Reinforcement Learning for Multi-Scale Molecular Modeling

Jun Zhang[1,†], Yao-Kun Lei[2] and Yi Isaac Yang[3,‡]
[1]*Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany*
[2] *Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, 100871 Beijing, China.*
3 Institute of Systems Biology, Shenzhen Bay Laboratory, 518055 Shenzhen, China
† jun.zhang@fu-berlin.de
‡ yangyi@szbl.ac.cn

Molecular simulations are widely applied in the study of chemical and bio-physical systems of interest. However, the accessible timescales of atomistic simulations are limited, and extracting equilibrium properties of systems containing rare events remains challenging. Two distinct strategies are usually adopted in this regard: either sticking to the atom level and performing enhanced sampling, or trading details for speed by leveraging coarse-grained models. Although both strategies are promising, either of them, if adopted individually, exhibits severe limitations. In this paper we propose a machine-learning approach to ally these two worlds. In our approach, simulations on different scales are executed simultaneously and benefit mutually from their cross-talks: Accurate coarse-grained (CG) models can be inferred from the fine-grained (FG) simulations; In turn, FG simulations can be boosted by the guidance of CG models. Our method grounds on unsupervised and reinforcement learning, defined by a variational and adaptive training objective, and allows end-to-end and online training of parametric models. Through multiple experiments, we show that our method is efficient and flexible, and performs well on challenging chemical and bio-molecular systems.

Keywords: Coarse graining, Machine learning, Imitation learning, Enhanced sampling, Artificial neural networks

I. INTRODUCTION

Molecular simulations, particularly all-atom and *ab initio* molecular dynamics (MD), have furthered our understanding of many chemical and bio-physical processes [1-3]. In molecular simulations, interactions between particles (e.g., atoms, residues or molecules) are described by a (potential) energy function $U(\mathbf{R})$ of the configuration $\mathbf{R}$. To investigate such systems, one is often not interested in the exact energy, but the free-energy, or the equilibrium distribution, of some reduced descriptors, e.g., collective variables or CG variables [4,5], $\mathbf{s}(\mathbf{R})$ as a function of $\mathbf{R}$:

$$p(\mathbf{s}) = \frac{\int e^{-\beta U(\mathbf{R})}\delta\left(\mathbf{s}-\mathbf{s}(\mathbf{R})\right)d\mathbf{R}}{\int e^{-\beta U(\mathbf{R})}d\mathbf{R}} = \frac{e^{-\beta F(\mathbf{s})}}{Z} \quad (1)$$

$$F(\mathbf{s}) = -\frac{1}{\beta}\left[\log p(\mathbf{s}) + \log Z\right] \quad (2)$$

where $Z = \int e^{-\beta U(\mathbf{R})}d\mathbf{R}$ is the partition function. Eq. (2) holds up to an arbitrary additive constant. Usually $\mathbf{s}$ is selected to be slowly changing variables governing the process of interest, and the rest of degrees of freedom (DOF) can be treated in the mean-field fashion [6,7]. Under this setting, $F(\mathbf{s})$ becomes a CG description of the original thermodynamic system, and simulations performed under $F(\mathbf{s})$ are generally much faster than those running on the FG potential (i.e., $U(\mathbf{R})$) because $\mathrm{Dim}(\mathbf{s}) \ll \mathrm{Dim}(\mathbf{R})$ the cost of losing finer details. Therefore, Eqs. (1-2) are also known as the principle of thermodynamic consistency for coarse graining [6,8]. However, practical implementation of Eqs. (1-2) is hindered by two major issues: (1) How can one approximate a reliable analytical form for the CG potential $F(\mathbf{s})$ given access to samples drawn from $U(\mathbf{R})$? (2) How can one obtain samples, which are further used to infer $F(\mathbf{s})$, from the equilibrium distribution of $U(\mathbf{R})$? The former is known as coarse graining problem, while the latter as importance sampling problem, and both are of particular importance in physics, chemistry and biology.

Conventionally, if $\mathbf{s}$ is low-dimensional (say, $\mathrm{Dim}(\mathbf{s}) \leq 3$), non-parametric methods like kernel density estimation (KDE) [9-11] can be adopted to infer $F(\mathbf{s})$, but they become quickly infeasible as $\mathrm{Dim}(\mathbf{s})$ increases. Artificial neural networks (ANNs) and deep learning may offer extra flexibility and expressivity to this end [12-14]. For instance, several recent work proposed supervised learning approaches to fit $F(\mathbf{s})$ by ANNs [15-17]. However, fitting $F(\mathbf{s})$ as a regression problem has several drawbacks: (1) it necessitates gridding the space of $\mathbf{s}$, which can be computationally prohibitive for large $\mathrm{Dim}(\mathbf{s})$; (2) it is rather data-inefficient in that calculating $F(\mathbf{s})$ at one point needs a large amount of samples from $U(\mathbf{R})$ at the neighborhood of $\mathbf{s}$ (i.e., $\delta\left(\mathbf{s}-\mathbf{s}(\mathbf{R})\right)$); (3) The optimization objective defined for regression learning is not suitable for fitting an

imbalanced distribution where the accuracy of areas with higher mass density should be more emphasized.

