Ahmadreza Rahbari

Thermodynamics of Industrially Relevant Systems: Method Development and Applications

Delft University of Technology

Thermodynamics of Industrially Relevant Systems: Method Development and Applications

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
donderdag, 16 april, 2020 om 12:30 uur

door

Ahmadreza RAHBARI
Mechanical Engineer, Norwegian University of Science and Technology,
Noorwegen
geboren te Teheran, Iran.

Dit proefschrift is goedgekeurd door de
promotor: Prof. dr. ir. T. J. H. Vlugt
promotor: Dr. D. Dubbeldam

Samenstelling promotiecommissie:

Rector Magnificus	Voorzitter
Prof. dr. ir. T.J.H. Vlugt	Technische Universiteit Delft, promotor
Dr. D. Dubbeldam	Universiteit van Amsterdam, promotor

Onafhankelijke leden:

Prof.dr. N. van der Vegt	Technische Universiteit Darmstadt
Prof.dr.-ing.habil. J. Vrabec	Technische Universiteit Berlin
Prof.dr. J.H. van Esch	Technische Universiteit Delft
Prof.dr.ir. W. de Jong	Technische Universiteit Delft
Dr. O.A. Moultos	Technische Universiteit Delft

This work was sponsored by NWO Exacte Wetenschappen (Physical Sciences) for the use of supercomputer facilities, with financial support from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research, NWO). TJHV acknowledges NWO-CW for a VICI grant.

Keywords: Statistical thermodynamics, Molecular simulations, Chemical equilibrium, Free energy calculations, (Partial) molar properties, Hydrogen, Formic acid, Ammonia, Water, Methanol

Printed by: Ridderprint B.V.
Cover Design by Ahmadreza Rahbari

ISBN: 978-94-6366-259-8

An electronic version of this dissertation is available at
http://repository.tudelft.nl/

Dedicated to my beloved parents
Alireza & Firouzeh and my dear sister
Azade

Contents

Chapter 1

Introduction

1.1 Classical Thermodynamics

Thermodynamics originates from the Greek words "Therme" (heat) and "dynamis" (force) literally meaning heat-force, or power developed from heat, and it refers to the science of heat, energy and work transfer in interaction with matter. The concept of energy may be vague at first, however it can be simply viewed as the ability to cause physical or chemical changes or to cause motion via interaction with matter [1]. Using the fundamental laws of thermodynamics, one can convert energy from one form to another, and determine the limits of what can be accomplished in terms of efficiency [2]. The first law of thermodynamics is the law of conservation of energy, and it states that energy cannot be "lost" or "created", but that it can be converted from one form to another (except for nuclear reactions where matter is converted to energy) [3]. It is important to note that the first law of thermodynamics does not predict whether a process can occur spontaneously (*e.g.* an object that moves up against the gravity field of the earth by extracting heat from the surrounding will not occur spontaneously). The second law of thermodynamics predicts which processes are spontaneous, possible and impossible. For more details about the fundamental thermodynamic laws, the reader is referred to Refs. [1–4].

Although the laws of thermodynamics have existed since the beginning of our universe, thermodynamics began officially as a science in the nineteenth century with the construction of heat engines [1, 3], to efficiently convert the thermal energy of hot bodies into work [3]. The scope of thermodynamics changed from designing heat engines to designing other mechanical equipments including turbines, compressors, heat pumps etc. [3]. To design this equipment, predictions based on thermodynamics of pure substances such as water (steam), different refrigerants, ammonia etc. play an important role. This area of thermodynamics is mostly central to mechanical engineers, while for chemical engineers studying systems of multicomponent (reacting) mixtures is more central (reacting, *e.g.* chemical reactors or non-reacting, *e.g.* separators). Producing value-added chemicals, for petrochemical, pharmaceutical, biological industries etc., involves chemical reactions of complex mixtures, separation and purification etc. [5]. Therefore, thermodynamic descriptions of multicomponent systems are crucial for process design and product development in chemical process design. One of the major challenges in chemical process design is lack of experimental data for multicomponent mixtures (*e.g.* ternary mixtures or multicomponent mixtures in general) at the specific temperatures or pressures [5]. At extreme conditions, high pressures, or using toxing materials, conducting experiments become more difficult. To describe and predict the behaviour of multicomponent mixtures, performing extrapolation based on limited sets of experimental data, or using thermodynamic models/relations are required. Chemical analytic methods may also be used to predict properties of (toxic) components at extreme conditions (high temperatures or pressures) before designing and performing expensive experiments [5]. In chemical engineering science, thermodynamic systems are mostly treated as macroscopic models to describe the overall behaviour of the system [3]. It turns out that the important system properties (such as density, volume, enthalpy etc.) can be evaluated relatively well without studying the molecular or atomic structure of the system. This is usually referred to as classical thermodynamics. One of the most commonly used tools in classical thermodynamics are cubic Equations of State (EoS) *e.g.* for natural gas components. Cubic EoS are widely used in industry [6, 7]. One

of the most commonly used cubic EoS are the Peng-Robinson (PR) and Soave-Redlich-Kwong (SRK) EoS [7, 8]. The performance of the cubic EoS are satisfactory for gasses (unless the pressure is too high). However, molar volumes of the liquid phase and fugacity coefficients at high pressures obtained from cubic EoS (with conventional mixing rules) are known to deviate significantly from experiments, and therefore modifications are required [9–12]. Up until now, more than 220 modifications of mixing rules for pure components and extensions to mixtures with the PR-EoS have been reported in literature [13].

It is well-known that classical thermodynamic models may fail to describe the phase equilibrium of mixtures involving polar components, even when using fitted parameters. For example, for the PR EoS, more than 220 different mixing rules for pure components and extensions to mixtures were reported [13]. This clearly shows that classical treatment of some systems is most likely not sufficient for analysis and process design of complex systems. Another important approach in studying behaviour of a system is based on the microscopic approach to thermodynamics. In the microscopic view of thermodynamics, the constituent molecules/atoms of the system are taken into account, see Fig. 1.1. The macroscopic/bulk properties of a substance depend on the behaviour of the molecules. This is usually referred to as Molecular Thermodynamics [14, 15]. Statistical molecular thermodynamics fills the gap between the quantum mechanics of individual molecules/atoms and classical thermodynamics of bulk phases. In an ensemble of molecules, the population of all possible molecular states (*i.e.* positions and momenta) leads to the mathematical formulation of the so-called ensemble partition function. For a simple system such as the 2D Ising model [16], an exact analytical solution is obtained from the ensemble partition function. For more complex systems, one could use more physically based theory with approximations to numerically calculate thermodynamic data of a system. Molecular simulation [14, 17] and Statistical Association Fluid Theory (SAFT) type EoS are examples of such approaches [18–20]. For more information on these topics the reader is referred to Refs. [4, 14, 15].

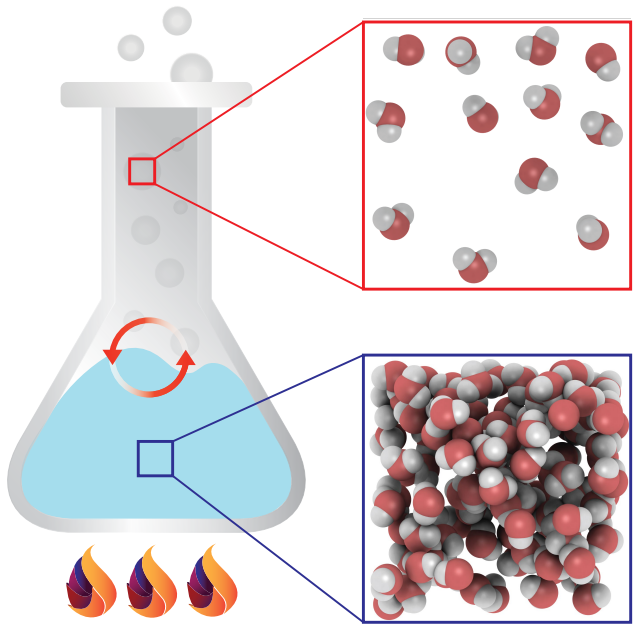


Figure 1.1: Illustration of classical approach to thermodynamics (left) and the statistical thermodynamics approach (right). Using principles of statistical mechanics, one can predict bulk properties of a substance (*e.g.* boiling point, heat capacity, enthalpy of vaporization etc.) based on the molecular structure and intermolecular interactions between the molecules in the system.

1.2 Molecular Simulation

Molecular simulations aim at describing the macroscopic properties of materials based on our knowledge of the properties of their constituent particles (molecules, atoms, etc.) and the interactions between them [14, 21, 22], see Fig. 1.2. Macroscopic properties such as pressure, heat capacity, heat of vaporization, chemical potential, diffusion coefficient and viscosity are all examples of important thermodynamic and transport properties which can be predicted from the collective behaviour of molecules. Assuming that the

intermolecular interactions are known, molecular simulations results can be very precise. This means that in principle, the error bars of computed properties can become infinitely small, if simulations are run infinitely long. Therefore, the precision is limited by the available computational power (and simulation technique). A comparison of computer processors over the past 60 years shows a 1 trillion fold increase in performance [23]. Due to (1) increased computational power since the early 1950s [24, 25], and (2) development of advanced calculation techniques, molecular simulations have become a very powerful tool to study complex molecular systems and to make predictions about properties of new substances [26]. Some of the distinct advantages or applications of modern molecular simulations can be mentioned as follows:

(a) One can obtain a molecular understanding of how properties of molecules result in the properties of a material. For example, molecular simulations can be used to understand/explain the correlation between negative enthalpy of mixing of water-methanol and the molecular structure of the mixture [27–29].

(b) Predict properties of materials without having to synthesize the materials first. For example, Ionic Liquids, ILs, (a class of liquid salts at room temperature) can be made from thousands of different combinations of anions and cations [30]. Molecular simulations can help screen many different possible combinations with different properties before synthesizing the IL for a specific purpose [31].

(c) Performing measurements/tests using dangerous materials or at extreme conditions can be accessed without problems/dangers. For example, Vapour-Liquid Equilibrium (VLE) calculations at high pressures [32], capture and degradation of nerve agents [33, 34], etc.

(d) Testing theories [14] and developing/improving simulation techniques. In this thesis, advanced simulation techniques are developed for studying multicomponent mixtures. These theories/methodologies were used to investigate the VLEs of pure components (water, methanol,

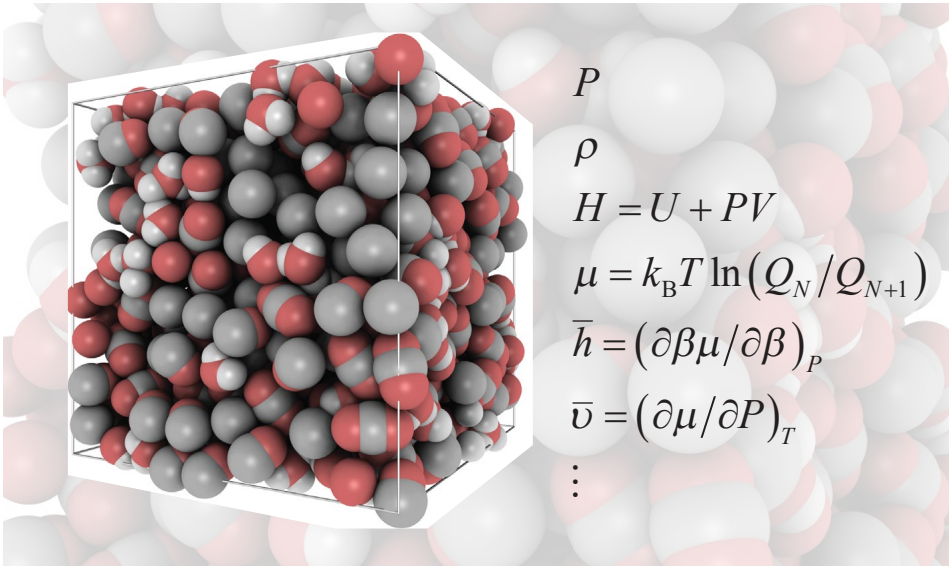


Figure 1.2: Macroscopic properties can be obtained using molecular simulations taking into account the properties of constituent particles in a system and interactions between them. Thermodynamic quantities such as the pressure P, density ρ and the total enthalpy are more straightforward to compute. In this thesis, advanced simulation techniques are developed/improved especially to compute chemical potentials μ, partial molar enthalpies $\bar{h}$ and partial molar volumes $\bar{v}$. Partial molar properties can be evaluated numerically by taking the derivatives of the chemical potential with respect to temperature and pressure: $\bar{h} = (\partial\beta\mu/\partial\beta)_P$ and $\bar{v} = (\partial\mu/\partial P)_T$ [8].

carbon dioxide, hydrogen sulfide, etc.), thermodynamic properties of multicomponent mixtures such as water-methanol mixtures, and reactive mixtures of ammonia, hydrogen and nitrogen.

Two main computational techniques used in molecular simulations are Monte Carlo (MC) and Molecular Dynamics (MD) [14, 15]. MD simulations are performed by numerically integrating classical equations of motions (Newton's second law), starting from initial particle configurations and momenta. MD simulations provide us with a powerful numerical approach to study the time evolution of a classical system of N particles with volume V [14, 15]. The results from MD simulations yield representative trajectories of the motion of a collection of a collection of molecules. This allows for computation of transport properties such as diffusion coefficients, viscosities, and thermal conductivities [35]. Conventional MD simulations in the NVE ensemble are deterministic, meaning that the results depend fully on the initial conditions (positions and momenta of the particles). Performing MD simulations at constant temperature can be either stochastic (*e.g.* using the Andersen thermostat) or deterministic (*e.g.* using the Nosé-Hoover thermostat) [14, 15]. It should be noted that by performing sufficiently long simulations and sampling a sufficient region of phase space, thermodynamic consistency is expected between the results obtained from deterministic and stochastic methods. For details, the reader is referred to Refs. [14, 15]. A MD simulation is performed for a fixed period of time to study time averages of processes, typically in maximum accessible time scale of 1 microsecond. To date, processes which require a longer time scale (*e.g.* charging the battery of your Tesla model 3 [36]) cannot be directly simulated with MD.

MC is a numerical technique for solving multidimensional integrals, or the ratio between the integrals [14, 21, 22]. MC simulations are stochastic in nature which implies an inherent randomness in the result. However, with sufficient and thorough sampling of phase space, thermodynamically consistent results are also expected from MC simulations [14, 15]. The MC simulation technique used in molecular simulations, and in this thesis, is the MC importance-sampling introduced by Metropolis, Rosenbluth and

co-workers in 1954 [22, 24]. As an illustrative example, the concepts used in MC importance-sampling (from now on referred to as MC) can be briefly explained using the classical formulation of the canonical ensemble partition function for a closed system of N particles at equilibrium [14]

$$Q = \frac{1}{\mathrm{h}^{\mathrm{d}N} N!} \int \mathrm{d}\mathbf{p}^N \mathrm{d}\mathbf{r}^N \exp\left[-\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)/k_B T\right] \tag{1.1}$$

in which h is the Planck's constant, d is the dimensionality of the system, $\mathbf{r}$ and $\mathbf{q}$ are coordinates and the momenta of all N particles in the system, k_B is the Boltzmann constant and T is the temperature in K. $\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)$ is the Hamiltonian of the system which is sum of the kinetic energy $K(\mathbf{p}^N, \mathbf{r}^N)$ and the potential energy $U(\mathbf{r}^N)$. To compute an ensemble average of any observable X in the ensemble of Eq. 1.1, one can use:

$$\langle X \rangle = \frac{\int \mathrm{d}\mathbf{p}^N \mathrm{d}\mathbf{r}^N X\left(\mathbf{p}^N, \mathbf{r}^N\right) \exp\left[-\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)/k_B T\right]}{\int \mathrm{d}\mathbf{p}^N d\mathbf{r}^N \exp\left[-\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)/k_B T\right]} \tag{1.2}$$

where $\langle X \rangle$ denotes an ensemble average. In Eq. 1.2, the momenta $\mathbf{p}^N$ can be integrated out [14]. To compute the ensemble average $\langle X \rangle$ one can write:

$$\langle X \rangle = \frac{\int \mathrm{d}\mathbf{r}^N X\left(\mathbf{r}^N\right) \exp\left[-U(\mathbf{r}^N)/k_B T\right]}{\int \mathrm{d}\mathbf{r}^N \exp\left[-U(\mathbf{r}^N)/k_B T\right]} \tag{1.3}$$

It follows from Eq. 1.3 that the dimension time is not involved in conventional MC simulations which means that transport properties cannot be calculated in conventional MC simulations [14]. Evaluating the integrals of Eq. 1.3 directly using a MC scheme is not possible in general [14, 22], however it is possible to evaluate the ratio of the two integrals in Eq. 1.2 [22]. In principle, it is possible to generate random configurations to sample the ratio in Eq. 1.2, however random configurations will nearly always have unfavourable energies ($U \approx \infty$) due to particle overlaps, leading to zero statistical weight for both the nominator and denominator. To avoid these sampling problems, in the 50s of the previous century, Metropolis, Rosenbluth and co-workers introduced a new sampling scheme [24]. By generating

configurations proportional to their Boltzmann weight $\exp\left[-U(\mathbf{r}^N)/k_\mathrm{B}T\right]$, the average of observable X can be calculated as:

$$\langle X \rangle \approx \frac{1}{L}\sum_{i=1}^{L} X(\mathbf{r}^N) \tag{1.4}$$

where L is the total number of times X is sampled. Since $\langle X \rangle$ is calculated from configurations proportional to the Boltzmann factor, it is calculated as an unweighed average and the $\frac{\text{“0”}}{\text{“0”}}$ sampling problem is avoided. For more details, the reader is referred to [14, 21, 22, 24, 26]. The same principle is used for sampling averages in open ensembles in which the number of molecules fluctuates [14, 17]. Sampling averages in any ensemble should be performed on a series of well-equilibrated system configurations at the desired conditions [14, 22]. Analogous to performing measurements in experiments, sampling ensemble averages in a simulation is affected by statistical uncertainties [14, 26]. Some ensemble averages are straightforward to compute in classical MC simulations *i.g.* density, pressure and energy of a single phase substance. However, describing phase equilibrium of single or multicomponent systems in MC simulations is by no means a trivial task, especially for systems including complex molecules. Phase equilibrium means that the temperature T, pressure P and the chemical potential of each component μ are uniform/equal in every phase at coexistence [2, 37]. While temperature and pressure are easy to compute or impose in a MC simulations, computing chemical potentials and derivatives of the chemical potential requires special simulation techniques [38–41]. In 1987, Panagiotopoulos introduced the Gibbs Ensemble (GE) to directly determine vapour-liquid phase coexistence properties using Monte Carlo (MC) simulations [42–44]. Simulations in the GE rely on sufficient molecular exchanges between the phases to sample phase coexistence correctly. With the increase in computational power, more complex systems are frequently studied which contain large and complex molecules. For complex molecules, or high density systems with strong/directional intermolecular interactions, the probability of molecule exchanges between the phases becomes very small. It is well-known that molecule exchanges in the conventional GE is

very inefficient for complex or large molecules which means that phase equilibrium cannot be reached within a reasonable simulation time, if possible at all [38, 45]. In addition, to sample configurations of large molecules with the correct statistical weight, additional trial moves are needed to change the configuration of the molecules [46]. The standard rotational trial move of a large molecule (around a bond) would most likely result in intermolecular or intramolecular overlaps. To overcome these sampling issues, Configurational Biased Monte Carlo (CBMC) was introduced in the 1990s which is a technique for (re)growing a molecule segment by segment [46–52]. CBMC enabled the computation of VLE of complex molecules, especially alkanes, and adsoprtion of alkanes in porous materials [14, 46, 48, 50–52]. The past decades, several improvements of the CBMC technique have been developed [53–58]. It should be noted that the efficiency of CBMC and related algorithms drops for long molecule chains or high densities [48]. Free energy calculations are also very inefficient in simulations of complex systems in the conventional GE. For low density systems, the chemical potentials of components in each phase can be obtained from GE simulations using a variation of Widom's Test Particle Insertion (WTPI) method [40]. Due to the above-mentioned sampling issues, Widom type methods perform poorly for high density systems [20, 38, 59]. Therefore, it is important to improve and develop simulation techniques to improve molecule exchange efficiency between the phases and to calculate μ. Computation of μ is also an independent and important check on chemical equilibrium [14].

1.3 Applications

In this thesis, we have followed up on recent advances on molecular simulation techniques for open ensembles (in which the number of molecules fluctuates). Using several extensions/improvements, we have applied the state of the art MC simulation techniques to systems of industrial importance, and we have considered the following systems:

(a) Hydrogen as an energy carrier: to date, hydrogen is one of the most popular alternatives for energy storage. A recently developed Elec-

trochemical Hydrogen Compressor by HyET BV [60] can compress hydrogen to pressures up to 1000 bar. Due to the inner workings of the compressor, the compressed hydrogen gas is saturated with water. Due to the limitations issued by ISO standard, the water content in the compressed gas should not exceed 5 μmol. To the best of our knowledge, the only experimental high pressure data ($P > 300$ bar) for $H_2O - H_2$ phase coexistence is from 1927 [61]. In this thesis, we have used advanced molecular simulation techniques to describe the VLE of $H_2O - H_2$ at pressures beyond 300 bar.

(b) Ammonia synthesis: ammonia is a useful chemical commodity and has received lots of attention both in academia and industry [62–68]. It is also a promising alternative medium for energy storage and transportation [69–72]. Industrial ammonia synthesis is carried out using the Haber-Bosch process with heterogeneous iron or ruthenium catalysts at high temperatures (623 K - 873 K) and at a pressure range of 20 MPa to 40 MPa [73–75]. Due to the limitations of cubic equations of state to model this reaction at high pressures [9–11], we have used molecular simulations to model this reaction, compute the chemical potentials, fugacity coefficients and partial molar properties of [NH_3, N_2, H_2] at equilibrium. The reaction enthalpy is calculated using various methods.

(c) Water and aqueous mixtures: Water is one of the most important molecules, central to life, and probably one of the most studied molecule in molecular simulation studies [76–80]. From a chemistry perspective, it is a simple molecule formed by two hydrogen atoms and one oxygen atom. Liquid water has unique thermodynamic properties and complex behavior as a pure substance and in mixtures [77, 81–83]. Methanol is an industrial solvent which plays an important role in different industrial applications [84]. Here, we study the thermodynamic properties of aqueous mixtures of methanol. Aqueous mixtures of methanol are investigated frequently in academia [27, 85–103] and are of practical importance in industrial applications [104–112]. To

compute the mixture properties at different conditions, different force field combinations of water and methanol are considered in molecular simulation studies [27, 79, 85, 86, 88–91, 99, 103]. To the best of our knowledge, chemical potentials and activity coefficients of water and methanol are not reported in molecular simulation studies for different/recent force field combinations of water and methanol. Because of the low vapor pressures of both water and methanol at ambient conditions, direct measurement or using macroscopic engineering models to determine activity coefficients are difficult [113–117].

(d) Formic acid (FA) as a CO or H_2 carrier: FA is the simplest C_1 carboxylic acid that can be made from carbon dioxide [118]. FA is non-toxic and can be stored in liquid phase between 281.55 K and 373.15 K and in aqueous solutions. FA is considered as a candidate for hydrogen storage via its decomposition to hydrogen and carbon dioxide [119–137]. FA is observed as an intermediate in the Water Gas Shift (WGS) reaction [138–140].

$$H_2 + CO_2 \rightleftharpoons HCOOH \rightleftharpoons H_2O + CO \tag{R1}$$

FA can also be considered as a carbon monoxide carrier as well, via its decomposition to water and carbon monoxide [133, 134]. The WGS reaction is an important step down-stream of the steam reforming reaction to adjust the composition of the syngas (the H_2:CO ratio) for different applications. Alternatively, FA can be used in the steam reforming of methane instead of a down-stream WGS reaction, to change the composition of the syngas in the steam reforming reaction. In this thesis, we used computational chemistry principles [4] to calculate the standard state chemical potentials of components in R1 to study combination of FA decompostision and the WGS reaction. Due to the low pressure $P \leq 25$ bar, the PR-EoS was used to model the reaction equilibria. At higher pressures, we may use the advanced simulation techniques described in this thesis to solve the reaction equilibria.

1.4 Outline of this Thesis

The Continuous Fractional Component Monte Carlo (CFCMC) simulation methodology considers an expanded ensemble to solve the problem of low insertion/deletion acceptance probabilities in open ensembles [38, 45, 141]. It allows for a direct calculation of the chemical potential by binning of the coupling parameter λ and using the probabilities $p(\lambda = 0)$ and $p(\lambda = 1)$, [38, 142]. In chapter 2, the combination of the CFCMC method with the GE (CFCGE MC) [38] is used to compute chemical potentials of coexisting gas and liquid phases for water, methanol, hydrogen sulfide and carbon dioxide for the temperature range T=220 K to T=375 K. The chemical potentials are also computed using Widom's test particle insertion (WTPI) method in the conventional GE. It is shown that the WTPI method fails to accurately compute the chemical potentials of water and methanol in the liquid phase at low temperatures. For the CFCGE MC method, the statistical uncertainty for computed chemical potentials of water and methanol in the liquid phase are considerably smaller compared to the WTPI method. For the water models considered in this study, computed excess chemical potentials based on three-site models are in better agreement with the chemical potentials computed from an empirical equation of state from the NIST database, based on experimental data [143–145]. To check whether certain orientation of test particle are energetically unfavourable, orientational biasing is applied for water during test particle insertion. A two-dimensional Overlapping Distribution Method (ODM) in the NVT ensemble is derived for this purpose. It is shown that failure of the WTPI method for systems with a strong hydrogen bonding network does not depend on orientation of the test molecule in that system. It is observed for all systems in this study that the WTPI method breaks down when the void fraction of the system drops below approximately 0.50. In the 2016 version of the CFCMC method [38], one relies on extrapolation to $\lambda \to 0$ and $\lambda \to 1$ to compute the chemical potential. In chapter 3, It is shown that extrapolation to $p(\lambda = 0)$ and $p(\lambda = 1)$ leads to systematic errors when the distribution $p(\lambda)$ is steep. Systematically incorrect values of the computed chemical potential may lead to a false impression of preci-

sion while missing accuracy. An alternative binning scheme is introduced which improves the accuracy of computed chemical potentials without the drawbacks of extrapolation. It is also shown that in simulations of multiple component systems, fractional molecules are very weakly correlated and that calculations of chemical potentials are not affected. It is found that the statistics of Boltzmann averages in systems with multiple fractional molecules is poor. Good agreement is found between CFCMC averages (uncorrected for the bias) and Boltzmann averages when the number of fractional molecules is less than 1% of the total number of all molecules. It is shown that in dense systems, biased averages have a smaller uncertainty compared to Boltzmann averages.

It is very common for molecular simulations to run for days or weeks, especially simulations of systems of complex molecules. One of the bottlenecks encountered in terms of simulation time is computation of intermolecular Coulombic interactions within the system. The Ewald summation [14, 15, 146–148] is the most widely used and accepted method to compute electrostatic interactions in molecular simulation and scales as $\mathcal{O}(N^2)$, or at best $\mathcal{O}(N \log N)$. In chapter 4, the performance of the spherical cutoff methods in MC and MD simulations is compared to the Ewald summation. The radial distribution functions obtained from the Ewald summation and the Damped-Shifted Force (DSF) method [149] are in excellent agreement. Numerical artifacts appear at the cutoff radius when the original Wolf method was used to calculate the electrostatic interactions. In this chapter, different combinations of water and methanol force fields are used to study the thermodynamic properties of aqueous methanol mixtures. The influence of the Wolf and DSF methods on the thermodynamic properties of aqueous methanol mixtures are compared to the Ewald summation. Simulation results show that the numerical artifacts of the original Wolf method have little effect on energy calculations (thermodynamic properties) in aqueous methanol mixtures. Knowledge of partial molar properties (first order derivatives of the chemical potential) plays a central role in studying multicomponent systems. These properties are used to study both the equilibrium and non-equilibrium properties of multicomponent systems, *e.g.* reaction enthalpy, or heat flux in a mixture, etc. [150]. In chapter 5,

we develop an alternative method for calculating partial molar excess enthalpies and partial molar volumes of components in Monte Carlo (MC) simulations. This method combines the original idea of Frenkel, Ciccotti, and co-workers [150] with the CFCMC technique [38]. The method is tested for a system of LJ particles at different densities. As an example of a realistic system, the ammonia synthesis reaction is considered at different pressures ranging from $P = 10$ MPa to $P = 80$ MPa. The compositions at equilibrium are obtained using an efficient implementation of the CFCMC method in the reaction ensemble [142]. Partial molar properties of [NH_3, N_2, H_2] mixtures at chemical equilibrium are computed at different pressures ranging from $P = 10$ MPa to $P = 80$ MPa. Results obtained from MC simulations are compared to those obtained from the PC-SAFT Equation of State (EoS) [18] and the Peng-Robinson EoS [151]. Excellent agreement is found between the results obtained from MC simulations and PC-SAFT EoS, and significant differences were found for PR-EoS modeling. The reaction enthalpy is computed from the partial molar properties of ammonia, nitrogen and hydrogen. We find that the ammonia synthesis reaction is much more exothermic at higher pressures. Based on the recent work of Josphson and co-workers [152], we show that in the reaction ensemble, the reaction enthalpy can be computed directly by simple linear regression of the enthalpy as a function of the number of reactant molecules. The numerical results by Josephson and co-workers strongly suggest that for multicomponent systems, fluctuations and multiple linear regression lead to identical values for thermodynamic derivatives.

In chapter 6, we rigorously prove the equivalence of ensemble fluctuations and multiple linear regression for computing thermodynamic derivatives in open ensembles of an n-component system. In the grand-canonical ensemble, multiple linear regression can be used to obtain the heat of adsorption even around sharp inflection points where the fluctuation approach is known to fail. This provides a conceptually simple and computationally efficient approach to obtain thermodynamic properties from fluctuations in multicomponent systems. It is concluded that multiple linear regression is thermodynamically consistent with fluctuations both in constant-volume and constant-pressure ensembles.

In chapter 7, we develop an alternative method by combining umbrella sampling with CFCMC method to estimate the chemical potential of a component over an appreciable temperature and pressure range from a single simulation, at constant mixture composition. Partial molar enthalpies and partial molar volumes are calculated using the estimated values of the chemical potential in the same simulation. As a proof of principle, our method is tested for a Lennard-Jones (LJ) system at reduced temperature of $T^* = 2$ and reduced pressure of $P^* = 6$. For the temperature range between $T^* = 1.82$ and $T^* = 6.05$, and pressure range between $P^* = 5.95$ and $P^* = 6.05$, excellent agreement is observed between the estimated chemical potentials from umbrella sampling and those obtained from independent simulations. For the LJ system, the partial molar properties obtained from umbrella sampling are identical to the partial molar properties obtained from the CFCMC method in chapter 5 and the original method of Frenkel, Ciccotti, and co-workers [150, 153]. As an example of a complex system, we tested our method for mixtures of water-methanol at different compositions ranging from $x_{\mathrm{MeOH}} = 0.2$ to $x_{\mathrm{MeOH}} = 1$, at $T = 298$ K and $P = 1$ bar. For different mixture compositions, excellent agreement is found between the chemical potentials of water and methanol from umbrella sampling and those obtained from independent simulations, in the temperature range between $T = 288$ K to $T = 308$ K. As an example of a strong hydrogen bond forming system, our method was applied to a mixture of ammonia, nitrogen, and hydrogen at chemical equilibrium at $T = 573$ K and $P = 800$ bar. It was observed that partial molar properties of ammonia, nitrogen and hydrogen obtained from umbrella sampling and the CFCMC method are in excellent agreement. We investigated the limitation of our method for pure methanol ($N = 410$ molecules) in the liquid phase. It is observed that the estimated chemical potentials from umbrella sampling are in excellent agreement with the reference values obtained from independent simulations, for $\Delta T = \pm 15$ K and $\Delta P = 100$ bar, with respect to the system which is simulated. For larger systems this range becomes smaller since the relative fluctuations of energy and volume become smaller. The advantages of using our method are explained in detail in this chapter.

In chapter 8, the molecular simulation techniques described in previous chapters are used to study solubility of water in hydrogen at high pressure. Hydrogen is one of the most popular alternatives for energy storage. Due to its low volumetric energy density, hydrogen should be compressed for practical storage and transportation purposes. Recently, Electrochemical Hydrogen Compressors (EHC) have been developed that are capable of compressing hydrogen up to $P = 1000$ bar, and have the potential of reducing compression costs to from 6 kWh/kg to 3 kWh/kg [154]. As EHC compressed hydrogen is saturated with water, the maximum water content in gaseous hydrogen should meet the fuel requirements issued by the International Organization for Standardization (ISO) when refuelling Fuel Cell Electric Vehicles (FCEV). The ISO 14687-2:2012 standard [155] has limited the water concentration in hydrogen gas to 5 μmol water per mol hydrogen fuel mixture. Knowledge on the vapor liquid equilibrium of $H_2O - H_2$ mixtures is crucial for designing a method to remove H_2O from compressed H_2. To the best of our knowledge, the only experimental high pressure data ($P > 300$ bar) for $H_2O - H_2$ phase coexistence is from 1927 [61]. In this chapter, we have used molecular simulation and thermodynamic modelling to study the phase coexistence of the $H_2O - H_2$ system for temperatures between $T = 283$ K to $T = 423$ K and pressures between $P = 10$ bar and $P = 1000$ bar. It is shown that the PR-EoS and SRK-EoS with van der Waals mixing rules fail to accurately predict the equilibrium coexistence compositions of the liquid and gas phase, with or without fitted binary interaction parameters. We have shown that the solubility of water in compressed hydrogen is adequately predicted using force field based molecular simulations. In appendix A.11, a detailed overview of available experimental VLE and solubility data for the $H_2O - H_2$ system at high pressures are presented.

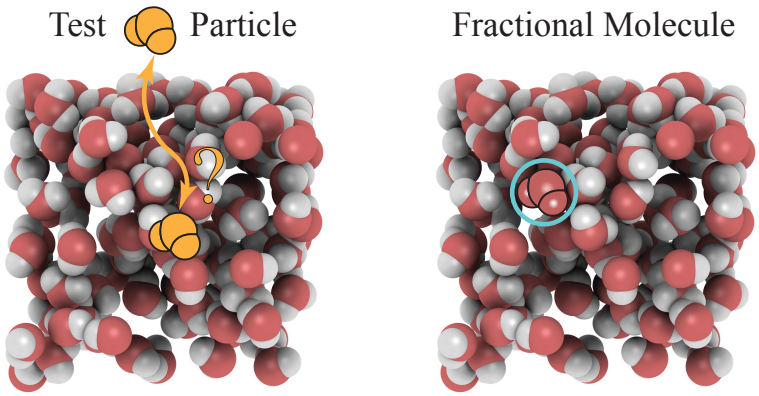
Syngas is a crucial building block in the chemical process industry to produce fuels and other important chemicals. It is used for the production of hydrocarbons, acetic acid, oxo-alcohols and other chemicals. Depending on the target product and stoichiometry of the reaction, an optimum (molar) ratio between hydrogen and carbon monoxide (H_2:CO) in the syngas is required. Different technologies are available to control the H_2:CO molar ratio

in the syngas. The combination of Steam Reforming of Methane (SRM) and the WGS reaction is the most established approach for syngas production. In chapter 9, to adjust the H_2:CO ratio, formic acid is considered as a source for both hydrogen and carbon monoxide. Using thermochemical equilibrium calculations, we show that the syngas composition can be controlled by co-feeding formic acid (FA) into the SRM process. The H_2:CO molar ratio can be adjusted to a value between one and three by adjusting the concentration of FA in the reaction feed. At steam reforming conditions, typically above 900 K, FA can decompose to water and carbon monoxide, and/or to hydrogen and carbon dioxide. Our results show that co-feeding FA into the SRM process can adjust the H_2:CO molar ratio in a single step. This can potentially be an alternative to the WGS process.

Chapter 2

CFCMC Simulations in the Gibbs Ensemble

This chapter is based on the following paper: Rahbari, A.; Poursaeidesfahani, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H.; Chemical Potentials of Water, Methanol, Carbon Dioxide, and Hydrogen Sulfide at Low Temperatures using Continuous Fractional Component Gibbs Ensemble Monte Carlo, Molecular Simulation, 2018, 44, 405-414, Ref. [59]



2.1 Introduction

Knowledge on Vapor-Liquid Equilibrium (VLE)/reaction equilibria and chemical potentials is important for process design and modelling [14, 39, 59]. The past decades, force field based molecular simulation has been developed as an attractive alternative for experiments, to accurately describe the behavior of matter, and to obtain reliable thermodynamic and transport properties [156–162]. Molecular simulations are used extensively for studying phase equilibria of pure and multicomponent systems [163–167], describing the behavior of guest molecules inside porous media [168–171], and reaction equilibria [171–177] etc. Many thermodynamic properties such as the density and pressure of a system are straightforward to compute in molecular simulations [42, 142, 167, 178].

In his pioneering work in 1987, Panagiotopoulos introduced the Gibbs Ensemble (GE) to directly determine the phase coexistence properties using Monte Carlo (MC) simulations [42–44]. The GE is reliable, and fast method to obtain relatively accurate critical properties for most systems with relatively small system sizes [167, 179, 180]. The finite size effects in the GE are small unless conditions close to the critical point are considered [179, 180]. In the GE, sufficient molecular exchanges between the phases leads to equal chemical potentials (which are directly related to activity/fugacity coefficients). Although chemical potentials of different component types are not strictly needed for calculating the coexistence densities, the equality of chemical potentials is an independent and important condition for phase equilibrium [14]. It can also be used to detect programming errors and errors in the implementation of the simulation technique. At high densities, successful molecule insertions depend strongly on occurrence of spontaneous cavities large enough to accommodate the inserted molecule. As a result of rare occurrence of such cavities at high densities, the acceptance probabilities of molecule insertions/deletions are very low in dense systems or systems with strong/directional intermolecular interactions, *e.g.* for water at ambient conditions [59, 181]. This is a major drawback of simulations of dense systems in the GE. Another drawback is that computing the excess

chemical potential in the GE using insertion/deletion methods [40, 182–185] suffers severely from molecule overlaps or random cavity formation.

The Widom's Test Particle Insertion (WTPI) method [40] is the most commonly used method for determining chemical potentials of components by sampling the interaction energy of a test molecule inserted at a randomly selected position in the system. The formulation of the WTPI method in the GE was first developed by Frenkel and smit [39] which takes into account the density fluctuations of each phase. It is well-known that methods based on the WTPI method perform poorly for high density systems, even when combined with CBMC or related methods [14, 181, 186, 187]. Coskuner and Deiters have used the WTPI method to compute, among other properties, the chemical potential of TIP5P water in the temperature range of T=300 K to T=320 K in the NPT ensemble [188]. However, the chemical potentials are not in agreement with the experimental Equation of State (EoS) [143, 144]. This deviation may stem from either inefficiency of the WTPI method at low temperatures [189–191] or the hydrogen bonding network of water. Other reasons may include inefficiency of the simulations or a force field limitation. Limitations of different simulation techniques involving (test) particle insertions/removals have led to development of more advanced MC techniques to improve molecule exchange efficiency and computation of chemical potential by combining particle insertions and removals [190, 192], or by gradual insertions/deletions in multiple MC steps such that the surrounding molecules can adjust to the molecule that is inserted or deleted [49, 187, 193]. In the past decades, the idea of gradual insertion/deletion was used for different systems, see the works of Mon et al. [193], Squire et al. [194], Mruzik et al. [195] and de Pablo from the 90s [196].

A few years ago, the Continuous Fractional Component Monte Carlo (CFCMC) technique was developed by Shi and Maginn [45, 141], leading to efficient molecule exchanges in open ensembles. The main new element developed by Shi and Maginn is that the gradual insertion of molecules is continuous, rather than in discrete stages [197, 198]. In the CFCMC method, a fractional molecule with scaled interactions with the surroundings is added to the ensemble. A coupling parameter λ is introduced as an extended variable in an expanded ensemble, and trial moves are carried out to change

the value of λ. The fractional molecule is distinguishable from the other "whole", or normal molecules. The value $\lambda = 0$ means that the fractional molecule does not interact with other molecules in the simulation box and acts as an "ideal gas" molecule. The value $\lambda = 1$ means that the fractional molecule is fully interacting with other molecules in the system, and thus acts as a "whole" molecule. To further increase the efficiency of molecule exchanges, an additional biasing potential $W(\lambda)$ can be used to ensure that the sampled probability distribution of λ is flat [45, 141, 199, 200]. Although the original method of Shi and Maginn [45, 141] significantly improves the efficiency of molecule exchanges, additional post-processing is required to compute the chemical potential or derivatives of the chemical potential.

Based on the earlier work of Maginn and co-workers, Vlugt and co-workers have developed a more efficient alternative of the CFCMC technique [38] in open ensembles (GC, reaction ensemble, Gibbs ensemble) with free energy calculations, in which molecule transfers are facilitated by CFCMC [38, 142]. The crucial difference between this method and the original implementation of Maginn and co-workers is that a single fractional molecule per component is used, and the fractional molecule of each component type can be in either one of the phases [38]. In addition to the conventional thermalization trial moves (translation, rotation and volume changes), three additional trial moves are associated with the fractional molecule of each component. Details on the trial moves associated with the fractional molecule in the GE are explained in detail in this chapter. During the simulation involving a fractional molecule, the probability distribution of $\lambda \in (0, 1)$ is sampled. It is shown in [38], that the ratio between $p(\lambda = 0)$ and $p(\lambda = 1)$ is directly related to the free energy difference of inserting a full additional molecule [38, 142]. The CFCMC method has been applied to the NPT/NVT ensemble [20], grand-canonical (GC) [201], Gibbs Ensemble (GE) [20, 38, 45, 141], and the reaction ensemble (RxMC) [142, 177]. In this chapter, the underlying reason and a criterion for the limitations of the WTPI method are investigated. We have chosen to simulate the VLE of four different components in the GE with similar coexistence liquid densities, namely water and methanol (with a hydrogen bonding network), and hydrogen sulfide and carbon dioxide (without a hydrogen bonding network).

To study the role of different models, we have used several force fields for water and methanol: SPC [202], TIP3P/EW [203], TIP4P/EW [204] and TIP5P/EW [205]) water, OPLS and TraPPE methanol [206, 207], hydrogen sulfide [208] and TraPPE carbon dioxide [209]. Number densities and void fractions of all systems in this work are compared as it may provide a criterion for limitation of the WTPI method. As an independent check, the Overlapping Distribution Method (ODM) is used in the NVT ensemble [14, 210] to check the reliability of the WTPI results at different system temperatures. As water has a strong hydrogen bonding network, a two-dimensional ODM in the NVT ensemble is used to test if certain orientations of the test molecule inside the hydrogen bonding network of water lead to inefficiency/failure of the WTPI method. In this method, different number of hydrogen bonds between the test molecule and its surrounding molecules corresponds to different orientations of the test molecule.

This chapter is organized as follows. In Section 2.2, the relevant equations used for sampling the chemical potentials using the WTPI method in the GE and th CFCGE MC method are provided, and the differences between the two methods are explained. Additionally, equations for the conventional ODM using orientational biasing and the two-dimensional ODM are presented and explained. Void fractions of every component are also defined and tabulated at different temperatures. Simulation details are provided in Section 2.3. Our findings are presented in section Section 2.4. Our results show that computing chemical potential using the WTPI method becomes quite inefficient when for systems with a void fraction smaller than 0.5, while the CFCGE MC method does not have this limitation and is more efficient at high densities. Our conclusions are summarized in Section 2.5.

