, we set  $q(t) = q(\tau) = q^*$ ; the predictions of the two models show good agreement for the saddle point and the two local minima.

### 3.5.4 Generalization of the Present Reduced Model Formulation

So far, we have constructed the reduced model (3.3) by assuming the matrix-valued kernel  $\Theta(t)$  is symmetry. In fact, this form can be generalized by introducing an anti-symmetry part, i.e.,

$$\Theta(t) = e^{-\alpha t} \sum_{k=0}^{N_\omega} (\Gamma_{k,1}^T \Gamma_{k,1} + \Gamma_{k,2}^T \Gamma_{k,2}) \cos(\omega_k t) + (\Gamma_{k,1}^T \Gamma_{k,2} - \Gamma_{k,2}^T \Gamma_{k,1}) \sin(\omega_k t), \quad (3.18)$$

where  $\Gamma_{k,1}$  and  $\Gamma_{k,2}$  are lower-triangular matrices representing the Fourier modes of  $\Theta(t)$ . The form is general non-symmetric except for  $t = 0$  and satisfies  $\Theta(-t) = \Theta(t)^T$ .

Similar to the symmetry form, we can model the fluctuation term  $\mathcal{R}_t$  as a noise in the form of  $\mathcal{R}_t = \phi_t^T \tilde{\mathbf{R}}(t)$ , where  $\tilde{\mathbf{R}}(t)$  is a Gaussian random process satisfying  $\langle \tilde{\mathbf{R}}(t) \tilde{\mathbf{R}}(\tau)^T \rangle = k_B T e^{-\alpha(t-\tau)} \Theta(t - \tau)$ . Similar to Prop. 3.2.1, we can show that this choice retains a consistent invariant density function.

In practice, we can generate the noise term  $\tilde{\mathbf{R}}(t)$  on  $[0, T]$  by

$$\tilde{\mathbf{R}}(t) = \beta^{-1/2} \sum_{k=0}^{2N} \left[ \tilde{\Theta}_{k,1}^{1/2} \cos(\omega_k t) \xi_k + \sin(\omega_k t) (\mathcal{Q}_{1,k}^{1/2} \xi_k + \mathcal{Q}_{2,k}^{1/2} \eta_k) \right],$$

$$\mathcal{Q}_{1,k} = \tilde{\Theta}_{k,2} \tilde{\Theta}_{k,1}^{-1} \tilde{\Theta}_{k,2}^T \quad \mathcal{Q}_{2,k} = \tilde{\Theta}_{k,1} - \tilde{\Theta}_{k,2} \tilde{\Theta}_{k,1}^{-1} \tilde{\Theta}_{k,2}^T,$$

where  $\beta^{-1} = k_B T$ ,  $\tilde{\Theta}_{k,1}$ ,  $\tilde{\Theta}_{k,2}$  are the Fourier cosine and sine modes on  $[-T, T]$  with  $\Theta(-t) = \Theta(t)^T$ ,  $\xi_k$  and  $\eta_k$  are independent Gaussian random vectors, and  $N$  is the total number of simulation step. Here  $\tilde{\mathbf{R}}(t)$  can still be generated using the Fast Fourier Transform algorithm (Cooley and Tukey, 1965) using  $O(N \log N)$  complexity. We will investigate this generalized formulation for model reduction in future studies.

## CHAPTER 4

### DEEP LEARNING-BASED NON-NEWTONIAN FLUID MODEL

In this chapter, we focus on a micro-to-macro model named DeePN<sup>2</sup> (Fang et al., 2022). A long standing problem in the modeling of non-Newtonian hydrodynamics of polymeric flows is the availability of reliable and interpretable hydrodynamic models that faithfully encode the underlying micro-scale polymer dynamics. The main complication arises from the long polymer relaxation time, the complex molecular structure and heterogeneous interaction. DeePN<sup>2</sup>, a deep learning-based non-Newtonian hydrodynamic model, has been proposed and has shown some success in systematically passing the micro-scale structural mechanics information to the macro-scale hydrodynamics for suspensions with simple polymer conformation and bond potential. The model retains a multi-scaled nature by mapping the polymer configurations into a set of symmetry-preserving macro-scale features. The extended constitutive laws for these macro-scale features can be directly learned from the kinetics of their micro-scale counterparts. In this paper, we develop DeePN<sup>2</sup> using more complex micro-structural models. We show that DeePN<sup>2</sup> can faithfully capture the broadly overlooked viscoelastic differences arising from the specific molecular structural mechanics without human intervention.

#### 4.1 Introduction

Accurate modeling of non-Newtonian hydrodynamics plays a central role in the modeling of the transport, diffusion, and synthesis processes in many scientific and engineering applications. Unlike simple fluids, non-Newtonian fluids may exhibit enormously complex flow behavior as a result of the micro-scale polymer dynamics. In particular, the polymer relaxation time often becomes comparable to the hydrodynamic time scale. As a result, the macro-scale fluid evolution can not be uniquely determined by the instantaneous flow field and the memory effect is generally important. To close the hydrodynamic equations, existing models are primarily based on the following two approaches. The first approach relies on empirical constitutive models (Larson, 1988; Owens and Phillips, 2002). Notable examples include the Hookean model (Oldroyd and Wilson, 1950; Lin et al., 2005), the FENE-P model (Peterlin, 1966; Bird et al., 1980), the Giesekus model (Giesekus,

1982), and the Phan-Thien and Tanner models (Thien and Tanner, 1977). Despite their popularity, the accuracy of these models is almost always in doubt. The second approach resorts to various sophisticated micro-macro coupling algorithms, e.g., by directly solving the Fokker-Planck equation, or sampling the polymer configuration via micro-scale simulations (Laso and Öttinger, 1993; Hulsen et al., 1997; Ren and E, 2005). While the effects of the polymer interaction can be carried over to the macro-scale model, the computational cost can be exceedingly large due to the retaining of the micro-scale description. Methods based on asymptotic analysis (Warner, 1972a) or the direct fitting of the strain-stress relationship (Zhao et al., 2018) are limited to simple flows such as the steady flow. Several semi-analytical approaches have been proposed (Grosso et al., 2000; Feng et al., 1998; Wang, 1997; Forest et al., 2003; Lielens et al., 1999; Yu et al., 2005; Hyon et al., 2008) using moment closure to approximate the micro-scale polymer configuration probability density function (PDF) and to derive the constitutive equations for the FENE dumbbell solution (Lielens et al., 1999; Yu et al., 2005; Hyon et al., 2008). However, these approaches are all based on restricted ansatz for the PDF and therefore are not reliable for more general flow regimes.

To construct truly reliable and interpretable hydrodynamic models with molecular-level fidelity, it is essential to be able to efficiently code the information from the micro-scale interaction into the macro-scale transport equations. Ideally, the construction should meet the following requirements:

- be interpretable;
- be reliable – it should be accurate for all kinds of practical situations that one might encounter;
- respect physical constraints, including symmetries and conservation laws;
- be numerically robust and efficient.

As a first step towards constructing models that meet these requirements, we developed a machine learning-based approach (Lei et al., 2020), “deep learning-based non-Newtonian hydrodynamic model” or DeePN<sup>2</sup>, that learns the non-Newtonian hydrodynamic model from the underlying micro-scale description of the dumbbell solution. Rather than approximating the closure with standard moments, DeePN<sup>2</sup> finds a set of encoders, i.e., a set of macro-scale features that best represent the micro-scale dumbbell structure. It also finds accurate closed-form equation for these



macro-scale features. The constructed model retains a clear physical interpretation and accurately captures the nonlinear viscoelastic responses, where the conventional Hookean and FENE-P models show limitations.

Beyond dumbbell suspensions, one major challenge towards constructing truly reliable hydrodynamic models arises from the heterogeneous polymer micro-structural mechanics. In this work, we aim to fill the gap by developing the generalized DeePN<sup>2</sup> model for multi-bead polymer molecules with arbitrary structure and interaction. Firstly, with the proper design of the generalized micro-macro encoders and the machine learning-based symmetry-preserving constitutive dynamics, we demonstrate that the heterogeneous molecular structural-induced interaction can be systematically encoded into the macro-scale hydrodynamics. Unlike moment closure approximations, the encoders are not designed to recover the high-dimensional configuration PDF. Instead, they take an interpretable form and are learned to probe the optimal approximation of the polymer stress and constitutive dynamics. This essential difference enables DeePN<sup>2</sup> to circumvent the high-dimensionality of the polymer configuration PDF. Secondly, the explicit form of the micro-macro encoders enables us to reliably learn the dynamics of the macro-scale features directly from the kinetic equations of their micro-scale analog. In this sense, this learning framework retains a multi-scaled nature where micro-scale interaction and physical constraints can be seamlessly inherited. Moreover, the learning only requires instantaneous micro-scale samples. This unique property differs from the common sophisticated data-driven approaches (Rudy et al., 2017; Schaeffer et al., 2018; Raissi et al., 2019b; Qin et al., 2019; Han et al., 2019b; Seryo et al., 2020; Huang et al., 2021), where time-derivative samples are often needed to learn the governing dynamics. This is particularly suited for multi-scale fluid models where accurate time-derivative samples may not be readily accessible. We demonstrate the power of the DeePN<sup>2</sup> model for polymer molecules of three distinct shapes with training samples collected from one-dimensional (1D) homogeneous shear flow. Numerical results show that the broadly overlooked heterogeneous molecular structural mechanics plays an important role in the rheology of non-Newtonian fluids, which, fortunately, can be faithfully encoded into DeePN<sup>2</sup>. The constructed model successfully captures the hydrodynamics with different viscoelastic responses for

a variety of 1D and 2D flows when compared with the micro-scale simulation results. The present work also paves the way towards constructing truly reliable non-Newtonian hydrodynamic models for general 3D flows.

## 4.2 Methods

### 4.2.1 Micro-scale and continuum hydrodynamic models

Let us start with the micro-scale description of the semi-dilute polymer suspension. We assume each molecule consists of  $N$  particles with the position vector  $\mathbf{q} = [\mathbf{q}_1; \mathbf{q}_2; \dots; \mathbf{q}_N]$ , where  $\mathbf{q}_i \in \mathbb{R}^3$  is the position of the  $i$ -th particle. The intramolecular potential energy  $V(\mathbf{q})$  takes the form

$$V(\mathbf{q}) = \sum_{j=1}^{N_b} V_b(|\mathbf{q}_{j_1} - \mathbf{q}_{j_2}|), \quad V_b(l) = -\frac{k_s}{2} l_0^2 \log \left[ 1 - \frac{l^2}{l_0^2} \right], \quad (4.1)$$

where  $N_b$  is the bond number and  $(j_1, j_2)$  represents the indices of beads associated with the  $j$ -th bond. Without loss of generality, the individual bond interaction  $V_b$  takes the form of the FENE potential (Warner, 1972b), where  $k_s$  is the spring constant and  $l_0$  is the maximum of the extension length. It is worth mentioning that the polymer molecule is not restricted to the dumbbell shape. Instead, it generally consists of multiple particles with arbitrary structure and bond connection. Fig. 4.1 shows a sketch of the polymer molecules with three different structures. As we will show, given the same form of the individual bond interaction  $V_b$ , the different polymer micro-structural mechanics leads to distinct non-Newtonian hydrodynamics.

In principle, the viscoelastic response of the system is determined by the full micro-scale

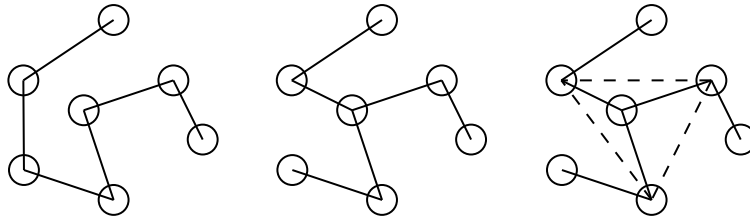


Figure 4.1 A sketch of 7-bead polymer molecules with chain-, star- and net-shaped structures (from left to right). The solid lines represent the FENE bond potential with the same interaction parameters. The dashed lines of the net-shaped molecule represent the three additional side chains connecting the polymer arms. While both the chain- and the star-shaped molecules are connected with six bonds; the suspensions exhibit different hydrodynamics due to the different micro-structural mechanics as shown below.

interaction. However, direct simulation for the full micro-scale interaction is often limited by the prohibited computational cost. Continuum hydrodynamics models based on various empirical constitutive models are often used, with the general form

$$\begin{aligned}\nabla \cdot \mathbf{u} &= 0, \\ \rho \frac{d\mathbf{u}}{dt} &= -\nabla p + \nabla \cdot (\boldsymbol{\tau}_s + \boldsymbol{\tau}_p) + \mathbf{f}_{\text{ext}},\end{aligned}\tag{4.2}$$

where  $\rho$ ,  $\mathbf{u}$  and  $p$  represent the fluid density, velocity and pressure field, respectively.  $\mathbf{f}_{\text{ext}}$  is the external body force and  $\boldsymbol{\tau}_s = \eta_s(\nabla\mathbf{u} + \nabla\mathbf{u}^T)$  is the solvent stress tensor with shear viscosity  $\eta_s$ .  $\boldsymbol{\tau}_p$  is the polymer stress tensor whose detailed form is generally unknown. To construct  $\boldsymbol{\tau}_p$ , the DeePN<sup>2</sup> model seeks the approximation in terms of a set of macro-scale features  $\mathbf{c}_1, \dots, \mathbf{c}_n$ , and simultaneously, the constitutive dynamics of these features, i.e.,

$$\boldsymbol{\tau}_p = \mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n),\tag{4.3a}$$

$$\frac{\mathcal{D}\mathbf{c}_i}{\mathcal{D}t} = \mathbf{H}_i(\mathbf{c}_1, \dots, \mathbf{c}_n), \quad i = 1, \dots, n,\tag{4.3b}$$

where  $\mathbf{G}$  and  $\mathbf{H}_i$  represent the stress and constitutive models, respectively.  $\frac{\mathcal{D}}{\mathcal{D}t}$  denotes the objective tensor derivative.

Eqs. (4.2) and (4.3) take the form similar to the conventional hydrodynamics. Instead of using empirical approximation to close the equation, we aim to construct a model directly from the micro-scale description (4.1) with the help of machine learning, such that the constructed model can naturally encode the molecular-specific interaction beyond empirical approximations with clear physical interpretation.

#### 4.2.2 DeePN<sup>2</sup> for arbitrary molecular structural mechanics

To learn Eq. (4.2) from the full model (4.1), one essential problem lies in how to seamlessly pass the micro-scale interaction to the continuum model. To bridge the scales, we learn a set of micro-to-macro encoders, denoted by  $\{\mathbf{b}_i(\mathbf{q})\}_{i=1}^n$ , such that the continuum modeling terms (e.g., the polymer stress  $\boldsymbol{\tau}_p$ ) can be well approximated in terms of the corresponding macro-scale features  $\{\mathbf{c}_i(\mathbf{q})\}_{i=1}^n$  via Eq. (4.3a), where  $\boldsymbol{\tau}_p := n_p \sum_j \langle \mathbf{q}_j \otimes \nabla_{\mathbf{q}_j} V(\mathbf{q}) \rangle$ ,  $\mathbf{c}_i = \langle \mathbf{b}_i(\mathbf{q}) \rangle$ ,  $n_p$  is the polymer number density and  $\langle \cdot \rangle$  denotes the average with respect to the configuration PDF. In particular, the features

$\mathbf{c}_i$  need to satisfy the proper invariant and symmetry conditions inherited from the encoders  $\mathbf{b}_i(\cdot)$  such that the constructed continuum model can strictly preserve frame-indifference condition:

$$\tilde{\boldsymbol{\tau}}_p = \mathbf{Q}\boldsymbol{\tau}_p\mathbf{Q}^T, \quad \mathbf{G}(\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_n) = \mathbf{Q}\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n)\mathbf{Q}^T, \quad (4.4)$$

where the superscript  $\tilde{\cdot}$  denotes the corresponding values under an arbitrary orthogonal transformation by  $\mathbf{Q} \in \text{SO}(3)$ .