An alternative view of inferring $p(\mathbf{s})$ is provided by statistical modeling and machine learning community, where density estimation of high-dimensional data has been a long-standing goal [14,18]. Particularly, a recent burst of work sparks new ideas to exploit deep learning to this end, giving rise to an active research field of generative learning and seminar models including variational auto-encoders [19], auto-regressive [20,21] and normalizing-flow-based models [22,23]. Nevertheless, these methods are all footed on certain simple prior distributions, hence the complexity of the distributions resulted from these methods is bound by the manifold structure of the prior distribution. Therefore, they are known to suffer from issues like mode-dropping (i.e., violating the ergodicity assumption in terms of statistical physics) and often assigning probability mass to areas unwarranted by the data [24].

On the other hand, generative learning with energy-based models (EBMs) can be dated back to even longer before [24-28], which in principle can fit arbitrarily complex distributions due to the flexibility and plasticity of the energy landscapes [29-31]. Based on EBMs we propose a variational approach to the CG problem, i.e., to infer analytic form for the complex free-energy surface without supervision. Furthermore, our new approach allows simulations on different scales, which are launched simultaneously, cross talk and benefit from each other. We will demonstrate that the inferred CG potential can in turn help enhance the sampling of the FG model in a reinforced and adaptive manner.

II. METHODS

1. Coarse graining with thermodynamic consistency

We propose a variational and parametric approach, called variational inference of free-energy (VIFE), to approximate the (possibly high-dimensional) free-energy surface (FES), $F(\mathbf{s})$, in a tractable manner. Specifically, we denote the approximate free-energy function and the associated probability distribution as $F_\theta(\mathbf{s})$ and $p_\theta(\mathbf{s})$ respectively (where θ are optimizable parameters), and define a strict divergence $D(p||p_\theta)$ between p and p_θ. A strict divergence, including but not limited to Kullback-Leibler divergence $D_{\mathrm{KL}}(p||p_\theta)$ [32] and Wasserstein divergence $D_{\mathrm{W}}(p||p_\theta)$ [33,34], satisfies the condition that $D(p||p_\theta) \geq 0$ where the equality holds if and only if $p = p_\theta$, hence can be used as a variational objective [35-37]. Note that VIFE is different from the variational methods for statistical mechanics where the reversed $D_{\mathrm{KL}}(p||p_\theta)$ is conventionally used [38-40]. Particularly, the gradient of $D_{\mathrm{KL}}(p||p_\theta)$ w.r.t. θ takes the following form (see SI or Ref. [36] for the derivation),

$$\nabla_\theta D_{\mathrm{KL}}\left(p \,\|\, p_\theta\right) = \left\langle \nabla_\theta F_\theta(\mathbf{s}) \right\rangle_{p(\mathbf{s})} - \left\langle \nabla_\theta F_\theta(\mathbf{s}) \right\rangle_{p_\theta(\mathbf{s})} \quad (3)$$

The gradient for $D_{\mathrm{W}}(p||p_\theta)$ has been separately derived in Targeted Adversarial Learning Optimized Sampling (TALOS) [41] and it shares dramatic similarity with Eq. (3). In this paper we perform experiments exclusively according to Eq. (3) and leave D_{W} for future research.

Now the problem is how to evaluate the two expectation values in Eq. (3). If we have access to, say, equilibrium MD samples on $U(\mathbf{R})$, the distribution of which is denoted as $p_{\mathrm{FG}}(\mathbf{s})$, we can approximate p by p_{FG}, and optimize F_θ w.r.t. a surrogate objective $D_{\mathrm{KL}}(p_{\mathrm{FG}}||p_\theta)$, minimizing which is equivalent to maximum likelihood estimation. The remaining task is to calculate $\langle \nabla_\theta F_\theta(\mathbf{s}) \rangle_{p_\theta(\mathbf{s})}$. Since the analytical forms of $F_\theta(\mathbf{s})$ and $\nabla_{\mathbf{s}} F_\theta(\mathbf{s})$ are both available, we can perform CG simulations on $F_\theta(\mathbf{s})$ to estimate $\langle \nabla_\theta F_\theta(\mathbf{s}) \rangle_{p_\theta(\mathbf{s})}$. Due to the simplicity of $F_\theta(\mathbf{s})$ (i.e., the number of DOFs in $\mathbf{s}$ is quite limited), the CG simulations are usually computationally economic and converge relatively fast, hence brute-force simulations via Monte Carlo (MC) sampling or Langevin dynamics generally suffice.

