January 2019

Molecular Exchange Monte Carlo. A Generalized Method For Identity Exchanges In Grand Canonical Monte Carlo Simulations

Mohammad Soroush Barhaghi
Wayne State University, m.soroush.b@gmail.com

MOLECULAR EXCHANGE MONTE CARLO. A GENERALIZED METHOD FOR IDENTITY EXCHANGES IN GRAND CANONICAL MONTE CARLO SIMULATIONS

by

MOHAMMAD SOROUSH BARHAGHI

THESIS

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

2019

MAJOR: COMPUTER SCIENCE

Approved by:

_____

Advisor Date

DEDICATION

I would like to dedicate to my work to my dear parents, Nasrin Akbari and Abdolhosein Soroush Barhaghi, and my siblings Mehrnaz and Mehrdad. Thank you for all your support and encouragement through my whole life. I never would have had the ambition and perseverance to pursue my graduate studies in the United States without you.

ACKNOWLEDGMENTS

I am deeply thankful for my advisors, Dr. Loren Schwiebert and Dr. Jeffrey Potoff for their guidance and support. Without their help and vision, I would not have the opportunity to become accomplished in my research. I would like to thank Dr. Korosh Torabi, Jason Mick, Younes Nejahi, and Niloofar Torabi for their limitless support.

In addition, it is my honor to invite Dr. Suzan Arslanturk and Dr. Jeffrey Potoff to serve on my defense committee.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

In Monte Carlo simulations in the grand canonical ensemble (GCMC), the chemical potential, volume and temperature are fixed ($\mu VT = \text{constant}$). Sampling of phase space is achieved through a variety of trial moves, such as displacement, and molecule insertion and deletion. For complex molecular typologies, additional trial moves, such as rigid body rotation and configurational-bias regrowth[1, 2], may be included to improve the sampling of conformational degrees of freedom. During the course of the simulation, the conjugate variables *N* (number of molecules) and *E* (potential energy) fluctuate. Because GCMC allows for the simulation of an open system, it has been used extensively to study the adsorption of gases in porous materials[3-6]. When combined with histogram-reweighting methods[7, 8], GCMC simulations provide precise predictions of vapor-liquid equilibria for pure fluids and mixtures[9, 10], and have been used to determine critical micelle concentrations for model surfactants[11].

1.1 Thesis Motivation

Perhaps the greatest challenge with GCMC simulations, however, is achieving a sufficient number of accepted molecular insertion/deletion moves to ensure adequate sampling of phase space. Therefore, significant effort has been expended to develop algorithms that improve the acceptance rate for molecule insertions and deletions. Biasing methods, such as rotational, energy and cavity-bias, were used to improve the efficiency of simulations for the adsorption of benzene and p-xylene in silicalite[12]. The introduction of configurational-bias Monte Carlo enabled the successful simulation of chain molecule adsorption in zeolites[13], which was followed by the coupled-decoupled[14] and reservoir methods[15, 16], which extended the complexity of systems that could be simulated to include molecules with branch points and rings.

These aforementioned biasing methods have greatly extended the complexity of systems that may be simulated with GCMC simulations, however, at high densities and low temperatures, the acceptance rate for molecule transfers is still unacceptably low due to the difficulty in finding a favorably sized cavity to insert a molecule. For example, in simulations of branched alkanes acceptance rates for molecule transfers at $0.7T_c$ were approximately 0.3%[9]. Others have sought to address these issues through the use of cavity-bias[17-19], to identify favorable locations to attempt molecule insertions, or continuous fractional component Monte Carlo[20, 21], and expanded ensembles[22, 23], where molecules are gradually inserted while the system is allowed to relax locally to minimize steric and energetic penalties due to molecule insertion.

For mixtures, a straightforward approach is to introduce a trial move where the identity of one molecule is changed to that of another[24]. The benefit of such a move is that steric overlaps are reduced significantly, leading to enhanced acceptance for the particle exchange. The identity exchange move has been used in many simulations of single particles in various ensembles, such as semi-grand[25, 26], Gibbs[24, 27, 28] and grand canonical[29-31]. The methodology has been extended to allow for the exchange of multiple solvent molecules with a polymer chain composed of solvent monomers without changing the coordinates of either polymer or solvent[31]. For the simulation of mixtures of colloids and solvent, it is necessary to swap a large colloid particle for multiple smaller solvent particles. By swapping multiple solvent particles, it is possible to create large enough voids such that a reasonable acceptance rate may be obtained for the insertion of colloid particles[29, 30]. For the exchange of a large particle with multiple small ones, Vink *et al*. used simple random insertions to determine the coordinates for the solvent particles. When inserting a large number of solvent particles, the potential for overlap increases, reducing the efficiency of the method. To address this issue, Kindt introduced the idea of “solvent repacking” for two-dimensional hard-disk and size

asymmetric three-dimensional Lennard-Jones systems, where configurational-bias was used to determine the positions of solvent particles in the large-small particle identity exchange[29, 32]. While a number of publications state that an identity exchange move was used for molecular systems[33-35], a detailed description of the algorithm and the acceptance criteria have not been published to date.

The previously described methods for identity exchange were generally applicable to only the special cases for which they were developed, e.g. single particle exchanges[26], a polymer composed of solvent monomers[31], or large hard particles or disks in a solvent of smaller hard particles[29, 30]. These methods are difficult to generalize to molecular systems of arbitrary molecular topology, and their computational performance is expected to be highly correlated with the type of system for which the move was originally developed. To address these issues, a generalized identity exchange move for simulations in the grand canonical ensemble, referred to as Molecular Exchange Monte Carlo (MEMC), is presented that works for systems of any molecular topology. Three different approaches for the insertion of the large molecule are presented. The result of this work has been published in Journal of Chemical Physics[36].

1.2 Thesis Organization

A derivation of acceptance criteria and the algorithms for performing the MEMC move is provided in chapter 2 for each of the three approaches. The simulation methodology is provided in chapter 3. The utility of the three methods and their computational efficiency is illustrated for selected binary mixtures in the chapter 4. The key findings of the work are summarized in chapter 5. The detailed computational procedure and mathematical calculations are included in the appendix A, and additional results are provided in the appendix B.

CHAPTER 2: METHODS

To describe the MEMC move in the grand canonical ensemble, it is helpful to consider the case of a large molecule that is exchanged with multiple smaller molecules. However, the methods may be applied without modification to the exchange of molecules of similar size. The original state is called the old state, while the state created by the attempted exchange move is called the new state. For a given configuration, with N_L large and N_S small molecules, an '*insertion move*" is an attempt to exchange one large molecule with N_{EX} small molecules inside a predefined exchange sub-volume V_{EX}, and a "*deletion move*" is an attempt to exchange N_{EX} small molecules for a large one. The exchange sub-volume is defined as an orthogonal box, where the length of the box in the x and y dimensions are set to the same values for simplicity and the z dimension is set independently. If desired, all three sub-volume box dimensions could be set independently. An orthogonal sub-volume is used instead of a cube or sphere to accommodate large molecules with different aspect ratios. Depending on the method used, the orientation of the exchange sub-volume z-axis may also be varied. Although not used in this work, it is also possible to optimize N_{EX} and V_{EX} "on the fly" during a simulation to maximize the acceptance rate.

The acceptance criterion for a molecular exchange move that satisfies the detailed balance equation is written as

$$\mathrm{K}(old \rightarrow new) = \mathrm{K}(new \rightarrow old) \tag{1}$$

where $\mathrm{K}(i \rightarrow j)$ is the flux of probability from state *i* to state *j*. The probability flux is equal to the product of the probability of finding the system in state *i*, the probability of generating a move that takes state *i* to state *j*, and the probability of accepting the move:

$$\mathrm{K}(old \rightarrow new) = \mathcal{N}(old) \times \alpha(old \rightarrow new) \times acc(old \rightarrow new) \tag{2}$$

Based on the detailed balance Eq. (1), the ratio of the probability of accepting the move from $old \rightarrow new$ to that of its reverse move $new \rightarrow old$ is:

$$\frac{acc(old \rightarrow new)}{acc(new \rightarrow old)} = \frac{\mathcal{N}(new)}{\mathcal{N}(old)} \times \frac{\alpha(new \rightarrow old)}{\alpha(old \rightarrow new)} \tag{3}$$

In the deletion move, where one large molecule is exchanged for N_{EX} small molecules, the ratio of the probability of being in the new configuration to the probability of being in the old configuration is

$$\frac{\mathcal{N}(new)}{\mathcal{N}(old)} = \frac{e^{-\beta U(new)} e^{\beta[\mu_L(N_L-1)+\mu_S(N_S+N_{EX})]}}{e^{-\beta U(old)} e^{\beta[N_L\mu_L+N_S\mu_S]}} = \frac{e^{\beta[N_{EX}\mu_S-\mu_L]}}{e^{\beta[U(new)-U(old)]}} \tag{4}$$

where $\beta = 1/k_BT$, μ_L and μ_S are the imposed chemical potentials of large and small molecules, respectively. $U(old)$ and $U(new)$ are the potential energies of the system in configuration old and configuration new, respectively.

For the insertion move, where N_{EX} small molecules are exchanged for one large molecule, the ratio of the probability of being in the new configuration to the probability of being in the old configuration is

$$\frac{\mathcal{N}(new)}{\mathcal{N}(old)} = \frac{e^{-\beta U(new)} e^{\beta[\mu_L(N_L+1)+\mu_S(N_S-N_{EX})]}}{e^{-\beta U(old)} e^{\beta[N_L\mu_L+N_S\mu_S]}} = \frac{e^{\beta[\mu_L-N_{EX}\mu_S]}}{e^{\beta[U(new)-U(old)]}} \tag{5}$$

The probability of generating the new state, for both insertion and deletion of the large molecule, is given by the product of the probability of locating the center of the exchange sub-volume at a particular point within the simulation box, the probability of choosing N_{EX} particular small molecules, the probability of choosing a particular large molecule, the probability of generating trial configurations for N_{EX} small molecules, and the probability of generating trial configurations for the large molecule,

$$\alpha(old \to new) = P_{sub-v}(old \to new) \times P_{pick-S}(old \to new) \times P_{pick-L}(old \to new) \times P_{pos-S}(old \to new) \times P_{pos-L}(old \to new) \quad (6)$$

Depending on how the center of the exchange sub-volume is located, the molecules to be exchanged are chosen, and how trial positions are generated, different algorithms to perform the MEMC move may be devised.

2.1 ME-1

For the large molecule insertion move, the exchange sub-volume V_{EX} with a random geometric center and a random orientation is defined within the simulation box. For a large molecule deletion move, the geometric center of V_{EX} is located at the geometric center (GC) of the selected large molecule and its z-axis is aligned with the backbone of the large molecule. See Figure 1 for more details.

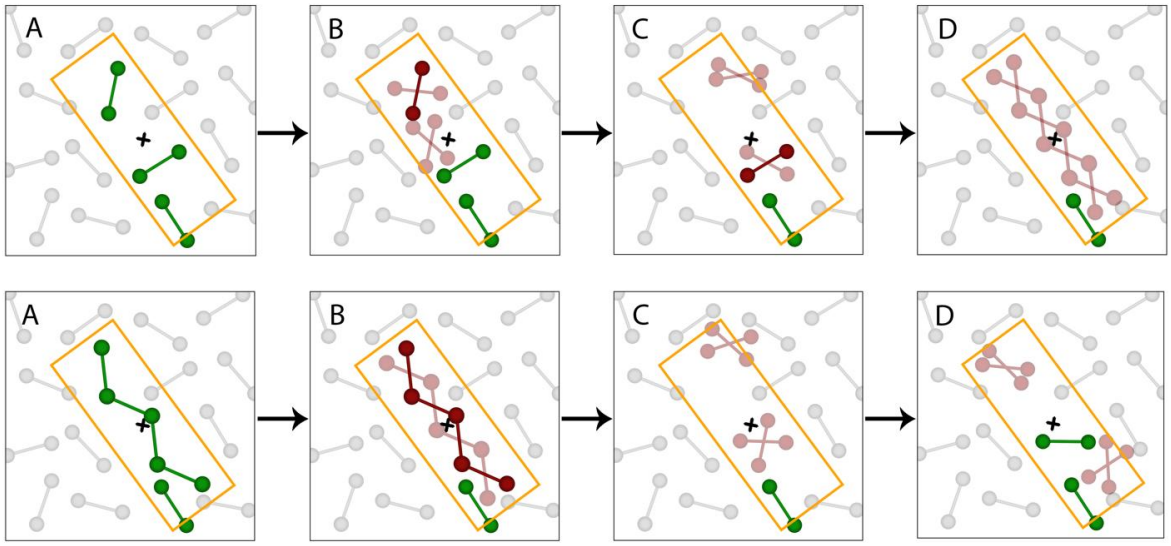

Figure 1: Schematic of the ME-1 algorithm. Selected or inserted molecule (green), trial position (light red), and actual position of the molecule (solid red). **Top row**, represents the exchange of two small molecules with one large molecule (insertion). The exchange sub-volume is defined as the orange box. (A) Identifying small molecules within the sub-volume with a random geometric center and orientation. (B) Generating CBMC trials (rotation and GC location) for one of the small molecules and then removing it. (C) Generating CBMC trials (rotation and GC location) for the second small molecule and then removing it. (D) Aligning the backbone of the large molecule with the sub-volume and performing CBMC rotational trials around its backbone. **Bottom row**, represents the exchange of a large molecule (deletion) with two small molecules. (A) Aligning the sub-volume with large molecule's backbone with geometric center placed at GC of the large molecule, and identifying the small molecules within the sub-volume. (B) Generating CBMC rotational trials around large molecule backbone and then removing it. (C) Generating CBMC trials (rotation and GC location) for the first small molecule and then inserting it into the sub-volume. (D) Generating CBMC trials (rotation and GC location) for the second small molecule and then inserting it into the sub-volume.

The algorithm for the insertion of a large molecule after deletion of small molecule(s) is as follows:

1. Define an orthogonal exchange sub-volume V_{EX}, with its geometric center located randomly within the simulation box of volume V (with the probability proportional to V^{-1}) and a random orientation. Determine the total number of small molecules within the exchange sub-volume ($N_{S,VEX}$) based on their geometric center.

2. Reject move if $N_{S,VEX} < N_{EX}$, otherwise continue.

3. Select N_{EX} small molecules out of $N_{S,VEX}$ found in the exchange sub-volume with the probability of $N_{EX}!\left(N_{S,VEX}-N_{EX}\right)!/N_{S,VEX}!$.

4. Repeat steps a and b for N_{EX} cycles ($i=1,2,\dots,N_{EX}$) to delete the selected small molecules.

   a. Generate $j-1$ random trial positions for the GC of the i^{th} small molecule within the exchange sub-volume V_{EX}. The original position of the GC of the i^{th} small molecule will be included as the j^{th} term.

   b. For each trial GC position p, generate k random trial orientations around the molecule's GC (except the j^{th} GC, where $k-1$ random trial orientations are generated and the original orientation of the molecule will be included as the k^{th} term) and calculate the Rosenbluth weight $W_{i,old}=\sum_{p=1}^{j}\sum_{r=1}^{k}exp\left(-\beta U_{i,p,r}\right)$, were $U_{i,p,r}$ is the interaction energy of the i^{th} molecule to be removed in position p and orientation r with all other molecules, excluding those removed in the earlier cycles of the move. Finally, remove the molecule from the simulation box. Calculate $P_{i,old}=\frac{exp\left(-\beta U_{i,j,k}\right)}{W_{i,old}}$, were $U_{i,j,k}$ is the interaction energy of the i^{th} small molecule at its original GC position and orientation with all other molecules remaining in the simulation box. $P_{i,old}$ is the probability of inserting the i^{th} small molecule back in its original configuration in the reverse move ($new \rightarrow old$).

5. Insert the GC of the large molecule at the geometric center of the exchange sub-volume V_{EX} and align the backbone of the large molecule with the z-axis of the exchange sub-volume. Generate k random trial orientations for the large molecule around its backbone (two-dimensional rotation). Calculate the Rosenbluth weight $W_{new}=\sum_{r=1}^{k}exp(-\beta U_r)$,

where U_r is the interaction energy of the inserted large molecule at orientation r with all other molecules in the simulation box.

6. Select one of the generated trial configurations with the probability $P_{new} = \frac{exp(-\beta U_r)}{W_{new}}$ and insert the large molecule.

The algorithm for the deletion of a large molecule and subsequent insertion of small molecule(s) is as follows:

1. Select a large molecule out of N_L large molecules within the simulation box with probability of $1/N_L$.
2. Define an orthogonal exchange sub-volume with its geometric center placed at the GC of the selected large molecule, and its z-axis aligned with the backbone of the large molecule. Determine the number of small molecules $N_{S,VEX}$ within the exchange sub-volume.
3. Generate $k-1$ random trial orientations around the large molecule's backbone. The original orientation will be included as the k^{th} term in the Rosenbluth weight. The Rosenbluth weight is calculated as $W_{old} = \sum_{r=1}^{k} exp(-\beta U_r)$, where U_r is the interaction energy of the large molecule in orientation r with all other molecules in the simulation box. Calculate the probability $P_{old} = \frac{exp(-\beta U_k)}{W_{old}}$, where U_k is the interaction energy of the large molecule at the original orientation with all other molecules in the simulation box. P_{old} is the probability of inserting the large molecule at its original configuration in the reverse move ($new \rightarrow old$). Then remove the large molecule from the simulation box.
4. Repeat the steps a→c for N_{EX} cycles ($i = 1, 2, \dots, N_{EX}$) to insert the small molecules with the probability of $N_{EX}!/V_{EX}^{N_{EX}}$.
   a. Generate j random trial positions for the GC of the i^{th} small molecule within V_{EX}.

b. For each trial position p, generate k random trial orientations around the molecule's GC (three-dimensional rotation) and calculate the Rosenbluth weight $W_{i,new} = \sum_{p=1}^{j}\sum_{r=1}^{k} exp(-\beta U_{i,p,r})$, where $U_{i,p,r}$ is the interaction energy of the i^{th} inserted small molecule at position p and orientation r with all the other molecules, including those added in the earlier cycles of the move.

c. Pick one of the generated trial configurations with probability $P_{i,new} = \frac{exp(-\beta U_{i,p,r})}{W_{i,new}}$ and insert the small molecule.

Based on the two algorithms described above, for the large molecule insertion, the ratio of the probability of generating the move $new\ (N_L+1, N_S-N_{EX}) \rightarrow old\ (N_L, N_S)$ to that of the reverse move is:

$$\frac{\alpha(new \rightarrow old)}{\alpha(old \rightarrow new)} = \frac{\frac{1}{N_L+1}}{\frac{1}{V}} \times \frac{\frac{N_{EX}!}{V_{EX}^{N_{EX}}}}{\frac{N_{EX}!\,(N_{S,VEX}-N_{EX})!}{N_{S,VEX}!}} \times \frac{\prod_{i=1}^{N_{EX}} P_{i,old}}{P_{new}} \tag{7}$$

Simplifying Eq. 7 and substituting into Eq. 3, produces the acceptance criteria for the large molecule insertion.

$$acc(old \rightarrow new) = min\left\{1\,, \frac{V}{N_L+1} \times \frac{N_{S,VEX}!}{V_{EX}^{N_{EX}}(N_{S,VEX}-N_{EX})!} \times \frac{W_{new}}{\prod_{i=1}^{N_{EX}} W_{i,old}} \times e^{\beta[\mu_L - N_{EX}\mu_S]}\right\} \tag{8}$$

For the large molecule deletion move, the ratio of the probability of generating the move $new\ (N_L-1, N_S+N_{EX}) \rightarrow old\ (N_L, N_S)$ to that of the reverse move is:

$$\frac{\alpha(new \rightarrow old)}{\alpha(old \rightarrow new)} = \frac{\frac{1}{V}}{\frac{1}{N_L}} \times \frac{\frac{N_{EX}!\, N_{S,VEX}!}{(N_{S,VEX}+N_{EX})!}}{\frac{N_{EX}!}{V_{EX}^{N_{EX}}}} \times \frac{\prod_{i=1}^{N_{EX}} P_{i,new}}{P_{old}} \tag{9}$$

Simplifying Eq. 9 and substituting into Eq. 3, produces the acceptance criteria for the large molecule deletion move.

$$acc(old \rightarrow new) = min\left\{1\,, \frac{N_L}{V} \times \frac{V_{EX}^{N_{EX}} \times N_{S,VEX}!}{(N_{S,VEX}+N_{EX})!} \times \frac{\prod_{i=1}^{N_{EX}} W_{i,new}}{W_{old}} \times e^{\beta[N_{EX}\mu_S - \mu_L]}\right\} \tag{10}$$

The energy difference between configuration new and old, $U(new) - U(old)$, does not appear directly in the acceptance criteria because their Boltzmann weight is already included in the probabilities used for selecting the position of the molecules.

The acceptance criterion derived for ME-1 is identical to the one introduced by Vink and Horbach[30]. This move performs well for binary mixtures with low concentrations of large molecules. However, the acceptance rate of the move decreases significantly as the concentration of large molecules increases, and the chance of finding N_{EX} small molecules in the exchange sub-volume becomes very low. To address this limitation, ME-2 was developed.

2.2 ME-2

In ME-1, for the insertion of a large molecule, the exchange sub-volume V_{EX} is defined with a random orientation and position. However, as the mole fraction of small molecules decreases, the required number of small molecules are frequently not available within the exchange sub-volume. Therefore, a large fraction of the attempted ME-1 moves will be rejected. In the ME-2 approach, the geometric center of V_{EX} is placed on the GC of a randomly selected small molecule. If the small molecule is monoatomic, the orientation of V_{EX} is assigned randomly, otherwise its z-axis is aligned with the backbone of the small molecule. The large

molecule deletion is identical to ME-1. An illustration of the ME-2 algorithm is provided in Figure 2.

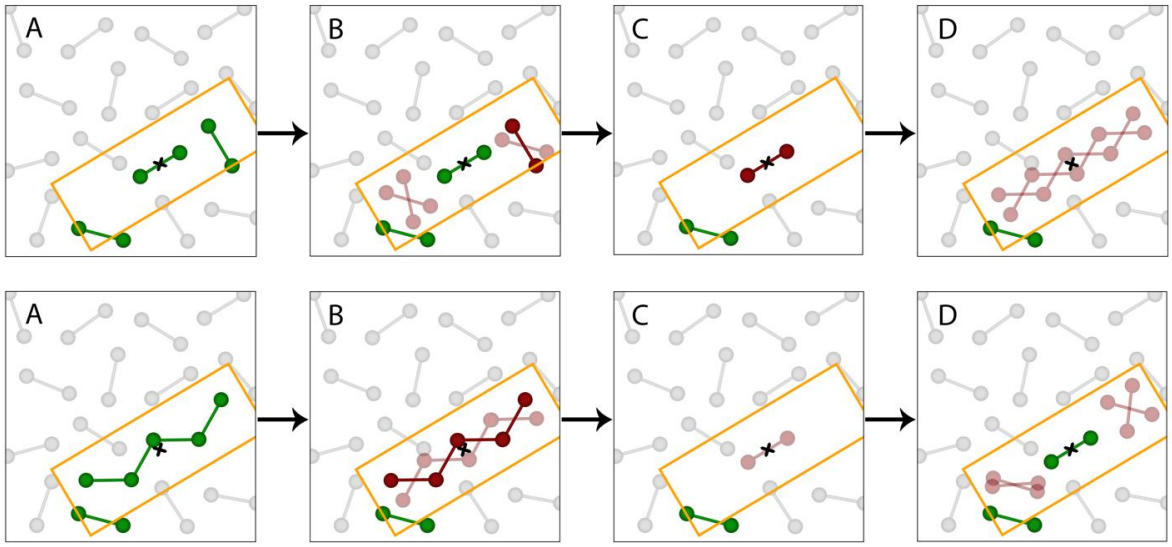

Figure 2: Schematic of the ME-2 algorithm. Selected or inserted molecule (green), trial position (light red), and actual position of the molecule (solid red). **Top row**, represents the exchange of two small molecules with one large molecule (insertion). The sub-volume is defined as the orange box. (A) Aligning the sub-volume with a randomly selected small molecule's backbone with geometric center placed at GC of the selected small molecule, and identifying the small molecules within the sub-volume. (B) Generating CBMC trials (rotation and GC location) for one of the small molecules and then removing it. (C) Generating CBMC rotational trials around selected small molecule and then removing it. (D) Aligning the backbone of the large molecule with the sub-volume and performing CBMC rotational trials around its backbone. **Bottom row** represents the exchange of one large molecule with two small molecules (deletion). (A) Aligning the sub-volume with large molecule's backbone with geometric center placed at GC of the large molecule, and identifying the small molecules within the sub-volume. (B) Generating CBMC rotational trials around large molecule backbone and then removing it. (C) Placing the GC of the first small molecule at the geometric center of the sub-volume and generate the CBMC rotational trials around its backbone and then inserting it into the sub-volume. (D) Generating CBMC trials (rotation and GC location) for the second small molecule and then inserting it into the sub-volume.

