

MACHINE LEARNING-BASED COARSE-GRAINED MODELS FOR MOLECULAR
SYSTEMS

By

Liyao Lyu

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

Multiscale modeling poses a formidable challenge in computational mathematics, particularly in integrating microscale interactions into meso- or macro-scale constitutive relations. While reduced-order models allow the simulation of extensive systems, their analytical formulations are generally unclosed. Take coarse-grained molecular dynamics as an example, the Mori-Zwanzig formalism decomposes the dynamics into deterministic, memory, and stochastic terms, but the explicit forms for these three terms are unknown. In this thesis, we present a series of studies to solve these problems. Firstly, the main challenge for the deterministic term comes from the high dimensionality and the presence of energy barriers of the free energy surface (FES). We propose a consensus sampling-based approach that reformulates the FES construction as a minimax problem. This framework simultaneously optimizes the function representation of the FES and the training set used to learn it. In particular, the maximization step establishes a stochastic interacting particle system to achieve the adaptive sampling of the max-residue regime by modulating the exploitation of the Laplace approximation of the current loss function and the exploration of the uncharted phase space; the minimization step updates the FES approximation with the new training set. By iteratively solving the minimax problem, the present method essentially achieves an adversarial learning of the FESs with unified tasks for both phase space exploration and posterior error-enhanced sampling. Besides, memory interactions are also important for predicting the collective transport and diffusion processes. To construct this, we introduce a machine-learning-based coarse-grained molecular dynamics model that captures the dissipative many-body contribution. The neural network representation is carefully designed to preserve the physical symmetries and the thermo-consistency. Unlike the common empirical reduced models, the present model is constructed based on the Mori-Zwanzig formalism and naturally inherits the heterogeneous state-dependent memory term rather than matching the mean-field metrics such as the velocity autocorrelation function. Finally, when applied to non-equilibrium systems, models based on the Mori-Zwanzig formalism face inherent challenges. A key issue lies in the Zwanzig projection, which relies on the marginal distribution of the system. We present a data-driven approach for constructing reduced models that

retain certain generalization abilities for non-equilibrium processes. Unlike the conventional CG models based on pre-selected CG variables (e.g., the center of mass), the present CG model seeks a set of auxiliary CG variables based on the time-lagged independent component analysis to minimize the entropy contribution of the unresolved variables. This ensures the distribution of the unresolved variables under a broad range of non-equilibrium conditions approaches the one under equilibrium. Through numerical validation, we demonstrate that our model can accurately predict viscoelastic behavior in various non-equilibrium flow regimes.

Copyright by
LIYAO LYU
2025

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Huan Lei. He is not only a brilliant scholar but also an exceptional mentor. Working with him has been an immensely enriching experience, helping me develop critical thinking and problem-solving skills. I am especially grateful for the intellectual freedom he provided, allowing me to explore and refine my ideas with confidence.

I am equally grateful to the members of my thesis committee, Prof. Andrew Christlieb, Prof. Di Liu, and Prof. Yimin Xiao, for their insightful critiques, constructive suggestions, and invaluable support. Prof. Andrew Christlieb is one of the most patient and approachable teachers I have had the privilege to learn from. I have gained a great deal from his expertise in hyperbolic conservation laws and tensor decomposition. I am also deeply appreciative of Prof. Di Liu for his guidance in machine learning, which provided me with valuable insights. Additionally, I sincerely enjoyed my discussions with Prof. Yimin Xiao on the analytical aspects of stochastic differential equations; these exchanges greatly enriched my understanding of the subject.

I would like to express my sincere gratitude to the professors whose courses greatly contributed to my academic growth. Prof. Jun Kitagawa's thorough instruction in partial differential equations helped me develop a careful approach to theoretical problems, while Prof. Hui-Chia Yu's engaging teaching style sparked my interest in numerical methods for chemical engineering. Prof. Rongrong Wang's clear explanations of numerical partial differential equations strengthened my understanding, and Prof. Daniel Appelö's expertise in numerical linear algebra helped me connect theory with practical applications. Their dedication to teaching has been invaluable in shaping my knowledge and perspective.

I would also like to sincerely thank my friends—Pei Ge, Zhiyuan She, Shijun Liang, Haishen Dai, Siyu Guo, Yue Zhao and many others—for their support and companionship throughout my PhD journey. This path can often be challenging and isolating, and their presence made it much more enjoyable and fulfilling. I deeply cherish the moments we shared, especially our dinners together, which brought warmth and laughter to this demanding period of my life.

Above all, I owe my deepest gratitude to my family for their unwavering support, encouragement, and sacrifices. Their patience during my countless hours of work, their belief in me during moments of doubt, and their unconditional love have been my foundation. This achievement is as much theirs as it is mine.

TABLE OF CONTENTS

CHAPTER 1	INTRODUCTION TO MULTISCALE MODELING AND MOLECULAR DYNAMICS	1
CHAPTER 2	DEVELOPING COARSE-GRAINED MOLECULAR DYNAMICS MODELS WITH NON-MARKOVIAN MEMORY EFFECTS	5
2.1	The Mori-Zwanzig Framework for Coarse-Graining	5
2.2	Mori-Zwanzig Formalism	7
2.3	Empirical Approaches to Coarse-Graining	9
2.4	Incorporating Physical Memory Kernels in Coarse-Grained Models	12
2.5	Symmetry-preserving neural network representation	13
2.6	Training	15
2.7	Numerical Result	15
2.8	Conclusion	18
CHAPTER 3	GENERALIZATION OF REDUCED MODELS FOR NON-EQUILIBRIUM DYNAMICS	21
3.1	Challenges in Modeling Non-Equilibrium Processes	21
3.2	Methodology for Generalizable Coarse-Grained Models	23
3.3	Numerical Results	33
3.4	Conclusion	42
CHAPTER 4	CONSENSUS-BASED ENHANCED SAMPLING FOR HIGH-DIMENSIONAL FREE ENERGY SURFACES	44
4.1	Challenges in Constructing High-Dimensional Free Energy Surfaces	44
4.2	Consensus-Based Sampling Methodology	47
4.3	Application of Consensus-Based Sampling to Biomolecular Systems	53
4.4	Conclusions	60
CHAPTER 5	CONCLUSION AND OUTLOOK	62
BIBLIOGRAPHY	64
APPENDIX A	SIMULATION DETAIL FOR NONEQUILIBRIUM CASE	76
APPENDIX B	SIMULATION DETAIL AND PROOF FOR CONSENSUS BASED SAMPLING	80

CHAPTER 1

INTRODUCTION TO MULTISCALE MODELING AND MOLECULAR DYNAMICS

Numerical simulation has emerged as an important tool for scientific understanding, complementing experimental and theoretical approaches. As was declared by Dirac back in 1929, the right physical principle for most of what we are interested in is already provided by the principles of quantum mechanics (QM), there is no need to look further. However, the ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe theoretically or computationally. This intricate challenge remained largely ignored until Anderson underscored in Anderson (1972), that the constructivist hypothesis falters when faced with the dual challenges of *scale* and *complexity*.

The vastness of *scale* leads to the well-known curse of dimensionality. While feeding atomic numbers of all involved atoms into QM can theoretically furnish a complete model, the problem's dimensionality swells by three with the addition of each atom. Consequently, these first principle models are currently only applied to simple cases and are inefficient for managing expansive system sizes and extended simulation timescales.

Model reduction shows promise, as they reduce degrees of freedom by omitting certain details, thereby facilitating larger-scale simulation. However, the reduced models are almost always obtained empirically, by guessing. Making the right guess often requires and represents far-reaching physical insight. In molecular dynamics (MD), the conservative interaction between particles is categorized into three distinct terms: two-particle pairwise, three-particle angular, and four-particle dihedral. On the mesoscale, as seen in Coarse-Grained Molecular Dynamics (CGMD), nonconservative friction forces are typically assumed to be isotropic in the Generalized Langevin Equation (GLE) or pairwise in Dissipative Particle Dynamics (DPD). At the macroscale, the derivation of macroscopic constitutive equations from kinetic equations necessitates the closure of the collision operator.

This guessing game can be quite hard and less productive because of the *complexity*. As we will show in the first project, the inherent complexity is often largely oversimplified by traditionally intuition-driven models, such as the GLE or DPD in coarse-grained molecular dynamics (CGMD)

problems. The recent emergence of machine learning, with its prowess in assimilating multimodal, multi-fidelity data and elucidating correlations among interrelated phenomena, offers a unique opportunity to address these challenges. However, simply inputting atomic coordinates into a network to fit interactions can lead to a cumbersome parameter space, complicating the training process. More critically, an exclusive reliance on machine learning risks sidelining the fundamental laws of physics, which may result in ill-posed problems or unphysical solutions.

Progress has been achieved by incorporating physical priors, such as physical symmetries, into the network ansatz. This not only enhances accuracy but also bolsters the transparency of the resulting models. A case in point is DeepMD, which constructs an extensive and symmetry-invariant network to capture the many-body nature of the conservative force field. However, accurate prediction of the dynamics in CGMD further relies on faithfully modeling a memory term that represents the energy-dissipation processes arising from the unresolved degrees of freedom (DoF).

The importance of the state-dependent non-Markovian friction tensor (force) for dynamic properties remains broadly overlooked, with the conventional state-independent GLE kernel being often used. In our first project Lyu and Lei (2023b), we will show that the state-dependence of memory terms can be critical for constructing a reliable extensive CGMD model that preserves the collective dynamics behavior. The commonly used velocity auto-correlation function (VACF) is a mean-field metric. Reproducing the VACF (as can be sufficient to capture in GLE) is insufficient to capture the crucial collective behavior such as the hydrodynamic mode, collective diffusion (Van-Hove function), et al. In contrast, we propose a physical information-embedded network-based model constructed directly from Zwanzig's formalism that faithfully retains the heterogeneous energy dissipation among the CG particles and accurately predicts the dynamics on the collective scale.

The complexity isn't confined solely to the dynamic equation; it also extends to the selection of dynamic coordinates Lyu and Lei (2024). A case in point is water models. Coordinates (or sites) ranging from one to six have been proposed to represent water, each aiming to capture distinct behaviors. The judicious selection of coordinates is rooted in physical insight, much like

the model-building previously mentioned. As the complexity of the molecular structure escalates, this task becomes progressively challenging. Data-driven coordinate selection seems a potential alternative to the traditional human intelligence-driven approach. In extensive CGMD systems, the center of mass (COM) remains the most prevalent coordinate. However, when modeling non-Newtonian fluids, which exhibit enormously complex flow behavior, capturing the deformation of polymers becomes crucial. This aspect is often oversimplified when relying solely on COM. The ongoing project intends to emphasize the importance of dynamic coordinate selection for faithfully capturing macroscopic rheology behavior.

Apart from dynamic equations and dynamic coordinates, one particular problem of these data-driven methods is sampling Lyu and Lei (2023a). Generating a good representative data set is in general not an easy task. For example, in molecular dynamics, a representative data set to construct a free energy surface (FES) includes most states of a molecule including ground states, meta-stable states, and some high-energy states (unstable states). In most cases, the molecules are stuck in a local minimum on the energy surface, and it is difficult to jump over the energy barriers. Enhance sampling thus is important for these problems to accelerate ergodic ability.

To increase the probability of rare events occurring in MD simulations, the simulation process can be interfered with using various methods, including raising the temperature, replica exchange, and adding bias potential. Traditional boosted sampling methods based on bias potential suffer from the curse of dimensionality. In the third project, we propose the Consensus-Based Enhanced Sampling (CES) by framing the FESs construction as a min-max problem. The minimization aspect focuses on optimizing the network parameters to minimize error at the sampled points, while the maximization aspect focuses on optimizing the distribution of the sampling to maximize the error within the dataset. By iteratively addressing the min-max problem, we are able to construct a high-dimensional energy surface.

Integrating deep learning with multiscale modeling holds potential, offering enhancements to models traditionally anchored in human intelligence. However, forging a seamless integration between machine learning and established scientific structural models remains elusive. The accuracy

and transparency of data-driven models continue to be viewed with skepticism by a broader community. We aspire that our work will contribute to bolstering the confidence in these approaches. While our primary focus has been on MD and CGMD scales, we also aim to extend our methods to the macro scale in the future.

CHAPTER 2

DEVELOPING COARSE-GRAINED MOLECULAR DYNAMICS MODELS WITH NON-MARKOVIAN MEMORY EFFECTS

© 2023 American Physical Society. Reproduced with permission from L. Lyu and H. Lei, *Phys. Rev. Lett.* **131**, 177301 (2023). <https://doi.org/10.1103/PhysRevLett.131.177301>

2.1 The Mori-Zwanzig Framework for Coarse-Graining

Accurately predicting the collective behavior of multi-scale physical systems is a long-standing problem that requires the integrated modeling of the molecular-level interactions across multiple scales Anderson (1972). However, for systems without clear scale separation, there often exists no such a set of simple collective variables by which we can formulate the evolution in an analytic and self-determined way. One canonical example is coarse-grained molecular dynamics (CGMD). While the reduced degrees of freedom (DoFs) enable us to achieve a broader range of the spatio-temporal scale, the construction of truly reliable CG models remains highly non-trivial. A significant amount of work Torrie and Valleau (1977a); Rosso et al. (2002); Maragliano and Vanden-Eijnden (2006); Izvekov and Voth (2005); Noid et al. (2008); Rudd and Broughton (1998); Lyubartsev and Laaksonen (1995); Shell (2008); Kumar et al. (1992); Nielsen et al. (2004); Laio and Parrinello (2002); Darve and Pohorille (2001) (see also review Noid (2013)), including recent machine learning (ML)-based approaches Behler and Parrinello (2007); Stecher et al. (2014a); John and Csányi (2017); Lemke and Peter (2017); Chmiela et al. (2017); Zhang et al. (2018a,b), have been devoted to constructing the conservative CG potential for retaining consistent static and thermodynamic properties. However, accurate prediction of the CG dynamics further relies on faithfully modeling a memory term that represents the energy-dissipation processes arising from the unresolved DoFs; the governing equations generally become non-Markovian on the CG scale. Moreover, such non-Markovian term often depends on the resolved variables in a complex way Satija et al. (2017); Luo et al. (2006); Best and Hummer (2010); Plotkin and Wolynes (1998); Straus et al. (1993); Morrone et al. (2012); Daldrop et al. (2017) where the analytic formulation is generally unknown. In particular, for extensive CGMD systems (i.e., the number of CG particles can be proportionally changed according

to the simulation size), the memory term often exhibits strong many-body effect and needs to satisfy various physical symmetry constraints among the CG particles. Existing approaches often rely on empirical models such as Brownian motion Einstein (1905), Langevin dynamics Kampen (2007), and dissipative particle dynamics (DPD) Hoogerbrugge and Koelman (1992); Español and Warren (1995). Despite their broad applications, studies Lei et al. (2010); Hijón et al. (2010); Yoshimoto et al. (2013) based on direct construction from full MD show that the empirical (e.g., pairwise additive) forms can be insufficient to capture the state-dependent energy-dissipation processes due to the many-body and non-Markovian effects. Recent efforts Lange and Grubmüller (2006); Wang et al. (2020b); Ceriotti et al. (2009); Baczewski and Bond (2013); Davtyan et al. (2015); Lei et al. (2016); Li et al. (2017); Russo et al. (2019); Jung et al. (2017); Lee et al. (2019); Ma et al. (2019, 2021); Klippenstein and van der Vegt (2021); Vroylandt et al. (2022); She et al. (2023); Xie et al. (2022) model the memory term based on the generalized Langevin equation (GLE) and its variants (see also review Klippenstein et al. (2021)). The velocity auto-correlation function (VACF) is often used as the target quantity for model parameterization. While it may serve as an appropriate measure for certain non-extensive systems Widder et al. (2022); Meyer et al. (2019), the VACF is essentially a metric of the background dissipation under mean-field approximation. For extensive CGMD systems, the homogeneous kernel overlooks the heterogeneity of the energy dissipation among the CG particles stemming from the many-body nature of the marginal probability density function of the CG variables. This limitation imposes a fundamental challenge for accurately modeling the local irreversible responses as well as the transport and diffusion processes on the collective scale.

This work aims to fill the gap with a new CG model that faithfully entails the state-dependent non-Markovian memory and the coherent noise for extensive MD systems. The model formulation can be loosely viewed as an extended dynamics of the CG variables joint with a set of non-Markovian features that embodies the many-body nature of the energy dissipation among the CG particles. Specifically, we treat each CG particle as an agent and seek a set of symmetry-preserving neural network (NN) representations that directly map its local environments to the non-Markovian friction interactions, and thereby circumvent the exhausting efforts of fitting the individual memory terms

with a unified empirical form. Different from the ML-based potential model Zhang et al. (2018b), the memory terms are represented by NNs in form of second-order tensors that strictly preserve the rotational symmetry and the positive-definite constraint. Coherent noise can be introduced satisfying the second fluctuation-dissipation theorem and retaining consistent invariant distribution. Rather than matching the VACF, the model is trained based on the Mori-Zwanzig (MZ) projection formalism such that the effects of the unresolved interactions can be seamlessly inherited. We emphasize that the construction is not merely for mathematical rigor. Numerical results of a polymer molecule system show that the CG models with empirical memory forms are generally insufficient to capture heterogeneous inter-molecular dissipation that leads to inaccurate cross-correlation functions among the particles. Fortunately, the present model can reproduce both the auto- and cross-correlation functions. More importantly, it accurately predicts the challenging collective dynamics characterized by the hydrodynamic mode correlation and the van Hove function Van Hove (1954) and shows the promise to predict the meso-scale transport and diffusion processes encoded with molecular-level fidelity.

2.2 Mori-Zwanzig Formalism

Let us consider a full MD system consisting of M molecules with a total number of N atoms. The phase space vector is denoted by $z = [\mathbf{q}, \mathbf{p}]$, where $\mathbf{q}, \mathbf{p} \in \mathbb{R}^{3N}$ represent the position and momentum vector, respectively. Given $z(0) = z_0$, the evolution follows $z(t) = e^{\mathcal{L}t} z_0$, where \mathcal{L} is the Liouville operator determined by the Hamiltonian $H(z)$. The CG variables are defined by representing each molecule as a CG particle, i.e., $\phi(z) = [\phi^Q(z), \phi^P(z)]$, where $\phi^Q(z) = [\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_M]$ and $\phi^P(z) = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_M]$ represent the center of mass and the total momentum of individual molecules, respectively. $\mathbf{Z}(t) = [\mathbf{Q}(t), \mathbf{P}(t)]$ denote the map $\phi(z(t))$ with $z(0) = z_0$. Using the Koopman operator Koopman (1931), $\mathbf{Z}(t)$ can be mapped from the initial values, i.e.,

$$\mathbf{Z}(t) = e^{\mathcal{L}t} \mathbf{Z}(0), \quad (2.1)$$

where \mathcal{L} is the Liouville operator determined by the full-model Hamiltonian $H(z)$. Below we derive the reduced model by choosing CG variables \mathbf{Z} as a linear mapping of the full phase-space

vector \mathbf{z} and we refer to Ref. Hijón et al. (2010); Darve et al. (2009) for discussions of the more general cases.

Following Zwanzig's approach, we define a projection operator as the conditional expectation with a fixed CG vector \mathbf{Z} , i.e., $\mathcal{P}_{\mathbf{Z}}f(\mathbf{z}) := \int \delta(\phi(\mathbf{z}) - \mathbf{Z})\rho_0(\mathbf{z})f(\mathbf{z}) d\mathbf{z}/\Omega(\mathbf{Z})$, where $\rho_0(\mathbf{z}) \propto e^{-\beta H(\mathbf{z})}$ represents the equilibrium density function and $\Omega(\mathbf{Z}) = \int \delta(\phi(\mathbf{z}) - \mathbf{Z})\rho_0(\mathbf{z}) d\mathbf{z}$. Also, we define an orthogonal operator $\mathcal{Q}_{\mathbf{Z}} = \mathbf{I} - \mathcal{P}_{\mathbf{Z}}$. Using Eq. (2.1), we have $\dot{\mathbf{Z}}(t) = e^{\mathcal{L}t}\mathcal{P}_{\mathbf{Z}}\mathcal{L}\mathbf{Z}(0) + e^{\mathcal{L}t}\mathcal{Q}_{\mathbf{Z}}\mathcal{L}\mathbf{Z}(0)$. In particular, we choose $\mathbf{Z} = [\mathbf{Q}, \mathbf{P}] = \mathbf{Z}(0)$. Using the Duhamel-Dyson identity, we can write the dynamics of $\mathbf{Z}(t)$ as

$$\dot{\mathbf{Z}}(t) = e^{\mathcal{L}t}\mathcal{P}_{\mathbf{Z}}\mathcal{L}\mathbf{Z}(0) + \int_0^t ds e^{\mathcal{L}(t-s)}\mathcal{P}_{\mathbf{Z}}\mathcal{L}e^{\mathcal{Q}_{\mathbf{Z}}\mathcal{L}s}\mathcal{Q}_{\mathbf{Z}}\mathcal{L}\mathbf{Z}(0) + e^{\mathcal{Q}_{\mathbf{Z}}\mathcal{L}t}\mathcal{Q}_{\mathbf{Z}}\mathcal{L}\mathbf{Z}(0). \quad (2.2)$$

Let us start with the mean-field term $\mathcal{P}_{\mathbf{Z}}\mathcal{L}\mathbf{Z}(0)$. For the present study, the CG variables are linear functions of \mathbf{z} . Therefore, we have $\mathcal{P}_{\mathbf{Z}}\mathcal{L}\mathbf{Q} = \mathcal{L}\mathbf{Q} = \mathbf{M}^{-1}\mathbf{P}$, i.e., $\mathcal{Q}_{\mathbf{Z}}\mathcal{L}\mathbf{Q} \equiv 0$. For $\mathcal{P}_{\mathbf{Z}}\mathcal{L}\mathbf{P}$ associated with the i -th CG particle, we have

$$\begin{aligned} \mathcal{P}_{\mathbf{Z}}\mathcal{L}\mathbf{P}_i &= \int \delta(\phi(\mathbf{z}) - \mathbf{Z})\rho_0(\mathbf{z})\mathcal{L}\mathbf{P}_i d\mathbf{z}/\Omega(\mathbf{Z}) \\ &= \int \delta(\phi(\mathbf{z}) - \mathbf{Z})\rho_0(\mathbf{z})\left(-\sum_{i \in \mathcal{N}_i} \nabla_{\mathbf{q}_i} H(\mathbf{z})\right) d\mathbf{z}/\Omega(\mathbf{Z}) \\ &= \int \delta(\phi(\mathbf{z}) - \mathbf{Z})\left(\beta^{-1} \sum_{i \in \mathcal{N}_i} \nabla_{\mathbf{q}_i}\right)\rho_0(\mathbf{z}) d\mathbf{z}/\Omega(\mathbf{Z}) \\ &= \beta^{-1}\nabla_{\mathbf{Q}_i} \int \delta(\phi^{\mathcal{Q}}\phi(\mathbf{q}) - \mathbf{Q})\rho_0(\mathbf{q}) d\mathbf{q} / \int \delta(\phi^{\mathcal{Q}}(\mathbf{q}) - \mathbf{Q})\rho_0(\mathbf{q}) d\mathbf{q} \\ &= -\nabla_{\mathbf{Q}_i} U(\mathbf{Q}), \end{aligned} \quad (2.3)$$

where \mathcal{N}_i represents the index set of the atoms that belongs to the i -th molecule, and the free energy $U(\mathbf{Q}) = -\beta^{-1} \ln \left[\int \delta(\phi^{\mathcal{Q}}(\mathbf{q}) - \mathbf{Q})\rho_0(\mathbf{q}) d\mathbf{q} \right]$.

For the memory term $\mathcal{P}_Z \mathcal{L} e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P$ associated with the i -th CG particle, we have

$$\begin{aligned}
\mathcal{P}_Z \mathcal{L} e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i &= \int \rho_0(z) \delta(\phi(z) - Z) \mathcal{L} e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i dz / \Omega(Z) \\
&= \int \rho_0(z) (\mathcal{L} \phi(z) \cdot \nabla_Z) \delta(\phi(z) - Z) e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i dz / \Omega(Z) \\
&= \int \rho_0(z) (Q_Z \mathcal{L} P \cdot \nabla_P) \delta(\phi(z) - Z) e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i dz / \Omega(Z) \quad (\text{by } Q_Z \mathcal{L} Q \equiv 0) \\
&= \nabla_P \cdot \int \rho_0(z) \delta(\phi(z) - Z) (Q_Z \mathcal{L} P) \otimes e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i dz / \Omega(Z) \\
&= \nabla_P \cdot \underbrace{\left(\int \rho_0(z) \delta(\phi(z) - Z) (Q_Z \mathcal{L} P) \otimes e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i dz / \Omega(Z) \right)}_{\tilde{K}_i(Z, s)} \\
&\quad - \tilde{K}_i(Z, s) \cdot \nabla_P (1/\Omega(Z)) \Omega(Z).
\end{aligned} \tag{2.4}$$

Furthermore, we take the assumption that the memory kernel only depends on the positions of the CG particles Q , i.e., $\nabla_P \cdot \tilde{K}(Z, s) \equiv 0$. Also, similar to the derivation in Eq. (2.3), we note that

$$\Omega(Z) \propto \int \delta(\phi^Q(q) - Q) \rho_0(q) \delta(\phi^P(q) - P) e^{-\beta P^T M^{-1} P / 2} dz \propto e^{-\beta P^T M^{-1} P / 2}. \tag{2.5}$$

Therefore, Eq. (2.4) can be further simplified as

$$\mathcal{P}_Z \mathcal{L} e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} P_i = -\beta \tilde{K}_i(Q, s) \cdot M^{-1} P. \tag{2.6}$$

With Eqs. (2.3) (2.6), we can show that the dynamics of $Z = [Q, P]$ can be written as

$$\begin{aligned}
\dot{Q} &= M^{-1} P \\
\dot{P} &= -\nabla U(Q) - \int_0^t K(Q(t-s), s) V(t-s) ds + R(t),
\end{aligned} \tag{2.7}$$

where $K(Q, s) = \beta \tilde{K}(Q, s)$ and $R(t) = e^{Q_Z \mathcal{L} t} Q_Z \mathcal{L} Z(0)$ is modeled as a random process representing the different initial condition z_0 with $\phi(z_0) = Z$.

2.3 Empirical Approaches to Coarse-Graining

Eq. (2.7) provides the starting point to derive the various CG models. Direct evaluation of $K(Q, t)$ imposes a challenge as it relies on solving the full-dimensional orthogonal dynamics $e^{Q_Z t}$. Some empirical models have been developed to represent the kernel historically.

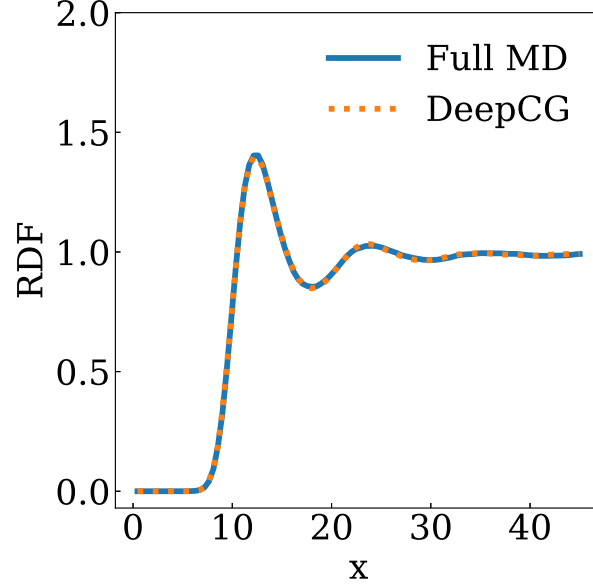


Figure 2.1 The RDF of the full MD and various CG models with the conservative CG potential $U(\mathbf{Q})$ constructed by the DeepCG model.

2.3.1 M-DPD

- **Formulation:** In the M-DPD model, the dynamics of the CG system is governed by

$$\begin{aligned}\dot{\mathbf{Q}}_i &= m_i^{-1} \mathbf{P}_i, \\ \dot{\mathbf{P}}_i &= \mathbf{F}_i^C - \sum_j \gamma_{ij} \mathbf{V}_j + \mathbf{R}(t).\end{aligned}\tag{2.8}$$

The computation of conservative force term \mathbf{F}^C is conducted by DeePCG Zhang et al. (2018b) method for all models considered in this paper. This approach enables the successful reproduction of the structural properties for each model as shown in Fig. 2.1.

- **Construction:** The friction tensor is calculated from random force correlation by constrained dynamics

$$\begin{aligned}\gamma_{ij}(\mathbf{Q}) &= \gamma_{\perp}(\mathbf{Q}_{ij})(\mathbf{1} - \mathbf{e}_{ij}\mathbf{e}_{ij}^T) + \gamma_{\parallel}(\mathbf{Q}_{ij})(\mathbf{e}_{ij}\mathbf{e}_{ij}^T), \\ \gamma_{\parallel}(\mathbf{Q}_{ij}) &= \frac{1}{k_B T} \int_0^{\infty} \mathbb{E}[\delta \mathbf{F}_i^{\parallel}(s)^T \cdot \delta \mathbf{F}_j^{\parallel}(0)] ds, \\ \gamma_{\perp}(\mathbf{Q}_{ij}) &= \frac{1}{k_B T} \int_0^{\infty} \mathbb{E}[\delta \mathbf{F}_i^{\perp}(s)^T \cdot \delta \mathbf{F}_j^{\perp}(0)] ds,\end{aligned}\tag{2.9}$$

where \mathbf{e}_{ij} is the unit vector represents along \mathbf{Q}_{ij} , the symbol \parallel represents the radial component along \mathbf{Q}_{ij} and \perp for perpendicular component. For further details, see Lei et al. (2010); Hijón

et al. (2010).