2. Reinforced VIFE

In a more common setting where p_{FG} is not available beforehand, we have to perform sampling according to $U(\mathbf{R})$ from scratch. Since $\mathrm{Dim}(\mathbf{R})$ is usually very high, i.e., $\mathrm{Dim}(\mathbf{R}) \gg \mathrm{Dim}(\mathbf{s})$, estimating ensemble averages over $p(\mathbf{s})$ through $p_{\mathrm{FG}}(\mathbf{s})$ is often infeasible for brute-force FG (e.g., all-atom or ab initio) simulations. A plethora of enhanced sampling methods have been developed trying to solve it, and there exists several excellent reviews on this topic [5,42-44]. Here we will show that VIFE, combined with reinforced imitation learning [45], provides a new solution to this problem. We name our new approach RE-VIFE, which exhibits several compelling merits: (1) RE-VIFE is able to handle high-dimensional CG variables while most of enhanced sampling methods become ineffective as $\mathrm{Dim}(\mathbf{s})$ grows; (2) RE-VIFE casts a well-defined optimization problem which allows cross-fertilization with deep learning; (3) RE-VIFE formulates an adaptive training objective which can be optimized variationally, hence ensuring the efficiency and convergence of the method.

Mathematically, if we have an optimal $F_\theta(\mathbf{s})$ which is equal to the ground-true $F(\mathbf{s})$, and perform sampling under $U\big((\mathbf{R})\big) - F_\theta\big(\mathbf{s}(\mathbf{R})\big)$, we will arrive at a uniform distribution over $\mathbf{s}$. Therefore, one is motivated to employ $F_\theta(\mathbf{s})$ as a bias potential so that one can achieve a flattened distribution over $\mathbf{s}$ which may originally involve high free-energy barriers [16]. However, the situation is complicated by the errors in F_θ: Since F_θ is optimized w.r.t. available FG samples (usually correspond to metastable states in MD simulations), its value can be very inaccurate in those under-sampled regions (e.g., transition regions). Therefore, directly inserting F_θ as the bias potential could be risky. We note here that similar problems where one has to deal with moving distributions and partial sampling are commonly

encountered and addressed in Reinforcement Learning (RL) [46]. Specifically, adding a bias potential is equivalent to manipulating the policy function in RL settings [47]. Inspired by recent development in reinforced imitation learning [45,48,49] and Actor-Critic RL [50,51], we introduce a two-timescale learning scheme, where a policy function, i.e., the bias potential $V_\phi(\mathbf{s})$ (with ϕ denoting optimizable parameters) is separately trained in addition to $F_\theta(\mathbf{s})$ which can now be viewed as a value function in the spirit of RL. As in variationally enhanced sampling (VES) [37] or TALOS [41], we can define a target distribution $p_{\mathrm{T}}(\mathbf{s})$ where the free-energy barrier is lowered (in other words, less-visited regions are more encouraged) according to $F_\theta(\mathbf{s})$ (see SI for more information about p_{T}), then optimize $V_\phi(\mathbf{s})$ by minimizing a strict divergence, for instance, $D_{\mathrm{KL}}(p_{\mathrm{T}}||p_\phi)$,

$$\nabla_\phi D_{\mathrm{KL}}\left(p_{\mathrm{T}} \,\|\, p_\phi\right) = \left\langle \nabla_\phi V_\phi(\boldsymbol{s}) \right\rangle_{p_{\mathrm{T}}} - \left\langle \nabla_\phi V_\phi(\boldsymbol{s}) \right\rangle_{p_\phi} \qquad (4)$$

where p_ϕ denotes the distribution induced by the bias potential. The separate parametrization allows us to use an imbalanced learning schedule for the value and policy functions. Particularly, we can train $F_\theta(\mathbf{s})$ based on the latest $p_{\mathrm{FG}}(\mathbf{s})$ with a higher rate, and update $V_\phi(\mathbf{s})$ in a more conservative manner. In terms of imitation learning, $F_\theta(\mathbf{s})$ plays the role of a leader that coins a moving target based on the current density estimation, while $V_\phi(\mathbf{s})$ learns to tune the policy in order to catch up (Fig. 1). A dramatic advantage of such leader-chaser scheme lies in the fact that F_θ along with p_{T} is constructed on the samples drawn from simulations under V_ϕ, thus p_{T} and p_ϕ always share substantial overlap, otherwise D_{KL} would fall victim to the notorious vanishing gradient issue [33]. Remarkably, such sort of separate parametrization and two-timescale updated rule are also adopted in the training of generative adversarial networks (GANs) [52,53] and AlphaGo-Zero [54] as well as other reinforcement and imitation learning settings [45,49].

