

Rigorous progress in coarse-graining

W. G. Noid,¹ R. J. Szukalo,¹ K. M. Kidder,¹
and M. C. Lesniewski,¹

¹Department of Chemistry, The Pennsylvania State University, University Park,
PA USA 16802; email: wnoid@chem.psu.edu

Xxxx. Xxx. Xxx. Xxx. 2024. AA:1–26

[https://doi.org/10.1146/\(\(please add article doi\)\)](https://doi.org/10.1146/((please add article doi)))

Copyright © 2024 by the author(s).
All rights reserved

Keywords

coarse-graining, effective potentials, soft materials, transferability, representability, statistical mechanics, fluctuations, machine learning

Abstract

Low resolution coarse-grained (CG) models provide remarkable computational and conceptual advantages for simulating soft materials. In principle, bottom-up CG models can reproduce all structural and thermodynamic properties of atomically detailed models that can be observed at the resolution of the CG model. This review discusses recent progress in developing theory and computational methods for achieving this promise. We first briefly review variational approaches for parameterizing interaction potentials and their relationship to machine learning methods. We then discuss recent approaches for simultaneously improving both the transferability and thermodynamic properties of bottom-up models by rigorously addressing the density- and temperature-dependence of these potentials. We also briefly discuss exciting progress in modeling high resolution observables with low-resolution CG models. More generally, we highlight the essential role of the bottom-up framework not only for fundamentally understanding the limitations of prior CG models, but also for developing robust computational methods that resolve these limitations in practice.

coarse-grained (CG): a lower resolution model that represents a system in coarse detail

1. Introduction

By judiciously representing systems in reduced detail, coarse-grained (CG) models enable efficient simulations on length- and time-scales that cannot be effectively addressed with atomically detailed models (1, 2). Conversely, by retaining key molecular features, CG models provide insight into fluctuations and interactions that are hidden from mean field and continuum models (3). Moreover, CG models provide profound conceptual advantages by focusing effort on the essential aspects of physical phenomena (4). Consequently, CG models continue to enjoy rapidly growing popularity for studying soft materials (5).

This review addresses “bottom-up” approaches that parameterize CG models based upon information from higher resolution models (6). We first briefly discuss variational approaches for parameterizing interaction potentials. We then focus on recent theoretical and computational advances for improving the accuracy and transferability of bottom-up models. In particular, we emphasize the density- and temperature-dependence of effective potentials and their impact upon the transferability and thermodynamic properties of CG models. Readers are referred to a number of excellent recent reviews for a much broader perspective on bottom-up coarse-graining approaches (7, 8, 9, 10, 11, 12, 13, 14, 15).

all-atom (AA): an atomically detailed high-resolution model

2. The bottom

Bottom-up models rely on an accurate high resolution model, which we shall refer to as an all-atom (AA) model. We denote the AA potential by $u(\mathbf{r}) \equiv u(\mathbf{r}; V)$, where $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_n)$ denotes the configuration for n atoms and V is the volume. In the canonical ensemble at a given volume and temperature, T , the AA model samples configurations according to

$$p_{\mathbf{r}}(\mathbf{r}) = \exp[-\beta u(\mathbf{r})]/z \quad 1.$$

$$z = \int_{V^n} d\mathbf{r} \exp[-\beta u(\mathbf{r})], \quad 2.$$

where $z = z(V, T)$ is the canonical configuration integral and $\beta = 1/k_B T$ (16). The configuration integral defines the excess (Helmholtz) free energy of the AA model:

$$a_{\text{xs}} \equiv a_{\text{xs}}(V, T) = -k_B T \ln V^{-n} z(V, T) = \bar{u} - T s_{\text{xs}}. \quad 3.$$

Here $\bar{u} \equiv \bar{u}(T, V) \equiv \langle u(\mathbf{r}) \rangle$ is the mean AA potential and the angular brackets denote an average according to $p_{\mathbf{r}}(\mathbf{r})$, while s_{xs} is the excess configurational entropy

$$s_{\text{xs}} \equiv s_{\text{xs}}(T, V) \equiv -k_B \int_{V^n} d\mathbf{r} p_{\mathbf{r}}(\mathbf{r}) \ln [p_{\mathbf{r}}(\mathbf{r})/q_{\mathbf{r}}(\mathbf{r})] \leq 0, \quad 4.$$

Kullback-Leibler (KL) divergence: a metric for quantifying the difference between probability densities. See associated sidebar.

where $q_{\mathbf{r}}(\mathbf{r}) = V^{-n}$ denotes the uniform distribution for n non-interacting atoms with a potential, $u_{\text{id}}(\mathbf{r}) = 0$. Eq 4 expresses s_{xs} as the Kullback-Leibler (KL) divergence (see KL divergence sidebar) between $p_{\mathbf{r}}$ and $q_{\mathbf{r}}$ (17). The AA free energy varies according to

$$da_{\text{xs}} = -\bar{p}_{\text{xs}}^{\text{th}} dV - s_{\text{xs}} dT, \quad 5.$$

where $\bar{p}_{\text{xs}}^{\text{th}} \equiv \bar{p}_{\text{xs}}^{\text{th}}(V, T) \equiv \langle p_{\text{xs}}(\mathbf{r}, V) \rangle$ is the thermodynamic excess pressure and

$$p_{\text{xs}}(\mathbf{r}, V) \equiv \frac{1}{3V} \sum_{i=1}^n \mathbf{f}_i \cdot \mathbf{r}_i - \left(\frac{\partial u}{\partial V} \right)_{\mathbf{r}}. \quad 6.$$

The first term in Eq. 6 gives the virial contribution in terms of the AA forces, $\mathbf{f}_i = -\partial u / \partial \mathbf{r}_i$. The second term accounts for any explicit density-dependence in the AA potential. The instantaneous internal pressure of the AA model is then

$$p_{\text{int}}(\mathbf{r}, V) = nk_B T / V + p_{\text{xs}}(\mathbf{r}, V). \quad 7.$$

KL divergence

Kullback and Leibler (KL) interpreted $\phi(x) \equiv \ln[p_1(x)/p_2(x)]$ as the information available at x for distinguishing two probability distributions, $p_1(x)$ and $p_2(x)$. (17) The KL divergence, or relative entropy,

$$D[p_1||p_2] = \int dx p_1(x) \ln[p_1(x)/p_2(x)], \quad \text{SB1}$$

is defined as the average of $\phi(x)$ with respect to $p_1(x)$. The Gibbs inequality implies that $D[p_1||p_2] \geq 0$ and only vanishes if $p_1 = p_2$. (18) Thus, $D[p_1||p_2]$ provides an important metric for quantifying the difference between p_1 and p_2 . If p_i is the canonical distribution, $p_i(x) = V^{-n} \exp[\beta a_i - \beta u_i(x)]$, then $D[p_1||p_2] = \beta(a_1 - a_2) - \beta \langle u_1(x) - u_2(x) \rangle_1$, where the average is performed with respect to p_1 . (19) Sanov's theorem implies another important property of the KL divergence. (20, 21) Suppose one obtains n_s statistically independent samples, $\{x_1, \dots, x_{n_s}\}$, from a probability distribution, $p_s(x)$. These n_s samples determine the empirical probability distribution, $\hat{p}(x) \equiv n_s^{-1} \sum_{i=1}^{n_s} \delta(x - x_i)$. Sanov's theorem implies that, as $n_s \rightarrow \infty$, the probability, or likelihood, of observing an empirical distribution, $\hat{p}(x)$, that deviates from the sampling distribution, $p_s(x)$, exponentially decays with n_s at a rate determined by $D[\hat{p}||p_s]$:

$$\mathbb{P}[\hat{p}] \approx \exp\{-n_s D[\hat{p}||p_s]\}. \quad \text{SB2}$$

In the context of parameterizing generative probabilistic models, one seeks to optimize the parameters, θ , for a model with a probability distribution, $p_M(x; \theta)$, such that it reproduces the “ground truth” target distribution, $p_T(x)$. In this case, one wishes to maximize the probability that sampling from the model, $p_s(x) = p_M(x; \theta)$, generates an empirical probability that matches the ground truth: $\hat{p}(x) = p_T(x)$. Thus, in this maximum likelihood approach, one should optimize θ by minimizing the KL divergence $D[p_T||p_M(\theta)]$.

3. Coarse-grained representation

Bottom-up approaches define the CG representation for a particular system by a mapping, $\mathbf{M}(\mathbf{r})$, that maps each AA configuration, $\mathbf{r}$, to a CG configuration, $\mathbf{R} = \mathbf{M}(\mathbf{r})$. This mapping usually determines the coordinates, $\mathbf{R}_I$, for each site, $I = 1, \dots, N$, as a linear combination of atomic coordinates with constant coefficients:

$$\mathbf{R}_I = \sum_{i=1}^n c_{Ii} \mathbf{r}_i. \quad 8.$$

If all of the atoms in a site are displaced by a vector, $\mathbf{v}$, then the site should also be displaced by the same vector. Consequently, the coefficients must be normalized $\sum_{i=1}^n c_{Ii} = 1$ for each site I (22). The mapping also defines a subensemble of AA configurations, $\mathbf{M}^{-1}(\mathbf{R}) =$

$\{\mathbf{r}|\mathbf{M}(\mathbf{r}) = \mathbf{R}\}$, that map to $\mathbf{R}$. We refer readers to several reviews that discuss recent insights into the CG mapping (8, 10, 23).

Bottom-up approaches typically focus on reproducing properties of the AA model at the resolution of the CG model, i.e., the properties in the mapped ensemble (23). We define the restricted configuration integral,

$$z_{\mathbf{R}}(\mathbf{R}) \equiv z_{\mathbf{R}}(\mathbf{R}; V, T) \equiv \int_{V^n} d\mathbf{r} \exp[-\beta u(\mathbf{r})] \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}), \quad 9.$$

as the Boltzmann weight for the subensemble $\mathbf{M}^{-1}(\mathbf{R})$. The mapped probability (density),

$$p_{\mathbf{R}}(\mathbf{R}) \equiv \int_{V^n} d\mathbf{r} p_{\mathbf{r}}(\mathbf{r}) \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) = z_{\mathbf{R}}(\mathbf{R})/z, \quad 10.$$

gives the probability that the AA model samples $\mathbf{M}^{-1}(\mathbf{R})$. The conditioned probability,

$$p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}) \equiv p_{\mathbf{r}}(\mathbf{r}) \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) / p_{\mathbf{R}}(\mathbf{R}) = e^{-\beta u(\mathbf{r})} \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) / z_{\mathbf{R}}(\mathbf{R}), \quad 11.$$

describes the distribution of AA configurations in $\mathbf{M}^{-1}(\mathbf{R})$.

The mapping defines the “exact” potential (24, 25, 26, 27) for bottom-up CG models:

$$W(\mathbf{R}) \equiv W(\mathbf{R}; V, T) \equiv -k_B T \ln V^{N-n} z_{\mathbf{R}}(\mathbf{R}; V, T). \quad 12.$$

Because $p_{\mathbf{R}}(\mathbf{R}) \propto e^{-\beta W(\mathbf{R})}$, CG simulations that employ $W(\mathbf{R})$ as an effective potential will perfectly reproduce the mapped probability distribution. The analogy between Eqs. 3 and 12 implies that $W(\mathbf{R})$ is the excess (Helmholtz) free energy of the AA model when it is viewed at the resolution of the CG representation (10, 3). Consequently, it follows that

$$W(\mathbf{R}) = E_W(\mathbf{R}) - TS_W(\mathbf{R}), \quad 13.$$

where

$$E_W(\mathbf{R}) \equiv \langle u(\mathbf{r}) \rangle_{\mathbf{R}} \quad 14.$$

and the subscripted brackets denote an average according to $p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R})$. The entropic contribution to W ,

$$S_W(\mathbf{R}) \equiv -k_B \int_{V^n} d\mathbf{r} p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}) \ln [p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}) / q_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R})] \leq 0, \quad 15.$$

is the excess configurational entropy associated with $\mathbf{M}^{-1}(\mathbf{R})$ (28). Here $q_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}) = V^{N-n} \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R})$ is the ideal conditioned distribution, such that S_W is ($-k_B$ times) the KL divergence between $p_{\mathbf{r}|\mathbf{R}}$ and $q_{\mathbf{r}|\mathbf{R}}$. The analogy between Eqs. 3 and 12 also implies (29)

$$dW = - \sum_{I=1}^N \bar{\mathbf{f}}_I \cdot (d\mathbf{R}_I)_V - \bar{p}_{\text{xs}} dV - S_W dT. \quad 16.$$

Here $(d\mathbf{R}_I)_V$ denotes variations in configuration at constant volume, i.e., $(d\mathbf{R}_I)_V = V^{1/3} d\hat{\mathbf{R}}_I$ where $\hat{\mathbf{R}}_I = \mathbf{R}_I / V^{1/3}$, and

$$\bar{\mathbf{f}}_I \equiv \bar{\mathbf{f}}_I(\mathbf{R}; V, T) \equiv \langle \mathbf{f}_I(\mathbf{r}) \rangle_{\mathbf{R}} \quad 17.$$

is the conditioned average of the AA net force, $\mathbf{f}_I(\mathbf{r}) = \sum_{i \in I} \mathbf{f}_i(\mathbf{r})$, on site I in the AA configuration $\mathbf{r}$. (30, 31, 22) Because the gradients of W are (conditioned) mean forces, W

is commonly referred to as the many-body potential of mean force (PMF) (24). Similarly, in Eq. 16, $\bar{p}_{\text{xs}}$ is the mean excess pressure of the AA model at the CG resolution, i.e.,

$$\bar{p}_{\text{xs}} \equiv \bar{p}_{\text{xs}}(\mathbf{R}; V, T) \equiv \langle p_{\text{xs}}(\mathbf{r}; V) \rangle_{\mathbf{R}} = \frac{1}{3V} \sum_{I=1}^N \bar{\mathbf{f}}_I \cdot \mathbf{R}_I - \left(\frac{\partial W}{\partial V} \right)_{\mathbf{R}}, \quad 18.$$

which reflects the mean-force virial and also the explicit density-dependence of the PMF. Consequently, if $W(\mathbf{R}; V, T)$ is known as a function of both configuration and thermodynamic conditions, then the CG model should be able to perfectly reproduce all structural and thermodynamic properties of the AA model at the resolution of the CG model.

potential of mean force (PMF): a potential from which average forces can be calculated as gradients. In this work, PMF refers to $W(\mathbf{R})$.

