

2014

A coarse-graining approach for molecular simulation that retains the dynamics of the all-atom reference system by implementing hydrodynamic interactions

Sergiy Markutsya
The Ames Laboratory

Monica H. Lamm
Iowa State University, mhlamm@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/cbe_pubs

 Part of the [Biological and Chemical Physics Commons](#), and the [Chemistry Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/cbe_pubs/289. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Chemical and Biological Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Chemical and Biological Engineering Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

A coarse-graining approach for molecular simulation that retains the dynamics of the all-atom reference system by implementing hydrodynamic interactions

Abstract

We report on a new approach for deriving coarse-grained intermolecular forces that retains the frictional contribution that is often discarded by conventional coarse-graining methods. The approach is tested for water and an aqueous glucose solution, and the results from the new implementation for coarse-grained molecular dynamics simulation show remarkable agreement with the dynamics obtained from reference all-atom simulations. The agreement between the structural properties observed in the coarse-grained and all-atom simulations is also preserved. We discuss how this approach may be applied broadly to any existing coarse-graining method where the coarse-grained models are rigorously derived from all-atom reference systems.

Disciplines

Biological and Chemical Physics | Chemistry | Physics

Comments

This article is published as Markutsya, Sergiy, and Monica H. Lamm. "A coarse-graining approach for molecular simulation that retains the dynamics of the all-atom reference system by implementing hydrodynamic interactions." *The Journal of chemical physics* 141, no. 17 (2014): 174107. doi:[10.1063/1.4898625](https://doi.org/10.1063/1.4898625). Posted with permission.

A coarse-graining approach for molecular simulation that retains the dynamics of the all-atom reference system by implementing hydrodynamic interactions

Sergiy Markutsya, and Monica H. Lamm

Citation: *The Journal of Chemical Physics* **141**, 174107 (2014); doi: 10.1063/1.4898625

View online: <http://dx.doi.org/10.1063/1.4898625>

View Table of Contents: <http://aip.scitation.org/toc/jcp/141/17>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Dynamic force matching: A method for constructing dynamical coarse-grained models with realistic time dependence](#)

The Journal of Chemical Physics **142**, 154104 (2015); 10.1063/1.4917454

[Dynamic force matching: Construction of dynamic coarse-grained models with realistic short time dynamics and accurate long time dynamics](#)

The Journal of Chemical Physics **145**, 224107 (2016); 10.1063/1.4971430

[Conservative and dissipative force field for simulation of coarse-grained alkane molecules: A bottom-up approach](#)

The Journal of Chemical Physics **140**, 134113 (2014); 10.1063/1.4870394

[A test of systematic coarse-graining of molecular dynamics simulations: Transport properties](#)

The Journal of Chemical Physics **139**, 094107 (2013); 10.1063/1.4819472

[Modeling real dynamics in the coarse-grained representation of condensed phase systems](#)

The Journal of Chemical Physics **125**, 151101 (2006); 10.1063/1.2360580

[Incorporation of memory effects in coarse-grained modeling via the Mori-Zwanzig formalism](#)

The Journal of Chemical Physics **143**, 243128 (2015); 10.1063/1.4935490

Scilight

Sharp, quick summaries **illuminating**
the latest physics research

Sign up for **FREE!**



A coarse-graining approach for molecular simulation that retains the dynamics of the all-atom reference system by implementing hydrodynamic interactions

Sergiy Markutsya¹ and Monica H. Lamm^{1,2,a)}

¹Ames Laboratory, Iowa State University, Ames, Iowa 50011, USA

²Department of Chemical and Biological Engineering, Iowa State University, Ames, Iowa 50011, USA

(Received 14 July 2014; accepted 7 October 2014; published online 4 November 2014)

We report on a new approach for deriving coarse-grained intermolecular forces that retains the frictional contribution that is often discarded by conventional coarse-graining methods. The approach is tested for water and an aqueous glucose solution, and the results from the new implementation for coarse-grained molecular dynamics simulation show remarkable agreement with the dynamics obtained from reference all-atom simulations. The agreement between the structural properties observed in the coarse-grained and all-atom simulations is also preserved. We discuss how this approach may be applied broadly to any existing coarse-graining method where the coarse-grained models are rigorously derived from all-atom reference systems. © 2014 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4898625>]

I. INTRODUCTION

Computer simulation is a powerful tool for investigating the fundamental principles of physics, chemistry, and biology. Molecular simulation, in particular, reveals microscopic details that are difficult, if not impossible, to observe solely by experimental or theoretical approaches. For example, classical molecular dynamics (MD) simulation¹ is commonly used to model the structure and dynamics of a system at the atom scale. Because the pairwise forces between atoms are calculated at every time step in MD, the system size that can be conveniently handled is about a million atoms for a time scale on the order of 100 ns. There are all-atom MD simulations that have reached the microsecond time scale^{2–4} and even the millisecond time scale⁵ but typically these are exceptional cases that model relatively smaller system sizes, $O(10\ 000)$ atoms. However, the phenomena of interest for many practical problems typically involve micrometer length scales (i.e., hundreds of millions of atoms) and occur on microsecond to millisecond time scales.⁶

To bridge MD simulations from the atomic scale to the mesoscale, coarse-grain (CG) models are developed. These CG models combine multiple atoms into one group (a CG bead), thereby reducing the degrees of freedom in the system by up to an order of magnitude. Once the CG beads have been defined, the effective forces (or potentials) between these CG beads are derived from all-atom simulation trajectories or fit to experimental data.

There are several coarse-graining methods that are used to derive effective CG forces (or potentials): (i) fitting free energy in the system,^{7,8} (ii) structure-based methods that reproduce a pre-defined target structure,^{9–12} and (iii) force matching approaches, where the instantaneous CG forces are fitted

to the forces from all-atom molecular simulations.^{13–16} These approaches yield CG models that, when implemented in MD simulation (CGMD), have successfully modeled the structure and thermodynamics of systems up to the mesoscale. The reduction in degrees of freedom eliminates much of the friction between CG beads and results in faster CG dynamics.¹⁷ This is advantageous for the efficient sampling of phase space but problematic if one wishes to calculate a time-dependent property from CGMD.

One way to address the problem of incorrect (fast) CG dynamics is to incorporate hydrodynamic interactions (HI) into the CG model, because HI can play a significant role in dynamic properties of a system. HI can be implemented by using equations of motion based on the generalized Langevin equation (GLE) formalism introduced by Mori and Zwanzig.^{22,23} The GLE is exact for a set of reduced variables but, as discussed by Higon *et al.*,²⁴ it is not fully explicit and there are several challenges that prohibit the use of the GLE in practical applications. Hence, one or more approximations are made to the GLE to obtain equations of motion in a form that is practical for applications.^{24–29}

For example, the Markovian multidimensional Langevin equation³⁰ obtained from the GLE is

$$\mathbf{M} \frac{d\mathbf{v}(t)}{dt} = \mathbf{F}(\mathbf{q}) - \boldsymbol{\zeta} \mathbf{v}(t) + \boldsymbol{\eta}(t), \quad (1)$$

where \mathbf{M} is the diagonal matrix of particle masses, \mathbf{q} are the generalized coordinates, $\boldsymbol{\zeta}$ is the friction tensor, \mathbf{v} is the velocity vector, \mathbf{F} is the force vector exerted on CG beads due to potential energy of interaction with other CG beads, and $\boldsymbol{\eta}$ is the noise force vector. The friction tensor $\boldsymbol{\zeta}$ is symmetric to satisfy symmetry in the Markovian noise correlation. The term $\boldsymbol{\zeta}$ is the $3N \times 3N$ friction tensor, and all the vectors are $3N$ -dimensional, where N is the number of CG beads in the system. The multidimensional nature of the Langevin equation in Eq. (1) introduces HI into the CG system, where

^{a)} Author to whom correspondence should be addressed. Electronic mail: mhlyamm@iastate.edu.

the disturbance of the velocity of one particle influences the friction exerted on other particles. In a practical sense, the Langevin equation with HI (Eq. (1)) has limited applicability due to the difficulty of defining a friction tensor for every system.