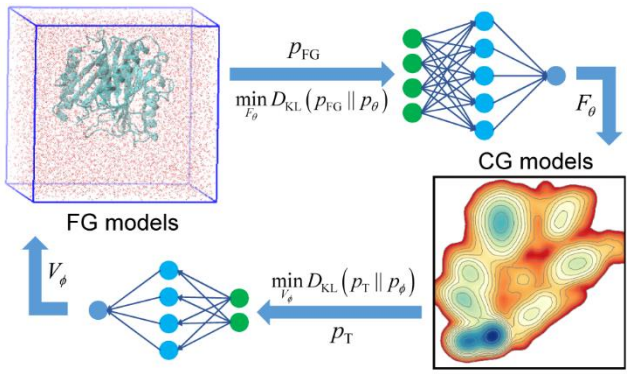

FIG. 1. Illustration of RE-VIFE. Given simulation samples from fine-grained (FG) models (p_{FG}), a coarse-grained (CG) potential (or free-energy) F_θ can be variationally approximated. In turn, a target distribution p_{T} can be defined based on CG simulations running over F_θ, according to which a bias potential V_ϕ can be variationally optimized to boost the FG simulation. This illustration consists of only one CG-FG cycle, and additional cycles corresponding to more scales can be added on top in the same manner.

In RE-VIFE, albeit $F_\theta(\mathbf{s})$ may be imperfect globally (especially at the beginning of training), it still gives relatively accurate density estimation of the recently visited regions, according to which $V_\phi(\mathbf{s})$ can be improved locally via Eq. (4) and gradually enhance the sampling efficiency. On the other hand, as the FG sampling is enhanced by $V_\phi(\mathbf{s})$, we can get a better approximation of $p(\mathbf{s})$ based on $p_\phi(\mathbf{s})$, thus gradually push $F_\theta(\mathbf{s})$ to the optimum according to Eq. (3). From this perspective, if we adopt KDE in replacement of $F_\theta(\mathbf{s})$, and use accumulative Gassians as $V_\phi(\mathbf{s})$, RE-VIFE turns equivalent to metadynamics. In this sense, RE-VIFE is a generalization of metadynamics to large Dim($\mathbf{s}$) and parametric bias functions.

3. Parametrization in (RE-)VIFE

In (RE-)VIFE, we train parametric models to approximate F_θ and V_ϕ, with θ and ϕ denoting optimizable parameters, and many functional forms are optional to this end. As in VES and TALOS, we can use certain orthonormal basis functions to expand the two functions in low-dimensional spaces [37,41]. With some caveats, we can also exploit ANNs to approximate F_θ and V_ϕ. Compared to other functional approximation methods, ANN is more flexible and expressive: On the one hand, ANNs allow us to cheaply approximate high-dimensional functions which would be intractable otherwise; On the other hand, the gradients or forces produced by ANNs are not necessarily smooth enough, which may cause instabilities for numerical integration in MD [55]. To attack the issue of irregular gradients while maximally harnessing the expressivity, we propose Neural Allocative Potentials (NAPs) as a new *ansatz* for constructing energy functions based on ANNs (see SI for more details). Additionally, in some cases where Cartesian coordinates are preferred to be directly chosen as $\mathbf{s}$, we recommend special neural networks which can remove the trans-rotational dependence in the coordinates as the parametric models [56-58].

III. RESULTS

We first tested VIFE on a 2-dimensional 3-well toy model [59]. The contour map of the potential energy surface (PES), $U(x,y)$, is shown in Fig. 2A. Langevin dynamics was first performed on $U(x,y)$ and yielded 50,000 samples, which were used as p_{FG}. We then choose the coarse-graining variable $\mathbf{s} = (x,y)$ (we desire to reproduce $U(x,y)$ instead of coarse-graining), and train a parametric model to infer $F_\theta(x,y)$ based on the available samples. NAP was adopted

as F_θ and the model was optimized via Eq. (3) (see SI for more details about model setup and training details).

