Estimating vapor-liquid coexistence properties of non-simulated force fields with histogram-free reweighting

Richard A. Messerly,[*,†] Mohammad Soroush Barhaghi,[‡] Jeffrey J. Potoff,[‡] and Michael R. Shirts[¶]

†*Thermodynamics Research Center, National Institute of Standards and Technology, Boulder, Colorado, 80305, United States*

‡*Department of Chemical Engineering and Materials Science, Wayne State University, Detroit, Michigan 48202, United States*

¶*Department of Chemical and Biological Engineering, University of Colorado Boulder, Colorado, 80309, United States*

E-mail: richard.messerly@nist.gov

Abstract

The prediction of accurate vapor-liquid coexistence properties with molecular simulation relies on well-parameterized force fields. Histogram reweighting (HR) is a standard approach for converting Grand Canonical Monte Carlo (GCMC) simulation output into vapor-liquid coexistence properties (saturated liquid density, $\rho_{\text{liq}}^{\text{sat}}$, saturated vapor density, $\rho_{\text{vap}}^{\text{sat}}$, saturated vapor pressures, $P_{\text{vap}}^{\text{sat}}$, and enthalpy of vaporization, ΔH_{v}). Due to

the abundance of experimental vapor-liquid coexistence data and the sensitivity of such properties to both short- and long-range non-bonded interactions, numerous force fields have been parameterized using $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and/or ΔH_{v}. Unfortunately, computing $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and/or ΔH_{v} for each proposed parameter set requires a large amount of GCMC simulations.

We demonstrate how a histogram-free reweighting approach can alleviate the computational burden of force field parameterization. Specifically, we utilize the Multistate Bennett Acceptance Ratio (MBAR) to estimate coexistence properties for non-bonded parameters without simulating each parameter set directly. MBAR thus greatly reduces the amount of GCMC simulations that are required for parameterizing the non-bonded potential.

First, we validate that GCMC-MBAR and GCMC-HR yield statistically indistinguishable results when computing coexistence properties for only the simulated force field. Second, we utilize GCMC-MBAR to optimize compound-specific parameters, where simulations are only performed with the initial parameter set. Third, we predict coexistence properties of force field j by simulating force field i, where i and j are common force fields from the literature. Fourth, we simulate cyclohexane with a common united-atom Lennard-Jones 12-6 potential (Transferable Potentials for Phase Equilibria, TraPPE) to predict coexistence properties over a wide range of Mie λ-6 (generalized Lennard-Jones) parameters (ϵ, σ, and λ). We provide guidelines for determining the accuracy of GCMC-MBAR predicted $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and/or ΔH_{v} values for non-simulated force fields.

1   Introduction

A key use of molecular simulation is the ability to accurately and efficiently estimate vapor-liquid coexistence properties, i.e., saturated liquid density ($\rho_{\text{liq}}^{\text{sat}}$), saturated vapor

density ($\rho_{\text{vap}}^{\text{sat}}$), saturated vapor pressures ($P_{\text{vap}}^{\text{sat}}$), and enthalpy of vaporization (ΔH_{v}). The accuracy of coexistence estimates depends on the underlying molecular model (a.k.a., force field, potential model, or Hamiltonian) while the computational efficiency depends primarily on the simulation methods, software, and hardware.

Several simulation approaches exist for computing vapor-liquid coexistence properties.[1] These include Gibbs Ensemble Monte Carlo (GEMC), two-phase molecular dynamics (2ϕMD),[2] isothermal-isochoric integration (ITIC),[3] and Grand Canonical Monte Carlo coupled with histogram reweighting (GCMC-HR).[1,4,5] The improved efficiency of these methods has greatly enabled the development of accurate force fields.[6–10] However, parameterization of non-bonded interactions with vapor-liquid coexistence calculations over a wide range of temperatures remains an arduous and time-consuming task. For example, recent studies have implemented an exhaustive grid-based search optimization which requires performing GCMC-HR simulations with hundreds of non-bonded parameter sets.[11–13]

The primary motivation for this work is to reduce the computational cost of optimizing non-bonded parameters with vapor-liquid saturation properties. This is achieved by substituting histogram reweighting with the Multistate Bennett Acceptance Ratio (MBAR),[14,15] a histogram-free reweighting schema. The proposed GCMC-MBAR method is identical to the traditional GCMC-HR approach except that it uses MBAR to reweight configurations rather than reweighting histograms. The benefit of this simple modification is that GCMC-MBAR can estimate coexistence properties for non-bonded parameter sets that have not been simulated directly.

In related studies, Messerly et al. demonstrate how to combine MBAR with ITIC (MBAR-ITIC) to optimize generalized Lennard-Jones (Mie λ-6) potentials.[16,17] For MBAR-ITIC, a series of constant number of molecules, constant volume, and constant temperature (NVT) simulations are performed along an isotherm and isochore(s) with a "refer-

ence" force field(s) (θ_{ref}). MBAR reweights configurations sampled with θ_{ref} to estimate the internal energy (U) and pressure (P) at each (T, ρ) state point for a non-simulated ("rerun") force field (θ_{rr}), with only single stored snapshots from simulations with (θ_{ref}) evaluated using simulations with (θ_{rr}). ITIC then converts the U and P estimates into vapor-liquid coexistence properties.[3,18]

The results from Messerly et al. demonstrate that MBAR-ITIC is most reliable in the local domain, i.e., when $\theta_{\text{rr}} \approx \theta_{\text{ref}}$.[16] Furthermore, MBAR-ITIC performs best for changes in the non-bonded well-depth parameter (ϵ) while it performs significantly worse for large changes in the non-bonded size (σ) and repulsive exponent (λ) parameters. This "poor overlap" when $\theta_{\text{rr}} \not\approx \theta_{\text{ref}}$ can be quantified by the number of effective snapshots ($K_{\text{snaps}}^{\text{eff}}$), which is essentially the number of non-negligible configurations that contribute to the estimated ensemble averages. Poor overlap (low $K_{\text{snaps}}^{\text{eff}}$) is especially problematic for MBAR-ITIC as a large number of snapshots is needed to obtain precise estimates of P in the liquid phase, due to large fluctuations in P at high densities.

Our initial hypothesis was that GCMC-MBAR would experience better overlap than was observed for MBAR-ITIC when $\theta_{\text{rr}} \not\approx \theta_{\text{ref}}$. There were two main reasons for this hypothesis/aspiration. First, as opposed to the fixed density NVT simulations used in ITIC, the fluctuating density in GCMC simulations samples from a wider range of configurations and energies. Second, ITIC requires larger box sizes (and, thereby, more molecules) than those typically utilized with GCMC. By utilizing fewer molecules, GCMC simulations experience larger energy fluctuations (on a percent basis) which improves the overlap between states. We also hypothesized that the impact of poor overlap would be less severe compared to MBAR-ITIC, where poor overlap leads to sporadic values of P and nonsensical coexistence estimates.

The method outlined in this study is similar in spirit to "Hamiltonian scaling" (HS), which has been applied to both GEMC[19] and GCMC simulations.[1,20–22] The HS approach

samples from multiple force fields (Hamiltonians) in a single simulation according to a weighted sampling probability. Vapor-liquid coexistence curves for each force field are estimated post-simulation by reweighting the configurations accordingly. For the GCMC implementation of Hamiltonian scaling (HS-GCMC), μ and T are not stationary during the simulation, rather the current value of μ and T depends on which force field is being sampled. Despite HS-GCMC proving to be a powerful tool to optimize force field parameters,[1,20–22] it has yet to gain widespread popularity. This is likely due to the added complexity of both the simulation protocol and the histogram post-processing. Also, HS requires that a decision be made *a priori* regarding which force fields are to be tested. By contrast, GCMC-MBAR does not require any modification of the simulation procedure, the post-processing is essentially unchanged, and the non-bonded parameter sets need not be selected prior to the simulations.

The outline for this study is the following. Section 2 provides details regarding the force fields, simulation set-up, and the HR/MBAR post-simulation analysis. Section 3 presents results of GCMC-MBAR for four scenarios in Sections 3.1 to 3.4. Section 3.1 validates that GCMC-MBAR and GCMC-HR yield indistinguishable coexistence estimates for a fixed force field. Section 3.2 applies GCMC-MBAR to a recently proposed ϵ-scaling approach. Section 3.3 shows how GCMC-MBAR can predict coexistence properties for force field j by reweighting configurations sampled with force field i. Section 3.4 demonstrates how GCMC-MBAR can be utilized to rapidly optimize the united-atom Mie λ-6 parameters for cyclohexane. Section 4 discusses some limitations and provides recommendations for future work. Section 5 reviews the primary conclusions.

2 Methods

2.1 Force fields

The force fields utilized in this study are Transferable Potentials for Phase Equilibria (TraPPE-UA, referred to simply as TraPPE[6,23,24]), Mie Potentials for Phase Equilibria (MiPPE),[8,12,13] and Nath, Escobedo, and de Pablo revised (NERD).[25,26] Each force field adopts a united-atom (UA) representation, where non-polar hydrogens are not modeled explicitly.

For computational efficiency, we utilize fixed bond lengths for each force field. Note that the original NERD publications utilized a harmonic bond potential.[25] The bond lengths for all bonds, compounds, and force fields investigated in this study are reported in Table 1.

Table 1: Equilibrium (fixed) bond lengths (r_{eq}).[12,13,23,26] CH_x and CH_y represent CH_3, $CH_2(sp^3)$, $CH(sp^3)$, or $C(sp^3)$ sites.

Bending sites	r_{eq} (nm)		
	TraPPE	MiPPE	NERD
CH_x-CH_y	0.154	0.154	0.154
$C(sp)$-CH_x	–	0.146	–
$CH{\equiv}CH$	–	0.121	–
$C{\equiv}CH$	–	0.121	–

Angular bending interactions for each force field are evaluated using a harmonic potential:

$$u^{bend} = \frac{k_\theta}{2} \left(\theta - \theta_{eq} \right)^2 \tag{1}$$

where u^{bend} is the bending energy, θ is the instantaneous bond angle, θ_{eq} is the equilibrium bond angle, and k_θ is the harmonic force constant. Table 2 provides the θ_{eq} and k_θ values for each angle type and force field. Note that the θ_{eq} and k_θ values are equivalent for the

6

three force fields, with the exception of the NERD CH_x-CH-CH_y θ_{eq} value.

Table 2: Equilibrium bond angles (θ_{eq}) and force constants (k_θ/k_B), where k_B is the Boltzmann constant.[12,13,23,26] CH_x and CH_y represent CH_3, $CH_2(sp^3)$, $CH(sp^3)$, or $C(sp^3)$ sites.

Bending sites	θ_{eq} (degrees)			k_θ/k_B (K/rad^2)
	TraPPE	MiPPE	NERD	
CH_x-CH_2-CH_y	114.0	114.0	114.0	62500
CH_x-CH-CH_y	112.0	112.0	109.5	62500
CH_x-C-CH_y	109.5	109.5	109.5	62500
CH_x-CH_2-C(sp)	–	112	–	62500
CH_x-C(sp)≡CH	–	180	–	30800
CH_x-C(sp)≡C	–	180	–	30800

Dihedral torsional interactions for each force field are determined using a cosine series:

$$u^{tors} = c_0 + c_1 \left(1 + \cos\phi\right) + c_2 \left(1 - \cos 2\phi\right) + c_3 \left(1 + \cos 3\phi\right) \tag{2}$$

where u^{tors} is the torsional energy, ϕ is the dihedral angle and c_n are the Fourier constants listed in Table 3. In accordance with Reference 27, we simulate cyclohexane using the TraPPE CH_x-CH_2-CH_2-CH_y torsional parameters instead of the TraPPE C-C-C-C six-member ring torsional parameters reported in Table 3 of Reference 24. This choice is made to better replicate the vapor-liquid coexistence densities reported in Reference 24. After investigating the dihedral distributions for cyclohexane obtained with the torsional parameters from Reference,[24] we suspect there is a sign error for at least the c_1 term.

Non-bonded interactions between sites located in two different molecules or separated by more than three bonds within the same molecule are calculated using a Mie λ-6 potential (of which the Lennard-Jones, LJ, 12-6 is a subclass):

$$u^{nb}(\epsilon, \sigma, \lambda; r) = \left(\frac{\lambda}{\lambda - 6}\right) \left(\frac{\lambda}{6}\right)^{\frac{6}{\lambda-6}} \epsilon \left[\left(\frac{\sigma}{r}\right)^\lambda - \left(\frac{\sigma}{r}\right)^6\right] \tag{3}$$

7

Table 3: Fourier constants (c_n/k_B) in units of K.[12,13,23,26] CH_x and CH_y represent CH_3, $CH_2(sp^3)$, $CH(sp^3)$, or $C(sp^3)$ sites.

Torsion sites	c_0/k_B	c_1/k_B	c_2/k_B	c_3/k_B
CH_x-CH_2-CH_2-CH_y	0.0	355.03	-68.19	791.32
CH_x-CH_2-CH-CH_y	-251.06	428.73	-111.85	441.27
CH_x-CH_2-C-CH_y	0.0	0.0	0.0	461.29
CH_x-CH-CH-CH_y	-251.06	428.73	-111.85	441.27
CH_x-CH_2-CH_2-C(sp)	94.88	162.00	-205.40	980.40
CH_x-CH_2-C(sp)≡C(sp)	0	0	0	0
CH_x-CH_2-C(sp)≡CH(sp)	0	0	0	0
CH_x-C(sp)≡C(sp)-CH_y	0	0	0	0

where u^{nb} is the non-bonded energy, σ is the distance (r) where $u^{nb} = 0$, $-\epsilon$ is the energy of the potential at the minimum $\left(\text{i.e., } u^{nb} = -\epsilon \text{ and } \frac{\partial u^{nb}}{\partial r} = 0 \text{ for } r = r_{min}\right)$, and λ is the repulsive exponent.