2.2 Methodology

The computation of chemical potential for a pure component in the GE is based on the original WTPI method [211], taking into account the fluctuations in the volume and the number of molecules in box i. Frenkel and Smit showed that the chemical potential in the GE equals [14, 39] :

$$\mu_{i,\,\mathrm{GE}} = -k_{\mathrm{B}}T \ln \left\langle \frac{V_i/\Lambda^3}{N_i+1} \exp\left[-\beta \Delta U_i^{+}\right] \right\rangle \tag{2.1}$$

$\beta = 1/(k_{\mathrm{B}}T)$ in which k_{B} is the Boltzmann constant and T is the temperature of the system. ΔU_i^{+} is the interaction energy of the test molecule with the rest of the molecules in box i. Λ is the thermal wavelength, V_i and N_i are the volume and number of molecules of box i respectively. To obtain the chemical potential in simulations in the NPT ensemble, Eq. 2.1 is used. One can split the chemical potential in Eq. 2.1 into an ideal gas part and excess part. The finite-size effect of the ideal gas part of Eq. 2.1 can be corrected using the approximation outlines in [164]. Details about obtaining chemical potential and derivatives of the chemical potential in the NPT ensemble are explained in detail in chapter 5. Due to overlaps between the test molecule and the existing molecules in a system, the potential energy change (ΔU_i^{+}) of the trial insertion move can become infinitely large ($\Delta U_i^{+} \to +\infty$), and the corresponding Boltzmann factor becomes almost equal to zero. Since the majority of trial insertion moves in a dense liquid phase contribute with almost zero statistical weight, the chemical potential computed using Eq. 2.1 may be questionable and has typically large standard deviations in high-density phases.

To circumvent the potential sampling problems of the WTPI method, Shing and Gubbins [190, 192] proposed an alternative way of obtaining chemical potential in the 80s of by combining particle insertions and removals. Similarly, Bennett [210] has introduced the ODM which is used in this work as an independent check to verify the validity of the WTPI method at different system densities/void fractions [14]. In the ODM, two separate simulations in the NVT ensemble based on the coexistence densi-

ties from GE simulations are performed for two separate systems 0 and 1 with N and $N+1$ particles, respectively. Here, the volumes of systems 0 and 1 are assumed to be the same, but this is not essential [14]. It is shown in Refs. [14, 212] that the excess chemical potential of system 0 equals

$$\mu^{\mathrm{ex}} = f_1\left(\Delta U\right) - f_0\left(\Delta U\right) \tag{2.2}$$

The functions $f_0(\Delta U)$ and $f_1(\Delta U)$ are defined as

$$\begin{array}{l} f_0\left(\Delta U\right) = RT\left[\ln p_0\left(\Delta U\right)\right] - \frac{\Delta U}{2} \\ f_1\left(\Delta U\right) = RT\left[\ln p_1\left(\Delta U\right)\right] + \frac{\Delta U}{2} \end{array} \tag{2.3}$$

Here, $k_{\mathrm{B}}T$ is written as RT since all chemical potentials and energies are reported in [kJ.mol^{-1}]. In Eq. 2.3, ΔU denotes the potential energy difference between systems 0 and 1. $p_0(\Delta U)$ is the probability distribution for the potential energy difference ΔU between system 0 and 1 while sampling configurations in system 0. The same definition applies for $p_1(\Delta U)$ in system 1. Since the ODM combines insertion and removal trial moves, it offers a better estimate of the chemical potential compared to the WTPI method [14, 212], and can be used as an independent check for the WTPI method. For systems with a strong hydrogen bonding network like water, the interaction energies of the test molecules may not only depend on the position, but also on the orientation with respect to the hydrogen bonding network. The number of hydrogen bonds that the test molecule forms with its adjacent molecules is related to the orientation of the test molecule. A geometrical definition of hydrogen bonds [213–218] was used in this chapter to count the number of hydrogen bonds between water molecules. The geometric criterion is shown in Fig. 2.1. We implemented this geometric definition in the RASPA software package [219, 220] and computed the average number of hydrogen bonds for the TIP4P/EW water model in the NVT ensemble. Excellent agreement with literature [218] was observed for the average number of hydrogen bonds for temperatures between T=300 K and T=600 K. The results are shown in Fig. 2.1.

To study the hydrogen bonding network of water and its effect on the performance of the WTPI method, a two-dimensional ODM is derived by

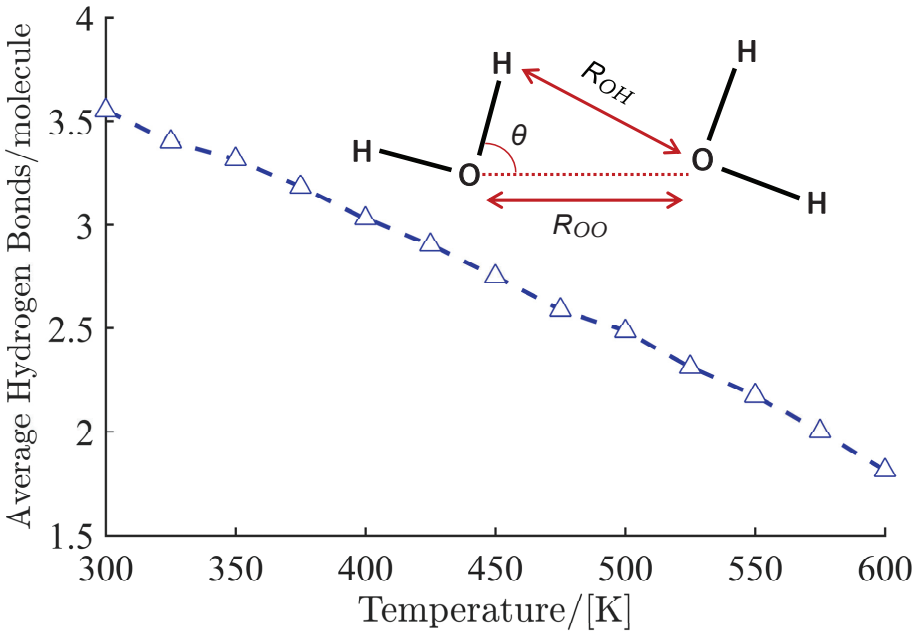


Figure 2.1: Triangles: Average number of hydrogen bonds for the TIP4P/EW water model computed in the *NVT* ensemble based on the coexistence densities of saturated liquid water from NIST database [144]. Error bars are smaller than symbols. The inset shows the hydrogen bond geometric criterion between two water molecules: $R_{\mathrm{OO}} < 3.5$ Å, $R_{\mathrm{OH}} < 2.5$ Å and $\theta < 30$ [213–218].

introducing the number of hydrogen bonds (H) as a second integer variable. The two-dimensional ODM can check the validity of the WTPI method for different numbers of formed hydrogen bonds between test molecule and its surrounding water molecules. If orientational biasing is included as part of the test molecule insertion and removal, the energy difference $-\beta\Delta U$ in Eqs. 2.2 and 2.3 should be replaced by the logarithm of the Rosenbluth weight of the test molecule [25, 179, 187, 212, 221]. The functions f_1 and

f_0 in the two-dimensional ODM including orientational biasing become

$$\begin{aligned} f_0(W,H) &= RT\left[\ln p_0(\ln W) + \tfrac{\ln W}{2}\right] \\ f_1(W,H) &= RT\left[\ln p_1(\ln W) - \tfrac{\ln W}{2}\right] \end{aligned} \tag{2.4}$$

H is the number of hydrogen bonds between the test molecule (in the trial insertions or trial removals) and its surrounding water molecules. Derivation of Eq. 2.4 is provided in the Supporting Information of Ref. [59]. The excess chemical potential for a given number of hydrogen bonds H equals

$$\mu^{\mathrm{ex}}(H) = f_1(W,H) - f_0(W,H) \tag{2.5}$$

In an attempt to quantify the limit at which the WTPI method breaks down, the number density and void fraction of all systems in this work are compared and shown in Table 2.1. The number density is defined as

$$\rho_N = \frac{N}{V} \tag{2.6}$$

The void fraction ϕ is defined as

$$\phi = 1 - \rho_N \cdot V_{\mathrm{mol}} \tag{2.7}$$

V_{mol} is the volume of the molecular model. To compute the volume, each interaction site in the molecule is considered to be a sphere with diameter σ. Therefore, the volume of a molecule with k interaction sites equals the sum of the volumes of all spheres minus the intersection volume between the spheres

$$V_{\mathrm{mol}} = \sum_{i=1}^{k} \frac{4}{3}\pi(\sigma/2)^3 - V_{\mathrm{intersection}} \tag{2.8}$$

$V_{\mathrm{intersection}}$ is the total intersection volume between the spheres. The value of σ of each molecule type can be found in the force field data in the Supporting Information of Ref. [59].

To ensure that phase equilibrium is reached in simulations in the GE, independent simulations are also performed using our formulation of the CFCGE MC method which has a more efficient molecule exchange [38] and allows for direct computation of the chemical potential. The molecule swap trial move in the GE is replaced by three types of trial moves to facilitate molecule exchange between the simulation boxes in the CFCGE MC method. *(1)*: swapping the fractional molecule to a randomly selected position in the other simulation box, while keeping the positions and orientations of other molecules constant; *(2)*: changing the value of scaling parameter λ while keeping the positions and orientations of all molecules including the fractional molecule constant; *(3)*: changing the identity of the fractional molecule with a whole molecule in the other box while keeping the value of λ, positions and orientations of other molecules constant. Fig. 2.2 shows a schematic representation of these types of trial moves. It is efficient to combine trial moves *(2)* and *(3)* into a single hybrid trial move, as trial move *(2)* has a high acceptance probability only at low values of λ, and trial move *(3)* has a high acceptance probability only at high values of λ. In this hybrid trial move, trial move *(2)* is only selected at low values of λ, and trial move *(3)* is only selected at high values of λ. This avoids the situations in which trial moves with a very low acceptance probability are selected. Since the value of λ does not change during this hybrid trial move, the probabilities of selecting this trial move and the reverse trial move are identical, and therefore the condition of detailed balance is not violated [14, 142]. For more details, the reader is referred to chapters 3 and 5. The acceptance rules of these moves are derived in Ref. [38]. It is also shown in Ref. [38] that by sampling the probability of λ approaching zero and one in each simulation box, one can compute the chemical potential as an ensemble average without any additional post processing. The chemical potential for a single component in box i in the CFCGE MC equals [38]

$$\mu_{i,\,\mathrm{CFCGE\,MC}} = \underbrace{-k_{\mathrm{B}}T \ln \left\langle \frac{V_i/\Lambda^3}{N_i+1} \right\rangle}_{\text{Ideal gas part}} \underbrace{-k_{\mathrm{B}}T \ln \left\langle \frac{p_i\,(\lambda \uparrow 1)}{p_i\,(\lambda \downarrow 0)} \right\rangle}_{\text{Excess part}} \tag{2.9}$$

$p_i(\lambda \uparrow 1)$ and $p_i(\lambda \downarrow 0)$ are the probability of λ approaching one and zero in box i. The chemical potential in Eq. 2.9 is split into an ideal gas part and excess part. It is shown in appendix A.1 that the computed chemical potentials obtained using Eqs. 2.1 and 2.9 are the same by definition. Eq. 2.9 can be extended to other ensembles. Applications of Eq. 2.9 to other ensembles are provided in chapters 3 to 5 and 8. One should be aware of potential pitfalls of the CFCMC method when performing extrapolation to compute the chemical potential in Eq. 2.9. Using chapter 3, we show that performing extrapolation on $p(\lambda)$ may lead to systematic error when the distribution is steep. To circumvent this issue, an alternative binning scheme is developed in chapter 3.

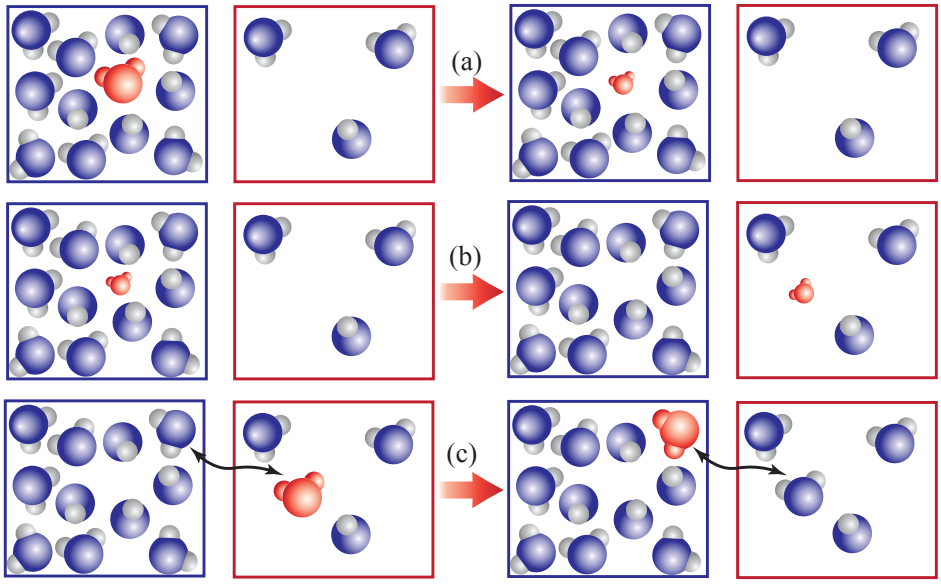


Figure 2.2: MC trial moves facilitating molecule exchanges in the CFCGE MC simulations [38]. The coupling parameter λ scales the interactions of the fractional molecule with its surroundings. In this figure, the fractional molecule is marked red. These types of trial moves are: (Swap Move): the fractional molecule is moved to a randomly selected position in the other box, while keeping the value of λ constant. (Change λ): random change of the coupling parameter λ while keeping the position of the fractional molecule constant. Trial moves that result in $\lambda < 0$ or $\lambda > 1$ are automatically rejected. (Change Identity): trial move to change the fractional molecule into a whole molecule and changing a randomly selected molecule in the other box into a fractional molecule, while keeping all positions fixed. For more details about the CFCGE MC method and the acceptance rules the reader is referred to the Ref. [38].

2.3 Simulation Details

The Lennard-Jones (LJ) interactions of the fractional molecule with the rest of the molecules can be scaled as follows [20, 38, 142, 222]:

$$u_{\mathrm{LJ}}\left(r,\lambda_{\mathrm{LJ}}\right)=\lambda_{\mathrm{LJ}}4\epsilon\left(\frac{1}{\left[\frac{1}{2}(1-\lambda_{\mathrm{LJ}})^{2}+\left(\frac{r}{\sigma}\right)^{6}\right]^{2}}-\frac{1}{\left[\frac{1}{2}(1-\lambda_{\mathrm{LJ}})^{2}+\left(\frac{r}{\sigma}\right)^{6}\right]}\right) \quad (2.10)$$

in which σ and ϵ are the LJ parameters and r is the intermolecular distance between two interaction sites. λ_{LJ} scales the LJ interactions of the fractional molecule. In principle, other thermodynamic pathways are possible to scale the LJ interactions of the molecule [223–225]. To protect the charges from overlapping, the scaling of the LJ and Coulombic interactions of the fractional molecule are decoupled. It is specifically important to switch on the repulsive LJ interactions before the Coulombic interactions, to protect the charges from overlapping [223–228]. Charge-overlaps can potentially lead to huge electrostatic potentials, inaccuracies and numerical instabilities [223, 224, 226–228]. First, the LJ interactions of the fractional molecule are switched on at $\lambda = 0$ and are fully interacting with the surrounding molecules when λ reaches a certain predefined threshold value of λ^*. Second, the Coulombic interactions of the fractional molecule are switched on at $\lambda = \lambda^*$ and are fully interacting with the surrounding molecules when $\lambda = 1$. In Fig. 2.3, it is shown how the LJ and Coulombic interactions are decoupled and scaled with separate coupling parameters $\lambda_{\mathrm{LJ}} \in [0, 1]$ and $\lambda_{\mathrm{Coul}} \in [0, 1]$ respectively. Different choices are possible for λ^* depending on the system. In this thesis, the selected values of λ^* are between $\lambda^* = 0.6$ and $\lambda^* = 0.8$. No attempts were made to choose the value of λ^* which leads to the most efficient simulation. Note that only the efficiency of the simulation depends on λ^* and not the result. Details on scaling electrostatic interactions of the fractional molecule are provided in chapter 4.

To compute the chemical potential of liquid and gas phases of water at equilibrium, MC simulations in the temperature range between T=210 K and T=375 K are performed in the CFCGE MC and GE. All simula-

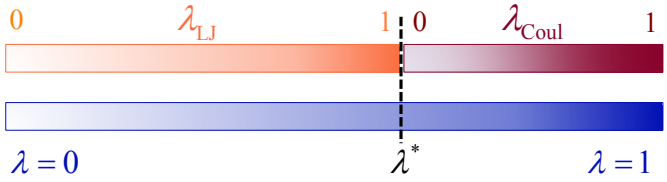


Figure 2.3: Coupling parameter λ to scale the interactions of fractional molecules. $\lambda_{LJ} \in [0, 1]$ is the coupling parameter used to scale the LJ interactions of the fractional molecule (Eq. 2.10). At $\lambda = \lambda^*$, the Coulombic interactions are switched on. $\lambda_{Coul} \in [0, 1]$ is the coupling parameter used to scale the Coulombic interactions of the fractional molecule (see chapter 4).

tions are performed using the RASPA software package [219, 220]. The chemical potentials are computed directly in the CFCGE MC simulations (Eq. 2.9). In the conventional GE, the WTPI method is used to compute the chemical potential of both phases (Eq. 2.1). Using the coexistence density of the liquid phase from GE simulations, the ODM in the NVT ensemble is performed to check independently at which densities the WTPI method breaks down. Four different rigid water models (SPC [202, 229], TIP3P/EW [203], TIP4P/EW [204] and TIP5P/EW [205]) were used for this study. Rigid methanol OPLS-UA [206] and flexible TraPPE force fields were selected for methanol. [207]. For carbon dioxide and hydrogen sulfide, TraPPE force field parameters were used [208]. Details about the force field parameters, truncation of intermolecular potentials, and tail corrections are provided in the Supporting Information of Ref. [59]. For all water models 10^5 equilibration cycles were carried out followed by $2.3 \cdot 10^6$ MC cycles. For methanol, carbon dioxide models $2 \cdot 10^5$ equilibration cycles were carried out followed by $1.2 \cdot 10^6$ MC cycles. Each MC cycle in RASPA consists of N Monte Carlo trial moves, where N equals the number of molecules. The Wang-Landau algorithm [199, 230] was used to compute the weight function in the CFCGE MC simulations.

For CFCGE MC, trial moves were selected with the following probabilities: probability of volume exchange between the boxes: 1.0%. The rest of the trial moves were selected with an equal probability of 19.8% including: translation, rotation, swap molecules, change value of lambda,

and change identity move. No biasing was used for molecule insertions or deletions using CFCGE MC. However, it is possible to combine orientational biasing with CFCGE MC for a higher computational efficiency [186]. For conventional GE simulations, trial moves were selected with the following probabilities: probability of volume exchange between the boxes: 1.2%. The rest of the trial moves were selected with an equal probability of 24.7% including: translation, rotation, swap. For insertion of molecules in the conventional GE, orientational biasing was used with 10 trial directions. For flexible methanol in CFCGE MC simulations, the probability of volume exchange was 0.8%. The probabilities of selecting other moves were evenly distributed between the aforementioned trial moves above and an additional molecule reinsertion trial move. For flexible methanol in GE simulations, probabilities of selecting volume exchange was 1.0%. The probabilities of selecting other moves were evenly distributed between the aforementioned moves above and an additional molecule reinsertion move.

2.4 Results

Simulation results show that computed chemical potentials of hydrogen sulfide and carbon dioxide using the WTPI method are in excellent agreement in the gas and liquid phase at coexistence. However, this is not the case for water and methanol. Fig. 2.4 compares the computed chemical potentials of water and methanol at coexistence using both methods. It is clearly shown that computed chemical potentials of water and methanol in the gas phase and the liquid phase at coexistence are only in excellent agreement when using the CFCGE MC method.

The inability to accurately compute the chemical potentials of the liquid phase at coexistence can be either due to inefficiency of the GE simulations to reach equilibrium, or the inefficiency of the WTPI for the chemical potential in the dense liquid phase. The molecule exchange efficiency in the conventional GE simulations is up to orders of magnitude lower than for the CFCGE MC simulations as shown in Fig. 2.5 (efficiencies defined in the caption). Despite the lower efficiency of molecule exchanges in the GE, the

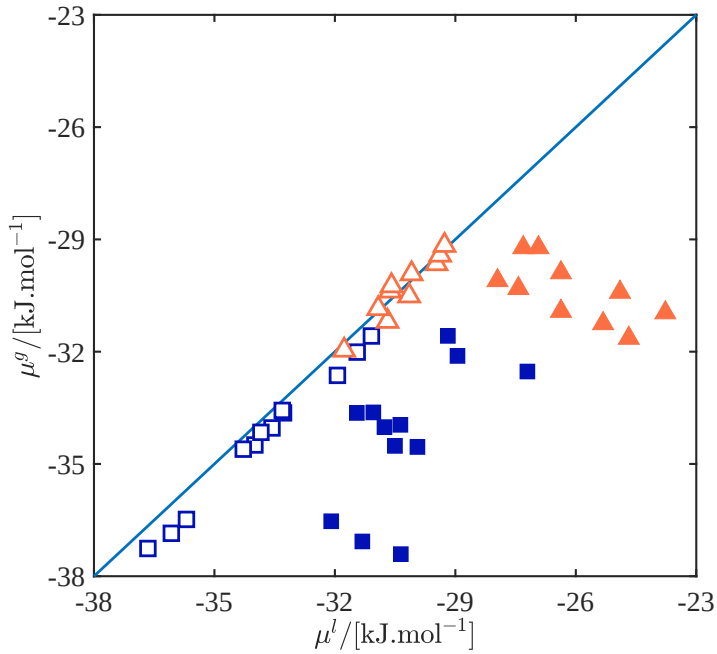


Figure 2.4: Comparison of chemical potentials of gas and liquid phases using the CFCGE MC method and the WTPI method in the GE. (open squares): computed chemical potentials of the water SPC, TIP3P/EW and TIP4P/EW models using the CFCMC GE method, (open triangles): computed chemicals potential of the methanol models OPLS-UA and TraPPE using CFCGE MC method, (closed squares): computed chemical potentials of water SPC, TIP3P/EW and TIP4P/EW models using the WTPI method, and (closed triangles): computed chemical potentials of the methanol models OPLS-UA and TraPPE using the WTPI method. The line is a guide to the eye to indicate the equal chemical chemical potentials between the gas and liquid phase at coexistence. Error bars are smaller than symbol sizes. The raw data are listed in the Supporting Information of Ref. [59].

coexistence densities obtained from CFCGE MC and GE simulations are in excellent agreement for all components (coexistence liquid densities are listed in the Supporting Information of Ref. [59].). Hence, both methods have converged to the same equilibrium densities. Since identical equilibrium densities are obtained using both methods, any difference between the

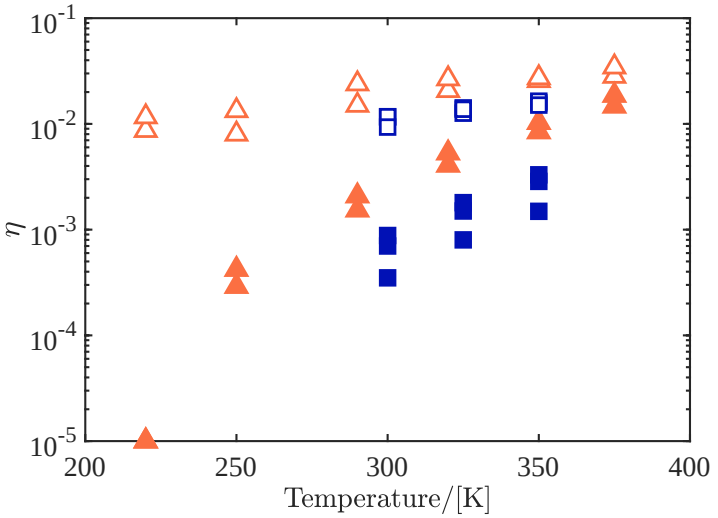


Figure 2.5: Molecule exchange efficiencies in the GE and CFCGE MC. Efficiency of the swap trial moves in the GE is defined as the ratio between accepted molecule insertions and the total attempted swap trial moves (closed symbols). The efficiency of the identity change move in the CFCGE MC is defined as ratio between accepted change identity trial moves and total number of trial moves to facilitate molecule transfers [38]. (squares): SPC, TIP3P/EW and TIP4P/EW water (CFCGE MC) (triangles): OPLS-UA and TraPPE methanol (CFCGE MC).

computed chemical potentials can only be attributed to inefficient computation of excess part of the chemical potential. The contribution of excess chemical potential in the dense liquid phase is significant due to strong interactions of the test molecule with its surroundings, while in the gas phase, the inserted test molecule has limited interactions with its surroundings which means that the chemical potential of the gas phase is mainly determined by the ideal part. Since chemical potentials in the liquid and gas phase are equal in CFCGE MC simulations, it can be concluded that the CFCGE MC method computes the excess part of the chemical potential correctly for water and methanol in the liquid phase, and the WTPI method does not. Computed chemical potentials of coexisting phases of

all systems using both methods are shown in Figs. 2.6 and 2.7 as a function of temperature. The inability of the WTPI method to compute the excess chemical potentials of water and methanol accurately may be due to the existence of a strong hydrogen bonding network. A two-dimensional overlapping distribution method in the NVT ensemble at the coexisting liquid densities was used to verify whether certain orientations of the test molecule can be energetically unfavorable such that the performance of the WTPI method is affected. As shown in Fig. 2.1, the number of hydrogen bonds that test molecule forms with its adjacent molecules follows from the orientation of the test molecule in the hydrogen bonding network. Fig. 2.8 shows results of implementing the two-dimensional ODM for TIP4P/EW water model at 300 K as a function of number of formed hydrogen bonds between the test molecule and its surrounding molecules. For $H \geqslant 4$, the overlap between functions $f_0(W, H)$ and $f_1(W, H)$ becomes smaller which results in poor statistics and therefore these are not shown in Fig. 2.8. It is shown in Fig. 2.8 that excess chemical potentials computed using the two-dimensional ODM and the WTPI method are not equal for any value of H. This observation does not depend on the of number of hydrogen bonds the test molecule forms with its surrounding molecules in the present system. The two-dimensional ODM method was also used to show that the chemical potential of water at T=500 K was computed accurately using the WTPI method, for all values of H. Simulation results of the two-dimensional ODM for TIP4P/EW water at T=500 K are provided in the Supporting Information of Ref. [59].

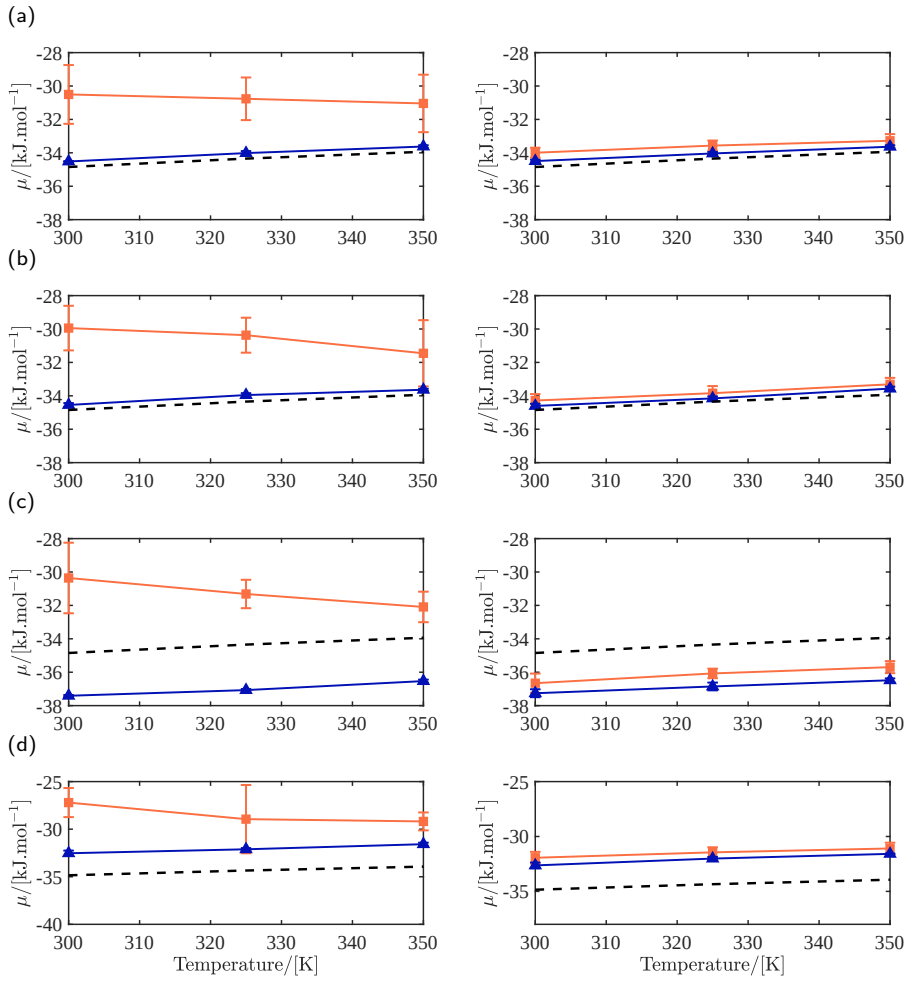


Figure 2.6: Chemical potentials of coexisting phases of water models for temperature range T=300 K to T=350 K: (a) SPC; (b) TIP3P/EW; (c) TIP4P/EW and; (d) TIP5P/EW (left: WTPI method in the conventional GE, right: CFCGE MC method). In all subfigures: (triangles): computed chemical potentials in the gas phase, (squares): computed chemical potentials in the liquid phase, (dashed lines): computed chemical potentials from the Helmholtz EoS based on empirical data [143]. The raw data are listed in the Supporting Information of Ref. [59].

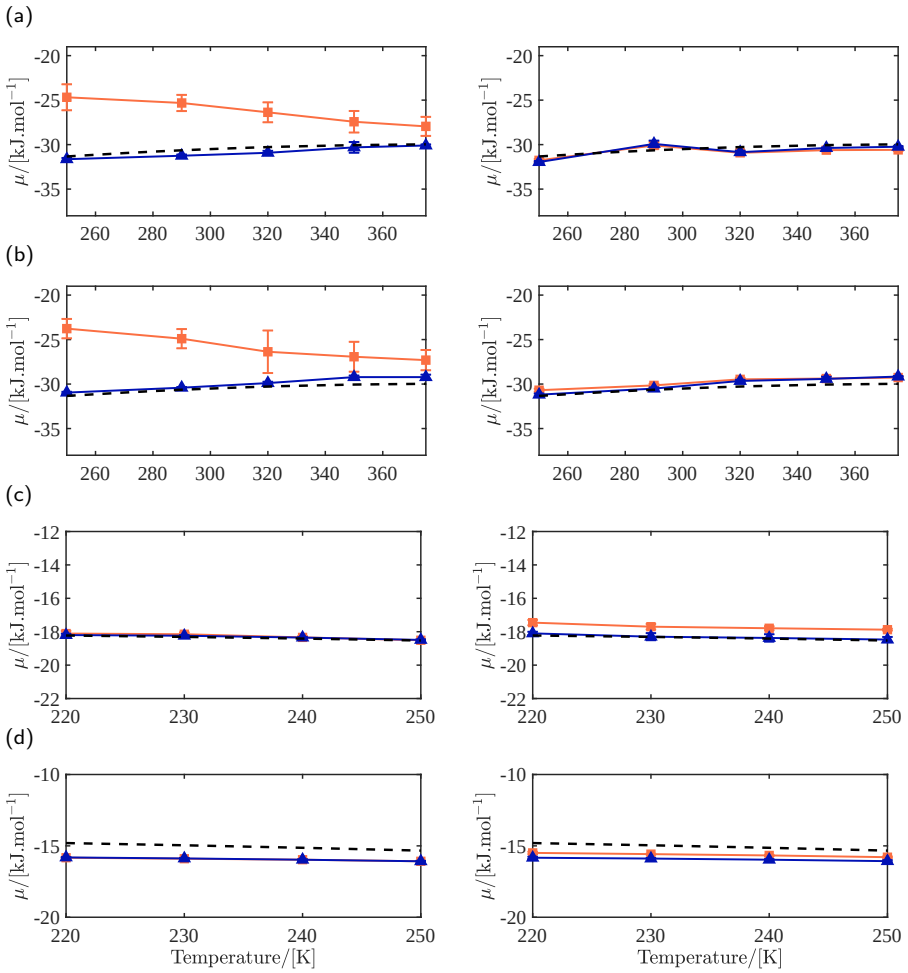


Figure 2.7: Chemical potentials of coexisting phases of methanol, hydrogen sulfide and carbon dioxide (liquid and gas) in the temperature range T=220 K to T=375 K: (a) methanol TraPPE; (b) methanol OPLS-UA; (c) H_2S TraPPE; (d) CO_2 TraPPE force field (left: WTPI method in the conventional GE, right: CFCGE MC method). In all subfigures: (triangles): computed chemical potential in the gas phase, (squares): computed chemical potentials in the liquid phase, (dashed line): computed chemical potentials from the empirical Helmholtz EoS based on experimental data [145, 231, 232]. The raw data are provided in the Supporting Information of Ref. [59].

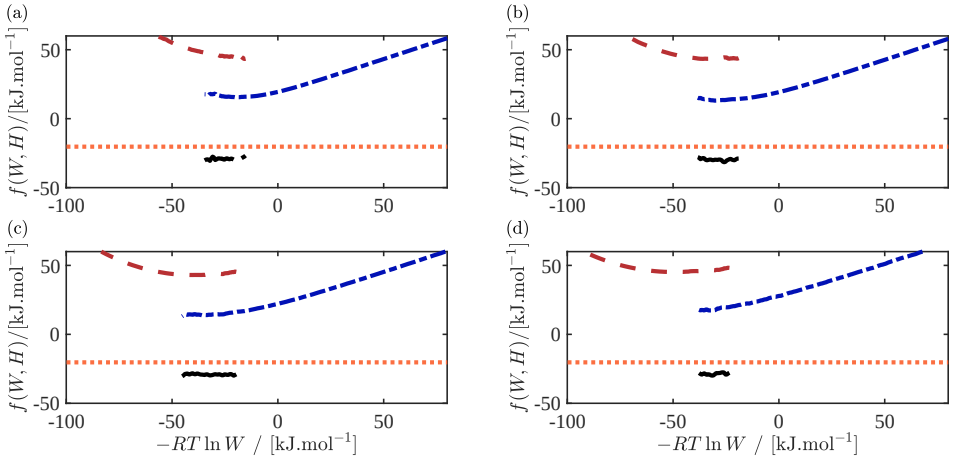


Figure 2.8: Two-dimensional overlapping distribution method (Eq. 2.5) applied to the TIP4P/EW liquid water model in the *NVT* ensemble with a void fraction of $\phi = 0.45$, based on the computed coexistence densities at T=300 K. The sampled hydrogen bond count in each subfigure equals: (a) $H = 0$; (b) $H = 1$; (c) $H = 2$; (d) $H = 3$. In all sub-figures: (dashed lines): $f_1(W,H)$, (dash-dotted lines): $f_0(W,H)$, (dotted lines): $\mu^{\mathrm{ex}}_{\mathrm{WTPI}}$, (lines): $f_0(W,H) - f_1(W,H)$.

Table 2.1 shows number densities and void fractions of all systems in the liquid phase at coexistence in the temperature range of T=250 K to T=350 K. It can be seen that water has the largest number density (around 0.03) between T=300 K to T=350 K, while number densities of other systems are much smaller (around 0.01) in the temperature range of T=220 K and T=350 K. Therefore, the number density cannot consistently offer a criterion for limitation of the WTPI method for water and methanol. For systems studied in this work, efficiency of the WTPI method appears to be correlated with the void fraction of the system. There is a clear distinction between void fractions of water and methanol systems (where WTPI method fails) compared to void fractions of carbon dioxide and hydrogen sulfide systems (where WTPI method works). It seems that for all systems in this study, the WTPI method fails when the void fraction of the system drops approximately below 0.50. This has also been independently tested

using the ODM in the NVT ensemble as can be seen in the Supporting Information of Ref. [59]. Fig. 2.9 shows the relative difference between the chemical potentials in the coexisting gas and liquid phases as a function of void fraction of the liquid system. It can be seen in Fig. 2.9 that as the void fraction drops below 0.5, the relative difference increases rapidly. A similar conclusion can be drawn for a a LJ liquid, which is shown in the Supporting Information of Ref. [59]. Providing a more accurate criterion on the limitations of the WTPI method requires studying several other components.

Although the failure of the WTPI method is explained for all systems based on the void fraction, the computed chemical potentials of water using the CFCGE MC method deviate from the empirical EoS in case of four-site and five-site models. This is most probably a limitation of the force field since chemical potentials of the two phases are equal using the CFCGE MC method. Fig. 2.10 shows computed excess chemical potentials of the aforementioned water models in the CFCGE MC together with those reported by Coskuner and Deiters [188] who used a five-site water model. Moreover, the excess chemical potential of water at different temperatures was computed using IAPWS empirical EoS [143] and shown in Fig. 2.10. Similar to the EoS, the excess chemical potentials of water obtained from CFCGE MC method increase linearly in the temperature range between T=300 K and T=350 K. Since the water models were not fitted to experimental chemical potential data, some deviation from the empirical data is expected depending on the model. It is shown in Fig. 2.10 that other test particle methods (Theodorou's deletion method [183], Widom's test particle deletion method [182]) fail to compute the chemical potentials of water accurately as well. Excess chemical potentials of methanol, hydrogen sulfide and carbon dioxide are also listed in the Supporting Information of Ref. [59].

Table 2.1: Number densities and void fractions of carbon dioxide, hydrogen sulfide, methanol and water in the coexisting liquid phase at different temperatures.

$T/[\mathrm{K}]$	$\rho_n/[\mathrm{\AA}^{-3}]$	$\phi/[\text{-}]$
CO_2		
220	0.0160	0.51
230	0.0153	0.52
240	0.0150	0.54
250	0.0143	0.55
H_2S		
210	0.0167	0.54
230	0.0160	0.56
250	0.0155	0.57
CH_3OH		
240	0.0158	0.44
270	0.0152	0.46
300	0.0147	0.48
350	0.0138	0.51
H_2O		
300	0.0332	0.44
325	0.0327	0.45
350	0.0322	0.46

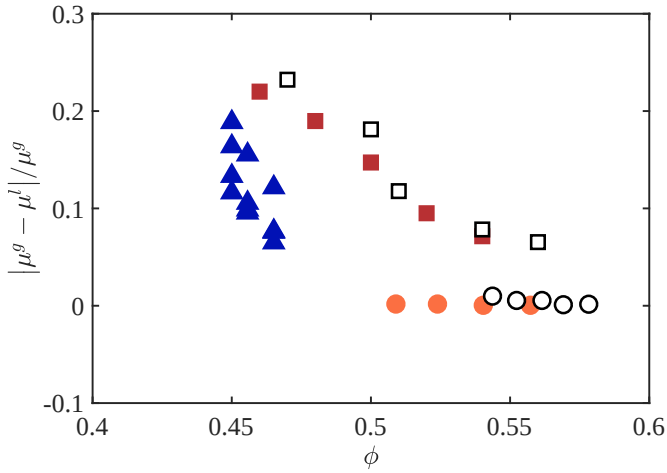


Figure 2.9: Relative difference between the chemical potentials in the coexisting gas and liquid phases using the WTPI method as a function of the void fraction. (closed triangles): Water (SPC, TIP3P/EW, TIP4P/EW, TIP4P/EW, TIP5P-EW); (closed squares): Methanol TraPPE; (open squares): Methanol OPLS-UA; (circles): carbon dioxide TraPPE; (open circles): hydrogen sulfide.

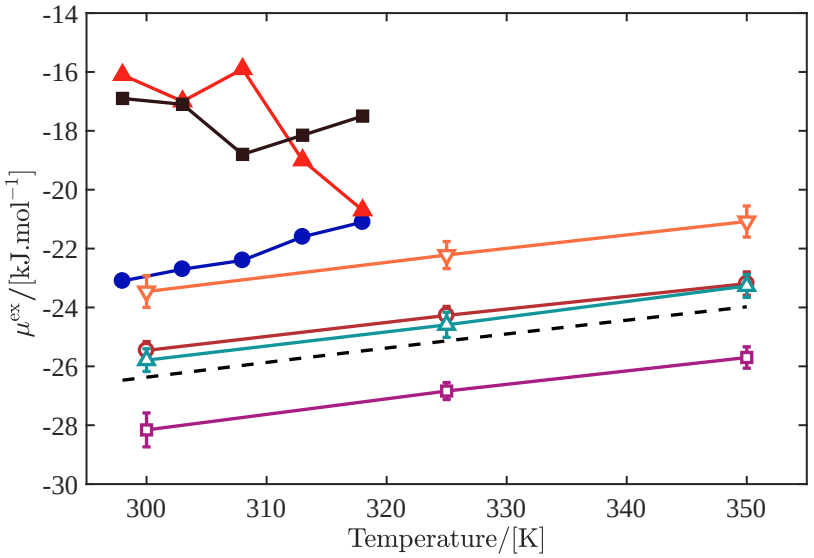


Figure 2.10: Comparison of computed excess chemical potentials of different water models for different methods at VL coexistence. Chemical potentials of water computed by Coskuner and Deiters [188] at different temperatures with a modified TIP5P model with different methods are shown with closed symbols: (circles): Widom's test particle insertion method, (upward-pointing triangles): Widom's test particle deletion method, (squares): Theodorou deletion method [183]. Excess chemical potentials of different water models using CFCGE MC simulations are shown with open symbols: (downward-pointing triangles): TIP5P-EW water model, (circles): SPC water model, (upward-pointing triangles): TIP3P/EW water model, (squares): TIP4P/EW, (dashed line): based on the empirical Helmholtz equation of state [143] provided by NIST, REFPROP [144]

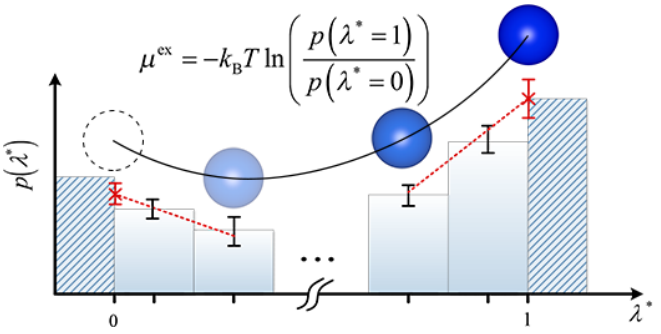
2.5 Conclusions

Despite the lower molecule exchange efficiency of the GE simulations, equal densities from CFCGE MC and GE simulations were obtained for all systems in this study. Computed chemical potentials using CFCGE MC for all systems in study are equal in both phases at coexistence. This is not the case for computed chemical potentials of water and methanol in the two phases using the WTPI method. However, for all cases the chemical potential of the gas phase using the WTPI method agrees well with the chemical potentials computed using the CFCGE MC method. Therefore, it is concluded that WTPI method is unable to compute the chemical potential of water and methanol accurately in the liquid phase. Since all densities computed in CFCGE MC are in very good agreement with the GE simulations in this study, any inefficiency of the WTPI method originates from sampling the excess part. By gradual molecule insertion and removal during the MC simulations, the CFCGE MC method has resolved the sampling issue of the WTPI method. Using the CFCGE MC method, one would be able to compute the chemical potentials of the two phases directly without any further calculations which provides an independent check for the condition of chemical equilibrium. Computed chemical potentials of water using three-site models are in better agreement with IAPWS empirical EoS [143, 145]. The importance of using models capable of accurately predicting the chemical potential of water is highlighted in chapter 8. Chemical potentials of methanol computed using both TraPPE and OPLS-UA force fields and hydrogen sulfide force fields are in excellent agreement with the empirical EoS [145, 231]. Slight deviations in the chemical potential of carbon dioxide relative to the experimental EoS [232] are observed. The inefficiency of the WTPI method in water with a strong hydrogen bonding network is independent on the the number of hydrogen bonds the test molecule forms with its surrounding molecules. For all systems in this study, it is shown consistently that efficiency of the WTPI method strongly depends on the void fraction of the system and reduces significantly for void fractions smaller than 0.50. More data may be needed to determine a more accurate limit for void fraction where the WTPI works correctly.