4. Configuration-dependence and structure

4.1. Approximate interaction potentials

In practice, Eq. 9 cannot be evaluated and the PMF must be approximated by a simpler interaction potential, U . In many cases, U may be expressed

$$U(\mathbf{R}) = \sum_{\zeta} \sum_{\lambda} U_{\zeta}(\psi_{\zeta}(\mathbf{R}_{\lambda})). \quad 19.$$

Here ζ indicates a particular type of interaction, U_{ζ} indicates a corresponding potential that depends upon a scalar variable ψ_{ζ} that is a function of the coordinates, $\mathbf{R}_{\lambda}$, for a particular set, λ , of CG sites. For instance, if $\zeta = 2$ indicates an intermolecular pair interaction, $U_{\zeta} = U_2$ is a pair potential, $\mathbf{R}_{\lambda} = \{\mathbf{R}_I, \mathbf{R}_J\}$ is the coordinates for the pair, $\lambda = \{I, J\}$, and $\psi_{\zeta}(\mathbf{R}_{\lambda}) = R_{IJ}(\mathbf{R}) = |\mathbf{R}_I - \mathbf{R}_J|$ is the distance between the pair in configuration, $\mathbf{R}$. The approximate potential determines the equilibrium distribution for the CG model:

$$P_{\mathbf{R}}(\mathbf{R}; U) \equiv \exp[-\beta U(\mathbf{R})]/Z[U], \quad 20.$$

where $Z[U] \equiv Z[V, T; U] \equiv \int_{V^N} d\mathbf{R} \exp[-\beta U(\mathbf{R})]$.

4.2. Relative entropy

Shell and coworkers pioneered a relative entropy (RE) variational principle (19, 32, 33) for optimizing the approximate interaction potential, U , by minimizing

$$S_{\text{rel}}[U] \equiv \int_{V^N} d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) \ln \left[\frac{p_{\mathbf{R}}(\mathbf{R})}{P_{\mathbf{R}}(\mathbf{R}; U)} \right] \geq S_{\text{rel}}[W] = 0. \quad 21.$$

Thus, the exact PMF is the global minimizer of $S_{\text{rel}}[U]$. More generally, given a space of approximate potentials, the minimizing potential, $U_{*} = \text{argmin}_U S_{\text{rel}}[U]$, corresponds to the equilibrium distribution, $P_{\mathbf{R}}(\mathbf{R}; U_{*})$, that is “closest” (as measured by the KL divergence, $D[p_{\mathbf{R}}||P_{\mathbf{R}}(U)]$) to the mapped distribution, $p_{\mathbf{R}}(\mathbf{R})$.

In practice, one typically minimizes $S_{\text{rel}}[U]$ with respect to the interaction potentials, $U_{\zeta}(x)$, that are included in Eq 19. It is convenient to re-express Eq 19 (34, 35)

$$U(\mathbf{R}) = \sum_{\zeta} \int dz U_{\zeta}(z) \hat{\rho}_{\zeta}(z; \mathbf{R}) \quad 22.$$

$$\hat{\rho}_{\zeta}(z; \mathbf{R}) \equiv \sum_{\lambda} \delta(\psi_{\zeta}(\mathbf{R}_{\lambda}) - z). \quad 23.$$

relative entropy (RE): another common name for the KL divergence (20).

radial distribution function (RDF): a correlation function, $g(r)$, describing the spatial density around particles that has been normalized with respect to the bulk density (16).

Inverse Monte Carlo (IMC): an early structure-based approach introduced in (34) for parameterizing CG models to accurately reproduce structural observables, including the RDF.

Iterative Boltzmann Inversion (IBI): a structure-based approach introduced in (36) for parameterizing CG models to accurately reproduce the mapped RDF.

Force-matching (FM): a variational approach for parameterizing approximate potentials to match gradients of a target potential.

Multiscale coarse-graining (MS-CG): a bottom-up approach (39, 40) that employs the FM variational principle to parameterize the approximate CG potential.

For each $U_\zeta(z)$, $\hat{p}_\zeta(z; \mathbf{R})$ serves as a corresponding conjugate operator that quantifies the number of times that $U_\zeta(z)$ contributes to $U(\mathbf{R})$. In this case,

$$\frac{\delta S_{\text{rel}}[U]}{\delta U_\zeta(z)} = \beta N_\zeta \{p_\zeta(z) - P_\zeta(z; U)\}, \quad 24.$$

where N_ζ is the number of instances of the ζ interaction in the model, while $p_\zeta(z)$ and $P_\zeta(z; U)$ give the probability (densities) that $\psi_\zeta = z$ in the mapped ensemble and in the equilibrium ensemble for the approximate CG model, respectively:

$$p_\zeta(z) \equiv \frac{1}{N_\zeta} \int_{V^N} d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) \hat{p}_\zeta(z; \mathbf{R}) \quad 25.$$

$$P_\zeta(z; U) \equiv \frac{1}{N_\zeta} \int_{V^N} d\mathbf{R} P_{\mathbf{R}}(\mathbf{R}; U) \hat{p}_\zeta(z; \mathbf{R}). \quad 26.$$

According to Eq 24, the relative entropy is minimized with respect to $U_\zeta(z)$ when $P_\zeta(z; U) = p_\zeta(z)$, i.e., when simulations with the CG model reproduce the mapped probability distribution for the corresponding degree of freedom. In particular, if $U_\zeta(z) = U_2(r)$ corresponds to a non-bonded pair potential, then $p_\zeta(r)/r^2$ and $P_\zeta(r)/r^2$ are proportional to the corresponding radial distribution function (RDF) in the mapped ensemble and in the simulated CG ensemble, respectively. Therefore, $S_{\text{rel}}[U]$ provides a variational framework for deriving and generalizing structure-based methods, (37, 32, 35) such as Iterative Boltzmann Inversion (IBI) (36) or the Inverse Monte Carlo (IMC) method (34), that optimize interaction potentials to reproduce conjugate structural correlation functions. In practice, $P_\zeta(z; U)$ cannot be determined without explicitly simulating the CG potential, U . Consequently, minimizing $S_{\text{rel}}[U]$ generally requires simulating a sequence of approximate potentials until the CG model reproduces the relevant mapped distributions (38).

4.3. Force-matching

Voth and coworkers (39, 40, 41) pioneered the multiscale coarse-graining (MS-CG) force-matching (FM) variational principle for optimizing U by minimizing

$$\chi^2[U] \equiv \left\langle \frac{1}{3N} \sum_{I=1}^N |\mathbf{f}_I(\mathbf{r}) - \mathbf{F}_{U;I}(\mathbf{M}(\mathbf{r}))|^2 \right\rangle. \quad 27.$$

Here $\mathbf{f}_I(\mathbf{r}) = \sum_{i \in I} \mathbf{f}_i(\mathbf{r})$ is the net AA force on site I , while $\mathbf{F}_{U;I}(\mathbf{R}) = -\nabla_I U(\mathbf{R})$ is the force specified by $U(\mathbf{R})$ and $\nabla_I = \partial/\partial \mathbf{R}_I$. The sidebar on least squares and conditioned means implies (22) that

$$\chi^2[U] = \chi^2[W] + \int_{V^N} d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) \frac{1}{3N} \sum_{I=1}^N |\bar{\mathbf{f}}_I(\mathbf{R}) - \mathbf{F}_{U;I}(\mathbf{R})|^2 \geq \chi^2[W], \quad 28.$$

where $\bar{\mathbf{f}}_I(\mathbf{R}) = -\nabla_I W(\mathbf{R})$ is the many-body mean force. Thus, the exact many-body PMF is the global minimizer of $\chi^2[U]$. More generally, given a space of approximate potentials, the minimizing MS-CG potential optimally approximates (in a least squares sense) the gradients and, thus, the configuration-dependence of the exact PMF (22).

Equation 28 has a beautiful geometric interpretation in a vector space of force fields. Each element of this vector space specifies a configuration-dependent, vector-valued force

Least squares and conditioned means

Let x denote a microscopic random variable with probability (density), $p(x)$. Let $f(x)$ be a microscopic observable with mean $\mu = \int dx p(x)f(x)$ and variance $\sigma^2 = \int dx p(x)|f(x) - \mu|^2$. For any constant c ,

$$\chi^2(c) \equiv \langle |f(x) - c|^2 \rangle = \sigma^2 + |\mu - c|^2 \quad \text{SB3.}$$

In Eq. SB3 the angular brackets denote an average with respect to $p(x)$ and the second equality follows from the simple identity $\langle f(x) - \mu \rangle = 0$. Because σ^2 is independent of c , we see that the mean, μ , minimizes χ^2 .

Now let us define a mapping, $M(x) : x \rightarrow X = M(x)$, that determines a “coarse-grained” description of x . This determines a marginal probability, $P(X) \equiv \int dx p(x)\delta(M(x) - X)$, and a conditioned probability, $p(x|X) \equiv p(x)\delta(M(x) - X)/P(X)$. For any CG function, $C(X)$, the reasoning that led to Eq. SB3 also implies

$$\langle |f(x) - C(M(x))|^2 \rangle_X = \tilde{\sigma}^2(X) + |\tilde{\mu}(X) - C(X)|^2. \quad \text{SB4.}$$

Here the subscripted angular brackets denote the conditioned average according to $p(x|X)$, $\tilde{\mu}(X) \equiv \langle f(x) \rangle_X$ is the conditioned mean, and $\tilde{\sigma}^2(X) \equiv \langle |f(x) - \tilde{\mu}(X)|^2 \rangle_X$ is the conditioned variance quantifying microscopic fluctuations in $f(x)$ over the subensemble of microstates, x , that map to the CG state, X . By averaging Eq. SB4 with respect to $P(X)$, we see that

$$\chi^2[C] \equiv \langle |f(x) - C(M(x))|^2 \rangle = \chi^2[\tilde{\mu}] + \int dX P(X) |\tilde{\mu}(X) - C(X)|^2 \geq \chi^2[\tilde{\mu}] \quad \text{SB5.}$$

Thus, among all CG functions $C(X)$, the conditioned mean, $\tilde{\mu}(X) \equiv \langle f(x) \rangle_X$, provides the optimal fit to the microscopic function, $f(x)$ (42). Moreover, Eq. SB5 may be interpreted as a statement of Pythagorean’s theorem for the distance between a fluctuating microscopic function, $f(x)$, and an arbitrary CG function, $C(X)$. Consequently, if one considers a particular subspace of CG functions, then the function within this subspace that minimizes $\chi^2[C]$ also is closest to the conditioned mean, $\tilde{\mu}(X)$.

on each CG site, $I = 1, \dots, N$. The AA potential, $u(\mathbf{r})$, specifies one particular force field, $\mathbf{f}_u \equiv \{\mathbf{f}_I(\mathbf{r})\}$. Within this space, there exists a vector subspace of CG force fields that depend only upon the CG representation of the AA configuration, $\mathbf{F} \equiv \{\mathbf{F}_I(\mathbf{M}(\mathbf{r}))\}$. The PMF, $W(\mathbf{R})$, specifies the CG force field, $\mathbf{F}_W \equiv \{\bar{\mathbf{f}}_I(\mathbf{M}(\mathbf{r}))\}$, that is defined by the mean forces, $\bar{\mathbf{f}}_I(\mathbf{R})$.¹ Given two force fields, $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$, we define an inner product $\odot$ (43, 44)

$$\mathbf{f}^{(1)} \odot \mathbf{f}^{(2)} \equiv \frac{1}{3N} \int_{V^n} d\mathbf{r} p_r(\mathbf{r}) \sum_{I=1}^N \mathbf{f}_I^{(1)}(\mathbf{r}) \cdot \mathbf{f}_I^{(2)}(\mathbf{r}). \quad 29.$$

In the case of two CG force fields, $\mathbf{F}^{(1)}$ and $\mathbf{F}^{(2)}$, Eq 29 simplifies to an average over the

¹From the perspective of the CG configuration space, $\mathbf{f}_I(\mathbf{r})$, is a fluctuating random function, while $\bar{\mathbf{f}}_I(\mathbf{R})$ is its deterministic conditioned mean. Consequently, $\bar{\mathbf{f}}_I(\mathbf{R})$ may be considered the projection of $\mathbf{f}_I(\mathbf{r})$ into the CG configuration space. Accordingly, $\mathbf{F}_W$ may be considered the projection of $\mathbf{f}_u$ into the subspace of CG force fields (42).

mapped ensemble

$$\mathbf{F}^{(1)} \odot \mathbf{F}^{(2)} = \frac{1}{3N} \int_{V^N} d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) \sum_{I=1}^N \mathbf{F}_I^{(1)}(\mathbf{R}) \cdot \mathbf{F}_I^{(2)}(\mathbf{R}). \quad 30.$$

This inner product then defines a norm for the space of force fields, $\|\mathbf{f}\|^2 \equiv \mathbf{f} \odot \mathbf{f}$.

