Modulating Coarse-Grained Dynamics by Perturbing Free-Energy-Landscapes

Ishan Nadkarni[1]†, Jinu Jeong[2]†, Bugra Yalcin[3] and Narayana. R. Aluru[1,3]*

1. Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin 78712, Texas, United States
2. Department of Mechanical Science and Engineering, University of Illinois at Urbana−Champaign, Urbana, Illinois, 61801, United States
3. Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin 78712, Texas, United States

†These authors contributed equally to this work.

* Correspondence to aluru@utexas.edu (N.A.)

Abstract: We introduce an approach to describe long-time dynamics of multiatomic molecules by modulating the Free-Energy-Landscape (FEL) to capture dominant features of the energy-barrier crossing dynamics of the All-Atom (AA) system. Notably, we establish that the self-diffusion coefficient of Coarse-Grained (CG) systems can be accurately delineated by enhancing conservative force fields with high-frequency perturbations. Using theoretical arguments, we show that these perturbations do not alter the lower-order distribution functions, thereby preserving the structure of the AA system after coarse-graining. We demonstrate the utility of this approach using molecular dynamics simulations of simple molecules in bulk with distinct dynamical characteristics with and without timescale separations as well as for inhomogeneous systems where a fluid is confined in a slit-like nanochannel. Additionally, we also apply our approach to more powerful many-body potentials optimized using Machine Learning (ML).

1. Introduction

Molecular Dynamics (MD) simulations have played a pivotal role as a "computational microscope[1]" in probing, understanding and engineering of atomic systems including biomolecules[2–6], bio-complexes[7–10], 2D materials[11–13] and other soft-matter systems[14–16] of utmost practical importance as well as elucidating important biophysical processes and new phenomenon inaccessible to experimental measurements. With unprecedented developments in hardware and the availability of unprecedented computational resources we find ourselves at the dawn of exascale computing where systems with millions[17,18] and billions[19–21] of atoms and also timescales spanning micro to milliseconds[22,23] have been made possible with distributed computing on specialized hardware. With the desire to breach even larger length and timescales in simulations, more efficient approaches to molecular modelling, probing only the essential and important aspects of a system have been investigated. One such pivotal technique that has emerged is coarse-graining[24], offering a powerful approach to modeling complex systems and studying their behavior across extended spatiotemporal scales. By reducing the degrees of freedom of the fine-grained atomistic system one seeks to model larger and more complex systems efficiently at reduced computational cost while also preserving the essential/important characteristics of the original system. Some noteworthy examples of largescale systems include coarse-grained modelling of SARS-CoV-2 virion[25], the silica polymerization reaction of million atoms over a microsecond timescale[26] and modelling transport in a nuclear pore involving around 200 million atoms[27].

Various top-down[24,28,29] and bottom-up[30–33] coarse-graining methodologies have been devised, considering the specific system and the targeted structural, thermodynamic or dynamic properties under investigation. Thermodynamic quantities can typically be represented as averages in the canonical ensemble and therefore can be derived as conditional averages of the

AA system like the Many-Body Potential of Mean Force (MB-PMF). The CG force-fields can then be effectively parameterized to approximate the MB-PMF[30–37].

While such a parameterization can successfully replicate certain experimental observables and thermodynamical quantities to some degree, the CG dynamics gets significantly altered and is seen to be much faster than the fine-grained atomistic system in most cases. This effect has been observed for a wide class of systems[38–42] and is typically attributed to reduced friction regulated by the "projected dynamics" of the removed degrees of freedom, as delineated in the framework by Mori and Zwanzig (MZ)[43–45]. Moreover, the conservative high-dimensional MB-PMF not being faithfully represented by an adequate basis set also leads to an incorrect description of the different dynamical timescales due to changes in relative energy-barriers of the Free-Energy-Landscape as discussed in the recent review[46]. These two sources of errors are non-trivially related and need to be considered in constructing a dynamically consistent model. An interesting theoretical connection between the two has quite recently been discussed in ref.[47,48].

Various methods have been developed to realistically model the time evolution of coarse-grained degrees of freedom, aiming to accurately describe the kinetics of fine-grained systems[47]. The most popular of these methods is the Generalized Langevin Equation (GLE), which can be rigorously derived from the Mori-Zwanzig (MZ) formalism. This formalism provides an exact model of the time evolution of coarse-grained degrees of freedom using appropriate projection operators. However, parameterizing the GLE can be very challenging due to the presence of non-conservative friction and random forces that depend on the projected trajectory of the removed degrees of freedom, which are not easily simulated[49]. To address this, various approximations have been proposed to make the problem more tractable, leading to the widespread use of the GLE in studying several coarse-grained systems[50–56]. Recent advancements[57–59] have made considerable progress in accurately obtaining memory kernels

in this context. Apart from GLE several other methods like Time Rescaling Approaches[60–63], Probabilistic Methods employing Bayesian Theory[64], Variational Inference[65–67], and Markov-State-Models[68–70], as well as Structure-Thermodynamic-Kinetic[71] relations have also been employed to model specific transport properties for a wide class of systems. Although some of these methods are heuristic in nature, they have proven to be quite useful in circumventing difficulties associated with GLE parameterization for complex systems and have provided a more feasible route to approach the problem of modelling the dynamics of CG systems.

An intriguing yet underexplored approach involves modulating the Free-Energy-Landscape of the coarse-grained system. This method aims to precisely capture the dominant barrier-crossing dynamics observed in the fine-grained model, facilitating a realistic depiction of pertinent transport properties of interest. This perspective is also supported by the Reaction Rate Theory[72], which relates the free energy barrier to the residence time, and hence, to dynamics of the system. The landscape paradigm has also been successfully used to study important phenomena like glass transition[73–78] and protein-folding[79,80].

A mathematically rigorous and particularly intriguing result was presented by Zwanzig[81] to understand protein dynamics, where the potential energy surfaces (PES) are found to be rough. Zwanzig proposed a theoretical model of an atom diffusing along a one-dimensional surface, consisting of a superposition of a slowly varying smooth part and a high-frequency rough part. The findings revealed that the addition of roughness to the originally smooth PES caused a significant slowdown in the liquid's diffusion compared to that over the smooth potential. Zwanzig's depiction of the impact of rough energy landscape features on the diffusive dynamics of the system holds computational significance in the context of coarse graining. It suggests the possibility of modulating diffusive dynamics by introducing high frequency landscape features, thereby rendering coarse-grained dynamics more akin to their AA counterparts.

We investigate this idea and demonstrate that the diffusive movement of atoms can be effectively regulated by introducing high-frequency perturbation to the conservative force field as proposed by Zwanzig, allowing us to accurately replicate self-diffusion coefficients consistent with those observed in fine-grained systems. By enriching conservative single site CG potentials, initially derived for matching low-order distributions such as the RDF, with high-frequency potential roughness features, we restore important energy landscape features that are otherwise smoothed out through coarse-graining. We showcase the methodology for bulk systems with different multiatomic molecules with and without timescale separations. Furthermore, we extend the approach to confined systems, such as fluids in slit-like nanochannels, where diffusion coefficients exhibit variations along different directions. We also demonstrate the utility of this method with flexible ML potentials which capture many-body effects, particularly applying this approach to two widely used ML potentials: 1) Deep Neural Network based "DeepCG" potential[82] and 2) Graph Neural Network[82–85] (GNN) (see Appendix B for the supporting information).

The remaining sections of the paper are organized as follows: In Section 2.1, we introduce the concept of the Free-Energy Landscape to examine how the characteristics of the energy landscape influence the diffusive dynamics of an atomic system, and also provide a quantitative perspective based on Zwanzig's model for diffusion in rough potentials. In Section 2.2, we present our methodology, which involves perturbing 1) pairwise potentials derived through (RE) relative entropy minimization and 2) many-body potentials obtained using ML methods. In Section 3 we present and discuss results for both bulk and confined systems featuring different multiatomic molecules.

2. Methods

2.1. Free - Energy - Landscape Picture of CG Dynamics.

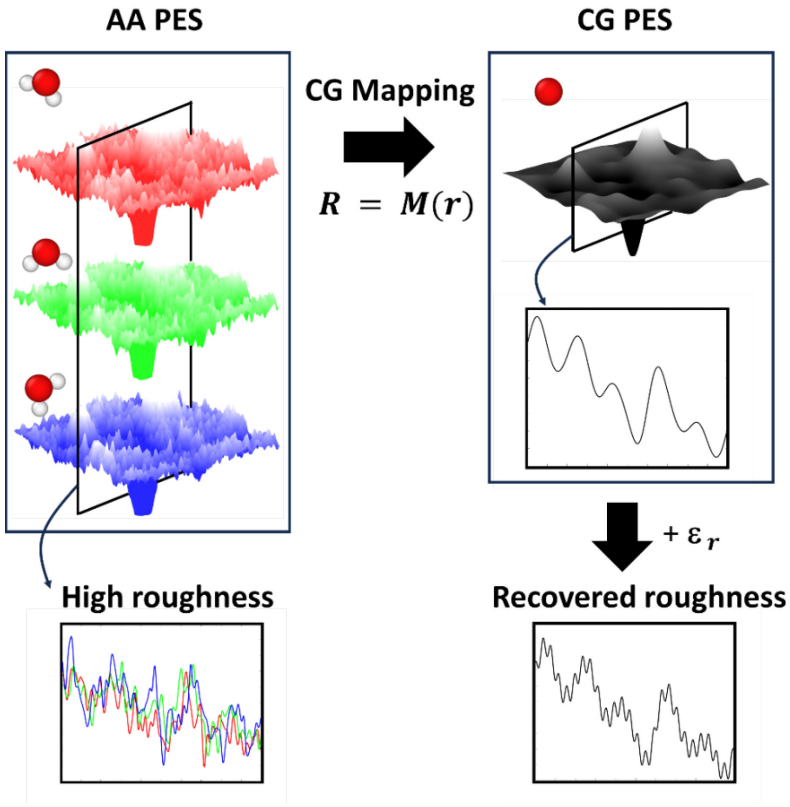

Figure 1: Visualization of an idealized schematic of a rugged All-Atom landscape (top-left) and a smoothened coarse-grained landscape (top-right). The sliced Potential Energy Surface at the bottom provides a qualitative representation of high-frequency "rough" features, which can be emulated by introducing perturbation to the CG potential.