Chapter 3

Improving the Accuracy of Computed Chemical Potentials in CFCMC Simulations

This chapter is based on the following paper: Rahbari, A.; Hens, R.; Dubbeldam , D.; Vlugt, T. J. H., Improving the accuracy of computing chemical potentials in CFCMC simulations, Molecular Physics, 2019, 117 ,3493-3508, Ref. [41]



3.1 Introduction

In the previous chapter, the CFCMC method was introduced and the advantages of combining CFCMC with the GE were demonstrated for water, methanol, carbon dioxide and hydrogen sulfide at low temperatures. As was shown in chapter 2, one of the key aspects of this method is the direct calculation of the chemical potential by binning of the coupling parameter λ and using the probabilities $p(\lambda = 0)$ and $p(\lambda = 1)$, which require extrapolation [38]. In this chapter, we show that that performing extrapolation leads to systematic errors when $p(\lambda)$ is steep. To eliminate this systematic error, we propose an alternative binning scheme which improves the accuracy of computed chemical potentials. We also investigate the difference between CFCMC averages (uncorrected for the bias) and Boltzmann averages including multiple fractional molecules. In addition, we investigate the use of multiple fractional molecules needed in simulations of multiple components, and show that these fractional molecules are very weakly correlated and that calculations of chemical potentials are not affected.

The main new element of the CFCMC method by Shi and Maginn is that λ has been changed from a discrete parameter into a continuous parameter [45, 141]. In principle, the intermediate λ states can be either continuous or discrete. Continuous and discrete intermediate stages for λ are both commonly used in expanded ensembles [45, 141, 197]. The advantage of having a continuous λ is that changes in λ (denoted by $\Delta\lambda$) can be adjusted to facilitate transfers between intermediate λ states. This eliminates the guesswork about how many intermediate stages are needed. When the number of intermediate stages is close to optimal, we do not expect much differences in the accuracy of the computed chemical potentials between continuous and discrete staging.

Since in CFCMC simulations, insertions/deletions are performed with fractional molecules, biasing of λ is used to improve molecule transfer efficiency [38, 45, 141, 142]. Adaptive computation of the weight function $W(\lambda)$ is performed iteratively to obtain a flat distribution of λ [199, 200]. This significantly improves the efficiency of CFCMC simulations. Using an optimum weight function in the simulations ensures smooth transitions

between $\lambda = 0$ and $\lambda = 1$. In CFCMC simulations, ensemble averages of thermodynamic properties can be computed, either Boltzmann averages or biased averages (uncorrected for the bias introduced by the biasing potential $W(\lambda)$). The Boltzmann average of any observable X is obtained from [38, 142]:

$$\langle X \rangle_{\text{Boltzmann}} = \frac{\langle X \exp\left[-W\left(\lambda\right)\right] \rangle_{\text{Biased}}}{\langle \exp\left[-W\left(\lambda\right)\right] \rangle_{\text{Biased}}} \tag{3.1}$$

Eq. 3.1 is used to transform the averages back from CFCMC simulations in which biasing is applied [38, 142]. Biased averages are obtained by taking the normal averages without correcting for the bias:

$$\langle X \rangle_{\text{Biased}} = \frac{\sum_{i=1}^{N_{\text{S}}} X}{N_{\text{S}}} \tag{3.2}$$

where N_{S} is the number of times the observable X is sampled. The NPT ensemble and an expanded version of this ensemble, the Continuous Fractional Component NPT (CFCNPT) ensemble, are used. Our conclusions for this ensemble can be easily extended to other ensembles. For simulations in the CFCNPT ensemble, the averages of Eq. 3.2 may be considered as approximations for averages in the CFCNPT ensemble, which in turn are approximations for averages in the conventional NPT ensemble. As the CFCNPT and conventional NPT ensemble have a different number of degrees of freedom [14, 20], ensemble averages in both ensembles are in principle different, but in practice these differences are small [38, 142].

As explained in chapter 2, the excess chemical potential can be computed by sampling the Boltzmann probability distribution of the coupling parameter, $p(\lambda)$. See appendix A.1.3 for derivation of the excess chemical potential in the CFCNPT ensemble. However, in the 2016 version of the CFCMC method [38] (as used in chapter 2), it is not possible to directly sample $p(\lambda = 0)$ and $p(\lambda = 1)$. Instead, one relies on extrapolation ($\lambda \downarrow 0$ and $\lambda \uparrow 1$) to compute the excess chemical potential. Therefore, it is necessary to use a binning scheme to sample the distribution $p(\lambda)$ which may affect

the accuracy of the method. In Ref. [38] it was proposed that in practice linear extrapolation of $p(\lambda)$ is sufficient to calculate the excess chemical potential using Eq. 2.9. A clear distinction needs to be made between "precise" and "accurate" computation of the excess chemical potential. The values for the computed excess chemical potential may be systematically wrong (inaccurate) with small error bars (precise). This leads to a false impression of precision while missing accuracy (large difference from the actual value). This sampling issue appears especially for systems in which the number of bins, N_b, is insufficient to capture the steepness of distribution $p(\lambda)$, leading to inaccurate extrapolation results. One could increase N_b to improve the accuracy of the extrapolation, however this leads to poor sampling of $p(\lambda)$ (less statistics per bin) and therefore loss of precision of the extrapolation. Therefore, it is not a priori clear which value to select for N_b for different systems. In this chapter, we investigate how the accuracy of the extrapolation scheme changes with N_b, and we develop a much more accurate scheme that allows a continuous coupling parameter $\lambda \in [0, 1]$ without having to use extrapolation for chemical potential calculations. The new scheme allows sampling a continuous coupling parameter including the states $\lambda = 0$ and $\lambda = 1$. This means that the chemical potential is obtained independent of any extrapolation scheme since the states $\lambda = 0$ and $\lambda = 1$ are directly sampled. We will show that this significantly improves the accuracy of computed values of μ^ex for systems with strong intermolecular interactions.

The effect of N_b on the accuracy and precision of our new binning scheme is also investigated. Simulations in the CFCMC ensemble with multiple fractional molecules may be used to study complex systems *e.g.* the multicomponent Gibbs ensemble chapter 8, the reaction ensemble chapter 5, and the reaction ensemble combined with phase equilibria [142, 173, 175, 176]. It is not recommended to include more fractional molecules in the system than required. However, in many cases it is necessary to use multiple fractional molecules [142]. Therefore, it is important to understand how multiple fractional molecules influence computed properties. For dense systems or systems in which N_frac fractional molecules are present, the multidimensional weight function $W(\lambda_1, \lambda_2, \ldots, \lambda_{N_\mathrm{frac}})$ becomes steeper with increasing N_frac. This results in difficulties when sampling Boltzmann

averages (Eq. 3.7). In principle, N_{frac} is the number of fractional molecule types in the simulation. This is important to consider when performing simulations in the reaction ensemble as fractional molecule types of reactants and reaction products are different [142]. For the rest of this chapter, all fractional types are considered the same, however the conclusions are transferable to the reaction ensemble [142].

Another drawback is the difficulty of computing the multidimensional weight function using an adaptive scheme such as the WL algorithm [199, 200]. To calculate the biasing function, a multidimensional histogram has to be filled until some flatness criterion is met, which can be difficult computationally. We find that splitting the multidimensional weight function into a sum of one dimensional weight functions can improve the calculation of the biasing function $W(\lambda)$ and sampling of Boltzmann averages. To the best of our knowledge, the effect of having multiple fractional molecules on the statistics of Boltzmann averages and biasing in CFCMC simulations are not systematically investigated/reported in literature.

In this chapter, three important points relevant to systems with multiple fractional molecules are investigated: *(1)* The correlation between λ's of different fractional molecules are investigated. *(2)* Sampling of Boltzmann averages using Eq. 3.1 is numerically difficult if the weight function is large. Due to this, sampling of the biased averages, Eq. 3.2, is an attractive alternative to Boltzmann averages in CFCMC simulations. Therefore, it is of interest to study the difference between the Boltzmann and biased averages for different systems. *(3)* The excess chemical potential is a thermodynamic property for any system state, independent of the number of the fractional molecules, N_{frac}. Therefore, it is important to check whether the value of the computed chemical potentials varies with N_{frac}.

3.2 Theory and Computational Methods

In chapter 2, a continuous coupling parameter $\lambda \in \langle 0, 1\rangle$, for each fractional molecule type was used in the partition function, corresponding to the 2016 version of the CFCMC [20, 38, 59, 142, 163, 233], and LJ interactions were scaled with λ (*i.e.* according to Eq. 2.10). Scaling of electrostatic interactions with λ is explained in detail in chapter 4. Using the 2016 version of CFCMC in chapter 2, it was not possible to directly sample the system states in which exactly $\lambda = 0$ or $\lambda = 1$. Here, we introduce a coupling parameter $\lambda^*(\lambda) \in [0, 1]$ to calculate the atomistic/molecular interactions, including system states when the interactions of the fractional molecule are completely switched on or off. *E.g.* for LJ interactions, this means that λ in Eq. 2.10 is replaced by λ^*, which is a function of λ. $\lambda^*(\lambda)$ is obtained from linear transformation of λ:

$$\lambda^*(\lambda) \equiv \begin{cases} 0, & \lambda < \frac{1}{N_\mathrm{b}} \\ \frac{N_\mathrm{b}\lambda - 1}{N_\mathrm{b} - 2}, & \frac{1}{N_\mathrm{b}} \leq \lambda \leq \frac{N_\mathrm{b}-1}{N_\mathrm{b}} \\ 1, & \lambda > \frac{N_\mathrm{b}-1}{N_\mathrm{b}} \end{cases} \tag{3.3}$$

in which N_b is the number of the bins. The transformation of Eq. 3.3 is also used for scaling the electrostatic interactions of the fractional molecule. It is important to note that the extended parameter in the partition function is still $\lambda \in \langle 0, 1\rangle$. Using the transformation of Eq. 3.3, only the interactions of the fractional molecule are scaled with $\lambda^*(\lambda) \in [0, 1]$ in an extra step. It follows directly from Eq. 3.3 that $\lambda^*(\lambda)$ is a continuous function at $\lambda = \frac{1}{N_\mathrm{b}}$ and $\lambda = \frac{N_\mathrm{b}-1}{N_\mathrm{b}}$. Scaling the interactions of the fractional molecule using $\lambda^*(\lambda)$ means that there are now two bins in λ space where interactions are completely switched on or off. Therefore, one can directly sample the probability of $\lambda^*(\lambda) = 0$ in the first bin, and the probability of $\lambda^*(\lambda) = 1$ in the last bin. The linear transformation of Eq. 3.3 is illustrated in Fig. 3.1 for $p(\lambda)$ and $p(\lambda^*(\lambda))$. The inset of this figure shows the function $\lambda^*(\lambda)$ in Eq. 3.3.

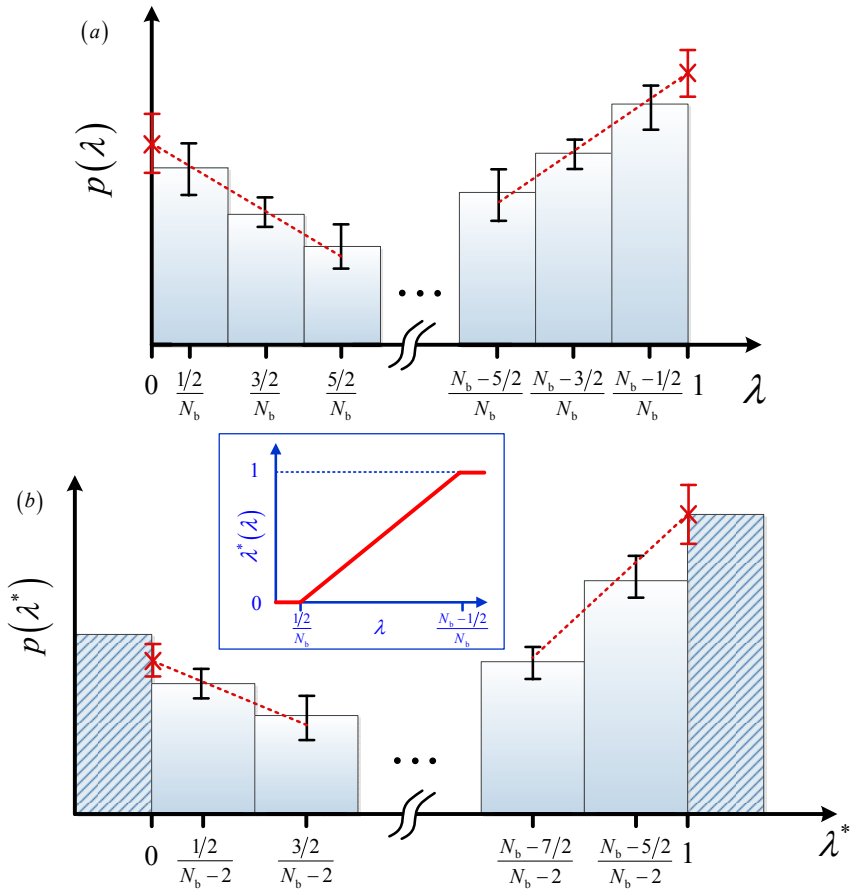


Figure 3.1: Linear transformation of the scaling parameter from λ (subfigure a) to λ^* (subfigure b). Based on the transformation of Eq. 3.3, the value $\lambda^*(\lambda)$ is set to zero for the first bin of $p(\lambda)$, and the value $\lambda^*(\lambda)$ equals one for the last bin. When the interaction parameter $\lambda^*(\lambda) = 0$, the fractional molecule behaves as an ideal gas, and when $\lambda^*(\lambda) = 1$, the fractional molecule behaves exactly as a whole molecule. The inset shows how λ^* depends on λ (Eq. 3.3).

As shown in Fig. 3.1(a), $p(\lambda)$ is constructed by sampling the probability of λ where the λ space is binned at equal distances; $[\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots, N_\mathrm{b} - \frac{5}{2}, N_\mathrm{b} - \frac{3}{2}, N_\mathrm{b} - \frac{1}{2}]$ in units of $\frac{1}{N_\mathrm{b}}$. The width of each bin, $\Delta\lambda$, equals $\frac{1}{N_\mathrm{b}}$ and value of λ assigned to each bin equals the middle of the bin, i.e. $\frac{i-1/2}{N_\mathrm{b}}$. Therefore, the value of λ in the first and last bins correspond to $\lambda = \frac{1}{2N_\mathrm{b}}$ and $\lambda = \frac{N_\mathrm{b}-1/2}{N_\mathrm{b}}$, respectively, and not to 0 or 1. To calculate μ^ex from $p(\lambda)$ instead of $p(\lambda^*(\lambda))$, one needs to perform a linear extrapolation on the first/last few points of $p(\lambda)$ [38]. The distribution $p(\lambda^*(\lambda))$ can be directly reconstructed from $p(\lambda)$ in a single step using Eq. 3.3. As shown in Fig. 3.1(b), $p(\lambda^*)$ is constructed using bins with the values of $[0, \frac{1}{2}, \frac{3}{2}, \ldots, N_\mathrm{b} - \frac{7}{2}, N_\mathrm{b} - \frac{5}{2}, 1]$ in units of $\frac{1}{N_\mathrm{b}-2}$. This grid is continuous but non-equidistant. Using the new binning scheme, μ^ex can be obtained directly using the probabilities of the first and the last bin, as shown in Fig. 3.1(b):

$$\mu^\mathrm{ex} = -\frac{1}{\beta} \ln \left(\frac{p(\lambda^*(\lambda) = 1)}{p(\lambda^*(\lambda) = 0)} \right) \tag{3.4}$$

In principle one could directly sample $p(\lambda = 0)$ and $p(\lambda = 1)$ without any biasing (and hence no binning is required). However, it is well-known that not applying a biasing function $W(\lambda)$ significantly reduces the efficiency of the simulation [38, 142]. In this chapter, we compare the differences between extrapolation, and direct sampling for calculating the chemical potential for different systems. The linear transformation of λ (Eq. 3.3) can be easily implemented in the original CFCMC algorithm. For instance, the partition function of a mixture of S different monoatomic components in the NPT ensemble expanded with a fractional molecule equals [20]

$$\begin{aligned} Q_{\mathrm{CFC}NPT} = \beta P \left[\prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \right] \times \frac{1}{\Lambda^3} \int_0^1 \mathrm{d}\lambda \int \mathrm{d}V V^{N+1} \exp\left[-\beta P V\right] \\ \times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s^\mathrm{A}_\mathrm{frac} \exp[-\beta U^\mathrm{A}_\mathrm{frac}(s^\mathrm{A}_\mathrm{frac}, s^N, \lambda^*(\lambda), V)] \end{aligned} \tag{3.5}$$

in which N is the total number of whole molecules which are distinguishable from the fractional molecule, S is the number of components, Λ_i is the

thermal wavelength of component i, Λ is the thermal wavelength of the fractional molecule, U is the total potential energy of the whole molecules, and $U_{\text{frac}}^{\text{A}}$ is the interaction potential of the fractional molecule with the surrounding molecules scaled with $\lambda^*(\lambda)$. No further changes are required for calculating the weight function $W(\lambda)$ and $p(\lambda)$ during the simulation [20, 59, 142]. Only at the end of the simulation, $p(\lambda)$ is transformed into $p(\lambda^*(\lambda))$ in a single step using Eq. 3.3. Note that the CFCNPT ensemble is used here as an example to explain the method. The linear transformation of the λ can be implemented in open ensembles in a similar manner. The linear transformation of $\lambda^*(\lambda)$ has several advantages: *(1)* The first bin of $p(\lambda)$ corresponds to system states where the interaction potential is completely switched off ($\lambda^*(\lambda) = 0$). At $\lambda^*(\lambda) = 0$, reinsertions of the fractional molecule at a randomly selected position [20] are always accepted since the energy difference between the old and new configurations is zero. It is important to note that the fractional molecule is part of the ensemble partition function and is never deleted from the system even when $\lambda^*(\lambda) = 0$. *(2)* The last bin of $p(\lambda)$, ($\lambda^*(\lambda) = 1$), corresponds to system states where the fractional molecule is interacting as a whole molecule. For $\lambda^* = 1$, identity changes of the fractional molecule [20] with a whole molecule are always accepted as the energy difference between the old and new configurations is zero. In the identity change trial moves, the fractional molecule is changed into a whole molecule of the same type, and a randomly selected whole molecule of the same molecule type is changed into a fractional molecule, while keeping the value of λ, positions and orientations of the molecules unchanged [20, 38, 142]. The identity change trial move can also serve as an independent check of the correctness of the simulation code and the bookkeeping. Essentially, the transformation of Eq. 3.3 allows rigorous sampling of the states $p(\lambda^*(\lambda) = 0)$ and $p(\lambda^*(\lambda) = 1)$ during the simulation without performing extrapolation. This method combines the benefits of free energy calculations in the CFCMC simulations with rigorous sampling of states in which $\lambda^* = 0$ and $\lambda^* = 1$ [20, 38, 197].

It is straightforward to extend the partition function of Eq. 3.5 to systems with multiple fractional molecules [20]. In CFCMC simulations with multiple fractional molecules, the biasing function W is a multidi-

mensional weight function [142] used to improve the efficiency of molecule insertions/removals and smooth transitions between $\lambda = 0$ and $\lambda = 1$ for every fractional molecule. However, calculating a multidimensional adaptive biasing function requires filling and flattening a multidimensional histogram during a random walk in $(\lambda_1, \lambda_2, \cdots)$ space, using a certain flatness criterion. Filling multidimensional histograms can be difficult with many fractional molecules in the system, *e.g.* using the WL algorithm [199, 200]. One could split the multidimensional biasing function into a series of one dimensional biasing functions. For a system in which N_{frac} fractional molecules are present, this leads to

$$W\left(\lambda_1, \lambda_2, \ldots \lambda_{N_{\text{frac}}}\right) \approx \sum_{i=1}^{N_{\text{frac}}} W_i\left(\lambda_i\right) \tag{3.6}$$

Filling multiple independent one-dimensional histograms is computationally more straightforward than filling a single multidimensional histogram. The biasing is then calculated for each λ_i independently. In Eq. 3.6, it is assumed that the λ_i's are independent coupling parameters. If there would be a strong correlation between λ's, the computed Boltzmann averages are still correct. However, the sampling of the distributions $p(\lambda_i)$ may be very inefficient due to neglected correlations between λ's (Eq. 3.6). By combining Eqs. 3.1 and 3.6, the Boltzmann average of any observable X is obtained as follows

$$\langle X \rangle_{\text{Boltzmann}} = \frac{\left\langle X \exp\left[- \sum_{i=1}^{N_{\text{frac}}} W_i\left(\lambda_i\right) \right] \right\rangle_{\text{CFC}NPT}}{\left\langle \exp\left[- \sum_{i=1}^{N_{\text{frac}}} W_i\left(\lambda_i\right) \right] \right\rangle_{\text{CFC}NPT}} \tag{3.7}$$

In many systems with strong intermolecular interactions or with multiple fractional molecules, the weight function $\sum_{i=1}^{N_{\text{frac}}} W_i\left(\lambda_i\right)$ is a large number, typically between 10^1 to 10^2 [20, 222]. This means that the exponents

in Eq. 3.7 are very small for such systems. This results in averaging over very small numbers, numerically close to zero, when sampling Boltzmann averages of Eq. 3.7. Therefore, taking Boltzmann averages for these systems may mostly lead to a $\frac{0}{0}$ numerical problem for ensemble averages like volume and energy. Except for excess chemical potentials, most ensemble averages hardly depend on the instantaneous values of λ's. In Refs. [38, 233], it was shown that the presence of multiple fractional molecules hardly influences the thermodynamic properties of the system however, the statistics of the Boltzmann averages are affected. To avoid the $\frac{0}{0}$ sampling problem of the Boltzmann averages, a possible solution is to sample biased averages as shown in Eq. 3.2. Here, we investigate how computed averages change with the number of fractional molecules. Preferably, one should use as few fractional molecules as possible in production runs. If no fractional molecules are required, it is recommended to use conventional ensembles instead of expanded ensembles.

It is not a priori clear whether fractional molecules are weakly or strongly correlated. The requirement for efficient splitting of the biasing, Eq. 3.6, is that λ_i's are independent. To validate this, we compute the pairwise correlation between different λ_i's as a function of the number of fractional molecules in the system, while keeping the number of whole molecules constant. The pairwise correlation between two (randomly) selected λ_i's in a simulation can be calculated by computing the correlation [234]:

$$\mathrm{Corr}\left(\lambda_1, \lambda_2\right) = \frac{\langle \lambda_1 \lambda_2 \rangle - \langle \lambda_1 \rangle \langle \lambda_2 \rangle}{\sqrt{\left[\langle \lambda_1^2 \rangle - \langle \lambda_1 \rangle^2\right]\left[\langle \lambda_2^2 \rangle - \langle \lambda_2 \rangle^2\right]}} \tag{3.8}$$

where λ_1 and λ_2 are the instantaneous values of two randomly selected coupling parameters during the single simulation. Eq. 3.8 can be applied to systems with and without biasing. In addition, we investigate how the presence of multiple fractional components influences the computed values of μ^{ex} and other thermodynamic properties such as the average volume, density and energy.

3.3 Simulation Details

As a proof of principle, the performance of the original binning scheme and the binning scheme of Eq. 3.3 are compared for a 2-atom model system consisting of two LJ molecules in one-dimensional phase space. Here, reduced units are used, so $\epsilon = 1$ and $\sigma = 1$. The 2-atom model system has two degrees of freedom, namely the interatomic distance r and λ. The partition function for this ensemble equals:

$$Q = \frac{1}{L} \int_0^L \mathrm{d}r \int_0^1 \mathrm{d}\lambda \exp[-\beta U(r,\lambda)] \tag{3.9}$$

where we selected $L = 3$, in units of σ, $\beta = 1/T^*$ in reduced units, and T^* is the reduced temperature. The interaction potential $U(r,\lambda)$ is a function of the distance $r \in [0,3]$ and $\lambda \in \langle 0,1\rangle$, obtained from Eq. 2.10. By performing long simulations, we can compute $p(\lambda)$ with brute-force sampling of λ and r. From the original binning scheme it follows that:

$$p(\lambda) = \frac{\int_0^L \mathrm{d}r \int_0^1 \mathrm{d}\lambda' \exp[-\beta U(r,\lambda')]\delta(\lambda - \lambda')}{\int_0^L \mathrm{d}r \int_0^1 \mathrm{d}\lambda' \exp[-\beta U(r,\lambda')]} \tag{3.10}$$

In the new binning scheme of Eq. 3.3, the term $-\beta U(r,\lambda)$ is replaced by $-\beta U(r,\lambda^*(\lambda))$ and after the simulation the distribution $p(\lambda)$ is converted to $p(\lambda^*(\lambda))$. Simulations are carried out at different temperatures between $T^* = 0.005$ and $T^* = 2$ in reduced units. For both binning schemes, the simulations at every temperature are repeated with different values of N_b ranging from 10 to 500. In each cycle, r and λ are randomly selected from uniform distributions, and the probability of λ is sampled using Eq. 3.10. To compare the simulation results, a reference value of $p(\lambda^* = 1)$ is obtained from direct sampling of the last bin from simulations carried out 10 times longer. Since the value of the last bin is directly sampled, no systematic errors are present in this reference value. To obtain $p(\lambda \uparrow 1)$ in the original binning scheme, linear extrapolation is carried out using the last 3 points of the λ grid. $p(\lambda^*(\lambda) = 1)$ is obtained by directly sampling the last bin

in the new binning scheme. For all the simulations, 10^8 random states of (r, λ) were generated to sample the probability of λ. To obtain the reference values for $p(\lambda) = 1$, 10^9 random states of (r, λ) were generated.

Simulations of SPC/E [235] water and TraPPE methanol [207] are performed in the CFCNPT ensemble [20] at $T = 323.15$ K and $p = 1$ bar. Both the original and the new binning scheme are used to compute excess chemical potentials. To investigate the effect of binning on chemical potential calculations, simulations are performed with different values of the number of bins, N_b, ranging from 5 to 100, for both binning schemes. All molecules are modelled as rigid objects, and the intermolecular potential consists only of LJ and Coulombic interactions. A cutoff radius of 14 Å is used for LJ interactions, and the DSF version of the Wolf method [149, 236–239] is used for handling electrostatic interactions. R_c and α were set to 14 Å and 0.12 Å^{-1}. For details on selecting R_c and α for water and methanol, the reader is referred to chapter 4 or Refs. [163, 222]. The LJ interactions of the fractional molecules are scaled using Eq. 2.10. The scaling of the Coulombic interactions of fractional molecules is described in chapter 4 or Ref. [222]. To protect the charges from overlapping, the LJ interactions of the fractional molecules are switched on before the electrostatics [224], see Fig. 2.3. Analytic tail corrections and periodic boundary conditions are applied [15]. The Lorentz–Berthelot mixing rule [14, 15] is used to calculate cross interactions. Force field parameters for SPC/E water and TraPPE methanol are provided in Tables 4.1 and 4.2. In CFCNPT ensemble simulations, beside thermalization trial moves [14] (volume changes, translations, rotations etc), three additional trial moves involving the fractional molecule were used to facilitate the gradual insertion/removal of molecules during the simulation: *(1)* Changes in λ: the coupling parameter λ is changed while keeping the positions of all molecules including the fractional molecule constant [142]. For changes in λ, it is required that λ is confined to the interval $[0, 1]$ [38, 142]. *(2)* Reinsertions: the fractional molecule is reinserted at a randomly selected position while keeping the positions of all the whole molecules and the value of λ constant [142]. *(3)* Identity changes: the fractional molecule is changed into a whole molecule of the same component, and a randomly

selected whole molecule of the same component is changed into a fractional molecule while keeping the value of λ constant [142]. These trial moves are accepted or rejected based on Metropolis acceptance rules (and automatically rejected when the new value of λ is outside the interval $[0, 1]$) [14]. It is efficient to combine trial moves *(2)* and *(3)* into a single hybrid trial move, as trial move *(2)* has a high acceptance probability only at low values of λ, and trial move *(3)* has a high acceptance probability only at high values of λ. It is possible to define a hybrid trial move in which trial moves of type (2) are only selected at low values of λ, and trial moves of type (3) are only selected at high values of λ. Therefore, trial moves (2) and (3) are only selected when the acceptance probabilities are high. In Refs [20, 142], it is shown that such a hybrid trial move obeys detailed balance. In practice, one uses a switching point at $\lambda = \lambda_s$ to select either a trial move of type *(2)* or *(3)*. To facilitate the sampling of λ, a weight function ($W(\lambda)$) was used to ensure that the sampled probability of λ is flat [45, 141]. In all simulations, maximum molecule displacements, maximum rotations, and maximum volume changes were adjusted to achieve on average 50% acceptance. We found that the hybrid trial moves significantly improve the sampling and reduce the error bars of the computed chemical potentials. Simulations in the CFCNPT ensemble of the SPC/E water [235] are started with with 10^5 equilibration cycles, followed by 4×10^6 production cycles. In each MC cycle, the number of trial moves equals the total number of molecules, with a minimum of 20. The trial moves are selected with the following probabilities: 1% volume changes, 35% translations, 30% rotations, 17% λ changes, 8.5% reinsertions of fractional molecules at randomly selected positions, and 8.5% identity changes of fractional molecules.

Simulations of LJ color mixtures ($\sigma = 1$ and $\epsilon = 1$) are carried out in the CFCNPT and NPT ensembles, at $T^* = 2$ and pressures between $P^* = 0.5$ and $P^* = 0.6$. For these systems, the LJ interactions were truncated and shifted at 2.5σ. In the CFCNPT simulations, 800 whole molecules are present. For every temperature and pressure, the simulations are repeated with different number of fractional molecules, $3 < N_{\text{frac}} < 50$ while keeping the number of whole molecules constant. In practice, when studying complex molecular systems, N_{frac} is nearly always below 5 [20, 38, 59, 142,

158, 163, 177, 222, 240]. Larger values of N_{frac} can be considered as an extreme situation to test the limits of the CFCMC method. The percentage of the fractional molecules in the CFCNPT simulations pr_{f}, changes between 0.125% to 6.25%. At each temperature and pressure, simulations are carried out with 10^5 equilibration cycles to equilibrate the system. From the equilibrated configurations, 10^6 productions runs are carried out to sample both Boltzmann averages, Eq. 3.7, and biased averages, Eq. 3.2. For the CFCNPT simulations, the trial moves in every MC step are selected with the following probabilities: 1% volume changes, 49% translations, 20% λ changes, 15% reinsertions of fractional molecules at a randomly selected position, and 15% identity changes of fractional molecules. The trial moves in simulations in the conventional NPT ensemble (i.e. without fractional molecules) are selected with probabilities: 1% volume changes and 99% translations. All trial moves are accepted or rejected based on the Metropolis acceptance rules [14].

3.4 Results

MC simulations are performed for the 2-atom model system in the ensemble of Eq. 3.9, between $T^* = 0.005$ and $T^* = 2$. The distributions $p(\lambda)$ and $p(\lambda^*)$ are sampled using Eq. 3.10. Linear extrapolation is performed on the last 3 bins of $p(\lambda)$ to calculate $p(\lambda \uparrow 1)$. The value of $p(\lambda^*(\lambda) = 1)$ is obtained by the direct sampling scheme. The results obtained for temperatures between $T^* = 0.005$ and $T^* = 0.05$ are shown in Table 3.1, and the raw data for temperatures between $T^* = 0.1$ and $T^* = 2$ are provided in the Supporting Information of Ref. [41]. In Table 3.1, it is shown that the reference values obtained from the direct sampling of the last bin are very similar, independent of the number of bins N_{b}. The results from the extrapolation scheme systematically deviate from the reference values for small N_{b}, while the uncertainties (standard deviation of the mean) are very small. This leads to a false impression of accuracy. Good agreement between the results based on the extrapolation scheme and the reference values is found with increasing N_{b}. However, a larger value of N_{b} leads to a

Table 3.1: Comparison of $p(\lambda = 1)$ for the 2-atom model system in the temperature range between $T^* = 0.005$ and $T^* = 0.05$, using different number of bins ranging from 10 to 500. N_b is the number of bins, $p_\mathrm{ext}(\lambda \uparrow 1)$ is obtained using the extrapolation scheme, and $p_\mathrm{dir}(\lambda = 1)$ is obtained using direct sampling. The reference values (i.e. from very long simulations) are denoted with $p_\mathrm{ref}(\lambda = 1)$. Numbers in brackets are uncertainties in the last digit, *i.e.*, 88.9(6) means 88.9 ± 0.6.

N_b	$p_\mathrm{ext}(\lambda \uparrow 1)$	$p_\mathrm{dir}(\lambda = 1)$	$p_\mathrm{ref}(\lambda = 1)$
$T^* = 0.005$			
5	5.4167(0)	199(3)	199.41
10	10.8333(0)	199(2)	199.50
20	21.6660(0)	200(2)	199.38
40	43.127(3)	199(2)	199.60
80	81.32(6)	199(3)	199.47
100	96.8(1)	200(3)	199.46
200	145.0(7)	199(3)	199.45
500	184(3)	200(4)	199.49
$T^* = 0.01$			
5	5.41670(0)	99.4(8)	99.46
10	10.83300(0)	99.3(7)	99.50
20	21.5620(8)	99.4(8)	99.50
40	40.64(2)	99.4(8)	99.43
80	65.2(1)	99.6(1)	99.55
100	72.4(2)	99(1)	99.41
200	88.9(6)	100(1)	99.51
500	97(2)	100(2)	99.54
$T^* = 0.05$			
5	5.3357(2)	19.41(5)	19.41
10	9.605(3)	19.41(5)	19.41
20	14.27(1)	19.40(5)	19.41
40	17.38(4)	19.41(6)	19.41
80	18.75(8)	19.41(1)	19.41
100	18.95(9)	19.4(1)	19.41
200	19.3(2)	19.4(1)	19.41
500	19.4(2)	19.4(2)	19.41

significant increase of the uncertainty of the results (between 1 to 4 orders of magnitude) for the extrapolation scheme. Therefore, it is difficult to a priori know what a sufficient N_b is for the extrapolation scheme. In sharp contrast to the extrapolation scheme, the magnitude of uncertainty does not change significantly with N_b for the direct sampling scheme. Excellent agreement is found between the results obtained from the direct sampling scheme and the reference values, independent of N_b. The simulation results clearly show that the direct sampling scheme is far less affected by the sampling issues pronounced in the extrapolation scheme. Therefore, the direct sampling scheme is recommended as the best method.

To map all results in a single plot, the corresponding bin size $\Delta\lambda = 1/N_\mathrm{b}$ is used as a scaling factor to scale $p(\lambda)$ at every temperature. The scaled probabilities are shown in Fig. 3.2. The advantage of this representation is that the results obtained at multiple temperatures can be shown in a single plot. As an alternative, a plot of $p(\lambda = 1)/p_\mathrm{ref}(\lambda = 1)$ versus N_b for different temperatures is provided in the Supporting Information of Ref. [41]. From Fig. 3.2, it is clear that the extrapolation scheme reaches its limitation for $\Delta\lambda \cdot p(\lambda = 1) \leq 1$. In Fig. 3.2 it is observed clearly that the performance of the extrapolation scheme depends both on $\Delta\lambda = 1/N_\mathrm{b}$ and the steepness of the distribution $p(\lambda)$. This means that in sharp contrast to the direct sampling, the accuracy of the extrapolation scheme strongly relies on $\Delta\lambda \cdot p(\lambda)$, especially when $p(\lambda)$ is steep. As shown in Fig. 3.2, the values for $\Delta\lambda \cdot p(\lambda = 1)$ obtained from direct sampling scheme are in excellent agreement with the reference values.

The relative uncertainties of $p(\lambda = 1)$ obtained from the simulations of the 2-atom model system are shown in Fig. 3.3, as a function of number of the bins. σ is the uncertainty of $p(\lambda = 1)$. It is clear that for small N_b, the relative uncertainties obtained using the extrapolation scheme are smaller compared to those obtained from the direct sampling. This indicates that the results obtained from the extrapolation scheme may be more precise but less accurate. For large N_b the relative uncertainties for both methods are very similar. Very similar results for $p(\lambda) = 1$ are obtained for both methods when N_b is large. Based on the results obtained from the 2-atom model system, it is obvious that the direct sampling scheme is the best

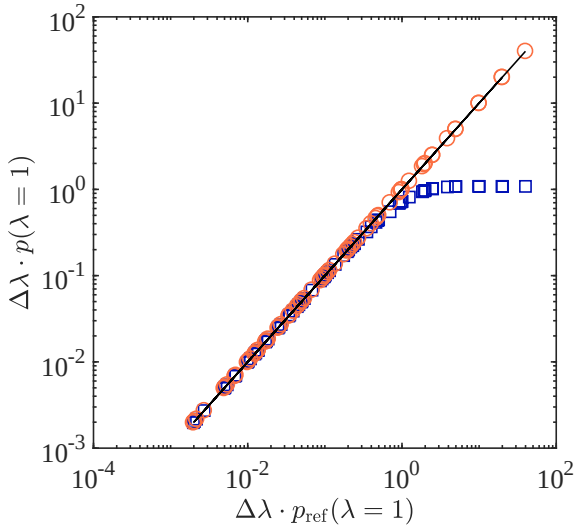


Figure 3.2: Comparison of scaled $p(\lambda = 1)$ for the 2-atom model system in the temperature range between $T^* = 0.005$ and $T^* = 2$, using different number of bins ranging from 10 to 500. To map all results for all temperatures in a single plot, for each system, the corresponding bin size $\Delta\lambda$ is used as a scaling factor for $p(\lambda = 1)$. Alternatively, a plot of $p(\lambda = 1)/p_{\mathrm{ref}}(\lambda = 1)$ versus N_{b} for different temperatures is provided in the Supporting Information of Ref. [41]. The vertical axis is used for the scaled probabilities obtained based on the extrapolation scheme (squares), and direct sampling (circles). The horizontal axis is used for the reference scaled probabilities $\Delta\lambda p_{\mathrm{ref}}(\lambda = 1)$ obtained from very long MC simulations, using direct sampling (thereby eliminating systematic errors). Raw data are listed in the Supporting Information of Ref. [41].

method with the least dependence on N_{b}. This is an important advantage as it may be difficult to a priori know the best value for N_{b} for the other schemes.

As an example of a system with strong LJ and electrostatic interactions, the excess chemical potential of SPC/E water is computed at $T = 323.15$ K and $P = 1$ bar, in the CFCNPT ensemble. The probabilities $p(\lambda \downarrow 0)$ and $p(\lambda \uparrow 1)$ are computed using extrapolation scheme, and $p(\lambda^* = 0)$ and $p(\lambda^* = 1)$ are obtained from the direct sampling. The excess chemical potential of water is computed by performing extrapolation using Eq. 2.9,

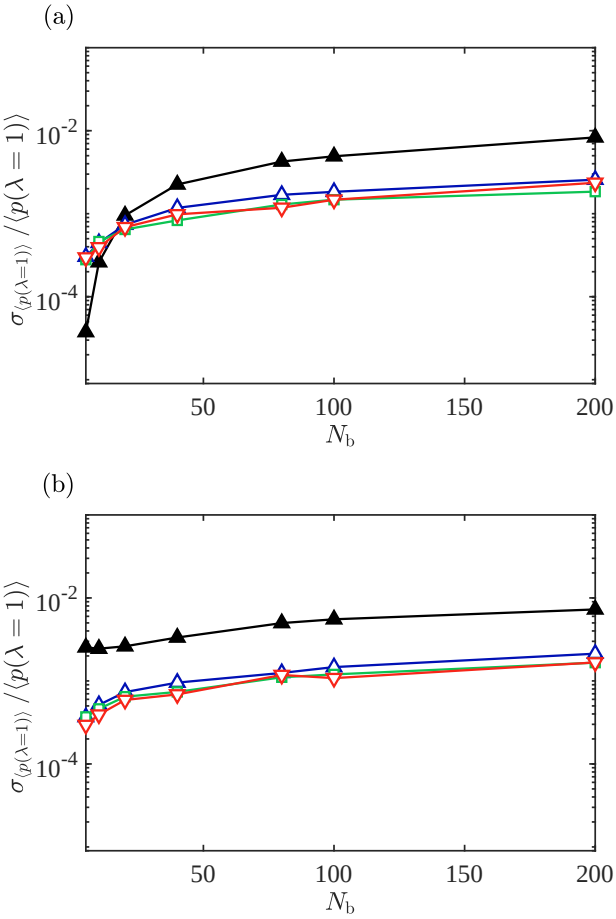


Figure 3.3: Relative uncertainty computed for the sampled $p(\lambda = 1)$ for the 2-atom model system, $\sigma_{p(\lambda=1)}/\langle p(\lambda = 1)\rangle$, using (a) linear extrapolation, Eq. 2.9, and (b) the direct sampling scheme, Eq. 3.3 as a function of number of the bins. The simulations are performed at reduced temperatures: $T^* = 0.05$ (filled triangles), $T^* = 0.5$ (upward-pointing triangles), $T^* = 1.0$ (squares) and $T^* = 1.5$ (down-ward pointing triangles).

and Eq. 3.4 for the direct sampling. The simulations in the CFCNPT ensemble are repeated for different N_b ranging from 5 to 100. The distributions $p(\lambda)$ are scaled with $\Delta\lambda = 1/N_\mathrm{b}$ and the results are shown in Fig. 3.4. Alternatively, plots of $p(\lambda)/p_\mathrm{ref}(\lambda)$ versus N_b for $p(\lambda = 0)$ and $p(\lambda = 1)$ are provided in the Supporting Information of Ref. [41]. Raw data for Fig. 3.4 are also provided in the Supporting Information of Ref. [41]. In Fig. 3.4(a), overall good agreement between all methods is observed except for one outlier for the extrapolation scheme for very few bins ($N_\mathrm{b} = 5$). This means selecting 5 bins for the entire λ space is not sufficient even for extrapolation to $\lambda \to 0$ where $p(\lambda)$ is relatively flat. The distributions $p(\lambda)$ and $p(\lambda^*)$ for water are shown in the Supporting Information of Ref. [41]. The choice of 5 bins may not be practical for CFCMC simulation, but it is considered here only to investigate the limitations of Eqs. 2.9 and 3.4. The scaled probabilities of $\lambda = 1$ are shown in Fig. 3.4(b). The performances of both methods to obtain $p(\lambda = 1)$ for water are very similar to what is observed for the 2-atom model system, as shown Figs. 3.2 and 3.4. The results of the extrapolation scheme deviate significantly from the reference values when $\Delta\lambda \cdot p(\lambda = 1) > 1$. The direct sampling scheme is clearly the best method to calculate $p(\lambda = 0)$ and $p(\lambda = 1)$. The excess chemical potential of water is calculated based on the extrapolation scheme using Eq. 2.9, and the direct sampling using Eq. 3.4.