Given these definitions, we see that $\chi^2[U] = \|\mathbf{f}_u - \mathbf{F}_U\|^2$ and $\chi^2[W] = \|\mathbf{f}_u - \mathbf{F}_W\|^2$, while the last term in Eq. 28 is $\|\mathbf{F}_W - \mathbf{F}_U\|^2$. Consequently, Eq. 28 is a statement of the Pythagorean theorem in the space of force fields (22, 45, 30)

$$\|\mathbf{f}_u - \mathbf{F}_U\|^2 = \|\mathbf{f}_u - \mathbf{F}_W\|^2 + \|\mathbf{F}_W - \mathbf{F}_U\|^2. \quad 31.$$

Because $\|\mathbf{F}_W - \mathbf{F}_U\|^2 \geq 0$ and only vanishes when $\mathbf{F}_W = \mathbf{F}_U$, Eq 31 implies that the mean force field, $\mathbf{F}_W$, is the CG force field closest to the AA force field, $\mathbf{f}_u$. Accordingly, $\mathbf{F}_W$ may be considered the geometric projection of $\mathbf{f}_u$ into the subspace of CG force fields. By minimizing $\chi^2[U] = \|\mathbf{f}_u - \mathbf{F}_U\|^2$, the FM variational principle determines the approximate CG force field that minimizes the distance $\|\mathbf{F}_W - \mathbf{F}_U\|^2$.

Eq. 19 defines a force field “basis set” for performing the FM variational calculation (22, 46). Specifically, for each interaction potential, $U_\zeta(z)$, that governs a mechanical degree of freedom, ψ_ζ , we define a force function, $F_\zeta(z) \equiv -dU_\zeta(z)/dz$, and a force field vector, $\mathcal{G}_\zeta(z) \equiv \{\mathcal{G}_{\zeta,I}(z; \mathbf{R})\}$, that specifies the direction of these forces:

$$\mathcal{G}_{\zeta,I}(z; \mathbf{R}) \equiv \sum_{\lambda} \left(\frac{\partial \psi_\zeta(\mathbf{R}_\lambda)}{\partial \mathbf{R}_I} \right) \delta(\psi_\zeta(\mathbf{R}_\lambda) - z). \quad 32.$$

The approximate force field, $\mathbf{F}_U$, may then be expressed in analogy to Eq 22

$$\mathbf{F}_U = \sum_{\zeta} \int dz F_\zeta(z) \mathcal{G}_\zeta(z), \quad 33.$$

where $\mathcal{G}_\zeta(z)$ acts as a basis vector and $F_\zeta(z)$ acts as the corresponding coefficient. The FM minimizing condition for the MS-CG force functions (47, 22) may be expressed

$$\sum_{\zeta'} \int dz' G_{\zeta\zeta'}(z, z') F_{\zeta'}(z') = b_\zeta(z), \quad 34.$$

where

$$G_{\zeta\zeta'}(z, z') \equiv \mathcal{G}_\zeta(z) \odot \mathcal{G}_{\zeta'}(z') = \frac{1}{3N} \left\langle \sum_{I=1}^N \mathcal{G}_{\zeta,I}(z; \mathbf{M}(\mathbf{r})) \cdot \mathcal{G}_{\zeta',I}(z'; \mathbf{M}(\mathbf{r})) \right\rangle \quad 35.$$

is a Gram matrix describing the inner product between basis vectors, (48) while

$$b_\zeta(z) \equiv \mathcal{G}_\zeta(z) \odot \mathbf{f}_u = \frac{1}{3N} \left\langle \sum_{I=1}^N \mathbf{f}_I(\mathbf{r}) \cdot \mathcal{G}_{\zeta,I}(z; \mathbf{M}(\mathbf{r})) \right\rangle \quad 36.$$

corresponds to the projection of the AA force field onto the basis vector, $\mathcal{G}_\zeta(z)$ (48). Importantly, $G_{\zeta\zeta'}(z, z')$ and $b_\zeta(z)$ are both expressed as averages over the AA canonical ensemble. Consequently, Eq 34 determines the optimal MS-CG potential directly from an AA simulation without requiring simulations of approximate CG potentials.

Because the many-body mean force field, $\mathbf{F}_W$, is the projection of the AA force field, $\mathbf{f}_u$, into the subspace of CG force fields, Eq 36 implies that $b_\zeta(z) = \mathcal{G}_\zeta(z) \odot \mathbf{F}_W$. Moreover, because the mean forces are related to gradients of the mapped probability distribution, i.e., $\bar{\mathbf{f}}_I(\mathbf{R}) = -\nabla_I W(\mathbf{R}) = k_B T \nabla_I p_R(\mathbf{R})/p_R(\mathbf{R})$, $b_\zeta(z)$ can be directly determined from structural correlations within the mapped ensemble (47, 43, 44). Therefore, Eq 34 allows one to perform the variational FM calculation directly from structural information within the mapped ensemble, i.e., one can perform force-matching without force information (45).

Eq 34 also has an instructive geometric interpretation (48, 45). By employing Eqs 30, 33, and 35, along with $b_\zeta(z) = \mathcal{G}_\zeta(z) \odot \mathbf{F}_W$, Eq. 34 may be re-expressed

$$\mathcal{G}_\zeta(z) \odot \mathbf{F}_U = \mathcal{G}_\zeta(z) \odot \mathbf{F}_W. \quad 37.$$

Thus, the FM variational principle determines the optimal MS-CG potential by geometrically projecting the many-body mean force field, $\mathbf{F}_W$, onto the CG force field basis set that is specified by Eq 19 for the approximate CG potential, U . By matching the projected mean forces - and thus the gradients of the exact many-body PMF - the MS-CG potential approximates the configuration-dependence of W . In fact, the force-balance relation, Eq 37, corresponds to a rigorous generalization of the Yvon-Born-Green (YBG) integral equation (18). The YBG equation has typically been interpreted as an exact relationship between the 2- and 3-body correlation functions of a simple liquid with a known pair potential. However, this generalized-YBG equation not only treats arbitrarily complex potentials, but also provides a direct (i.e., non-iterative) variational approach for solving the inverse problem of deducing the potentials that generated observed structural correlations (47, 43, 44, 45).

However, the force field basis set for physics-based CG models is usually highly incomplete and cannot represent the exact many-body PMF. Moreover, Eq 37 determines the MS-CG force functions to match projected mean forces based upon averages over the mapped distribution, $p_R(\mathbf{R})$, and not the distribution, $P_R(\mathbf{R}; U)$, for the approximate potential (47, 48). Consequently, simulations with physics-based MS-CG potentials are not guaranteed to reproduce any particular structural properties of the mapped ensemble (49, 50).

Recently, the MS-CG FM variational principle has been generalized in several exciting directions. In particular, Voth and coworkers have extended the FM variational principle to treat virtual sites, (51, 52, 53) as well as ultra-CG models in which the state of CG sites dynamically varies with configuration (54, 55, 56). Kalligiannaki et al. have generalized FM to treat nonlinear CG maps, (30) while Huang and Nguyen introduced a torque-matching variational principle to optimize orientation-dependent interactions between anisotropic CG sites (57). Furthermore, Dequidt and coworkers have developed a Bayesian stochastic trajectory matching approach that employs FM with time-averaged forces (58, 59).

4.4. Coarse-graining and generative models

Recently, a growing number of bottom-up models have employed machine learning (ML) tools, such as artificial neural networks (60, 61, 62, 63, 64) and kernel methods, (65, 66, 67) to represent much more flexible interaction potentials that more accurately approximate the many-body PMF. Perhaps more importantly, there exists a fundamental relationship between bottom-up coarse-graining and generative probabilistic ML methods (68). In particular, energy-based ML models determine an energy function, $E(x; \theta)$, such that the model distribution, $p_M(x; \theta) \equiv Z^{-1}(\theta) \exp[-E(x; \theta)]$, accurately approximates the target “ground truth” distribution, $p_T(x)$ (69). This is perfectly analogous to bottom-up methods

Yvon-Born-Green (YBG) equation: an integral equation that relates the potential and equilibrium structural correlations of a liquid (18).

machine-learning (ML): a class of modern computational approaches for modeling data.

parameterizing $U(\mathbf{R})$ such that $P_{\mathbf{R}}(\mathbf{R}; U)$ accurately approximates $p_{\mathbf{R}}(\mathbf{R})$.

As discussed in the KL divergence sidebar, the RE variational principle determines the potential, U , that maximizes the likelihood that the CG model will sample the mapped distribution (19, 35). Consequently, the RE variational principle corresponds to maximum likelihood methods that are commonly adopted to parameterize probabilistic ML models. However, these maximum likelihood methods are limited by the need to treat $Z[U]$. Even though $U(\mathbf{M}(\mathbf{r}))$ can be readily evaluated, $S_{\text{rel}}[U]$ cannot be calculated from the mapped ensemble because $Z[U]$ cannot be evaluated. In practice, minimizing $S_{\text{rel}}[U]$ requires a series of simulations with approximate potentials because $P_{\zeta}(z; U)$ cannot be determined without explicitly simulating U (38). In the case of CG models with computationally expensive ML potentials, this series of simulations can require significant computational resources.

The FM variational principle can also be re-expressed in terms of information theoretic quantities that are employed to parameterize modern generative probabilistic models. Kullback and Leibler interpreted $\Phi(\mathbf{R}; U) \equiv \ln p_{\mathbf{R}}(\mathbf{R})/P_{\mathbf{R}}(\mathbf{R}; U)$ as the information available in CG configuration, $\mathbf{R}$, for distinguishing $p_{\mathbf{R}}$ and $P_{\mathbf{R}}(U)$ (17). The gradients of Φ directly compare the forces due to the PMF and the approximate potential (70)

$$\nabla_I \Phi(\mathbf{R}; U) = \beta [\bar{\mathbf{f}}_I(\mathbf{R}) - \mathbf{F}_{U;I}(\mathbf{R})]. \quad 38.$$

Consequently, Eq. 28 may be expressed

$$\chi^2[U] = \chi^2[W] + \frac{1}{3N\beta^2} \int_{V^N} d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) \sum_{I=1}^N |\nabla_I \Phi(\mathbf{R}; U)|^2. \quad 39.$$

Thus, while the RE variational principle minimizes Φ , the FM variational principle minimizes the gradients of Φ .

The FM variational principle is closely related to a “score-matching” method that parameterizes the ML model distribution, $p_{\mathbf{M}}(x; \theta)$, to match the scores of the target distribution, $p_{\mathbf{T}}(x)$ (71). Given a probability function, $p(x)$, the (Stein) score, $s(x)$, is defined $s(x) \equiv \frac{d}{dx} \ln p(x)$ (72, 73). Importantly, the Stein score is independent of the normalization $Z(\theta)$ and, consequently, can be directly evaluated from x . In the case of equilibrium physical models and energy-based ML models, the Stein score corresponds to a force that drives the system to more likely configurations. By definition, gradients of Φ give the difference in the Stein scores of the mapped distribution, $p_{\mathbf{R}}(\mathbf{R})$, and the approximate distribution, $P_{\mathbf{R}}(\mathbf{R}; U)$ (70). Thus, the FM and score-matching methods both seek to minimize gradients of Φ . By eliminating Z , the FM and score-matching functionals can be minimized directly from the mapped ensemble without requiring multiple, potentially expensive, simulations. Consequently, many studies have employed the FM variational principle to parameterize computationally expensive ML potentials (61, 62, 63, 65, 66, 67, 64).

