Editors' Suggestion

Construction of Coarse-Grained Molecular Dynamics with Many-Body Non-Markovian Memory

Liyao Lyu[1] and Huan Lei[1,2,*]
[1]Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, Michigan 48824, USA
[2]Department of Statistics and Probability, Michigan State University, East Lansing, Michigan 48824, USA

(Received 5 April 2023; accepted 19 September 2023; published 23 October 2023)

We introduce a machine-learning-based coarse-grained molecular dynamics model that faithfully retains the many-body nature of the intermolecular dissipative interactions. Unlike the common empirical coarse-grained 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. Numerical results show that preserving the many-body nature of the memory term is crucial for predicting the collective transport and diffusion processes, where empirical forms generally show limitations.

DOI: 10.1103/PhysRevLett.131.177301

Accurately predicting the collective behavior of multiscale physical systems is a long-standing problem that requires the integrated modeling of the molecular-level interactions across multiple scales [1]. 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 enable us to achieve a broader range of the spatiotemporal scale, the construction of truly reliable coarse-grained (CG) models remains highly nontrivial. A significant amount of work [2–13] (see also review [14]), including recent machine learning (ML)-based approaches [15–21], 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 degrees of freedom; 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 [22–28] 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 [29], Langevin dynamics [30], and dissipative particle dynamics (DPD) [31,32]. Despite their broad applications, studies [33–35] 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 [36–51] model the memory term based on the generalized Langevin equation (GLE) and its variants (see also review [52]). The velocity autocorrelation function (VACF) is often used as the target quantity for model parametrization. While it may serve as an appropriate measure for certain nonextensive systems [53,54], 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 Letter 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 [21], 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 intermolecular 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 [55] and shows the promise to predict the mesoscale transport and diffusion processes encoded with molecular-level fidelity.

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 $\mathbf{z} = [\mathbf{q}, \mathbf{p}]$, where $\mathbf{q}, \mathbf{p} \in \mathbb{R}^{3N}$ represent the position and momentum vector, respectively. Given $\mathbf{z}(0) = \mathbf{z}_0$, the evolution follows $\mathbf{z}(t) = \mathrm{e}^{\mathcal{L}t}\mathbf{z}_0$, where $\mathcal{L}$ is the Liouville operator determined by the Hamiltonian $H(\mathbf{z})$. The CG variables are defined by representing each molecule as a CG particle, i.e., $\boldsymbol{\phi}(\mathbf{z}) = [\boldsymbol{\phi}^Q(\mathbf{z}), \boldsymbol{\phi}^P(\mathbf{z})]$, where $\boldsymbol{\phi}^Q(\mathbf{z}) = [\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_M]$ and $\boldsymbol{\phi}^P(\mathbf{z}) = [\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_M]$ 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 $\boldsymbol{\phi}[\mathbf{z}(t)]$ with $\mathbf{z}(0) = \mathbf{z}_0$. To construct the reduced model, we define the Zwanzig projection operator as the conditional expectation with a fixed CG vector $\boldsymbol{Z}$, i.e., $\mathcal{P}_{\boldsymbol{Z}} f(\mathbf{z}) \coloneqq \mathbb{E}[f(\mathbf{z})|\boldsymbol{\phi}(\mathbf{z}) = \boldsymbol{Z}]$ under conditional density proportional to $\delta(\boldsymbol{\phi}(\mathbf{z}) - \boldsymbol{Z})\mathrm{e}^{-\beta H(\mathbf{z})}$ and its orthogonal operator $\mathcal{Q}_{\boldsymbol{Z}} = \mathbf{I} - \mathcal{P}_{\boldsymbol{Z}}$. Using Zwanzig's formalism [56], the dynamics of $\mathbf{Z}(t)$ [see Ref. [57] and Supplemental Material (SM) [58]] can be written as

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

where $\mathbf{M}$ is the mass matrix and $\mathbf{V} = \mathbf{M}^{-1}\mathbf{P}$ is the velocity. $U(\mathbf{Q})$ is the free energy under $\boldsymbol{\phi}^Q(\mathbf{z}) \equiv \mathbf{Q}$. $\mathbf{K}(\mathbf{Q}, t) = \mathcal{P}_{\mathbf{Z}}[(e^{\mathcal{Q}_{\mathbf{Z}}\mathcal{L}t}\mathcal{Q}_{\mathbf{Z}}\mathcal{L}\mathbf{P})(\mathcal{Q}_{\mathbf{Z}}\mathcal{L}\mathbf{P})^T]$ is the memory representing the coupling between the CG and unresolved variables, and $\mathbf{R}(t)$ is the fluctuation force which can be modeled as a Gaussian random process satisfying the second fluctuation-dissipation theorem [57].

