

# Configuration-Sampling-Based Surrogate Models for Rapid Parameterization of Non-bonded Interactions

Richard A. Messerly,<sup>\*,†</sup> S. Mostafa Razavi,<sup>‡</sup> and Michael R. Shirts<sup>¶</sup>

<sup>†</sup>*Thermodynamics Research Center, National Institute of Standards and Technology, Boulder,  
Colorado, 80305*

<sup>‡</sup>*Department of Chemical and Biomolecular Engineering, The University of Akron, Akron, Ohio,  
44325*

<sup>¶</sup>*Department of Chemical and Biological Engineering, University of Colorado, Boulder,  
Colorado, 80309*

E-mail: richard.messerly@nist.gov

## Abstract

In this study, we present an approach for rapid force field parameterization and uncertainty quantification of the non-bonded interaction parameters for classical force fields. The accuracy of most thermophysical properties, and especially vapor-liquid equilibria

---

Contribution of NIST, an agency of the United States government; not subject to copyright in the United States.

(VLE), obtained from molecular simulation depends strongly on the non-bonded interactions. Traditionally, non-bonded interactions are parameterized to agree with macroscopic properties by performing large amounts of direct molecular simulation. Due to the computational cost of molecular simulation, surrogate models (i.e. efficient models that approximate direct molecular simulation results) are an essential tool for high-dimensional parameterization and uncertainty quantification of non-bonded interactions. The present study compares two different configuration-sampling-based surrogate models, namely, Multistate Bennett Acceptance Ratio (MBAR) and Pair Correlation Function Rescaling (PCFR). MBAR and PCFR are coupled with the Isothermal Isochoric (ITIC) thermodynamic integration method for estimating vapor-liquid saturation properties. We find that MBAR and PCFR are complementary in their roles. Specifically, PCFR is preferred when exploring distant regions of the parameter space while MBAR is better in the local domain.

## 1 Introduction

Molecular simulation is an invaluable tool in many fields of science and engineering. One of its many purposes is the efficient prediction of thermophysical properties such as saturated liquid density ( $\rho_l^{\text{sat}}$ ), saturated vapor density ( $\rho_v^{\text{sat}}$ ), saturated vapor pressure ( $P_v^{\text{sat}}$ ), enthalpy of vaporization, surface tension, viscosity, etc. The quantitative reliability of the estimated property values depends almost entirely on the force field (i.e. the molecular model) employed in the molecular simulation. For this reason, force field development is an important area in molecular modeling. Since the intramolecular and electrostatic contributions to the force field are often parameterized with *ab initio* results, this study focuses on parameterizing the non-bonded (intermolecular) van der Waals interactions with vapor-liquid equilibria (VLE) properties.

Recently, several groups have developed united-atom (UA) based force fields for predicting vapor-liquid equilibria. Potoff and Bernard-Brunel<sup>1</sup> demonstrated that the Mie  $\lambda$ -6 potential (i.e. a three parameter Lennard-Jones) provides considerable improvement at predicting both  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  as compared to UA LJ 12-6 models, such as TraPPE-UA<sup>2</sup> and NERD.<sup>3</sup> Subsequently, Hemmen and Gross<sup>4</sup> implemented the Mie  $\lambda$ -6 potential but introduced an additional fitting parameter by using anisotropic-united-atom (AUA) sites for terminal groups (TAMie). Although the Mie  $\lambda$ -6 potential has shown significant promise in providing highly accurate force fields for VLE, it should not be viewed as a panacea. In fact, the force fields developed by Shah et al.<sup>5</sup> (TraPPE-UA2) and Errington and Panagiotopoulos<sup>6</sup> for ethane appear to reproduce experimental VLE data just as reliably as the Potoff and TAMie models. Errington’s force field is a four-parameter AUA Exponential-6 model while TraPPE-UA2 is a three-parameter AUA LJ 12-6 model. In general, the increased accuracy of the Potoff, TAMie, Errington, and TraPPE-UA2 force fields compared to the TraPPE-UA and NERD models has come at the cost of additional model parameters.

Due to the increased accuracy and complexity (i.e. number of parameters) of modern force fields, sophisticated high-dimensional optimization,<sup>7</sup> multi-objective Pareto front,<sup>8-10</sup> and uncertainty quantification (UQ) methods<sup>11-14</sup> should play a key role in force field development. However, these methods are not tractable when molecular simulation is performed at each step of the algorithm, as this may necessitate  $O(10^2 \text{ to } 10^6)$  simulations.<sup>15</sup> For this reason, surrogate models are essential to render these methods computationally feasible. In this context, “surrogate model” is an umbrella term that covers all methods that predict simulation results without requiring direct simulation using the given force field. The obvious benefit of a surrogate model is the reduction of computational cost. Typical surrogate models are a few orders of magnitude faster than direct simulation. For this reason, UQ and Pareto front studies rely heavily on surrogate models to replace

molecular simulation.<sup>8-14</sup>

The first type of surrogate model is a mathematical or statistical model (sometimes referred to as a meta-model<sup>16-18</sup>). These require little or no understanding of the underlying physics or what each force field parameter means on a molecular level. Instead, they simply interpolate and smooth the simulation output in the corresponding parameter space.<sup>11,14,19</sup> Although meta-models are computationally cheap to evaluate, developing reliable meta-models can be an arduous task and may require large amounts of molecular simulations. In addition, while these surrogate models are typically reliable for interpolation, extrapolation is dubious for certain model forms.

The second type of surrogate model is an analytic equation-of-state model. Equation-of-state models relate the force field parameters to different types of engineering equations-of-state. The simplest equation-of-state surrogate model is likely the corresponding states model. Typically, a corresponding states model is a correlation fit to reduced properties (simulation output scaled by the force field parameters). Examples of this type of surrogate model can be found for the single-site Lennard-Jones fluid (both with tail corrections and truncated<sup>20-22</sup>) as well as the two-site Lennard-Jones plus point quadrupole.<sup>23</sup> More sophisticated equation-of-state surrogate models are PC-SAFT<sup>24</sup> and SAFT- $\gamma$ .<sup>25,26</sup> The PC-SAFT surrogate model has been used extensively to develop the TAMie force field by relating the Mie parameters to parameters in the PC-SAFT equation-of-state.<sup>4,27,28</sup>

In this study, we investigate a third type of surrogate model, namely, configuration-sampling-based surrogate models. Configuration-sampling-based surrogate models rely on atomic configurations that are sampled by simulating the reference force field(s) to predict observables for a non-simulated force field. The primary assumption is that the reference force field(s) has a distribution of configurations similar to the desired force field. For this reason, the reference force field(s) plays a significant role in the accuracy of this class of surrogate models and must be chosen judiciously. A key advantage

in configuration-sampling-based surrogate models is that they are compatible with any non-bonded functional form and can be used with both all-atom (AA) and coarse-grained (UA, AUA, etc.) force fields, whereas the SAFT- $\gamma$  and PC-SAFT surrogate models are limited to coarse-grained models. The two configuration-sampling-based surrogate models that we examine in this study are Multistate Bennett Acceptance Ratio (MBAR) and Pair Correlation Function Rescaling (PCFR). While MBAR is a well-established method, PCFR is a novel approach set forth in this study.

Section 2 discusses the methodology, starting with the force fields, the simulation conditions, data analysis, and surrogate model derivation. Section 3 compares the results for MBAR and PCFR. This comparison is made for simple systems, namely, united-atom representations of ethane, hexafluoroethane, propane, *n*-butane, and *n*-octane, but it is applicable to any compound and force field. Section 4 discusses recommendations and limitations regarding the implementation of these surrogate models. Section 4 also provides an algorithm for parameterization of non-bonded potentials which is demonstrated in Section 5. Finally, Section 6 summarizes the conclusions from this work.

## 2 Methods

### 2.1 Force Field

We emphasize that the methodology proposed in Sections 2.2-2.3 is applicable to any force field. Specifically, the configuration-sampling-based surrogate models can be applied to united-atom (UA) or all-atom (AA) based force fields and to LJ 12-6, Mie  $\lambda$ -6, Exp-6, or any other non-bonded functional form. In this study, however, we focus on a specific subset of force fields, namely, UA Mie  $\lambda$ -6. This model type was selected as it has received significant attention in recent years for development of accurate hydrocarbon

force fields.<sup>1,4,29-31</sup>

The Mie  $\lambda$ -6 is a three-parameter non-bonded central potential of the form:

$$u^{\text{vdw}}(\epsilon, \sigma, \lambda; r) = \left(\frac{\lambda}{\lambda - 6}\right) \left(\frac{\lambda}{6}\right)^{\frac{6}{\lambda-6}} \epsilon \left[ \left(\frac{\sigma}{r}\right)^\lambda - \left(\frac{\sigma}{r}\right)^6 \right] \quad (1)$$

where  $u^{\text{vdw}}$  is the van der Waals interaction,  $\sigma$  is the distance ( $r$ ) where  $u^{\text{vdw}} = 0$ ,  $-\epsilon$  is the energy of the potential at the minimum (i.e.  $u^{\text{vdw}} = -\epsilon$  and  $\frac{\partial u^{\text{vdw}}}{\partial r} = 0$  for  $r = r_{\text{min}}$ ), and  $\lambda$  is the repulsive exponent. Note that the Mie potential reduces to the LJ 12-6 potential for  $\lambda = 12$ . Therefore, Equation 1 can be considered a generalized Lennard-Jones where the repulsive exponent is a parameter. Although an attractive exponent of 6 has a strong theoretical basis,  $\lambda = 12$  is a historical artifact that was chosen primarily for computational purposes.<sup>32</sup> For the same reason (i.e. computational efficiency), a common practice is to use integer values of  $\lambda$  in Equation 1.

Non-bonded interactions between two different site types (i.e. cross-interactions) are determined using Lorentz-Berthelot combining rules<sup>32</sup> for  $\epsilon$  and  $\sigma$  with an arithmetic mean for the repulsive exponent ( $\lambda$ ) (as recommended by Potoff and Bernard-Brunel<sup>1</sup>):

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \quad (2)$$

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad (3)$$

$$\lambda_{ij} = \frac{\lambda_{ii} + \lambda_{jj}}{2} \quad (4)$$

where the  $ij$  subscript refers to cross-interactions and the subscripts  $ii$  and  $jj$  refer to same-site interactions. Section S.I.3 of the Supporting Information provides the TraPPE-UA and Potoff non-bonded parameters for both same-site interactions and cross-interactions.

We use the same intramolecular potential as the TraPPE-UA and Potoff force fields, which was in large part adopted from the well-known Optimized Potential for Liquid

Simulations (OPLS-UA) force field.<sup>33,34</sup> Specifically, the simulations performed in this study use fixed bond-lengths, harmonic angular potential, and Fourier series for the dihedral torsional interactions. To be consistent with these force fields, non-bonded interactions between united-atom sites in the same molecule are only included if they are separated by at least four neighboring sites (i.e. we exclude 1-2, 1-3, and 1-4 non-bonded intramolecular interactions). Section SII.3 of the Supporting Information provides the equations and parameters for intramolecular interactions.

## 2.2 Isothermal Isochoric Thermodynamic Integration

In this study, isothermal isochoric (ITIC) thermodynamic integration<sup>35</sup> is used to determine  $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$  for each force field and molecule. The equations for ITIC are:

$$\frac{A^{\text{dep}}}{R_g T^{\text{sat}}} = \int_0^{\rho_1^{\text{sat}}} \frac{Z-1}{\rho} \partial \rho |_{T=T^{\text{IT}}} + \int_{T^{\text{IT}}}^{T^{\text{sat}}} U^{\text{dep}} \partial \left( \frac{1}{R_g T} \right) |_{\rho=\rho_1^{\text{sat}}} \quad (5)$$

$$\rho_v^{\text{sat}} \approx \rho_1^{\text{sat}} \exp \left( \frac{A^{\text{dep}}}{R_g T^{\text{sat}}} + Z_1^{\text{sat}} - 1 - 2B_2 \rho_v^{\text{sat}} - \frac{3}{2} B_3 \rho_v^{\text{sat}^2} \right) \quad (6)$$

$$P_v^{\text{sat}} \approx (1 + B_2 \rho_v^{\text{sat}} + B_3 \rho_v^{\text{sat}^2}) \rho_v^{\text{sat}} R_g T^{\text{sat}} \quad (7)$$

$$Z_1^{\text{sat}} = \frac{P_v^{\text{sat}}}{\rho_1^{\text{sat}} R_g T^{\text{sat}}} \quad (8)$$

where  $A^{\text{dep}} \equiv A - A^{\text{ig}}$  is the Helmholtz free energy departure from ideal gas for temperature ( $T$ ) equal to the saturation temperature ( $T^{\text{sat}}$ ) and density ( $\rho$ ) equal to the saturated liquid density ( $\rho_1^{\text{sat}}$ ),  $U^{\text{dep}} \equiv U - U^{\text{ig}}$  is the internal energy departure,  $Z_1^{\text{sat}}$  is the saturated liquid compressibility factor ( $Z$ ),  $B_2$  is the second virial coefficient,  $B_3$  is the third virial coefficient,  $T^{\text{IT}}$  is the isothermal temperature, and  $R_g$  is the universal gas constant. (Note

that our definitions for  $A^{\text{dep}}$  and  $U^{\text{dep}}$  are slightly different than those used by Razavi,<sup>35</sup> where the energy departures are dimensionless, but are consistent with Elliott and Lira.<sup>36)</sup>

The ITIC equations are solved iteratively to ensure self-consistency. Specifically, the  $Z_1^{\text{sat}}$  value calculated in Equation 8 is then used to compute a new value of  $T^{\text{sat}}$  by interpolating the simulation results for  $Z$  along the corresponding isochore. The  $T^{\text{sat}}$  value in Equation 5 is updated and  $\frac{A^{\text{dep}}}{R_g T^{\text{sat}}}$  is recomputed. The new values for  $Z_1^{\text{sat}}$  and  $\frac{A^{\text{dep}}}{R_g T^{\text{sat}}}$  are then used in Equation 6 to solve for  $\rho_v^{\text{sat}}$ . Likewise,  $P_v^{\text{sat}}$  and, subsequently,  $Z_1^{\text{sat}}$  are recalculated using Equations 7 and 8, respectively. This process is repeated until the value of  $Z_1^{\text{sat}}$  (or alternatively  $T^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , or  $P_v^{\text{sat}}$ ) has converged to within a predefined tolerance.

The  $B_2$  and  $B_3$  values found in Equations 6-7 can be determined in several different ways: using experimental data or correlations fit to data (i.e. REFPROP,<sup>37</sup> ThermoData Engine (TDE),<sup>38</sup> etc.), calculated with Mayer-sampling Monte Carlo,<sup>39,40</sup> or obtained by extrapolating low density simulation results.<sup>35</sup> In this study, we utilize the  $B_2$  and  $B_3$  values from REFPROP, primarily for simplicity. This decision was made after first validating that the ITIC  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  values for  $n$ -alkanes (calculated using REFPROP  $B_2$  and  $B_3$ ) are consistent with the literature values reported using Gibbs Ensemble Monte Carlo (GEMC) and Grand Canonical Monte Carlo (GCMC) for the TraPPE-UA and Potoff force fields, respectively<sup>1,41</sup> (see Section SI.II.3 of the Supporting Information). We also verified that the ITIC results for  $n$ -octane do not differ significantly when using Mayer-sampling Monte Carlo  $B_2$  and  $B_3$  values reported by Schultz and Kofke<sup>39</sup> instead of the REFPROP  $B_2$  and  $B_3$  values (see Section SI.II.3 of the Supporting Information).