Eqs 38 and 39 provide important insight for developing CG potentials from mapped ensembles with metastable basins (69). Suppose that the CG configuration space space, $\mathbb{S}$, contains two basins, $\mathbb{S}_1$ and $\mathbb{S}_2$, that are separated by a rarely sampled barrier region, $\mathbb{S}_b$. Let π_{α} be the probability for each region, such that $\pi_1 + \pi_2 + \pi_b = 1$ and $\pi_b \approx 0$. The mapped distribution may then be expressed

$$p_{\mathbf{R}}(\mathbf{R}) = \pi_1 p_1(\mathbf{R}) + \pi_2 p_2(\mathbf{R}) + \pi_b p_b(\mathbf{R}) \quad 40.$$

where $p_{\alpha}(\mathbf{R})$ is normalized over $\mathbb{S}_{\alpha}$, and $p_{\alpha}(\mathbf{R}) = 0$ for $\mathbf{R} \notin \mathbb{S}_{\alpha}$. For each region, $\mathbb{S}_{\alpha}$, we define $W_{\alpha}(\mathbf{R}) = -k_B T \ln[V^N p_{\alpha}(\mathbf{R})]$, which only reflects sampling in $\mathbb{S}_{\alpha}$. We also define

$a_\alpha = a_{\text{xs}} - k_B T \ln \pi_\alpha$, where a_{xs} is the excess free energy of the AA model. The PMF in the two basins is

$$W(\mathbf{R}) = \begin{cases} W_1(\mathbf{R}) + a_1 & \text{if } \mathbf{R} \in \mathbb{S}_1 \\ W_2(\mathbf{R}) + a_2 & \text{if } \mathbf{R} \in \mathbb{S}_2 \end{cases} \quad 41.$$

Note that $\Delta a \equiv a_2 - a_1 = -k_B T \ln \pi_2 / \pi_1$ determines the relative depth of the two basins. Because the RE variational principle considers probabilities, it can estimate Δa based upon the ratio π_2 / π_1 without requiring information from the rarely sampled barrier region. However, the mean forces contain no information about a_1 or a_2 :

$$\bar{\mathbf{f}}(\mathbf{R}) = \begin{cases} -\nabla W_1(\mathbf{R}) & \text{if } \mathbf{R} \in \mathbb{S}_1 \\ -\nabla W_2(\mathbf{R}) & \text{if } \mathbf{R} \in \mathbb{S}_2 \end{cases} \quad 42.$$

Consequently, even in the limit of a complete basis set, the FM variational principle cannot determine Δa without information from $\mathbb{S}_b$.

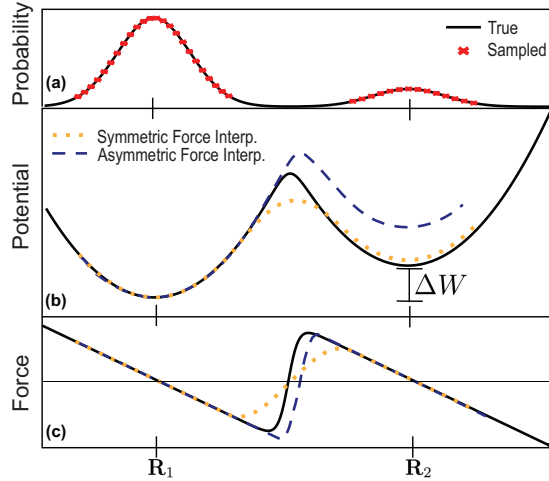


Figure 1

Gedanken of Zavadlav and coworkers illustrating difficulties that may arise when force-matching systems with distinct metastable states (74). From top to bottom, the solid black curves indicate the mapped probability distribution, p_R , the corresponding many-body PMF, $W = -k_B T \ln p_R$, and the corresponding mean force, $\bar{\mathbf{f}} = -\nabla W$. The red points in panel a indicate well-sampled regions of configuration space. The dashed blue and dotted orange curves in panel c present two interpolations of the mean force into the barrier region. The corresponding curves in panel b present the resulting approximations for W .

Figure 1 illustrates these considerations with the gedanken of Zavadlav and coworkers (74). Consider two CG configurations, $\mathbf{R}_1 \in \mathbb{S}_1$ and $\mathbf{R}_2 \in \mathbb{S}_2$, that correspond to the minima of two basins in the mapped ensemble with different depths, $\Delta W \equiv W(\mathbf{R}_2) - W(\mathbf{R}_1)$. If the approximate potential, U , reproduces this difference, then $\Delta U - \Delta W = k_B T \Delta \Phi = 0$. However, the FM variational principle optimizes the CG potential based upon local gradients, $\nabla \Phi$. Consequently, in order for the MS-CG potential to accurately approximate ΔW , it must reproduce the integrated mean force along a path between $\mathbf{R}_1$ and $\mathbf{R}_2$, i.e.,

$$k_B T \Delta \Phi = \Delta U - \Delta W = \int_{\mathbf{R}_1}^{\mathbf{R}_2} \{\bar{\mathbf{f}}(\mathbf{R}) - \mathbf{F}_U(\mathbf{R})\} \cdot d\mathbf{R} \approx 0. \quad 43.$$

This requires that the mapped ensemble contains sufficient information to accurately determine the mean force in the rarely sampled barrier region, $\mathbb{S}_b$, between $\mathbb{S}_1$ and $\mathbb{S}_2$. Consequently, while it is computationally efficient, the FM variational principle may become relatively data expensive for complex systems with multiple basins.

The flow-matching approach of Clementi, Noé, and coworkers appears a promising approach for combining the data efficiency of the RE variational principle with the computational efficiency of the FM variational principle (75). The flow-matching approach employs the RE variational principle to parameterize a normalizing flow (76, 77) that can be efficiently sampled to generate data for parameterizing U via the FM variational principle. Alternatively, it may be useful to employ contrastive learning (78) or classification based approaches (60, 68, 79) to parameterize CG models.

5. Density-dependence and the internal pressure

Historically, bottom-up approaches have focused on parameterizing U to reproduce the configuration-dependence of the PMF, while neglecting its density dependence. These CG models have typically calculated the internal pressure according to

$$P_{\text{int}}^{\phi}(\mathbf{R}, V) \equiv Nk_B T/V + \frac{1}{3V} \sum_{I=1}^N \mathbf{F}_I \cdot \mathbf{R}_I, \quad 44.$$

which assumes that the CG interaction potential does not explicitly depend on density, i.e., $(\partial U / \partial V)_{\mathbf{R}} = 0$. However, as illustrated in Fig. 2 (80), Eq. 44 often dramatically overestimates the internal pressure of the AA model (81, 82) because the effective potentials indeed vary significantly with density (83, 84, 85, 86, 87). In principle, one can adopt an “active approach” that treats this density dependence when calculating the internal pressure (88, 89, 90). In practice, though, bottom-up approaches have typically modified U such that Eq 44 reasonably approximates the internal pressure of the AA model. For instance, many studies modify IBI pair potentials with linear “ramp corrections” of the form $-A(1 - r/r_c)$ that are tuned to reproduce the average density of the AA model (36, 91).

5.1. Volume potentials

Das and Andersen (DA) derived a fundamentally new approach by treating the volume, V , as a CG variable (92). In this framework, the relevant mapped distribution is

$$p_{\text{RV}}(\mathbf{R}, V; p_{\text{ext}}) \propto p_{\text{R}}(\mathbf{R}; V) \exp[-\beta p_{\text{ext}} V], \quad 45.$$

where $p_{\text{R}}(\mathbf{R}; V)$ is the canonical mapped distribution defined by Eq. 10 and p_{ext} is the external pressure. In order to reproduce p_{RV} , the CG potential must reproduce not only the configuration-dependence, but also the volume-dependence of the modified PMF:

$$W_P(\mathbf{R}, V) \equiv W(\mathbf{R}; V) - (n - N)k_B T \ln(V/V_0), \quad 46.$$

where the first term is defined by Eq 12, the second term corresponds to the ideal contribution from the $n - N$ particles that have been eliminated from the CG model, and V_0 is an arbitrary dimensional constant. If $U(\mathbf{R}, V) = W_P(\mathbf{R}, V)$, then the CG model will perfectly reproduce both the mapped distribution, $p_{\text{RV}}(\mathbf{R}, V; p_{\text{ext}})$, and also the AA pressure at the

Das and Andersen (DA): pioneered a bottom-up approach for coarse-graining at constant pressure (92).

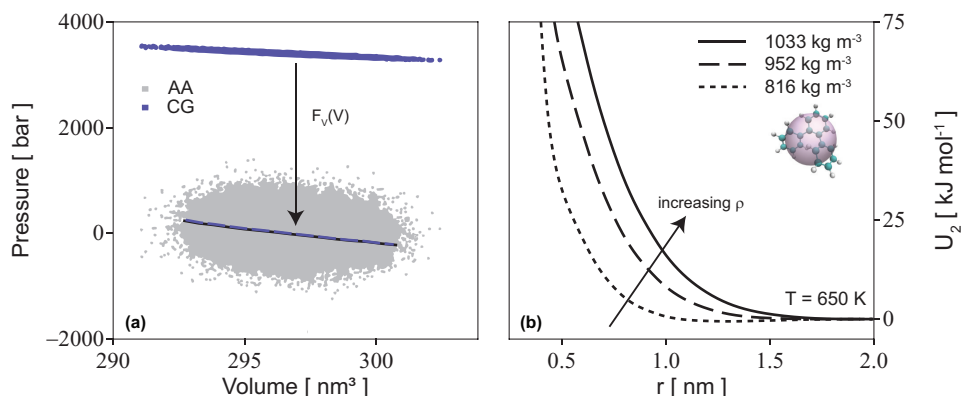


Figure 2

Impact of density upon AA and CG models for OTP (80). Panel a compares the internal pressure of the models. The gray points indicate the internal pressure, $p_{\text{int}}(\mathbf{r}, V)$, of configurations sampled by a constant NPT AA simulation, while the solid black curve indicates the AA equation of state. The blue points indicate the naïve internal pressure, $P_{\text{int}}^{\phi}(\mathbf{M}(\mathbf{r}), V)$, of the CG model obtained by evaluating Eq 44 for the sampled AA configurations. The dashed blue curve indicates the simulated pressure equation of state for the CG model when including the volume potential. Panel b presents the MS-CG pair potentials calculated at $T = 650$ K for the three indicated densities.

resolution of the CG model, i.e., $\bar{p}_{\text{int}}(\mathbf{R}, V) \equiv \langle p_{\text{int}}(\mathbf{r}, V) \rangle_{\mathbf{R}V}$, where the average is evaluated with respect to $p_{\text{R}}(\mathbf{r}|\mathbf{R}; V)$ over the subensemble of AA configurations that map to $(\mathbf{R}, V)$.

DA approximated $W_P(\mathbf{R}, V)$ with the effective potential, $U(\mathbf{R}, V) = U_{\text{R}}(\mathbf{R}) + U_V(V)$, where $U_{\text{R}}(\mathbf{R})$ is a density-independent interaction potential, while $U_V(V)$ is a configuration-independent “volume potential” (92). Because $U_V(V)$ does not generate forces on the CG sites, it does not impact the canonical configuration distribution at a fixed volume. Rather, it directly contributes $F_V(V) = -dU_V(V)/dV$ to the internal pressure, as already recognized in the early work of Dijkstra et al (93, 27). Consequently, simulating volume potentials only requires modifying the barostat equation of motion to account for $F_V(V)$.

DA represented the volume potential by

$$U_V(V) = N \left\{ \psi_1 (V/\bar{V}_{\text{AA}}) + \psi_2 (1 - V/\bar{V}_{\text{AA}})^2 \right\}, \quad 47.$$

where $\bar{V}_{\text{AA}}$ is the average volume of the AA model, while ψ_1 and ψ_2 are parameters that directly contribute to the internal pressure and compressibility, respectively, of the CG model at the reference state point. Given the fixed interaction potential, U_{R} , they optimized U_V by minimizing a “pressure-matching” functional:

$$\chi_V^2[U_V; U_{\text{R}}] = \langle |\Delta p_{\text{int}}(\mathbf{r}, V) - F_V(V)|^2 \rangle_{p_{\text{ext}}}, \quad 48.$$

where $\Delta p_{\text{int}}(\mathbf{r}, V) = p_{\text{int}}(\mathbf{r}, V) - P_{\text{int}}^{\phi}(\mathbf{M}(\mathbf{r}), V)$, while the angular brackets denote an average over the AA isothermal-isobaric ensemble.