For systems with large CG beads compared to the size of solvent molecules, the position and velocity Markovian Langevin equation (Eq. (1)) can be reduced to the position-only Brownian dynamic (BD) equation.³¹ In Brownian dynamics with HI, a diffusion tensor \mathbf{D} (diffusion tensor and friction tensor related as $\mathbf{D} \sim \boldsymbol{\zeta}^{-1}$) is typically obtained with Oseen or Rotne-Prager approximations.^{32,33} Brownian dynamics with HI has been successfully applied to study such phenomena as protein folding,³⁴ the stretching and orientation of polymers under flow,³⁵ and DNA dynamics.³⁶ The application of BD with HI is limited to dilute systems with large solute/solvent size separation.

For the majority of applications, a simple Langevin equation with diagonal friction tensor ζ_i is used³⁷

$$m_i \frac{dv_i(t)}{dt} = F_i - \zeta_i v_i(t) + \eta(t), \quad (2)$$

where m_i is the mass of particle i , F_i is the force exerted on particle i due to potential energy interaction with other particles, $v_i(t)$ is the velocity of particle i , and $\eta(t)$ is the noise term with the properties obtained from the second fluctuation-dissipation theorem⁴⁸

$$\langle \eta_i(t) \rangle = 0, \quad (3)$$

$$\langle \eta_i(t_0) \eta_j(t_0 + t) \rangle = 2k_B T_{\text{ref}} \zeta_i \delta_{ij} \delta(t), \quad (4)$$

where k_B is the Boltzmann constant, and T_{ref} is the reference temperature. Equation (2) was originally derived to describe the dynamics of a Brownian particle across time scales^{39,40} and it has been successfully used to model the dynamics of many other systems.⁴¹⁻⁴³ Unfortunately, by replacing the complete friction tensor (Eq. (1)) with a diagonal friction tensor, Eq. (2) is not able to reproduce HI any more. Although, even without HI, it is possible to correctly reproduce the dynamic properties of some CG systems by choosing an appropriate diagonal friction tensor. There are multiple methods described in literature that address this problem. For example, one may address the problem of fast CG dynamics by rescaling the time during the simulation. A simple scaling factor for converting the CG time to the reference system time scale can be obtained by comparing the diffusion coefficients from the CGMD and reference MD simulations.^{18,19} For supercooled liquids, applying temperature-scaling to the CGMD simulation has been shown to better reproduce the all-atom dynamics when compared to time-scaling, although accuracy in the structure predicted by the temperature-scaling approach was compromised.²⁰ Rescaling approaches are simple to implement for homogeneous systems but become more difficult for heterogeneous systems, where the scale factor must necessarily be a tensor.²¹

Another approach to obtain accurate CG dynamics is to put the dynamics from the missing degrees of freedom back into the system by using equations of motion based on the generalized Langevin equation formalism. The advantage

of the GLE approach for CGMD is that it provides a basis for parameterizing the equations of motion as a function of the coarse-grained model itself. For example, Izvekov and Voth derived friction coefficients for one- and two-site coarse-grained models of methanol by using a GLE equation formulated for autocorrelation functions.²⁵ The force-velocity and velocity-velocity correlation functions for the coarse-grained beads were computed directly from all-atom MD trajectories for methanol. These correlation functions and the diffusion coefficient calculated from the all-atom MD simulation were used to parameterize the GLE-based equations of motion. The resulting GLE simulation with the CG model predicted a diffusion coefficient in good agreement with the diffusion coefficient obtained with the all-atom model. This strategy is straightforward for homogeneous systems with one or two types of coarse-grained beads but it may be problematic for heterogeneous systems with several CG bead types. Specifically, for heterogeneous systems with many types of CG beads the number of unique terms for the friction tensor significantly increases and this requires one to compute the autocorrelation function for each unique pair type of CG beads in the system. Accurate calculation of the autocorrelation functions, in turn, requires good statistical sampling of the all-atom MD trajectory. Hence, this approach may be computationally prohibitive for systems with many types of CG beads.

In the spirit of developing a universal coarse-graining approach for molecular systems, Shell has shown that relative entropy between a CG model and the all-atom reference system is the fundamental quantity to be minimized when deriving a CG model.⁴⁴ When this approach is applied to simple liquids, the resulting CG models are capable of retaining the information needed to accurately model the thermodynamic properties and the diffusion coefficients of the reference system.⁴⁵ The relative entropy framework has been further generalized by Espanol and Zuniga⁴⁶ to address the dynamic regime in a way that circumvents the need for calculating components of the friction tensor. The minimum relative entropy approach is a promising strategy but there remain obstacles, such as practical implementations for solution of the optimization problem,⁴⁷ that currently prevent it from being widely adopted.

In this article, a rigorous and self-consistent method for deriving and implementing Langevin dynamics with hydrodynamic interactions for CGMD simulations (Eq. (1)) is presented. This method is based on the well-known Langevin equation formalism discussed above. However, compared to previous implementations of the Langevin equation for CGMD, this paper describes how to compute the friction tensor directly during the derivation of the CG force field, using routine post-processing calculations. Examples are shown to demonstrate that the method is applicable to homogeneous and heterogeneous systems alike. The application of this approach to other coarse-graining methods is straightforward.

The paper is organized as follows: in Sec. II, a self-consistent method for deriving Langevin equations with hydrodynamic interactions using CG force fields and implementing them in CGMD simulations to obtain accurate dynamic properties is described. The details of the all-atom MD

and CGMD simulations are given in Sec. III. In Sec. IV, two examples are used to demonstrate the effectiveness of the approach for retaining dynamics. In Sec. V, the paper concludes with a discussion of how this self-consistent approach may be broadly applied with other coarse-graining methods.

II. THEORY

In this section, we briefly review the second fluctuation-dissipation theorem and describe the central idea that motivates our approach for deriving the parameters for the multidimensional Langevin equation (Eq. (1)) from data collecting during the coarse-graining procedure.

A. The second fluctuation-dissipation theorem relates the friction tensor to the noise force vector

The well-known second fluctuation-dissipation theorem³⁸ governs the relationship between the frictional forces and the random forces in the multidimensional Langevin equation (Eq. (1)). Berendsen's⁴⁸ formulation of the second fluctuation-dissipation theorem is particularly convenient for practical applications and shall be used here. Briefly, one first recognizes that the noise force vector can be expressed as linear combinations of independent normalized white noise functions $\boldsymbol{\eta}^0(t)$,

$$\boldsymbol{\eta}(t) = \mathbf{B}\boldsymbol{\eta}^0(t), \quad (5)$$

where \mathbf{B} is a $3N \times 3N$ matrix. The vector of normalized white noise functions, also known as Gaussian random noise $\boldsymbol{\eta}^0(t)$ with the following properties:

$$\langle \boldsymbol{\eta}^0(t) \rangle = 0, \quad (6)$$

$$\langle \boldsymbol{\eta}^0(t_0)\boldsymbol{\eta}^0(t_0 + t) \rangle = \mathbf{I}\delta(t), \quad (7)$$

where \mathbf{I} is a $3N \times 3N$ unit matrix. The friction term and the noise terms are then related through the second fluctuation-dissipation theorem as

$$\mathbf{B}\mathbf{B}^T = 2\zeta k_B T_{\text{ref}}. \quad (8)$$

Substituting Eqs. (5) and (8) into Eq. (1) and rearranging the two rightmost terms yields

$$\mathbf{M} \frac{d\mathbf{v}(t)}{dt} = \mathbf{F}(\mathbf{q}) + \mathbf{B} \left(\boldsymbol{\eta}^0(t) - \frac{\mathbf{B}^T \mathbf{v}(t)}{2k_B T_{\text{ref}}} \right). \quad (9)$$

For practical application in numerical simulations, it is convenient to rewrite Eq. (9) in discrete form using a simple Euler scheme

$$\begin{aligned} & \mathbf{M}[\mathbf{v}(t + \Delta t) - \mathbf{v}(t)] \\ &= \mathbf{F}(\mathbf{q})\Delta t + \mathbf{B} \left(\sqrt{\Delta t} \mathbf{N}(0, 1) - \frac{\mathbf{B}^T \mathbf{v}(t)}{2k_B T_{\text{ref}}} \Delta t \right), \end{aligned} \quad (10)$$

where $\mathbf{N}(0, 1)$ is the vector of random numbers generated with the standard normal distribution, and Δt is the computational time step.