The computed excess chemical potentials of water are compared to a reference simulation, 10 times longer, where the excess chemical potential is obtained from Eq. 3.4. The relative difference between both methods and the reference are shown in Fig. 3.5. It can be seen in Fig. 3.5 that the accuracy of the extrapolation scheme improves with increasing N_b, while the accuracy of direct sampling scheme is hardly influenced by a change in N_b. However, very large values of N_b makes computing μ^ex more difficult as the statistics of the computed occupancy of the bins of $p(\lambda)$ are reduced. As shown in the Supporting Information of Ref. [41], the free energy barrier as a function of λ that the system needs to overcome is about 12 $k_\mathrm{B}T$ at $T = 323$ K. Using fewer bins for the weight function increases the free energy barrier between adjacent λ bins, which affects the statistics. Increasing the number of bins results in decreasing the free energy barrier between adjacent bins.

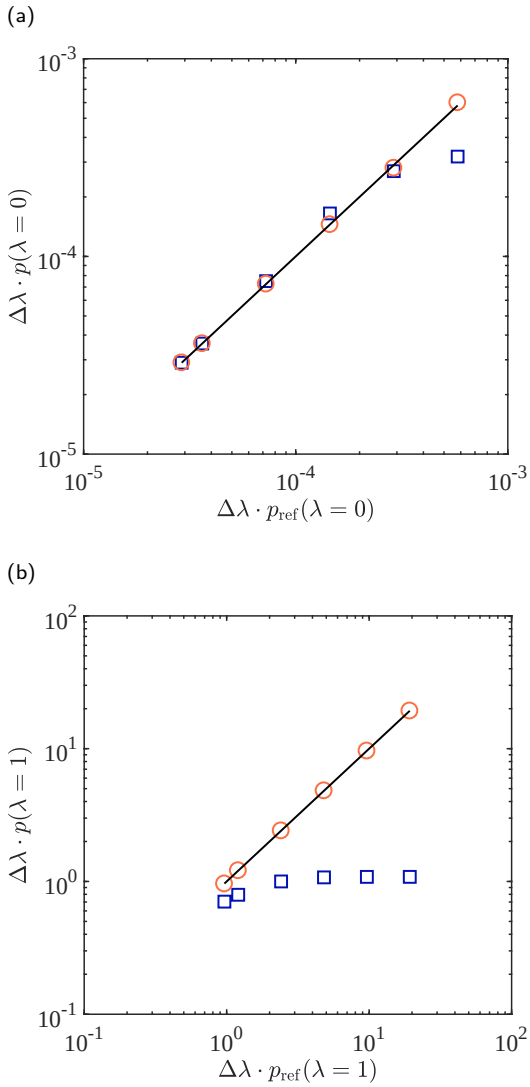


Figure 3.4: Comparison of the scaled probability distributions (a): $p(\lambda) = 0$ and (b): $p(\lambda = 1)$ for the SPC/E water at $T = 323$ K and $P = 1$ bar, using different number of bins ranging from 5 to 100. To map all results in a single plot, for each system, the corresponding bin size $\Delta\lambda$ is used as a scaling factor for $p(\lambda = 0)$ and $p(\lambda = 1)$. Alternatively, plots of $p(\lambda)/p_{\mathrm{ref}}(\lambda)$ versus N_{b} for $p(\lambda = 0)$ and $p(\lambda = 1)$ are provided in the Supporting Information of the Ref. [41]. The vertical axis is used for the scaled probabilities obtained based on the extrapolation scheme (squares), and direct sampling (circles). The horizontal axis is used for the reference scaled probabilities $p_{\mathrm{ref}}(\lambda = 0)$ and $\Delta\lambda p_{\mathrm{ref}}(\lambda = 1)$ obtained longer MC simulations, using direct sampling. Raw data are listed in the Supporting Information of Ref. [41].

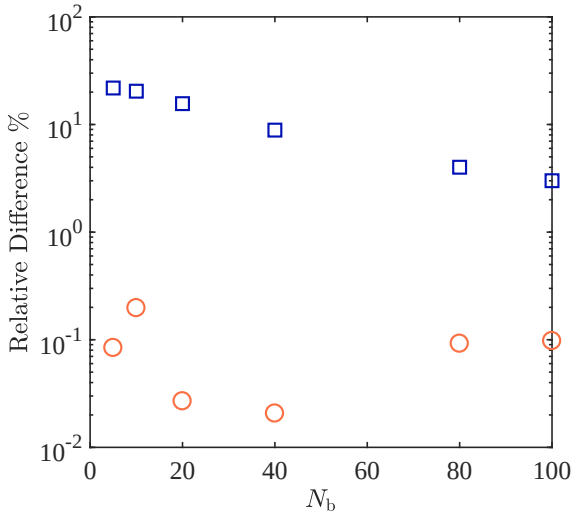


Figure 3.5: Relative difference (in percent) in the computed excess chemical potential of SPC/E water at $T = 323$ K and $P = 1$ bar using the extrapolation scheme Eq. 2.9 (squares), and the direct sampling Eq. 3.4 (circles). The chemical potential obtained using direct sampling from longer MC simulations is considered as the reference value for the chemical potential. The raw data are provided in the Supporting Information of Ref. [41].

However, increasing the number of bins also decreases the statistics of the computed occupancy of bins. The relative difference with respect to the reference values observed for the direct sampling scheme is about one to two orders of magnitude smaller compared to those for the extrapolation scheme. The results clearly show that the direct sampling scheme outperforms the extrapolation scheme. The Boltzmann probability distribution of $p(\lambda)$ for water and methanol in equimolar water-methanol mixture at $T = 323.15$ K and $P = 1$ bar are also shown the Supporting Information of Ref. [41].

To investigate the correlation between the fractional molecules, simulations in the CFCNPT ensemble of LJ color mixtures with multiple fractional molecules are carried out at a reduced temperature of $T^* = 2$ and a reduced pressure of $P^* = 6$. The simulations are repeated by keeping the number of whole molecules constant (800) while changing N_{frac} between 3 and 50. The instantaneous λ's for two randomly selected fractional molecules are recorded every 100 MC cycles. Simulations are performed both with and without biasing. Calculation of the optimal biasing leads to a flat distribution in λ space during the simulation, the so-called observed distribution denoted by $p_{\text{obs}}(\lambda_i)$. It is expected that the average $\langle\lambda_i\rangle$ is close to 0.5 when an optimum biasing is used. Eq. 3.8 is used to calculate the covariance between the two randomly selected coupling parameters and the results are shown in Table 3.2. The correlation between λ_1 and λ_2 is very weak independent of the biasing. The averages $\langle\lambda_1\rangle$ and $\langle\lambda_2\rangle$ are around 0.5 for the simulations when biasing is used. The correlation between for λ_1 and λ_2 is very weak for all the systems studied, independent of the number of the fractional molecules present in the system. The changes in the correlation between λ_1 and λ_2 appear to be very small and random with respect to changes in N_{frac}. λ_1 and λ_2 are also weakly correlated when no biasing is used ($W(\lambda) = 0$), as shown in Table 3.2. Obviously, the average $\langle\lambda_i\rangle \neq 0.5$ when no biasing is used (except for ideal gas). It is clear that the coupling parameters are not correlated in simulations in the CFCNPT ensemble, independent of the weight function $W(\lambda)$. No significant change in the correlation between λ_1 and λ_2 is observed when varying N_{frac}.

As an example of an atomistic system with electrostatic interactions, in an equimolar water-methanol mixture of water-methanol the correlation

Table 3.2: Correlations between two randomly selected fractional molecules in a LJ color mixture at $T^* = 2$ and $P^* = 6$. The system has a constant number of 800 whole LJ molecules, and N_{frac} is the number of the fractional molecules in each simulation. In simulations with non-zero biasing, an independent biasing is calculated for each fractional molecule. The biasing is set such that the observed probability distribution of every fractional molecule, $p_{\text{obs}}(\lambda)$, is flat.

N_{frac}	$\langle\lambda_1\rangle$	$\langle\lambda_2\rangle$	$\langle\lambda_1^2\rangle$	$\langle\lambda_2^2\rangle$	$\langle\lambda_1\lambda_2\rangle$	$\lvert\text{Corr}\,(\lambda_1,\lambda_2)\rvert$
when $p_{\text{obs}}(\lambda)$ is flat						
3	0.520	0.496	0.355	0.329	0.254	0.046
5	0.512	0.499	0.341	0.331	0.254	0.010
8	0.522	0.500	0.353	0.333	0.257	0.048
10	0.495	0.502	0.327	0.338	0.253	0.061
20	0.518	0.488	0.353	0.323	0.252	0.014
50	0.514	0.528	0.347	0.359	0.273	0.019
$W(\lambda) = 0$						
3	0.160	0.157	0.073	0.069	0.026	0.010
5	0.138	0.145	0.057	0.062	0.020	0.010
8	0.144	0.152	0.057	0.063	0.023	0.042
10	0.159	0.145	0.072	0.061	0.022	0.027
20	0.154	0.144	0.065	0.060	0.021	0.019
50	0.151	0.153	0.067	0.065	0.023	0.007

between the fractional molecules of water and methanol is studied. The results are obtained by performing simulations in the CFCNPT ensemble. Coupling parameters λ_1 and λ_2 are assigned to the fractional molecules of water and methanol, respectively. It is clear from Table 3.3 that the coupling parameters for water and methanol are very weakly correlated or essentially uncorrelated. In the simulation of water-methanol with non-zero biasing, the averages $\langle\lambda_1\rangle$ and $\langle\lambda_2\rangle$ are close to 0.5. This is due to the fact that the observed $p(\lambda)$ for water and methanol is flat. The values for $\langle\lambda_1\rangle$ and $\langle\lambda_2\rangle$ are very close to 1 when the weight function $W(\lambda)$ is zero. This is due to the fact that the interactions between the fractional

Table 3.3: Correlations between the fractional molecules of SPC/E water and TraPPE methanol at $T = 323.15$ K and $P = 1$ bar. λ_1 is the fractional molecule of SPC/E water, and λ_2 is the fractional molecule of TraPPE rigid methanol. In simulations with non-zero biasing, an independent biasing is calculated for each fractional molecule. The biasing is set such that the observed probability distribution of every fractional molecule, $p_{\text{obs}}(\lambda)$, is flat.

N_{frac}	$\langle\lambda_1\rangle$	$\langle\lambda_2\rangle$	$\langle\lambda_1^2\rangle$	$\langle\lambda_2^2\rangle$	$\langle\lambda_1\lambda_2\rangle$	$\lvert\text{Corr}\,(\lambda_1,\lambda_2)\rvert$
when $p_{\text{obs}}(\lambda)$ is flat						
2	0.514	0.458	0.350	0.295	0.237	0.020
$W(\lambda) = 0$						
2	0.989	0.980	0.978	0.961	0.969	0.039

molecules and the whole molecules are most favorable when the value of the coupling parameters are close to 1. Figures for $p(\lambda)$ and $W(\lambda)$ for water-methanol simulations are provided in the Supporting Information of Ref. [41]. Since the fractional molecules are not correlated, we can verify that the approximation of Eq. 3.6 is valid.

To investigate the effect of biasing on sampling Boltzmann averages (Eq. 3.7) two LJ color mixtures are considered in which 1 and 5 fractional molecules are present, respectively. The optimum biasing is calculated using the Wang-Landau algorithm at $P^* = 6$ and $T^* = 2$. During the simulations, the instantaneous weight factor $\exp\left[-\sum_{i=1}^{N_{\text{frac}}} W_i\left(\lambda_i\right)\right]$ for both systems is recorded every 100 cycles and the results are shown in Fig. 3.6. The instantaneous weight factor is the statistical weight of a sample system state. It is shown in Fig. 3.6 that the statistical weight for the system including 5 fractional molecules fluctuates mostly between 10^{-9} and 10^{-4}, and quite rarely, between 10^{-4} and 10^{-2}. Multiplying any observable X by such a small number (weight) results in very small numbers, or practically "zero", resulting in poor statistics for $\langle X\rangle_{\text{Boltzmann}}$. The sum of weights in the denominator of Eq. 3.7 is also a very small number close to zero. This is the aforementioned numerical problem of $\frac{0}{0}$ when computing Boltzmann averages using Eq. 3.7. For the system including a single fractional molecule,

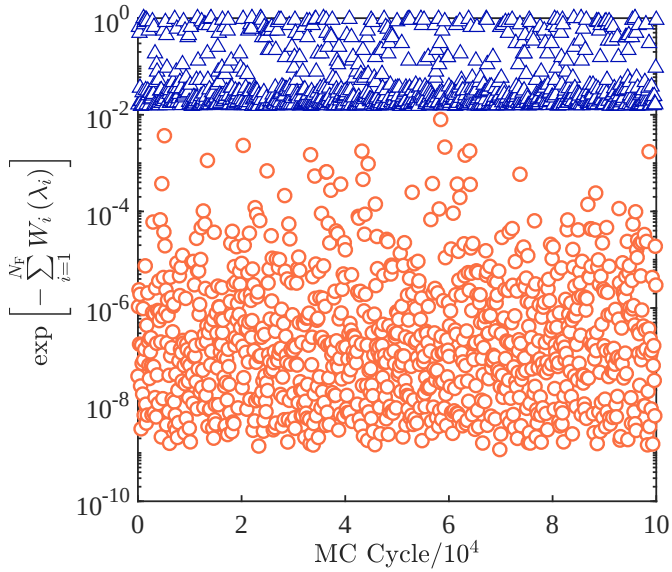


Figure 3.6: Instantaneous weight factor, $\exp\left[-\sum_{i=1}^{N_F} W_i(\lambda_i)\right]$, for a LJ system with 1 fractional molecule (triangles) and 5 fractional molecules (circles), at $T^* = 2$ and $P^* = 6$. $W_i(\lambda_i)$ is set such that $p_{\text{obs}}(\lambda_i)$ for every fractional molecule is flat.

the weight fluctuates between 10^{-2} to 10^{0} during the simulation. Based on Fig. 3.6, it can be concluded that the uncertainty in the Boltzmann average of any observable X increases with the increase in the number of fractional molecules.

One possible solution to circumvent the sampling of Boltzmann averages in simulations with multiple fractional molecules, is to directly sample the averages without removing the biasing (Eq. 3.2). $\langle X \rangle_{\text{biased}}$ is the average of observable X in simulations in the CFCMC ensemble. To compare the statistics of the Boltzmann and biased averages, we have selected the average volume of the system in the CFCNPT simulations. The Boltzmann and biased ensemble averages of volume obtained from the CFCNPT ensemble simulations, with pr_{f} between 0.125% and 6.25%, are calculated for $P^* = 1$

and $P^* = 6$. pr_f is the ratio between the number of the fractional molecules with respect to the number of the whole molecules, expressed as a percentage. The Boltzmann average of volume in the NPT is computed for the same number of whole molecules at the same temperature and pressure as a reference value. The relative uncertainty of the volume of every system is shown in Fig. 3.7 as a function of pr_f. The normalized uncertainty of the volume obtained from the NPT ensemble simulations is shown on the vertical axis. It is shown in Fig. 3.7(a) that at a number density $\langle\rho^*\rangle_{NPT} = 0.43$, the normalized uncertainties of the Boltzmann and biased averages of the volume are very similar for $pr_\mathrm{f} \leq 1.25\%$. Good agreement is observed with the results from the NPT ensemble simulations. However, poor statistics are observed for increasing pr_f. As shown in Fig. 3.7(b), the differences between the averages obtained from Eqs. 3.2 and 3.7 are more pronounced at higher densities ($\langle\rho^*\rangle_{NPT} = 0.80$). It is observed in Fig. 3.7(b) that the sampling of the Boltzmann average of the volume is significantly affected with increasing pr_f, in sharp contrast to the biased averages. The uncertainty of the biased average of volume does not change significantly for increasing pr_f. This is due to the aforementioned $\frac{0}{0}$ sampling problem. Excellent agreement is observed between the relative uncertainties of biased averages of volume and the results obtained from the NPT ensemble. Raw data for Fig. 3.7 are provided in Tables 3.4 and 3.5.

The results in Fig. 3.7 show that the biased average of volume in the CFCNPT ensemble simulations can be statistically more precise compared to the Boltzmann average of the volume. Therefore, it is instructive to investigate the difference between the Boltzmann and biased averages obtained from the CFCNPT simulations, with multiple fractional molecules, and the Boltzmann averages obtained from the conventional NPT ensemble. This may provide guidelines for how many fractional molecules are allowed before the Boltzmann/biased averages significantly deviate from those obtained from the conventional NPT ensemble simulations. Note that by increasing the number of fractional molecules, we are investigating the performance of the CFCMC method in extreme cases. In most practical applications N_frac is usually smaller than five which means that pr_f is significantly smaller than 1% [20, 38, 59, 142, 158, 163, 177, 181, 222, 240]. For this percentage

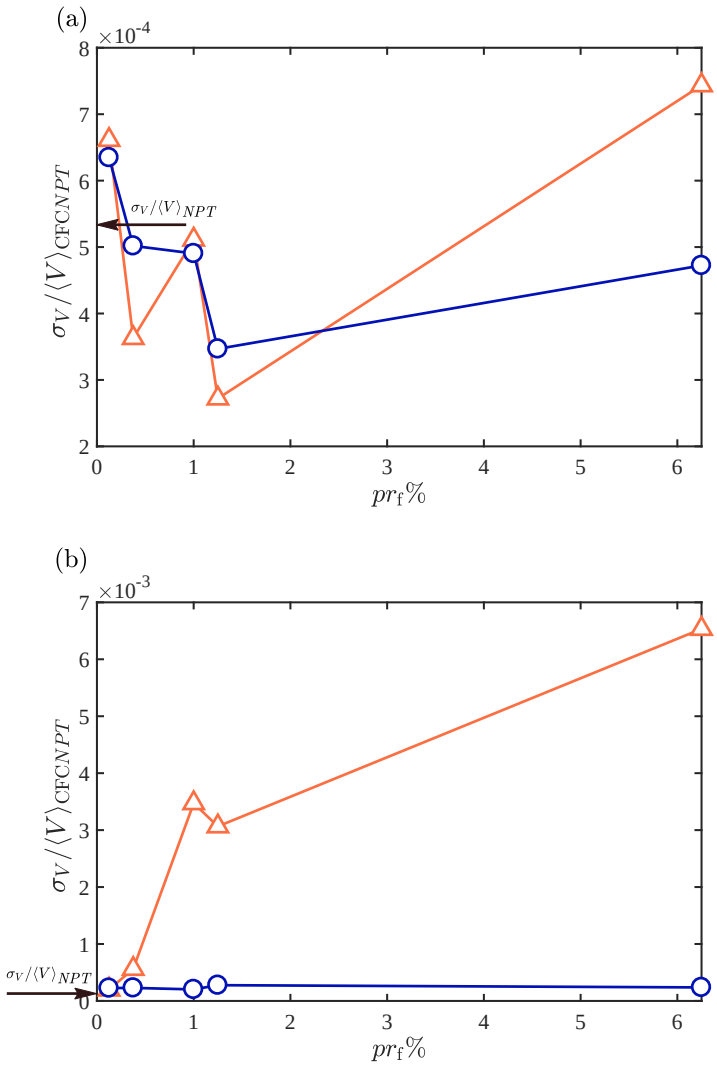


Figure 3.7: Relative uncertainty of the Boltzmann averages of the volume, $\sigma/\langle V\rangle$, (circles) and the biased averages of volume (triangles) in the CFC*NPT* ensemble, at (a) $T^* = 2$, $P^* = 1$ and (b) $T^* = 2$, $P^* = 6$. pr_f is the ratio between the number of the fractional molecules with respect to the number of the whole molecules (constant 800), expressed as a percentage. The relative uncertainty of V is defined as the ratio of the uncertainty of the volume σ_V to the mean volume $\langle V\rangle$. The arrows on the left indicate the value of the relative uncertainty of volume obtained from the *NPT* simulations, on the vertical axes. Raw data are provided in Tables 3.4 and 3.5

.

Table 3.4: Relative difference between the biased averages obtained from the CFC*NPT* simulations and Boltzmann averages obtained from the *NPT* simulations of different LJ color-mixtures, at $T^* = 2$ and reduced pressures between $P^* = 0.5$ to $P^* = 6$. pr_{f} is the ratio between the number of the fractional molecules with respect to the number of the whole molecules (constant 800), expressed as a percentage. The ensemble averages for energy and volume obtained from the *NPT* ensemble simulations, in reduced units, equal $-1130.7(8)$ and $3037(2)$ for $P^* = 0.5$, $-1820(1)$ and $1850(1)$, for $P^* = 1$, $-3127.7(8)$ and $998.6(8)$ for $P^* = 6$, respectively. In the table below, η_X is the difference of the biased average $\langle X\rangle_{\mathrm{CFC}NPT}$ with respect to $\langle X\rangle_{NPT}$, expressed as a percentage. Numbers in brackets are uncertainties in the last digit, *i.e.*, -0.37(3) means -0.37 $\pm$ 0.03.

pr_{f}	$\langle E\rangle_{\mathrm{CFC}NPT}$	$\langle V\rangle_{\mathrm{CFC}NPT}$	η_E	η_ρ	η_V
$P^* = 0.5$					
0.125	-1130(1)	3042(3)	0.07	0.15	0.15
0.375	-1128.9(8)	3051(2)	0.15	0.47	0.47
1	-1126(1)	3075(3)	0.41	1.23	1.25
1.25	-1124.3(8)	3086(2)	0.56	1.58	1.61
6.25	-1103.4(7)	3276(2)	2.41	7.31	7.88
$P^* = 1$					
0.125	-1819(1)	1853(1)	0.06	0.17	0.17
0.375	-1819(1)	1857.4(9)	0.10	0.40	0.40
1	-1816(1)	1869.4(9)	0.24	1.03	1.04
1.25	-1814.1(7)	1874.9(7)	0.35	1.33	1.34
6.25	-1791(1)	1975.5(9)	1.61	6.35	6.78
$P^* = 6$					
0.125	-3126(1)	998.9(2)	0.04	0.03	0.03
0.375	-3127(1)	1001.2(2)	0.02	0.26	0.26
1	-3126(1)	1005.7(2)	0.07	0.71	0.71
1.25	-3125(2)	1007.6(3)	0.10	0.89	0.90
6.25	-3117(1)	1043.4(2)	0.36	4.29	4.49

of fractional molecules, very good estimations for conventional ensemble averages are obtained from CFCMC simulations [20, 38, 59, 142]. For instance, CFCGE simulations of binary or ternary mixtures include at most two or three fractional molecules. For a reactive system of A + B $\rightleftharpoons$ C + D in the liquid phase where component A is volatile, three fractional molecules are required *i.e.* two fractional molecules of reactant molecules (A and B) or reaction products (C and D) and a fractional molecule of the type A in the gas phase. We also investigate how the excess chemical potential calculations are affected when pr_{f} increases. For these systems, the excess chemical potential of a randomly selected fractional molecule, μ_1^{ex} and the average of all the chemical potentials of all fractional molecules, $\langle \mu_i^{\mathrm{ex}} \rangle$ are shown in Table 3.5. As shown in this table, the uncertainty of μ_1^{ex} increases as the number of the fractional molecules in the system increases. This is because the simulation time is divided to perform random walks multidimensional λ space. However, the statistics of the chemical potential averaged over all fractional molecules $\langle \mu_i^{\mathrm{ex}} \rangle$ does not depend strongly on the number of fractional molecules in the color mixture. This is due to the fact that all the intermolecular interactions in the color mixture are similar. Therefore, the chemical potentials of all the LJ molecules in this simulation are equal.

The relative difference for ensemble averages of the energy, density and the volume in the CFC*NPT* simulations are compared to Boltzmann averages obtained from the *NPT* simulations. The results are provided in Tables 3.4 and 3.5. At $p^* = 6$, the relative difference for the Boltzmann average of energy increases significantly with the increase in pr_{f}, in sharp contrast to the error associated with the biased average of energy. This shows once more that the sampling issue of Boltzmann averages in dense systems with multiple fractional molecules is more pronounced (because of larger biasing). It can be seen in Table 3.5 that the Boltzmann averages obtained from the CFC*NPT* ensemble where $pr_{\mathrm{f}} \leq 1\%$ are very similar to those obtained from the *NPT* ensemble. For $pr_{\mathrm{f}} = 1\%$, the relative difference for the Boltzmann averages density and volume in the CFC*NPT* ensemble are about 1% or smaller. We consider 1% as a typical uncertainty from simulations (also differences between experimental data and force

Table 3.5: Relative differences between Boltzmann averages obtained from CFC*NPT* simulations and Boltzmann averages obtained from *NPT* simulations of different LJ color-mixtures, at $T^* = 2$ and reduced pressures between $P^* = 0.5$ to $P^* = 6$. $pr_{\rm f}$ is the ratio between the number of the fractional molecules with respect to the number of the whole molecules (constant 800), expressed as a percentage. The ensemble averages for energy and volume obtained from the *NPT* ensemble simulations, in reduced units, equal $-1130.7(8)$ and $3037(2)$ for $P^* = 0.5$, $-1820(1)$ and $1850(1)$, for $P^* = 1$, $-3127.7(8)$ and $998.6(8)$ for $P^* = 6$, respectively. In the table below, η_X is the difference of the Boltzmann average $\langle X \rangle_{\rm CFC}{}_{NPT}$ with respect to $\langle X \rangle_{NPT}$, expressed as a percentage. For color-mixtures with multiple fractional molecules, $\mu_1^{\rm ex}$ is the excess chemical potential of a randomly selected fractional molecule, in reduced units, and $\langle \mu_i^{\rm ex} \rangle$ is the excess chemical potential averaged over all the fractional molecules. Numbers in brackets are uncertainties in the last digit, *i.e.*, -0.37(3) means -0.37 $\pm$ 0.03.

$pr_{\rm f}$	$\langle E \rangle_{\rm CFC}{}_{NPT}$	$\langle V \rangle_{\rm CFC}{}_{NPT}$	η_E	η_ρ	η_V	$\mu_1^{\rm ex}$	$\langle \mu_i^{\rm ex} \rangle$
$P^* = 0.5$							
0.125	-1130.2(9)	3042(2)	0.04	0.15	0.15	-0.37(3)	-0.37(3)
0.375	-1129.4(6)	3051(2)	0.11	0.45	0.45	-0.40(2)	-0.38(2)
1.00	-1127.1(6)	3074(2)	0.32	1.22	1.23	-0.37(5)	-0.38(3)
1.25	-1125.5(5)	3085(2)	0.46	1.55	1.58	-0.37(5)	-0.37(3)
6.25	-1109.1(9)	3272(2)	1.95	7.18	7.74	-0.4(2)	-0.34(1)
$P^* = 1$							
0.125	-1820(2)	1852(1)	0.03	0.13	0.13	0.07(2)	0.07(2)
0.375	-1818(1)	1857(1)	0.09	0.40	0.40	0.08(2)	0.07(3)
1.00	-1816(1)	1869(1)	0.22	1.04	1.05	0.07(6)	0.07(2)
1.25	-1816(1)	1873(1)	0.23	1.25	1.26	0.0(1)	0.07(2)
6.25	-1800(2)	1970(1)	1.14	6.08	6.48	0.1(1)	0.08(2)
$P^* = 6$							
0.125	-3126.7(9)	998.9(2)	0.03	0.03	0.03	6.39(5)	6.39(5)
0.375	-3125(2)	999(1)	0.07	0.09	0.09	6.41(5)	6.43(3)
1.00	-3113(13)	1002(3)	0.47	0.30	0.30	6.32(6)	6.41(3)
1.25	-3126(13)	1001(3)	0.06	0.24	0.24	6.32(6)	6.38(3)
6.25	-3085(42)	1033(7)	1.38	3.30	3.42	6.4(2)	6.38(2)

field-based simulations are typically also of that order). This applies to normal averages *e.g.* density, volume etc, but not chemical potentials. The chemical potentials computed by CFCMC and without fractional molecule are usually identical [38]. The relative error for the Boltzmann averages density and volume decreases to 0.3% by increasing the pressure to $P^* = 6$. As shown in Table 3.4, good agreement is observed between the biased averages from the CFCNPT simulations and the Boltzmann averages from the NPT simulations for $pr_{\mathrm{f}} \leq 1\%$ (typical differences are around 1%). The relative difference between the biased averages density and volume obtained from the simulations at $P^* = 1$ and $P^* = 6$ are smaller than 1%. For $P^* = 0.5$, relative difference smaller than 1% are obtained for $pr_{\mathrm{f}} \leq 0.375\%$. It can be seen from Tables 3.4 and 3.5 that the errors associated with the biased averages are smaller compared to the Boltzmann averages, especially at high densities. Therefore, it is possible to use biased averages in systems where $pr_{\mathrm{f}} \leq 1\%$. The advantage is that the statistics of biased averages may be better, depending on the system density, compared to Boltzmann averages. In practice, N_{frac} is nearly always below 5 even for studying complex molecular systems [20, 38, 59, 142, 158, 163, 177, 222, 240]. This means that for a system of 500 molecules (a relatively small system size), pr_{f} would nearly always be smaller than 1%.

3.5 Conclusions

An alternative binning scheme is presented to compute the excess chemical potential in CFCMC simulations. This scheme is developed to overcome sampling issues of the excess chemical potential associated with the linear extrapolation to $\lambda \downarrow 0$ and $\lambda \uparrow 1$ used in CFCMC simulations in chapter 2. The drawback of linear extrapolation is that precise values obtained for the excess chemical potential may provide a false impression of accuracy. Increasing the number of bins may improve the accuracy of the extrapolation scheme, however, this leads to poor sampling (larger uncertainty) of $p(\lambda)$ for a fixed simulation time. It is a priori unclear what the optimum number of bins should be for a certain system. In the alternative binning

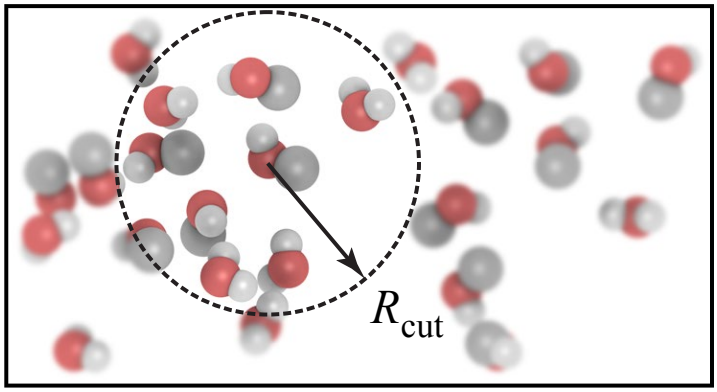
scheme, the first and the last bins are directly used to sample the probability of the interaction parameters $\lambda^* = 0$ (ideal gas behavior) and $\lambda^* = 1$ (fully scaled interactions), respectively. The excess chemical potential is computed by sampling the beginning and end states of λ rigorously (the direct sampling scheme). This method can be implemented in a single step in existing CFCMC codes by performing linear transformation of λ (Eq. 3.3) when calculating the interaction potential of the fractional molecule with the surroundings. In sharp contrast to linear extrapolation, the accuracy and precision of this alternative binning scheme does not strongly depend on the number of the bins. As an example of a system with strong intermolecular interactions, we have computed the excess chemical potential of SPC/E water using both methods. We observed that the excess chemical potential is underestimated for SPC/E water using linear extrapolation to $\lambda \to 0$ and $\lambda \to 1$, since $p(\lambda)$ is steep close to $\lambda = 1$. Generally, this steepness is observed for dense systems or systems with large molecules or with strong intermolecular interactions. We found that the direct sampling scheme is the best method for chemical potential calculations. Very weak or no correlation was found between the fractional molecules in multicomponent systems. This allows one to effectively split a multidimensional weight function into a series of one dimensional weight functions for every fractional molecule. Using this approach, filling a multidimensional histogram of the weight function is avoided, which is computationally not efficient, and a flatness criterion can be applied to each histogram separately. In systems where multiple fractional molecules are present, the weight function is typically large, which leads to the aforementioned $\frac{\text{“0”}}{\text{“0”}}$ numerical problem associated with poor sampling of Boltzmann averages. Our solution is to use biased averages instead of Boltzmann averages. To have similar ensemble averages compared to those obtained from the conventional ensembles, it is recommended that the number of the fractional molecules does not exceed 1% of the total number of molecules. The threshold may be system dependent. In may practical applications, the percentage of fractional molecules is much lower than 1%. To investigate the limits of the CFCMC method, systems with higher percentage of fractional molecules were considered in this work. We have shown that increasing the number

of the fractional molecules does not affect the value/accuracy of the excess chemical potential of each fractional molecule.

Chapter 4

Truncating Long-Range Interactions in Aqueous Methanol Solutions

This chapter is based on the following paper: Rahbari, A.; Hens, R.; Jamali, S. H.; Ramdin, M.; Dubbeldam , D.; Vlugt, T. J. H.; Effect of truncating electrostatic interactions on predicting thermodynamic properties of water-methanol systems, Molecular Simulation, 2019, 45, 336-350, Ref. [222].



4.1 Introduction

In chapters 2 and 3, the CFCMC technique was introduced as an efficient way of solving the problem of low insertion/deletion acceptance probabilities and efficiently computing the chemical potential. In chapter 3, we developed an alternative binning scheme to improve the accuracy of computed chemical potentials. In this chapter, we use the CFCMC method to compute the chemical potentials and activity coefficients for aqueous methanol solutions. Mixing enthalpies for aqueous methanol solutions are also calculated from simulations in the conventional NPT ensemble. We have chosen aqueous methanol solutions since water and methanol are very common compounds in the chemical industry. Aqueous mixtures of methanol are investigated frequently in academia [27, 85–103] and are of practical importance in industrial applications [104–112]. Depending on the application, different mixture properties are required for process design. To compute the mixture properties at different conditions, different force field combinations of water and methanol are considered in few molecular simulation studies [27, 79, 85, 86, 88–91, 99, 103]. To the best of our knowledge, chemical potentials and activity coefficients of water and methanol are not reported in molecular simulation studies for different/recent force field combinations of water and methanol. Because of the low vapor pressures of water and methanol at ambient conditions, it is also experimentally challenging to determine the activity coefficients. Water is a flexible and polarizable molecule [241–243]. This means that the electronic structure of the water molecule undergoes deformation due to the electric field induced by the surrounding water or other polar molecules [244]. To account for polarization effects in the recent years, polarizable force fields for water have been developed [241, 243–251]. Some properties of water, such as vapor pressure, critical properties, dielectric constant, and virial coefficient are most accurately predicated by considering polarization effects [243]. However, the performance of different polarizable force fields to compute chemical potentials and activity coefficients of water is not fully investigated in literature. Although the water molecule is flexible and polarizable [241], including polarization effects in MC simulations significantly increases the computational costs [252–254].

In many studies, water is considered as an explicit solvent/medium and is not the main focus. Therefore, the majority of water models in literature are rigid with point charges, and polarization is ignored [76–78, 241]. These simplified models for water are computationally advantageous and the reproduced bulk properties of water are usually in good agreement with experiments at ambient conditions [76, 241]. Since water models are usually fitted to a limited set of experimental data, no water model can simultaneously reproduce all thermophysical properties in good agreement with experiments [76–78]. Therefore, reproducing thermophysical properties of water depends strongly on the choice of experimental data used to fit the intermolecular interaction parameters [255]. Deviations from experimental data may also arise because of inherent limitations in the molecule models of water [77, 78, 256, 257]. The most popular classical force fields of water include three-point site interaction models [235, 258, 259], four-point interaction site models [77], five-point interaction site models [205, 260]. In this chapter, we focus on popular rigid non-polarizable models of water including: TIP3P [259], SPC/E [235], OPC [256], TIP4P/2005 [255], and TIP4P/EW [261]. These models are frequently used in benchmark studies [76–78, 256]. For methanol, we have chosen the TraPPE [207] and OPLS/2016 [262] force fields. The TraPPE force field is widely used, and the OPLS/2016 force field is recently published. The OPLS/2016 force field it is claimed to be a more accurate force field for methanol [262]. The force fields for methanol are rigid and non-polarizable. These models are different in bond geometries and/or point charge distributions, and predicting different properties of the mixture depends on different combinations of these force fields [86, 87]. In this chapter, we compare the performance of the spherical cutoff methods in MC and MD simulations to the Ewald summation [146]. The advantage of spherical cutoff methods compared to the Ewald summation is the reduced simulation time. However, we show that numerical artifacts may appear (in the molecular structure of the system around the cutoff) due to truncating long-range electrostatic interactions. For a system of rigid molecules with point charges, the potential energy is the sum of pairwise interaction potentials consisting of LJ and Coulombic

interactions:

$$E = \frac{1}{2} \sum_{i} \sum_{j \neq i} \left[4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \right] \tag{4.1}$$

where the double summation is over all the interaction sites. r_{ij} is the distance between atoms i and j, ϵ_0 is the dielectric constant, q_i is the partial charge of atom i, σ_{ij} and ϵ_{ij} are the LJ parameters between atoms i and j, obtained from σ_i, σ_j, ϵ_i and ϵ_j. Proper treatment of LJ and Coulombic interactions are essential for an accurate description of molecular structure, and thermodynamic properties of water and methanol [14, 15, 146, 147, 149, 163, 236, 237, 263, 264]. In molecular simulation, periodic boundary conditions are applied to simulate bulk phases and compute structural and thermodynamic properties [14]. In systems with periodic boundary conditions, the long-range electrostatic interactions decay slowly with r^{-1}, and the treatment of the electrostatic interactions becomes computationally demanding. Direct summation of Coulombic interactions is conditionally convergent which means that the results depend on the order of summation [15, 265]. Various molecular simulation studies investigated the accuracy and scalability of different methods for treating long-range Coulombic interactions [147, 266, 267]. These methods can be divided into Ewald-based and cutoff-based methods. In this paper, we investigate to what extent different electrostatic methods influence thermophysical properties of water-methanol mixtures, and whether numerical artifacts are observed. First, a brief description on electrostatic methods is provided. The Ewald summation [14, 15, 146–148] is the most widely used and accepted method to compute electrostatic interactions in molecular simulation. In the Ewald summation, the electrostatic interactions are split into effective short-range interactions, evaluated in real space, and a Fourier series to account for the long-range contribution of the electrostatic interactions, evaluated in reciprocal space [14, 15, 146, 264]. The high computational costs of the Ewald summation have lead to efforts to develop faster alternatives [163, 236, 238, 264, 265, 267]. In the past decades, more Ewald-based algorithms have been developed to reduce the time complexity to $\mathcal{O}(N \log N)$ by optimizing the reciprocal space sum-

mation [268, 269]. Particle mesh algorithms such as Staggered Mesh Ewald method (StEM) [268], Particle-Particle Particle-Mesh (PPPM) [270], and Particle-Mesh Ewald (PME) [269] are based on the Ewald method and scale as $\mathcal{O}(N \log N)$ [147, 271]. These algorithms allow efficient parallel implementations, and are especially advantageous to use in MD simulations [147]. When the Ewald summation is used in Monte Carlo (MC) simulations, one only need consider the charged atoms that are changed in a trial move. This can be achieved by storing the Fourier part in memory in an efficient way [265, 272]. For a faster computation of the electrostatic interactions, truncation or spherical cutoff-based methods are proposed as an alternative to the Ewald-type methods [149, 236, 267, 273–275]. Cutoff-based methods are based on the idea that the effective electrostatic potential of condensed phases has rather short-ranged behavior [149, 163, 236, 238, 265, 267, 276, 277], and the effect of the long range interactions beyond the cutoff radius becomes negligible due to the screening of charges [147, 149, 238, 274, 278]. In the gas phase, the screening of charges is weak compared to the liquid phase [149, 163, 236, 238]. Compared to the Ewald-type methods, the cutoff-based methods are much simpler to implement and scale as $(\mathcal{O}(N))$. The performance of the cutoff-based methods are promising especially for bulk homogeneous systems, while in systems with an interface, severe numerical problems may be expected [264]. A common criticism about cutoff-based methods is that numerical artifacts may arise due to the truncation of long-range interactions [147, 264, 266, 279–281]. These numerical artifacts can affect the structure of the liquid. For example, two atoms of opposite charges (positive-negative) prefer to be located within the cutoff radius, while two atoms with the same charge (positive-positive or negative-negative) are preferentially located outside the cutoff radius. As the concept of using a cutoff radius for electrostatic interactions is due to simulation efficiency reasons rather than a physical effect, changes in the liquid structure due to an imposed cutoff radius are nonphysical and should be avoided. To observe whether anomalies occur in the liquid structure, the Radial Distribution Function (RDF) of the liquid can be studied. Depending on the extent of the numerical artifacts, energy calculations are affected. Mark et al. [278] reported artificial structuring of water when using atom-

based cutoff methods developed by Steinbach et al. [274] to compute the electrostatic interactions. The slow convergence issue appeared especially in truncation and shifting schemes in which the truncation sphere inside the cutoff was not electroneutral [147, 236, 237]. Wolf et al. realized that the error in computing the intramolecular electrostatic interactions was related to the net charge inside the cutoff sphere and enforced charge neutrality within the cutoff radius [236, 264]. By using a charge-neutralized damped pair potential, fast convergence is achieved and effects of possible artifacts due to the truncation of electrostatic interactions are minimized [149, 236–238, 264]. Using the so-called Wolf method, the damped shifted pairwise electrostatic potential for a system of N charges is obtained from [163, 236, 238, 265]

$$E^{\mathrm{Wolf}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i \\ r_{ij}<R_{\mathrm{c}}}}^{N} q_i q_j \left[\frac{\operatorname{erfc}(\alpha r_{ij})}{r_{ij}} - \frac{\operatorname{erfc}(\alpha R_{\mathrm{c}})}{R_{\mathrm{c}}}\right] - \left[\frac{\operatorname{erfc}(\alpha R_{\mathrm{c}})}{2R_{\mathrm{c}}} + \frac{\alpha}{\sqrt{\pi}}\right]\sum_{i=1}^{N} q_i^2 \tag{4.2}$$

where r_{ij} is the distance between the partial charges i and j, $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ is the complementary error function. In principle, other damping functions may be used [236, 263]. R_{c} is the cutoff radius, α is the Wolf damping parameter, and its value determines how fast the complementary error function approaches zero with increasing r_{ij}. The second term on the right hand side of Eq. 4.2 is the energy associated with the so-called self term [236, 238]. The Wolf method has been used in different studies including water and ionic systems [267, 282–284]. For a nice and readable derivation of the Wolf method for molecular systems, the reader is referred to the paper by Waibel and Gross [238]. Applying the Wolf method to a molecular system, the charge-charge interactions within each molecule should be excluded from the total sum of the electrostatic energy [238] because these are not part of intermolecular interactions. Initially, we

consider a system without intramolecular Coulombic interactions, *e.g.* for rigid molecules. The total electrostatic energy equals [142, 163, 265]:

$$
\begin{aligned}
E^{\mathrm{Wolf}} = & \frac{1}{2}\sum_{i=1}^{N_\mathrm{m}}\sum_{a=1}^{N_\mathrm{a}^i}\sum_{\substack{j=1\\ j\neq i}}^{N_\mathrm{m}}\sum_{b=1}^{N_\mathrm{a}^j}\underset{r_{iajb}<R_\mathrm{c}}{q_{ia}q_{jb}}\left[\frac{\mathrm{erfc}\left(\alpha r_{iajb}\right)}{r_{iajb}}-\frac{\mathrm{erfc}\left(\alpha R_\mathrm{c}\right)}{R_\mathrm{c}}\right]+ \\
& \frac{1}{2}\sum_{i=1}^{N_\mathrm{m}}\sum_{a=1}^{N_\mathrm{a}^i}\sum_{\substack{b=1\\ b\neq a}}^{N_\mathrm{a}^i}\underset{r_{iaib}<R_\mathrm{c}}{q_{ia}q_{ib}}\left[\frac{\mathrm{erfc}\left(\alpha r_{iaib}\right)}{r_{iaib}}-\frac{\mathrm{erfc}\left(\alpha R_\mathrm{c}\right)}{R_\mathrm{c}}\right]- \\
& \frac{1}{2}\sum_{i=1}^{N_\mathrm{m}}\sum_{a=1}^{N_\mathrm{a}^i}\sum_{\substack{b=1\\ b\neq a}}^{N_\mathrm{a}^i}\frac{q_{ia}q_{ib}}{r_{iaib}}- \\
& \left[\frac{\mathrm{erfc}\left(\alpha R_\mathrm{c}\right)}{2R_\mathrm{c}}+\frac{\alpha}{\sqrt{\pi}}\right]\sum_{i=1}^{N_\mathrm{m}}\sum_{a=1}^{N_\mathrm{a}^i}q_{ia}^2
\end{aligned}
\tag{4.3}
$$

In Eq. 4.3, N_m is the number of molecules, N_a^i is the number of atoms in molecule i, indices i and j are used to count the number of molecules, indices a and b are used to count atoms within molecules. q_{ia} is the partial charge of atom a in molecule i, r_{iajb} is the distance between interaction sites a and b. The first and second terms on the right hand side of Eq. 4.3 are used to compute the pairwise electrostatic interactions between all partial charges in the system, which are screened electrostatic interactions between different molecules and electrostatic interactions inside molecules, respectively. For rigid molecules, we are only interested in intermolecular interactions, and therefore we have to exclude electrostatic interactions within molecules, i.e. charge-charge interactions between the atoms within the molecules. This is achieved by the third term on the right hand side of Eq. 4.3, the so-called exclusion term, similar to the exclusion term in the Ewald summation [285]. The fourth term on the right hand side of Eq. 4.3 is the self-interaction term which is similar to that of an atomic system Eq. 4.2. For non-rigid molecules,

the intramolecular Coulombic interactions should be added to Eq. 4.3 by:

$$E_{\text{Intra}} = \frac{1}{2} \sum_{i=1}^{N_{\text{m}}} \sum_{a=1}^{N_{\text{a}}^i} \sum_{\substack{b=1 \\ b \neq a}}^{N_{\text{a}}^i} c_{ab} \frac{q_{ia} q_{ib}}{r_{iaib}} \tag{4.4}$$

in which c_{ab} is a scaling parameter for electrostatic interactions between atom a and atom b. c_{ab} equals zero if atoms a and b do not have any intramolecular electrostatic interactions. Several force fields have intramolecular interactions that are scaled [286–290], so the value of c_{ab} can be non-integer. Note that a cutoff radius is not applied for the intramolecular interactions of Eq. 4.4, and that these interactions are not screened with a $\frac{\text{erfc}(\alpha r)}{r}$ term. The reason is that for a system consisting of an isolated molecule with intramolecular interactions but no intermolecular interactions, the correct result is obtained, i.e. a direct pairwise $\frac{1}{r}$ summation over the intramolecular Coulombic interactions. It is important to note that in sharp contrast to the terms for intramolecular Coulombic interactions and intramolecular Coulombic exclusions proposed by Gross et al. [238], in this work no cutoff is imposed for these interactions. It is expected that this avoids potential artifacts for systems including molecules larger than the cutoff radius.