ITIC requires on the order of 10 NVT (constant number of molecules, constant volume, constant temperature) simulations along a supercritical isotherm, i.e.  $T^{\text{IT}} > T_c$  (where  $T_c$  is the experimental critical temperature). As recommended by Razavi,<sup>35</sup> we use a value around 1.2 for the isotherm reduced temperature ( $T_r^{\text{IT}} \equiv \frac{T^{\text{IT}}}{T_c}$ ), i.e.  $T^{\text{IT}} \approx 1.2T_c$ . Two or three additional NVT simulations are performed along different isochores that intersect



the saturated liquid curve ( $\rho = \rho_1^{\text{sat}}$ ). Thus, each of the ITIC state points corresponds to either a liquid or supercritical fluid. The specific state points simulated in this study are provided in Section SI.I.1 of the Supporting Information. Section SI.II.3 of the Supporting Information also contains an example of the ITIC data analysis.

One advantage of ITIC is that all the simulations are performed in the NVT ensemble. In fact, although ITIC requires roughly three simulations for each saturation temperature, the total simulation time is typically comparable to the traditional VLE methods, i.e. GEMC or GCMC. This is primarily because NVT systems converge quickly, as they do not require expensive particle insertion/deletion or volume fluctuation moves. NVT simulations should also foster reproducibility, as there are fewer user decisions that can introduce error (although the choice of thermostat can be important<sup>42</sup>). By contrast, it was demonstrated recently that simulation practitioners struggled to generate reproducible results in the NPT ensemble.<sup>43</sup>

Furthermore, the NVT ensemble is amenable to both molecular dynamics (MD) and Monte Carlo (MC) simulations. Thus, practitioners can implement ITIC with their preferred simulation software. We use GROMACS,<sup>44</sup> as it is an extremely fast, parallelized, graphics processing unit (GPU) optimized, open-source, MD code. (Sample input (.mdp, .top, .gro) files are provided in Section SI.I.4 of the Supporting Information.) Obtaining VLE properties from MD simulations has some advantages. For example, MD methods are ideal for highly branched compounds whereas traditional GEMC or GCMC methods may struggle to reach equilibrium due to the low acceptance rate of particle insertions.<sup>45</sup> For the same reason, GEMC and GCMC are typically limited to  $T_r^{\text{sat}} > 0.7$  or  $0.6$  (depending on the molecular structure and  $\rho_1^{\text{sat}}$ ) whereas ITIC can provide accurate VLE estimates for  $T_r^{\text{sat}} \approx 0.45$ . However, while GEMC and GCMC can be used for  $T_r^{\text{sat}} \approx 0.95$ , one disadvantage of ITIC is that Equations 6-7 require higher-order virial expansion terms ( $B_4, B_5$ , etc.) for  $T_r^{\text{sat}} > 0.85$ .<sup>35</sup> In this study, we use ITIC to obtain  $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$  for

$$0.45 < T_r^{\text{sat}} < 0.85.$$

The primary reason we employ ITIC is to calculate saturated properties ( $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$ ) from  $U^{\text{dep}}$  and  $Z$  at specified state points ( $\rho - T$ ). Internal energies and pressures (compressibility factors) are intimately related to the non-bonded interactions and the atomic configurations. Therefore,  $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$  can be predicted for any non-bonded interactions by combining configuration-sampling-based surrogate models for predicting  $U^{\text{dep}}$  and  $Z$  with the ITIC analysis, Equations 5-8. Converting  $U^{\text{dep}}$  and  $Z$  to  $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$  is important because large amounts of evaluated experimental VLE data are available through databases, such as the Thermodynamics Research Center (TRC) source database, whereas experimental data for  $U^{\text{dep}}$  and  $Z$  are scarce.

## 2.3 Surrogate Models

ITIC requires  $U^{\text{dep}}$  and  $Z$  for each state point ( $\rho$ - $T$ ) to calculate  $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$ . Thus, predicting  $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$  for a given molecule and force field necessitates predicting  $U^{\text{dep}}$  and  $Z$ . Therefore, the aim of the surrogate models presented in this section is to predict  $U^{\text{dep}}$  and  $Z$  for a given state point with as little direct simulation as possible.

The configuration-sampling-based surrogate models (MBAR and PCFR) are well-suited for the task at hand as energies and pressures are calculated directly from the coordinates of interacting particles. In such surrogate models, we carry out a set of simulations at the ITIC conditions using the reference force field(s), and the configurations obtained from the reference force field(s) are used to estimate the VLE properties for a non-simulated force field.

### 2.3.1 Multistate Bennett Acceptance Ratio

Importance sampling is a commonly used statistical technique for computing averages of properties for one model by reweighting configurations sampled with another model based on the ratio of probabilities for the two models. In chemistry and chemical physics, importance sampling from one or a set of simulations to another set of simulation conditions can be implemented using the MBAR algorithm.<sup>46,47</sup>

With MBAR the expectation  $\langle O(\theta) \rangle$  for force field  $(\theta)$  of any given observable  $(O)$  can be expressed as:

$$\langle O(\theta) \rangle = \sum_{n=1}^N O(\mathbf{x}_n; \theta) W_n(\theta) \quad (9)$$

where  $\mathbf{x}_n$  are configurations sampled from one or more reference force fields  $(\theta_{\text{ref}})$ ,  $O(\mathbf{x}_n; \theta)$  is the observable value using force field  $\theta$  with configurations  $\mathbf{x}_n$ , and  $W_n(\theta)$  is the weight of the  $n^{\text{th}}$  configuration using force field  $\theta$ , calculated by using:

$$W_n(\theta) = \frac{\exp[\hat{f}(\theta) - u(\mathbf{x}_n; \theta)]}{\sum_{k=1}^K N_k \exp[\hat{f}(\theta_{\text{ref},k}) - u(\mathbf{x}_n; \theta_{\text{ref},k})]} \quad (10)$$

where the reduced free energies  $(\hat{f}(\theta))$  are calculated with:

$$\hat{f}(\theta) = -\ln \sum_{n=1}^N \frac{\exp[-u(\mathbf{x}_n; \theta)]}{\sum_{k=1}^K N_k \exp[\hat{f}(\theta_{\text{ref},k}) - u(\mathbf{x}_n; \theta_{\text{ref},k})]} \quad (11)$$

where  $K$  is the number of reference force fields,  $N = \sum_k N_k$  is the total number of snapshots for all  $K$  reference force fields,  $N_k$  are the total number of snapshots from the  $k^{\text{th}}$  reference force field,  $\theta_{\text{ref},k}$  is the  $k^{\text{th}}$  reference (i.e. simulated) force field, and  $u(\mathbf{x}_n; \theta) = \beta U(\mathbf{x}_n; \theta)$  is the reduced potential energy evaluated with  $\theta$  for configuration

$\mathbf{x}_n$  where  $\beta = \frac{1}{k_B T}$  and  $k_B$  is the Boltzmann constant.

Note that  $\hat{f}(\theta_{\text{ref},k})$  is required to evaluate the denominator of Equations 10-11. The values for  $\hat{f}(\theta_{\text{ref},k})$  are obtained by solving a system of  $K$  equations for self-consistency. Specifically, an initial guess for  $\hat{f}(\theta_{\text{ref},k})$  is used to evaluate Equation 11 with  $\theta = \theta_{\text{ref},k}$  to obtain updated values of  $\hat{f}(\theta_{\text{ref},k})$ . This process is repeated until the values for  $\hat{f}(\theta_{\text{ref},k})$  converge to within a desired tolerance. Although solving the MBAR system of equations for self-consistency may require several iterations, fortunately, once this process has been performed  $\hat{f}(\theta)$  (for an arbitrary  $\theta$ ) is evaluated without further iteration.

For the specific case of predicting  $U^{\text{dep}}$  and  $Z$  for a non-sampled force field, expressed by the set of force field parameters  $\theta$ , the MBAR-based estimators for the departure internal energy and compressibility can be written as:

$$\langle U^{\text{dep}}(\theta) \rangle = \sum_{n=1}^N U^{\text{dep}}(\mathbf{x}_n; \theta) W_n(\theta) \quad (12)$$

$$\langle Z(\theta) \rangle = \sum_{n=1}^N Z(\mathbf{x}_n; \theta) W_n(\theta) \quad (13)$$

where the energies and forces are computed using force field  $\theta$  for each configuration ( $\mathbf{x}_n$ ) to determine  $U^{\text{dep}}(\mathbf{x}_n; \theta)$  and  $Z(\mathbf{x}_n; \theta)$  (from the virial pressure<sup>32</sup>), respectively, while the weights ( $W_n(\theta)$ ) are again calculated using Equations 10-11.

The performance of MBAR depends strongly on good phase space overlap, meaning that the configurations sampled by the reference force field(s) must represent a significant portion of the “true” configurations that the non-simulated force field would sample.<sup>48</sup> If the configurational overlap is small, the MBAR estimates are often dominated by a few configurations, which are likely not representative of the ensemble that would be generated by direct simulation of force field  $\theta$ . The amount of overlap can be quantified

by the number of effective samples ( $N_{\text{eff}}$ ),<sup>49</sup> using Kish’s formula:

$$N_{\text{eff}} = \frac{(\sum_n W_n)^2}{\sum_n W_n^2} \quad (14)$$

which reduces to  $N_{\text{eff}} = (\sum_n W_n^2)^{-1}$  when the weights are normalized. This has the property that when the weights are equal,  $N_{\text{eff}} = N$ , when all but one weight is negligible,  $N_{\text{eff}} \approx 1$ , and behaves appropriately for intermediate cases. In the case of poor overlap ( $N_{\text{eff}} \approx 1$ ), the predicted values of MBAR will demonstrate a strong bias and the uncertainties will likely be underestimated by the MBAR covariance matrix.

We discuss two methods for overcoming poor phase space overlap, namely, pair correlation function rescaling (PCFR) and configuration mapping. Configuration mapping is an attempt to *predict* how the probabilities of each configuration would change when changing from the reference force field ( $\theta_{\text{ref}}$ ) to the non-simulated force field ( $\theta$ ). The clearest example of this is the theory of corresponding states.<sup>50</sup> For the single-site Lennard-Jones fluid, the NVT simulation results of  $\theta_{\text{ref}}$  can be exactly weighted to predict the simulation properties for a different value of  $\sigma$  at a new volume,  $V = V_{\text{ref}} \frac{\sigma^3}{\sigma_{\text{ref}}^3}$ , where  $V_{\text{ref}}$  is the initial volume sampled using  $\theta_{\text{ref}}$ . This is because the energies (and thus the Boltzmann weights) will be exactly the same if we simply scale all of the coordinates by  $\frac{\sigma}{\sigma_{\text{ref}}}$ . Therefore, we can reuse (at least for that volume) all of the configurations to calculate new properties.

If there is more than one length scale, however, the same corresponding states tactics do not work. A general formalism for creating new, more suitable overlap from the old configurations is to use configuration mapping.<sup>51,52</sup> This approach can drastically increase the efficiency of calculating properties using configurations sampled from similar models as in the case of similar rigid water models or dipolar molecule lengths.<sup>51</sup> One important key, of course, is finding a simple transformation rule. The coordinate transformation free

energy equations are true for any transformation, but will only be useful if they increase the overlap between the non-sampled and reference configurations.

In the case of rigid water models and dipolar molecules, a linear transformation between shapes which preserves the center of mass has been shown to work very well.<sup>51</sup> In the case of small molecule hydrocarbons, however, the simplest transformation would be one which scales the centers of mass while keeping the intramolecular distances the same. This, however, gives properties of the new model at a new volume, which is not useful if we are performing canonical simulations. It is not immediately clear if there is a simple remapping between simulations with different  $\sigma$  at the same volume, but it seems unlikely, since the configurations must shift radial distributions in an inherently coupled way. We defer the development of a more general mapping formalism to a future study. We instead propose an alternative method, namely, PCFR. In essence, PCFR is an *approximate* way to map the coordinates for constant volume. We derive PCFR in the following section.

### 2.3.2 Pair Correlation Function Rescaling

PCFR is a method to predict  $U^{\text{dep}}$  and  $Z$  (the two necessary quantities for ITIC) for a non-simulated force field. Similar to MBAR, PCFR makes use of the configurations sampled from direct simulation of a reference force field ( $\theta_{\text{ref}}$ ) to predict these properties without direct simulation of a modified force field ( $\theta$ ). To derive the PCFR equations, we first assume that  $\theta_{\text{ref}}$  and  $\theta$  have the same intramolecular and Coulombic potentials. The departure internal energy and compressibility factors for a non-simulated force field can then be expressed as:

$$\langle U^{\text{dep}}(\theta) \rangle = \langle U^{\text{dep}}(\theta_{\text{ref}}) \rangle + U^{\text{vdw}}(\theta) - U^{\text{vdw}}(\theta_{\text{ref}}) \quad (15)$$

$$\langle Z(\theta) \rangle = \langle Z(\theta_{\text{ref}}) \rangle + Z^{\text{vdw}}(\theta) - Z^{\text{vdw}}(\theta_{\text{ref}}) \quad (16)$$

where  $\langle \dots \rangle$  denotes an ensemble average and  $U^{\text{vdw}}$  and  $Z^{\text{vdw}}$  are the van der Waals (i.e. the non-bonded, non-Coulombic) contributions to  $U$  and  $Z$ , respectively. Subsequently, for pair-wise additive central potentials we obtain the energy and compressibility equations that relate the pair correlation function (PCF) to  $U^{\text{vdw}}$  and  $Z^{\text{vdw}}$  for a polyatomic molecule:<sup>32</sup>

$$U^{\text{vdw}}(\theta) = 2\pi\rho R_g \sum_{i=1}^{N_s-1} \sum_{j>i}^{N_s} \int_0^\infty u_{ij}^{\text{vdw}}(\theta; r_{ij}) g_{ij}(\theta; r_{ij}) r_{ij}^2 \partial r_{ij} \quad (17)$$

$$Z^{\text{vdw}}(\theta) = -\frac{2\pi\rho}{3T} \sum_{i=1}^{N_s-1} \sum_{j>i}^{N_s} \int_0^\infty \frac{\partial u_{ij}^{\text{vdw}}(\theta; r_{ij})}{\partial r_{ij}} g_{ij}(\theta; r_{ij}) r_{ij}^3 \partial r_{ij} \quad (18)$$

where  $\rho$  is the number density (units of molecules per volume),  $N_s$  is the number of sites in a molecule,  $u_{ij}^{\text{vdw}}(\theta; r_{ij})$  is the van der Waals potential (units of kelvin, i.e. energy divided by the Boltzmann constant) between sites  $i$  and  $j$  with force field  $\theta$ , and  $g_{ij}(\theta; r_{ij})$  is the site-site radial distribution function (or “pair correlation function”) between sites  $i$  and  $j$  obtained with  $\theta$ . Substitution of Equations 17 and 18 into Equations 15 and 16, respectively, and combining the two integrals for  $U^{\text{vdw}}$  and  $Z^{\text{vdw}}$  gives:

$$\begin{aligned} \langle U^{\text{dep}}(\theta) \rangle &= \langle U^{\text{dep}}(\theta_{\text{ref}}) \rangle \\ &+ 2\pi\rho R_g \sum_{i=1}^{N_s-1} \sum_{j>i}^{N_s} \int_0^\infty [u_{ij}^{\text{vdw}}(\theta; r_{ij}) g_{ij}(\theta; r_{ij}) - u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij}) g_{ij}(\theta_{\text{ref}}; r_{ij})] r_{ij}^2 \partial r_{ij} \quad (19) \end{aligned}$$

$$\begin{aligned} \langle Z(\theta) \rangle &= \langle Z(\theta_{\text{ref}}) \rangle \\ &- \frac{2\pi\rho}{3T} \sum_{i=1}^{N_s-1} \sum_{j>i}^{N_s} \int_0^\infty \left[ \frac{\partial u_{ij}^{\text{vdw}}(\theta; r_{ij})}{\partial r_{ij}} g_{ij}(\theta; r_{ij}) - \frac{\partial u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})}{\partial r_{ij}} g_{ij}(\theta_{\text{ref}}; r_{ij}) \right] r_{ij}^3 \partial r_{ij} \quad (20) \end{aligned}$$

Equations 19 and 20 are exact. Unfortunately, while  $g_{ij}(\theta_{\text{ref}}; r_{ij})$  is obtained from direct simulation of  $\theta_{\text{ref}}$ , the PCF for the non-simulated force field ( $g_{ij}(\theta; r_{ij})$ ) is unknown *a priori*. PCFR attempts to estimate  $g_{ij}(\theta; r_{ij})$  and, thereby, predict the properties  $U^{\text{dep}}$  and  $Z$  for different non-bonded interactions.