2.3.2 NM-GLE

- **Formulation:** In the standard GLE model, the dynamics of the CG system are governed by

$$\begin{aligned}\dot{\mathbf{Q}} &= \mathbf{M}^{-1} \mathbf{P}, \\ \dot{\mathbf{P}} &= \mathbf{F}^C - \int_0^t \mathbf{K}(t-s) \mathbf{V}(s) ds + \mathbf{R}(t).\end{aligned}\tag{2.10}$$

where the over-simplified kernel \mathbf{K} is only time-independent and homogeneous among the CG particles, as discussed in the main context.

- **Construction:** The memory kernel depends on solving the Volterra equation

$$\mathbf{G}(t) = - \int_0^t \mathbf{K}(t-s) \mathbf{C}(s) ds,\tag{2.11}$$

where $\mathbf{C}(t) = \mathbb{E}[\mathbf{V}(t)\mathbf{V}(0)^T]$ and $\mathbf{G}(t) = \mathbb{E}[\delta\mathbf{F}(t)\mathbf{V}(0)^T]$ are the correlation matrices. In practice, the velocity correlation matrix $\mathbf{C}(t)$ is often simplified to VACF $c(t) = \mathbb{E}[\mathbf{V}(t)^T \cdot \mathbf{V}(0)]$ for individual CG particles under the mean-field approximation. As a result, kernel \mathbf{K} reduces to a scalar function depending only on time, which overlooks the state dependency. With the obtained kernel in proper form, the system can be cast into an extended dynamics driven by white noise, e.g., see Ceriotti et al. (2009); Lei et al. (2016); Wang et al. (2020b); She et al. (2023).

2.3.3 NM-DPD

- **Formulation:** In the NM-DPD model, the dynamics of the CG system are governed by

$$\begin{aligned}\dot{\mathbf{Q}}_i &= m_i^{-1} \mathbf{P}_i, \\ \dot{\mathbf{P}}_i &= \mathbf{F}_i^C - \sum_j \int_0^t \mathbf{K}_{ij}(t-s) \mathbf{V}_{ij}(s) ds + \sum_j \mathbf{R}_{ij}(t),\end{aligned}\tag{2.12}$$

where the kernel \mathbf{K}_{ij} further depends on the pair-wise distance Q_{ij} .

- **Construction:** The memory kernel depends on solving the Volterra integral equation

$$\mathbf{A}_{ij}(t) = - \int_0^t \mathbf{K}_{ij}(t-s) \mathbf{B}_{ij}(s) ds,\tag{2.13}$$

where the correlation functions are computed as two orthogonal parts, i.e.,

$$\begin{aligned}\mathbf{A}_{ij}(t) &= \mathbb{E}[\delta\mathbf{F}_{ij}^{\parallel}(t) \cdot \mathbf{V}_{ij}^{\parallel}(0)^T] + \mathbb{E}[\delta\mathbf{F}_{ij}^{\perp}(t) \cdot \mathbf{V}_{ij}^{\perp}(0)^T], \\ \mathbf{B}_{ij}(t) &= \mathbb{E}[\mathbf{V}_{ij}^{\parallel}(t) \cdot \mathbf{V}_{ij}^{\parallel}(0)^T] + \mathbb{E}[\mathbf{V}_{ij}^{\perp}(t) \cdot \mathbf{V}_{ij}^{\perp}(0)^T].\end{aligned}\tag{2.14}$$

As a result, the orthogonal parts of the friction term can be solved separately, and extended dynamics can be also applied for simulation, e.g., see Li et al. (2017).

However, as shown below, such empirical forms are limited to capturing the state-dependence that turns out to be crucial for the dynamics on the collective scale, and motivates the present model retaining the many-body nature of $\mathbf{K}(\mathbf{Q}, t)$.

2.4 Incorporating Physical Memory Kernels in Coarse-Grained Models

To elaborate the essential idea, let us start with the Markovian approximation $\mathbf{K}(\mathbf{Q}, t) \approx -\mathbf{\Gamma}(\mathbf{Q})\delta(t)$, where $\mathbf{\Gamma}(\mathbf{Q}) = \mathbf{\Xi}(\mathbf{Q})\mathbf{\Xi}(\mathbf{Q})^T$ is the friction tensor preserving the semi-positive definite condition, and $\mathbf{\Xi}(\mathbf{Q})$ needs to retain the translational, rotational, and permutational symmetry, i.e.,

$$\begin{aligned}\mathbf{\Xi}_{ij}(\mathbf{Q}_1 + \mathbf{b}, \dots, \mathbf{Q}_M + \mathbf{b}) &= \mathbf{\Xi}_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) \\ \mathbf{\Xi}_{ij}(\mathcal{U}\mathbf{Q}_1, \dots, \mathcal{U}\mathbf{Q}_M) &= \mathcal{U}\mathbf{\Xi}_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M)\mathcal{U}^T \\ \mathbf{\Xi}_{\sigma(i)\sigma(j)}(\mathbf{Q}_{\sigma(1)}, \dots, \mathbf{Q}_{\sigma(M)}) &= \mathbf{\Xi}_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M),\end{aligned}\tag{2.15}$$

where $\mathbf{\Xi}_{ij} \in \mathbb{R}^{3 \times 3}$ represents the friction contribution of j -th particle on i -th particle, $\mathbf{b} \in \mathbb{R}^3$ is a translation vector, \mathcal{U} is a unitary matrix, and $\sigma(\cdot)$ is a permutation function.

To inherit the many-body interactions, we map the local environment of each CG particle into a set of generalized coordinates, i.e., $\hat{\mathbf{Q}}_i^k = \mathbf{Q}_i + \sum_{l \in \mathcal{N}_i} f^k(Q_{il})\mathbf{Q}_{il}$, where $f : \mathbb{R} \rightarrow \mathbb{R}^K$ is an encoder function to be learned, and $\mathcal{N}_i = \{l | Q_{il} < r_c\}$ is the neighboring index set of the i -th particle within a cut-off distance r_c . Accordingly, $\hat{\mathbf{Q}}_{ij} \in \mathbb{R}^{3 \times K}$ represents a set of features that encode the inter-molecular configurations beyond the pairwise approximation. The k -th column $\hat{\mathbf{Q}}_{ij}^k = \hat{\mathbf{Q}}_i^k - \hat{\mathbf{Q}}_j^k$ preserves the translational and permutational invariance, by which we represent $\mathbf{\Xi}_{ij}$ by

$$\mathbf{\Xi}_{ij} = \sum_{k=1}^K h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) \hat{\mathbf{Q}}_{ij}^k \otimes \hat{\mathbf{Q}}_{ij}^k + h_0(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) \mathbf{I}\tag{2.16}$$

where $h : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^{K+1}$ are encoder functions which will be represented by NNs. For $i = j$, we have $\mathbf{\Xi}_{ii} = -\sum_{j \in \mathcal{N}_i} \mathbf{\Xi}_{ij}$ based on the Newton's third law.

Eq. (2.16) entails the state-dependency of the memory term $K(\mathbf{Z}, t)$ under the Markovian approximation. To incorporate the non-Markovian effect, we embed the memory term within an extended Markovian dynamics Ceriotti et al. (2009) (see also Ref. She et al. (2023)). Specifically, we seek a set of non-Markovian features $\zeta := [\zeta_1, \zeta_2, \dots, \zeta_n]$, and construct the joint dynamics of $[\mathbf{Z}, \zeta]$ by imposing the many-body form of the friction tensor between \mathbf{P} and ζ , i.e.,

$$\begin{aligned}\dot{\mathbf{Q}} &= \mathbf{M}^{-1} \mathbf{P} \\ \dot{\mathbf{P}} &= -\nabla U(\mathbf{Q}) + \Xi(\mathbf{Q}) \zeta \\ \dot{\zeta} &= -\Xi(\mathbf{Q})^T \mathbf{V} - \Lambda \zeta + \xi(t),\end{aligned}\tag{2.17}$$

where $\Xi = [\Xi^1 \Xi^2 \dots \Xi^n]$ and each sub-matrix takes the form (2.16) constructed by $\{\mathbf{f}^i(\cdot), \mathbf{h}^i(\cdot)\}_{i=1}^n$ respectively. $\Lambda = \hat{\Lambda} \otimes \mathbf{I}$ represents the coupling among n features, where $\mathbf{I} \in \mathbb{R}^{3N \times 3N}$ is the identity matrix and $\hat{\Lambda} \in \mathbb{R}^{n \times n}$ needs to satisfy the Lyapunov stability condition $\hat{\Lambda} + \hat{\Lambda}^T \geq 0$. Therefore, we write $\hat{\Lambda} = \hat{\mathbf{L}} \hat{\mathbf{L}}^T + \hat{\mathbf{L}}^a$, where $\hat{\mathbf{L}}$ is a lower triangular matrix and $\hat{\mathbf{L}}^a$ is an anti-symmetry matrix which will be determined later. By choosing the white noise $\xi(t)$ following

$$\langle \xi(t) \xi(t') \rangle = \beta^{-1} (\Lambda + \Lambda^T) \delta(t - t'),\tag{2.18}$$

Eq. (2.17) retains the consistent invariant distribution

$$\rho(\mathbf{Q}, \mathbf{P}, \xi) \propto \exp[-\beta(U(\mathbf{Q}) + \mathbf{P}^T \mathbf{M}^{-1} \mathbf{P}/2 + \zeta^T \zeta/2)].$$

2.5 Symmetry-preserving neural network representation

Preserving the physical symmetry constraints is crucial for both the accuracy and the generalization ability of the constructed ML-models. Besides the conservative potential $U(\mathbf{Q})$, the constructed memory term will need to satisfy the translation- and permutation-invariance, as well as the rotation-symmetries. Let \mathcal{T}_b , \mathcal{R}_U , and \mathcal{P}_σ denote the translation, rotation, and permutation operator whose actions on a general function $\mathcal{F}(\mathbf{Q}_1, \dots, \mathbf{Q}_M)$ defined by

$$\begin{aligned}\mathcal{T}_b \mathcal{F}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) &:= \mathcal{F}(\mathbf{Q}_1 + \mathbf{b}, \dots, \mathbf{Q}_M + \mathbf{b}), \\ \mathcal{R}_U \mathcal{F}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) &:= \mathcal{F}(\mathbf{Q}_1 \mathcal{U}, \dots, \mathbf{Q}_M \mathcal{U}), \\ \mathcal{P}_\sigma \mathcal{F}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) &:= \mathcal{F}(\mathbf{Q}_{\sigma(1)}, \dots, \mathbf{Q}_{\sigma(M)}),\end{aligned}\tag{2.19}$$

where $\mathbf{b} \in \mathbb{R}^3$ is a position vector, $\mathbf{U} \in \mathbb{R}^{3 \times 3}$ is an orthogonal matrix and σ is an arbitrary permutation of the set of indices. The components of the constructed memory will need to satisfy the symmetry constraints

$$\begin{aligned}\mathcal{T}_b \Xi_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) &= \Xi_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) \\ \mathcal{R}_U \Xi_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) &= \mathbf{U} \Xi_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) \mathbf{U}^T \\ \mathcal{P}_\sigma \Xi_{ij}(\mathbf{Q}_1, \dots, \mathbf{Q}_M) &= \Xi_{\sigma(i)\sigma(j)}(\mathbf{Q}_{\sigma(1)}, \dots, \mathbf{Q}_{\sigma(M)}),\end{aligned}\tag{2.20}$$

Proposition 2.5.1. *The representation $\Xi_{ij} = \sum_{k=1}^K h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) (\hat{\mathbf{Q}}_{ij}^k) (\hat{\mathbf{Q}}_{ij}^k)^T + h_0(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) \mathbf{I}$ preserves the symmetry conditions (2.20), where $\hat{\mathbf{Q}}_i^k = \mathbf{Q}_i + \sum_{l \in \mathcal{N}_i} f^k(Q_{il}) \mathbf{Q}_{il}$ represents the local environment-determined features (generalized coordinate) for the i -th particle, $\mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}^K$ and $\mathbf{h} : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}^{K+1}$ are two encoder functions.*

Proof. We note that $\mathcal{T}_b \mathbf{Q}_{ij} = \mathcal{T}_b \mathbf{Q}_i - \mathcal{T}_b \mathbf{Q}_j = \mathbf{Q}_{ij}$, $\mathcal{T}_b \mathbf{Q}_{ij} = \|\mathcal{T}_b \mathbf{Q}_i - \mathcal{T}_b \mathbf{Q}_j\| = \mathbf{Q}_{ij}$, $\mathcal{R}_U \mathbf{Q}_{ij} = \mathbf{U} \mathbf{Q}_{ij}$, $\mathcal{R}_U \mathbf{Q}_{ij} = \mathbf{Q}_{ij}$, $\mathcal{P}_\sigma \mathbf{Q}_{ij} = \mathbf{Q}_{\sigma(i)\sigma(j)}$, and $\mathcal{P}_\sigma \mathbf{Q}_{ij} = \mathbf{Q}_{\sigma(i)\sigma(j)}$. Therefore, for arbitrary indices i and k , the feature $\hat{\mathbf{Q}}_i^k$ satisfy the following symmetry conditions

$$\begin{aligned}\mathcal{T}_b \hat{\mathbf{Q}}_i^k &= \mathcal{T}_b \mathbf{Q}_i + \sum_{l \in \mathcal{N}_i} f^k(\mathcal{T}_b \mathbf{Q}_{il}) \mathcal{T}_b \mathbf{Q}_{il} = \hat{\mathbf{Q}}_i^k + \mathbf{b} \\ \mathcal{R}_U \hat{\mathbf{Q}}_i^k &= \mathcal{R}_U \mathbf{Q}_i + \sum_{l \in \mathcal{N}_i} f^k(\mathcal{R}_U \mathbf{Q}_{il}) \mathcal{R}_U \mathbf{Q}_{il} = \mathbf{U} \hat{\mathbf{Q}}_i^k \\ \mathcal{P}_\sigma \hat{\mathbf{Q}}_i^k &= \mathcal{P}_\sigma \mathbf{Q}_i + \sum_{l \in \mathcal{N}_{\sigma(i)}} f^k(\mathcal{P}_\sigma \mathbf{Q}_{il}) \mathcal{P}_\sigma \mathbf{Q}_{il} = \hat{\mathbf{Q}}_{\sigma(i)}^k,\end{aligned}\tag{2.21}$$

where we have used the fact that $\sum_l f(r_l) \mathbf{r}_l$ is permutational invariant for the last equation.

Therefore, we have $\mathcal{T}_b \hat{\mathbf{Q}}_{ij} = \mathcal{T}_b \hat{\mathbf{Q}}_i - \mathcal{T}_b \hat{\mathbf{Q}}_j = \hat{\mathbf{Q}}_{ij}$, $\mathcal{T}_b \hat{\mathbf{Q}}_{ij} = \|\mathcal{T}_b \hat{\mathbf{Q}}_i - \mathcal{T}_b \hat{\mathbf{Q}}_j\| = \hat{\mathbf{Q}}_{ij}$, $\mathcal{R}_U \hat{\mathbf{Q}}_{ij} = \mathbf{U} \hat{\mathbf{Q}}_{ij}$, $\mathcal{R}_U \hat{\mathbf{Q}}_{ij} = \hat{\mathbf{Q}}_{ij}$, $\mathcal{P}_\sigma \hat{\mathbf{Q}}_{ij} = \hat{\mathbf{Q}}_{\sigma(i)\sigma(j)}$, and $\mathcal{P}_\sigma \hat{\mathbf{Q}}_{ij} = \hat{\mathbf{Q}}_{\sigma(i)\sigma(j)}$. Thus, for arbitrary indices i, j and k , the encoder functions $h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}^T)$ satisfy the following symmetry condition

$$\begin{aligned}\mathcal{T}_b h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) &= h_k((\mathcal{T}_b \hat{\mathbf{Q}}_{ij})^T \mathcal{T}_b \hat{\mathbf{Q}}_{ij}) = h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) \\ \mathcal{R}_U h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) &= h_k((\mathcal{R}_U \hat{\mathbf{Q}}_{ij})^T \mathcal{R}_U \hat{\mathbf{Q}}_{ij}) = h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) \\ \mathcal{P}_\sigma h_k(\hat{\mathbf{Q}}_{ij}^T \hat{\mathbf{Q}}_{ij}) &= h_k(\hat{\mathbf{Q}}_{\sigma(i)\sigma(j)}^T \hat{\mathbf{Q}}_{\sigma(i)\sigma(j)}).\end{aligned}\tag{2.22}$$

Plugging Eq. (2.22) into the definition of Ξ_{ij} yields (2.20). \square

2.6 Training

Eq. (2.17) departs from the common CG models by retaining both the heterogeneity and non-Markovianity of the energy dissipation process. Rather than matching the mean-field metrics such as the homogeneous VACF, we learn the embedded memory $\Xi(\mathbf{Q}(t))e^{\Lambda(t-s)}\Xi(\mathbf{Q}(s))^T$ based on the MZ form. However, directly solving the orthogonal dynamics $e^{\mathbf{Q}_Z \mathbf{L} t}$ is computationally intractable. Alternatively, we introduce the constrained dynamics $\tilde{\mathbf{z}}(t) = e^{\mathcal{R}t} \mathbf{z}(0)$ following Ref. Hijón et al. (2010). Based on the observation $\mathcal{P}\mathbf{Q} = \mathcal{P}\mathcal{R} \equiv 0$, we sample the MZ form from $\tilde{\mathbf{z}}(t)$, i.e., $\mathbf{K}_{MZ}(\mathbf{Z}, t) = \mathcal{P}_Z[(e^{\mathcal{R}t} \mathbf{Q}_Z \mathcal{L} \mathbf{P})(\mathbf{Q}_Z \mathcal{L} \mathbf{P})^T]$ and the memory of the CG model reduces to $\mathbf{K}_{CG}(\mathbf{Z}, t) = \Xi(\mathbf{Q})e^{\Lambda t}\Xi(\mathbf{Q})^T$. This enables us to train the CG models in terms of the encoders $\{f^i(\cdot), h^i(\cdot)\}_{i=1}^n$ and matrices $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}^a$ by minimizing the empirical loss

$$L = \sum_{l=1}^{N_s} \sum_{j=1}^{N_t} \left\| \mathbf{K}_{CG}(\mathbf{Z}^{(l)}, t_j) - \mathbf{K}_{MZ}(\mathbf{Z}^{(l)}, t_j) \right\|^2, \quad (2.23)$$

where l represents the different CG configurations.

2.7 Numerical Result

To demonstrate the accuracy of the present model, we consider a full micro-scale model of a star-shaped polymer melt system similar to Ref. Hijón et al. (2010), where each molecule consists of 73 atoms. The atomistic interactions are modeled by the Weeks-Chandler-Andersen potential and the Hookean bond potential. The full system consists of 486 molecules in a cubic domain $90 \times 90 \times 90$ with periodic boundary conditions. The Nosé-Hoover thermostat Nosé (1984); Hoover (1985) is employed to equilibrate the system with $k_B T = 4.0$ and micro-canonical ensemble simulation is conducted during the production stage. Below we compare different dynamic properties predicted by the full MD and the various CG models. For fair comparisons, we use the same CG potential $U(\mathbf{Q})$ constructed by the DeePCG scheme Zhang et al. (2018b) for all the CG models; the differences in dynamic properties solely arise from the different formulations of the memory term.

Let us start with the VACF which has been broadly used in CG model parameterization and validation. As shown in Fig. 2.2, the predictions from the present model (NM-MB) show good agreement with the full MD results. In contrast, the CG model with the memory term represented by

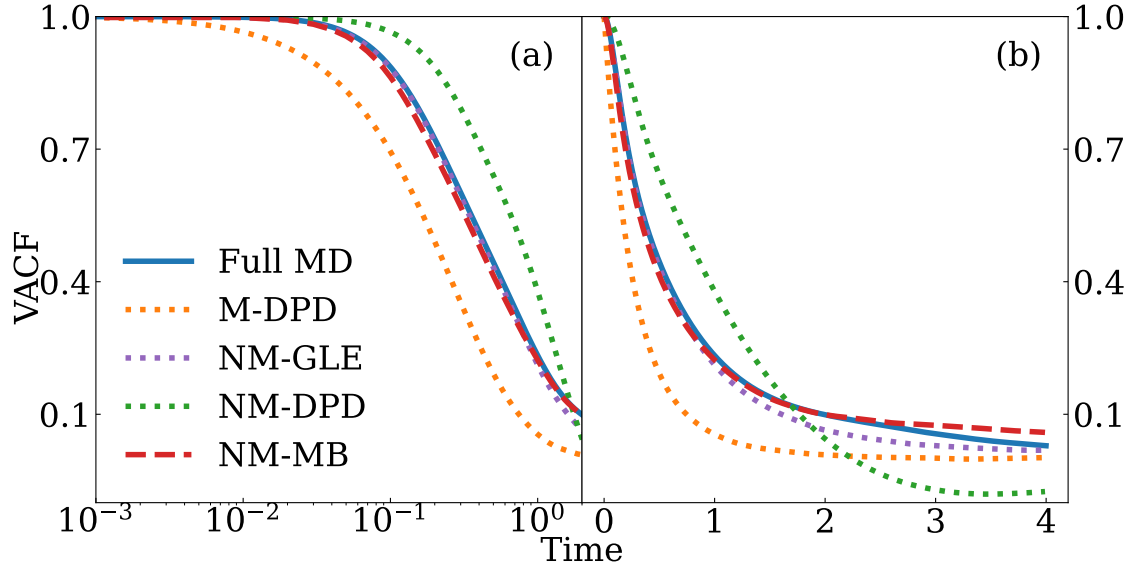


Figure 2.2 The VACF of the full MD and CG models with various memory formulations in (a) semi-log scale (b) original scale. “M” and “NM” represent Markovian and Non-Markovian; GLE, DPD, and MB represent state-independent, pairwise, and the present (NM-MB) model retaining the many-body effects, respectively.

the pairwise decomposition and Markovian approximation (i.e., the standard M-DPD form) yields apparent deviations. The form of the pairwise decomposition with non-Markovian approximation (NM-DPD) shows improvement at a short time scale but exhibits large deviations at an intermediate scale. Such limitations indicate pronounced many-body effects in the energy dissipation among the CG particles. Alternatively, if we set the VACF as the target quantity, we can parameterize the empirical model such as GLE by matching the VACF predicted by the full MD. Indeed, the prediction from the constructed GLE recovers the MD results. However, as shown below, this form over-simplifies the heterogeneity of the memory term and leads to inaccurate predictions on the collective scales.

Fig. 2.3 shows the velocity cross-correlation function (VCCF) between two CG particles, i.e., $C^{xx}(t; r_0) = \mathbb{E}[\mathbf{V}_i(0) \cdot \mathbf{V}_j(t) | Q_{ij}(0) = r_0]$, where r_0 represents the initial distance. Similar to VACF, the present model (NM-MB) yields good agreement with the full MD results. However, the predictions from other empirical models, including the GLE form, show apparent deviations. Such limitations arise from the inconsistent representation of the local energy dissipation and can be understood as following. The VACF represents the energy dissipation on each particle as a

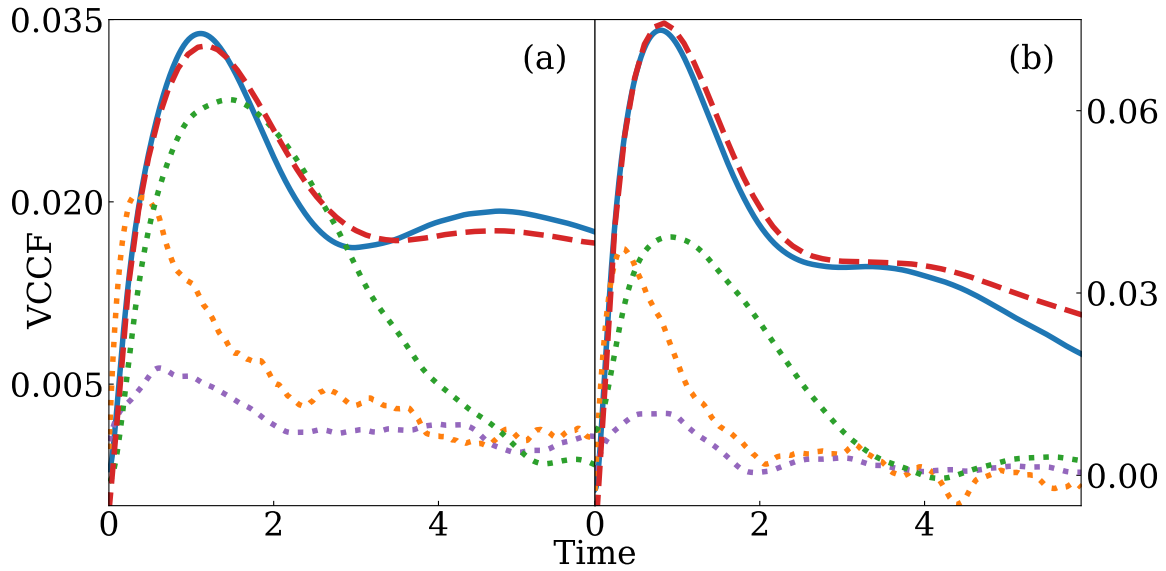


Figure 2.3 The VCCF $C^{xx}(t; r_0)$ predicted by the full MD and different CG models with initial distance (a) $10 < r_0 < 11$ and (b) $14 < r_0 < 15$. Same line legend as Fig. 2.2.

homogeneous background heat bath; it is essentially a mean-field metric and can not characterize the dissipative interactions among the particles. Hence, the reduced models that only recover the VACF could be insufficient to retain the consistent local momentum transport and the correlations among the particles.

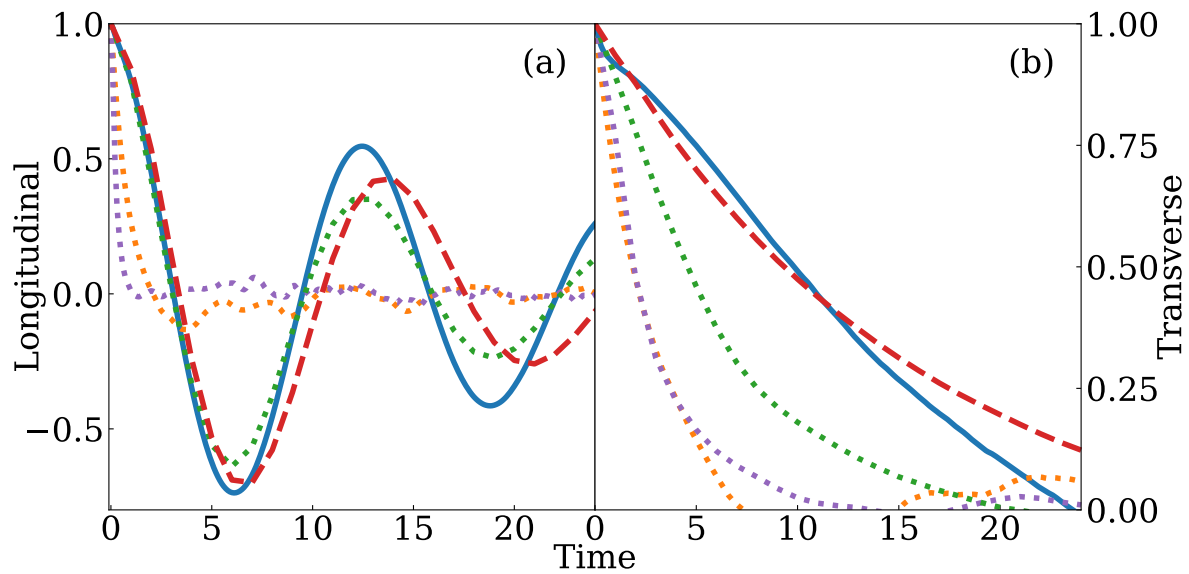


Figure 2.4 (a) Longitudinal and (b) Transverse hydrodynamic modes predicted by MD and different CG models. Same line legend as Fig. 2.2.

Furthermore, the various empirical models for local energy dissipations can lead to fundamentally different transport processes on the collective scale. Fig. 2.4 shows the normalized correlations of the longitudinal and transverse hydrodynamic modes Hansen and McDonald (1990), i.e., $C_L(t) = \langle \tilde{u}_1(t)\tilde{u}_1(0) \rangle$ and $C_T(t) = \langle \tilde{u}_2(t)\tilde{u}_2(0) \rangle$, where $\tilde{\mathbf{u}} = 1/M \sum_{j=1}^M \mathbf{V}_j e^{i\mathbf{k}\cdot\mathbf{Q}_j}$, \mathbf{k} is the wave vector, and the subscripts 1 and 2 represent the direction parallel and perpendicular to \mathbf{k} , respectively. Similar to the VCCF, the prediction from the present model (NM-MB) agrees well with the MD results while other models show apparent deviations. In particular, the prediction from the GLE model shows strong over-damping due to the ignorance of the inter-molecule dissipations.

Finally, we examine the diffusion process on the collective scale. Fig. 2.5 shows the van Hove function that characterizes the evolution of the inter-particle structural correlation defined by $G(r, t) \propto \frac{1}{M^2} \sum_{j \neq i}^M \delta(\|\mathbf{Q}_i(t) - \mathbf{Q}_j(0)\| - r)$. At $t = 0$, $G(r, t)$ reduces to the standard radial distribution function where all the CG models can recover such initial conditions. However, for $t > 0$, predictions from the models with the pairwise decomposition (NM-DPD) and the GLE form show apparent deviations. Specifically, at an early stage near $t = 50$, the neighboring particles begin to artificially jump into the region near the reference particle, violating the fluid-structure thereafter. In contrast, the present model (NM-MB) shows consistent predictions of the structure evolution over a long period until $t = 1000$, when the initial fluid structure ultimately diffuses into a homogeneous state.

2.8 Conclusion

To conclude, this work reveals a caveat in constructing reliable CGMD models that preserve the collective dynamics. Unlike the empirical forms, we developed a CG model that faithfully accounts for the broadly overlooked many-body nature of the non-Markovian memory term for extensive MD systems. While the significance of preserving the many-body nature of the conservative force field on static properties has been gradually recognized, the caveat on the memory term seems to remain under-explored. We show that retaining the heterogeneity and the strong correlation of the local energy dissipation is crucial for accurately predicting the cross-correlation among the CG particles, which, however, can not be fully characterized by the mean-field metrics such as VACF.