In the literature, Eq. (8) is often used to determine the matrix \mathbf{B} from a known friction tensor. The solution of this square-root operation is not unique and methods such as

the Choleski decomposition,⁵³ the Chebyshev polynomial approximation to the decomposition,⁵⁰ or the ‘‘truncated expansion ansatz’’ approach⁵¹ are used. With the absence of a unified approach to determine the non-diagonal friction tensor components, this approach has limited practical application. Conversely, if one could first determine the matrix \mathbf{B} , and then solve Eq. (8) for the friction tensor ζ , the solution would yield unique values for the friction tensor. This idea, that the friction tensor may be obtained *from* the random noise forces, is central to the new implementation presented in this paper and it will be discussed further below.

B. The random noise force vector in the Langevin equation can be derived from the random force fluctuations observed during coarse-graining

The next key step in the development of our approach is to show that the matrix \mathbf{B} required to obtain the random and frictional terms in Eq. (10) can be derived from the random force fluctuations observed during coarse-graining procedure. We start with the supposition that the random noise force vector $\boldsymbol{\eta}(t)$ in the multidimensional Langevin equation (Eq. (1)) can be directly computed during the coarse-graining procedure. Generally, $\boldsymbol{\eta}(t)$ represents the stochastic fluctuation forces that are exerted on particles by the surrounding (and implicitly modeled) fluid. In the context considered here, the coarse-grained sites are the ‘‘particles’’ and the degrees of freedom removed from the all-atom reference system during coarse-graining are analogous to the ‘‘surrounding fluid.’’ These random fluctuation forces may be directly observed during the coarse-graining procedure as follows. Typically, during coarse-graining, only the average pairwise CG force versus separation distance is computed. To directly observe the effect of the random fluctuation forces one may, additionally, compute the probability distribution function (PDF) of the pairwise CG force versus separation distance.

To illustrate this concept, we provide a brief example using data collected during a coarse-graining procedure for a system of bulk water. Our intent in this section is to focus on outcomes; a complete description of the all-atom reference simulation and the coarse-graining procedure will be provided in Sec. III. Fig. 1(a) shows the average pairwise force versus separation distance between two CG water sites. This average pairwise force is a typical example of the outcome that one would obtain from any conventional coarse-graining approach. The PDF of the pairwise CG force between two CG water sites, at a separation distance of 3.4 Å, is shown in Fig. 1(b). This distribution of forces about the average value is a direct result of the degrees of freedom that were eliminated during the coarse-graining procedure. To quantify this impact directly, Fig. 1(c) shows the difference between the average pairwise CG force and the PDF of the pairwise CG force, again at a separation distance of 3.4 Å.

When the pairwise force PDFs are calculated (as for the example shown in Fig. 1) the actual force, $\tilde{\mathbf{F}}$, for each CG bead i due to interaction with all other CG beads j can be written as

$$\tilde{\mathbf{F}}(\mathbf{q}, t) = \mathbf{F}(\mathbf{q}) + \Delta\mathbf{F}(\mathbf{q}, t), \quad (11)$$

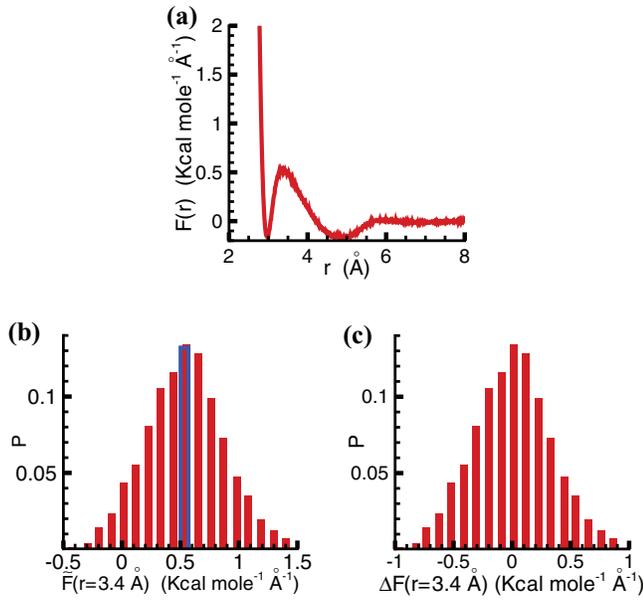


FIG. 1. (a) The magnitude of the average effective force $F(r)$ between CG water molecules as the outcome from conventional coarse-graining for bulk water. Simulation details are presented in Sec. III. (b) A representative PDF of magnitude of the actual force $\tilde{F}(r)$ at separation $r = 3.4 \text{ Å}$. The blue bar represents the average value. (c) The PDF of magnitude of the error fluctuation force $\Delta F(r, t)$ at separation $r = 3.4 \text{ Å}$. A separation r represents the distance between centers of two CG beads calculated along the line connected their centers and is calculated as $r = |\mathbf{r}_{ij}| = |\mathbf{q}_i - \mathbf{q}_j|$, where \mathbf{q}_i and \mathbf{q}_j are coordinates of the centers of i th and j th beads.

where the first term $\mathbf{F}(\mathbf{q})$ represents the contribution from the average pairwise CG forces and the second term $\Delta\mathbf{F}(\mathbf{q}, t)$ denotes the contribution from the random fluctuation forces that account for the eliminated degrees of freedom. Thus, the first term is computed as the force that acts on i th CG bead due to interaction with all others CG beads due to CG average pairwise force at time t (see Fig. 1(a)). The random fluctuation force, $\Delta\mathbf{F}(\mathbf{q}, t)$, is then computed as the difference between the actual pairwise force $\tilde{\mathbf{F}}(\mathbf{q}, t)$ and the average pairwise force $\mathbf{F}(\mathbf{q})$ at some instance of time t (see Fig. 1(c)). Hence, the dependence on time is included in the random force contribution because each component of the fluctuation force at any t has a different value, even when the relative distances at different times are exactly the same. We emphasize that this definition for $\Delta\mathbf{F}(\mathbf{q}, t)$ and its physical meaning are distinctly different than the instantaneous difference between all-atom and coarse-grained forces used in previous work.²⁵

Comparing Eq. (11) multiplied by the computational time step Δt to the Markovian Langevin equation (Eq. (10)) provides a direct link between the coarse-graining procedure and the friction and noise parameters that must be defined to accurately model the dynamics of the coarse-grained system. Clearly, the average pairwise CG force $\mathbf{F}(\mathbf{q})$ in Eqs. (10) and (11) is the same. The random fluctuation force $\Delta\mathbf{F}(\mathbf{q}, t)$ in Eq. (11) is the time dependent contribution of the “surrounding fluid,” that is equivalent to the second term in Eq. (10) that consists of the random and friction contributions. There is no established way to split the fluctuation force $\Delta\mathbf{F}(\mathbf{q}, t)$ into the random and friction part. This issue can be resolved for cases where the random term

$\sqrt{\Delta t} \mathbf{N}(0, 1) \gg (\mathbf{B}^T \mathbf{v}(t)/2k_B T_{\text{ref}}) \Delta t$ (see Appendix A for more details). When such conditions hold, the random fluctuation force $\Delta\mathbf{F}(\mathbf{q}, t)$ is related to the random term in the Langevin equation (Eq. (10))

$$\Delta\mathbf{F}(\mathbf{q}, t)\Delta t \sim \mathbf{B}\sqrt{\Delta t}\mathbf{N}(0, 1). \quad (12)$$

For exact equivalence between the random fluctuation force and the random white noise, a coefficient of proportionality $\sqrt{\Delta t_0}/\sqrt{\Delta t}$ is used to satisfy the assumption applied to the derivation of Eq. (1) that the friction tensor ζ (and corresponding matrix \mathbf{B}) is time independent. Thus, Eq. (12) is rewritten as

$$\sqrt{\frac{\Delta t_0}{\Delta t}} \Delta\mathbf{F}(\mathbf{q}, t)\Delta t = \sqrt{\Delta t_0 \Delta t} \Delta\mathbf{F}(\mathbf{q}, t) = \mathbf{B}\sqrt{\Delta t}\mathbf{N}(0, 1), \quad (13)$$

where Δt_0 is a characteristic time to be defined later. This equation establishes a direct route for determining the components of matrix \mathbf{B} , used in the Langevin equation (10), based on knowledge of the random fluctuation force, $\Delta\mathbf{F}(\mathbf{q}, t)$, that is computed during the coarse-graining procedure.