Training of VIFE lasted for 100 iterations, and we plotted the estimated D_{KL} between p_θ and p_{FG} against training progress (Fig. 2B). It can be seen that D_{KL} between the two distributions quickly diminished, and the optimization of the CG model converged within 50 iterations, demonstrating the efficiency and efficacy of VIFE. Note that Eq. (3) can be viewed as the gradient for the following effective loss function,

$$L(\theta) \doteq \left\langle F_\theta(\boldsymbol{s}) \right\rangle_{p(\boldsymbol{s})} - \left\langle F_\theta(\boldsymbol{s}) \right\rangle_{p_\theta(\boldsymbol{s})} \tag{5}$$

hence we also showed the evolution of $L(\theta)$ during training. Intriguingly, we observed in Fig. 1B that there is an ups-and-downs pattern in the curve of $L(\theta)$, with the oscillating amplitude diminishes throughout training. This is an expected phenomenon because Eq. (5) is reminiscent of the training objective of Wasserstein-GAN [33], which solves a saddle-point (or mini-max) problem thus entails an oscillating loss curve. Besides D_{KL}, for EBMs, another important indicator of the quality of models is whether or not the energies of FG and CG samples are identically (or similarly) distributed over F_θ. Therefore, we examined the two distributions accordingly (Fig. S1), and found that the they almost overlap perfectly, proving the resulting F_θ is of high quality.

We then drew the contour plot of F_θ in Fig. 2C. Compared to the original PES (Fig. 2A), F_θ preserves all the three local minima correctly, showing no signs of mode-dropping. This is a particular advantage of EBMs over other generative methods, since ergodicity is vital to investigation of thermodynamic systems. Besides, the overall contour and landscape of F_θ also resemble $U(x,y)$, especially on regions with higher densities. In contrast, deviations seem to be more significant in the transition regions. This arises from the fact that we only collected 50,000 FG samples (most of which locate around the local minima), and the transition regions are covered by limited number of given FG samples (as shown in Fig. 2D). If more samples over the transition regions can be obtained, we would expect F_θ can be further improved, and such consideration is a strong motivation for the birth of RE-VIFE.

Now that we obtained the analytical form of F_θ based merely on samples from $U(x,y)$ (rather than the mathematical form of $U(x,y)$), we can manage to locate free-energy minima and cluster samples. We exploited Nelder-Mead algorithm [60] to minimize the FG samples over F_θ, and observed that all the samples finally fall into three distinct local minima, shown by different symbols in Fig. 2C. We also colored FG samples according to their final minimizers, as shown in Fig. 2D, and found that the noisy samples were indeed assigned to different metastable states quite reasonably. This result implies important potential application of VIFE in identifying free-energy minima (or metastable states) and clustering noisy high-dimensional samples, which is demanded by many mechanism analysis and kinetic modeling methods.

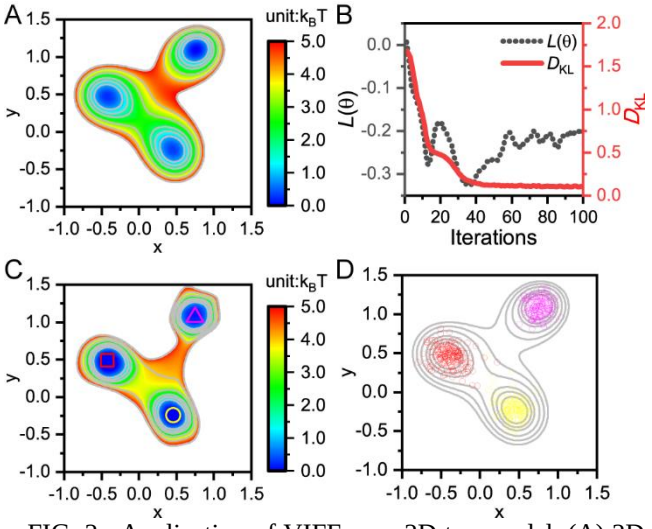

FIG. 2. Application of VIFE on a 2D toy model. (A) 2D potential energy surface (PES), $U(x,y)$ of the toy model. (B) Evolution of effective loss function $L(\theta)$ defined in Eq. (5) (black dotted line) and KL-divergence $D_{\mathrm{KL}}(p_{\mathrm{FG}}||p_\theta)$ (red solid line) against training iterations by VIFE. (C) The contour map corresponds to the optimized F_θ by VIFE, the three symbols (magenta triangle, red square and yellow circle) represent three free-energy minima found in F_θ. (D) Hollow circles are representative samples produced by Langevin dynamics simulations on $U(x,y)$ colored by which local minimum they are minimized to over F_θ. Colors are in line with the symbols in panel C. The contour map of $U(x,y)$ is shown in grey as background.