To construct the encoder  $\mathbf{b}(\cdot)$ , we note that the micro-scale potential  $V(\mathbf{q})$  is translational and rotational invariant. Accordingly, let  $\mathbf{r}^*(\mathbf{q}) \in \mathbb{R}^{3N-6}$  (we consider the general case  $N \geq 3$  here) denote the translational-rotational-invariant configuration vector and  $\mathbf{r}(\mathbf{q}) \in \mathbb{R}^{3N-3}$  denote the translational-invariant configuration vector consisting of  $N - 1$  linearly independent position vectors. Since  $N_b \geq N - 1$  for all molecules, one straightforward choice is the first  $N - 1$  bond connection vectors, i.e.,

$$\begin{aligned} \mathbf{r} &= [\mathbf{r}_1; \mathbf{r}_2; \dots; \mathbf{r}_{N-1}], \quad \mathbf{r}_j = \mathbf{q}_{j1} - \mathbf{q}_{j2}, \quad 1 \leq j \leq N - 1, \\ \mathbf{r}^* &= [|\mathbf{r}_1|, |\mathbf{r}_2|, |\mathbf{r}_{12}|, |\mathbf{r}_3|, |\mathbf{r}_{13}|, |\mathbf{r}_{23}|, |\mathbf{r}_4|, |\mathbf{r}_{24}|, |\mathbf{r}_{34}|, \dots, |\mathbf{r}_{N-1}|, |\mathbf{r}_{(N-2)(N-1)}|], \end{aligned} \quad (4.5)$$

where  $\mathbf{r}_{jk} := \mathbf{r}_j - \mathbf{r}_k$ . We note that this form applies to general molecular structures;  $\mathbf{r}$  determines the molecular structure up to translations. Specifically,  $\mathbf{r}^*$  represents the  $3N - 6$  degrees of freedom after eliminating translational and rotational degrees of freedom, and  $\mathbf{r}$  suffices to fully determine the translational invariant polymer configuration and strictly retains the rotational symmetry in accordance with  $\mathbf{q}$ , i.e.,

$$\mathbf{r}_j(\mathbf{Q}\mathbf{q}) = \mathbf{Q}\mathbf{r}_j(\mathbf{q}), \quad \mathbf{r}^*(\mathbf{Q}\mathbf{q}) = \mathbf{r}^*(\mathbf{q}).$$

To preserve rotational symmetry, one straightforward approach is to represent  $\mathbf{b}(\cdot)$  in the linear space spanned by  $\{\mathbf{r}_j\}_{j=1}^{N-1}$ . However, this choice yields the trivial macro-scale feature, i.e.,  $\langle \mathbf{r}_j \rangle \equiv 0$ , due to the rotational symmetry. Alternatively, we construct the following second-order tensor

$$\begin{aligned} \mathbf{c}_i &= \langle \mathbf{b}_i(\mathbf{r}) \rangle, \quad \mathbf{b}_i = \mathbf{f}_i \mathbf{f}_i^T, \quad 1 \leq i \leq n, \\ \mathbf{f}_i &= g_i(\mathbf{r}^*) \sum_{j=1}^{N-1} w_{ij} \mathbf{r}_j, \end{aligned} \quad (4.6)$$

where  $[w_{ij}]_{1 \leq i \leq n, 1 \leq j \leq N-1}$  are the weights and  $\{g_i(\cdot)\}_{i=1}^n$  is a set of scalar functions that encodes the polymer intramolecular interaction. Both terms will be learned from the micro-scale description and represented by deep neural networks (DNNs). Rotational symmetries can be naturally inherited, i.e.,  $\tilde{\mathbf{c}} = \langle \mathbf{b}(\tilde{\mathbf{r}}) \rangle \equiv \mathbf{Q}\mathbf{c}\mathbf{Q}^T$ . Compared with the special form for dumbbell molecules in Ref. (Lei et al., 2020), Eq. (4.6) provides a general form of  $\mathbf{c}$  applicable to multi-bead molecules of arbitrary structure since  $\mathbf{r}$  and  $\mathbf{r}^*$  fully determine the  $3N - 3$  translational invariant polymer configuration. In the remaining of the paper, we will abuse the notation and denote  $\mathbf{b}(\mathbf{q})$  as  $\mathbf{b}(\mathbf{r})$ .

Besides the polymer stress model (4.3a), the remaining task to close Eq. (4.2) is the construction of the constitutive dynamics (4.3b) of the macro-scale features  $\{\mathbf{c}_i\}_{i=1}^n$ . There are two issues to deal with: the proper form of the objective time derivative of  $\mathbf{c}_i$  and the accurate estimation of their time evolution. In the literature, the objective tensor derivative, denoted by  $\frac{D\mathbf{c}_i}{Dt}$ , is often chosen to take some heuristic forms (e.g. the convected (Oldroyd and Wilson, 1950) and corotational (Zaremba, 1903) forms). Moreover, the time-series samples collected from the micro-scale simulations are generally super-imposed with pronounced sampling error; direct estimation of the time derivative as was done in (Rudy et al., 2017; Raissi et al., 2019b; Seryo et al., 2020) will end with noisy data. Fortunately, both challenges are addressed in DeePN<sup>2</sup> using an explicit micro-macro correspondence. The dynamics of  $\mathbf{c}_i$  can be derived from the its micro-scale correspondence  $\mathbf{b}_i(\mathbf{r})$  in the form of the micro-scale configuration  $\mathbf{r}$ , i.e.,

$$\begin{aligned} \frac{d}{dt}\mathbf{c}_i - \kappa : \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle &= \frac{k_B T}{\gamma} \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b}_i \right\rangle \\ &\quad - \frac{1}{\gamma} \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V(\mathbf{r}_1, \dots, \mathbf{r}_{N_b}) \cdot \nabla_{\mathbf{r}_j} \mathbf{b}_i \right\rangle, \end{aligned} \quad (4.7)$$

where  $\kappa := \nabla \mathbf{u}^T$ ,  $\gamma$  is the friction coefficient and  $\mathbf{r}_j$  is the connection vector as defined in Eq. (4.5) for  $j > N - 1$ . We abuse the notation and denote  $V(\mathbf{q})$  as  $V(\mathbf{r}_1, \dots, \mathbf{r}_{N_b}) = \sum_{j=1}^{N_b} V_b(r_j)$ . The molecular structure and interaction are specified via  $\mathbf{A} \in \mathbb{R}^{N_b \times N_b}$ , which is defined by

$$\mathbf{A} = \mathbf{S}\mathbf{S}^T, \quad S_{jk} = \begin{cases} +1, & k = j_1, \\ -1, & k = j_2, \\ 0, & \text{else} \end{cases} \quad 1 \leq j \leq N_b, \quad 1 \leq k \leq N, \quad (4.8)$$

where  $j_1$  and  $j_2$  are the same notations as those in Eq. (4.1). We note that Eq. (4.7) only requires the first  $(N - 1)$  rows of  $\mathbf{A}$  since the polymer configuration can be fully determined by  $\mathbf{r}_1, \dots, \mathbf{r}_{N-1}$ . As a special case, if the molecule takes the chain shape,  $\mathbf{A}$  recovers the standard Rouse matrix (Bird et al., 1987; Rouse, 1953).

Eq. (4.7) defines the dynamics for the features  $\{\mathbf{c}_i\}_{i=1}^n$ , derived from their micro-scale correspondences. In particular, given the proposed form of the encoder functions (4.6), we can show that the two combined terms of the left-hand-side of Eq. (4.7) strictly preserve rotational symmetry (see Section 4.2.3). This leads to an important observation that the two combined terms provide the generalized form for the macro-scale objective tensor derivative  $\frac{\partial \mathbf{c}_i}{\partial t}$ . Unlike the heuristic choices in empirical models, the new form retains a clear micro-scale physical interpretation. Furthermore, all the modeling terms in the form of  $\langle \cdot \rangle$  can be directly evaluated using samples collected from the micro-scale simulations under the corresponding flow condition. This enables us to avoid estimating the time derivative values from the noise-prone time-series samples. Accordingly, the macro-scale constitutive dynamics takes the form

$$\frac{d\mathbf{c}_i}{dt} - \boldsymbol{\kappa} : \mathcal{E}_i = \frac{k_B T}{\gamma} \mathbf{H}_{1,i}(\mathbf{c}_1, \dots, \mathbf{c}_n) - \frac{1}{\gamma} \mathbf{H}_{2,i}(\mathbf{c}_1, \dots, \mathbf{c}_n), \quad (4.9)$$

where the individual terms will be represented by proper neural networks and parameterized by matching their micro-scale correspondences, i.e.,

$$\begin{aligned} \mathcal{E}_i(\mathbf{c}_1, \dots, \mathbf{c}_n) &= \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b}_i \right\rangle, \\ \mathbf{H}_{1,i}(\mathbf{c}_1, \dots, \mathbf{c}_n) &= \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b}_i \right\rangle, \\ \mathbf{H}_{2,i}(\mathbf{c}_1, \dots, \mathbf{c}_n) &= \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V(\mathbf{r}_1, \dots, \mathbf{r}_{N-1}) \cdot \nabla_{\mathbf{r}_j} \mathbf{b}_i \right\rangle. \end{aligned} \quad (4.10)$$

#### 4.2.3 Rotational frame-indifference of the constitutive dynamics for the multi-bead encoder function

We consider a polymer molecule consisting of  $N$  particles. Let  $\mathbf{r} = [\mathbf{r}_1; \mathbf{r}_2; \dots; \mathbf{r}_{N-1}]$  denote the polymer configuration, so that there exists an invertible linear transformation between  $[\mathbf{r}; \sum_{i=1}^N \mathbf{q}_i/N]$

and  $[\mathbf{q}_1; \mathbf{q}_2; \dots; \mathbf{q}_N]$ , where  $\mathbf{q}_i$  is the position of the  $i$ -th particle. In fact, there are multiple choices for  $\mathbf{r}$ , including the one we have applied in Eq. (4.5), where  $\mathbf{r}$  consists of  $(N-1)$  edges of a spanning tree in the bead-bond structure.

We consider a second-order tensor taking the general form

$$\mathbf{b} = \mathbf{f}^{(1)}(\mathbf{r})\mathbf{f}^{(2)}(\mathbf{r})^T, \quad \mathbf{f}^{(1)}(\mathbf{r}) = \sum_{j=1}^{N-1} g_j^{(1)}(\mathbf{r}^*)\mathbf{r}_j, \quad \mathbf{f}^{(2)}(\mathbf{r}) = \sum_{j=1}^{N-1} g_j^{(2)}(\mathbf{r}^*)\mathbf{r}_j, \quad (4.11)$$

where  $\mathbf{r}^*$  is a translational-rotational-invariant vector and  $g^{(1)}$  and  $g^{(2)}$  are two scalar functions. We note that the encoder in the form of Eq. (4.11) is more general than Eq. (4.6).

In this section and the next, we consider two frames: frame 1 is static inertial, and frame 2 is rotating with respect to frame 1 with an time dependent orthogonal transformation  $\mathbf{Q}(t)$ . Let  $\tilde{\mathbf{x}}, \tilde{\mathbf{v}}, \tilde{\mathbf{b}}$  and  $\mathbf{x}, \mathbf{v}, \mathbf{b}$  denote the positions, velocities, and second-order tensors in frame 1 and 2 respectively. They have the following relations:

$$\tilde{\mathbf{x}} = \mathbf{Q}\mathbf{x}, \quad \tilde{\mathbf{v}} = \mathbf{Q}\mathbf{v} + \dot{\mathbf{Q}}\mathbf{x}, \quad \tilde{\mathbf{b}} = \mathbf{Q}\mathbf{b}\mathbf{Q}^T. \quad (4.12)$$

The material derivatives in both frames are

$$\left. \frac{d}{dt} \right|_{\text{frame 1}} := \frac{\partial}{\partial t} + \tilde{\mathbf{v}} \cdot \nabla_{\tilde{\mathbf{x}}}, \quad \left. \frac{d}{dt} \right|_{\text{frame 2}} := \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}}. \quad (4.13)$$

**Proposition 4.2.1.** With  $\mathbf{b}$  defined by Eq. (4.11), we have

$$\begin{aligned} \left. \frac{d}{dt} \right|_{\text{frame 1}} \mathbf{b} - \kappa : \left\langle \sum_{j=1}^{N-1} \mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} \otimes \mathbf{b} \right\rangle &= \frac{k_B T}{\gamma} \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b} \right\rangle \\ &\quad - \frac{1}{\gamma} \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V_p(\mathbf{r}) \cdot \nabla_{\mathbf{r}_j} \mathbf{b} \right\rangle, \end{aligned} \quad (4.14)$$

obeys rotational symmetry.

**Proof.** Let us choose the vector  $\mathbf{r}^* = [|\mathbf{r}_1|, |\mathbf{r}_2|, |\mathbf{r}_{12}|, |\mathbf{r}_3|, |\mathbf{r}_{13}|, |\mathbf{r}_{23}|, \dots, |\mathbf{r}_{N-2,N-1}|]$ . Denote by  $r_i^*$  the  $i$ -th element of  $\mathbf{r}^*$  and  $\mathbf{r}_i^*$  the corresponding the 3-dimensional vector, i.e.,  $r_6^* = |\mathbf{r}_{23}|$  and  $\mathbf{r}_6^* = \mathbf{r}_{23}$ . Following Eq. (4.11),  $\mathbf{b}$  consists of

$$\mathbf{b} = \sum_{j,k=1}^{N-1} \mathbf{b}_{jk}, \quad \mathbf{b}_{jk} = g(\mathbf{r}^*)\mathbf{r}_j\mathbf{r}_k^T, \quad (4.15)$$

where  $g(\mathbf{r}^*)$  denotes  $g_j^{(1)}(\mathbf{r}^*)g_k^{(2)}(\mathbf{r}^*)$  for simplicity. With this general form, we have

$$\frac{d}{dt} \langle \tilde{\mathbf{b}}_{jk} \rangle \Big|_{\text{frame 1}} = \dot{\mathbf{Q}} \langle \mathbf{b}_{jk} \rangle \mathbf{Q}^T + \mathbf{Q} \langle \mathbf{b}_{jk} \rangle \dot{\mathbf{Q}}^T + \mathbf{Q} \frac{d}{dt} \langle \mathbf{b}_{jk} \rangle \Big|_{\text{frame 2}} \mathbf{Q}^T. \quad (4.16)$$

Moreover, we note that

$$\begin{aligned} \tilde{\kappa} : & \left( \sum_{i=1}^{N-1} \tilde{\mathbf{r}}_i \otimes \nabla_{\tilde{\mathbf{r}}_i} \otimes \tilde{\mathbf{b}}_{jk} \right) \\ &= \sum_{i=1}^{N-1} \left[ \left( \mathbf{Q} \kappa \mathbf{Q}^T + \dot{\mathbf{Q}} \mathbf{Q}^T \right) \cdot \mathbf{Q} \mathbf{r}_j \right] \cdot \mathbf{Q} \otimes \nabla_{\mathbf{r}_i} \otimes \left( \mathbf{Q} \mathbf{b}_{jk} \mathbf{Q}^T \right) \\ &= \sum_{i=1}^{N-1} (\kappa \cdot \mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} \left( \mathbf{Q} \mathbf{b}_{jk} \mathbf{Q}^T \right) + (\mathbf{Q}^T \dot{\mathbf{Q}} \mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} \left( \mathbf{Q} \mathbf{b}_{jk} \mathbf{Q}^T \right) \\ &= \sum_{i=1}^{N-1} \mathbf{Q} (\kappa \cdot \mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} \mathbf{b}_{jk} \mathbf{Q}^T + \mathbf{Q} \left( \mathbf{Q}^T \dot{\mathbf{Q}} \mathbf{b}_{jk} + \mathbf{b}_{jk} \dot{\mathbf{Q}}^T \mathbf{Q} \right) \mathbf{Q}^T \\ &\quad + \mathbf{Q} \left( \sum_{i=1}^{N-1} \mathbf{r}_i^T (\dot{\mathbf{Q}}^T \mathbf{Q}) \nabla_{\mathbf{r}_i} g(\mathbf{r}^*) \right) \mathbf{r}_j r_k^T \mathbf{Q}^T \\ &= \sum_{i=1}^{N-1} \mathbf{Q} (\kappa \cdot \mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} \mathbf{b}_{jk} \mathbf{Q}^T + \dot{\mathbf{Q}} \mathbf{b}_{jk} \mathbf{Q}^T + \mathbf{Q} \mathbf{b}_{jk} \dot{\mathbf{Q}}^T, \end{aligned} \quad (4.17)$$

where we have used  $\mathbf{r}_i^T (\dot{\mathbf{Q}}^T \mathbf{Q}) \mathbf{r}_i \equiv 0$  since  $\dot{\mathbf{Q}}^T \mathbf{Q}$  is anti-symmetric. Eq. (4.16) and Eq. (4.17) shows that the combination of the two terms on the left-hand-side of Eq. (4.14) rigorously preserve the rotational symmetry, i.e.,

$$\left( \frac{d}{dt} \langle \tilde{\mathbf{b}} \rangle - \tilde{\kappa} : \sum_{i=1}^{N-1} \langle \tilde{\mathbf{r}}_i \otimes \nabla_{\tilde{\mathbf{r}}_i} \otimes \tilde{\mathbf{b}} \rangle \right) \Big|_{\text{frame 1}} \equiv \mathbf{Q} \left( \frac{d}{dt} \langle \mathbf{b} \rangle - \kappa : \sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle \right) \Big|_{\text{frame 2}} \mathbf{Q}^T.$$

It is straightforward to prove rotational symmetry for the other terms in Eq (4.14).