C. Direct implementation

As seen from Eq. (13) the random fluctuation force $\Delta\mathbf{F}(\mathbf{q}, t)$ is proportional to the standard normal distribution. This fact motivates a comparison of the PDF for $\Delta\mathbf{F}(\mathbf{q}, t)$ to a normal distribution. Fig. 2 shows that the PDF data for bulk water are represented well by the normal distribution $\mathbf{N}(0, \mathbf{F}_0^2(\mathbf{q}))$, with zero mean and variance $\mathbf{F}_0^2(\mathbf{q})$. The goodness of fit is quantified by computation of χ^2 .^{49,52} A value of $\chi^2 = O(10^{-2})$ is obtained for the PDF data in Fig. 2. Thus, the PDF for the random fluctuation force may be written as

$$\Delta\mathbf{F}(\mathbf{q}, t) = \mathbf{N}(0, \mathbf{F}_0^2(\mathbf{q})) = \mathbf{F}_0(\mathbf{q})\mathbf{N}(0, 1). \quad (14)$$

Combining Eqs. (13) and (14) the matrix \mathbf{B} then becomes

$$\mathbf{B} = \sqrt{\Delta t_0} \mathbf{F}_0(\mathbf{q}), \quad (15)$$

where the standard deviation of the random term $\mathbf{F}_0(\mathbf{q})$ is a $3N \times 3N$ matrix, and the components of this standard deviation have the property $F_{0,ij} = -F_{0,ji}$. Hence, the matrix \mathbf{B} is skew-symmetric and $-\mathbf{B} = \mathbf{B}^T$. Thus, by substituting Eq. (15)

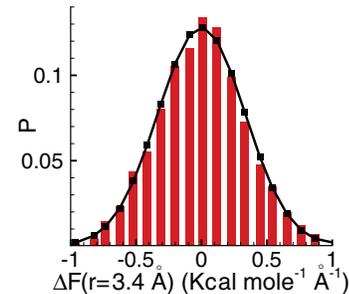


FIG. 2. The PDF for a magnitude of the random fluctuation force $\Delta F(r, t)$ for CG water–CG water beads at separation $r = 3.4 \text{ Å}$ (represented by red bars) and the corresponding normal distribution fitting $\mathbf{N}(0, \mathbf{F}_0^2)$ (represented by black squares connected by line for better visual perception) with the standard deviation $F_0(r = 3.4 \text{ Å}) = 0.333 \text{ Kcal mole}^{-1} \text{Å}^{-1}$. Simulation details are presented in Sec. III.

into Eq. (10) the Langevin equation velocities can be updated

$$\begin{aligned} \mathbf{M}[\mathbf{v}(t + \Delta t) - \mathbf{v}(t)] \\ = \mathbf{F}(\mathbf{q})\Delta t \\ + \mathbf{F}_0(\mathbf{q})\mathbf{N}(0, 1)\sqrt{\Delta t_0\Delta t} + \frac{\mathbf{F}_0(\mathbf{q})\mathbf{F}_0(\mathbf{q})\mathbf{v}(t)}{2k_B T_{\text{ref}}}\Delta t_0\Delta t. \end{aligned} \quad (16)$$

The values for standard deviation matrix $\mathbf{F}_0(\mathbf{q})$ in Eq. (16) are obtained by fitting the random fluctuation force $\Delta\mathbf{F}(\mathbf{q}, t)$ for each type of CG bead pair to a normal distribution, and the random fluctuation force $\Delta\mathbf{F}(\mathbf{q}, t)$ is obtained from the block-average force matching coarse-graining procedure. According to this method, all N atomistic configurations split into m equi-size blocks of n configurations each, such that $N = m*n$. The forces shown in Fig. 1 are computed from all the m blocks. Using this method the PDFs of the CG forces were calculated based on the ‘‘pre-average’’ actual CG forces $\tilde{\mathbf{F}}(\mathbf{q})$ obtained from each m block. As a result, the fluctuation forces may have incorrect magnitude since they were build based on pre-averaged values of the CG force-distance curves rather than based on the raw forces. Thus, a value of a characteristic time Δt_0 is selected to recover the correct magnitude of the random fluctuation force $\mathbf{F}_0(\mathbf{q})$.

In the most general case, a CG system contains several different types of CG beads. Thus, values of the standard deviation matrix $F_{0,ij}^{A_i,A_j}(r_{ij})$ and the characteristic time $\Delta t_0^{A_i,A_j}$ will be different for different CG type pairs, and the final Langevin equation with hydrodynamic interaction is written in components form as

$$\begin{aligned} m_i[v_i(t + \Delta t) - v_i(t)] \\ = F_i(r)\Delta t + \sum_j F_{0,ij}^{A_i,A_j}(r_{ij})\sqrt{\Delta t_0^{A_i,A_j}}\Delta t N_j(0, 1) \\ + \frac{1}{2k_B T_{\text{ref}}}\sum_j \sum_k F_{0,ik}^{A_i,A_k}(r_{ik})\sqrt{\Delta t_0^{A_i,A_k}} \\ \times F_{0,kj}^{A_j,A_k}(r_{kj})\sqrt{\Delta t_0^{A_k,A_j}} v_j(t)\Delta t, \end{aligned} \quad (17)$$

where A_i and A_j represent types of i th and j th CG beads correspondingly.

Once the characteristic times $\Delta t_0^{A_i,A_j}$ are determined (see details in Appendix B), the PDF-based CGMD approach with hydrodynamic interactions can be implemented.

III. METHODS

The performance of the new PDF-based CGMD method with hydrodynamic interactions (Eq. (17)) was tested on the two systems: (i) water with 1048 TIP3P⁵⁴ water molecules in a periodic cubic box with side length 3.2 nm, and (ii) a single glucose molecule solvated with 1064 TIP3P water molecules in a periodic cubic box with side length 3.2 nm (a dilute glucose solution). The following paragraphs describe the procedures used for the all-atom MD simulations, the derivation of the CG model by force matching, and the CGMD and PDF-based CGMD simulations.

The all-atom molecular dynamics simulations were carried out using GROMACS.⁵⁵ The glucose molecule was modeled using a modified CHARMM36⁵⁶ force field obtained from Oak Ridge National Laboratory.⁵⁷ To speed up forces calculations, the cutoff radius of 8.5 Å was chosen. The long-range electrostatic interactions were treated with the particle mesh Ewald (PME) method.⁵⁸ The cutoff radius for electrostatic interactions was 12.0 Å. The LINCS algorithm was used to fix the bond lengths to hydrogen atoms of glucose molecule and SETTLE algorithm⁵⁹ was used for water molecules.⁶⁰ The system was then equilibrated using constant pressure simulations. The production runs were carried out at constant temperature $T = 298$ K using a Nosé-Hoover thermostat. Time step in the all-atom molecular dynamics simulations was 2.0 fs. Time evolution of the system for 20 ns was studied with configurations and force data saved every 1 ps.