Next, we proposed to construct a CG model for a prototypical bio-molecular system, alanine dipeptide (Ala2) in explicit water. Specifically, the backbone torsional angles were chosen to be the coarse-graining variables (Fig. 3A), that is, $\mathbf{s} = (\phi, \varphi)$, and our goal is to infer a reasonable FES, $F_\theta(\mathbf{s})$. However, different from the previous toy model, isomerization of (ϕ, φ) involves relatively high barrier, thus brute-force FC simulations of Ala2 converge too slowly to obtain an accurate estimate of $\langle \nabla_\theta F_\theta(\mathbf{s}) \rangle_{p_{\mathrm{FG}}}$. Therefore, we adopted RE-VIFE to enhance the sampling over p_{FG}. Technically, we simultaneously launched two simulations: one FG all-atom MD simulation under a bias potential $V_\phi(\mathbf{s})$, and a CG Langevin dynamics simulation over $F_\theta(\mathbf{s})$. Both F_θ and V_ϕ are initialized to be zero everywhere. After a period of FG sampling biased by V_ϕ (40 ps for Ala2), we reweighted the yielded FG samples to represent p_{FG}, and optimize F_θ according to Eq. (3). Based on the newly-optimized F_θ, a well-tempered target distribution p_{T} (see SI for more details) is established, w.r.t. which the bias potential V_ϕ was optimized. V_ϕ was then fed into the FG simulations of Ala2 which yields an updated set of samples

representing p_{FG}. This procedure constitutes one iteration of RE-VIFE, and the entire process continued until convergence criteria were met.

We tracked how the energies of FG simulation samples and CG ones distribute on F_θ (Fig. 3B), provided that the similarity (or overlap) between these two distributions is a good indicator of convergence. We trained the FG and CG models by RE-VIFE for 8 ns (or equivalently, 200 iterations), and found that the two distributions overlap well and that both spread for a relatively wide range (which implies no mode-dropping).

Noteworthy, in such a short simulation length, it is impossible for brute-force MD to produce equilibrium samples that cover all important metastable states. To illustrate how RE-VIFE helps enhance the sampling of FG models, we presented the 1-ns trajectories for torsion ϕ and φ produced by vanilla MD in contrast to those produced by simulations biased by V_ϕ in Fig. 3C. It can be seen that isomerization of torsion ϕ is fairly frequent in biased MD but hardly found in vanilla MD. Similarly, rotation of torsion φ is also boosted significantly by V_ϕ.

One may wonder how V_ϕ looks like and why it is able to boost the FG sampling so efficiently. By drawing the contour map of the bias potential (Fig. 3D), we observed that the optimized V_ϕ appears is complementary to the ground-true FES of $\mathbf{s} = (\phi, \varphi)$ (a reference FES of $\mathbf{s}$ was presented in Fig. S2). We also superimposed some randomly selected FG samples produced under V_ϕ over the contour map, demonstrating an excellent coverage over both free-energy minima and transition regions. Therefore, samples from V_ϕ are better representatives of p_{FG} and can be reliably used to optimize the CG models. The final CG model (F_θ) optimized via RE-VIFE, which can be regarded as a variationally approximated FES for Ala2 in explicit water is shown in Fig. 3E. It can be found that F_θ not only captures all known metastable states of Ala2 w.r.t. (ϕ, φ), but also quantitatively agrees well with the reference FES (Fig S2). This example demonstrates that simulations on multiple scales can be bridged by RE-VIFE and that CG models can be reliably inferred even without access to FG samples *a priori*.

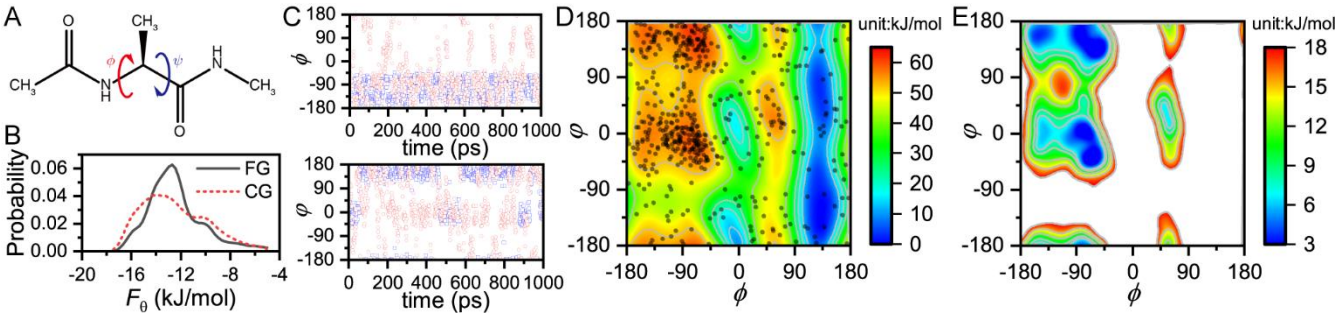

FIG. 3. Application of RE-VIFE on Ala2 in explicit water. (A) Chemical structure of Ala2 and two coarse-graining variables: the torsions ϕ and φ. (B) Distributions of F_θ for all-atom MD simulation samples (black solid line) and for CG simulation samples (red dotted line). (C) 1-ns simulation trajectories projected on torsion ϕ (upper panel) and torsion φ (lower panel). Blue squares correspond to vanilla MD, red circles to MD biased by V_ϕ. (D) The contour map of V_ϕ optimized via RE-VIFE. Superimposed black dots are representative samples produced by the enhanced MD simulation under V_ϕ. (E) Contour map of F_θ optimized via RE-VIFE.