Consider a schematic view of the All-Atom and smoothed CG energy landscape as shown in Figure 1. The dynamics of a single atom can be thought of as a point traversing along the 3N-dimensional Potential Energy Surface (PES) with a 3N-dimensional temperature-dependent velocity, where N is the total number of atoms in the atomistic system. A “slice” of such a landscape along the atom’s trajectory would reveal different topological features that characterize dynamical processes associated with multiple timescales of the AA system.

However, coarse-graining, which reduces the system's degrees of freedom, results in a smoothed energy landscape, as illustrated in the sliced CG landscape in Figure 1. This perspective has also been widely used for studying protein folding[86,87] and discovering dominant pathways[88] that control the kinetics of the folding process. Another interesting application based on this idea has been in describing the diffusive behavior near the glass transition temperature[73-76].

While such a qualitative picture of the influence of potential landscape on dynamics is intriguing, it fails to quantitatively describe the dependence of the rate of diffusion of atoms on the roughness of the landscape. Zwanzig developed a simplified model to determine such a relation by considering the motion of a single particle transversing a 1-D potential comprising of a smooth potential $U_0(x)$ on top of which a high frequency rough part $U_1(x)$ is added so that the net potential is $U(x) = U_0(x) + U_1(x)$. Using the expression for the Mean First-Passage Time, it was shown that the self-diffusion coefficient can be represented as,

$$D_{eff} = \frac{D_0}{\langle e^{\beta U_1}\rangle\langle e^{-\beta U_1}\rangle} \tag{1}$$

Here D_0, D_{eff} denote the diffusion coefficients under the potential $U_0(x)$ and $U(x)$ respectively. $\langle ... \rangle$ denotes the local averaging operation along a characteristic length of roughness along the x-axis. For a perturbation of the form $U_1(x) = A * \cos(\omega x)$, the expression for the averages can be analytically computed in terms of the modified Bessel function $I_0(A/k_{\mathrm{B}}T)$ so that,

$$\left\langle e^{\beta U_1}\right\rangle = \left\langle e^{-\beta U_1}\right\rangle = I_0\left(\frac{A}{k_{\mathrm{B}}T}\right) \tag{2}$$

At lower temperature, where $A/k_{\mathrm{B}}T$ is large, the Bessel function can be approximated as an exponential. Consequently, the effective coefficient can be re-expressed as,

$$D_{eff} = D_0 * \exp\left[-\frac{2A}{k_B T}\right] \quad (3)$$

This elegant formulation highlights the direct influence of the roughness on the diffusion coefficient. This expression although simple, has certain limitations and many efforts have been made to address these[89–92]. Nevertheless, this concept remains highly useful from a coarse-graining standpoint, because it suggests that by introducing sinusoidal perturbations in a potential landscape and altering the barrier depths of these perturbations, one could potentially control the diffusive dynamics of a system. Moreover, in Appendix A.1 of the supporting information we provide theoretical arguments along with numerical analysis to show that such a high-frequency perturbation does not change the low order structure of the liquid like the Radial Distribution Function (RDF).

2.2. Perturbing Conservative Potentials

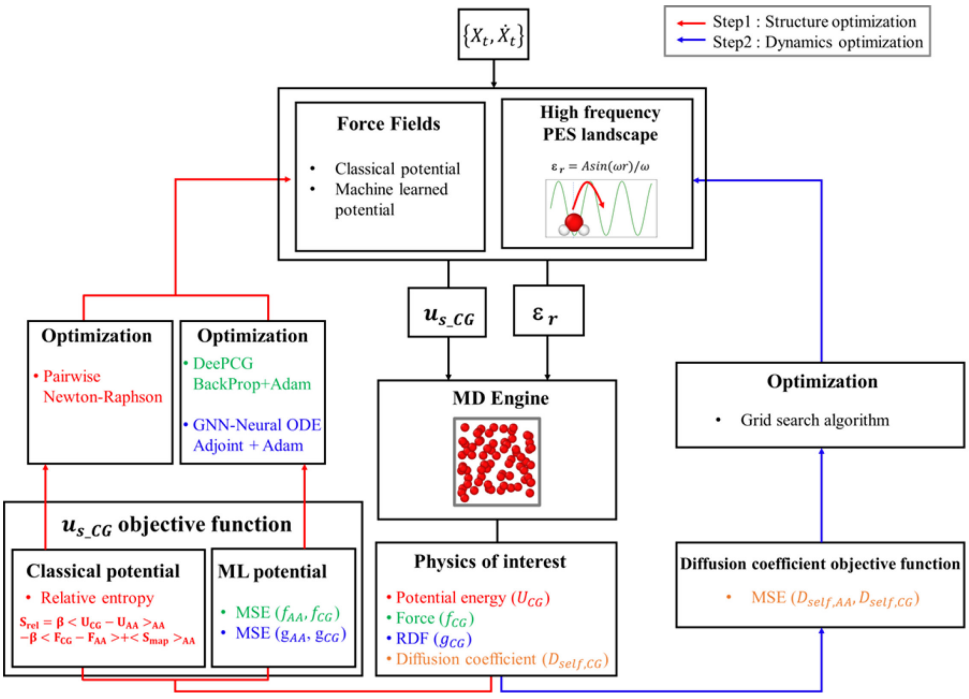

Figure 2: Workflow for parameterizing a coarse-grained potential to match the structure and reproduce long-time diffusion of the full atomistic model. We sequentially perform the optimization to first derive the interatomic CG potential u_{s_CG} (red arrow) and then optimize

the perturbation $\varepsilon_r(A,\omega)$ which is added (blue arrow) to u_{s_CG} to match the diffusion coefficient of the AA system. We start by defining the initial phase space configuration and run an MD engine with either a classical-pairwise or a neural network potential. In the structure optimization procedure, we compute physical quantities of interest required to compute gradients of each loss-functions: S_{rel}, $MSE\ (f_{AA}, f_{CG})$ and, $MSE\ (g_{AA}, g_{CG})$, based on the methods used to derive u_{s_CG}. Once the gradients are computed, the potential parameters are updated using different optimization procedures depending on the coarse-graining method and the type of interatomic potential used. The converged potential u_{s_CG} reproduces structure, however, it shows accelerated dynamics. To correct dynamics, we perturb u_{s_CG} with $\varepsilon_r(A,\omega)$, and carry out a grid search over parameters A and ω to find the optimal parameters which reproduce the self-diffusion coefficient of the corresponding AA system.

In this section, we present our approach to emulate the long-time dynamical behavior along with the distribution function of the AA system as shown in Figure 2. Leveraging the theoretical discussions outlined in Appendix A.1 of the supporting information, we establish that the structure of the coarse-grained system remains unaltered, when subjected to high-frequency perturbations in the interatomic potential. We introduce sinusoidal perturbations to the conservative potential as follows:

$$u_{pert\_CG} = u_{s\_CG} + \varepsilon_r(A,\omega) \tag{4}$$

where the perturbation can be chosen to be of the form,

$$\varepsilon_r(A,\omega) = A\ sin(\omega r)/\omega \tag{5}$$

u_{s_CG} is the interaction potential obtained using methods such as RE minimization, that result in a CG system with RDF matching that of the AA system. All the molecules are coarse-grained into single beads. The unperturbed CG system is referred to as s-CG throughout the paper. ε_r

is the perturbation which is a function of interatomic distance r and parameters A and ω. These parameters are optimized to give the right dynamics without varying the structure. Initially, we derive the structure-matching (s-CG) pairwise potentials using the RE minimization framework, and subsequently, we expand upon this approach to include neural network potentials. Our rationale for exploring various interaction potentials, primarily pairwise and many-body potentials, is twofold. Firstly, the utilization of simple pairwise potentials allow a more fundamental analysis of the impact of perturbations on the s-CG system. This analysis would otherwise be challenging with neural network-based potentials due to their high-dimensional representations. Secondly, by employing machine learning (ML)-based potentials, we demonstrate the application of this methodology with force fields acquired through diverse optimization procedures. Moreover, these potentials are well-suited for modeling different systems targeting various static properties, such as angular distributions, which cannot be easily replicated using a simple pairwise basis.

2.3.a. Perturbations to Relative Entropy Optimized Pairwise Potentials

We employ the perturbation strategy to the s-CG interatomic potentials using the RE minimization approach[35–37,93–95]. For a canonical ensemble it can be defined as follows,

$$S_{\mathrm{rel}} = \beta\langle U_{\mathrm{CG}} - U_{\mathrm{AA}}\rangle_{\mathrm{AA}} - \beta(A_{\mathrm{CG}} - A_{\mathrm{AA}}) + \left\langle S_{\mathrm{map}} \right\rangle_{\mathrm{AA}} \tag{6}$$

Where $\beta = 1/k_bT$ and k_b is the Boltzmann constant. $U_{\mathrm{CG/AA}}$ and $A_{\mathrm{CG/AA}}$ are the potential energies and configurational part of Helmholtz free energy respectively, and $\left\langle S_{\mathrm{map}} \right\rangle_{\mathrm{AA}}$ is the entropy of mapping. Numerically minimizing S_{rel} maximizes the overlap between the CG and AA ensemble, and thus yields CG potentials that accurately reproduce the structure of the fine-grained atomistic system. By introducing high-frequency perturbations to the RE minimized potentials, we show that the faster diffusion of the s-CG model can be corrected, and the long-time diffusion behavior of the AA system can be captured without disturbing the structure

obtained by the RE minimization procedure. To understand the effect of perturbations defined in Equation (5) on the diffusion behavior, we carry out an extensive grid search covering different values of A and ω. By varying the two perturbation parameters, we tune the degree of "roughness" that is added to the s-CG landscape, where the optimum values better capture the FEL characteristics that get smoothed out during coarse-graining. The grid parameters are chosen from the following table:

Table 1. Range of values of Amplitude and Frequency parameters.