The non-bonded Mie λ-6 force field parameters for TraPPE, MiPPE, and NERD are provided in Table 4. MiPPE reports a "generalized" (MiPPE-gen) and "short/long" (MiPPE-SL) CH and C parameter set. The "short" and "long" parameters are implemented when the number of carbons in the backbone is ≤ 4 and > 4, respectively. Also note that the NERD force field has several different CH_3 non-bonded parameter sets.

Non-bonded parameters between two different site types (i.e., cross-interactions) are determined using Lorentz-Berthelot combining rules[28] for ϵ and σ and an arithmetic mean for the repulsive exponent λ (as recommended in Reference 8):

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{4}$$

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{5}$$

$$\lambda_{ij} = \frac{\lambda_{ii} + \lambda_{jj}}{2} \tag{6}$$

Table 4: Non-bonded (Mie λ-6) parameters for TraPPE,[6,23,24] MiPPE,[8,12,13] and NERD.[25,26]

United-atom	ϵ/k_{B} (K)	σ (nm)	λ
TraPPE			
CH_3	98	0.375	12
$CH_2(sp^3)$	46	0.395	12
$CH(sp^3)$	10	0.468	12
$C(sp^3)$	0.5	0.640	12
CH_2 (cyclohexane)	52.5	0.391	12
MiPPE			
CH_3	121.25	0.3783	16
$CH_2(sp^3)$	61	0.399	16
$CH(sp^3)$, gen.	15	0.46	16
$C(sp^3)$, gen.	1.2	0.61	16
$CH(sp^3)$, short	15	0.47	16
$C(sp^3)$, short	1.45	0.61	16
$CH(sp^3)$, long	14	0.47	16
$C(sp^3)$, long	1.2	0.62	16
$CH(sp)$	148.5	0.357	28
$C(sp)$ (1-alkyne)	206	0.2875	16
$C(sp)$ (2-alkyne)	118	0.312	16
CH_2 (cyclohexane)[a]	69.5	0.391	16
NERD			
CH_3 (general)	104.00	0.3910	12
CH_3 (2-methylpropane)	78.23	0.3880	12
CH_3 (2,2-dimethylpropane)	74.50	0.3910	12
CH_3 (methyl side chain)	70.00	0.3850	12
CH_3 (ethyl side chain)	83.00	0.3820	12
$CH_2(sp^3)$	45.80	0.3930	12
$CH(sp^3)$	39.70	0.3850	12
$C(sp^3)$	17.00	0.3910	12
[a] This work. See Section 3.4			

where the ij subscript refers to cross-interactions and the subscripts ii and jj refer to same-site interactions.

2.2 Simulation set-up

The results presented in Sections 3.1 and 3.2 are obtained by reprocessing simulation output that were analyzed in previous studies with histogram reweighting.[12,13] New simulation results are provided in Sections 3.3 and 3.4 for 2-methylpropane, 2,2-dimethylpropane, 2,2-dimethylbutane, 3,3-dimethylhexane, 3-methyl-3-ethylpentane, 2,2,4-trimethylhexane, 2,3-dimethylbutane, 2,3,4-trimethylpentane, and cyclohexane.

Each compound is simulated with Grand Canonical Monte Carlo (GCMC), i.e., constant chemical potential (μ), volume (V), and temperature (T). A series of nine simulations are performed, two in the vapor phase, six in the liquid phase, and one near critical which acts as the "bridge" between the vapor and liquid phases. A single simulation is performed at each state point (μ, V, and T), unless otherwise stated. A low-density (less than twenty molecules) initial configuration is utilized for the vapor phase simulations, while the bridge and liquid phase simulations are initialized with a high-density (around 150 molecules) configuration.

The system volume is constant with respect to μ and T for a given compound. The cubic box side length is 3 nm for 2-methylpropane, 2,2-dimethylpropane, 2,3-dimethylbutane, and cyclohexane, 3.5 nm for 2,3,4-trimethylpentane and 3,3-dimethylhexane, and 4 nm for 2,2,4-trimethylhexane and 3-methyl-3-ethylpentane. The prescribed μ, T, and V values for the branched alkanes are the same as those utilized in Mick et al. and vary somewhat between force fields.[12] All simulated state points are reported in Supporting Information.

Each GCMC simulation performed in this study consists of an equilibration stage of 2×10^7 Monte Carlo steps (MCS). The production stage is 4×10^7 MCS for vapor simulations and 2.5×10^7 MCS for the liquid and "bridge" simulations. Snapshots (i.e., number of molecules, internal energy, and optionally the xyz coordinates) are stored every 200 MCS to reduce the correlation between sequential configurations. Thus, the number of snapshots (K_{snaps}) is 2×10^5 for vapor simulations and 1.25×10^5 for liquid and "bridge"

10

simulations.

The type of Monte Carlo move implemented for each step is selected randomly with a 30%, 10%, and 60% probability of performing a displacement, rotation, and particle swap move, respectively. Displacement and rotation moves are required to thermally equilibrate the system at the simulation temperature while particle swap moves ensure that μ is equal to the prescribed value. Cyclohexane simulations also employ crank-shaft moves to sample internal configurations, which can be challenging for ring molecules.[29] The move probabilities for cyclohexane are 30%, 10%, 40%, and 20% for displacement, rotation, particle swap, and crank-shaft moves, respectively. All simulations utilize coupled-decoupled configurational-bias Monte Carlo (CBMC)[23] to enhance the insertion acceptance rate, with 100 angle trials, 30 dihedral trials, 10 initial site trials, and 4 subsequent site trials. The move probabilities are consistent with those of Mick et al. and Soroush Barhaghi et al.,[12,13] while the CBMC dihedral trials and initial site trials differ slightly.

All simulations are performed using GPU optimized Monte Carlo (GOMC)[30] version 2.31, compiled using GCC with OpenMP enabled. CPU GCMC simulations are run on a Linux 4.4.0-112-generic x86_64 on an Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20GHz machine. Initial configurations are generated with Packmol,[31] while psfgen is used to generate the coordinate (*.pdb) and connectivity (*.psf) files.[32] Example GOMC input files with corresponding shell and Python scripts for preparing, running, and analyzing simulations are provided at https://github.com/ramess101/MBAR_GCMC.

2.3   GCMC-MBAR compared with GCMC-HR

Converting the GCMC simulation output into phase coexistence properties requires significant post-processing through reweighting. Histogram reweighting (HR) and, more generally, configuration reweighting is an important tool in many fields of molecular simulation. In fact, it has long since been known that it is possible to estimate properties for

11

state j by reweighting configurations that were sampled with state i.[1,33–35] For example, umbrella sampling simulations are often processed using the weighted histogram analysis method (WHAM) to compute free energy differences between states.[36] WHAM (or HR) is essentially an approximation of MBAR and, therefore, MBAR should be favored whenever a histogram-free approach is feasible.[37] In this study, we implement MBAR in Python 2.7 through the *pymbar* package.[38]

Before demonstrating how to compute vapor-liquid coexistence with GCMC-MBAR, we review the traditional GCMC-HR approach. We also discuss the steps of this procedure that are the same for both GCMC-HR and GCMC-MBAR. We refer the interested reader to the literature for derivations and more detailed discussion of the GCMC-HR equations (cf. Reference 1).

2.3.1 Histogram reweighting

The ensemble average for a given observable (O, e.g., U and $\frac{N}{V}$) is computed with HR according to

$$\langle O(\mu, \beta) \rangle = \sum_U \sum_N O \times Pr(N, U; \mu, \beta) \tag{7}$$

where $\langle \dots \rangle$ denotes an ensemble average and $Pr(N, U; \mu, \beta)$ is the probability of observing N particles with internal energy U for a given chemical potential (μ) and inverse temperature ($\beta \equiv \frac{1}{k_\mathrm{B}T}$, where k_B is the Boltzmann constant). The double summation is computed numerically where U and N are discritized into a 2-dimensional histogram. The probability is obtained with HR from

$$Pr(N, U; \mu, \beta) = \frac{\sum_{i=1}^{R} Pr_i(N, U) \exp(-\beta U + \beta \mu N)}{\sum_{i=1}^{R} K_{\mathrm{snaps},i} \exp(-\beta_i U + \beta_i \mu_i N + \hat{f}_i)} \tag{8}$$

where $Pr_i(N, U)$ is the probability of observing N particles and U within the histogram bin width, R is the number of runs (where the ith run corresponds to β_i and μ_i), $K_{\text{snaps},i}$ is the number of observations ("snapshots") for run i (i.e., $K_{\text{snaps},i} = \sum_{N,U} Pr_i(N, U)$), and $\hat{f}_i$ is an estimate for the reduced free energy, which is calculated with the relationship

$$\hat{f}(\mu, \beta) = -\ln \sum_U \sum_N \frac{\sum_{i=1}^{R} Pr_i(N, U) \exp(-\beta U + \beta \mu N)}{\sum_{i=1}^{R} K_{\text{snaps},i} \exp(-\beta_i U + \beta_i \mu_i N + \hat{f}_i)} \tag{9}$$

where $\hat{f}_i \equiv \hat{f}(\mu_i, \beta_i)$. Note that because $\hat{f}_i$ can also be viewed simply as a constant, the literature typically adopts the notation C_i (or more specifically, $-C_i$) instead of $\hat{f}_i$. For a clear comparison with the MBAR expressions, we prefer $\hat{f}_i$.

2.3.2 Histogram-free reweighting

Equations 7, 8, and 9 only allow for reweighting simulations at a different β and μ. By contrast, MBAR can also be applied to reweight simulations for different force field parameters (θ). The analogous MBAR equation to Equation 7 is

$$\langle O(\theta, \mu, \beta) \rangle = \sum_{n=1}^{K_{\text{snaps}}^{\text{tot}}} O(\boldsymbol{x}_n, N_n; \theta, \mu, \beta) \times W_n(\theta, \mu, \beta) \tag{10}$$

where $\boldsymbol{x}_n, N_n$ are configurations sampled from $i = 1 \ldots R$ simulations at inverse temperature (β_i), chemical potential (μ_i), and force field parameters (θ_i), and $K_{\text{snaps}}^{\text{tot}} = \sum_{i=1}^{R} K_{\text{snaps},i}$ is the total number of snapshots for all R runs. $W_n(\theta, \beta, \mu)$ is the weight of the nth configuration in a simulation with arbitrary μ, β, and θ. W_n is computed with the following expression (analogous to Equation 8)

$$W_n(\theta, \beta, \mu) = \frac{\exp[\hat{f}(\theta, \beta, \mu) - u(\boldsymbol{x}_n, N_n; \theta, \beta, \mu)]}{\sum_{i=1}^{R} K_{\text{snaps},i} \exp[\hat{f}(\theta_i, \beta_i, \mu_i) - u(\boldsymbol{x}_n, N_n; \theta_i, \beta_i, \mu_i)]} \tag{11}$$

where $u(\boldsymbol{x}_n, N_n; \theta, \beta, \mu)$ is the reduced potential energy evaluated with θ, β, and μ for configuration $\boldsymbol{x}_n, N_n$. The reduced free energy is computed with an expression analogous to Equation 9

$$\hat{f}(\theta, \beta, \mu) = -\ln \sum_{n=1}^{K_{\text{snaps}}^{\text{tot}}} \frac{\exp[-u(\boldsymbol{x}_n, N_n; \theta, \beta, \mu)]}{\sum\limits_{i=1}^{R} K_{\text{snaps},i} \exp[\hat{f}(\theta_i, \beta_i, \mu_i) - u(\boldsymbol{x}_n, N_n; \theta_i, \beta_i, \mu_i)]} \tag{12}$$

For the grand canonical ensemble, the reduced potential energy is

$$u(\boldsymbol{x}_n, N_n; \theta, \beta, \mu) = \beta U(\boldsymbol{x}_n, N_n; \theta) - \beta \mu N_n \tag{13}$$

2.3.3   Comparison between HR and MBAR

Both HR and MBAR require solving a system of $R - 1$ nonlinear equations for self-consistency (Equations 8 and 9 for HR and Equations 11 and 12 for MBAR). Specifically, initial guesses for $\hat{f}_i$ are updated iteratively until convergence. There is provably only one solution,[15] so as long as certain criteria are met that will be discussed below. Thus, although a range of different solver methods exist, the only difference is efficiency and numerical stability.

The similarities between MBAR (Equations 11 and 12) and HR (Equations 8 and 9) are readily apparent after substituting Equation 13 into Equations 11 and 12. Indeed, the difference between HR and MBAR is primarily that of bookkeeping, although the histogram-free approach of MBAR does have some benefits when varying force field parameters, as discussed below.

The key advantage of MBAR over HR is that by changing from a sum over histograms to a sum over snapshots, we are free to perform simulations with other conditions besides μ and β. For example, we can carry out simulations at different force field parameters

14

(θ_i). In the histogram context, we cannot easily separate out snapshots in the same U histogram that were performed with different force field parameters. Furthermore, as MBAR performs a sum over snapshots, we can reevaluate the configurational energy, $U(\boldsymbol{x}_n, N_n; \theta_{\mathrm{rr}})$, with a range of different "rerun" parameter sets (θ_{rr}).