#### 4.2.4 Symmetry-preserving DNN models

To complete the DeePN<sup>2</sup> model, we need to specify the DNN models. These DNN models should also strictly preserve rotational symmetry. Different from the rotational-invariant scalar stress model considered in Ref. (Zhou et al., 2021), the second-order tensors  $\mathbf{G}$ ,  $\mathbf{H}_{1,i}$ ,  $\mathbf{H}_{2,i}$  need to satisfy the symmetry condition (4.4) and the fourth-order tensors  $\mathcal{E}_i$  need to retain the objectivity of  $\frac{D\mathbf{c}_i}{Dt}$ . However, there does not exist such a reference frame in which these symmetry constraints can be satisfied by the macro-scale modeling terms.



To handle this problem, we consider the eigen-space of the feature  $\mathbf{c}_1$  with a fixed form of the encoder  $\mathbf{b}_1(\cdot)$ , e.g., by setting  $g_1(\cdot) = w_{1,\cdot} \equiv 1$  and let other  $\mathbf{b}_i(\cdot)$  involved in the training. Consider the eigen-decomposition  $\mathbf{c}_1 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$  has distinct eigenvalues, where  $\mathbf{U}$  is the matrix whose columns are the eigenvectors of  $\mathbf{c}_1$ .  $\mathbf{U}$  is not unique due to the non-uniqueness of the eigenvectors. Without loss of generality, we further assume that the first element of  $\mathbf{u}_1$  to be positive. With the following lemma, we show that the general form of  $\mathbf{U}$  can be always written as  $\mathbf{U}^{(j)} := \mathbf{U}S^{(j)}$  with  $j = 1, \dots, 4$ , where  $S^{(j)}$  is given by

$$S^{(1)} = \begin{pmatrix} +1 & & \\ & +1 & \\ & & +1 \end{pmatrix}, S^{(2)} = \begin{pmatrix} +1 & & \\ & -1 & \\ & & +1 \end{pmatrix}, S^{(3)} = \begin{pmatrix} +1 & & \\ & +1 & \\ & & -1 \end{pmatrix}, S^{(4)} = \begin{pmatrix} +1 & & \\ & -1 & \\ & & -1 \end{pmatrix}.$$

**Lemma 4.2.2.** For a symmetry matrix  $M \in \mathbb{R}^{3 \times 3}$ , let  $S_M$  denote the set of matrices with the transformation of  $S^{(j)}$ , i.e.,  $S_M := \{S^{(1)}MS^{(1)}, \dots, S^{(4)}MS^{(4)}\}$ . For any  $M^{(j)} := S^{(j)}MS^{(j)} \in S_M$ ,  $S^{(k)}M^{(j)}S^{(k)} \in S_M$ ,  $1 \leq j, k \leq 4$ . Furthermore,  $S_M$  can be constructed by  $M^{(j)}$ , i.e.,  $S_M \equiv \{S^{(1)}M^{(j)}S^{(1)}, \dots, S^{(4)}M^{(j)}S^{(4)}\}$ .

**Proof.** By applying  $S^{(j)}$  to  $M$ , it is easy to see that the diagonal part of  $M^{(j)}$  remains the same. Since  $M^{(j)}$  is also symmetric, we only need to check the upper-triangular part, taking the four possible operations

$$\begin{pmatrix} * & + & + \\ & * & + \\ & & * \end{pmatrix} \begin{pmatrix} * & - & + \\ & * & - \\ & & * \end{pmatrix} \begin{pmatrix} * & + & - \\ & * & - \\ & & * \end{pmatrix} \begin{pmatrix} * & - & + \\ & * & - \\ & & * \end{pmatrix},$$

where “+” represents that the element remains the same and “-” represents a sign change. We see that number of “-” operations is either 0 or 2. Starting from any of the above choice for  $M^{(j)}$ , all of the four operators yields either 0 or 2 “-” operations. Therefore,  $S^{(k)}M^{(j)}S^{(k)} \in S_M$ . Furthermore, if the upper triangular part of  $M$  has distinct absolute values, then  $\forall M^{(j)}, S^k M^j S^k \neq S^{k'} M^j S^{k'}$  with  $k \neq k'$ , hence  $S_M$  can be constructed by  $M^j$ . Otherwise, if some upper triangular entries of  $M$  share the same absolute value, we can draw the same conclusion accordingly.

Now we consider the matrix whose columns are the eigenvectors of  $\tilde{\mathbf{c}}_1 = \mathbf{Q}\mathbf{c}_1\mathbf{Q}^T$ , denoted by  $\tilde{\mathbf{U}}$ . We can write  $\tilde{\mathbf{U}} = \mathbf{Q}\mathbf{U}\mathbf{S}^{(j)}$ , where  $j \in \{1, 2, 3, 4\}$ . Accordingly, the DNN input of  $\mathbf{c}_i$  takes the form

$$\tilde{\mathbf{U}}^T \tilde{\mathbf{c}}_i \tilde{\mathbf{U}} = \left( \mathbf{Q}\mathbf{U}\mathbf{S}^{(j)} \right)^T \mathbf{Q}\mathbf{c}_i\mathbf{Q}^T \left( \mathbf{Q}\mathbf{U}\mathbf{S}^{(j)} \right) = \mathbf{S}^{(j)} \mathbf{U}^T \mathbf{c}_i \mathbf{U} \mathbf{S}^{(j)}.$$

Let  $M = \mathbf{U}^T \mathbf{c}_i \mathbf{U}$ , by using Lemma 4.2.2, it is easy to see that  $S_{\mathbf{U}^T \mathbf{c}_i \mathbf{U}}$  can be constructed by taking  $j = 1, \dots, 4$ .

**Proposition 4.2.3.** Let  $\mathbf{U}$  be the matrix whose columns are the eigenvectors of  $\mathbf{c}_1$ . Let the DNN input be  $\hat{\mathbf{c}}_i^{(j)} = \mathbf{S}^{(j)} \mathbf{U}^T \mathbf{c}_i \mathbf{U} \mathbf{S}^{(j)}$ . The following form of  $\tau_p$

$$\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \frac{1}{4} \sum_{j=1}^4 \mathbf{U}^{(j)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j)}, \dots, \hat{\mathbf{c}}_n^{(j)}) \mathbf{U}^{(j)T}, \quad \mathbf{U}^{(j)} = \mathbf{U} \mathbf{S}^{(j)}. \quad (4.18)$$

satisfies the rotational symmetry constraint (4.4).

Finally, to account for the swap of the eigenvectors when the eigenvalues cross over, we consider the 6 permutations of the three eigenvalues of  $\mathbf{c}_1$ , i.e.,

$$\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \frac{1}{24} \sum_{k=0}^5 \sum_{j=1}^4 \mathbf{U}^{(j,k)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j,k)}, \dots, \hat{\mathbf{c}}_n^{(j,k)}) \mathbf{U}^{(j,k)T}, \quad (4.19)$$

where  $k$  represents the rank of permutation (e.g., in lexicographical order) and  $\mathbf{U}^{(j,k)}$  is a variation of  $\mathbf{U}^{(j)}$  with corresponding column permutation.

During simulation, the eigenvalues of  $\mathbf{c}_1$  may cross each other. To account for this, we consider all the 6 permutations of the three eigenvalues, i.e.,

$$\mathbf{G}(\mathbf{c}_1, \dots, \mathbf{c}_n) = \frac{1}{24} \sum_{k=0}^5 \sum_{j=1}^4 \mathbf{U}^{(j,k)} \hat{\mathbf{G}}(\hat{\mathbf{c}}_1^{(j,k)}, \dots, \hat{\mathbf{c}}_n^{(j,k)}) \mathbf{U}^{(j,k)T}, \quad (4.20)$$

where  $k$  represents the rank of permutation (e.g., in lexicographical order) and  $\mathbf{U}^{(j,k)}$  is a variation of  $\mathbf{U}^{(j)}$  with corresponding column permutation. Furthermore, to avoid the eigenvector degeneracy, we set a threshold value  $\epsilon$  for the eigenvalues. When two eigenvalues approach each other, e.g.,  $|\lambda_2 - \lambda_3| < \epsilon$ , we freeze all the eigenvectors until  $|\lambda_2 - \lambda_3| \geq \epsilon$ . In this work, we take  $\epsilon = 10^{-3}$ , and we refer to Section 4.3.6 for detailed numerical studies.

Eq. (4.20) provides the rotation-symmetric form for the second-order stress tensor  $\mathbf{G}$ , where  $\hat{\mathbf{G}}$  is represented by DNNs. The constitutive model terms  $\mathbf{H}_{1,i}$  and  $\mathbf{H}_{2,i}$  can be constructed in a similar

manner. Finally, we can show the fourth-order tensors  $\{\mathcal{E}_i\}_{i=1}^n$  associated with the encoders (4.6) can be constructed in the form

$$\kappa : \mathcal{E}_i = \kappa \mathbf{c}_i + \mathbf{c}_i \kappa^T + \kappa : \left( \sum_{j=1}^9 \mathbf{E}_{1,i}^{(j)} \otimes \mathbf{E}_{2,i}^{(j)} \right), \quad (4.21)$$

where  $\mathbf{E}_{1,i}^{(j)}$  and  $\mathbf{E}_{2,i}^{(j)}$  are second-order tensors which respect the symmetry condition (4.4) and can be constructed in the form of Eq. (4.20) (see Prop. 4.2.4). The constructed DeePN<sup>2</sup> model takes the form similar to the general hydrodynamic equations (4.2) and (4.3), where some of the model terms are represented by DNNs in the form of Eqs. (4.20) and (4.21).

**Proposition 4.2.4.** The following ansatz of  $\langle \sum_{i=1}^{N-1} \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle$  ensures that the dynamic of evolution of  $\mathbf{c}$  retains rotational invariance.

$$\begin{aligned} \sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle &= \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) (\mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} + \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k}) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle \\ &\quad + \sum_{k=1}^9 \mathbf{E}_1^{(k)}(\mathbf{c}) \otimes \mathbf{E}_2^{(k)}(\mathbf{c}), \end{aligned} \quad (4.22)$$

where  $\mathbf{c} = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ ,  $\tilde{\mathbf{c}} = (\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_n)$ , and  $\mathbf{E}_1$  and  $\mathbf{E}_2$  satisfy

$$\tilde{\mathbf{E}}_1 := \mathbf{E}_1(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{E}_1(\mathbf{c}) \mathbf{Q}^T, \quad \tilde{\mathbf{E}}_2 := \mathbf{E}_2(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{E}_2(\mathbf{c}) \mathbf{Q}^T. \quad (4.23)$$

**Proof.** Without loss of generality, we represent the fourth order tensor by the following two bases

$$\begin{aligned} \mathbf{F}_1(\mathbf{c}) \otimes \mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_3(\mathbf{c}) + \mathbf{F}_3(\mathbf{c}) \otimes (\mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_1(\mathbf{c}))^{T_{\{2,3\}}}, \quad \mathbf{F}_1(\mathbf{c}), \mathbf{F}_3(\mathbf{c}) \in \mathbb{R}^3, \mathbf{F}_2(\mathbf{c}) \in \mathbb{R}^{3 \times 3}, \\ \mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c}), \quad \mathbf{E}_1(\mathbf{c}), \mathbf{E}_2(\mathbf{c}) \in \mathbb{R}^{3 \times 3}, \end{aligned} \quad (4.24)$$

where the super-script  $T_{\{2,3\}}$  represents the transpose between the 2nd and 3rd indices; also  $\mathbf{F}_1$ ,  $\mathbf{F}_2$ ,  $\mathbf{F}_3$ ,  $\mathbf{E}_1$  and  $\mathbf{E}_2$  satisfy the symmetry conditions

$$\begin{aligned} \mathbf{F}_1(\tilde{\mathbf{c}}) &= \mathbf{Q} \mathbf{F}_1(\mathbf{c}), \quad \mathbf{F}_3(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{F}_3(\mathbf{c}), \\ \mathbf{E}_1(\tilde{\mathbf{c}}) &= \mathbf{Q} \mathbf{E}_1(\mathbf{c}) \mathbf{Q}^T, \quad \mathbf{E}_2(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{E}_2(\mathbf{c}) \mathbf{Q}^T, \quad \mathbf{F}_2(\tilde{\mathbf{c}}) = \mathbf{Q} \mathbf{F}_2(\mathbf{c}) \mathbf{Q}^T. \end{aligned} \quad (4.25)$$

For the term  $\mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c})$ , we have

$$\kappa : \mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c}) = \text{Tr}(\kappa \mathbf{E}_1(\mathbf{c})) \mathbf{E}_2(\mathbf{c}) \quad (4.26)$$

and

$$\begin{aligned}
\tilde{\kappa} : \tilde{\mathbf{E}}_1 \otimes \tilde{\mathbf{E}}_2|_{\text{frame } 1} &= \left( \mathbf{Q} \kappa \mathbf{Q}^T + \dot{\mathbf{Q}} \mathbf{Q}^T \right) : \left( \mathbf{Q} \mathbf{E}_1(\mathbf{c}) \mathbf{Q}^T \otimes \tilde{\mathbf{E}}_2 \right) \\
&= \text{Tr}(\kappa \mathbf{E}_1(\mathbf{c})) \tilde{\mathbf{E}}_2 + \text{Tr}(\dot{\mathbf{Q}} \mathbf{Q}^T \mathbf{Q} \mathbf{E}_1(\mathbf{c}) \mathbf{Q}^T) \tilde{\mathbf{E}}_2 \\
&= \text{Tr}(\kappa \mathbf{E}_1(\mathbf{c})) \tilde{\mathbf{E}}_2 \\
&\equiv \mathbf{Q} \left( \kappa : \mathbf{E}_1(\mathbf{c}) \otimes \mathbf{E}_2(\mathbf{c})|_{\text{frame } 2} \right) \mathbf{Q}^T,
\end{aligned} \tag{4.27}$$

where we have used  $\text{Tr}(\dot{\mathbf{Q}} \mathbf{Q}^T) \equiv 0$ .