Subsequently, Dunn and Noid developed a self-consistent pressure-matching approach that quantitatively reproduces the AA equation of state by minimizing the relative entropy with respect to $U_V(V)$ (82, 94). This self-consistent pressure-matching method has accurately modeled the internal pressure for a variety of liquids, (94, 95, 87) as illustrated in

ortho-terphenyl (OTP): a glass-forming liquid consisting of a central benzene ring with two benzene ring substituents in the ortho- position.

local density (LD): a microscopic density defined by the local environment of a given molecule.

dissipative particle dynamics (DPD): a popular coarse-graining approach that originally developed from computational fluid dynamics (104).

Fig. 2 for ortho-terphenyl (OTP) (80). Moreover, the resulting volume potential can be accurately predicted as a function of solution composition (94) and temperature (96, 87).

5.2. Local density potentials

Volume potentials provide a simple and robust means for modeling the cohesive free energy density of homogeneous fluids. Local density (LD) potentials appear similarly promising for modeling liquid interfaces and other inhomogeneous systems. Pagonabarraga and Frenkel pioneered LD potentials (97) for the many-body dissipative particle dynamics (DPD) method (98, 99). While early DPD models employed purely repulsive pair potentials that generated quadratic pressure equations of state, (100) Pagonabarraga and Frenkel (97) employed LD potentials as a local free energy that could be tuned to model liquid-vapor coexistence and other non-trivial thermodynamic behavior. Independently, Wolynes and coworkers introduced LD potentials to describe many-body solvation effects in CG protein models (101). Since the work of Allen and Rutledge, (102, 103) a growing number of studies have demonstrated that LD potentials provide similar advantages for bottom-up models.

The local density, ρ_I , around site I is defined

$$\rho_I(\mathbf{R}) = \sum_{J(\neq I)} \bar{w}(R_{IJ}), \quad 49.$$

where $\bar{w}$ is a non-increasing weighting function that vanishes beyond a finite distance, r_c . LD potentials generate pair-additive forces and, thus, provide similar computational scaling to conventional pair-additive potentials (97). In particular, if all interactions are described with a pair potential, U_2 , and a LD potential, U_ρ , then the net force on molecule I due to molecule J is along the intermolecular vector, $\hat{\mathbf{R}}_{IJ} = (\mathbf{R}_I - \mathbf{R}_J) / R_{IJ}$, with a magnitude

$$F_{IJ}(\mathbf{R}) = F_2(R_{IJ}) + [F_\rho(\rho_I) + F_\rho(\rho_J)] \bar{w}'(R_{IJ}), \quad 50.$$

where $F_2(r) = -dU_2(r)/dr$ and $F_\rho(\rho) = -dU_\rho(\rho)/d\rho$. Note that Eq. 50 implies that U_2 and U_ρ are not independent. For any constant, c , if one simultaneously transforms $U_2(r) \rightarrow U_2(r) + c\bar{w}(r)$ and $U_\rho(\rho) \rightarrow U_\rho(\rho) - c\rho/2$, then $F_{IJ}(\mathbf{R})$ is left invariant (105).

When the local density is defined over a sufficiently long and slowly varying length-scale, then it is reasonable to simply equate the LD potential with the volume potential that is obtained from self-consistent pressure-matching, i.e., $U_\rho(\rho) = N^{-1}U_V(N/\rho)$. Simulations with this LD potential quite accurately reproduce both the local structure and pressure-density equation of state of bulk methanol, but generate significant artifacts at liquid-vapor interfaces (106). Conversely, if the local density is defined over a sufficiently short length-scale, then LD potentials provide remarkable accuracy and transferability for describing both homogeneous liquids and also their interfaces (105). For instance, Fig. 3 presents the results of DeLyser and Noid for a film of liquid methanol adsorbed against a solid substrate (107). DeLyser and Noid represented each molecule with a single site and employed the MS-CG variational principle to simultaneously determine an effective external potential, V_{ext} , as well as pair and LD potentials to describe intermolecular interactions. Interestingly, the MS-CG equation for the external field generalizes the first YBG equation (107, 18). Figure 3 demonstrates that the MS-CG model quite accurately describes the liquid film, including the liquid-solid and liquid-vapor interfaces, as well as the RDF and internal pressure of bulk methanol. Subsequently, DeLyser and Noid (108) introduced

potentials that depend upon the square gradient of the local density:

$$U(\mathbf{R}) = \sum_{(I,J)} U_2(R_{IJ}) + \sum_{I=1}^N \{U_\rho(\rho_I) + U_\nabla(\rho_I) |\nabla_I \rho_I|^2\}, \quad 51.$$

which appears similar to a van der Waals density functional (18). We anticipate that Eq. 51 may prove particularly useful for bottom-up models of highly inhomogeneous systems, as well as for phenomenological top-down models of exotic phases with many interfaces.

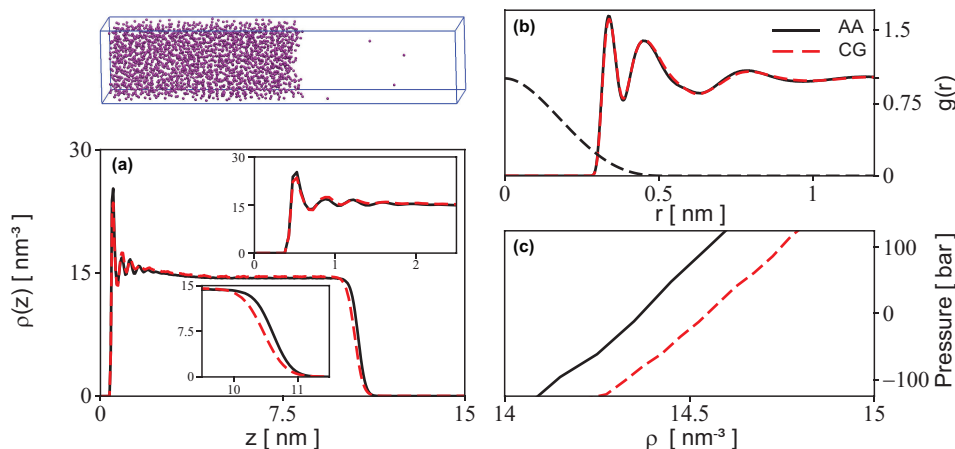


Figure 3

AA (solid black) and 1-site MS-CG models (dashed red) for a film of adsorbed methanol. Panel (a) presents the simulated density profile. The insets highlight the solid-liquid and liquid-vapor interfaces. Panels (b) and (c) present the RDF and pressure-density equation of state from constant pressure simulations of the bulk fluid. The dashed curve in panel (b) presents the shape of the LD weighting function, $\bar{w}$. Adapted with permission from Ref. 107; M. DeLyser and W. G. Noid. *J Chem Phys* **153** 224103; [10.1063/5.0030103](https://doi.org/10.1063/5.0030103) copyright 2020 AIP Publishing.

While our group and the Voth group (109) have primarily employed LD potentials to model liquids and liquid interfaces, other groups have employed LD potentials to accurately model a wide range of complex systems. For instance, Shell and coworkers have demonstrated that bottom-up LD potentials can accurately reproduce many-body solvation effects that drive the collapse of hydrophobic polymers, (110) the aggregation of hydrophobic solutes, (110) and the phase separation of immiscible liquids (111). Similarly, bottom-up LD potentials can accurately describe the structure and thermodynamic properties of polymer melts (112, 113, 114). Consequently, LD potentials appear very promising for accurately modeling thermodynamic properties, interfacial structures, and many-body solvation forces, while providing outstanding transferability between bulk and interfacial environments. Furthermore, we anticipate that LD potentials may provide a unique connection between structure-based bottom-up approaches and DPD models for nonequilibrium energy transfer and chemical reactions (115, 116, 117).

6. Temperature-dependence and energy/entropy

The exact PMF depends upon temperature according to $(\partial W/\partial T)_{\mathbf{R},V} = -S_W$, where $S_W(\mathbf{R})$ quantifies the excess configurational entropy associated with $\mathbf{M}^{-1}(\mathbf{R})$ (28). This not only limits the transferability of approximate interaction potentials, but also complicates the description of energetic and entropic quantities. Because $W = E_W - TS_W$, the PMF cannot be employed to model atomic energetics (3, 84, 10). Similarly, atomic entropies cannot be determined without accounting for the mapping entropy (19, 70, 118)

$$S_{\text{map}} = \int d\mathbf{R} p_{\mathbf{R}}(\mathbf{R}) S_W(\mathbf{R}). \quad 52.$$

These considerations are important for all CG models (119). For instance, top-down models that reproduce configurational-dependent free energies necessarily map entropic contributions from the missing atomic degrees of freedom into the CG potential (120). Consequently, CG models will generally fail to properly distinguish energetic and entropic driving forces unless the entropic contributions to the CG potential are properly treated (121, 122).

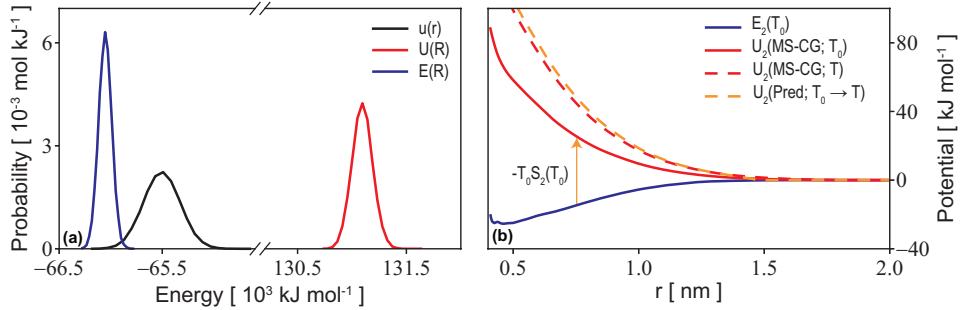


Figure 4

Entropic consequences of coarse-graining for a 1-site model of OTP (10, 123). Panel a compares the distribution of intermolecular energies (black) obtained from an AA simulation at $T_0 = 400$ K with the distribution obtained by evaluating U (red) and E (blue) for the configurations obtained by a corresponding CG simulation that modeled interactions with U . Panel b compares the MS-CG pair potentials, U_2 , calculated for T_0 (solid red) and for $T = 650$ K (dashed red) with the energetic pair operator, E_2 , (solid blue) calculated via energy matching at T_0 , as well as the pair potential predicted for T via the dual approach (dashed orange).

Figure 4 illustrates the practical ramifications of these considerations for a one-site MS-CG model of OTP (123). The left panel compares the AA distribution of intermolecular energies with the distribution obtained by simply evaluating the MS-CG potential, U , for configurations sampled by a corresponding CG simulation. The latter distribution completely fails to describe the cohesive energy of the AA model due to the entropic contributions to U . These entropic contributions correspond to the temperature-dependence of the corresponding MS-CG pair potentials, $U_2(T)$, as illustrated in the right panel (124, 125, 126, 87, 127, 128).

Accordingly, Voth and coworkers proposed calculating MS-CG potentials, $U(T)$, for a range of temperatures and then approximating the entropic component of the PMF, S_W , based upon their observed temperature-dependence, i.e., $S_W \approx -\Delta U(T)/\Delta T$ (129). Moreover, they quite reasonably estimated atomic entropies by evaluating this approximation S for sampled CG configurations (127).

6.1. Dual and observable-based approaches

Conversely, Lebold and Noid (LN) introduced a “dual approach” for modeling energetics and predicting temperature-dependent potentials (130). In contrast to V , the AA potential, $u(\mathbf{r})$, is not a CG observable (109). However, the conditioned mean of the AA potential, $E_W(\mathbf{R}) \equiv \langle u(\mathbf{r}) \rangle_{\mathbf{R}}$, is a CG observable that describes energetics at the resolution of the CG model. Moreover, if both $W(\mathbf{R})$ and $E_W(\mathbf{R})$ are known at a given temperature, then $S_W(\mathbf{R}) = (E_W(\mathbf{R}) - W(\mathbf{R})) / T = -\partial W / \partial T$ determines the temperature-dependence of W .

Lebold and Noid (LN): introduced a dual approach for modeling energetics and predicting temperature-dependent potentials (130).

In practice, neither $W(\mathbf{R})$ nor $E_W(\mathbf{R})$ can be exactly calculated. Consequently, LN employed structure-based methods to determine a potential, $U(\mathbf{R})$, that approximated $W(\mathbf{R})$ at a reference temperature, T_0 . Following the arguments of the least-squares sidebar, LN determined an operator, $E(\mathbf{R})$, that accurately approximated $E_W(\mathbf{R})$ by minimizing

$$\chi_E^2[E] \equiv \langle |u(\mathbf{r}) - E(\mathbf{M}(\mathbf{r}))|^2 \rangle, \quad 53.$$

which repurposes an early energy-matching method (131). Moreover, LN employed the simple approximation $S(\mathbf{R}) \equiv (E(\mathbf{R}) - U(\mathbf{R})) / T_0 \approx S_W(\mathbf{R})$ to predict the temperature-dependence of $U(\mathbf{R})$, i.e., $(\partial U / \partial T)_{\mathbf{R}, V} \approx -S$. LN initially validated this dual approach for an IBI model for water and a MS-CG model for methanol (130).