There are a number of possible approximations that could be made. The approach we recommend (and use in Section 3) is to perform a density expansion for the PCF:

$$g_{ij}(\theta; r, \rho, T) = g_{0,ij}(\theta; r, T) + \rho g_{1,ij}(\theta; r, T) + \rho^2 g_{2,ij}(\theta; r, T) + \dots \quad (21)$$

where the zeroth order term ( $g_0$ ) is known analytically to be:

$$g_{0,ij}(\theta; r_{ij}) \equiv \exp\left(\frac{-u_{ij}(\theta; r_{ij})}{T}\right) \quad (22)$$

By substituting Equation 21 into Equations 17-18 and separating the integration by each term of  $g_{ij}$ , we can express  $U^{\text{vdw}}$  and  $Z^{\text{vdw}}$  as a virial expansion:

$$U^{\text{vdw}}(\theta) = \sum_{h=0}^{\infty} \rho^h U_h^{\text{vdw}} \quad (23)$$

$$Z^{\text{vdw}}(\theta) = \sum_{h=0}^{\infty} \rho^h Z_h^{\text{vdw}} \quad (24)$$

where  $U_h^{\text{vdw}}$  and  $Z_h^{\text{vdw}}$  are obtained, in principle, by integrating Equations 17-18 after substituting  $g_{h,ij}$  for  $g_{ij}$ . By assuming that the higher order ( $h > 0$ ) contributions to  $U^{\text{vdw}}$  and  $Z^{\text{vdw}}$  for  $\theta$  and  $\theta_{\text{ref}}$  are equal (or at least negligibly different)  $U^{\text{dep}}(\theta)$  and  $Z(\theta)$  can be



estimated using:

$$\begin{aligned} \langle U^{\text{dep}}(\theta) \rangle &\approx \langle U^{\text{dep}}(\theta_{\text{ref}}) \rangle \\ &+ 2\pi\rho R_g \sum_{i=1}^{N_S-1} \sum_{j>i}^{N_S} \int_0^\infty [u_{ij}^{\text{vdw}}(\theta; r_{ij})g_{0,ij}(\theta; r_{ij}) - u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})g_{0,ij}(\theta_{\text{ref}}; r_{ij})] r_{ij}^2 \partial r_{ij} \quad (25) \end{aligned}$$

$$\begin{aligned} \langle Z(\theta) \rangle &\approx \langle Z(\theta_{\text{ref}}) \rangle \\ &- \frac{2\pi\rho}{3T} \sum_{i=1}^{N_S-1} \sum_{j>i}^{N_S} \int_0^\infty \left[ \frac{\partial u_{ij}^{\text{vdw}}(\theta; r_{ij})}{\partial r_{ij}} g_{0,ij}(\theta; r_{ij}) - \frac{\partial u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})}{\partial r_{ij}} g_{0,ij}(\theta_{\text{ref}}; r_{ij}) \right] r_{ij}^3 \partial r_{ij} \quad (26) \end{aligned}$$

The primary advantage of this assumption is that, because  $g_{0,ij}(\theta; r_{ij})$  is calculated with Equation 22, the integrals in Equations 25-26 can be evaluated without performing any additional simulations. However, the PCFR results presented in Section 3 suggest that Equations 25-26 adequately approximate  $U^{\text{dep}}$  but not  $Z$ , i.e.  $U_h^{\text{vdw}}(\theta) \approx U_h^{\text{vdw}}(\theta_{\text{ref}})$  while  $Z_h^{\text{vdw}}(\theta) \not\approx Z_h^{\text{vdw}}(\theta_{\text{ref}})$  for  $h > 0$  when  $\theta \not\approx \theta_{\text{ref}}$ .

PCFR is orders of magnitude faster than MBAR since PCFR only requires a numerical integration of Equations 25-26 while MBAR requires energy and force “rerun” calculations. Although “rerun” calculations are typically orders of magnitude faster than direct simulation, MBAR may still be too costly if an optimization or uncertainty quantification (UQ) method requires on the order of  $10^4$  to  $10^6$  MBAR evaluations. However, the computational cost of MBAR is approximately the same as PCFR when basis functions are implemented.<sup>48</sup> Basis functions can be constructed if a linear relationship exists between the non-bonded parameters and the non-bonded energies and forces. For example, basis functions are amenable to Mie  $\lambda$ -6 potentials because the energies are linearly dependent on  $r^{-\lambda}$  and  $r^{-6}$  while the forces are linearly dependent on  $r^{(-\lambda-1)}$  and  $r^{-7}$  (see Section SI.IV in the Supporting Information). By contrast, basis functions are not compatible with

the Exponential-6 function due to the non-linearity of the exponential term.

Alternative PCFR approaches that we investigated, but found less reliable than Equations 25-26, are included in the Appendix. One such method is the “constant PCF” approach. “Constant PCF” assumes that the pair correlation function for the reference force field is equal to the PCF for the non-simulated force field. As demonstrated in the Appendix, this is mathematically identical to assuming the weights of each configuration are equal in Equations 12-13. Therefore, a comparison between PCFR and MBAR with “constant PCF” (or “equal weights”) quantifies the improvement due to rescaling and reweighting configurations, respectively. For this reason, we include the “constant PCF” results in Section 3 to provide a common basis for MBAR and the recommended PCFR approach (Equations 25-26). Furthermore, “constant PCF” can provide valuable insight due to its conceptual and mathematical simplicity.

### 3 Results

In this section we compare MBAR and PCFR (Equations 25-26) based on their ability to predict the thermodynamic properties  $U^{\text{dep}}$ ,  $Z$ ,  $\rho_1^{\text{sat}}$ , and  $P_v^{\text{sat}}$ . First, we compare the predicted  $U^{\text{dep}}$  and  $Z$  values with direct molecular simulation results. Subsequently, we combine ITIC with MBAR and PCFR and the predicted values for  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  are compared with direct molecular simulation results.

The systems considered are the united-atom representations of ethane ( $\text{C}_2\text{H}_6$ ), hexafluoroethane ( $\text{C}_2\text{F}_6$ ), propane ( $\text{C}_3\text{H}_8$ ), *n*-butane ( $\text{C}_4\text{H}_{10}$ ), and *n*-octane ( $\text{C}_8\text{H}_{18}$ ). A detailed comparison of the surrogate models is only presented for ethane. In the case of ethane, we perform two types of test for MBAR and PCFR.

For the first test, the non-simulated force field uses the same non-bonded function (i.e. the LJ 12-6) as the reference force field. We refer to this test as the “constant model”

test. For the second test, the non-simulated force field uses a different non-bonded function (i.e. the Mie  $\lambda$ -6 or Exp-6) than the reference force field. We refer to this test as the “perturbed model” test. Since this study focuses on the Mie  $\lambda$ -6 potential, the “perturbed model” test corresponds to varying the value of  $\lambda$ . For both tests, we compare how well MBAR performs with only a single reference and with multiple references.

The discussion for hexafluoroethane and the longer  $n$ -alkanes in Section 3.2 is limited but complements the ethane discussion in Section 3.1. Specifically, Section 3.2 compares how well MBAR and PCFR predict the Potoff  $\lambda$ -6 potential when the TraPPE-UA LJ 12-6 potential is used as a single reference force field. Hexafluoroethane is included since it provides an example for the “perturbed model” test where  $\lambda \gg 12$ . The longer  $n$ -alkanes are included to demonstrate that the results are not specific to two-site molecules.

### 3.1 Ethane

In this section, we compare the accuracy of MBAR and PCFR when predicting  $U^{\text{dep}}$ ,  $Z$ ,  $\rho_1^{\text{sat}}$ , and  $P_v^{\text{sat}}$  for ethane. First, in Section 3.1.1 we investigate how well these surrogate models predict the quantities of interest when there is only a single reference force field, namely, the TraPPE-UA LJ 12-6 model. Section 3.1.1 allows for a fair comparison between PCFR and MBAR since PCFR is not compatible with multiple references. More importantly, our initial goal is to determine if the optimal force field parameters can be obtained by performing direct simulations with only a single reference force field. This would enable rapid reparameterization of the non-bonded interactions for any existing force field without performing additional molecular simulations. Unfortunately, the results in section 3.1.1 demonstrate that this ideal scenario is not obtainable. For this reason, in Section 3.1.2 we investigate the improvement that is possible for MBAR when using multiple references. Then, in Section 3.1.3 we demonstrate how well MBAR and PCFR perform when used in connection with ITIC.

### 3.1.1 Single Reference

Figure 1 presents the results from the “constant model” single reference test. Specifically, we use the TraPPE-UA LJ 12-6 model for ethane as our reference force field from which we sample configurations. We predict the departure internal energy ( $U^{\text{dep}}$ ) and compressibility factor ( $Z$ ) for a wide range of  $\text{CH}_3$  LJ parameters ( $88 < \epsilon/\text{K} < 108, 0.365 < \sigma/\text{nm} < 0.385$ ) using the various configuration-sampling-based surrogate models (MBAR, PCFR, and “constant PCF”). Note that the reference  $\epsilon$ ,  $\sigma$ , and  $\lambda$  are at the center of the investigated parameter space, i.e.  $\epsilon_{\text{ref}} = 98 \text{ K}$ ,  $\sigma_{\text{ref}} = 0.375 \text{ nm}$ , and  $\lambda_{\text{ref}} = 12$ . Figure 1 Panels a)-b) and c)-d) compare the direct simulation results for the departure internal energy ( $U^{\text{dep}} \equiv U - U^{\text{ig}}$ ) and compressibility factor ( $Z$ ), respectively, with those predicted from the various surrogate models. Figure 1 includes parity plots as well as embedded deviation plots. Panels a) and c) compare each surrogate model while Panels b) and d) present the results only for MBAR.

Note that the MBAR and PCFR deviations are smaller than the “constant PCF” deviations. This shows the improvement due to reweighting (MBAR) or “rescaling” (PCFR) the reference configurations. Panel a) demonstrates that PCFR is superior to MBAR for estimating  $U^{\text{dep}}$ , while Panel c) shows that neither method is particularly robust for estimating  $Z$ . However, Panels b) and d) demonstrate that with a sufficient number of effective samples ( $N_{\text{eff}}$ ) the MBAR estimates for  $U^{\text{dep}}$  and  $Z$  are typically reliable. For example, note that in Panels a) and c) the MBAR estimates follow the parity line very closely for larger values of  $N_{\text{eff}}$ . This makes intuitive sense because  $N_{\text{eff}}$  is intimately related to the degree of “overlap”, i.e. the probability that configurations sampled by the reference force field would be sampled by the non-simulated force field. To quantify this observation, the insets in Panels b) and d) suggest that for  $N_{\text{eff}} > 50$  (or  $\log_{10}(N_{\text{eff}}) > 1.7$ ) the percent deviation in  $U^{\text{dep}}$  and the deviation in  $Z$  are less than 1 % and 0.3, respectively.

Figure 2 presents the “perturbed model” single reference test results for ethane. Fig-

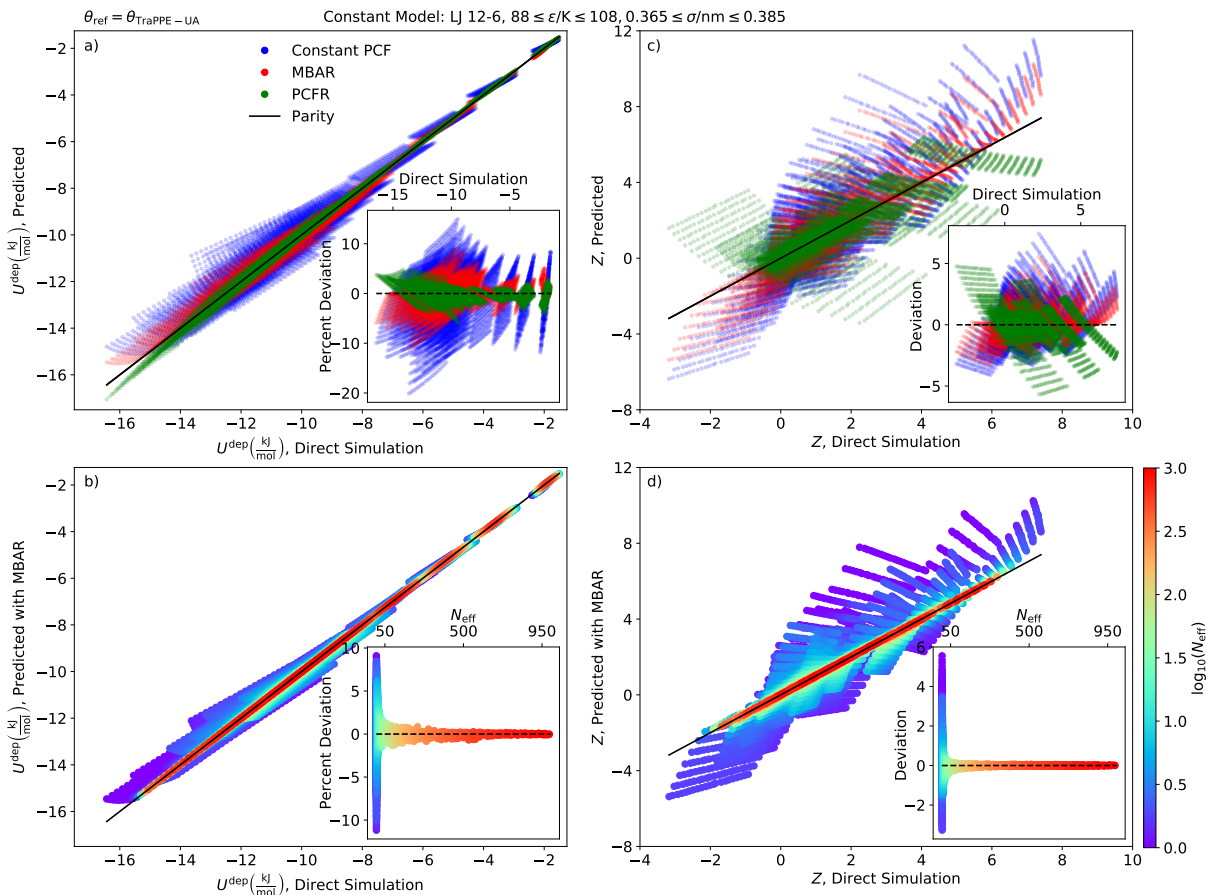


Figure 1: PCFR performs well for  $U^{\text{dep}}$  but not for  $Z$  while MBAR accurately predicts both properties when  $N_{\text{eff}} > 50$  for “constant model” test of ethane. The parity plots compare direct simulation values with those predicted using a constant PCF (equal weights), MBAR, and PCFR for  $\lambda_{\text{ref}} = \lambda = 12$ . Panels a) and b) correspond to departure internal energy ( $U^{\text{dep}}$ ) while Panels c) and d) correspond to compressibility factor ( $Z$ ). Deviation plots are embedded, where percent deviation and deviation are plotted for  $U^{\text{dep}}$  and  $Z$ , respectively. Panels b) and d) focus on the MBAR results where the color map corresponds to the number of effective samples (on a  $\log_{10}$  scale).