The algorithm for the insertion of a large molecule after deletion of small molecule(s) is as follows:

1. Select one molecule out of N_S small molecules in the simulation box with the probability of $1/N_S$. This molecule will be the last molecule to be removed from the system.

2. Define V_{EX} with its geometric center placed at the GC of the small molecule selected in step 1. The z-axis of the exchange sub-volume is aligned with the backbone of the small molecule. If the small molecule is monoatomic, the orientation of V_{EX} is assigned randomly. Determine the number of small molecules $N_{S,VEX}$ within V_{EX} ($N_{S,VEX}$ includes the molecule selected in step 1).
3. Reject the move if $N_{S,VEX} < N_{EX}$, otherwise continue.
4. Select $N_{EX} - 1$ small molecules out of $N_{S,VEX} - 1$, with probability $(N_{EX} - 1)!\,(N_{S,VEX} - N_{EX})!/(N_{S,VEX} - 1)!$.
5. Repeat steps a and b of the large molecule insertion move of ME-1 for $\mathrm{N_{EX}} - 1$ cycles ($\mathrm{i} = 1, 2, \ldots, \mathrm{N_{EX}} - 1$) to delete the selected small molecules.
6. For the last small molecule to be deleted, generate $k - 1$ random trial orientations around its backbone. If the small molecule is monoatomic, orientations are generated around its GC. The original orientation will be included as the k^{th} term in the Rosenbluth weight. The Rosenbluth weight is calculated from $W_{N_{EX},old} = \sum_{r=1}^{k} exp(-\beta U_{N_{EX},r})$, where $U_{N_{EX},r}$ is the interaction energy of the last small molecule in orientation r with all other molecules in the simulation box. Finally, remove the last small molecule from the simulation box and calculate $P_{N_{EX},old} = \frac{exp(-\beta U_{N_{EX},k})}{W_{N_{EX},old}}$, where $U_{N_{EX},k}$ is the interaction energy of the last small molecule at its original configuration with all other molecules remaining in the simulation box. $P_{i,old}$ is the probability of inserting the i^{th} small molecule back at its original configuration in the reverse move ($new \rightarrow old$).
7. Insert the large molecule according to steps 5 and 6 of ME-1.

The algorithm for the deletion of a large molecule and subsequent insertion of small molecule(s) is as follows:

1. Follow steps 1-4 of the ME-1 large molecule deletion move.
2. Insert the GC of the first small molecule at the geometric center of V_{EX} and align its backbone with the z-axis of the exchange sub-volume. Generate k random trial orientations around its backbone. If small molecules are monoatomic, the orientation is assigned randomly around its GC. Calculate the Rosenbluth weight $W_{1,new} = \sum_{r=1}^{k} exp(-\beta U_{1,r})$, where $U_{1,r}$ is the interaction energy of the first small molecule inserted at orientation r with all other molecules in the simulation box.
3. Select one of the trial orientations with the probability $P_{1,new} = \frac{exp(-\beta U_{1,r})}{W_{1,new}}$.
4. Repeat steps a→c of the large molecule deletion move of ME-1 for $N_{EX} - 1$ cycles ($i = 2, \dots, N_{EX}$) to insert the small molecules with probability $(N_{EX} - 1)!/V_{EX}^{(N_{EX}-1)}$.

Based on the two algorithms described above, for the large molecule insertion move, the ratio of the probability of generating move $new\ (N_L + 1, N_S - N_{EX}) \rightarrow old\ (N_L, N_S)$ to that of the reverse move is:

$$\frac{\alpha(new \rightarrow old)}{\alpha(old \rightarrow new)} = \frac{\frac{1}{N_L + 1}}{\frac{1}{N_S}} \times \frac{\frac{(N_{EX} - 1)!}{V_{EX}^{(N_{EX}-1)}}}{\frac{(N_{EX} - 1)!\,(N_{S,VEX} - N_{EX})!}{(N_{S,VEX} - 1)!}} \times \frac{\prod_{i=1}^{N_{EX}} P_{i,old}}{P_{new}} \tag{11}$$

Simplifying Eq. 11 and substituting into Eq. 3 results in the acceptance criterion for the large molecule insertion move:

$$acc(old \rightarrow new) = min\left\{1\,, \frac{N_S}{N_L + 1} \times \frac{(N_{S,VEX} - 1)!}{V_{EX}^{(N_{EX}-1)}(N_{S,VEX} - N_{EX})!} \times \frac{W_{new}}{\prod_{i=1}^{N_{EX}} W_{i,old}} \times e^{\beta[\mu_L - N_{EX}\mu_S]}\right\} \tag{12}$$

For the large molecule deletion move, the ratio of the probability of generating configuration $new\ (N_L - 1, N_S + N_{EX}) \rightarrow old\ (N_L, N_S)$ to that of the reverse move is:

$$\frac{\alpha(new \rightarrow old)}{\alpha(old \rightarrow new)} = \frac{\frac{1}{(N_S + N_{EX})}}{\frac{1}{N_L}} \times \frac{\frac{(N_{EX} - 1)!\, N_{S,VEX}!}{(N_{S,VEX} + N_{EX} - 1)!}}{\frac{(N_{EX} - 1)!}{V_{EX}^{(N_{EX}-1)}}} \times \frac{\prod_{i=1}^{N_{EX}} P_{i,new}}{P_{old}} \tag{13}$$

Simplifying Eq. 13 and substituting into Eq. 3 results in the acceptance criterion for the large molecule deletion move.

$$acc(old \rightarrow new) = min\left\{1\ , \frac{N_L}{(N_S + N_{EX})} \times \frac{V_{EX}^{(N_{EX}-1)} \times N_{S,VEX}!}{(N_{S,VEX} + N_{EX} - 1)!} \times \frac{\prod_{i=1}^{N_{EX}} W_{i,new}}{W_{old}} \times e^{\beta[N_{EX}\mu_S - \mu_L]}\right\} \tag{14}$$

If $N_{EX} = 1$, the acceptance criteria given in Eqs. 13 and 14 simplifies to that of the standard identity-exchange acceptance move[26].

$$acc(\ N_L \rightarrow\ N_L + 1) = min\left\{1\ , \frac{N_S}{(N_L + 1)} \times \frac{W_{new}}{W_{old}} \times e^{\beta[\mu_L - \mu_S]}\right\} \tag{15}$$

$$acc(\ N_L \rightarrow\ N_L - 1) = min\left\{1\ , \frac{N_L}{(N_S + 1)} \times \frac{W_{new}}{W_{old}} \times e^{\beta[\mu_S - \mu_L]}\right\} \tag{16}$$

2.3 ME-3

For the large molecule insertion move in ME-2, the large molecule is inserted as a rigid body and its backbone is aligned with the z-axis of the V_{EX}. This move performs well for large molecules with a straight backbone. However, the acceptance rate decreases for a large molecule with nonlinear geometry as it becomes significantly more difficult to fit a complex

rigid body into the void space created after deleting the small molecule(s). Therefore, a modification to ME-2 was developed to address this limitation.

In the ME-3 algorithm, a predefined atom of the large molecule is first placed at the geometric center of V_{EX} and the molecule is built segment by segment using the coupled-decoupled configurational-bias Monte Carlo (CBMC) algorithm[14]. For the large molecule deletion move, the exchange sub-volume is defined with a random orientation, with its geometric center placed at the same predefined atom of the large molecule to be deleted. Next, the Rosenbluth weight W_{old} of the large molecule is calculated. Insertion and deletion of N_{EX} small molecules are identical to the ME-2 method. Figure 3 illustrates the ME-3 algorithm.

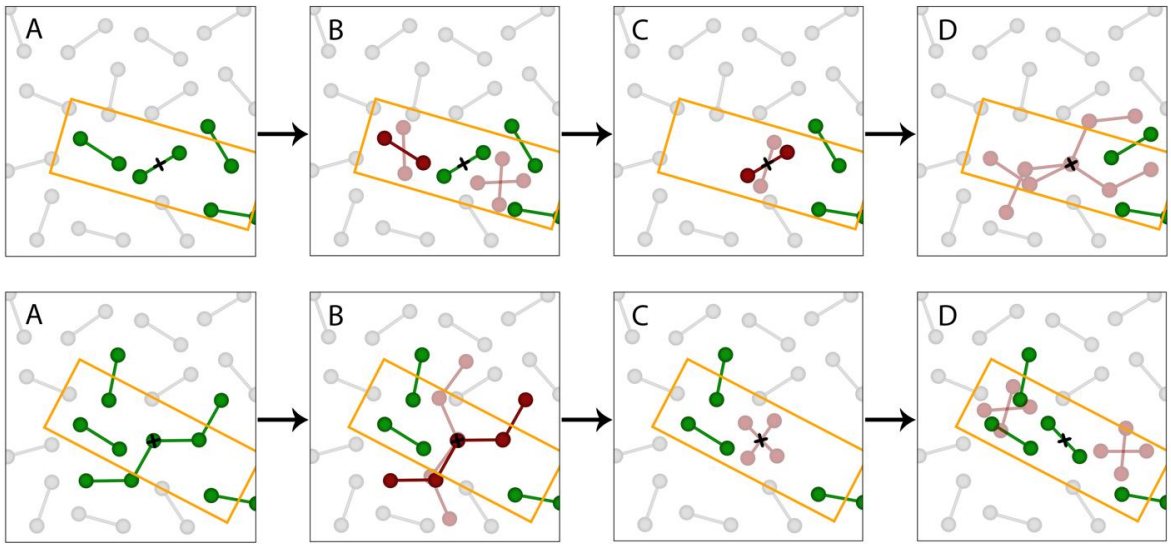

Figure 3: Schematic of the ME-3. Selected or inserted molecule (green), trial position (light red), and actual position of the molecule (solid red). **Top row**, represents the exchange of two small molecules with one large molecule (insertion). The sub-volume is defined as the orange box. (A) Defining the sub-volume with a random orientation, where its geometric center is placed at a randomly selected small molecule's GC, and identifying the small molecules within the sub-volume. (B) Generating CBMC trials (rotation and GC location) for one of the small molecules and then removing it. (C) Generating CBMC rotational trials around its GC of the selected small molecule and then removing it. (D) Placing the predefined atom of the large molecule at the geometric center of the sub-volume and growing the large molecule using coupled-decoupled CBMC technique. **Bottom row**, represents the exchange of a large molecule with two small molecules (deletion). (A) Defining the sub-volume with a random orientation with geometric center placed at the predefined atom of the large molecule, and identifying the small molecules within the sub-volume. (B) Generating coupled-decoupled CBMC trials and then removing it. (C) Placing the GC of the first small molecule at the geometric center of the sub-volume, generating CBMC rotational trials around its GC and then

inserting it into the sub-volume. (D) Generating CBMC trials (rotation and GC location) for the second small molecule and then inserting it into the sub-volume.

The algorithm for the insertion of a large molecule after deletion of small molecule(s) is as follows:

1. Select one molecule out of N_S small molecules in the simulation box with probability $1/N_S$. This molecule will be the last molecule to be removed from the system.
2. Define an orthogonal exchange sub-volume V_{EX} with a random orientation and its geometric center placed at the GC of the small molecule selected above. Then determine the number of small molecules $N_{S,VEX}$ within V_{EX} ($N_{S,VEX}$ includes the molecule selected in step 1).
3. Repeat steps 3-6 of the ME-2 method to delete N_{EX} small molecules from simulation box.
4. Insert the predefined atom of the large molecule at the center of V_{EX} and perform coupled-decoupled configurational-bias Monte Carlo to grow the large molecule segment by segment. Calculate the Rosenbluth weight W_{new}.
5. Insert the large molecule by selecting one of the generated trial configurations with the probability P_{new}.

The algorithm for the deletion of a large molecule and subsequent insertion of small molecule(s) is as follows:

1. Within the simulation box of volume V, pick one large molecule out of N_L with probability of $1/N_L$.
2. Define an orthogonal exchange sub-volume V_{EX} with a random orientation and place its geometric center at the predefined atom of the selected large molecule. Determine the number small molecules $N_{S,VEX}$ within the exchange sub-volume.

3. Perform coupled-decoupled CBMC for the large molecule and calculate the Rosenbluth weight W_{old} and P_{old}.
4. Repeat steps 2-4 of ME-2 to insert N_{EX} small molecules within V_{EX}.

The forward to reverse probability ratios for generating the large molecule insertion and the large molecule deletion moves are identical to those given in Eq. 11 and 13, respectively. The acceptance criteria for the ME-3 algorithm is identical to that of ME-2 and are given by Eq. 12 and 14.

CHAPTER 3: SIMULATION METHODOLOGY

The three molecular exchange algorithms described in Chapter 2 were implemented in the development version of GPU Optimized Monte Carlo[37] (GOMC), which is available to the public on GitHub[38]. GOMC is an object-oriented Monte Carlo simulation engine, capable of performing simulations in canonical, isobaric-isothermal, and grand canonical ensembles, as well as Gibbs ensemble Monte Carlo. GOMC is designed for the simulation of complex molecular topologies and supports a variety of potential functions, such as Lennard-Jones and Mie potentials. Coulomb interactions are also supported via the Ewald summation method[39]. GOMC is capable of parallel computation, either on multicore CPUs or GPUs.

Phase diagrams were determined from histogram-reweighting Monte Carlo simulations in the grand-canonical ensemble[38]. A cubic box size of 25 Å × 25 Å × 25 Å was used for methane+ethane, methane+propane, methane+n-butane, and water+impurity. For perfluorobutane+n-butane and methane+n-pentane, a box size of 30 Å × 30 Å × 30 Å was used, while for 2,2,4-trimethylpentane+neopentane a box size of 40 Å × 40 Å × 40 Å was used. Initial configurations were generated with Packmol[40]. Psfgen was used to generate coordinate (*.pdb) and connectivity (*.psf) files[41]. Potentials were truncated at 10 Å and analytical tail corrections were applied[42]. To enhance the acceptance rate for molecule insertions, the coupled-decoupled configurational-bias Monte Carlo (CBMC) algorithm was used[14]. For all liquid phase simulations, unless otherwise noted in the Results and Discussion, configurational-bias parameters were: 100 angle trials, 100 dihedral trials, 10 trial locations for the first site, and 8 trial locations for secondary sites. For standard GCMC simulations, a move ratio of 20% displacements, 10% rotations, 10% regrowth, and 60% molecule transfers was used. For simulations that included the molecular exchange move, 30% molecular exchanges were performed with a corresponding reduction in the percentage of attempted molecule transfers.

Uncertainties used in the calculation of the statistical efficiency of the methods were calculated as the standard deviation determined from five unique simulation trajectories, each started from a unique initial configuration and random number seed. All simulations, except those used to generate phase diagrams, were run for $2x10^7$ Monte Carlo steps (MCS), without equilibration period. Simulations used to generate phase diagrams were run for $5x10^7$ MCS with a $5x10^6$ MCS equilibration period. Every 200-500 MCS, the instantaneous state of the system (N_1, N_2, E) was saved as a histogram. Every one million MCS, the natural log of distribution of large particle $\ln(P_N)$ for each simulation was determined, and the standard deviation and efficiency were calculated for each binary system for a variety of compositions along the vapor-liquid coexistence curve. Calculations were performed on one core of an Intel Xeon E5-4627v4 2.6 GHz CPU.

The efficiency was computed using the calculated standard deviation and the CPU time.

$$\eta = (\sigma^2 s)^{-1} \tag{17}$$

where σ is average uncertainty in natural log of large particle distribution and s is the CPU time in seconds.

CHAPTER 4: RESULTS AND DISCUSSION

In this chapter, a number of examples are provided to illustrate the effect of molecular exchange moves on the statistical sampling in grand canonical histogram reweighting Monte Carlo simulations. Mixtures simulated include perfluorobutane+n-butane, and methane +ethane, +propane, +n-butane, and +n-pentane. Additional calculations were performed to generate pure fluid phase diagrams for water and 2,2,4-trimethylpentane to demonstrate the utility of the method and to provide comparisons to prior work[43-45]. For binary mixture phase diagrams, all calculations were performed at temperatures below $0.7T_c$. For pure fluid phase diagrams, calculations were performed from the critical temperature to $0.44T_c - 0.51T_c$. Performing grand canonical Monte Carlo simulations, using standard configurational-bias methods[14], below $0.7T_c$ is a challenging task, and therefore a good test to evaluate the improvement in sampling of phase space provided by the proposed algorithms.

4.1 Methane+n-alkane

Methane+n-alkane systems are well studied and extensive experimental data may be found in the literature[46-53]. In general, the determination of vapor-liquid coexistence for these systems at temperatures above $0.7T_c$ can be done using standard configurational-bias methods in grand canonical or Gibbs ensemble Monte Carlo simulations[9, 54-56]. However, below $0.7T_c$, acceptance rates for the insertion of n-alkanes into a liquid phase drops to approximately 0.1%, which necessitates long simulations to obtain convergence of the simulations. In this section, the effect of the three ME algorithms on the convergence of grand canonical Monte Carlo simulations is assessed for mixtures of methane +ethane, +propane, +n-butane, and +n-pentane, and the effectiveness of performing a two for one exchange is evaluated.

The methane+n-butane mixture is presented first as an example of the validation process used in the development of the molecular exchange methods. Grand canonical Monte

Carlo (GCMC) simulations were performed for a variety of temperatures, chemical potentials, and move ratios using both standard configurational-bias insertions/deletions and the ME-1, ME-2, and ME-3 methods. Probability distributions of states sampled during the simulation were collected and compared to reference distributions determined using standard configuration-bias insertions. An example of this is shown in Figure 4, for gas $(\mu_{butane} = -2960, \mu_{methane} = -2000)$ and liquid $(\mu_{butane} = -2840, \mu_{methane} = -2000)$ phase simulations at 277 K. As expected, the probability distributions produced by the ME-3 algorithm are an exact match to the reference distributions. Additional data for the ME-1 and ME-2 algorithms are presented in the Appendix B, Figures S1 and S2.

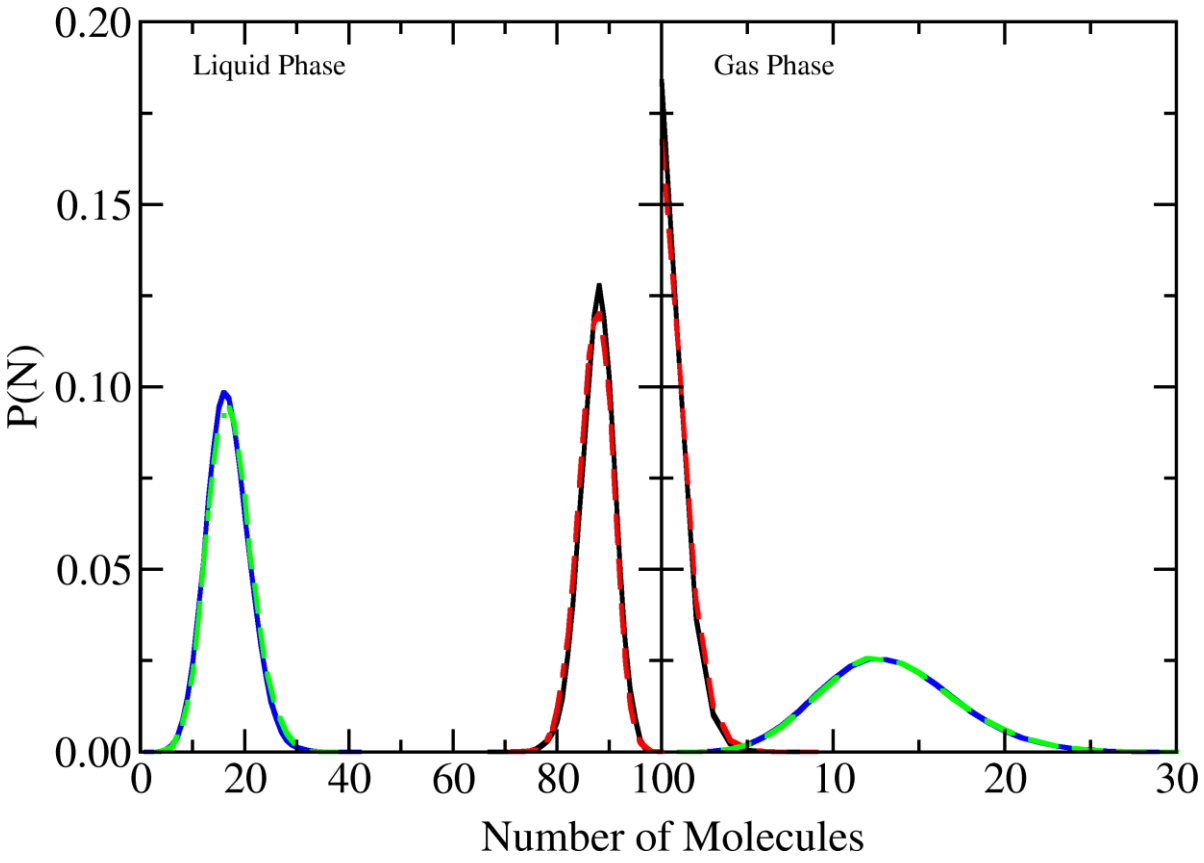

Figure 4: Probability distributions predicted from gas ($\mu_{butane} = -2960, \mu_{methane} = -2000$) and liquid ($\mu_{butane} = -2840, \mu_{methane} = -2000$) phase GCMC simulations of methane+n-butane at 277 K. Solid lines denote the probability distributions for n-butane (black) and methane (blue) using standard configurational-bias insertions and deletions. Dashed lines denote the probability distributions for n-butane (red) and methane (green) using the ME-3 algorithm.

In Figure 5, the pressure vs. composition diagram for methane+n-butane at 277 K, predicted using both the coupled-decoupled configurational-bias method[14] and the ME-3 algorithm, is shown. Interactions between molecules were described with Optimized Mie Potentials for Phase Equilibria[54]. In addition to showing excellent agreement with experimental data[53], the ME-3 algorithm produced results that are statistically indistinguishable from standard configurational-bias insertions, providing further validation of the method.

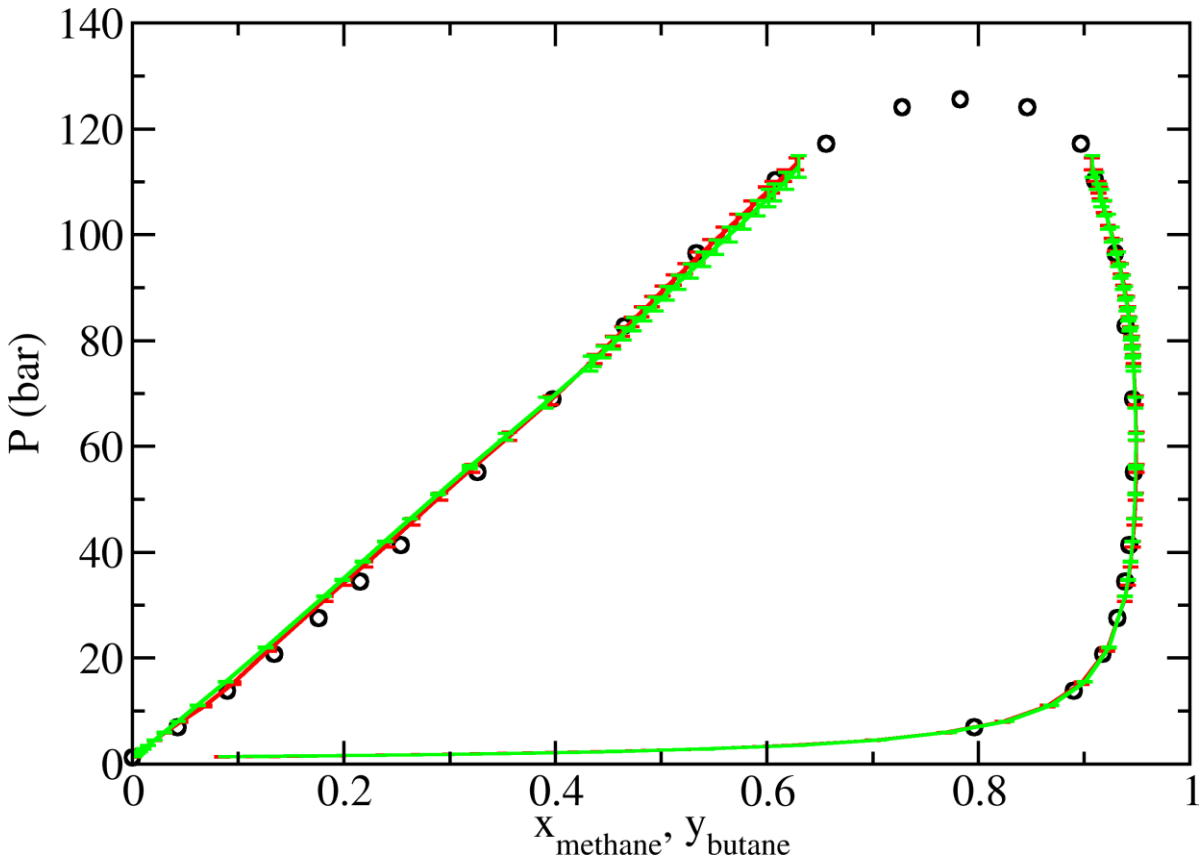

Figure 5: Pressure composition diagram for methane+n-butane at 277 K predicted from GCMC+histogram reweighting simulations using Mie potentials[54]. Experimental data (circles)[53], standard configurational-bias insertions (red lines), ME-3 algorithm (green lines).