Coarse-grained force fields were derived based on the all-atom data set of trajectories and forces for glucose molecule and water using the multiscale coarse-graining method^{13,14} based on force matching. A detailed description of the force matching method can be found elsewhere.^{13,15,16} Briefly, for a given configuration from the reference all-atom molecular dynamics simulation, the positions of N coarse-grained beads and the net forces $\mathbf{F}_i^{\text{ref}}$ acting on them are computed. Recognizing that the net force on a particular coarse-grained bead i is due to the sum over all the effective forces between pairs of coarse-grained beads leads to the construction of an objective function

$$\sum_j^N \mathbf{f}_{ij}(\mathbf{q}_i, \mathbf{q}_j, g_1, g_2, \dots, g_m) = \mathbf{F}_i^{\text{ref}}, \quad (18)$$

where \mathbf{f}_{ij} is the force on the i th coarse-grained bead due to the j th coarse-grained bead, and \mathbf{q}_i and \mathbf{q}_j are the position vectors for the i th and j th coarse-grained bead, respectively. The analytical function for \mathbf{f}_{ij} is not known *a priori* so cubic splines are chosen to systematically write the pairwise force as a linear function of m unknown parameters (g_1, g_2, \dots, g_m). Hence, a system of N linear equations with m unknowns is obtained. This system of equations is over-determined ($N > m$) and can be solved using the singular value decomposition (SVD) method.⁶¹ There are two ways the SVD method can be applied to the over-determined system. It can be applied to all the atomistic configurations, then outcome is the CG force versus separation distance. Another way is to split all atomistic configurations into the equi-size blocks and apply SVD to each of these blocks. In second case, the CG force versus separation distance is obtained by averaging outcomes of all the equi-sized blocks. In the present article, the block-average force-matching method is used as described in Sec. II and the standard deviations versus separation ($F_{0,ij}^{A_i,A_j}(r_{ij})$) are obtained for all pair types of CG beads as shown on Fig. 3.

The components of the standard deviation of the random term are computed during force calculation in CG MD code as $F_{\alpha 0,ij}^{A_i,A_j} = F_{0,ij}^{A_i,A_j} \frac{r_{\alpha,ij}}{|\mathbf{r}_{ij}|}$, where $\alpha = x, y, z$.

The coarse-grain glucose molecule is represented with single CG bead and each coarse-grain water molecule is represented with single CG bead (W) as described in previous

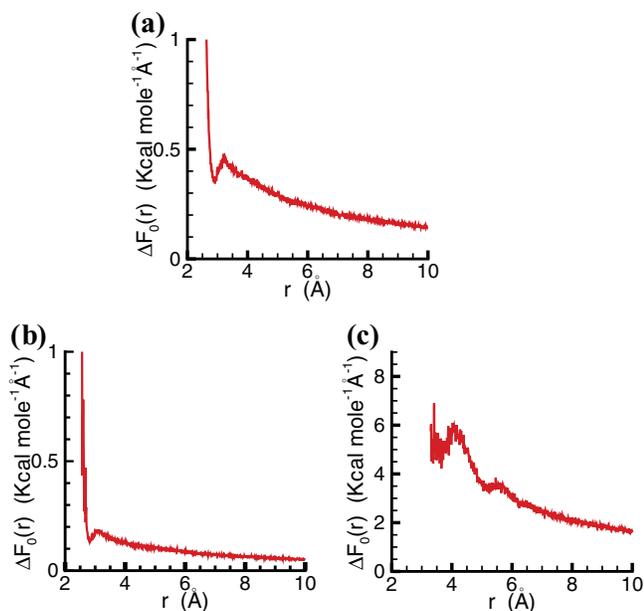


FIG. 3. The standard deviations $F_0(r)$ as function of the separation extracted from the random fluctuation forces $\Delta\mathbf{F}(\mathbf{q}, t)$ for (a) CG water molecules for the system of bulk water, (b) CG water molecules for the system of single glucose in water, (c) CG glucose-CG water.

work.⁶² All CG beads are defined at the center of mass of the atoms that correspond to the beads. In the CG model for bulk water, there are non-bonded interactions between CG water beads. For CG model of glucose in CG water, there are non-bonded interactions between glucose bead and water beads, and between water beads.

All conventional CGMD simulations were carried out using LAMMPS.^{63,64} Conventional CGMD simulations used the average pairwise CG force and the Newtonian equation of motion,

$$\mathbf{M} \frac{d\mathbf{v}(t)}{dt} = \mathbf{F}(\mathbf{q}). \quad (19)$$

PDF-based CGMD simulations were performed using an in-house modification for LAMMPS where the average pairwise CG force, random force, and frictional force were incorporated according to Eq. (17). Before PDF-based CGMD runs can be started the characteristic time Δt_0 needs to be calculated for every unique pair type of CG beads in the system, as described in Appendix B. For both conventional CGMD and PDF-based CGMD runs, the system was equilibrated for 1 ns followed by a 30 ns production run at constant temperature $T = 298$ K. A Nosé-Hoover thermostat was used for temperature control in the conventional CGMD runs. In the PDF-based CGMD simulations for the system with bulk water, Eq. (17) serves as a Langevin thermostat for the temperature control. A computational time step $\Delta t = 0.5$ fs is used for this system. For the system with single glucose molecule solvated with water molecules, Eq. (17) was applied to the CG bead of glucose molecule and CG beads of water molecules and serves as a Langevin thermostat for the temperature control. A computational time step $\Delta t = 2.0$ fs is used for this system. To speed up forces calculations for all CGMD simulations, the cutoff radius of 8.5 Å was chosen. Configurations were

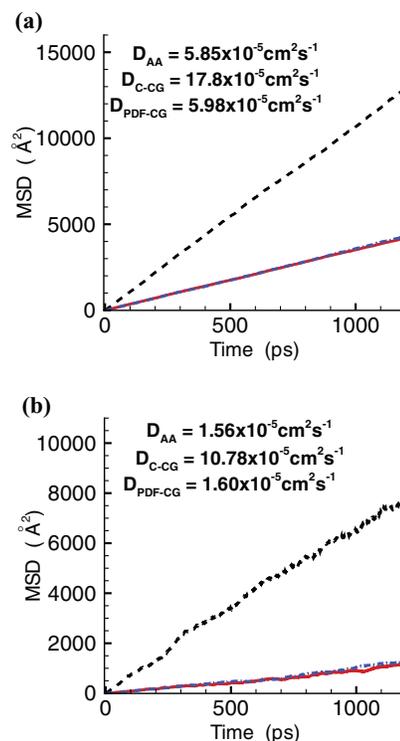


FIG. 4. The MSD of: (a) a water molecule in bulk water, and (b) a glucose molecule in bulk water computed from all-atom MD, conventional CGMD, and PDF-based CGMD. Diffusion coefficients calculated using Eq. (20) are identified as D_{AA} (solid red line), D_{C-CG} (black dashed line), and D_{PDF-CG} (blue dashed-dotted line), respectively.

saved every 1 ps during the production run. The initial configuration and dimension of the simulation box were identical to those used in the all-atom molecular dynamics simulations.

IV. RESULTS

A. Diffusion coefficients

The impact of the PDF-based CGMD approach on dynamic properties is clearly observed by comparing diffusion coefficients calculated from all-atom MD, conventional CGMD, and PDF-based CGMD. Fig. 4 shows the mean-square displacement (MSD) versus time and the diffusion coefficient, D , calculated from the MSD data according to Einstein's relation⁶⁵

$$D = \frac{\langle (\mathbf{q}(0) - \mathbf{q}(t))^2 \rangle}{(2d)t}, \quad (20)$$

where $\langle (\mathbf{q}(0) - \mathbf{q}(t))^2 \rangle$ is the MSD in d -dimensional space (here, $d = 3$) during time interval t .

Einstein's equation is valid only in the diffusive regime where $t \gg \tau_p$. Here, $\tau_p = m/\xi$ is the momentum relaxation time of molecule with mass m , and friction coefficient ξ . The MSD of the center of mass of each CG water molecule (bulk water) and the MSD of the center of mass of the CG glucose molecule (aqueous glucose solution) was computed for each of these time intervals. The production runs were split into $t = 1.2$ ns time intervals and the final MSD data are obtained by averaging over all the time intervals. Diffusion coefficients

were calculated by fitting MSD versus time with a linear function (the first 100 ps of MSD data were excluded from the diffusion calculations because these data did not satisfy the diffusive regime (linearity) requirement).