Equation (1) provides the starting point to derive the various CG models. Direct evaluation of $\mathbf{K}(\mathbf{Q}, t)$ imposes a challenge as it relies on solving the full-dimensional orthogonal dynamics $e^{\mathcal{Q}_{\mathbf{Z}}\mathcal{L}t}$. Further simplification $\mathbf{K}(\mathbf{Q}, t) \approx \theta(t)$ leads to the common GLE with a homogeneous kernel. Alternatively, the pairwise approximation $[\mathbf{K}(\mathbf{Q}, t)]_{ij} \approx \gamma(Q_{ij})\delta(t)$ or $\gamma(Q_{ij})\theta(t)$ leads to the standard DPD (M-DPD) and non-Markovian variants (NM-DPD), respectively. 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 $\mathbf{K}(\mathbf{Q}, t)$.

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

$$\begin{aligned}
&\boldsymbol{\Xi}_{ij}(\mathbf{Q}_1 + \mathbf{b}, \ldots, \mathbf{Q}_M + \mathbf{b}) = \boldsymbol{\Xi}_{ij}(\mathbf{Q}_1, \ldots, \mathbf{Q}_M),\\
&\boldsymbol{\Xi}_{ij}(\mathcal{U}\mathbf{Q}_1, \ldots, \mathcal{U}\mathbf{Q}_M) = \mathcal{U}\boldsymbol{\Xi}_{ij}(\mathbf{Q}_1, \ldots, \mathbf{Q}_M)\mathcal{U}^T,\\
&\boldsymbol{\Xi}_{\sigma(i)\sigma(j)}(\mathbf{Q}_{\sigma(1)}, \ldots, \mathbf{Q}_{\sigma(M)}) = \boldsymbol{\Xi}_{ij}(\mathbf{Q}_1, \ldots, \mathbf{Q}_M), \quad (2)
\end{aligned}$$

where $\boldsymbol{\Xi}_{ij} \in \mathbb{R}^{3\times 3}$ represents the friction contribution of jth particle on ith 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 $\mathbf{f}$: $\mathbb{R} \to \mathbb{R}^K$ is an encoder function to be learned, and $\mathcal{N}_i = \{l|Q_{il} < r_c\}$ is the neighboring index set of the ith particle within a cutoff distance r_c. Accordingly, $\hat{\mathbf{Q}}_{ij} \in \mathbb{R}^{3\times K}$ represents a set of features that encode the intermolecular configurations beyond the pairwise approximation. The kth 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 $\boldsymbol{\Xi}_{ij}$ by

$$\boldsymbol{\Xi}_{ij} = \sum_{k=1}^{K} h_k(\hat{\mathbf{Q}}_{ij}^T\hat{\mathbf{Q}}_{ij})\hat{\mathbf{Q}}_{ij}^k \otimes \hat{\mathbf{Q}}_{ij}^k + h_0(\hat{\mathbf{Q}}_{ij}^T\hat{\mathbf{Q}}_{ij})\mathbf{I}, \quad (3)$$

where $\mathbf{h}$: $\mathbb{R}^{K\times K} \to \mathbb{R}^{K+1}$ are encoder functions which will be represented by NNs. For $i = j$, we have $\boldsymbol{\Xi}_{ii} = -\sum_{j\in\mathcal{N}_i}\boldsymbol{\Xi}_{ij}$ based on the Newton's third law. We refer to SM [58] for the proof of the symmetry constraint (2).