The θ dependence of Equations 10, 11, and 12 can be removed by performing all R simulations with a single reference force field (θ_{ref}) and only computing W_n and $\hat{f}$ for $\theta = \theta_i = \theta_{\mathrm{ref}}$. With this simplification, these two sets of equations (HR: Equations 7, 8, and 9, MBAR: Equations 10, 11, and 12) can be seen as equivalent in the limit of infinitesimal histogram bin widths. In the zero bin width limit, no histogram contains more than 1 snapshot and, therefore, U and N for each histogram can be taken to be the $U(\boldsymbol{x}_n, N_n)$ and N_n of the single observation in that histogram, while histograms with no snapshots can be omitted. Thus, $Pr_i(N, U)$ is either 1 or 0, and the sum over all histograms becomes a sum over snapshots conducted in all R simulation runs. Equations 8 and 9 then reduce to Equations 11 and 12, respectively.

2.3.4 Computing saturation properties

For both HR and MBAR, the pressure is computed from

$$P(\theta, \beta, \mu) = \frac{k_{\mathrm{B}}T}{V} \ln \Xi + B = \hat{f}(\theta, \beta, \mu) + B \tag{14}$$

where Ξ is the grand canonical partition function and B is an additive constant. For both HR and MBAR, B is determined by fitting a straight-line to $\ln \Xi$ with respect to N at very low densities. At these low densities, the system is assumed to behave as an ideal gas and, therefore, the slope is unity and $B = \frac{k_{\mathrm{B}}T}{V} \times b$, where b is the y-intercept from the straight-line regression.

Having determined B, the saturated vapor pressure $(P_{\mathrm{vap}}^{\mathrm{sat}})$ is computed with Equation

14 at the desired saturation temperature (T^{sat}) and corresponding saturation chemical potential (μ^{sat}). μ^{sat} is obtained by equating the pressures in the vapor and liquid phases at a fixed value of T^{sat}. This is done by integrating Pr (HR) or W_n (MBAR) for the two phases separately, i.e., by dividing the snapshots into low and high density regimes. For example, the equality of pressures is satisfied for HR when

$$\sum_{U} \sum_{N > N_{\text{c}}} Pr(N, U; \mu^{\text{sat}}, \beta^{\text{sat}}) = \sum_{U} \sum_{N \leq N_{\text{c}}} Pr(N, U; \mu^{\text{sat}}, \beta^{\text{sat}}) \tag{15}$$

where N_{c} is an estimate for the number of molecules at the critical density, which serves to distinguish between snapshots that are in the vapor or liquid phases. The analogous MBAR equation is

$$\sum_{n=1}^{K_{\text{snaps}}^{\text{liq}}} W_n(\theta, \beta^{\text{sat}}, \mu^{\text{sat}}) = \sum_{n=1}^{K_{\text{snaps}}^{\text{vap}}} W_n(\theta, \beta^{\text{sat}}, \mu^{\text{sat}}) \tag{16}$$

where $K_{\text{snaps}}^{\text{liq}}$ and $K_{\text{snaps}}^{\text{vap}}$ are the number of liquid and vapor snapshots, respectively.

By solving Equations 15 (HR) or 16 (MBAR) for μ^{sat}, the vapor and liquid saturation densities and energies ($\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $U_{\text{liq}}^{\text{sat}}$, and $U_{\text{vap}}^{\text{sat}}$) are also computed with a modified version of Equations 7 (HR) or 10 (MBAR) where only snapshots from the desired phase are included in the weighted average. For example, when computing $\rho_{\text{liq}}^{\text{sat}}$ and $U_{\text{liq}}^{\text{sat}}$, the double summation in Equation 7 (HR) is performed only for $N > N_{\text{c}}$ and the sum in Equation 10 (MBAR) is only over $K_{\text{snaps}}^{\text{liq}}$ liquid snapshots.

Having computed the pressure, internal energies, and densities for the saturated vapor and saturated liquid, the enthalpy of vaporization is calculated with

$$\Delta H_{\text{v}} = \bar{U}_{\text{vap}}^{\text{sat}} - \bar{U}_{\text{liq}}^{\text{sat}} + P_{\text{vap}}^{\text{sat}}(\bar{V}_{\text{vap}}^{\text{sat}} - \bar{V}_{\text{liq}}^{\text{sat}}) \tag{17}$$

where $\bar{U}$ and $\bar{V}$ denote molar energy and molar volume, respectively.

2.3.5 Number of effective snapshots

The performance of HR and MBAR depends primarily on good phase space overlap. For HR, good overlap means that the different sets of T and μ sample configurations and densities that are representative of the vapor and liquid phases at T^{sat} and μ^{sat}. For MBAR, an additional requirement is that the configurations sampled with θ_{ref} also represent feasible configurations for θ_{rr}.[16,39] The amount of overlap can be quantified by the number of effective snapshots $(K_{\text{snaps}}^{\text{eff}})$,[40] using Kish's formula:

$$K_{\text{snaps}}^{\text{eff}} = \frac{\left(\sum_n W_n\right)^2}{\sum_n W_n^2} \tag{18}$$

which reduces to $K_{\text{snaps}}^{\text{eff}} = (\sum_n W_n^2)^{-1}$ when the weights are normalized. This has the property that when the weights are equal, $K_{\text{snaps}}^{\text{eff}} = K_{\text{snaps}}^{\text{tot}}$, when all but one weight is negligible, $K_{\text{snaps}}^{\text{eff}} \approx 1$, and behaves appropriately for intermediate cases. Messerly et al. propose a heuristic that MBAR-ITIC is reliable when $K_{\text{snaps}}^{\text{eff}} > 50$. We test whether this is a reasonable heuristic for GCMC-MBAR as well.

2.4 Basis functions

When applying GCMC-MBAR to different force field parameter sets $(\theta_{\text{rr}} \neq \theta_{\text{ref}})$ it is necessary to recompute the internal energy for each snapshot (Equation 13). GCMC-HR typically requires millions of snapshots for precise estimates of $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and ΔH_{v} over a wide range of T^{sat}. The naive approach when $\theta_{\text{rr}} \neq \theta_{\text{ref}}$ is to store the molecular configurations $(\boldsymbol{x}_n)$ at each snapshot and then recompute $U(\boldsymbol{x}_n, N_n; \theta_{\text{rr}})$. Although this "rerun" process is orders of magnitude faster than performing direct GCMC simulations with θ_{rr}, the naive approach is memory intensive and computationally expensive. Fortu-

17

nately, basis functions can greatly accelerate the energy recomputation step.[16,39] Section 3.4 utilizes basis functions to rapidly recompute the non-bonded energies for $\theta_{\mathrm{rr}} \neq \theta_{\mathrm{ref}}$.

Basis functions are applicable whenever the energy can be separated linearly with respect to the force field parameters. For example, the Mie λ-6 non-bonded energy is separated into a repulsive and attractive term that can be expressed as

$$u^{\mathrm{nb}}(C_6, C_\lambda; r) = C_\lambda r^{-\lambda} - C_6 r^{-6} \tag{19}$$

where C_6 and C_λ are proportional to $\epsilon \sigma^6$ and $\epsilon \sigma^\lambda$, respectively. Therefore, the total non-bonded energy between all α and β sites ($U_{\alpha\beta}^{\mathrm{nb}}$) is simply

$$U_{\alpha\beta}^{\mathrm{nb}}(C_6, C_\lambda) = C_\lambda \sum_{i \neq j} r_{ij}^{-\lambda} - C_6 \sum_{i \neq j} r_{ij}^{-6} = C_\lambda \Psi_{\mathrm{r}} + C_6 \Psi_{\mathrm{a}} \tag{20}$$

where $\Psi_{\mathrm{r}}(\equiv \sum_{i \neq j} r_{ij}^{-\lambda})$ is the repulsive basis function, $\Psi_{\mathrm{a}}(\equiv \sum_{i \neq j} r_{ij}^{-6})$ is the attractive basis function and, for simplicity, $\sum_{i \neq j}$ denotes a sum over all unique pairwise interactions. Note that, because C_λ depends on λ, a separate repulsive basis function is required for each value of λ. For this reason, we adopt the common practice of limiting λ to integer values.

With Equation 20, the total non-bonded internal energy for all interaction sites of all $K_{\mathrm{snaps}}^{\mathrm{tot}}$ snapshots can be recomputed for any ϵ and σ with linear algebra instead of computing $u^{\mathrm{nb}}(\theta_{\mathrm{rr}}; r_{ij})$ for each unique pairwise interaction. Storing Ψ_{r} and Ψ_{a} for $K_{\mathrm{snaps}}^{\mathrm{tot}}$ snapshots also greatly reduces the memory storage load compared to storing $K_{\mathrm{snaps}}^{\mathrm{tot}}$ configurational snapshots (which must be full-precision for reliable "rerun" results).

Although it is preferable to output Ψ_{r} and Ψ_{a} at runtime, including this capability would require significant modification of GOMC. Instead, we utilize the GOMC "rerun" feature to indirectly compute Ψ_{r} and Ψ_{a} post-simulation. Specifically, we recompute $U_{\alpha\beta}^{\mathrm{nb}}$

for $\sigma_{\mathrm{rr}} \neq \sigma_{\mathrm{ref}}$, such that $C_{6,\mathrm{ref}} \neq C_{6,\mathrm{ref}}$ and $C_{\lambda,\mathrm{rr}} \neq C_{\lambda,\mathrm{ref}}$. We then perform an additional "rerun" calculation for each value of $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$. Finally, we solve the following system of equations for Ψ_{a} and $\Psi_{\mathrm{r},\lambda}$ (where λ corresponds to each value of λ_{rr}):

2.5 ϵ-scaling

Recently, Weidler and Gross proposed an ϵ-scaling approach for converting the Transferable Anisotropic Mie (TAMie) parameters into individualized (compound-specific) parameters (iTAMie).[41] The philosophy for individualized parameters is that some compounds have sufficient reliable experimental data to refine the force field parameters for a specific molecule. However, refitting all non-bonded parameters simultaneously would likely lead to an underspecified optimization. To avoid overfitting, Weidler and Gross optimize a single adjustable parameter (ψ) that scales all the ϵ values according to

$$\epsilon_{ii}^{\mathrm{ind}} = \psi \times \epsilon_{ii}^{\mathrm{tran}} \tag{21}$$

where $\epsilon_{ii}^{\mathrm{ind}}$ is the individualized ϵ value for united-atom ii, $\epsilon_{ii}^{\mathrm{tran}}$ is the corresponding transferable ϵ value, and ψ is a fixed value for a given compound.

GCMC-MBAR is ideally suited for this "ϵ-scaling" approach for at least two reasons. First, MBAR is most reliable when extrapolating in ϵ rather than σ and/or λ.[16] Second, because $\epsilon_{\mathrm{rr}} = \psi \times \epsilon_{\mathrm{ref}}$ while $\sigma_{\mathrm{rr}} = \sigma_{\mathrm{ref}}$ and $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}}$, recomputing the total non-bonded energy for each snapshot is simply

$$U_{\mathrm{rr}}^{\mathrm{nb,tot}} = \psi \times U_{\mathrm{ref}}^{\mathrm{nb,tot}} \tag{22}$$

where $U_{\mathrm{rr}}^{\mathrm{nb,tot}}$ and $U_{\mathrm{ref}}^{\mathrm{nb,tot}}$ are the total non-bonded energy with θ_{rr} and θ_{ref}, respectively. Therefore, ϵ-scaling does not require storing and recomputing configurations or basis

functions.

Section 3.2 applies the GCMC-MBAR ϵ-scaling approach to convert the MiPPE-SL semi-transferable parameters into individualized parameters (iMiPPE) for 8 branched alkanes and 11 alkynes. For consistency with the original MiPPE-SL optimization, we use the same scoring function for the branched alkanes and alkynes as Mick et al. and Soroush Barhaghi et al., respectively

$$
\begin{aligned}
S = \frac{1}{N_{\text{exp}}} \Big[& w_0 \sum_{j=0}^{N_{\text{exp}}} APD(\rho_{\text{liq}}^{\text{sat}}(T_j^{\text{sat}})) + w_1 \sum_{j=0}^{N_{\text{exp}}} APD(\rho_{\text{vap}}^{\text{sat}}(T_j^{\text{sat}})) \\
& + w_2 \sum_{j=0}^{N_{\text{exp}}} APD(P_{\text{vap}}^{\text{sat}}(T_j^{\text{sat}})) + w_3 \sum_{j=0}^{N_{\text{exp}}} APD(\Delta H_{\text{v}}(T_j^{\text{sat}})) \\
& + w_4 \sum_{j=0}^{N_{\text{exp}}} \frac{dAPD(\rho_{\text{liq}}^{\text{sat}}(T_j^{\text{sat}}))}{dT} + w_5 \sum_{j=0}^{N_{\text{exp}}} \frac{dAPD(\rho_{\text{vap}}^{\text{sat}}(T_j^{\text{sat}}))}{dT} \\
& + w_6 \sum_{j=0}^{N_{\text{exp}}} \frac{dAPD(P_{\text{vap}}^{\text{sat}}(T_j^{\text{sat}}))}{dT} + w_7 \sum_{j=0}^{N_{\text{exp}}} \frac{dAPD(\Delta H_{\text{v}}(T_j^{\text{sat}}))}{dT} \Big]
\end{aligned} \tag{23}
$$

where S is the scoring function, N_{exp} is the number of (pseudo-)experimental data points, w_x are the weights for property x, T_j^{sat} are the saturation temperatures for data point j, and the absolute percent deviation (APD) is defined as

$$
APD(X) = \left| \frac{X_{\text{sim}} - X_{\text{exp}}}{X_{\text{exp}}} \right| \times 100\% \tag{24}
$$

where X_{sim} and X_{exp} correspond to the respective simulation and (pseudo-) experimental values for property X (e.g., $\rho_{\text{liq}}^{\text{sat}}$).