The self-diffusion coefficient calculated for water in the all-atom MD simulation is $5.85 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$, which is in agreement with a previously reported self-diffusion coefficient for TIP3P water.⁶⁶ As expected, the diffusion coefficient calculated from conventional CGMD ($17.8 \times 10^{-5} \text{ cm}^2 \text{ s}^{-1}$) is higher than those calculated from the reference all-atom MD simulations, while the diffusion coefficients for PDF-based CGMD and the reference all-atom simulations agree remarkably well. The same great match of the diffusion coefficients for PDF-based CGMD and the reference all-atom simulations is observed for the system of single glucose molecule solvated with water molecules.

B. Bulk viscosity

Another transport property of interest is bulk viscosity, η_v , which can be obtained from the correlation of pressure fluctuations in a system at equilibrium using the Green-Kubo equation

$$\eta_v = \frac{V}{k_B T_{\text{ref}}} \int_0^\infty \langle \delta \mathcal{P}(t) \delta \mathcal{P}(0) \rangle dt, \quad (21)$$

where V is the volume, $\mathcal{P}(t)$ is the instantaneous pressure, computed as the average of the diagonal elements of the stress operator, $\langle \mathcal{P} \rangle$ is the average pressure of the system, and $\delta \mathcal{P}(t) = \mathcal{P}(t) - \langle \mathcal{P} \rangle$ is the pressure fluctuation. The diagonal elements of the stress operator can be found as

$$\mathcal{P}_{\alpha\alpha} = \frac{1}{V} \left(\sum_i \frac{p_{i\alpha} p_{i\alpha}}{m_i} + \sum_i q_{i\alpha} f_{i\alpha} \right), \quad (22)$$

where the first term on the right hand side is called a kinetic term, and second term on the right hand side is called a virial term (or potential term), $\alpha = x, y, z$, $p_{i\alpha}$ is the momentum of atom i , $q_{i\alpha}$ is the position of atom i , and $f_{i\alpha}$ is the systematic force due to potential interactions of atom i and all other atoms. For the purposes of direct comparison, the correlation of pressure fluctuations was calculated and compared for all-atom MD, conventional CG, and PDF-based CG. Pressure was obtained from GROMACS MD package for all-atom MD simulations and LAMMPS MD package for CG MD simulations. For the system of bulk water, the total pressure for all-atom MD simulations was used. For both conventional CG MD and PDF-based CGMD simulations, the pressure calculated based on systematic force was used. Pressure was recorded at every 2 fs for a total time of 8 ns after system reached an equilibrium. The total trajectory was split into 4 ps sub-trajectories (2000 sub-trajectories in total), and the averaging was performed for these sub-trajectories.

Comparison of the pressure correlation function for all-atom MD, conventional CGMD, and PDF-based CGMD is shown in Fig. 5. It can be seen that the correlation of the pressure fluctuation computed from conventional CGMD has much smaller value and decays with a shorter characteristic timescale (time when autocorrelation function reaches zero

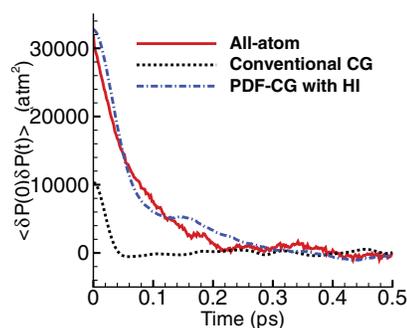


FIG. 5. The correlation of the pressure fluctuation for all-atom MD (red solid line), conventional CG (black dashed line), PDF-based CG (blue dashed-dotted line).

value is 0.06 ps) when compared to all-atom MD (0.22 ps). The pressure correlation function computed for PDF-based CGMD has similar magnitude and decays with a timescale much closer to the all-atom MD data. These results suggest that the bulk viscosities (when computed using Eq. (21)) obtained from the PDF-based CGMD and all-atom MD will also be in good agreement. It is possible to conclude that implementation of HI into Langevin dynamics reproduces pressure fluctuations of the system much better without pressure correction term when comparing with CG methods with no HI.^{14,67,68}

C. Radial distribution functions

Structural properties can be characterized by calculating the radial distribution function $g(r)$. To make direct comparisons between the coarse-grained model and the all-atom model, the configurations from the all-atom molecular dynamics simulation are first reduced to the CG mapping before calculating the radial distribution function. Because the radial distribution function represents the average structure for a system at equilibrium, it is expected that $g(r)$ obtained from the PDF-based CGMD method would, at least, reproduce the $g(r)$ observed from the conventional CGMD method. Fig. 6 shows excellent agreement between the radial distribution functions computed from conventional CGMD and PDF-based CGMD.

V. DISCUSSION

There are two reasons for the dramatic improvement in dynamic properties with PDF-based CGMD. First, using the Langevin equation of motion to compute the trajectory includes the frictional forces between CG beads and the random forces due to the bombardment of a CG bead by others. This is essential if one aims to obtain the correct dynamic properties from CGMD simulations. Second, using PDF-based CG pair forces instead of average effective pair CG forces allows for a more accurate model of the friction coefficient when coupled together with Eq. (17). In general, the friction term in the Langevin equation (Eq. (17)) is a tensor with components that include coupling (space) and memory (time) information.⁴⁸ It is difficult to determine these components and typically the

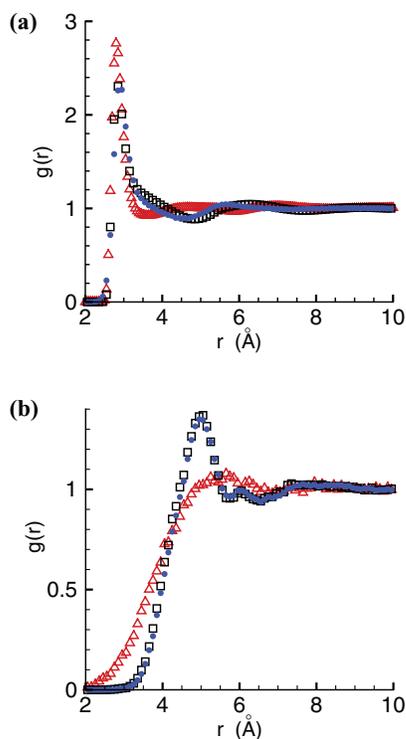


FIG. 6. The radial distribution functions for the system of glucose molecule in aqueous solution determined from all-atom MD (red triangles), conventional CGMD (black squares), and PDF-based CGMD (blue dots). (a) $g(r)$ calculated between CG water–CG water, (b) $g(r)$ calculated between CG glucose–CG water.

friction tensor is approximated by a diagonal matrix where all the diagonal elements are of the same value for the same type of particles, and non-diagonal elements are zero (Eq. (2)).²⁵ Thus, both coupling and memory information is excluded. In the PDF-based CGMD approach, the coupling information is retained because unique components for the friction tensor are derived from the random force contribution, which itself is directly obtained from different PDFs for different pairs of CG beads.

In this work, PDF-based CGMD is demonstrated by using the force matching method for coarse-graining. Generally, this PDF-based approach can be used with any coarse-graining method where the effective CG forces or potentials can be cast into the form of Eq. (11).⁶⁹ As an example, the output from the REACH⁷⁰ coarse-graining method takes the form of Eq. (11) written for pairwise force constants. Coarse-graining methods that are based on structure fitting^{9–12} can also conveniently report their potentials according to Eq. (11) and use PDF-based CGMD. Coarse-grained potentials, such as MARTINI,¹⁸ that are derived by fitting parameters to predetermined functions would not be direct candidates for the PDF-based CGMD method. However, this method can be applied in a hybrid coarse-graining scheme, such as the one introduced by Kim and Lamm⁷¹ where the MARTINI force field is combined with the force-matching method. Work on a straightforward approach that would allow derivation of the PDF for MARTINI and similar CG force fields is underway.