Equation (3) entails the state dependency of the memory term $\mathbf{K}(\mathbf{Z}, t)$ under the Markovian approximation. To incorporate the non-Markovian effect, we embed the memory term within an extended Markovian dynamics [38] (see also Ref. [50]). Specifically, we seek a set of non-Markovian features $\boldsymbol{\zeta} \coloneqq [\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \ldots, \boldsymbol{\zeta}_n]$, and construct the joint dynamics of $[\mathbf{Z}, \boldsymbol{\zeta}]$ by imposing the many-body form of the friction tensor between $\mathbf{P}$ and $\boldsymbol{\zeta}$, i.e.,

$$\dot{\mathbf{Q}} = \mathbf{M}^{-1}\mathbf{P},$$
$$\dot{\mathbf{P}} = -\nabla U(\mathbf{Q}) + \Xi(\mathbf{Q})\boldsymbol{\zeta},$$
$$\dot{\boldsymbol{\zeta}} = -\Xi(\mathbf{Q})^T\mathbf{V} - \boldsymbol{\Lambda}\boldsymbol{\zeta} + \boldsymbol{\xi}(t), \tag{4}$$

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

$$\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}(t')\rangle = \beta^{-1}(\boldsymbol{\Lambda} + \boldsymbol{\Lambda}^T)\delta(t-t'), \tag{5}$$

Eq. (4) retains the consistent invariant distribution $\rho(\mathbf{Q},\mathbf{P},\boldsymbol{\xi}) \propto \exp\{-\beta[U(\mathbf{Q}) + \mathbf{P}^T\mathbf{M}^{-1}\mathbf{P}/2 + \boldsymbol{\zeta}^T\boldsymbol{\zeta}/2]\}$ (see proof in SM [58]).

Equation (4) 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 $\Xi[\mathbf{Q}(t)]e^{\boldsymbol{\Lambda}(t-s)}\Xi[\mathbf{Q}(s)]^T$ based on the MZ form. However, directly solving the orthogonal dynamics $e^{\mathcal{Q}_Z\mathcal{L}t}$ is computationally intractable. Alternatively, we introduce the constrained dynamics $\tilde{\mathbf{z}}(t) = e^{\mathcal{R}t}\mathbf{z}(0)$ following Ref. [34]. Based on the observation $\mathcal{P}\mathcal{Q} = \mathcal{P}\mathcal{R} \equiv 0$, we sample the MZ form from $\tilde{\mathbf{z}}(t)$, i.e., $\mathbf{K}_{MZ}(\mathbf{Z},t) = \mathcal{P}_Z[(e^{\mathcal{R}t}\mathcal{Q}_Z\mathcal{L}\mathbf{P})(\mathcal{Q}_Z\mathcal{L}\mathbf{P})^T]$ and the memory of the CG model reduces to $\mathbf{K}_{CG}(\mathbf{Z},t) = \Xi(\mathbf{Q})e^{\boldsymbol{\Lambda}t}\Xi(\mathbf{Q})^T$. This enables us to train the CG models in terms of the encoders $\{\mathbf{f}^i(\cdot), \mathbf{h}^i(\cdot)\}_{i=1}^n$ and matrices $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}^a$ by minimizing the empirical loss

$$L = \sum_{l=1}^{N_s}\sum_{j=1}^{N_t}\left\|\mathbf{K}_{CG}(\mathbf{Z}^{(l)},t_j) - \mathbf{K}_{MZ}(\mathbf{Z}^{(l)},t_j)\right\|^2, \tag{6}$$

where l represents the different CG configurations (see SM [58] for details in training).

To demonstrate the accuracy of the present model, we consider a full microscale model of a star-shaped polymer melt system similar to Ref. [34], where each molecule consists of 73 atoms. The atomistic interactions are modeled by the Weeks-Chandler-Anderse 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 [63,64] is employed to equilibrate the system with $k_BT = 4.0$ and microcanonical ensemble simulation is conducted during the production stage (see SM [58] for details). 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 [21] for all the CG models; the differences in dynamic properties solely arise from the different formulations of the memory term.

