Multistate Bennett Acceptance Ratio replaces histogram reweighting for vapor-liquid coexistence calculations; Multistate Bennett Acceptance Ratio to enable rapid force field parameterization; Multistate Bennett Acceptance Ratio as a substitute for histogram reweighting when optimizing non-bonded parameters

Richard Messerly[1], Mohammad S. Barhaghi[2], Jeffrey J. Potoff[2], and Michael R. Shirts[3]

[1]*Thermodynamics Research Center (TRC), National Institute of Standards and Technology (NIST), Boulder, Colorado, 80305, United States*
[2]*Department of Chemical Engineering and Materials Science, Wayne State University, Detroit, Michigan 48202, United States*
[3]*Department of Chemical and Biological Engineering, University of Colorado, Boulder, Colorado, 80309, United States*

ABSTRACT

Key points

1. MBAR is equivalent to standard histogram reweighting approaches in the limit of zero bin width
2. MBAR allows for estimating VLE properties of any force field
3. Scaling epsilon values post-simulation is straightforward with MBAR
4. Basis functions are an efficient means for re-computing energy of Mie potentials

Outline

1 Introduction

1. Accurate and efficient computation of vapor-liquid coexistence is an important but challenging task for molecular simulation
2. Reweighting simulation outputs between different states is a well-known and powerful tool, e.g., histogram reweighting of GCMC results
3. Force field parameterization with VLE data is an arduous and time-consuming task
4. Hamiltonian scaling (histogram reweighting for multiple force fields) allows for estimating VLE properties of multiple force fields from single set of simulations
5. Messerly et al. demonstrated how MBAR can be combined with ITIC to predict VLE properties. Several Weakness of ITIC is need large systems, which is not ideal for MBAR

6. Gross demonstrated benefits of ε-scaling for "individualized", i.e., compound-specific parameter sets
7. In this study, we demonstrate that:
   (a) MBAR yields indistinguishable results from histogram reweighting (HR)
   (b) Scaling epsilon is straightforward by scaling U with MBAR
   (c) MBAR can estimate VLE properties for multiple force fields simultaneously
   (d) Basis functions allow for rapid computation of VLE for non-simulated Mie parameter sets

2 Methods

2.1 Force fields

1. Simulations are performed for united-atom generalized Lennard-Jones (a.k.a., Mie λ-6) force fields
2. We investigate the TraPPE, Potoff-generalized, Potoff (S/L), and NERD force fields
3. Details of force fields

2.2 Simulation set-up

1. Simulations performed by Mick et al. are reanalyzed using MBAR
2. Additional simulations are performed in GCMC ensemble using GPU optimized Monte Carlo (GOMC)
3. Simulation specifications, i.e., box size, number of steps, type of moves, etc.
4. State points (chemical potentials and temperatures) simulated are same as those utilized in Mick et al.

2.3 Multistate Bennett Acceptance Ratio

1. Traditionally, histogram reweighting (HR) has been applied with GCMC to calculate vapor-liquid coexistence properties
2. Present histogram reweighting equations
3. Discuss how to compute phase equilibria by equating pressures
4. Discuss how to compute heat of vaporization
5. In this study, we demonstrate how to compute VLE using MBAR-GCMC
6. Procedure is identical to that utilized for HR but using the MBAR equations
7. Present MBAR equations
8. MBAR for $\theta = \theta_{\text{ref}}$ is mathematically equivalent to histogram reweighting in the limit of zero bin width
9. MBAR-GCMC allows for prediction of multiple force fields from single simulation without modifying force fields mid-simulation (i.e., Hamiltonian scaling approach)
10. MBAR uncertainties are computed using bootstrap resampling

2.4 Basis functions

1. When applying MBAR to different parameter sets, $\theta \neq \theta_{\text{ref}}$, it is necessary to recompute energies
2. Basis functions accelerate the recompute energy step by storing the repulsive and attractive contributions that can be scaled by ε and σ
3. Basis functions are computed from GOMC using the recompute feature for different ε and σ and solving system of equations

3 Results

1. We validate that MBAR and HR are indistinguishable by re-analyzing the simulation results of Mick et al. and Barhaghi et al. utilizing MBAR
2. Epsilon scaling for all the compounds that Mohammad has U and N values for (branched alkanes and alkynes) and which have good experimental data
3. We estimate Potoff generalized and NERD VLE from TraPPE simulations, Potoff S/L from Potoff generalized, and TraPPE from Potoff generalized
4. For $\lambda_{\text{ref}} = 12$ and $\lambda_{\text{rr}} = 16$, MBAR-GCMC predicts vapor density, vapor pressure, and heat of vaporization more accurately than liquid density
5. For $\lambda_{\text{ref}} = 12$ and $\lambda_{\text{rr}} = 12$, i.e., computing NERD from TraPPE simulations, MBAR-GCMC predicts all four properties accurately
6. We present how basis functions allow for rapid computation of wide range of parameter sets:
   (a) *n*-hexane
   (b) 2-methylpropane
   (c) 2,2-dimethylpropane
   (d) cyclopentane or cyclohexane
7. We provide supporting information with basis functions for several branched alkanes with TraPPE and Potoff force fields

3.1 Figures

1. Percent deviation between MBAR and HR results for rholiq, rhovap, Psat, and DeltaHv
2. Comparison between MBAR bootstrapping and analytical uncertainties and HR uncertainties
3. Scaling of epsilon post-simulation for branched alkanes and alkynes
4. Prediction of VLE for $\lambda_{\text{ref}} \neq \lambda_{\text{rr}}$
5. Prediction of VLE for $\lambda_{\text{ref}} = \lambda_{\text{rr}} = 12$
6. Prediction of VLE for $\lambda_{\text{ref}} = \lambda_{\text{rr}} = 16$
7. Two-D scans of scoring functions for $\varepsilon - \sigma$ of CH3 (a) and CH2 (b) for *n*-hexane

8. Two-D scans of scoring functions for $\varepsilon - \sigma$ of CH3 (a) and CH (b) for 2-methylpropane
9. Two-D scans of scoring functions for $\varepsilon - \sigma$ of CH3 (a) and C (b) for 2,2-dimethylpropane
10. Two-D scans of scoring functions for $\varepsilon - \sigma$ of CH2 for cyclopentane or cyclohexane (reference is TraPPE)

4 Discussion/Limitations/Future work

1. We recommend that future GCMC-VLE studies report the snapshots of N and U and/or basis functions to recompute U as this allows for future force field optimization
2. Improvements are possible with multiple θ or simulating a range of μ values

5 Conclusions

6 Acknowledgments

Mostafa and J. Richard Elliott provided valuable insights.

7 Supporting Information

7.1 MBAR VLE estimates

Provide tables of MBAR estimates

7.2 Basis functions

1. Validation that basis functions give accurate energies

7.3 Raw data

1. Comparison of 2-D histograms for TraPPE and Potoff. MBAR overlap, possible? Probably not without rerunning the simulations.