In Table 1, the acceptance rate for molecule transfers as a function of composition is presented for each methane+n-alkane binary mixture. Calculations were performed for liquid phase simulations along the coexistence curve at 186 K (methane+ethane), 213 K (methane+propane), 225 K (methane+n-butane), and 273 K (methane+n-pentane). The systems exhibit similar general trends, with acceptance rates climbing as the critical point of the mixture is reached. For $x_{methane} < 0.5$, acceptances rates for the insertion of the larger n-alkane using configurational-bias were less than 1%. When performing a one to one exchange, ME-3 was found to produce the largest improvement in acceptance rates for the large molecule, producing improvements of 2X for methane+n-pentane at $x_{methane} = 0.7$ to 70X for methane+ethane at $x_{methane} = 0.1$. The ME-2 algorithm also produced significant enhancement in the

acceptances rates for the insertion of the longer n-alkane, while the ME-1 algorithm was found to yield significantly lower acceptance rates than traditional configurational-bias insertions. Because the ME-2 algorithm uses a rigid swap and the GC of the large molecule is placed at the geometric center of the exchange sub-volume, only a fraction of the sub-volume is guaranteed to be empty. In most of the ME-2 exchanges, it is likely that some atoms from the large molecule will overlap with existing molecules, lowering acceptance rates compared to ME-3. The ME-3 algorithm uses the same initial placement for the central atom as ME-2, but grows the rest of the large molecule, allowing it to find more energetically favorable configurations than are possible through a rigid molecule insertion, leading to greater acceptance rates for the exchange move. As expected, the more similar the large and small molecule were in terms of excluded volume, the greater the success of the molecular exchange. It is also interesting to note that even for the highly asymmetric mixture of methane+n-pentane, acceptance rates for molecule transfers were improved substantially through the inclusion of the molecular exchange move.

The molecular exchange algorithm allows for trial moves where any number of small molecules may be exchanged for one large molecule. An example of this is shown in Table 1, where acceptance rates are presented for exchange of two methanes with one n-butane or n-pentane ($N_{EX} = 2$). For the ME-3 algorithm, acceptance rates are always lower than the one for one exchange, although, this difference decreases as the chain length of the large molecule increases. Part of the decrease in the acceptance rate stems from the reduced probability of finding two methane molecules in the sub-volume to exchange at low methane concentrations. For ME-2, acceptance rates are slightly lower for the exchange of two methanes with one n-butane, compared to the one for one exchange. However, for the exchange of two methanes with one n-pentane, slight improvements in the acceptance rates were observed. The ME-1 algorithm shows a slight improvement in acceptance rates for the exchange of two methanes

with one n-butane or n-pentane, although in all cases, acceptance rates for the ME-1 algorithm are lower than configurational-bias insertions.

While size of the sub-volume does not have an effect on the acceptance rates for the ME-2 and ME-3 algorithms for a one to one exchange, it was found to have an effect on the acceptance rates for the two to one exchange, as shown in Table 1. Increasing the size of the sub-volume increases the probability that a second small molecule will be found within the sub-volume, leading to an increased overall acceptance rate for the MEMC move. Therefore, it is possible to optimize acceptance rates for the two to one exchange ratio by performing a series of short simulations for a range of sub-volume box lengths, and by using a heuristic that the sub-volume should be large enough to contain the entire large molecule. For methane+n-butane, the optimum exchange sub-volume size for a two for one exchange was found to be 8.8 Å × 8.8 Å × 11.8 Å for ME-3 and ME-2, while for ME-1 it was 5 Å × 5 Å × 8 Å.

Table 1: n-alkane insertion/removal acceptance percentages in GCMC liquid phase simulations of methane+n-alkane mixtures for CBMC, ME-1, ME-2, and ME-3 methods.

Binary system	Sub-volume size (Å)	N_{EX}	x_{CH_4}	CBMC	ME-1	ME-2	ME-3
methane+ethane	5 × 5 × 6	1	0.1	0.33	0.11	11.68	23.62
			0.5	1.47	0.96	16.20	33.33
			0.9	8.3	4.18	24.09	47.84
methane+propane	5 × 5 × 7	1	0.1	0.08	0.05	3.42	4.13
			0.4	0.38	0.40	5.67	7.21
			0.8	5.18	3.36	13.56	18.36
methane+n-butane	5 × 5 × 8	1	0.1	0.14	0.025	0.835	2.373
			0.3	0.33	0.099	1.207	3.421
			0.6	2.52	0.948	3.378	8.128
	5 × 5 × 8	2	0.1	0.14	0.019	0.196	0.362
			0.3	0.33	0.144	0.557	0.928
			0.6	2.52	1.262	2.288	3.160
	8.8 × 8.8 × 11.8	2	0.1	0.14	0.022	0.398	0.984
			0.3	0.33	0.086	0.821	1.860
			0.6	2.52	0.621	2.682	5.252
methane+n-pentane	5 × 5 × 9	1	0.1	0.064	0.007	0.209	0.824
			0.5	0.397	0.116	0.638	2.163
			0.7	2.461	0.666	1.72	4.814
	5 × 5 × 9	2	0.1	0.639	0.006	0.086	0.189
			0.5	0.397	0.270	0.736	1.160
			0.7	2.461	1.332	2.389	3.170
	8.8 × 8.8 × 13	2	0.1	0.639	0.008	0.145	0.455
			0.5	0.397	0.102	0.675	1.806
			0.7	2.461	0.473	2.054	4.133

A more detailed analysis of the statistical uncertainty and efficiency for an exchange ratio of one to one is provided in Figure 6 for the methane+n-butane mixture. A direct

comparison between the efficiencies obtained for the one to one and one to two exchange ratios are presented in the Appendix B, Figure S3. Uncertainties were determined from probability distributions collected from liquid phase grand canonical Monte Carlo simulations performed along the vapor-liquid coexistence curve. For all mole fractions investigated, the ME-3 algorithm shows the fastest convergence of the n-particle probability distribution, converging in approximately half the number of Monte Carlo steps of ME-2. Both the ME-3 and ME-2 algorithms produce similarly converged probability distributions after $2x10^7$ MCS, with average uncertainties of approximately 0.05. The ME-1 algorithm and configurational-bias insertions show similar convergence properties. However, with $2x10^7$ MCS each produced uncertainties that were approximately double those of the ME-3 and ME-2 methods.

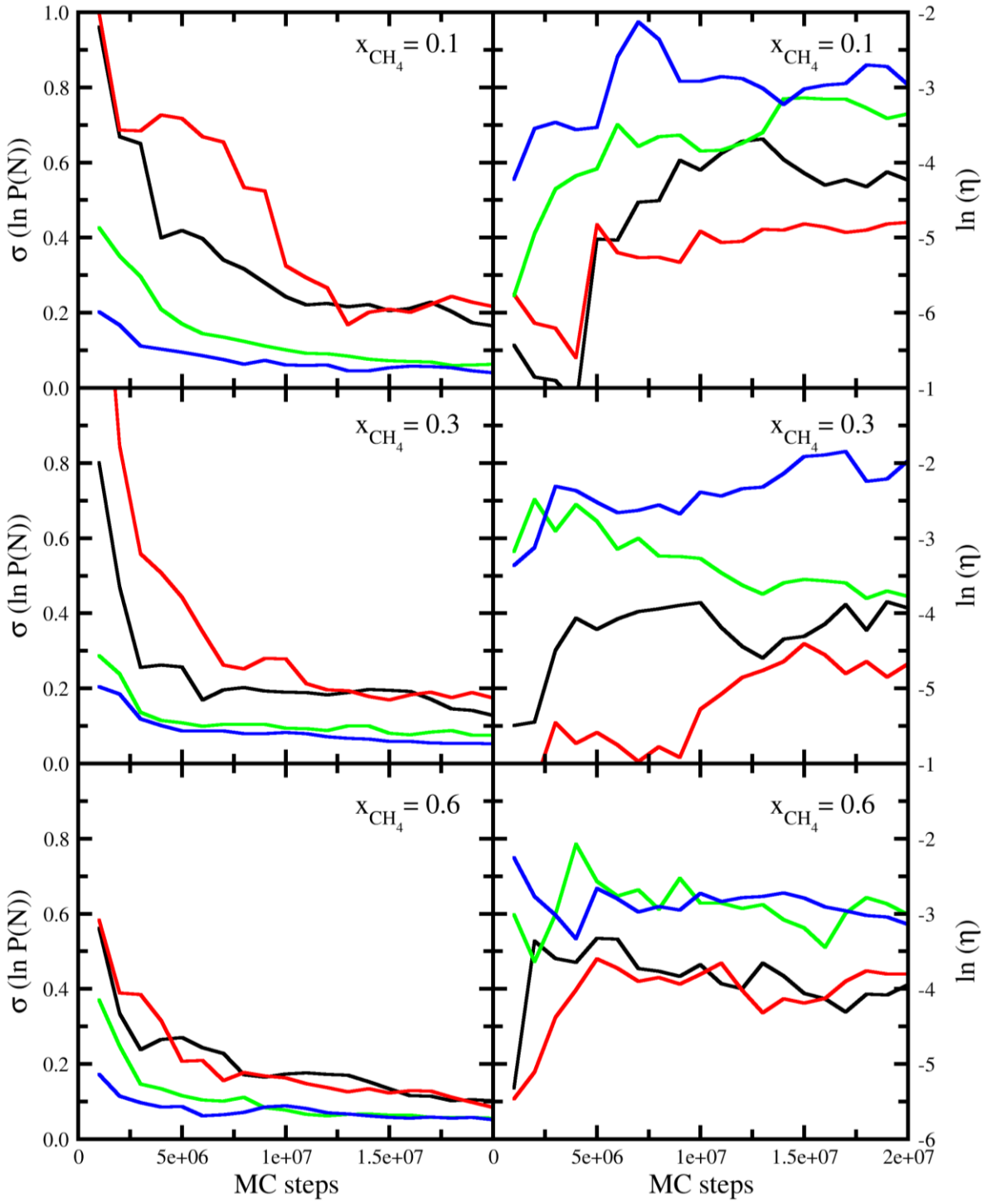

Figure 6: Efficiency and standard deviation in methane+n-butane at 255 K. Lines represent the efficiency and average uncertainty in probability distributions generated from GCMC simulations. Standard configurational-bias insertions (black), ME-1 (red), ME-2 (green), and ME-3 (blue). The MEMC move was performed with the exchange ratio of one butane with one methane.

In Figure 7, the probability distributions resulting from GCMC simulations with the various ME methods using an exchange ratio of one to one are presented for $x_{methane} = 0.3$, while data for other mole fractions are given in Appendix B, Figures S4 and S5. The probability distributions resulting from GCMC simulations with the various ME methods using an exchange ratio of one to two are presented in Figure S6-8 for a range of mole fractions. All MEMC methods converge to the same distribution. ME-3 shows rapid convergence, and within only $5x10^6$ MCS the correct distribution is obtained. The ME-2 algorithm shows slightly slower convergence compared to ME-3, but is still more efficient that ME-1 or configurational-bias trial insertions.

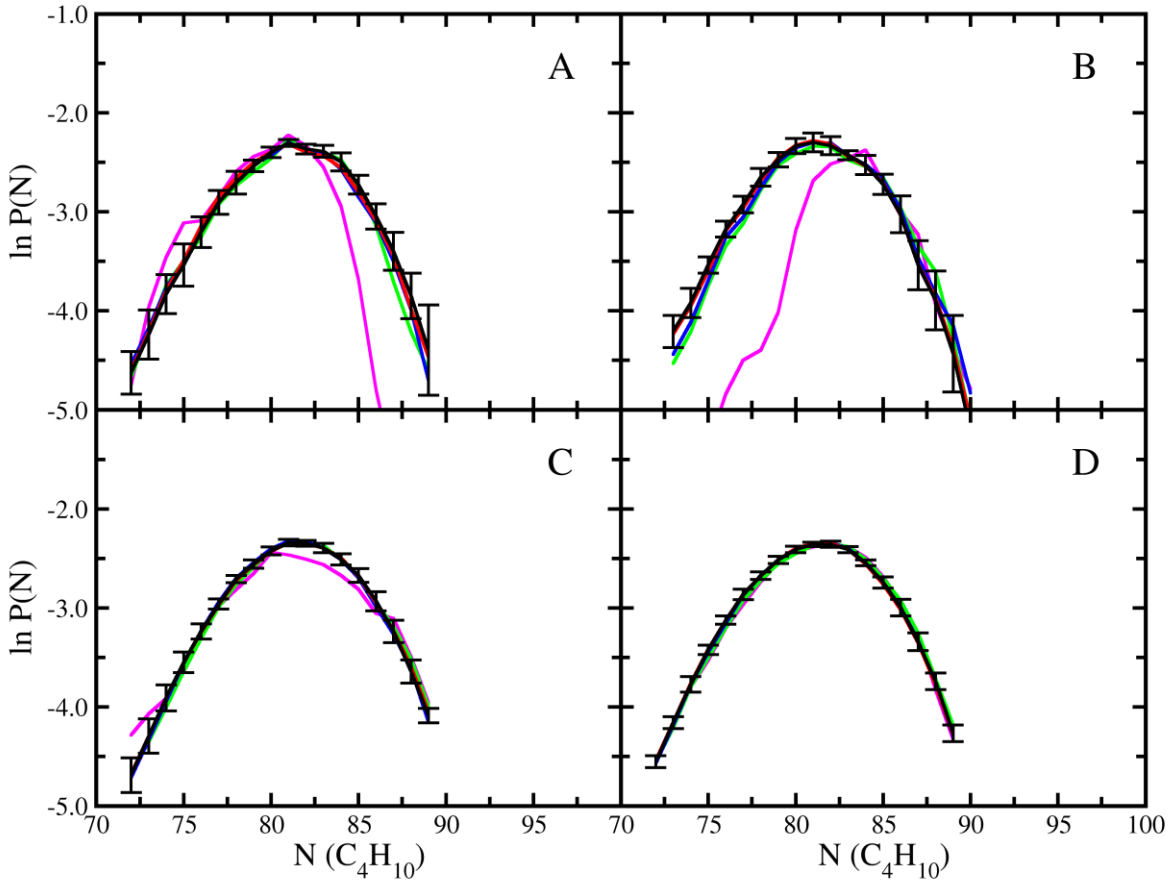

Figure 7: Probability distributions for methane+n-butane at 255 K and $x_{methane} = 0.3$. After simulations of: $1x10^6$ MCS (magenta), $5x10^6$ MCS (green), $1x10^7$ MCS (blue), $1.5x10^7$ MCS (red), and $2x10^7$ MCS (black) (A) Standard configurational-bias insertions, (B) ME-1 (C) ME-2 and (D) ME-3.

4.2 Perfluorobutane+n-butane

The perfluorobutane+n-butane system is an interesting case study because of its large deviations from Raoult's law, despite the fact that perfluorobutane and n-butane have very similar normal boiling points (270.96 K for C_4F_{10} and 272.61 K for C_4H_{10}) and both are non-polar with similar molecular geometries. This system has been modeled in the past with SAFT-VR[57], PC-SAFT[58] and GC-SAFT-VR[59], which showed close agreement with experimental data[60]. Gibbs ensemble Monte Carlo simulations using an identity exchange move have been used to study liquid-liquid equilibria for n-heptane+perfluoheptane[61], otherwise, grand canonical and Gibbs ensemble methods have rarely been applied to these kinds of mixtures. This is due, in part, to the difficulty in achieving an adequate number of accepted molecule transfers. For example, at 260 K, acceptance rates for the insertion of perfluorobutane in the neat liquid phase was approximately 0.075%.

In Figure 8, the pressure vs. composition diagram for perfluorobutane+n-butane at 260 K, predicted using the ME-3 algorithm and the Mie potentials developed by our group[54], is shown. The force field for perfluorobutane was modified slightly from the original work to use a more accurate seven term cosine series, which is described in detail in the Appendix A. Using standard Lorentz-Berthelot combining rules[62, 63] and no adjustable parameters for the cross interaction, very good agreement was achieved with experiment. The largest deviation results from the limitation in the united-atom force field for perfluorobutane, which over-predicts the vapor pressure at 260 K by approximately 0.1 bar.

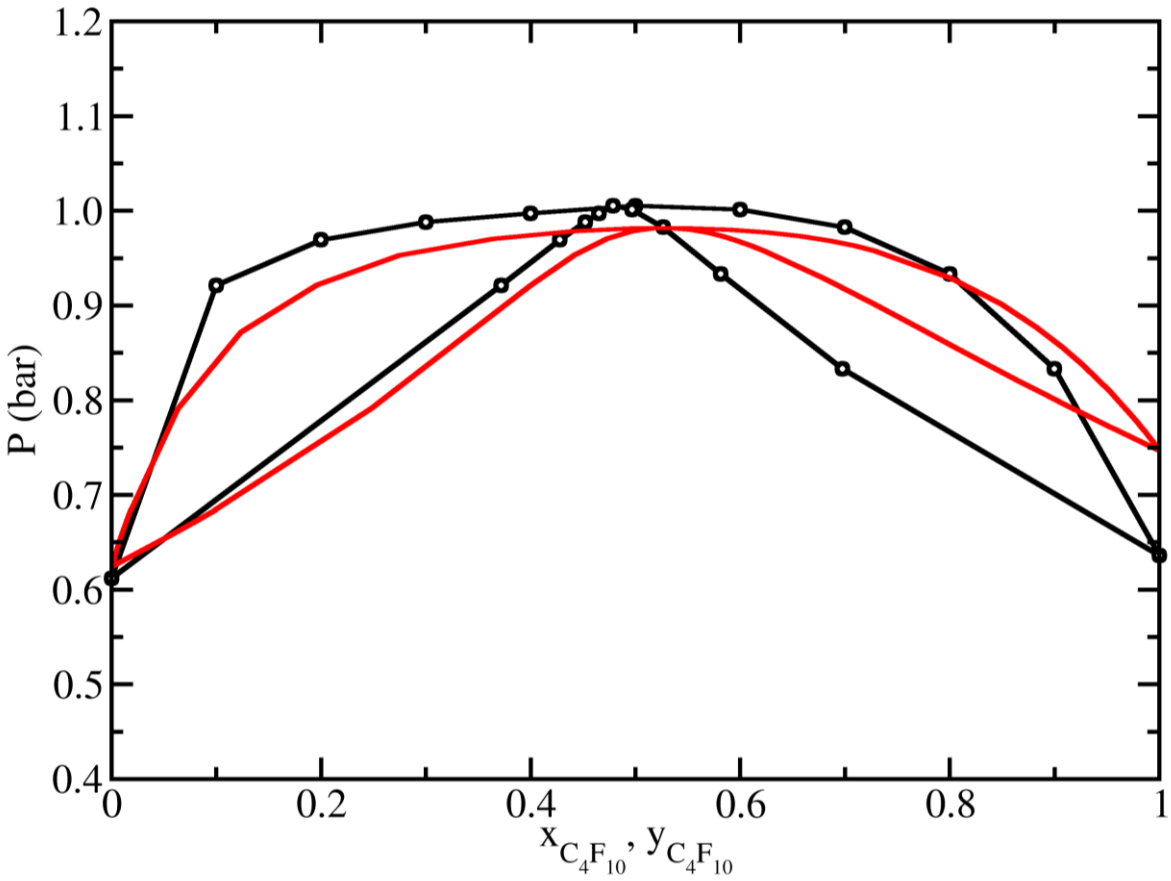

Figure 8: Pressure-composition diagram for perfluorobutane+n-butane at 259.95 K. The predictions from GCMC+histogram reweighting simulations using the ME-2 algorithm are given by (red line) while experiment data[60] are represented by (black circles). The line connecting the experimental data points is provided as a guide to the eye.

To evaluate the effectiveness of the molecular exchange move with a one to one exchange ratio and an exchange sub-volume of 6 Å × 6 Å × 9 Å, acceptance rates, uncertainties in the probability distributions, and efficiencies produced from the grand canonical Monte Carlo simulations were determined for liquid phase simulations at selected points along the coexistence curve. The effect of various simulation parameters on the performance of the CBMC and MEMC acceptance rates and efficiencies were also evaluated for liquid phase simulations containing 50 mol% n-butane, and are shown in Figure S9. Using the coupled-decoupled configurational-bias method[14], the probability of successfully inserting one perfluorobutane into a simulation box containing 10 mol%, 50 mol%, and 90 mol% of n-butane

was 0.073%, 0.026%, and 0.011%, respectively. The ME-1 algorithm increased acceptance rates approximately 4 times that of standard trial insertions for $x_{butane} > 0.50$, however, for lower concentrations of n-butane, no improvement was observed. For the ME-2 algorithm, acceptance rates of 4.92%, 4.17%, and 3.15% were obtained, while for ME-3, acceptance rates were 3.52%, 2.73 %, and 1.69%, respectively. For this system, the ME-2 algorithm produces the best acceptance rates because it works by aligning the backbone of perfluorobutane with the cavity left by the leaving n-butane. Acceptance rates were slightly lower for ME-3 since it grows the molecule using coupled-decoupled configurational-bias without requiring the backbone of the molecule to be aligned with the cavity created by the molecule that was removed.

The efficiency of the various molecular exchange algorithms is shown in Figure 9 as a function of Monte Carlo step for $x_{butane} = 0.1, 0.5, \text{and } 0.9$. Uncertainties shown are the average over uncertainties for each histogram bin in the probability distribution. Both the ME-2 and ME-3 algorithms show that convergence of the probability distributions was achieved within 10 million MCS, while for ME-1 and configurational-bias insertions, convergence was not achieved within 20 million MCS. Depending on composition, ME-3 provides efficiencies that are between 12 and 200 times greater than configurational-bias insertions for the insertion of perfluorobutane. Based on the trajectory of the uncertainties, it is unlikely that convergence of the probability distributions using standard Monte Carlo insertions would ever occur. Despite the fact that the ME-2 method provides slightly better acceptance rates for the molecular exchange move, at most compositions, ME-3 produces slightly faster convergence and better efficiencies. By growing the inserted molecule with coupled-decoupled configurational-bias[14], larger rearrangements take place in the system, even though more of the trial moves are rejected than in ME-2.