Figure 2 is analogous to Figure 1, where Figures 1 and 2 present the “constant model” (i.e.  $\lambda_{\text{ref}} = \lambda$ ) and “perturbed model” (i.e.  $\lambda_{\text{ref}} \neq \lambda = 16$ ) test results, respectively. In both tests we use the TraPPE-UA LJ 12-6 model as our reference force field from which we sample configurations. For the “perturbed model” test, the parameter space over which the comparison is performed is shifted to higher values of  $\epsilon$  ( $108 < \epsilon/K < 128$ ) since Potoff and

Bernard-Brunel<sup>1</sup> demonstrated that for ethane the optimal  $\epsilon_{\text{CH}_3}$  increases with increasing  $\lambda_{\text{CH}_3}$ . (Note that the Potoff CH<sub>3</sub> parameter set of  $\epsilon = 121.25$  K,  $\sigma = 0.3783$  nm, and  $\lambda = 16$  is near the center of the investigated parameter space.) We again compare the surrogate model estimates of  $U^{\text{dep}}$  and  $Z$  with those obtained by performing direct simulations with each Mie 16-6 parameter set.

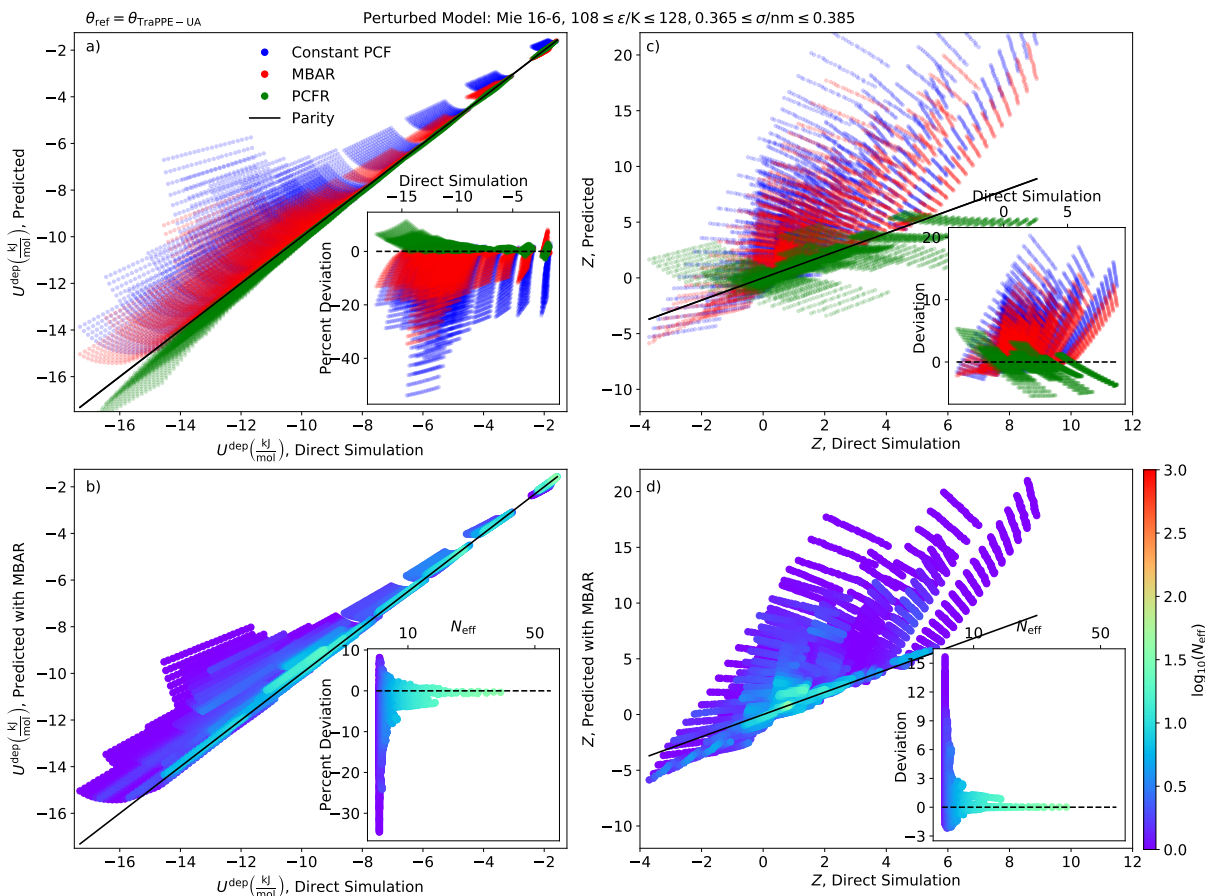


Figure 2: Similar to the “constant model” results presented in Figure 1, PCFR accurately predicts  $U^{\text{dep}}$  but not  $Z$  for the “perturbed model” test of ethane. MBAR performs poorly because very few state points have  $N_{\text{eff}} > 50$  for  $\lambda_{\text{ref}} \neq \lambda = 16$ . The panels and symbols are the same as Figure 1.

Panel a) of Figure 2 demonstrates that, similar to the LJ 12-6 results, the PCFR method is superior to MBAR at predicting  $U^{\text{dep}}$ , while Panel c) shows that neither PCFR or MBAR is particularly accurate at predicting  $Z$ . Panels b) and d) help explain the poor perfor-

mance of MBAR. Note that the maximum number of effective samples for this entire parameter space is less than 100, while for most systems  $N_{\text{eff}} \ll 50$ . This poor “overlap” causes a single (slightly more favorable) configuration to have a weight equal to one while all other configurations have a weight of zero. When  $N_{\text{eff}} \approx 1$  typically even the non-zero weight configuration is still very unfavorable.

The “constant PCF” results in Figure 2 Panels a) and c) provide further insight into why MBAR performs poorly when  $\lambda_{\text{ref}} \neq \lambda = 16$ . Note the strong positive bias relative to the parity line for  $U^{\text{dep}}$  and  $Z$  “constant PCF” and MBAR. This is because the “softer” LJ 12-6 potential samples close-range configurations that result in extremely high energies and forces when recomputed with the “harder” Mie 16-6 potential. Therefore, MBAR reweighting cannot produce accurate estimates because few (if any) of the reference configurations represent a reasonable state that would be sampled from the Mie 16-6 potential. By contrast, PCFR rescales these close-range interactions to avoid unreasonable energies and forces. This is an important advantage of PCFR when  $\lambda_{\text{ref}} \neq \lambda$ .

The data depicted in Figure 3 help quantify over what range of the  $\epsilon$ ,  $\sigma$ , and  $\lambda$  parameter space MBAR is reliable (i.e.  $N_{\text{eff}} > 50$ ) when using a single reference parameter set. The different color contours correspond to the “constant model” ( $\lambda = 12$ ) and the “perturbed model” ( $\lambda = 13$  to 18) tests. The contours in Figure 3 represent the average number of effective samples ( $\bar{N}_{\text{eff}}$ ), i.e. the mean  $N_{\text{eff}}$  for the 19 different ITIC state points. Multiple contours are included for  $\lambda = 12$  while only a single contour is provided for  $\lambda = 13$  to 18 for visual clarity.

Figure 3 demonstrates that  $N_{\text{eff}}$  (and, thus, the “overlap” between force fields) depends strongly on  $\sigma$  and the repulsive exponent ( $\lambda$ ) with much less dependence on  $\epsilon$ . Specifically, the average  $N_{\text{eff}}$  is greater than 50 for  $\sigma = \sigma_{\text{ref}} = 0.375$  nm,  $\lambda = \lambda_{\text{ref}} = 12$ , and  $88 < \epsilon/\text{K} < 113$ , covering a range of  $\pm 15\%$   $\epsilon_{\text{ref}}$ . By contrast,  $\bar{N}_{\text{eff}} > 50$  for  $\epsilon = \epsilon_{\text{ref}} = 98$  K,  $\lambda = \lambda_{\text{ref}} = 12$ , and  $0.3675 < \sigma/\text{nm} < 0.3725$ , a range of only  $\pm 0.7\%$   $\sigma_{\text{ref}}$ . In addition, the

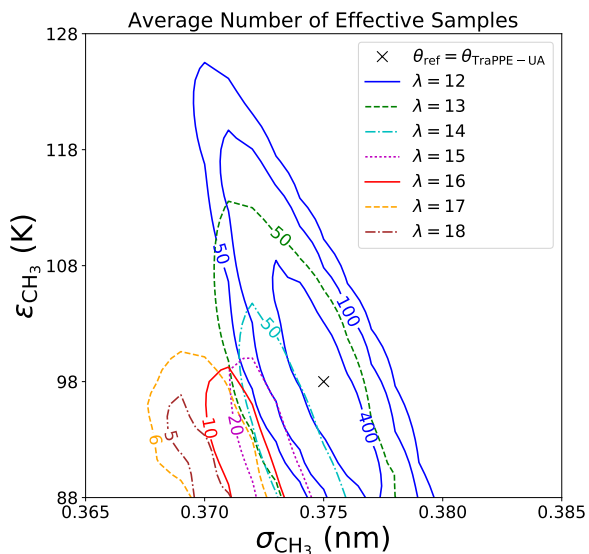


Figure 3: Average number of effective samples depends strongly on  $\sigma$  and  $\lambda$  for ethane. Specifically, large differences in  $\sigma$  from the reference lead to  $N_{\text{eff}} < 50$  while the entire  $\lambda = 15$  to 18 parameter space has  $N_{\text{eff}} \ll 50$ . Contours represent the average  $N_{\text{eff}}$  for  $\lambda = 12$  to 18. A single reference parameter set was used (TraPPE-UA, i.e.  $\lambda_{\text{ref}} = 12$ ).

$\bar{N}_{\text{eff}} > 50$  region for  $\lambda = 13, 14$  is smaller than for  $\lambda = 12$  and is shifted to  $\sigma < \sigma_{\text{ref}}$  and  $\epsilon < \epsilon_{\text{ref}}$ . This is especially troubling, since several studies have demonstrated that with increasing  $\lambda$  the optimal  $\epsilon$  increases significantly and the optimal  $\sigma$  rarely decreases.<sup>1,29,30</sup> Therefore, the  $\bar{N}_{\text{eff}} > 50$  regions for  $\lambda = 13, 14$  are not particularly useful as they do not overlap with the expected optimal  $\epsilon$  and  $\sigma$ . Even more concerning,  $\bar{N}_{\text{eff}} \ll 50$  for the entire joint  $\epsilon$  and  $\sigma$  parameter space for  $\lambda = 15$  to 18. For example,  $\bar{N}_{\text{eff}} \approx 1$  for the Potoff model ( $\epsilon = 121.25$  K,  $\sigma = 0.3783$  nm,  $\lambda = 16$ ). These results suggest that the TraPPE-UA LJ 12-6 configurations cannot be reweighted to optimize  $\epsilon$  and  $\sigma$  for  $\lambda > 12$ .

These observations make some intuitive sense as the configurations sampled from  $\sigma = 0.375$  nm should be significantly different than those sampled from  $\sigma \gg 0.375$  nm or  $\sigma \ll 0.375$  nm. Likewise, larger values of  $\lambda$  have a steeper, more repulsive barrier which increases the “effective hard-sphere diameter” and significantly reduces the probability of close-range interactions. These results are convincing evidence that MBAR requires



more than a single reference for large changes in  $\sigma$  (greater than 0.0025 nm) and when  $\lambda \neq \lambda_{\text{ref}}$ .

### 3.1.2 Multiple References

MBAR is intrinsically designed for multiple reference force fields, which is one of its primary advantages. With additional (and judiciously chosen) references, the reliability of MBAR estimates improves dramatically, as demonstrated in the following discussion. By contrast, PCFR has no obvious way to combine the prediction of multiple references in a meaningful manner. We attempted to incorporate multiple references into PCFR by performing a weighted average of independent PCFR estimates where the weights depend on the difference between  $\sigma$  and  $\sigma_{\text{ref}}$ . However, we did not find this approach to be beneficial. Therefore, we do not discuss PCFR in this section and focus solely on MBAR.

The results presented in Section 3.1.1 demonstrate that MBAR is limited in how far it can extrapolate in parameter space. As discussed previously, MBAR does not provide accurate  $U^{\text{dep}}$  and  $Z$  estimates when  $N_{\text{eff}} \ll 50$ , which is typically the case when  $\sigma$  and/or  $\lambda$  are significantly different from the  $\sigma_{\text{ref}}$  and  $\lambda_{\text{ref}}$ . One obvious solution to increase  $N_{\text{eff}}$  and, thereby, improve MBAR is to sample additional configurations from multiple references. These references should be chosen to cover a wide range of parameter space while using as few as possible to limit the increase in the computational cost of additional molecular simulations.

Figure 4 presents the results when multiple reference force fields are utilized with MBAR. Specifically, we performed direct simulations for nine equally spaced values of  $\sigma$  from 0.365-0.385 nm for the “constant model”  $\lambda = 12$  with  $\epsilon = 98$  K. The spacing corresponds to one reference  $\sigma$  every 0.0025 nm since we observed that  $N_{\text{eff}} > 50$  when  $\sigma = \sigma_{\text{ref}} \pm 0.0025$  nm (recall Figure 1). For the “perturbed model”  $\lambda = 16$  test, we use two

different  $\epsilon$  values of 98 and 118 K. We use the same range of  $\sigma$  as the “constant model” test for  $\epsilon_{\text{ref}} = 98$  K to provide a fair comparison between the two tests for multiple references. For  $\epsilon_{\text{ref}} = 118$  K (the center value of the Mie 16-6 parameter space), we modify the  $\sigma$  range slightly by including three additional larger  $\sigma_{\text{ref}}$  values (0.3875, 0.3900, and 0.3925 nm). Recall that in Figure 3 the maximum  $N_{\text{eff}}$  for  $\lambda = 16$  was observed when  $\sigma < \sigma_{\text{ref}}$ . Thus, the expanded  $\sigma_{\text{ref}}$  range with  $\epsilon_{\text{ref}} = 118$  K results help elucidate whether MBAR can adequately predict the Mie 16-6 parameter space if the LJ 12-6 references are chosen judiciously. Panels a) and d), b) and e), and c) and f) provide the results for the “constant model”, “perturbed model” with  $\epsilon_{\text{ref}} = 98$  K, and “perturbed model” with  $\epsilon_{\text{ref}} = 118$  K, respectively.

For the “constant model” test (Panels a) and d)), due to the increased number of references most of the LJ 12-6 parameter sets have good overlap and, thus,  $N_{\text{eff}} > 50$ . However, it should be noted that there are a few exceptions where  $N_{\text{eff}} < 50$  in Panels a) and d). In each case, these correspond to high density, saturated liquids with  $\sigma = 0.385$  nm and  $\epsilon > 107$  K. By comparing Figure 4 Panels a) and d) with Figure 1 Panels b) and d), respectively, additional  $\sigma_{\text{ref}}$  values significantly improve the MBAR “constant model” results.