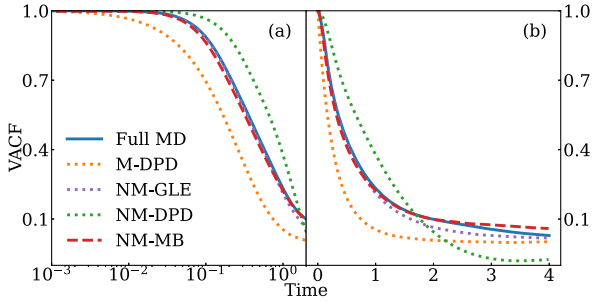

FIG. 1. The VACF of the full MD and CG models with various memory formulations in (a) semilog scale (b) original scale. "M" and "NM" represent Markovian and Non-Markovian; GLE, DPD, and MB represent state-independent, pairwise, and the present (NM-MB) model retaining the many-body effects, respectively. See SM [58] for the details of M-DPD, NM-GLE, and NM-DPD models.

Let us start with the VACF which has been broadly used in CG model parametrization and validation. As shown in Fig. 1, 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 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 timescale 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 parametrize 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 oversimplifies the heterogeneity of the memory term and leads to inaccurate predictions on the collective scales.

Figure 2 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

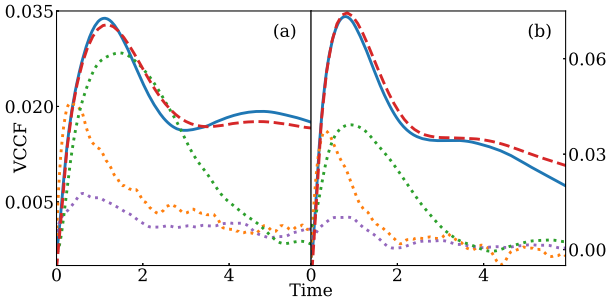

FIG. 2. The VCCF $C^{xx}(t;r_0)$ predicted by the full MD and different CG models with initial distance (a) $10 < r_0 < 11$ and (b) $14 < r_0 < 15$. Same line legend as Fig. 1.

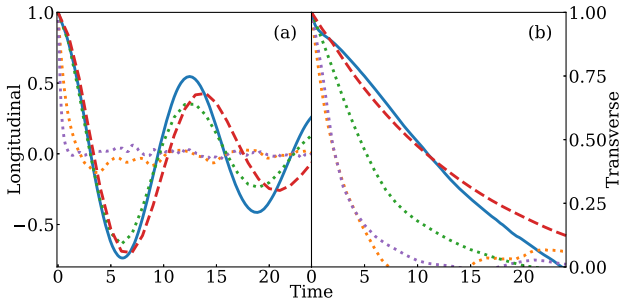

FIG. 3. (a) Longitudinal and (b) transverse hydrodynamic modes predicted by MD and different CG models. Same line legend as Fig. 1.

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 follows. The VACF represents the energy dissipation on each particle as a homogeneous background heat bath; it is essentially a mean-field metric and cannot 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.

Furthermore, the various empirical models for local energy dissipations can lead to fundamentally different transport processes on the collective scale. Figure 3 shows the normalized correlations of the longitudinal and transverse hydrodynamic modes [65], 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 intermolecule dissipations.

Finally, we examine the diffusion process on the collective scale. Figure 4 shows the van Hove function that characterizes the evolution of the interparticle structural correlation defined by $G(r,t) \propto (1/M^2) \sum_{j\neq i}^{M} \delta(\|\mathbf{Q}_i(t) - \mathbf{Q}_j(0)\| - r)$. 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

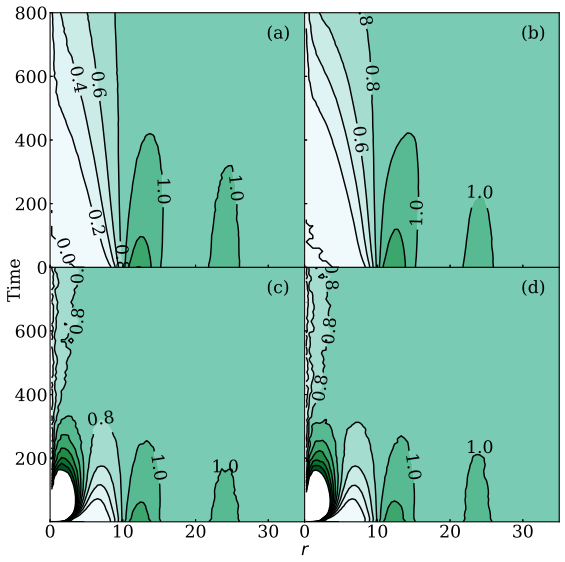

FIG. 4. The van Hove function $G(r,t)$ predicted by (a) full MD, (b) the present NM-MB model, (c) NM-DPD model, and (d) GLE model. It depicts the diffusive process of the radius distribution function (x axis) over time (y axis).

until $t = 1000$, when the initial fluid structure ultimately diffuses into a homogeneous state.