For the term  $\mathbf{F}_1(\mathbf{c}) \otimes \mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_3(\mathbf{c}) + \mathbf{F}_3(\mathbf{c}) \otimes (\mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_1(\mathbf{c}))^{T_{\{2,3\}}}$ , we have

$$\kappa : \mathbf{F}_1(\mathbf{c}) \otimes \mathbf{F}_2(\mathbf{c}) \otimes \mathbf{F}_3(\mathbf{c}) = \mathbf{F}_2(\mathbf{c})^T \kappa \mathbf{F}_1(\mathbf{c}) \mathbf{F}_3(\mathbf{c})^T \tag{4.28}$$

and

$$\tilde{\kappa} : \tilde{\mathbf{F}}_1 \otimes \tilde{\mathbf{F}}_2 \otimes \tilde{\mathbf{F}}_3 = \mathbf{Q} \mathbf{F}_2(\mathbf{c})^T \kappa \mathbf{F}_1(\mathbf{c}) \mathbf{F}_3(\mathbf{c})^T \mathbf{Q}^T + \mathbf{Q} \mathbf{F}_2(\mathbf{c})^T \mathbf{Q}^T \dot{\mathbf{Q}} \mathbf{F}_1(\mathbf{c}) \mathbf{F}_3(\mathbf{c})^T \mathbf{Q}^T. \tag{4.29}$$

On the other hand, we note that

$$\frac{d\tilde{\mathbf{b}}}{dt}|_{\text{frame } 1} = \dot{\mathbf{Q}} \mathbf{b} \mathbf{Q}^T + \mathbf{Q} \mathbf{b} \dot{\mathbf{Q}}^T + \mathbf{Q} \frac{d\mathbf{b}}{dt}|_{\text{frame } 2} \mathbf{Q}^T. \tag{4.30}$$

To ensure the rotational symmetry of  $\frac{\partial \mathbf{b}}{\partial t}$ , we have

$$\mathbf{F}_2 \equiv \mathbf{I}, \quad \sum_i \mathbf{F}_1^{(i)} \otimes \mathbf{I} \otimes \mathbf{F}_3^{(i)} = \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) \mathbf{r}_j \otimes \mathbf{I} \otimes \mathbf{r}_k \right\rangle. \tag{4.31}$$

Hence, we have

$$\begin{aligned}
\frac{d}{dt} \tilde{\mathbf{c}} - \tilde{\kappa} : \left( \sum_i \tilde{\mathbf{F}}_1^{(i)} \otimes \tilde{\mathbf{F}}_2^{(i)} \otimes \tilde{\mathbf{F}}_3^{(i)} + \tilde{\mathbf{F}}_3^{(i)} \otimes \left( \tilde{\mathbf{F}}_2^{(i)} \otimes \tilde{\mathbf{F}}_1^{(i)} \right)^{T_{\{2,3\}}} \right) \Big|_{\text{frame } 1} \\
\equiv \mathbf{Q} \left( \frac{d}{dt} \mathbf{c} - \kappa : \left( \sum_i \mathbf{F}_1^{(i)} \otimes \mathbf{F}_2^{(i)} \otimes \mathbf{F}_3^{(i)} + \mathbf{F}_3^{(i)} \otimes \left( \mathbf{F}_2^{(i)} \otimes \mathbf{F}_1^{(i)} \right)^{T_{\{2,3\}}} \right) \right) \Big|_{\text{frame } 2} \mathbf{Q}^T.
\end{aligned} \tag{4.32}$$

Furthermore, using Eq. (4.31), we obtain

$$\begin{aligned}
&\sum_i \mathbf{F}_1^{(i)} \otimes \mathbf{F}_2^{(i)} \otimes \mathbf{F}_3^{(i)} + \mathbf{F}_3^{(i)} \otimes \left( \mathbf{F}_2^{(i)} \otimes \mathbf{F}_1^{(i)} \right)^{T_{\{2,3\}}} \\
&= \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) (\mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} + \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k}) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle.
\end{aligned} \tag{4.33}$$

Accordingly, the remaining part of  $\sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle$  is expanded by

$$\left\langle \sum_{i=1}^{N-1} \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \sum_{j,k=1}^{N-1} g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle = \sum_{i=1}^9 \mathbf{E}_1^{(i)}(\mathbf{c}) \otimes \mathbf{E}_2^{(i)}(\mathbf{c}). \quad (4.34)$$

Combining Eq. (4.32), (4.33) and (4.34), we conclude that the decomposition

$$\begin{aligned} \sum_{i=1}^{N-1} \langle \mathbf{r}_i \otimes \nabla_{\mathbf{r}_i} \otimes \mathbf{b} \rangle &= \sum_{j,k=1}^{N-1} \left\langle g_j^{(1)}(\mathbf{r}^*) g_k^{(2)}(\mathbf{r}^*) (\mathbf{r}_j \otimes \nabla_{\mathbf{r}_j} + \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k}) \otimes \mathbf{r}_j \mathbf{r}_k^T \right\rangle \\ &\quad + \sum_{k=1}^9 \mathbf{E}_1^{(k)}(\mathbf{c}) \otimes \mathbf{E}_2^{(k)}(\mathbf{c}) \end{aligned} \quad (4.35)$$

ensures the objectivity of the time-derivative of  $\mathbf{c}$ .

### 4.3 Numerical results

The present DeePN<sup>2</sup> model is trained using micro-scale samples collected from the homogeneous shear flow. We demonstrate the model accuracy and generalization ability by considering various flows in comparison with the results of the micro-scale simulations for the suspensions with three different polymer structural models as shown in Fig. 4.1. As we will see, the micro-scale structure does play an important role in the viscoelastic response. We will use this to examine the DeePN<sup>2</sup> model fidelity.

#### 4.3.1 Micro-scale model of the polymer solutions

In the present study, we consider suspensions with three different polymer structures as shown in Fig. 4.1. Each polymer molecule consists of  $N = 7$  beads connected with  $N_b$  FENE bonds, i.e.,

$$V(\mathbf{q}) = \sum_{j=1}^{N_b} V_b(|\mathbf{q}_{j_1} - \mathbf{q}_{j_2}|), \quad V_b(l) = -\frac{k_s}{2} l_0^2 \log \left[ 1 - \frac{l^2}{l_0^2} \right], \quad (4.36)$$

where  $k_s$  represents the spring constant and  $l_0$  is the maximum of the extension length. The chain- and star-shaped molecules have  $N_b = 6$  bonds with the same bond parameters  $k_s = 0.1$  and  $l_0 = 2.3$  (in reduced unit). The net-shaped molecule is similar to the star-shaped molecule with the same parameters for the first 6 bonds; 3 additional bonds connect the side chain particles with  $k_s = 0.1$  and  $l_0 = 3.7$ . The polymer number density of the three suspensions is  $n_p = 0.3$ . The solvent is modeled by the dissipative particle dynamics (DPD) (Hoogerbrugge and Koelman, 1992; Groot and

Warren, 1997) with number density  $n_s = 4.0$ . The pairwise interaction between particle  $i$  and  $j$  takes the standard form

$$\mathbf{F}_{ij} = \mathbf{F}_{ij}^C + \mathbf{F}_{ij}^D + \mathbf{F}_{ij}^R,$$

$$\mathbf{F}_{ij}^C = \begin{cases} a(1.0 - r_{ij}/r_c)\mathbf{e}_{ij}, & r_{ij} < r_c \\ 0, & r_{ij} > r_c \end{cases},$$

$$\mathbf{F}_{ij}^D = \begin{cases} -\gamma w^D(r_{ij})(\mathbf{v}_{ij} \cdot \mathbf{e}_{ij})\mathbf{e}_{ij}, & r_{ij} < r_c \\ 0, & r_{ij} > r_c \end{cases}, \quad \mathbf{F}_{ij}^R = \begin{cases} \sigma w^R(r_{ij})\xi_{ij}\mathbf{e}_{ij}, & r_{ij} < r_c \\ 0, & r_{ij} > r_c \end{cases},$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ ,  $r_{ij} = |\mathbf{r}_{ij}|$ ,  $\mathbf{e}_{ij} = \mathbf{r}_{ij}/r_{ij}$ , and  $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$ ,  $\xi_{ij}$  are independent identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance.  $\gamma$  and  $\sigma$  are related with the system temperature by the second fluctuation-dissipation theorem (Español and Warren, 1995) as  $\sigma^2 = 2\gamma k_B T$ , where  $k_B T$  is set to 0.25. The detailed parameters are given in Tab. 4.1.

Table 4.1 Parameters (in reduced unit) of the micro-scale model of the polymer solution (S-solvent, P-polymer).

	$a$	$\gamma$	$\sigma$	$k$	$r_c$
S-S	4.0	5.0	1.58	0.25	1.0
S-P	0.0	40.0	4.47	0.0	1.0
P-P	4.0	0.01	0.071	1.0	0.7

### 4.3.2 Collecting training samples

Collecting training samples is one of the most important steps in the construction of DeePN<sup>2</sup>. To obtain reliable models, we need to ensure that the training sample set is representative enough of all the practical situations that the model is intended for. In the present study, we collect the training samples in shear flow with shear rate  $\dot{\gamma} \in [0, 0.09]$ . Since the training of the DeePN<sup>2</sup> model only requires discrete polymer configurations rather than time-series samples, one convenient approach is to consecutively increase the shear rate and collect the discrete configurations during the shear extension and relaxation process, where the inclusion of the relaxation process can facilitate the sampling of polymer configuration phase space due to the viscoelastic hysteresis effect. 32000 samples are collected where each sample consists of 5000 polymer configurations, which will be employed to evaluate the constitutive dynamics terms  $\langle \cdot \rangle$ . Due to the permutation symmetry of the

the particle label, the effective number of configurations per sample is  $1 \times 10^4$  for the chain-shaped molecule and  $3 \times 10^4$  for the star- and net-shaped molecules.

### 4.3.3 Training procedure

The DeePN<sup>2</sup> model is constructed via the training of the NN representations of the encoder mappings  $\{g_j(\mathbf{r}^*)\}_{j=1}^n$ , stress model  $\mathbf{G}$ , evolution dynamics  $\{\mathbf{H}_{1,j}\}_{j=1}^n$ ,  $\{\mathbf{H}_{2,j}\}_{j=1}^n$  and the 4th order tensors  $\{\mathcal{E}_j\}_{j=1}^n$  of the objective tensor derivatives. In this study, we choose  $n = 3$  encoders and fix  $g_1(\mathbf{r}^*) \equiv 1$ . For the chain-shaped molecule, we set  $w_{1,i} = 1 - i/N$ ,  $1 \leq i \leq N - 1$  and  $\sum_i w_{1,i} \mathbf{r}_i$  represents the orientation between the free-end particle and the center of mass. For the star- and net-shaped molecules, we set  $w_{1,1} = 1$  and  $w_{1,i} = 0$  for  $i \geq 2$ . All terms are represented by the fully connected NN. The number of hidden layers are set to be (120, 120, 120), (300, 300, 300), (400, 400, 400), (450, 450, 450), (560, 560, 560), respectively. The activation function is taken to be the hyperbolic tangent. We emphasize that the mappings  $\{g_j(\mathbf{r}^*)\}_{j=1}^n$  and weights  $w \in \mathbb{R}^{n \times (N-1)}$  involve in the training process for the joint learning of the encoders  $\{\mathbf{b}_j(\mathbf{r})\}_{j=1}^n$  defined in Eq. (4.6) and the macro-scale features  $\{\mathbf{c}_j\}_{j=1}^n$ , although they do not appear explicitly in the macro-scale hydrodynamic equations.

The DNNs are trained by the Adam stochastic gradient descent method (Kingma and Ba, 2015) for 20 epochs, using 5 samples per batch size. The initial learning rate is  $2.8 \times 10^{-4}$  and decay rate is 0.75 per 20000 steps.

Similar to Ref. (Lei et al., 2020), the loss function is defined by

$$L = \lambda_G L_G + \lambda_{H_1} L_{H_1} + \lambda_{H_2} L_{H_2} + \lambda_{\mathcal{E}} L_{\mathcal{E}},$$

where  $\lambda_G = 0.2$ ,  $\lambda_{H_1} = 0.1$ ,  $\lambda_{H_2} = 0.6$  and  $\lambda_{\mathcal{E}} = 0.1$  are hyperparameters. For each training batch of  $m$  training samples,  $L_G$ ,  $L_{H_1}$ ,  $L_{H_2}$ ,  $L_{\mathcal{E}}$  of the system are given by

$$\begin{aligned}
L_G &= \sum_{l=1}^m \sum_{i=1}^n \left\| \mathbf{G}_i(\mathbf{c}^{(l)}) - \left\langle \sum_{k=1}^{N_b} \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k} V \right\rangle^{(l)} \right\|^2 \\
L_{H_1} &= \sum_{l=1}^m \sum_{i=1}^n \left\| \mathbf{H}_{1,i}(\mathbf{c}^{(l)}) - \left\langle \sum_{j,k=1}^{N-1} A_{jk} \nabla_{\mathbf{r}_j} \cdot \nabla_{\mathbf{r}_k} \mathbf{b}_i \right\rangle^{(l)} \right\|^2 \\
L_{H_2} &= \sum_{l=1}^m \sum_{i=1}^n \left\| \mathbf{H}_{2,i}(\mathbf{c}^{(l)}) - \left\langle \sum_{j=1}^{N-1} \sum_{k=1}^{N_b} A_{jk} \nabla_{\mathbf{r}_k} V \cdot \nabla_{\mathbf{r}_j} \mathbf{b}_i \right\rangle^{(l)} \right\|^2 \\
L_{\mathcal{E}} &= \sum_{l=1}^m \sum_{i=1}^n \left\| \sum_{s=1}^9 \mathbf{E}_{1,i}^{(s)}(\mathbf{c}^{(l)}) \otimes \mathbf{E}_{2,i}^{(s)}(\mathbf{c}^{(l)}) - \left\langle \sum_{k=1}^{N-1} \mathbf{r}_k \otimes \nabla_{\mathbf{r}_k} g_i^2 \otimes \sum_{j,j'=1}^{N-1} w_{ij} w_{ij'} \mathbf{r}_j \mathbf{r}_{j'}^T \right\rangle^{(l)} \right\|^2,
\end{aligned} \tag{4.37}$$

where  $\|\cdot\|^2$  denotes the total sum of squares of the entries in the tensor, and  $\mathbf{c}^{(l)} = (\mathbf{c}_1^{(l)}, \dots, \mathbf{c}_n^{(l)})$ .

#### 4.3.4 Numerical Result of Reverse Poiseuille flow

First, we consider the reverse Poiseuille flow in a  $60 \times 100 \times 60$  domain (in reduced unit) with the opposite body force  $\mathbf{f}_{\text{ext}} = (0.016, 0, 0)$  applied to each half of the domain divided by the plane  $y = 50$  starting from  $t = 0$ . At  $t = 800$ , the external force is removed. The relaxation process of the flow field is recorded until the total simulation time  $t = 1600$ . For all the three systems, the predictions from DeePN<sup>2</sup> agree well with the micro-scale simulations results, as shown in Fig. 4.2. In particular, the flow velocity fields of the three systems are nearly identical at the initial stage  $t \in [0, 200]$ , as the development of the flow field is dominated by the solvent and the near-equilibrium responses of the polymer molecules in this regime. Starting from  $t = 250$ , the velocity fields of the three systems exhibit distinct evolution processes. The velocity of the chain-shaped molecule suspension exhibits the largest oscillation and the longest development stage during  $t \in [250, 800]$ . In contrast, the velocity of the star-shaped molecule suspension exhibits moderate oscillation and shows an apparent increase during  $t \in [400, 800]$ , indicating that the polymer elastic energy reaches a plateau earlier than the chain-shaped system. Moreover, the velocity of the net-shaped molecule suspension exhibits the smallest oscillation, indicating that the three additional side-chains further affect the rheological properties of the polymer suspension.

Such differences can also be studied by examining the polymer stress development. As shown



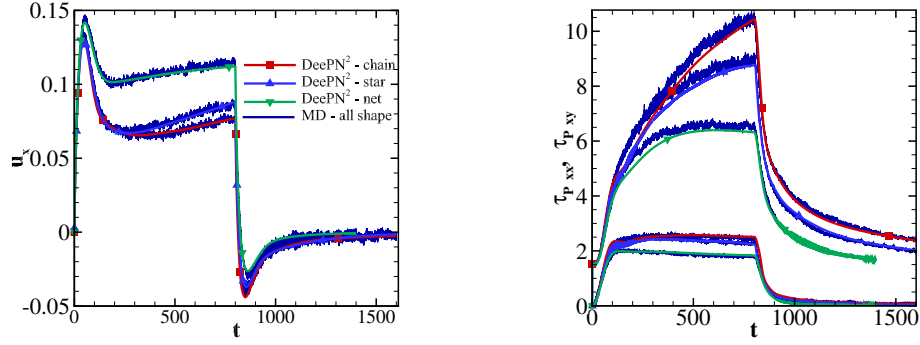


Figure 4.2 The velocity  $u_x$  (left) and polymer stress  $\tau_p$  (right) of the reverse Poiseuille flow ( $y = 6$ ) of the polymer suspensions of three different molecule structures shown in Fig. 4.1.  $\tau_p$  is normalized by polymer number density  $n_p$ , i.e., it is the stress energy per polymer (the same for the remaining figures). With the same FENE bond, the polymer suspensions exhibit different flow responses due to the different molecule structural mechanics. The dark blue lines with rough oscillations denote the micro-scale simulation results; the solid lines with symbols denote the DeePN<sup>2</sup> predictions.

in Fig. 4.2, the value of  $\tau_{p,xx}$  for the chain-shaped molecule suspension keeps increasing through the development stage  $t \in [0, 800]$  while for the star-shaped molecule,  $\tau_{p,xx}$  shows only a moderate increase. In contrast, the net-shaped molecule suspension reaches steady state at about  $t = 400$ . Moreover, the steady value of the shear stress  $\tau_{p,xy}$  of the chain-shaped molecule is also larger than the star-shaped and the net-shaped molecules, indicating the largest restored elastic energy. This result is also consistent with the larger velocity oscillation from the minimal values to 0 during the relaxation process with  $t \in [800, 1000]$ .