	$A_{min} - A_{max}$ (kcal/mol/Å)	$\omega_{min} - \omega_{max}$ ($Å^{-1}$)	**ΔA (kcal/mol/Å)**	**$\Delta\omega$ ($Å^{-1}$)**
H_2O	0 - 8	0 - 500	0.2	10
CO_2	0 - 10	0 - 400	0.2	10
CH_4	0 - 2	0 - 500	0.1	10

We apply our methodology to two distinct types of systems: liquids in bulk and confinement. For bulk systems, we consider a cube of side 4 nm with periodic boundary condition and examine three multiatomic molecules at 300K: SPC/F[96] water (H_2O), carbon dioxide (CO_2), and methane (CH_4). For confinement, we focus on SPC/F water confined within graphene walls of size 2.5 nm by 2.5 nm and separated by 8 nm width. Force fields for bonded (bonds and angles) and non-bonded interactions are obtained from the GROMOS force-field available in the ATB Repository[97]. The MD simulations were performed in the LAMMPS package in the NVT ensemble with a timestep of 1 fs. The temperature was maintained using the Nosé–Hoover thermostat[98]. A cutoff of 1.2 nm was used for short range van der Waals and electrostatic interactions. The long-range electrostatic interactions were modelled using the particle mesh Ewald algorithm with an accuracy value of 10^{-4}. These molecules are selected

based on their distinct dynamical characteristics. Previous studies have shown[99,100] that water lacks timescale separation, whereas there is a clear separation between the evolution of coarse-grained degrees of freedom and removed degrees of freedom for CO_2. Specifically, in the case of H_2O, the Velocity Autocorrelation Function (VACF) and Random-Force Autocorrelation Function (RACF) decay at a similar rate, whereas for CO_2, the VACF decays much faster than the RACF. Systems, with and without timescale separations have traditionally been presented in studies assessing the validity of various approximations to the GLE. For instance, the Markovian assumption invoked to simplify the terms in GLE is known to work well for systems with timescale separation and give discrepancies otherwise[101,102]. Therefore, we demonstrate our approach towards SPC/F water, where a clear timescale separation doesn't exist and show that the long-term dynamical properties can be reproduced accurately.

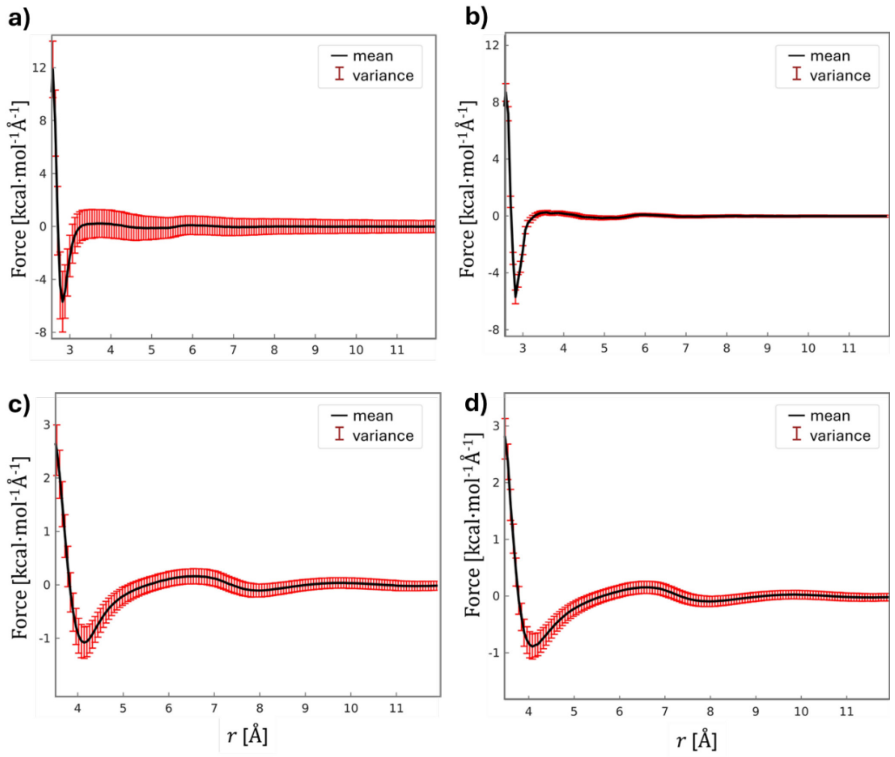

Figure 3: Stochastic force distribution on the Center of Mass for AA (a, c) and s-CG (b, d) system for H_2O (a, b) and CH_4 (c, d). The black line is the mean-force, and the red error bars indicate the force variance at a radial distance.

The qualitative evidence of the difference in energy landscape between the AA and s-CG landscapes can also be seen by observing the distribution of the force acting on the Center of Mass of the AA molecules and the single bead CG atoms as shown in Figure 3a-d. A similar force distribution-based analysis has been previously done by ref. [103] to add hydrodynamic interactions to the CG model. It can clearly be observed that for water there is a significant reduction in the stochasticity of the force in the CG system compared to the AA system after the coarse-graining procedure. The AA model for SPC/F water shows significant force fluctuations due to a more rugged FEL along which the water molecules travel. In the case of CH_4, the difference in the distribution of stochastic force is much less compared to SPC/F water. We attribute this to the geometry of the CH_4 molecule being spherically symmetric, and this symmetry is retained even after the coarse-graining process. Similar observations have been recently made in ref.[104], where the influence of "geometric roughness" has been studied on the diffusion behavior of the fine-grained system. Since the AA and s-CG landscape show similar characteristic features for CH_4, we show that the amount of perturbation needed to correct for the dynamics in s-CG system is relatively much smaller than for CO_2 and H_2O.

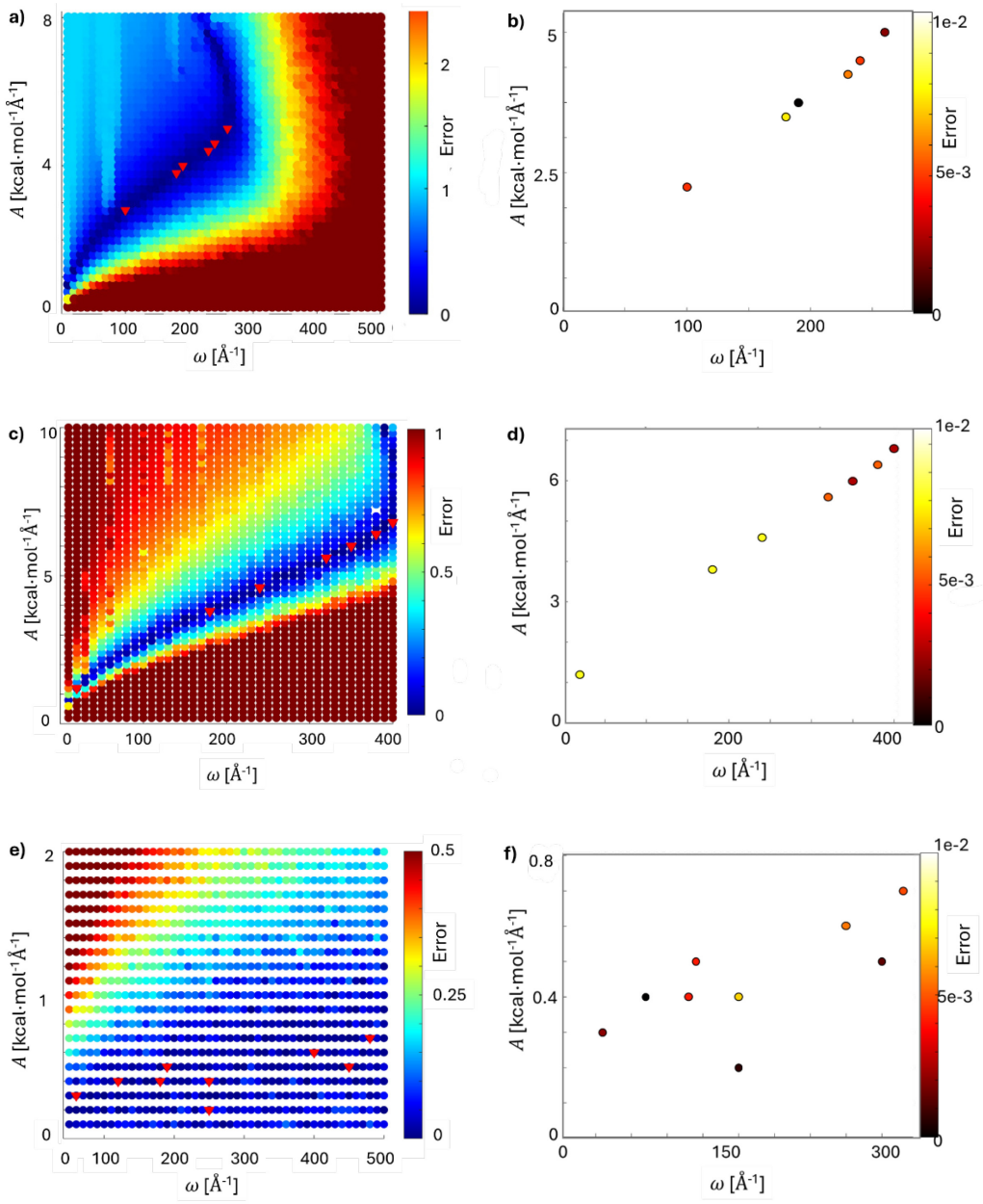

Figure 4: Diffusion coefficient loss manifolds (left) and optimum parameters (right) for H_2O (a, b), CO_2 (c, d), and CH_4 (e, f), that give the smallest error in diffusion coefficient.

Figure 4a-f illustrates the diffusion coefficient error manifold and optimal parameters ($\% \, err = \frac{D_{CG} - D_{AA}}{D_{AA}}$) for the three multiatomic molecules. Here D_{AA} and D_{CG} are the self-diffusion coefficients computed for the AA and the perturbed CG system using the Einstein relation. The optimum parameters plotted on the right in Figure 4 give errors of less than 1% for H_2O, CO_2, and 0.1% for CH_4. The loss manifolds describe the combined effect of the

amplitude and frequency of perturbation on the long-time diffusion of the resulting perturbed CG model. Moreover, the error manifolds for the three molecules are seen to be very different, indicating differences in their dominant FEL barriers that influence the long-time diffusion.