IV. CONCLUDING REMARKS

Despite that all-atom and *ab initio* MD have assisted scientists gain insights over many important physical, chemical and biological processes, their applications to complex systems containing rare events are limited, because experimentally related timescales of such systems (like protein folding and chemical reactions) are well beyond the reachable scope for even the most powerful supercomputers. Two distinct strategies are separately developed to combat this issue: either to perform enhanced sampling over the atomistic model, or to leverage CG models at the cost of losing atomic details. However, traditional enhanced sampling methods can neither scale well to large system sizes, nor transfer well to different system types due to the requirement of system-specific expert knowledge. On the other end, although different attempts exist to build CG models incorporating atom-level knowledge, transitioning from atomistic models to CG models still remains challenging.

In this paper we developed a machine-learning approach, (RE-)VIFE, to connect FG and CG models. In (RE-)VIFE, simulations on different scales can benefit from each other: CG models are optimized w.r.t. the FG simulations hence incorporating information on finer scales; In turn, FG simulations are enhanced under the guidance of CG models. Mathematically, (RE-)VIFE belongs to the realm of unsupervised and reinforcement learning. The variational and self-adaptive training objective allows end-to-end and online training of parametric models like ANNs.

In (RE-)VIFE, CG models can be variationally inferred based merely on equilibrium FG samples, thus involving less artifacts and computational cost than existing methods. This feature allows researchers to fully exploit the available atomistic simulations in order to construct transferrable CG models. More importantly, (RE-)VIFE also allows one to construct CG models even without access to FG samples *a priori*. We remark here that although the CG models obtained this way are able to reproduce the thermodynamic properties of finer-scale models, the dynamics is not necessarily correct. In terms of Langevin dynamics, our CG models only provide a reliable description of the drifting field, but the diffusion field remains unknown. Inferring diffusion field based on available drifting field can be a very interesting direction for further study.

In the presented examples, (RE-)VIFE exclusively compose of one CG-FG cycle over merely two scales, because we assume that sampling over CG models can be readily achieved with simple simulation techniques. While this is true for most cases because CG models are designed for computational tractability, some CG models for very large or complex systems may entail heavy computation and may also suffer from the sampling issues like FG models. In these scenarios, one can build a coarser-grained model (denoted as CG2) over the CG model, and the original CG model becomes relatively finer-grained compared to CG2, the sampling of CG model can thus be boosted by CG2. Therefore, by adding another cycle consisting of CG2-CG on top of the existing CG-FG cycle, one indeed extends (RE-)VIFE to three scales. Following this line, one can naturally develop CG models over a cascade of scales via (RE-)VIFE, and we leave this idea for future research.

ACKNOWLEDGEMENTS

This research was supported by Alexander von Humboldt Foundation.