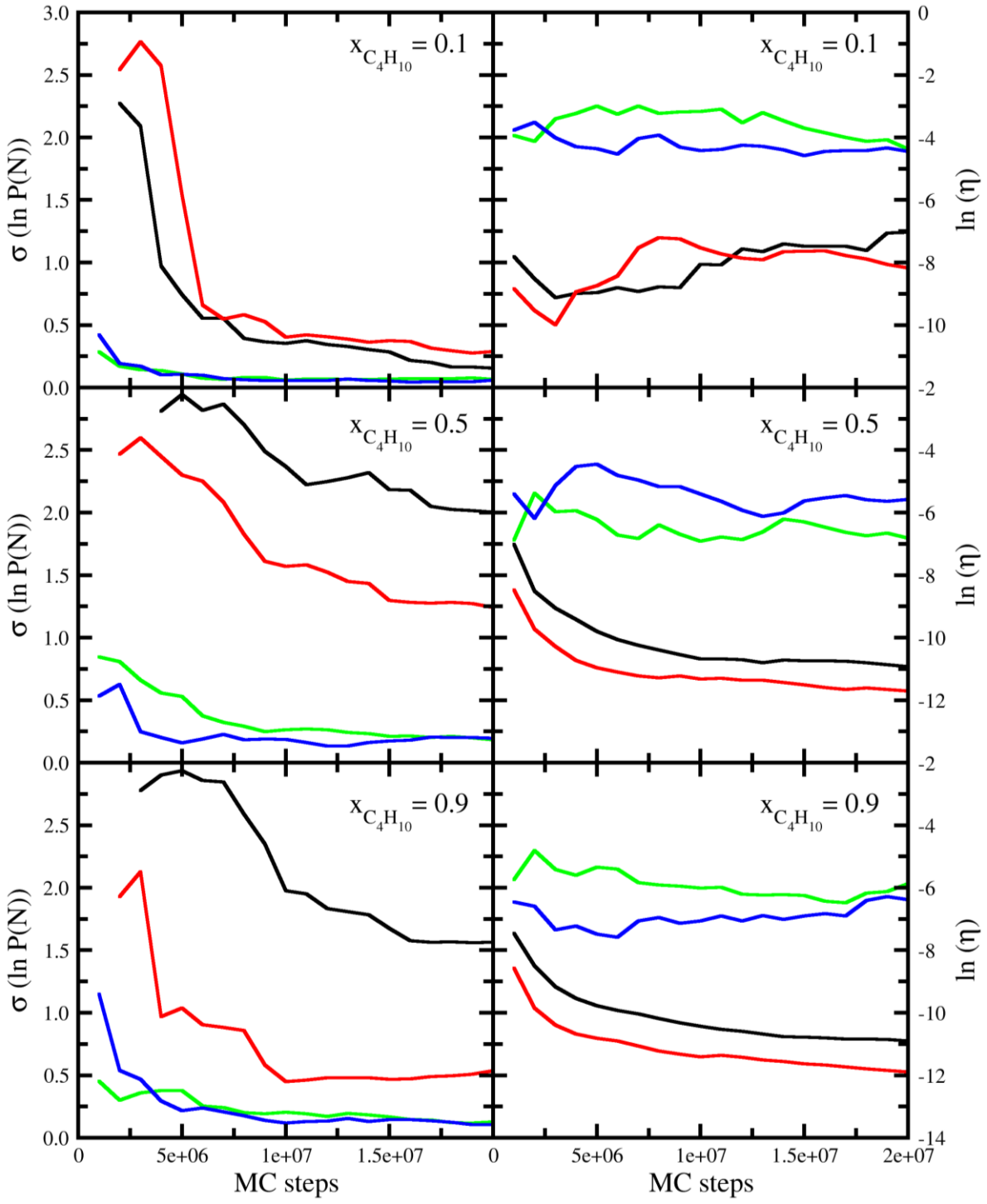

Figure 9: Efficiency and standard deviation in the perfluorobutane+n-butane binary mixture at 259.95 K. Lines represent the efficiency and average uncertainty in the perfluorobutane probability distribution; standard configurational-bias insertions(black), ME-1 (red), ME-2 (green), and ME-3 (blue). The MEMC moves were performed with an exchange ratio of one to one.

In Figure 10, the probability distributions resulting from GCMC simulations with the various ME methods are presented for $x_{butane} = 0.5$, while data for $x_{butane} = 0.1$ and 0.9 are given in Appendix B, Figures S10 and S11. The figure shows rapid convergence of the probability distributions for the ME-2 and ME-3 methods, while ME-1 and standard GCMC have not converged in 20 million MCS, although, the uncertainties calculated for ME-1 are approximately half those of standard GCMC. In Figure 11, heat maps are presented for the particle numbers and potential energies sampled during a liquid phase GCMC simulation. The heat maps illustrate how simulations with only configurational-bias insertions/deletions may become trapped in metastable states, resulting in poor sampling. Inclusion of the ME-3 algorithm produced a short equilibration period and a much broader sampling of the N_1, N_2, E phase space.

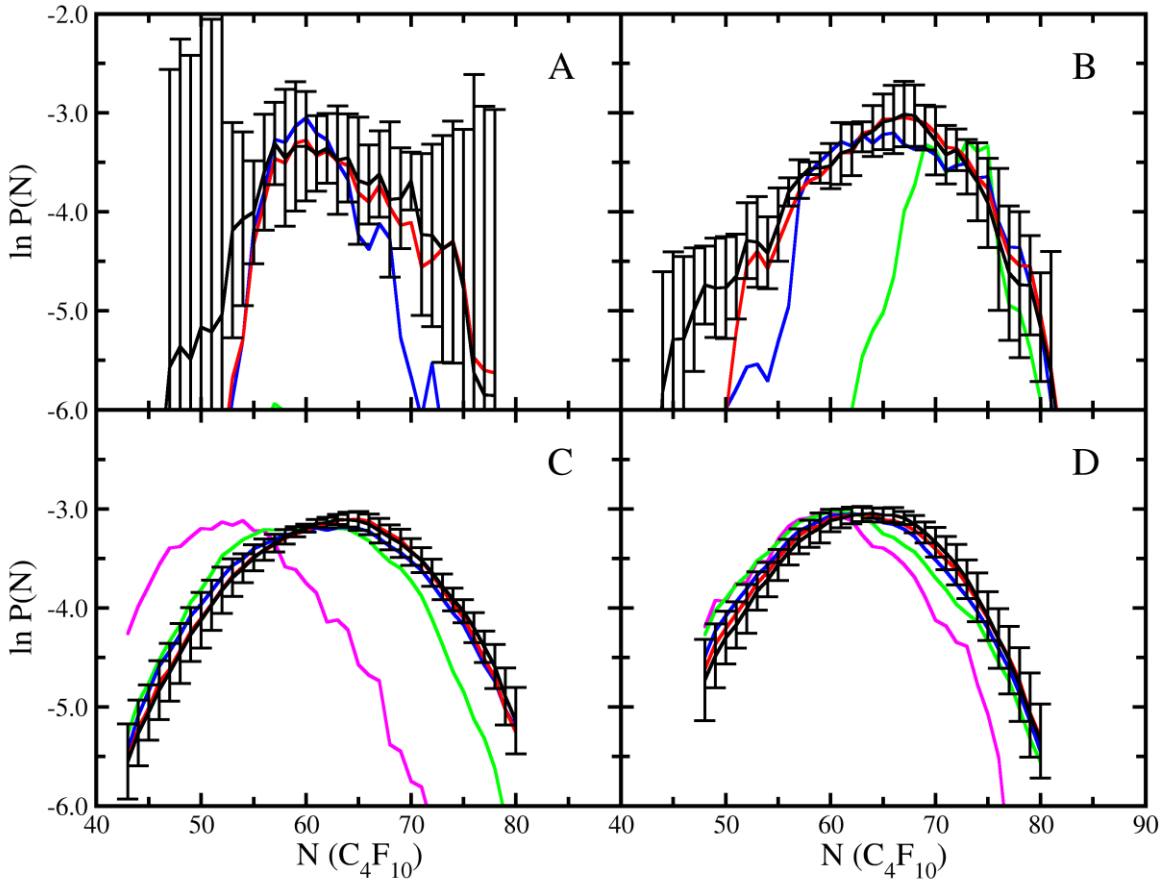

Figure 10: Molecule probability distribution for perfluorobutane+n-butane at $x_{butane} = 0.5$ and 259.95 K. After simulations of: 1x10^6 MCS (magenta), 5x10^6 MCS (green), 1x10^7 MCS (blue), 1.5x10^7 MCS (red), and 2x10^7 MCS (black) (A) Standard configurational-bias insertions, (B) ME-1 (C) ME-2 and (D) ME-3.

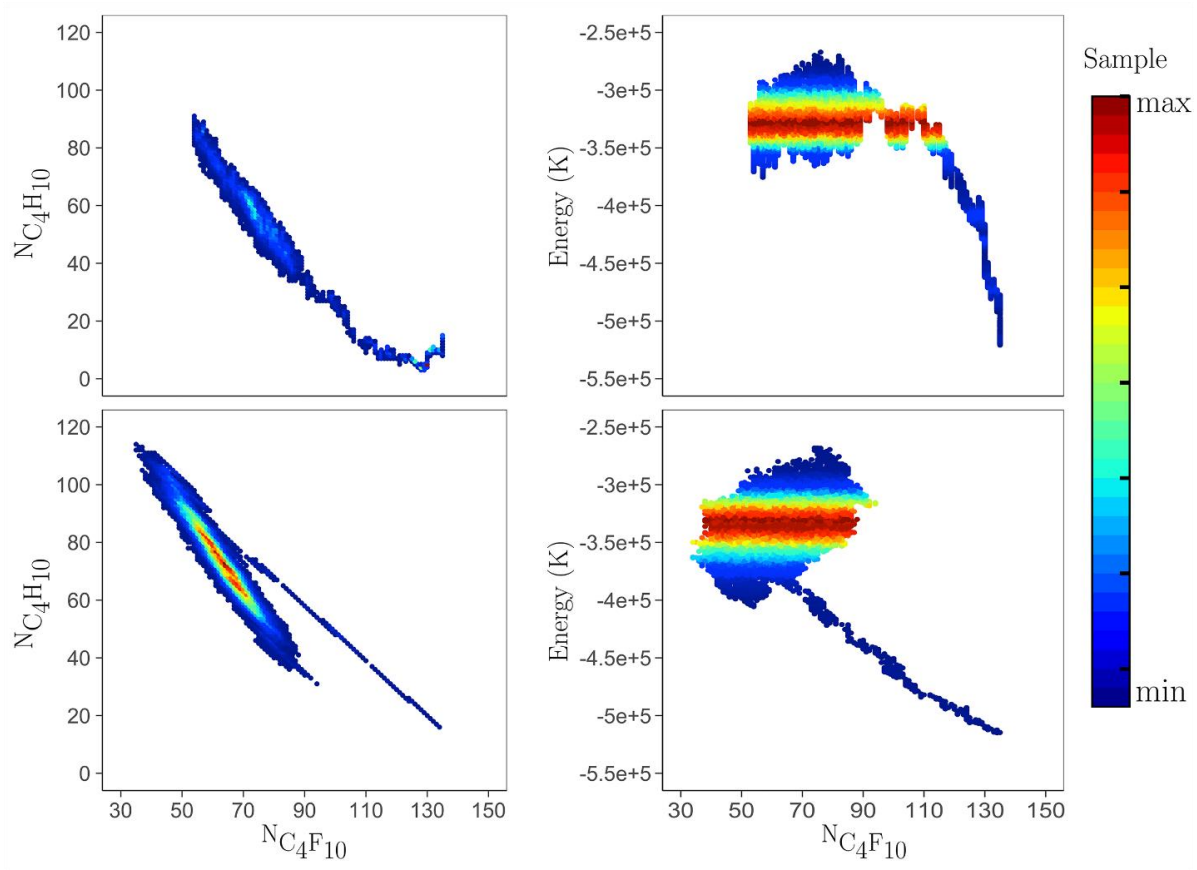

Figure 11: Heat maps of particle numbers (left panel) and potential energies (right panel) sampled during liquid phase grand canonical Monte Carlo simulations of perfluorobutane+n-butane at 259.95 K. Upper figures correspond to GCMC simulations with standard configurational-bias insertions/deletions, while the bottom figures correction to GCMC simulations with the ME-3 algorithm.

4.3 Water

In order to compare the performance of the MEMC move with other advanced sampling techniques, such as CBMC swap + identity switch[43](IS), continuous fractional component Monte Carlo (CFCMC)[44, 45], and configurational-bias continuous fractional component Monte Carlo (CB-CFCMC)[44], the vapor-liquid coexistence curve for SPC/E water[64] was predicted from the critical temperature to $0.44T_c$. To enhance the acceptance rate for insertions and deletions of water and to provide a uniform basis for comparison, the strategy of Bai and Siepmann was used[43]. For regular CBMC swaps, oxygen is inserted first, followed by the two hydrogen atoms. 16 trials were used for the first atom and 8 trials for all remaining atoms.

Simulations were performed as a mixture that contained approximately 0-10 “impurity” molecules, where the impurity molecule had an identical geometry to the SPC/E water model, but with partial charges reduced by a factor of 2 and the oxygen atom Lennard-Jones epsilon reduced by a factor of 4 compared to SPC/E water. Swap moves were performed only for impurity molecules, while the MEMC move is performed to exchange the impurity with water and vice versa. Move frequencies were adjusted to yield approximately to the same number of accepted molecule transfers for the swap and MEMC moves. Due to the poor performance of the ME-1 method in prior calculations, only the performance of the ME-2 and ME-3 methods were evaluated. An exchange ratio of one to one was used for all calculations.

The phase diagram for SPC/E water predicted from GCMC simulations using the ME-2 or ME-3 algorithm is shown in Figure 12, with a comparison to prior simulations[65]. Additional information on vapor pressure is provided in Appendix B, Figure S12. Excellent agreement was observed, validating both the MEMC algorithms and the simulation code used to perform the calculations.

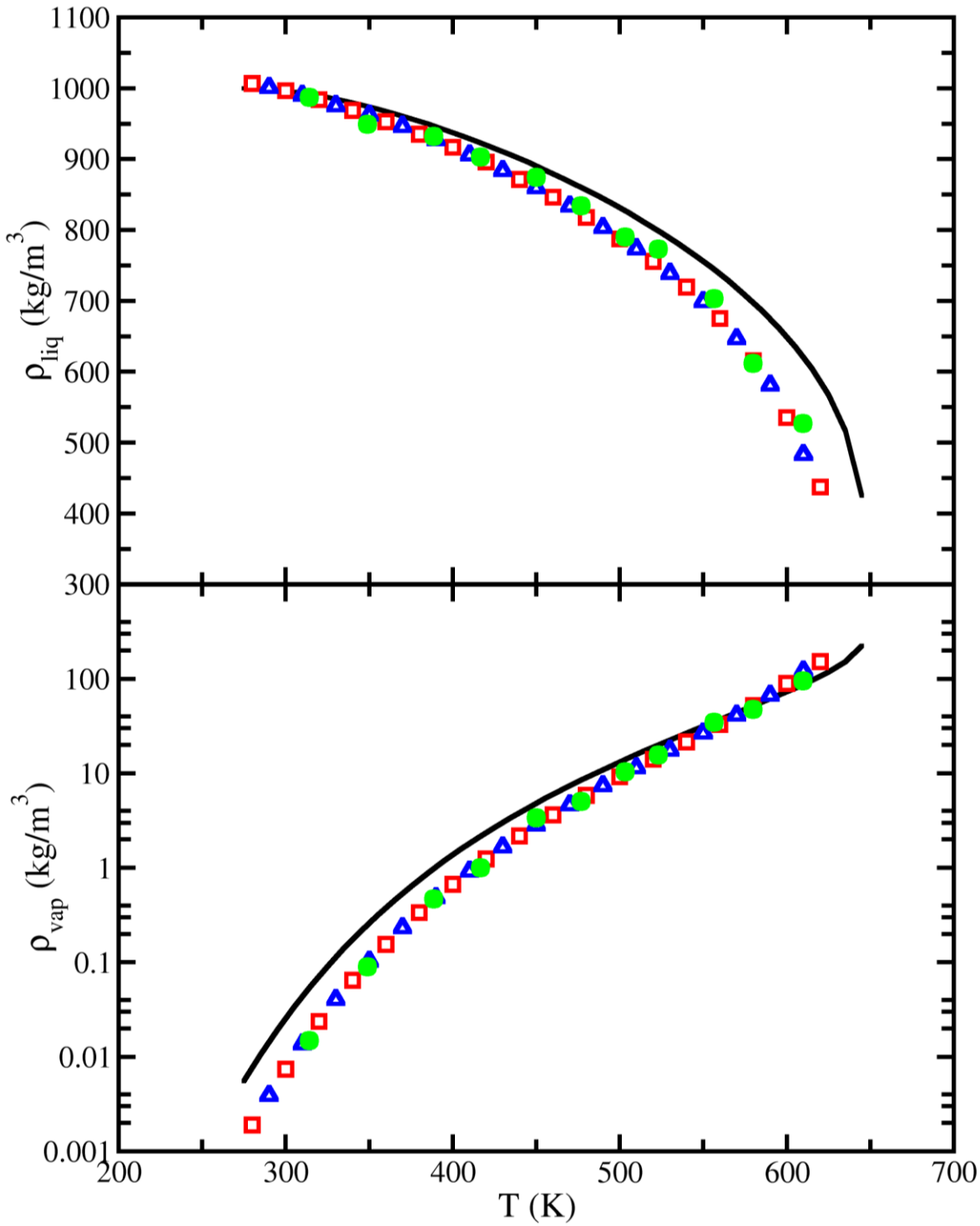

Figure 12: Vapor-liquid coexistence curve for SPC/E water predicted from GCMC+histogram reweighting simulations. NIST Chemistry WebBook[66] (solid lines), values obtained by Boulougouris et al.[65] (green circles), ME-2 algorithm (red squares), and ME-3 algorithm (blue triangles).

To compare the performance of MEMC with other methods, the effective number of molecule transfers was calculated. The effective number of molecule transfers was defined as

the insertion of an impurity molecule by the swap move and its conversion to a regular water molecule by the MEMC move, or the conversion of regular water to impurity via MEMC and then deletion of impurity by the swap move. Exchanges of impurity to water and back to impurity were not counted. The effective acceptance rate was calculated from the effective number of molecule transfers divided by the sum of attempted swap and MEMC moves. The results of these calculations are summarized in Table 2, with comparisons to the work of Bai and Siepmann[43], and Torres-Knoop *et al*.[44]. At 283 K, the effective acceptance rates for the ME-2 and ME-3 algorithms are 7.6 and 1.4 times greater, respectively, than the IS algorithm[43]. While the S+IS method reuses atomic coordinates of the molecule to be removed, the MEMC methods perform multiple trial orientations to insert the water molecule. In ME-2, first the center of the sub-volume was placed at the geometric center of the impurity, second the z-axis of the sub-volume was aligned with the O-H bond of impurity, and then multiple rotational trials were performed around the z-axis of the sub-volume. Aligning the O-H bond of water and the sub-volume allows some of the original hydrogen bonding to be maintained, while finding an energetically favorable position for the oxygen atom through rotational trials around the z-axis of the sub-volume, leading to significant improvements in the effective acceptance. In the ME-3 method, the oxygen atom of water was placed at the geometric center of the impurity molecule, and multiple rotational trials were performed on a sphere to find the most energetically favorable position. In order to maintain the hydrogen bonding formed by the impurity molecule, a large number of rotational trials are required, leading to a significant decrease in the acceptance efficiency compared to ME-2 method.

Compared to the original CFCMC method of Shi and Maginn[45], at 280 K, the ME-2 method exhibits twice the effective acceptance rate, while the ME-3 method is approximately 40% lower. The continuous fractional component Monte Carlo (CFCMC) and configurational-bias continuous fractional component Monte Carlo (CB-CFCMC) methods of Torres-Knoop

et al.[44] produced the largest acceptance rates of all methods. At 280 K, CFCMC and CB-CFCMC had acceptance rates that were 2.25 and 3.6 times larger, respectively, than the ME-2 method.

Table 2: Comparison of Swap + MEMC move acceptance percentages with standard CBMC, S+IS[43], CFCMC[44, 45], and CB-CFCMC[44] for SPC/E water.

T (K)	$\%P_{Imp-acc}$ $(CBMC)$		$\%P_{Switch-acc}$			$\%P_{Effective-acc}$			$\%P_{water-acc}$ $(CBMC)$			$\%P_{water-acc}$ $(CFCMC)$		$\%P_{water-acc}$ $(CB-CFCMC)$
	This work	Bai *et al.*	ME-2	ME-3	IS	ME-2	ME-3	S+IS	This work	Bai *et al.*	Torres-Knoop *et al.*	Shi. *et al.*	Torres-Knoop *et al.*	Torres-Knoop *et al.*
280	5.7	-	5.70	0.59	-	2.73	0.51	-	0.063	-	0.027	1.38	6.16	9.86
283	5.9	4.3	6.07	0.61	1.4	2.94	0.53	0.36	0.076	0.061	-	-	-	-
313	6.3	-	6.74	0.98	-	3.35	0.83	-	0.167	-	0.068	1.00	7.49	11.7
343	6.8	7.8	6.61	1.10	3.1	3.28	0.91	0.73	0.35	0.37	-	-	-	-
348	7.0	-	6.47	1.28	-	2.94	1.07	-	0.423	-	0.155	2.18	9.52	14.93
375	9.8	-	8.67	2.11	-	4.55	1.71	-	0.761	-	0.286	-	10.14	16.53
473	20.5	22	14.84	6.31	7.3	8.48	4.84	2.2	3.989	3.5	1.374	1.98	15.17	21.82
500	23	-	15.95	7.49	-	9.29	5.62	-	5.556	-	1.964	-	15.23	21.5

The acceptance efficiency was defined as the effective number of molecules transferred, divided by the total CPU time spent on swap and MEMC moves. In order to have a fair comparison between the acceptance efficiency of MEMC and S+IS, CFCMC, and CB-CFCMC methods, this quantity was normalized with respect to the acceptance efficiency of the standard CBMC method, minimizing the impact of CPU choice on the relative performance of the algorithms. The results of these calculations are listed in Table 3. At 280 K, the ME-2 method outperformed S+IS by 3.8 times, while the S+IS method is 23.9% better than ME-3. The performance of CFCMC and CB-CFCMC is 5-6 times greater than ME-2, although, it should be noted that the acceptance rates reported by Torres-Knoop *et al*. for standard swaps

of water were approximately 2.4 times lower than those reported in this work, or Bai and Siepmann[43].

Table 3: Comparison of relative acceptance efficiency for the MEMC, S+IS[43], CFCMC[44] and CB-CFCMC[44] methods.

T (K)	ME-2	ME-3	S+IS	CFCMC[44]	CB-CFCMC[44]
280	38.8	7.61	-	243.47	195.28
283	34.1	6.49	10	-	-
313	19.33	4.91	-	97.07	85.27
343	11.04	3.32	3.45	-	-
348	7.97	3.02	-	52.18	42.69
375	6.39	2.47	-	33.16	27.59
473	2.08	1.25	1.23	7.74	6.85
500	1.65	1.04	-	6.52	5.18

4.4 2,2,4-Trimethylpentane

As mentioned earlier, achieving a statistically valid number of molecule insertions in low temperature ($T < 0.7T_c$) simulations of branched alkanes can be challenging. Here, 2,2,4-trimethylpentane is used as an example to highlight how the MEMC move can significantly extend the range of temperatures where GCMC simulations may be used to predict vapor-liquid coexistence for a highly branched molecule. In this case, neopentane is used as the impurity molecule based on its similar structure to part of 2, 2, 4-trimethylpentane. This also illustrates the general nature of the MEMC algorithm, which does not require the molecules to be exchanged to be an integer numbers of each other. In Figure 13, the vapor-liquid coexistence curve for 2,2,4-trimethylpentane, using ME-2 algorithm and GCMC+histogram reweighting Monte Carlo simulations, is presented. Additional data for the ME-3 algorithms is presented in Appendix B, Figure S13. Using the ME-2 or ME-3 algorithms, it is possible to predict vapor-liquid coexistence to 280 K ($0.51T_c$), while prior simulations using only coupled-decoupled

configurational bias were limited to 390 K ($0.7T_c$). In Table 4, a detailed comparison is presented for the acceptance rates for direct swaps of neopentane and 2,2,4-trimethylpentane, MEMC moves, effective acceptance rates and effective acceptance rates per CPU time. Effective acceptance rate and acceptance efficiency is calculated using a similar method explained in the water section. The results of additional calculations performed with different CBMC parameters are given in Appendix B, Table S6. At all temperatures, the combination of impurity swap plus ME-2 or ME-3 method outperforms standard configurational-bias Monte Carlo. At 280 K, the relative acceptance efficiency (impurity swap+MEMC/standard CBMC) was 409 for ME-2 and 154 for ME-3. ME-2 is more effective than ME-3 for branched molecules because it inserts the entire molecule at the same time and aligns the backbone of the molecule to be inserted with the backbone of the molecule to be removed. ME-3 regrows the entire molecule using coupled-decoupled CBMC, however, many of these regrowths fail because they are unable to satisfy the internal molecular constraints due to the bond bending and torsional potentials[67]. In future work, it may be possible to improve the performance of the ME-3 algorithm for branched molecules by inclusion of the Jacobian-Gaussian scheme[68] for generating bending angle trials in the CBMC growth.

Table 4: Comparison of acceptance rates for swaps of the impurity molecule (neopentane), identity exchange via the MEMC algorithm, and swaps performed with standard configurational-bias Monte Carlo for 2,2,4-trimethylpentane.