Next, we implement the methodology for SPC/F water confined in a slit-like graphene nanochannel of 8 nm width. The water molecules inside the nanochannel interact with both the neighboring water molecules, as well as the walls. Close to the wall, the effect of wall-fluid interactions is dominant compared to the fluid-fluid interactions and the value of the first peak in density is almost entirely determined by the wall-fluid potential[105]. As a molecule moves away from the wall the fluid-fluid forces start dominating, and the structure starts becoming more homogeneous. Such confinement effects cause interesting features not only in the structure but also in the transport properties[106]. Particularly, the diffusion properties along the longitudinal and transverse directions happen to be significantly different, where the self-diffusion coefficient along the unconstrained direction is seen to be higher than the constrained direction[107]. Unlike bulk, the forces on the water molecule can be decomposed into 2 parts, wall-fluid and fluid-fluid. Figure 5a-b shows the total and wall-fluid force distribution along the width of the nanochannel in the AA confined water system. While the total force shows considerable force variation, the wall-fluid force component shows a very small variance along the width of the channel. This is due to the specific single site interaction potential of SPC/F water with the carbon atoms of the wall. Therefore, for this system the high-frequency perturbations are added only to the fluid-fluid component of the s-CG interaction potential. The wall-fluid component, which is important for obtaining the correct density in the channel, remains unperturbed. Let $u_{s\_CG} = u^{WF}{}_{s\_CG} + u^{FF}{}_{s\_CG}$ is the total interaction potential and $u^{WF}{}_{s\_CG}$, $u^{FF}{}_{s\_CG}$ the corresponding wall-fluid and fluid-fluid components that are simultaneously optimized in the RE minimization framework as done in Ref [95]. We then perturb the fluid-fluid potential without changing the wall-fluid potential so that $u^{FF}{}_{pert\_CG} =$

$u^{FF}{}_{s\_CG} + \varepsilon^{FF}$, where ε^{FF} is the perturbation made to the RE optimized fluid-fluid interaction potential.

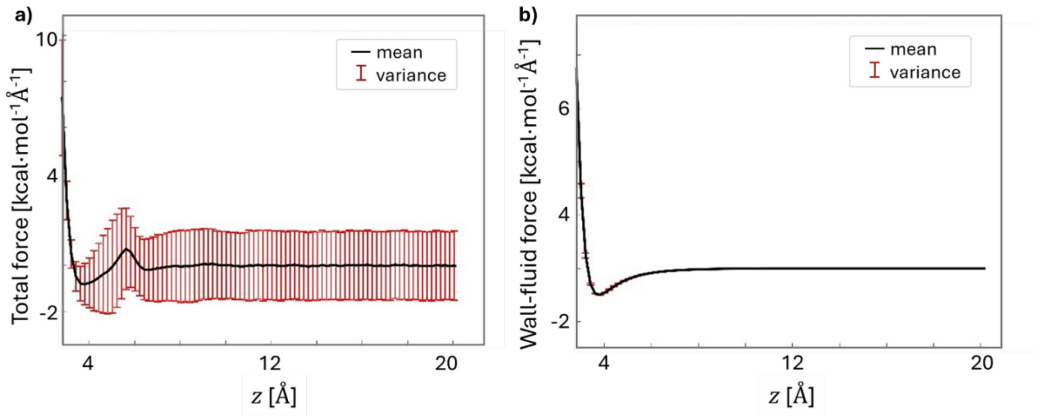

Figure 5: a) Total force and b) wall-fluid component of total force distribution for All-Atom water confined in a slit-like nanochannel. The black line is the mean-force, and the red error bars indicate the force variance at the location.

2.3.b. Perturbations to Many-Body Neural Network Potentials

A) DeepCG derived potentials

We apply the perturbation strategy to DeePCG[82] which is a popular neural-network based many-body potential that can be trained on fine-grained atomistic models in a symmetry-preserving manner. In the DeePCG framework, an atomic configuration is transformed into a descriptor, D_i, which is expressed as

$$D_i = \{D_{i1}, \dots, D_{ij}, \dots D_{i|Neigh|}\}, j \in Neigh \quad (7)$$

Here, D_{ij} conveys local atomic information of atom j observed from atom i: $D_{ij} = \{1/R_{ij}, x_{ij}/R_{ij}^2, y_{ij}/R_{ij}^2, z_{ij}/R_{ij}^2\}$. $Neigh$ indicates the set of neighbors of atom i, within the cutoff distance, r_{cutoff}. D_i is then fed into four hidden layers that are fully connected layers with the hyperbolic tangent activation function. The final layer is a fully connected layer with linear activation. This feed forward process can be written as

$$U_i = W_5\psi\left(W_4\psi\left(W_3\psi\left(W_2\psi\left(W_1\psi(D_i)\right)\right)\right)\right) \tag{8}$$

Here, W_1, W_2, W_3, W_4, and W_5 are matrices having the size of $|D| \times 128, 128 \times 64$, 64×32, 32×16, and 16×1, where $|D|$ refers to the size of the vectorized descriptor. The DeePCG-predicted force, $f_{CG}(R_i)$, is computed by taking the derivative of neural network output with respect to position.

$$f_{CG}(R_i) = -\nabla U_{CG}(R_i) \tag{9}$$

The loss function of DeepCG is mean-squared error between predicted CG force and AA mean force, which can be written as

$$L_{DPCG} = \frac{1}{|\mathcal{B}|}\sum_{i}^{\mathcal{B}}\left(f_{CG}(R_i) - f(R_i)\right)^2 \tag{10}$$

where $\mathcal{B}$, R_i, $f_{CG}(R_i)$, and $f(R_i)$ indicate minibatch, atomic configuration in CG space, DeePCG-predicted force, and net force acting on a water molecule for a given atomic configuration (R_i). We use a minibatch size of four and ADAM[108] optimizer with the initial learning rate being $1e^{-3}$ and exponentially reducing down until $1e^{-6}$.

After training the baseline DeePCG model, we conduct a comprehensive grid search, as outlined in Section 2.3.a. As shown in Figure 6, the loss manifold for the DeepCG potential happens to be qualitatively similar to the loss manifold of RE minimized pairwise potentials as shown in Figure 4. However, we observe subtle differences in the spectrum of frequencies which result in accurate diffusion behavior for both potentials. For example, the number of frequencies corresponding to the same error limit of less than 1% happens to be more for the DeepCG potential. Such differences, we believe, are indicative of the changes in relative FEL barriers caused by using different basis representations to the actual MB-PMF. Apart from the DeepCG potential, we also train a GNN potential to match the RDF directly using Adjoint-

state method[83]. The details of constructing the GNN potential are presented in Appendix B of the supporting information.

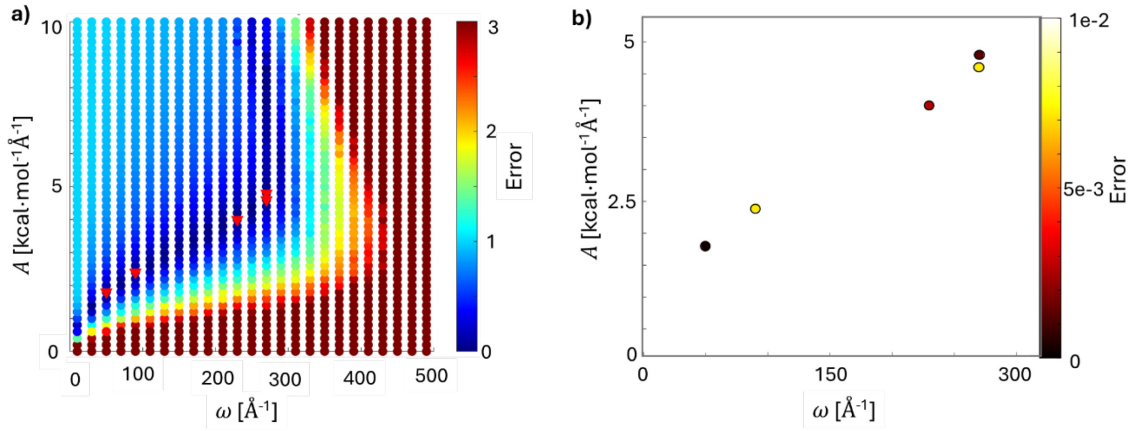

Figure 6: Diffusion coefficient loss manifold (left) and optimum parameters (right) corresponding to DeepCG potential for water.

3. Results

3.1. Structure and long-time dynamics of perturbed potentials.

We compare the structure and dynamics of the perturbed pairwise potentials optimized using the methodology outlined in Section 2.3.a. Pairwise potentials optimized using RE accurately reproduce the RDF for H_2O, CO_2, and CH_4, and further perturbing the s-CG potentials, we reproduce the long-time diffusion coefficients of these CG systems. The parameters $\omega = [100, 180, 120]$ $\mathring{A}^{-1}$ and $A = [2.8, 3.8, 0.4]$ $kcal{\cdot}mol^{-1}\mathring{A}^{-1}$ are used to perturb H_2O, CO_2, and CH_4 respectively. The lowest frequencies which give accurate structure and dynamics are chosen since lower timestep would be needed if the frequency of perturbation is too high. Figure 7a-b shows the comparison between the RDF and MSD of the AA and CG systems. The high-frequency perturbations added to the s-CG potentials hinder CG diffusion and bring it closer to the AA diffusion. For both CH_4 and CO_2, we see that the s-CG system is significantly accelerated, while in the case of CH_4 the speed-up is less. As explained in Section 2.3.a, this is

a result of the differences in the force stochasticity/distribution that a molecule experiences as it traverses the PES. Consequently, the perturbation needed to correct the dynamics for CH_4 is significantly smaller than that for H_2O and CO_2.

