Coarse-grained Potential Derivation With Multi-State Iterative Boltzmann Inversion

Timothy C. Moore[1], Christopher R. Iacovella[1], Shan Guo[1], Clare McCabe[1,2]

[1]Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, TN, 37235-1604
[2]Department of Chemistry, Vanderbilt University, Nashville, TN, 37235-1604

NIAMS

VANDERBILT School of Engineering

Overview

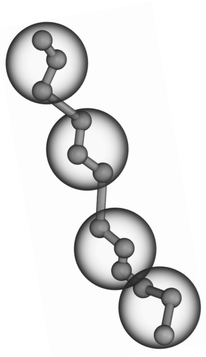
Iterative Boltzmann inversion (IBI) has become a popular method to derive coarse-grained (CG) potentials due to its straightforward nature and general applicability. The method optimizes a potential to match target properties from an atomistic simulation mapped to the CG level. For non-bonded potentials, target data is structural, taking the form of the radial distribution function (RDF). Though widely used, the potentials derived with IBI:

- Are generally non-unique
- Are often heavily state dependent
- May include artifacts from intermediate and long-range structural correlations

Here, we propose an extension to the IBI method to include target data from multiple states, adding constraints to the potential optimization process. We show that adding these constraints results in potentials that are less state-dependent and more representative of the underlying potential and therefore applicable over a wider range of state points.

Coarse-Grained Models

CG models reduce the number of degrees of freedom in a system by reducing the number interaction sites, resulting in a model that is computationally less expensive than the equivalent fully atomistic model.



Superposition of fully atomistic and coarse-grained models of n-dodecane. A 3:1 mapping is used, where each CG bead represents 3 heavy atoms.

Effective bond and angle potentials are typically derived from the bond and angle distributions in an atomistic trajectory. *E.g.,* for a bond length distribution *p(r)*, a Gaussian curve is fitted:

$$p(r) = \frac{A}{\omega\sqrt{\pi/2}} \exp\left[\frac{-2(r-r_{eq})^2}{\omega^2}\right]$$

from which a harmonic potential is derived through a Boltzmann inversion:

$$V(r) = -k_B T \ln p(r) = K_r(r-r_{eq})^2$$

Iterative Boltzmann Inversion

Single-state

A numerical non-bonded potential is iteratively updated until the trial RDF matches, within some tolerance, the target data obtained from atomistic simulations mapped to the CG level. The potentials are updated according to:

$$V_{i+1}(r) = V_i(r) - \alpha k_B T \ln\left[\frac{g_i(r)}{g_t(r)}\right]$$

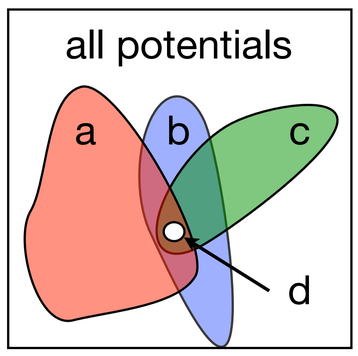
where $V_i(r)$ is the interaction potential after step i, α is a scaling factor to prevent large fluctuations in the updated potential, k_B the Boltzmann constant, T the temperature, $g_i(r)$ the trial RDF, and $g_t(r)$ the target RDF.

This method will typically yield a potential that matches the target RDF well; however, the derived potentials tend to be state dependent and non-unique. The resulting potentials derived via this method may not be representative of the "true" underlying potential, which may result in significant artifacts when considering other properties or other state points not included in the original optimization.

Iterative Boltzmann Inversion

Multi-state extension

Single-state IBI does not guarantee a unique solution, but rather a solution from the region of *potential phase space* that encompasses solution potentials, i.e., those that result in an RDF that match the target. Different thermodynamic states have different regions of phase space that contain solution potentials. In multi-state IBI, we assume that the "true" underlying potential lies within the intersection of these regions of solution phase spaces. Ultimately, the underlying idea of multi-state IBI is to add additional constraints to the optimization process, in order to provide a unique, more generally applicable solution.