T (K)	$\%P_{Imp-acc}$	$\%P_{Switch-acc}$		$\%P_{Effective-acc}$		$\%P_{acc}$	Effective acceptance per CPU time (s^{-1})			Relative acceptance efficiency	
	swap	ME-2	ME-3	ME-2	ME-3	CBMC	CBMC	ME-2	ME-3	ME-2	ME-3
280	0.013	0.89	0.03	0.013	0.008	0.00008	0.0003	0.109	0.041	409.2	153.7
330	0.10	2.21	0.15	0.096	0.057	0.0008	0.0026	0.917	0.288	356.9	112.0
390	0.85	5.69	0.55	0.653	0.274	0.022	0.0769	5.727	1.135	74.5	14.8
450	4.09	9.84	1.27	2.645	0.837	0.225	0.838	24.12	3.497	28.8	4.17
510	13.50	21.07	2.89	6.613	1.894	1.026	4.120	55.74	7.210	13.5	1.75

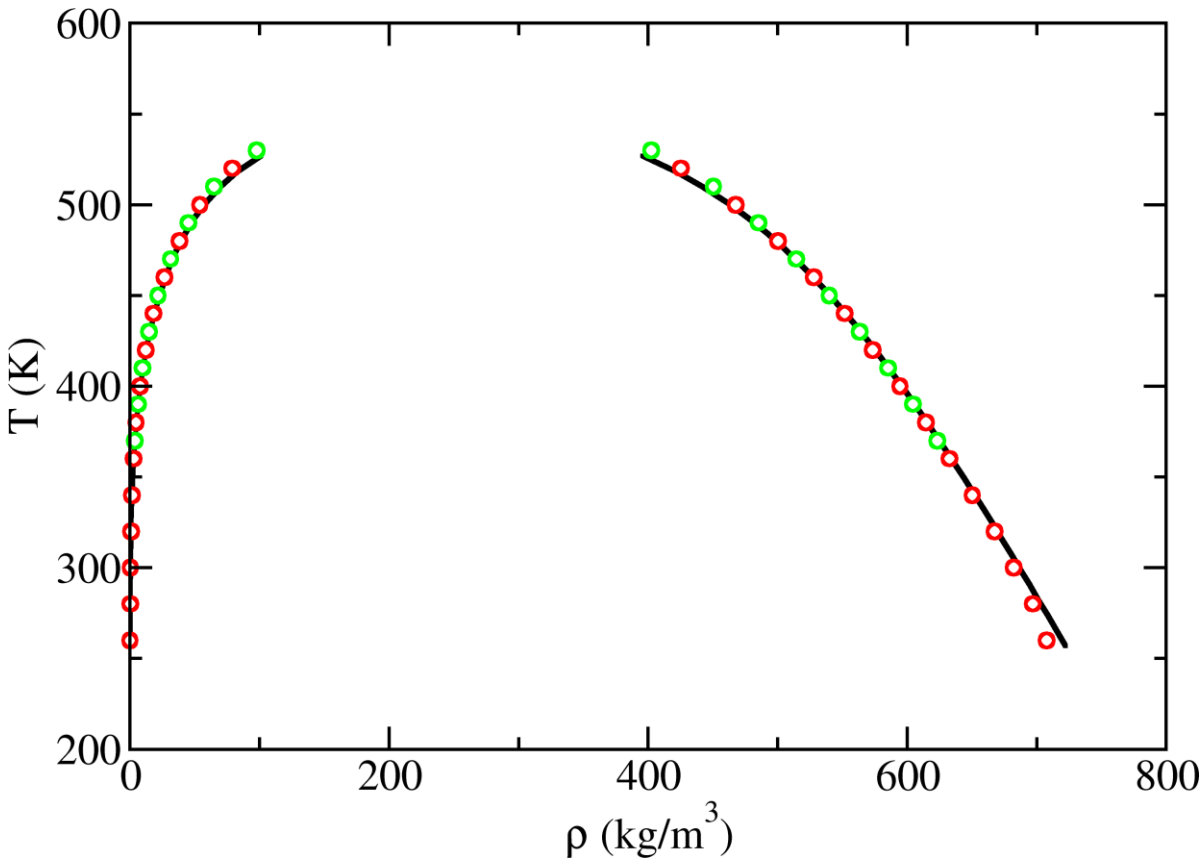

Figure 13: Vapor-liquid coexistence curve for 2,2,4-trimethylpentane predicted from GCMC+histogram reweighting simulations using Mie potentials[9]. Experimental data (solid lines)[69], ME-2 algorithm (red circles), and prior calculations using only configurational-bias Monte Carlo (green circles)[9].

CHAPTER 5: CONCLUSIONS

In this work, three variants of the molecular exchange method were developed, which could be used to evaluate the efficiency of various aspects of the algorithms. Locating the exchange sub-volume randomly (ME-1) was found to have the lowest efficiency, since frequently, no small molecules were found in the sub-volume that could be used for the molecular exchange, resulting in immediate rejection of the move. The ME-1 method is suitable only for systems that are very dilute with respect to the concentration of the large molecule. By identifying a small molecule at random first, placing the center of the sub-volume at the geometric center of the small molecule (ME-2), and aligning the backbone of the large molecule to be inserted with the small molecule to be removed, acceptance rates for the exchange move increased substantially. For water, the acceptance efficiency of the ME-2 method was found to be nearly 40 times greater than standard configurational-bias insertions, while for 2,2,4-trimethylpentane a 410 times improvement in acceptance efficiency was achieved. In the latter case, this was due to the use of a rigid-body insertion in ME-2, which eliminated the need to regrow the molecule in place. Finally, the inclusion of coupled-decoupled configurational-bias methods[14] to grow sections of the molecule from a central atom (ME-3) placed at the center of the sub-volume resulted in the greatest improvement in statistical efficiency compared to standard configurational-bias insertions for linear molecules without strong directional interactions. Improvements in efficiency of up to 200 times were observed for the perfluorobutane+n-butane system.

The algorithms presented in this work are notable because they were designed to work for any molecular topology over a wide range of compositions. Substantial performance gains were observed for ME-2 and ME-3 for all systems and compositions studied. As shown through the various case studies, however, each method has its strengths and weaknesses. For linear non-polar molecules, ME-3 is generally more efficient than ME-2, while ME-2 offers better

performance for small polar molecules, such as water, and highly branched molecules. Each algorithm has been implemented, and is now available, in the open-source Monte Carlo simulation engine GOMC, which is available to the public at GitHub[38].

APPENDIX A

In this section, the detailed computational procedures and mathematical methods of Molecular Exchange Monte Carlo (MEMC) move are provided.

Defining the exchange sub-volume vectors and transformation matrix $\boldsymbol{T_{VEX}}$:

An exchange sub-volume is a rectangular cuboid defined by three mutually orthogonal vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$. Vector $\boldsymbol{c}$ is either defined by the backbone orientation of the selected molecule or randomly defined according to a uniform distribution. For a given vector $\boldsymbol{c}$, vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are generated based on the following *Gram-Schmidt* algorithm.

1- Set $\boldsymbol{a}$ and $\boldsymbol{b}$ to two independent vectors, such as $\boldsymbol{i}$ and $\boldsymbol{j}$. (if $\boldsymbol{c}$ was in the same plane as $\boldsymbol{a}$ and $\boldsymbol{b}$, set either of $\boldsymbol{a}$ or $\boldsymbol{b}$ to $\boldsymbol{k}$).

2- $\boldsymbol{e_3} = \frac{\boldsymbol{c}}{|\boldsymbol{c}|}$

3- $\boldsymbol{b} = \boldsymbol{b} - (\boldsymbol{b}.\boldsymbol{e_3})\boldsymbol{e_3}$

4- $\boldsymbol{e_2} = \frac{\boldsymbol{b}}{|\boldsymbol{b}|}$

5- $\boldsymbol{a} = \boldsymbol{a} - (\boldsymbol{a}.\boldsymbol{e_3})\boldsymbol{e_3} - (\boldsymbol{a}.\boldsymbol{e_2})\boldsymbol{e_2}$

6- $\boldsymbol{e_1} = \frac{\boldsymbol{a}}{|\boldsymbol{a}|}$

where $|\boldsymbol{a}|$ is the norm of vector $\boldsymbol{a}$, and $(\boldsymbol{a}.\boldsymbol{b})$ represent scalar product of the two vectors.

To perform MEMC operations such as, counting the number of small molecules in sub-volume V_{EX}, inserting small molecules in V_{EX}, and aligning small and large molecules backbones with z-axis of the sub-volume, we need to define a new coordinate system based on the three unit vectors $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, and $\boldsymbol{e_3}$. To transform the coordinates from the simulation box reference frame

to the one defined by $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, and $\boldsymbol{e_3}$, we apply the transformation matrix $\boldsymbol{T_{VEX}}^{-1}$ and for the inverse transformation we apply $\boldsymbol{T_{VEX}}$ as defined below:

$$\boldsymbol{T_{VEX}} = \begin{bmatrix} e_{11} & e_{21} & e_{31} \\ e_{12} & e_{22} & e_{32} \\ e_{13} & e_{23} & e_{33} \end{bmatrix} \tag{S1}$$

$$\boldsymbol{T_{VEX}}^{-1} = \boldsymbol{T_{VEX}}^{T} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \tag{S2}$$

Defining a 2D random rotation matrix $\boldsymbol{R_z}$ about the backbone of the molecule:

In an MEMC move, the backbone of the molecule is aligned with $\boldsymbol{e_3}$. To perform random rotation around the backbone, a rotation matrix $\boldsymbol{R_z}$ is defined according to the following procedure:

1- Set θ to a random number between 0 and 1.

2- $\theta = \theta \times 2 \times \pi$

3- $\theta = \theta - \pi$

$$\boldsymbol{R_z} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{S3}$$

Defining a 3D random rotation matrix $\boldsymbol{R_s}$:

In the MEMC move, to perform rotation on a sphere uniformly, the fast random rotation matrices algorithm by Arvo is used. To construct the rotation matrix, perform the following steps.

1- Set θ to a random number between 0 and 1.

2- $\theta = \theta \times 2 \times \pi$

3- $\theta = \theta - \pi$

4- Set φ to a random number between 0 and 1.

5- $\varphi = \varphi \times 2 \times \pi$

6- Set r to a random number between 0 and 1.

7- Construct the 2D rotation $\boldsymbol{R_z}$, using θ.

8- Define $\boldsymbol{v}$ as

$$\boldsymbol{v} = \begin{bmatrix} \sqrt{r}\sin\varphi \\ \sqrt{r}\cos\varphi \\ \sqrt{1-r} \end{bmatrix} \quad \text{(S4)}$$

9- Defining the *Householder matrix* $\boldsymbol{H} = \boldsymbol{I} - 2\,\boldsymbol{v}\,\boldsymbol{v}^T$

10- The final rotation matrix can be expressed as

$$\boldsymbol{R_s} = -\boldsymbol{H}\,\boldsymbol{R_z} = 2\,\boldsymbol{v}\,\boldsymbol{v}^T\boldsymbol{R_z} - \boldsymbol{R_z} \quad \text{(S5)}$$

Defining the random orientation vector $\boldsymbol{c}$ for the exchange sub-volume $\boldsymbol{V_{EX}}$:

To generate a random orientation for the exchange sub-volume V_{EX}, we generate the vector $\boldsymbol{c}$ according to the following algorithm:

1- $\boldsymbol{c} = \boldsymbol{k}$

2- Construct the 3D rotation matrix $\boldsymbol{R_s}$

3- $\boldsymbol{c} = \boldsymbol{R_s}\,\boldsymbol{c}$

Finding the number of small molecules within the sub-volume V_{EX}:

To count the number of small molecules inside the V_{EX}, with the geometric center defined as vector $\boldsymbol{r_c}$ and dimensions of $w \times w \times l$, the following steps are performed. Repeat steps 1-3 for all the small molecules within the simulation box.

1- Calculate the minimum image distance between the geometric center of the sub-volume and GC of the molecule: $\boldsymbol{\Delta r = r_c - r_{GC}}$.

2- Transform the vector to the sub-volume coordinate system: $\boldsymbol{\Delta r' = {T_{VEX}}^{-1} \Delta r}$

3- If $\Delta r_1' < 0.5w$ and $\Delta r_2' < 0.5w$ and $\Delta r_3' < 0.5l$, the molecule is located within the sub-volume.

Finding a random location for GC of small molecule, within the sub-volume V_{EX}:

1- Set u_1, u_2, and u_3 to a random number between 0 and 1, independently.

2- $x_{GC} = u_1 \times w - 0.5\, w$, $y_{GC} = u_2 \times w - 0.5\, w$, $z_{GC} = u_3 \times l - 0.5\, l$

3- Transform the GC coordinate vector $\boldsymbol{r_{GC}}$, to the sub-volume coordinate system: $\boldsymbol{r'_{GC} = T_{VEX}\, r_{GC}}$

4- Shift the $\boldsymbol{r'_{GC}}$ to the geometric center of the sub-volume $\boldsymbol{r_c}$: $\boldsymbol{r''_{GC} = r'_{GC} + r_c}$

Generate Rotational trial around GC:

1- Construct the 3D rotation matrix $\boldsymbol{R_s}$

2- Repeat the following steps, for all atoms in the molecule ($i = 0, 1, \dots, n$)

   a. Shift the atom i to the origin with respect of its GC: $\boldsymbol{r'_i = r_i - r_{GC}}$

   b. Rotate the atom i around origin: $\boldsymbol{r''_i = R_s\, r'_i}$

   c. Shift the atom i back to its location: $\boldsymbol{r'''_i = r''_i + r_{GC}}$

Generate Rotational trial around the molecule's backbone:

To generate the rotational trial around backbone of the molecule, the molecule's backbone must be aligned with the predefined sub-volume V_{EX} system coordinate, $\boldsymbol{T_{VEX}}$. To align the molecule with the V_{EX}, the transformation matrix of molecule system coordinate $\boldsymbol{T_M}$ is defined as follow:

1- Shift the molecule coordinates to the origin with respect to its GC.

2- Calculate the minimum image vector of two specific atoms of the molecule $\boldsymbol{\Delta r}$ that represent the orientation of the molecule's backbone.

3- Set $\boldsymbol{c}$ to this vector: $\boldsymbol{c = \Delta r}$

4- Construct transformation matrix $\boldsymbol{T_M}$ of the molecule using the *Gram-Schmidt* algorithm.

5- Transform the molecule coordinates to the simulation box coordinate system, where $\boldsymbol{c}$ is aligned with the z-axis. Repeat the following step for all atoms in the molecule ($i = 0, 1, \dots, n$)

   a. $\boldsymbol{r_i' = {T_M}^{-1} r_i}$

Once the molecule coordinates are transformed, rotational trials around the z-axis are generated, molecule coordinates are transformed to V_{EX} system coordinate, and shifted to the geometric center of the sub-volume $\boldsymbol{r_c}$, as follows:

6- Construct the 2D rotational matrix $\boldsymbol{R_Z}$.

7- Repeat the following step for all atoms in the molecule ($i = 0, 1, \dots, n$)

   a. $\boldsymbol{r_i'' = R_Z r_i'}$

   b. $\boldsymbol{r_i''' = T_{VEX} r_i''}$

   c. $\boldsymbol{r_i''' = r_i''' + r_c}$

Forcefield:

The Mie potential is defined as:

$$U(r_{ij}) = C_n \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{n_{ij}} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \quad \text{(S6)}$$

where r_{ij}, ε_{ij}, and σ_{ij} are the separation, well depth, and collision diameter, respectively, for the pair of interaction sites *i* and *j*. The constant C_n is a normalization factor used such that the minimum of the potential remains at $-\varepsilon_{ij}$ for all n_{ij}.

$$C_n = \left(\frac{n_{ij}}{n_{ij} - 6} \right) \left(\frac{n_{ij}}{6} \right)^{6/(n_{ij}-6)} \quad \text{(S7)}$$

For the 12-6 potential, C_n reduces to the familiar value of 4. Parameters governing interactions between unlike interaction sites were determined using the Lorentz-Berthelot combining rules[70, 71].

$$\sigma_{ij} = (\sigma_{ii} + \sigma_{jj})/2 \quad \text{(S8)}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad \text{(S9)}$$

To determine repulsion exponents for cross interactions, an arithmetic average was used.

$$n_{ij} = (n_{ii} + n_{jj})/2 \quad \text{(S10)}$$

Mie potential has been optimized for noble gases[72, 73], linear and branched alkane[9, 54], n-alkyne[74]. All non-bonded parameters used in this work are listed in Table S1.

Table S1: Non-bonded parameters for n-alkanes, perfluoro-alkanes[54], branched alkanes[9], and SPC/E water[64].

Pseudo-atom	$\varepsilon_i/k_b(K)$	σ_i (Å)	n_i	q_i
CH_4	161.00	3.740	14	0.00
CH_3	121.25	3.783	16	0.00
CH_2	61.00	3.990	16	0.00
CH (C_N > 4, S/L)	14.00	4.700	16	0.00
C ($C_N \leq 4$, S/L)	1.45	6.100	16	0.00
C (C_N > 4, S/L)	1.20	6.200	16	0.00
CF_3	155.75	4.475	36	0.00
CF_2	72.20	4.750	44	0.00
O	78.21	3.167	12	-0.8476
H	0.00	0.00	0.00	0.4238

Fixed bond lengths for n-alkanes, perfluoro-alkane[54], branched alkanes[9], and SPC/E water[64] were used to connect pseudo-atoms and are listed in Table S2. Angle bending was described using a harmonic potential

$$U_{bend} = \frac{k_\theta}{2}(\theta - \theta_0)^2 \quad \text{(S11)}$$

where θ_0 is the measured bond angle, θ_0 is the equilibrium bond angle, and k_θ is the force constant. Equilibrium bond angles and force constants are listed in Table S2.

Table S2: Bonded parameters for n-alkanes, perfluoro-alkane[54], branched alkanes[9], and SPC/E water[64].

Bond type	Bond length/ Å	Angle type	θ_0/degree	K_θ/ K.rad^{-2}
CH_2-CH_3	1.54	$CH_3-CH_2-CH_2$	114	62500
CH_2-CH_2	1.54	$CH_2-CH_2-CH_2$	114	62500
CH_2-CH_2	1.54	$C-CH_2-CH$	114	62500
$CH-CH_3$	1.54	$CH_3-CH-CH_3$	112	62500
$CH-CH_2$	1.54	$CH_3-CH-CH_2$	112	62500
$C-CH_3$	1.54	CH_3-C-CH_3	109.47	62500
$C-CH_2$	1.54	CH_3-C-CH_2	109.47	62500
CF_2-CF_3	1.54	$CF_3-CF_2-CF_2$	114	62500
CF_2-CF_2	1.54	$CF_2-CF_2-CF_2$	114	62500
$O-H$	1.00	$H-O-H$	109.47	Fixed

Dihedral energies were governed by a cosine series

$$U_{tors} = \sum_{n=0}^{N} c_n[1 + cos(n\varphi - \delta_n)] \quad (12)$$

where φ is the dihedral angle, c_n are dihedral force constants, n is the multiplicity, and δ_n is the phase shift. These constants are listed in Table S3. Fourier constants for alkanes were taken from OPLS-UA[75, 76] and for perfluoroalkanes, more accurate seven term cosine series were used.

Table S3: Torsional parameters for n-alkanes, perfluoro-alkane[54], branched alkanes[9], and SPC/E water[64].

torsion	n	c_n/ (K)	δ_n
CH_x—(CH_2)—(CH_2)—CH_2	1	335.03	0
	2	-68.19	π
	3	791.32	0
CH_x—(CH_2)—(CH)—CH_y	0	-251.06	0
	1	428.73	0
	2	-111.85	π
	3	441.27	0
CH_x—(CH_2)—(C)—CH_y	3	461.29	0
CF_x—(CF_2)—(CF_2)—CF_y	0	-1577.68	0
	1	791.61	0
	2	333.65	0
	3	854.01	0
	4	349.25	0
	5	211.51	0
	6	117.66	0
	7	-83.44	0

APPENDIX B

In this section, the numerical results and additional data are provided.

Table S4: Selected phase coexistence data for perfluorobutane(1)+n-butane(2) predicted by grand canonical Monte Carlo simulations using Mie potentials[54]. Uncertainty in data are presented by the numbers in parenthesis.

P (bar)	x_1	y_1
0.68(3)	0.02(1)	0.10(5)
0.87(2)	0.12(1)	0.34(2)
0.92(1)	0.20(2)	0.40(1)
0.97(1)	0.36(2)	0.472(8)
0.981(5)	0.50(3)	0.517(5)
0.980(3)	0.59(3)	0.553(3)
0.963(5)	0.709(7)	0.610(1)
0.928(6)	0.802(2)	0.680(1)
0.855(4)	0.908(3)	0.806(1)
0.747(3)	1.000	1.000

Table S5: Vapor-liquid coexistence data predicted from GCMC+histogram reweighting simulations using ME-2 method for SPC/E water.

T (K)	ρ_l (kg/m^3)	ρ_v (kg/m^3)	P (bar)	ΔH_v (kJ/mol)
600	535(3)	90(2)	102.7(7)	15.6(2)
580	615(2)	52.1(8)	74.7(2)	21.4(2)
560	675(2)	33.0(1)	53.54(9)	25.76(3)
540	720(1)	21.51(6)	37.54(6)	29.09(3)
520	756(1)	14.11(5)	25.71(5)	31.77(4)
500	787(1)	9.17(4)	17.12(4)	34.01(3)
480	818(1)	5.86(2)	11.04(2)	36.01(6)
460	846(2)	3.64(1)	6.84(2)	37.97(8)
440	871(1)	2.173(5)	4.05(2)	39.69(4)
420	896(2)	1.241(3)	2.26(1)	41.31(6)
400	917(2)	0.669(4)	1.19(1)	42.92(4)
380	935(1)	0.336(3)	0.576(5)	44.29(5)
360	953(2)	0.155(1)	0.255(3)	45.59(3)
340	969(1)	0.065(1)	0.101(1)	46.74(4)
320	984(1)	0.024	0.035	47.94(4)
300	996(2)	0.007	0.010	49.19(5)
280	1007(5)	0.002	0.003	50.4(1)

The acceptance rate of inserting or removing neopentane was 68 times lower than the acceptance rate for exchanging neopentane with 2,2,4-trimethylpentane and vice versa via the ME-2 algorithm. This shows that insertion of neopentane is the rate limiting step in the process. In order to improve the acceptance rate for insertions of neopentane, CBMC angle and dihedral trials were increased to 500, and the number of CBMC trials for the first atom and remaining atoms were increased to 16 and 10, respectively. In Table S6, a detailed comparison is presented for the acceptance rates for direct swaps of neopentane and 2,2,4-trimethylpentane, MEMC moves, effective acceptance rates and effective acceptance rates per CPU time.

Table S6: Comparison of acceptance rates for swaps of the impurity molecule (neopentane), identity exchange via the MEMC algorithm, and swaps performed with standard configurational-bias Monte Carlo for 2,2,4-trimethylpentane.

T (K)	$\%P_{Imp-acc}$	$\%P_{Switch-acc}$		$\%P_{Effective-acc}$		$\%P_{acc}$	Effective acceptance per CPU time (1/s)			Relative acceptance efficiency	
	swap	ME-2	ME-3	ME-2	ME-3	CBMC	CBMC	ME-2	ME-3	ME-2	ME-3
280	0.017	1.19	0.05	0.039	0.099	0.0002	0.0005	0.244	0.042	520.2	90.0
330	0.168	2.60	0.18	0.183	0.077	0.0023	0.0047	1.312	0.250	278.7	53.3
390	1.30	5.89	0.71	0.878	0.393	0.036	0.0745	5.504	1.023	73.9	13.7
450	6.05	10.26	1.56	3.466	1.067	0.338	0.745	22.40	2.779	30.1	3.73
510	18.0	21.76	3.43	7.887	2.337	1.510	3.520	49.02	5.576	13.9	1.58

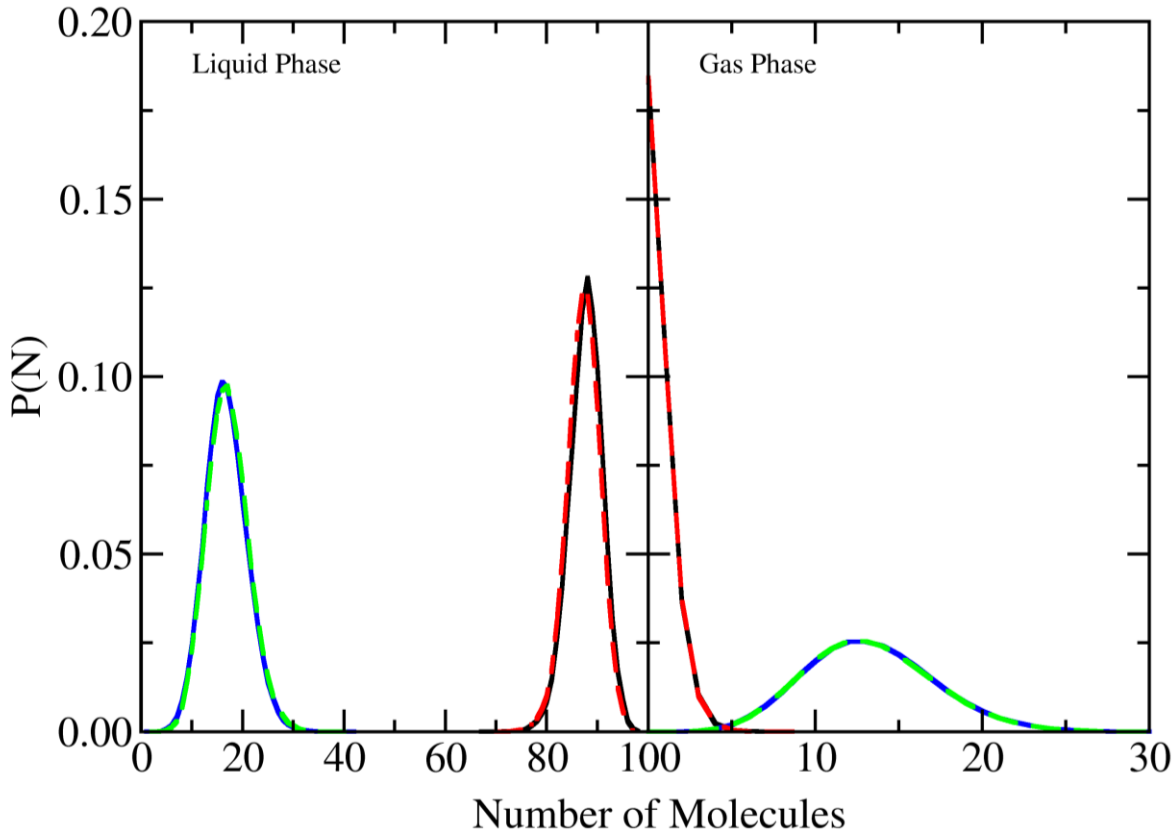

Figure S1: Probability distributions predicted from gas ($\mu_{butane} = -2960, \mu_{methane} = -2000$) and liquid ($\mu_{butane} = -2840, \mu_{methane} = -2000$) phase GCMC simulations of methane+n-butane at 277 K. Solid lines denote the probability distributions for n-butane (black) and methane (blue) using standard configurational-bias insertions and deletions. Dashed lines denote the probability distributions for n-butane (red) and methane (green) using the ME-1 algorithm.