The different rheological properties of the three polymer suspensions can be understood as follows. Although both the chain-shaped and star-shaped molecules have 6 identical FENE bonds, the chain-shaped molecule is less symmetric than the star-shaped molecule. Accordingly, it shows larger dispersion in the  $\mathbb{R}^{18}$  configuration space, and hence, is more flexible than the star-shaped molecule. The elastic response time of the chain-shaped molecule suspension is longer than that of the star-shaped molecule suspension; larger elastic energy can be restored during the relaxation stage. On the other hand, the net-shaped molecule is more rigid than the star-shaped molecule due to the additional bond interaction.

Another important feature of non-Newtonian fluids is the hysteresis effect. Classical models such

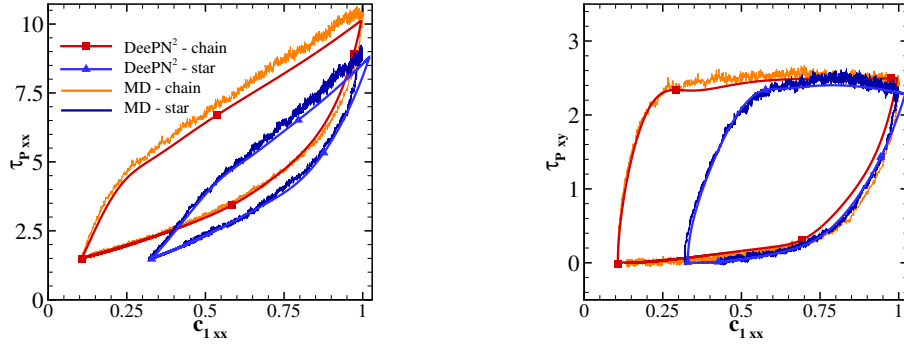


Figure 4.3 The evolution of the polymer stress  $\tau_p$  and conformation tensor  $\mathbf{c}_1$  obtained from the reverse Poiseuille flow ( $y = 6$ ) of the polymer suspensions. The clockwise loops represent the development and relaxation processes. For the visualization, the conformation tensor component  $c_{1,xx}$  is rescaled by the maximum value obtained from the micro-scale simulation.

as Hookean and FENE-P cannot capture such effects (Doyle et al., 1998; Lielens et al., 1998). Fig. 4.3 shows the evolution of the polymer stress and conformation tensor for the chain- and star-shaped molecule suspensions. The clockwise loops show the hysteresis effects during the development and relaxation processes; the non-unique stress values indicate that linear and mean field approximations are insufficient in describing the viscoelastic response of the system. In contrast, these effects are accurately captured with the DeePN<sup>2</sup> model. Similar to Fig. 4.2, the chain-shaped molecule suspension shows more pronounced hysteresis effect due to the larger dispersion in the configuration space, reflected as the larger “loop area” than the results for star-shaped molecule suspension.

#### 4.3.5 Numerical Result of Womersley flow

Next, we investigate the Womersley flow (Womersley, 1955) by applying the opposite oscillating body force  $\mathbf{f}_{\text{ext}} = (\pm f_0 \cos(2\pi\omega t), 0, 0)$  to each half of the domain along the  $z$ -direction, where we set  $f_0 = 0.012$  and  $\omega = 1/3000$ . Fig. 4.4 shows the velocity development of the star- and net-shaped molecule suspensions. Similar to the reverse Poiseuille flow, the net-shaped molecule suspension shows less pronounced viscoelastic responses, reflected as the slower decay near  $t \in [200, 400]$  and the larger oscillation due to the less elastic energy storage. For comparison, we also show the prediction from the conventional FENE-P model. The parameters are chosen to match the dynamics of the orientation tensor (the vector between two free-end particles) near equilibrium. As expected,

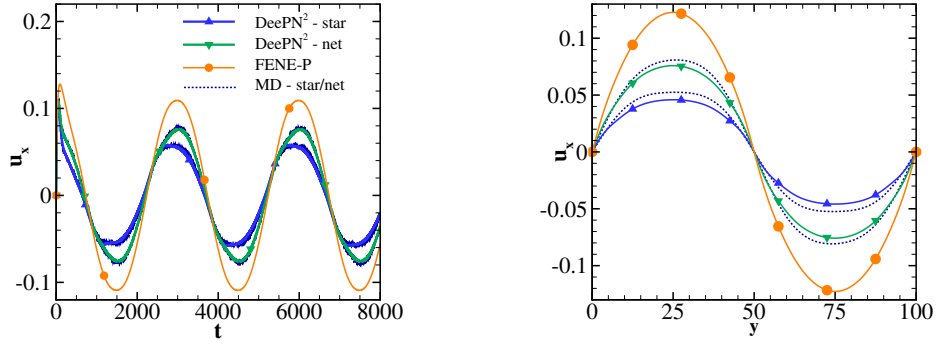


Figure 4.4 The oscillating Womersley flow of the star- and net-shaped molecule suspensions predicted from the micro-scale simulation, DeePN<sup>2</sup> and the FENE-P model. The FENE-P model parameters are chosen to match the dynamics of the orientation tensor (the vector between two free-end particles) near equilibrium. Left: the velocity evolution  $u_x(y, t)$  at  $y = 6$ . Right: the velocity profile  $u_x(y, t)$  at  $t = 6450$ .

the FENE-P model shows limitations for predicting the flow responses of the two suspensions.

The distinct viscoelastic responses of the different suspensions can be further elucidated by examining the elongation flow. We impose the traceless flow gradient  $\nabla \mathbf{u} = \text{diag}(\dot{\epsilon}, -\dot{\epsilon}, 0)$  where the strain rate  $\dot{\epsilon}$  is set to be  $4 \times 10^{-4}$ . Fig. 4.5 shows the stress development of the chain- and star-shaped molecule suspensions. The micro-scale simulations are imposed by the generalized uniaxial extension flow boundary conditions (Nicholson and Rutledge, 2016; Murashima et al., 2018). Compared with the shear flow, the elongation flow yields larger extension and longer processes, as was shown in experimental studies (Smith et al., 1999); the steady state is achieved at about  $t = 2.5 \times 10^3$  and  $t = 10^4$  for the star- and chain-shaped molecule suspensions, respectively. Moreover, the steady stress value  $\tau_{p_{xx}}$  of the chain-shaped molecule suspension is much larger than the value of the star-shaped molecule suspension. Such differences are also due to the larger flexibility of the chain-shaped molecule, which produces a stronger extension under external flow. DeePN<sup>2</sup> successfully captures the different responses and shows good agreement with the micro-scale simulations for both cases.

Finally, we consider the Taylor-Green vortex flow (Taylor, 1934; Thomases and Shelley, 2007) in a  $100 \times 100 \times 160$  domain (in reduced unit) of the micro-scale simulation. The external force

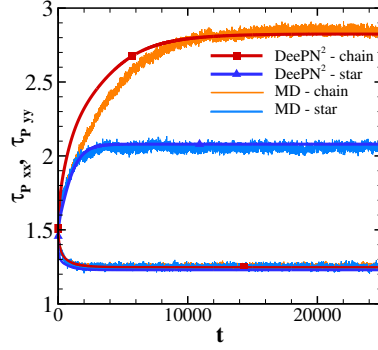


Figure 4.5 The elongation flow of the chain- and star-shaped molecule suspensions predicted from the micro-scale simulation and DeePN<sup>2</sup>. With the same bond potential and strain rate, the chain-shaped molecule suspension yields larger elongation stress. The lines with rough oscillations denote the micro-scale simulation results; the solid lines with symbols denote the DeePN<sup>2</sup> predictions.

$\mathbf{f}_{\text{ext}} = (f_x, f_y, 0)$  is applied to the domain following

$$f_x(x, y) = -2f_0 \sin\left(\frac{2\pi x}{L}\right) \cos\left(\frac{2\pi y}{L}\right), \quad f_y(x, y) = 2f_0 \cos\left(\frac{2\pi x}{L}\right) \sin\left(\frac{2\pi y}{L}\right),$$

where  $L = 100$  and  $f_0 = 6 \times 10^{-3}$ . Periodic boundary conditions are imposed along all of the three directions. The force field imposes an elongation to the flow field along the x-direction and a compression along the y-direction. The flow near the center  $(L/2, L/2)$  resembles the planar elongation flow. Four vortices appear at  $(L/2 \pm L/4, L/2 \pm L/4)$ . Figure. 4.6(a-b) shows the steady-state velocity field. Compared with the star-shaped molecule suspension, the velocity field of the chain-shaped molecule suspension shows larger deviation from the symmetric structure of the Newtonian flow (i.e.,  $\propto [-\sin(2\pi x/L) \cos(2\pi y/L), \cos(2\pi x/L) \sin(2\pi y/L)]$ ) due to the larger polymer stress across the flow regime. Furthermore, the two suspensions yield different velocity magnitude, as shown in Fig. 4.6(c). Fig. 4.6(d) shows the velocity development at  $(75, 49)$ . The velocities of both suspensions achieve a similar maximum value near  $t = 30$  and decay along with the polymer stress development. However, the star-shaped molecule suspension reaches the steady state much earlier with a larger velocity than the chain-shaped molecule suspension.

Fig. 4.7 (a-b) shows the steady-state stress field for the two suspensions. We see that the chain-shaped molecule suspension exhibits larger polymer stress variation along the elongation and contraction directions, reflected in the larger loop area in Fig. 4.7(b). Such difference is also

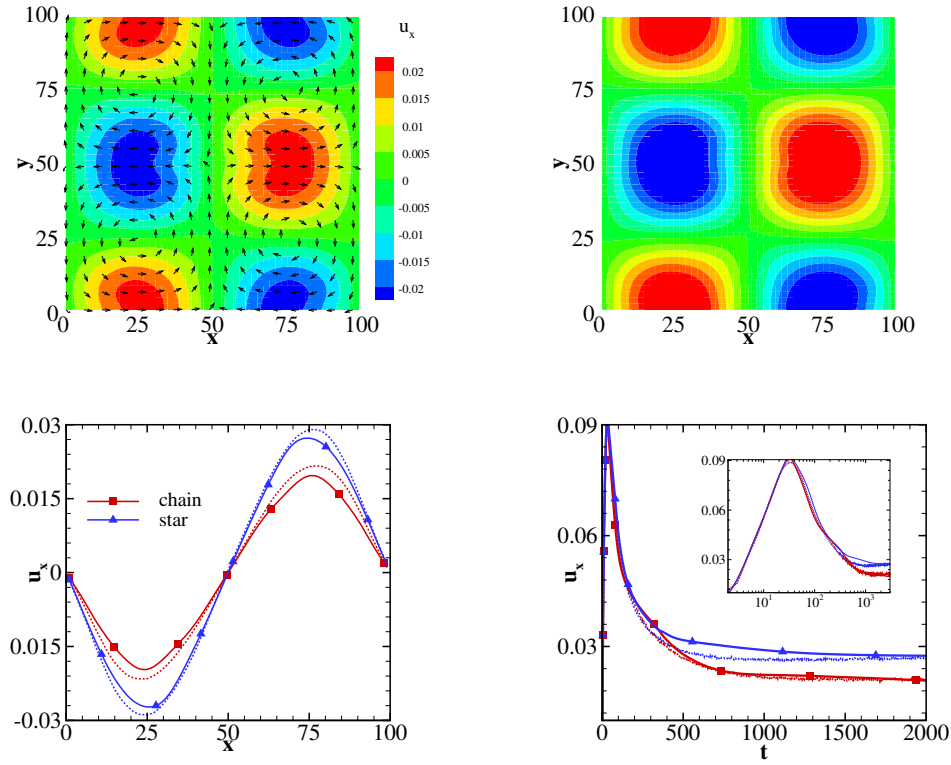


Figure 4.6 The velocity field of the Taylor-Green vortex flow of the chain- and star-shaped molecule suspensions predicted from the micro-scale simulations and DeePN<sup>2</sup>. (a-b) The 2D steady-state velocity field of the chain- and star-shaped molecule suspensions from the micro-scale simulations. The velocity field of the chain-shaped system yields more pronounced deviations from the symmetric Newtonian flow due to the more pronounced polymer stress across the flow regime. (c) The steady-state 1D velocity profile  $u_x(x, y = 49)$ . The solid and dashed lines represent the predictions from the micro-scale simulations and the DeePN<sup>2</sup> model, respectively. (d) The time history of  $u_x(x = 75, y = 49)$ .

consistent with the more pronounced asymmetric velocity field shown in Fig. 4.6(a-b). In addition, we also examine the transient states where the flow undergoes intricate and heterogeneous process. Fig. 4.7(c) shows the stress development at point (49, 35), where  $\tau_{p_{xx}}$  and  $\tau_{p_{yy}}$  cross over during the evolution. During the initial stage,  $\tau_{p_{yy}}$  increases along with the flow development towards to the stagnation point. At  $t > 150$ ,  $\tau_{p_{yy}}$  decreases due to the compression along the y-direction. Meanwhile,  $\tau_{p_{xx}}$  increases and achieves a steady state slightly larger than  $\tau_{p_{yy}}$  for the star-shaped solution. On the other hand, the chain-shaped solution ends up with a significantly larger value of  $\tau_{p_{xx}}$  due to the larger molecule flexibility and further extension along the x-direction. The different viscoelastic responses are also reflected in the stress development at point (49, 49). As shown in

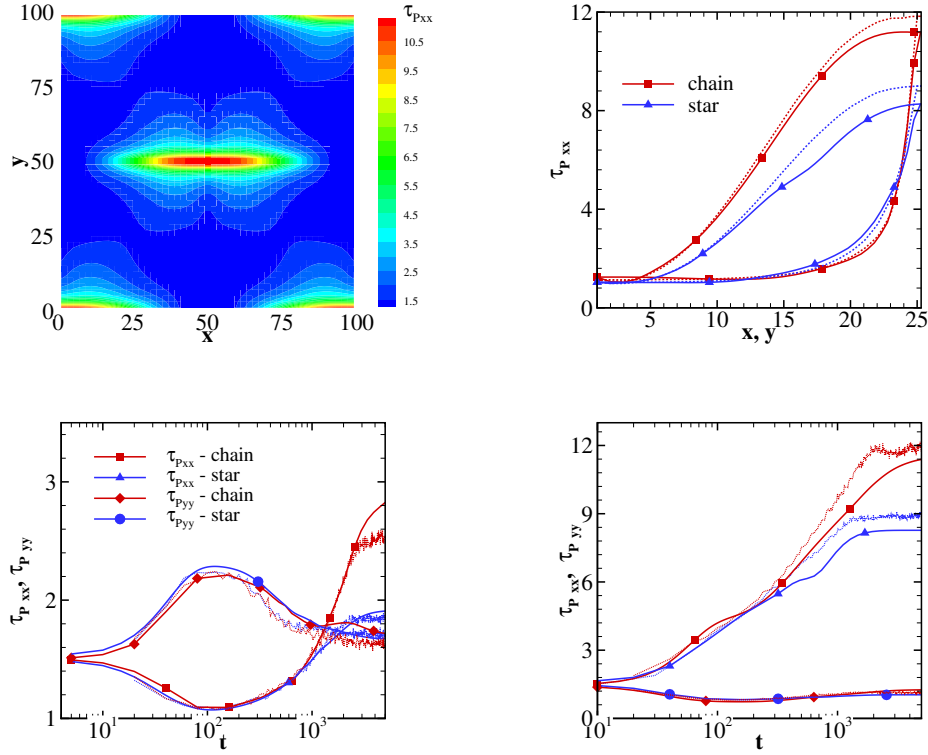


Figure 4.7 The stress field of the Taylor-Green vortex flow of the chain- and star-shaped molecule suspensions predicted from the micro-scale simulations and DeePN<sup>2</sup>. (a) The 2D steady-state stress field of the chain-shaped molecule suspension from the micro-scale simulations. (b) The 1D steady-state stress profiles  $\tau_{p_{xx}}(x, y = 49)$  and  $\tau_{p_{xx}}(x = 49, y)$ . The chain-shaped molecule suspension yields larger stress variations (i.e., the “loop area”) along the flow domain. (c-d) The stress evolution of  $\tau_{p_{xx}}(t)$  and  $\tau_{p_{yy}}(t)$  at the points (49, 35) and (49, 49), respectively. The dashed and the solid lines denote the micro-scale simulations and the DeePN<sup>2</sup> predictions, respectively.