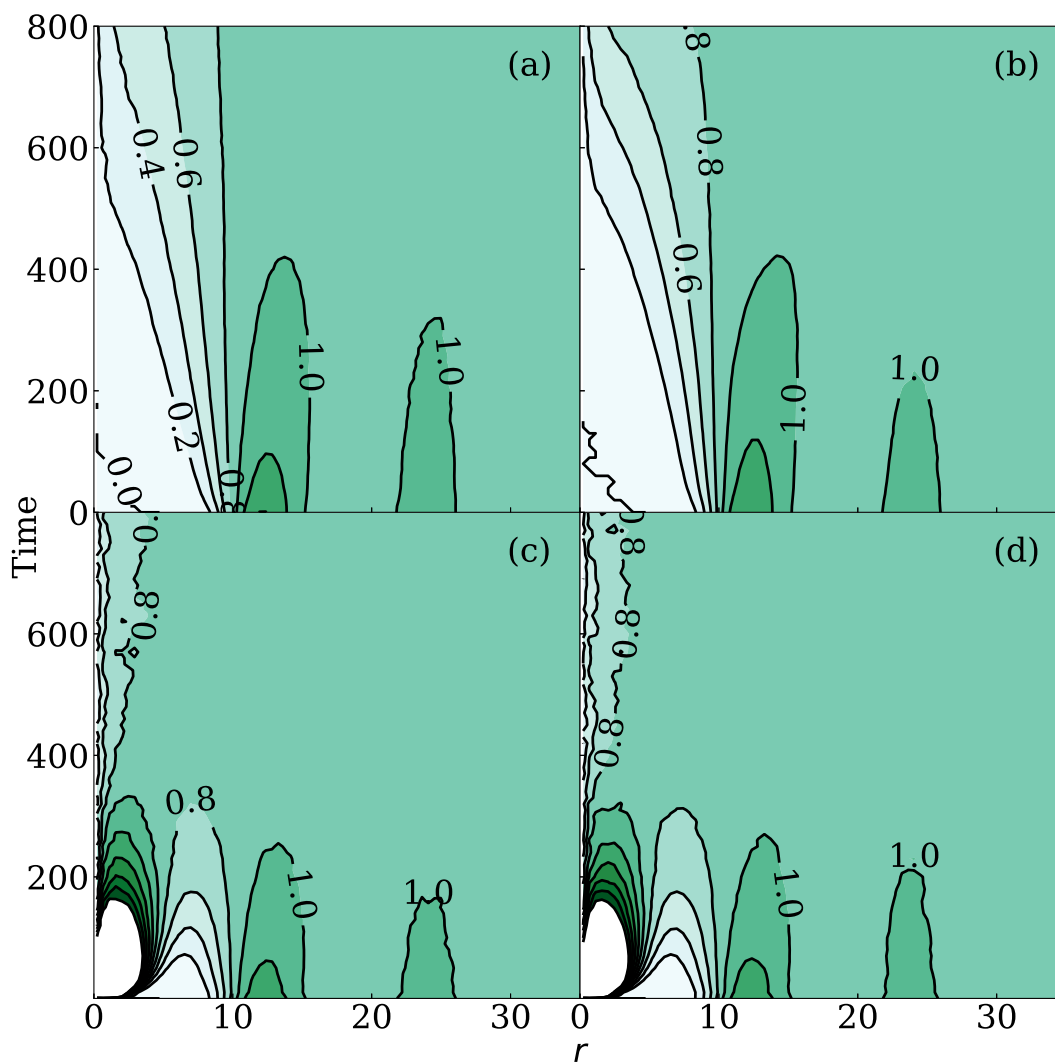


Figure 2.5 The van Hove function predicted by (a) full MD (b) the present NM-MB model (c) NM-DPD model (d) GLE model. It depicts the time evolution (y-axis) of the radius distribution function (x-axis).

More importantly, the memory form representing the inter-molecule energy dissipations may play a profound role in the transport and diffusion processes on the collective scale. In particular, the present model accurately predicts the hydrodynamic mode correlation and the van Hove function where empirical forms show limitations, and therefore, shows the promise to accurately predict the emergent phenomena relevant to hydrodynamic transport and diffusive processes on the collective scale.

CHAPTER 3

GENERALIZATION OF REDUCED MODELS FOR NON-EQUILIBRIUM DYNAMICS

3.1 Challenges in Modeling Non-Equilibrium Processes

Predictive modeling of the collective behaviors of multi-scale physical systems poses a persistent challenge for both fundamental science advancement and various applications Anderson (1972). While the canonical molecular dynamics enables us to faithfully account for the micro-scale interactions, numerical simulations often show limitations for systems without clear scale separation, where the computational cost could become prohibitive to reach the resolved scale of interest. This motivates the construction of various coarse-grained molecular dynamics (CGMD) models. By picking a set of collective variables (CVs), the CG models seek the reduced dynamics with less degrees of freedom to achieve a broader range of spatio-temporal scales. The constructed CG models are generally governed by a conservative CG potential representing the free energy of the resolved CVs and a memory term (along with a coherent noise) representing the energy dissipation that arises from the coupling with unresolved variables. Extensive research efforts Rosso et al. (2002); Maragliano and Vanden-Eijnden (2006); Izvekov and Voth (2005); Noid et al. (2008); Rudd and Broughton (1998); Lyubartsev and Laaksonen (1995); Shell (2008); Nielsen et al. (2004); Laio and Parrinello (2002); Darve and Pohorille (2001); Soper (1996); Reith et al. (2003); Nielsen et al. (2004); Das and Andersen (2012), including machine learning (ML) based approaches Behler and Parrinello (2007); Stecher et al. (2014a); John and Csányi (2017); Lemke and Peter (2017); Chmiela et al. (2017); Zhang et al. (2018b); Ge et al. (2023); van der Oord et al. (2023), have been dedicated to developing the CG potential to preserve the marginal density distribution of the CG variables and hence the various static properties. To retain the CG dynamics, the memory term needs to be properly introduced. Several approaches have been developed based on the direct approximations Darve et al. (2009); Lei et al. (2010); Hijón et al. (2010); Yoshimoto et al. (2013); Hudson and Li (2020) of the Mori-Zwanzig (MZ) projection formalism Mori (1965); Zwanzig (1973) and data-driven parameterization Lange and Grubmüller (2006); Ceriotti et al. (2009); Coifman et al. (2008); Crosskey and Maggioni (2017); Baczewski and Bond (2013); Ma et al. (2016); Lei et al.

(2016); Russo et al. (2019); Jung et al. (2017); Lee et al. (2019); Klippenstein and van der Vegt (2021); Vroylandt et al. (2022); She et al. (2023); Xie et al. (2022); Ge et al. (2024) of empirical forms (e.g., dissipative particle dynamics Hoogerbrugge and Koelman (1992); Español and Warren (1995), the generalized Langevin dynamics Zwanzig (2001)). A recent work Lyu and Lei (2023b) proposed a symmetry-preserving representation of the memory term that accounts for both non-Markovianity and the many-body nature among the CG variables, which proves to be crucial for diffusion and transport processes on a resolved scale.

Ideally, by accurately constructing the conservative free energy and the memory term, the CG models will enable us to quantify the propagation of the micro-scale interactions and therefore probe the collective behaviors across multiple scales. However, the validity and generalization ability for practical applications remains under-explored. In particular, for extensive MD systems (i.e., the number of molecules proportional to the system size), the CG variables are often chosen a priori such as the centers of mass (COMs) of individual molecules; the CG model is generally constructed such that certain dynamic properties (e.g, the mean square displacement, velocity correlation functions) under equilibrium can be properly reproduced. On the other hand, the model's applicability for non-equilibrium processes remains questionable. For instance, if we coarse grain a polymer melt system using the COMs of individual molecules, we should not expect the CG model can capture the visco-elastic responses arising from the molecule deformation under an external flow field. From a model reduction perspective, this limitation arises from the choice of the CG projection operator, which is generally defined with respect to the (marginal) equilibrium density of the full MD system. Accordingly, reduced dynamics can recover the full MD prediction only when the underlying distribution is close to equilibrium. However, this caveat seems to be broadly overlooked in existing CGMD models that use equilibrium dynamic properties as the metric, and therefore poses fundamental challenges for predicting the non-equilibrium processes in real applications.

In principle, we may directly construct the reduced model for a non-equilibrium process by defining a projection operator with respect to a specific probability measure. However, the reduced dynamics typically relies on a time-dependent projection operator (e.g., see Ref. Kawasaki and

Gunton (1973); Willis and Picard (1974); Baxevari et al. (2023)), which, as a result, generates a non-stationary memory term that may not be transferred to a different scenario. In this work, we present a new approach for constructing reliable CGMD models applicable to non-equilibrium processes. Rather than using pre-selected CG variables such as the COMs, we seek a set of auxiliary CG variables as the generalized coordinates of each molecule for the optimal representation of both the intra- and inter-molecular interactions. The key observation is that by systematically introducing these auxiliary CG variables, the conditional distribution of unresolved variables under various non-equilibrium conditions approaches that under equilibrium conditions. This warrants the generalization ability and enables us to transfer the empirical approximation of the non-equilibrium processes into the construction of the Zwanzig’s projection dynamics with respect to a probability density distribution with augmented conditional variables. To construct the reduced dynamics, we note that both the free energy and the memory term exhibit the many-body nature. We generalize the symmetry-preserving neural network representation of the state-dependent memory developed in our previous work Lyu and Lei (2023b) to model both the intra- and inter-energy dissipation. For each CG variable, we further introduce a number of non-Markovian features, allowing a coherent white noise term to be naturally imposed that satisfies the second fluctuation-dissipation theorem and preserves the invariant distribution. We demonstrate the proposed method by constructing the CGMD models of a polymer melt system. Numerical results show that the CG model based on the pre-selected COMs can recover the dynamic properties near equilibrium but is insufficient to predict the reduced dynamics under external flow conditions. Conversely, the present model guarantees that the unresolved orthogonal dynamics remains near equilibrium, ensuring the applicability of the projection formalism, and therefore, accurately predict non-equilibrium processes under various flow field conditions.

3.2 Methodology for Generalizable Coarse-Grained Models

3.2.1 Preliminaries

Let us consider a full MD system consisting of M molecules with a total number of N atoms. For simplicity, we assume each molecule consists of N_m atoms, and the mass is set to be unit. The

full model is governed by

$$\dot{z} = S\nabla H(z) \quad z(0) = z_0, \quad (3.1)$$

where $H(z)$ is the Hamiltonian, $\mathbf{z} = [\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^M, \mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^M]$ is the phase space vector and S is the symplectic matrix. Specifically, $\mathbf{q}^I = [\mathbf{q}_1^I, \dots, \mathbf{q}_{N_m}^I]$ and $\mathbf{p}^I = [\mathbf{p}_1^I, \dots, \mathbf{p}_{N_m}^I] \in \mathbb{R}^{N_m \times 3}$ represent the position and momentum vectors of the I -th molecule for $I = 1, 2, \dots, M$. In this work, the superscript indices in capital letter I and J ranging from 1 to M label the molecule index, while the subscript indices n and l ranging from 1 to N_m label the individual particles within the molecule. To construct the reduced model, we define the CG variables of each molecule $\mathbf{Z}^I = [\mathbf{Q}^I, \mathbf{P}^I]$ via a map $\phi : \mathbb{R}^{N_m \times 6} \rightarrow \mathbb{R}^{m \times 6}$, i.e.,

$$\mathbf{Q}^I = \phi^Q(\mathbf{q}^I) \quad \mathbf{P}^I = \phi^P(\mathbf{q}^I, \mathbf{p}^I), \quad (3.2)$$

where $\mathbf{Q}^I \in \mathbb{R}^{m \times 3}$ and $\mathbf{P}^I \in \mathbb{R}^{m \times 3}$ represent the generalized coordinates and momenta of the I -th molecule, respectively. Let $\mathbf{Q} = [Q_1, Q_2, \dots, Q_{mM}]$ and $\mathbf{P} = [P_1, P_2, \dots, P_{mM}]$ denote the CG variables of the full system. Furthermore, we use $\mathbf{Q}^I = [Q_{(I-1)m+1}, \dots, Q_{Im}]$ to denote the position vector of the I -th molecule, where $\mathbf{Q}_j^I = Q_\mu$ and $\mu = (I-1)m + j$, similarly for \mathbf{P}^I . In this work, the Greek letters α, β, μ, ν ranging from 1, \dots , mM label the global indices of CG coordinate.

In particular, one natural choice is to define $\mathbf{Q}^I \in \mathbb{R}^3$ as the COM and $\mathbf{P}^I \in \mathbb{R}^3$ the total momentum of the I -th molecule (e.g., see Refs. Lei et al. (2010); Hijón et al. (2010); Yoshimoto et al. (2013)), which essentially eliminates the intra-molecular DOFs. The dynamic evolution is governed by $\dot{\mathbf{Z}}(t) = \mathcal{L}\mathbf{Z}(t)$, where the Louville operator $\mathcal{L}\phi(z) := -((\nabla H(z_0))^T S \nabla z_0)\phi(z)$ depends on the full phase space vector z . To derive the reduced dynamics in terms of $\mathbf{Z}(t)$, we define the Zwanzig's projection operator as a condition expectation with respect to $\mathbf{Z}(0) = \mathbf{Z}$, i.e.,

$$\mathcal{P}_{\mathbf{Z}} f := \mathbb{E}[f(z) | \phi(z) = \mathbf{Z}] = \int \delta(\phi(\mathbf{z}) - \mathbf{Z}) \rho_0(\mathbf{z}) f(\mathbf{z}) d\mathbf{z} / \int \delta(\phi(\mathbf{z}) - \mathbf{Z}) \rho_0(\mathbf{z}) d\mathbf{z}, \quad (3.3)$$

where $\rho_0(z) \propto \exp[-\beta H(z)]$ represents the equilibrium Boltzmann distribution. Accordingly, we may project the evolution dynamics $\dot{\mathbf{Z}}(t) = \mathcal{L}\mathbf{Z}(t)$ on the sub-space of the CG variables. Using the Duhamel-Dyson formula, the reduced dynamics takes the form

$$\dot{\mathbf{Z}}(t) = e^{\mathcal{L}t} \mathcal{P}_{\mathbf{Z}} \mathcal{L} \mathbf{Z}(0) + \int_0^t ds e^{\mathcal{L}(t-s)} \mathcal{P}_{\mathbf{Z}} \mathcal{L} e^{Q_Z \mathcal{L} s} Q_Z \mathcal{L} \mathbf{Z}(0) + e^{Q_Z \mathcal{L} t} Q_Z \mathcal{L} \mathbf{Z}(0), \quad (3.4)$$

where $\mathbf{Q}_Z = \mathbf{I} - \mathcal{P}_Z$. With some further assumptions, the reduced dynamics can be simplified into the following integro-differential equation

$$\begin{aligned}\dot{\mathbf{Q}} &= \mathbf{M}^{-1} \mathbf{P} \\ \dot{\mathbf{P}} &= -\nabla U(\mathbf{Q}) + \int_0^t \mathbf{K}(\mathbf{Q}(s), t-s) \dot{\mathbf{Q}}(s) ds + \mathbf{R}(t),\end{aligned}\tag{3.5}$$

where \mathbf{M} is the mass matrix of the CG variables, $U(\mathbf{Q})$ is the conservative free energy, and $\mathbf{K}(\mathbf{Q}, t)$ is the memory term assumed to be independent of \mathbf{P} and $\mathbf{R}(t)$ is the fluctuation force.

In principle, with proper construction of the individual modeling terms, Eq. (3.5) provides an accurate CGMD model of the full dynamics (3.1). However, its applicability for practical applications, especially the generalization ability for modeling non-equilibrium processes remains questionable. This caveat is rooted in the definition of the projection operator in Eq. (3.3), where the conditional expectation is defined with respect to the marginal density of the equilibrium distribution $\rho_0(\mathbf{z})$. Accordingly, the derived reduced dynamics (3.5) assumes that the initial distribution of the full model satisfies $\rho(\mathbf{z}|\mathbf{Z}) \propto \delta(\phi(\mathbf{z}) - \mathbf{Z})\rho_0(\mathbf{z})$. Hence, Eq. (3.5) is valid only when the underlying distribution of the unresolved variables is close to equilibrium. Unfortunately, this condition is generally not guaranteed when we model non-equilibrium processes (e.g., full model (3.1) in the presence of an external field) using the CG variables like the COMs of individual molecules, which poses a fundamental challenge for the transferability of the reduced model (2.2) and (3.5) in real applications.

3.2.2 Construction of the auxiliary CG variables

Following the above discussion, let $H_e(\mathbf{z})$ and $\rho_e(\mathbf{z})$ denote the Hamiltonian and an initial (e.g., the steady-state) distribution under an external field. Theoretically, we may define the Zwanzig's projection operator with respect to $\rho_e(\mathbf{z})$ and derive the reduced dynamics under this external field. However, the constructed reduced dynamics could be valid only for this specific condition but lack the generalization ability for other external field conditions.

To address this issue, we transfer the exhausting effort of constructing specific reduced models for individual non-equilibrium processes to pursuing the following question: how to ensure the conditional projection operator defined by Eq. (3.3) remains valid for a range of external conditions?

A key observation is that by introducing a set of auxiliary CG variables $\mathbf{Z}^I = [\mathbf{Z}_1^I, \mathbf{Z}_2^I, \dots, \mathbf{Z}_m^I]$ for the individual molecules $I = 1, 2, \dots, M$, the conditional probability density function (PDF) under external field approaches that under the equilibrium distribution, i.e.,

$$\frac{1}{\mathcal{Z}_e} \prod_{I=1}^M \prod_{j=1}^m \delta(\phi(\mathbf{z}^I)_j - \mathbf{Z}_j^I) \rho_e(\mathbf{z}) \approx \frac{1}{\mathcal{Z}_0} \prod_{I=1}^M \prod_{j=1}^m \delta(\phi(\mathbf{z}^I)_j - \mathbf{Z}_j^I) \rho_0(\mathbf{z}). \quad (3.6)$$

Intuitively, this approximation can be understood as follows: As we increase the number of resolved CG variables, the entropy contribution of the unresolved ones decreases. Accordingly, the free energy of different external conditions approaches the equilibrium case, and the reduced model (3.4) and (3.5) in the form of \mathbf{Z}^I can be applied to non-equilibrium processes. This is somewhat similar to the work Lei et al. (2020); Yu et al. (2005) on modeling the constitutive closure of polymer solution, where a set of generalized conformation tensors are introduced to represent the subgrid polymer configurations under various flow conditions.

To construct the auxiliary CG variables, we seek a linear map $\phi(\cdot)$ in the form of a matrix $\mathbf{fW} \in \mathbb{R}^{N_m \times m}$, i.e.,

$$\phi^Q(\mathbf{q}^I)_j = \frac{\sum_{n=1}^{N_m} w_{nj} \mathbf{q}_n^I}{\sum_{n=1}^{N_m} w_{nj}}, \quad \phi^P(\mathbf{p}^I)_j = \sum_{i=1}^m M_{ji} \frac{\sum_{n=1}^{N_m} w_{ni} \mathbf{p}_n^I}{\sum_{n=1}^{N_m} w_{ni}}, \quad (3.7)$$

where the mass matrix is defined as $M_{ji} = \sum_{n=1}^{N_m} w_{nj} w_{ni}$, and the normalization ensures that the transitional and rotational symmetry of the CG variables, i.e.,

$$\begin{aligned} \phi(\mathbf{q}^I + \mathbf{c}, \mathbf{p}^I) &= [\phi^Q(\mathbf{q}^I + \mathbf{c}), \phi^P(\mathbf{p}^I)] = [\phi^Q(\mathbf{q}^I) + \mathbf{c}, \phi^P(\mathbf{p}^I)] \\ \phi(\mathbf{q}^I \mathcal{U}, \mathbf{p}^I \mathcal{U}) &= [\phi^Q(\mathbf{q}^I \mathcal{U}), \phi^P(\mathbf{p}^I \mathcal{U})] = [\phi^Q(\mathbf{q}^I) \mathcal{U}, \phi^P(\mathbf{p}^I) \mathcal{U}], \end{aligned}$$

for any $\mathbf{c} \in \mathbb{R}^3$ and $\mathcal{U} \in SO(3)$. As shown below, Eq. (3.7) can be loosely viewed as seeking the principal normal modes of the full dynamics. We note that it is possible to construct $\phi(\cdot)$ as a non-linear encoder to seek optimal CG representations of the inter- and intra- molecular interactions. In this work, we focus on developing a general CG framework of introducing auxiliary CG variables to enhance the generalization ability and examining Eq. (3.6) as a valid metric for modeling

non-equilibrium processes. We restrict $\phi(\cdot)$ to be linear; the model reduction in terms of nonlinear CG variables will be investigated in future work.

Furthermore, we impose the following two constraints to the weight matrix \mathbf{W} , i.e.,

$$0 < w_{ni} < 1, \quad \sum_{i=1}^m w_{ni} = 1 \quad \text{for all } n = 1, \dots, N_m. \quad (3.8)$$

We emphasize that these constraints do not impose further restrictions on constructing the CG variables. Specifically, for any mapping defined by a matrix $\mathbf{W}_1 \in \mathbb{R}^{N_m \times (m-1)}$, there exists an equivalent mapping defined by a matrix $\mathbf{W}_2 \in \mathbb{R}^{N_m \times m}$ which satisfies the above constraints (3.8). We refer to the Appendix for the detailed discussion. Furthermore, we note that the second constraint ensures that the COM of the full molecule is consistent with that of the CG variables, i.e., $\sum_{j=1}^m \left(\sum_{n=1}^{N_m} w_{nj} \right) \mathbf{Q}_j^I = \sum_{j=1}^m \sum_{n=1}^{N_m} w_{nj} \mathbf{q}_n^I = \sum_{n=1}^{N_m} \mathbf{q}_n^I$, which enables us to establish a fair comparison of the different CG models on the same metric in Sec. 3.3. Moreover, Newton's third law can be naturally imposed in Eq. (3.20).

To learn CG mapping $\phi(\cdot)$, we aim to find a weight matrix \mathbf{W} such that condition (3.6) holds, which, unfortunately, cannot be easily checked a priori. Alternatively, we propose to learn the CG variables with the longest relaxation of the time correlation function. Specifically, we collect training samples under an external field and seek \mathbf{W} by solving the optimization problem

$$\max_{\mathbf{W}} \frac{\sum_{\tau} \sum_{t=0}^{t_c} \sum_{I,j}^{M,m} \mathbf{V}_j^I(\tau) \mathbf{V}_j^I(\tau+t)^T}{\sum_{\tau} \sum_{I,j}^{M,m} \mathbf{V}_j^I(\tau) \mathbf{V}_j^I(\tau)^T} \quad (3.9)$$

under constraint (3.8), where $\mathbf{V}_j^I = \frac{\sum_{n=1}^{N_m} w_{nj} \mathbf{p}_n^I}{\sum_{n=1}^{N_m} w_{nj}}$, t_c is a hyperparameter representing the cut-off time when the correlation function decays to 0 and τ is randomly selected from a long MD trajectory under the steady state. Intuitively, Eq. (3.9) yields the eigenmodes of the largest (negative) eigenvalue of the operator \mathcal{L} under the linear approximation Crommelin and Vanden-Eijnden (2011). We note that this is somewhat similar to the time-lagged independent component analysis (TICA) Pérez-Hernández et al. (2013); Molgedey and Schuster (1994); Schwantes and Pande (2013) and dynamic mode decomposition (DMD) Schmid (2010); Chen et al. (2012); Kutz et al. (2016) to identify the dominant dynamics from the simulation data. Furthermore, we emphasize that Eq. (3.9)

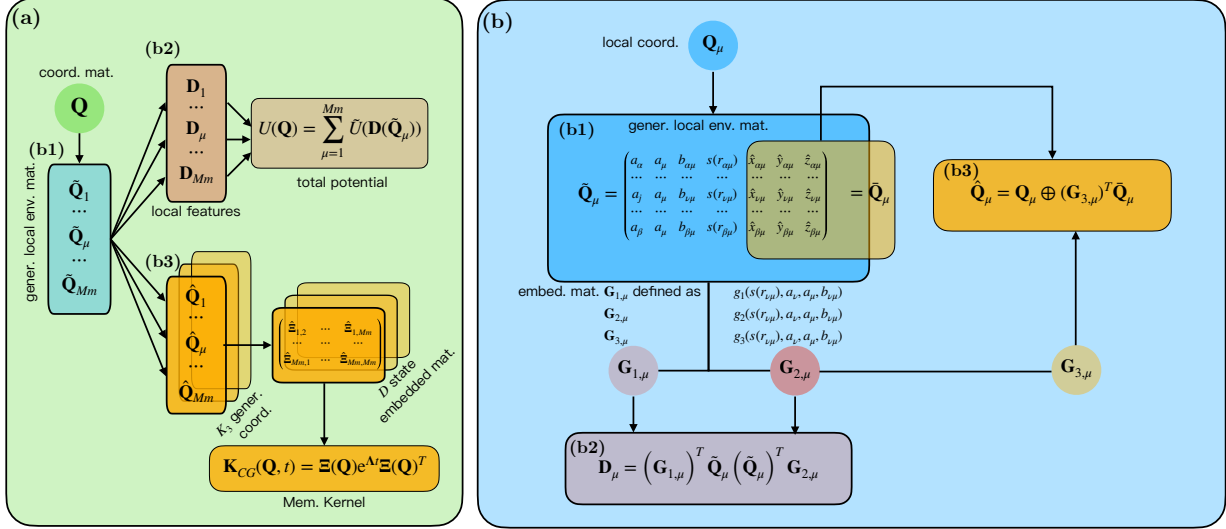


Figure 3.1 Diagram illustrating the neural network architecture for the CG conservative force and the memory kernel. (a) The construction of CG potential function U and memory kernel \mathbf{K} . Initially, \mathbf{Q} is converted into a local environment matrix $\{\tilde{\mathbf{Q}}_\mu\}_{\mu=1}^{Mm}$. Sub-networks, illustrated in (b), map $\tilde{\mathbf{Q}}_\mu$ to a local feature \mathbf{D}_μ and generalized coordinate $\hat{\mathbf{Q}}_\mu$. Finally, the total potential is constructed by Eq. (3.10), i.e. $U = \sum_{\mu=1}^{Mm} \tilde{U}(\mathbf{D}_\mu)$. The total memory kernel \mathbf{K} is constructed with the state-dependent component of the memory kernel derived from $\hat{\mathbf{Q}}_\mu$ using Eq. (2.16) and the time-dependent component Λ . (b) The sub-networks map $\tilde{\mathbf{Q}}_\mu$ to a local feature \mathbf{D}_μ and generalized coordinate $\hat{\mathbf{Q}}_\mu$. The neighbour of the μ -th CG coordinate is denoted by $\mathcal{N}_\mu = \{\alpha, \dots, \nu, \dots, \beta\}$. (b1) The k -th row of generalized local environment matrix embeds the relative information between μ -th coordinate and its k -th neighbor (labeled as the ν -th coordinate), including type embedding $a_\nu, a_\mu, b_{\nu\mu}$, and distance information. (b2) The $K_1 \times K_2$ symmetry preserving feature \mathbf{D}_μ is constructed by $(\mathbf{G}_{1,\mu})^T, \hat{\mathbf{Q}}_\mu, \hat{\mathbf{Q}}_\mu^T$ and $\mathbf{G}_{2,\mu}$. (b3) The generalized coordinate is constructed from the local environment embedding matrix $\mathbf{G}_{3,\mu}$ and relative position $\tilde{\mathbf{Q}}_\mu$, which is the last three columns of $\tilde{\mathbf{Q}}_\mu$.

should not be viewed as the equivalent condition of Eq. (3.6). Rather, it is essentially an empirical approach for learning CG mappings such that the metric (3.6) satisfies approximately.

3.2.3 Construction of the symmetry-preserving free energy function

With the CG mapping ϕ , the construction of the CG model proceeds with learning the free energy $U(\mathbf{Q})$ and the memory kernel $\mathbf{K}(\mathbf{Q}, t)$. To retain the extensive structure, the free energy $U(\mathbf{Q})$ is decomposed into the local potential of individual CG coordinates, i.e.,

$$U(\mathbf{Q}) = \sum_{\mu=1}^{Mm} \tilde{U}(\mathbf{D}(\tilde{\mathbf{Q}}_\mu)), \quad (3.10)$$

where $\tilde{U}(\cdot)$ represents the local potential, $\tilde{\mathbf{Q}}_\mu$ represents the local environment determined by the relative position between the CG coordinates \mathbf{Q}_μ and its neighbors \mathbf{Q}_ν in $\{\nu \mid \nu \in \mathcal{N}_\mu\}$. The detailed

form will be specified following Eq. (3.12). $D(\cdot)$ represents the encoder functions that map \tilde{Q}_μ into symmetry-preserving features. \mathcal{N}_μ denotes the CG coordinate indices ν in the neighbour of CG coordinate indices μ , such that $r_{\mu\nu} < r_c$. We define N_μ as the cardinality of the set \mathcal{N}_μ .