In accordance with the work of Mick et al. and Soroush Barhaghi et al., the target values (X_{exp}) are computed with pseudo-experimental correlations. The alkyne correlations are from the Design Institute for Physical Properties (DIPPR)[42] while the branched alkane correlations are from the National Institute of Standards and Technology (NIST)

Reference Fluid Properties (REFPROP) database.[43]

The weights (w_x) are equal to (0.6135, 0.0123, 0.2455, 0.0245, 0.0613, 0.0061, 0.0245, 0.0123) and (0.757, 0, 0.152, 0, 0.076, 0, 0.015, 0) for branched alkanes and alkynes, respectively. This weighting is the same as Mick et al. and Soroush Barhaghi et al. The $\rho_{\text{liq}}^{\text{sat}}$ weights are greatest due to both the high precision of experimental data and the high importance that the MiPPE force field reproduce this property. The $P_{\text{vap}}^{\text{sat}}$ weights are second greatest, demonstrating the importance of this property but the slightly higher experimental uncertainties. The $\rho_{\text{vap}}^{\text{sat}}$ and ΔH_{v} weights are 0 for alkynes as the DIPPR database does not report $\rho_{\text{vap}}^{\text{sat}}$ and the DIPPR ΔH_{v} uncertainties are relatively large for most alkynes.

3 Results

Four different applications for GCMC-MBAR are demonstrated in this study, where slightly different types of simulation output are required. Section 3.1 demonstrates that GCMC-MBAR yields consistent results to those previously reported using histogram reweighting when the force field parameters do not change, i.e., $\theta_{\text{rr}} = \theta_{\text{ref}}$. The same simulation output as GCMC-HR is used in this application, namely, a $2 \times K_{\text{snaps}}^{\text{tot}}$ array containing the number of molecules (N) and the internal energy (U) for all $K_{\text{snaps}}^{\text{tot}}$ snapshots. Section 3.2 demonstrates how these same data can be used with MBAR to perform a one-dimensional optimization with ϵ-scaling. Section 3.3 investigates how well GCMC-MBAR performs when predicting vapor-liquid coexistence properties for force field j from configurations sampled with force field i, where i and j are common force fields from the literature. In this case, a $3 \times K_{\text{snaps}}^{\text{tot}}$ array is required, where the additional column is the internal energy computed with force field j. Section 3.4 presents a case study for optimizing Mie λ-6 non-bonded parameters for cyclohexane. Basis functions are employed as a

computationally efficient method for predicting vapor-liquid coexistence properties with hundreds of ϵ_{CH_2}, σ_{CH_2}, and λ_{CH_2} values that are unknown at runtime.

3.1 Validation of GCMC-MBAR for constant θ

Section 2.3 demonstrated that MBAR and HR are mathematically equivalent in the limit of zero bin width when $\theta_{rr} = \theta_{ref}$. Figure 1 provides numerical validation that GCMC-MBAR and GCMC-HR yield indistinguishable vapor-liquid coexistence properties. The evidence for this conclusion is that the median percent deviation is approximately zero and the largest deviations are within a few percent. Note that the uncertainties are largest near the critical point, i.e., reduced temperatures $(T_r) \approx 1$. The percent deviations shown in Figure 1 are averaged over the 31 branched alkanes studied by Mick et al. and the 11 alkynes studied by Soroush Barhaghi et al.

The GCMC-HR values used in Figure 1 were not recomputed in this study but were taken from the literature.[12,13] For a fair comparison between GCMC-HR and GCMC-MBAR, the GCMC-MBAR values were computed using the same raw simulation data as Mick et al. and Soroush Barhaghi et al. However, for simplicity, we only reprocess one of the five replicate simulations. Tabulated results are provided as Supporting Information.

3.2 Application of GCMC-MBAR to ϵ-scaling

As discussed in Section 2.5, ϵ-scaling is used to obtain individualized (compound-specific) ϵ values (scaled by ψ) for well-studied compounds, i.e., those with large amounts of reliable experimental data.[41] In this section we demonstrate that GCMC-MBAR is an efficient tool for performing ϵ-scaling. Simulations are performed with the MiPPE-SL force field and coexistence properties are computed over a range of ψ values, where $\epsilon_{rr} = \psi \times \epsilon_{ref}$.
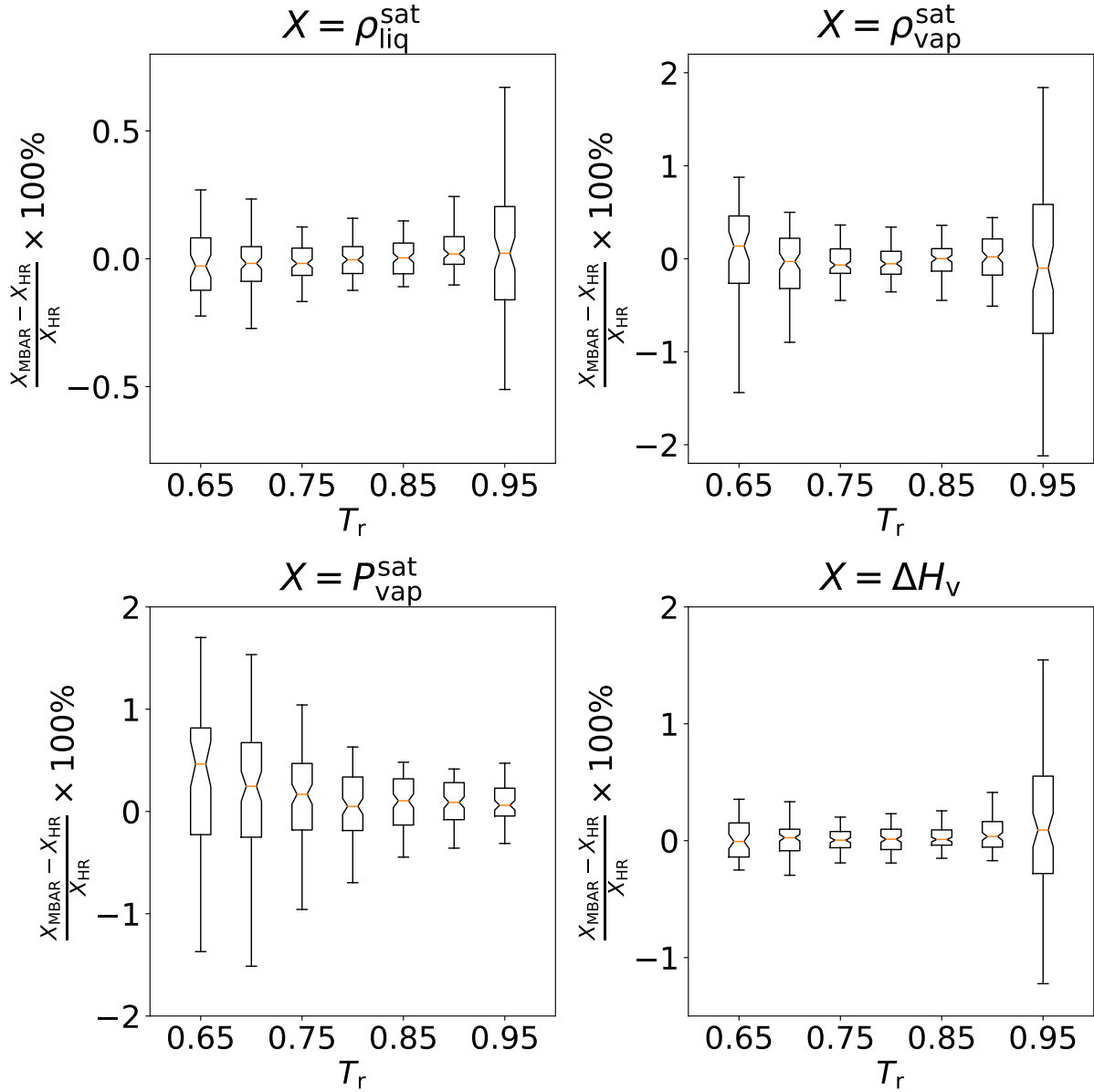
Figure 1: Percent deviations between coexistence properties computed using histogram reweighting (HR) and Multistate Bennett Acceptance Ratio (MBAR). The HR and MBAR results are in good agreement, i.e., within a few percent and an average percent deviation of approximately 0%. Top-left, top-right, bottom-left, and bottom-right panels correspond to saturated liquid density, saturated vapor density, saturated vapor pressure, and enthalpy of vaporization, respectively. Middle line denotes the median deviation, boxes depict the first and third quartiles, and whiskers represent the range that contains 95% of the data.

Figure 2 presents the ϵ-scaling results for 8 branched alkanes and 11 alkynes using the MiPPE-SL force field as θ_{ref} ($\psi = 1$). Figure 2 shows that the alkynes require a greater degree of scaling than the branched alkanes. Weigler et al. tend to characterize the individualization as being useful when the scaling is greater than 0.4% (i.e., $|1 - \psi| > 0.004$). With this rationale, 3-methylpentane is the only branched alkane that merits ϵ-scaling with the MiPPE-SL force field. Similarly, the TAMie force field also found $\psi \approx 1$ for all branched alkanes, except 3-methylpentane. Although ψ values for iTAMie were not reported for alkynes, the largest ψ value for olefins, ethers, and ketones was ≈ 1.01. A truly transferable force field should have $\psi \approx 1$ for all compounds. Therefore, the transferability of the MiPPE force field appears to be slightly poorer for 2-pentyne and 2-hexyne, which have an optimized $\psi > 1.01$.

It is also interesting that only 3 out of 19 compounds require $\psi < 1$. Thus, the well-depths appear to be slightly underestimated by the MiPPE force field. By contrast, this trend was not observed in Reference 41 for TAMie. Also, note that branched alkanes have a pronounced minimum in S with respect to ψ, whereas the minimum is more gradual for the alkynes. We attribute this to the fact that the alkynes do not include ΔH_{v} (a property that depends strongly on ϵ) in the scoring function, i.e., $w_4 = 0$ and $w_7 = 0$.

3.3 Performance of GCMC-MBAR when $\theta_{\text{rr}} \not\approx \theta_{\text{ref}}$

A more demanding test of GCMC-MBAR than ϵ-scaling is to vary several non-bonded parameters simultaneously, including σ and λ. Because it is not possible to visualize a parameter space of greater than three dimensions, we perform this analysis of GCMC-MBAR using the TraPPE, NERD, MiPPE-gen, and MiPPE-SL force fields. Specifically, we utilize GCMC-MBAR to predict coexistence properties for the NERD and MiPPE-SL force fields using configurations sampled from TraPPE and MiPPE-gen, respectively (see Figure 3). We also use GCMC-MBAR to predict coexistence properties for the TraPPE
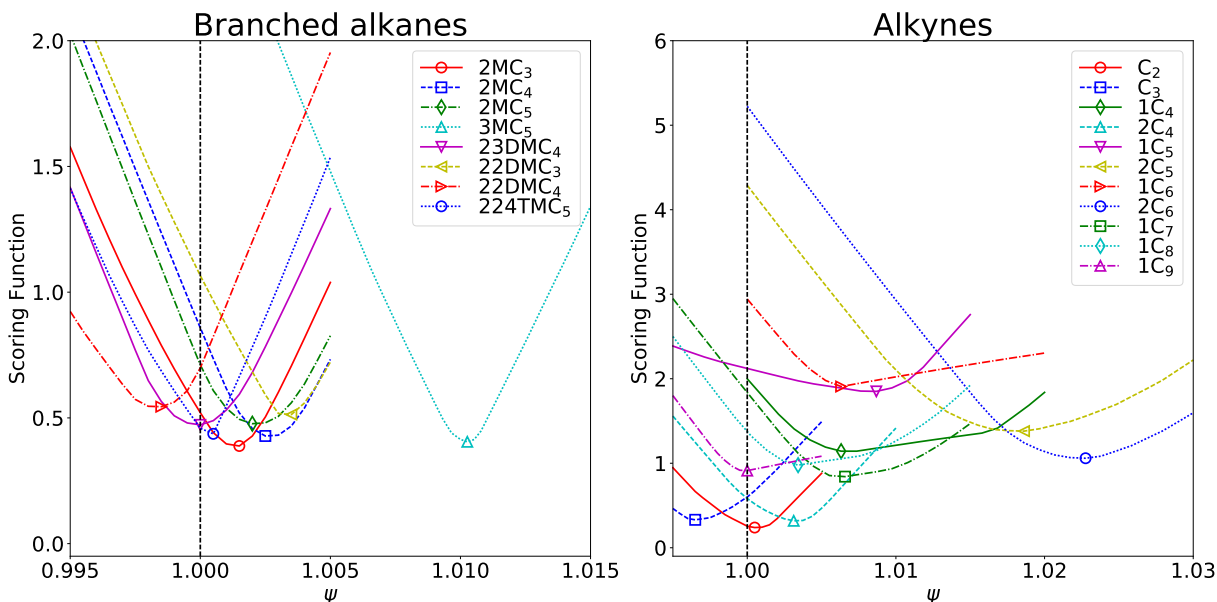
Figure 2: One dimensional optimization with ϵ-scaling (ψ) of MiPPE-SL for select branched alkanes (left) and alkynes (right). MBAR enables prediction of scoring function over range of ψ from configurations that were sampled with $\psi = 1$ (dashed line). Open symbols correspond to the optimal ψ value for a given compound.

force field by sampling configurations with MiPPE-gen, and vice versa (see Figure 4).