To conclude, this Letter reports a caveat in constructing reliable CGMD models that retain consistent 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, cannot be fully characterized by the mean-field metrics such as VACF. More importantly, the memory form representing the intermolecule 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.

The work is supported by the National Science Foundation under Grant No. DMS-2110981 and the ACCESS program through allocation MTH210005. H. L. acknowledges partial support from the Department of Energy Center for Hieratical and Robust Modeling of Non-Equilibrium Transport under Grant No. DOE-DE-SC0023164.

*leihuan@msu.edu

[1] P. W. Anderson, Science **177**, 393 (1972).
[2] G. Torrie and J. Valleau, J. Comput. Phys. **23**, 187 (1977).
[3] L. Rosso, P. Mináry, Z. Zhu, and M. E. Tuckerman, J. Chem. Phys. **116**, 4389 (2002).
[4] L. Maragliano and E. Vanden-Eijnden, Chem. Phys. Lett. **426**, 168 (2006).
[5] S. Izvekov and G. A. Voth, J. Phys. Chem. B **109**, 2469 (2005).
[6] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, J. Chem. Phys. **128**, 244114 (2008).
[7] R. E. Rudd and J. Q. Broughton, Phys. Rev. B **58**, R5893 (1998).
[8] A. P. Lyubartsev and A. Laaksonen, Phys. Rev. E **52**, 3730 (1995).
[9] M. S. Shell, J. Chem. Phys. **129**, 144108 (2008).
[10] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, J. Comput. Chem. **13**, 1011 (1992).
[11] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein, J. Phys. Condens. Matter **16**, R481 (2004).
[12] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).
[13] E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
[14] W. G. Noid, J. Chem. Phys. **139**, 090901 (2013).
[15] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[16] T. Stecher, N. Bernstein, and G. Csányi, J. Chem. Theory Comput. **10**, 4079 (2014).
[17] S. T. John and G. Csányi, J. Phys. Chem. B **121**, 10934 (2017).
[18] T. Lemke and C. Peter, J. Chem. Theory Comput. **13**, 6213 (2017).
[19] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, Sci. Adv. **3**, e1603015 (2017).
[20] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018).
[21] L. Zhang, J. Han, H. Wang, R. Car, and W. E, J. Chem. Phys. **149**, 034101 (2018).
[22] R. Satija, A. Das, and D. E. Makarov, J. Chem. Phys. **147**, 152707 (2017).
[23] G. Luo, I. Andricioaei, X. S. Xie, and M. Karplus, J. Phys. Chem. B **110**, 9363 (2006).
[24] R. B. Best and G. Hummer, Proc. Natl. Acad. Sci. U.S.A. **107**, 1088 (2010).
[25] S. S. Plotkin and P. G. Wolynes, Phys. Rev. Lett. **80**, 5015 (1998).
[26] J. B. Straus, J. M. Gomez Llorente, and G. A. Voth, J. Chem. Phys. **98**, 4082 (1993).
[27] J. A. Morrone, J. Li, and B. J. Berne, J. Phys. Chem. B **116**, 378 (2012).
[28] J. O. Daldrop, B. G. Kowalik, and R. R. Netz, Phys. Rev. X **7**, 041065 (2017).
[29] A. Einstein, Ann. Phys. (Berlin) **17**, 549 (1905).
[30] N. V. Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 2007).
[31] P. J. Hoogerbrugge and J. M. V. A. Koelman, Europhys. Lett. **19**, 155 (1992).
[32] P. Español and P. Warren, Europhys. Lett. **30**, 191 (1995).