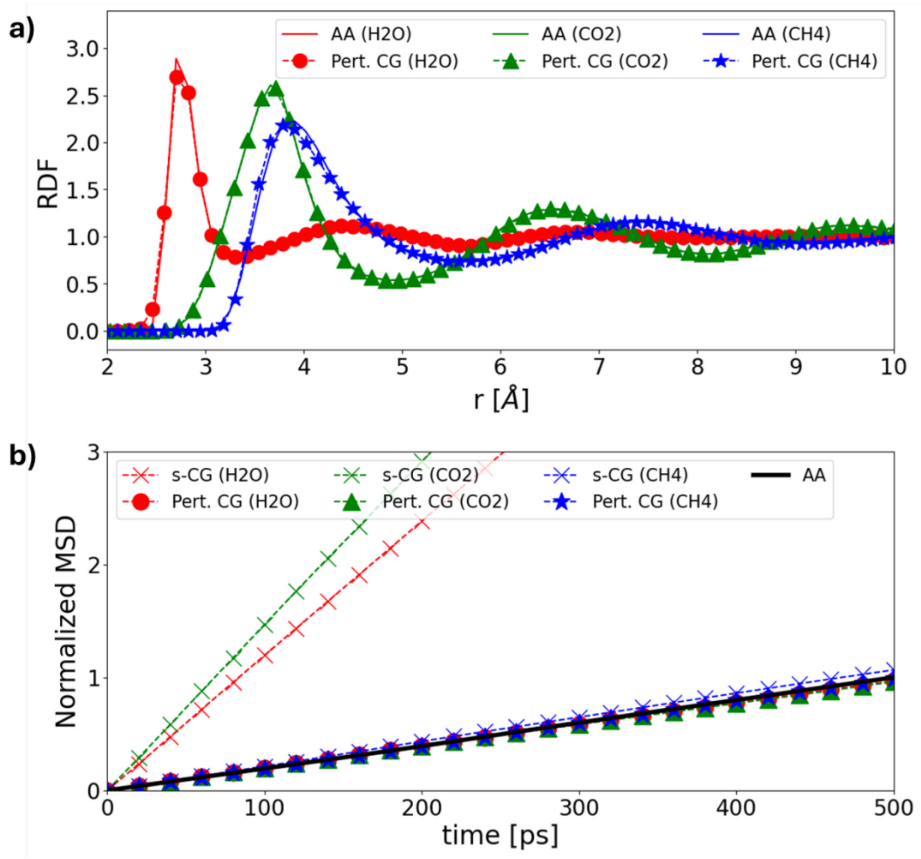

Figure 7: a) Radial Distribution Function and b) Normalized MSD for AA and CG systems for H_2O, CO_2 and CH_4.

Next, we demonstrate the methodology for inhomogeneous systems, where the RE minimization approach has previously been used to derive potentials for confined water[95]. We followed the same approach and obtained CG force fields to reproduce the density profile in the confined channel. It is well known that the self-diffusion of water in confinement is anisotropic and the self-diffusion coefficients are different for the transverse and longitudinal direction. Figure 8 compares the density and MSD of the perturbed CG and the AA system along the transverse and longitudinal directions obtained by perturbing the RE minimized

potentials with parameters, $\omega = 200\ \mathring{A}^{-1}$ and $A = 4.9\ kcal{\cdot}mol^{-1}\mathring{A}^{-1}$. By perturbing the force field as mentioned in Section 2.3.a, the dynamics is seen to be closer to the AA system.

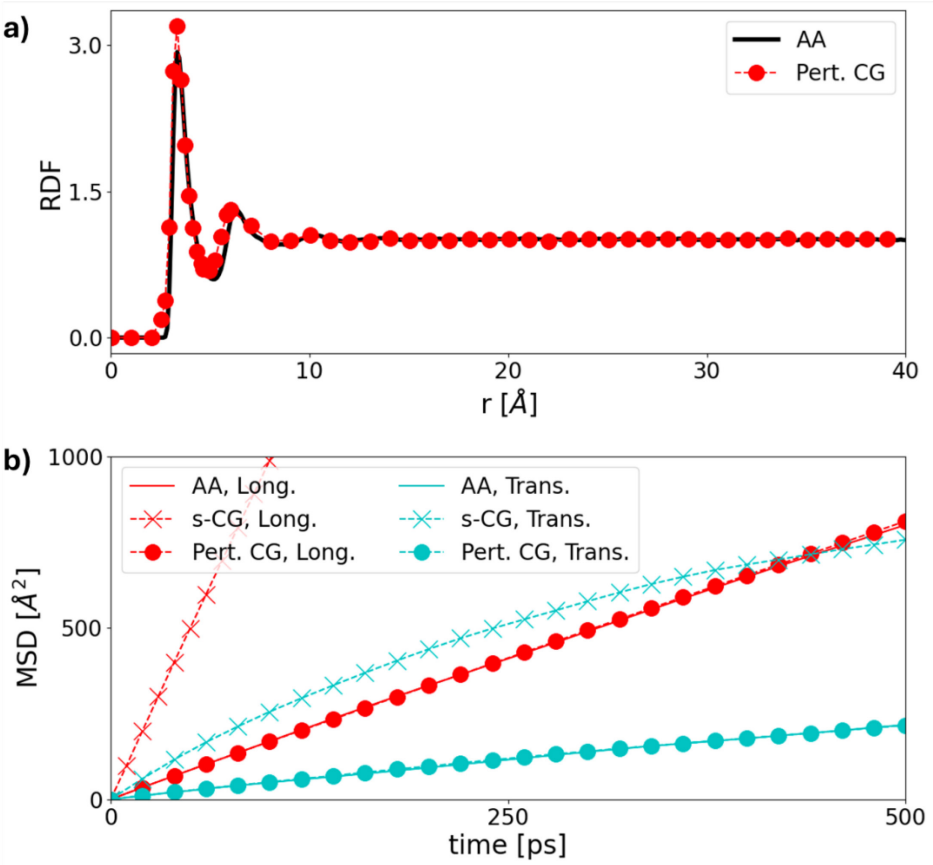

Figure 8: a) Density variation, b) Longitudinal and Transverse MSD for AA and CG water in confinement.

Next, we present the application of the methodology to a many-body potential derived for the SPC/F water system used earlier. The results illustrated in Figure 9 indicate that, in addition to accurately replicating the RDF, the DeepCG potential also captures the angular distribution function (ADF). The ADF describes the probability of finding a particular angle between a central atom and two neighboring withing a given cutoff radius, thereby providing insights into angular structure between triplet of atoms. The many-body DeepCG potential allows for more accurate modeling of the MB-PMF enabling the capture of higher-order correlations that are often missed by pairwise potentials. By introducing high-frequency perturbations with

parameters ω = 230 $Å^{-1}$ and A = 4.0 kcal·$mol^{-1}Å^{-1}$ to the DeepCG potential, we observe a correction in the self-diffusion behavior of the s-CG system, bringing its dynamics closer to that of the All-Atom system. Similar results are obtained for the GNN potential optimized using Adjoint-state method, as shown in Figure S4 of the supporting information.

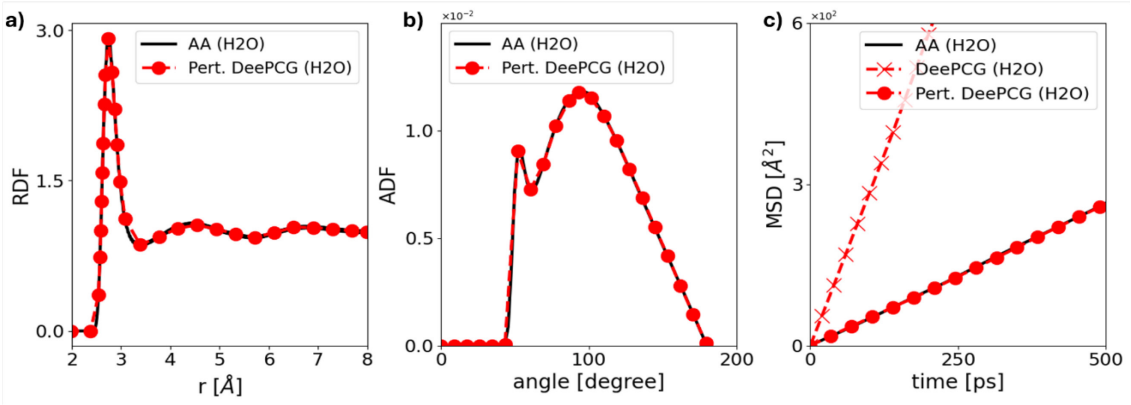

Figure 9: a) Radial Distribution Function, b) Angular Distribution Function and c) Mean Squared Displacement for AA and DeepCG systems.

4. Conclusion

In this study, we explored the idea of perturbing the Free-Energy-Landscape to capture characteristic features of the fine-grained atomistic system, which are absent in the simplistic CG force fields that primarily capture average statistical properties like the MB-PMF. The investigation demonstrated that by perturbing structure-matching CG potentials, we could reproduce the long-time diffusion behavior of the All-Atom system. We theoretically established that the gradient of relative entropy between the perturbed CG and the AA system remains unchanged for sufficiently high-frequency values, thereby ensuring local structure invariance to such perturbations. The impact of these perturbation parameters in the CG system, interacting via relative entropy-optimized pairwise potentials, was scrutinized through molecular dynamics simulations employing an extensive grid search. This exploration unveiled distinctive features of the energy landscapes of H_2O, CO_2, and CH_4 thereby allowing us to

optimize the perturbation parameters to match the RDF and self-diffusion coefficients for all three molecules. Extending our methodology to inhomogeneous systems, we demonstrated the ability of a perturbed CG system in replicating the density as well as the diffusion of water molecules parallel and perpendicular to the nanochannel. While pairwise potentials aptly predict key thermodynamic quantities for most atomic systems, they falter in modelling situations dominated by multibody effects. Expanding our approach to more potent many-body potentials, we showcased its compatibility with popular neural network potentials optimized using different frameworks - specifically, DeepCG and a GNN Potential. The landscape paradigm thus provides a reasonable approach for parameterizing coarse-grained systems, to match long-time dynamics through the modeling of prevailing energy-barriers.

ASSOCIATED CONTENT

The supporting information is available free of charge at: -

Theoretical and Numerical analysis of structure invariance to high-frequency perturbations is presented in Appendix A along with details of the Graph Neural Network based potential in Appendix B.

Data availability: -

The RE optimized and Machine Learnt Potential files used for coarse-graining procedure in this work can be accessed at https://github.com/multinanogroup/PES_perturbation

AUTHOR INFORMATION:

Corresponding Author

*Email - aluru@utexas.edu

Notes

The authors declare no competing interests**.**

ACKNOWLEDGEMENTS:

The work on deep learning was supported by the Center for Enhanced Nanofluidic Transport (CENT), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences (Award No. DE-SC0019112). All other aspects of this work were supported by the National Science Foundation under Grant No 2137157. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing access to the Lonestar6 resource that has contributed to the research results reported within this paper. We also acknowledge the use of the Extreme Science and Engineering Discovery Environment (XSEDE) Stampede2 at the Texas Advanced Computing Centre through Allocation No. TG-CDA100010.