Figure 4 presents the results of this dual approach for the 1-site MS-CG model of OTP (123). The solid blue curve in Fig. 4b presents the resulting pair energy function, $E_2(T_0)$. Clearly, E_2 is far more attractive than the purely repulsive MS-CG pair potential, U_2 . In fact, OTP solidifies if one models interactions with E_2 . Thus, the entropic contributions to U_2 appear essential for modeling the liquid phase.

The dotted blue curve in Fig. 4a presents the distribution of energies that is obtained by evaluating the energetic operator, $E(\mathbf{R})$, for the configurations that are sampled by CG simulations with the MS-CG potential, $U(\mathbf{R})$. This distribution is necessarily more narrow than the AA potential distribution because it averages over the subensemble, $\mathbf{M}^{-1}(\mathbf{R})$, of AA configurations that map to each CG configuration, $\mathbf{R}$. Nevertheless, this distribution quite accurately reproduces the mean of the AA potential distribution and certainly recovers the cohesion that stabilizes the condensed phase. Furthermore, the inferred entropic contribution to the MS-CG potential, $S_2 = (E_2 - U_2) / T$, very accurately predicts the temperature dependence of the MS-CG potential over a range of 250 K.

Figure 4 suggests a fundamental inconsistency may arise in treating thermodynamic energies with reduced CG models (10). If one determines an accurate energetic operator, $E(T)$, for a range of temperatures, then one should be able to reproduce the thermodynamic energy of the AA model at each temperature, $\bar{u}(T)$. This would imply that the CG model also should reproduce the (excess) specific heat, $c_V = \partial \bar{u} / \partial T$, of the AA model. However, the specific heat is related to the variance, σ^2 , in the AA potential fluctuations according to, $\sigma^2 = k_B T^2 c_V$ (16). This suggests a contradiction: the CG model should not be able to reproduce c_V because it cannot reproduce fluctuations due to AA degrees of freedom that are missing from the CG resolution. The sidebar on hidden fluctuations provides the key to resolving this suggested contradiction. Specifically, if one defines $\tilde{\sigma}^2(\mathbf{R}) = \langle |u(\mathbf{r}) - E_W(\mathbf{M}(\mathbf{r}))|^2 \rangle_{\mathbf{R}}$ as the variance in energetic fluctuations within $\mathbf{M}^{-1}(\mathbf{R})$, then the total atomic variance, σ^2 can be reproduced by accounting for both the variance, Σ^2 , in E_W over the mapped ensemble and also the average of $\tilde{\sigma}^2(\mathbf{R})$ from the missing degrees of freedom according to Eq. SB6 (132). Moreover, because W is a free energy and, thus, a cumulant generating function, $\tilde{\sigma}^2(\mathbf{R})$ can be related to a configuration-dependent specific

heat, $C_W \equiv (\partial E_W / \partial T)_{\mathbf{R}, V}$:

$$\left(\frac{\partial^2 (-\beta W(\mathbf{R}))}{\partial (-\beta)^2} \right)_{\mathbf{R}} = k_B T^2 C_W(\mathbf{R}) = \tilde{\sigma}^2(\mathbf{R}). \quad 54.$$

More generally, one can define an operator $A(\mathbf{R}) \equiv \langle a(\mathbf{r}) \rangle_{\mathbf{R}}$ to model any microscopic observable $a(\mathbf{r})$ at the resolution of the CG model (109). These operators can be optimized by simple least squares fitting, as illustrated in the dual approach, or by minimizing a corresponding relative entropy (133). In recent years, CG operators have been developed for modeling the pressure, surface tension, and specific heat (130, 132, 109, 134, 135).

Hidden fluctuations

We continue the arguments of the preceding sidebar and define, $C(X) = \mu$, such that $\chi^2[\mu] = \sigma^2$. Eq. SB5 then allows us to decompose the total microscopic variance, σ^2 :

$$\sigma^2 = \Sigma^2 + \int dX P(X) \tilde{\sigma}^2(X) \quad \text{SB6.}$$

where $\Sigma^2 \equiv \int dX P(X) |\tilde{\mu}(X) - \mu|^2$ quantifies the fluctuations that are visible at the CG representation, while the second term in Eq. SB6 quantifies the fluctuations that have been eliminated by coarse-graining.

6.2. Microcanonical formalism and modeling fluctuations

Pretti and Shell (PS): introduced an approach for modeling a microcanonical partition function of CG coordinates.

Pretti and Shell (PS) have recently introduced an elegant microcanonical framework for treating the entropic consequences of coarse-graining (136). Somewhat in analogy to the DA framework, PS proposed directly modeling the joint configuration-energy distribution

$$p_{\text{RE}}(\mathbf{R}, E; T) \equiv \int_{V^n} d\mathbf{r} p_{\text{r}}(\mathbf{r}; T) \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) \delta(u(\mathbf{r}) - E), \quad 55.$$

that is specified by the AA potential and the CG mapping. The central quantity in this approach is not the PMF but rather the configuration-energy density of states

$$\omega(\mathbf{R}, E) \propto \int_{V^n} d\mathbf{r} \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R}) \delta(u(\mathbf{r}) - E), \quad 56.$$

which determines the exact distribution of energies sampled by the AA model in the subensemble $\mathbf{M}^{-1}(\mathbf{R})$. If the function $\omega(\mathbf{R}, E)$ is exactly known for all $\mathbf{R}$ and E , then one can determine the exact PMF at every temperature via Laplace transform:

$$e^{-\beta W(\mathbf{R}; T)} = \int dE e^{-\beta E} \omega(\mathbf{R}, E). \quad 57.$$

PS proposed approximating $\omega(\mathbf{R}, E)$ as a convolution of densities of state, $\Omega_{\zeta}(E_{\zeta}, z)$, associated with each term in Eq. 19. This convolution approximation leads to a corresponding additive approximation for the PMF (136). PS proposed approximating Ω_{ζ} with a Gaussian,

$$\Omega_{\zeta}(E_{\zeta}, z) \propto \frac{1}{\sqrt{2\pi\gamma_{\zeta}(z)}} \exp \left[-\frac{(E_{\zeta} - E_{\zeta\infty}(z))^2}{2\gamma_{\zeta}(z)} \right], \quad 58.$$

where $E_{\zeta\infty}(z)$ and $\gamma_{\zeta}(z)$ are temperature-independent functions that specify the mean and variance in the distribution of potential energies associated with the ζ interaction when $\psi_{\zeta}(\mathbf{R}_{\lambda}) = z$. This then determines an approximate density of states, $\Omega(\mathbf{R}, E) \approx \omega(\mathbf{R}, E)$, and a corresponding approximation $P_{\text{RE}}(\mathbf{R}, E) \propto e^{-\beta E} \Omega(\mathbf{R}, E)$ for $p_{\text{RE}}(\mathbf{R}, E)$. PS optimized $E_{\zeta\infty}(z)$ and $\gamma_{\zeta}(z)$ by minimizing the KL divergence between p_{RE} and P_{RE} .

By analytically evaluating the relevant convolutions, PS determined temperature-dependent approximations $E(\mathbf{R}) = \sum_{\zeta} \sum_{\lambda} E_{\zeta}(\psi_{\zeta}(\mathbf{R}_{\lambda}))$ and $U(\mathbf{R}) = \sum_{\zeta} \sum_{\lambda} U_{\zeta}(\psi_{\zeta}(\mathbf{R}_{\lambda}))$ for $E_{\text{W}}(\mathbf{R})$ and $W(\mathbf{R})$, respectively. Specifically,

$$E_{\zeta}(z) \equiv E_{\zeta}(z; T) = E_{\zeta\infty}(z) - \beta \gamma_{\zeta}(z) \quad 59.$$

$$U_{\zeta}(z) \equiv U_{\zeta}(z; T) = E_{\zeta\infty}(z) - \frac{1}{2} \beta \gamma_{\zeta}(z). \quad 60.$$

While prior studies have extrapolated $U(T)$ as a power series in $\delta T = T - T_0$ about a finite reference temperature, T_0 (129, 124, 125, 87, 130), the microcanonical formalism linearly extrapolates U as a function of $\beta = 1/k_B T$ from $T = \infty$. In particular, as $T \rightarrow \infty$, the effective interaction potential becomes purely energetic and temperature-independent $U_{\zeta}(z; T) \rightarrow E_{\zeta\infty}(z)$. In fact, this is consistent with $S_{\text{W}}(\mathbf{R})$ being an excess configurational entropy that should vanish when $T \rightarrow \infty$ and $p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R})$ equals the uniform distribution $q_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R}) = V^{-(n-N)} \delta(\mathbf{M}(\mathbf{r}) - \mathbf{R})$ (28). PS employed the microcanonical formalism to parameterize one-site CG models that accurately reproduced the structure and energetic distributions of LJ tetramers and liquid water across a rather wide range of temperatures (136).

More generally, this suggests the intriguing possibility of employing bottom-up models to describe arbitrary AA observables, $a(\mathbf{r})$. Back-mapping approaches that “invert” the coarse-grained mapping provide one obvious approach for modeling AA observables. Traditional back-mapping approaches determine a representative configuration $\mathbf{r}_{\mathbf{R}} = \mathbf{M}^+(\mathbf{R})$ from the subensemble, $\mathbf{M}^{-1}(\mathbf{R}) \equiv \{\mathbf{r}|\mathbf{M}(\mathbf{r}) = \mathbf{R}\}$, that maps to $\mathbf{R}$ (137, 62). However, recent studies have demonstrated remarkable progress in more rigorously sampling the conditioned distribution, $p_{\mathbf{r}|\mathbf{R}}(\mathbf{r}|\mathbf{R})$, for this subensemble (138, 139). In particular, Rotskoff and coworkers defined the notion of an “invertible coarse-graining” by the combination $(\mathbf{M}, U, T)$ of a CG mapping, $\mathbf{M}$, an approximate CG potential, U , and a normalizing flow, T , for approximately sampling AA configurations according to $p_{\mathbf{r}|\mathbf{R}}$ (64). Furthermore, they introduced an intriguing “weak” consistency criterion for optimizing an invertible coarse-graining, $(\mathbf{M}, U, T)$, in order to reproduce AA averages for a given family, $\mathcal{F} = \{a(\mathbf{r})\}$, of AA observables. Conversely, Jackson and coworkers have highlighted the promise of “bypassing” back-mapping by directly modeling the distribution, $p_{a|\mathbf{R}}(a|\mathbf{R})$, of the observable in $\mathbf{M}^{-1}(\mathbf{R})$ (140). Specifically, Jackson and coworkers employed ML tools to model the distribution of electronic energies for the ensemble of quantum states that map to a given CG configuration (141, 140). The PS microcanonical framework and the “predictive framework” of Koutsourelakis and coworkers (142) similarly bypass backmapping. These exciting ideas promise to radically extend the scope and utility of bottom-up CG methods.

7. Concluding thoughts

By employing information from accurate AA models, the bottom-up framework anchors CG models in reality. Bottom-up models are often referred to as “structure-based” models because early studies primarily focused on reproducing structural properties (34, 36). It was quickly realized, though, that structure-based potentials not only provided limited

transferability, but also provided a poor description of thermodynamic properties, which was termed the “representability” problem (83, 91). The bottom-up formalism elucidates the origin of the transferability and representability problems and reveals that they are two weeds stemming from the same root (29). Moreover, the bottom-up formalism guides the development of robust computational methods for rigorously addressing these problems in practice. For instance, the DA framework beautifully addresses the density-dependence of the PMF in order to accurately model the internal pressure (92). LD potentials appear a promising approach for extending this framework to inhomogeneous systems. Similarly, the dual (130) and microcanonical frameworks (136) provide predictive approaches for treating the temperature-dependence of effective potentials in order to model energetic and entropic quantities. More generally, recent studies have reported exciting bottom-up approaches for modeling AA observables that would seem beyond the scope of CG models (138, 139, 64, 141, 140, 142). Furthermore, the bottom-up framework provides insight for coarse-graining complex systems that are characterized by multiple conformational states (74, 75). Thus, the bottom-up framework provides powerful insight for fundamentally understanding and practically resolving the limitations of existing CG models. We anticipate this framework will continue rapidly propelling the field into the future.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of former group members who made this review possible, including Wayne Mullinax, Joseph Rudzinski, Chris Ellis, Nick Dunn, Tommy Foley, Kathryn Lebold, and Michael DeLyser. WGN also gratefully acknowledges very fruitful work with G. A. Voth, H. C. Andersen, and M. Scott Shell that deeply influenced this perspective. WGN acknowledges financial support from the National Science Foundation (Grant Nos. MCB-1053970, CHE-1565631, CHE-1856337 and CHE-2154433). KMK acknowledges the financial support of a Marie Skłodowska Curie science achievement graduate scholarship in Chemistry from Penn State. Figures 2 and 4 employed VMD. (143) VMD is developed with NIH support by the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign. Finally, the authors gratefully acknowledge the late Phillip Geissler for the invitation to contribute this review and for the inspiration of his beautiful work.