There are some aspects to consider before implementing PDF-based CGMD. As with any coarse-graining method, the

accuracy of the CGMD simulation depends on the smoothness of the effective CG force (or potential), which is achieved by sampling a large number of independent all-atom trajectories (more than 20 000 are used here). Moreover, for PDF-based CGMD additional computation time is required for the friction and random terms in Eq. (17). Especially for the friction term it is required to do multiplication of two $3N \times 3N$ matrices and $3N$ -dimensional vector $\mathbf{F}_0(\mathbf{q})\mathbf{F}_0(\mathbf{q})\mathbf{v}(t)$. This might take a significant computation time if directly calculated, since $\mathbf{F}_0(\mathbf{q})\mathbf{F}_0(\mathbf{q})$ requires $O(N^3)$ operations. However, it reduces to $O(N^2)$ by changing order of multiplications as $\mathbf{F}_0(\mathbf{q})\mathbf{F}^*$, where $\mathbf{F}^* = \mathbf{F}_0(\mathbf{q})\mathbf{v}(t)$. The cost of these computations can be reduced even further, since $3N$ vector \mathbf{F}^* can be calculated on fly during regular force calculation routine. For the systems considered here, the CPU time decreased by a factor of 8 for conventional CGMD and a factor of 1.8 for PDF-based CGMD. In present work, the HI is implemented for the CG systems with coarse-grained explicit water molecules. If the PDF-based CGMD would be applied to the CG systems with implicit solvent, a significant reduction in CPU time compared to all-atom MD would be observed. Another caveat is that this correction is applicable for pairs of CG beads that belong to molecules of similar mass (solvent-solvent pairs, solute-solute pairs) or when the CG bead i belongs to a molecule of much greater mass than CG bead j does (for example, CG bead i belongs to a solute molecule and CG bead j is a solvent molecule). Otherwise, if CG bead i belongs to a molecule with significantly smaller mass than the molecule in which CG bead j resides, the separation of time scales implied by the Langevin equation will be not satisfied.

VI. CONCLUSIONS

In summary, this work introduces an approach for systematically deriving the frictional and random fluctuation contributions to the forces of interactions between CG beads in a coarse-grained molecular system. It is shown that the random fluctuation forces arise naturally as an output of the coarse-graining procedure, and, furthermore, can be used with second fluctuation-dissipation theorem to compute the friction tensor components. This procedure for parameterizing the Langevin equation is in contrast to other examples^{24,25,28,42,43} where the frictional force is first determined and then used to derive the random fluctuation force. The PDF-based CG approach is tested for water and a dilute glucose solution, and the results show remarkable agreement with the dynamics obtained from reference all-atom simulations. The agreement between the structural properties such as bulk viscosity and the radial distribution functions is also preserved. Investigation about how this approach complements the relative entropy concept^{44–46} is a topic for future study.

ACKNOWLEDGMENTS

This research was sponsored by the (U.S.) Department of Energy (DOE) SciDAC program through the Office of Advanced Scientific Computing Research (ASCR) and Biological and Environmental Research (BER), and was performed at the Ames Laboratory, FWP AL-08-330-039. Ames

Laboratory is managed by Iowa State University for the US-DOE under Contract No. DE-AC02-07CH11358. The authors gratefully acknowledge G. A. Voth for providing the multi-scale coarse-graining (MS-CG) software, J. C. Smith for providing the parameters for the CHARMM36 potential used in the all-atom simulations, and M. S. Gordon, R. D. Vigil, and B. H. Shanks for discussion.

APPENDIX A: ESTIMATION OF SCALE SEPARATIONS

In this section, the method of evaluation of terms in brackets of Eq. (17) is discussed. This evaluation allows to select conditions when an approximation made in Eq. (12) is valid. The average magnitude of the first term in brackets of Eq. (10) was calculated as

$$\text{Term 1} = \langle \sqrt{\Delta t} \mathbf{N}(0, 1) \rangle = \sqrt{\Delta t} \sqrt{\frac{1}{N} \sum_i^N [N_i(0, 1)]^2}. \quad (\text{A1})$$

Second term in brackets of Eq. (10) is evaluated as

$$\begin{aligned} \text{Term 2} &= \left\langle \frac{\mathbf{B}^T \mathbf{v}(t)}{2k_B T_{\text{ref}}} \Delta t \right\rangle \\ &= \frac{\Delta t}{2k_B T_{\text{ref}}} \times \sqrt{\frac{1}{N} \sum_i^N \left(\sum_{j, j \neq i}^N F_{0,ij}^{A_i, A_j} \Delta t_0^{A_i, A_j} v_j \right)^2}. \end{aligned} \quad (\text{A2})$$

Because Eq. (12) is written based on assumption Term1 \gg Term2, it can be seen that this requirement can be satisfied by tuning the computational time step Δt . For two systems considered in present work, the computational time steps were chosen to $\Delta t = 0.5$ fs for the bulk water (the ratio Term1/Term2 = 6), and for a system of a single glucose molecule solvated with water molecules the computational time step is $\Delta t = 2.0$ fs, and the ratio Term1/Term2 = 24.

APPENDIX B: CHARACTERISTIC TIME Δt_0 CALCULATION

Before Eq. (17) can be used the characteristic time(s) Δt_0 needs to be determined first. In general, this time is different for different types of pairs of CG beads. For example, the characteristic time for a CG water δ CG water pair, $\Delta t_0^{W,W}$, is different from the characteristic time for a CG glucose CG water pair, $\Delta t_0^{G,W}$.

For both systems of bulk water and non-homogeneous system of single glucose molecule in environment of water molecules, the characteristic times are obtained in the iterative manner to match the corresponding diffusion coefficients obtained from CG MD and all-atom MD. The algorithm for this procedure is as following: select some initial values of the characteristic time(s) for the solute and solvent molecules represented in the system. In present work for the system of bulk water, only one characteristic time $\Delta t_0^{W,W}$ is presented and for the system of glucose molecule in the environment of water molecules two characteristic times exist: $\Delta t_0^{W,W}$ and $\Delta t_0^{G,W}$. Then, perform CG MD simulations with chosen char-

acteristic times for the time necessary to generate the MSD versus time plot (to extract MSD for water it is enough to perform simulation for up to 2 ns, while for the glucose molecule in water, the meaningful MSD data can be obtained after performing CG MD simulation for at least 20 ns). Next, compare diffusion coefficients obtained from CG MD and from all-atom MD and correct the characteristic time(s) if necessary. Thus, if the diffusion coefficient predicted with CG MD is larger than those obtained from all-atom MD, then the characteristic time should be increased (characteristic time is proportional to the friction tensor that has inverse proportionality to the diffusion coefficient), and vice versa. Typically, there are three to four iterations are necessary to obtain the characteristic times for the system. It was found that for the system of bulk water the characteristic time for water-water interactions is $\Delta t_0^{W,W} = 720$ fs. And for the system of single glucose molecule in the environment of water molecules the characteristic time for water-water interaction is $\Delta t_0^{W,W} = 130$ fs and the characteristic time for glucose-water interaction is $\Delta t_0^{G,W} = 12$ fs.

By determining these characteristic time scales the definition of all parameters for the PDF-based CGMD approach (Eq. (17)) is complete.

¹D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, San Diego, 2002), pp. 63–108.

²M. R. Walsh, C. A. Koh, E. D. Sloan, A. K. Sum, and D. T. Wu, *Science* **326**, 1095 (2009).

³D. P. Hurst, A. Grossfield, D. L. Lynch, S. Feller, T. D. Romo, K. Gawrisch, M. C. Pitman, and P. H. Reggio, *J. Biol. Chem.* **285**, 17954 (2010).

⁴T. D. Romo, A. Grossfield, and M. C. Pitman, *Biophys. J.* **98**, 76 (2010).

⁵D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wrighers, *Science* **330**, 346 (2010).