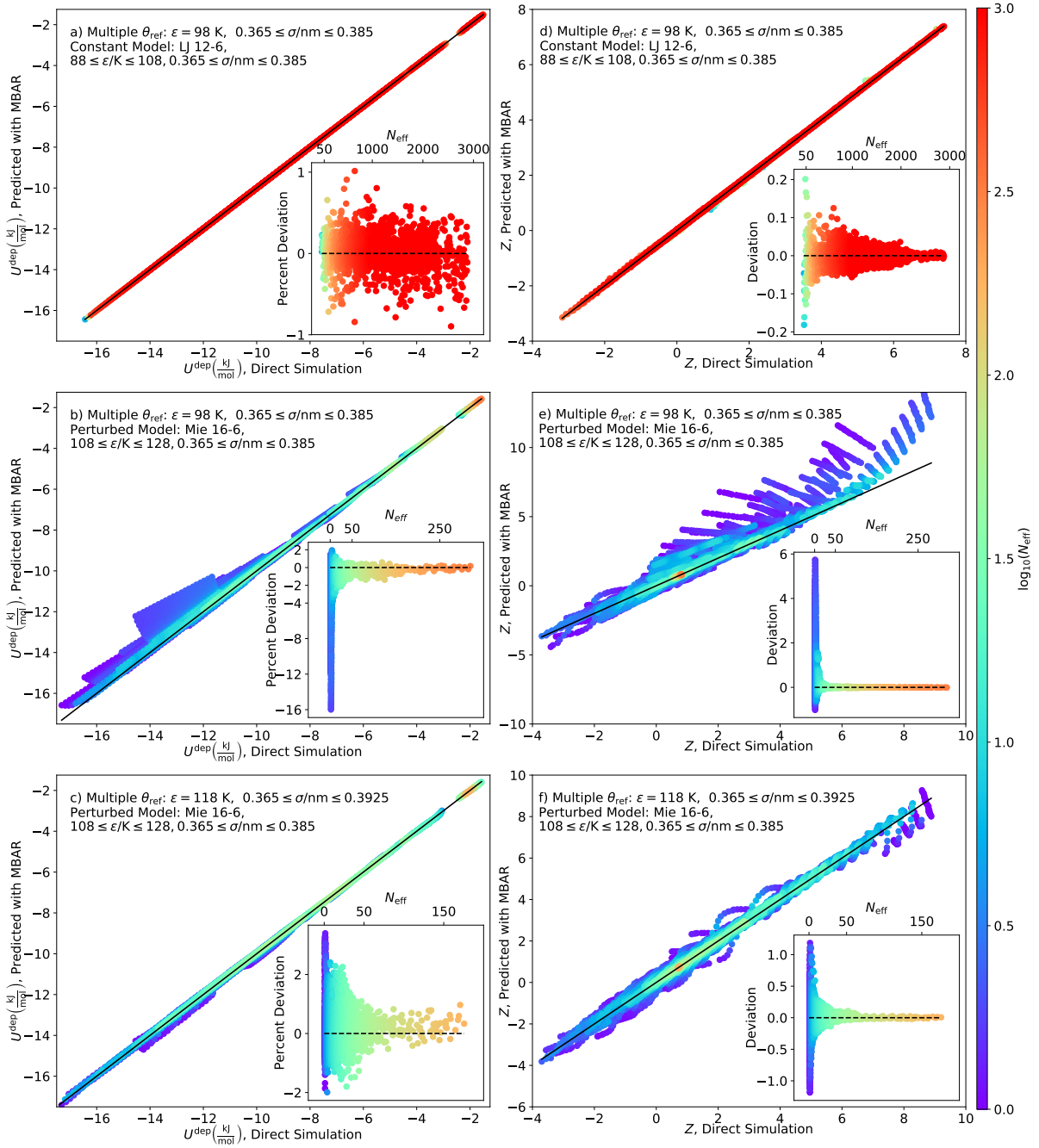


Figure 4: Multiple  $\sigma_{\text{ref}}$  values significantly improve the “constant model” (i.e.  $\lambda_{\text{ref}} = \lambda = 12$ ) results for ethane (see Panels a) and d)). The “perturbed model” (i.e.  $\lambda_{\text{ref}} \neq \lambda = 16$ ) results are only moderately improved when  $\epsilon_{\text{ref}} = 98$  K (see Panels b) and e)) while  $\epsilon_{\text{ref}} = 118$  K provides substantial improvement (see Panels c) and f)), although large deviations still exist at several state points where  $N_{\text{eff}} \ll 50$ . Panels a)-c) and d)-f) correspond to  $U^{\text{dep}}$  and  $Z$ , respectively. The parity plots, embedded deviation plots, and color map are the same as Figure 1 Panels b) and d).

By contrast, MBAR still performs quite poorly for the “perturbed model” (Mie 16-6) potential despite the increased number of references (see Figure 4 Panels b), c), e) and f)). In Panels b) and e) this is somewhat expected since  $\epsilon_{\text{ref}} = 98$  K, while the non-simulated  $\epsilon$  values range from 108-128 K. As seen in Panels c) and f), one way to improve the performance of MBAR for the perturbed model is to use  $\epsilon_{\text{ref}} = 118$  K. However, notice that  $N_{\text{eff}}$  in Panels c) and f) are still much less than those for the constant model results (Panels a) and d)). Therefore, sampling directly from  $\lambda_{\text{ref}} = 16$  with multiple  $\sigma_{\text{ref}}$  (and a reasonable value of  $\epsilon_{\text{ref}}$ ) is the only way to ensure that  $N_{\text{eff}} > 50$ . In Section 4 we propose how PCFR can be used to assist in choosing  $\epsilon_{\text{ref}}$  and  $\sigma_{\text{ref}}$  for a given  $\lambda_{\text{ref}}$ .

### 3.1.3 Vapor-Liquid Equilibria

The comparison between MBAR and PCFR in Sections 3.1.1-3.1.2 has focused on  $U^{\text{dep}}$  and  $Z$ . However, for the purpose of force field optimization, the properties of interest are most likely  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  ( $U^{\text{dep}}$  could also be of interest as a substitute for enthalpy of vaporization and/or heat capacity). For this reason, in this section we compare how well MBAR and PCFR predict saturation properties when used in conjunction with ITIC.

Figure 5 plots the root-mean-square (RMS) with respect to REFPROP values for the “constant model” test. Panels a) and b) plot the RMS for  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$ , respectively.  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  are computed with ITIC using the  $U^{\text{dep}}$  and  $Z$  values obtained from either direct simulation or the surrogate models. Both the single and multiple reference results are included for MBAR. For visual clarity, only the first two contours are included for the single reference ( $\theta_{\text{ref}} = \theta_{\text{TraPPE}}$ ) results of MBAR and PCFR.

The primary conclusion from Figure 5 is that the “MBAR with multiple references” RMS contours for  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$  are nearly identical to those obtained from direct simulation. By contrast, “MBAR with a single reference” shows good agreement for RMS  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$  for only a small range of  $\sigma$  values. PCFR has the wrong shape for

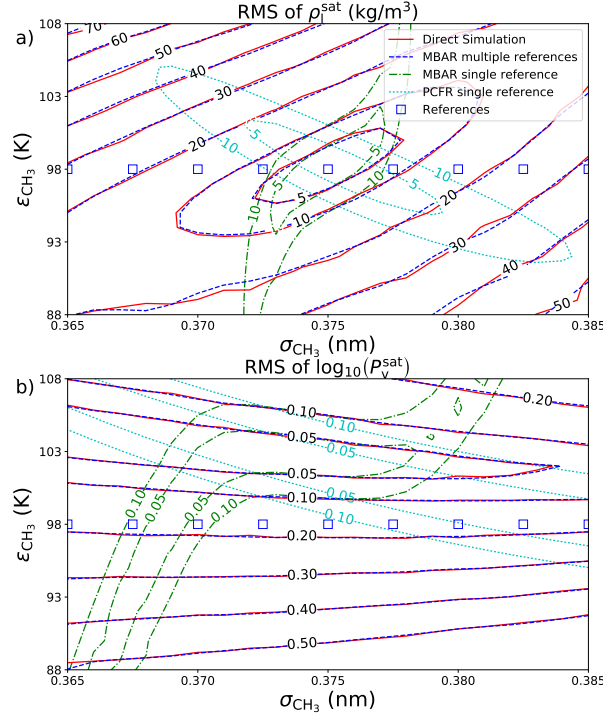


Figure 5: MBAR with multiple references predicts ethane  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  more accurately than PCFR and MBAR with a single reference for  $\lambda_{\text{ref}} = \lambda = 12$ . Contours of the root-mean-square (RMS) relative to REFPROP values are plotted for each method and compared with direct simulation. Panels a)-b) show RMS for  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$ , respectively. Note that the “single reference” is at the center of the plot, i.e.  $\epsilon = 98$  K,  $\sigma = 0.375$  nm. Direct simulations were performed on a 21x21 grid equally spaced between 88-108 K and 0.365-0.385 nm.

the RMS  $\rho_1^{\text{sat}}$  contours while the RMS  $\log_{10}(P_v^{\text{sat}})$  contours for PCFR are in satisfactory agreement. The incorrect shape of the RMS  $\rho_1^{\text{sat}}$  contours for PCFR is especially concerning for Bayesian and other UQ methods that depend on having the correct local behavior near the optimum.

Figure 6 plots the “perturbed model” test results, i.e.  $\lambda_{\text{ref}} = 12$  while  $\lambda = 16$ , using the same format as the “constant model” test results in Figure 5. The references are not depicted in Figure 6 because  $\lambda_{\text{ref}} \neq \lambda = 16$ . Only the “MBAR multiple references” results are presented for  $\epsilon_{\text{ref}} = 118$  K since  $\epsilon_{\text{ref}} = 98$  K with multiple  $\sigma_{\text{ref}}$  values showed very little improvement over a single reference when  $\lambda_{\text{ref}} = 12$ .

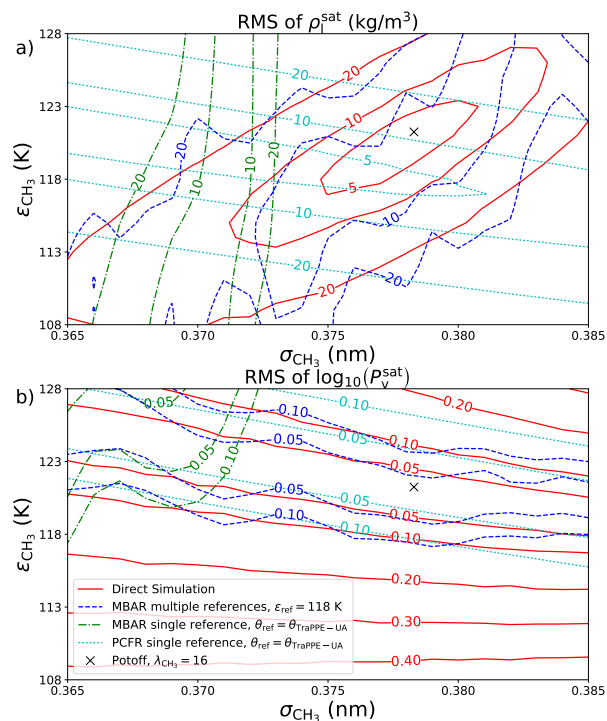


Figure 6: MBAR with multiple references provides reasonable (although quite noisy) estimates for ethane  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$  for  $\lambda_{\text{ref}} \neq \lambda = 16$ . Similar to the “constant model” results presented in Figure 5, PCFR produces an incorrect trend for  $\rho_1^{\text{sat}}$  while the  $\log_{10}(P_v^{\text{sat}})$  contours are reasonably accurate and follow the correct trend. MBAR with a single reference produces inaccurate estimates of  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$ . Contours of the root-mean-square (RMS) relative to REFPROP values are plotted for each method and compared with direct simulation. Panels a)-b) show RMS for  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$ , respectively. The “multiple references” are the same as Figure 4 Panels c) and f), namely,  $\epsilon_{\text{ref}} = 118$  K,  $\lambda_{\text{ref}} = 12$ , with 11  $\sigma_{\text{ref}}$  values evenly spaced between 0.365-0.3925 nm. The “single reference” is the TraPPE-UA force field ( $\lambda_{\text{ref}} = 12$ ). The Potoff parameters are included only as a visual reference. Direct simulations were performed on a 21x21 grid equally spaced between 108-128 K and 0.365-0.385 nm with  $\lambda = 16$ .

Figure 6 demonstrates that the  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$  RMS contours for “MBAR multiple references” are similar to those from direct simulation while the “MBAR single reference” contours are completely unreliable for the “perturbed model” test. However, the “MBAR multiple references” contours are extremely noisy, which would render parameterization quite difficult. More importantly, the moderate improvement (compared to “MBAR single reference”) was primarily achieved by using  $\epsilon_{\text{ref}} = 118$  K instead of the TraPPE-UA

value of 98 K. Unfortunately, the optimal  $\epsilon_{\text{ref}}$  value for a different  $\lambda_{\text{ref}}$  is not known *a priori*. For these reasons, regardless of what values are used for  $\epsilon_{\text{ref}}$  and  $\sigma_{\text{ref}}$ , sampling multiple LJ 12-6 references is not a recommended approach for predicting  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  when  $\lambda \neq \lambda_{\text{ref}}$ .

Although the reliability of MBAR is greatly diminished when  $\lambda \neq \lambda_{\text{ref}}$ , the performance of PCFR is similar for both the “perturbed model” and the “constant model” tests. For example, the PCFR RMS contours in Figure 6 are very similar to the PCFR RMS contours in Figure 5. Specifically, although the  $\rho_1^{\text{sat}}$  RMS contours have an incorrect trend with respect to  $\epsilon$  and  $\sigma$ , the  $\log_{10}(P_v^{\text{sat}})$  RMS contours are in close agreement with direct simulation, especially when  $\sigma \approx \sigma_{\text{ref}}$ . Furthermore, the PCFR contours are smooth (i.e. not noisy) with respect to  $\epsilon$  and  $\sigma$ . This suggests that PCFR can be a useful tool for parameterizing different Mie  $\lambda$ -6 potentials from a single LJ 12-6 reference. In fact, one of our key recommendations provided in Section 4 is that the RMS of  $\log_{10}(P_v^{\text{sat}})$  (or, alternatively,  $U^{\text{dep}}$ ) predicted by PCFR be used as an initial objective function when perturbing the value of  $\lambda$ .

The poor prediction of  $\rho_1^{\text{sat}}$  for the “perturbed model” is expected since ITIC  $\rho_1^{\text{sat}}$  depends primarily on the isochore  $Z$  values (see discussion in Section 2.2). Thus, neither “MBAR single reference” or “PCFR single reference” provide reliable  $\rho_1^{\text{sat}}$  contours because neither method accurately predicts  $Z$  for the “perturbed model” and single reference test (see Figure 2, Panels c-d)). The large fluctuations in the “MBAR multiple references” contours are likely the result of the inherent randomness in  $Z$  when  $N_{\text{eff}} \approx 1$  (see Figure 4 Panel f)). The near-saturation, isochore state points that directly impact  $T^{\text{sat}}$  are more likely to have  $N_{\text{eff}} \approx 1$  because  $N_{\text{eff}}$  decreases with increasing density and decreasing temperature.

Elucidating the reason why “MBAR multiple references” and “PCFR single reference” provide reasonable estimates of  $P_v^{\text{sat}}$  for the “perturbed model” is more complicated, because ITIC  $P_v^{\text{sat}}$  depends strongly on both  $Z$  and  $U^{\text{dep}}$ . This is due to the exponential

relationship between  $A^{\text{dep}}$  and  $P_v^{\text{sat}}$  (see Equations 6-7) where  $A^{\text{dep}}$  is computed from the isotherm  $Z$  and isochore  $U^{\text{dep}}$  values (see Equation 5). In addition, the isochore  $Z$  values can have a significant impact on  $P_v^{\text{sat}}$  through  $T^{\text{sat}}$  in Equations 5-7. However, because “PCFR single reference” does not provide accurate estimates of  $Z$ , the accuracy of  $P_v^{\text{sat}}$  appears to depend primarily on the reliability of  $U^{\text{dep}}$  (see Figure 2 Panels a) and b)). Since “MBAR multiple references” predicts  $U_{\text{dep}}$  to within 1-2 %, the noise in the corresponding  $\log_{10}(P_v^{\text{sat}})$  contours can likely be attributed to large (random) deviations for  $Z$  at certain state points (see Figure 4 Panels c) and f)). By contrast, the smooth “PCFR single reference”  $\log_{10}(P_v^{\text{sat}})$  contours suggest that the “PCFR single reference” deviations in  $Z$  are more systematic (i.e. less random) than those from “MBAR multiple references.”

### 3.2 Additional Compounds

In this section we perform the “perturbed model” and single reference test for additional compounds. Specifically, we verify that MBAR struggles to predict  $U^{\text{dep}}$ , while PCFR can provide reasonable estimates of  $U^{\text{dep}}$  when  $\lambda > \lambda_{\text{ref}}$ . The  $Z$  results are not included because we have already concluded that neither MBAR or PCFR is capable of predicting  $Z$  in this case. The “perturbed model” test is performed by sampling from the TraPPE-UA LJ 12-6 model ( $\theta_{\text{ref}} = \theta_{\text{TraPPE}}$ ) while the non-simulated force field is the Potoff Mie  $\lambda$ -6 model ( $\theta = \theta_{\text{Potoff}}$ ).