Fig. 4.7(d), the chain-shaped solution exhibits longer evolution of  $\tau_{p_{xx}}$  and larger steady value than the star-shaped solution. DeePN<sup>2</sup> successfully captures such micro-structure-induced rheological differences and shows good agreement with the micro-scale simulation results.

#### 4.3.6 Validation of the rotational-symmetry preserving NN representation

To validate the performance of the proposed DNN representation, we check the accuracy of the modeling terms given a set of conformation tensors  $\mathbf{c}_1, \dots, \mathbf{c}_n$  under different unitary transformations. Fig. 4.8 shows the relative error under each transformation. The DNN representation (4.18) yields the same results under all the transformation. In contrast, the DNN without accounting for the four transformations yields significant error due to the non-uniqueness of the eigenvectors of  $\mathbf{c}_1$ .

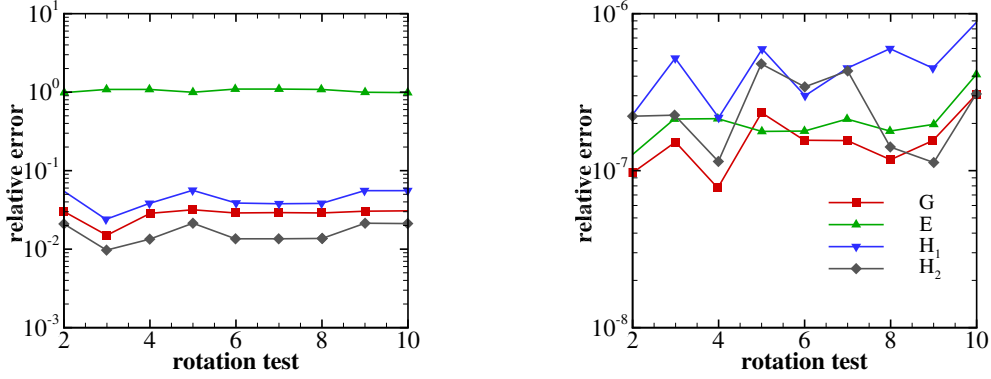


Figure 4.8 The relative  $l_\infty$  error of the model prediction under randomly chosen orthogonal transformations without (left) and with (right) accounting for the four eigen-space transformations in Eq. (4.18).

In addition, we examine the 2D Taylor-Green vortex flow where the evolution of  $\mathbf{c}_1$  becomes degenerate at certain points. Fig. 4.9 shows the stress evolution at (45, 37). At  $t = 1080$ , the eigenvalues  $\lambda_2$  and  $\lambda_3$  cross over. Concurrently, the prediction of the polymer stress  $\tau_p$  from the model without considering the swap of  $\mathbf{u}_2$  and  $\mathbf{u}_3$  shows apparent deviations near the regime as shown in Fig. 4.9. In contrast, the prediction from the model retaining the eigenvalue permutation trained by Eq. (4.20) shows good agreement with the MD results.

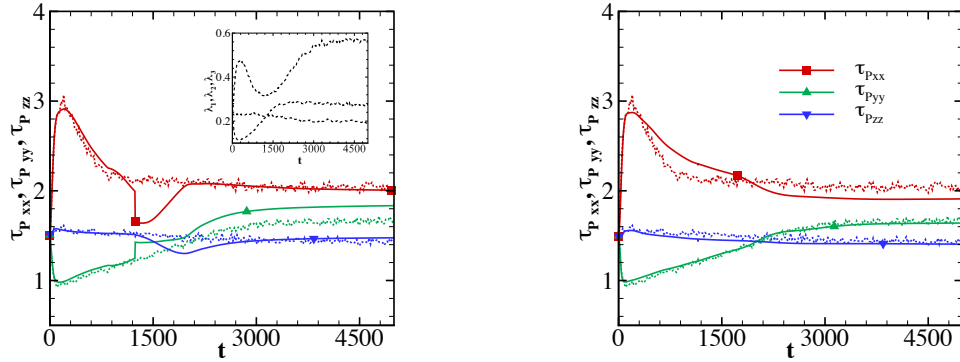


Figure 4.9 Stress evolution of the Taylor-Green vortex flow at position (45, 37) of the chain-shaped molecule suspension. Left: prediction without considering the swap of eigenvectors when the two eigenvalues approaches near  $t = 1255$  as shown in the inset plot. Right: predictions from the model retaining the eigenvalue permutation trained by Eq. (4.20). The dashed and the solid lines denote the micro-scale simulations and the DeePN<sup>2</sup> predictions, respectively.

#### 4.4 Discussion

We have developed a general machine-learning based model, DeePN<sup>2</sup>, for describing the non-Newtonian hydrodynamics for polymer solutions with arbitrary molecular structure and interaction. The constructed model retains a clear physical interpretation and faithfully encodes the micro-scale structural information into the macro-scale hydrodynamics, where conventional models based on empirical closures generally show limitations. In particular, for the chain- and star-shaped molecule suspensions with the same bead number and bond interaction, DeePN<sup>2</sup> successfully captures the different viscoelastic responses arising from the different molecular structural symmetry (i.e., the effective rigidity) in the configuration space without additional human intervention. Unlike the direct evaluation or moment-closure representations of the configurational PDF, the present DeePN<sup>2</sup> model directly learns a set of micro-to-macro mappings to probe the optimal approximations of the constitutive dynamics in terms of the macro-scale features, and thereby circumventing the numerical challenges due to the high-dimensionality of the polymer configuration space. This multi-scaled nature enables us to learn the constitutive dynamics of the macro-scale features directly from the kinetic equations of their micro-scale counterparts using only discrete rather than the time-derivative samples commonly used in the machine learning-based models of complex dynamic problems.

One thing we have not investigated systematically is the generation of training samples. For DeePN<sup>2</sup> to be truly reliable, the training samples should be representative enough for all the practical situations that one might encounter. However, due to the cost associated with generating such training samples, we would also like the training set to be as small as possible. This calls for an adaptive procedure for generating the training sample, such as the concurrent learning procedure discussed in (E et al., 2021). The present DeePN<sup>2</sup> models are trained with samples collected from homogeneous shear flow. Even though the numerical predictions show good agreement with micro-scale simulations for a variety of flows, one should not expect this to be generally the case. Further work on sampling is needed to make sure that one can produce truly reliable DeePN<sup>2</sup> models. Furthermore, instead of the general form (4.6), a specific design of the encoders  $\mathbf{b}(\cdot)$  accounting for the molecule symmetry and rigidity may facilitate the extraction of the macro-scale features  $\mathbf{c}$ . In



addition, more accurate micro-scale kinetic models accounting for the heterogeneous hydrodynamic interactions and non-Markovianity (Lei et al., 2016; Lei and Li, 2021) can be used to construct the macro-scale constitutive dynamics. Finally, the adaptive choice of the number of features and the enhanced sampling of the discrete micro-scale configurations may further improve the performance of the DeePN<sup>2</sup> model. We leave these issues for future work.

## CHAPTER 5

### CONCLUSION AND OUTLOOK

This dissertation presents advancements in multi-scale modeling by utilizing machine learning to construct reliable and structure-preserved reduced models with interpretable micro-scale to macro and meso-scale mapping. We explored two main directions in reduced modeling: (1) micro-to-meso data-driven stochastic reduced model, and (2) micro-to-macro deep-learning-based non-Newtonian hydrodynamic model.

The first part on micro-to-meso modeling developed a data-driven method to learn stochastic reduced models of complex systems that retain a state-dependent memory beyond the standard GLE with a homogeneous kernel. The main idea is to seek a generalized representation of the state-dependent memory which satisfies the second fluctuation-dissipation theorem with a coherent colored noise. To parameterize the representation, traditional two-point correlation functions are not enough for extracting state-dependent information where we use the state-dependent three-point correlation functions instead. Efficient training is achieved by constructing the state-dependent encoders using a set of sparse bases, whose correlations can be efficiently precomputed. The convolution part in simulation can be efficiently evaluated using the fast convolution algorithm. Numerical results demonstrate the limitation of the standard GLE and the essential role of the broadly overlooked state-dependency nature in predicting molecule kinetics related to conformation relaxation and transition.

The second part on micro-to-macro modeling developed a deep learning-based non-Newtonian hydrodynamic model, which focused on learning accurate non-Newtonian hydrodynamic models from micro-scale polymer kinetics. The model aims to close the empirical constitutive hydrodynamics form with micro-scale level fidelity. First, we establish a micro-macro correspondence via a set of encoders for the micro-scale polymer configurations and their macro-scale counterparts, a set of conformation tensors. Instead of directly approximating the momenta of distribution, we use neural networks to construct the nonlinear encoders. Then the dynamics (including generalized objective tensor derivative) of conformation tensors can be directly derived from the micro-scale

model by the Fokker-Planck equation, where the molecular structural mechanics are automatically included. Here the dynamics are only related to instantaneous samples instead of time series. Finally, we use the symmetry-preserving neural networks to parameterize the dynamics based on the conformation tensors. By learning the model only based on reverse Poiseuille flow, we demonstrate its accuracy and generalization ability by considering various flows in comparison with the results of the micro-scale simulations.

Despite these advancements, several challenges remain open for future research.

### 5.1 High Dimensional Stochastic Reduced Model

Although our work (Ge et al., 2024) shows high promising results, it only works for 1D reduced systems due to the curse of dimensionality. The main challenge arises from the state-dependent features, which are directly related to the dimension. Traditional methods are limited for such situations. However, neural networks provide a powerful tool to construct high-dimensional functions efficiently. By carefully designing the training process and the formulation of the state-dependent features, we can develop a practical high-dimensional stochastic reduced model. Furthermore, the high-dimensional stochastic reduced model can also help build a more accurate DeePN<sup>2</sup> model with state-dependent memory accounted.

### 5.2 Variational-informed Structure-preserving Macro-scale Reduced Model

Structure-preserving is important in multi-scale modeling, especially for constructing stable and reliable reduced models. Some existing methods such as deep learning-based coarse-grained molecular dynamics Zhang et al. (2018b) follow some physical constraints. One way to impose energy stability is by pre-building the constitutive dynamics with a generalized extendable energy functional structure. One possible formulation is the GENERIC formalism Grmela and Öttinger (1997) governed by the coupling of a reversible and an irreversible process:

$$\frac{d\mathcal{X}}{dt} = \mathcal{L} \frac{\delta E}{\delta \mathcal{X}} + \mathcal{M} \frac{\delta S}{\delta \mathcal{X}} \quad (5.1)$$

$$\mathcal{L} \frac{\delta S}{\delta \mathcal{X}} = \mathcal{M} \frac{\delta E}{\delta \mathcal{X}} \equiv \mathbf{0} \quad (5.2)$$

where  $\mathcal{X} : \Omega \rightarrow \mathbb{R}^d$  is the field variables,  $E[\mathcal{X}] = \int_{\Omega} \hat{E}(\mathcal{X}) d\Omega$  is the energy,  $S[\mathcal{X}] = \int_{\Omega} \hat{S}(\mathcal{X}) d\Omega$  is the entropy.  $L$  is the Poisson matrix satisfied  $\mathcal{L} = -\mathcal{L}^T$ ,  $\mathcal{M}$  is the friction matrix satisfied  $\mathcal{M} \succ 0$ . The degeneracy condition Eq. (5.2) ensures the energy conservation  $dE/dt \equiv 0$  and the entropy production  $dS/dt \geq 0$  and therefore the free energy stability.

Our main idea is to seek a generalized extendable energy variational form of our DeePN<sup>2</sup> model where  $\mathcal{X} = (\rho, \mathbf{u}, \mathbf{c}_1, \dots, \mathbf{c}_n)$ .  $\rho$  is the density,  $\mathbf{u}$  is the velocity,  $\{\mathbf{c}_i\}_{i=1}^n$  is our conformation tensor. By doing so, we can use existing numerical schemes such as the scalar auxiliary variable approach Shen et al. (2018) and invariant energy quadratization method Yang (2016) to ensure energy stability.

## BIBLIOGRAPHY

- Abrams, J. B. and Tuckerman, M. E. (2008). Efficient and direct generation of multidimensional free energy surfaces via adiabatic dynamics without coordinate transformations. The Journal of Physical Chemistry B, 112(49):15742–15757.
- Allen, E. C. and Rutledge, G. C. (2008). A novel algorithm for creating coarse-grained, density dependent implicit solvent models. The Journal of Chemical Physics, 128(15):154115.
- Allen, E. C. and Rutledge, G. C. (2009). Evaluating the transferability of coarse-grained, density-dependent implicit solvent models to mixtures and chains. The Journal of Chemical Physics, 130(3):034904.
- Ayaz, C., Scalfi, L., Dalton, B. A., and Netz, R. R. (2022a). Generalized langevin equation with a non-linear potential of mean force and non-linear memory friction from a hybrid projection scheme. Physical Review E, 105:054138.
- Ayaz, C., Tepper, L., Brünig, F. N., Kappler, J., Daldrop, J. O., and Netz, R. R. (2021). Non-markovian modeling of protein folding. Proceedings of the National Academy of Sciences, 118(31):e2023856118.
- Ayaz, C., Tepper, L., and Netz, R. R. (2022b). Markovian embedding of generalized langevin equations with a nonlinear friction kernel and configuration-dependent mass. Turkish Journal of Physics, 46(6):194 – 205.
- Baczewski, A. D. and Bond, S. D. (2013). Numerical integration of the extended variable generalized Langevin equation with a positive Prony representable memory kernel. The Journal of chemical physics, 139(4):044107.
- Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. The Journal of Physical Chemistry, 97(40):10269–10280.
- Berezhkovskii, A. and Szabo, A. (2011). Time scale separation leads to position-dependent diffusion along a slow coordinate. The Journal of Chemical Physics, 135(7):074108.
- Berkowitz, M., Morgan, J., and McCammon, J. A. (1983). Generalized Langevin dynamics simulations with arbitrary time-dependent memory kernels. J. Chem. Phys., 78:3256.
- Berne, B. J., Weeks, J. D., and Zhou, R. (2009). Dewetting and hydrophobic interaction in physical and biological systems. Annual Review of Physical Chemistry, 60(1):85–103.
- Berressem, F., Scherer, C., Andrienko, D., and Nikoubashman, A. (2021). Ultra-coarse-graining of homopolymers in inhomogeneous systems. Journal of Physics: Condensed Matter, 33(25):254002.

- Best, R. B. and Hummer, G. (2006). Diffusive model of protein folding dynamics with kramers turnover in rate. Phys. Rev. Lett., 96:228104.
- Best, R. B. and Hummer, G. (2010). Coordinate-dependent diffusion in protein folding. Proceedings of the National Academy of Sciences, 107(3):1088–1093.
- Bird, R. B., Curtiss, C. F., Armstrong, R. C., and Hassager, O. (1987). Dynamics of Polymeric Liquids, Volume 2: Kinetic Theory, 2nd Edition. Wiley, 2nd edition.
- Bird, R. B., Dotson, P. J., and Johnson, N. (1980). Polymer solution rheology based on a finitely extensible bead—spring chain model. Journal of Non-Newtonian Fluid Mechanics, 7(2):213 – 235.
- Buff, F. P., Lovett, R. A., and Stillinger, F. H. (1965). Interfacial density profile for fluids in the critical region. Phys. Rev. Lett., 15:621–623.
- Carmeli, B. and Nitzan, A. (1983). Theory of activated rate processes: Position dependent friction. Chemical Physics Letters, 102(6):517–522.
- Cerioti, M., Bussi, G., and Parrinello, M. (2009). Langevin equation with colored noise for constant-temperature molecular dynamics simulations. Physical review letters, 102(2):020601.
- Chan, H., Cherukara, M. J., Narayanan, B., Loeffler, T. D., Benmore, C., Gray, S. K., and Sankaranarayanan, S. K. R. S. (2019). Machine learning coarse grained models for water. Nature Communications, 10(1):379.
- Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. Nature, 437(7059):640–647.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. Mathematics of Computation, 19(90):297–301.
- Cossio, P., Hummer, G., and Szabo, A. (2015). On artifacts in single-molecule force spectroscopy. Proceedings of the National Academy of Sciences, 112(46):14248–14253.
- Daldrop, J. O., Kowalik, B. G., and Netz, R. R. (2017). External potential modifies friction of molecular solutes in water. Physical Review X, 7(4):041065.
- Dalton, B. A., Ayaz, C., Kiefer, H., Klimek, A., Tepper, L., and Netz, R. R. (2023). Fast protein folding is governed by memory-dependent friction. Proceedings of the National Academy of Sciences, 120(31):e2220068120.
- Darve, E. and Pohorille, A. (2001). Calculating free energies using average force. The Journal of Chemical Physics, 115(20):9169–9183.