[33] H. Lei, B. Caswell, and G. E. Karniadakis, Phys. Rev. E **81**, 026704 (2010).
[34] C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni, Faraday Discuss. **144**, 301 (2010).
[35] Y. Yoshimoto, I. Kinefuchi, T. Mima, A. Fukushima, T. Tokumasu, and S. Takagi, Phys. Rev. E **88**, 043305 (2013).
[36] O. F. Lange and H. Grubmüller, J. Chem. Phys. **124**, 214903 (2006).
[37] S. Wang, Z. Ma, and W. Pan, Soft Matter **16**, 8330 (2020).
[38] M. Ceriotti, G. Bussi, and M. Parrinello, Phys. Rev. Lett. **102**, 020601 (2009).
[39] A. D. Baczewski and S. D. Bond, J. Chem. Phys. **139**, 044107 (2013).
[40] A. Davtyan, J. F. Dama, G. A. Voth, and H. C. Andersen, J. Chem. Phys. **142**, 154104 (2015).
[41] H. Lei, N. A. Baker, and X. Li, Proc. Natl. Acad. Sci. U.S.A. **113**, 14183 (2016).
[42] Z. Li, H. S. Lee, E. Darve, and G. E. Karniadakis, J. Chem. Phys. **146**, 014104 (2017).
[43] A. Russo, M. A. Durán-Olivencia, I. G. Kevrekidis, and S. Kalliadasis, arXiv:1903.09562.
[44] G. Jung, M. Hanke, and F. Schmid, J. Chem. Theory Comput. **13**, 2481 (2017).
[45] H. S. Lee, S.-H. Ahn, and E. F. Darve, J. Chem. Phys. **150**, 174113 (2019).
[46] L. Ma, X. Li, and C. Liu, J. Comput. Phys. **380**, 170 (2019).
[47] Z. Ma, S. Wang, M. Kim, K. Liu, C.-L. Chen, and W. Pan, Soft Matter **17**, 5864 (2021).
[48] V. Klippenstein and N. F. A. van der Vegt, J. Chem. Phys. **154**, 191102 (2021).
[49] H. Vroylandt, L. Goudenège, P. Monmarché, F. Pietrucci, and B. Rotenberg, Proc. Natl. Acad. Sci. U.S.A. **119**, e2117586119 (2022).
[50] Z. She, P. Ge, and H. Lei, J. Chem. Phys. **158**, 034102 (2023).
[51] P. Xie, R. Car, and W. E, arXiv:2211.06558.
[52] V. Klippenstein, M. Tripathy, G. Jung, F. Schmid, and N. F. van der Vegt, J. Phys. Chem. B **125**, 4931 (2021).
[53] C. Widder, F. Koch, and T. Schilling, J. Chem. Phys. **157**, 194107 (2022).
[54] H. Meyer, P. Pelagejcev, and T. Schilling, Europhys. Lett. **128**, 40001 (2019).
[55] L. Van Hove, Phys. Rev. **95**, 249 (1954).
[56] R. Zwanzig, J. Stat. Phys. **9**, 215 (1973).
[57] H. Vroylandt and P. Monmarché, J. Chem. Phys. **156**, 244105 (2022).
[58] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.131.177301, for the derivation of the reduced dynamics of the CG variables, the setup of the full MD simulations, the introduction of coherent noise, the construction of the conservative CG potential, the proof of the symmetry constraints of the NN-based model, the details of the training process, and an overview of common CGMD models, which includes Refs. [59–62].
[59] B. O. Koopman, Proc. Natl. Acad. Sci. U.S.A. **17**, 315 (1931).
[60] T. Kinjo and S. A. Hyodo, Phys. Rev. E **75**, 051109 (2007).

[61] E. Darve, J. Solomon, and A. Kia, Proc. Natl. Acad. Sci. U.S.A. **106**, 10884 (2009).
[62] D. Kingma and J. Ba, *Proceedings of the International Conference on Learning Representations (ICLR)* (Academic Press, Cambridge, 2015).
[63] S. Nosé, Mol. Phys. **52**, 255 (1984).
[64] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[65] J. P. Hansen and I. McDonald, *Theory of Simple Liquids* (Academic, London, 1990).