The reason to emphasize the “perturbed model” results is because a plethora of LJ 12-6 parameters exist in the literature. Improved force field accuracy likely necessitates a systematic conversion of these two-parameter LJ models to a model with additional parameters (i.e. the three-parameter Mie  $\lambda$ -6, Exp-6, extended Lennard-Jones,<sup>35</sup> etc.). There are at least two primary reasons to focus on the single reference test. First, the number of direct simulations is minimized by using a single reference. Second, our goal is to verify the reliability of PCFR, which is designed only for a single reference.



The additional compounds considered in this section are hexafluoroethane, propane, *n*-butane, and *n*-octane. Hexafluoroethane (C<sub>2</sub>F<sub>6</sub>) provides an extreme test case where the perturbed repulsive barrier is significantly different from the LJ 12-6. Specifically, the Potoff model utilizes a Mie 36-6 potential for CF<sub>3</sub> sites. The *n*-alkanes provide insight into how the surrogate models scale with additional sites (note that  $\lambda = 16$  for the Potoff CH<sub>3</sub>, CH<sub>2</sub>, and cross-interactions). The parameters for each force field and the simulation conditions for each compound are provided in Sections SI.I.1-SI.I.3 of the Supporting Information.

Figure 7 Panels a) and b) provide parity plots of  $U^{\text{dep}}$  for hexafluoroethane and the *n*-alkanes, respectively. MBAR and PCFR estimates are compared with direct simulation results. Panel a) also includes “constant PCF” for comparison. Percent deviation plots are included as insets for PCFR in Panel a) and for MBAR and PCFR in Panel b).

Panel a) demonstrates that MBAR (and “constant PCF”) are completely inadequate for providing reasonable estimates of  $U^{\text{dep}}$  for hexafluoroethane. There are two primary reasons for the poor performance of MBAR. First, the  $\epsilon$  value for Potoff (155.75 K) is nearly twice that of TraPPE-UA (87 K). The second, and more important, reason is the extreme difference in  $\lambda$ . The LJ 12-6 potential is much softer than the Mie 36-6 potential. Therefore, none of the configurations sampled with TraPPE-UA represents a reasonable configuration for the Potoff model. This is manifested by the fact that  $N_{\text{eff}} \approx 1$  for every state point (not depicted). By contrast, PCFR provides relatively accurate estimates of  $U^{\text{dep}}$  for hexafluoroethane. As discussed in Section 3.1.1, this is because PCFR effectively “rescales” the configurations to avoid infeasible energies and forces.

Panel b) demonstrates that the MBAR single reference and “perturbed model” results observed for ethane are similar to those for larger *n*-alkanes. Specifically, the MBAR deviations from direct simulation are significantly positive for each compound. By contrast, the PCFR percent deviations increase in magnitude with increasing chain-length. There

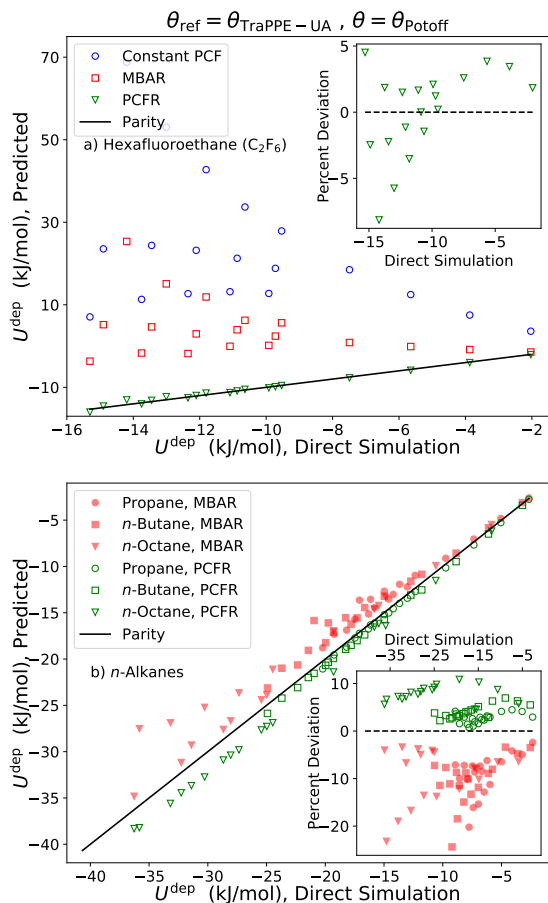


Figure 7: MBAR performs poorly for large changes in  $\lambda$  ( $\lambda = 36$  for hexafluoroethane) while PCFR provides reasonable estimates of  $U_{\text{dep}}$ , although the reliability of PCFR diminishes with increasing chain-length of  $n$ -alkanes. Panels a) and b) correspond to hexafluoroethane and  $n$ -alkanes, respectively. Panel a) compares PCFR with the “constant PCF” and MBAR results while Panel b) only includes the MBAR and PCFR results. The parity plots compare direct simulation values with those predicted using the different surrogate models. Percent deviation plots are included as insets for PCFR in Panel a) and for MBAR and PCFR in Panel b).

are at least two likely explanations for the poor performance of PCFR for  $n$ -octane. The first reason is that the error for each interaction site accumulates and, therefore, the overall error increases with the number of sites. The second reason is that PCFR assumes that the PCF of all site-types, regardless of molecular topology, can be scaled based on the non-bonded interactions. However, it makes intuitive sense that the PCF for  $\text{CH}_3$  sites (which

are located at a terminal position) would be more sensitive to the non-bonded interactions than the PCF for a  $\text{CH}_2$  site with several neighboring sites. Therefore, PCFR should primarily be utilized for smaller compounds and/or when non-bonded parameters are modified for fewer site-types.

## 4 Recommendations and Limitations

MBAR provides accurate estimates of  $U^{\text{dep}}$  and  $Z$  (and, thereby,  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$ ) when there is sufficient configurational overlap between the reference force field(s) and the non-simulated force field. We recommend using the number of effective samples ( $N_{\text{eff}}$ ) to quantify the overlap. Specifically, we recommend that MBAR be utilized if  $N_{\text{eff}} > 50$ . Multiple reference parameter sets are necessary to ensure that  $N_{\text{eff}} > 50$  over a large region of parameter space.

The key limitations to implementing MBAR are to determine the best reference parameters and to minimize the number of references required. More reference force fields necessitates more direct simulations. Therefore, the reference force fields should be chosen judiciously such that adequate, but not excessive, sampling of the parameter space is achieved in the region of most importance. For example, an adaptive sampling algorithm can determine which additional parameters will reduce the MBAR uncertainties.<sup>48</sup> However, this approach attempts to reduce the uncertainty in the entire parameter space while we are only interested in the parameter space near the optimum.

Rather than sampling hundreds of  $\epsilon_{\text{ref}}$  and  $\sigma_{\text{ref}}$  sets for  $\lambda_{\text{ref}} = 12$ , we recommend sampling a few different  $\epsilon_{\text{ref}}$  and  $\sigma_{\text{ref}}$  sets for each value of  $\lambda$ . Unfortunately, the optimal  $\epsilon_{\text{ref}}$  for a given  $\lambda$  is not known *a priori* and increases with increasing  $\lambda$ . For this reason, a key recommendation from this study is that PCFR be utilized to determine the reference parameters for MBAR when the non-bonded potential form (particularly the repulsive

barrier, i.e.  $\lambda$ ) is modified.

Specifically, in the case of converting from a single LJ 12-6 reference force field to a Mie  $\lambda$ -6 potential, the optimal  $\epsilon_{\text{ref}}$  for different values of  $\lambda$  is determined by minimizing the RMS of  $\log_{10}(P_{\text{v}}^{\text{sat}})$  with  $\sigma = \sigma_{\text{ref}}$ . We also found that minimizing the RMS of  $U^{\text{dep}}$  using PCFR could be a reliable approach, as shown in Section SI.III of the Supporting Information. We recommend maintaining  $\sigma$  constant during this preliminary optimization since the PCFR estimates for  $P_{\text{v}}^{\text{sat}}$  (and  $U^{\text{dep}}$ ) are most reliable when  $\sigma \approx \sigma_{\text{ref}}$ . Subsequently, multiple references are sampled for the desired  $\lambda$  using this ‘‘PCFR-optimal’’  $\epsilon_{\text{ref}}$ .

Since  $P_{\text{v}}^{\text{sat}}$  (and  $U^{\text{dep}}$ ) is fairly insensitive to  $\sigma$  (or  $r_{\text{min}}$ ), we recommend that a constraint be applied to  $\sigma_{\text{ref}}$  (or  $r_{\text{min,ref}}$ ) to reduce the number of reference parameters. For example, the TraPPE-UA LJ 12-6 model has been well optimized to match  $\rho_1^{\text{sat}}$  (which depends strongly on  $\sigma$ ), therefore we could constrain the reference  $\sigma$  (and  $r_{\text{min}}$ ) values to be within a certain ‘‘trust region’’, say  $\pm 1\%$ . Alternatively, we recommend that  $\sigma_{\text{ref}} \geq 0.99 \times \sigma_{\text{TraPPE}}$  and  $r_{\text{min,ref}} \leq 1.01 \times r_{\text{min,TraPPE}}$  for  $\lambda > 12$  and vice versa for  $\lambda < 12$ . This empirical recommendation is based on the fact that most united-atom sites follow the trend that  $\sigma_{\text{Potoff}} > \sigma_{\text{TraPPE}}$  and  $r_{\text{min,Potoff}} < r_{\text{min,TraPPE}}$  where  $\lambda_{\text{Potoff}} > 12$ .

To facilitate future implementation, we recommended the following algorithm (referred to as PCFR-MBAR-ITIC) for rapid parameterization of non-bonded interactions aimed at accurate prediction of vapor-liquid equilibria:

1. Perform molecular simulations using:
  - (a) NVT ensemble (either MD or MC)
  - (b) ITIC conditions (19  $\rho$ - $T$  state points: nine for isotherm, two for each of the five isochores)
  - (c) Reference force field(s) ( $\theta_{\text{ref}}$ ), initial reference is taken from literature (important to have reasonably optimized  $\sigma_{\text{ref}}$ )

2. Store independent configurations or basis functions (see Section SI.IV in the Supporting Information)
3. Determine additional references:
  - (a)  $\sigma_{\text{ref}}$ : evenly spaced within “trust region” (reasonable range for optimal  $\sigma$  based on  $\sigma_{\text{ref}}$ )
  - (b)  $\epsilon_{\text{ref}}$ : minimize the RMS of  $\log_{10}(P_{\text{v}}^{\text{sat}})$  (or  $U^{\text{dep}}$ ) predicted by PCFR and ITIC
4. Repeat Steps 1-2 using additional references found in Step 3
5. Optimize force field parameters:
  - (a) Predict  $U^{\text{dep}}$  and  $Z$  with MBAR
  - (b) Calculate  $\rho_{\text{l}}^{\text{sat}}$ ,  $\rho_{\text{v}}^{\text{sat}}$  and  $P^{\text{sat}}$  with ITIC
  - (c) Define an objective function that depends on  $\rho_{\text{l}}^{\text{sat}}$ ,  $\rho_{\text{v}}^{\text{sat}}$ ,  $P^{\text{sat}}$ ,  $U^{\text{dep}}$  and/or  $Z$
6. Determine if additional references are needed based on  $N_{\text{eff}}$  near optimum
7. Repeat Steps 5-6 until parameters converge to within a desired tolerance
8. Quantify uncertainty in non-bonded parameters

The computational bottleneck for configuration-sampling-based, statistical,<sup>11,14,16–19</sup> and equation-of-state<sup>4,20–28</sup> surrogate models is the molecular simulation step. Specifically, the “real time to solution” of the PCFR-MBAR-ITIC algorithm depends primarily on Step 1, namely, the real time to perform the direct molecular simulations. The post-simulation optimization time, i.e. the real time required to complete Steps 3 and 5, is negligible in comparison. By performing the additional reference simulations in parallel, i.e. Step 4, the “real time to solution” is approximately the same as the real time to perform a single

NVT ensemble simulation, which depends on the computer hardware, simulation software, force field complexity, cut-off distance, number of molecules, size of compound, etc. For example, the results presented in Section 5 were obtained in approximately one week of real time.

Note that the PC-SAFT algorithm proposed in Reference 24 determines additional reference parameter sets sequentially by iteratively finding a new proposed optimum and, therefore, molecular simulations must be performed in serial. This is because the PC-SAFT equation-of-state surrogate model is limited to a single reference parameter set, similar to PCFR. Information from previous iterations is effectively lost. By contrast, MBAR and meta-models<sup>16-18</sup> are able to utilize information from several reference parameter sets and, thus, molecular simulations can be performed in parallel. Although this would appear to be a significant algorithmic advantage for PCFR-MBAR-ITIC, the PC-SAFT approach typically finds the optimal parameter set within 2 to 3 iterations of direct simulation. Therefore, the “real time to solution” is comparable for all three surrogate model classes. Furthermore, the PCFR-MBAR-ITIC algorithm relies on starting with a reasonable reference parameter set, see Step 1c, whereas the PC-SAFT approach demonstrates rapid convergence even with a poor initial guess.<sup>24</sup>

## 5 Algorithm Application

We apply the PCFR-MBAR-ITIC algorithm from Section 4 to optimize transferable CH<sub>3</sub> and CH<sub>2</sub> Mie  $\lambda$ -6 parameters for *n*-alkanes. Although we do not explicitly define an objective function, we provide contours of the RMS deviations from the REFPROP  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$  values to help visualize the optimal region. We perform this analysis sequentially by assuming the CH<sub>3</sub> parameters obtained for ethane are transferable to larger *n*-alkanes. The CH<sub>2</sub> parameter analysis is performed independently for propane,

*n*-butane, and *n*-octane to investigate the transferability of the CH<sub>2</sub> parameters.

For Step 1, we simulate ethane using the TraPPE-UA LJ 12-6 force field as our initial reference. Additional reference points (i.e. Step 3) are determined by minimizing the  $\log_{10}(P_v^{\text{sat}})$  RMS predicted by PCFR. Figure 8 depicts the results from this analysis. The lines in Figure 8 represent the “PCFR-optimal”  $\epsilon_{\text{CH}_3}$  values for the corresponding  $\sigma_{\text{CH}_3}$  and  $\lambda_{\text{CH}_3}$ . The points in Figure 8 are the  $\epsilon_{\text{ref,CH}_3}$  and  $\sigma_{\text{ref,CH}_3}$  values determined in Step 3 for integer values of  $\lambda_{\text{ref,CH}_3} = 13$  to 18. Specifically,  $\epsilon_{\text{ref,CH}_3}$  for a given  $\lambda_{\text{ref,CH}_3}$  corresponds to the “PCFR-optimal”  $\epsilon_{\text{CH}_3}$  value for  $\sigma_{\text{CH}_3} = \sigma_{\text{TraPPE,CH}_3}$ . The  $\sigma_{\text{ref,CH}_3}$  values are evenly spaced such that  $\sigma_{\text{ref,CH}_3} \geq 0.99 \times \sigma_{\text{TraPPE,CH}_3}$  and  $r_{\text{min,ref,CH}_3} \leq 1.01 \times r_{\text{min,TraPPE,CH}_3}$ . The spacing between neighboring  $\sigma_{\text{ref,CH}_3}$  values is no more than 0.0025 nm since this is the range over which we found MBAR to be reliable (i.e.  $\bar{N}_{\text{eff}} > 50$ , recall Figure 3). Note that the “Potoff” points are included in Figure 8 for comparison with the “PCFR-optimal” curves for  $\lambda_{\text{CH}_3} = 14, 16$ , and 18. “Potoff” points for  $\lambda_{\text{CH}_3} = 14, 18$  are approximations extracted from Figure 2 of Reference [1].<sup>1</sup>

The “PCFR-optimal” lines in Figure 8 follow a reasonable, smooth trend and are in good quantitative agreement with the “Potoff” values for a given  $\lambda_{\text{CH}_3}$ . This suggests that  $\epsilon_{\text{ref,CH}_3}$  and  $\sigma_{\text{ref,CH}_3}$  obtained in Step 3 are near the optimum for the respective  $\lambda_{\text{CH}_3}$  values and, therefore, should adequately sample the relevant region of parameter space. To substantiate this statement we perform Steps 4-5 for the  $\lambda_{\text{CH}_3} = 16$  reference values plotted in Figure 8. The configurations sampled from these additional references are then reweighted using MBAR to predict  $U^{\text{dep}}$  and  $Z$ . The ITIC equations convert the estimated  $U^{\text{dep}}$  and  $Z$  values to  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$ . Figure 9 shows the high level of agreement of direct simulation and MBAR RMS values relative to REFPROP for the same range of  $\epsilon_{\text{CH}_3}$  and  $\sigma_{\text{CH}_3}$  as plotted previously in Figure 6.