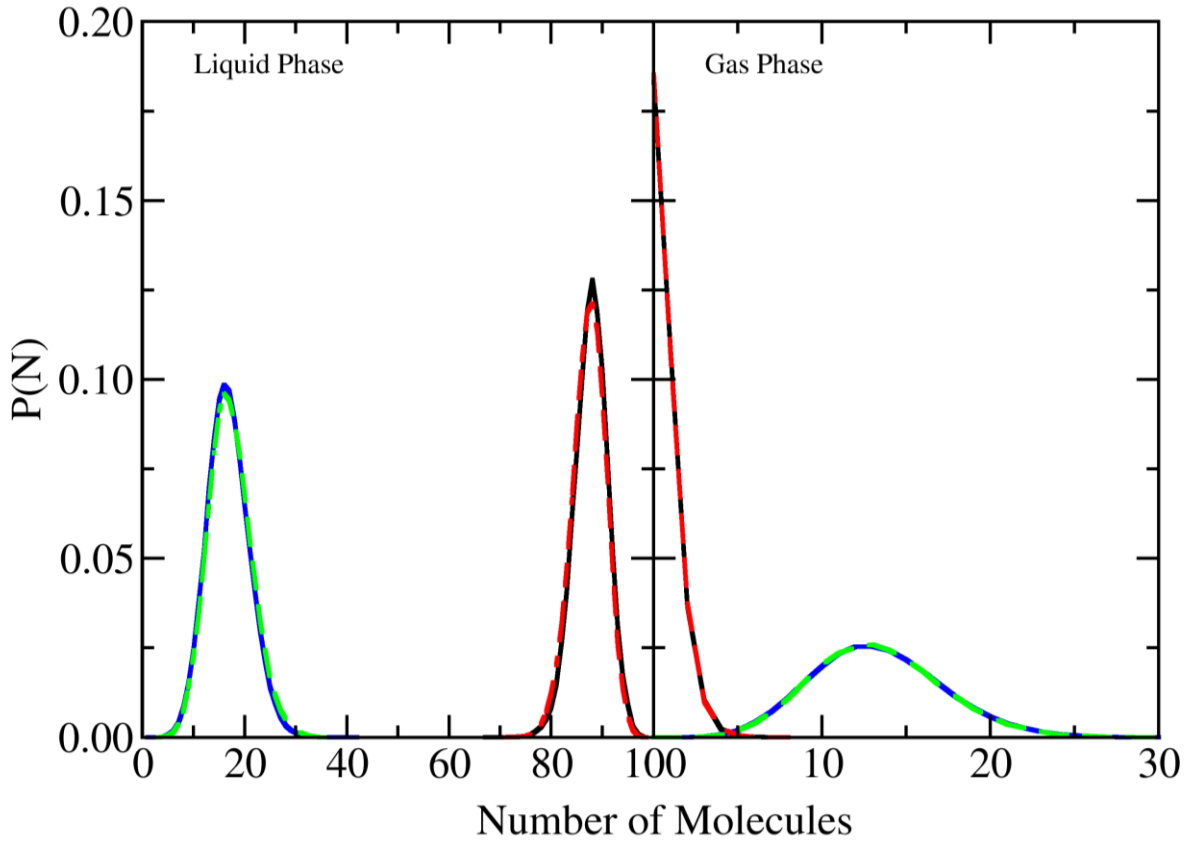

Figure S2: Probability distributions predicted from gas ($\mu_{butane} = -2960, \mu_{methane} = -2000$) and liquid ($\mu_{butane} = -2840, \mu_{methane} = -2000$) phase GCMC simulations of methane+n-butane at 277 K. Solid lines denote the probability distributions for n-butane (black) and methane (blue) using standard configurational-bias insertions and deletions. Dashed lines denote the probability distributions for n-butane (red) and methane (green) using the ME-2 algorithm.

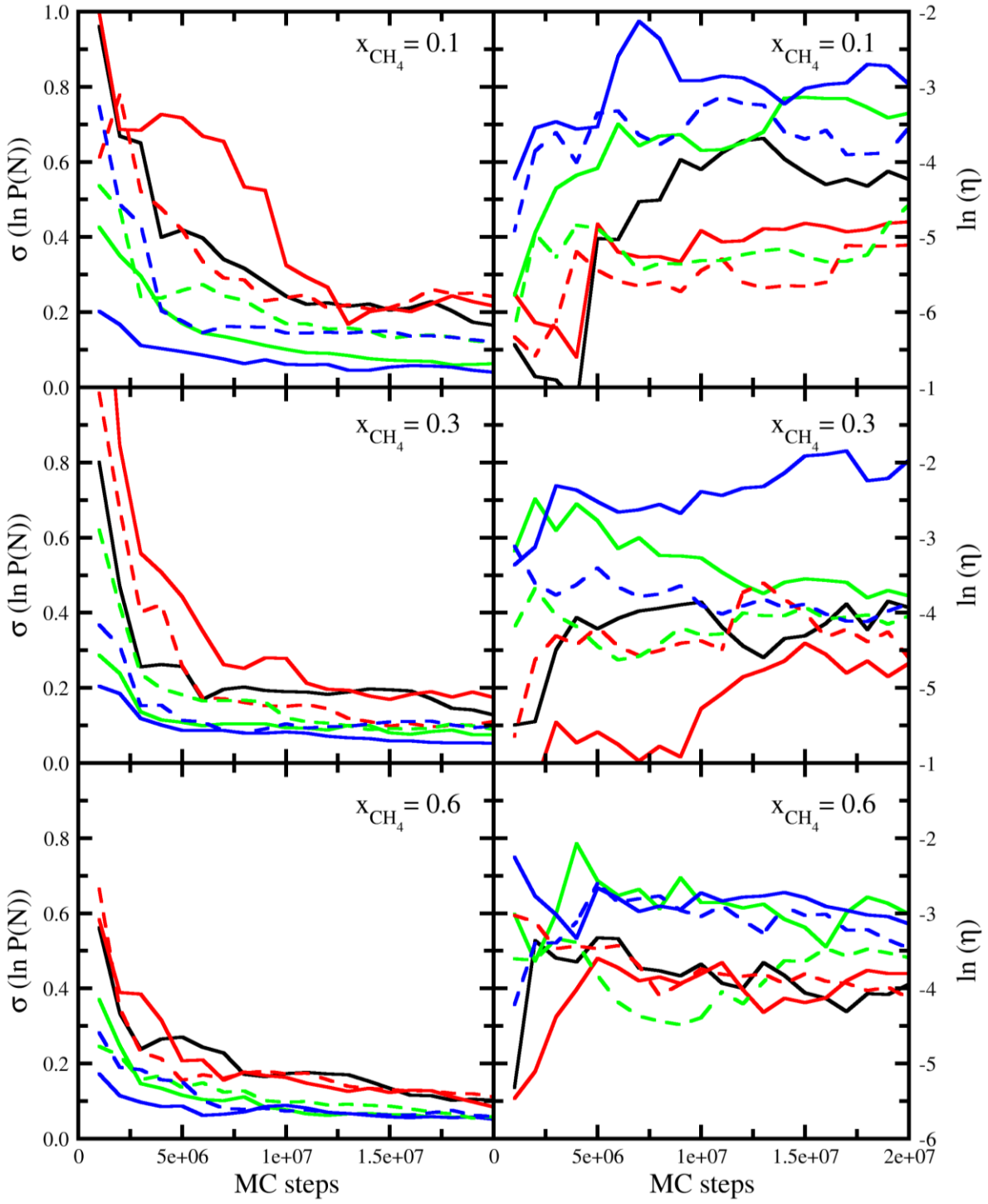

Figure S3: Efficiency and standard deviation in methane+n-butane binary mixture at 255 K. Lines represent the efficiency and uncertainty in n-butane distribution probability; standard CBMC method (black), ME-1 (red), ME-2 (green), and ME-3 (blue). MEMC move with exchanging one n-butane with one methane represented by solid lines, exchanging one n-butane with two methane molecules represented by dashed lines.

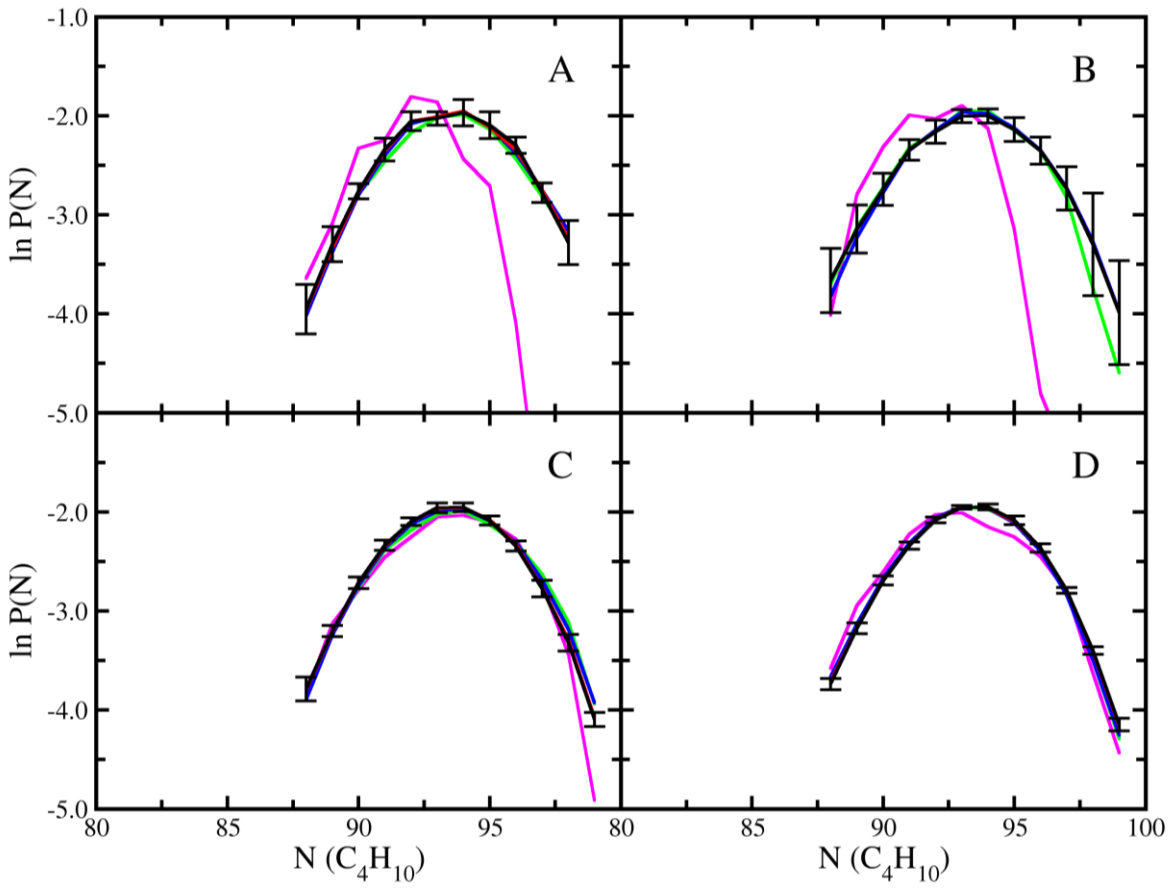

Figure S4: Molecule probability distribution in methane+n-butane binary mixture system at $x_{methane} = 0.1$ and 255 K. Lines in magenta, green, blue, red, and black represent the probability distribution of n-butane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using ME-1, ME-2, and ME-3 method with exchanging ratio of one methane with one n-butane, respectively.

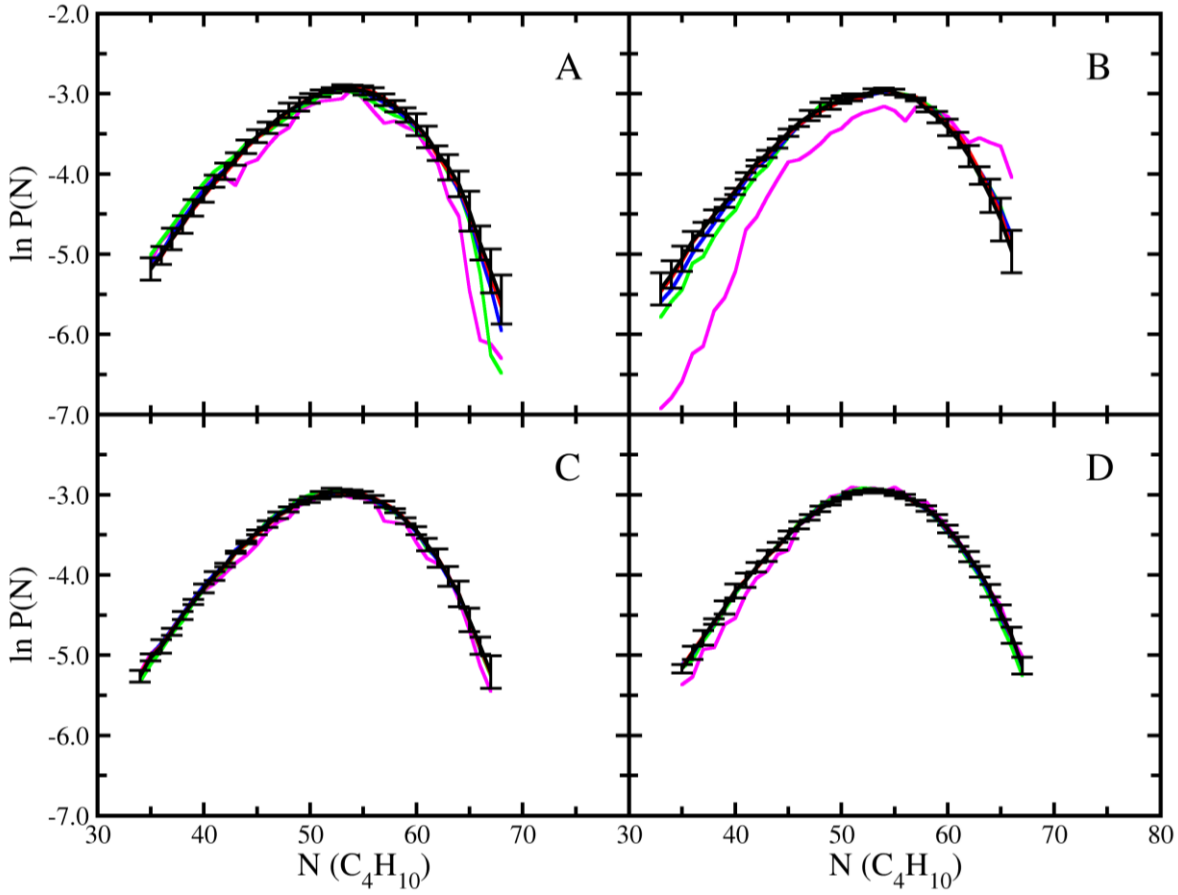

Figure S5: Molecule probability distribution in methane+n-butane binary mixture system at $x_{methane} = 0.6$ and 255 K. Lines in magenta, green, blue, red, and black represent the probability distribution of n-butane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using ME-1, ME-2, and ME-3 method with exchanging ratio of one methane with one n-butane, respectively.

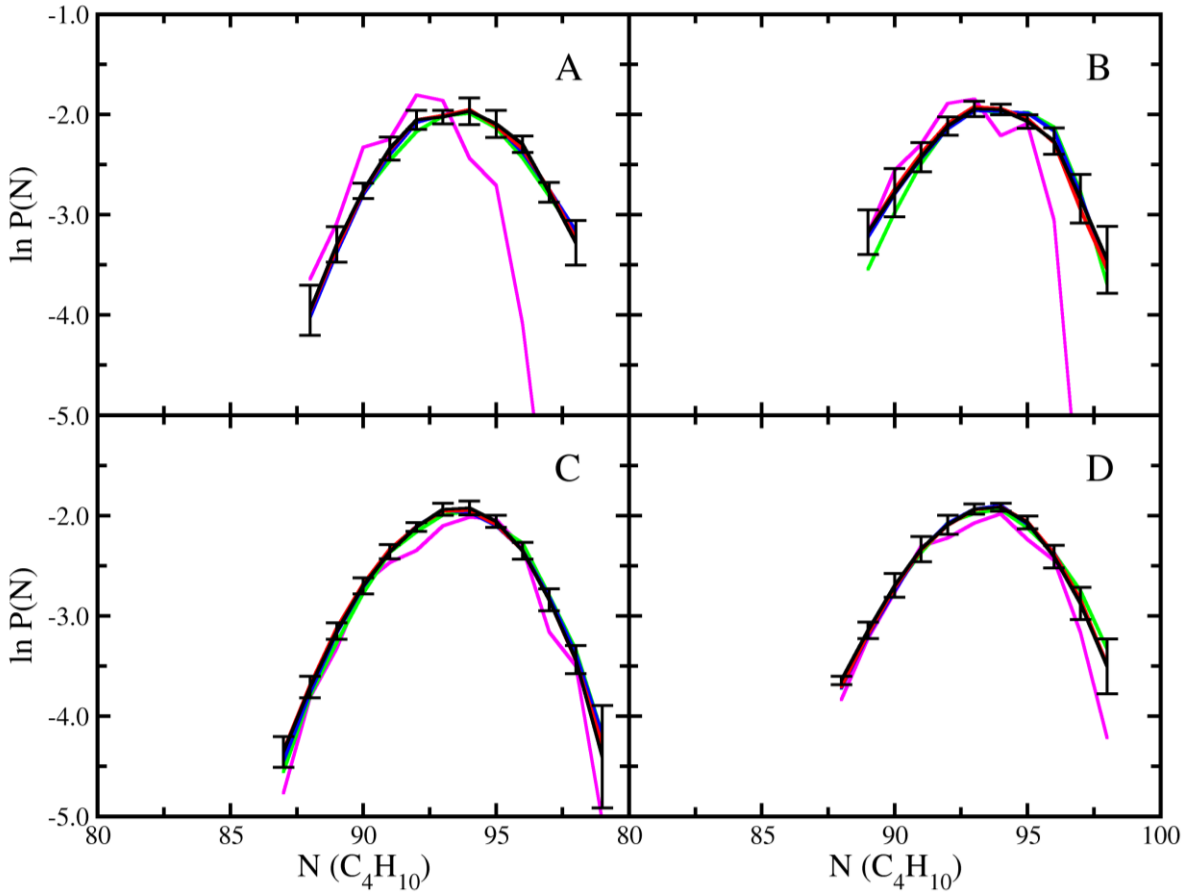

Figure S6: Molecule probability distribution in methane+n-butane binary mixture system at $x_{methane} = 0.1$ and 255 K. Lines in magenta, green, blue, red, and black represent the probability distribution of n-butane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using ME-1, ME-2, and ME-3 method with exchanging ratio of two methane molecules with one n-butane, respectively.

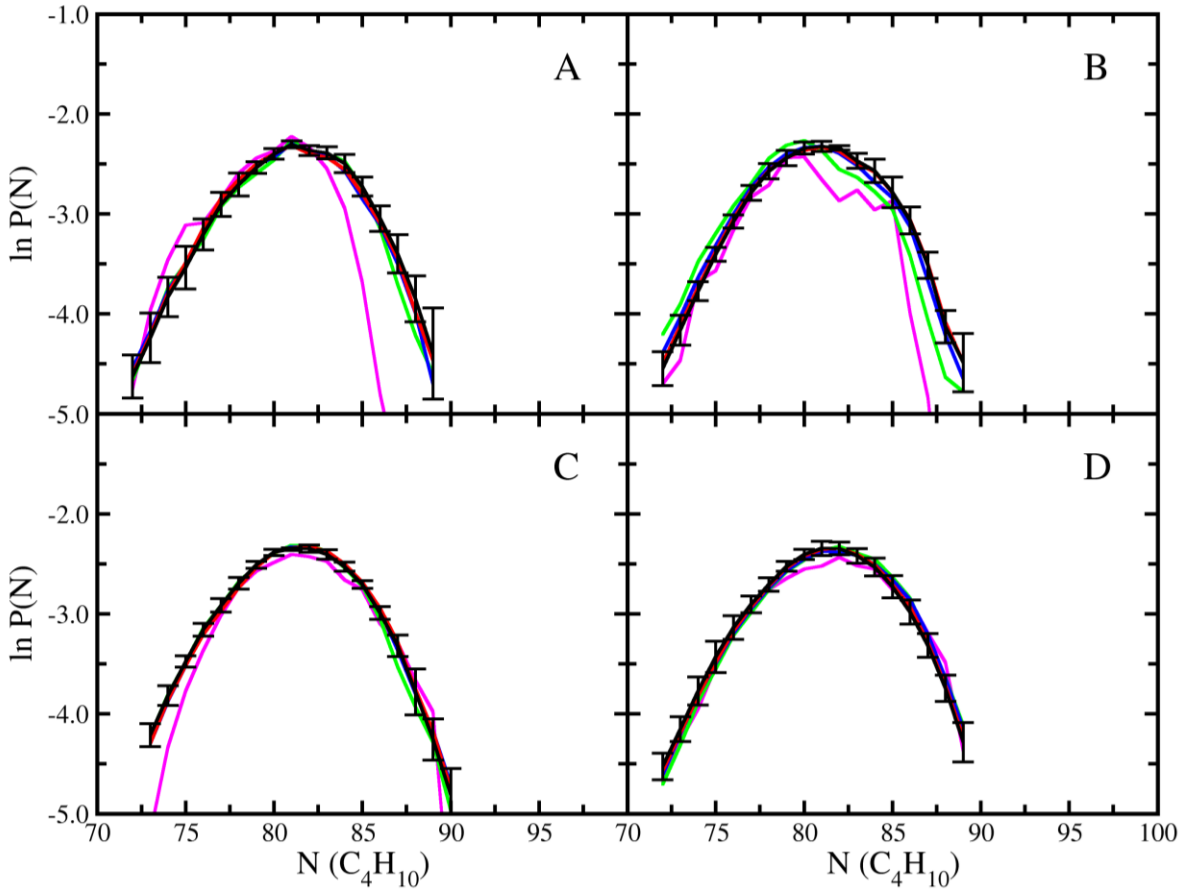

Figure S7: Molecule probability distribution in methane+n-butane binary mixture system at $x_{methane} = 0.3$ and 255 K. Lines in magenta, green, blue, red, and black represent the probability distribution of n-butane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using ME-1, ME-2, and ME-3 method with exchanging ratio of two methane molecules with one n-butane, respectively.