Fig. 3.1 sketches the neural network representation of $U(\mathbf{Q})$, where a structure similar to Ref. Zhang et al. (2018c) is used to preserve the translational and rotational symmetry constraints. To impose the permutational symmetry, we note that $U(\mathbf{Q})$ remains invariant with the index permutation among the individual molecule, i.e.,

$$U(\mathbf{Q}^{\sigma(1)}, \mathbf{Q}^{\sigma(2)}, \dots, \mathbf{Q}^{\sigma(M)}) = U(\mathbf{Q}^1, \mathbf{Q}^2, \dots, \mathbf{Q}^M), \quad (3.11)$$

where $\{\sigma(I)\}_{I=1}^M$ represents an index permutation among moleculars. However, $U(\mathbf{Q})$ does not necessarily remain invariant under an index permutation among the generalized CG coordinates, i.e.,

$$\begin{aligned} U(\mathbf{Q}_1^1, \mathbf{Q}_2^1, \dots, \mathbf{Q}_1^2, \mathbf{Q}_2^2 \dots) &\neq U(\mathbf{Q}_1^1, \mathbf{Q}_2^2, \dots, \mathbf{Q}_1^2, \mathbf{Q}_2^1 \dots), \\ U(\mathbf{Q}_1^1, \mathbf{Q}_2^1, \dots, \mathbf{Q}_1^2, \mathbf{Q}_2^2 \dots) &\neq U(\mathbf{Q}_2^1, \mathbf{Q}_1^1, \dots, \mathbf{Q}_1^2, \mathbf{Q}_2^1 \dots), \end{aligned} \quad (3.12)$$

where $\mathbf{Q}_i^I = \mathbf{Q}_{(I-1)m+i}$. This complication can not be remedied by introducing the particle type information as done in Ref. Zhang et al. (2018c). This is because, for a specific CG coordinate, its contribution to $U(\mathbf{Q})$ could be either inter- or intra-molecular interaction associated with other CG coordinates. Therefore, we define the local environment of the μ -th coordinate by $\tilde{Q}_\mu = (a_\nu, a_\mu, b_{\nu\mu}, s(r_{\nu\mu}), \hat{x}_{\nu\mu}, \hat{y}_{\nu\mu}, \hat{z}_{\nu\mu}) \in \mathbb{R}^{N_\mu \times 7}$, where relative position $\mathbf{r}_{\nu\mu} = \mathbf{Q}_\nu - \mathbf{Q}_\mu = (x_{\nu\mu}, y_{\nu\mu}, z_{\nu\mu})$ denotes the relative position to its neighbor, $\hat{x}_{\nu\mu} = s(r_{\nu\mu})x_{\nu\mu}/r_{\nu\mu}$, and similar for $\hat{y}_{\nu\mu}$ and $\hat{z}_{\nu\mu}$. $s(r)$ is a preselected function defined by

$$s(r) = \begin{cases} \frac{r}{r_c^2} & \text{if } I = J, \\ \frac{1}{r} & \text{if } r < r_{cs} \text{ and } I \neq J, \\ \frac{1}{2r} \left[1 + \cos \left(\frac{\pi(r-r_{cs})}{r_c-r_{cs}} \right) \right] & \text{if } r_{cs} \leq r < r_c \text{ and } I \neq J, \\ 0 & \text{else,} \end{cases}$$

where I and J represent the molecule indices of the μ -th and its neighbor coordinate. The motivation of this definition is the interactions between molecules are intended to steadily decline to zero as

the distance grows, whereas intra-molecular interactions are expected to intensify with increased distance. In particular, if the two coordinates belong to the same molecule, their intra-molecular interactions will always be accounted for. Otherwise, $s(r)$ is a smooth differentiable function that decays to 0 beyond r_c . r_{cs} is a hyper-parameter and is set to be $r_c - 1$ in this study. The a_ν, a_μ and $b_{\nu\mu}$ are parameters to represent the type of atom and the type of interaction. Specifically, $b_{\nu\mu}$ is set to be 1 if the μ -th coordinate and its neighbor ν -th coordinate are part of the same molecule I , which means that there exist $i, j \in \{1, \dots, m\}$ such that $\mu = (I - 1)m + i$ and $\nu = (I - 1)m + j$. Otherwise, it is set to be -1 . In addition, we also introduce a_ν, a_μ to represent the type of the μ -th CG coordinate and the type of the ν -th CG coordinate within the neighborhood. In practice, a_μ, a_ν are set to be $i, j \in \{1, \dots, m\}$ defined in $b_{\nu\mu}$.

With the local environment $\tilde{\mathbf{Q}}_\mu$, we construct the local features $\mathbf{D} \in \mathbb{R}^{K_1 \times K_2}$ by

$$\begin{aligned} \mathbf{D}(\tilde{\mathbf{Q}}_\mu) &= (\mathbf{G}_{1,\mu})^T \tilde{\mathbf{Q}}_\mu (\tilde{\mathbf{Q}}_\mu)^T \mathbf{G}_{2,\mu} \\ &:= \left(\sum_{\nu \in \mathcal{N}_\mu} \mathbf{g}_1(s(r_{\nu\mu}), a_\nu, a_\mu, b_{\nu\mu}) \tilde{\mathbf{Q}}_{\nu\mu} \right) \left(\sum_{\nu \in \mathcal{N}_\mu} \mathbf{g}_2(s(r_{\nu\mu}), a_\nu, a_\mu, b_{\nu\mu}) \tilde{\mathbf{Q}}_{\nu\mu} \right)^T, \end{aligned} \quad (3.13)$$

where $\mathbf{G}_{1,\mu} \in \mathbb{R}^{N_\mu \times K_1}$, $\mathbf{G}_{2,\mu} \in \mathbb{R}^{N_\mu \times K_2}$. The k -th row of $\mathbf{G}_{1,\mu}$ is represented as a neural network $\mathbf{g}_1(s(r_{\nu\mu}), a_\nu, a_\mu, b_{\nu\mu})$ that embeds the relative information between the μ -th coordinate and its k -th neighbor (labeled as the ν -th coordinate) in \mathbb{R}^4 to a feature space in \mathbb{R}^{K_1} . Similarly, $\mathbf{G}_{2,\mu} = \mathbf{g}_2(s(r_{\nu\mu}), a_\nu, a_\mu, b_{\nu\mu})$ is represented by a neural network mapping \mathbb{R}^4 to \mathbb{R}^{K_2} .

To construct the free energy function $U(\mathbf{Q})$, we conduct restrained MD simulations (see Sec. 3.2.4 for details) and train the neural network representations $\tilde{U}(\cdot)$, $\mathbf{g}_1(\cdot)$ and $\mathbf{g}_2(\cdot)$ by minimizing the empirical loss between the MD and CG model, i.e.,

$$L_U = \sum_{s=1}^S \left\| \nabla U(\mathbf{Q}^{(s)}) + \mathcal{F}(\mathbf{Q}^{(s)}) \right\|^2, \quad (3.14)$$

where the superscript s represents various configurations and the \mathcal{F} represents the conservative force term sampled from the full MD simulation; see Sec. 3.2.5 for details.

3.2.4 Construction of the symmetry-preserving memory function

Besides the free energy function $U(\mathbf{Q})$, the reduced dynamics further depends on the memory function $\mathbf{K}(\mathbf{Q}, t)$. In particular, recent work Lyu and Lei (2023b) shows that the memory function could exhibit a strong many-body nature. Empirical approximations such as the standard GLE with a homogeneous kernel and the dissipative particle dynamics with a pairwise decomposition generally show limitations in modeling the heterogeneous energy dissipation among the CG particles. To predict the collective dynamics, it is crucial to accurately model both the non-Markovian and many-body nature of the memory function.

Following the neural network structure proposed in Ref. Lyu and Lei (2023b), we encode the many-body dissipative interactions among the CG particles by introducing coordinate features $\widehat{\mathbf{Q}}_\mu = [\widehat{\mathbf{Q}}_\mu^1, \widehat{\mathbf{Q}}_\mu^2, \dots, \widehat{\mathbf{Q}}_\mu^{K_3}] \in \mathbb{R}^{K_3 \times 3}$, where the feature $\widehat{\mathbf{Q}}_\mu$ is defined by

$$\begin{aligned}\widehat{\mathbf{Q}}_\mu &= \mathbf{Q}_\mu \oplus (\mathbf{G}_{3,\mu})^T \bar{\mathbf{Q}}_\mu \\ &:= \mathbf{Q}_\mu \oplus \sum_{\nu \in \mathcal{N}_\mu} \mathbf{g}_3(s(r_{\nu\mu}), a_\nu, a_\mu, b_{\nu\mu}) \bar{\mathbf{Q}}_{\nu\mu},\end{aligned}\tag{3.15}$$

where $\mathbf{G}_{3,\mu} \in \mathbb{R}^{N_\mu \times K_3}$, where $\mathbf{g}_3 : \mathbb{R}^4 \rightarrow \mathbb{R}^{K_3}$ is a neural network that encodes the dissipative interactions beyond the pairwise form. $\bar{\mathbf{Q}}_\mu \in \mathbb{R}^{N_\mu \times 3}$ denotes the relative position between the μ -th coordinate and its local neighbor, whose k -th row (corresponding to ν -th coordinate) is defined as $\bar{\mathbf{Q}}_{\nu\mu} := (\hat{x}_{\nu\mu}, \hat{y}_{\nu\mu}, \hat{z}_{\nu\mu})$. Accordingly, we can construct the state-dependent memory in the form of $\Xi(\mathbf{Q}(t))e^{\Lambda(t-s)}\Xi(\mathbf{Q}(s))^T$, with $\Xi = [\Xi^1 \Xi^2 \dots \Xi^D]$ and

$$\Xi_{\mu\nu}^d = \sum_{k=1}^{K_3} h_k^d(\widehat{\mathbf{Q}}_{\mu\nu} \widehat{\mathbf{Q}}_{\mu\nu}^T) \widehat{\mathbf{Q}}_{\mu\nu}^k \otimes \widehat{\mathbf{Q}}_{\mu\nu}^k + h_0^d(\widehat{\mathbf{Q}}_{\mu\nu} \widehat{\mathbf{Q}}_{\mu\nu}^T) \mathbf{I},\tag{3.16}$$

where $h : \mathbb{R}^{K_3 \times K_3} \rightarrow \mathbb{R}^{D(K_3+1)}$ and $d = 1, \dots, D$. Λ takes an extendable form $\Lambda = \widehat{\Lambda} \otimes \mathbf{I}$, where $\mathbf{I} \in \mathbb{R}^{3M \times 3M}$ is the identity matrix and $\widehat{\Lambda} \in \mathbb{R}^{D \times D}$ specifies the coupling among the features. It takes the form $\widehat{\Lambda} = \widehat{\mathbf{L}} \widehat{\mathbf{L}}^T + \widehat{\mathbf{L}}^a$, where $\widehat{\mathbf{L}}$ is a lower triangular matrix and $\widehat{\mathbf{L}}^a$ is an anti-symmetry matrix; it satisfies the Lyapunov stability condition $\widehat{\Lambda} + \widehat{\Lambda}^T \geq 0$.

We can show that Ξ in the form of Eq. (3.16) strictly preserve the translational, rotational, and

permutational symmetry constraints, i.e.,

$$\begin{aligned}
\Xi_{\mu\nu}^d(Q_1 + \mathbf{b}, \dots, Q_{mM} + \mathbf{b}) &= \Xi_{\mu\nu}^d(Q_1, \dots, Q_{mM}) \\
\Xi_{\mu\nu}^d(Q_1 \mathbf{U}, \dots, Q_{mM} \mathbf{U}) &= \mathbf{U} \Xi_{\mu\nu}^d(Q_1, \dots, Q_{mM}) \mathbf{U}^T \\
\Xi_{\mu\nu}^d(Q^{\sigma(1)}, Q^{\sigma(2)}, \dots, Q^{\sigma(M)}) &= \Xi_{\sigma(\mu)\sigma(\nu)}^d(Q^1, Q^2, \dots, Q^M)
\end{aligned} \tag{3.17}$$

for any $\mathbf{c} \in \mathbb{R}^3$, $\mathbf{U} \in SO(3)$, and index permutation $\{\sigma(I)\}_{I=1}^M$ between the molecules, where $\mathbf{Q}^I = [\mathbf{Q}_{(I-1)m+1}, \dots, \mathbf{Q}_{Im}]$.

To construct $\mathbf{K}(\mathbf{Q}, t)$, we learn the memory embedded in Eq. (3.21) in the form of $\Xi(\mathbf{Q}(t))e^{\Lambda(t-s)}\Xi(\mathbf{Q}(s))^T$ directly from the Zwanzig's formalism so that the many-body nature can be naturally inherited. We estimate the memory term $\mathbf{K}_{MZ}(\mathbf{Q}, t)$ for a collection of CG configurations. This enables us to train the memory function in terms of the encoders $\{g_3(\cdot), h\}$ and matrices $\{\hat{\mathbf{L}}, \hat{\mathbf{L}}^a\}$ by minimizing the empirical loss

$$L_K = \sum_{s=1}^{N_s} \sum_{j=1}^{N_t} \left\| \mathbf{K}_{CG}(\mathbf{Q}^{(s)}, t_j) - \mathbf{K}_{MZ}(\mathbf{Q}^{(s)}, t_j) \right\|^2, \tag{3.18}$$

where s represents the different CG configurations.

3.2.5 Training and evaluation of the CGMD model

Following Ref. Hijón et al. (2010), we use the restrained dynamics $\tilde{\mathbf{z}}(t) = e^{\mathcal{R}t} \mathbf{z}(0)$ to approximate the orthogonal dynamics $e^{\mathcal{Q}ZLt}$ based on the observation $\mathcal{P}\mathcal{Q} = \mathcal{P}\mathcal{R} \equiv 0$. Accordingly, the conservative term can be approximated by $\mathcal{F}(\mathbf{Q}) = \mathcal{P}_Z \mathcal{L} \mathbf{P}$ and the memory term of Zwanzig's form (see Eq. (3.4)) can be approximated by $\mathbf{K}_{MZ}(\mathbf{Q}, t) = \mathcal{P}_Z \left[(e^{\mathcal{R}t} \mathcal{Q}_Z \mathcal{L} \mathbf{P})(\mathcal{Q}_Z \mathcal{L} \mathbf{P})^T \right]$ and the memory of the CG model reduces to $\mathbf{K}_{CG}(\mathbf{Q}, t) = \Xi(\mathbf{Q})e^{\Lambda t}\Xi(\mathbf{Q})^T$. To collect the training samples, we establish the restrained dynamics following

$$\begin{aligned}
\dot{\mathbf{q}}^I &= \mathbf{p}^I - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{p}^I \\
\dot{\mathbf{p}}^I &= \mathbf{f}^I - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{f}^I
\end{aligned} \tag{3.19}$$

where $\mathbf{f}^I = [\mathbf{f}_1^I, \dots, \mathbf{f}_{N_m}^I] \in \mathbb{R}^{N_m \times 3}$ is the force of the individual atoms belonging to the molecule I . The total force of each CG variable is defined as $\mathbf{F}_j^I = \sum_{n=1}^{N_m} w_{nj} \mathbf{f}_n^I$, which is equivalently denoted by \mathbf{F}_μ with $\mu = (I-1)m + j$ and $\mu = 1, \dots, mM$.

Under the restrained dynamics (3.19), the orthogonal fluctuation term corresponds to the random force of each CG variable, i.e., $\mathbf{Q}_Z \mathcal{L} \mathbf{P} = [\delta \mathbf{F}_1, \delta \mathbf{F}_2, \dots, \delta \mathbf{F}_{mM}]$, where $\delta \mathbf{F}_\mu = \mathbf{F}_\mu - \langle \mathbf{F}_\mu \rangle$ represents the random force of the μ -th CG coordinate and $\langle \cdot \rangle$ represents the condition expectation whose generalization ability will be examined in Sec. 3.3. Due to the constraint (3.8), the Newton's third law is naturally satisfied, i.e.,

$$\sum_{\mu} {}^{mM} \mathbf{F}_{\mu} = \sum_{I,i} {}^{M,m} \sum_n {}^{N_m} w_{ni} \mathbf{f}_n^I = \sum_{I,i} {}^{M,m} \sum_n {}^{N_m} w_{ni} \sum_{J,l} {}^{M,N_m} \mathbf{f}_{nl}^{IJ} = \sum_{I,n} {}^{M,N_m} \sum_{J,l} {}^{M,N_m} \mathbf{f}_{nl}^{IJ} \equiv 0, \quad (3.20)$$

where \mathbf{f}_{nl}^{IJ} represents the force between the n -th atom of the I -th molecule and the l -th atom of the J -th molecule. Therefore, $\delta \mathbf{F}_{\mu} = -\sum_{\nu \neq \mu} \delta \mathbf{F}_{\nu}$ and $\Xi_{\mu\mu} = -\sum_{\nu \neq \mu} \Xi_{\mu\nu}$.

The constructed reduced model can be simulated by the extended dynamics of $[\mathbf{Q}, \mathbf{P}, \zeta]$ in the form of

$$\begin{aligned} \dot{\mathbf{Q}} &= \mathbf{M}^{-1} \mathbf{P} \\ \dot{\mathbf{P}} &= -\nabla U(\mathbf{Q}) + \Xi(\mathbf{Q}) \zeta \\ \dot{\zeta} &= -\Xi(\mathbf{Q})^T \dot{\mathbf{Q}} - \Lambda \zeta + \xi(t), \end{aligned} \quad (3.21)$$

where we introduce the white noise term following $\langle \xi(t) \xi(t') \rangle = \beta^{-1} (\Lambda + \Lambda^T) \delta(t - t')$. With this choice, we can show that Eq. (3.21) satisfies the second fluctuation-dissipation theorem. We refer to Ref. Lyu and Lei (2023b) for the detailed proof.

3.3 Numerical Results

3.3.1 The full atomistic model

To validate our method, we consider a full micro-scale model of a polymer melt system. Each polymer molecule consists of $N_m = 106$ atoms. The central atom is connected to 3 arms and each arm consists of 5 atoms. At the end of each arm, there are 6 additional arms connected to the end atom and each additional arm consists of 5 atoms. The pairwise atomistic interaction is chosen to be the Weeks-Chandler-Andersen potential

$$V_p(r) = \begin{cases} V_{LJ}(r) - V_{LJ}(R_c), & r < R_c \\ 0, & r \geq R_c \end{cases} \quad V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (3.22)$$

where $\epsilon = 1.0$ is the dispersion energy, $R_c = 2^{1/6}\sigma$ and $\sigma = 2.415$ is the hardcore distance. The bond interaction is chosen to be the FENE potential

$$V_b(r) = -0.5KR_0^2 \ln \left[1 - \left(\frac{r}{R_0} \right)^2 \right], \quad (3.23)$$

where stiffness is set to be $K = 1$ and the distance is to be $R_0 = 5$. The complete system consists of $M = 300$ molecules in a $90 \times 90 \times 90$ domain in the reduced unit. The periodic boundary condition is imposed along each direction. The system is equilibrated under the Nosé-Hoover thermostat with temperature $k_B T = 4.0$. In this work, the external force field is imposed on the x - y plane and the temperature is defined by the velocity in the z direction. To collect the training samples, we impose an external force field f_x along the x -direction to generate the reverse Poiseuille flow, i.e.,

$$f_x(y) = \begin{cases} f_0 & \text{if } 0 < y < 45, \\ -f_0 & \text{else,} \end{cases} \quad (3.24)$$

where $f_0 = 0.01$ is the force magnitude. The full model is simulated for 1×10^6 steps with a time step $dt = 1 \times 10^{-3}$ to achieve the steady state, followed by 5×10^5 steps to collect the statistical samples in Eq (3.9), with t_c set to be 5. To construct the CG models, we learn $m = 3$ and $m = 4$ CVs from the training set. Throughout the rest of this paper, we will denote the reduced model that only uses the standard COMs by CGCOM, and the reduced models with 3 and 4 auxiliary CVs (per each molecule) by CG3 and CG4, respectively. In particular, for the CG3 model, the obtained 3 CG variables are equivalent which uniformly divides the individual molecule into three parts. In contrast, the obtained CG variables are inequivalent for the CG4 model.

3.3.2 Generalization ability of the CG models

As discussed in Sec. 3.2.1 and Sec. 3.2.2, the generalization ability of the CG models for the non-equilibrium processes relies on the assumption that the conditional PDF (i.e., $\propto \delta(\phi^Q(\mathbf{q}) - \mathbf{Q})\rho_e(\mathbf{q}))$) used for defining the CG projection operator (3.3) remains nearly the same, i.e., Eq. (3.6) holds for various external flow conditions. In practice, the direct evaluation of this high-dimensional PDF becomes computationally intractable. Instead, we relax this metric by examining the second moments of the atomistic particle distribution conditional to fixed CG

variables under various non-equilibrium conditions. Specifically, we examine the variation of the particle position in the orthogonal direction defined by $\mathbf{q}_{\perp, \mathbf{W}}^I = \mathbf{q}^I - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{q}^I$, where $\mathbf{q}^I \in \mathbb{R}^{N_m \times 3}$ represents the atomistic positions of the I -th molecule and \mathbf{W} is the weight matrix constructed by Eq. (3.9). The second-moment covariance matrix is defined by

$$\mathbf{C}(y) = \mathbb{E}^{\text{neq}} \left[\mathbf{q}_{\perp, \mathbf{W}}^I \mathbf{q}_{\perp, \mathbf{W}}^{I T} \mid y_c^I = y \right],$$

where $\mathbf{C}(y) \in \mathbb{R}^{N_m \times N_m}$ and the conditional expectation \mathbb{E}^{neq} is evaluated over the full molecules $I = 1, 2, \dots, M$ under the stationary distribution generated by the external force field (3.24). The condition $y_c^I = y$ restricts the COMs of the sampling molecule along the y -direction.

In particular, the external force generates the reverse Poiseuille flow along the x -direction, where the flow shear rate $\frac{\partial u_x}{\partial y}$ and the external stress vary along the y -direction. Specifically, the shear rate is close to 0 at $y = L/4$ and $y = 3L/4$, and achieves the largest value at $y = 0$, $y = L/2$, and $y = L$. Therefore, the variation of $\mathbf{C}(y)$ with respect to y provides a computationally accessible metric for Eq. (3.6). If the conditional PDF is close to the equilibrium distribution, $\mathbf{C}(y)$ should be homogeneous. Conversely, the large variation implies the deviation from the equilibrium distribution.

Figure 3.2 shows the difference $\|\mathbf{C}(y) - \mathbf{C}(y')\|_F$ of the various CG models at different locations, where y and y' correspond to the horizontal and vertical axis, respectively, and $\|\cdot\|_F$ is the Frobenius norm. It is clear that when the COMs are the only CG variables, the unresolved (i.e., orthogonal) degrees of freedom exhibit heterogeneous second moment along the y -direction, implying pronounced variation of the conditional PDF under different external fields. On the other hand, the variation decreases significantly when auxiliary CG variables are introduced into the CG3 and CG4 models, showing the promise of certain generalization ability for non-equilibrium processes.

The discrepancy in the conditional PDF will further lead to inconsistent reduced modeling terms under various non-equilibrium processes. To probe this effect, we examine the pairwise conservative force between the CG coordinates. We emphasize that the CG force generally exhibits a many-body nature Ge et al. (2023); here we use the projection along the pairwise direction Lei et al. (2010); Hijón et al. (2010) as a metric to quantify the generalization ability of the free energy

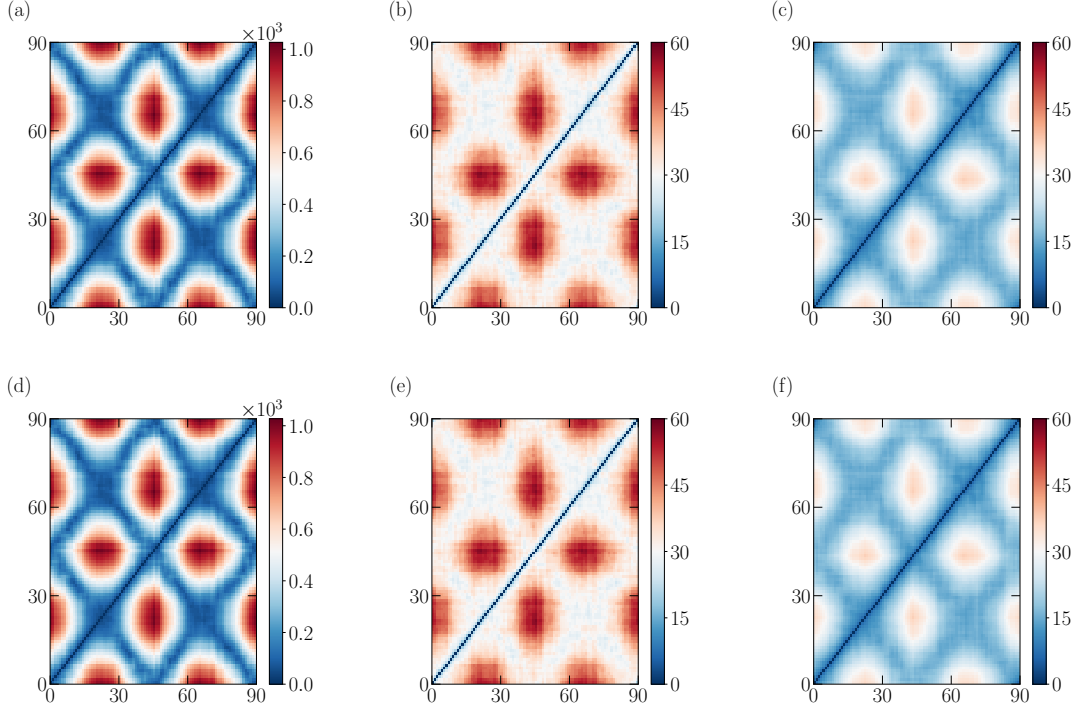


Figure 3.2 The Frobenius norm of the second-moment difference $\|C(y) - C(y')\|_F$ of the various CG models under the steady reverse Poiseuille flow generated by external force $f_0 = 0.01$ (upper) and $f_0 = 0.005$ (lower), where y and y' are represented by the horizontal and vertical axis, respectively. (a) and (d) CGCOM; (b) and (e) CG3; (c) and (f) CG4.

$U(Q)$. Specifically, let F_{ij}^{IJ} denote the force between the CG coordinates Q_i^I and Q_j^J . For the equilibrium state, the pairwise conservative force is defined as $F_{ij}^{\text{eq}}(Q) = \mathbb{E}^{\text{eq}} \left[F_{ij}^{IJ} \cdot \mathbf{e}_{ij}^{IJ} \mid Q_{ij}^{IJ} = Q \right]$, where $\mathbf{e}_{ij}^{IJ} = Q_{ij}^{IJ} / |Q_{ij}^{IJ}|$ and $\mathbb{E}^{\text{eq}}[\cdot]$ is the expectation over the molecule index $1 \leq I, J \leq M$ under the equilibrium state, and $1 \leq i, j \leq m$ is kept to represent the CG coordinate index within each molecule. For the non-equilibrium state, the pairwise conservative force is defined as $F_{ij}^{\text{neq}}(Q, y) = \mathbb{E}^{\text{neq}} \left[F_{ij}^{IJ} \cdot \mathbf{e}_{ij}^{IJ} \mid y_i^I = y_j^J = y, Q_{ij}^{IJ} = Q \right]$, where the condition is taken that both Q_i^I and Q_j^J share a specific value along the y -direction.

Figure 3.3 shows the difference $\left| F_{ij}^{\text{eq}}(Q) - F_{ij}^{\text{neq}}(Q, y) \right|$ for the different CG models. For the standard CGCOM model where Q is represented by the COMs, the two CG force terms agree well at $y = L/4 = 22.5$ and $y = 3L/4 = 67.5$, where the shear rate is close to zero and the atomistic particle distribution is near equilibrium. In contrast, the two force terms show pronounced differences at $y = 0, 2/L$, and L where the shear rate is large and the non-equilibrium distribution deviates from

the equilibrium distribution. This is consistent with the specific pattern of the second-moment difference in Fig. 3.2(a) and (d). The large difference implies the heterogeneity of the conditional PDF $\delta(\phi^Q(\mathbf{q}) - \mathbf{Q})\rho_e(\mathbf{Q})$ under various external conditions. Such limitation has a clear physical interpretation: the intra-molecular interactions may significantly affect the visco-elastic response as well as the collective dynamics, which, however, cannot be captured by the COMs. Fortunately, the inconsistency can be alleviated by introducing auxiliary CG variables into the reduced model. As shown in Fig. 3.3, the conservative force difference $|F_{ij}^{\text{eq}}(\mathbf{Q}) - F_{ij}^{\text{neq}}(\mathbf{Q}, y)|$ decreases to $O(0.1)$ for the CG3 and CG4 models, showing the improvement in the generalization ability of the free energy term $U(\mathbf{Q})$.

Finally, we study the generalization ability of the memory term $\mathbf{K}(\mathbf{Q}, t)$. Similar to $U(\mathbf{Q})$, we note that the memory term $\mathbf{K}(\mathbf{Q}, t)$ generally exhibits a many-body nature Lyu and Lei (2023b); here we examine the variation of the pairwise fluctuation force $\delta\mathbf{F}_{ij}^{IJ} = \mathbf{F}_{ij}^{IJ} - F_{ij}(\mathbf{Q}_{ij}^{IJ})\mathbf{e}_{ij}^{IJ}$. It can be loosely viewed as a metric of $\mathbf{K}(\mathbf{Q}, t = 0) = \mathbb{E}[\delta\mathbf{F} \otimes \delta\mathbf{F}]$. Specifically, we evaluate the variation of $\delta\mathbf{F}_{ij}^{IJ}$ under the equilibrium state

$$K_{ij}^{\text{eq}}(\mathbf{Q}) = \mathbb{E}^{\text{eq}} \left[\delta\mathbf{F}_{ij}^{IJT} \left(\mathbf{I} - \mathbf{e}_{ij}^{IJT} \mathbf{e}_{ij}^{IJ} \right) \delta\mathbf{F}_{ij}^{IJ} \middle| \mathbf{Q}_{ij}^{IJ} = \mathbf{Q} \right],$$

and the nonequilibrium state

$$K_{ij}^{\text{neq}}(\mathbf{Q}, y) = \mathbb{E}^{\text{neq}} \left[\delta\mathbf{F}_{ij}^{IJT} \left(\mathbf{I} - \mathbf{e}_{ij}^{IJT} \mathbf{e}_{ij}^{IJ} \right) \delta\mathbf{F}_{ij}^{IJ} \middle| \mathbf{Q}_{ij}^{IJ} = \mathbf{Q}, y_i^I = y_j^J = y \right],$$

where the expectation is over the molecule index $1 \leq I, J \leq M$. Fig. 3.4 shows the difference $|K_{ij}^{\text{neq}}(\mathbf{Q}, y) - K_{ij}^{\text{eq}}(\mathbf{Q})|$ for various CG models. Similar to the conservative CG force, the variation of the fluctuation force exhibits pronounced heterogeneity for the CGCOM model and achieves the largest discrepancy at the locations of large shear rate ($y = 0, L/2, L$). In contrast, the discrepancy becomes much smaller for the CG3 and CG4 models with auxiliary CG variables.