Fennell and Gezelter [149] found that the damped shifted potential proposed by Wolf et al, Eq. 4.3, results in force discontinuity at the cutoff radius in MD simulations. This is undesirable as it may lead to energy

drifts [14]. To remove this discontinuity, the electrostatic potential proposed by Fennell and Gezelter (the so-called DSF method) is calculated using

$$
\begin{aligned}
E^{\mathrm{DSF}} = & \frac{1}{2} \sum_{i=1}^{N_\mathrm{m}} \sum_{a=1}^{N_\mathrm{a}^i} \sum_{\substack{j=1 \\ j \neq i}}^{N_\mathrm{m}} \sum_{b=1}^{N_\mathrm{a}^j} q_{ia} q_{jb} \Bigg[\frac{\operatorname{erfc}(\alpha r_{iajb})}{r_{iajb}} - \frac{\operatorname{erfc}(\alpha R_\mathrm{c})}{R_\mathrm{c}} \quad + \\
& \qquad r_{iajb} < R_\mathrm{c} \\
& \left(\frac{\operatorname{erfc}(\alpha R_\mathrm{c})}{R_\mathrm{c}^2} + \frac{2\alpha}{\sqrt{\pi}} \frac{\exp\left(-\alpha^2 R_\mathrm{c}^2\right)}{R_\mathrm{c}} \right) (r_{iajb} - R_\mathrm{c}) \Bigg] + \\
& \frac{1}{2} \sum_{i=1}^{N_\mathrm{m}} \sum_{a=1}^{N_\mathrm{a}^i} \sum_{\substack{b=1 \\ b \neq a}}^{N_\mathrm{a}^i} q_{ia} q_{ib} \left[\frac{\operatorname{erfc}(\alpha r_{iaib})}{r_{iaib}} - \frac{\operatorname{erfc}(\alpha R_\mathrm{c})}{R_\mathrm{c}} \right] - \qquad (4.5) \\
& \qquad r_{iajb} < R_\mathrm{c} \\
& \frac{1}{2} \sum_{i=1}^{N_\mathrm{m}} \sum_{a=1}^{N_\mathrm{a}^i} \sum_{\substack{b=1 \\ b \neq a}}^{N_\mathrm{a}^i} \frac{q_{ia} q_{ib}}{r_{iaib}} - \\
& \left[\frac{\operatorname{erfc}(\alpha R_\mathrm{c})}{2R_\mathrm{c}} + \frac{\alpha}{\sqrt{\pi}} \right] \sum_{i=1}^{N_\mathrm{m}} \sum_{a=1}^{N_\mathrm{a}^i} q_{ia}^2
\end{aligned}
$$

Similar to Eq. 4.3, the intramolecular Coulombic interactions are not included here, and these interactions need to be taken into account according to Eq. 4.4. The shift in Eq. 4.5 (first summation on the right hand side) is only applied to intermolecular interactions but not to intramolecular interactions. For simplicity, Eqs. 4.3 to 4.5 are formulated for single-component systems. Extending these equations to multicomponent systems is trivial: *(1)* The summation in the first terms of Eqs. 4.3 and 4.5 should be over all intermolecular interactions between all the atoms of all molecules; *(2)* the summations in other terms of Eqs. 4.3 and 4.5 (and the summation in Eq. 4.4) should be over all molecules. The extra term on the right hand side of Eq. 4.5 compared to Eq. 4.3 makes both the potential and its

first derivative (the force) continuous at the cutoff radius [149]. Although the continuity of the force at the cutoff radius is of primary interest in MD simulations, the electrostatic potential derived in Eq. 4.5 is slightly different compared to the original Wolf method. This small difference in the electrostatic interaction potential can potentially change the computed structure and/or other properties of the system. In different studies of ionic liquids [149, 240, 269, 291, 292], it was found that the electrostatic energies and forces obtained using the DSF method are in excellent agreement with the conventional Ewald/smooth PME method. Instead of using an atom-based spherical cutoff method, one could also apply a group-based or charge-group cutoff method [147, 149, 274]. In this method, atoms are assigned to charge-neutral groups. If the distance between the geometric centers of the charge-neutral groups is smaller than a certain cutoff, all atoms belonging to the charge-neutral groups interact, otherwise no atomic interaction between the two groups is taken into account. Therefore the computational time is reduced by ignoring the pairwise electrostatic interactions between charge-neutral groups which are further away from the cutoff radius. Another advantage is that the leading $\frac{1}{r}$ term for Coulombic interactions reduces to a higher order term that decays faster [274]. Gross et al. [238] found no differences between a group-based cutoff radius and an atom-based cutoff radius for the Wolf method, and therefore we have not considered this further in this paper.

This chapter is organized as follows. In Section 4.2, expressions to compute the following thermodynamic properties: excess mixing enthalpy, chemical potentials and activity coefficients are described. Simulation details, force field parameters, and scaling of the intramolecular interactions are described in Section 4.3. Simulation results are presented in Section 4.4. It is shown that the liquid structure obtained using the Wolf method, Eq. 4.3, and the DSF method, Eq. 4.5, is different especially near the cutoff radius. However, computing the excess chemical potentials of water and methanol obtained using the Wolf method and the DSF method yield very similar results. It is also shown that the activity coefficients of water and methanol in water-methanol mixtures computed based on TIP4P/2005 and TraPPE

force fields show the best agreement with experimental data. Our conclusions are summarized in Section 4.5.

4.2 Theory

The enthalpy of a system (pure component or a mixture) can be directly computed from simulations in the NPT ensemble [14, 15]

$$H = \langle U \rangle_{NPT} + P \langle V \rangle_{NPT} \tag{4.6}$$

$\langle U \rangle_{NPT}$ is the ensemble average internal energy of the system, and $\langle V \rangle_{NPT}$ is the ensemble average volume in the NPT ensemble. For the water-methanol mixture, the excess enthalpy of the mixing is obtained from

$$h^{\text{ex}}_{\text{mix}} = h_{\text{mix}} - (1 - x_{\text{MeOH}}) h_{\text{H}_2\text{O}} - x_{\text{MeOH}} h_{\text{MeOH}} \tag{4.7}$$

h^{ex} is the excess enthalpy of mixing with respect to pure liquid, h_{mix} is the enthalpy of the water-methanol mixture, h_{MeOH} is the enthalpy of pure methanol, $h_{\text{H}_2\text{O}}$ is the enthalpy of pure water, and x_{MeOH} is the mole fraction of methanol. The Continuous Fractional Component NPT (CFCNPT) ensemble is used [20] to compute the chemical potentials of water and methanol (see Eq. 2.9). The expression for the partition function of this expanded ensemble is provided in Eq. 3.5. Expressions for the chemical potential and derivatives of the chemical potential in the CFCNPT ensemble are provided in appendix A.1. For improved accuracy of the computed chemical potentials, the reader is referred to chapter 3. In appendix A.10, it is shown that the activity coefficient of component A, γ_{A}, in a mixture is obtained according to [293, 294]

$$\gamma_{\text{A}} = \frac{\rho_{\text{A}}}{x_{\text{A}} \rho_{0\text{A}}} \cdot \exp\left[\beta \left(\mu^{\text{ex}}_{\text{A}} - \mu^{\text{ex}}_{0\text{A}}\right)\right] \tag{4.8}$$

ρ_{A} and $\rho_{0\text{A}}$ are the number densities of component A in the mixture and the reference number density of the pure solvent, respectively. $\mu^{\text{ex}}_{\text{A}}$ is the excess chemical potential of component A in a mixture with mole fraction

of x_A, and μ_{0A}^{ex} is the excess chemical potential of A in the pure fluid A. The excess chemical potentials in Eq. 4.8 are referenced with respect to the ideal gas. The derivation of Eq. 4.8 is provided in appendix A.10.

4.3 Simulation Details

4.3.1 Monte Carlo Simulations

Different water-methanol mixtures with compositions ranging from $x_{MeOH} = 0$ to $x_{MeOH} = 1$ are simulated at $P = 1$ bar and $T = 298$ K, both in the conventional NPT ensemble [14, 15] and the CFCNPT ensemble [20]. All MC simulations were performed using our in-house code which is verified to produce the same results as the RASPA software package [219, 220] in various works [142, 163]. For water, the TIP3P [258], SPC/E [235], OPC [256], TIP4P/2005 [255], and TIP4P/EW [261] force fields, and for methanol the TraPPE [207] and OPLS/2016 [262] force fields are used. All force field parameters are provided in Tables 4.1 and 4.2. For every mixture composition and every water-methanol force field combination, the Wolf and the DSF methods, Eqs. 4.3 and 4.5, were both used to compute the electrostatic interactions. To obtain the parameters for the Wolf method for pure water and pure methanol, independent simulations were performed in the NVT ensemble, using SPC/E [235] and OPLS/2016 [262] force fields, close to the experimental densities at $T = 298$ K [3, 295]. For dense liquids such as water and methanol at ambient conditions, it is sufficient to plot Fig. 4.1 for a single configuration [163]. For the single equilibrated configuration, the electrostatic energies were calculated for different values of cutoff radii ranging from $R_c = 10$ Å to $R_c = 15$ Å, as a function of α. The relative difference in electrostatic energies, for water and methanol, were compared to the results obtained from the Ewald summation, and the results are shown in Fig. 4.1. It is shown in Fig. 4.1 that the relative difference between the electrostatic energies between these methods is within 0.5% for the cutoff radii ranging from 10 Å to 15 Å, and α ranging from 0.1 Å^{-1} to 0.15 Å^{-1}. This means

Table 4.1: Force field parameters for water used in this study. All molecules are considered rigid. ϵ is reported in units of K, σ and r are reported in units of Å.

Force field	TIP3P [258]	SPCE/E [235]	OPC [256]	TIP4P/2005 [255]	TIP4P/EW [261]
ϵ_{OO}	76.500	78.175	107.086	93.196	81.899
σ_{OO}	3.1506	3.1660	3.1666	3.1589	3.1644
q_{O}	-0.8340	-0.8476	-1.3582	-	-
q_{H}	0.41700	0.42380	0.67910	0.55640	0.52422
q_{L}	-	-	-	-1.11280	-1.04844
r_{OH}	0.9572	1.0000	0.8724	0.9572	0.9572
r_{OL}	-	-	0.1594	0.1546	0.1250

Table 4.2: Force field parameters for methanol used in this study. All molecules are considered rigid. ϵ is reported in units of K, σ and r are reported in units of Å.

Force field	OPLS/2016 [262]	TraPPE [207]
ϵ_{OO}	97.775	93.000
σ_{OO}	3.1659	3.0200
$\epsilon_{\mathrm{CH_3CH_3}}$	110.450	98.000
$\sigma_{\mathrm{CH_3CH_3}}$	3.6449	3.75
q_{O}	-0.6544	-0.70000
q_{H}	0.49980	0.43500
$q_{\mathrm{CH_3}}$	0.1546	0.2650
r_{OH}	0.9450	0.9450
$r_{\mathrm{CH_3O}}$	1.43	1.43

that the results obtained from the Wolf method in this (α,R_{c}) range, are consistent with the energetics from the Ewald summation. For all simulations, Wolf parameters, R_{c} and α were set to 14 Å and 0.12 $\mathrm{Å}^{-1}$, respectively, and a cutoff radius of 14 Å was used for LJ interactions. The same parameters were used for the DSF method Eq. 4.5. All molecules are rigid and the interactions between the molecules only consist of LJ and Coulombic interactions. Periodic boundary conditions were used. LJ potentials were truncated but not shifted, and analytic tail corrections and the Lorentz-Berthelot mixing rules were applied [14, 15]. Simulations in the NPT ensemble were performed to compute the excess mixing enthalpies

of water-methanol mixtures (with respect to pure liquid) based on Eq. 4.7. Simulations in the CFCNPT ensemble were performed to compute the excess chemical potentials of water and methanol (with respect to ideal gas phase) using Eq. 2.9. To improve the accuracy of the computed chemical potentials, it is recommended to use the binning scheme developed in chapter 3. The activity coefficients of water and methanol were obtained using Eq. 4.8. Each simulation in the NPT ensemble was carried out with 10^5 equilibration cycles and 4×10^6 production cycles. In each cycle, the number of MC steps equals the total number of molecules, and trial moves were selected with the following probabilities: 1% volume changes, 49.5% translations, and 49.5% rotations. Simulations in the CFCNPT ensemble were performed to compute the chemical potentials and activity coefficients of water and methanol at different mixture compositions. Each simulation in the CFCNPT ensemble was carried out with 10^5 equilibration cycles and 4×10^6 production cycles. To facilitate the sampling of λ, a weight function ($W(\lambda)$) was used to make the sampled probability of λ flat [45, 141]. During equilibration, the WL algorithm was used to construct the weight function [199, 200]. In each MC step, trial moves were selected with the following probabilities: 1% volume changes, 35% translations, 30% rotations, 17% λ changes, 8.5% reinsertions, and 8.5% identity changes. Details on specific trial moves associated with the fractional molecules are provided in Section 3.3.

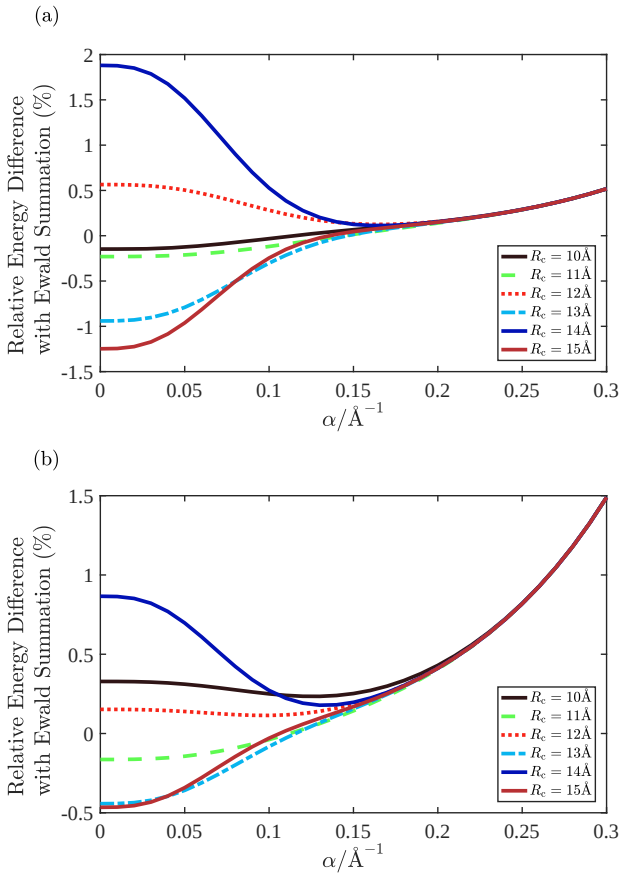


Figure 4.1: Relative differences in computed electrostatic energies between the Wolf method, Eq. 4.3, and the Ewald summation for (a) water and (b) methanol. The parameters for the Ewald summation are calculated based on relative precision of 10^{-6} [163]. The SPC/E [235] and OPLS/2016 [262] force fields were used to obtain the densities of water and methanol at $T = 298$ K and $P = 1$ bar. Individual configurations were obtained at constant densities of 1000 kg.m^{-3} and 748 kg.m^{-3} for water and methanol, respectively.

Scaling of the Lennard-Jones and Electrostatic Interactions

In CFC*NPT* simulations, the LJ interactions of the fractional molecule are scaled using Eq. 2.10. The Coulombic interactions for the DSF potential are scaled as

$$
\begin{aligned}
&E_{\mathrm{Coul}}^{\mathrm{DSF}}\left(r,\lambda_{\mathrm{Coul}}\right)= \\
&\frac{1}{2}\sum_{i=1}^{N_{\mathrm{m}}}\sum_{a=1}^{N_{\mathrm{a}}^{i}}\sum_{\substack{j=1\\ j\neq i}}^{N_{\mathrm{m}}}\sum_{b=1}^{N_{\mathrm{a}}^{j}}\lambda_{\mathrm{Coul}}q_{ia}q_{jb}\left[\frac{\operatorname{erfc}\left(\alpha\left[r_{iajb}+r_{*}\right]\right)}{r_{iajb}+r_{*}}-\frac{\operatorname{erfc}\left(\alpha\left[R_{\mathrm{c}}+r_{*}\right]\right)}{R_{\mathrm{c}}+r_{*}}\right. \quad + \\
&\qquad_{r_{iajb}<R_{\mathrm{c}}} \\
&\left.\left(\frac{\operatorname{erfc}\left(\alpha\left[R_{\mathrm{c}}+r_{*}\right]\right)}{\left(R_{\mathrm{c}}+r_{*}\right)^{2}}+\frac{2\alpha}{\sqrt{\pi}}\frac{\exp\left(-\alpha^{2}\left[R_{\mathrm{c}}+r_{*}\right]^{2}\right)}{R_{\mathrm{c}}+r_{*}}\right)\left(r_{iajb}-R_{\mathrm{c}}\right)\right]+ \\
&\frac{1}{2}\sum_{i=1}^{N_{\mathrm{m}}}\sum_{a=1}^{N_{\mathrm{a}}^{i}}\sum_{\substack{b=1\\ b\neq a}}^{N_{\mathrm{a}}^{i}}\lambda_{\mathrm{Coul}}q_{ia}q_{ib}\left[\frac{\operatorname{erfc}\left(\alpha\left[r_{iaib}+r_{*}\right]\right)}{r_{iaib}+r_{*}}-\frac{\operatorname{erfc}\left(\alpha\left[R_{\mathrm{c}}+r_{*}\right]\right)}{R_{\mathrm{c}}+r_{*}}\right]- \\
&\qquad_{r_{iaib}<R_{\mathrm{c}}} \\
&\frac{1}{2}\sum_{i=1}^{N_{\mathrm{m}}}\sum_{a=1}^{N_{\mathrm{a}}^{i}}\sum_{\substack{b=1\\ b\neq a}}^{N_{\mathrm{a}}^{i}}\frac{\lambda_{\mathrm{Coul}}q_{ia}q_{ib}}{r_{iaib}+r_{*}}-\left[\frac{\operatorname{erfc}\left(\alpha R_{\mathrm{c}}\right)}{2R_{\mathrm{c}}}+\frac{\alpha}{\sqrt{\pi}}\right]\sum_{i=1}^{N_{\mathrm{m}}}\sum_{a=1}^{N_{\mathrm{a}}^{i}}\lambda_{\mathrm{Coul}}q_{ia}^{2}
\end{aligned}
\tag{4.9}
$$

where $r_{*}=A\left(1-\lambda_{\mathrm{Coul}}\right)^{2}$, and $A=\frac{1}{2}$Å. λ_{Coul} is the coupling parameter for Coulombic interactions as shown in Fig. 2.3. By ignoring the second line in Eq. 4.9, a similar expression is obtained for the scaled Coulombic interaction of the fractional molecule using the Wolf method (Eq. 4.3).

4.3.2 Molecular Dynamics Simulations

All MD simulations were performed with the LAMMPS software package [296]. Periodic boundary conditions were used. All molecules were kept rigid using the SHAKE algorithm [15]. LJ potentials were truncated and analytic tail corrections were applied to compute the energy and pressure of the system [15]. The Lorentz-Berthelot mixing rules were used for

non-bonded LJ interactions [14, 15]. To compute densities and enthalpies of water-methanol mixtures, MD simulations are performed in the NPT ensemble. The Nosé-Hoover thermostat and barostat are used in all MD simulations performed in this work [15]. To calculate the RDFs, the ensemble average densities obtained from NPT simulations were used to fix the densities in the NVT ensemble simulations. The temperature of the system is regulated by using the Nosé-Hoover thermostat [15]. The length of each simulation for computing thermodynamic properties, i.e. densities and enthalpies, and RDFs in the NPT and NVT ensembles are 5 ns and 10 ns, respectively. A time step of 1 fs is used to integrate the equations of motion. The specifications of the force fields used in MD simulations are the same as the specifications used in MC simulations. For the Ewald summation method, long-range electrostatic interactions are computed with a relative precision of 10^{-6} [15].

4.4 Results

4.4.1 Electrostatics

The simplest way to characterize the structure of the liquid phase is by calculating its RDF [4, 14, 15]. To investigate how different methods for handling the electrostatics may influence the structure of the liquid, a binary mixture of water-methanol (50%-50%) was considered as a representative case. The RDF obtained using the Ewald summation, from MD simulations, was considered as a reference case. For this comparison, the TIP4P/2005 [255] and TraPPE [207] force fields were considered. The densities of the binary water-methanol mixtures were obtained from independent MC and MD simulations, based on different treatments of the long-range electrostatic interactions. The relative difference in densities obtained from MD and MC simulations was below 0.2%, see the Supporting Information of Ref. [222]. In Fig. 4.2, three different RDFs are shown for water-water, water-methanol, and methanol-methanol, in a 50%-50% water-methanol mixture. The oxygen atoms in water and methanol represent the molecules. Excellent agreement between the RDFs in Fig. 4.2 shows that

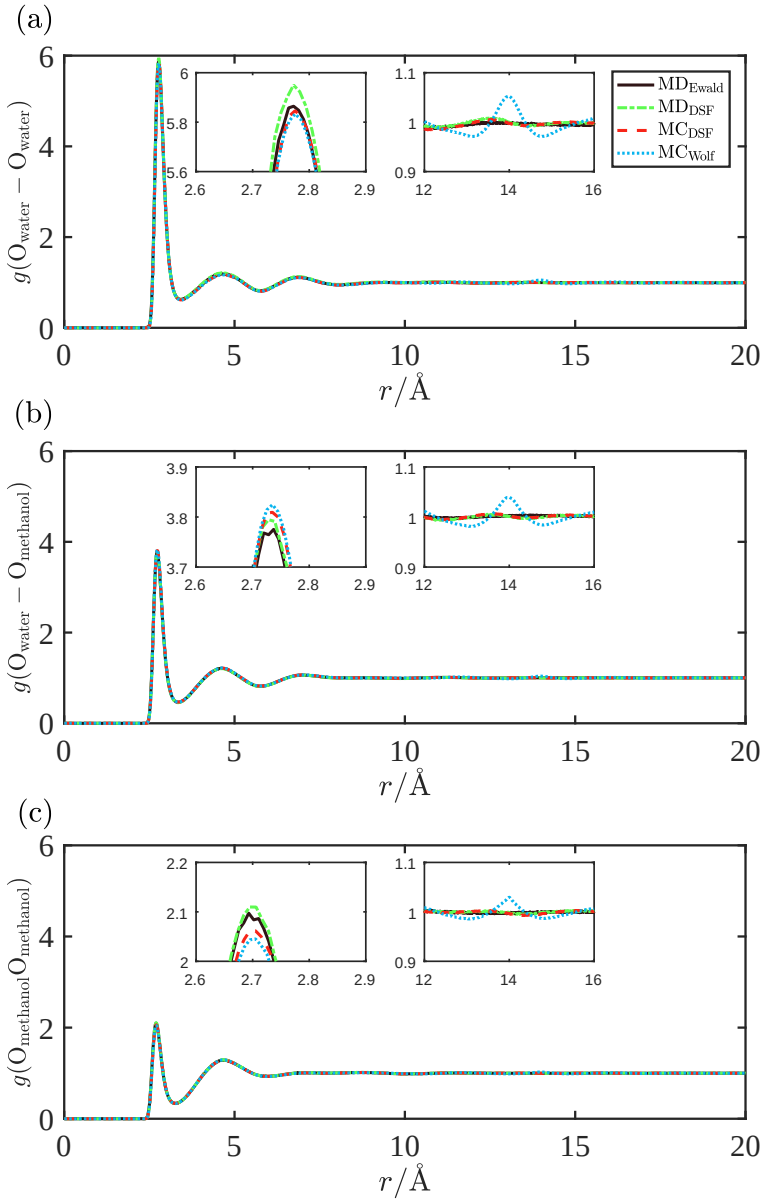


Figure 4.2: Radial distribution functions of water-methanol mixtures (50%-50%), at $T = 298$ K and $P = 1$ bar, for: (a) water-water (b) water-methanol (c) methanol-methanol. The TIP4P/2005 [255] and TraPPE [207] force fields were used to compute the density of water-methanol mixtures in MD and MC simulations. The relative difference in densities obtained from MD and MC simulations was 0.2%. To compute the long-range electrostatic interactions, the Ewald and DSF methods were used in MD simulations. In the MC simulations, the Wolf and DSF methods (Eqs. 4.3 and 4.5) were used.

all electrostatic methods, both in MD and MC simulations can capture the same probability distributions for the first and second coordination shells. The small difference observed at the first coordination shells in Fig. 4.2(a) to Fig. 4.2(c) is practically negligible. This can be due to the difference in the computed densities obtained from independent simulations, and/or statistical noise in the simulations [15]. As shown in Fig. 4.2, all the RDFs have converged to unity after distance of 10 Å. For systematic discussion of the RDFs obtained from experiments and molecular simulations, the reader is referred to Refs. [27, 86]. In all the RDFs obtained using the Wolf method, a numerical artifact is observed at $R_c = 14$ Å. This Gaussian-shaped artifact is most noticeable for $g(O_{H_2O}\text{-}O_{H_2O})$ in Fig. 4.2(a), with less than 5% deviation from 1. The artificial structuring at $R_c = 14$ Å indicates a non-physical behavior at the cutoff [274, 278]. This is due to the discontinuity in intermolecular electrostatic interactions between the molecules inside and outside the cutoff sphere. To investigate to what extent the numerical artifact of the Wolf method affects the computed thermodynamic properties, excess mixing enthalpies of water-methanol mixtures obtained from MD and MC simulations were compared based on different treatments of electrostatics: the Ewald, DSF and the Wolf methods. The water-methanol mixtures were defined based on the TIP4P/2005 [255] and TraPPE [207] force fields. The results are compared in Fig. 4.3. From the MD simulation results, it is clear that the Ewald and DSF methods yield identical excess mixing enthalpies at different mole fractions of methanol, x_{MeOH}. Therefore, the excess mixing enthalpies based on the DSF method, from MD simulations, were considered as reference. At $x_{MeOH} = 0.9$, excess mixing enthalpies obtained from MD and MC simulations are equal within statistical uncertainty. At mole fractions $x_{MeOH} = 0.5$ and $x_{MeOH} = 0.7$, the computed excess mixing enthalpies using the Wolf method, from MC simulations, are equal, within statistical uncertainty, to those obtained using the DSF method, from MD simulations. At mole fraction $x_{MeOH} = 0.3$, the results using the DSF method, from MC simulations are in excellent agreement with the MD simulation results. Furthermore, the MD and MC simulations yield marginally different results at $x_{MeOH} = 0.1$ and identical results at $x_{MeOH} = 0.9$. This means that there is no clear distinction be-

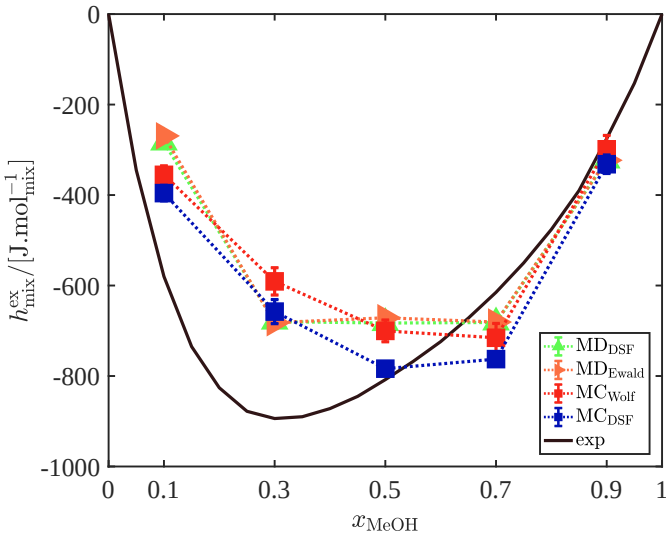


Figure 4.3: Excess enthalpies of mixing for water-methanol mixtures based on the TIP4P/2005 [255] and TraPPE [207] force fields at $T = 298$ K and $P = 1$ bar. To compute the electrostatic energies, the DSF and Ewald [146] methods were used in MD simulations. In MC simulations, the Wolf and DSF methods (Eqs. 4.3 and 4.5) were used to treat the electrostatic interactions. The solid line indicates experimental values for the excess mixing enthalpy [95]. Dotted lines are a guide to the eye. Error bars are smaller than symbol sizes. Raw data are listed in the Supporting Information of Ref. [222].

tween the Wolf and DSF methods in calculating excess mixing enthalpies of water-methanol mixtures. To further investigate the numerical artifact of the Wolf method, excess chemical potentials and activity coefficients of water and methanol are computed using the Wolf and DSF methods, in MC simulations. The experimental values for activity coefficients of water and methanol are taken from Ref. [117]. Experimental excess chemical potentials of water and methanol, μ^{ex}, at different mole fractions were calculated using Eq. 2.9. The experimental density data are provided in Ref. [102]. The excess chemical potentials of pure water and pure methanol, μ_0^{ex}, were computed from empirical equations of state [145, 231, 232], using the REFPROP software [297]. For this comparison, the TIP4P/2005 [255] and TraPPE [207] force fields were used for simulations in the CFCNPT ensemble. The computed chemical potentials and activity coefficients of water and methanol are shown in Figs. 4.4 and 4.5, respectively. It is important to note that at low concentrations of water and methanol, i.e. $x_{\text{MeOH}} = 0.1$ and 0.9, larger error bars are observed for the excess chemical potentials. This is due to the smaller number of molecules of one of the species (either water or methanol), limiting the number of identity changes of the fractional molecule. Therefore, the sampling of the excess chemical potential becomes more difficult. Since activity coefficients are computed directly from the excess chemical potentials (Eq. 4.8), only at low concentrations of water and methanol, the values of the activity coefficients display scatter. Considering larger error bars at low mole fractions, it can be seen that the results from the Wolf and DSF methods are in excellent agreement. This suggests that the artifact of the Wolf method observed in the RDFs has a minor effect on thermodynamic properties of water-methanol mixtures. For the rest of the paper, the results based on the DSF method in MC simulations are presented, and the raw data corresponding the DSF method are presented in the Supporting Information of Ref. [222]. The results obtained based on the Wolf method, Eq. 4.3 are not considered further in Section 9.4 due to this artifact. The corresponding properties computed using the Wolf method are listed in the Supporting Information of Ref. [222].

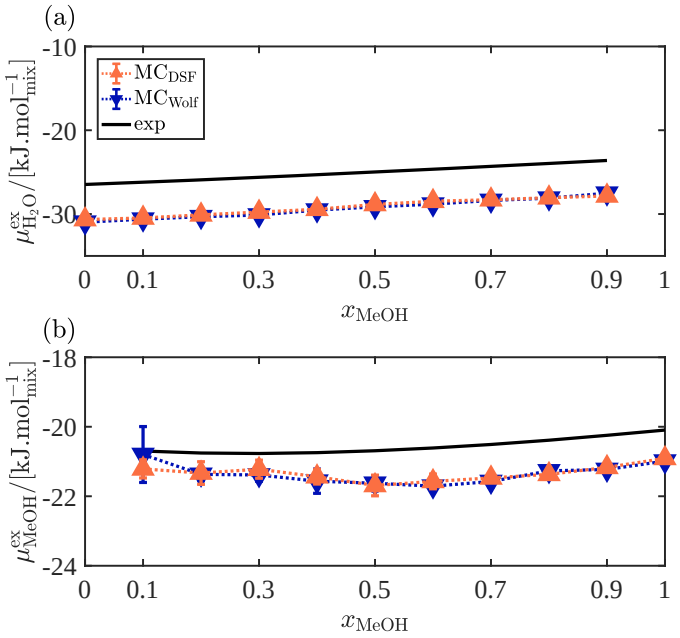


Figure 4.4: Excess chemical potentials of: (a) water, (b) methanol, with respect to the ideal gas phase, in water-methanol mixtures obtained from MC simulations in the CFC*NPT* ensemble [20], at $T = 298$ K and $P = 1$ bar. The Wolf and the DSF methods (Eqs. 4.3 and 4.5) were used to calculate the electrostatic interactions. The TIP4P/2005 [255] and TraPPE [207] force fields were used. Error bars are smaller than symbol sizes. Raw data are listed in Tables S5 and S15 of the Supporting Information of Ref. [222].

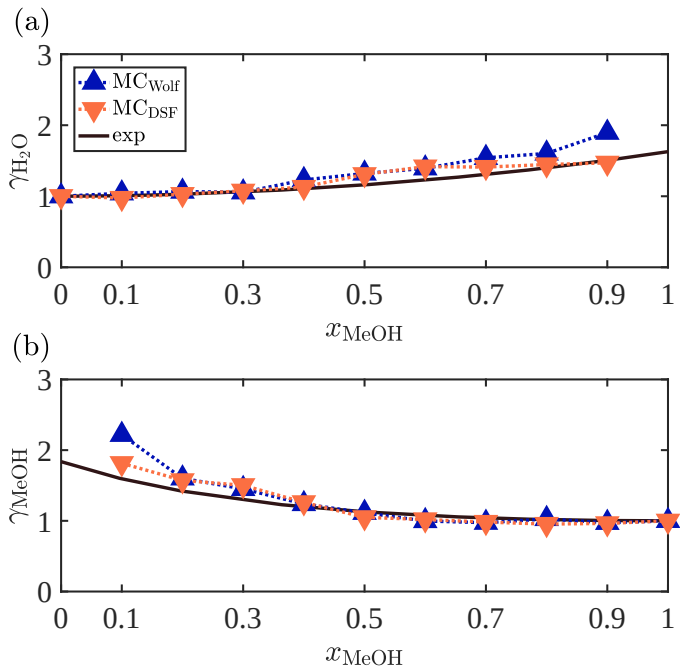


Figure 4.5: Activity coefficients of: (a) water, (b) methanol in water-methanol mixtures obtained from MC simulations in the CFC*NPT* ensemble, at $T = 298$ K and $P = 1$ bar. The Wolf [236] and the DSF [149] methods were used to calculate the electrostatic interactions. The TIP4P/2005 [255] and TraPPE [207] force fields were used. The line indicates experimental values for the activity coefficients [101]. Raw data are listed in the Supporting Information of Ref. [222].

4.4.2 Thermodynamic properties of water-methanol mixtures

Few molecular simulation studies investigated the properties of water-methanol mixtures using the OPLS/2016 force field [262] and popular rigid, non-polarizable force fields for water [85–87]. In some cases, non-Lorentz-Berthelot mixing rules [87] were applied to improve the predicted thermodynamic properties of water-methanol compared to experiments. To the best of our knowledge, a comparative analysis of these force field combinations to compute the activity coefficients and excess chemical potentials of water and methanol is missing. To compute the chemical potentials and activity coefficients of water and methanol, TIP3P [258], SPC/E [235], OPC [256], TIP4P/2005 [255], TIP4PEW [261] force fields for water, and the OPLS/2016 [262] and TraPPE [207] force fields for methanol are considered. The excess mixing enthalpies of water-methanol mixtures were computed for all water-methanol force field combinations, using the DSF method (Eq. 4.5) to calculate the electrostatic interactions. The results are shown in Fig. 4.6 as a function of x_{MeOH}, and the raw data are listed in the Supporting Information of Ref. [222]. From Fig. 4.6, it is clear that computing the excess mixing enthalpies for water-methanol mixtures using the TraPPE force field for methanol provides considerably better results compared to the OPLS/2016 force field. The sign of the excess mixing enthalpy is predicted correctly and its parabolic shape is reproduced with partial success for all four-site water force fields [86], for $x_{\mathrm{MeOH}} > 0.5$. None of the water-methanol force field combinations can precisely reproduce experimental excess enthalpies and the location of the minimum for $x_{\mathrm{MeOH}} < 0.5$. Different experimental studies suggest that the unique thermodynamic behavior of water-methanol mixtures arises from incomplete mixing of the species at molecular level [27–29]. Segregation of water and methanol, and formation of clusters, is reported in neutron diffraction experiments [27–29] and molecular simulation studies [86] for the whole concentration range. In aqueous methanol solutions, it is observed that hydrophobic methyl groups tend to cluster together, while the hydrophilic hydroxyl groups are pushed further apart and oriented more towards water-rich regions [27]. This leads

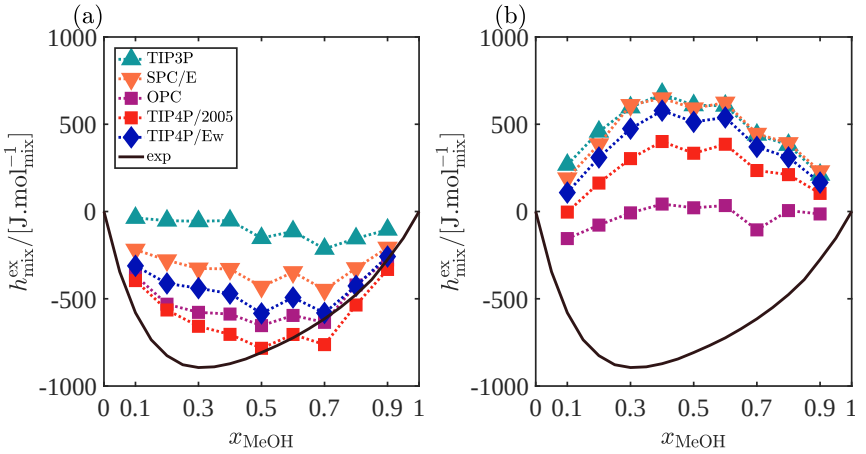


Figure 4.6: Excess mixing enthalpies for water-methanol mixtures defined by: (a) TraPPE [207] and (b) OPLS/2016 [262] force fields at $T = 298$ K and $P = 1$ bar. The TIP3P [259], SPC/E [235], OPC [256], TIP4P/2005 [255], TIP4P/EW [261] force fields were considered for water. The DSF method (Eq. 4.5) was used to treat the electrostatic interactions. The solid line indicates experimental values for the excess mixing enthalpy [95]. Dotted lines are a guide to the eye. Raw data are listed in the Supporting Information of Ref. [222].

to a reduction in the extent of the methanol-methanol hydrogen bonding network compared to pure methanol and an addition in the extent of the water-methanol hydrogen bonding network [27, 28]. In contrast to methanol, no significant change in the local structure of water is observed in neutron diffraction studies [28]. These observations suggest that water-methanol hydrogen bonding network has a strong influence on the behavior of the excess properties of water-methanol mixtures. Based on the significant deviation between simulation results and experiments, for $x_{\mathrm{MeOH}} < 0.5$, it can be concluded that the selected force field combinations, cannot reproduce the actual clustering/orientation of methanol molecules in aqueous mixtures. In this work, the TIP4P/2005-TraPPE potential outperforms the other force field combinations in predicting the excess mixing enthalpy and its shape. For water-methanol mixtures defined by the OPLS/2016 force field, the sign of the excess mixing enthalpy is not reproduced, except

partially for the OPC-OPLS/2016 potential. A comparative analysis of the TraPPE and OPLS/2016 force field parameters, in Table 4.2, shows that partial charges on the oxygen and methyl sites are larger for the TraPPE force field. Similarly, in Table 4.1, it is shown that the four-site water models have larger partial charges on the oxygen or dummy site. Clearly, increasing partial charges plays an important role in producing the excess molar enthalpies of water-methanol mixtures closer to the experimental values. This is in agreement with the work of Dopazo-Paz et al. [86]. Excess chemical potentials with respect to the ideal gas phase and activity coefficients of water and methanol, were calculated using Eqs. 2.9 and 4.8. The results are shown in Figs. 4.7 and 4.8, and the raw data are provided in the Supporting Information of Ref. [222]. In Fig. 4.7, the computed activity coefficients of water and methanol for all force field combinations were plotted as a function of x_{MeOH}. Overall, activity coefficients of methanol obtained based on the OPLS/2016 and TraPPE force fields are in good agreement with the experiments for $x_{\mathrm{MeOH}} > 0.5$. For the OPLS/2016 force field, significant deviation from experiments is observed for $x_{\mathrm{MeOH}} < 0.5$. In contrast to the OPLS/2016 force field, the computed activity coefficients of methanol for the TraPPE force field are considerably closer to the experimental results. For water, the predicted activity coefficients are in good agreement for $x_{\mathrm{MeOH}} < 0.3$. The predicted activity coefficients of water in mixtures defined by the OPLS/2016 force field deviate significantly from experimental data for $x_{\mathrm{MeOH}} > 0.3$, except for the TIP3P force field. It is clear from Fig. 4.7 that the activity coefficients of different water models, obtained in combination with the TraPPE force field, are in better agreement with the experiments. In Fig. 4.7, it can be seen that the TIP4P2005-TraPPE potential outperforms other force field combinations to predict the activity coefficients closest to the experimental values. Among water-methanol mixtures defined by the OPLS/2016 force field, the activity coefficients obtained from TIP3P-OPLS/2016 force fields deviate less from the experiments. As the excess chemical potential and activity coefficient are related (Eq. 4.8), it is important that both properties agree well with experiments. It is expected that the performance of a force field combination should be the same in predicting both the excess chemical potentials and

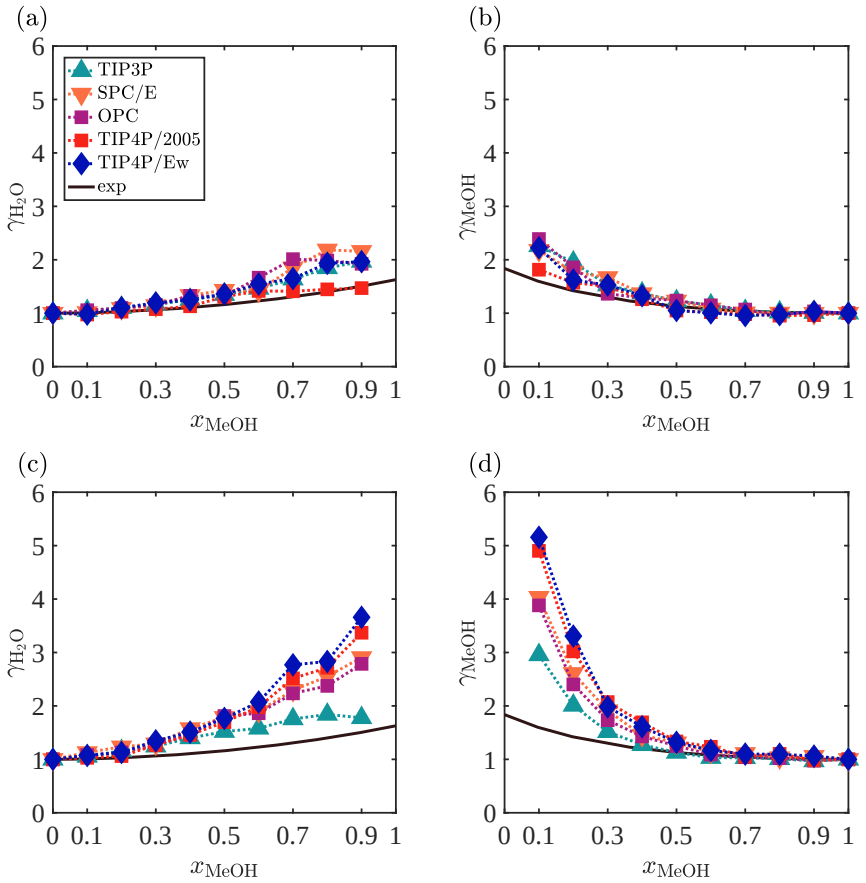


Figure 4.7: Activity coefficients of water and methanol in water-methanol mixtures for different combinations of water-methanol force fields, at $T = 298$ K and $P = 1$ bar. In subfigures (a) and (b); the TraPPE force field was used for methanol and in subfigures (c) and (d); the OPLS/2016 force field was used for methanol. The TIP3P [259], SPC/E [235], OPC [256], TIP4P/2005 [255], TIP4P/EW [261] force fields were considered for water. The DSF method (Eq. 4.5) was used to treat the electrostatic interactions. The solid lines indicate experimental values for the activity coefficients [101]. Dotted lines are a guide to the eye. Raw data are listed in the Supporting Information of Ref. [222].

activity coefficients of water and methanol. Since water and methanol force fields are not fitted to experimental chemical potentials, some deviation is expected depending on the force field [59]. In Fig. 4.8, it is shown that all force fields predict the excess chemical potentials with some deviation/shift with respect to the experimental values. Since computing the activity coefficients depends only on the difference between the excess chemical potentials, see Eq. 4.8, a constant shift between the predicted excess chemical potentials and the experimental data does not introduce an error in computing the activity coefficients. This is because term $(\mu_{\mathrm{A}}^{\mathrm{ex}} - \mu_{0\mathrm{A}}^{\mathrm{ex}})$ in Eq. 4.8 remains constant. This is especially the case for methanol when $x_{\mathrm{MeOH}} > 0.5$. For $x_{\mathrm{MeOH}} < 0.5$, the calculated excess chemical potentials deviate significantly from the experiments, for methanol OPLS/2016. Clearly, this leads to a considerable error in computing the activity coefficients. It can be seen in Fig. 4.8 that the TIP4P/2005 and TIP3P show the best performance combined with the TraPPE and OPLS/2016 force fields, respectively. For pure components, the excess chemical potential of pure water predicted by the TIP3P force field has the best agreement with the empirical equation of state [232]. The excess chemical potential of water predicated by the OPC force field deviates the most from experimental data. For pure methanol, the computed excess chemical potential using the TraPPE force field agrees best with the experimental equation of state [145, 231].