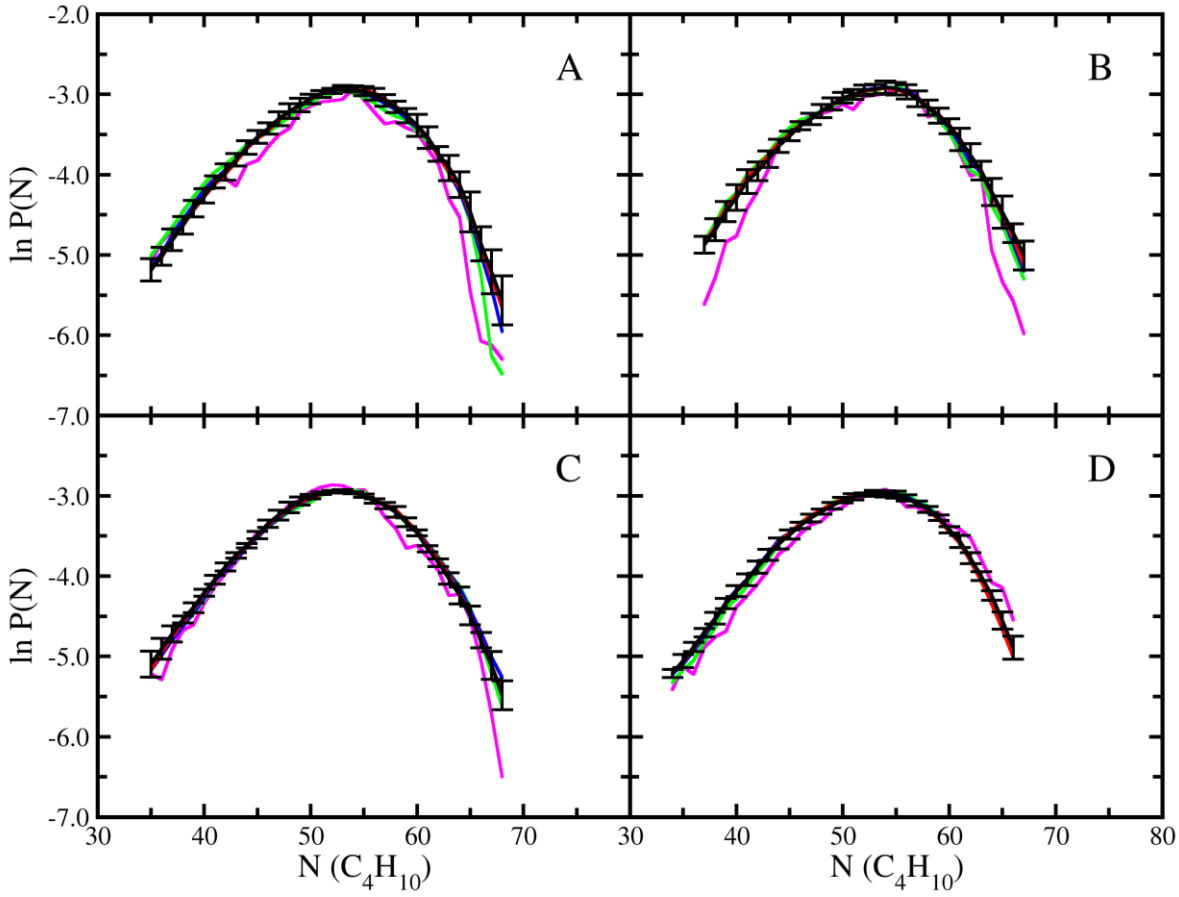

Figure S8: Molecule probability distribution in methane+n-butane binary mixture system at $x_{methane} = 0.6$ and 255 K. Lines in magenta, green, blue, red, and black represent the probability distribution of n-butane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using ME-1, ME-2, and ME-3 method with exchanging ratio of two methane molecules with one n-butane, respectively.

In Figure S9, the effect of CBMC parameters on perfluorobutane insertion/deletion acceptance and acceptance efficiency of standard CBMC and ME methods are provided. For perfluorobutane+butane with an exchange ratio of 1, both ME-2 and ME-3 are independent from first site atom trials, while ME-1 and standard CBMC are dependent to this variable. The maximum acceptance 0.09% and acceptance efficiency 0.91 (1/sec) for standard CBMC is achieved at 18 trials for the first atom site and 12 trials for remaining atoms. In the ME-1 method, increasing both variables would lead to increases in acceptance but decreases in the acceptance efficiency. Using 2 trials for the GC position and 1 trial for molecular rotation

results in an acceptance rate of 0.008% and acceptance efficiency of 0.45 (1/sec). In the case of ME-2, by increasing the number of secondary site trials, the acceptance increases while acceptance efficiency decreases. The maximum acceptance efficiency of 114 (1/sec) is achieved by using 1 trial for molecular rotation, which leads to 1.65% acceptance. The behavior of the ME-3 method is similar to the standard CBMC method, where the maximum acceptance of 3.85% and acceptance efficiency of 26.3 (1/sec) was achieved by using 18 trials for the first atom and 12 trials for the remaining atoms. Comparing acceptance efficiency of ME methods with standard CBMC using the optimum CBMC parameters, ME-2 and ME-3 are 120 and 28 more efficient, respectively. For ME-1, acceptance efficiency decreases by a factor of 2.

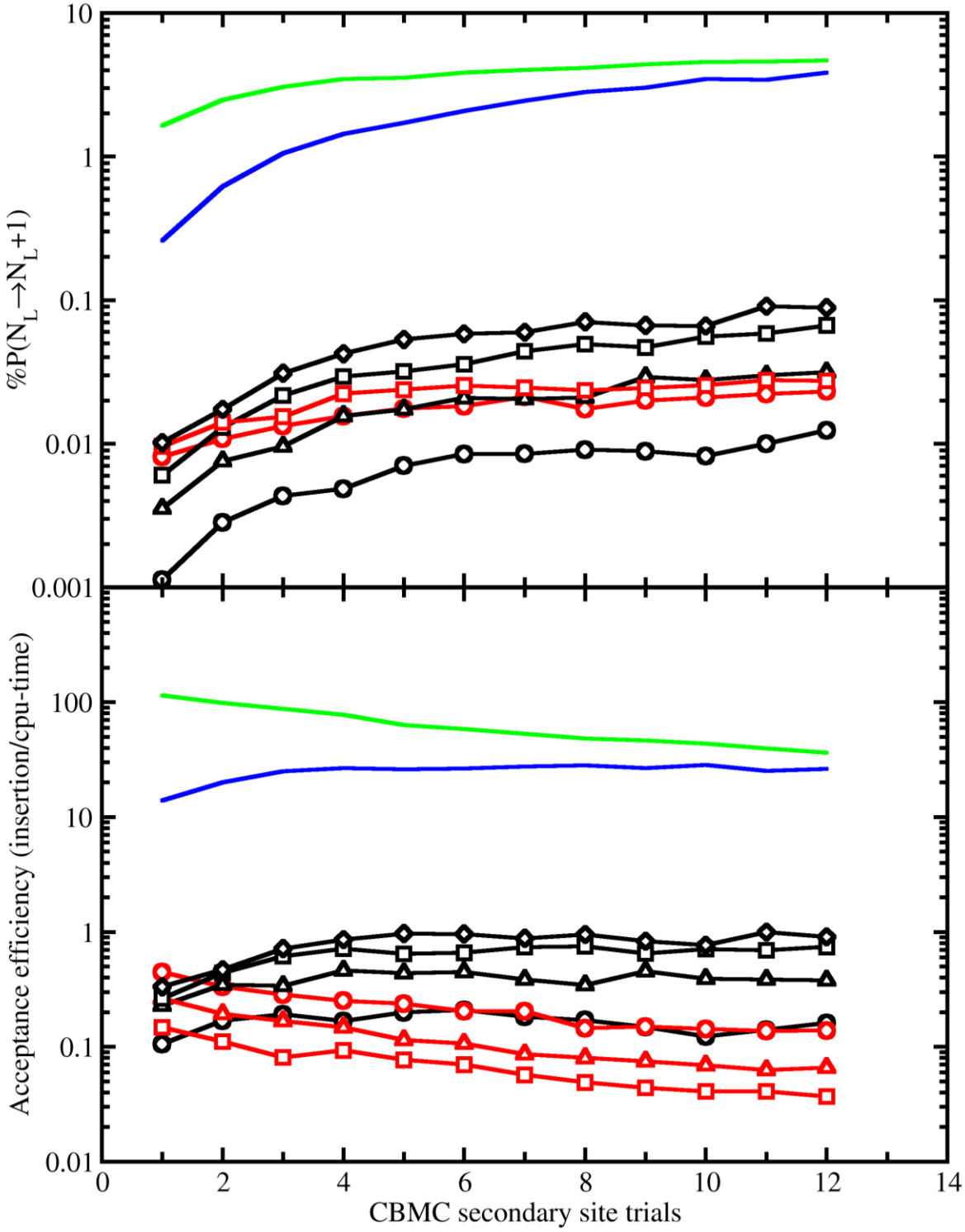

Figure S9: Acceptance and acceptance efficiency in perfluorobutane+n-butane binary mixture at 259.95 K and composition of 0.5. Lines represents acceptance and acceptance efficiency of perfluorobutane insertion in various CBMC trials for the site. Standard CBMC (black), ME-1 (red), ME-2 (green), and ME-3 (blue), 2 trials(circle), 6 trials (squares), 12 trials (triangles), 18 trials (diamonds). MEMC moves were performed with an exchange ratio of one to one.

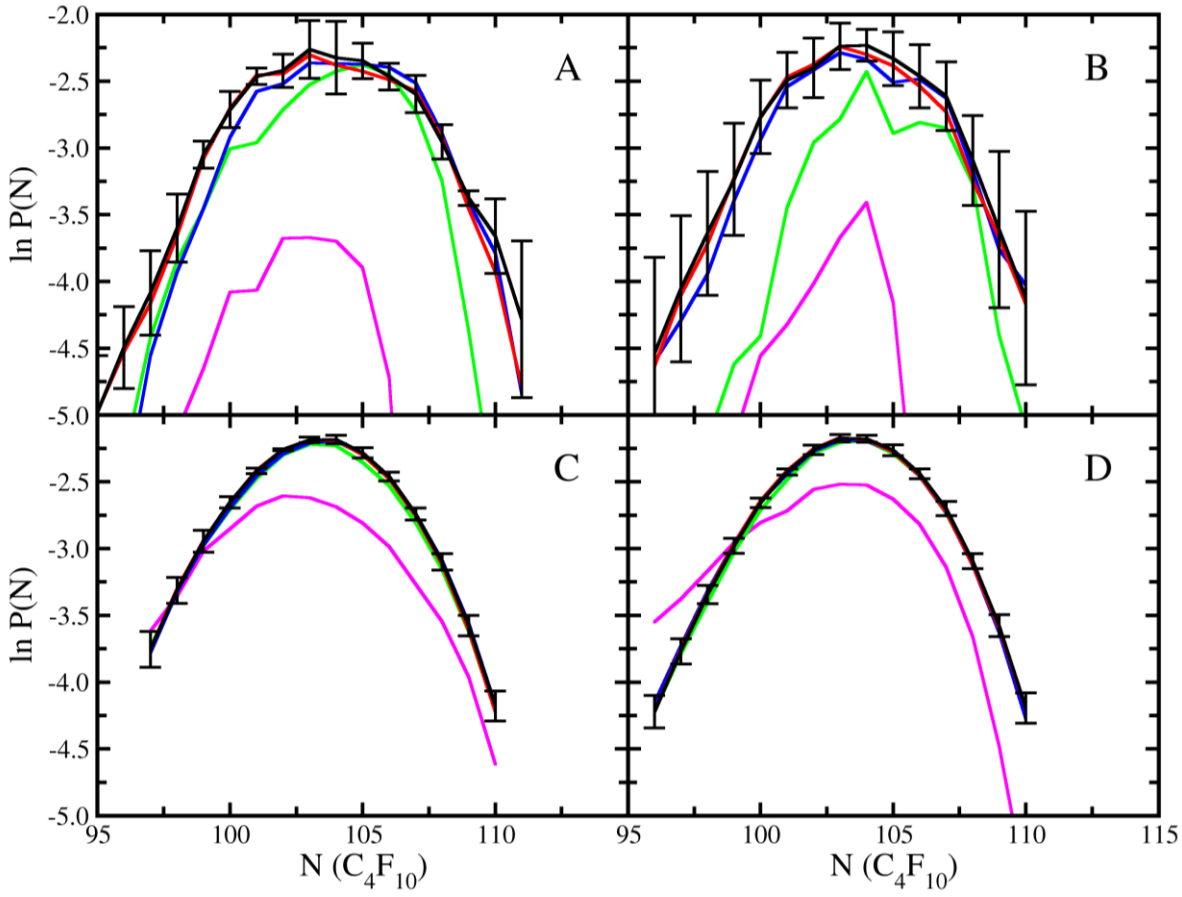

Figure S10: Molecule probability distribution in perfluorobutane+n-butane binary mixture system at $x_{butane} = 0.1$ and 259.95 K. Lines in magenta, green, blue, red, and black represent the probability distribution of perfluorobutane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using the ME-1, ME-2, and ME-3 method with exchanging ratio of one n-butane with one perfluorobutane, respectively.

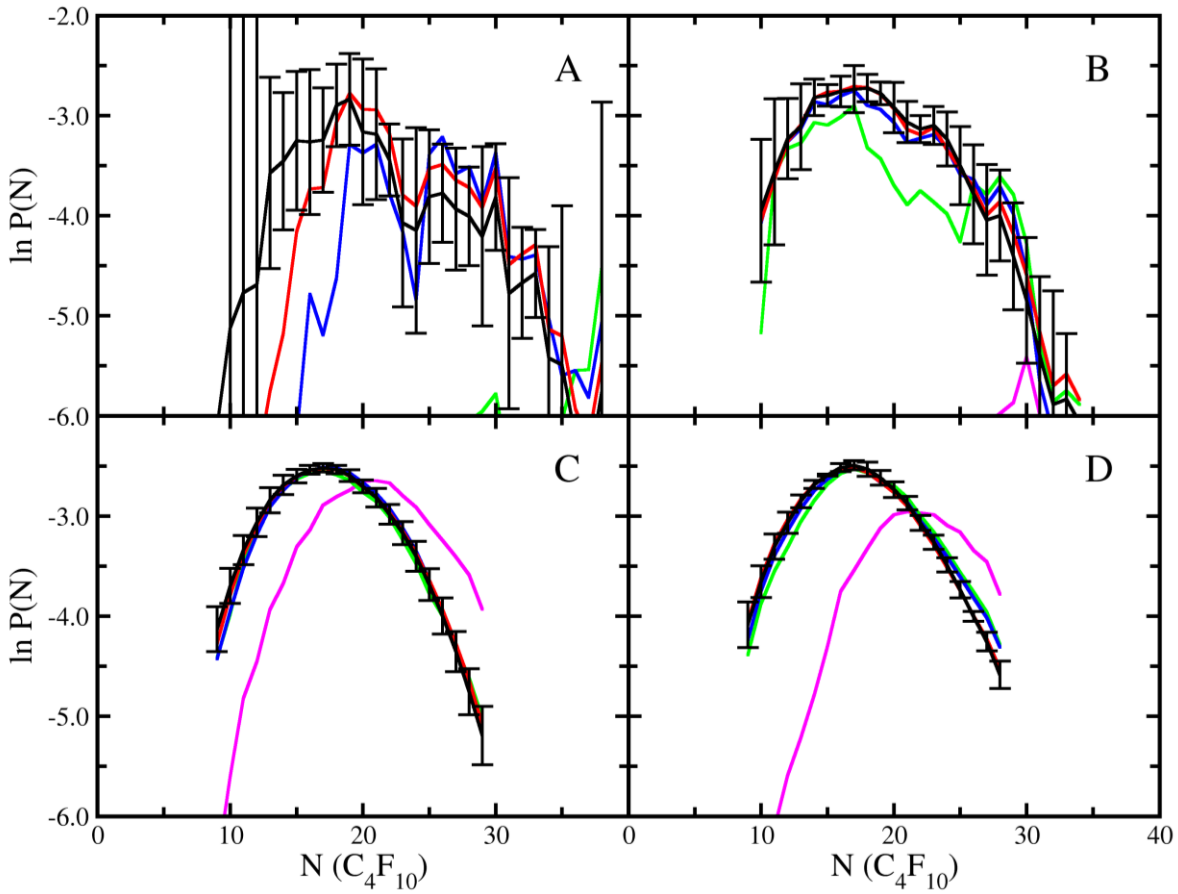

Figure S11: Molecule probability distribution in perfluorobutane+n-butane binary mixture system at $x_{butane} = 0.9$ and 259.95 K. Lines in magenta, green, blue, red, and black represent the probability distribution of perfluorobutane after 1, 5, 10, 15, and 20 million MC steps, respectively. (A) represent probability distribution using standard insertion and deletion with coupled-decoupled CBMC technique, (B), (C), and (D) represent probability distribution using ME-1, ME-2, and ME-3 method with exchanging ratio of one n-butane with one perfluorobutane, respectively.

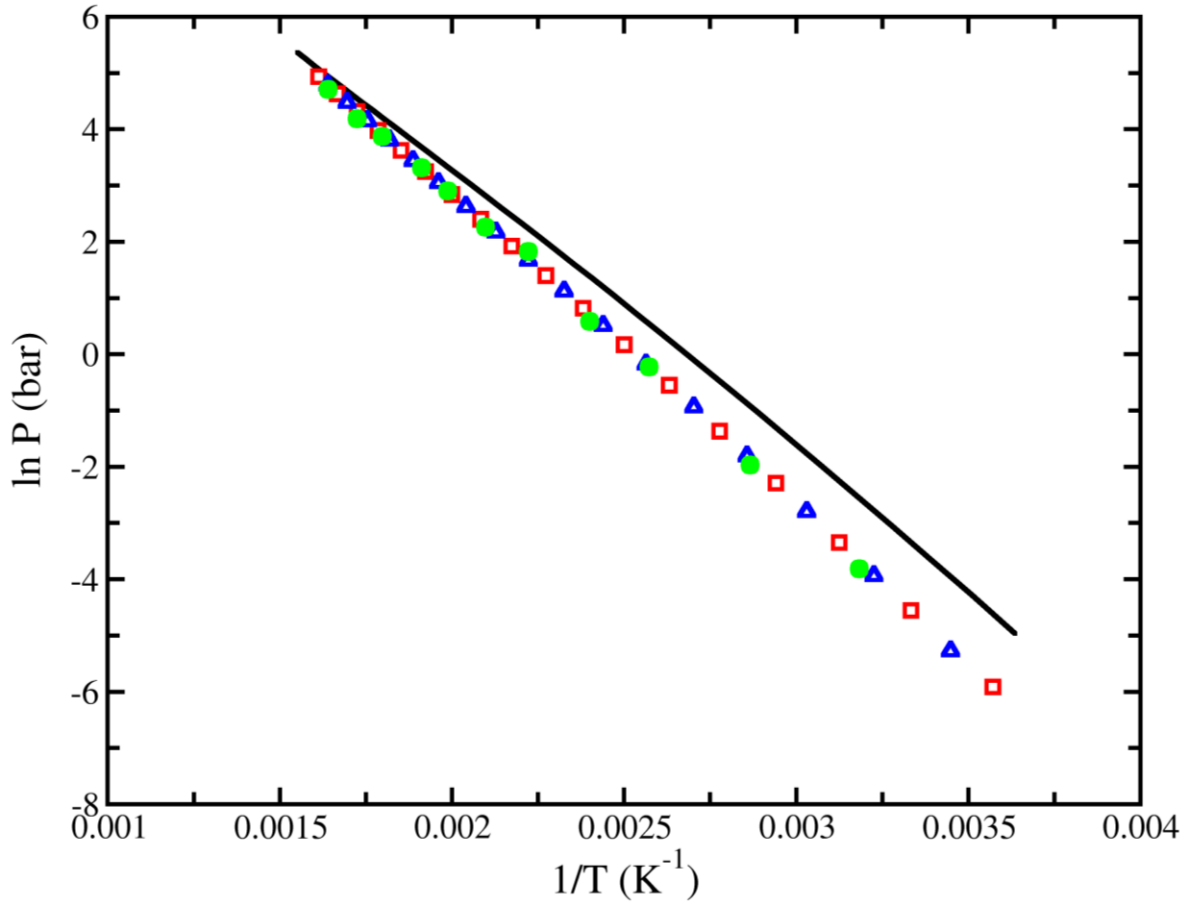

Figure S12: Clausius-Clapeyron plot for SPC/E water predicted from GCMC+histogram reweighting simulations. NIST Chemistry WebBook[66] (solid lines), values obtained by Boulougouris et. al.[65], (green circles) ME-2 algorithm (red squares), and ME-3 algorithm (blue triangles).

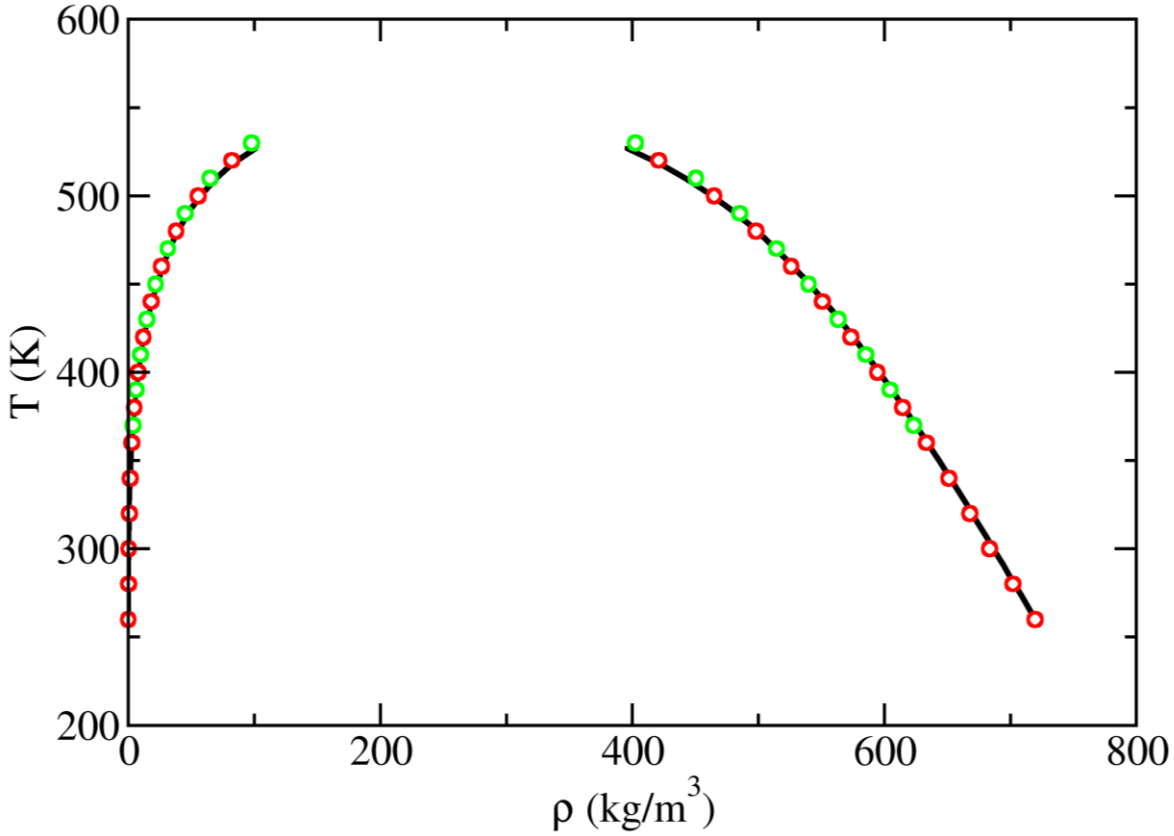

Figure S13: Vapor-liquid coexistence curve for 2,2,4-trimethylpentane predicted from GCMC+histogram reweighting simulations using Mie potentials[9]. Experimental data (solid lines)[69], ME-3 algorithm (red circles), and prior calculations using only configurational-bias Monte Carlo (green circles)[9].