3.3.3 Non-equilibrium flows

The inconsistent reduced modeling terms shown in Sec. 3.3.2 reveal a caveat of using the standard CGCOM model for non-equilibrium processes. We examine this effect by simulating various non-equilibrium flows.

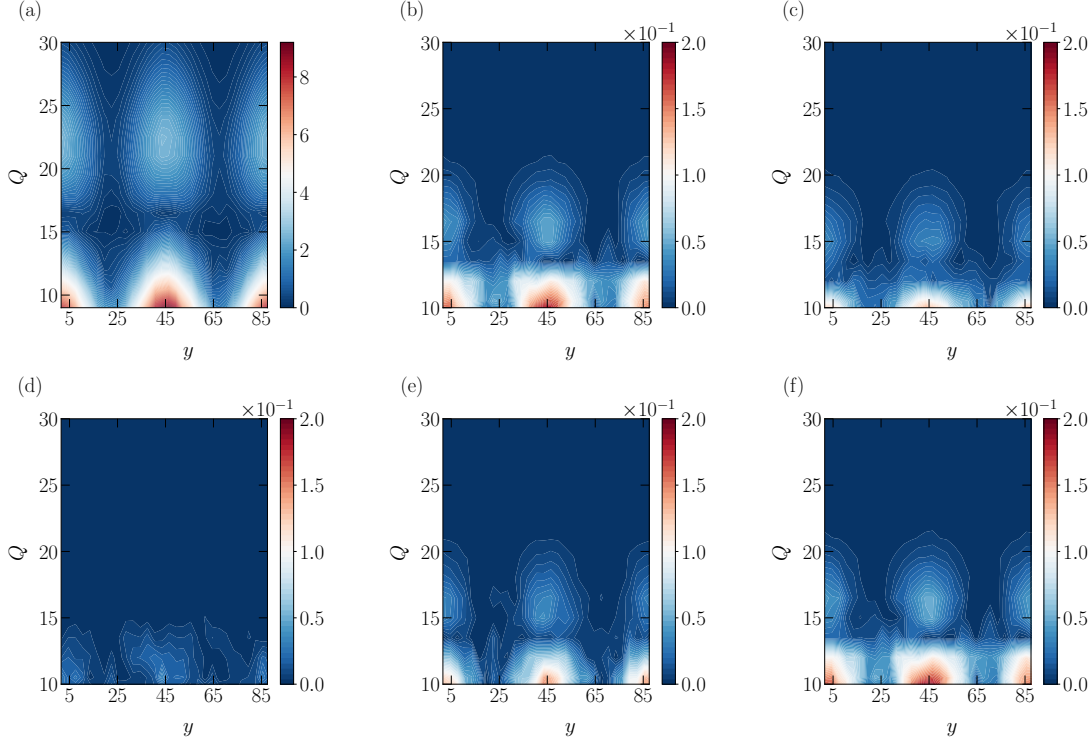


Figure 3.3 The pairwise conservative force difference $\left| F_{ij}^{\text{neq}}(\mathbf{Q}) - F_{ij}^{\text{eq}}(\mathbf{Q}, y) \right|$ for various CG models, which loosely quantifies the generalization ability of the CG free energy $U(\mathbf{Q})$. (a) CGCOM; (b) CG3 with $(i, j) = (1, 1)$; (c-f) CG4 with $(i, j) = (1, 1), (2, 2), \dots, (4, 4)$.

To establish a fair comparison, we construct the many-body form of the free energy $U(\mathbf{Q})$ and the memory term $\mathbf{K}(\mathbf{Q}, t)$ with the symmetry-preserving neural network representations presented in Sec. 3.2.3 and Sec. 3.2.4 for the CGCOM, CG3 and CG4 models, respectively. To verify $U(\mathbf{Q})$, we sample the radial distribution function of the inter-molecular COMs for the CGCOM model and the intra-molecular CG coordinates for the CG3 and CG4 models. To verify $\mathbf{K}(\mathbf{Q}, t)$, we sample the velocity auto-correlation function of the CG variables. As shown in the Appendix, the predictions of the three CG models show good agreement with the full MD results, which verify their efficacy for modeling equilibrium processes.

Next, we consider the non-equilibrium reverse Poiseuille flow generated by an external force (3.24) with $f_0 = 0.01$. Fig. 3.5 shows the velocity development $u_x(y, t)$. Compared with the full MD results, the prediction of the standard CGCOM overestimates the instantaneous velocity magnitude by three times. Furthermore, it can not capture the development oscillation near $t = 60$. This

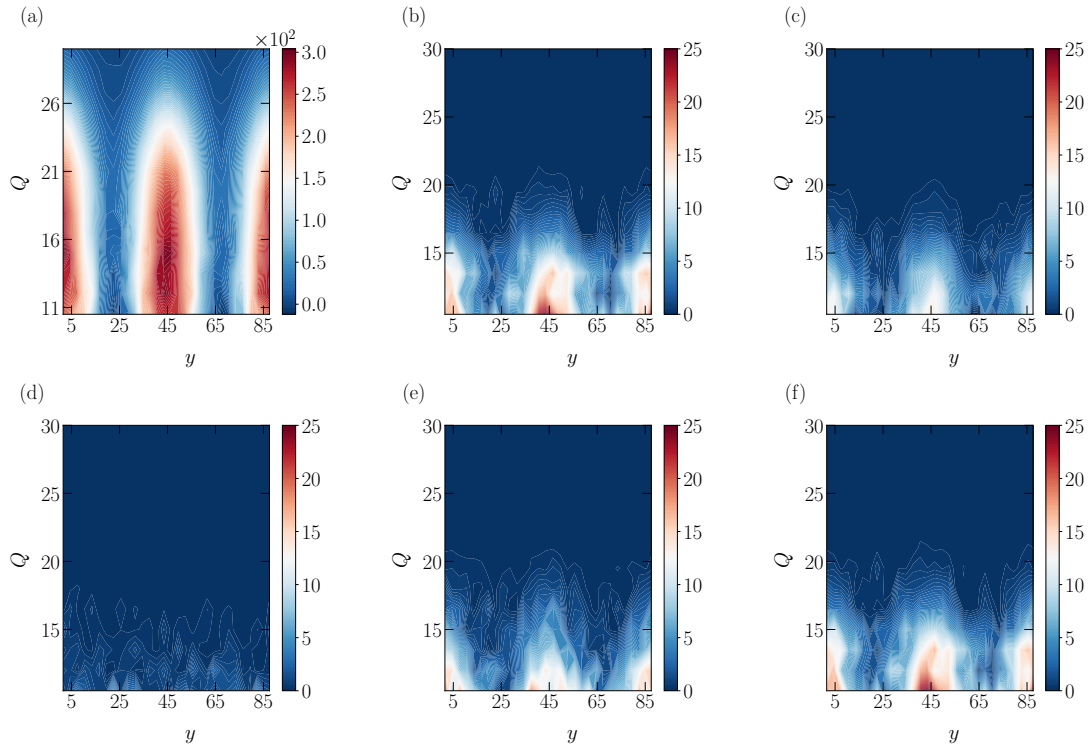


Figure 3.4 The difference of the variation of the fluctuation force $|K_{\text{neq}}^{ij}(Q) - K_{\text{eq}}^{ij}(Q, y)|$ for different CG models, which loosely quantifies the generalization ability of the CG memory term $\mathbf{K}(Q, t)$ at $t = 0$. The sub-figure labels are the same as Fig. 3.3.

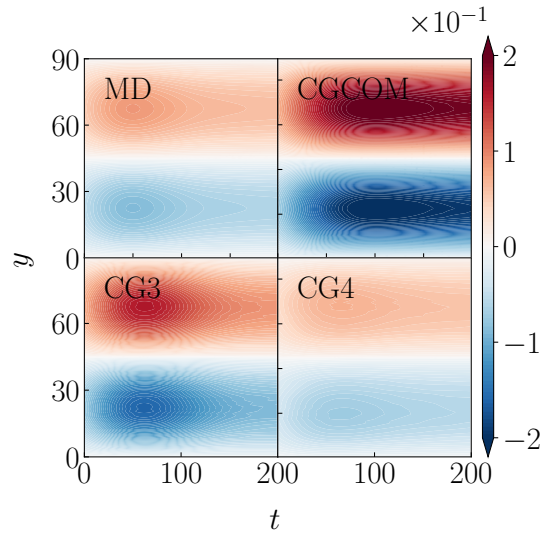


Figure 3.5 The development of the reverse Poiseuille flow under the external force $f_0 = 0.01$ predicted by the full MD and three CG models.

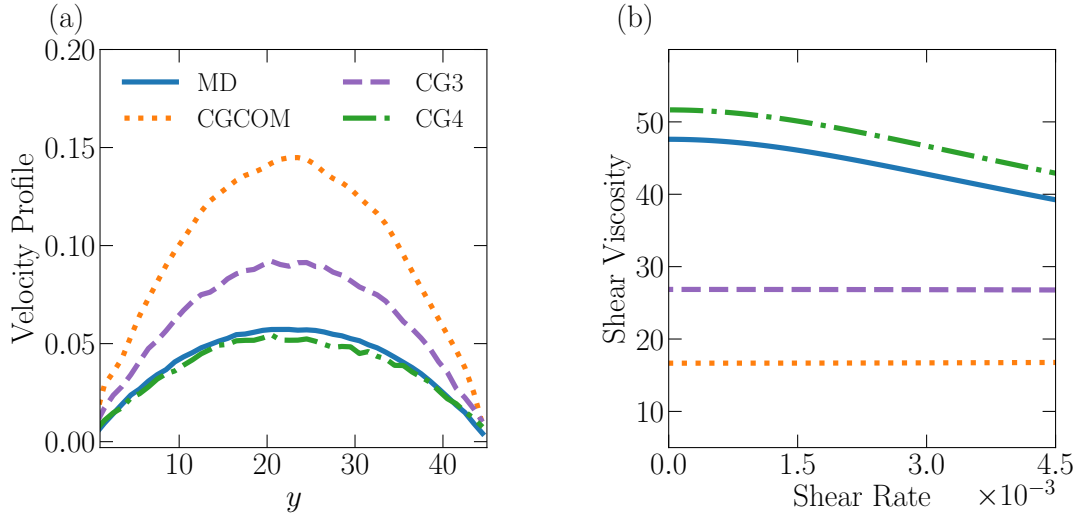


Figure 3.6 The steady state reverse Poiseuille flow under external force $f_0 = 0.01$ predicted by the full MD and the CG models. (a) The stationary velocity profile $u_x(y)$ (b) The shear-rate-dependent viscosity.

limitation is rooted in the choice of the COMs as the CG variables, which ignores the intra-molecular interactions and therefore can not capture the complex visco-elastic responses under non-equilibrium processes. The prediction of the CG3 model shows significant improvement to the CGCOM model but overestimates the magnitude of the velocity oscillation during the flow development stage. On the other hand, the prediction of the CG4 model shows a good agreement with the full MD results throughout the development stage.

Fig. 3.6 shows the steady state velocity profile $u_x(y)$ predicted by the different models. Similar to the development stage, the prediction of the CG4 model shows a good agreement with the MD results. However, the prediction of the CGCOM and the CG3 models show apparent discrepancies due to their inefficacy in modeling the intra-molecular interactions. To quantify the difference, we compute the shear-rate-dependent viscosity from the steady velocity profile. The predictions of the CGCOM and the CG3 model are nearly independent of the shear rate, which indicates that they can not capture the visco-elastic responses associated with the molecular conformation change under external shear flow. In contrast, the CG4 model can faithfully reproduce the shear-rate-dependent viscosity of the full MD results.

Finally, we examine the generalization ability of the CG models for other flow fields. While

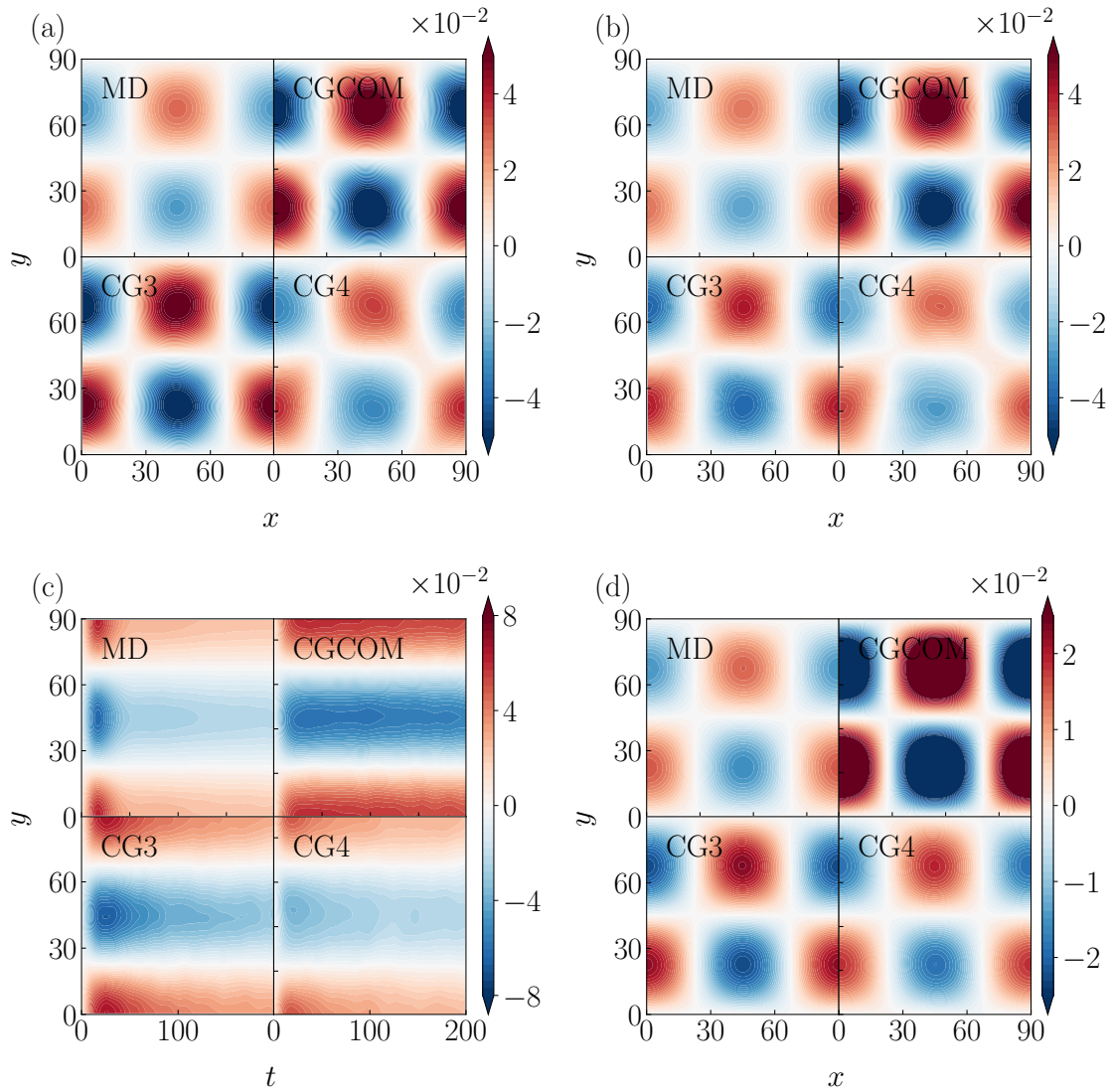


Figure 3.7 The velocity field of a 2D Taylor vortex predicted by the full MD and the CG models. (a-b) The contour of the instantaneous velocity $u_y(x, y, t)$ at $t = 50$ (a) and $t = 100$ (b). (c) The development of velocity $u_y(x = 25, y, t)$ (d) The steady velocity field $u_y(x, y)$.

the training samples of the CG models are collected in the reverse Poiseuille flow, we validate the constructed models with the Taylor vortex generated by the external force field

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

where $f_0 = 1 \times 10^{-2}$ and $L = 90$. Fig. 3.7 shows the prediction of the 2D velocity contour from various models. Similar to the previous example, the CGCOM and CG3 models overestimate the magnitude of the velocity field due to their insufficiency in modeling the intra-molecular interactions arising from the molecular conformation change under the external flow field. In contrast, the prediction of the CG4 model yields show good agreement with the full MD model.

Although the CG4 model performs well in the vortex flow, we should not view it as the “correct” (as opposed to the CGCOM) model for all non-equilibrium processes. Our main point is that the conditional PDF associated with the CG projection needs to be consistent to ensure the model’s generalization ability, which, unfortunately, can not be guaranteed for non-equilibrium processes. This issue may severely limit a CG model’s applicability to practical problems. On the other hand, properly introducing auxiliary CG variables may mitigate this inconsistency, and significantly improve the model’s generalization ability.

3.4 Conclusion

This work presents a caveat in constructing reliable CG molecular dynamics models for non-equilibrium processes. Specifically, a CG model’s generalization ability relies on its consistency in the conditional PDF of the phase space vector (or equivalently, the CG projector operator) under various non-equilibrium conditions. This criterion is determined by the proper choice of the CG variables a priori and can not be remedied by constructing more accurate reduced modeling (the free energy and the memory) terms. Although the Zwanzig’s projection formalism can in principle provide the rigorous reduced dynamics, it is valid only when the distribution of the CG variables is consistent with the one associated with the CG projection operator. Unfortunately, this metric seems to be broadly overlooked in most existing studies on CGMD modeling. While the previous works (e.g. Refs. Lei et al. (2010); Hijón et al. (2010)) use dynamic properties such as the velocity auto-correlation functions and the mean squared displacement to validate the CG models, these

quantities are essentially defined under the marginal density distribution near-equilibrium; their applications to non-equilibrium processes are generally un-warranted.

To alleviate the above challenge, this work proposes to learn a set of auxiliary CG variables based on the TICA to minimize the entropy contribution of the unresolved variables so that the conditional PDF of the full phase space vector under various conditions approaches the equilibrium case (i.e., Eq. (3.6)). We verify this generalization metric by examining the distribution of the second moment, the fluctuation force, and its variation under various external conditions. We show that the common CG model that uses the molecule's COMs as the CG variables is generally insufficient to retain the consistent conditional PDF. In contrast, the present models with auxiliary CG variables show significant improvement. Furthermore, the crucial role of this metric is reflected in modeling the non-equilibrium reverse Poiseuille flow and the Taylor vortex. In particular, the prediction of the standard CGCOM model exhibits large discrepancies from the full MD results due to its inefficacy in modeling the intramolecular interactions under external flows different from the equilibrium conditions. Conversely, the present model can faithfully recover the complex visco-elastic responses and therefore yields consistent predictions with the full MD results.

Finally, we note that the learning of the auxiliary CG variables in the present study remains somewhat empirical. They are constructed based on the TICA and hence take a linear form of the full atomistic coordinates. In practice, we may learn the CG variables as nonlinear encoders of the molecular conformation to achieve a more efficient representation of the intra-molecular interactions (e.g., see Ref. Fang et al. (2022)). We leave this for future study.

CHAPTER 4

CONSENSUS-BASED ENHANCED SAMPLING FOR HIGH-DIMENSIONAL FREE ENERGY SURFACES

4.1 Challenges in Constructing High-Dimensional Free Energy Surfaces

Molecular dynamics (MD) provides an essential tool Frenkel and Smit (2023) to access the micro-scale insights of complex processes in many scientific applications, including chemical engineering, material synthesis, and drug design. To further predict the collective behaviors, it is often desirable to construct the free energy surfaces (FESs) with respect to a set of functional-relevant collective variables (CVs). However, the accurate construction of FES remains a practical challenge due to the high dimensionality and the prevalence of energy barriers. Several ingenious methods based on importance sampling have been developed, such as umbrella sampling Torrie and Valleau (1977b), histogram reweighting Kumar et al. (1992), metadynamics Laio and Parrinello (2002); Barducci et al. (2008); Grafke and Laio (2024), variational enhanced sampling Valsson and Parrinello (2014); Shaffer et al. (2016); Bonati et al. (2019), weighted ensemble and its variants Zhang et al. (2007); Ahn et al. (2017, 2021), hyperdynamics Voter (1997), transition matrix based unbiasing method Wu and Noé (2014), orthogonal space random walk Zheng et al. (2008) and adaptive biasing force Darve and Pohorille (2001); Darve et al. (2008); Lelièvre et al. (2008); Chipot and Lelièvre (2011). Despite the efficacy in enhancing the phase space exploration, these methods are based on iterative estimation of the biased probability density functions (PDFs) (or biased force), where the efficiency generally diminishes as the number of CVs increases. Strategies such as bias-exchange Piana and Laio (2007) and parallel bias metadynamics Pfaendtner and Bonomi (2015); Prakash et al. (2018) have been developed to alleviate this issue. Both methods factorize the biasing potential as a product of several low-dimensional functions and can work effectively when the underlying FESs admit such structures which, however, may not always be the case for high-dimensional problems. Several other methods Sugita and Okamoto (1999); Marinari and Parisi (1992); Gao (2008); Yang and Qin Gao (2009); Kim et al. (2006); Hamelberg et al. (2004); Miao et al. (2015); Lu and Vanden-Eijnden (2013); Yu et al. (2016) achieve enhanced sampling

irrespective of pre-defined CVs (see review Yang et al. (2019)). However, the construction of explicit high-dimensional FESs remains a challenging problem.

Alternative to the histogram methods, the temperature-accelerated molecular dynamics Maragliano and Vanden-Eijnden (2006) and the adiabatic free energy dynamics Rosso et al. (2002); Abrams and Tuckerman (2008) are developed by introducing an extended dynamics of the CVs with an artificially high temperature to overcome the energy barriers. The methods avoid dealing with the biased potentials and have shown the promise of efficient exploration of the high-dimensional CV space Abrams and Vanden-Eijnden (2010). On the other hand, to further construct the FES, the sampling points are deposited in a greedy way Maragliano and Vanden-Eijnden (2008); the effectiveness for high-dimensional problems remains under-explored.

With the recent advancements in machine learning, kernel methods and deep neural networks (DNNs) have demonstrated promising results in representing multi-dimensional FES Stecher et al. (2014b); Mones et al. (2016); Schneider et al. (2017); Bonati et al. (2019); Sidky and Whitmer (2018). However, as also noted in Ref. Cendagorta et al. (2020), one essential problem for constructing high-dimensional FES lies in how to optimize the training set. Ideally, to efficiently approximate a high-dimensional function, the sampling should enable certain adaptivity based on the posterior error so that the training set can be updated according to the particular underlying FES. However, such posterior error-based sampling remains largely open for FES construction since the approximation residual can neither be easily queried nor sampled within the full phase space. In practice, the training samples are often collected either as pre-defined grid points or in a heuristic manner. As the number of CVs increases, such choices may induce pronounced approximation error and suffer from performance degradation, since most phase space regimes essentially become thermodynamically inaccessible. Remarkably, the DeepVes Bonati et al. (2019) enables the efficient sampling of the CV space by jointly constructing the bias potential and the target distribution. On the other hand, the accurate reconstruction of explicit FES further requires the estimation of high-dimensional PDFs from the samples. Recent efforts Noé et al. (2019); Gabrié et al. (2022) seek the direct approximation of the transition operator instead of estimating PDFs using generative models, when representative

configurations are known a priori. In addition, the reinforced dynamics Zhang et al. (2018d); Wang et al. (2022) (see also Ref. van der Oord et al. (2023)) proposes an efficient approach to impose adaptivity by using an uncertainty indicator as an indirect measure of the construction error to bias MD simulations, which relies on calculating the standard deviation of the predictions from multiple DNNs trained on the same dataset.

In this work, we present a consensus-based enhanced sampling (CES) method to efficiently construct high-dimensional FES for complex MD systems. A unique feature is that the method enables adaptive posterior error-based sampling such that the FES approximation and the sample points can be simultaneously optimized. Also, unlike the reinforced dynamics, the method does not rely on the training of multiple DNN representations. The main idea is to formulate the construction as a min-max problem, where the max-problem seeks a residue-based distribution to establish adaptive sampling in the vicinity of the explored phase space regime, and the min-problem optimizes the DNN parameters for the FES representation. Iteratively optimizing both the training set and the DNN representation achieves an adversarial construction of the FES pertaining to the kinetic transition regime. For the maximization step, it is worth mentioning that the query of the thermodynamically accessible phase space is non-trivial. In particular, existing approaches Tang et al. (2023); Bao et al. (2020) in the spirit of adversarial generative models can not be directly applied, since the global residual error is unknown a priori. Instead, we establish a consensus-based sampling Carrillo et al. (2022) (see also Refs. Pinnau et al. (2017); Carrillo et al. (2018, 2021); Chen et al. (2022)) in the form of a stochastic interacting particle system governed by a McKean stochastic differential equation. A quadratic potential is adaptively constructed to probe the local maximum error regime by exploiting the Laplace approximation under a low-temperature limit. Meanwhile, a coherent noise term is introduced to efficiently explore the full CV space under a high-temperature limit and yield the target sampling points for the maximization problem, by which the minimization of the DNN residue can be solved jointly. We demonstrate the effectiveness of the proposed method by constructing the FES of several biomolecule systems with the number of CVs up to 30. Fig. 4.1 sketches the workflow of the proposed method.

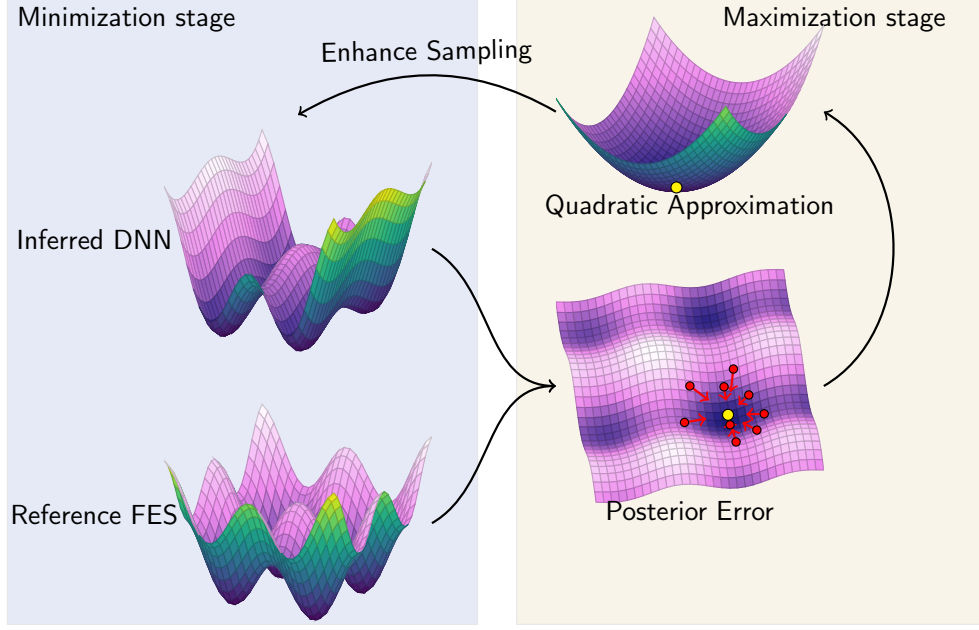


Figure 4.1 The workflow of the present CES-based method. In the minimization step, given a collection of sampling points, the reference force (i.e., the gradient of the underlying FES) can be calculated using the restrained dynamics; a comparison with the force inferred from the DNN approximation yields a posterior error. In the maximization step, the posterior error determines a residue-based distribution with entropy regularization. An interacting particle system following McKean stochastic dynamics is used to achieve an adaptive sampling of the max-residue regime by modulating the exploitation of the Laplace approximation of the current residue-based distribution and the exploration of the uncharted phase space. The FES can be accurately reconstructed after several iterations of the minimization and maximization step.

4.2 Consensus-Based Sampling Methodology

4.2.1 Free energy and mean forces

We consider a full model with micro-scale coordinates $\mathbf{r} \in \mathbb{R}^N$ whose dynamics is governed by potential $U(\mathbf{r}) : \mathbb{R}^N \rightarrow \mathbb{R}$ under temperature T . Suppose we are interested in CVs $\mathbf{s}(\mathbf{r}) : \mathbb{R}^N \rightarrow \Gamma$ with $\Gamma \subset \mathbb{R}^M$, the FES $A(\mathbf{z})$ of the CVs is defined by

$$A(\mathbf{z}) = -\frac{1}{\beta} \ln \rho(\mathbf{z}), \quad (4.1)$$

where $\beta = 1/k_B T$ is the inverse of the thermal temperature,

$$\rho(\mathbf{z}) = \frac{1}{Z} \int \exp(-\beta U(\mathbf{r})) \delta(\mathbf{s}(\mathbf{r}) - \mathbf{z}) d\mathbf{r} \quad (4.2)$$

is the marginal density for $\mathbf{s}(\mathbf{r}) = \mathbf{z}$, $Z = \int \exp(-\beta U(\mathbf{r})) d\mathbf{r}$ is the partition function, and $\delta(\cdot)$ represents the Dirac delta function, which we refer readers to Stoltz et al. (2010) for more detailed

explanation. For high-dimensional CVs, the direct estimation of $\rho(\mathbf{z})$ often becomes numerically challenging. An alternative approach is to fit the mean force $\mathbf{F}(\mathbf{z}) := -\nabla A(\mathbf{z})$ at various sample points, which can be estimated via the restrained dynamics Allen and Tildesley (2017) by introducing a harmonic term into the full potential, i.e.,

$$U_k(\mathbf{r}, \mathbf{z}) = U(\mathbf{r}) + \frac{k}{2}(\mathbf{s}(\mathbf{r}) - \mathbf{z})^T(\mathbf{s}(\mathbf{r}) - \mathbf{z}), \quad (4.3)$$

where k represents the magnitude of the restrained potential. As shown in Ref Maragliano and Vanden-Eijnden (2006), the mean force can be computed by $\mathbf{F}(\mathbf{z}) = \lim_{k \rightarrow \infty} \mathbf{F}^k(\mathbf{z})$, where $\mathbf{F}^k(\mathbf{z})$ is defined by

$$\mathbf{F}^k(\mathbf{z}) = \frac{1}{Z_k(\mathbf{z})} \int k(\mathbf{s}(\mathbf{r}) - \mathbf{z}) \exp(-\beta U_k(\mathbf{r}, \mathbf{z})) d\mathbf{r}, \quad (4.4)$$

and can be sampled as the local first-moment estimation.