REFERENCES:

(1) Lee, E. H.; Hsin, J.; Sotomayor, M.; Comellas, G.; Schulten, K. Discovery Through the Computational Microscope. *Structure* **2009**, *17* (10), 1295–1306. https://doi.org/10.1016/j.str.2009.09.001.

(2) Tozzini, V. Coarse-Grained Models for Proteins. *Curr. Opin. Struct. Biol.* **2005**, *15* (2), 144–150. https://doi.org/10.1016/j.sbi.2005.02.005.

(3) Stevens, M. J. Coarse-Grained Simulations of Lipid Bilayers. *J. Chem. Phys.* **2004**, *121* (23), 11942–11948. https://doi.org/10.1063/1.1814058.

(4) *Multiscale compression-induced restructuring of stacked lipid bilayers: From buckling delamination to molecular packing | PLOS ONE*. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0275079 (accessed 2023-12-24).

(5) Knotts, T. A., IV; Rathore, N.; Schwartz, D. C.; de Pablo, J. J. A Coarse Grain Model for DNA. *J. Chem. Phys.* **2007**, *126* (8), 084901. https://doi.org/10.1063/1.2431804.

(6) Srinivas, G.; Cheng, X.; Smith, J. C. A Solvent-Free Coarse Grain Model for Crystalline and Amorphous Cellulose Fibrils. *J. Chem. Theory Comput.* **2011**, *7* (8), 2539–2548. https://doi.org/10.1021/ct200181t.

(7) Zhao, G.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C.; Zhang, P. Mature HIV-1 Capsid Structure by Cryo-Electron Microscopy and All-Atom Molecular Dynamics. *Nature* **2013**, *497* (7451), 643–646. https://doi.org/10.1038/nature12162.

(8) Roos, W. H.; Bruinsma, R.; Wuite, G. J. L. Physical Virology. *Nat. Phys.* **2010**, *6* (10), 733–743. https://doi.org/10.1038/nphys1797.

(9) Sanbonmatsu, K. Y. Computational Studies of Molecular Machines: The Ribosome. *Curr. Opin. Struct. Biol.* **2012**, *22* (2), 168–174. https://doi.org/10.1016/j.sbi.2012.01.008.

(10) Chandler, D. E.; Strümpfer, J.; Sener, M.; Scheuring, S.; Schulten, K. Light Harvesting by Lamellar Chromatophores in Rhodospirillum Photometricum. *Biophys. J.* **2014**, *106* (11), 2503–2510. https://doi.org/10.1016/j.bpj.2014.04.030.

(11) Feng, J.; Graf, M.; Liu, K.; Ovchinnikov, D.; Dumcenco, D.; Heiranian, M.; Nandigana, V.; Aluru, N. R.; Kis, A.; Radenovic, A. Single-Layer MoS2 Nanopores as Nanopower Generators. *Nature* **2016**, *536* (7615), 197–200. https://doi.org/10.1038/nature18593.

(12) Heiranian, M.; Farimani, A. B.; Aluru, N. R. Water Desalination with a Single-Layer MoS2 Nanopore. *Nat. Commun.* **2015**, *6* (1), 8616. https://doi.org/10.1038/ncomms9616.

(13) Noh, Y.; Aluru, N. R. Phonon-Fluid Coupling Enhanced Water Desalination in Flexible Two-Dimensional Porous Membranes. *Nano Lett.* **2022**, *22* (1), 419–425. https://doi.org/10.1021/acs.nanolett.1c04155.

(14) Baschnagel, J.; Binder, K.; Doruker, P.; Gusev, A. A.; Hahn, O.; Kremer, K.; Mattice, W. L.; Müller-Plathe, F.; Murat, M.; Paul, W.; Santos, S.; Suter, U. W.; Tries, V. Bridging the Gap Between Atomistic and Coarse-Grained Models of Polymers: Status and Perspectives. In *Viscoelasticity,*

Atomistic Models, Statistical Chemistry; Advances in Polymer Science; Springer: Berlin, Heidelberg, 2000; pp 41–156. https://doi.org/10.1007/3-540-46778-5_2.

(15) Richard Wilson, M. Molecular Simulation of Liquid Crystals: Progress towards a Better Understanding of Bulk Structure and the Prediction of Material Properties. *Chem. Soc. Rev.* **2007**, *36* (12), 1881–1888. https://doi.org/10.1039/B612799C.

(16) Padding, J. T.; Louis, A. A. Hydrodynamic Interactions and Brownian Forces in Colloidal Suspensions: Coarse-Graining over Time and Length Scales. *Phys. Rev. E* **2006**, *74* (3), 031402. https://doi.org/10.1103/PhysRevE.74.031402.

(17) Zimmerman, M. I.; Bowman, G. SARS-CoV-2 Simulations Go Exascale to Capture Spike Opening and Reveal Cryptic Pockets Across the Proteome. *Biophys. J.* **2021**, *120* (3), 299a. https://doi.org/10.1016/j.bpj.2020.11.1909.

(18) Jia, W.; Wang, H.; Chen, M.; Lu, D.; Lin, L.; Car, R.; E, W.; Zhang, L. Pushing the Limit of Molecular Dynamics with Ab Initio Accuracy to 100 Million Atoms with Machine Learning. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*; SC '20; IEEE Press: Atlanta, Georgia, 2020; pp 1–14.

(19) Guo, Z.; Lu, D.; Yan, Y.; Hu, S.; Liu, R.; Tan, G.; Sun, N.; Jiang, W.; Liu, L.; Chen, Y.; Zhang, L.; Chen, M.; Wang, H.; Jia, W. Extending the Limit of Molecular Dynamics with *Ab Initio* Accuracy to 10 Billion Atoms. In *Proceedings of the 27th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*; ACM: Seoul Republic of Korea, 2022; pp 205–218. https://doi.org/10.1145/3503221.3508425.

(20) Jung, J.; Kobayashi, C.; Kasahara, K.; Tan, C.; Kuroda, A.; Minami, K.; Ishiduki, S.; Nishiki, T.; Inoue, H.; Ishikawa, Y.; Feig, M.; Sugita, Y. New Parallel Computing Algorithm of Molecular Dynamics for Extremely Huge Scale Biological Systems. *J. Comput. Chem.* **2021**, *42* (4), 231–241. https://doi.org/10.1002/jcc.26450.

(21) *Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations - Jung - 2019 - Journal of Computational Chemistry - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.25840?casa_token=ugAnf_8CAKMAAAAA%3ATOS6SZ6LGqzZoSa_K7KpKdzspu472dZKigyk85v2b_7vH98xlfzGMDlJJ2La7U_G_dJPoFEvbKc6ig (accessed 2023-12-24).

(22) Shaw, D. E.; Dror, R. O.; Salmon, J. K.; Grossman, J. P.; Mackenzie, K. M.; Bank, J. A.; Young, C.; Deneroff, M. M.; Batson, B.; Bowers, K. J.; Chow, E.; Eastwood, M. P.; Ierardi, D. J.; Klepeis, J. L.; Kuskin, J. S.; Larson, R. H.; Lindorff-Larsen, K.; Maragakis, P.; Moraes, M. A.; Piana, S.; Shan, Y.; Towles, B. Millisecond-Scale Molecular Dynamics Simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*; SC '09; Association for Computing Machinery: New York, NY, USA, 2009; pp 1–11. https://doi.org/10.1145/1654059.1654126.

(23) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science* **2011**, *334* (6055), 517–520. https://doi.org/10.1126/science.1208351.

(24) *Perspective: Coarse-grained models for biomolecular systems | The Journal of Chemical Physics | AIP Publishing*. https://pubs.aip.org/aip/jcp/article/139/9/090901/315544 (accessed 2023-12-24).

(25) Yu, A.; Pak, A. J.; He, P.; Monje-Galvan, V.; Casalino, L.; Gaieb, Z.; Dommer, A. C.; Amaro, R. E.; Voth, G. A. A Multiscale Coarse-Grained Model of the SARS-CoV-2 Virion. *Biophys. J.* **2021**, *120* (6), 1097–1104. https://doi.org/10.1016/j.bpj.2020.10.048.

(26) Carvalho, A. P.; Santos, S. M.; Pérez-Sánchez, G.; Gouveia, J. D.; Gomes, J. R. B.; Jorge, M. Sticky-MARTINI as a Reactive Coarse-Grained Model for Molecular Dynamics Simulations of Silica Polymerization. *Npj Comput. Mater.* **2022**, *8* (1), 1–13. https://doi.org/10.1038/s41524-022-00722-w.

(27) *Percolation transition prescribes protein size-specific barrier to passive transport through the nuclear pore complex | Nature Communications*. https://www.nature.com/articles/s41467-022-32857-1 (accessed 2023-12-24).

(28) *Full article: Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants*. https://www.tandfonline.com/doi/full/10.1080/08927020601054050?casa_token=gVlJRY2GlBQAAAA:-K6-mh_aZvzJPVa4228hD0K5-rxgoUlkKh6plDd3roWzqzBi8-pvP4TMEfx-N0CHz7o2cTSqThkN (accessed 2023-12-24).

(29) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Moore, P. B.; Klein, M. L. Simulations of Phospholipids Using a Coarse Grain Model. *J. Phys. Chem. B* **2001**, *105* (40), 9785–9792. https://doi.org/10.1021/jp011637n.

(30) Jin, J.; Pak, A. J.; Durumeric, A. E. P.; Loose, T. D.; Voth, G. A. Bottom-up Coarse-Graining: Principles and Perspectives. *J. Chem. Theory Comput.* **2022**, *18* (10), 5759–5791. https://doi.org/10.1021/acs.jctc.2c00643.

(31) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3* (5), 300–313. https://doi.org/10.1063/1.1749657.

(32) Izvekov, S.; Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. *J. Phys. Chem. B* **2005**, *109* (7), 2469–2473. https://doi.org/10.1021/jp044629q.