LITERATURE CITED

References

1. Guenza M, Dinpajoo M, McCarty J, Lyubimov I. 2018. Accuracy, transferability, and efficiency of coarse-grained models of molecular liquids. *J. Phys. Chem. B* 122(45):10257–10278
2. Gartner TE, Jayaraman A. 2019. Modeling and simulations of polymers: A roadmap. *Macromolecules* 52(3):755–786
3. Muller M, Katsov K, Schick M. 2006. Biological and synthetic membranes: What can be learned from a coarse-grained description? *Phys. Rep.* 434(5-6):113–176

4. Deserno M. 2009. Mesoscopic membrane physics: Concepts, simulations, and selected applications. *Macromol. Rapid Comm.* 30(9-10):752–771
5. Marrink SJ, Monticelli L, Melo MN, Alessandri R, Tieleman DP, Souza PCT. 2022. Two decades of Martini: Better beads, broader scope. *WIREs Computational Molecular Science*
6. Noid WG. 2013. Perspective: coarse-grained models for biomolecular systems. *J. Chem. Phys.* 139(9):090901
7. Rudzinski JF. 2019. Recent progress towards chemically-specific coarse-grained simulation models with consistent dynamical properties. *Comput.* 7(3):42
8. Giulini M, Rigoli M, Mattiotti G, Menichetti R, Tarenzi T, et al. 2021. From System Modeling to System Analysis: The Impact of Resolution Level and Resolution Distribution in the Computer-Aided Investigation of Biomolecules. *Front. Mol. Biosci.* 8:676976
9. Dhamankar S, Webb MA. 2021. Chemically specific coarse-graining of polymers: Methods and prospects. *Journal of Polymer Science* 59(22):2613–2643
10. Kidder KM, Szukalo RJ, Noid WG. 2021. Energetic and entropic considerations for coarse-graining. *Eur. Phys. J. B* 94:153
11. Sun T, Minhas V, Korolev N, Mirzoev A, Lyubartsev AP, Nordenskiöld L. 2021. Bottom-Up Coarse-Grained Modeling of DNA. *Frontiers in Molecular Biosciences* 8:645527
12. Jin J, Pak AJ, Durumeric AEP, Loose TD, Voth GA. 2022. Bottom-up Coarse-Graining: Principles and Perspectives. *J. Chem. Theory Comput.* 18(10):5759–5791
13. Schilling T. 2022. Coarse-grained modelling out of equilibrium. *Physics Reports* 972:1–45
14. Durumeric AE, Charron NE, Templeton C, Musil F, Bonneau K, et al. 2023. Machine learned coarse-grained protein force-fields: Are we there yet? *Curr. Opin. Struct. Biol.* 79:102533
15. Schmid F. 2023. Understanding and Modeling Polymers: The Challenge of Multiple Scales. *ACS Polymers Au* 3(1):28–58
16. Tuckerman. ME. 2013. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford, Great Britain: Oxford University Press
17. Kullback S, Leibler RA. 1951. On information and sufficiency. *Ann. Math. Stat.* 22(1):79–86
18. Hansen JP, McDonald IR. 1990. *Theory of Simple Liquids*. San Diego, CA USA: Academic Press, 2nd ed.
19. Shell MS. 2008. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *J. Chem. Phys.* 129:144108
20. Cover TM, Thomas JA. 2006. *Elements of Information Theory*. Wiley Interscience, 2nd ed.
21. Touchette H. 2009. The large deviation approach to statistical mechanics. *Physics Reports* 478(1-3):1–69
22. Noid WG, Chu JW, Ayton GS, Krishna V, Izvekov S, et al. 2008. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.* 128:244114
23. Noid WG. 2023. Perspective: Advances, challenges, and insight for predictive coarse-grained models. *J. Phys. Chem. B* 127:4174–4207
24. Kirkwood JG. 1935. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* 3(5):300–313
25. Liwo A, Czaplewski C, Pillardy J, Scheraga HA. 2001. Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field. *J. Chem. Phys.* 115:2323 – 2347
26. Akkermans RLC, Briels WJ. 2001. A structure-based coarse-grained model for polymer melts. *J. Chem. Phys.* 114(2):1020–1031
27. Likos CN. 2001. Effective interactions in soft condensed matter physics. *Phys. Rep.* 348(4–5):267 – 439
28. Foley TT, Shell MS, Noid WG. 2015. The impact of resolution upon entropy and information in coarse-grained models. *J. Chem. Phys.* 143:243104
29. Dunn NJH, Foley TT, Noid WG. 2016. van der waals perspective on coarse-graining: progress toward solving representability and transferability problems. *Acc. Chem. Res.* 49(12):2832–

30. Kalligiannaki E, Harmandaris V, Katsoulakis MA, Plechac P. 2015. The geometry of generalized force matching and related information metrics in coarse-graining of molecular systems. *J. Chem. Phys.* 143(8):084105
31. Ciccotti G, Kapral R, Vanden-Eijnden E. 2005. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem* 6:1809–14
32. Chaimovich A, Shell MS. 2011. Coarse-graining errors and numerical optimization using a relative entropy framework. *J. Chem. Phys.* 134(9):094112
33. Shell MS. 2016. Coarse-graining with the relative entropy. John Wiley & Sons, Inc., 395–441
34. Lyubartsev AP, Laaksonen A. 1995. Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach. *Phys. Rev. E* 52:3730–3737
35. Noid WG. 2013. Systematic methods for structurally consistent coarse-grained models. *Methods Mol Biol* 924:487–531
36. Müller-Plathe F. 2002. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *ChemPhysChem* 3:754 – 769
37. Murtola T, Karttunen M, Vattulainen I. 2009. Systematic coarse graining from structure using internal states: Application to phospholipid/cholesterol bilayer. *J. Chem. Phys.* 131:055101
38. Carmichael SP, Shell MS. 2012. A new multiscale algorithm and its application to coarse-grained peptide models for self-assembly. *J. Phys. Chem. B* 116(29):8383–93
39. Izvekov S, Voth GA. 2005. A multiscale coarse-graining method for biomolecular systems. *J. Phys. Chem. B* 109:2469 – 2473
40. Izvekov S, Voth GA. 2005. Multiscale coarse graining of liquid-state systems. *J. Chem. Phys.* 123:134105
41. Lu L, Voth GA. 2012. The multiscale coarse-graining method. *Adv. Chem. Phys.* 149:47–81
42. Chorin AJ. 2003. Conditional expectations and renormalization. *Multiscale Model. Simul.* 1:105–118
43. Mullinax JW, Noid WG. 2009. A generalized Yvon-Born-Green theory for molecular systems. *Phys. Rev. Lett.* 103:198104
44. Mullinax JW, Noid WG. 2010. A generalized Yvon-Born-Green theory for determining coarse-grained interaction potentials. *J. Phys. Chem. C* 114:5661–5674
45. Rudzinski JF, Noid WG. 2015. A generalized-yvon-born-green method for coarse-grained modeling. *Eur. Phys. J.: Spec. Top.* 224:2193–2216
46. Noid WG, Liu P, Wang YT, Chu JW, Ayton GS, et al. 2008. The multiscale coarse-graining method. II. Numerical implementation for molecular coarse-grained models. *J. Chem. Phys.* 128:244115
47. Noid WG, Chu JW, Ayton GS, Voth GA. 2007. Multiscale coarse-graining and structural correlations: Connections to liquid state theory. *J. Phys. Chem. B* 111:4116–4127
48. Rudzinski JF, Noid WG. 2012. The role of many-body correlations in determining potentials for coarse-grained models of equilibrium structure. *J. Phys. Chem. B* 116(29):8621–8635
49. Rühle V, Junghans C, Lukyanov A, Kremer K, Andrienko D. 2009. Versatile object-oriented toolkit for coarse-graining applications. *J. Chem. Theory Comput.* 5(12):3211–3223
50. Das A, Lu L, Andersen HC, Voth GA. 2012. The multiscale coarse-graining method. X. Improved algorithms for constructing coarse-grained potentials for molecular systems. *J. Chem. Phys.* 136(19):194115
51. Pak AJ, Dannenhoffer-Lafage T, Madsen JJ, Voth GA. 2019. Systematic Coarse-Grained Lipid Force Fields with Semiexplicit Solvation via Virtual Sites. *J. Chem. Theory Comput.* 15(3):2087–2100
52. Jin J, Han Y, Voth GA. 2019. Coarse-graining involving virtual sites: Centers of symmetry coarse-graining. *J. Chem. Phys.* 150(15):154103
53. Sahrman PG, Loose TD, Durumeric AEP, Voth GA. 2022. Utilizing Machine Learning to Greatly Expand the Range and Accuracy of Bottom-Up Coarse-Grained Models Through

- Virtual Particles. ArXiv:2212.04530 [physics]
54. Dama JF, Sinitskiy AV, McCullagh M, Weare J, Roux B, et al. 2013. The theory of ultra-coarse-graining. 1. general principles. *J. Chem. Theory Comput.* 9:2466–80
 55. Dama JF, Jin J, Voth GA. 2017. The theory of ultra-coarse-graining. 3. coarse-grained sites with rapid local equilibrium of internal states. *J. Chem. Theory Comput.* 13(3):1010–1022 PMID: 28112956
 56. Jin J, Han Y, Voth GA. 2018. Ultra-Coarse-Grained Liquid State Models with Implicit Hydrogen Bonding. *J. Chem. Theory Comput.* 14(12):6159–6174
 57. Nguyen H, Huang D. 2022. Systematic bottom-up molecular coarse-graining via force and torque matching using anisotropic particles. *J. Chem. Phys.* 156:184118
 58. Dequidt A, Solano Canchaya JG. 2015. Bayesian parametrization of coarse-grain dissipative dynamics models. *J. Chem. Phys.* 143(8):084122
 59. Nkpesu Mbitou RL, Goujon F, Dequidt A, Latour B, Devémy J, et al. 2022. Consistent and Transferable Force Fields for Statistical Copolymer Systems at the Mesoscale. *J. Chem. Theory Comput.* 18(11):6940–6951
 60. Lemke T, Peter C. 2017. Neural Network Based Prediction of Conformational Free Energies - A New Route toward Coarse-Grained Simulation Models. *J. Chem. Theory Comput.* 13(12):6213–6221
 61. Zhang L, Han J, Wang H, Car R, E W. 2018. DeePCG: Constructing coarse-grained models via deep neural networks. *J. Chem. Phys.* :10
 62. Wang W, Gómez-Bombarelli R. 2019. Coarse-graining auto-encoders for molecular dynamics. *npj Comput. Mat.* 5(1):125
 63. Wang J, Olsson S, Wehmeyer C, Pérez A, Charron NE, et al. 2019. Machine learning of coarse-grained molecular dynamics force fields. *ACS Cent. Sci.* 5(5):755–767
 64. Chennakesavalu S, Toomer DJ, Rotskoff GM. 2023. Ensuring thermodynamic consistency with invertible coarse-graining. *The Journal of Chemical Physics* 158(12):124126
 65. John ST, Csányi G. 2017. Many-Body Coarse-Grained Interactions Using Gaussian Approximation Potentials. *J. Phys. Chem. B* 121(48):10934–10949
 66. Scherer C, Scheid R, Andrienko D, Bereau T. 2020. Kernel-Based Machine Learning for Efficient Simulations of Molecular Liquids. *J. Chem. Theory Comput.* 16(5):3194–3204
 67. Wang J, Chmiela S, Müller KR, Noé F, Clementi C. 2020. Ensemble learning of coarse-grained molecular dynamics force fields with a kernel approach. *J. Chem. Phys.* 152(19):194106
 68. Durumeric AEP, Voth GA. 2019. Adversarial-residual-coarse-graining: Applying machine learning theory to systematic molecular coarse-graining. *J. Chem. Phys.* 151(12):124110
 69. Song Y, Kingma DP. 2021. How to train your energy-based models
 70. Rudzinski JF, Noid WG. 2011. Coarse-graining entropy, forces, and structures. *J. Chem. Phys.* 135(21):214101
 71. Hyvarinen A. 2005. Estimation of Non-Normalized Statistical Models by Score Matching. *J. Mach. Learn. Res.* 6:695–709
 72. Stein C. 1972. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proc. Sixth Berkeley Symp. Math. Stat. Prob.*, pp. 583–602
 73. Liu Q, Lee J, Jordan M. 2016. A kernelized Stein discrepancy for goodness-of-fit tests. In *International conference on machine learning*, pp. 276–284. PMLR
 74. Thaler S, Stupp M, Zavadlav J. 2022. Deep coarse-grained potentials via relative entropy minimization. *The Journal of Chemical Physics* 157(24):244103
 75. Köhler J, Chen Y, Krämer A, Clementi C, Noé F. 2023. Flow-Matching: Efficient Coarse-Graining of Molecular Dynamics without Forces. *J. Chem. Theory Comput.* 19(3):942–952
 76. Papamakarios G, Nalisnick E, Rezende DJ, Mohamed S, Lakshminarayanan B. 2021. Normalizing Flows for Probabilistic Modeling and Inference. *J. Mach. Learn. Res.* 22:1–64
 77. Kobyzev I, Prince SJ, Brubaker MA. 2021. Normalizing Flows: An Introduction and Review of Current Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 43(11):3964–3979