⁶M. L. Klein and W. Shinoda, *Science* **321**, 798 (2008).

⁷S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, *J. Phys. Chem. B* **111**, 7812 (2007).

⁸L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, *J. Chem. Theory Comput.* **4**, 819 (2008).

⁹A. P. Lyubartsev and A. Laaksonen, *Phys. Rev. E* **52**, 3730 (1995).

¹⁰F. Müller-Plathe, *ChemPhysChem* **3**, 754 (2002).

¹¹D. Reith, M. Pütz, and F. Müller-Plathe, *J. Comput. Chem.* **24**, 1624 (2003).

¹²T. Murtola, E. Falck, M. Patra, M. Karttunen, and I. Vattulainen, *J. Chem. Phys.* **121**, 9156 (2004).

¹³S. Izvekov and G. A. Voth, *J. Phys. Chem. B* **109**, 6573 (2005).

¹⁴S. Izvekov and G. A. Voth, *J. Chem. Phys.* **123**, 134105 (2005).

¹⁵W. 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).

¹⁶W. G. Noid, P. Liu, Y. Wang, J.-W. Chu, G. S. Ayton, S. Izvekov, H. C. Andersen, and G. A. Voth, *J. Chem. Phys.* **128**, 244115 (2008).

¹⁷G. A. Voth, *Coarse-graining of Condensed Phase and Biomolecular Systems* (CRC Press, Boca Raton, 2009), pp. 1–4.

¹⁸S. J. Marrink, A. H. de Vries, and A. E. Mark, *J. Phys. Chem. B* **108**, 750 (2004).

¹⁹J. Wong-Ekkabut, S. Baoukina, W. Triampo, I. M. Tang, D. P. Tieleman, and L. Monticelli, *Nat. Nanotechnol.* **3**, 363 (2008).

²⁰J. B. Accary and V. Teboul, *J. Chem. Phys.* **136**, 094502 (2012).

²¹V. A. Harmandaris and K. Kremer, *Soft Matter* **5**, 3920 (2009).

²²H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965).

²³R. Zwanzig, *J. Stat. Phys.* **9**, 215 (1973).

²⁴C. Hijon, P. Espanol, E. Vanden-Eijnden, and R. Delgado-Buscalioni, *Faraday Discuss.* **144**, 301 (2010).

²⁵S. Izvekov and G. A. Voth, *J. Chem. Phys.* **125**, 151101 (2006).

²⁶J. R. Silbermann, M. Schoen, and S. H. L. Klapp, *Phys. Rev. E* **78**, 011201 (2008).

²⁷X. Li, *Int. J. Numer. Methods Eng.* **83**, 986 (2010).

²⁸I. Lyubimov and M. G. Guenza, *Phys. Rev. E* **84**, 031801 (2011).

- ²⁹S. Matysiak and C. Clementi, *J. Chem. Phys.* **128**, 024503 (2008).
- ³⁰J. H. Weiner and R. E. Forman, *Phys. Rev. B* **10**, 315 (1974).
- ³¹D. L. Ermak and J. A. McCammon, *J. Chem. Phys.* **69**, 1352 (1978).
- ³²H. Yamakawa, *Modern Theory of Polymer Solutions* (Harper and Row, New York, 1971).
- ³³J. Rotne and S. Prager, *J. Chem. Phys.* **50**, 4831 (1969).
- ³⁴T. Frembsgen-Kesner and A. H. Elcock, *J. Chem. Theory Comput.* **5**, 242 (2009).
- ³⁵D. Petera and M. Muthukumar, *J. Chem. Phys.* **111**, 7614 (1999).
- ³⁶G. Chirico and J. Langowski, *Macromolecules* **25**, 769 (1992).
- ³⁷D. L. Ermak and H. Buckholz, *J. Comput. Phys.* **35**, 169 (1980).
- ³⁸R. Kubo, *Rep. Prog. Phys.* **29**, 255 (1966).
- ³⁹P. Langevin, *C. R. Acad. Sci. (Paris)* **146**, 530 (1908).
- ⁴⁰D. S. Lemons and A. Gythiel, *Am. J. Phys.* **65**, 1079 (1997).
- ⁴¹S. H. Northrup, J. O. Boles, and J. C. Reynolds, *Science* **241**, 67 (1988).
- ⁴²S. Markutsya, R. O. Fox, and S. Subramaniam, *Ind. Eng. Chem. Res.* **51**, 16116 (2012).
- ⁴³R. Moussavi-Baygi, Y. Jamali, R. Karimi, and M. R. K. Mofrad, *PLoS Comput. Biol.* **7**, e1002049 (2011).
- ⁴⁴M. S. Shell, *J. Chem. Phys.* **129**, 144108 (2008).
- ⁴⁵M. S. Shell, *J. Chem. Phys.* **137**, 084503 (2012).
- ⁴⁶P. Espanol and I. Zuniga, *Phys. Chem. Chem. Phys.* **13**, 10538 (2011).
- ⁴⁷I. Bilonis and N. Zabarar, *J. Chem. Phys.* **133**, 044313 (2013).
- ⁴⁸H. J. C. Berendsen, *Simulating the Physical World* (Cambridge University Press, Cambridge, England, 2007).
- ⁴⁹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 2nd ed. (Cambridge University Press, New York, 2002), pp. 664–666.
- ⁵⁰M. Fixman, *Macromolecules* **19**, 1204 (1986).
- ⁵¹T. Geyer and U. Winter, *J. Chem. Phys.* **130**, 1149051 (2009).
- ⁵²J. R. Taylor, *An Introduction to Error Analysis* (University Science Books, Mill Valley, CA, 1992), pp. 261–285.
- ⁵³W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, 2002), p. 99.
- ⁵⁴W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, *J. Chem. Phys.* **79**, 926 (1983).
- ⁵⁵H. Bekker *et al.*, in *Conference Proceedings of Physics Computing'92* (World Scientific Publishing Co., Singapore, 1993), pp. 252–261.
- ⁵⁶O. Guvench, E. R. Hatcher, R. M. Venable, R. W. Pastor, and A. D. MacKerell, Jr., *J. Chem. Theory Comput.* **5**, 2353 (2009).
- ⁵⁷G. Srinivas, X. Cheng, and J. C. Smith, *J. Chem. Theory Comput.* **7**, 2539 (2011).
- ⁵⁸T. Darden, D. York, and L. G. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- ⁵⁹S. Miyamoto and P. A. Kollman, *J. Comput. Chem.* **13**, 952 (1992).
- ⁶⁰B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, *J. Comput. Chem.* **18**, 1463 (1997).
- ⁶¹C. L. Lawson and R. L. Hanson, *Solving Least Squares Problems* (Prentice-Hall, Englewood Cliffs, 1974).
- ⁶²S. Markutsya, Y. Kholod, A. Devarajan, T. L. Windus, M. S. Gordon, and M. H. Lamm, *Theor. Chem. Acc.* **131**, 1162 (2012).
- ⁶³S. J. Plimpton *et al.*, *LAMMPS Molecular Dynamics Simulator*, 2012, see <http://lammps.sandia.gov/index.html>.
- ⁶⁴S. J. Plimpton, *J. Comput. Phys.* **117**, 1 (1995).
- ⁶⁵A. Einstein, *Ann. Phys.* **17**, 549 (1905).
- ⁶⁶P. Mark and L. Nilsson, *J. Phys. Chem. A* **105**, 9954 (2001).
- ⁶⁷A. Das and H. C. Andersen, *J. Chem. Phys.* **132**, 164106 (2010).
- ⁶⁸M. Praprotnik, L. D. Site, and K. Kremer, *Annu. Rev. Phys. Chem.* **59**, 545 (2008).
- ⁶⁹For coarse-graining methods that fit potentials, forces for Eq. (11) can be obtained by taking the derivative with respect to distance of separation.
- ⁷⁰K. Moritsugu and J. C. Smith, *Biophys. J.* **93**, 3460 (2007).
- ⁷¹S. H. Kim and M. H. Lamm, *Phys. Rev. E* **84**, 025701(R) (2011).