In principle, given a collection of sample points \mathbf{z} , $A(\mathbf{z})$ can be re-constructed (up to a constant) by matching the mean force $-\nabla A(\mathbf{z})$ at the individual points. However, for complex MD systems, the sampling over the phase space could be highly non-trivial due to the prevalence of local energy minima; the training set is often determined a priori as pre-selected points or in a greedy manner. As the number of CVs increases, the empirical random samples may introduce pronounced discretization error of the loss function over the full phase space. To efficiently construct $A(\mathbf{z})$ in the thermodynamically accessible regime, it is desirable to simultaneously optimize the training set and the FES approximation through certain adaptive sampling based on the posterior residual error. This motivates the present method illustrated below.

4.2.2 Min-Max formulation

Let $A_{\mathcal{N}}(\mathbf{z})$ denote the DNN representation of the FES $A(\mathbf{z})$, which is parameterized by minimizing the loss function

$$\mathcal{L}_{\mathcal{N}}(\mathbf{z}) = \|\nabla_{\mathbf{z}} A_{\mathcal{N}}(\mathbf{z}) + \mathbf{F}(\mathbf{z})\|^2 \quad (4.5)$$

for $\mathbf{z} \in \Gamma$. To solve the problem, we introduce a sampling distribution $q(\mathbf{z})$ and define the weighted loss

$$(\mathcal{L}_{\mathcal{N}}, q) = \int_{\Gamma} \mathcal{L}_{\mathcal{N}}(\mathbf{z}) q(\mathbf{z}) d\mathbf{z}. \quad (4.6)$$

A desired distribution intends to maximize the discrepancy in the dataset for a given network $A_{\mathcal{N}}(\mathbf{z})$. Accordingly, we define the maximum problem as

$$G[\mathcal{L}_{\mathcal{N}}] = \max_q(\mathcal{L}_{\mathcal{N}}, q). \quad (4.7)$$

Since $(\mathcal{L}_{\mathcal{N}}, q)$ is always non-negative, a good approximation of the original free energy surface $A_{\mathcal{N}}(\mathbf{z})$ (up to a constant) can be obtained by solving the following problems

$$\min_{A_{\mathcal{N}}} \max_q(\mathcal{L}_{\mathcal{N}}, q). \quad (4.8)$$

Proposition 4.2.1. *Assuming that there exists a solution of $\mathcal{L}_{\mathcal{N}}(\mathbf{z}) = 0$ for $\mathbf{z} \in \Gamma$, then A^* is a solution if and only if it solves (4.8).*

Proof. Suppose A^* is the solution for $\mathcal{L}_{\mathcal{N}}(\mathbf{z}) = 0$, it satisfies $(\mathcal{L}_{\mathcal{N}^*}, q) = 0$ for any $q \in W$, i.e., $G[\mathcal{L}_{\mathcal{N}^*}] = 0$. Therefore, A^* is a solution for the minimax problem (4.8). On the other hand, if \hat{A} is the minimizer for problem (4.8) but not the solution for $\mathcal{L}_{\mathcal{N}}(\mathbf{z}) = 0$, then there exists $\hat{q} \in V$ such that $(\mathcal{L}_{\hat{\mathcal{N}}}, \hat{q}) > 0$. However, $(\mathcal{L}_{\mathcal{N}^*}, q) = 0$ for all q , which contradicts the assumption that \hat{A} is the minimizer. \square

Proposition 4.2.1 shows that the direct construction of FES $A(\mathbf{z})$ can be reformulated as an adversarial learning of the optimal solution $A_{\mathcal{N}}(\mathbf{z})$ for the min-max problem (4.8). Accordingly, the training consists of two components: the minimization part optimizes the DNN representation with the current training set; the maximization part explores the regime of the largest residue for the current DNN representation and essentially establishes an adaptive sampling of the training set based on the posterior error.

To numerically solve the max-problem, certain regularization needs to be introduced. Otherwise, the max-problem will simply yield a delta measure, i.e., $\delta(\mathbf{z} - \mathbf{z}^*)$, where $\mathbf{z}^* = \arg \max \mathcal{L}_{\mathcal{N}}(\mathbf{z})$. Since the sampling needs to simultaneously identify the max-residual regime and explore the uncharted phase space, we introduce the entropy-based regularization Wang et al. (2020a); Gao et al. (2022) (see also Refs. Gulrajani et al. (2017); Miyato et al. (2018); Tang et al. (2023) for gradient-based

regularization), and the max-problem is reformulated by

$$\min_q \int (-\mathcal{L}_{\mathcal{N}}(\mathbf{z}) + \kappa_h^{-1} \ln q(\mathbf{z}))q(\mathbf{z})d\mathbf{z}. \quad (4.9)$$

The problem is convex for a PDF q with a unique global minimum at $q^*(\mathbf{z}) = \exp(-\kappa_h \mathcal{L}_{\mathcal{N}}^-(\mathbf{z}))/Z^*$, where $Z^* = \int \exp(-\kappa_h \mathcal{L}_{\mathcal{N}}^-(\mathbf{z}))d\mathbf{z}$ and $\mathcal{L}_{\mathcal{N}}^-(\mathbf{z}) = -\mathcal{L}_{\mathcal{N}}(\mathbf{z})$. The parameter κ_h is a Lagrangian multiplier, balancing the focus between the peak concentration and the scope of exploration, and is somewhat similar to the inverse temperature in statistical mechanics. An elevated κ_h^{-1} induces a distribution closer to a uniform distribution. Conversely, a diminished κ_h^{-1} induces a concentrated distribution near the max-residue point.

However the analytical formula of $q^*(\mathbf{z})$ further depends on the accurate FES, the direct sampling remains challenging for high-dimensional CVs. Inspired by the consensus-based sampling method Carrillo et al. (2022), we establish a stochastic particle system governed by the McKean stochastic differential equation. By properly constructing the conservative potential and noise terms, the particle distribution provides a good approximation of q^* as illustrated below.

4.2.3 Exploitation and exploration in the max-problem

To approximate the target distribution $q^*(\mathbf{z})$, particularly in the vicinity of the max-residual point \mathbf{z}^* , we exploit Laplace's principle in the large deviations theory, i.e.,

$$\lim_{\kappa \rightarrow \infty} \left(-\frac{1}{\kappa} \log \left(\int \exp(-\kappa f(\mathbf{z})) d\rho^*(\mathbf{z}) \right) \right) = f(\mathbf{z}^*) \quad (4.10)$$

holds true for any compactly supported probability measure ρ^* , where $\mathbf{z}^* \in \text{supp}(\rho^*)$ uniquely minimizes the function f . This enables us to identify the max-residual point from a collection of samples $\{\mathbf{z}^i\}_{i=1}^{N_w}$ by the first-order momentum \mathbf{m} under the weighted density function $p(\mathbf{z})$, i.e.,

$$\mathbf{m} = \int \mathbf{z} p(\mathbf{z}) d\mathbf{z} \approx \sum_{i=1}^{N_w} \mathbf{z}^i \hat{p}(\mathbf{z}^i) \quad \hat{p}(\mathbf{z}) = \frac{\exp(-\kappa_l \mathcal{L}_{\mathcal{N}}^-(\mathbf{z}))}{\sum_{i=1}^{N_w} \exp(-\kappa_l \mathcal{L}_{\mathcal{N}}^-(\mathbf{z}^i))} \quad (4.11)$$

where κ_l^{-1} represents a low temperature limit. However, the integration is subject to the so-called curse of dimensionality as the number of CVs increases. Instead, we treat sampler \mathbf{z}^i as a random walker \mathbf{z}_t^i governed by the following McKean stochastic differential equation

$$\dot{\mathbf{z}}_t^i = -\frac{1}{\gamma} \nabla_{\mathbf{z}} G(\mathbf{z}_t^i) + \sqrt{\frac{2}{\kappa_h \gamma}} \xi_i(t), \quad (4.12)$$

where $G(\mathbf{z}_t) = \frac{1}{2}(\mathbf{z}_t - \mathbf{m}_t)^T V_t^{-1}(\mathbf{z}_t - \mathbf{m}_t)$ denotes an adaptively constructed conservative potential function. The formulations of \mathbf{m}_t, V_t are specified in (4.13) with the rationale discussed in the next section. Consequently, $G(\mathbf{z}_t)$ navigates the random walkers (i.e., individual particles) towards \mathbf{m}_t , which represents the region of large residual error. The second term in Eq. (4.12) represents a stochastic term where γ represents the friction coefficient and $\xi(t)$ represents the standard Gaussian white noise characterized by zero mean and covariance $\mathbb{E}[\xi_i(t)\xi_j(t')] = \delta_{ij}\delta(t - t')$.

The coupling of the conservative and stochastic terms maintains a relatively high temperature κ_h^{-1} , and a large friction coefficient γ is applied such that the distribution is almost always Gaussian during the evolution. As shown in the following section, the distribution of walkers $q_t(\mathbf{z})$ converges to $\propto \exp(-\kappa_h G(\mathbf{z}))$ characterized by \mathbf{m}_∞ and V_∞ . Accordingly, the balance between exploitation and exploration is controlled using two temperatures κ_l^{-1} and κ_h^{-1} . As κ_l^{-1} decreases, the random walker distribution concentrates near the max-residual points of the current model, reflecting the role of exploitation. Conversely, as κ_h^{-1} increases, the distribution smoothens progressively, enhancing the exploration of the uncharted regions.

4.2.4 Adaptive parameter estimation and FES construction

In this section, we show that the sampling distribution governed by Eq. (4.12) converges to a steady state as an approximation of q^* with the proper choices of \mathbf{m}_t and V_t given by

$$\begin{aligned} \mathbf{m}_t &= \sum_{i=1}^{N_w} \mathbf{z}_t^i \hat{p}(\mathbf{z}_t^i), \\ V_t &= \kappa_t \sum_{i=1}^{N_w} (\mathbf{z}_t^i - \mathbf{m}_t)(\mathbf{z}_t^i - \mathbf{m}_t)^T \hat{p}(\mathbf{z}_t^i), \end{aligned} \tag{4.13}$$

where $\hat{p}(\mathbf{z})$ is defined by (4.11). In particular, with $\kappa_t = \kappa_l + \kappa_h$, (4.12) converges to the distribution based on the target residues.

Proposition 4.2.2. *Suppose $\mathcal{L}_N^-(\mathbf{z})$ takes a local quadratic approximation in form of $\frac{1}{2}(\mathbf{z}-\mu)^T \Sigma^{-1}(\mathbf{z}-\mu)$, $q_t \rightarrow \frac{\exp(-\kappa_h \mathcal{L}_N^-(\mathbf{z}))}{\int \exp(-\kappa_h \mathcal{L}_N^-(\mathbf{z})) d\mathbf{z}}$ as $t \rightarrow \infty$, given $\kappa_t = \kappa_l + \kappa_h$.*

We refer to SI for the proof. The stochastic dynamics (4.12) and (4.13) differs from the regular Langevin dynamics; it is nonlocal and similar to the one in Carrillo et al. (2022) except that the

parameter κ_h appears in the target distribution that modulates the exploration of the sampling dynamics. The quadratic assumption of the loss function $\mathcal{L}_{\mathcal{N}}^-(\mathbf{z})$ is due to the fact that we are mainly interested in the regime near the max-residue point. Under the low-temperature limit, the local regime can be well characterized by the first and second moments following Laplace’s principle. While \mathbf{m} identifies the extremal point, V recognizes the anisotropic nature among the different CVs. In this study, for the sake of computational efficiency, we further simplify V by only considering the diagonal entries denoted as \mathbf{v} . Moreover, we utilize a moving average to ensure stable estimation

$$\begin{aligned}\mathbf{m}_{t+1} &= \beta_1 \mathbf{m}_t + (1 - \beta_1) \sum_{i=1}^{N_w} \mathbf{z}_t^i \hat{p}(\mathbf{z}_t^i), \\ \mathbf{v}_{t+1} &= \beta_2 \mathbf{v}_t + (1 - \beta_2) \kappa_t \sum_{i=1}^{N_w} (\mathbf{z}_t^i - \mathbf{m}) \odot (\mathbf{z}_t^i - \mathbf{m}) \hat{p}(\mathbf{z}_t^i),\end{aligned}\tag{4.14}$$

where β_1 and β_2 are hyper-parameters. To maintain unbiased estimation, normalization is implemented as follows: $\mathbf{m} = \frac{\mathbf{m}_t}{1 - \beta_1^t}$ and $\mathbf{v} = \frac{\mathbf{v}_t}{1 - \beta_2^t}$.

With the training samples obtained from the aforementioned maximization step, the DNN representation of the FES is optimized using the Adam stochastic gradient descent method Kingma and Ba (2015) for the minimization step. The loss function of the updated DNN representation $A_{\mathcal{N}}(\mathbf{z})$, in turn, navigates the consensus-based adaptive sampling for the updated maximization step. The min-max problem is solved iteratively to achieve comprehensive sampling of the full phase space.

So far, the construction process is based on the function approximation of $A(\mathbf{z})$ over the full regime. However, we note that the kinetic processes of a molecular system are generally characterized by the local minima and saddles points on the FES. On the other hand, the regimes of high free energy are less relevant. To accurately construct these thermodynamically accessible regimes, we modify the loss function \mathcal{L} as

$$\mathcal{L}_{\mathcal{N}}(\mathbf{z}) = \frac{\|\nabla_{\mathbf{z}} A_{\mathcal{N}}(\mathbf{z}) + \mathbf{F}(\mathbf{z})\|^2}{\|\mathbf{F}(\mathbf{z})\|^2 + e},\tag{4.15}$$

for all the biomolecule systems except the toy example of the 1D Rastrigin function. e is a small value to regularize the denominator. We note that a similar formulation is used in Ref. Maragliano and

Vanden-Eijnden (2008) to quantify the model accuracy after the FES is constructed. Alternatively, in this study, $\mathcal{L}_{\mathcal{N}}(\mathbf{z})$ directly involves the construction process for the adaptive sampling of the training set. A detailed algorithm is presented in Algorithm 4.1.

Algorithm 4.1 Consensus-based enhance sampling.

Require: Initial sampling point \mathbf{z}_0^i , for $i = 1, \dots, N_w$

Require: Initial DNN parameter θ_0

Require: The number of training iterations N_{train}

Require: The number of data collected N_{data} in each training iteration

$j \leftarrow 0, t \leftarrow 0$

$T \leftarrow \lceil \frac{N_{data}}{N_w} \rceil$

while $j < N_{train}$ **do**

while $t \leq T$ **do**

 calculate the mean force \mathbf{F}_t^i at \mathbf{z}_t^i

 calculate the predicted force $\mathcal{F}_{\theta}(\mathbf{z}_t^i) = \nabla_{\mathbf{z}} A_{\mathcal{N}}(\mathbf{z}_t^i; \theta_j)$

$L^i \leftarrow \mathcal{L}_{\mathcal{N}}(\mathbf{z}_t^i)$

$w^i \leftarrow \frac{\exp(\kappa_l L^i)}{\sum_i \exp(\kappa_l L^i)}$

$\mathbf{m}_{t+1} \leftarrow \beta_1 \mathbf{m}_t + (1 - \beta_1) \sum_i \mathbf{z}_t^i w^i$

$\mathbf{v}_{t+1} \leftarrow \beta_2 \mathbf{v}_t + (\kappa_l + \kappa_h)(1 - \beta_2) \sum_i (\mathbf{z}_t^i - \mathbf{m}_t)^2 w^i$

$\mathbf{m} \leftarrow \frac{\mathbf{m}_{t+1}}{1 - \beta_1^t}$

$\mathbf{v} \leftarrow \frac{\mathbf{v}_{t+1}}{1 - \beta_2^t}$

$\mathbf{z}_{t+1}^i \leftarrow \mathbf{z}_t^i - \frac{\alpha}{\gamma} (\mathbf{z}_t^i - \mathbf{m}) \div \mathbf{v} + \sqrt{\frac{2\alpha}{\gamma \kappa_h}} \eta_t, \eta_t \sim \mathcal{N}(0, 1)$

$t \leftarrow t + 1$

end while

 Save the training dataset $\mathcal{D}_j = \{\mathbf{z}_t^i, \mathbf{F}_t^i\}_{t=0}^T$

 Optimize θ_{j+1} using the generated training set \mathcal{D}_l for $l = 0, \dots, j$.

$j \leftarrow j + 1$

end while

4.3 Application of Consensus-Based Sampling to Biomolecular Systems

4.3.1 One-dimensional Rastrigin function

To illustrate the essential idea of the present method, we start with the 1-dimensional Rastrigin function:

$$f(x) = x^2 - \cos(2\pi x), \quad x \in [-3, 3]. \quad (4.16)$$

Instead of a neural network, we simply use a piecewise polynomial function $f_{\theta}(x)$ for this 1D

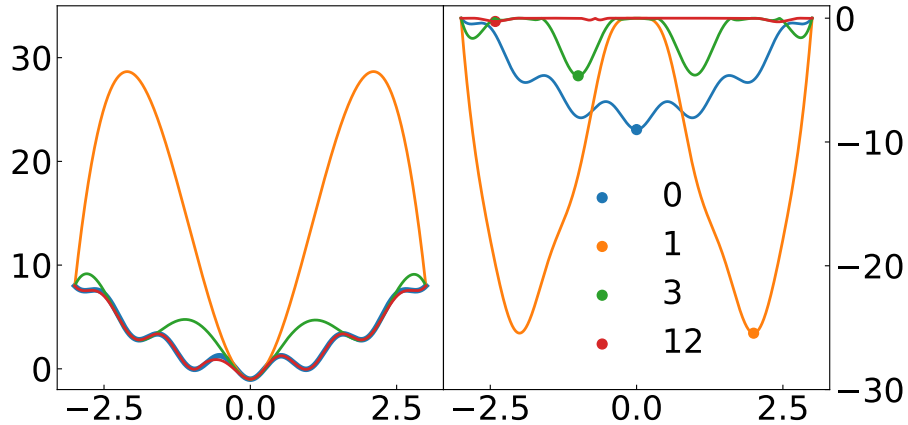


Figure 4.2 Adaptive sampling and construction of the 1D Rastrigin function. Left: The reference function $f(x)$ and the constructed approximations $f_{\theta_i}(x)$ obtained at different iteration steps. The relative l_2 error is less than 6×10^{-3} after 12 iterations. Right: The residual function $|f(x) - f_{\theta_i}(x)|$. The symbols represent the locations identified by each adaptive sampling (i.e., the maximization) step where new points will be added for the next training (i.e., the minimization) step.

problem, where θ represents the fitting parameters. Accordingly, the residual is directly defined as $|f - f_{\theta_i}|$. Initially, we set $f_{\theta_0}(x) \equiv 8$ with consistent boundary condition. We use the proposed sampling method with 10 walkers to estimate the first and second moments, yielding $m_1 = 1 \times 10^{-4}$ and $V_1 = 0.022$ and therefore the first and second derivative $f'(m_1) = 0$ and $f''(m_1) = \frac{1}{V_1} = 40.97$. This value is extremely close to the actual minimizer at 0 (i.e., the max-residue point) and the second derivative $f''(m_1) = 41.47$. Accordingly, we add new data points near $x = 1 \times 10^{-4}$ into the training set consisting of the boundary points $f(-3) = f(3) = 8$, which enables us to construct an updated approximation $f_{\theta_1}(x)$. Similarly, with the approximation $f_{\theta_{i-1}}(x)$, we conduct another sampling process (i.e., the maximization step) and include the obtained samples near m_i into the training set, yielding an updated approximation $f_{\theta_i}(x)$. As shown in Fig. 4.2, for each iteration, the sampling step can pinpoint the regime where the error is most pronounced. Furthermore, as shown in the Supplementary Information (SI), the second momentum near the maximum region can be estimated correctly as well. The underlying function $f(x)$ can be fully recovered after 12 iterations.

4.3.2 Two-dimensional FES

We use the alanine dipeptide (Ace-Ala-Nme), referred to as Ala2, as a benchmark problem. The molecule is solvated in 383 TIP3P water molecules similar to Ref. Zhang et al. (2018d). The full MD system is simulated in a canonical ensemble under temperature 300K using the Amber99-SB force field Hornak et al. (2006) with a time step is 2 fs. We refer to the SI for the simulation details.

We choose the CVs as the torsion angles: ϕ (C, N, C $_{\alpha}$, C) and ψ (N, C $_{\alpha}$, C, N). For comparison, we also construct the FESs using the VES Bonati et al. (2019) and RiD Zhang et al. (2018d) method with the same setup parameters presented therein. For the CES method, we use 10 walkers to explore the configuration space. The initial points of the walkers at the k th iteration are chosen to be the final 10 points of a biased dynamics with the bias potential $-A_{\mathcal{N}}(\mathbf{z}; \theta_{k-1})$. In the sampling stage, the inverse of the low and high temperatures are set to be $\kappa_l = 10$ and $\kappa_h = 1$, respectively. For each sample point, restrained dynamics is conducted with $\kappa = 500$ for 5000 steps to compute the average force. The timestep α for the dynamics of the random walkers (4.12) is set to be 0.1. Fig. 4.3 shows the FESs constructed by the three different methods and the reference obtained by the metadynamics Laio and Parrinello (2002) using a long simulation time. As presented in Fig. 4.3e, the CES method yields smaller approximation error and meanwhile requires lower computational cost. The better performance is not unexpected since the sample points can be adaptively optimized based on the construction error.

It is worth mentioning that the VES method was initially developed to achieve efficient enhanced sampling in the high-dimensional phase space; it is somewhat unfair to use it for direct FES construction since the estimation of the probability density function from the obtained samples could be numerically challenging for high-dimensional cases. On the other hand, the present CES method enables the explicit construction of FES if the analytical form is needed. Also, we note that the accuracy of the RiD method can be further improved using more iterations at the cost of larger simulation and training overhead. Below, we test the methods in more complex systems.

Table 4.1 The accuracy of the constructed 2D FES (the Ala2 molecule) and computational time (in hours, the same below) for the VES, RiD, and CES methods. The l_2 and l_∞ error are computed up to 40 KJ/mol. The reference solution is constructed by the metadynamics. The simulation time of the CES method is multiplied by 10 since 10 walkers are used.

Method	Accuracy		Time	
	l_2 error	l_∞ error	Simulation	Train
VES	5.39	21.03	47.5	
RiD	3.15	11.04	17.98	0.22 (GPU)
CES	1.88	10.68	0.23×10	0.18 (CPU)
				0.13 (GPU)

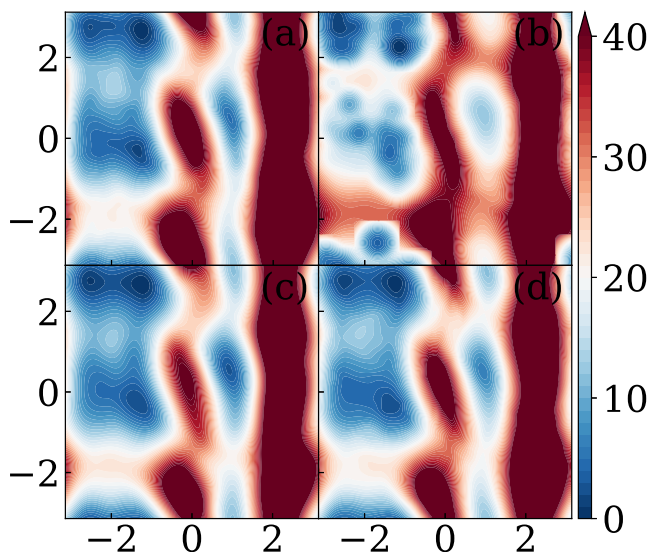


Figure 4.3 The 2D FES for the Ala2 molecule on the $\phi - \psi$ plane constructed by (a) Metadynamics (reference) (b) VES (c) RiD (d) CES (the present method). The accuracy and the computational cost are shown in Table 4.1.

4.3.3 Two-dimensional FES

4.3.4 Three-dimensional FES

Next, we consider a s-(1)-phenylethyl (s1pe) peptoid in an aqueous environment similar to Refs. Weiser and Santiso (2019); Wang et al. (2020a). The full system consists of one biomolecule and 546 water molecules in a $(2.9nm)^3$ dodecahedron box. The CHARMM general force field (CGenFF) Weiser and Santiso (2019) is used for the biomolecule and the TIP3P model Jorgensen et al. (1983) is used for the water molecules. The system temperature is set to be 298 K with a time step is 2 fs.

We refer to the SI for details.

The CVs are the three torsion angles ω spanned by atoms (C_α , C, N, C_α), ϕ , and ψ , where the latter two are the same as the Ala2 molecule. The FES is constructed by both the RiD and CES methods. The setup of the RiD method is the same as Wang et al. (2020a). For the CES method, we use 20 walkers and set the inverse of the low and high temperatures to be $\kappa_l = 10$ and $\kappa_h = 2$, respectively. The initial points of these walkers at each iteration are chosen in the same method as the two-dimensional problem. The timestep α for the dynamics of the random walkers (4.12) is set to be 0.1. For each sample point, restrained dynamics with $\kappa = 500$ is conducted for 10000 steps to compute the mean force.

For visualization, we project the constructed FES onto a two-dimensional plane and fix the third variable. Fig. 4.4 shows the projected FES on the $\omega - \phi$ and $\phi - \psi$ plane obtained from the CES and the RiD methods. For each projection, the reference is constructed as a 2D FES using the metadynamics Laio and Parrinello (2002). Similar to the previous 2D case, the present CES method yields higher accuracy with lower computational cost.

4.3.5 Nine-dimensional FES

Furthermore, we consider a more complex molecule, the peptoid trimer (s1pe)₃, solvated in a (4.2 nm)³ dodecahedron box with 1622 TIP3P water molecules. The force field and other simulation setups are similar to the s1pe molecule.

The chosen CVs consist of the 9 torsion angles ω , ϕ , ψ that are defined in s1pe case, associated with the different C_α atoms and denoted as $\omega_1, \phi_1, \psi_1, \omega_2, \phi_2, \psi_2, \omega_3, \phi_3, \psi_3$. We use 64 walkers for this case and the initial conditions of the walkers at each iteration are chosen with the same method as the previous cases. The inverse of the low and the high temperature are set to be $\kappa_l = 100$ and $\kappa_h = 2$, respectively. The timestep α of the dynamics of the random walkers is set to be 0.1. The FES is constructed by the CSE method using 28 iterations of sampling and training, which requires $225.53 \times 64 = 14434.35$ CPU hours for simulation and 6.06 GPU hours for training. For comparison, the RiD method uses 17900 CPU hours for simulation and 15.44 GPU hours for training. Similar to the above 3D problem, the constructed 9D FES is projected onto various 2D

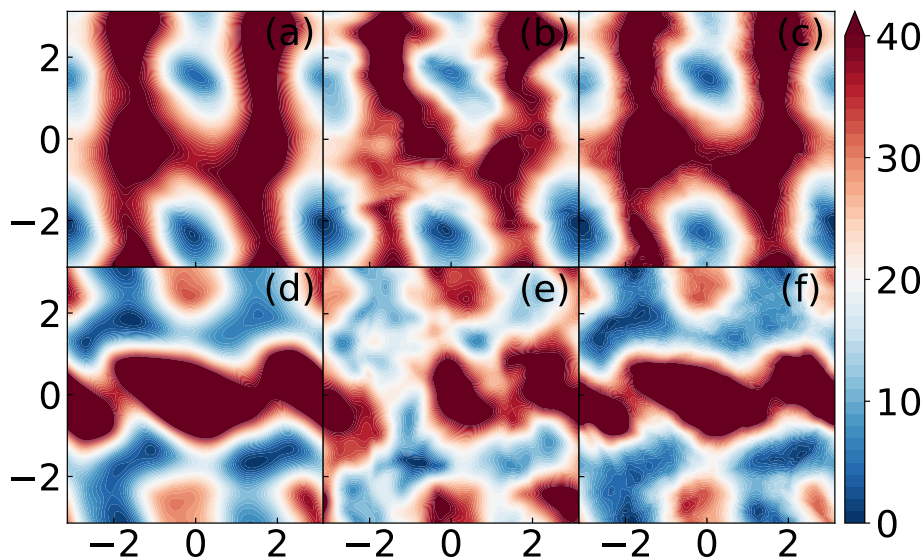


Figure 4.4 The 3D FES for molecule s1pe projected on the $\omega - \phi$ with $\psi = 1.5$ (first row) and $\phi - \psi$ planes with $\omega = 1.5$ (second-row). (a-d) The 2D FES constructed using metadynamics with the third variable restrained (reference); (b-e) projection of the 3D FES constructed by the RiD method; (c-f) projection of the 3D FES constructed by the present CES method. The CES method requires 4.81×20 CPU hours for sampling and 0.84 GPU hours for training, whereas the RiD method uses 423.33 CPU hours and 8 GPU hours, respectively. For a more detailed quantitative analysis, we refer to the SI.

planes with the remaining variables fixed. For each 2D projection, the reference is constructed as a 2D FES using metadynamics. Fig. 4.5 shows the projection on the $\omega_1 - \phi_1$ and $\omega_1 - \psi_1$ plane (see SI for visualization of other 2D projections). Compared with the Rid method, the present CES method yields higher accuracy with lower computational cost.