Note that all three non-bonded parameters (ϵ, σ, and λ) for all four united-atom types (CH$_3$, CH$_2$, CH, and C) are different between the TraPPE and MiPPE-gen force fields. The TraPPE and NERD ϵ and σ values are different for all four united-atom types while $\lambda = 12$ for both force fields. The MiPPE-gen and MiPPE-SL force fields only differ in the ϵ and/or σ values for the CH and C sites. However, the difference in ϵ and σ values for MiPPE-gen and MiPPE-SL is significantly smaller than that between TraPPE and NERD. Therefore, the MiPPE-gen $\Rightarrow$ MiPPE-SL results correspond to $\theta_{\text{rr}} \approx \theta_{\text{ref}}$, which is important when fine-tuning a pre-optimized force field.

Figures 3 and 4 compare the GCMC-MBAR predicted values for $\theta_{\text{rr}} \neq \theta_{\text{ref}}$ to the literature GCMC-HR values obtained by direct simulation. Figure 3 contains $\lambda_{\text{rr}} = \lambda_{\text{ref}}$ while Figure 4 corresponds to $\lambda_{\text{rr}} \neq \lambda_{\text{ref}}$. As observed in Figures 3 and 4, MBAR is extremely reliable at predicting vapor phase properties ($\rho_{\text{vap}}^{\text{sat}}$ and $P_{\text{vap}}^{\text{sat}}$). GCMC-MBAR is

remarkably accurate at predicting liquid phase properties ($\rho_{\mathrm{liq}}^{\mathrm{sat}}$ and ΔH_{v}, which depends on both phases) when $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}}$, even for the fairly significant differences in the TraPPE and NERD σ values. However, Figure 4 shows that GCMC-MBAR is less reliable for liquid phase properties when $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$. This undesirable behavior can be explained by the low number of effective snapshots in the liquid phase.

Figure 5 demonstrates that, as expected, $K_{\mathrm{snaps}}^{\mathrm{eff}}$ is typically much greater in the vapor phase than in the liquid phase. Messerly et al. report that MBAR-ITIC is reliable if $K_{\mathrm{snaps}}^{\mathrm{eff}} > 50$. Applying this heuristic to GCMC-MBAR helps qualify why the liquid properties are poorly estimated in some systems while the vapor properties are much more accurate. Specifically, $K_{\mathrm{snaps}}^{\mathrm{eff}}$ is less than 50 in the liquid phase when $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$ (TraPPE $\Leftrightarrow$ MiPPE-gen) while it is typically greater than 50 for $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}}$ (MiPPE-gen $\Rightarrow$ MiPPE-SL and TraPPE $\Rightarrow$ NERD). This demonstrates that $K_{\mathrm{snaps}}^{\mathrm{eff}} > 50$ is a good indication that the GCMC-MBAR estimates are reliable.

The poor overlap in the liquid phase when varying λ is consistent with the MBAR-ITIC results.[16] By contrast, GCMC-MBAR provides considerable improvement in predicting $\rho_{\mathrm{vap}}^{\mathrm{sat}}$ and $P_{\mathrm{vap}}^{\mathrm{sat}}$ for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$ compared to what was previously observed for MBAR-ITIC. Therefore, similar to MBAR-ITIC, optimizing Mie λ-6 parameters requires performing direct simulations for each value of λ. However, as opposed to MBAR-ITIC, GCMC-MBAR can help determine where these additional simulations should be performed, i.e., GCMC-MBAR can localize the likely optimal ϵ and σ values for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$.

We recommend the following algorithm for optimizing Mie λ-6 parameters with GCMC-MBAR:

1. Simulate reference force field (θ_{ref}), e.g., TraPPE

2. Compute basis functions for various values of λ

3. Optimize ϵ and σ for each λ ($\epsilon_{\mathrm{opt}}^{\lambda}$, $\sigma_{\mathrm{opt}}^{\lambda}$) by:

(a) Estimating $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and ΔH_{v} with GCMC-MBAR

(b) Minimizing scoring function (S)

4. Perform additional simulations with each $\epsilon_{\text{opt}}^{\lambda}$, $\sigma_{\text{opt}}^{\lambda}$, and λ from Step 3

5. Determine overall optimal parameter set (θ_{opt}) by:

(a) Estimating $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and ΔH_{v} with GCMC-MBAR for $\lambda_{\text{rr}} = \lambda_{\text{ref}}$

(b) Minimizing scoring function (S)

6. Repeat Steps 4 and 5 until $\min(K_{\text{snaps}}^{\text{eff}}) \gg 50$ for θ_{opt}

The stopping criterion is based on the minimum number of effective snapshots, which always corresponds to the liquid phase at lower values of T^{sat}. If the optimal parameter set has $\min(K_{\text{snaps}}^{\text{eff}}) \gg 50$ the GCMC-MBAR estimates should be reliable over the entire range of saturation temperatures. Any additional iterations would be ill-advised and unnecessary as the optimization would primarily be fitting to the uncertainty in the simulation output. Section 3.4 presents an application of this approach.

3.4 Case Study: Optimizing cyclohexane Mie λ-6 parameters

We have demonstrated that GCMC-MBAR is accurate when $K_{\text{snaps}}^{\text{eff}} \gg 50$, which is typically the case in the vapor phase and in the liquid phase when $\lambda_{\text{rr}} = \lambda_{\text{ref}}$. In this section, we present how GCMC-MBAR can rapidly optimize the Mie λ-6 parameters when starting with a Lennard-Jones 12-6 potential (namely, TraPPE). We have chosen cyclohexane for this case study as MiPPE does not yet have non-bonded parameters for this compound, while the TraPPE force field does. Also, because cyclohexane consists of a single united-atom site type (CH_2), it is a convenient molecule for representing the scoring function in 2-dimensions (ϵ_{CH_2} and σ_{CH_2} for a given value of λ_{CH_2}).

Figure 6 depicts the scoring function for different values of $\epsilon_{\mathrm{CH_2}}$, $\sigma_{\mathrm{CH_2}}$, and $\lambda_{\mathrm{CH_2}}$ as a heat map. The optimal parameter set for each value of $\lambda_{\mathrm{CH_2}}$ (depicted as a white star in Figures 6 and 5) corresponds to the lowest values of S. Similar figures have been reported in the literature using GCMC-HR.[12,13] The key difference is that the heat maps reported in the literature were obtained by performing GCMC simulations with each value of ϵ, σ, and λ. By contrast, the results shown in Figure 6 were obtained by performing GCMC simulations with a single parameter set, namely, the TraPPE parameters (depicted as an "X" in Figures 6 and 5). MBAR reweights these same configurations for all other parameter sets. Furthermore, $U(\theta_{\mathrm{rr}})$ is computed with basis functions, enabling the GCMC-MBAR recompute step to be extremely fast.

In this optimization example (an exhaustive 2-dimensional grid search over integer values of λ), each proposed parameter set could be simulated directly in parallel. Therefore, the real time to solution for GCMC-MBAR would be similar to that of the GCMC-HR approach utilized by Mick et al. and Soroush Barhaghi et al. (although GCMC-MBAR would still reduce the total CPU time). In general, however, higher dimensional optimization algorithms are performed in sequence, where each iteration proposes new parameter set(s). In this scenario, GCMC-MBAR (with basis functions) is orders of magnitude faster than the literature GCMC-HR approach, which would require performing new GCMC simulations for each iteration.

Note that the TraPPE force field utilizes a Lennard-Jones 12-6 potential (i.e., $\lambda_{\mathrm{TraPPE}} = 12$) and, therefore, the results in the top-left panel are for the case where $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}} = 12$, while the other panels correspond to $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$. Figure 7 shows that, as expected, $K_{\mathrm{snaps}}^{\mathrm{eff}}$ in the liquid phase are much lower for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$ than for $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}}$. In addition, the smooth contours in Figure 6 for $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}} = 12$ and the wide range of parameters over which $\bar{K}_{\mathrm{snaps}}^{\mathrm{eff,liq}} \gg 50$ suggests that GCMC-MBAR is highly reliable for optimizing $\epsilon_{\mathrm{CH_2}}$ and $\sigma_{\mathrm{CH_2}}$ for a fixed value of $\lambda_{\mathrm{CH_2}}$. Even more remarkable, considering the erroneous prediction of

$\rho_{\mathrm{liq}}^{\mathrm{sat}}$ and ΔH_{v} for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$ (see Figure 4), is that GCMC-MBAR predicts smooth contours for $\lambda \neq 12$ despite lower values of $\bar{K}_{\mathrm{snaps}}^{\mathrm{eff,liq}}$. This characteristic would be important when implementing gradient descent optimization schemes.

Having minimized the scoring function for each value of $\lambda_{\mathrm{CH_2}}$, we perform additional simulations with the optimal parameter sets ($\epsilon_{\mathrm{opt}}^{\lambda}$, $\sigma_{\mathrm{opt}}^{\lambda}$) serving as new reference parameter sets. Because the optimal $\epsilon_{\mathrm{CH_2}}$ and $\sigma_{\mathrm{CH_2}}$ parameters for $\lambda_{\mathrm{CH_2}} = 12$ already have $\min(K_{\mathrm{snaps}}^{\mathrm{eff}}) \gg 50$, we only consider $\lambda_{\mathrm{CH_2}} \geq 14$. Figure 8 presents the results from this second iteration of the two-dimensional optimization for each $\lambda_{\mathrm{CH_2}}$. Note that the first iteration optimal parameter sets (black "X"s) are quite similar to the second iteration (white stars). Figure 9 verifies that there is sufficient overlap $(\min(K_{\mathrm{snaps}}^{\mathrm{eff}}) \gg 50)$ between the first and second iterations to consider the optimization converged. Table 5 summarizes the parameter sets obtained at each stage of the optimization.

Table 5: Optimal Mie λ-6 cyclohexane parameters for $\lambda_{\mathrm{CH_2}} = 12, 14, 16, 18$, and 20. Superscript denotes the iteration stage of the optimization.

$\lambda_{\mathrm{CH_2}}$	$\epsilon^{\langle 0 \rangle}_{\mathrm{CH_2}}/k_B$ (K)	$\sigma^{\langle 0 \rangle}_{\mathrm{CH_2}}$ (nm)	$\epsilon^{\langle 1 \rangle}_{\mathrm{CH_2}}/k_B$ (K)	$\sigma^{\langle 1 \rangle}_{\mathrm{CH_2}}$ (nm)	$\epsilon^{\langle 2 \rangle}_{\mathrm{CH_2}}/k_B$ (K)	$\sigma^{\langle 2 \rangle}_{\mathrm{CH_2}}$ (nm)
12	52.5	0.391	53.0	0.393	–	–
14	–	–	61.5	0.393		
16	–	–	70.0	0.389	69.5	0.391
18	–	–	77.0	0.389		
20	–	–	84.0	0.388		

The overall optimal parameter set is reported in Table 4 as the MiPPE cyclohexane parameters. Note that $\lambda_{\mathrm{CH_2}} = 16$ is once again found to be the optimal repulsive exponent value. It is important to verify that the overall optimal parameter set does indeed provide accurate predictions of vapor-liquid coexistence properties. For this reason, we perform direct GCMC simulations with the overall optimal parameter set to compare the new MiPPE force field with the initial reference (TraPPE) and the first iteration parameter sets.

Figure 10 is provided to quantify the improved accuracy achieved for the two itera-

tions by comparing the percent deviations between predicted and pseudo-experimental (REFPROP) values of $\rho_{\text{liq}}^{\text{sat}}$, $\rho_{\text{vap}}^{\text{sat}}$, $P_{\text{vap}}^{\text{sat}}$, and ΔH_{v} for the zeroth iteration (TraPPE), first iteration for $\lambda_{\text{CH}_2} = 14, 16, 18$, and 20, and final iteration parameter sets. Even the first iteration parameters demonstrate considerable improvement compared to the zeroth iteration (TraPPE) for predicting $\rho_{\text{vap}}^{\text{sat}}$ and $P_{\text{vap}}^{\text{sat}}$ without significantly deminishing the accuracy for $\rho_{\text{liq}}^{\text{sat}}$ and ΔH_{v}. The second (final) iteration is clearly the optimal parameter set as it is the most accurate for $\rho_{\text{vap}}^{\text{sat}}$ and $P_{\text{vap}}^{\text{sat}}$ with similar performance for $\rho_{\text{liq}}^{\text{sat}}$ and ΔH_{v} as the first iteration.