This process is then repeated for propane, *n*-butane, and *n*-octane to examine the CH<sub>2</sub> parameter space. For simplicity and visualization purposes, we only investigate

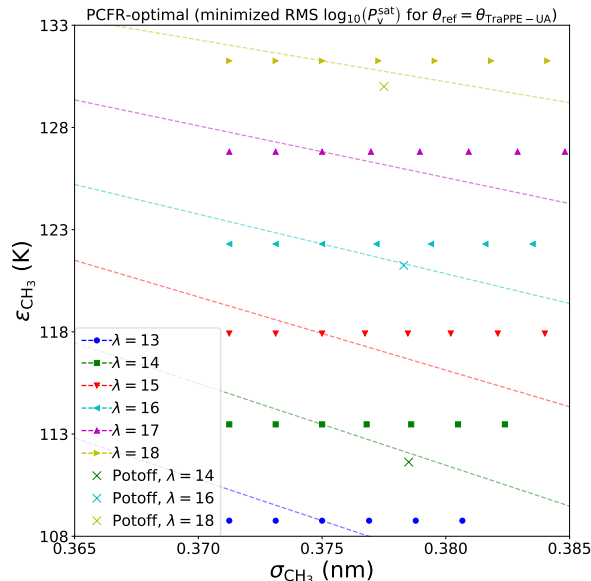


Figure 8: Ethane “PCFR-optimal”  $\epsilon_{\text{ref}}$  values for  $\text{CH}_3$  follow a reasonable trend and are in good quantitative agreement with “Potoff” points. The dashed lines are obtained by minimizing  $\text{RMS } \log_{10}(P_v^{\text{sat}})$  over a range of  $\sigma$  for a given  $\lambda$ . The different symbols correspond to the recommended  $\epsilon$  and  $\sigma$  reference sets for the respective values of  $\lambda$ .

$\lambda_{\text{CH}_2} = 16$  and use the Potoff  $\text{CH}_3$  parameter values, since these appear to be near optimal (see Figure 9). The TraPPE-UA force field is again used as the single reference for PCFR. The PCFR-optimal  $\epsilon_{\text{ref,CH}_2}$  values for propane and  $n$ -butane are obtained by minimizing their respective RMS of  $\log_{10}(P_v^{\text{sat}})$  with  $\sigma_{\text{CH}_2} = \sigma_{\text{TraPPE,CH}_2} = 0.395 \text{ nm}$ ,  $\lambda_{\text{CH}_2} = 16$ ,  $\epsilon_{\text{CH}_3} = 121.25 \text{ K}$ ,  $\sigma_{\text{CH}_3} = 0.3783 \text{ nm}$ , and  $\lambda_{\text{CH}_3} = 16$ . Since PCFR is less reliable for larger molecules, we use the optimal  $\epsilon_{\text{CH}_2}$  from  $n$ -butane for the  $\epsilon_{\text{ref,CH}_2}$  of  $n$ -octane. We simulate eight evenly spaced  $\sigma_{\text{ref,CH}_2}$  values between  $0.99 \times \sigma_{\text{TraPPE,CH}_2}$  and  $1.01 \times r_{\text{min,TraPPE,CH}_2}$  with constant  $\epsilon_{\text{ref,CH}_2}$ . Finally, MBAR and ITIC are implemented to predict  $\rho_1^{\text{sat}}$  and  $P_v^{\text{sat}}$  over a wide range of  $\epsilon_{\text{CH}_2}$  and  $\sigma_{\text{CH}_2}$  with  $\lambda_{\text{CH}_2} = 16$  and using the Potoff  $\text{CH}_3$  parameter values. Figure 10 compares the direct simulation and MBAR RMS values relative to REF-PROP for the joint  $\epsilon_{\text{CH}_2}$  and  $\sigma_{\text{CH}_2}$  parameter space. For clarity, only the propane direct simulation results are plotted and only one contour level is included for  $n$ -butane and  $n$ -octane.



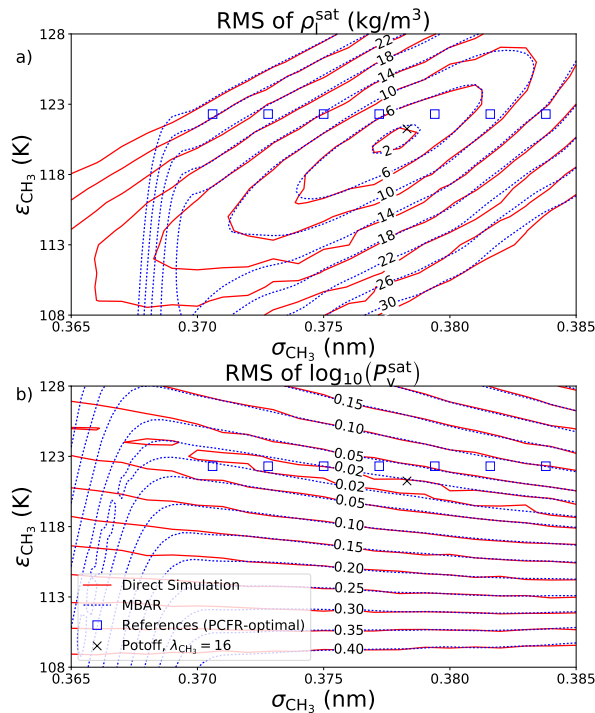


Figure 9: Contours of RMS deviation from REFPROP ethane values predicted by MBAR using PCFR-optimal references are in excellent agreement with direct simulation over a wide region of the CH<sub>3</sub> parameter space. Panels a) and b) plot the RMS for  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$ , respectively. “References (PCFR-optimal)” are obtained as described in Step 3 for  $\lambda = 16$  as found in Figure 8. The Potoff CH<sub>3</sub> parameters are included only as a visual reference. Direct simulations were performed on a 21x21 grid equally spaced between 108-128 K and 0.365-0.385 nm.

The MBAR contours for propane in Figure 10 appear to be in excellent agreement with direct simulation over the relevant region of parameter space. Note that the MBAR contours in Figures 9-10 are smoother than direct simulation because MBAR effectively includes uncorrelated replicate simulations from each reference, whereas the direct simulation contours were obtained without any replicate simulations.

The good agreement between the MBAR and direct simulation contours (and the smoothness) near the optimum is important for quantifying parameter uncertainties in Step 8 (although the “optimal”  $\epsilon$  and  $\sigma$  values depend on how the objective function is defined). Although Bayesian inference methods (such as Markov Chain Monte Carlo) are

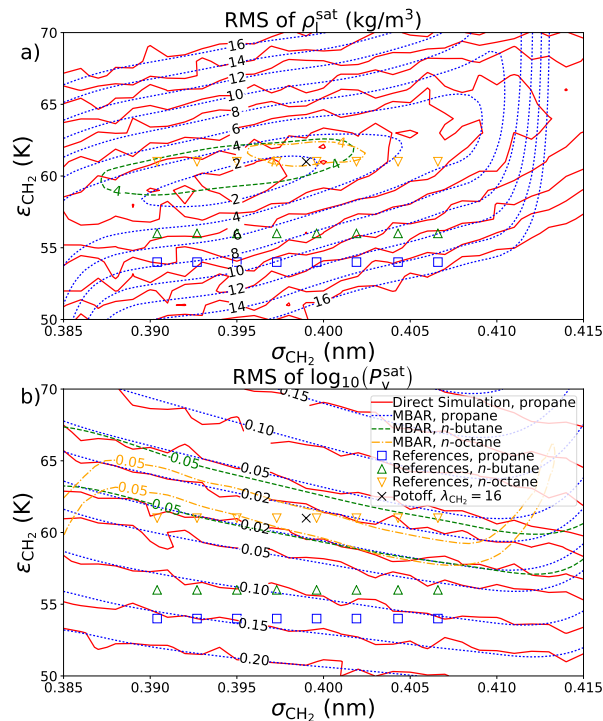


Figure 10: Contours of RMS deviation from REFPROP propane values predicted by MBAR using PCFR-optimal references are in excellent agreement with direct simulation over a wide region of the  $\text{CH}_2$  parameter space. RMS contours for propane,  $n$ -butane, and  $n$ -octane are slightly different. Panels a) and b) plot the RMS for  $\rho_1^{\text{sat}}$  and  $\log_{10}(P_v^{\text{sat}})$ , respectively. Reference values are obtained as described in Step 3 for  $\lambda_{\text{CH}_2} = 16$ . The  $\text{CH}_3$  parameters are from the Potoff force field, i.e.  $\epsilon_{\text{CH}_3} = 121.25$  K,  $\sigma_{\text{CH}_3} = 0.3783$  nm, and  $\lambda_{\text{CH}_3} = 16$ .  $\text{CH}_3$ - $\text{CH}_2$  cross-interactions are calculated using Lorentz-Berthelot combining rules for  $\epsilon$  and  $\sigma$  and an arithmetic mean for  $\lambda$ . The Potoff  $\text{CH}_2$  parameters are included only as a visual reference. Direct simulations were performed on a  $21 \times 31$  grid equally spaced between 50-70 K and 0.385-0.415 nm.

beyond the scope of this article, it is worth mentioning that these UQ methods are feasible because MBAR (with basis functions) is several orders of magnitude less expensive than direct simulation. Also, by utilizing basis functions the computational cost to perform a UQ analysis, and to generate Figure 10, is the same for  $n$ -octane and  $n$ -propane (see Section SI.IV in the Supporting Information).

Although not depicted here, we observe that the average number of effective samples ( $\bar{N}_{\text{eff}} \gg 50$ ) for most of the parameter space investigated (Step 6). Therefore, we conclude

that no additional references are needed near the optimum.

## 6 Conclusions

The development of accurate and reliable surrogate models is essential for rapid, high-dimensional parameterization and uncertainty quantification of force fields. The non-bonded potential has been the focus of this study since it is typically parameterized to reproduce vapor-liquid equilibria data. Configuration-sampling-based surrogate models, such as Multistate Bennett Acceptance Ratio (MBAR) and Pair Correlation Function Rescaling (PCFR) are well suited for estimating internal energies and pressures (compressibility) for a non-simulated non-bonded potential. Isothermal Isochoric (ITIC) thermodynamic integration converts the MBAR and PCFR outputs into saturation properties ( $\rho_1^{\text{sat}}$ ,  $\rho_v^{\text{sat}}$ , and  $P_v^{\text{sat}}$ ).

We performed several tests to determine the range of reliability of MBAR and PCFR for parameterizing a Mie  $\lambda$ -6 potential. MBAR provides accurate estimates of  $U^{\text{dep}}$ ,  $Z$ ,  $\rho_1^{\text{sat}}$ , and  $P_v^{\text{sat}}$  when  $N_{\text{eff}} > 50$ . Unfortunately,  $N_{\text{eff}} \ll 50$  when  $\lambda \not\approx \lambda_{\text{ref}}$  and/or  $\sigma \not\approx \sigma_{\text{ref}}$ . Therefore, MBAR should primarily be used when several  $\sigma_{\text{ref}}$  values are sampled from  $\lambda = \lambda_{\text{ref}}$ . By contrast, PCFR accurately predicts  $U^{\text{dep}}$  and  $P_v^{\text{sat}}$  from a single reference force field for a wide range of  $\lambda$  values (and a moderate range of  $\sigma$  values). However, PCFR has the wrong trend for  $Z$  and  $\rho_1^{\text{sat}}$  with respect to  $\epsilon$  and  $\sigma$ .

Therefore, MBAR and PCFR are complementary tools. PCFR is a useful exploratory tool for proposing optimal regions of parameter space while MBAR is for final optimization and parameter uncertainty quantification. We developed a method that utilizes PCFR to rapidly determine which reference parameters should be sampled to improve the MBAR results. We demonstrated that this algorithm is a reliable and efficient means for parameterizing a Mie  $\lambda$ -6 potential for  $n$ -alkanes. We expect the same to be true in

general for other potentials, such as the Exponential-6. Future studies will utilize these surrogate models (MBAR and PCFR) to rapidly optimize accurate force fields with meaningful estimates of parameter uncertainty.

## Acknowledgments

The authors would like to acknowledge J. Richard Elliott for his collaboration and guidance for implementing the ITIC approach. This research was performed while R.A.M. held a National Research Council (NRC) Postdoctoral Research Associateship at the National Institute of Standards and Technology (NIST).

## Supporting Information

Supporting Information provides simulation set-up specifications in Section SI.I, details for the ITIC analysis in Section SI.II, an alternative PCFR-optimal approach in Section SI.III, and a thorough discussion of basis functions in Section SI.IV. This information is available free of charge via the Internet at <http://pubs.acs.org>. The MBAR-PCFR-ITIC code is available at [https://github.com/ramess101/MBAR\\_ITIC](https://github.com/ramess101/MBAR_ITIC).

## Appendix

We now provide a brief summary of some other PCFR methods to estimate  $g_{ij}(\theta; r_{ij})$  that we have investigated. First, the simplest PCFR approach is to assume that the pair correlation functions for the two force fields are equal (referred to as “constant PCF” in Section 3):

$$g_{ij}(\theta; r_{ij}) = g_{ij}(\theta_{\text{ref}}; r_{ij}) \quad (27)$$

For very small differences between  $\theta$  and  $\theta_{\text{ref}}$  this assumption may be a good approximation. Applying Equation 27 to Equations 19 and 20 yields the expressions:

$$\begin{aligned} \langle U^{\text{dep}}(\theta) \rangle &\approx \langle U^{\text{dep}}(\theta_{\text{ref}}) \rangle \\ &+ 2\pi\rho R_g \sum_{i=1}^{N_S-1} \sum_{j>i}^{N_S} \int_0^\infty [u_{ij}^{\text{vdw}}(\theta; r_{ij}) - u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})] g_{ij}(\theta_{\text{ref}}; r_{ij}) r_{ij}^2 \partial r_{ij} \end{aligned} \quad (28)$$

$$\begin{aligned} \langle Z(\theta) \rangle &\approx \langle Z(\theta_{\text{ref}}) \rangle \\ &- \frac{2\pi\rho}{3T} \sum_{i=1}^{N_S-1} \sum_{j>i}^{N_S} \int_0^\infty \left[ \frac{\partial u_{ij}^{\text{vdw}}(\theta; r_{ij})}{\partial r_{ij}} - \frac{\partial u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})}{\partial r_{ij}} \right] g_{ij}(\theta_{\text{ref}}; r_{ij}) r_{ij}^3 \partial r_{ij} \end{aligned} \quad (29)$$

A connection exists between Equations 28-29 and the MBAR equations for calculating expectation values of  $U^{\text{dep}}$  and  $Z$  (Equations 12-13). The assumption that the pair correlation functions (i.e. configurations) for the two force fields are equal (Equation 27) is mathematically equivalent to assigning each sampled configuration an equal probability in the MBAR equations, i.e.  $W_n(\theta) = \frac{1}{N}$  for all  $\mathbf{x}_n$ . With the weights equal, Equations 10-11 are no longer needed to calculate  $U^{\text{dep}}$  and  $Z$  as Equations 12-13 reduce to:

$$\langle U^{\text{dep}}(\theta) \rangle = \frac{1}{N} \sum_{n=1}^N U^{\text{dep}}(\mathbf{x}_n; \theta) \quad (30)$$

$$\langle Z(\theta) \rangle = \frac{1}{N} \sum_{n=1}^N Z(\mathbf{x}_n; \theta) \quad (31)$$

Although Equations 28-29 (“constant PCF”) and 30-31 (“equal weights”) are equivalent, the use of pair correlation functions in Equations 28-29 significantly reduces the book-keeping compared to MBAR when  $N$  configurations (i.e. coordinates of every interaction-

site for each molecule) are stored. As shown in Figures 1, 2, and 7 “constant PCF” (i.e. Equations 28-29 or 30-31) has a strong bias in  $U^{\text{dep}}$  and  $Z$ .