The different performances could be possibly due to the distinct ways of imposing the sampling adaptivity for the two methods. Specifically, the Rid method trains a replica of DNNs on the same sample set and uses the standard deviation of multiple DNNs' predictions as an indirect measure of the construction error. The significant sampling error of the mean force (i.e., $-\nabla A(\mathbf{z})$) could be overlooked by using the standard deviation as the uncertainty indicator, and therefore may further propagate into the construction error. For instance, a biased sampling of mean force may lead to a consistent biased prediction among multiple DNNs. As a result, even if the standard deviation of multiple DNNs' predictions for a sample point is small, the construction error could be still

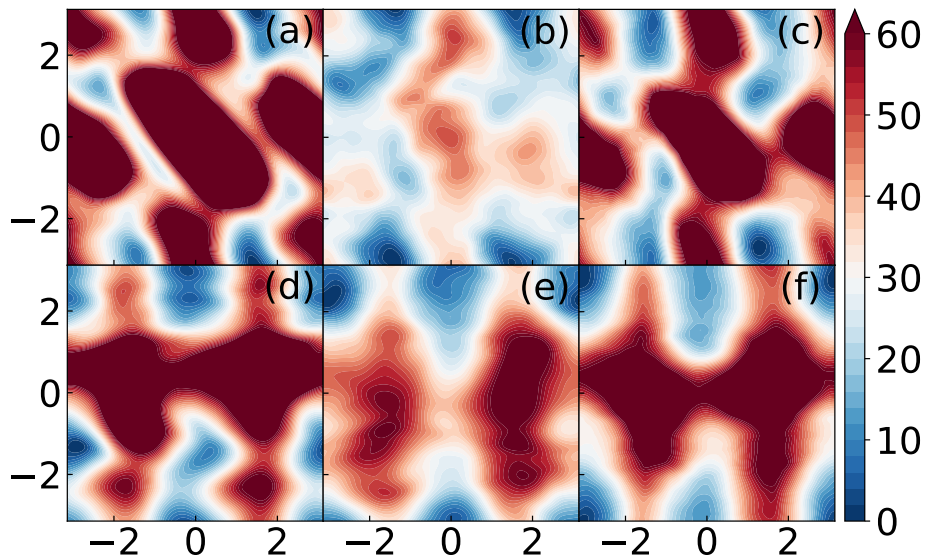


Figure 4.5 The 9D FES for molecule (s1pe)₃ projected on the $\omega_1 - \phi_1$ (first row) and $\omega_1 - \psi_1$ plane (second row). (a-d) The 2D FES constructed using metadynamics with the remaining variable restrained (reference); (b-e) projection of the 9D FES constructed by the RiD method using 17900 CPU hours for sampling and 15.44 GPU hours for training; (c-f) projection of the 9D FES constructed by the CES method using 14434.35 CPU hours and 6.06 GPU hours, respectively.

pronounced in that region. Instead, the CES method directly uses the construction error to impose the adaptivity and therefore achieve enhanced sampling of the regions lacking accuracy.

4.3.6 Thirty-dimensional FES

Finally, we consider the polyaniline-15 (Ace-(Ala)₁₅-Nme), denoted as Ala16, solvated in 2258 water molecules. The full system is simulated temperature 300K in a (4.62 nm)³ dodecahedron box using the Amber99-SB force field with a time step is 2 fs. We refer to the SI for the simulation details.

The chosen CVs consist of torsion angles ϕ and ψ , defined in Ala2 case, associated with the different C_αs, denoted as $\{\phi_i, \psi_i\}_{i=1}^{15}$. We use 64 walkers for this case and the initial value of the walkers at each iteration is given by the biased simulation before. The inverse of the low and high temperatures are set to be $\kappa_l = 20$ and $\kappa_h = 5$, respectively. The timestep α for dynamics of the random walkers Eq. (4.12) is set to be 0.1. The FES is constructed using 100 iterations of sampling and training. Similar to the previous case, the obtained FES is plotted on a two-dimensional plane

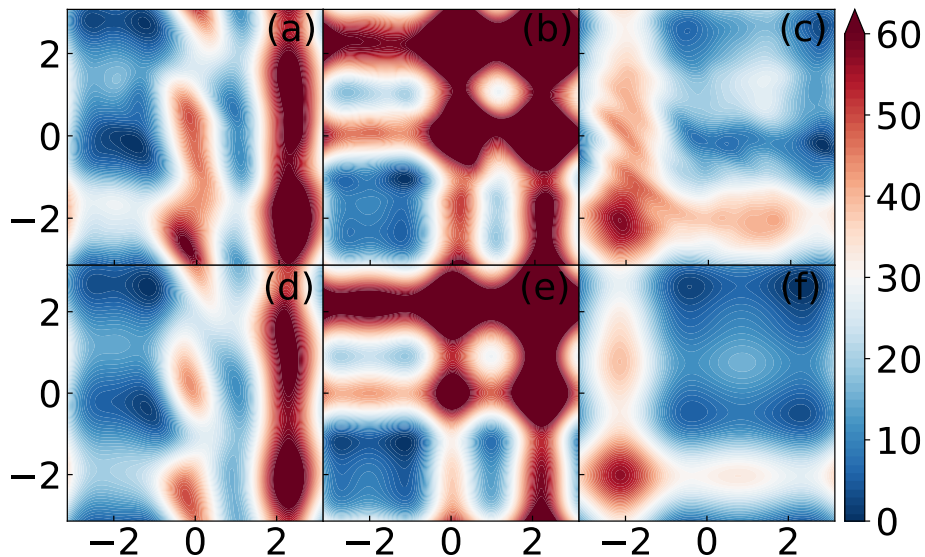


Figure 4.6 The 3D FES for molecule Ala16 projected on various 2D planes. (a-d) $\phi_1 - \psi_1$ (b-e) $\phi_2 - \phi_3$ (c-f) $\psi_2 - \psi_3$. (a-b-c) 2D FES constructed by metadynamics (reference) (d-e-f) The 3D FES constructed by the present CES method.

while the remaining variables are fixed and the reference is constructed as a 2D FES using the metadynamics. Fig. 4.6 shows the projection on the $\phi_1 - \psi_1$, $\phi_2 - \phi_3$ and $\psi_2 - \psi_3$ plane. For all the cases, the projection of the 30-dimensional FES shows good agreement with the 2-dimensional reference solution. We have also examined the projection on other planes; the prediction shows good agreement with the reference solution as well. We refer to the SI for details.

4.4 Conclusions

We have presented a consensus-based approach for constructing high-dimensional FESs by reformulating the construction task as a minimax optimization problem. Rather than seeking the direct fitting, this method essentially establishes an adversarial learning of FESs by simultaneously optimizing the target function approximation and the training set. While the common approaches mainly focus on the efficient exploration of the phase space in the presence of local minima, the present method further accounts for the discretization error that has been broadly overlooked in FES construction. Adaptive sampling of the max-residue regime is achieved through the consensus-based sampling of a posteriori residue-induced distribution in the form of a stochastic particle system in

the CV space. Given the fact the sampling only relies on the first and second-moment estimation, the method could be particularly efficient for high-dimensional problems. While the numerical results of biomolecular systems have demonstrated the effectiveness for FES construction, the present framework of unifying the residual minimization and max-residue enhanced sampling is quite general for model reduction of complex systems, e.g., the stochastic reduced dynamics (e.g., see Ref. She et al. (2023)) with a state-dependent memory for multiscale physical systems.

CHAPTER 5

CONCLUSION AND OUTLOOK

In this thesis, we aim to address the challenges of mesoscale modeling by tackling the problem from three key perspectives: model representation, collective variable selection, and data collection. By integrating these techniques, we strive to develop a comprehensive framework for constructing faithful models that accurately capture phenomena across different scales.

In model representation part, we propose a machine learning-based memory kernel to model the many-body nature of dissipative interactions. This approach is designed to respect physical constraints, such as symmetry, and ensures thermodynamic consistency by strictly satisfying the second fluctuation-dissipation theorem. By embedding these fundamental principles into the model, we aim to achieve accurate and physically meaningful representations of complex systems.

In collective variable selection part, we introduce a data-driven method to identify collective variables that capture the principal differences between distributions under various nonequilibrium conditions. By treating these differences as CVs, our approach can generalize across different external force fields, making it adaptable to a wide range of nonequilibrium scenarios. This step is crucial for reducing the dimensionality of the problem while preserving the essential dynamics of the system.

In data collection section, to address the challenges of sampling in high-dimensional phase spaces, we reformulate the problem as a minimax optimization. This approach balances exploration of uncharted regions of the phase space with exploitation of areas where the model exhibits high error. By iteratively refining the training data, we ensure that the sampling process is both comprehensive and focused on regions critical for improving model accuracy. By integrating these three perspectives, we aim to develop a unified framework for mesoscale modeling that bridges the gap between microscale interactions and macroscale dynamics.

These challenges are not confined to modeling problems between the microscale and mesoscale. We also aim to extend our framework to address modeling challenges between the mesoscale and macroscale, such as non-Fourier heat conduction and the diffusion of active materials. In these

applications, we strive to develop models that are not only physically consistent and accurate but also mathematically well-posed. While this is relatively straightforward for ordinary differential equations or stochastic differential equations for models we construct here, it becomes significantly more challenging for partial differential equations or stochastic partial differential equations. We hope to address this by carefully designing the model structure to guarantee existence, uniqueness, and stability of solutions, which are essential for reliable simulations and predictions.

Recent advancements in generative models, such as generative adversarial networks (GANs), variational autoencoders (VAEs), and diffusion models, have demonstrated remarkable capabilities in modeling and sampling from high-dimensional distributions. Despite the explicit connection with the sampling problems we presented ahead, a natural and intriguing question arises: Can these generative models be extended to generate high-dimensional stochastic processes? Extending generative models to generate high-dimensional stochastic processes that simultaneously satisfy physical symmetries and thermodynamic consistency is a challenging but highly promising direction.

BIBLIOGRAPHY

- Abrams, C. F. and Vanden-Eijnden, E. (2010). Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. Proceedings of the National Academy of Sciences, 107(11):4961–4966.
- 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.
- Ahn, S.-H., Grate, J. W., and Darve, E. F. (2017). Efficiently sampling conformations and pathways using the concurrent adaptive sampling (cas) algorithm. The Journal of Chemical Physics, 147(7).
- Ahn, S.-H., Ojha, A. A., Amaro, R. E., and McCammon, J. A. (2021). Gaussian-accelerated molecular dynamics with the weighted ensemble method: A hybrid method improves thermodynamic and kinetic sampling. Journal of Chemical Theory and Computation, 17(12):7938–7951.
- Allen, M. P. and Tildesley, D. J. (2017). Computer simulation of liquids. Oxford university press.
- Anderson, P. W. (1972). More is different: Broken symmetry and the nature of the hierarchical structure of science. Science, 177(4047):393–396.
- 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.
- Bao, G., Ye, X., Zang, Y., and Zhou, H. (2020). Numerical solution of inverse problems by weak adversarial networks. Inverse Problems, 36(11):115003.
- Barducci, A., Bussi, G., and Parrinello, M. (2008). Well-tempered metadynamics: a smoothly converging and tunable free-energy method. Physical review letters, 100(2):020603.
- Baxevari, G., Harmandaris, V., Kalligiannaki, E., and Tsantili, I. (2023). Bottom-up transient time models in coarse-graining molecular systems. Multiscale Modeling & Simulation, 21(4):1746–1774.
- Behler, J. and Parrinello, M. (2007). Generalized neural-network representation of high-dimensional potential-energy surfaces. Phys. Rev. Lett., 98:146401.
- Best, R. B. and Hummer, G. (2010). Coordinate-dependent diffusion in protein folding. Proceedings of the National Academy of Sciences, 107(3):1088–1093.
- Bonati, L., Zhang, Y.-Y., and Parrinello, M. (2019). Neural networks-based variationally enhanced sampling. Proceedings of the National Academy of Sciences, 116(36):17641–17647.

- Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. The Journal of chemical physics, 126(1):014101.
- Carrillo, J. A., Choi, Y.-P., Totzeck, C., and Tse, O. (2018). An analytical framework for consensus-based global optimization method. Mathematical Models and Methods in Applied Sciences, 28(06):1037–1066.
- Carrillo, J. A., Hoffmann, F., Stuart, A. M., and Vaes, U. (2022). Consensus-based sampling. Studies in Applied Mathematics, 148(3):1069–1140.
- Carrillo, J. A., Jin, S., Li, L., and Zhu, Y. (2021). A consensus-based global optimization method for high dimensional machine learning problems. ESAIM: Control, Optimisation and Calculus of Variations, 27:S5.
- Cendagorta, J. R., Tolpin, J., Schneider, E., Topper, R. Q., and Tuckerman, M. E. (2020). Comparison of the performance of machine learning models in representing high-dimensional free energy surfaces and generating observables. The Journal of Physical Chemistry B, 124(18):3647–3660.
- 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.
- Chen, J., Jin, S., and Lyu, L. (2022). A consensus-based global optimization method with adaptive momentum estimation. Communications in Computational Physics, 31(4):1296–1316.
- Chen, K. K., Tu, J. H., and Rowley, C. W. (2012). Variants of dynamic mode decomposition: boundary condition, koopman, and fourier analyses. Journal of nonlinear science, 22:887–915.
- Chipot, C. and Lelièvre, T. (2011). Enhanced sampling of multidimensional free-energy landscapes using adaptive biasing forces. SIAM Journal on Applied Mathematics, 71(5):1673–1695.
- Chmiela, S., Tkatchenko, A., Sauceda, H. E., Poltavsky, I., Schütt, K. T., and Müller, K.-R. (2017). Machine learning of accurate energy-conserving molecular force fields. Science advances, 3(5):e1603015.
- Coifman, R. R., Kevrekidis, I. G., Lafon, S., Maggioni, M., and Nadler, B. (2008). Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems. Multiscale Modeling & Simulation, 7(2):842–864.
- Crommelin, D. and Vanden-Eijnden, E. (2011). Diffusion estimation from multiscale data by operator eigenpairs. Multiscale Modeling & Simulation, 9(4):1588–1623.
- Crosskey, M. and Maggioni, M. (2017). Atlas: a geometric approach to learning high-dimensional stochastic systems near manifolds. Multiscale Modeling & Simulation, 15(1):110–156.
- 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.
- Darve, E. and Pohorille, A. (2001). Calculating free energies using average force. The Journal of Chemical Physics, 115(20):9169–9183.
- Darve, E., Rodríguez-Gómez, D., and Pohorille, A. (2008). Adaptive biasing force method for scalar and vector free energy calculations. The Journal of Chemical Physics, 128(14).
- 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. (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.
- Einstein, A. (1905). On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. Annalen der Physik, 17:549–560.
- Español, P. and Warren, P. (1995). Statistical mechanics of dissipative particle dynamics. Europhysics Letters, 30(4):191–196.
- Fang, L., Ge, P., Zhang, L., E, W., and Lei, H. (2022). DeePN²: A deep learning-based non-Newtonian hydrodynamic model. Journal of Machine Learning, 1:114–140.
- Frenkel, D. and Smit, B. (2023). Understanding molecular simulation: from algorithms to applications. Elsevier.
- Gabrié, M., Rotskoff, G. M., and Vanden-Eijnden, E. (2022). Adaptive monte carlo augmented with normalizing flows. Proceedings of the National Academy of Sciences, 119(10):e2109420119.
- Gao, X., Xu, Z. Q., and Zhou, X. Y. (2022). State-dependent temperature control for langevin diffusions. SIAM Journal on Control and Optimization, 60(3):1250–1268.
- Gao, Y. Q. (2008). An integrate-over-temperature approach for enhanced sampling. The Journal of Chemical Physics, 128(6).
- Ge, P., Zhang, L., and Lei, H. (2023). Machine learning assisted coarse-grained molecular dynamics modeling of meso-scale interfacial fluids. J. Chem. Phys., 158:064104.
- 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.

- Grafke, T. and Laio, A. (2024). Metadynamics for transition paths in irreversible dynamics. Multiscale Modeling & Simulation, 22(1):125–141.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. (2017). Improved training of Wasserstein GANs. Advances in neural information processing systems, 30.
- Hamelberg, D., Mongan, J., and McCammon, J. A. (2004). Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. The Journal of Chemical Physics, 120(24):11919–11929.
- Hansen, J. P. and McDonald, I. (1990). Theory of Simple Liquids. Academic, London.
- Hess, B., Bekker, H., Berendsen, H. J., and Fraaije, J. G. (1997). Lincs: A linear constraint solver for molecular simulations. Journal of Computational Chemistry, 18(12):1463–1472.
- 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.
- 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.
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins: Structure, Function, and Bioinformatics, 65(3):712–725.
- Hudson, T. and Li, X. H. (2020). Coarse-graining of overdamped Langevin dynamics via the Mori–Zwanzig formalism. Multiscale Modeling & Simulation, 18(2):1113–1135.
- 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.
- 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, G., Hanke, M., and Schmid, F. (2017). Iterative reconstruction of memory kernels. Journal of Chemical Theory and Computation, 13(6):2481–2488.
- Kampen, N. V. (2007). Stochastic processes in physics and chemistry. North Holland.

- Kawasaki, K. and Gunton, J. D. (1973). Theory of nonlinear transport processes: Nonlinear shear viscosity and normal stress effects. Physical Review A, 8(4):2048.
- Kim, J., Straub, J. E., and Keyes, T. (2006). Statistical-temperature monte carlo and molecular dynamics algorithms. Physical review letters, 97(5):050601.
- 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.
- Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. Journal of Computational Chemistry, 13(8):1011–1021.
- Kutz, J. N., Brunton, S. L., Brunton, B. W., and Proctor, J. L. (2016). Dynamic mode decomposition: data-driven modeling of complex systems. SIAM.
- Laio, A. and Parrinello, M. (2002). Escaping free-energy minima. Proceedings of the National Academy of Sciences, 99(20):12562–12566.
- Lange, O. F. and Grubmüller, H. (2006). Collective Langevin dynamics of conformational motions in proteins. J. Chem. Phys., 124:214903.
- 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):174113.
- 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. Physical Review E, 81(2).
- Lei, H., Wu, L., and E, W. (2020). Machine learning based non-newtonian fluid model with molecular fidelity. Physics Review E, 102:043309.
- Lelièvre, T., Rousset, M., and Stoltz, G. (2008). Long-time convergence of an adaptive biasing force

- method. Nonlinearity, 21(6):1155.
- 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., Lee, H. S., Darve, E., and Karniadakis, G. E. (2017). Computing the non-Markovian coarse-grained interactions derived from the Mori–Zwanzig formalism in molecular systems: Application to polymer melts. The Journal of Chemical Physics, 146(1):014104.
- Lindahl, Abraham, Hess, and van der Spoel (2019). Gromacs 2019.2 source code.
- Lu, J. and Vanden-Eijnden, E. (2013). Infinite swapping replica exchange molecular dynamics leads to a simple simulation patch using mixture potentials. The Journal of Chemical Physics, 138(8).
- 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. (2023a). Consensus-based construction of high-dimensional free energy surface. arXiv preprint arXiv:2311.05009.
- Lyu, L. and Lei, H. (2023b). Construction of coarse-grained molecular dynamics with many-body non-markovian memory. Physical Review Letters, 131(17):177301.
- Lyu, L. and Lei, H. (2024). On the generalization ability of coarse-grained molecular dynamics models for non-equilibrium processes. arXiv preprint arXiv:2409.11519.
- 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. (2016). Derivation and approximation of coarse-grained dynamics from Langevin dynamics. Arxiv preprint, page 1605.04886.
- Ma, L., Li, X., and Liu, C. (2019). Coarse-graining langevin dynamics using reduced-order techniques. Journal of Computational Physics, 380:170–190.
- Ma, Z., Wang, S., Kim, M., Liu, K., Chen, C.-L., and Pan, W. (2021). Transfer learning of memory kernels for transferable coarse-graining of polymer dynamics. Soft Matter, 17(24):5864–5877.
- 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(1-3):168–175.
- Maragliano, L. and Vanden-Eijnden, E. (2008). Single-sweep methods for free energy calculations. The Journal of Chemical Physics, 128(18):184110.

- Marinari, E. and Parisi, G. (1992). Simulated tempering: a new monte carlo scheme. Europhysics letters, 19(6):451.
- Meyer, H., Pelagejcev, P., and Schilling, T. (2019). Non-markovian out-of-equilibrium dynamics: A general numerical procedure to construct time-dependent memory kernels for coarse-grained observables. Europhysics Letters, 128(4):40001.
- Miao, Y., Feher, V. A., and McCammon, J. A. (2015). Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. Journal of Chemical Theory and Computation, 11(8):3584–3595.
- Miyamoto, S. and Kollman, P. A. (1992). Settle: An analytical version of the shake and rattle algorithm for rigid water models. Journal of Computational Chemistry, 13(8):952–962.
- Miyato, T., Kataoka, T., Koyama, M., and Yoshida, Y. (2018). Spectral normalization for generative adversarial networks. arXiv preprint arXiv:1802.05957.
- Molgedey, L. and Schuster, H. G. (1994). Separation of a mixture of independent signals using time delayed correlations. Physical review letters, 72(23):3634.
- 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.
- Mori, H. (1965). Transport, collective motion, and Brownian motion. Progress of Theoretical Physics, 33(3):423–455.
- 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.
- 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., Olsson, S., Köhler, J., and Wu, H. (2019). Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. Science, 365(6457):eaaw1147.
- 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.
- Parrinello, M. and Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied Physics, 52(12):7182–7190.
- Pérez-Hernández, G., Paul, F., Giorgino, T., De Fabritiis, G., and Noé, F. (2013). Identification of slow molecular order parameters for markov model construction. The Journal of chemical physics, 139(1).
- Pfaendtner, J. and Bonomi, M. (2015). Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics. Journal of chemical theory and computation, 11(11):5062–5067.
- Piana, S. and Laio, A. (2007). A bias-exchange approach to protein folding. The journal of physical chemistry B, 111(17):4553–4559.
- Pinnau, R., Totzeck, C., Tse, O., and Martin, S. (2017). A consensus-based model for global optimization and its mean-field limit. Mathematical Models and Methods in Applied Sciences, 27(01):183–204.
- Plotkin, S. S. and Wolynes, P. G. (1998). Non-Markovian configurational diffusion and reaction coordinates for protein folding. Phys. Rev. Lett., 80:5015–5018.
- Prakash, A., Fu, C. D., Bonomi, M., and Pfaendtner, J. (2018). Biasing smarter, not harder, by partitioning collective variables into families in parallel bias metadynamics. Journal of chemical theory and computation, 14(10):4985–4990.
- Reith, D., Pütz, M., and Müller-Plathe, F. (2003). Deriving effective mesoscale potentials from atomistic simulations. Journal of Computational Chemistry, 24(13):1624–1636.
- 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.
- 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.
- 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.
- 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.
- Schmid, P. J. (2010). Dynamic mode decomposition of numerical and experimental data. Journal of

fluid mechanics, 656:5–28.

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. Physical review letters, 119(15):150601.

Schwantes, C. R. and Pande, V. S. (2013). Improvements in markov state model construction reveal many non-native interactions in the folding of ntl9. Journal of chemical theory and computation, 9(4):2000–2009.

Shaffer, P., Valsson, O., and Parrinello, M. (2016). Enhanced, targeted sampling of high-dimensional free-energy landscapes using variationally enhanced sampling, with an application to chignolin. Proceedings of the National Academy of Sciences, 113(5):1150–1155.

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.

Shell, M. S. (2008). The relative entropy is fundamental to multiscale and inverse thermodynamic problems. The Journal of Chemical Physics, 129(14).

Sidky, H. and Whitmer, J. K. (2018). Learning free energy landscapes using artificial neural networks. The Journal of chemical physics, 148(10):104111.

Soper, A. (1996). Empirical potential monte carlo simulation of fluid structure. Chemical Physics, 202(2):295–306.

Stecher, T., Bernstein, N., and Csányi, G. (2014a). Free energy surface reconstruction from umbrella samples using gaussian process regression. Journal of Chemical Theory and Computation, 10(9):4079–4097.

Stecher, T., Bernstein, N., and Csányi, G. (2014b). Free energy surface reconstruction from umbrella samples using gaussian process regression. Journal of chemical theory and computation, 10(9):4079–4097.

Stoltz, G., Rousset, M., et al. (2010). Free energy computations: A mathematical perspective. World Scientific.

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.

Sugita, Y. and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. Chemical physics letters, 314(1-2):141–151.

Tang, K., Zhai, J., Wan, X., and Yang, C. (2023). Adversarial adaptive sampling: Unify pinn and optimal transport for the approximation of pdes. arXiv preprint arXiv:2305.18702.

- Torrie, G. and Valleau, J. (1977a). Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. Journal of Computational Physics, 23(2):187–199.
- Torrie, G. M. and Valleau, J. P. (1977b). Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. Journal of Computational Physics, 23(2):187–199.
- Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C., and Bussi, G. (2014). Plumed 2: New feathers for an old bird. Computer physics communications, 185(2):604–613.
- Valsson, O. and Parrinello, M. (2014). Variational approach to enhanced sampling and free energy calculations. Phys. Rev. Lett., 113:090601.
- van der Oord, C., Sachs, M., Kovács, D. P., Ortner, C., and Csányi, G. (2023). Hyperactive learning for data-driven interatomic potentials. npj Computational Materials, 9(1):168.
- Van Hove, L. (1954). Correlations in space and time and born approximation scattering in systems of interacting particles. Phys. Rev., 95:249–262.
- Voter, A. F. (1997). Hyperdynamics: Accelerated molecular dynamics of infrequent events. Physical Review Letters, 78(20):3908.
- 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.
- Wang, D., Wang, Y., Chang, J., Zhang, L., Wang, H., and E, W. (2022). Efficient sampling of high-dimensional free energy landscapes using adaptive reinforced dynamics. Nature Computational Science, 2(1):20–29.
- Wang, H., Zariphopoulou, T., and Zhou, X. Y. (2020a). Reinforcement learning in continuous time and space: A stochastic control approach. The Journal of Machine Learning Research, 21(1):8145–8178.
- Wang, S., Ma, Z., and Pan, W. (2020b). Data-driven coarse-grained modeling of polymers in solution with structural and dynamic properties conserved. Soft Matter, 16(36):8330–8344.
- Weiser, L. J. and Santiso, E. E. (2019). A cgenff-based force field for simulations of peptoids with both cis and trans peptide bonds. Journal of computational chemistry, 40(22):1946–1956.
- Widder, C., Koch, F., and Schilling, T. (2022). Generalized langevin dynamics simulation with non-stationary memory kernels: How to make noise. The Journal of Chemical Physics, 157(19):194107.
- Willis, C. and Picard, R. (1974). Time-dependent projection-operator approach to master equations for coupled systems. Physical Review A, 9(3):1343.

- Wu, H. and Noé, F. (2014). Optimal estimation of free energies and stationary densities from multiple biased simulations. Multiscale Modeling & Simulation, 12(1):25–54.
- Xie, P., Car, R., and E, W. (2022). Ab Initio Generalized Langevin Equations. arXiv preprint arXiv:2211.06558.
- Yang, L. and Qin Gao, Y. (2009). A selective integrated tempering method. The Journal of Chemical Physics, 131(21).
- Yang, Y. I., Shao, Q., Zhang, J., Yang, L., and Gao, Y. Q. (2019). Enhanced sampling in molecular dynamics. The Journal of Chemical Physics, 151(7).
- Yoshimoto, Y., Kinefuchi, I., Mima, T., Fukushima, A., Tokumasu, T., and Takagi, S. (2013). Bottom-up construction of interaction models of non-markovian dissipative particle dynamics. Phys. Rev. E, 88:043305.
- 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.
- Yu, T.-Q., Lu, J., Abrams, C. F., and Vanden-Eijnden, E. (2016). Multiscale implementation of infinite-swap replica exchange molecular dynamics. Proceedings of the National Academy of Sciences, 113(42):11744–11749.
- Zhang, B. W., Jasnow, D., and Zuckerman, D. M. (2007). Efficient and verified simulation of a path ensemble for conformational change in a united-residue model of calmodulin. Proceedings of the National Academy of Sciences, 104(46):18043–18048.
- 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.
- Zheng, L., Chen, M., and Yang, W. (2008). Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. Proceedings of the National Academy of Sciences, 105(51):20227–20232.

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

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

APPENDIX A

SIMULATION DETAIL FOR NONEQUILIBRIUM CASE

A.1 Additional result of CGMD model under equilibrium state

To establish a fair comparison of the various CG models, we examine the static and dynamic properties that have been widely used as benchmark problems for model validation. Fig. A.1 shows the radial distribution function between the CG coordinates. For all three CG models, their predictions agree well with the full MD results, which verifies the accuracy of the CG free energy function $U(\mathbf{Q})$. Fig. A.2 shows the normalized velocity auto-correlation function (VACF) of the CG variables. Similar to the static properties, the prediction of the CG models agrees well with the full MD results, which verifies the accuracy of the many-body memory term $\mathbf{K}(\mathbf{Q}, t)$.

Despite of the good agreement, we emphasize that these properties are essentially defined under the marginal density near equilibrium. The applicability to non-equilibrium processes is generally unwarranted, and further relies on the consistency in the conditional probability density function associated with the CG projection operator (i.e., Eq. (3.6)) as discussed in Sec. 3.2.

A.2 CG variable training

In this section, we show that Eq. 3.8 does not impose further constraints on constructing the CG variables. We examine the equivalence of two CG mapping. For instance, when the centers of mass (COMs) are utilized as the only CG variables, the information they encompass is the same as the COMs plus a constant or the COMs multiplied by a constant.

Definition A.2.1. For two CG map ϕ and $\check{\phi}$, if there exist a map \mathbb{T} , such that $\phi(\mathbf{q}, \mathbf{p}) = \mathbb{T}(\check{\phi}(\mathbf{q}, \mathbf{p}))$, we say $\phi \leq \check{\phi}$. Also, if $\phi \leq \check{\phi}$ and $\check{\phi} \leq \phi$, we say that the two maps are equivalent.

In other words, if $\phi \leq \check{\phi}$, any properties that can be measured from the CG coordinate ϕ , denote as $\mathcal{F}(\phi)$, can also be measured from $\check{\phi}$ as $\mathcal{F}(\mathbb{T}(\check{\phi}))$. If $\phi \leq \check{\phi}$ and a CG model $\frac{d}{dt}\check{\phi} = \check{\mathcal{L}}\check{\phi}$ is constructed using CG coordinate $\check{\phi}$, the CG model for ϕ can be constructed by the chain rule $\frac{d}{dt}\phi = \frac{\delta T}{\delta \check{\phi}} \check{\mathcal{L}}\check{\phi}$ if $\check{\mathcal{L}}$ is deterministic or Itô's lemma if $\check{\mathcal{L}}$ includes stochastic term.