(33) Noid, W. G.; Chu, J.-W.; Ayton, G. S.; Krishna, V.; Izvekov, S.; Voth, G. A.; Das, A.; Andersen, H. C. The Multiscale Coarse-Graining Method. I. A Rigorous Bridge between Atomistic and Coarse-Grained Models. *J. Chem. Phys.* **2008**, *128* (24), 244114. https://doi.org/10.1063/1.2938860.

(34) *Deriving effective mesoscale potentials from atomistic simulations - Reith - 2003 - Journal of Computational Chemistry - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.10307?casa_token=-BYoO00GNMcAAAAA%3AuaMuyy_8u2c9gQLOPHsdJ6MGt4V9dVpM3eWUKBWLM37fW5MzQJYcm9Jvq_I92WHNbIPBLUQmsaAVog (accessed 2023-12-24).

(35) Shell, M. S. The Relative Entropy Is Fundamental to Multiscale and Inverse Thermodynamic Problems. *J. Chem. Phys.* **2008**, *129* (14), 144108. https://doi.org/10.1063/1.2992060.

(36) Chaimovich, A.; Shell, M. S. Coarse-Graining Errors and Numerical Optimization Using a Relative Entropy Framework. *J. Chem. Phys.* **2011**, *134* (9), 094112. https://doi.org/10.1063/1.3557038.

(37) *COARSE-GRAINING WITH THE RELATIVE ENTROPY - Shell - 2016 - Advances in Chemical Physics - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119290971.ch5 (accessed 2023-12-24).

(38) Depa, P. K.; Maranas, J. K. Speed up of Dynamic Observables in Coarse-Grained Molecular-Dynamics Simulations of Unentangled Polymers. *J. Chem. Phys.* **2005**, *123* (9), 094901. https://doi.org/10.1063/1.1997150.

(39) Depa, P.; Chen, C.; Maranas, J. K. Why Are Coarse-Grained Force Fields Too Fast? A Look at Dynamics of Four Coarse-Grained Polymers. *J. Chem. Phys.* **2011**, *134* (1), 014903. https://doi.org/10.1063/1.3513365.

(40) Salerno, K. M.; Agrawal, A.; Peters, B. L.; Perahia, D.; Grest, G. S. Dynamics in Entangled Polyethylene Melts. *Eur. Phys. J. Spec. Top.* **2016**, *225* (8–9), 1707–1722. https://doi.org/10.1140/epjst/e2016-60142-7.

(41) Fritz, D.; Herbers, C. R.; Kremer, K.; Vegt, N. F. A. van der. Hierarchical Modeling of Polymer Permeation. *Soft Matter* **2009**, *5* (22), 4556–4563. https://doi.org/10.1039/B911713J.

(42) Hoyt, J. J.; Asta, M.; Sadigh, B. Test of the Universal Scaling Law for the Diffusion Coefficient in Liquid Metals. *Phys. Rev. Lett.* **2000**, *85* (3), 594–597. https://doi.org/10.1103/PhysRevLett.85.594.

(43) Mori, H. A Continued-Fraction Representation of the Time-Correlation Functions. *Prog. Theor. Phys.* **1965**, *34* (3), 399–416. https://doi.org/10.1143/PTP.34.399.

(44) Zwanzig, R. Memory Effects in Irreversible Thermodynamics. *Phys. Rev.* **1961**, *124* (4), 983–992. https://doi.org/10.1103/PhysRev.124.983.

(45) Kirkwood, J. G. The Statistical Mechanical Theory of Transport Processes I. General Theory. *J. Chem. Phys.* **2004**, *14* (3), 180–201. https://doi.org/10.1063/1.1724117.

(46) Rudzinski, J. F. Recent Progress towards Chemically-Specific Coarse-Grained Simulation Models with Consistent Dynamical Properties. *Computation* **2019**, *7* (3), 42. https://doi.org/10.3390/computation7030042.

(47) Glatzel, F.; Schilling, T. The Interplay between Memory and Potentials of Mean Force: A Discussion on the Structure of Equations of Motion for Coarse-Grained Observables. *Europhys. Lett.* **2022**, *136* (3), 36001. https://doi.org/10.1209/0295-5075/ac35ba.

(48) Schilling, T. Coarse-Grained Modelling out of Equilibrium. *Phys. Rep.* **2022**, *972*, 1–45. https://doi.org/10.1016/j.physrep.2022.04.006.

(49) Hijón, C.; Español, P.; Vanden-Eijnden, E.; Delgado-Buscalioni, R. Mori–Zwanzig Formalism as a Practical Computational Tool. *Faraday Discuss* **2010**, *144*, 301–322. https://doi.org/10.1039/B902479B.

(50) *Three Routes to the Friction Matrix and Their Application to the Coarse-Graining of Atomic Lattices - Kauzlarić - 2011 - Macromolecular Theory and Simulations - Wiley Online Library*. https://onlinelibrary.wiley.com/doi/full/10.1002/mats.201100014?casa_token=HbK62ThG01MAAAAA%3AtOHsxA-oLPFqo7NCVt1fmAYiRGXF8Czk9M8A8v4q0armJjXYw_b2aUcEQmnOhIqCsBwIN-kwaO77cA (accessed 2023-12-24).

(51) Li, Z.; Bian, X.; Li, X.; Karniadakis, G. E. Incorporation of Memory Effects in Coarse-Grained Modeling via the Mori-Zwanzig Formalism. *J. Chem. Phys.* **2015**, *143* (24), 243128. https://doi.org/10.1063/1.4935490.

(52) Li, Z.; Lee, H. S.; Darve, E.; Karniadakis, G. E. Computing the Non-Markovian Coarse-Grained Interactions Derived from the Mori–Zwanzig Formalism in Molecular Systems: Application to Polymer Melts. *J. Chem. Phys.* **2017**, *146* (1), 014104. https://doi.org/10.1063/1.4973347.

(53) te Vrugt, M.; Wittkowski, R. Mori-Zwanzig Projection Operator Formalism for Far-from-Equilibrium Systems with Time-Dependent Hamiltonians. *Phys. Rev. E* **2019**, *99* (6), 062118. https://doi.org/10.1103/PhysRevE.99.062118.

(54) Izvekov, S. Mori-Zwanzig Theory for Dissipative Forces in Coarse-Grained Dynamics in the Markov Limit. *Phys. Rev. E* **2017**, *95* (1), 013303. https://doi.org/10.1103/PhysRevE.95.013303.

(55) *Generalized Langevin Equations for Systems with Local Interactions | Journal of Statistical Physics*. https://link.springer.com/article/10.1007/s10955-020-02499-y (accessed 2023-12-24).

(56) *Dissipative Particle Dynamics (DPD): An Overview and Recent Developments | Archives of Computational Methods in Engineering*. https://link.springer.com/article/10.1007/s11831-014-9124-x (accessed 2023-12-24).

(57) Klippenstein, V.; van der Vegt, N. F. A. Bottom-Up Informed and Iteratively Optimized Coarse-Grained Non-Markovian Water Models with Accurate Dynamics. *J. Chem. Theory Comput.* **2023**, *19* (4), 1099–1110. https://doi.org/10.1021/acs.jctc.2c00871.

(58) Klippenstein, V.; Wolf, N.; van der Vegt, N. F. A. A Gauss–Newton Method for Iterative Optimization of Memory Kernels for Generalized Langevin Thermostats in Coarse-Grained Molecular Dynamics Simulations. *J. Chem. Phys.* **2024**, *160* (20), 204115. https://doi.org/10.1063/5.0203832.

(59) Tripathy, M.; Klippenstein, V.; van der Vegt, N. F. A. Dynamical Coarse-Grained Models of Molecular Liquids and Their Ideal and Non-Ideal Mixtures. *J. Chem. Phys.* **2023**, *159* (9), 094904. https://doi.org/10.1063/5.0163097.

(60) Padding, J. T.; Briels, W. J. Systematic Coarse-Graining of the Dynamics of Entangled Polymer Melts: The Road from Chemistry to Rheology. *J. Phys. Condens. Matter* **2011**, *23* (23), 233101. https://doi.org/10.1088/0953-8984/23/23/233101.

(61) Rosenfeld, Y. A Quasi-Universal Scaling Law for Atomic Transport in Simple Fluids. *J. Phys. Condens. Matter* **1999**, *11* (28), 5415. https://doi.org/10.1088/0953-8984/11/28/303.

(62) Dzugutov, M. A Universal Scaling Law for Atomic Diffusion in Condensed Matter. *Nature* **1996**, *381* (6578), 137–139. https://doi.org/10.1038/381137a0.

(63) Veldhorst, A. A.; Dyre, J. C.; Schrøder, T. B. Scaling of the Dynamics of Flexible Lennard-Jones Chains. *J. Chem. Phys.* **2014**, *141* (5), 054904. https://doi.org/10.1063/1.4888564.

(64) Willers, C.; Kamps, O. Efficient Bayesian Estimation of the Generalized Langevin Equation from Data. *J. Comput. Phys.* **2024**, *497*, 112626. https://doi.org/10.1016/j.jcp.2023.112626.

(65) Harmandaris, V.; Kalligiannaki, E.; Katsoulakis, M.; Plecháč, P. Path-Space Variational Inference for Non-Equilibrium Coarse-Grained Systems. *J. Comput. Phys.* **2016**, *314*, 355–383. https://doi.org/10.1016/j.jcp.2016.03.021.

(66) Hernández, C. X.; Wayment-Steele, H. K.; Sultan, M. M.; Husic, B. E.; Pande, V. S. Variational Encoding of Complex Dynamics. *Phys. Rev. E* **2018**, *97* (6), 062412. https://doi.org/10.1103/PhysRevE.97.062412.

(67) *Incorporating physical constraints in a deep probabilistic machine learning framework for coarse-graining dynamical systems - ScienceDirect*. https://www.sciencedirect.com/science/article/pii/S0021999120304472?casa_token=Zp1K1-g5gfYAAAAA:y8tWn1gf7366WuwD9ZFZ7_NtmIrOWO6HgqIIuxDmQxQU1IbwqX5L7A_Wq4o_4AEMu98LXn5hAg (accessed 2023-12-24).

(68) Bowman, G. R.; Pande, V. S.; Noé, F. *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*; Springer Science & Business Media, 2013.