[1] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications* (Elsevier, 2001), Vol. 1.
[2] M. E. Tuckerman, Journal of Physics: Condensed Matter **14**, R1297 (2002).
[3] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw, Science **334**, 517 (2011).
[4] G. A. Voth, *Coarse-graining of condensed phase and biomolecular systems* (CRC press, 2008).
[5] G. Fiorin, M. L. Klein, and J. Hénin, Molecular Physics **111**, 3345 (2013).
[6] W. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, The Journal of chemical physics **128**, 244114 (2008).
[7] G. M. Torrie and J. P. Valleau, Journal of Computational Physics **23**, 187 (1977).
[8] J. G. Kirkwood, The Journal of Chemical Physics **3**, 300 (1935).
[9] B. W. Silverman, *Density estimation for statistics and data analysis* (Routledge, 2018).
[10] S. J. Sheather and M. C. Jones, Journal of the Royal Statistical Society: Series B (Methodological) **53**, 683 (1991).
[11] A. Laio and M. Parrinello, Proceedings of the National Academy of Sciences **99**, 12562 (2002).
[12] I. A. Basheer and M. Hajmeer, Journal of microbiological methods **43**, 3 (2000).
[13] Y. LeCun, Y. Bengio, and G. Hinton, nature **521**, 436 (2015).
[14] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning* (MIT press, 2016).
[15] E. Schneider, L. Dai, R. Q. Topper, C. Drechsel-Grau, and M. E. Tuckerman, Physical review letters **119**, 150601 (2017).
[16] L. Zhang, H. Wang, and W. E, The Journal of chemical physics **148**, 124113 (2018).
[17] J. Wang, S. Olsson, C. Wehmeyer, A. Pérez, N. E. Charron, G. De Fabritiis, F. Noé, and C. Clementi, ACS central science (2019).
[18] V. N. Vapnik, IEEE transactions on neural networks **10**, 988 (1999).
[19] D. P. Kingma and M. Welling, arXiv preprint arXiv:1312.6114 (2013).
[20] A. v. d. Oord, N. Kalchbrenner, and K. Kavukcuoglu, arXiv preprint arXiv:1601.06759 (2016).
[21] A. Van den Oord, N. Kalchbrenner, L. Espeholt, O. Vinyals, and A. Graves, in *Advances in neural information processing systems*2016), pp. 4790.
[22] L. Dinh, D. Krueger, and Y. Bengio, arXiv preprint arXiv:1410.8516 (2014).
[23] D. J. Rezende and S. Mohamed, arXiv preprint arXiv:1505.05770 (2015).
[24] Y. Du and I. Mordatch, arXiv preprint arXiv:1903.08689 (2019).
[25] P. Dayan, G. E. Hinton, R. M. Neal, and R. S. Zemel, Neural computation **7**, 889 (1995).
[26] R. Salakhutdinov, A. Mnih, and G. Hinton, in *Proceedings of the 24th international conference on Machine learning* (ACM, 2007), pp. 791.
[27] A. Mnih and G. Hinton, in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* (IEEE, 2005), pp. 1302.
[28] R. Salakhutdinov and G. Hinton, in *Artificial intelligence and statistics*2009), pp. 448.
[29] J. J. Hopfield, Proceedings of the national academy of sciences **79**, 2554 (1982).
[30] D. J. MacKay and D. J. Mac Kay, *Information theory, inference and learning algorithms* (Cambridge university press, 2003).
[31] D. Wales, *Energy landscapes: Applications to clusters, biomolecules and glasses* (Cambridge University Press, 2003).

[32] S. Kullback and R. A. Leibler, The annals of mathematical statistics **22**, 79 (1951).
[33] M. Arjovsky, S. Chintala, and L. Bottou, in *International conference on machine learning*2017), pp. 214.
[34] J. Wu, Z. Huang, J. Thoma, D. Acharya, and L. Van Gool, in *Proceedings of the European Conference on Computer Vision (ECCV)*2018), pp. 653.
[35] M. S. Shell, The Journal of chemical physics **129**, 144108 (2008).
[36] A. Chaimovich and M. S. Shell, The Journal of chemical physics **134**, 094112 (2011).
[37] O. Valsson and M. Parrinello, Physical review letters **113**, 090601 (2014).
[38] H. B. Callen, (AAPT, 1998).
[39] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, Journal of the American Statistical Association **112**, 859 (2017).
[40] D. Wu, L. Wang, and P. Zhang, Physical Review Letters **122**, 080602 (2019).
[41] J. Zhang, Y. I. Yang, and F. Noé, (2019).
[42] C. Abrams and G. Bussi, Entropy **16**, 163 (2014).
[43] Y. Okamoto, Journal of Molecular Graphics and Modelling **22**, 425 (2004).
[44] R. C. Bernardi, M. C. Melo, and K. Schulten, Biochimica et Biophysica Acta (BBA)-General Subjects **1850**, 872 (2015).
[45] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, and D. Hassabis, Trends in cognitive sciences (2019).
[46] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction* (MIT press, 2018).
[47] R. J. Williams, Machine learning **8**, 229 (1992).
[48] S. Ross, G. Gordon, and D. Bagnell, in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*2011), pp. 627.
[49] T. Anthony, Z. Tian, and D. Barber, in *Advances in Neural Information Processing Systems*2017), pp. 5360.
[50] I. Grondman, L. Busoniu, G. A. Lopes, and R. Babuska, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **42**, 1291 (2012).
[51] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, in *International conference on machine learning*2016), pp. 1928.
[52] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, in *Advances in neural information processing systems*2014), pp. 2672.
[53] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, in *Advances in Neural Information Processing Systems*2017), pp. 6626.
[54] D. Silver *et al.*, Nature **550**, 354 (2017).
[55] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, arXiv preprint arXiv:1802.05957 (2018).
[56] J. Behler and M. Parrinello, Physical review letters **98**, 146401 (2007).
[57] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, The Journal of Chemical Physics **148**, 241722 (2018).
[58] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, Physical review letters **120**, 143001 (2018).
[59] P. Tiwary and B. Berne, The Journal of chemical physics **147**, 152701 (2017).
[60] S. Singer and J. Nelder, Scholarpedia **4**, 2928 (2009).