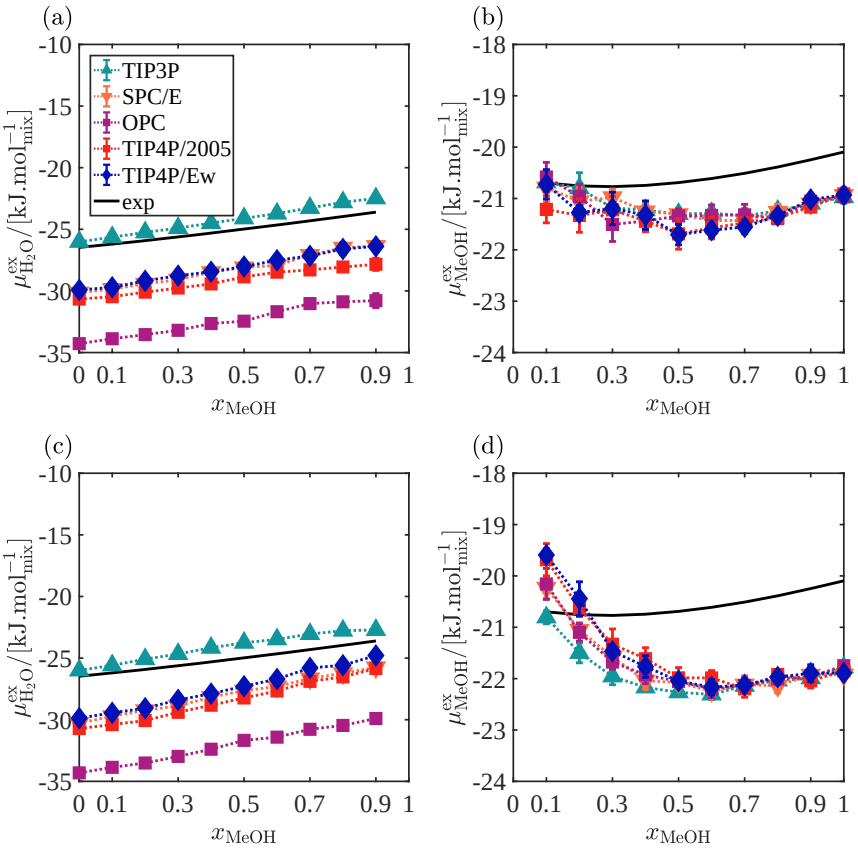


Figure 4.8: Excess chemical potentials of water and methanol, with respect to the ideal gas phase, for different combinations of water-methanol force fields, at $T = 298$ K and $P = 1$ bar. In subfigures (a) and (b); the TraPPE force field was used for methanol and in subfigures (c) and (d); the OPLS/2016 force field was used for methanol. The TIP3P [259], SPC/E [235], OPC [256], TIP4P/2005 [255], TIP4P/EW [261] force fields were considered for water. Error bars are smaller than symbol sizes. The solid lines indicate experimental values for the chemical potentials [145, 231, 232, 297]. Dotted lines are a guide to the eye. Raw data are listed in the Supporting Information of Ref. [222].

4.5 Conclusions

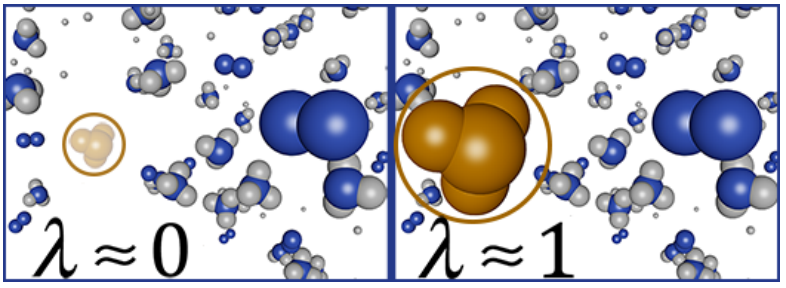
To predict and reproduce the thermodynamic properties of water, a large number of force fields have been published in literature. The most popular force fields for water are rigid non-polarizable potentials. The performance of these force fields depend on the experimental data used for fitting the force field parameters. These force fields can be combined with other force fields to calculate thermodynamic properties of aqueous solutions. In this work, different force field combinations for water-methanol mixtures were considered to compute excess mixing enthalpies, excess chemical potentials and activity coefficients of water and methanol. In MC simulations, spherical cutoff-based methods are computationally more efficient compared to the Ewald-type methods since computation of electrostatic interactions is reduced to the molecules inside the cutoff sphere. To investigate the accuracy of two spherical cutoff-based methods, i.e., the Wolf and DSF methods (Eqs. 4.3 and 4.5), RDFs and excess mixing enthalpies of aqueous solutions of methanol were computed and compared to the results obtained from the Ewald summation method. The RDFs and excess mixing enthalpies obtained from the Ewald summation and DSF methods were in excellent agreement. We observed numerical artifacts at the cutoff radius in RDFs in simulations using the Wolf method. Based on the RDFs, it can be concluded that some orientational correlation exists between between the molecules inside and outside the cutoff sphere. This may imply that the dielectric properties of water are not treated correctly [281]. However, the good agreement between the excess mixing enthalpies, activity coefficients, and chemical potentials computed based on the DSF and Wolf methods suggests that these numerical artifacts have a small effect on thermodynamic properties. By using the DSF method, we investigated the performance of the TraPPE and OPLS/2016 force fields for methanol combined with five water models: TIP3P, SPC/E, OPC, TIP4P/2005, and TIP4P/EW. For these force field combinations, we computed excess mixing enthalpies, chemical potentials and activity coefficients of water and methanol. All these properties are reproduced better by the TraPPE force field compared to the OPLS/2016 force field. The predicted properties of water-methanol mixture

defined by TIP4P/2005-TraPPE force fields show the best agreement with experimental data compared to other force fields combinations.

Chapter 5

Computation of Partial Molar Properties Using Continuous Fractional Component Monte Carlo

This chapter is based on the following paper: Rahbari, A.; Hens, R.; Nikolaidis, I. K. ; Poursaeidesfahani, A.; Ramdin, M.; Economou, I.G.; Moultos, O. A. ; Dubbeldam, D.; Vlugt, T. J. H., Computation of Partial Molar Properties using Continuous Fractional Component Monte Carlo, 116, 3331–3344, 2018, Ref. [20]



5.1 Introduction

In the previous chapters, we have shown the advantages of the CFCMC method for efficient molecule insertion/deletion and free energy calculations. In chapter 3, we proposed a new binning scheme to improve the accuracy of the computed chemical potentials in the 2016 version of the CFCMC [38]. In this chapter, we develop a new formulation to directly compute the partial molar properties using CFCMC. Partial molar excess enthalpies and partial molar volumes are key properties in studying thermodynamics of multicomponent fluid mixtures [3, 5, 298]. Knowledge of these quantities is central to process design of chemical and biochemical processes [11, 299–305], including separation systems [306], chemisorption processes [173, 304, 305], equilibrium and non-equilibrium reactive systems [173, 304, 305]. Unfortunately, partial molar properties are computationally difficult to calculate and are experimentally difficult to measure at extreme conditions [150, 153, 307–310]. At present, application of computer simulations to calculate partial molar properties is limited and more work is needed in this field [30, 152].

Partial molar properties are first order derivatives of the chemical potential [8, 14, 150, 153]. The partial molar enthalpy of component A in a multicomponent mixture equals

$$\bar{h}_{\mathrm{A}} = \left(\frac{\partial H}{\partial N_{\mathrm{A}}}\right)_{T,P,N_{i\neq \mathrm{A}}} = \left(\frac{\partial \beta \mu_{\mathrm{A}}}{\partial \beta}\right)_{P,N_i} \tag{5.1}$$

For convenience, in this chapter, partial molar properties are considered per molecule instead of per mole. In Eq. 5.1, H is the enthalpy of the system, N_i denotes the number of molecules (or mole) of component i, μ_{A} is the chemical potential of component A, P is the imposed pressure, T is the temperature, $\beta = 1/(k_{\mathrm{B}}T)$, and k_{B} is the Boltzmann constant. The partial molar volume of component A equals

$$\bar{v}_{\mathrm{A}} = \left(\frac{\partial V}{\partial N_{\mathrm{A}}}\right)_{T,P,N_{i\neq \mathrm{A}}} = \left(\frac{\partial \mu_{\mathrm{A}}}{\partial P}\right)_{T,N_i} \tag{5.2}$$

in which V is the volume of the mixture. In molecular simulation, chemical potentials and partial molar properties cannot be computed directly as a function of atomic positions and/or momenta of the molecules in the system [14, 38, 39, 150, 153], and special molecular simulation techniques are required. To date, different molecular simulation techniques have been used to compute partial molar properties: *(1)* Numerical Differentiation (ND): in a multicomponent mixture, a partial molar property of component A is computed directly by numerically differentiating the total property of the mixture at constant temperature and pressure with respect to the number of molecules of component A, while keeping the number of molecules of all other components constant [5, 311, 312]. This requires several independent and long simulations. Therefore, it is not well suited for multicomponent mixtures. Moreover, the accuracy of the numerical differentiation depends strongly on the uncertainty of the computed total property [150, 153]. *(2)* From fluctuations in the number of particles using multiple linear regression [152]. The method can be applied to the reaction ensemble which allows for direct computation of enthalpy of the reaction [152]. This is discussed further in Section 5.6. *(3)* Kirkwood-Buff (KB) integrals: Schnell *et al.* have used KB integrals to compute the partial molar enthalpies for mixtures of gases or liquids [298, 311, 313–317]. This method uses transformations between ensembles and it is numerically difficult to compute partial molar enthalpies. However, the computation of partial molar volumes using KB integrals is straightforward [318]. *(4)* Direct Method: in their pioneering work in 1987, Frenkel, Ciccotti, and co-workers used the WTPI method [211] to compute partial molar properties of components in a single MC simulation in the *NPT* ensemble [150, 153]. Due to the inefficiency of the WTPI method for high density systems, application of this method is rather limited [59, 150, 191, 210, 319, 320]. *(5)* Difference Method (DM): to avoid sampling issues of the WTPI method, an alternative approach was proposed by Frenkel, Ciccotti, and co-workers which uses identity changes between two molecule types [150, 153]. From this, partial molar properties of binary systems could be computed. However, if the two molecules are very different in size or have very different interactions with surrounding molecules, identity changes often lead to unfavorable

configurations in phase-space, resulting again in poor statistics. *(6)* Using multiple linear regression in open constant-pressure ensembles, as proposed by Josephson, Siepmann, and co-workers. This is performed by fitting extensive thermodynamic properties (*e.g.* the enthalpy and volume) as a function of the instantaneous number of molecules of each component [152]. This requires an ensemble in which the number of molecules of each component fluctuates (*e.g.* the NPT version of the Gibbs ensemble [44], reaction ensemble [321], or grand-canonical ensemble [14]). For details, the reader is referred to Ref. [152], Section 5.6, and chapter 6. *(7)* Using a combination of the CFCMC technique with umbrella sampling [14, 322]. With this method, one can perform a simulation of a mixture at a certain pressure and temperature, and accurately compute the chemical potential at other temperatures and pressures close to the simulation conditions [323]. Partial molar volumes and enthalpies are obtained directly from the estimated chemical potentials. This is the topic of chapter 7.

In this chapter, we combine the original idea of Frenkel, Ciccotti, and co-workers [150, 153], and the CFCMC method (see chapters 2 and 3) to compute partial molar properties by gradual insertion/removal of molecules. This avoids the well-known drawbacks of the WTPI method at high densities. As a test case, partial molar properties in the NPT ensemble and the expanded NPT ensemble are computed for a 50%-50% binary LJ color mixture, *i.e.* a mixture where all interactions are identical. Since the WTPI method works efficiently for this mixture, the results are used to verify our method. Next, our method is applied to a case of industrial relevance, *i.e.* the Haber-Bosch process for ammonia production [324]. The reason to select this mixture as a realistic case study is that ammonia is a useful chemical commodity and has received lots of attention both in academia and industry [62–68]. It is also a promising alternative medium for energy storage and transportation [69–72]. Industrial ammonia synthesis is carried out using the Haber-Bosch process with heterogeneous iron or ruthenium catalysts at high temperatures (623 K - 873 K) and at a pressure range of 20 MPa to 40 MPa [73–75]. In industrial applications, the ammonia synthesis and many other gas phase reactions are mostly modeled with cubic Equations of State (EoS) because of their simplicity [12, 325, 326].

Ammonia is a molecule that forms hydrogen bonds, but this phenomenon cannot be modeled using a standard cubic EoS [7]. Moreover, limitations of using a cubic EoS in studying thermodynamic properties of mixtures at high pressures are well-known [9–11]. Therefore, due to the hydrogen bonding of ammonia, and the elevated pressures at which the ammonia synthesis reaction takes place, it is of interest to study the pressure dependency of partial molar properties of the mixture using physically based models (*i.e.* molecular simulation and PC-SAFT [18, 19, 327]), and compare the results to those obtained from a cubic EoS. In this work, partial molar properties for the [NH_3, N_2, H_2] mixture at chemical equilibrium, based on the Haber-Bosch reaction [62, 324], are computed at $T = 573$ K and a pressure range of 10 MPa to 80 MPa [63, 66, 328, 329].

This chapter is organized as follows. In Section 5.2, expressions for partial molar properties derived by Frenkel, Ciccotti, and co-workers are reviewed [150, 153]. Expressions for partial molar properties in the expanded version of the NPT ensemble are introduced. In Section 5.3, our recent work on the reaction ensemble [142] is reviewed and applied for the Haber-Bosch process to obtain the composition of the reacting mixture at chemical equilibrium. The composition of the [NH_3, N_2, H_2] mixture at equilibrium is used to compute the partial molar properties. Simulation details and an overview of the systems considered in our simulations are summarized in Section 5.4. Our simulation results are presented in Section 5.5. It is shown that the computed partial molar properties for a binary LJ color mixture obtained using both methods are identical. Partial molar properties for [NH_3, N_2, H_2] mixtures at chemical equilibrium are computed as a function of pressure. Based on these results, the reaction enthalpy of the Haber-Bosch process is computed using MC simulations and EoS modeling. It is shown that the results obtained from MC simulations and PC-SAFT EoS modeling are in excellent agreement. The results obtained from PR-EoS modeling deviate from those obtained from MC simulations and PC-SAFT EoS modeling at high pressures. This leads to a relative difference of up to 8% in calculated reaction enthalpies at 80 MPa. Our conclusions are summarized in Section 5.7.

5.2 Theory

As shown in chapter 2, the chemical potential of a component can be calculated using the WTPI method [211] by sampling the interaction energy of a test molecule of the same type, inserted at a randomly selected position in the system (Eq. 2.1). In 1987, Frenkel, Ciccotti, and co-workers extended the WTPI method to compute first order derivatives of the chemical potential, namely the partial molar excess enthalpy and the partial molar volume [150, 153]. These authors have shown that the partial molar excess enthalpy of a component A in the conventional NPT ensemble of a multicomponent mixture using WTPI method equals

$$\begin{aligned}\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = & \\ & -\frac{1}{\beta}+\frac{\left\langle(\Delta U^{\mathrm{A}+}+U(s^N,V)+PV)V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}{\left\langle V\exp[-\beta\Delta U^{\mathrm{A}+}]\right\rangle_{N_i,P,T}} \\ & -\left\langle U(s^N,V)+PV\right\rangle_{N_i,P,T}\end{aligned} \tag{5.3}$$

The brackets $<\cdots>_{N_i,P,T}$ denote an ensemble average in the NPT ensemble in which the number of molecules of each component i is constant. s are the scaled coordinates of molecules in the system, N is the total number of molecules, and $U(s^N,V)$ is the total energy of the system. For an ideal gas, the partial molar excess enthalpy of Eq. 5.3 equals zero, since there are no interactions between ideal gas molecules. This is shown analytically in appendix A.1. The partial molar volume of component A equals

$$\bar{v}_{\mathrm{A}}=\frac{\left\langle V^2\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}{\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}-\left\langle V\right\rangle_{N_i,P,T} \tag{5.4}$$

A detailed derivation of Eqs. 5.3 and 5.4 is provided in appendix A.1, and the Supporting Information of Ref. [20]. Although Eqs. 2.1, 5.3 and 5.4 are correct, their application is rather limited because of the inefficient sampling of the WTPI method at high densities. Ensemble averages computed using the WTPI method strongly depend on the spontaneous occurrence

of cavities large enough to accommodate the test molecule. These spontaneous cavities occur very rarely at high densities which renders the WTPI method essentially inefficient. To circumvent sampling problems of the WTPI method, the CFCMC technique is used to compute the ensemble averages of Eqs. 2.1, 5.3 and 5.4 without relying on test particle insertions/removals. An expanded version of the conventional NPT ensemble is introduced in Eq. 3.5 by adding a so-called fractional molecule (CFCNPT). Implementation details regarding the scaling of the interaction potential of fractional molecules are explained in chapters 2 to 4. The partition function of Eq. 3.5 can be extended to mixtures of polyatomic molecules by simply multiplying it by the ideal gas partition function of each polyatomic molecule (excluding the translational part) [4, 14, 173]. This changes only the reference state or the ideal gas contribution of the partial molar properties and not the excess part [4, 14, 298].

In appendix A.1, it is shown that by expanding the NPT with a fractional molecule of component A, the partial molar enthalpy of component A can be computed in the CFCNPT ensemble using

$$\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = -\frac{1}{\beta} + \langle H\left(\lambda_{\mathrm{A}} \uparrow 1\right)\rangle_{\mathrm{CFC}NPT} - \frac{\langle H/V\left(\lambda_{\mathrm{A}} \downarrow 0\right)\rangle_{\mathrm{CFC}NPT}}{\langle 1/V\left(\lambda_{\mathrm{A}} \downarrow 0\right)\rangle_{\mathrm{CFC}NPT}} \tag{5.5}$$

$\langle H(\lambda_{\mathrm{A}} \uparrow 1)\rangle_{\mathrm{CFC}NPT}$ is ensemble average enthalpy of the system in the limit at which λ_{A} approaches one. $\langle H/V\ (\lambda_{\mathrm{A}} \downarrow 0)\rangle_{\mathrm{CFC}NPT}$ is ensemble average of the ratio between total enthalpy and the volume of the system in the limit at which λ_{A} approaches zero. In appendix A.1, it is shown that Eqs. 5.3 and 5.5 yield identical results.

The expression for the partial molar volume of component A in the CFCNPT ensemble equals

$$\bar{v}_{\mathrm{A}} = \langle V(\lambda_{\mathrm{A}} \uparrow 1)\rangle_{\mathrm{CFC}NPT} - \langle 1/V\ (\lambda_{\mathrm{A}} \downarrow 0)\rangle^{-1}_{\mathrm{CFC}NPT} \tag{5.6}$$

$\langle V(\lambda_{\mathrm{A}} \uparrow 1)\rangle_{\mathrm{CFC}NPT}$ is ensemble average of volume when λ_{A} approaches one, and $\langle 1/V\ (\lambda_{\mathrm{A}} \downarrow 0)\rangle_{\mathrm{CFC}NPT}$ is ensemble average of the inverse volume when λ_{A} approaches zero. In appendix A.1, it is shown that the computed ensemble average of Eq. 5.6 is equal to that computed in the conventional

NPT ensemble (Eq. 5.4). Furthermore, it is shown that the partial molar volume of an ideal gas molecule obtained from Eqs. 5.4 and 5.6 results in RT/P (as expected). Using Eqs. 2.9, 5.5 and 5.6, one can compute the chemical potential, partial molar excess enthalpy, and partial molar volume of a component in a single simulation without relying on the WTPI method or identity changes. A potential drawback of using Eqs. 5.5 and 5.6 is that the partial molar (excess) properties are obtained by subtracting two large numbers (at $\lambda = 1$ and at $\lambda = 0$) with a (relatively) small difference. This may induce large error bars, similar to the ND method explained earlier. In Section 5.5, it is shown how this can be avoided.

The partial molar excess enthalpy of component A in a mixture can also be computed using EoS modeling. In this chapter partial molar excess enthalpies and partial molar volumes of NH_3, N_2 and H_2 are computed using the PR-EoS [151, 330] and the PC-SAFT EoS [18, 19, 331–333]. Details about PR-EoS and PC-SAFT EoS are provided in appendix A.5. The expression for the partial molar enthalpy of component A, relative to the standard reference state equals [3]

$$\bar{h}_{\mathrm{A}}(T,P) = \bar{h}^{\circ}_{f,\mathrm{A}} + \left[\bar{h}_{\mathrm{A}}(T,P_{\mathrm{ref}}) - \bar{h}_{\mathrm{A}}(T_{\mathrm{ref}},P_{\mathrm{ref}})\right] + \left[\bar{h}^{\mathrm{ex}}_{\mathrm{A}}(T,P) - \bar{h}^{\mathrm{ex}}_{\mathrm{A}}(T,P_{\mathrm{ref}})\right] \quad (5.7)$$

T_{ref} and P_{ref} are the standard reference state temperature and pressure defined at 298 K and 1 bar, respectively. $h^{\circ}_{f,\mathrm{A}}$ is the formation enthalpy of component A at the standard reference state (T_{ref}, P_{ref}), which can be found in literature [3, 334]. The second term on the right hand side of Eq. 5.7 is associated with the enthalpy difference at (T, P_{ref}) at constant composition relative to the reference state. The last term on the right hand side of Eq. 5.7 is associated with the excess enthalpy difference between states (T, P) and (T, P_{ref}). This accounts for departure from ideal gas behavior relative to the standard reference pressure [3]. $h^{\mathrm{ex}}_{\mathrm{A}}(T,P)$ can be obtained from molecular simulation (Eqs. 5.5 and 5.6), EoS modeling, or literature. At high temperatures and a pressure of 1 bar, $h^{\mathrm{ex}}_{\mathrm{A}}(T,P_{\mathrm{ref}})$ is considered to be zero (ideal gas behavior). For details, the reader is referred to A.6. The reaction enthalpy of the Haber-Bosch process (per mole of N_2)

at state (T, P) is calculated using

$$\Delta\bar{h}(T,P) = 2\bar{h}_{\mathrm{NH_3}}(T,P) - \bar{h}_{\mathrm{N_2}}(T,P) - 3\bar{h}_{\mathrm{H_2}}(T,P) \tag{5.8}$$

Details are provided in appendix A.7.

5.3 Reaction Ensemble

The Reaction ensemble Monte Carlo (RxMC) technique [321, 328, 335, 336] is used in MC simulations to obtain equilibrium distribution of reactants and reaction products. Beside thermalization trial moves (translation, rotation, etc.), additional trial moves are performed to insert/delete reactants or reaction products. In RxMC simulations, the chemical reactions should be specified in advance and the chemical reactions are mimicked as MC trial moves. By using the correct acceptance rules, chemical equilibrium is obtained [321, 328, 335, 336]. RxMC is purely thermodynamic and reaction kinetics is not considered. This means that the simulations are not affected by the height of the activation energy barrier. To sample the correct phase equilibria using RxMC, one relies on sufficient molecule exchange between the reactants and reaction products [336]. It is well-known that RxMC struggles with the insertions/deletion of molecules at low temperatures and high densities [177, 328]. For a detailed review of RxMC techniques, the reader is referred to Ref. [328]. In the formulation introduced by Rosch and Maginn [177] which combines the CFCMC technique with the Reaction Ensemble, it is not possible to directly calculate the chemical potentials [142, 177]. We recently introduced a more efficient application of the CFCMC in the Reaction Ensemble (Rx/CFC) [142]. The main ingredient in our formulation of Rx/CFC is that the fractional molecules of either reactants or reaction products are present, and that the trial moves mimicking chemical reactions always involve fractional molecules [142]. In the Rx/CFC simulations, a coupling parameter λ is introduced as an extended variable for each component participating in the reaction. To perform the reaction in a continuous and gradual manner, three additional trial moves involving the fractional molecules are introduced (similar to the application of the

CFCMC in the GE as described in chapter 2): *(1)* changing the value of λ: the value of λ is changed while the positions and orientations of all other molecules are fixed. *(2)* reaction for the fractional molecules: while the number of whole molecules is kept constant, a trial move is performed to remove the fractional molecules of reactants and insert the fractional molecules of reaction products at randomly selected positions, with randomly selected orientations. In this trial move, λ, and the positions and orientations of all other molecules are fixed. *(3)* reaction for the whole molecules: the fractional molecules of reactants are transformed into whole molecules, and randomly selected reaction product molecules are transformed into fractional molecules. In this trial move, λ, and the positions and orientations of all molecules are fixed. As an illustrative example, the trial moves associated with the fractional molecules in Rx/CFC simulations of the Haber-Bosch reaction, $N_2 + 3\,H_2 \rightleftharpoons 2\,NH_3$ are shown in Fig. 5.1.

For the formulation of the partition function of the Rx/CFC, and acceptance rules for the fractional molecules, the reader is referred to Ref. [142]. The RxMC method requires the ideal gas partition functions of all reactant and reaction product molecules, a list of all possible chemical reactions in the system, and an appropriate force field accurately describing interactions between molecules [328]. The ideal gas partition function can be obtained from quantum calculations [4, 337, 338] or standard thermochemical tables *e.g.* the JANAF tables [334]. Computing ideal gas partition functions using quantum packages is well-established [338]. However, due to lack of experimental data, it is not always straightforward to evaluate the accuracy of the quantum calculations, especially for large molecules [172, 339]. Ideal gas partition functions are directly related to the standard chemical potential of pure components [334, 340]. In appendix A.2, it is explained in detail how the ideal gas partition functions can be obtained from standard thermochemical tables and computational chemistry (experiments or calculations). In thermochemical tables, energies are normally referenced relative to a zero of energy. Therefore, a consistent choice of energy reference is central to calculating the partition functions, which is also explained in appendix A.2. To the obtain the equilibrium distribution for the ammonia synthesis reaction, the ideal gas partition functions for nitrogen, hydrogen and ammonia

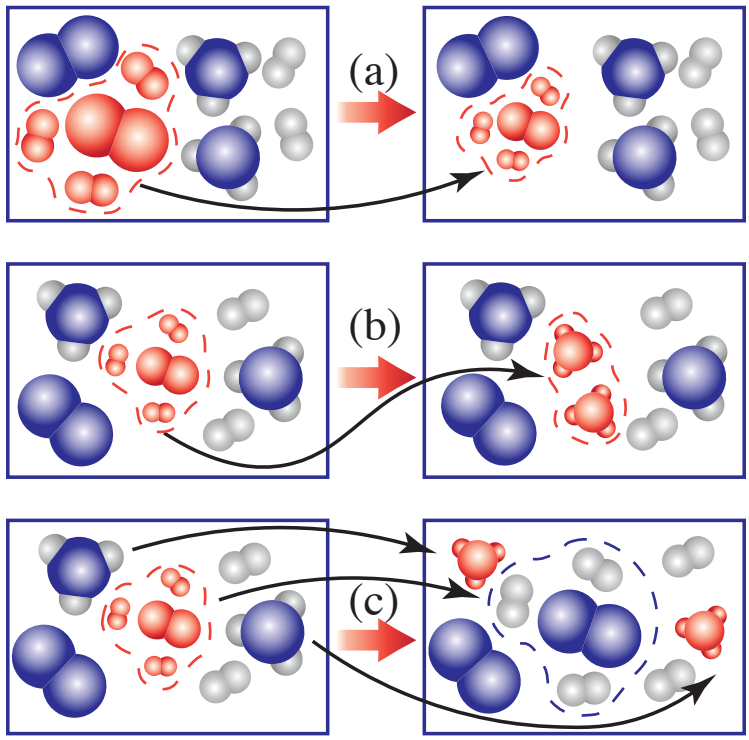


Figure 5.1: Trial moves associated with fractional molecules in Rx/CFC simulations of the Haber-Bosch reaction $N_2 + 3\,H_2 \rightleftharpoons 2\,NH_3$. (a) changing the value of λ: the positions and orientations of all other molecules are fixed. Here the interactions of the reactants ($N_2 + 3\,H_2$) are scaled. (b) reaction for the fractional molecules: the fractional molecules of the reactants are removed and the fractional molecules of reaction products are inserted. The positions and orientations of other molecules are kept fixed. (c) reaction for the whole molecules: the fractional molecules of the reactants ($N_2 + 3\,H_2$) are transferred into whole molecules, and randomly selected reaction products ($2\,NH_3$) are transformed into fractional molecules.

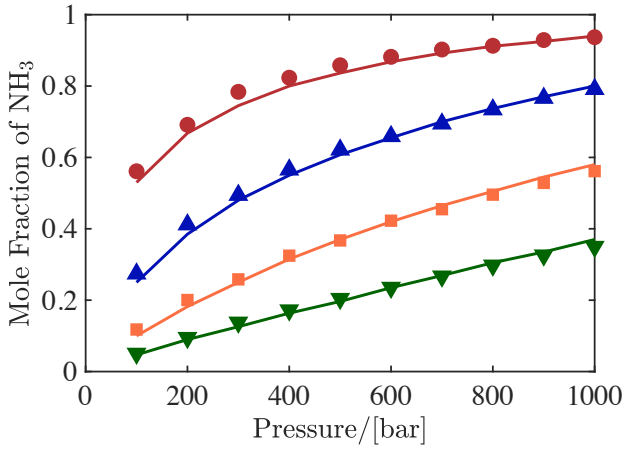


Figure 5.2: Mole fractions of ammonia at equilibrium obtained from Rx/CFC simulations (symbols) and experiments (solid lines) [341] at 573K (red), 673K (blue), 773K (orange) and, 873K (green) as a function of pressure. All simulations start with a random configuration of 120 N_2, 360 H_2 molecules, and no ammonium molecules.

are calculated in appendix A.3 using JANAF tables [334], quantum calculations using Gaussian09 and experimental thermochemical data [4, 340]. The results are provided in Table A5 in the appendix. Details of simulations of the ammonia synthesis reaction using Rx/CFC are provided in Ref. [142]. In that paper, it is shown that by using Rx/CFC, the efficiency of the reaction trial moves are significantly improved in systems with high density, compared to the conventional reaction ensemble or the method proposed by Rosch and Maginn [177]. The equilibrium mixture compositions of the ammonia synthesis reaction are obtained from the Rx/CFC simulations in the temperature range between $T = 573$ K and $T = 873$ and a pressure range of 10 MPa to 100 MPa. The results are shown in Fig. 5.2. Excellent agreement is observed between experiments and molecular simulations results. The equilibrium mixture compositions obtained at $T = 573$ K and pressures between 10 MPa and 80 MPa are used in this chapter to calculate partial molar properties of nitrogen, hydrogen and ammonia in the CFCNPT ensemble, and calculate the reaction enthalpy at high pressures.

5.4 Simulation Details

As a proof of principle, MC simulations were performed to compute the partial molar properties of a binary LJ color mixture (composition: 50%-50%), both in the conventional NPT ensemble using the WTPI method and in the CFCNPT ensemble. Details on performing simulations in the NPT and CFCNPT ensembles are provided in chapter 3. All simulations were carried out at a reduced temperature of $T^* = 2$ and reduced pressures between $P^* = 0.1$ and $P^* = 9$, leading to average reduced densities between $\rho^* = 0.052$ and $\rho^* = 0.880$. The binary color mixture contained 200 LJ molecules. In the conventional NPT ensemble, 6×10^6 production cycles were carried out. To sample partial molar properties of each component, ten trial insertions per cycle were performed. For simulations in the CFCNPT ensemble, 50×10^6 production cycles were carried out for the same binary mixture at the same reduced temperature and reduced pressures. Each trial move was selected with the following probabilities: 1% volume changes, 33% translations, 33% λ changes, and 33% hybrid trial moves chapter 3. For the hybrid trial moves, switching points at $\lambda_s = 0.3$ and $\lambda = 0.8$ are defined.

The equilibrium compositions of the [NH_3, N_2, H_2] mixture at 573 K and at various pressures from 10 MPa to 80 MPa were obtained by performing the reaction ensemble simulations of the Haber-Bosch process as described in Section 5.3. The computed equilibrium compositions are in excellent agreement with experimental data [63, 65, 66]. All molecules are rigid, and a combination of LJ and electrostatic interactions is used for the force fields. For simulation details of the reaction ensemble, the reader is referred to Ref. [142]. Equilibrium compositions were then used to initiate the NPT and CFCNPT simulations of the [NH_3, N_2, H_2] mixture. The equilibrium compositions were also used as input for PR-EoS modeling and PC-SAFT EoS modeling. Simulation details corresponding to each method are summarized below:

(a) CFCNPT ensemble: simulations were carried out to compute partial molar properties of NH_3, N_2 and H_2 at 573 K and eight pressures between $P = 10$ MPa and $P = 80$ MPa. To compute partial molar

properties of each component, separate simulations were performed in the CFCNPT ensemble, in which one fractional molecule of that component was added to the system. This was repeated for all eight pressures, leading to a total number of 24 independent simulations. The starting mixture compositions for each pressure are listed in the Supporting information of Ref. [20]. At each pressure, six independent simulations were carried out where 2×10^5 equilibration cycles were performed to compute the weight function $W(\lambda)$ using the WL algorithm [199, 200]. Starting with equilibrated configurations and weight functions, 3.2×10^6 production cycles were carried out. This leads to six data points per pressure per component. For the rest of the simulation details in the CFCNPT ensemble, the reader is referred to chapter 3. Extrapolation was used to evaluate Eqs. 2.9, 5.5 and 5.6 at $\lambda = 0$ and $\lambda = 1$. As shown in chapter 3, it is expected that the systematic error due to extrapolation is avoided by using the new binning scheme.

(b) ND method: NPT ensemble simulations of the [NH_3, N_2, H_2] mixture were carried out to compute the partial molar properties of each component at 573 K and a pressure range of $P = 10$ MPa to $P = 80$ MPa. To compute the partial molar properties of NH_3, N_2 or H_2 (component A) at each pressure, NPT ensemble simulations of the mixture were performed by changing the number of the molecules of component A with respect to that of the equilibrium mixture (N_A), while keeping the number of all other molecules in mixture constant. Seven mixture compositions were used with $N_A \pm 1$, N_A, $N_A \pm 3$ and $N_A \pm 5$ molecules around the composition of the equilibrium mixture. Independent NPT simulations were performed for every mixture composition and pressure to compute the total enthalpy (H) and the volume (V) of the mixture as a function of N_A. First order polynomials were fitted to the H and V as a function of N_A. The slopes of these lines were calculated to obtain the partial molar excess enthalpy ($\bar{h}_A^{ex} = (\partial H^{ex}/\partial N_A)_{T,P,N_{i\neq A}}$) and the partial molar volume

($\bar{v}_{\mathrm{A}} = (\partial V/\partial N_{\mathrm{A}})_{T,P,N_{i\neq \mathrm{A}}}$) of component A, respectively. Details on the simulations in the NPT ensemble are provided in chapter 3.

(c) EoS modeling: The PC-SAFT and Peng-Robinson EoS were used to the compute partial molar excess enthalpies and partial molar volumes of the $[\mathrm{NH_3}, \mathrm{N_2}, \mathrm{H_2}]$ mixture at the same temperature and pressure range. The same mixture compositions were used for these calculations. For PC-SAFT EoS modeling, ammonia is treated as an associating molecule with 4 association sites [333]. The Binary Interaction Parameters (BIPs) are set to zero both for the PR-EoS and PC-SAFT EoS. For additional details, the reader is referred to the Supporting Information of Ref. [20].

5.5 Results

To illustrate the trial moves for the fractional molecule, in Fig. 5.3a the acceptance probabilities for reinsertions and identity changes are shown as a function of λ, for the LJ system at $\langle\rho^*\rangle = 0.052$, $\langle\rho^*\rangle = 0.433$, and $\langle\rho^*\rangle = 0.880$. Reinsertion trial moves of fractional molecules are always accepted when λ approaches zero. This is the case for all densities/pressures, which is due to the very limited interactions of the fractional molecule at low values of λ. For the system at the highest reduced density ($\langle\rho^*\rangle = 0.880$), reinsertion attempts are mostly rejected for $\lambda > 0.3$. This is due to overlaps between the reinserted fractional molecule and whole molecules. The acceptance probabilities of attempted identity changes as a function of λ are shown in Fig. 5.3b. In sharp contrast to reinsertions, molecule exchanges are mostly accepted when the value of λ is close to one. This is expected as a fractional molecule with nearly fully scaled interactions behaves almost as a whole molecule. Therefore, the energy difference associated with this trial move is small at values of λ close to one. For the system with the highest density, identity changes are mostly rejected when $\lambda < 0.3$. A whole molecule within an equilibrated system already has favorable interactions with the surrounding molecules. For small values of λ, exchanging a whole molecule

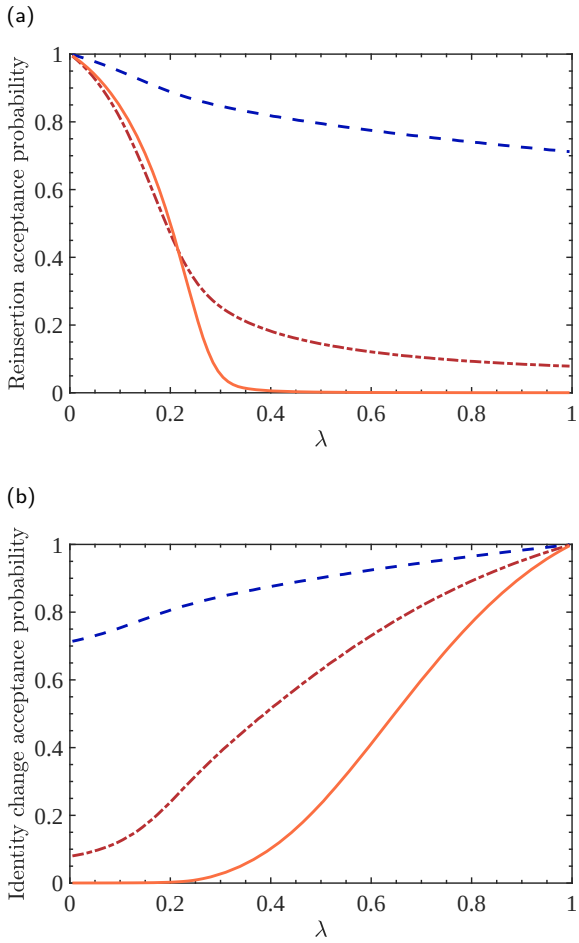


Figure 5.3: (a) Acceptance probabilities for reinserting and (b) identity changes of the fractional molecule of a binary LJ color mixture (50%-50%) consisting of 200 molecules at $T^* = 2$ and different reduced pressures. In both subfigures: $P^* = 0.1$ and $\langle \rho^* \rangle_{\mathrm{CFC}NPT} = 0.052$ (dashed blue line), $P^* = 1$ and $\langle \rho^* \rangle_{\mathrm{CFC}NPT} = 0.433$ (dash-dotted red line), $P^* = 9$ and $\langle \rho^* \rangle_{\mathrm{CFC}NPT} = 0.880$ (solid orange line).

with the fractional molecule results in formation of a cavity which has an unfavorable energy. As a result, the energy difference associated with this trial move is high at low values of λ. It can be concluded that defining a switch will ensure a high acceptance probability for the hybrid trial move. It is found that the same value can be used for the switching point ($\lambda_s = 0.3$) in the simulations of the [NH_3, N_2, H_2] system. We feel that this is a coincidence.