This section has shown that GCMC-MBAR is capable of converting a pre-tuned Lennard-Jones 12-6 potential (TraPPE) into a Mie λ-6 potential (MiPPE) without performing hundreds of simulations. Specifically, with only two stages of direct simulation we considered hundreds of ϵ_{CH_2}, σ_{CH_2}, and λ_{CH_2} parameter sets.

4   Discussion/Limitations/Future work

The results presented in this study were obtained by performing simulations with only a single reference force field. As shown in previous studies, a logical approach for improving the performance of MBAR is to include additional reference force fields.[16,17] For example, in Section 3.4, we did not utilize the snapshots from $\lambda_{\text{ref}} = 14$ when computing properties for $\lambda_{\text{rr}} = 16$ and vice versa. Although this would increase the number of effective snapshots, a single reference provides sufficient overlap when λ_{CH_2} is constant.

As molecular insertion moves are frequently rejected in high density systems, GCMC simulations are typically not reliable at low saturation temperatures $(T < 0.65 T_{\text{c}})$. Because ITIC does not suffer from this low-temperature limitation, we recommend combining the MBAR-ITIC and GCMC-MBAR methods when predicting vapor-liquid coexistence properties from near-triple-point to near-critical-point conditions.

Although GCMC-HR is a standard approach for computing vapor-liquid coexistence, HR has also been applied to GEMC simulations (GEMC-HR).[44] Therefore, while the present study presents how MBAR can be applied to GCMC simulations, an analogous GEMC-MBAR approach is worth investigating in future work.

5 Conclusions

This study demonstrates how the Multistate Bennett Acceptance Ratio can replace the traditional histogram reweighting approach for estimating vapor-liquid coexistence properties from Grand Canonical Monte Carlo simulations. MBAR and HR are mathematically equivalent in the limit of infinitesimal bin widths when the coexistence properties are computed for a single force field. However, the primary benefit of MBAR is the ability to estimate properties for force fields that are not simulated directly. This greatly accelerates non-bonded parameterization without requiring large amounts of simulation. We demonstrate this capability by performing a one-dimensional ϵ-scaling optimization of several branched alkanes and alkynes. We then show how GCMC-MBAR can rapidly re-optimize a Lennard-Jones 12-6 potential (TraPPE) into a family of Mie λ-6 potentials for cyclohexane. The optimized Mie 16-6 parameters form the most recent contribution to the Mie Potentials for Phase Equilibria (MiPPE) force field.

6 Acknowledgments

References

(1) Panagiotopoulos, A. Z. Monte Carlo methods for phase equilibria of fluids. *Journal of Physics: Condensed Matter* **2000**, *12*, R25.

(2) Fern, J. T.; Keffer, D. J.; Steele, W. V. Measuring Coexisting Densities from a Two-Phase Molecular Dynamics Simulation by Voronoi Tessellations. *The Journal of Physical Chemistry B* **2007**, *111*, 3469–3475.

(3) Razavi, S. M.; Messerly, R. A.; Elliott, J. R. Coexistence Calculation Using the Isothermal-Isochoric Integration Method. **2018**, *Pending publication.*

(4) Potoff, J. J.; Errington, J. R.; Panagiotopoulos, A. Z. Molecular simulation of phase equilibria for mixtures of polar and non-polar components. *Molecular Physics* **1999**, *97*, 1073–1083.

(5) Stubbs, J. M.; Potoff, J. J.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 6. United-Atom Description for Ethers, Glycols, Ketones, and Aldehydes. *The Journal of Physical Chemistry B* **2004**, *108*, 17596–17605.

(6) Martin, M. G.; Siepmann, J. I. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.

(7) Hemmen, A.; Gross, J. Transferable Anisotropic United-Atom Force Field Based on the Mie Potential for Phase Equilibrium Calculations: n-Alkanes and n-Olefins. *J. Phys. Chem. B* **2015**, *119*, 11695–11707.

(8) Potoff, J. J.; Bernard-Brunel, D. A. Mie Potentials for Phase Equilibria Calculations: Applications to Alkanes and Perfluoroalkanes. *J. Phys. Chem. B* **2009**, *113*, 14725–14731.

(9) Ungerer, P.; Beauvais, C.; Delhommelle, J.; Boutin, A.; Rousseau, B.; Fuchs, A. H. Optimization of the anisotropic united atoms intermolecular potential for n-alkanes. *J. Chem. Phys.* **2000**, *112*, 5499–5510.

(10) Messerly, R. A.; KnottsIV, T. A.; Wilding, W. V. Uncertainty quantification and propagation of errors of the Lennard-Jones 12-6 parameters for n-alkanes. *J. Chem. Phys.* **2017**, *146*, 194110.

(11) Mick, J. R.; Soroush Barhaghi, M.; Jackman, B.; Rushaidat, K.; Schwiebert, L.; Potoff, J. J. Optimized Mie potentials for phase equilibria: Application to noble gases and their mixtures with n-alkanes. *J. Chem. Phys.* **2015**, *143*, 114504.

(12) Mick, J. R.; Soroush Barhaghi, M.; Jackman, B.; Schwiebert, L.; Potoff, J. J. Optimized Mie Potentials for Phase Equilibria: Application to Branched Alkanes. *J. Chem. Eng. Data* **2017**, *62*, 1806–1818.

(13) Barhaghi, M. S.; Mick, J. R.; Potoff, J. J. Optimised Mie potentials for phase equilibria: application to alkynes. *Molecular Physics* **2017**, *115*, 1378–1388.

(14) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J. Chem. Theory Comput.* **2007**, *3*, 26–41.

(15) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.

(16) Messerly, R. A.; Razavi, S. M.; Shirts, M. R. Configuration-Sampling-Based Surrogate Models for Rapid Parameterization of Non-Bonded Interactions. *Journal of Chemical Theory and Computation* **2018**, *14*, 3144–3162.

(17) Messerly, R. A.; Shirts, M. R.; Kazakov, A. F. Uncertainty quantification confirms unreliable extrapolation toward high pressures for united-atom Mie λ-6 force field. *The Journal of Chemical Physics* **2018**, *149*, 114109.

(18) Razavi, S. M. Optimization of a Transferable Shifted Force Field for Interfaces and Inhomogenous Fluids using Thermodynamic Integration. M.Sc. thesis, The University of Akron, 2016.

(19) Kiyohara, K.; Spyriouni, T.; Gubbins, K. E.; Panagiotopoulos, A. Z. Thermodynamic scaling Gibbs ensemble Monte Carlo: a new method for determination of phase coexistence properties of fluids. *Molecular Physics* **1996**, *89*, 965–974.

(20) Errington, J. R.; Panagiotopoulos, A. Z. Phase equilibria of the modified Buckingham exponential-6 potential from Hamiltonian scaling grand canonical Monte Carlo. *The Journal of Chemical Physics* **1998**, *109*, 1093–1100.

(21) Errington, J. R.; Panagiotopoulos, A. Z. A new intermolecular potential model for the n-alkane homologous series. *J. Phys. Chem. B* **1999**, *103*, 6314–6322.

(22) Errington, J. R.; Panagiotopoulos, A. Z. New intermolecular potential models for benzene and cyclohexane. *The Journal of Chemical Physics* **1999**, *111*, 9731–9738.

(23) Martin, M. G.; Siepmann, J. I. Novel Configurational-Bias Monte Carlo Method for Branched Molecules. Transferable Potentials for Phase Equilibria. 2. United-Atom Description of Branched Alkanes. *The Journal of Physical Chemistry B* **1999**, *103*, 4508–4517.

(24) Keasler, S. J.; Charan, S. M.; Wick, C. D.; Economou, I. G.; Siepmann, J. I. Transferable Potentials for Phase EquilibriaâĂŞUnited Atom Description of Five- and Six-Membered Cyclic Alkanes and Ethers. *The Journal of Physical Chemistry B* **2012**, *116*, 11234–11246.

(25) Nath, S. K.; Escobedo, F. A.; de Pablo, J. J. On the simulation of vapor-liquid equilibria for alkanes. *J. Chem. Phys.* **1998**, *108*, 9905–9911.

(26) Nath, S. K.; Khare, R. New forcefield parameters for branched hydrocarbons. *The Journal of Chemical Physics* **2001**, *115*, 10837–10844.

(27) Yiannourakou, M.; Ungerer, P.; Lachet, V.; Rousseau, B.; Teuler, J.-M. United atom forcefield for vapor-liquid equilibrium (VLE) properties of cyclic and polycyclic compounds from Monte Carlo simulations. *Fluid Phase Equilibria* **2019**, *481*, 28 – 43.

(28) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Clarendon Press ; Oxford University Press: Oxford England New York, 1987; pp xix, 385 p.

(29) Shah, J. K.; Maginn, E. J. A general and efficient Monte Carlo method for sampling intramolecular degrees of freedom of branched and cyclic molecules. *The Journal of Chemical Physics* **2011**, *135*, 134121.

(30) Nejahi, Y.; Soroush Barhaghi, M.; Mick, J.; Jackman, B.; Rushaidat, K.; Li, U.; Schwiebert, L.; Potoff, J. GOMC: GPU Optimized Monte Carlo for the simulation of phase equilibria and physical properties of complex fluids. *SoftwareX*

(31) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry 30*, 2157–2164.

(32) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33 – 38.

(33) McDonald, I. R.; Singer, K. Machine Calculation of Thermodynamic Properties of a Simple Fluid at Supercritical Temperatures. *The Journal of Chemical Physics* **1967**, *47*, 4766–4772.

(34) Card, D. N.; Valleau, J. P. Monte Carlo Study of the Thermodynamics of Electrolyte Solutions. *The Journal of Chemical Physics* **1970**, *52*, 6232–6240.

(35) Wood, W. W. Monte Carlo Calculations for Hard Disks in the IsothermalâĂŘIsobaric Ensemble. *The Journal of Chemical Physics* **1968**, *48*, 415–434.

(36) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **1992**, *13*, 1011–1021.

(37) Duarte Ramos Matos, G.; Kyu, D. Y.; Loeffler, H. H.; Chodera, J. D.; Shirts, M. R.; Mobley, D. L. Approaches for Calculating Solvation Free Energies and Enthalpies Demonstrated with an Update of the FreeSolv Database. *Journal of Chemical & Engineering Data* **2017**, *62*, 1559–1569.

(38) https://github.com/choderalab/pymbar.

(39) Naden, L. N.; Shirts, M. R. Rapid Computation of Thermodynamic Properties Over Multidimensional Nonbonded Parameter Spaces using Adaptive Multistate Reweighting. *J. Chem. Theory Comput.* **2016**, *12*, 1806–1823.

(40) Dybeck, E. C.; König, G.; Brooks, B. R.; Shirts, M. R. Comparison of Methods To Reweight from Classical Molecular Simulations to QM/MM Potentials. *J. Chem. Theory Comput.* **2016**, *12*, 1466–1480.

(41) Weidler, D.; Gross, J. Individualized force fields for alkanes, olefins, ethers and ketones based on the transferable anisotropic Mie potential. *Fluid Phase Equilibria* **2018**,

(42) Rowley, R. L.; Wilding, W. V.; Oscarson, J. L.; KnottsIV, T. A.; Giles, N. F. *DIPPR Data Compilation of Pure Chemical Properties*; Design Institute for Physical Properties; AIChE: New York, NY, 2013.

(43) Lemmon, E. W.; Bell, I. H.; Huber, M. L.; McLinden, M. O. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 10.0, National Institute of Standards and Technology. 2018; `https://www.nist.gov/srd/refprop`.