- Darve, E., Solomon, J., and Kia, A. (2009). Computing generalized Langevin equations and generalized fokker-planck equations. Proc. Natl. Acad. Sci., 106(27):10884–10889.
- Das, A. and Andersen, H. C. (2010). The multiscale coarse-graining method. v. isothermal-isobaric ensemble. The Journal of Chemical Physics, 132(16):164106.
- Das, A. and Andersen, H. C. (2012). The multiscale coarse-graining method. ix. a general method for construction of three body coarse-grained force fields. The Journal of chemical physics, 136(19):194114.
- Davtyan, A., Dama, J. F., Voth, G. A., and Andersen, H. C. (2015). Dynamic force matching: A method for constructing dynamical coarse-grained models with realistic time dependence. J. Chem. Phys., 142.
- DeLyser, M. and Noid, W. G. (2020). Bottom-up coarse-grained models for external fields and interfaces. The Journal of Chemical Physics, 153(22):24103.
- DeLyser, M. R. and Noid, W. G. (2017). Extending pressure-matching to inhomogeneous systems via local-density potentials. The Journal of chemical physics, 147(13):134111.
- DeLyser, M. R. and Noid, W. G. (2019). Analysis of local density potentials. The Journal of Chemical Physics, 151(22):224106.
- DeLyser, M. R. and Noid, W. G. (2022). Coarse-grained models for local density gradients. The Journal of Chemical Physics, 156(3):034106.
- Deutch, J. M. and Oppenheim, I. (1971). Molecular theory of brownian motion for several particles. The Journal of Chemical Physics, 54(8):3547–3555.
- Dinpajoo, M. and Guenza, M. G. (2017). On the density dependence of the integral equation coarse-graining effective potential. The Journal of Physical Chemistry B, 122(13):3426–3440.
- Doyle, P. S., Shaqfeh, E. S., McKinley, G. H., and Spiegelberg, S. H. (1998). Relaxation of dilute polymer solutions following extensional flow. Journal of Non-Newtonian Fluid Mechanics, 76(1):79–110.
- Dunn, N. J. H. and Noid, W. G. (2015). Bottom-up coarse-grained models that accurately describe the structure, pressure, and compressibility of molecular liquids. The Journal of Chemical Physics, 143(24):243148.
- Dunn, N. J. H. and Noid, W. G. (2016). Bottom-up coarse-grained models with predictive accuracy and transferability for both structural and thermodynamic properties of heptane-toluene mixtures. The Journal of Chemical Physics, 144(20):204124.
- E, W., Han, J., and Zhang, L. (2021). Machine-learning-assisted modeling. Physics Today,

74(7):36–41.

- E, W., Lei, H., Xie, P., and Zhang, L. (2023). Machine learning-assisted multi-scale modeling. Journal of Mathematical Physics, 64(7):071101.
- E, W. and Vanden-Eijnden, E. (2010). Transition-path theory and path-finding algorithms for the study of rare events. Annual Review of Physical Chemistry, 61(1):391–420.
- Empereur-mot, C., Capelli, R., Perrone, M., Caruso, C., Doni, G., and Pavan, G. M. (2022). Automatic multi-objective optimization of coarse-grained lipid force fields using swarmcg. The Journal of Chemical Physics, 156(2):024801.
- Español, P. and Warren, P. (1995). Statistical mechanics of dissipative particle dynamics. Europhysics Letters, 30(4):191–196.
- Evans, R. (1979). The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. Advances in Physics, 28(2):143–200.
- Fang, L., Ge, P., Zhang, L., E, W., and Lei, H. (2022). DeePN<sup>2</sup>: A deep learning-based non-newtonian hydrodynamic model. Journal of Machine Learning, 1(1):114–140.
- Feng, J., Chaubal, C. V., and Leal, L. G. (1998). Closure approximations for the doi theory: Which to use in simulating complex flows of liquid-crystalline polymers? Journal of Rheology, 42(5):1095–1119.
- Forest, G. M., Zhou, R., and Wang, Q. (2003). Full-tensor alignment criteria for sheared nematic polymers. Journal of Rheology, 47(1):105–127.
- Galvelis, R. and Sugita, Y. (2017). Neural network and nearest neighbor algorithms for enhancing sampling of molecular dynamics. Journal of Chemical Theory and Computation, 13(6):2489–2500.
- Ge, P., Zhang, L., and Lei, H. (2023). Machine learning assisted coarse-grained molecular dynamics modeling of meso-scale interfacial fluids. The Journal of Chemical Physics, 158(6). 064104.
- Ge, P., Zhang, Z., and Lei, H. (2024). Data-driven learning of the generalized langevin equation with state-dependent memory. Phys. Rev. Lett., 133:077301.
- Giesekus, H. (1982). A simple constitutive equation for polymer fluids based on the concept of deformation-dependent tensorial mobility. Journal of Non-Newtonian Fluid Mechanics, 11(1):69 – 109.
- Glatzel, F. and Schilling, T. (2022). The interplay between memory and potentials of mean force: A discussion on the structure of equations of motion for coarse-grained observables. Europhysics Letters, 136(3):36001.



- Grmela, M. and Öttinger, H. C. (1997). Dynamics and thermodynamics of complex fluids. i. development of a general formalism. Physical Review E, 56(6):6620.
- Grogan, F., Lei, H., Li, X., and Baker, N. A. (2020). Data-driven molecular modeling with the generalized Langevin equation. J. Comput. Phys., 418:109633–109641.
- Groot, R. D. and Warren, P. B. (1997). Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. Journal of Chemical Physics, 107(11):4423–4435.
- Grosso, M., Maffettone, P., Halin, P., Keunings, R., and Legat, V. (2000). Flow of nematic polymers in eccentric cylinder geometry: influence of closure approximations. Journal of Non-Newtonian Fluid Mechanics, 94(2):119–134.
- Han, J., Ma, C., Ma, Z., and E, W. (2019a). Uniformly accurate machine learning-based hydrodynamic models for kinetic equations. Proceedings of the National Academy of Sciences, 116(44):21983–21991.
- Han, J., Ma, C., Ma, Z., and E, W. (2019b). Uniformly accurate machine learning-based hydrodynamic models for kinetic equations. Proceedings of the National Academy of Sciences, 116(44):21983–21991.
- Hänggi, P. (1997). Generalized langevin equations: A useful tool for the perplexed modeller of nonequilibrium fluctuations? In Schimansky-Geier, L. and Pöschel, T., editors, Stochastic Dynamics, pages 15–22, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Haynes, G. R., Voth, G. A., and Pollak, E. (1993). A theory for the thermally activated rate constant in systems with spatially dependent friction. Chemical Physics Letters, 207(4):309–316.
- Haynes, G. R., Voth, G. A., and Pollak, E. (1994). A theory for the activated barrier crossing rate constant in systems influenced by space and time dependent friction. The Journal of Chemical Physics, 101(9):7811–7822.
- Hijón, C., Español, P., Vanden-Eijnden, E., and Delgado-Buscalioni, R. (2010). Mori–Zwanzig formalism as a practical computational tool. Faraday discuss., 144:301–322.
- Hinczewski, M., von Hansen, Y., Dzubiella, J., and Netz, R. R. (2010). How the diffusivity profile reduces the arbitrariness of protein folding free energies. The Journal of Chemical Physics, 132(24):245103.
- Hoogerbrugge, P. J. and Koelman, J. M. V. A. (1992). Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. Europhys. Lett., 19(3):155–160.
- Hoover, W. G. (1985). Canonical dynamics: equilibrium phase-space distributions. Physical Review A, 31(3):1695–1697.

- Huang, J., Ma, Z., Zhou, Y., and Yong, W.-A. (2021). Learning thermodynamically stable and galilean invariant partial differential equations for non-equilibrium flows. Journal of Non-Equilibrium Thermodynamics, 46(4):355–370.
- Hulsen, M., van Heel, A., and van den Brule, B. (1997). Simulation of viscoelastic flows using brownian configuration fields. Journal of Non-Newtonian Fluid Mechanics, 70(1):79 – 101.
- Hummer, G., Garde, S., García, A. E., Paulaitis, M. E., and Pratt, L. R. (1998). Hydrophobic effects on a molecular scale. The Journal of Physical Chemistry B, 102(51):10469–10482.
- Hummer, G., Garde, S., García, A. E., Pohorille, A., and Pratt, L. R. (1996). An information theory model of hydrophobic interactions. Proceedings of the National Academy of Sciences, 93(17):8951–8955.
- Hyon, Y., Du, Q., and Liu, C. (2008). An enhanced macroscopic closure approximation to the micro-macro fene model for polymeric materials. Multiscale Modeling & Simulation, 7(2):978–1002.
- Izvekov, S., Chung, P. W., and Rice, B. M. (2010). The multiscale coarse-graining method: Assessing its accuracy and introducing density dependent coarse-grain potentials. The Journal of Chemical Physics, 133(6):064109.
- Izvekov, S. and Voth, G. A. (2005). A multiscale coarse-graining method for biomolecular systems. The Journal of Physical Chemistry B, 109(7):2469–2473.
- Jin, J. and Voth, G. A. (2018). Ultra-coarse-grained models allow for an accurate and transferable treatment of interfacial systems. Journal of Chemical Theory and Computation, 14(4):2180–2197.
- John, S. T. and Csányi, G. (2017). Many-body coarse-grained interactions using gaussian approximation potentials. The Journal of Physical Chemistry B, 121(48):10934–10949.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics, 79(2):926–935.
- Jung, B. and Jung, G. (2023). Dynamic coarse-graining of linear and non-linear systems: Mori–Zwanzig formalism and beyond. The Journal of Chemical Physics, 159(8):084110.
- Jung, G., Hanke, M., and Schmid, F. (2017). Iterative reconstruction of memory kernels. Journal of Chemical Theory and Computation, 13(6):2481–2488.
- Kingma, D. and Ba, J. (2015). Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR).
- Klippenstein, V., Tripathy, M., Jung, G., Schmid, F., and van der Vegt, N. F. (2021). Introducing memory in coarse-grained molecular simulations. The Journal of Physical Chemistry B,

125(19):4931–4954.

- Klippenstein, V. and van der Vegt, N. F. A. (2021). Cross-correlation corrected friction in (generalized) langevin models. The Journal of Chemical Physics, 154(19):191102.
- Koopman, B. O. (1931). Hamiltonian systems and transformation in Hilbert space. Proceedings of the National Academy of Sciences, 17(5):315–318.
- Kramers, H. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. Physica, 7(4):284–304.
- Krishnan, R., Singh, S., and Robinson, G. W. (1992). Space-dependent friction in the theory of activated rate processes: The hamiltonian approach. The Journal of Chemical Physics, 97(8):5516–5521.
- Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992a). The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. Journal of Computational Chemistry, 13(8):1011–1021.
- Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992b). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. Journal of Computational Chemistry, 13(8):1011–1021.
- Laio, A. and Parrinello, M. (2002). Escaping free-energy minima. Proceedings of the National Academy of Sciences, 99(20):12562–12566.
- Landau, L. D. and Lifshitz, E. M. (1987). Fluid Mechanics, Second Edition: Volume 6 (Course of Theoretical Physics). Course of theoretical physics / by L. D. Landau and E. M. Lifshitz, Vol. 6. Butterworth-Heinemann, 2 edition.
- Lange, O. F. and Grubmüller, H. (2006). Collective Langevin dynamics of conformational motions in proteins. J. Chem. Phys., 124:214903.
- Larini, L., Lu, L., and Voth, G. A. (2010). The multiscale coarse-graining method. vi. implementation of three-body coarse-grained potentials. The Journal of Chemical Physics, 132(16):164107.
- Larson, R. G. (1988). Constitutive Equations for Polymer Melts and Solutions. Butterworth-Heinemann Press.
- Laso, M. and Öttinger, H. C. (1993). Calculation of viscoelastic flow using molecular models: the connffessit approach. Journal of Non-Newtonian Fluid Mechanics, 47:1 – 20.
- Lee, H. S., Ahn, S.-H., and Darve, E. F. (2019). The multi-dimensional generalized Langevin equation for conformational motion of proteins. The Journal of Chemical Physics, 150(17).

- Lei, H., Baker, N. A., and Li, X. (2016). Data-driven parameterization of the generalized Langevin equation. Proc. Natl. Acad. Sci., 113(50):14183–14188.
- Lei, H., Caswell, B., and Karniadakis, G. E. (2010). Direct construction of mesoscopic models from microscopic simulations. Phys. Rev. E, 81:026704.
- Lei, H. and Li, X. (2021). Petrov–Galerkin methods for the construction of non-Markovian dynamics preserving nonlocal statistics. The Journal of Chemical Physics, 154(18):184108.
- Lei, H., Mundy, C. J., Schenter, G. K., and Voulgarakis, N. K. (2015). Modeling nanoscale hydrodynamics by smoothed dissipative particle dynamics. J. Chem. Phys., 142(19):194504.
- Lei, H., Wu, L., and E, W. (2020). Machine learning based non-newtonian fluid model with molecular fidelity. Physics Review E, 102:043309.
- Lemke, T. and Peter, C. (2017). Neural network based prediction of conformational free energies - a new route toward coarse-grained simulation models. Journal of Chemical Theory and Computation, 13(12):6213–6221.
- Li, Z., Kovachki, N. B., Azizzadenesheli, K., liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A. (2021). Fourier neural operator for parametric partial differential equations. In International Conference on Learning Representations.
- Lielens, G., Halin, P., Jaumain, I., Keunings, R., and Legat, V. (1998). New closure approximations for the kinetic theory of finitely extensible dumbbells. Journal of Non-Newtonian Fluid Mechanics, 76(1):249–279.
- Lielens, G., Keunings, R., and Legat, V. (1999). The fene-l and fene-ls closure approximations to the kinetic theory of finitely extensible dumbbells. Journal of Non-Newtonian Fluid Mechanics, 87(2):179 – 196.
- Lin, F.-H., Liu, C., and Zhang, P. (2005). On hydrodynamics of viscoelastic fluids. Communications on Pure and Applied Mathematics, 58(11):1437–1471.
- Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E. (2021). Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. Nature machine intelligence, 3(3):218–229.
- Lum, K., Chandler, D., and Weeks, J. D. (1999). Hydrophobicity at small and large length scales. The Journal of Physical Chemistry B, 103(22):4570–4577.
- Luo, G., Andricioaei, I., Xie, X. S., and Karplus, M. (2006). Dynamic distance disorder in proteins is caused by trapping. The Journal of Physical Chemistry B, 110(19):9363–9367.
- Lyu, L. and Lei, H. (2023). Construction of coarse-grained molecular dynamics with many-body

- non-markovian memory. Phys. Rev. Lett., 131:177301.
- Lyubartsev, A. P. and Laaksonen, A. (1995). Calculation of effective interaction potentials from radial distribution functions: A reverse monte carlo approach. Phys. Rev. E, 52:3730–3737.
- Ma, L., Li, X., and Liu, C. (2019). Coarse-graining langevin dynamics using reduced-order techniques. Journal of Computational Physics, 380:170–190.
- Maragliano, L. and Vanden-Eijnden, E. (2006). A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. Chemical Physics Letters, 426(1):168–175.
- Maragliano, L. and Vanden-Eijnden, E. (2008). Single-sweep methods for free energy calculations. The Journal of Chemical Physics, 128(18):184110.
- Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. The Journal of Chemical Physics, 101(5):4177–4189.
- Miller, T. F., Vanden-Eijnden, E., and Chandler, D. (2007). Solvent coarse-graining and the string method applied to the hydrophobic collapse of a hydrated chain. Proceedings of the National Academy of Sciences, 104(37):14559–14564.
- Miyamoto, S. and Kollman Peter, A. (2004). Settle: An analytical version of the shake and rattle algorithm for rigid water models. Journal of Computational Chemistry, 13(8):952–962.
- Molinero, V. and Moore, E. B. (2009). Water modeled as an intermediate element between carbon and silicon. The Journal of Physical Chemistry B, 113(13):4008–4016.
- Mones, L., Bernstein, N., and Csányi, G. (2016). Exploration, sampling, and reconstruction of free energy surfaces with gaussian process regression. Journal of Chemical Theory and Computation, 12(10):5100–5110.
- Moore, J. D., Barnes, B. C., Izvekov, S., Lísal, M., Sellers, M. S., Taylor, D. E., and Brennan, J. K. (2016). A coarse-grain force field for rdx: Density dependent and energy conserving. The Journal of Chemical Physics, 144(10):104501.
- Mori, H. (1965). Transport, collective motion, and Brownian motion. Progress of Theoretical Physics, 33(3):423–455.
- Morris, J. P., Fox, P. J., and Zhu, Y. (1997). Modeling low reynolds number incompressible flows using sph. J. Comput. Phys., 136(1):214–226.
- Morrone, J. A., Li, J., and Berne, B. J. (2012). Interplay between hydrodynamics and the free energy surface in the assembly of nanoscale hydrophobes. The Journal of Physical Chemistry B, 116(1):378–389.