Proposition A.2.2. For any linear map ϕ defined by a matrix $\mathbf{W} \in \mathbb{R}^{N_m \times (m-1)}$, there exists a map

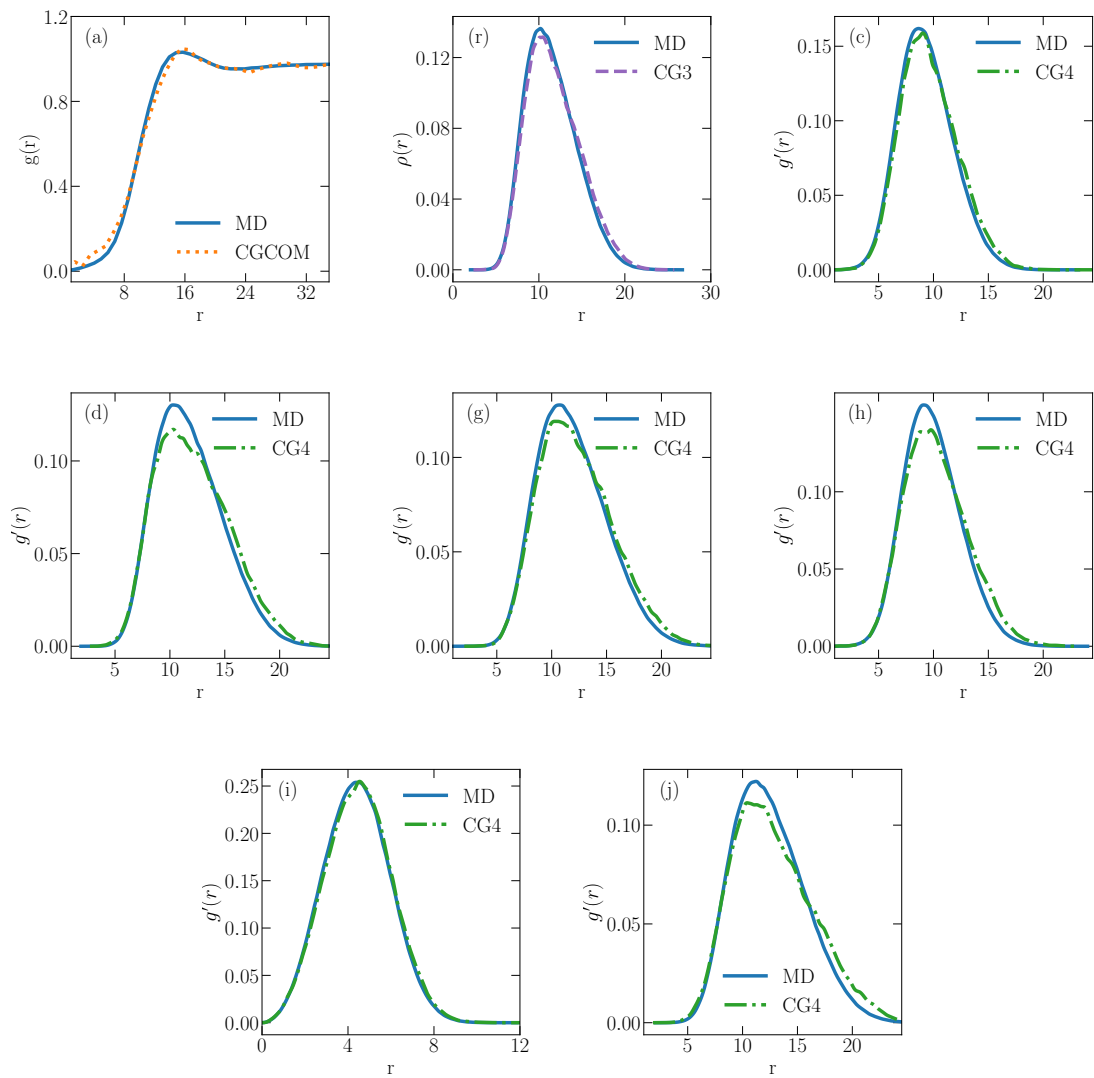


Figure A.1 The distribution of the pairwise distance between CG coordinates in the CGMD model compared with the Full MD model. (a) The radius distribution function $g(r)$ of the CGCOM model. (b) The distribution of the pairwise distance $g'(r)$ between two CG coordinates within the same molecular of the CG3 model. (c)-(j) The distribution of distance $g'(r)$ between two CG coordinate (1-2,1-3,1-4,2-3,2-4,3-4, respectively) within the same molecular of CG4 model.

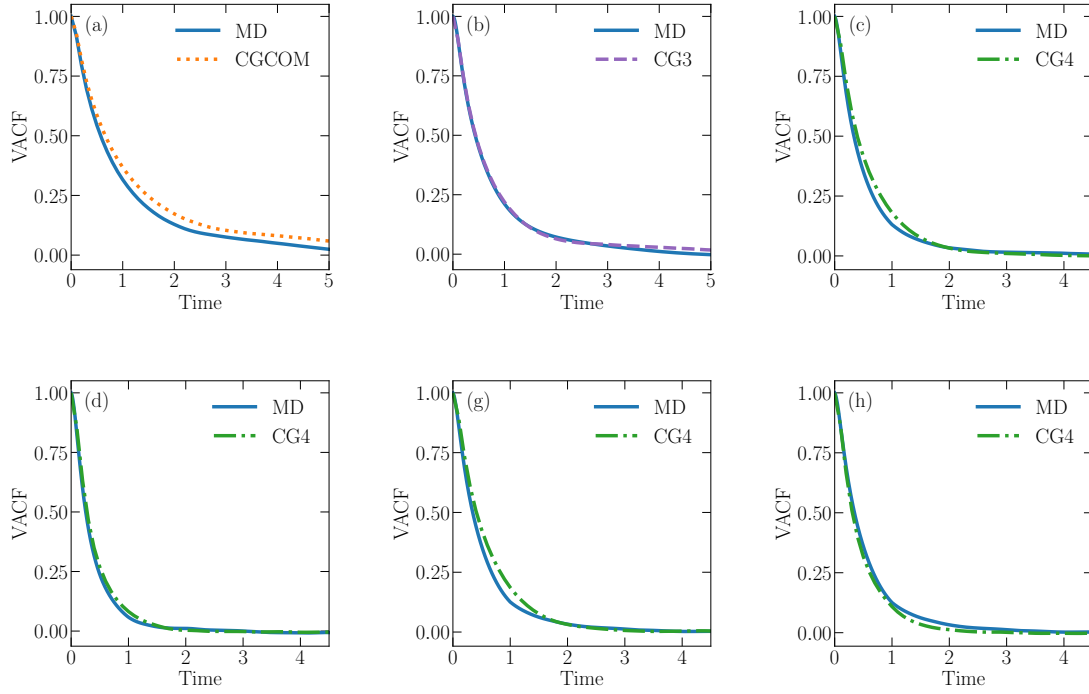


Figure A.2 The VACF of CVs in different CG models and their comparison with Full MD. (a) VACF of CVs in CG3 model and the comparison with Full MD. (b)-(g) VACF of CVs (1,2,3,4, respectively) in CG4 model and the comparison with Full MD.

$\check{\phi}$, defined by a matrix $\check{\mathbf{W}} \in \mathbb{R}^{N_m \times m}$, subject to the following constraints

$$0 < \check{w}_{ni} < 1, \quad \sum_{i=1}^m \check{w}_{ni} = 1. \quad (\text{A.1})$$

Proof. For any \mathbf{W} with entry w_{ij} , $\check{\mathbf{W}}$ can be constructed as

$$\check{w}_{ni} = \begin{cases} \frac{w_{ni}^1 - \tilde{w}}{\bar{w}}, & 0 < i < m, \\ 1 - \sum_{i=1}^{m-1} \check{w}_{ni}, & i = m, \end{cases} \quad (\text{A.2})$$

where $\tilde{w} = \min_{n,i} w_{ni}$ and $\bar{w} = \max_i \sum_{i=1}^{m-1} (w_{ni} - \tilde{w})$. It's obvious that $\check{w}_{ni} \in [0, 1]$ for $0 < i < m$.

Since

$$\sum_{i=1}^{m-1} \check{w}_{ni} = \frac{\sum_{i=1}^{m-1} (w_{ni} - \tilde{w})}{\bar{w}} \in [0, 1],$$

we have $\check{w}_{nm} \in [0, 1]$. Also, $\sum_i \check{w}_{ni} = 1$ for every n , hence $\check{\mathbf{W}}$ satisfies the constraints. Since

$$\sum_i^m \left(\sum_n^{N_m} \check{w}_{ni} \right) \check{\mathbf{Q}}_i^I = \sum_i^m \sum_n^{N_m} \check{w}_{ni} \mathbf{q}_n^I = \sum_n^{N_m} \mathbf{q}_n^I, \quad (\text{A.3})$$

we have

$$\begin{aligned}
\mathbf{Q}_i^I &= \frac{\sum_n^{N_m} w_{ni} \mathbf{q}_{I,n}}{\sum_n^{N_m} w_{ni}} \\
&= \frac{\sum_n^{N_m} (\bar{w} \check{w}_{ni} + \tilde{w}) \mathbf{q}_{I,n}}{\sum_n^{N_m} (\bar{w} \check{w}_{ni} + \tilde{w})} \\
&= \frac{\bar{w} (\sum_n^{N_m} \check{w}_{ni}) \check{\mathbf{Q}}_i^I + \tilde{w} \sum_j^m (\sum_n^{N_m} \check{w}_{nj}) \check{\mathbf{Q}}_j^I}{\sum_n^{N_m} (\bar{w} \check{w}_{ni} + \tilde{w})}.
\end{aligned} \tag{A.4}$$

Therefore, \mathbf{Q}^I is a linear transformation of $\check{\mathbf{Q}}^I$ for all $I = 1, \dots, M$. Similarly, \mathbf{P}^I can be shown as a linear transformation of $\check{\mathbf{P}}^I$. □

APPENDIX B

SIMULATION DETAIL AND PROOF FOR CONSENSUS BASED SAMPLING

B.1 Proof of Proposition 2

The sampling distribution is governed by the following McKean stochastic differential equation

$$\dot{\mathbf{z}}_t^i = -\frac{1}{\gamma} \nabla_{\mathbf{z}} G(\mathbf{z}_t^i) + \sqrt{\frac{2}{\kappa_h \gamma}} \xi_i(t), \quad (\text{B.1})$$

where $G(\mathbf{z}_t) = \frac{1}{2}(\mathbf{z}_t - \mathbf{m}_t)^T V_t^{-1}(\mathbf{z}_t - \mathbf{m}_t)$ and γ is the friction coefficient. We want to show it converges to a steady state as an approximation of q^* with the proper choices of \mathbf{m}_t and V_t given by

$$\begin{aligned} \mathbf{m}_t &= \sum_{i=1}^{N_w} \mathbf{z}_t^i \hat{p}(\mathbf{z}_t^i), \\ V_t &= \kappa_t \sum_{i=1}^{N_w} (\mathbf{z}_t^i - \mathbf{m}_t)(\mathbf{z}_t^i - \mathbf{m}_t)^T \hat{p}(\mathbf{z}_t^i), \end{aligned} \quad (\text{B.2})$$

where $\hat{p}(\mathbf{z}) = \frac{\exp(-\kappa_l \mathcal{L}_N^-(\mathbf{z}))}{\sum_{i=1}^{N_w} \exp(-\kappa_l \mathcal{L}_N^-(\mathbf{z}^i))}$ and $\kappa_t = \kappa_l + \kappa_h$.

Proposition 2. Suppose $\mathcal{L}_N^-(\mathbf{z})$ takes a local quadratic approximation in form of $\frac{1}{2}(\mathbf{z} - \mu)^T \Sigma^{-1}(\mathbf{z} - \mu)$, $q_t \rightarrow \frac{\exp(-\kappa_h \mathcal{L}_N^-(\mathbf{z}))}{\int \exp(-\kappa_h \mathcal{L}_N^-(\mathbf{z})) d\mathbf{z}}$ as $t \rightarrow \infty$, given $\kappa_t = \kappa_l + \kappa_h$.

Proof. Let $q_\infty(\mathbf{z})$ denote the invariant distribution of the McKean SDE (B.1). Then $q_\infty(\mathbf{z})$ must be the invariant distribution of the following McKean stochastic dynamics

$$\dot{\mathbf{z}} = -\frac{1}{\gamma} V_{\kappa_l, \infty}^{-1}(\mathbf{z} - \mathbf{m}_{\kappa_l, \infty}) + \sqrt{\frac{2}{\gamma \kappa_h}} \xi(t), \quad (\text{B.3})$$

where $\mathbf{m}_{\kappa_l, \infty}$ and $\kappa_l^{-1} V_{\kappa_l, \infty}$ are the mean and the covariance matrix of the re-weighted density $\propto q_\infty(\mathbf{z}) e^{-\kappa_l \mathcal{L}_N^-(\mathbf{z})}$. With the fluctuation-dissipation relation for Eq. (B.3), we can show $q_\infty(\mathbf{z})$ follows the Gaussian distribution with mean $\mathbf{m}_{\kappa_l, \infty}$ and covariance matrix $\kappa_h^{-1} V_{\kappa_l, \infty}$.

Since $\mathcal{L}_N^-(\mathbf{z}) = \frac{1}{2}(\mathbf{z} - \mu)^T \Sigma^{-1}(\mathbf{z} - \mu)$ is quadratic, the re-weighted density of a Gaussian distribution $q(\mathbf{z}) \sim \mathcal{N}(\mathbf{m}, V)$ remains Gaussian, i.e.,

$$q(\mathbf{z}) e^{-\kappa_l \mathcal{L}_N^-(\mathbf{z})} \propto \mathcal{N}(\mathbf{m}_{\kappa_l}, V_{\kappa_l}),$$

where \mathbf{m}_{κ_l} and V_{κ_l} are defined by

$$\begin{aligned}\mathbf{m}_{\kappa_l} &= (V^{-1} + \kappa_l \Sigma^{-1})^{-1} (\kappa_l \Sigma^{-1} \mu + V^{-1} \mathbf{m}), \\ V_{\kappa_l} &= (V^{-1} + \kappa_l \Sigma^{-1})^{-1}.\end{aligned}\tag{B.4}$$

In particular, we choose $\mathbf{m} = \mathbf{m}_{\kappa_l, \infty}$ and $V = \kappa_h^{-1} V_{\kappa_l, \infty}$, then Eq. (B.4) yields

$$\begin{aligned}\mathbf{m}_{\kappa_l, \infty} &= (\kappa_h V_{\kappa_l, \infty}^{-1} + \kappa_l \Sigma^{-1})^{-1} (\kappa_l \Sigma^{-1} \mu + \kappa_h V_{\kappa_l, \infty}^{-1} \mathbf{m}_{\kappa_l, \infty}), \\ \kappa_t^{-1} V_{\kappa_l, \infty} &= (\kappa_h V_{\kappa_l, \infty}^{-1} + \kappa_l \Sigma^{-1})^{-1}.\end{aligned}$$

It is easy to show that by choosing $\kappa_t = \kappa_l + \kappa_h$, $\mathbf{m}_{\kappa_l, \infty}$ and $V_{\kappa_l, \infty}$ recovers μ and Σ , respectively, and the invariant density takes the form

$$q_{\infty}(\mathbf{z}) \sim \mathcal{N}\left(\mu, \kappa_h^{-1} \Sigma\right).$$

□

B.2 Simulation setup

All the MD simulations are performed using the package GROMACS 2019.2 Lindahl et al. (2019) and open-source, community-developed PLUMED library Tribello et al. (2014). The simulation is carried out on Intel(R) Xeon(R) Platinum 8260 CPU.

B.2.1 ala2 (the two-dimensional problem)

The Ace-Ala-Nme (ala2) molecule is modeled by the Amber99SB force field Hornak et al. (2006). The molecules are dissolved in 383 water molecules in a periodic simulation cell. The cut-off radius of the van der Waals interaction is 0.9 nm. The Coulomb interaction is treated with the smooth particle mesh Ewald method with a real space cutoff of 0.9 nm and a reciprocal space grid spacing of 0.12 nm. The system is integrated with the leap-frog scheme at time step 2 fs. The temperature of the system is set to 300 K by a velocity-rescale thermostat Bussi et al. (2007) with a relaxation time of 0.2 ps. The Parrinello-Rahman barostat Parrinello and Rahman (1981) with a relaxation time scale of 1.5 ps and a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ is coupled to the system to control the pressure to 1 bar. The hydrogen atom is constrained by the LINCS algorithm Hess et al. (1997) and the H–O bond and H–O–H angle of water molecules are constrained by the SETTLE algorithm Miyamoto and Kollman (1992).

B.2.2 s1pe (three-dimensional problem)

The s-(1)-phenylethyl (s1pe) molecule is modeled by the CHARMM general force field (CGenFF) Weiser and Santiso (2019). The molecule is dissolved in 546 TIP3P water molecules in a (2.69 nm)³ dodecahedron box. The cut-off radius of the van der Waals interaction is 1 nm. The Coulomb interaction is treated with the smooth particle mesh Ewald method with a real space cutoff of 1 nm and a reciprocal space grid spacing of 0.12 nm. The system is integrated with the leap-frog scheme at time step 2 fs. The temperature of the system is set to 298 K by a velocity-rescale thermostat Bussi et al. (2007) with a relaxation time of 0.2 ps. The Parrinello-Rahman barostat pressure couple, the hydrogen atom constraint, the H–O bond constraint, and the H–O–H angle of water molecules constraint are the same as the previous one.

B.2.3 (s1pe)₃ (the nine-dimensional problem)

The (s1pe)₃ molecule is modeled by the CHARMM general force field (CGenFF) Weiser and Santiso (2019). The molecule is solvated in a (4.2 nm)³ dodecahedron box with 1622 TIP3P water molecules. Other simulation setups are similar to the s1pe molecule.

B.2.4 ala16 (the thirty-dimensional problem)

The Ace-(Ala)₁₅-Nme (ala16) molecule is modeled by the Amber99SB force field Hornak et al. (2006). The molecule is solvated in 2258 water molecules in a (4.62 nm)³ dodecahedron box. Other simulation setups are similar to the ala2 molecule.

B.3 Training

The training data is collected during the sampling process by restricted dynamics. The initial 10% steps of the restricted dynamics are used as equilibrium and the rest 90% steps are used to calculate the mean force. The FES are parameterized as a fully connected neural network. The depth and width of the NN are shown in Supplementary Table B.1. The NNs are trained by Adam for 100000 steps with a learning rate 1×10^{-3} . For each training step, 5000 sampling points are randomly selected from the data set. All the training process is carried out on v100 with 32768 MB memory.

Table B.1 The depth and width of the NN used to parameterize the FES of different molecules.

molecule	depth	width
ala2	3	48
s1pe	4	64
(s1pe) ₃	4	512
ala16	4	640

B.4 Additional result

B.4.1 Additional result of the 1D Rastrigin function

Supplementary Table B.2 presents the estimations of the first and second momentum during the construction of the 1D Rastrigin function. The first momentum, predicted to be the maximum residual points in **Proposition 2**, is confirmed to be numerically accurate. The second momentum is validated through the comparison between f'' and $f''_{\theta_{i-1}} \pm V^{-1}$, with the sign influenced by the residual's sign. Numerical results verify the accurate prediction of the second momentum, i.e., $\left| f''(x^*) - f''_{\theta_i}(x^*) \right| \approx V^{-1}$ for x^* at the max-residual point.

Table B.2 The first and second momentum estimation made during the construction of 1D Rastrigin function. The maximum residual point of $|f(x) - f_{\theta_i}(x)|$ is computed by mesh points with mesh size 1×10^{-4} . The first momentum is estimated from the equilibrium distribution of CES dynamics. The second derivative of f is computed at the first momentum estimated point. The second derivative estimation is made from $f''_{\theta_{i-1}} \pm V^{-1}$, where V is second momentum estimated from the equilibrium distribution of CES dynamics and sign is determined by the sign of $f - f_{\theta_i}$.

iteration i	0	1	2	3	4	5	6	7	8	9	10	11	12
maximum residual point	0.000	2.000	-2.000	0.996	-0.996	2.805	-2.806	1.500	-1.502	0.503	-0.501	2.416	-2.413
first momentum estimated	-0.000	2.000	-2.000	0.996	-0.995	2.805	-2.806	1.499	-1.503	0.502	-0.501	2.414	-2.413
second derivative of f	41.478	41.478	41.478	41.468	41.462	15.273	15.539	-37.478	-37.473	-37.474	-37.478	-31.901	-31.688
second derivative estimated	40.974	37.442	36.916	38.283	39.519	11.841	13.380	-36.796	-36.271	-37.515	-37.285	-31.521	-31.189

B.4.2 Additional result of the three-dimensional FES (s1pe)

Supplementary Figure B.1-B.3 shows the additional 2D projections of the 3D FES (molecule s1pe) constructed by the Rid and presented CES method. For each case, the reference is constructed as a 2D FES using the metadynamics. The CES method yields a better agreement with the reference. The computational cost and accuracy of CES and RiD is shown in Supplementary Table B.3. It shows that the CES method has better accuracy with less computational cost.

Table B.3 The accuracy of the constructed 3D FES (the s1pe molecule) and computational time for RiD and CES methods. The l_2 and l_∞ error are calculated up to 40 KJ/mol. For each case, the FES is projected onto a 2D plane with the third variable fixed; the reference solution is constructed as a 2D FES by the metadynamics. The simulation time of the CES method is multiplied by 20 since 20 walkers are used.

Method	Restrained	Accuracy		Time	
		l_2 error	l_∞ error	Sampling	Train
RiD	$\psi = 1.5$	5.76	25.72	423.33	8 (GPU)
	$\omega = 1.5$	12.04	49.13		
CES	$\psi = 1.5$	2.44	11.21	4.81×20	0.84 (GPU)
	$\omega = 1.5$	3.89	28.80		

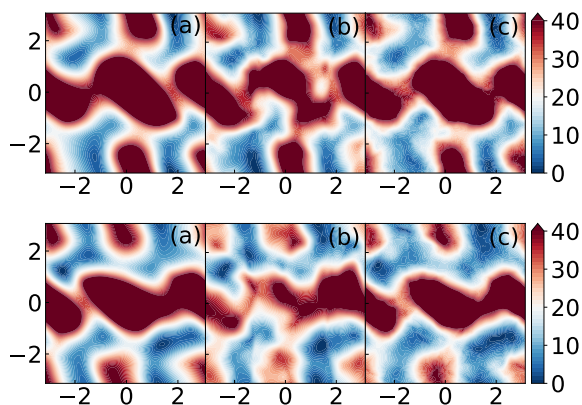


Figure B.1 The 3D FES for the s1pe molecule projected on the $\phi - \psi$ plane with the third variable $\omega = -1$ (first row) and $\omega = -2$ (second row). (a) 2D FES by metadynamics (reference) (b) RiD (c) CES.

B.4.3 Additional results of the nine-dimensional FES (s1pe)₃

The additional 2D projections of the 9D FES for molecule (s1pe)₃ are presented in Supplementary Figure B.4. Similar to the 3D case, the FES constructed by the present CES method shows a better agreement with the reference constructed as a 2D FES using metadynamics.

B.4.4 Additional results of the thirty-dimensional FES (ala16)

The additional 2D projections of the 30D FES for molecule ala16 are presented in Supplementary Figure B.5. We note that the projection on the $\phi_5 - \psi_5$ plane is significantly different from other projections such as the $\phi_1 - \psi_1$ plane in the main context. While the ala16 molecule consists of 15 sequential alanine residues, the FES for individual $\phi - \psi$ projections shows different features. The numerical results of the present CES method show good agreement with the reference.

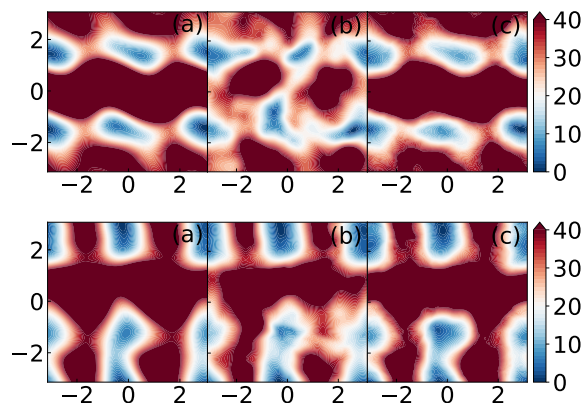


Figure B.2 The 3D FES of the s1pe molecule on the $\omega - \psi$ plane with the third variable $\phi = 0$ (first row) and $\phi = -1$ (second row). (a) 2D FES by metadynamics (reference) (b) RiD (c) CES.

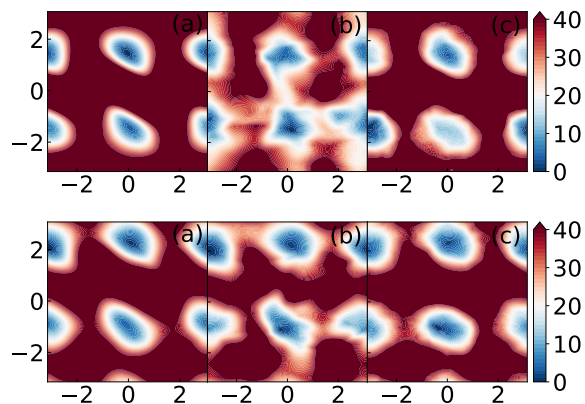


Figure B.3 The 3D FES of the s1pe molecule on the $\omega - \phi$ plane with the third variable $\psi = 0$ (first row) and $\psi = -1$ (second row). (a) 2D FES by metadynamics (reference) (b) RiD (c) CES.

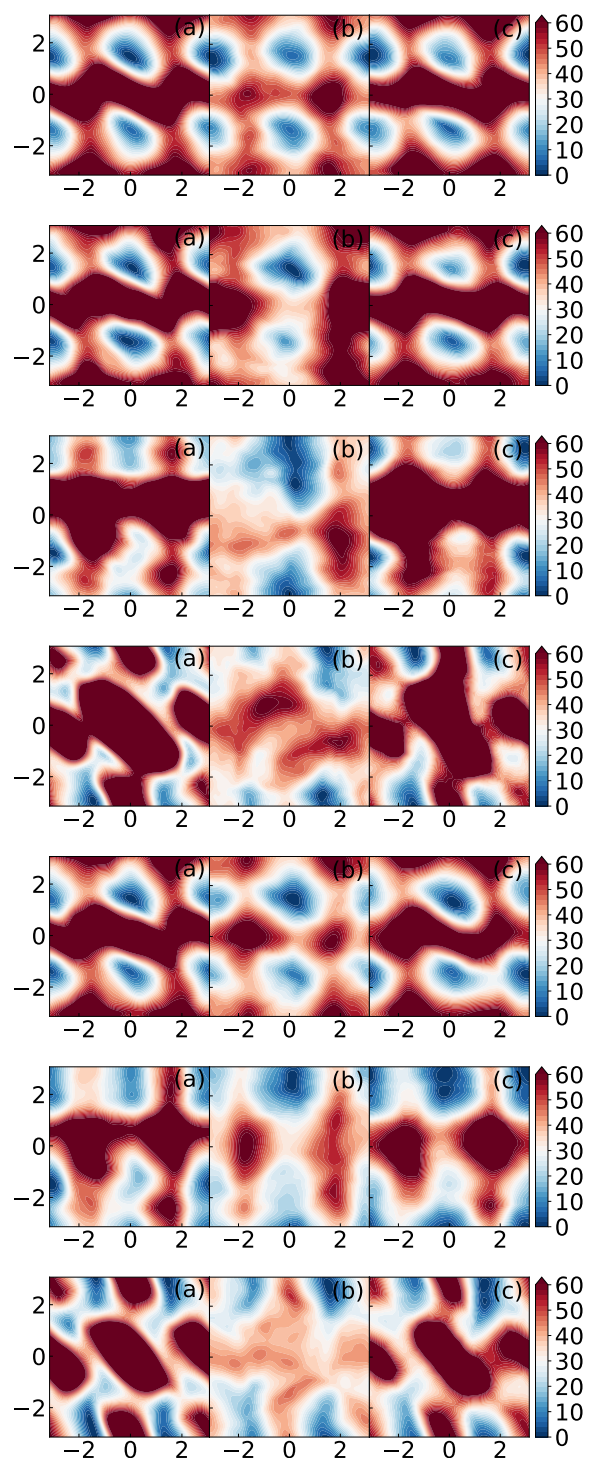


Figure B.4 The 9D FES for the (s1pe)₃ molecule projected on the $\omega_1 - \phi_1, \omega_2 - \phi_2, \omega_2 - \psi_2, \phi_2 - \psi_2, \omega_3 - \phi_3, \omega_3 - \psi_3, \phi_3 - \psi_3$ plane from top to bottom. (a) 2D FES by metadynamics (reference) (b) RiD (c) CES.

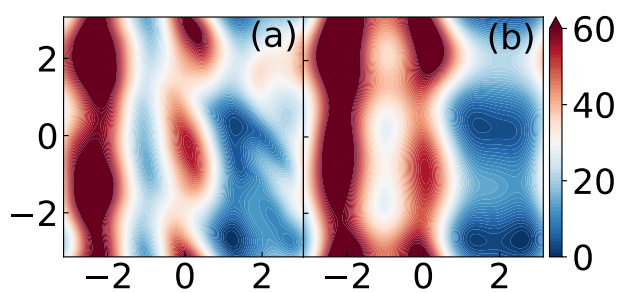


Figure B.5 The 30D FES of the ala16 projected on the $\phi_5 - \psi_5$ plane. (a) 2D FES constructed by metadynamics (reference) (b) 30D FES constructed by the CES method.