Representation of potential phase space. Each region (a, b and c) represents a region of solution potentials for a particular state. The overlap of the regions (and hence most representative potential) shown as d.

A single potential is updated to match structural data over *N* number of states through:

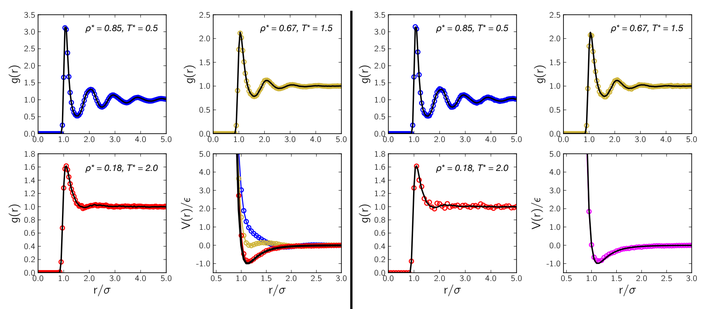
$$V_{i+1}(r) = V_i(r) - \frac{1}{N}\sum_{\{s\}} \alpha_s(r)\, k_B T_s \ln\left[\frac{g_i^s(r)}{g_t^s(r)}\right]$$

where the sum is taken over all states, and the "*s*" script denotes the property at state *s*. The scaling factor $\alpha_s(r)$ is now a weighting factor for state *s*, allowing more or less emphasis to be put on this state in the potential update. The parameter α is set to be a linear function of *r* such that the potential smoothly approaches zero at the cutoff.

Applications

Lennard-Jonesium Fluid

Both single-state IBI and multi-state IBI were performed on a system of particles interacting through a Lennard-Jones (LJ) potential. No coarse-graining was performed, enabling us to asses the ability of multi-state IBI to recover the known underlying potential from 3 unique states, varying from a gaseous phase to a dense liquid.



Results of single-state IBI for LJ particles. 3 separate potentials were derived, one for each state.

Results of multi-state IBI for LJ particles. A single potential was derived using target data from multiple states.

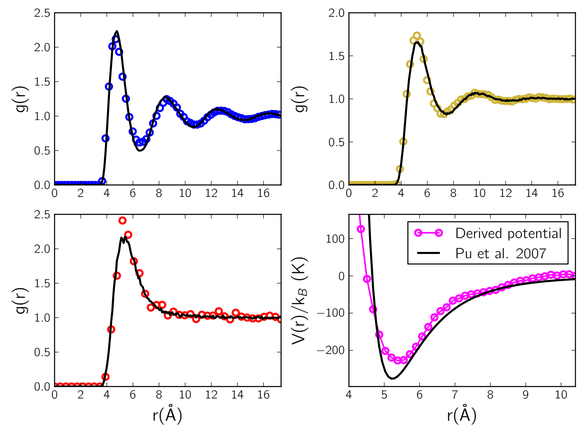
Multi-state IBI is able to reproduce the LJ potential with near perfect accuracy. Single-state optimizations yield 3 unique potentials, where only one potential is actually representative of the true potential; in practice, the true underlying potential is not known, which may lead to issues in determining whether single-state potentials are reasonable. These results highlight two important facts:

- Potentials derived using data from a single state are highly state dependent
- Multi-state IBI is able to derive a single potential that converges to the known true potential, simultaneously matching target data at multiple states

Applications

Propane

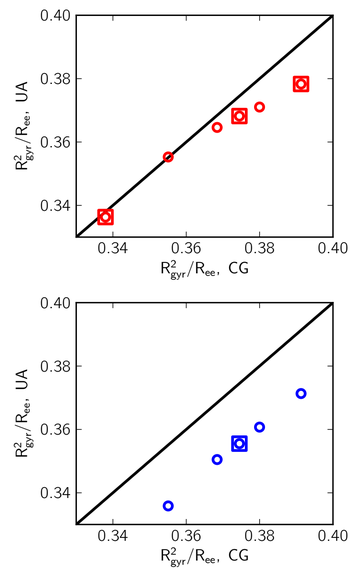
A CG model was developed for propane using a united atom model as the target data. The resulting model is a single-site model which can be compared to the LJ system mapped to have the same critical point as propane.