To validate our final expressions for the partial molar excess enthalpy and partial molar volume (Eqs. 5.5 and 5.6), the values of the partial molar properties obtained using simulations in the CFCNPT ensemble and simulations in the NPT ensemble (as proposed by Frenkel, Ciccotti, and co-workers [150, 153]) are compared in Fig. 5.4. The excellent agreement between the results shows that Eqs. 5.5 and 5.6 are implemented correctly. For values of λ close to one and zero, the quantities in Eqs. 8 and 9 are well behaved. For a typical example, the reader is referred to Figs. A1 and A2 in appendix A.1. The error introduced by extrapolating is smaller than the error bars from the independent simulations. Computed excess chemical potentials using both methods (Eqs. 2.1 and 2.9) are in excellent agreement as well. The raw data of Fig. 5.4 and the excess chemical potentials are listed in the Supporting Information of Ref. [20]. The partial molar excess enthalpies of the LJ system at a reduced temperature of $T^* = 2$ and reduced pressures between $P^* = 0.1$ and $P^* = 9$ are computed using both methods, and the results are compared in Fig. 5.4a. At low pressures (low densities), excellent agreement is found, and the error bars are small. In Fig. 5.4a, it is shown that the values of partial molar excess enthalpies approach zero at low pressures which indicates (and confirms) the ideal gas behavior. However, there is a clear distinction between the computed partial molar excess enthalpies at high pressures (high densities) using these two methods. The performance of the method proposed by Frenkel, Ciccotti, and co-workers [150, 153] strongly depends on the sampling efficiency of the WTPI method, and it is well known that WTPI becomes less efficient at high densities [59, 319]. Indeed, the values of partial molar excess enthalpies computed using the WTPI method display scatter as the pressure increases, and the error bars are significantly larger compared to those obtained from

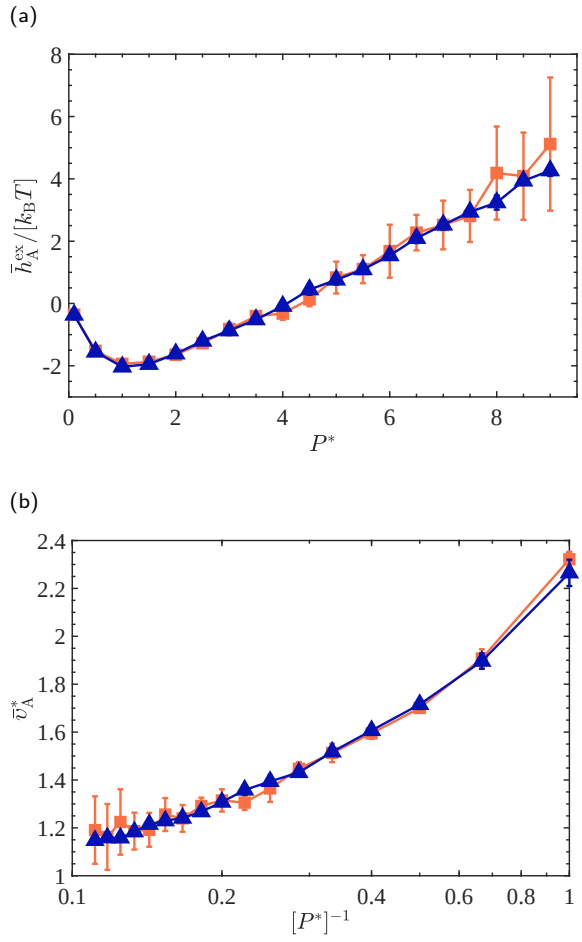


Figure 5.4: (a) Computed partial molar excess enthalpies (Eqs. 5.3 and 5.5) and (b) partial molar volumes (Eqs. 5.4 and 5.6) of a LJ molecule in a binary color mixture consisting of 200 molecules (50%-50%) at $T^* = 2$, reduced pressures between $P^* = 0.1$ and $P^* = 9$, and reduced densities ranging from $\langle\rho^*\rangle_{\mathrm{CFC}NPT} = 0.052$ to $\langle\rho^*\rangle_{\mathrm{CFC}NPT} = 0.880$. For an ideal gas, a horizontal line is expected in Fig. 5.4b. In both subfigures: computed properties in the CFC*NPT* ensemble (blue triangles), computed properties using the WTPI method in the conventional *NPT* ensemble (orange squares) as proposed by Frenkel, Ciccotti, and co-workers [150, 153]. Some error bars are smaller than the symbol size. Raw data are listed in Supporting Information of Ref. [20].

CFCNPT simulations. CFCNPT simulations provide better statistics for computing the partial molar excess enthalpies as the density of the system increases, and the magnitude of the error bars remains almost the same for the whole pressure range. Average partial molar volumes computed using both methods are shown in Fig. 5.4b. Similarly, this comparison shows that computation of the partial molar volumes using the WTPI method at high pressures results in poor statistics. Average partial molar volumes computed using CFCNPT simulations have considerably smaller error bars at high pressures.

It is instructive to compare the efficiency of the CFCNPT method with the ND method. This was tested for the LJ systems. To compute the partial molar properties using the ND method, two simulations are performed in the NPT ensemble (with $N-1$ and $N+1$ molecules respectively). The partial molar properties in the CFCNPT ensemble are obtained by computing the quantities in Eqs. 5.5 and 5.6 at $\lambda = 0$ and $\lambda = 1$ in a single simulation. Essentially, in both methods the derivatives of Eq. 5.1 are computed using two data points. This means that to obtain the same accuracy in the computed partial molar properties, a single simulation performed in the CFCNPT ensemble inevitably needs to be longer than each of the NPT simulations. We have verified numerically that the error bars are very similar for both methods for a given amount of CPU time (data not provided here). The advantage of the CFCMC approach is that one can compute the excess chemical potential from the same simulation (Eq. 2.9).

In Fig. 5.5, the partial molar excess enthalpies and the partial molar volumes of NH_3 are plotted as a function of pressure. The raw data with error bars are listed in in the Supporting Information of Ref. [20]. The results are presented from the four methods discussed in Section 5.4, namely the PR-EoS modeling, the PC-SAFT EoS modeling, the ND method, and the CFCNPT simulations. Fig. 7.1a shows that the results of the CFCNPT ensemble simulations are in excellent agreement with those obtained from the ND method. This is used as an independent check to validate our method for systems other than a LJ system. The values of partial molar excess enthalpies of NH_3 are negative, and they decrease with increasing pressure. Excellent agreement is also found between the results from MC

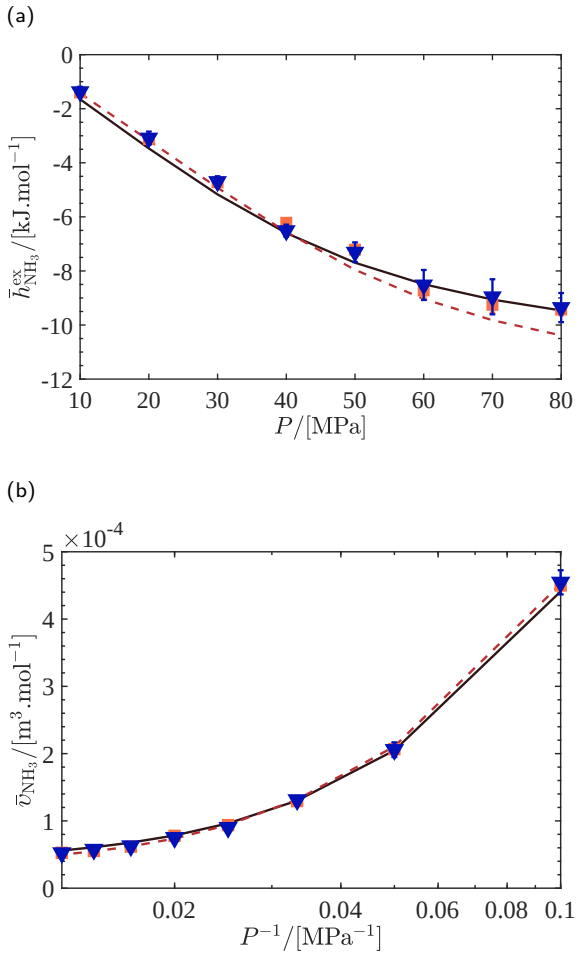


Figure 5.5: (a) Computed partial molar excess enthalpies of NH_3 and (b) computed partial molar volumes of NH_3 in a $[NH_3, N_2, H_2]$ equilibrium mixture at 573 K and pressure range of $P = 10$ MPa to $P = 80$ MPa. The compositions of the mixtures are obtained from equilibrium simulations of the Haber-Bosch reaction using Rx/CFC [142], and are listed in the Supporting Information of Ref. [20]. In both subfigures: computed properties using the PR-EoS (solid black line), computed properties using the PC-SAFT (dashed red line), computed properties using the ND method (orange squares), computed properties in the CFC*NPT* ensemble (blue triangles) using Eqs. 5.5 and 5.6. Zero BIPs were used for the EoS modeling. Raw data are listed in the Supporting Information of Ref. [20].

simulations and EoS modeling for pressures up to 50 MPa. At pressures higher than 50 MPa, the results obtained from PR-EoS and PC-SAFT EoS deviate slightly from each other (0.88 $kJ \cdot mol^{-1}$ at $P = 80$ MPa). For pressures between 50 MPa and 80 MPa, computed partial molar excess enthalpies of NH_3 obtained from MC simulations agree better with those obtained from the PR-EoS. Computed partial molar volumes of NH_3 using all methods are shown in Fig. 7.1b. The results from MC simulations and EoS modeling are in excellent agreement for the whole pressure range.

In Fig. 5.6a, computed partial molar excess enthalpies of N_2 are shown as a function of pressure. In sharp contrast to NH_3, the values of the partial molar excess enthalpies of N_2 are positive and increase with increasing pressure. In addition, the difference between the PC-SAFT EoS and the PR-EoS is more obvious, specifically for pressures higher than 50MPa. As the pressure increases, better agreement is found between the results obtained from the PC-SAFT EoS and CFC*NPT* simulations. In Fig. 5.6b, the computed partial molar volumes of N_2 are plotted as a function of pressure. Overall, very good agreement between all methods is observed. Raw data are listed in the Supporting Information of Ref. [20]. Computed partial molar excess enthalpies of H_2 increase also with increasing pressure, as shown in Fig. 5.7a. The results obtained from the PR-EoS deviate the most from the other methods. This difference contributes directly to a significant deviation in the calculated reaction enthalpy (Fig. 5.8) using the PR-EoS, specifically at high pressures. The partial molar excess enthalpies obtained from MC simulations and PC-SAFT EoS are in excellent agreement for pressures up to 70 MPa. Partial molar volumes of H_2 computed using different methods are in excellent agreement as shown in Fig. 5.7b.

Excess chemical potentials of NH_3, N_2 and N_2 are computed using EoS modeling, CFC*NPT* simulations, and the results are compared to those obtained from Rx/CFC simulations [142] in the Supporting Information of Ref. [20]. Computed excess chemical potentials using the different methods are within 1 $kJ.mol^{-1}$.

Our results from the PR-EoS were obtained using zero BIPs, and slight improvement is expected when using non-zero BIPs [342]. The non-zero BIPs for the PR-EoS were taken from the Aspen Plus software (version

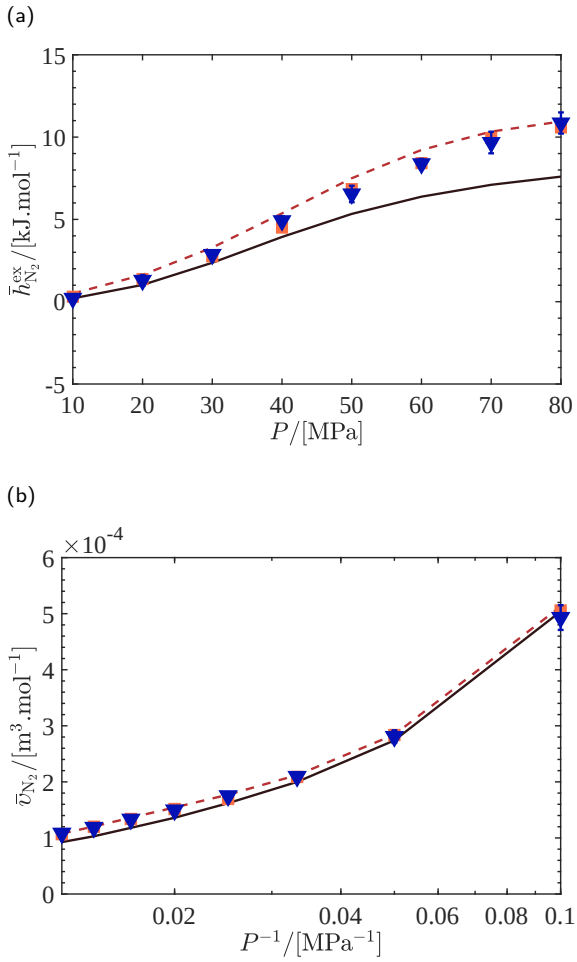


Figure 5.6: (a) Computed partial molar excess enthalpies of N_2 and (b) computed partial molar volumes of N_2 in a $[NH_3, N_2, H_2]$ equilibrium mixture at 573 K and pressure range of $P = 10$ MPa to $P = 80$ MPa. The compositions of the mixtures are obtained from equilibrium simulations of the Haber-Bosch reaction using Rx/CFC [142], and are listed in the Supporting Information of Ref. [20]. In both subfigures: computed properties using the PR-EoS (solid black line), computed properties using the PC-SAFT (dashed red line), computed properties using the ND method (orange squares), computed properties in the CFC*NPT* ensemble (blue triangles) using Eqs. 5.5 and 5.6. Zero BIPs were used for the EoS modeling. Raw data are listed in the Supporting Information of Ref. [20].

8.8) [343]. The partial molar excess enthalpies of NH_3, N_2 and N_2 were computed using PR-EoS modeling with non-zero BIPs, and the results are presented in the Supporting Information of Ref. [20]. It is shown that the difference between the partial molar excess enthalpies obtained from the PR-EoS and MC simulations/PC-SAFT become smaller at low and medium pressures, but only for N_2 and H_2. At high pressures large differences between the results obtained from the PR-EoS (using non-zero BIPs) and the other methods remain an issue. No changes were observed for the computed partial molar volumes of all components, using non-zero BIPs. Using BIPs enhances the performance of the EoS mainly in VLE calculations and not elsewhere [342, 344].

The reaction enthalpies of the Haber-Bosch process are computed at temperature of 573 K and a pressure range of $P = 10$ MPa to $P = 80$ MPa using Eqs. 5.7 and 5.8. The formation enthalpies and the ideal gas contributions are obtained from the data provided in Tables A7 and A9. The partial molar enthalpies in Eq. 5.7 are computed using the four methods discussed in Section 5.2. The computed reaction enthalpies are shown in Fig. 5.8. The reaction enthalpy of the ammonia synthesis reaction at 573 K and standard reference pressure of $P_{\text{ref}} = 1$ bar is 102.07 kJ per mole of N_2. This is indicated in Fig. 5.8 as a reference. Excellent agreement is observed between the reaction enthalpies computed using MC simulations and the PC-SAFT EoS for pressures up to 80 MPa. The reaction enthalpy computed using the PR-EoS deviates from the other methods as pressure increases (up to 8% at 80 MPa). This is associated with the well-known limitations of cubic EoS. At high pressures, volumetric estimates, fugacity coefficients and other related derivative thermodynamic properties calculated using the PR-EoS are known to be inaccurate [9–12]. Although it is one the most widely used cubic EoS in industry [345], certain drawbacks are associated with using cubic EoS. A cubic EoS cannot accurately estimate the properties of a fluid for a full temperature and pressure range. Moreover, the density dependency of the co-volume term is not known, and many different modifications of the attractive term have been proposed in literature [12, 346]. In contrast to cubic EoS, the PC-SAFT EoS is more physically based and takes into account association interactions [327]. Therefore, for associating mixtures

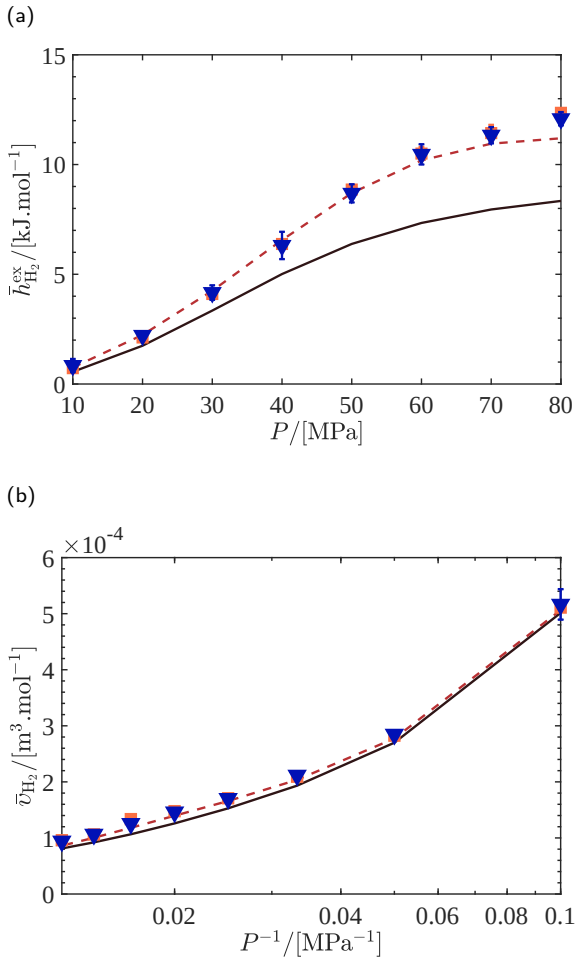


Figure 5.7: (a) Computed partial molar excess enthalpies of H_2 and (b) computed partial molar volumes of H_2 in a $[NH_3, N_2, H_2]$ equilibrium mixture at 573 K and pressure range of $P = 10$ MPa to $P = 80$ MPa. The compositions of the mixtures are obtained from equilibrium simulations of the Haber-Bosch reaction using Rx/CFC [142], and are listed in the Supporting Information of Ref. [20]. In both subfigures: computed properties using the PR-EoS (solid black line), computed properties using the PC-SAFT (dashed red line), computed properties using the ND method (orange squares), computed properties in the CFC*NPT* ensemble (blue triangles) using Eqs. 5.5 and 5.6. Zero BIPs were used for the EoS modeling. Raw data are listed in the Supporting Information of Ref. [20].

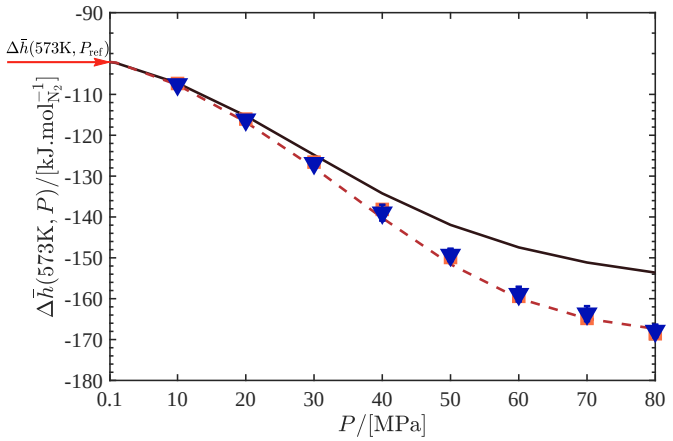


Figure 5.8: Computed reaction enthalpy of the Haber-Bosch process per mole of N_2 at 573 K and pressure range of $P = 10$ MPa to $P = 80$ MPa. The arrow on the left indicates the value of the reaction enthalpy at standard reference pressure ($P_{\mathrm{ref}} = 1$ bar). The compositions of the mixtures are obtained from equilibrium simulations of the Haber-Bosch reaction using Rx/CFC [142], and are listed in the Supporting Information of Ref. [20]. Different methods used to compute enthalpy of reaction: PR-EoS (solid black line). PC-SAFT (dashed red line), ND method (orange squares), CFC*NPT* ensemble (blue triangles). Zero BIPs were used for the EoS modeling. Raw data are listed in the Supporting Information of Ref. [20].

(including the mixture studied in this work), the results from PR-EoS modeling are expected to be less accurate compared to those obtained from the PC-SAFT EoS modeling and molecular simulations, especially at high pressures [347].

In Fig. 5.8, it is shown that the contributions of the partial molar excess enthalpies to the reaction enthalpy of the Haber-Bosch process become significant at high pressures. Not including the contribution of the partial molar excess enthalpies results in differences of 24% to 64% relative to the reaction enthalpy at the reference pressure, in the pressure range of 30 MPa to 80 MPa. From Fig. 5.2, one can observe that at chemical equilibrium, the mole fraction of NH_3 increases when the pressure increases. As NH_3 molecules show association behavior [7], this results in favorable NH_3-NH_3 interactions. This is reflected by the negative partial molar excess enthalpy

Table 5.1: Computed partial molar excess enthalpies, partial molar volumes and excess chemical potentials of NH_3, N_2 and H_2 in the [NH_3, N_2, H_2] equilibrium mixture at $T = 573$ K and pressure $P = 50$ MPa, obtained from CFC*NPT* ensemble simulations. The first three columns show the results from simulations where only one fractional molecule per component is used in each simulation. The last three columns show the results from simulations where only a single simulation is performed with three fractional molecules present at the same time (one of each component). The composition of the mixture is obtained from equilibrium simulations of the Haber-Bosch reaction using Rx/CFC [142] and is listed in the Supporting Information of Ref. [20]. Partial molar enthalpies, partial molar volumes and excess chemical potentials are reported in units of [$\mathrm{kJ.mol^{-1}}$], [$\mathrm{m^3.kmol^{-1}}$] and [$\mathrm{kJ.mol^{-1}}$], respectively. Numbers in brackets are uncertainties in the last digit, *i.e.*, 3.4(2) means 3.4 ± 0.2.

	CFCMC (1 fractional molec.)			CFCMC (3 fractional molec.)		
Component	$\bar{h}^{\mathrm{ex}}$	$\bar{v}$	μ^{ex}	$\bar{h}^{\mathrm{ex}}$	$\bar{v}$	μ^{ex}
NH_3	-7.3(4)	0.075(5)	-2.39(2)	-6.9(9)	0.079(9)	-2.37(2)
N_2	6.5(6)	0.150(4)	2.139(9)	7.3(9)	0.156(9)	2.13(3)
H_2	8.7(5)	0.146(4)	2.36(1)	8.7(7)	0.147(6)	2.36(3)

of NH_3 at high pressures, see Fig. 5.5. In sharp contrast to NH_3, the N_2-N_2 and H_2-H_2 interactions become less favorable at high pressures as indicated by a positive partial molar excess enthalpy, see Figs. 5.6 and 5.7. The net result of the pressure behavior of the partial molar excess enthalpies is that the reaction enthalpy becomes more exothermic at high pressures.

Instead of running three different simulations for the [NH_3, N_2, H_2] equilibrium mixture, (each with a single fractional molecule of a different type), it is possible to run a single simulation with three fractional molecules at the same time (one of each component). For sufficiently large systems, the fractional molecules do not influence each other, and the structure of the fluid is not disturbed by the presence of the fractional molecules [233]. Therefore, one may compute the partial molar properties of all components in a single CFC*NPT* simulation using the same method explained in Section 5.2. The drawback is that more production cycles are needed to obtain results with the same accuracy. More cycles are needed because the number of trial moves related to the coupling parameter are now distributed between three

fractional molecules, instead of one. In Table 5.1, a comparison is made between this procedure and three separate simulations for the equilibrium mixture at $P = 50$ MPa. Good agreement is found between both methods.

5.6 Reaction Enthalpy from Linear Regression

Recently, Josephson, Siepmann, and co-workers showed that partial molar properties can also be obtained from fluctuations in the number of particles of each component in the system, using multiple linear regression [152]. The method can be extended to the reaction equilibria in a closed system [152]. Based on the work of Josephson, Siepmann, and co-workers, we show that it is possible to compute the reaction enthalpy of the ammonia synthesis reaction in the CFCMC version of the reaction ensemble, using linear regression. Based on the fluctuations in the system, the relation between the total enthalpy of the system (H) and the number of nitrogen molecules ($N_{\mathrm{N_2}}$) can be written as

$$H = aN_{\mathrm{N_2}} + b + \epsilon \tag{5.9}$$

in which ϵ is statistical noise due to the fluctuations in the composition and energy, a and b are the slope and intercept of the fitted regression line. From linear regression we know that the slope is obtained from [265, 348, 349]

$$a = \frac{\langle H N_{\mathrm{N_2}} \rangle - \langle H \rangle \langle N_{\mathrm{N_2}} \rangle}{\left\langle N_{\mathrm{N_2}}^2 \right\rangle - \langle N_{\mathrm{N_2}} \rangle^2} \tag{5.10}$$

in which the brackets $\langle \cdots \rangle$ denote ensemble average in the reaction ensemble. In this section, we verify that a is the reaction enthalpy per mole of nitrogen ($a = \Delta \bar{h}_{\mathrm{sim}}$). The subscript "sim" means computed reaction enthalpy from the the simulation. It should be noted that Eq. 5.10 can also be written for ammonia or hydrogen. Since the fluctuations of reactants or reaction products depend on the extent of the reaction, this will result in identical reaction enthalpies. To verify Eq. 5.10, simulations of the ammonia synthesis

reaction at $P = 400$ bar and $P = 800$ bar are performed in the reaction ensemble in a similar manner as explained in Ref. [142]. The reaction enthalpies $\Delta\bar{h}_{\text{sim}}$ obtained using Eq. 5.10 are compared to the reaction enthalpies obtained from simulations in the CFCNPT ensemble. For this comparison, we did not need to include contributions from the formation enthalpies of components ($\bar{h}_f^\circ$), or the enthalpy differences relative to the reference temperature ($T_{\text{ref}} = 298.15$K) at $P_{\text{ref}} = 0.1$MPa in $\Delta\bar{h}_{\text{sim}}$. For details, see Eqs. 5.7 and 5.8. Excellent agreement is observed between the results from the CFCNPT ensemble simulations and direct computation of the reaction enthalpy using linear regression. It is important to note that the enthalpy Eq. 5.10 is the total enthalpy, and it includes contributions from fractional molecules [349].

It is observed that for single component systems, linear regression using Eq. 5.9 is equivalent to calculating thermodynamic derivatives such as reaction enthalpy or heat of adsorption [350]. For a single-component adsorption in the grand-canonical ensemble, one can fit a linear regression model to simulation data to obtain thermodynamic derivatives [350, 351]. In case of single-component adsorption, it follows from statistical mechanics that

$$(\partial N/\partial(\beta\mu))_{V,T} = \left\langle N^2 \right\rangle - \langle N \rangle^2 \tag{5.11}$$

and

$$(\partial U/\partial(\beta\mu))_{V,T} = \langle Un \rangle - \langle U \rangle \langle n \rangle \tag{5.12}$$

in which N is the number of molecules, μ is the chemical potential, V is the volume, U is the potential energy of the adsorbed phase and $\beta = (k_{\text{B}}T)^{-1}$ [15, 265, 349–351]. The brackets $< \cdots >$ denote ensemble averages. The partial derivative of the total energy of the adsorbed phase with respect to the number of molecules is calculated by combining Eqs. 5.11 and 5.12 [15]:

$$\left(\frac{\partial U}{\partial N}\right)_{V,T} = \left(\frac{(\partial U/\partial(\beta\mu))_{V,T}}{(\partial N/\partial(\beta\mu))_{V,T}}\right) = \frac{\langle UN \rangle - \langle U \rangle \langle N \rangle}{\langle N^2 \rangle - \langle N \rangle^2} \tag{5.13}$$

Table 5.2: Reaction enthalpy of the ammonia synthesis reaction at $T = 573$ K and $P = 400$ bar and $P = 800$ bar, computed from simulations in the CFC*NPT* ensemble [20] and the RxMC simulation using Eq. 5.10. In this table, the reaction enthalpy $\Delta\bar{h}_{\text{sim}}$ is written in terms of ideal, $\Delta\bar{h}^{\text{id}}$, and excess, $\Delta\bar{h}^{\text{ex}}$, contributions. The excess contribution of the reaction enthalpy in the CFC*NPT* ensemble is obtained using $\sum \nu_i \bar{h}_i^{\text{ex}}$ in which ν_i is the stoichiometric coefficient of component i. The excess contribution of the reaction enthalpy in the RxMC is obtained using $\Delta\bar{h}_{\text{sim}} - \Delta\bar{h}^{\text{id}}$. $\Delta\bar{h}_{\text{sim}}$ is obtained directly from Eq. 5.10, and $\Delta\bar{h}_i^{\text{id}} = \sum \nu_i \bar{h}_i^{\text{id}}$, where $\bar{h}_i^{\text{id}} = 1/\beta$. This is due to the fact that the thermal wavelength Λ in our simulation is set to unity, see Eq. A13. Enthalpies are in units of $[\text{kJ.mol}^{-1}]$.

		RxMC	CFC*NPT* [20]
P/[bar]	$\Delta\bar{h}^{\text{id}}$	$\Delta\bar{h}^{\text{ex}}$	$\Delta\bar{h}^{\text{ex}}$
400	9.53	36.1(6)	-36(2)
800	9.53	64(1)	-65(2)

In this case, fitting the simulation data (*i.e.* the instantaneous value of U as a function of N) to the simple linear regression model ($U = b_0 + b_1 N + \epsilon$, in which b_0 and b_1 are constants and ϵ represents noise) leads to the identical expression for the slope (b_1) as in Eq. 5.13 [352]. In the next chapter, we rigorously prove the equivalence of thermodynamic fluctuations and multiple linear regression for computing thermodynamic derivatives in open ensembles.

5.7 Conclusions

An alternative method is presented to compute partial molar excess enthalpies and partial molar volumes in the *NPT* ensemble combined with Continuous Fractional Component Monte Carlo (CFC*NPT*). To compute partial molar properties of component A in a mixture, the *NPT* ensemble of the mixture is expanded with a fractional molecule of type A. Computation of partial molar properties in the CFC*NPT* ensemble does not have the drawbacks of Widom-like test particle methods, since particle insertions/removals take place in a gradual manner. Three additional trial moves associated with the fractional molecule are used: *(1)* changing the

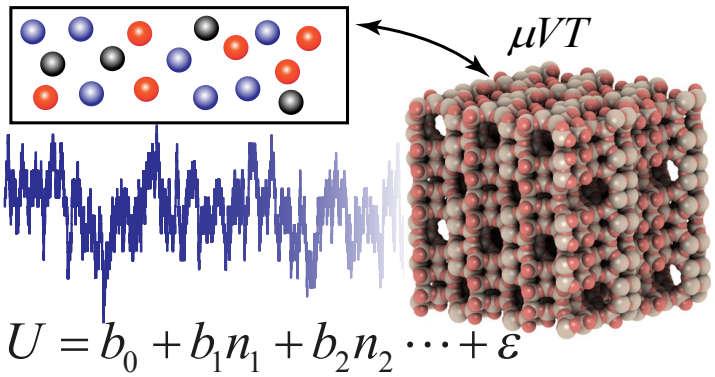
coupling parameter of the fractional molecule (λ), *(2)* reinsertion of the fractional molecule at a randomly selected position, *(3)* changing the identity of the fractional molecule with a randomly selected molecule of the same type. The latter two trial moves can be efficiently combined into a hybrid trial move which significantly enhances the sampling of partial molar properties. As a proof of principle, this method is compared to the original method of Frenkel, Ciccotti, and co-workers [150, 153] for a binary LJ color mixture at constant composition and different conditions. Partial molar properties obtained using both methods are in excellent agreement. Our method is also applied to an industrially relevant system: mixtures of NH_3, N_2 and H_2 at chemical equilibrium. We also compared our method to the Numerical Differentiation (ND) method. This provides an independent check for the results obtained from CFCNPT simulations. Excellent agreement is found between the results obtained from the ND method and the CFCNPT ensemble simulations. It would be interesting to investigate how the method works for associating large molecules with complex internal degrees of freedom and strong intramolecular interactions. For such systems, conformations with $\lambda = 0$ and $\lambda = 1$ will be very different. The PR-EoS and PC-SAFT EoS are also used to compute the partial molar properties of NH_3, N_2 and H_2 at the same mixture compositions and conditions. Excellent agreement was found between the results obtained from the molecular simulations and the PC-SAFT EoS modeling. The results obtained from the PR-EoS deviate from the other methods at high pressures. It is shown that the contribution of the partial molar enthalpies in calculating the reaction enthalpy of the Haber-Bosch process is significant at high pressures (up to 64% at a pressure of 80 MPa, relative to the reaction enthalpy at a pressure of 1 bar). It is observed that at high pressures, the contribution of the partial molar excess enthalpies is not negligible for this process, leading to a more exothermic process at high pressures. It is expected that partial molar properties at high pressures are more accurate using a physically based EoS such as PC-SAFT or advanced MC techniques compared to a cubic EoS. However, cubic EoS are widely used to study other industrially important applications due to their simplicity. For example the methanol synthesis reaction which is carried out at elevated pressures up to 10 MPa

[325, 326, 353]. To better understand these processes, it is important to use methods which can accurately model the nonideal behavior of the system. In addition, we have verified in this theis that the reaction enthalpy can be obtained directly from simulations in the RxMC ensemble using linear regression. Excellent agreement is observed between the results from the CFCNPT ensemble simulations and direct computation of the reaction enthalpy using linear regression.

Chapter 6

Computation of Thermodynamic Derivatives Using Multiple Linear Regression

Rahbari, A. ; Vlugt, T. J. H.; et al. Manuscript in preparation.



6.1 Introduction

Based on the work of Josephson, Siepmann, and co-workers [152], it was shown in chapter 5 that multiple linear regression can be used to obtain partial molar properties in open constant-pressure ensembles. For a single-component adsorption in the grand-canonical ensemble, it is known that fitting a linear regression model to simulation data leads to identical expression for $(\partial U/\partial N)_{V,T}$ as in thermodynamic fluctuations [350, 351]. In this chapter, we rigorously prove that obtaining thermodynamic derivatives in open ensembles using multiple linear regression and thermodynamic fluctuations lead to identical expressions. However, multiple linear regression provides conceptually and computationally a simple way of computing thermodynamic derivatives in multicomponent systems. In the grand-canonical ensemble, multiple linear regression can be used to obtain the heat of adsorption even around sharp inflection points where the fluctuation approach is known to fail. It was shown in chapter 5 that for single component adsorption in the grand-canonical ensemble, $(\partial U/\partial N)_{V,T}$ can obtained from thermodynamic fluctuations using Eqs. 5.11 and 5.12 or linear regression using Eq. 5.13. It is important to note that $(\partial U/\partial N)_{V,T} \neq U/N$ [354]. This is due to changes in adsorption behavior as a function of N, as already adsorbed molecules change the adsorption energy of newly adsorbed molecules. For example, this is the case when there is an inflection in the adsorption isotherm [355–357]. At the inflection point, the adsorbed molecules start to fill a new adsorption site which leads to a change in $(\partial U/\partial N)_{V,T}$. Therefore, when fitting a simple linear regression model to simulation data in the grand-canonical ensemble, the intercept should always be included in the regression model [354]. This is shown schematically in Fig. 6.1(a) for the adsorption of 2-methylpentane in MFI-type zeolite with two different adsorption sites [358]. For binary and multicomponent systems, analytic expressions for $(\partial U/\partial N_i)_{V,T,N_{j\neq i}}$ from fluctuations are complicated [350]. Therefore, for multicomponent systems, comparing the expressions for partial derivatives obtained from fluctuations and multiple linear regression becomes tedious. Here, we show that for a multicompo-

nent system, identical thermodynamic derivatives are obtained from *(1)* fitting simulation data to a multiple linear regression model [152, 348], and *(2)* fluctuations [350]. It is important to note that thermodynamic derivatives obtained from regression may not necessarily be equal to those obtained from fluctuations [349, 350]. For example, for single component adsorption in the grand-canonical ensemble, a polynomial regression of the second order ($U = b_0 + b_1 N + b_2 N^2 + \epsilon$) leads to a different expression for $(\partial U/\partial N_i)_{V,T,N_{j\neq i}}$ than in Eq. 5.13. Derivatives of a fitted function are not necessarily the same as derivatives of the function itself (*e.g.* fitting $y = x^2$ with linear regression leads to different derivatives). For the general case of k independent variables $x_1, x_2, \cdots, x_k$, the dependent variable y can be fitted with the following multiple linear regression model:

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_k x_{ik} + \epsilon \tag{6.1}$$

in which i indicates an observation from the sample [348]. Solving Eq. 6.1 using the least squares method leads to the following set of equations [348]

$$\begin{array}{lllll}
b_0 N_{\mathrm{S}} & +b_1 \sum_{i=1}^{N_{\mathrm{s}}} x_{1i} & + b_2 \sum_{i=1}^{N_{\mathrm{s}}} x_{2i} & + \cdots + b_k \sum_{i=1}^{N_{\mathrm{s}}} x_{ki} & = \sum_{i=1}^{N_{\mathrm{s}}} y_i \\
b_0 \sum_{i=1}^{N_{\mathrm{s}}} x_{1i} & +b_1 \sum_{i=1}^{N_{\mathrm{s}}} x_{1i}^2 & + b_2 \sum_{i=1}^{N_{\mathrm{s}}} x_{1i}x_{2i} & + \cdots + b_k \sum_{i=1}^{N_{\mathrm{s}}} x_{1i}x_{ki} & = \sum_{i=1}^{N_{\mathrm{s}}} x_{1i}y_i \\
b_0 \sum_{i=1}^{N_{\mathrm{s}}} x_{2i} & +b_1 \sum_{i=1}^{N_{\mathrm{s}}} x_{2i}x_{1i} & + b_2 \sum_{i=1}^{N_{\mathrm{s}}} x_{2i}^2 & + \cdots + b_k \sum_{i=1}^{N_{\mathrm{s}}} x_{2i}x_{ki} & = \sum_{i=1}^{N_{\mathrm{s}}} x_{2i}y_i \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
b_0 \sum_{i=1}^{N_{\mathrm{s}}} x_{ki} & +b_1 \sum_{i=1}^{N_{\mathrm{s}}} x_{ki}x_{1i} & +b_2 \sum_{i=1}^{N_{\mathrm{s}}} x_{ki}x_{2i} & + \cdots + b_k \sum_{i=1}^{N_{\mathrm{s}}} x_{ki}^2 & = \sum_{i=1}^{N_{\mathrm{s}}} x_{ki}y_i
\end{array} \tag{6.2}$$

in which N_{s} is the sample size. The least squares estimates of the parameters $b_0, b_1, \cdots, b_k$ are obtained by solving Eq. 6.2 using any appropriate

method [348]. It can be shown that the intercept b_0 is obtained from [348]

$$b_0 = \langle y \rangle - b_1 \langle x_1 \rangle - b_2 \langle x_2 \rangle \cdots - b_k \langle x_k \rangle \tag{6.3}$$

in which $< \cdots >$ denotes the sample mean. The numerical results by Josephson et al. [152] strongly suggest that for multicomponent systems, fluctuations and multiple linear regression lead to identical values for thermodynamic derivatives. Here, we show rigorously that this is indeed the case for multicomponent systems. Therefore, multiple linear regression provides a conceptually simple way of computing thermodynamic derivatives in multicomponent systems.