(44) Boulougouris, G. C.; Peristeras, L. D.; Economou, I. G.; Theodorou, D. N. Predicting fluid phase equilibrium via histogram reweighting with Gibbs ensemble Monte Carlo simulations. *The Journal of Supercritical Fluids* **2010**, *55*, 503 – 509, 100th year Anniversary of van der Waals' Nobel Lecture.
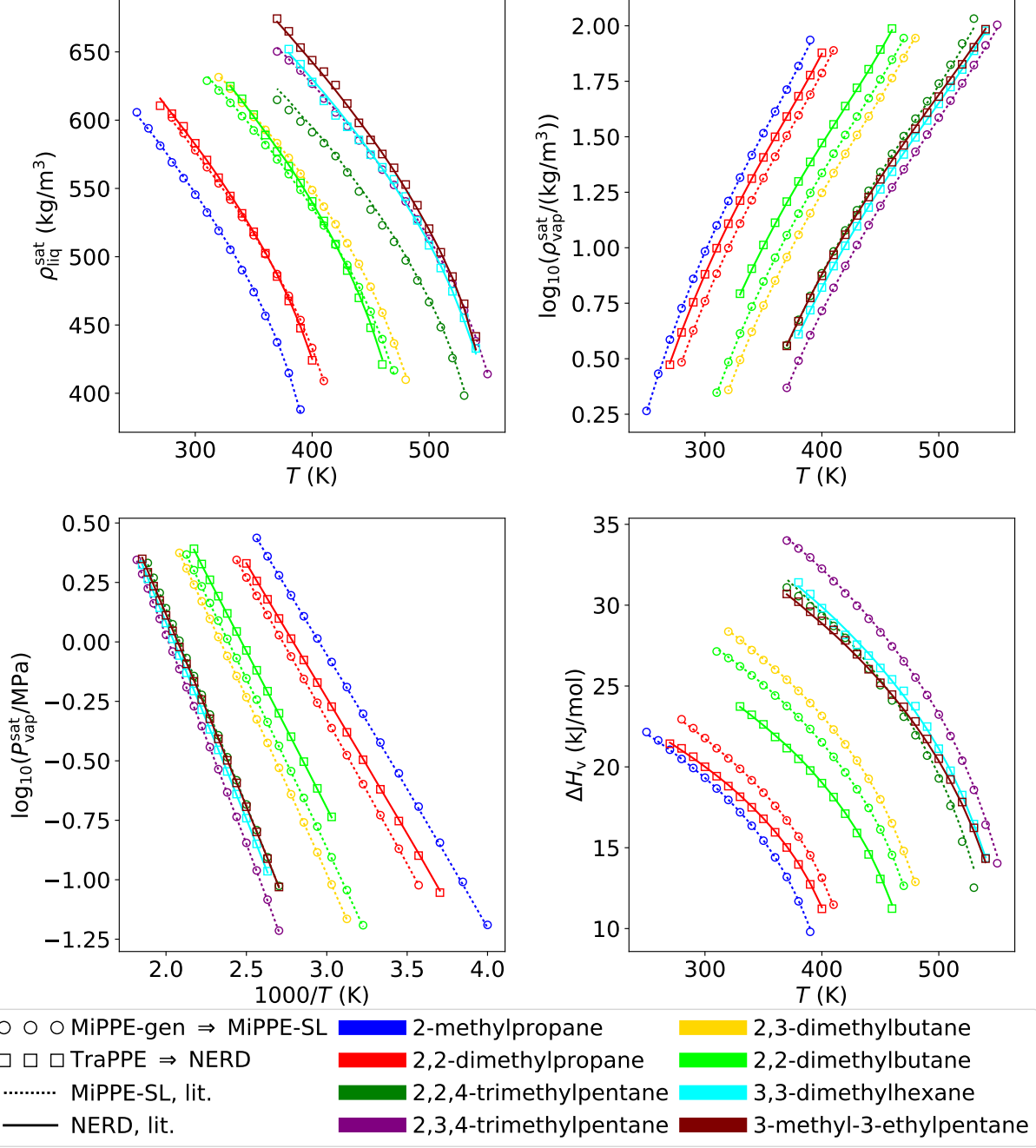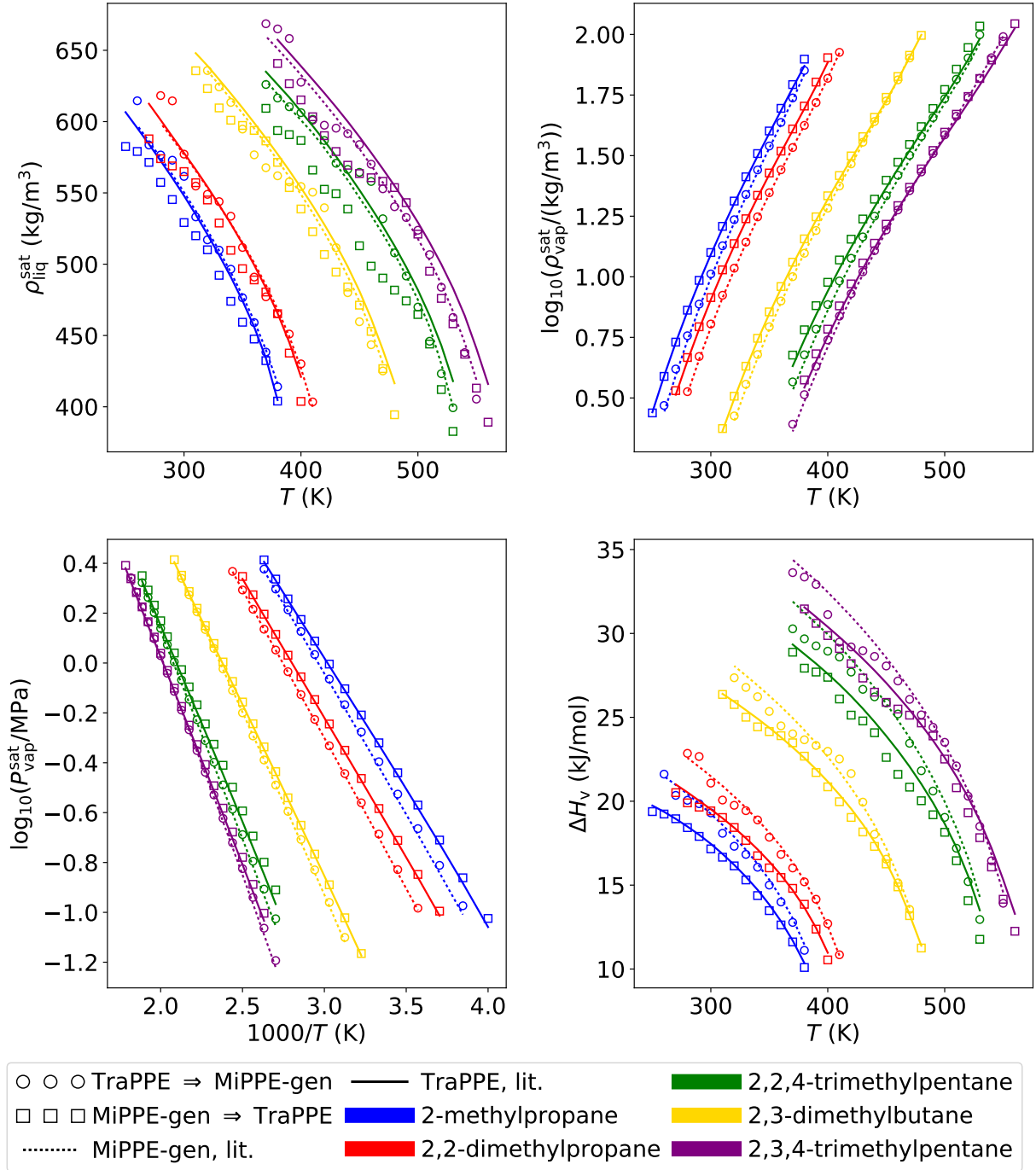
Figure 3: Comparison between MBAR-GCMC estimates (symbols, $\theta_{rr} \neq \theta_{ref}$) and MBAR-HR literature values (lines) with a constant repulsive exponent, i.e., $\lambda_{rr} = \lambda_{ref}$. MBAR predicts both liquid and vapor properties accurately for $\lambda_{rr} = \lambda_{ref}$. MBAR-GCMC estimates for the NERD and MiPPE-SL force fields are computed using configurations sampled from TraPPE and MiPPE-gen, respectively. Top-left, top-right, bottom-left, and bottom-right panels correspond to saturated liquid density, saturated vapor density, saturated vapor pressure, and enthalpy of vaporization, respectively.