(69) Mardt, A.; Pasquali, L.; Wu, H.; Noé, F. VAMPnets for Deep Learning of Molecular Kinetics. *Nat. Commun.* **2018**, *9* (1), 5. https://doi.org/10.1038/s41467-017-02388-1.

(70) Rudzinski, J. F.; Bereau, T. Concurrent Parametrization against Static and Kinetic Information Leads to More Robust Coarse-Grained Force Fields. *Eur. Phys. J. Spec. Top.* **2016**, *225* (8), 1373–1389. https://doi.org/10.1140/epjst/e2016-60114-5.

(71) Rudzinski, J. F.; Bereau, T. Structural-Kinetic-Thermodynamic Relationships Identified from Physics-Based Molecular Simulation Models. *J. Chem. Phys.* **2018**, *148* (20), 204111. https://doi.org/10.1063/1.5025125.

(72) Hänggi, P.; Talkner, P.; Borkovec, M. Reaction-Rate Theory: Fifty Years after Kramers. *Rev. Mod. Phys.* **1990**, *62* (2), 251–341. https://doi.org/10.1103/RevModPhys.62.251.

(73) Goldstein, M. Viscous Liquids and the Glass Transition: A Potential Energy Barrier Picture. *J. Chem. Phys.* **1969**, *51* (9), 3728–3739. https://doi.org/10.1063/1.1672587.

(74) Stillinger, F. H.; Weber, T. A. Hidden Structure in Liquids. *Phys. Rev. A* **1982**, *25* (2), 978–989. https://doi.org/10.1103/PhysRevA.25.978.

(75) Angell, A. Liquid Landscape. *Nature* **1998**, *393* (6685), 521–523. https://doi.org/10.1038/31110.

(76) Sastry, S.; Debenedetti, P. G.; Stillinger, F. H. Signatures of Distinct Dynamical Regimes in the Energy Landscape of a Glass-Forming Liquid. *Nature* **1998**, *393* (6685), 554–557.

(77) Schulz, M. Energy Landscape, Minimum Points, and Non-Arrhenius Behavior of Supercooled Liquids. *Phys. Rev. B* **1998**, *57* (18), 11319–11333. https://doi.org/10.1103/PhysRevB.57.11319.

(78) Debenedetti, P. G.; Stillinger, F. H. Supercooled Liquids and the Glass Transition. *Nature* **2001**, *410* (6825), 259–267. https://doi.org/10.1038/35065704.

(79) Chavez, L. L.; Onuchic, J. N.; Clementi, C. Quantifying the Roughness on the Free Energy Landscape: Entropic Bottlenecks and Protein Folding Rates. *J. Am. Chem. Soc.* **2004**, *126* (27), 8426–8432. https://doi.org/10.1021/ja049510+.

(80) Yu, H.; Dee, D. R.; Liu, X.; Brigley, A. M.; Sosova, I.; Woodside, M. T. Protein Misfolding Occurs by Slow Diffusion across Multiple Barriers in a Rough Energy Landscape. *Proc. Natl. Acad. Sci.* **2015**, *112* (27), 8308–8313. https://doi.org/10.1073/pnas.1419197112.

(81) Zwanzig, R. Diffusion in a Rough Potential. *Proc. Natl. Acad. Sci.* **1988**, *85* (7), 2029–2030. https://doi.org/10.1073/pnas.85.7.2029.

(82) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. DeePCG: Constructing Coarse-Grained Models via Deep Neural Networks. *J. Chem. Phys.* **2018**, *149* (3), 034101. https://doi.org/10.1063/1.5027645.

(83) Chen, R. T. Q.; Rubanova, Y.; Bettencourt, J.; Duvenaud, D. K. Neural Ordinary Differential Equations. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc., 2018; Vol. 31.

(84) Wang, W.; Axelrod, S.; Gómez-Bombarelli, R. Differentiable Molecular Simulations for Control and Learning. arXiv December 23, 2020. https://doi.org/10.48550/arXiv.2003.00868.

(85) Doerr, S.; Majewski, M.; Pérez, A.; Krämer, A.; Clementi, C.; Noe, F.; Giorgino, T.; De Fabritiis, G. TorchMD: A Deep Learning Framework for Molecular Simulations. *J. Chem. Theory Comput.* **2021**, *17* (4), 2355–2363. https://doi.org/10.1021/acs.jctc.0c01343.

(86) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* **1997**, *48* (1), 545–600. https://doi.org/10.1146/annurev.physchem.48.1.545.

(87) Nelson Onuchic, J.; Nymeyer, H.; García, A. E.; Chahine, J.; Socci, N. D. The Energy Landscape Theory of Protein Folding: Insights into Folding Mechanisms and Scenarios. In *Advances in Protein Chemistry*; Protein folding mechanisms; Academic Press, 2000; Vol. 53, pp 87–152. https://doi.org/10.1016/S0065-3233(00)53003-4.

(88) *The nature of protein folding pathways | PNAS*. https://www.pnas.org/doi/abs/10.1073/pnas.1411798111 (accessed 2023-12-24).

(89) Banerjee, S.; Biswas, R.; Seki, K.; Bagchi, B. Diffusion on a Rugged Energy Landscape with Spatial Correlations. *J. Chem. Phys.* **2014**, *141* (12), 124105. https://doi.org/10.1063/1.4895905.

(90) Gray, T. H.; Yong, E. H. Effective Diffusion in One-Dimensional Rough Potential-Energy Landscapes. *Phys. Rev. E* **2020**, *102* (2), 022138. https://doi.org/10.1103/PhysRevE.102.022138.

(91) Gray, T. H.; Yong, E. H. An Effective One-Dimensional Approach to Calculating Mean First Passage Time in Multi-Dimensional Potentials. *J. Chem. Phys.* **2021**, *154* (8), 084103. https://doi.org/10.1063/5.0040071.

(92) Duncan, A. B.; Kalliadasis, S.; Pavliotis, G. A.; Pradas, M. Noise-Induced Transitions in Rugged Energy Landscapes. *Phys. Rev. E* **2016**, *94* (3), 032107. https://doi.org/10.1103/PhysRevE.94.032107.

(93) Mashayak, S. Y.; Jochum, M. N.; Koschke, K.; Aluru, N. R.; Rühle, V.; Junghans, C. Relative Entropy and Optimization-Driven Coarse-Graining Methods in VOTCA. *PLOS ONE* **2015**, *10* (7), e0131754. https://doi.org/10.1371/journal.pone.0131754.

(94) Thaler, S.; Stupp, M.; Zavadlav, J. Deep Coarse-Grained Potentials via Relative Entropy Minimization. *J. Chem. Phys.* **2022**, *157* (24), 244103. https://doi.org/10.1063/5.0124538.

(95) Mashayak, S. Y.; Aluru, N. R. Thermodynamic State-Dependent Structure-Based Coarse-Graining of Confined Water. *J. Chem. Phys.* **2012**, *137* (21), 214707. https://doi.org/10.1063/1.4769297.

(96) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible Simple Point-Charge Water Model with Improved Liquid-State Properties. *J. Chem. Phys.* **2006**, *124* (2), 024503. https://doi.org/10.1063/1.2136877.

(97) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J. Chem. Theory Comput.* **2011**, *7* (12), 4026–4037. https://doi.org/10.1021/ct200196m.

(98) Nosé, S. An Extension of the Canonical Ensemble Molecular Dynamics Method. *Mol. Phys.* **1986**, *57* (1), 187–191. https://doi.org/10.1080/00268978600100141.

(99) Sanghi, T.; Aluru, N. R. Tarun-Thermal Noise in Confined Fluids. *J. Chem. Phys.* **2014**, *141* (17), 174707. https://doi.org/10.1063/1.4900501.

(100) Wallqvist, A.; Teleman, O. Properties of Flexible Water Models. *Mol. Phys.* **1991**, *74* (3), 515–533. https://doi.org/10.1080/00268979100102391.

(101) Yoshimoto, Y.; Kinefuchi, I.; Mima, T.; Fukushima, A.; Tokumasu, T.; Takagi, S. Bottom-up Construction of Interaction Models of Non-Markovian Dissipative Particle Dynamics. *Phys. Rev. E* **2013**, *88* (4), 043305. https://doi.org/10.1103/PhysRevE.88.043305.

(102) Español, P.; de la Torre, J. A.; Duque-Zumajo, D. Solution to the Plateau Problem in the Green-Kubo Formula. *Phys. Rev. E* **2019**, *99* (2), 022126. https://doi.org/10.1103/PhysRevE.99.022126.

(103) Markutsya, S.; Lamm, M. H. A Coarse-Graining Approach for Molecular Simulation That Retains the Dynamics of the All-Atom Reference System by Implementing Hydrodynamic Interactions. *J. Chem. Phys.* **2014**, *141* (17), 174107. https://doi.org/10.1063/1.4898625.

(104) Meinel, M. K.; Müller-Plathe, F. Loss of Molecular Roughness upon Coarse-Graining Predicts the Artificially Accelerated Mobility of Coarse-Grained Molecular Simulation Models. *J. Chem. Theory Comput.* **2020**, *16* (3), 1411–1419. https://doi.org/10.1021/acs.jctc.9b00943.

(105) Sarman, S. The Influence of the Fluid–Wall Interaction Potential on the Structure of a Simple Fluid in a Narrow Slit. *J. Chem. Phys.* **1990**, *92* (7), 4447–4455. https://doi.org/10.1063/1.457755.

(106) Tsimpanogiannis, I. N.; Moultos, O. A.; Franco, L. F. M.; Spera, M. B. de M.; Erdős, M.; Economou, I. G. Self-Diffusion Coefficient of Bulk and Confined Water: A Critical Review of Classical Molecular Simulation Studies. *Mol. Simul.* **2019**, *45* (4–5), 425–453. https://doi.org/10.1080/08927022.2018.1511903.

(107) Gordillo, M. C.; Martí, J. High Temperature Behavior of Water inside Flat Graphite Nanochannels. *Phys. Rev. B* **2007**, *75* (8), 085406. https://doi.org/10.1103/PhysRevB.75.085406.

(108) Kingma, D. P.; Ba, J. *Adam: A Method for Stochastic Optimization*. arXiv.org. https://arxiv.org/abs/1412.6980v9 (accessed 2023-12-28).

TOC Graphic-