- Murashima, T., Hagita, K., and Kawakatsu, T. (2018). Elongational viscosity of weakly entangled polymer melt via coarse-grained molecular dynamics simulation. Nihon Reoroji Gakkaishi, 46(5):207–220.
- Nicholson, D. A. and Rutledge, G. C. (2016). Molecular simulation of flow-enhanced nucleation in n-eicosane melts under steady shear and uniaxial extension. The Journal of Chemical Physics, 145(24):244903.
- Nielsen, S. O., Lopez, C. F., Srinivas, G., and Klein, M. L. (2004). Coarse grain models and the computer simulation of soft materials. Journal of Physics: Condensed Matter, 16(15):R481–R512.
- Noé, F., Tkatchenko, A., Müller, K.-R., and Clementi, C. (2020). Machine learning for molecular simulation. Annual Review of Physical Chemistry, 71(1):361–390.
- Noid, W. G. (2013). Perspective: coarse-grained models for biomolecular systems. J. Chem. Phys., 139(9):090901.
- Noid, W. G., Chu, J.-W., Ayton, G. S., Krishna, V., Izvekov, S., Voth, G. A., Das, A., and Andersen, H. C. (2008). The multiscale coarse-graining method. I. a rigorous bridge between atomistic and coarse-grained models. J. Chem. Phys., 128(24):244114.
- Nosé, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. Molecular Physics, 52(2):255–268.
- Ogorodnikov, V. A. and Prigarin, S. M. (1996). Numerical Modelling of Random Processes and Fields: Algorithms and Applications. De Gruyter, Berlin, Boston.
- Oldroyd, J. G. and Wilson, A. H. (1950). On the formulation of rheological equations of state. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 200(1063):523–541.
- Owens, R. G. and Phillips, T. N. (2002). Computational Rheology. Imperial College Press.
- Pagonabarraga, I. and Frenkel, D. (2001). Dissipative particle dynamics for interacting systems. The Journal of Chemical Physics, 115(11):5015–5026.
- Patel, A. J., Varilly, P., and Chandler, D. (2010). Fluctuations of water near extended hydrophobic and hydrophilic surfaces. The Journal of Physical Chemistry B, 114(4):1632–1637.
- Peterlin, A. (1966). Hydrodynamics of macromolecules in a velocity field with longitudinal gradient. Journal of Polymer Science Part B: Polymer Letters, 4(4):287–291.
- Plotkin, S. S. and Wolynes, P. G. (1998). Non-Markovian configurational diffusion and reaction coordinates for protein folding. Phys. Rev. Lett., 80:5015–5018.

- Posch, H. A., Balucani, U., and Vallauri, R. (1984). On the relative dynamics of pairs of atoms in simple liquids. Physica A, 123:516–534.
- Qin, T., Wu, K., and Xiu, D. (2019). Data driven governing equations approximation using deep neural networks. Journal of Computational Physics, 395:620–635.
- Raissi, M., Perdikaris, P., and Karniadakis, G. (2019a). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378:686–707.
- Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2019b). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378:686–707.
- Rein ten Wolde, P., Sun, S. X., and Chandler, D. (2001). Model of a fluid at small and large length scales and the hydrophobic effect. Phys. Rev. E, 65:011201.
- Reith, D., Pütz, M., and Müller-Plathe, F. (2003). Deriving effective mesoscale potentials from atomistic simulations. Journal of Computational Chemistry, 24(13):1624–1636.
- Ren, W. and E, W. (2005). Heterogeneous multiscale method for the modeling of complex fluids and micro-fluidics. Journal of Computational Physics, 204(1):1 – 26.
- Rosso, L., Mináry, P., Zhu, Z., and Tuckerman, M. E. (2002). On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. The Journal of Chemical Physics, 116(11):4389–4402.
- Rouse, P. E. (1953). A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. The Journal of Chemical Physics, 21(7):1272–1280.
- Rowlinson, J. J. S. and Widom, B. (2002). Molecular theory of capillarity, volume 8. Courier Dover Publications.
- Rudd, R. E. and Broughton, J. Q. (1998). Coarse-grained molecular dynamics and the atomic limit of finite elements. Phys. Rev. B, 58:R5893–R5896.
- Rudy, S. H., Brunton, S. L., Proctor, J. L., and Kutz, J. N. (2017). Data-driven discovery of partial differential equations. Science Advances, 3(4).
- Russo, A., Durán-Olivencia, M. A., Kevrekidis, I. G., and Kalliadasis, S. (2019). Deep learning as closure for irreversible processes: A data-driven generalized Langevin equation. arXiv preprint arXiv:1903.09562.
- Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. Journal of

Computational Physics, 23(3):327–341.

Sanyal, T. and Shell, M. S. (2016). Coarse-grained models using local-density potentials optimized with the relative entropy: Application to implicit solvation. The Journal of chemical physics, 145(3):034109.

Sanyal, T. and Shell, M. S. (2018). Transferable coarse-grained models of liquid-liquid equilibrium using local density potentials optimized with the relative entropy. The Journal of Physical Chemistry B, 122(21, SI):5678–5693.

Satija, R., Das, A., and Makarov, D. E. (2017). Transition path times reveal memory effects and anomalous diffusion in the dynamics of protein folding. The Journal of Chemical Physics, 147(15):152707.

Satija, R. and Makarov, D. E. (2019). Generalized langevin equation as a model for barrier crossing dynamics in biomolecular folding. The Journal of Physical Chemistry B, 123(4):802–810.

Schädle, A., López-Fernández, M., and Lubich, C. (2006). Fast and oblivious convolution quadrature. SIAM Journal on Scientific Computing, 28(2):421–438.

Schaeffer, H., Tran, G., and Ward, R. (2018). Extracting sparse high-dimensional dynamics from limited data. SIAM Journal on Applied Mathematics, 78(6):3279–3295.

Schneider, E., Dai, L., Topper, R. Q., Drechsel-Grau, C., and Tuckerman, M. E. (2017). Stochastic neural network approach for learning high-dimensional free energy surfaces. Phys. Rev. Lett., 119:150601.

Serrano, M. and Español, P. (2001). Thermodynamically consistent mesoscopic fluid particle model. Phys. Rev. E, 64:046115.

Seryo, N., Sato, T., Molina, J. J., and Taniguchi, T. (2020). Learning the constitutive relation of polymeric flows with memory. Phys. Rev. Research, 2:033107.

Shahidi, N., Chazirakis, A., Harmandaris, V., and Doxastakis, M. (2020). Coarse-graining of polyisoprene melts using inverse monte carlo and local density potentials. The Journal of Chemical Physics, 152(12):124902.

She, Z., Ge, P., and Lei, H. (2023). Data-driven construction of stochastic reduced dynamics encoded with non-markovian features. The Journal of Chemical Physics, 158(3):034102.

Shen, J., Xu, J., and Yang, J. (2018). The scalar auxiliary variable (sav) approach for gradient flows. Journal of Computational Physics, 353:407–416.

Shinoda, W., DeVane, R., and Klein, M. L. (2008). Coarse-grained molecular modeling of non-ionic surfactant self-assembly. Soft Matter, 4(12):2454–2462.



- Singh, D., Mondal, K., and Chaudhury, S. (2021). Effect of memory and inertial contribution on transition-time distributions: Theory and simulations. The Journal of Physical Chemistry B, 125(17):4536–4545.
- Singh, S., Krishnan, R., and Robinson, G. (1990). Theory of activated rate processes with space-dependent friction. Chemical Physics Letters, 175(4):338–342.
- Smith, D. E., Babcock, H. P., and Chu, S. (1999). Single-polymer dynamics in steady shear flow. Science, 283(5408):1724–1727.
- Soper, A. (1996). Empirical potential monte carlo simulation of fluid structure. Chemical Physics, 202(2):295–306.
- Stecher, T., Bernstein, N., and Csányi, G. (2014). Free energy surface reconstruction from umbrella samples using gaussian process regression. Journal of Chemical Theory and Computation, 10(9):4079–4097.
- Straub, J. E., Berne, B. J., and Roux, B. (1990). Spatial dependence of time-dependent friction for pair diffusion in a simple fluid. The Journal of Chemical Physics, 93(9):6804–6812.
- Straub, J. E., Borkovec, M., and Berne, B. J. (1987). Calculation of dynamic friction on intramolecular degrees of freedom. The Journal of Physical Chemistry, 91(19):4995–4998.
- Straub, J. E., Borkovec, M., and Berne, B. J. (1988). Molecular dynamics study of an isomerizing diatomic in a lennard-jones fluid. The Journal of Chemical Physics, 89(8):4833–4847.
- Straus, J. B., Gomez Llorente, J. M., and Voth, G. A. (1993). Manifestations of spatially dependent friction in classical activated rate processes. The Journal of Chemical Physics, 98(5):4082–4097.
- Tarjus, G. and Kivelson, D. (1991). Solvent effect on activated rate processes: On the validity of the gle approach. Chemical Physics, 152(1):153–167.
- Taylor, G. I. (1934). The formation of emulsions in definable fields of flow. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 146(858):501–523.
- Thien, N. P. and Tanner, R. I. (1977). A new constitutive equation derived from network theory. Journal of Non-Newtonian Fluid Mechanics, 2(4):353–365.
- Thomases, B. and Shelley, M. (2007). Emergence of singular structures in oldroyd-b fluids. Physics of Fluids, 19(10):103103.
- Torrie, G. and Valleau, J. (1977). Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. Journal of Computational Physics, 23(2):187–199.

- Voth, G. A. (1992). A theory for treating spatially-dependent friction in classical activated rate processes. The Journal of Chemical Physics, 97(8):5908–5910.
- Vroylandt, H. (2022). On the derivation of the generalized langevin equation and the fluctuation-dissipation theorem. Europhysics Letters, 140(6):62003.
- Vroylandt, H., Goudenège, L., Monmarché, P., Pietrucci, F., and Rotenberg, B. (2022). Likelihood-based non-markovian models from molecular dynamics. Proceedings of the National Academy of Sciences, 119(13):e2117586119.
- Vroylandt, H. and Monmarché, P. (2022). Position-dependent memory kernel in generalized Langevin equations: Theory and numerical estimation. The Journal of Chemical Physics, 156(24):244105.
- Wagner, J. W., Dannenhoffer-Lafage, T., Jin, J., and Voth, G. A. (2017). Extending the range and physical accuracy of coarse-grained models: Order parameter dependent interactions. The Journal of chemical physics, 147(4):044113.
- Wang, J., Olsson, S., Wehmeyer, C., Pérez, A., Charron, N. E., de Fabritiis, G., Noé, F., and Clementi, C. (2019). Machine learning of coarse-grained molecular dynamics force fields. ACS Central Science, 5(5):755–767.
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and testing of a general amber force field. Journal of Computational Chemistry, 25(9):1157–1174.
- Wang, Q. (1997). Comparative studies on closure approximations in flows of liquid crystal polymers: I. elongational flows. Journal of Non-Newtonian Fluid Mechanics, 72(2):141 – 162.
- Wang, S., Ma, Z., and Pan, W. (2020). Data-driven coarse-grained modeling of polymers in solution with structural and dynamic properties conserved. Soft Matter, 16(36):8330–8344.
- Warner, H. R. (1972a). Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. Industrial & Engineering Chemistry Fundamentals, 11(3):379–387.
- Warner, H. R. (1972b). Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. Industrial & Engineering Chemistry Fundamentals, 11(3):379–387.
- Willard, A. P. and Chandler, D. (2010). Instantaneous liquid interfaces. The Journal of Physical Chemistry B, 114(5):1954–1958.
- Womersley, J. R. (1955). Method for the calculation of velocity, rate of flow and viscous drag in arteries when the pressure gradient is known. The Journal of Physiology, 127(3):553–563.
- Xie, P., Car, R., and E, W. (2022). Ab Initio Generalized Langevin Equations. arXiv preprint arXiv:2211.06558.

- Yang, X. (2016). Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends. Journal of Computational Physics, 327:294–316.
- Yu, P., Du, Q., and Liu, C. (2005). From micro to macro dynamics via a new closure approximation to the fene model of polymeric fluids. Multiscale Modeling & Simulation, 3(4):895–917.
- Zaremba, S. (1903). Sur une forme perfectionnee de la theorie de la relaxation. Bull. Int. Acad. Sci. Cracovie, pages 594–614.
- Zavadlav, J., Marrink, S. J., and Praprotnik, M. (2018). Multiscale simulation of protein hydration using the swinger dynamical clustering algorithm. Journal of Chemical Theory and Computation, 14(3):1754–1761.
- Zhang, L., Han, J., Wang, H., Car, R., and E, W. (2018a). Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics. Phys. Rev. Lett., 120:143001.
- Zhang, L., Han, J., Wang, H., Car, R., and E, W. (2018b). Deepcg: Constructing coarse-grained models via deep neural networks. The Journal of Chemical Physics, 149(3):034101.
- Zhang, L., Han, J., Wang, H., Saidi, W., Car, R., and E, W. (2018c). End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, Advances in Neural Information Processing Systems 31, pages 4436–4446. Curran Associates, Inc.
- Zhang, L., Wang, H., and E, W. (2018d). Reinforced dynamics for enhanced sampling in large atomic and molecular systems. The Journal of Chemical Physics, 148(12):124113.
- Zhang, Y., Gao, C., Liu, Q., Zhang, L., Wang, H., and Chen, M. (2020). Warm dense matter simulation via electron temperature dependent deep potential molecular dynamics. Physics of Plasmas, 27(12):122704.
- Zhao, L., Li, Z., Caswell, B., Ouyang, J., and Karniadakis, G. E. (2018). Active learning of constitutive relation from mesoscopic dynamics for macroscopic modeling of non-newtonian flows. Journal of Computational Physics, 363:116 – 127.
- Zhou, X.-H., Han, J., and Xiao, H. (2021). Frame-independent vector-cloud neural network for nonlocal constitutive modeling on arbitrary grids. Computer Methods in Applied Mechanics and Engineering.
- Zhu, Y. and Venturi, D. (2020). Generalized langevin equations for systems with local interactions. Journal of Statistical Physics.
- Zwanzig, R. (1961). Statistical mechanics of irreversibility. Lectures in Theoretical Physics, 3:106–141.

Zwanzig, R. (1973). Nonlinear generalized Langevin equations. J. Stat. Phys., 9:215 – 220.

Zwanzig, R. (1992). Diffusion past an entropy barrier. The Journal of Physical Chemistry, 96(10):3926–3930.

Zwanzig, R. (2001). Nonequilibrium Statistical Mechanics. Oxford University Press.