Figure 4: Comparison between MBAR-GCMC estimates (symbols, $\theta_{rr} \neq \theta_{ref}$) and MBAR-HR literature values (lines) with a non-constant repulsive exponent, i.e., $\lambda_{rr} \neq \lambda_{ref}$. MBAR predicts only vapor properties accurately for $\lambda_{rr} \neq \lambda_{ref}$. MBAR-GCMC estimates for the TraPPE force field are computed using configurations sampled from MiPPE-gen, and vice versa. Top-left, top-right, bottom-left, and bottom-right panels correspond to saturated liquid density, saturated vapor density, saturated vapor pressure, and enthalpy of vaporization, respectively.
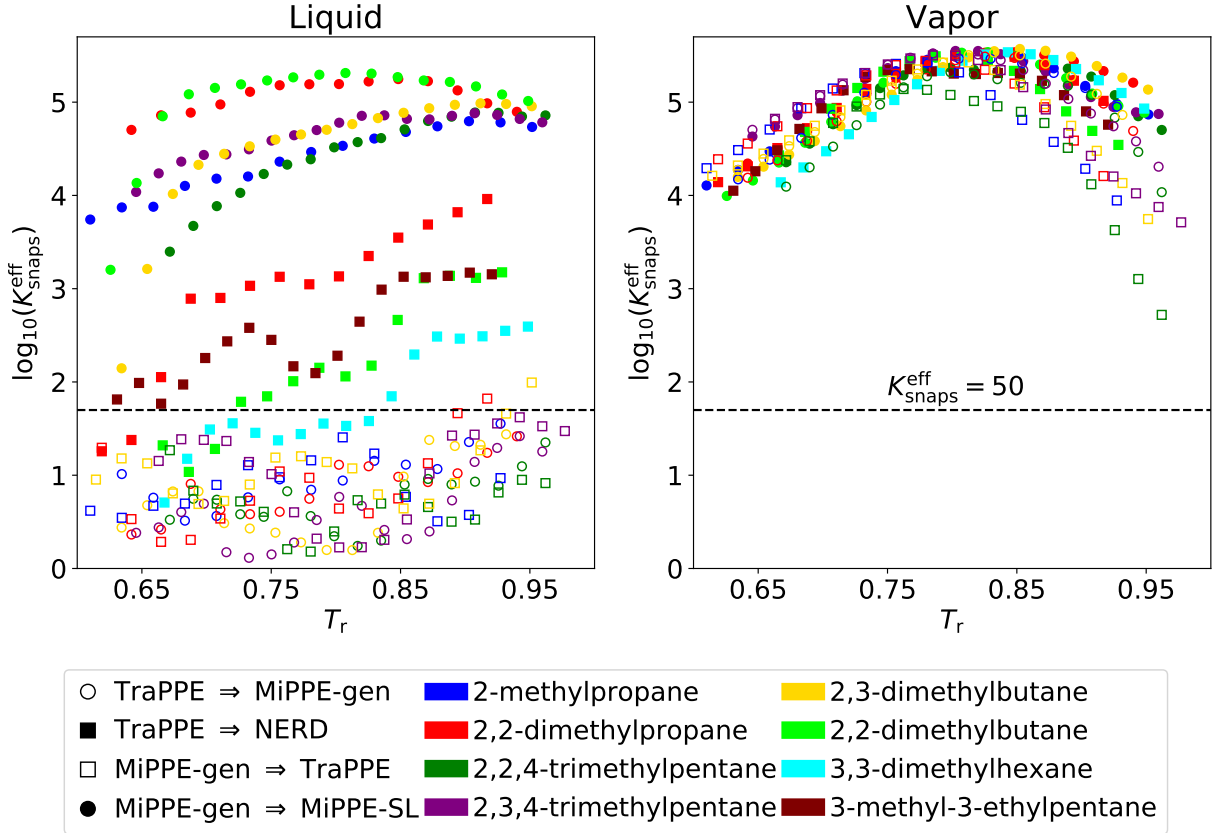
Figure 5: Number of effective snapshots ($K_{\mathrm{snaps}}^{\mathrm{eff}}$) in the liquid (left panel) and vapor (right panel) phases. Good overlap ($K_{\mathrm{snaps}}^{\mathrm{eff}} \gg 50$) is achieved in the vapor phase for each system while poor overlap in liquid phase ($K_{\mathrm{snaps}}^{\mathrm{eff}} < 50$) is observed for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$. The amount of overlap explains why GCMC-MBAR is highly reliable in both phases for $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}}$ (see Figure 3), while GCMC-MBAR is not reliable in the liquid phase for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$ (see Figure 4). Color scheme is the same as Figures 3 and 4. Closed and open symbols correspond to $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}}$ and $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$, respectively.
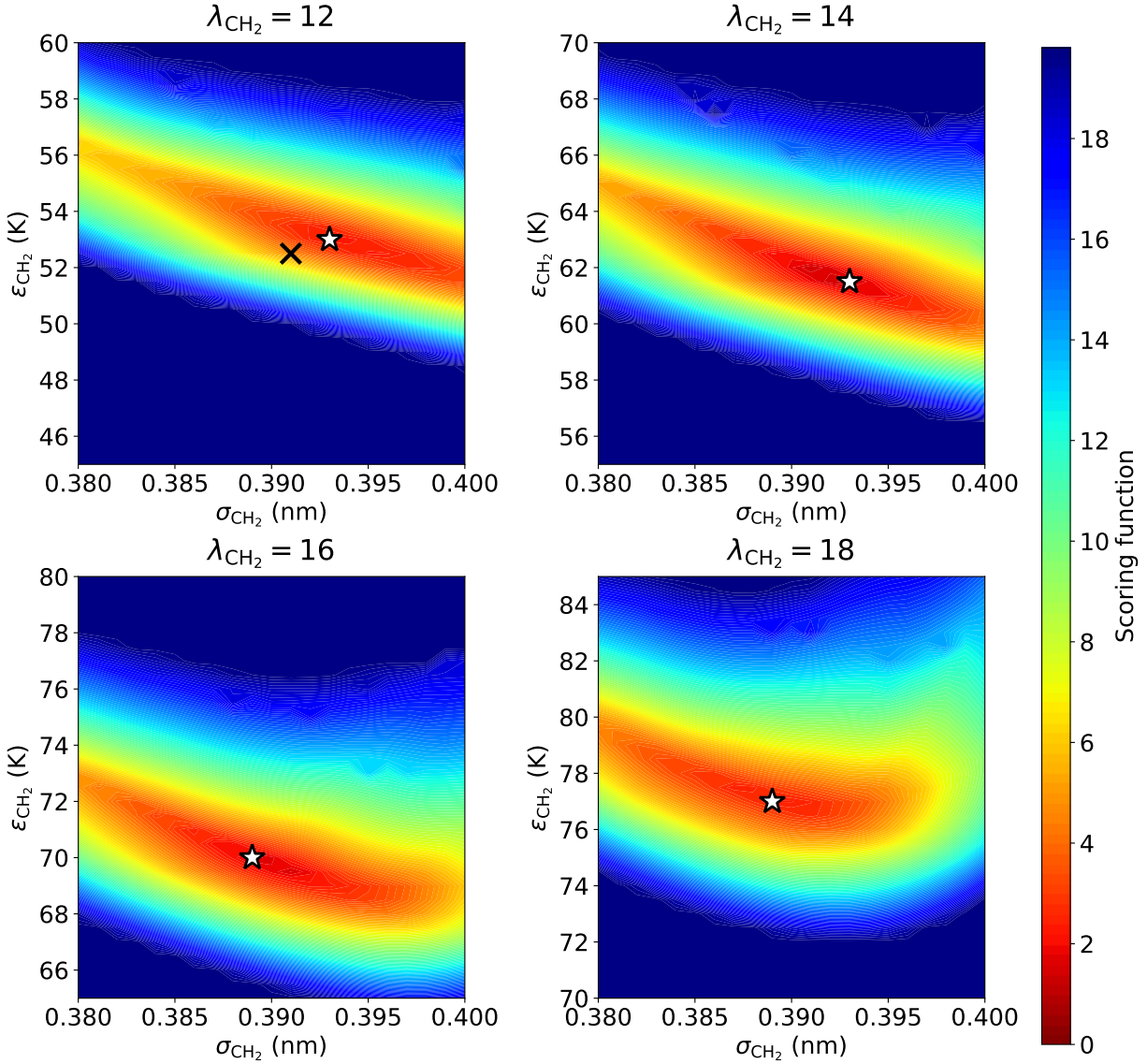
Figure 6: Scoring function values with respect to ϵ_{CH_2} and σ_{CH_2} for cyclohexane. GCMC-MBAR enables rapid optimization of Mie λ-6 parameters from a single reference force field ($\theta_{ref} = \theta_{TraPPE}$, depicted as a black "X"). Top-left, top-right, bottom-left, and bottom-right panels correspond $\lambda_{CH_2} = 12$, $\lambda_{CH_2} = 14$, $\lambda_{CH_2} = 16$, $\lambda_{CH_2} = 18$, respectively. White stars represent the optimal parameter set, i.e., the lowest value of S, for each λ_{CH_2}.
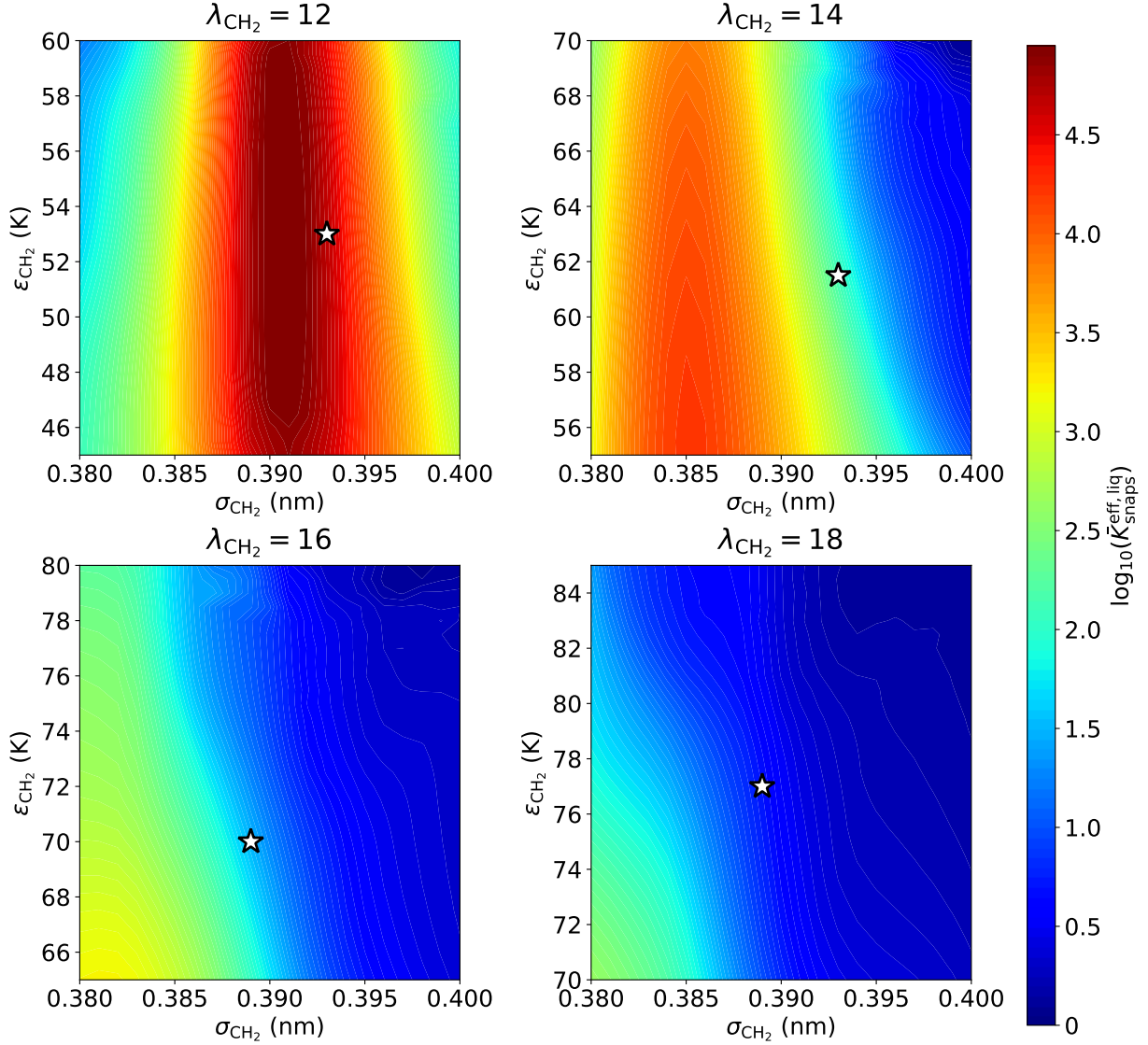
Figure 7: Average number of effective snapshots in the liquid phase ($\bar{K}_{\mathrm{snaps}}^{\mathrm{eff,liq}}$) with respect to $\epsilon_{\mathrm{CH_2}}$ and $\sigma_{\mathrm{CH_2}}$ for cyclohexane. $K_{\mathrm{snaps}}^{\mathrm{eff,liq}} \gg 50$ over a wide range of parameters when $\lambda_{\mathrm{rr}} = \lambda_{\mathrm{ref}} = 12$ (top-left panel), while $K_{\mathrm{snaps}}^{\mathrm{eff,liq}}$ is typically less than 50 for $\lambda_{\mathrm{rr}} \neq \lambda_{\mathrm{ref}}$ (other panels). Top-left, top-right, bottom-left, and bottom-right panels correspond $\lambda_{\mathrm{CH_2}} = 12$, $\lambda_{\mathrm{CH_2}} = 14$, $\lambda_{\mathrm{CH_2}} = 16$, $\lambda_{\mathrm{CH_2}} = 18$, respectively. White stars are the same as Figure 6.

Figure 8: Refined scoring function values with respect to ϵ_{CH_2} and σ_{CH_2} for cyclohexane. Pseudo-optimal parameter sets from first iteration serve as reliable reference parameter sets for refining the optimization. Top-left, top-right, bottom-left, and bottom-right panels correspond $\lambda_{CH_2} = 14$, $\lambda_{CH_2} = 16$, $\lambda_{CH_2} = 18$, and $\lambda_{CH_2} = 20$, respectively. White star represents the overall optimal parameter set (MiPPE), i.e., the lowest value of S between all four λ_{CH_2} values. Black "X" depicts the single reference force field for each λ_{CH_2}, i.e., $\lambda_{rr} = \lambda_{ref}$.

Figure 9: Minimum number of effective snapshots $(\min(K_{snaps}^{eff}))$ with respect to ϵ_{CH_2} and σ_{CH_2} for cyclohexane. Optimization has converged as $\min(K_{snaps}^{eff}) \gg 50$ for the optimal ϵ_{CH_2}, σ_{CH_2}, λ_{CH_2} parameter set. Top-left, top-right, bottom-left, and bottom-right panels correspond $\lambda_{CH_2} = 14$, $\lambda_{CH_2} = 16$, $\lambda_{CH_2} = 18$, and $\lambda_{CH_2} = 12$, respectively. White star represents the optimal parameter set, i.e., the lowest value of S, for a given λ_{CH_2}
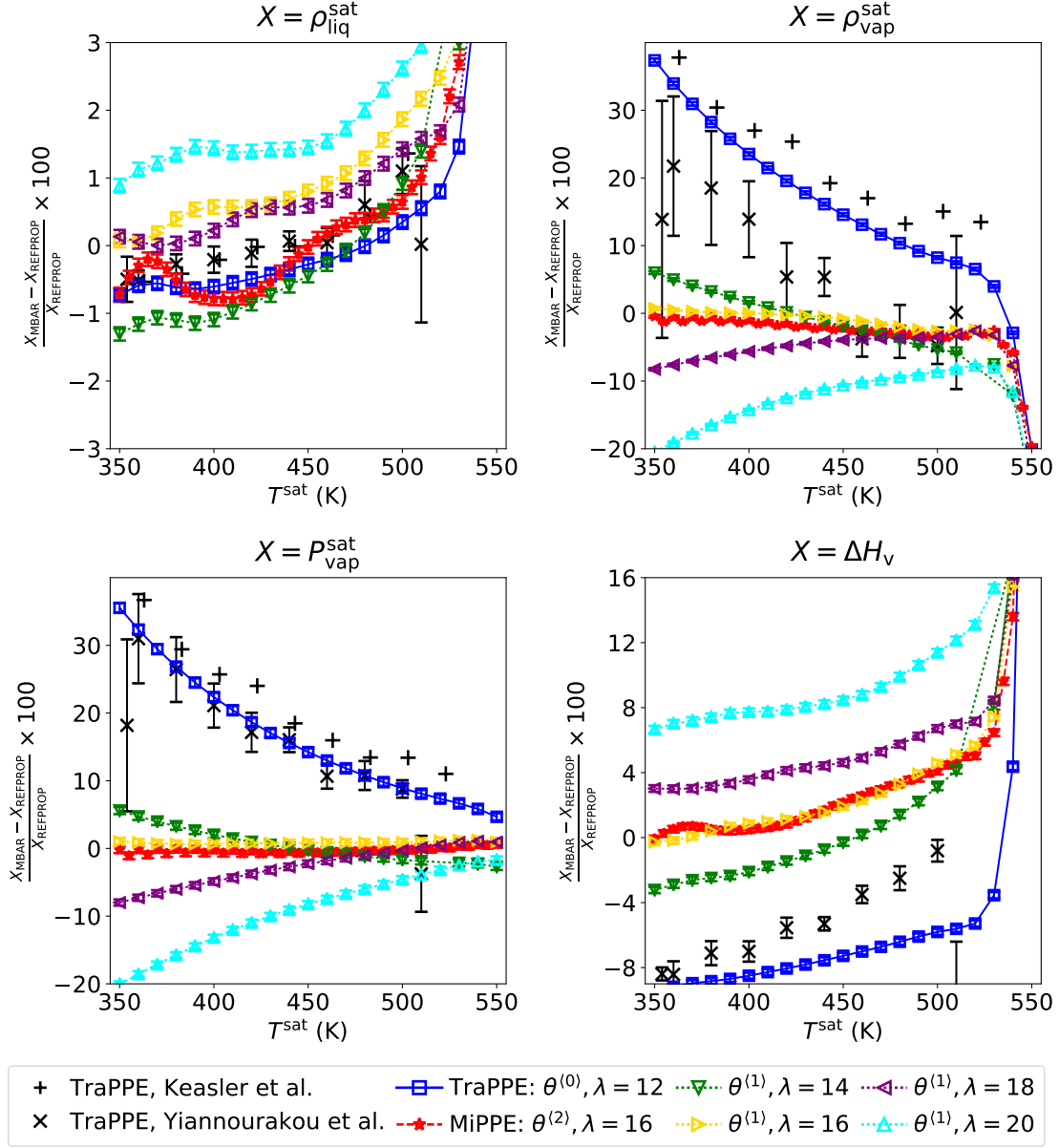
Figure 10: Percent deviations between coexistence properties computed with GCMC-MBAR (for $\theta_{rr} = \theta_{ref}$) and REFPROP. The first iteration parameter sets provide significant improvement while the final optimized parameter set is the most accurate. Top-left, top-right, bottom-left, and bottom-right panels correspond to saturated liquid density, saturated vapor density, saturated vapor pressure, and enthalpy of vaporization, respectively. When not depicted, uncertainties are smaller than one symbol size. TraPPE uncertainties are not provided by Keasler et al. Yiannourakou et al. obtained uncertainties for TraPPE using block averaging, but it is unclear what type of uncertainties are reported, i.e., standard error or 95% confidence intervals. Uncertainties for reference force fields (including TraPPE) are 95% confidence intervals obtained with bootstrap re-sampling of a single set of GCMC simulations. Uncertainties for the MiPPE force field are 95% confidence intervals from 20 independent replicate sets of GCMC simulations.