78. Ding X, Zhang B. 2022. Contrastive Learning of Coarse-Grained Force Fields. *J. Chem. Theory Comput.* 18(10):6334–6344
79. Durumeric AEP, Voth GA. 2021. Explaining classifiers to understand coarse-grained models. ArXiv:2109.07337 [physics], doi:10.48550/ARXIV.2109.07337
80. Szukalo RJ, Noid W. 2020. Investigation of coarse-grained models across a glass transition. *Soft Mater.* 18:1–15
81. Guenza M. 2015. Thermodynamic consistency and other challenges in coarse-graining models. *Eur. Phys. J.: Spec. Top.* 224:2177–2191
82. Dunn NJH, Noid WG. 2015. Bottom-up coarse-grained models that accurately describe the structure, pressure, and compressibility of molecular liquids. *J. Chem. Phys.* 143(24):243148
83. Louis AA. 2002. Beware of density dependent pair potentials. *J. Phys.: Condens. Matter* 14:9187–9206
84. Johnson ME, Head-Gordon T, Louis AA. 2007. Representability problems for coarse-grained water potentials. *J. Chem. Phys.* 126:144509
85. Chaimovich A, Shell MS. 2009. Anomalous waterlike behavior in spherically-symmetric water models optimized with the relative entropy. *Phys. Chem. Chem. Phys.* 11(12):1901–1915
86. Izvekov S, Chung PW, Rice BM. 2010. The multiscale coarse-graining method: Assessing its accuracy and introducing density dependent coarse-grain potentials. *J. Chem. Phys.* 133:064109
87. Lebold KM, Noid WG. 2019. Systematic study of temperature and density variations in effective potentials for coarse-grained models of molecular liquids. *J. Chem. Phys.* 150(1):014104
88. Stillinger FH, Sakai H, Torquato S. 2002. Statistical mechanical models with effective potentials: Definitions, applications, and thermodynamic consequences. *J. Chem. Phys.* 117(1):288–296
89. D’Adamo G, Pelissetto A, Pierleoni C. 2013. Predicting the thermodynamics by using state-dependent interactions. *J. Chem. Phys.* 138(23):234107
90. Dinpajoo M, Guenza MG. 2018. On the Density Dependence of the Integral Equation Coarse-Graining Effective Potential. *J. Phys. Chem. B* 122(13):3426–3440
91. Wang H, Junghans C, Kremer K. 2009. Comparative atomistic and coarse-grained study of water: What do we lose by coarse-graining? *Eur. Phys. J. E* 28(2):221–229
92. Das A, Andersen HC. 2010. The multiscale coarse-graining method. V. Isothermal-isobaric ensemble. *J. Chem. Phys.* 132:164106
93. Dijkstra M, van Roij R, Evans R. 1998. Phase behavior and structure of binary hard-sphere mixtures. *Phys. Rev. Lett.* 81(11):2268–2271
94. Dunn NJH, Noid WG. 2016. Bottom-up coarse-grained models with predictive accuracy and transferability for both structural and thermodynamic properties of heptane-toluene mixtures. *J. Chem. Phys.* 144:204124
95. Dunn NJH, Lebold KM, DeLyser MR, Rudzinski JF, Noid WG. 2018. BOCS: Bottom-up Open-source Coarse-graining Software. *J. Phys. Chem. B* 122(13):3363–3377
96. Rosenberger D, van der Vegt NFA. 2018. Addressing the temperature transferability of structure based coarse graining models. *Phys. Chem. Chem. Phys.* 20:6617–6628
97. Pagonabarraga I, Frenkel D. 2001. Dissipative particle dynamics for interacting systems. *J. Chem. Phys.* 115(11):5015–5026
98. Warren PB. 2003. Vapor-liquid coexistence in many-body dissipative particle dynamics. *Phys. Rev. E* 68(6):066702
99. Ghoufi A, Malfreyt P. 2010. Calculation of the surface tension from multibody dissipative particle dynamics and Monte Carlo methods. *Phys. Rev. E* 82(1):016706
100. Groot RD, Warren PB. 1997. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* 107(11):4423–4435
101. Papoian GA, Ulander J, Eastwood MP, Luthey-Schulten Z, Wolynes PG. 2004. Water in protein structure prediction. *Proc. Natl. Acad. Sci. U.S.A.* 101(10):3352–3357

102. Allen EC, Rutledge GC. 2008. A novel algorithm for creating coarse-grained, density dependent implicit solvent models. *J. Chem. Phys.* 128:154115
103. Allen EC, Rutledge GC. 2009. Evaluating the transferability of coarse-grained, density-dependent implicit solvent models to mixtures and chains. *J. Chem. Phys.* 130:034904
104. Español P, Warren PB. 2017. Perspective: Dissipative particle dynamics. *J. Chem. Phys.* 146(15):150901
105. DeLyser MR, Noid WG. 2019. Analysis of local density potentials. *J. Chem. Phys.* 151(22):224106
106. DeLyser MR, Noid WG. 2017. Extending pressure-matching to inhomogeneous systems via local-density potentials. *J. Chem. Phys.* 147:134111
107. DeLyser MR, Noid WG. 2020. Bottom-up coarse-grained models for external fields and interfaces. *J. Chem. Phys.* 153(22):224103
108. DeLyser MR, Noid WG. 2022. Coarse-grained models for local density gradients. *J. Chem. Phys.* 156:034106
109. Wagner JW, Dama JF, Durumeric AEP, Voth GA. 2016. On the representability problem and the physical meaning of coarse-grained models. *J. Chem. Phys.* 145(4):044108
110. Sanyal T, Shell MS. 2016. Coarse-grained models using local-density potentials optimized with the relative entropy: Application to implicit solvation. *J. Chem. Phys.* 145(3):034109
111. Sanyal T, Shell MS. 2018. Transferable coarse-grained models of liquid-liquid equilibrium using local density potentials optimized with the relative entropy. *J. Phys. Chem. B* 122(21):5678–5693
112. Agrawal V, Peralta P, Li Y, Oswald J. 2016. A pressure-transferable coarse-grained potential for modeling the shock hugoniot of polyethylene. *J. Chem. Phys.* 145(10):104903
113. Shahidi N, Chazirakis A, Harmandaris V, Doxastakis M. 2020. Coarse-graining of polyisoprene melts using inverse Monte Carlo and local density potentials. *J. Chem. Phys.* 152(12):124902
114. Berressem F, Scherer C, Andrienko D, Nikoubashman A. 2021. Ultra-coarse-graining of homopolymers in inhomogeneous systems. *J. Phys.: Condens. Matter* 33(25):254002
115. Español P, Serrano M, Pagonabarraga I, Zúñiga I. 2016. Energy-conserving coarse-graining of complex molecules. *Soft Matter* 12(21):4821–4837
116. Avalos JB, Lísal M, Larentzos JP, Mackie AD, Brennan JK. 2019. Generalised dissipative particle dynamics with energy conservation: density- and temperature-dependent potentials. *Phys. Chem. Chem. Phys.* 21(45):24891–24911
117. Lísal M, Larentzos JP, Avalos JB, Mackie AD, Brennan JK. 2022. Generalized Energy-Conserving Dissipative Particle Dynamics with Reactions. *J. Chem. Theory Comput.* 18(4):2503–2512
118. McCarty J, Clark AJ, Copperman J, Guenza MG. 2014. An analytical coarse-graining method which preserves the free energy, structural correlations, and thermodynamic state of polymer melts from the atomistic to the mesoscale. *J. Chem. Phys.* 140(20):204913
119. Riniker S, Allison JR, van Gunsteren WF. 2012. On developing coarse-grained models for biomolecular simulation: a review. *Phys. Chem. Chem. Phys.* 14(36):12423–12430
120. Baron R, Trzesniak D, de Vries AH, Elsener A, Marrink SJ, van Gunsteren WF. 2007. Comparison of thermodynamic properties of coarse-grained and atomic-level simulation models. *ChemPhysChem* 8(3):452–461
121. Wu Z, Cui Q, Yethiraj A. 2011. Driving Force for the Association of Hydrophobic Peptides: The Importance of Electrostatic Interactions in Coarse-Grained Water Models. *J. Phys. Chem. Lett.* 2(14):1794–1798
122. Jarin Z, Newhouse J, Voth GA. 2020. Coarse-grained Force Fields from the Perspective of Statistical Mechanics: Better Understanding the Origins of a MARTINI Hangover. *J. Chem. Theory Comput.* 17(2):1170–1180
123. Szukalo RJ, Noid WG. 2021. Investigating the energetic and entropic components of effective potentials across a glass transition. *J. Phys.: Condens. Matter* 33(15):154004

124. Izvekov S. 2011. Towards an understanding of many-particle effects in hydrophobic association in methane solutions. *J. Chem. Phys.* 134(3):034104
125. Farah K, Fogarty AC, Böhm MC, Müller-Plathe F. 2011. Temperature dependence of coarse-grained potentials for liquid hexane. *Phys. Chem. Chem. Phys.* 13(7):2894–902
126. Mukherjee B, Delle Site L, Kremer K, Peter C. 2012. Derivation of coarse grained models for multiscale simulation of liquid crystalline phase transitions. *J. Phys. Chem. B* 116(29):8474–8484
127. Jin J, Pak AJ, Voth GA. 2019. Understanding entropy in coarse-grained systems: Addressing issues of representability and transferability. *J. Phys. Chem. Lett.* 10(16):4549–4557
128. Jin J, Yu A, Voth GA. 2020. Temperature and phase transferable bottom-up coarse-grained models. *J. Chem. Theory Comput.* 16(11):6823–6842
129. Lu L, Voth GA. 2011. The multiscale coarse-graining method. VII. Free energy decomposition of coarse-grained effective potentials. *J. Chem. Phys.* 134(22):224107
130. Lebold KM, Noid WG. 2019. Dual approach for effective potentials that accurately model structure and energetics. *J. Chem. Phys.* 150(23):234107
131. Tóth G. 2007. Effective potentials from complex simulations: a potential-matching algorithm and remarks on coarse-grained potentials. *J. Phys.: Condens. Matter* 19(33):335222
132. Lebold KM, Noid WG. 2019. Dual-potential approach for coarse-grained implicit solvent models with accurate, internally consistent energetics and predictive transferability. *J. Chem. Phys.* 151(16):164113
133. Wagner JW, Dannenhoffer-Lafage T, Jin J, Voth GA. 2017. Extending the range and physical accuracy of coarse-grained models: Order parameter dependent interactions. *J. Chem. Phys.* 147(4):044113
134. Dannenhoffer-Lafage T, Wagner JW, Durumeric AEP, Voth GA. 2019. Compatible observable decompositions for coarse-grained representations of real molecular systems. *J. Chem. Phys.* 151(13):134115
135. Jin J, Pak AJ, Voth GA. 2019. Understanding Missing Entropy in Coarse-Grained Systems: Addressing Issues of Representability and Transferability. *J. Phys. Chem. Lett.* 10:4549–4557
136. Pretti E, Shell MS. 2021. A microcanonical approach to temperature-transferable coarse-grained models using the relative entropy. *J. Chem. Phys.* 155:094102
137. Wassenaar TA, Pluhackova K, Böckmann RA, Marrink SJ, Tieleman DP. 2014. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* 10(2):676–690
138. Stieffenhofer M, Wand M, Bereau T. 2020. Adversarial Reverse Mapping of Equilibrated Condensed-Phase Molecular Structures. *Machine Learning: Science and Technology* 1:045014
139. Wang W, Xu M, Cai C, Miller BK, Smidt T, et al. 2022. Generative coarse-graining of molecular conformations
140. Maier JC, Jackson NE. 2022. Bypassing backmapping: Coarse-grained electronic property distributions using heteroscedastic Gaussian processes. *The Journal of Chemical Physics* 157(17):174102
141. Sivaraman G, Jackson NE. 2022. Coarse-Grained Density Functional Theory Predictions via Deep Kernel Learning. *J. Chem. Theory Comput.* 18(2):1129–1141
142. Schöberl M, Zabarar N, Koutsourelakis PS. 2017. Predictive coarse-graining. *J. Comp. Phys.* 333:49–77
143. Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* 14:33–38