Rather than assuming the PCFs for the two force fields are equal with respect to the absolute distance,  $r_{ij}$ , (Equation 27) we can assume  $g_{ij}$  is constant with respect to a reduced distance,  $r_{ij}^*$ :

$$g_{ij}(\theta; r_{ij}^*) \approx g_{ij}(\theta_{\text{ref}}; r_{ij}^*) \quad (32)$$

where the reduced distance ( $r_{ij}^*$ ) may be defined as  $\frac{r}{\sigma}$  or  $\frac{r}{r_{\text{min}}}$ , for example. This is equivalent to the MBAR configuration mapping example discussed in Section 2.3.1 for the single-site LJ system. The advantage of Equation 32 is that the positions of the peaks in  $g_{ij}$  depend strongly on  $\sigma$  (and  $r_{\text{min}}$ ). However, this approach is only rigorous if the system volume is also scaled, which is not straightforward for a multi-site molecule. For a constant volume, the location of the first peak and the peaks heights depend strongly on  $\sigma$ .

In addition, Equation 32 does not account for changes in the well depth ( $\epsilon$ ) or the steepness of repulsive barrier ( $\lambda$  for the Mie  $\lambda$ -6 potential). The value of  $\epsilon$  changes the height of the first peak for  $g_{ij}$ , while  $\lambda$  changes the slope of the initial ascent for  $g_{ij}$ . In order to account for changes in  $\lambda$  and/or  $\epsilon$ , we can rescale the PCF by estimating the potential of mean force ( $w$ , or PMF) as:

$$w_{ij}(\theta; r_{ij}) \approx w_{ij}(\theta_{\text{ref}}; r_{ij}) + [u_{ij}^{\text{vdw}}(\theta; r_{ij}) - u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})] \quad (33)$$

where  $w$  (or PMF) is defined as:

$$w_{ij}(\theta; r_{ij}) \equiv \frac{-\ln(g_{ij}(\theta; r_{ij}))}{T} \quad (34)$$

Equation 34 is rearranged to yield:

$$g_{ij}(\theta; r_{ij}) \equiv \exp\left(\frac{-w_{ij}(\theta; r_{ij})}{T}\right) \quad (35)$$

and an estimation of the PCF is obtained by substituting Equation 33 into Equation 35:

$$g_{ij}(\theta; r_{ij}) \approx g_{ij}(\theta_{\text{ref}}; r_{ij}) \exp\left(\frac{-\Delta u_{ij}^{\text{vdw}}(\theta, \theta_{\text{ref}}; r_{ij})}{T}\right) \quad (36)$$

where  $\Delta u_{ij}^{\text{vdw}}(\theta, \theta_{\text{ref}}; r_{ij}) = u_{ij}^{\text{vdw}}(\theta; r_{ij}) - u_{ij}^{\text{vdw}}(\theta_{\text{ref}}; r_{ij})$ . Note that  $r_{ij}$  does not need to be a scaled distance ( $r_{ij}^*$ ), because  $\Delta u_{ij}^{\text{vdw}}$  already accounts for the difference in  $\sigma$  (or  $r_{\text{min}}$ ). Equation 36 is then substituted into Equations 19-20. In our experience, Equation 36 is better than Equations 27 and 32, although the rescaled PCF still suffers from some deficiencies. Future development of additional PCFR methods is an ongoing task.

In many instances, Equations 25-26 yield similar results to Equations 19-20 and 36. This is not surprising, since Equation 36 can also be viewed as multiplying the PCF by the ratio of the zeroth order radial distribution functions ( $g_0$ ) of  $\theta$  and  $\theta_{\text{ref}}$ . However, Equations 25-26 are simpler and have some computational benefits. For example, Equations 25-26 are numerically stable since they do not depend on the bin widths of the  $g_{ij}(\theta_{\text{ref}}; r_{ij})$  histogram. Therefore, in this study the PCFR results are obtained using Equations 25-26.

## References

- (1) 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.
- (2) 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.
- (3) 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.
- (4) 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.
- (5) Shah, M. S.; Siepmann, J. I.; Tsapatsis, M. Transferable potentials for phase equilibria. Improved united-atom description of ethane and ethylene. *AIChE J.* **2017**, *63*, 5098–5110.
- (6) 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.
- (7) Rhinehart, R. R.; Su, M.; Manimegalai-Sridhar, U. Leapfrogging and synoptic Leapfrogging: A new optimization approach. *Comput. Chem. Eng.* **2012**, *40*, 67 – 81.
- (8) Stöbener, K.; Klein, P.; Reiser, S.; Horsch, M.; Kufer, K.-H.; Hasse, H. Multicriteria optimization of molecular force fields by Pareto approach. *Fluid Ph. Equilibria* **2014**, *373*, 100 – 108.
- (9) Werth, S.; Stöbener, K.; Klein, P.; Kufer, K.-H.; Horsch, M.; Hasse, H. Molecular modelling and simulation of the surface tension of real quadrupolar fluids. *Chem. Eng. Sci.* **2015**, *121*, 110 – 117, 2013 Danckwerts Special Issue on Molecular Modelling in Chemical Engineering.
- (10) Stöbener, K.; Klein, P.; Horsch, M.; Kufer, K.; Hasse, H. Parametrization of two-center Lennard-Jones plus point-quadrupole force field models by multicriteria optimization. *Fluid Ph. Equilibria* **2016**, *411*, 33 – 42.



- (11) Cailliez, F.; Pernot, P. Statistical approaches to forcefield calibration and prediction uncertainty in Molecular Simulation. *J. Chem. Phys.* **2011**, *134*, 054124.
- (12) Rizzi, F.; Najm, H. N.; Debusschere, B. J.; Sargsyan, K.; Salloum, M.; Adalsteins-son, H.; Knio, O. M. Uncertainty Quantification in MD Simulations. Part II: Bayesian Inference of Force-Field Parameters. *Multiscale Model. Sim.* **2012**, *10*, 1460–1492.
- (13) Angelikopoulos, P.; Papadimitriou, C.; Koumoutsakos, P. Bayesian uncertainty quantification and propagation in molecular dynamics simulations: A high performance computing framework. *J. Chem. Phys.* **2012**, *137*, 144103.
- (14) 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.
- (15) Hülsmann, M.; Vrabec, J.; Maaß, A.; Reith, D. Assessment of numerical optimization algorithms for the development of molecular models. *Comput. Phys. Commun.* **2010**, *181*, 887 – 905.
- (16) Laurent, L.; Le Riche, R.; Soulier, B.; Boucard, P.-A. An Overview of Gradient-Enhanced Metamodels with Applications. *Archives of Computational Methods in Engineering* **2017**, *24*, 1–46.
- (17) Hülsmann, M.; Kirschner, K. N.; Krämer, A.; Heinrich, D. D.; Krämer-Fuhrmann, O.; Reith, D. In *Foundations of Molecular Modeling and Simulation: Select Papers from FOMMS 2015*; Snurr, R. Q., Adjiman, C. S., Kofke, D. A., Eds.; Springer Singapore: Singapore, 2016; pp 53–77.
- (18) Hülsmann, M.; Reith, D. SpaGrOW - A Derivative-Free Optimization Scheme for Intermolecular Force Field Parameters Based on Sparse Grid Methods. *Entropy* **2013**, *15*, 3640.

- (19) Cailliez, F.; Bourasseau, A.; Pernot, P. Calibration of forcefields for Molecular Simulation: Sequential design of computer experiments for building cost-efficient kriging metamodels. *J. Comput. Chem.* **2014**, *35*, 130–149.
- (20) Lotfi, A.; Vrabec, J.; Fischer, J. Vapour liquid equilibria of the Lennard-Jones fluid from the NpT plus test particle method. *Mol. Phys.* **1992**, *76*, 1319–1333.
- (21) Thol, M.; Rutkai, G.; Span, R.; Vrabec, J.; Lustig, R. Equation of State for the Lennard-Jones Truncated and Shifted Model Fluid. *Int. J. Thermophys.* **2015**, *36*, 25–43.
- (22) Thol, M.; Rutkai, G.; Köster, A.; Lustig, R.; Span, R.; Vrabec, J. Equation of State for the Lennard-Jones Fluid. *J. Phys. Chem. Ref. Data* **2016**, *45*, 023101.
- (23) Stoll, J.; Vrabec, J.; Hasse, H.; Fischer, J. Comprehensive study of the vapour-liquid equilibria of the pure two-centre Lennard-Jones plus point-quadrupole fluid. *Fluid Ph. Equilibria* **2001**, *179*, 339 – 362.
- (24) van Westen, T.; Vlugt, T. J. H.; Gross, J. Determining Force Field Parameters Using a Physically Based Equation of State. *J. Phys. Chem. B* **2011**, *115*, 7872–7880.
- (25) Papaioannou, V.; Calado, F.; Lafitte, T.; Dufal, S.; Sadeqzadeh, M.; Jackson, G.; Adjiman, C. S.; Galindo, A. Application of the SAFT- $\gamma$  Mie group contribution equation of state to fluids of relevance to the oil and gas industry. *Fluid Ph. Equilibria* **2016**, *416*, 104 – 119, Special Issue: SAFT 2015.
- (26) Avendaño, C.; Lafitte, T.; Adjiman, C. S.; Galindo, A.; Müller, E. A.; Jackson, G. SAFT- $\gamma$  Force Field for the Simulation of Molecular Fluids: 2. Coarse-Grained Models of Greenhouse Gases, Refrigerants, and Long Alkanes. *J. Phys. Chem. B* **2013**, *117*, 2717–2733.

- (27) Hemmen, A.; Panagiotopoulos, A. Z.; Gross, J. Grand Canonical Monte Carlo Simulations Guided by an Analytic Equation of State-Transferable Anisotropic Mie Potentials for Ethers. *J. Phys. Chem. B* **2015**, *119*, 7087–7099.
- (28) Weidler, D.; Gross, J. Transferable Anisotropic United-Atom Force Field Based on the Mie Potential for Phase Equilibria: Aldehydes, Ketones, and Small Cyclic Alkanes. *Ind. Eng. Chem. Res.* **2016**, *55*, 12123–12132.
- (29) 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.
- (30) Mick, J. R. Force Field Development with GOMC, A Fast New Monte Carlo Molecular Simulation Code. Ph.D. thesis, Wayne State University, 2016.
- (31) Hoang, H.; Delage-Santacreu, S.; Galliero, G. Simultaneous Description of Equilibrium, Interfacial, and Transport Properties of Fluids Using a Mie Chain Coarse-Grained Force Field. *Ind. Eng. Chem. Res.* **2017**, *56*, 9213–9226.
- (32) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Clarendon Press ; Oxford University Press: Oxford England New York, 1987; pp xix, 385 p.
- (33) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *J. Am. Chem. Soc.* **1984**, *106*, 6638–6646.
- (34) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (35) Razavi, S. M. Optimization of a Transferable Shifted Force Field for Interfaces and In-

- homogenous Fluids using Thermodynamic Integration. M.Sc. thesis, The University of Akron, 2016.
- (36) Elliott, J. R.; Lira, C. T. *Introductory Chemical Engineering Thermodynamics, Second Edition*; Prentice Hall: Upper Saddle River, New Jersey, 2012.
- (37) Lemmon, E. W.; Huber, M. L.; McLinden, M. O. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 9.1, National Institute of Standards and Technology. 2013; <https://www.nist.gov/srd/refprop>.
- (38) Frenkel, M.; Chirico, R. D.; Diky, V.; Yan, X.; Dong, Q.; Muzny, C. ThermoData Engine (TDE): Software Implementation of the Dynamic Data Evaluation Concept. *Journal of Chemical Information and Modeling* **2005**, *45*, 816–838.
- (39) Schultz, A. J.; Kofke, D. A. Virial coefficients of model alkanes. *J. Chem. Phys.* **2010**, *133*, 104101.
- (40) Schultz, A. J.; Barlow, N. S.; Chaudhary, V.; Kofke, D. A. Mayer Sampling Monte Carlo calculation of virial coefficients on graphics processors. *Mol. Phys.* **2013**, *111*, 535–543.
- (41) Eggimann, B.; Bai, P.; Bliss, A.; Chen, Q.; Chen, T.; Corest-Morales, A.; Fetisov, E.; Haldoupis, E.; Harwood, D.; Lindsey, R.; Arachchi, T.; Shah, M.; Stern, H.; Struk, K.; Sung, J.; Sunnarborg, A.; Xue, B.; Siepmann, J. I. T-UA No. 2 ethane. TraPPE Validation Database, University of Minnesota: Minneapolis, MN. <http://www.chem.umn.edu/groups/siepmann/trappe/>, <http://www.chem.umn.edu/groups/siepmann/trappe/> (accessed 2015 June 11).
- (42) Shirts, M. R. Simple Quantitative Tests to Validate Sampling from Thermodynamic Ensembles. *J. Chem. Theory Comput.* **2013**, *9*, 909–926.

- (43) Schappals, M.; Mecklenfeld, A.; Kröger, L.; Botan, V.; Köster, A.; Stephan, S.; García, E. J.; Rutkai, G.; Raabe, G.; Klein, P.; Leonhard, K.; Glass, C. W.; Lenhard, J.; Vrabec, J.; Hasse, H. Round Robin Study: Molecular Simulation of Thermodynamic Properties from Models with Internal Degrees of Freedom. *J. Chem. Theory Comput.* **2017**, *13*, 4270–4280.
- (44) Abraham, M.; van der Spoel, D.; Lindahl, E.; B.Hess,; the GROMACS development team, GROMACS User Manual version 2016.3, [www.gromacs.org](http://www.gromacs.org) (2017).
- (45) Morales, A. D. C.; Economou, I. G.; Peters, C. J.; Siepmann, J. I. Influence of simulation protocols on the efficiency of Gibbs ensemble Monte Carlo simulations. *Mol. Simul.* **2013**, *39*, 1135–1142.
- (46) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.
- (47) Shirts, M. R. Reweighting from the mixture distribution as a better way to describe the Multistate Bennett Acceptance Ratio. <https://arxiv.org/abs/1704.00891>
- (48) 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.
- (49) 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.
- (50) Strett, W. B.; Staveley, L. A. K. Calculation on a Corresponding States Basis of the Volume Change on Mixing Simple Liquids. *J. Chem. Phys.* **1967**, *47*, 2449–2454.

- (51) Paliwal, H.; Shirts, M. R. Multistate reweighting and configuration mapping together accelerate the efficiency of thermodynamic calculations as a function of molecular geometry by orders of magnitude. *J. Chem. Phys.* **2013**, *138*, 154108.
- (52) Moustafa, S. G.; Schultz, A. J.; Kofke, D. A. Very fast averaging of thermal properties of crystals by Molecular Simulation. *Phys. Rev. E* **2015**, *92*, 043303.