Good agreement between the derived and target structural data is seen at all states, with the derived potential demonstrating similar behavior to the published LJ propane model of Pu *et al*.[2]

Dodecane

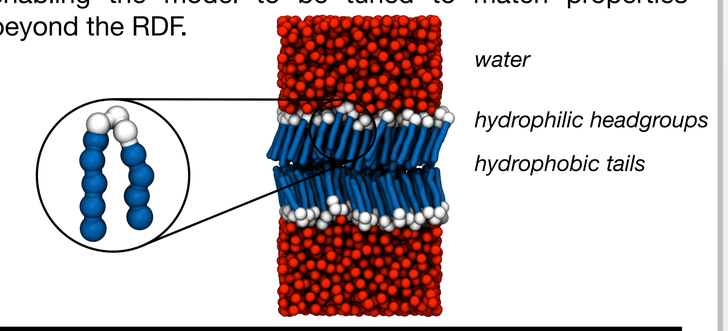
Showing the transferability of the potentials, middle bead potentials were derived for *n*-dodecane, using end bead potentials taken from from *n*-hexane optimizations. The square of the radius of gyration normalized by the end-to-end length was compared to the atomistic data mapped to the CG level, for both single- and multi-state IBI. Single-state optimization was found to lack quantitative agreement, but provides consistent deviation. Multi-state optimization demonstrates quantitative agreement over a range of states the potential was not originally optimized against.



Structural metrics from CG and united atom n-dodecane simulations, using potentials from multi-state (top) and single-state (bottom) IBI. Each point represents a certain state point. Points with an exscribed square represent states used in the potential derivation. A point lying on the y=x line represents a perfect match between the united atom and CG simulations.

Lipid bilayers

Using tail potentials from fatty acid simulations and a water potential from previous work,[3] we optimized the headgroup bead potentials for a 2-tailed skin lipid using fluid and crystalline structure of pure lipids. Varying the weighting factors, $\alpha_s(r)$, used in the potential derivation alters the bilayer properties, *e.g.,* area per lipid (APL), enabling the model to be tuned to match properties beyond the RDF.



water

hydrophilic headgroups

hydrophobic tails

$\alpha_1/\alpha_2/\alpha_3$	Thickness(Å)	Chain tilt (deg)	APL (Å²)
0.7/0.4/0.4	42.7	15.7	43.4
0.8/0.3/0.3	42.0	17.9	44.2
Atomistic	44.1	20.5	42.1

Bilayer properties from various forcefields. α_i is the α value for state i used in the potential derivation. State 1: crystal; state 2: room temperature/experimental density; state 3: 450K/experimental density.

References and Acknowledgements

1. Reith, D.; Pütz, M.; Müller-Plathe, F. *Journal of Computational Chemistry* 2003, *24*, 1624– 1636.
2. Pu, Q.; Leng, Y.; Zhao, X.; Cummings, P. T. *Nanotechnology* 2007, *18*, 424007.
3. K. R. Hadley and C. McCabe, J. Phys. Chem. B, 2010, 114, 4590– 4599.

We would like to thank the National Institute of Arthritis and Musculoskeletal and Skin Diseases for funding this work.

Conclusion

We extend the IBI method to yield less state-dependent potentials by including target data from multiple states. We show the multi-state extension is able to recover a known potential where the single-state method cannot. A potential is derived for a single-site model of propane, showing good agreement with a Lennard-Jones particle mapped to the critical point. Structural properties of n-dodecane are better estimated with potentials derived using the multi-state extension than the single-state method. Finally, adjusting relative weighting factors of the states used can alter a system's properties, allowing dynamic adjustment of the weighting factors to match bulk properties of a system of interest.