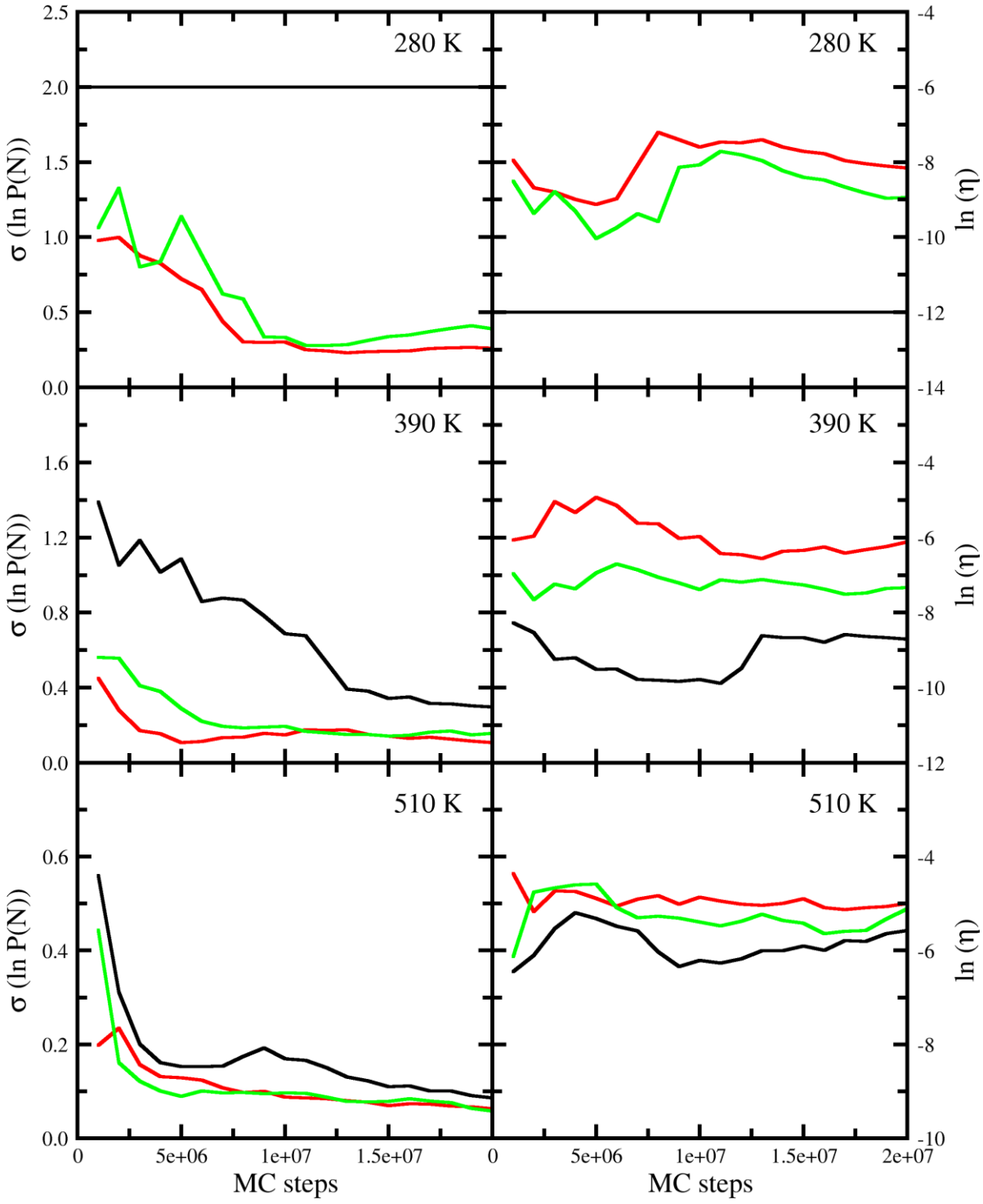

Figure S14: Standard deviation (left panel) and efficiency (right panel) for GCMC simulations for 2,2,4-trimethylpentane in the liquid phase. Configurational-bias insertions (black), ME-2 (red) and ME-3 (green).

REFERENCES

[1] de Pablo, J. J., Laso, M., Siepmann, J. I. and Suter, U. W. Continuum-Configurational-Bias Monte-Carlo Simulations of Long-Chain Alkanes. *Mol Phys*, 80, 1 (Sep 1993), 55-63.

[2] Siepmann, J. I. and Frenkel, D. Configurational Bias Monte-Carlo - a New Sampling Scheme for Flexible Chains. *Mol Phys*, 75, 1 (Jan 1992), 59-70.

[3] Wilmer, C. E. and Snurr, R. Q. Towards rapid computational screening of metal-organic frameworks for carbon dioxide capture: Calculation of framework charges via charge equilibration. *Chem Eng J*, 171, 3 (Jul 15 2011), 775-781.

[4] Kim, J., Lin, L. C., Martin, R. L., Swisher, J. A., Haranczyk, M. and Smit, B. Large-Scale Computational Screening of Zeolites for Ethane/Ethene Separation. *Langmuir*, 28, 32 (Aug 14 2012), 11923-11928.

[5] Banerjee, D., Simon, C. M., Plonka, A. M., Motkuri, R. K., Liu, J., Chen, X. Y., Smit, B., Parise, J. B., Haranczyk, M. and Thallapally, P. K. Metal-organic framework with optimally selective xenon adsorption and separation. *Nat Commun*, 7 (Jun 2016).

[6] Wilmer, C. E., Leaf, M., Lee, C. Y., Farha, O. K., Hauser, B. G., Hupp, J. T. and Snurr, R. Q. Large-scale screening of hypothetical metal-organic frameworks. *Nat Chem*, 4, 2 (Feb 2012), 83-89.

[7] Ferrenberg, A. M. and Swendsen, R. H. Optimized Monte-Carlo Data-Analysis. *Phys Rev Lett*, 63, 12 (Sep 18 1989), 1195-1198.

[8] Ferrenberg, A. M. and Swendsen, R. H. New Monte-Carlo Technique for Studying Phase-Transitions. *Phys Rev Lett*, 61, 23 (Dec 5 1988), 2635-2638.

[9] Mick, J. R., Barhaghi, M. S., Jackman, B., Schwiebert, L. and Potoff, J. J. Optimized Mie Potentials for Phase Equilibria: Application to Branched Alkanes. *J Chem Eng Data*, 62, 6 (Jun 2017), 1806-1818.

[10] Errington, J. R. and Panagiotopoulos, A. Z. A new intermolecular potential model for the n-alkane homologous series. *J Phys Chem B*, 103, 30 (Jul 29 1999), 6314-6322.

[11] Floriano, M. A., Caponetti, E. and Panagiotopoulos, A. Z. Micellization in model surfactant systems. *Langmuir*, 15, 9 (Apr 27 1999), 3143-3151.

[12] Snurr, R. Q., Bell, A. T. and Theodorou, D. N. Prediction of Adsorption of Aromatic-Hydrocarbons in Silicalite from Grand-Canonical Monte-Carlo Simulations with Biased Insertions. *J Phys Chem-Us*, 97, 51 (Dec 23 1993), 13742-13752.

[13] Smit, B. Grand-Canonical Monte-Carlo Simulations of Chain Molecules - Adsorption-Isotherms of Alkanes in Zeolites. *Mol Phys*, 85, 1 (May 1995), 153-172.

[14] Martin, M. G. and Siepmann, J. I. Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes. *J Phys Chem B*, 103, 21 (May 27 1999), 4508-4517.

[15] Errington, J. R. and Panagiotopoulos, A. Z. New intermolecular potential models for benzene and cyclohexane. *J Chem Phys*, 111, 21 (Dec 1 1999), 9731-9738.

[16] Macedonia, M. D. and Maginn, E. J. A biased grand canonical Monte Carlo method for simulating adsorption using all-atom and branched united atom models. *Mol Phys*, 96, 9 (May 10 1999), 1375-1390.

[17] Stapleton, M. R. and Panagiotopoulos, A. Z. Application of Excluded Volume Map Sampling to Phase-Equilibrium Calculations in the Gibbs Ensemble. *J Chem Phys*, 92, 2 (Jan 15 1990), 1285-1293.

[18] Ortiz, V., Maury-Evertsz, J. R. and Lopez, G. E. Parallel tempering-cavity-bias algorithm in the Gibbs ensemble. *Chem Phys Lett*, 368, 3-4 (Jan 17 2003), 452-457.

[19] Mezei, M. A Cavity-Biased (T,V,Mu) Monte-Carlo Method for the Computer-Simulation of Fluids. *Mol Phys*, 40, 4 (1980), 901-906.

[20] Shi, W. and Maginn, E. J. Continuous fractional component Monte Carlo: An adaptive biasing method for open system atomistic simulations. *J Chem Theory Comput*, 3, 4 (Jul-Aug 2007), 1451-1463.

[21] Torres-Knoop, A., Balaji, S. P., Vlugt, T. J. H. and Dubbeldam, D. A Comparison of Advanced Monte Carlo Methods for Open Systems: CFCMC vs CBMC. *J Chem Theory Comput*, 10, 3 (Mar 2014), 942-952.

[22] Escobedo, F. A. and Martinez-Veracoechea, F. J. Optimization of expanded ensemble methods. *J Chem Phys*, 129, 15 (Oct 21 2008).

[23] Escobedo, F. A. and dePablo, J. J. Expanded grand canonical and Gibbs ensemble Monte Carlo simulation of polymers. *J Chem Phys*, 105, 10 (Sep 8 1996), 4391-4394.

[24] Panagiotopoulos, A. Z. Exact Calculations of Fluid-Phase Equilibria by Monte-Carlo Simulation in a New Statistical Ensemble. *Int J Thermophys*, 10, 2 (Mar 1989), 447-457.

[25] Mehta, M. and Kofke, D. A. Implementation of the Gibbs Ensemble Using a Thermodynamic Model for One of the Coexisting Phases. *Mol Phys*, 79, 1 (May 1993), 39-52.

[26] Kofke, D. A. and Glandt, E. D. Monte-Carlo Simulation of Multicomponent Equilibria in a Semigrand Canonical Ensemble. *Mol Phys*, 64, 6 (Aug 20 1988), 1105-1131.

[27] Panagiotopoulos, A. Z., Quirke, N., Stapleton, M. and Tildesley, D. J. Phase-Equilibria by Simulation in the Gibbs Ensemble - Alternative Derivation, Generalization and Application to Mixture and Membrane Equilibria. *Mol Phys*, 63, 4 (Mar 1988), 527-545.

[28] Depablo, J. J. and Prausnitz, J. M. Phase-Equilibria for Fluid Mixtures from Monte-Carlo Simulation. *Fluid Phase Equilibr*, 53 (Dec 1989), 177-189.

[29] Kindt, J. T. Grand canonical Monte Carlo using solvent repacking: Application to phase behavior of hard disk mixtures. *J Chem Phys*, 143, 12 (Sep 28 2015).

[30] Vink, R. L. C. and Horbach, J. Grand canonical Monte Carlo simulation of a model colloid-polymer mixture: Coexistence line, critical behavior, and interfacial tension. *J Chem Phys*, 121, 7 (Aug 15 2004), 3253-3258.

[31] Wijmans, C. M., Smit, B. and Groot, R. D. Phase behavior of monomeric mixtures and polymer solutions with soft interaction potentials. *J Chem Phys*, 114, 17 (May 1 2001), 7644-7654.

[32] Guo, Z. and Kindt, J. T. Gibbs ensemble Monte Carlo with solvent repacking: phase coexistence of size–asymmetrical binary Lennard-Jones mixtures. *Molecular Simulation*, 44, 4 (2018/03/04 2018), 300-308.

[33] Martin, M. G. and Siepmann, J. I. Predicting multicomponent phase equilibria and free energies of transfer for alkanes by molecular simulation. *Journal of the American Chemical Society*, 119, 38 (Sep 24 1997), 8921-8924.

[34] Chen, B., Siepmann, J. I. and Klein, M. L. Vapor-liquid interfacial properties of mutually saturated water/1-butanol solutions. *Journal of the American Chemical Society*, 124, 41 (Oct 16 2002), 12232-12237.

[35] Wick, C. D., Siepmann, J. I. and Theodorou, D. N. Microscopic origins for the favorable solvation of carbonate ether copolymers in CO2. *Journal of the American Chemical Society*, 127, 35 (Sep 7 2005), 12338-12342.

[36] Barhaghi, M. S., Torabi, K., Nejahi, Y., Schwiebert, L. and Potoff, J. J. Molecular exchange Monte Carlo: A generalized method for identity exchanges in grand canonical Monte Carlo simulations. *J Chem Phys*, 149, 7 (Aug 21 2018), 072318.

[37] Nejahi, Y., Soroush Barhaghi, M., Mick, J., Jackman, B., Rushaidat, K., Li, Y., Schwiebert, L. and Potoff, J. GOMC: GPU Optimized Monte Carlo for the simulation of phase equilibria and physical properties of complex fluids. *SoftwareX*, 9 (2019/01/01/ 2019), 20-27.

[38] GOMC: http://github.com/GOMC-WSU. *GOMC: http://github.com/GOMC-WSU* (

[39] Ewald, P. P. The calculation of optical and electrostatic grid potential. *Annalen Der Physik*, 64, 3 (Feb 1921), 253-287.

[40] Martinez, L., Andrade, R., Birgin, E. G. and Martinez, J. M. PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J Comput Chem*, 30, 13 (Oct 2009), 2157-2164.

[41] Humphrey, W., Dalke, A. and Schulten, K. VMD: Visual molecular dynamics. *J Mol Graph Model*, 14, 1 (Feb 1996), 33-38.

[42] Frenkel, D. and Smit, B. *Understanding molecular simulation: from algorithms to applications*. Academic Press, 2002.

[43] Bai, P. and Siepmann, J. I. Assessment and Optimization of Configurational-Bias Monte Carlo Particle Swap Strategies for Simulations of Water in the Gibbs Ensemble. *J Chem Theory Comput*, 13, 2 (Feb 2017), 431-440.

[44] Torres-Knoop, A., Burtch, N. C., Poursaeidesfahani, A., Balaji, S. P., Kools, R., Smit, F. X., Walton, K. S., Vlugt, T. J. H. and Dubbeldam, D. Optimization of Particle Transfers in the Gibbs Ensemble for Systems with Strong and Directional Interactions Using CBMC, CFCMC, and CB/CFCMC. *J Phys Chem C*, 120, 17 (May 5 2016), 9148-9159.

[45] Shi, W. and Maginn, E. J. Improvement in Molecule Exchange Efficiency in Gibbs Ensemble Monte Carlo: Development and Implementation of the Continuous Fractional Component Move. *J Comput Chem*, 29, 15 (Nov 30 2008), 2520-2530.

[46] Chu, T. C., Chen, R. J. J., Chappelear, P. S. and Kobayashi, R. Vapor-Liquid-Equilibrium of Methane-Normal-Pentane System at Low-Temperatures and High-Pressures. *J Chem Eng Data*, 21, 1 (1976), 41-44.

[47] Wichterle, I. and Kobayashi, R. Vapor-Liquid Equilibrium of Methane-Ethane System at Low-Temperatures and High-Pressures. *J Chem Eng Data*, 17, 1 (1972), 9-+.

[48] Wichterle, I. and Kobayashi, R. Vapor-Liquid Equilibrium of Methane-Propane System at Low-Temperatures and High-Pressures. *J Chem Eng Data*, 17, 1 (1972), 4-+.

[49] Beaudoin, J. M. and Kohn, J. P. Multiphase and Volumetric Equilibria of Methane-N-Decane Binary System at Temperatures between -36 Degrees and 150 Degrees C. *J Chem Eng Data*, 12, 2 (1967), 189-&.

[50] Poston, R. S. and Mcketta, J. J. Vapor-Liquid Equilibrium in Methane-N-Hexane System. *J Chem Eng Data*, 11, 3 (1966), 362-&.

[51] Chang, H. L., Hurt, L. J. and Kobayash.R Vapor-Liquid Equilibria of Light Hydrocarbons at Low Temperatures and High Pressures - Methane-N-Heptane System. *Aiche J*, 12, 6 (1966), 1212-&.

[52] Kohn, J. P. and Bradish, W. F. Multiphase and Volumetric Equilibria of the Methane-n-Octane System at Temperatures between -110 and 150 C. *J Chem Eng Data*, 9, 1 (1964), 5-8.

[53] Elliot, D. G., Chen, R. J. J., Chappele.Ps and Kobayash.R Vapor-Liquid-Equilibrium of Methane-N-Butane System at Low-Temperatures and High-Pressures. *J Chem Eng Data*, 19, 1 (1974), 71-77.

[54] Potoff, J. J. and Bernard-Brunel, D. A. Mie Potentials for Phase Equilibria Calculations: Application to Alkanes and Perfluoroalkanes. *J Phys Chem B*, 113, 44 (Nov 5 2009), 14725-14731.

[55] Potoff, J. J., Errington, J. R. and Panagiotopoulos, A. Z. Molecular simulation of phase equilibria for mixtures of polar and non-polar components. *Mol Phys*, 97, 10 (Nov 20 1999), 1073-1083.

[56] Chen, B. and Siepmann, J. I. Transferable potentials for phase equilibria. 3. Explicit-hydrogen description of normal alkanes. *J Phys Chem B*, 103, 25 (Jun 24 1999), 5370-5379.

[57] McCabe, C., Galindo, A., Gil-Villegas, A. and Jackson, G. Predicting the high-pressure phase equilibria of binary mixtures of perfluoro-n-alkanes plus n-alkanes using the SAFT-VR approach. *J Phys Chem B*, 102, 41 (Oct 8 1998), 8060-8069.

[58] Aparicio, S. Phase equilibria in perfluoroalkane plus alkane binary systems from PC-SAFT equation of state. *J Supercrit Fluid*, 46, 1 (Aug 2008), 10-20.

[59] Haley, J. D. and McCabe, C. Predicting the phase behavior of fluorinated organic molecules using the GC-SAFT-VR equation of state. *Fluid Phase Equilibr*, 440 (May 25 2017), 111-121.

[60] Simons, J. H. and Mausteller, J. W. The Properties of Normal-Butforane and Its Mixtures with Normal-Butane. *J Chem Phys*, 20, 10 (1952), 1516-1519.

[61] Zhang, L. and Siepmann, J. I. Pressure dependence of the vapor-liquid-liquid phase behavior in ternary mixtures consisting of n-alkanes, n-perfluoroalkanes, and carbon dioxide. *J Phys Chem B*, 109, 7 (Feb 24 2005), 2911-2919.

[62] Lorentz, H. A. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Annalen der Physik*, 248, 1 (1881), 127-136.

[63] Berthelot, D. Sur le melange des gaz. *C. R. Hebd. Seanc. Acad. Sci. (Paris)*, 126 (1898), 1703-1855.

[64] Berendsen, H. J. C., Grigera, J. R. and Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J Phys Chem-Us*, 91, 24 (Nov 19 1987), 6269-6271.

[65] Boulougouris, G. C., Economou, I. G. and Theodorou, D. N. Engineering a molecular model for water phase equilibrium over a wide temperature range. *J Phys Chem B*, 102, 6 (Feb 5 1998), 1029-1035.

[66] Lemmon, E. W., O., M. M. and Friend, D. G. Thermophysical Properties of Fluid Systems. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom*

and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, *http://webbook.nist.gov*

[67] Martin, M. G. and Frischknecht, A. L. Using arbitrary trial distributions to improve intramolecular sampling in configurational-bias Monte Carlo. *Mol Phys*, 104, 15 (Aug 10 2006), 2439-2456.

[68] Sepehri, A., Loeffler, T. D. and Chen, B. Improving the Efficiency of Configurational-Bias Monte Carlo: A Jacobian-Gaussian Scheme for Generating Bending Angle Trials for Linear and Branched Molecules. *J Chem Theory Comput*, 13, 4 (Apr 2017), 1577-1583.

[69] Smith, B. D. and Srivastava, R. *Thermodynamic Data for Pure Compounds: Part A Hydrocarbons and Ketones*. Elsevier, Amsterdam, 1986.

[70] Lorentz, H. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Annalen der Physik*, 248, 1 (1881), 127-136.

[71] Berthelot, D. Hebd. *Seances Acad. Sci*, 126 (1898), 1703.

[72] Mick, J. R., Soroush Barhaghi, M., Jackman, B., Rushaidat, K., Schwiebert, L. and Potoff, J. J. Optimized Mie potentials for phase equilibria: Application to noble gases and their mixtures with n-alkanes. *J Chem Phys*, 143, 11 (Sep 21 2015), 114504.

[73] Mick, J. R., Barhaghi, M. S. and Potoff, J. J. Prediction of Radon-222 Phase Behavior by Monte Carlo Simulation. *J Chem Eng Data*, 61, 4 (Apr 2016), 1625-1631.

[74] Barhaghi, M. S., Mick, J. R. and Potoff, J. J. Optimised Mie potentials for phase equilibria: application to alkynes. *Mol Phys*, 115, 9-12 (2017), 1378-1388.

[75] Jorgensen, W. L., Madura, J. D. and Swenson, C. J. Optimized intermolecular potential functions for liquid hydrocarbons. *Journal of the American Chemical Society*, 106, 22 (1984), 6638-6646.

[76] Jorgensen, W. L., Maxwell, D. S. and Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118, 45 (1996), 11225-11236.

ABSTRACT

MOLECULAR EXCHANGE MONTE CARLO. A GENERALIZED METHOD FOR IDENTITY EXCHANGES IN GRAND CANONICAL MONTE CARLO SIMULATIONS

by

MOHAMMAD SOROUSH BARHAGHI

August 2019

Advisor: Dr. Loren Schwiebert

Major: Computer Science

Degree: Master of Science

A generalized identity exchange algorithm is presented for Monte Carlo simulations in the grand canonical ensemble. The algorithm, referred to as Molecular Exchange Monte Carlo (MEMC), may be applied to multicomponent systems of arbitrary molecular topology, and provides significant enhancements in the sampling of phase space over a wide range of compositions and temperatures. Three different approaches are presented for the insertion of large molecules, and the pros and cons of each method are discussed. The performance of the algorithms is highlighted through grand canonical Monte Carlo histogram-reweighting simulations performed on several systems, including 2,2,4-trimethylpentane+neopentane, butane+perfluorobutane, methane+n-alkanes, and water+impurity. Relative acceptance efficiencies of up to 400 times that of standard configurational-bias Monte Carlo are obtained for molecule transfers.

AUTOBIOGRAPHICAL STATEMENT

MOHAMMAD SOROUSH BARHAGHI
5050 Anthony Wayne Dr. Detroit, MI, 48201
(313)980-8605

Education:

Ph.D. Chemical Engineering and Materials Science
Wayne State University, Detroit, US, 2014 – Present
Advisor: Dr. Jeffrey Potoff

M.S. Computer Science
Wayne State University, Detroit, US, 2017 – 2019
Advisor: Dr. Loren Schwiebert

B.S. Chemical Engineering
University of Tehran, Tehran, Iran, 2008 – 2012
Advisor: Dr. Omid Tavakoli

Publication:

1. Richard A. Messerly, Mohammad Soroush Barhaghi, Jeffrey J. Potoff, and Michael R. Shirts, “Histogram-free reweighting with grand canonical Monte Carlo: Post-simulation optimization of non-bonded potentials for phase equilibria,” J. Chem. Eng. Data, (2019) 0021-9568.
2. Younes Nejahi, Mohammad Soroush Barhaghi, Jason Mick, Brock Jackman, Kamel Rushaidat, Yuanzhe Li, Loren Schwiebert, and Jeffrey J. Potoff, “GOMC: GPU Optimized Monte Carlo for the simulation of phase equilibria and physical properties of complex fluids,” J. SoftwareX, 9 (2019) 20-27.
3. Mohammad Soroush Barhaghi and Jeffrey J. Potoff, “Prediction of phase equilibria and Gibbs free energies of transfer using molecular exchange Monte Carlo in the Gibbs ensemble,” J. Fluid Phase Equilibria, 486 (2019) 106-118.
4. Mohammad Soroush Barhaghi, Korosh Torabi, Younes Nejahi, Loren Schwiebert, Jeffrey J. Potoff, “Molecular exchange Monte Carlo: A generalized method for identity exchanges in grand canonical Monte Carlo simulations,” J. Chem. Phys., 149 (2018) 072318.
5. Mohammad Soroush Barhaghi, Jason R. Mick, and Jeffrey J. Potoff, “Optimized Mie potentials for phase equilibria: application to alkynes,” J. Mol. Phys., 115 (2017) 1378-1388.
6. Jason R. Mick, Mohammad Soroush Barhaghi, Brock Jackman, Loren Schwiebert, and Jeffrey J. Potoff, “Optimized Mie Potentials for Phase Equilibria: Application to Branched Alkanes,” J. Chem. Eng. Data, 62 (2017) 1806-1818.
7. Jason R. Mick, Mohammad Soroush Barhaghi, and Jeffrey J. Potoff, "Prediction of Radon-222 Phase Behavior by Monte Carlo Simulation, " J. Chem. Eng. Data, 61 (2016) 1625-1631.
8. Jason R. Mick, Mohammad Soroush Barhaghi, Brock Jackman, Kamel Rushaidat, Loren Schwiebert, and Jeffrey J. Potoff, "Optimized Mie potentials for phase equilibria: Application to noble gases and their mixtures with n-alkanes," J. Chem. Phys., 143 (2015) 114504.