6.2 Fluctuations in the Grand-Canonical Ensemble

For multicomponent adsorption in the grand-canonical ensemble, the derivative $(\partial U/\partial N_i)_{V,T,N_{j\neq i}}$ is obtained from [350]

$$\left(\frac{\partial U}{\partial N_i}\right)_{V,T,N_{j\neq i}} = \sum_{i=1}^{k} \left(\frac{\partial U}{\partial(\beta\mu_k)}\right)_{V,T,\mu_{j\neq k}} \left(\frac{\partial(\beta\mu_k)}{\partial N_i}\right)_{V,T,N_{j\neq i}} \tag{6.4}$$

where k is the number of components. The second term on the right hand side of Eq. 6.4 is the change in chemical potential of component i while the numbers of molecules of all other components are fixed. Using a matrix notation, Eq. 6.4 can be written as

$$\begin{bmatrix} \left(\frac{\partial U}{\partial N_1}\right)_{N_{j\neq 1}} \\ \left(\frac{\partial U}{\partial N_2}\right)_{N_{j\neq 2}} \\ \vdots \\ \left(\frac{\partial U}{\partial N_k}\right)_{N_{j\neq k}} \end{bmatrix} = \underbrace{\begin{bmatrix} \left(\frac{\partial(\beta\mu_1)}{\partial N_1}\right)_{N_{j\neq 1}} & \left(\frac{\partial(\beta\mu_1)}{\partial N_2}\right)_{N_{j\neq 2}} & \cdots & \left(\frac{\partial(\beta\mu_1)}{\partial N_k}\right)_{N_{j\neq k}} \\ \left(\frac{\partial(\beta\mu_2)}{\partial N_1}\right)_{N_{j\neq 1}} & \left(\frac{\partial(\beta\mu_2)}{\partial N_2}\right)_{N_{j\neq 2}} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \left(\frac{\partial(\beta\mu_k)}{\partial N_1}\right)_{N_{j\neq 1}} & \cdots & \cdots & \left(\frac{\partial(\beta\mu_k)}{\partial N_k}\right)_{N_{j\neq k}} \end{bmatrix}}_{\mathbf{M}} \begin{bmatrix} \left(\frac{\partial U}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq 1}} \\ \left(\frac{\partial U}{\partial(\beta\mu_2)}\right)_{\mu_{j\neq 2}} \\ \vdots \\ \left(\frac{\partial U}{\partial(\beta\mu_k)}\right)_{\mu_{j\neq k}} \end{bmatrix}$$

(6.5)

in which we have defined matrix $\mathbf{M}$. The constant temperature and volume notation is dropped to make equations more compact. In the grand-canonical ensemble, the derivatives $(\partial U/\partial(\beta\mu_i))_{\mu_{j\neq i}}$ in Eq. 6.5 can expressed as the covariance between U and N_i [349–351]:

$$\left(\frac{\partial U}{\partial(\beta\mu_i)}\right)_{\mu_{j\neq i}} = f\left(U, N_i\right) \tag{6.6}$$

where $f\left(U, N_i\right)$ equals $\langle UN_i\rangle - \langle U\rangle\langle N_i\rangle$. The terms $f\left(U, N_i\right)$ can be calculated directly from simulations in the grand-canonical ensemble. However, the derivatives $(\partial(\beta\mu_i)/\partial N_l)_{N_{j\neq l}}$ in the matrix $\mathbf{M}$ (Eq. 6.5) cannot be directly calculated as ensemble averages, as the chemical potential imposed instead of adsorbed molecules. The inverse of the matrix $\mathbf{M}$ is obtained directly from the reciprocals of the derivatives in $\mathbf{M}$ [359]. In Refs. [350, 359] it is shown that

$$\begin{bmatrix} \left(\frac{\partial(\beta\mu_1)}{\partial N_1}\right)_{N_{j\neq1}} & \left(\frac{\partial(\beta\mu_1)}{\partial N_2}\right)_{N_{j\neq2}} & \cdots & \left(\frac{\partial(\beta\mu_1)}{\partial N_k}\right)_{N_{j\neq k}} \\ \left(\frac{\partial(\beta\mu_2)}{\partial N_1}\right)_{N_{j\neq1}} & \left(\frac{\partial(\beta\mu_2)}{\partial N_2}\right)_{N_{j\neq2}} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \left(\frac{\partial(\beta\mu_k)}{\partial N_1}\right)_{N_{j\neq1}} & \cdots & \cdots & \left(\frac{\partial(\beta\mu_k)}{\partial N_k}\right)_{N_{j\neq k}} \end{bmatrix} \underbrace{\begin{bmatrix} \left(\frac{\partial N_1}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq1}} & \left(\frac{\partial N_1}{\partial(\beta\mu_2)}\right)_{\mu_{j\neq2}} & \cdots & \left(\frac{\partial N_1}{\partial(\beta\mu_k)}\right)_{\mu_{j\neq k}} \\ \left(\frac{\partial N_2}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq1}} & \left(\frac{\partial N_2}{\partial(\beta\mu_2)}\right)_{\mu_{j\neq2}} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \left(\frac{\partial N_k}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq1}} & \cdots & \cdots & \left(\frac{\partial N_k}{\partial(\beta\mu_k)}\right)_{\mu_{j\neq k}} \end{bmatrix}}_{\mathbf{M}^{-1}} = \mathbf{I} \tag{6.7}$$

where $\mathbf{I}$ is the identity matrix. The advantage of this formulation is that the elements of the matrix $\mathbf{M}^{-1}$ can be calculated directly as a function of ensemble averages in this ensemble [350, 359]: where $\mathbf{I}$ is the identity matrix. The advantage of this formulation is that the elements of the matrix $\mathbf{M}^{-1}$ can be calculated directly as a function of ensemble averages in this ensemble [350, 359]:

$$\left(\frac{\partial N_l}{\partial(\beta\mu_i)}\right)_{\mu_{j\neq i}} = f\left(N_i, N_l\right) \tag{6.8}$$

where $f(N_i, N_l)$ denotes the covariance $\langle N_i N_l \rangle - \langle N_i \rangle \langle N_l \rangle$. By multiplying both sides of Eq. 6.5 with the inverse matrix $\mathbf{M}^{-1}$ we have

$$\begin{bmatrix} \left(\frac{\partial N_1}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq 1}} & \left(\frac{\partial N_1}{\partial(\beta\mu_2)}\right)_{\mu_{j\neq 2}} & \cdots & \left(\frac{\partial N_1}{\partial(\beta\mu_k)}\right)_{\mu_{j\neq k}} \\ \left(\frac{\partial N_2}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq 1}} & \left(\frac{\partial N_2}{\partial(\beta\mu_2)}\right)_{\mu_{j\neq 2}} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \left(\frac{\partial N_k}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq 1}} & \cdots & \cdots & \left(\frac{\partial N_k}{\partial(\beta\mu_k)}\right)_{\mu_{j\neq k}} \end{bmatrix} \begin{bmatrix} \left(\frac{\partial U}{\partial N_1}\right)_{N_{j\neq 1}} \\ \left(\frac{\partial U}{\partial N_2}\right)_{N_{j\neq 2}} \\ \vdots \\ \left(\frac{\partial U}{\partial N_k}\right)_{n_{j\neq k}} \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial U}{\partial(\beta\mu_1)}\right)_{\mu_{j\neq 1}} \\ \left(\frac{\partial U}{\partial(\beta\mu_2)}\right)_{\mu_{j\neq 2}} \\ \vdots \\ \left(\frac{\partial U}{\partial(\beta\mu_k)}\right)_{\mu_{j\neq k}} \end{bmatrix} \tag{6.9}$$

Rewriting Eq. 6.9 in a compact form using Eqs. 6.6 and 6.8 leads to

$$\begin{bmatrix} f(N_1, N_1) & f(N_2, N_1) & \cdots & f(N_k, N_1) \\ f(N_1, N_2) & f(N_2, N_2) & \cdots & f(N_k, N_2) \\ \vdots & & & \vdots \\ f(N_1, N_k) & f(N_2, N_k) & \cdots & f(N_k, N_k) \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{bmatrix} = \begin{bmatrix} f(U, N_1) \\ f(U, N_2) \\ \vdots \\ f(U, N_k) \end{bmatrix} \tag{6.10}$$

To make equations more compact, the terms $d_{i\in[1,k]}$ are used to denote the derivatives $(\partial U/\partial N_i)_{N_{j\neq i}}$. We can show that Eq. 6.10 is identical to the set of equations for multiple linear regression (Eq. 6.2) [348]. Writing out the l^{th} row of Eq. 6.10 leads to

$$d_1 f(N_1, N_l) + d_2 f(N_2, N_l) \cdots + d_k f(N_k, N_l) = f(U, N_l) \tag{6.11}$$

Combining Eqs. 6.6, 6.8 and 6.11 leads to

$$\begin{aligned} \langle U \rangle \langle N_l \rangle - [d_1 \langle N_1 \rangle \langle N_l \rangle + d_2 \langle N_2 \rangle \langle N_l \rangle \cdots + d_l \langle N_l \rangle \langle N_l \rangle \cdots + d_k \langle N_k \rangle \langle N_l \rangle] + \\ [d_1 \langle N_1 N_l \rangle + d_2 \langle N_2 N_l \rangle \cdots + d_l \langle N_l N_l \rangle \cdots + d_k \langle N_k N_l \rangle] = \langle U N_k \rangle \end{aligned} \tag{6.12}$$

Rearranging this equation leads to

$$\begin{aligned} \langle N_l \rangle \underbrace{[\langle U \rangle - (d_1 \langle N_1 \rangle + d_2 \langle N_2 \rangle \cdots + d_k \langle N_k \rangle)]}_{d_0} + d_1 \langle N_1 N_l \rangle + d_2 \langle N_2 N_l \rangle \\ \cdots + d_l \langle N_l N_l \rangle \cdots + d_k \langle N_k N_l \rangle = \langle U N_k \rangle \end{aligned} \tag{6.13}$$

In Eq. 6.13, we have defined the term d_0. For a sample size of N_s, this leads to the following independent linear equation

$$N_\mathrm{s} d_0 + d_1 \sum_{i=1}^{N_\mathrm{s}} N_{1i} + d_2 \sum_{i=1}^{N_\mathrm{s}} N_{2i} \cdots + d_k \sum_{i=1}^{N_\mathrm{s}} N_{ki} = \sum_{i=1}^{N_\mathrm{s}} U_i \tag{6.14}$$

Multiplying Eq. 6.13 with the sample size N_s and rearranging leads to

$$d_0 \sum_{i=1}^{N_\mathrm{s}} N_{li} + d_1 \sum_{i=1}^{N_\mathrm{s}} N_{li} N_{1i} + d_2 \sum_{i=1}^{N_\mathrm{s}} N_{li} N_{2i} \cdots + d_l \sum_{i=1}^{N_\mathrm{s}} N_{li}^2 \cdots + d_k \sum_{i=1}^{N_\mathrm{s}} N_{li} N_{ki} = \sum_{i=1}^{N_\mathrm{s}} U_i N_{li} \tag{6.15}$$

Eqs. 6.14 and 6.15 are identical to the equations for least-square multiple linear regression as in Eq. 6.2. Therefore, it is shown that identical expressions are obtained for *(1)* thermodynamic derivatives using least square multiple linear regression, and *(2)* fluctuations. For two-component mixtures, it is still possible to solve Eq. 6.4 without resorting to matrix algebra. As a service to the reader, the solution of Eq. 6.4 for a two component system is provided here. Using multiple linear regression ($U = b_0 + b_1 N_1 + b_2 N_2 + \epsilon$), the partial derivatives in the grand-canonical ensemble can be written as

$$\left(\frac{\partial U}{\partial N_1}\right)_{V,T,N_2} = \frac{\left(\langle U N_1 \rangle - \langle U \rangle \langle N_1 \rangle\right)\left(\left\langle N_2^2 \right\rangle - \langle N_2 \rangle^2\right) - \left(\langle U N_2 \rangle - \langle U \rangle \langle N_2 \rangle\right)\left(\langle N_1 N_2 \rangle - \langle N_1 \rangle \langle N_2 \rangle\right)}{\left(\left\langle N_1^2 \right\rangle - \langle N_1 \rangle^2\right)\left(\left\langle N_2^2 \right\rangle - \langle N_2 \rangle^2\right) - \left(\langle N_1 N_2 \rangle - \langle N_1 \rangle \langle N_2 \rangle\right)^2} \tag{6.16}$$

and

$$\left(\frac{\partial U}{\partial N_2}\right)_{V,T,N_1} = \frac{\left(\langle U N_2 \rangle - \langle U \rangle \langle N_2 \rangle\right)\left(\left\langle N_1^2 \right\rangle - \langle N_1 \rangle^2\right) - \left(\langle U N_1 \rangle - \langle U \rangle \langle N_1 \rangle\right)\left(\langle N_1 N_2 \rangle - \langle N_1 \rangle \langle N_2 \rangle\right)}{\left(\left\langle N_1^2 \right\rangle - \langle N_1 \rangle^2\right)\left(\left\langle N_2^2 \right\rangle - \langle N_2 \rangle^2\right) - \left(\langle N_1 N_2 \rangle - \langle N_1 \rangle \langle N_2 \rangle\right)^2}$$

(6.17)

Similar to Eq. 6.3, the intercept is obtained from

$$b_0 = \langle U \rangle - b_1 \langle N_1 \rangle - b_2 \langle N_2 \rangle \tag{6.18}$$

Although we have shown that the expressions for the partial derivatives obtained from thermodynamic fluctuations and multiple linear regression are identical, the performance of the methods may vary, especially around inflection points. A typical example is provided in Fig. 6.1(a). In this figure, each block corresponds to a different adsorption site [354]. As can be observed in Fig. 6.1(a), the adsorption energy U changes almost linearly with N in each block. However, computation of fluctuations may become problematic around the inflection points, where the slope $(\partial U/\partial N)$ changes due to contributions from adsorption at a new adsorption site [354], since the total energy U is calculated and not the energy for each adsorption site. It is possible to approach the inflection points from both sides by fitting the simulation data on every side using linear regression. This is shown in Fig. 6.1(b). The slope $(\partial U/\partial N)_{V,T}$ is obtained both from linear regression of raw simulation data (from independent NVT simulations) using Eq. 6.2, and thermodynamic fluctuations in the grand-canonical ensemble using Eqs. 5.11 and 5.12. It can be clearly seen that the slope computed using thermodynamic fluctuations show scatter around the inflection point. Sample snapshots of distribution of 2-methylpentane in the zeolite channels are shown in Fig. 6.1(c) and Fig. 6.1(d).

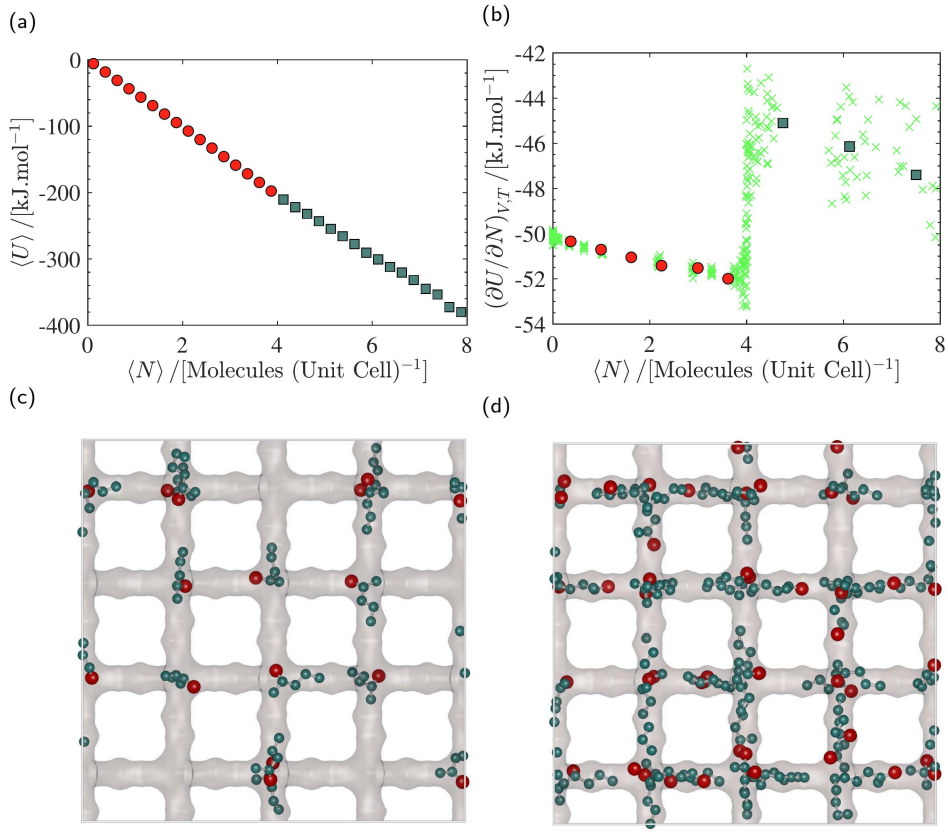


Figure 6.1: (a) The total energy as a function of loading for grand-canonical simulations of 2-methylpentane in MFI-type zeolite. For each adsorption site (block), a different symbol is used: circles for loadings smaller than 4 molecules per unit cell (block 1), and squares for higher loadings (block 2). Each block corresponds to an adsorption site in the unit cell. Due to the contribution from each adsorption site, the slope $(\partial U/\partial N)_{V,T}$ is different in each block [354]. The inflection point is at $N = 4$. Details on the simulations and the force field can be found in Ref. [358]; (b) the slope $(\partial U/\partial N)_{V,T}$ corresponding to each block is computed using linear regression of raw data obtained from independent simulations in the *NVT* ensemble (circles and squares), using Eq. 6.2, and thermodynamic fluctuations in the grand-canonical ensemble (crosses), using Eqs. 5.11 and 5.12; (c) snapshot of the MFI-type zeolite with 2-methylpentane at a loading corresponding to block 1 ; (d) snapshot of the MFI-type zeolite with 2-methylpentane at a loading corresponding to block 2.

6.3 Application to Constant Pressure Ensembles

It is not practical to perform simulations in which only the pressure, temperature and the chemical potential are fixed (μPT ensemble) [14], as the pressure, temperature, and chemical potential are all intensive variables. This ensemble can be used although its size will increase or decrease in the simulation [360]. To have an upper bound on the system size, at least one extensive variable should be fixed (*e.g.* in the grand-canonical ensemble, the volume is fixed). To constrain the system size at constant pressure, there are several possibilities: *(1)* Performing simulation in the NPT version of the Gibbs Ensemble (where the total number of molecules is fixed) [14, 44]; *(2)* Performing simulations in the NPT version of the reaction ensemble (total number of atoms is fixed) [142, 177, 321, 328]. The reaction ensemble can also be considered a grand-canonical ensemble in which the chemical potentials of reactants and reaction products are imposed in such a way that chemical equilibrium is obtained [142, 321, 328]. To calculate the partial derivatives of thermodynamic properties at constant pressure, one can fit the multiple linear regression model (Eq. 6.2) to the simulation data [152]. Since the partial derivatives in Eqs. 6.6 and 6.8 do not change when the pressure is kept constant instead of the volume, the resulting expressions from fluctuations (Eq. 6.4) do not change either. At constant pressure, it is possible to show that the intercept from the linear regression model will be zero [152]. As an example, one can write the instantaneous total volume as the sum of the composition-weighted partial molar volumes (k component types) [152]

$$V = N_1\bar{v_1} + N_2\bar{v_2}\cdots + N_k\bar{v_k} \tag{6.19}$$

in which $\bar{v_i}$ is the partial molar volume of component i. By fitting the least squares multiple linear regression model, ($V = N_0 + N_1\bar{v_1} + N_2\bar{v_2} + \cdots + N_k\bar{v_k} + \epsilon$) in which ϵ is the noise, the expression for the intercept follows from Eq. 6.3:

$$N_0 = \langle V\rangle - \bar{v_1}\langle N_1\rangle - \bar{v_2}\langle N_2\rangle - \cdots - \bar{v_k}\langle N_k\rangle \tag{6.20}$$

From Eqs. 6.19 and 6.20 it follows directly that the intercept N_0 is zero. For application of linear regression in the reaction ensemble, the reader is refereed to Section 5.6, or Ref. [152].

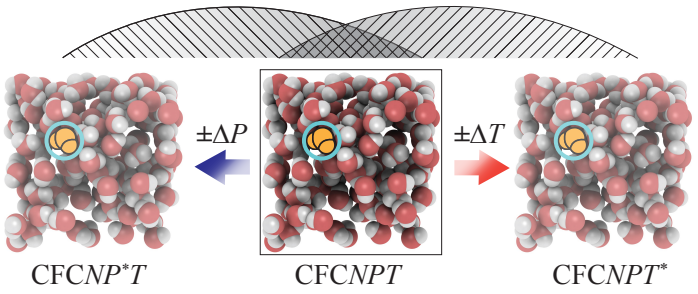
6.4 Conclusions

We have shown that in the grand-canonical ensemble, expressions for thermodynamic derivatives obtained from least squares multiple linear regression are identical to the expressions obtained from fluctuations. This provides a conceptually simple and computationally efficient approach to obtain thermodynamic properties from fluctuations in multicomponent systems. Identical expressions are only obtained only if higher order terms are not included in the linear regression model. Multiple linear regression is thermodynamically consistent with fluctuations both in constant-volume and constant-pressure ensembles. In the grand-canonical ensemble, multiple linear regression can be used to obtain the heat of adsorption even around sharp inflection points where the fluctuation approach is known to fail. We also show that the reaction enthalpy can be obtained directly from a single simulation in the reaction ensemble (see chapter 5), by fitting the enthalpy as a function of the number of reactant molecules with simple linear regression.

Chapter 7

Multiple Free Energy Calculations from a Single CFCMC Simulation Using Umbrella Sampling

This chapter is based on the following paper: Rahbari, A.; Hens, R.; Moultos, O.A.; Dubbeldam, D.; Vlugt, T. J. H., Multiple free energy calculations from single state point Continuous Fractional Component Monte Carlo simulation using umbrella sampling, Journal of Computational Theory and Computation, 2020, in press. DOI:10.1021/acs.jctc.9b01097



7.1 Introduction

In chapter 5, we introduced an alternative method to obtain partial molar properties in the CFC*NPT* ensemble without the drawbacks of the WTPI method. In this chapter, we introduce an alternative method to perform free energy calculations for mixtures at multiple temperatures and pressures from a single simulation, by combining umbrella sampling [14, 15, 322] and the CFCMC method. One can perform a simulation of a mixture at a certain pressure and temperature, and accurately compute the chemical potential at other pressures and temperatures close to the simulation conditions. This method has the following advantages: *(1)* Accurate estimates of the chemical potential as a function of pressure and temperature are obtained from a single state simulation without additional post-processing. This can potentially reduce the number of simulations of a system for free energy calculations for a specific temperature and/or pressure range. *(2)* Partial molar volumes and enthalpies are obtained directly from the estimated chemical potentials.

Umbrella sampling is a well-known method developed by Torrie et al. [322] from which the free energy difference between the system of interest and a reference system can be obtained [14, 17, 322]. By introducing a biasing function $\mathcal{W}\left(r^N\right)$, the ensemble average $\langle A \rangle$ in the canonical ensemble is calculated from an ensemble where the configuration space is sampled proportional to $\Pi\left(r^N\right) \propto \exp\left[-\beta U\left(r^N\right) + \mathcal{W}\left(r^N\right)\right]$:

$$\langle A \rangle = \frac{\left\langle A \exp\left[-\mathcal{W}\left(r^N\right)\right]\right\rangle_{\Pi}}{\left\langle \exp\left[-\mathcal{W}\left(r^N\right)\right]\right\rangle_{\Pi}} \tag{7.1}$$

in which $\langle \cdots \rangle_{\Pi}$ denotes an ensemble average in the ensemble Π. The biasing function $\mathcal{W}\left(r^N\right)$ is only a function of the configuration space. The derivation is provided in Ref. [17]. In this chapter, we show that by combining Eq. 7.1 with CFCMC, the chemical potentials can be accurately estimated for an appreciable temperature and pressure range from a single state simulation in the CFC*NPT* ensemble. To estimate the chemical potential at a temperature or pressure different than that of the simula-

tion, sufficient overlap is required between the configuration space of the two systems [14, 17]. This depends also on the system size. The relative density and energy fluctuations become smaller with the increase in system size, which may reduce the overlap between the configuration spaces of two systems. Therefore, we investigated the limitations of combining umbrella sampling with the CFCMC method when estimating chemical potentials for various temperature and pressure intervals. The combination of umbrella sampling with CFCMC method offers the following advantages: *(1)* Accurate estimates of the chemical potential as a function of pressure or temperature are computed from a single simulation. *(2)* Partial molar properties are obtained directly from a single simulation. By definition, partial molar properties can be obtained by numerically evaluating the expressions $\bar{h} = (\partial(\beta\mu)/\partial\beta)_P$ and $\bar{v} = (\partial\mu/\partial P)_T$, using the estimated chemical potentials at different pressures and temperatures [8, 20]. *(3)* Partial molar enthalpies obtained from umbrella sampling can be used as an independent check for the CFCMC method in Eqs. 5.5 and 5.6, both for programming bugs or an independent check whether the phase space is sufficiently sampled.

In Section 7.2, the combination of umbrella sampling and free energy calculations in CFCMC simulations is explained. In Section 7.3, we provide a detailed overview of the systems considered in our simulations. In Section 7.4, our simulation results are presented. In this section, the chemical potentials for a binary LJ color mixture are estimated at multiple temperatures and pressures by performing single state simulations in the CFCNPT ensemble. The results from umbrella sampling are in excellent agreement with the ones obtained from independent simulations. In addition, partial molar properties are estimated by numerically evaluating $\bar{h} = (\partial(\beta\mu)/\partial\beta)_P$ and $\bar{v} = (\partial\mu/\partial P)_T$, based on results from umbrella sampling. The computed partial molar properties for the LJ mixture are in excellent agreement with the results from the CFCMC method (Eqs. 5.5 and 5.6) and the WTPI method (Eqs. 5.3 and 5.4). As an example of a molecular system, we applied our method to different mixtures of water and methanol at standard conditions ($T = 298$ K and $P = 1$ bar) and compared the results with those obtained from independent CFCNPT simulations. Accurate estimates of

the chemical potential of water and methanol are obtained using umbrella sampling for a temperature difference $\Delta T = \pm 10$ K, for $N = 470$ molecules. Excellent agreement is observed between partial molar properties of water and methanol obtained from umbrella sampling and those obtained using Eqs. 5.5 and 5.6. We also applied our method to a mixture of ammonia, nitrogen, and hydrogen at $T = 573$ K and $P = 800$ bar at chemical equilibrium, and compared the partial molar properties obtained from our method to those obtained from the CFCMC method. The limitations of the method are tested for by decreasing overlap in configuration space for pure methanol. In Section 7.5, our conclusions are presented.

7.2 Theory

Similar to Eq. 7.1, we combine umbrella sampling with the CFCMC method, to estimate the probability distribution $p(\lambda)$ at (T, $P^\star$) while performing a simulation at ($T^\star$, $P^\star$) in the CFCNPT ensemble. This distribution is calculated using:

$$p\left(\lambda\right)|_\beta = c \cdot \left\langle \delta(\lambda^{'} - \lambda) \exp\left[(\beta^\star - \beta)\, H\right] \right\rangle_{\beta^\star} \tag{7.2}$$

in which $H = U + PV$, U is the internal energy of the system and c is a normalization constant. In the ensemble at $\beta^\star$, the value of the scaling parameter equals $\lambda^{'}$. In appendix A.8, one can find the derivation of Eq. 7.2. To estimate the excess part of the chemical potential in Eq. 2.9 at β, the values $p(\lambda = 1)|_\beta$ and $p(\lambda = 0)|_\beta$ are obtained using Eq. 7.2. Similar to Eq. 7.2, it can be shown that the number density can be calculated at T^* using

$$\rho_N|_\beta = \frac{\langle \rho_N \exp\left[(\beta^\star - \beta)\, H\right] \rangle_{\beta^\star}}{\langle \exp\left[(\beta^\star - \beta)\, H\right] \rangle_{\beta^\star}} \tag{7.3}$$

in which $\rho_N|_\beta$ is the number density at β. The number of fractional molecules is not included in computing number densities [142]. If the number of fractional molecules are much smaller compared to the whole molecules (less than 1%), Boltzmann averages are not affected [41]. By com-

bining Eqs. 2.9, 7.2 and 7.3, the chemical potential at $(T, P^\star)$ is estimated while performing a single simulation at $(T^\star, P^\star)$. In a similar manner, the distribution $p(\lambda)$ at $(T^\star, P)$ can be estimated while performing a simulation at $(T^\star, P^\star)$, keeping the number of molecules constant. This distribution is obtained by introducing a bias to the ensemble average:

$$p\left(\lambda\right)|_P = c \cdot \left\langle \delta(\lambda^{'} - \lambda) \exp\left[\beta V \left(P^\star - P\right)\right] \right\rangle_{P^\star} \tag{7.4}$$

The derivation of Eq. 7.4 can be found in appendix A.8. By computing the values $p\left(\lambda = 0\right)|_P$ and $p\left(\lambda = 1\right)|_P$ based on Eq. 7.4, one can obtain the excess chemical potential at pressure P while performing a simulation at pressure $P^\star$. For the ideal part of the chemical potential in Eq. 2.9, the number density is estimated at P using umbrella sampling:

$$\rho_N|_P = \frac{\langle \rho_N \exp\left[\beta V \left(P^\star - P\right)\right] \rangle_{P^\star}}{\langle \exp\left[\beta V \left(P^\star - P\right)\right] \rangle_{P^\star}} \tag{7.5}$$

By combining Eqs. 2.9, 7.4 and 7.5, the chemical potential at $(T^\star, P)$ is obtained while performing a single state simulation at $(T^\star, P^\star)$. The derivation of Eq. 7.5 is very similar to derivation of Eq. 7.4.

7.3 Simulation Details

To demonstrate the feasibility of our method, MC simulations were carried out to compute the chemical potentials of a LJ color mixture (50%-50%) in the CFCNPT ensemble, containing $N = 200$ molecules. All simulations were carried out at $T^* = 2$ and pressures between $P^* = 0.1$ and $P^* = 8$ (the symbol * denotes reduced units). For the LJ system, Eqs. 7.2 and 7.3 were used to compute the chemical potential at temperatures between $T^* = 1.82$ and $T^* = 2.22$ and a pressure of $P^* = 6$ from a single simulation at $T^* = 2,\ P^* = 6$. Similarly, Eqs. 7.4 and 7.5 were used to compute the chemical potentials at $T^* = 2$ and pressures between $P^* = 5.95$ and $P^* = 6.05$ in the same simulation. To compute the reference values for the chemical potentials, independent simulations were performed for each temperature

and pressure. The estimated chemical potentials obtained from umbrella sampling were used to numerically evaluate the partial molar properties using the expressions $\bar{h} = (\partial(\beta\mu)/\partial\beta)_P$ and $\bar{v} = (\partial\mu/\partial P)_T$. As a reference, the partial molar properties were also computed using the WTPI method using Eqs. 5.3 and 5.4, and the CFCMC method using Eqs. 5.5 and 5.6. Simulation details for the LJ system are provided in chapter 5.

To test this method for a more complex system, aqueous methanol mixtures with mole fractions of methanol between $x_{\text{MeOH}} = 0.2$ and $x_{\text{MeOH}} = 0.8$ were simulated in the CFCNPT ensemble. For water, the TIP4P/2005 [255] and for methanol the TraPPE [207] force fields were used. In all simulations of water-methanol mixtures, a fractional molecule of methanol and a fractional molecule of water were present. In chapter 3, we showed that the correlation between scaling parameters of different fractional molecules is weak and independent of the weight function. It is therefore computationally advantageous to separate the multidimensional weight function into a series of one dimensional weight functions: $W(\lambda_1, \lambda_2) \approx W(\lambda_1) + W(\lambda_2)$. This is due to the fact that filling one-dimensional histograms is more efficient than filling two separate multidimensional histograms [41]. Chemical potentials of water and methanol between $T = 288$ K and $T = 308$ K were estimated by performing umbrella sampling from a single simulation of water-methanol mixture at $T = 298$ K and $P = 1$ bar. As a reference, the chemical potentials of water and methanol were computed independently at every pressure and temperature. Subsequently, partial molar properties of water and methanol were obtained from the estimated chemical potentials using umbrella sampling. The results were compared to the partial molar properties obtained using Eqs. 5.5 and 5.6. To investigate the limitations of umbrella sampling in CFCMC simulations, umbrella sampling was applied to estimate the chemical potentials of pure methanol for a wide pressure and temperature range. The results were compared to results obtained from independent simulations of pure methanol in the CFCNPT ensemble, at every temperature and pressure. Simulation details for water-methanol mixtures are provided in chapter 4.

To compute partial molar properties of ammonia, nitrogen and hydrogen in their mixture at $T = 573$ K and $P = 800$ bar, simulations were performed in the CFCNPT ensemble. Similar to chapter 5, separate simulations were performed in which only a single fractional molecule was present. This was repeated for the three components, leading to three independent simulations. For simulation details, the reader is referred to chapter 5. The starting mixture composition ($N_{NH_3} = 407$, $N_{N_2} = 7$, and $N_{H_2} = 20$) was obtained from simulations of the Haber-Bosch process in the reaction ensemble as described in Ref. [142].

7.4 Results and Discussion

A simulation in the CFCNPT ensemble was performed for a LJ system at $T^* = 2$ and $P^* = 6$. Using umbrella sampling, the chemical potentials were computed for temperatures ranging from $T^* = 1.82$ to $T^* = 2.22$ and pressures ranging from $P^* = 5.95$ to $P^* = 6.05$. The values of the chemical potential were compared to those obtained from independent CFCNPT simulations for all temperatures and pressures. The results are shown in Table 7.1. In Table 7.1(a), excellent agreement is observed between the computed chemical potentials obtained using umbrella sampling from a single simulation at ($T^* = 2,\ P^* = 6$) and the chemical potentials calculated from independent simulations at $P^* = 6$ and temperatures between $T^* = 1.82$ and $T^* = 2.22$. In this work, the error bars (uncertainties) are obtained by computing the standard deviation of the mean from five independent simulations. It should be noted that the uncertainty of the computed chemical potential associated with umbrella sampling increases with the increase in the temperature difference $|\Delta T^*|$. Since the overlap between the configuration spaces decreases with the increase in temperature or pressure difference, it impairs the sampling. The overlap between the configuration spaces also decreases with the increase in system size, as the relative fluctuations of energy and volume become smaller for larger systems. In Table 7.1(b), the estimated chemical potentials from umbrella sampling from a single simulation at ($T^* = 2,\ P^* = 6$) are compared to the chemical

Table 7.1: Chemical potentials for a binary (50%-50%) LJ mixture obtained from umbrella sampling (Eqs. 7.2 and 7.4) and independent simulations in the CFC*NPT* ensemble (Eq. 2.9). To obtain estimates of the chemical potentials at (a) different temperatures and (b) pressures, umbrella sampling is performed at $T^* = 2$ and $P^* = 6$ in a single simulation. Boltzmann averages at $T^* = 2$ and $P^* = 6$ are highlighted in gray. Numbers in brackets indicate uncertainties.

(a)		Umbrella sampling at $T^* = 2$			Independent CFCNPT simulations		
T^*	P^*	μ_A^{Ig}	μ_A^{ex}	μ_A	μ_A^{Ig}	μ_A^{ex}	μ_A
1.82	6	-1.6125(6)	6.10(5)	4.48(6)	-1.6151(1)	6.09(1)	4.47(1)
1.92	6	-1.7379(3)	6.28(1)	4.54(2)	-1.7344(2)	6.27(1)	4.54(1)
2	6	-1.8316(1)	6.41(1)	4.58(1)	-	-	-
2.13	6	-1.9904(3)	6.63(1)	4.63(1)	-1.9935(3)	6.61(1)	4.62(1)
2.22	6	-2.1104(8)	6.80(3)	4.69(3)	-2.10789(2)	6.77(1)	4.66(1)

(b)		Umbrella sampling at $P^* = 6$			Independent CFCNPT simulations		
T^*	P^*	μ_A^{Ig}	μ_A^{ex}	μ_A	μ_A^{Ig}	μ_A^{ex}	μ_A
2	5.95	-1.8359(2)	6.350(7)	4.514(7)	-1.8358(3)	6.535(5)	4.517(5)
2	5.98	-1.8335(2)	6.385(6)	4.551(6)	-	-	-
2	6	-1.8318(2)	6.408(6)	4.576(6)	-	-	-
2	6.02	-1.8302(2)	6.432(6)	4.601(6)	-	-	-
2	6.05	-1.8278(2)	6.467(6)	4.639(6)	-1.8277(2)	6.47(1)	4.65(1)

potentials obtained from independent simulations at $T^* = 2$ and pressures between $P^* = 5.95$ and $P^* = 6.05$. Excellent agreement is observed between the computed chemical potentials estimated using umbrella sampling and the results obtained from independent simulations for the entire pressure range. From the results presented in Table 7.1, it can be observed that accurate estimates of the chemical potentials are obtained from a single simulation combining the CFCMC method with umbrella sampling. While the statistical uncertainties of the estimated chemical potentials obtained from umbrella sampling are larger compared to those obtained from independent simulations, the differences in uncertainties are not significant. The chemical potentials obtained from umbrella sampling are used to compute the partial molar properties for the LJ mixture, by numerically differentiating

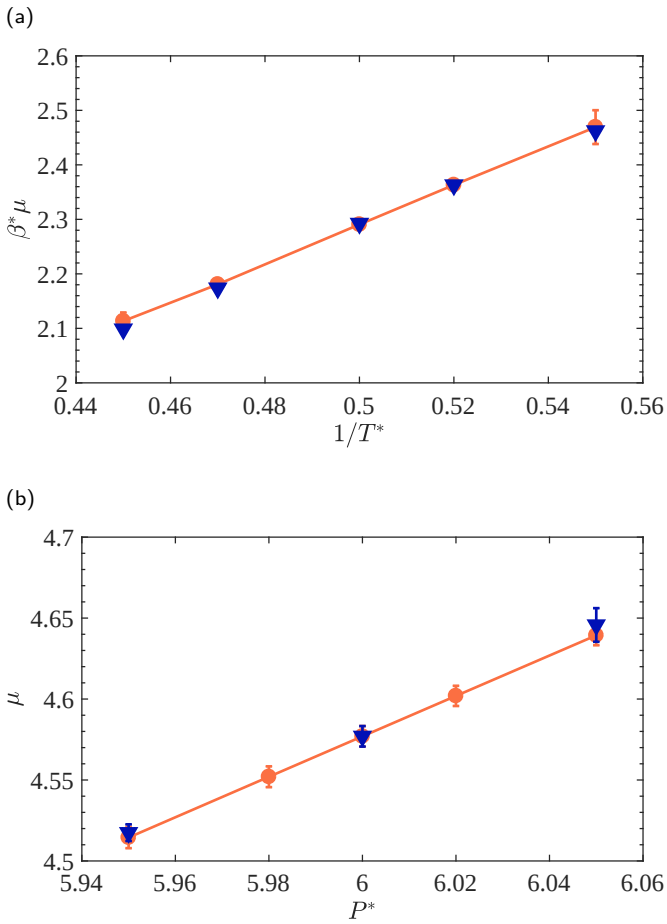


Figure 7.1: Plots showing (a) $\beta^*\mu$–β^* and (b) μ–P^* for a LJ binary color mixture consisting of 200 molecules. Downward-pointing triangles indicate results from obtained from independent CFC*NPT* simulations at pressures between $P^* = 5.94$ and $P^* = 6.05$ and temperatures between $T^* = 1.82$ and $T^* = 2.22$. Circles indicate results obtained by performing umbrella sampling from a single simulation in the CFC*NPT* ensemble at $T^* = 2$ and $P^* = 6$. Lines indicate that the data are obtained from a single simulation. Error bars are smaller than the symbol sizes. Raw data are listed in Table 7.1.

$(\partial(\beta\mu)/\partial\beta)_P$ and $(\partial\mu/\partial P)_T$. For instance, the chemical potentials in Table 7.1 are used to plot $\beta\mu$–β and μ–P for $T^* = 2$ and $P^* = 6$, as shown is in Fig. 7.1. The data points connected with a line in Fig. 7.1 indicate that the data points were obtained from a single simulation combining CFCMC and umbrella sampling, and the individual data points were obtained from independent CFCNPT simulations. Calculating the slopes in Fig. 7.1, at $T^* = 2$ and $P^* = 6$, leads to the values for the partial molar volumes and enthalpies. A central difference scheme with high order approximation $\mathcal{O}(h^4)$ was used to compute the partial molar enthalpies in Fig. 7.1(a) and partial molar volumes in Fig. 7.1(b). The partial molar excess enthalpy was computed by subtracting the ideal gas contribution from the total partial molar enthalpy. It is shown in appendix A.1 (Eq. A13) that the ideal part of the partial molar enthalpy of a LJ particle, $\bar{h}^{\text{id}}$ equals $5/(2\beta)$. It is noteworthy that the contribution of the thermal wavelength Λ in $\bar{h}^{\text{id}}$ equals $3/(2\beta)$. Since Λ is set to unity in our simulations, the contribution of the thermal wavelength cancels out when numerically evaluating $(\partial(\beta\mu)/\partial\beta)_P$. Therefore, the partial molar excess enthalpy is obtained from umbrella sampling of the chemical potential using $\bar{h}^{\text{ex}} = (\partial(\beta\mu)/\partial\beta)_P - 1/\beta$. The partial molar properties for the LJ system obtained from umbrella sampling and the CFCMC method (Eqs. 5.5 and 5.6) and the WTPI method (Eqs. 5.3 and 5.4) are provided in Table 7.2. Excellent agreement is observed between the partial molar properties from CFCNPT simulations, and the results obtained from umbrella sampling. Similar uncertainties are observed between the two methods. For the LJ system, umbrella sampling works equally well compared to the CFCMC method. In addition, accurate estimates of the chemical potential at different pressures and temperatures are obtained without any extra computational power or post processing.

Other realistic systems with complex intermolecular interactions considered here are different aqueous methanol mixtures with mole fractions of methanol between $x_{\text{MeOH}} = 0.2$ and $x_{\text{MeOH}} = 1$. These systems were simulated in the CFCNPT ensemble. For a mixture composition $x_{\text{MeOH}} = 0.8$, umbrella sampling was used to estimate the chemical potentials of water and methanol at temperatures between $T = 288$ K and $T = 308$ K while running a single CFCNPT simulation at $P = 1$ bar and $T = 298$ K. In addi-

Table 7.2: Densities, chemical potentials, partial molar excess enthalpies and volumes at $T^* = 2$ and pressures between $P^* = 0.1$ and $P^* = 8$ computed for a binary LJ color mixture (200 molecules). Three different methods are used at each P^*, in order to compare the results for the partial molar properties: the WTPI method (Eqs. 5.3 and 5.4) in the *NPT* ensemble [150, 153], the CFCMC method (in the CFC*NPT* ensemble, Eqs. 5.5 and 5.6) [20], and Umbrella sampling (Eqs. 7.2 and 7.4). The chemical potentials are calculated using Eq. 2.9 from independent simulations. Numbers in brackets indicate uncertainties.

			CFC*NPT*		Umbrella sampling		WTPI method	
P^*	$\langle\rho^*\rangle$	μ_A	$\bar{h}_A^{ex}$	$\bar{v}_A$	$\bar{h}_A^{ex}$	$\bar{v}_A$	$\bar{h}_A^{ex}$	$\bar{v}_A$
0.1	0.052	-7.460(7)	-0.44(4)	18.6(3)	-0.44(5)	19(2)	-0.361(5)	19.29(4)
2	0.584	-1.075(8)	-1.6(1)	1.72(3)	-1.6(1)	1.72(3)	-1.64(4)	1.70(1)
4	0.722	1.957(5)	-0.06(5)	1.39(1)	-0.06(5)	1.39(1)	-0.3(2)	1.36(5)
6	0.800	4.581(9)	1.7(2)	1.26(2)	1.7(2)	1.26(2)	1.6(8)	1.23(5)
8	0.856	7.001(6)	3.2(1)	1.17(1)	3.2(1)	1.17(1)	4(1)	1.2(1)

tion, independent simulations in the CFC*NPT* simulations were performed to obtain the chemical potentials as a reference. The results are shown in Table 7.3. Accurate estimates of the chemical potentials of water and methanol are obtained between $T = 288$ K and $T = 308$ K, from a single CFC*NPT* simulation $T = 298$ K ($\Delta T = \pm 10$ K). The relative differences between the estimates of the chemical potentials and the chemical potentials obtained from independent simulations are well below 1%. For each temperature, the differences between the absolute values of the chemical potentials in Table 7.3 are significantly smaller than 1 kcal/mol = 4.184 kJ/mol, which is typically considered as benchmark in computational chemistry literature [361]. The results show that umbrella sampling can provide an accurate estimate of the chemical potentials of water and methanol for an appreciable temperature range (around $\Delta T = \pm 10$ relative to the simulation temperature). The raw data in Table 7.3 were used to plot $\beta\mu$ as a function of β for the water-methanol mixture ($N = 470$), $x_{MeOH} = 0.8$, at temperatures between $T = 288$ K and $T = 308$ K. The results are shown in Fig. 7.2. The lines indicate data obtained from a single simulation. It can be seen in Fig. 7.2 that excellent agreement is observed between the

Table 7.3: Chemical potentials $\left[\text{kJ.mol}^{-1}\right]$ of water-methanol mixtures ($x_{\text{MeOH}} = 0.8$) obtained from umbrella sampling and independent simulations in the CFC*NPT* ensemble at temperatures between $T = 288$ K and $T = 308$ K. The values of the chemical potentials are estimated using umbrella sampling at $T = 298$ K and $P = 1$ bar. The corresponding Boltzmann averages are highlighted in gray. Numbers in brackets indicate uncertainties.

		Umbrella sampling at $T = 298$ K			Independent CFC*NPT* Simulations		
T		μ^{Ig}	μ^{ex}	μ	μ^{Ig}	μ^{ex}	μ
288	MeOH	-10.312(4)	-21.6(7)	-32.0(6)	-10.295(3)	-21.6(1)	-31.9(1)
	H_2O	-13.635(4)	-27.7(7)	-41.4(7)	-13.615(3)	-28.1(2)	-41.7(2)
290	MeOH	-10.383(4)	-21.5(4)	-31.8351	-	-	-
	H_2O	-13.728(4)	-27.7(6)	-41.4(6)	-	-	-
292	MeOH	-10.456(3)	-21.3(1)	-31.7(3)	-10.447(3)	-21.2(1)	-31.6(1)
	H_2O	-13.822(3)	-27.6(5)	-41.4(6)	-13.813(3)	-27.8(2)	-41.6(2)
294	MeOH	-10.530(3)	-21.1(2)	-31.7(3)	-	-	-
	H_2O	-13.919(3)	-27.4(3)	-41.4(4)	-	-	-
298	MeOH	-10.682(2)	-20.9(1)	-31.5(1)	-	-	-
	H_2O	-14.117(2)	-27.2(3)	-41.3(2)	-	-	-
302	MeOH	-10.838(3)	-20.6(1)	-31.5(1)	-	-	-
	H_2O	-14.320(3)	-27.0(2)	-41.3(2)	-	-	-
304	MeOH	-10.918(3)	-20.5(1)	-31.4(1)	-10.915(1)	-20.6(1)	-31.5(1)
	H_2O	-14.423(3)	-26.9(3)	-41.3(3)	-14.418(1)	-26.6(2)	-41.1(2)
306	MeOH	-10.998(4)	-20.4(2)	-31.3(2)	-	-	-
	H_2O	-14.527(4)	-26.8(4)	-41.3(4)	-	-	-
308	MeOH	-11.078(4)	-20.2(2)	-31.3(2)	-11.069(5)	-20.4(1)	-31.4(1)
	H_2O	-14.632(4)	-26.6(5)	-41.2(5)	-14.619(5)	-26.5(1)	-41.1(1)

results obtained from umbrella sampling and the reference values from independent simulations in the CFC*NPT* at every temperature. The error bars associated with the estimated chemical potentials for water and methanol increase with increasing in $|\Delta T|$ relative to $T = 298$ K. This indicates that umbrella sampling becomes inefficient and inaccurate for a larger temperature difference. It should be emphasized that this temperature range is not a priori known and is system and system size dependent. it is explained later in this chapter how a reasonable temperature or pressure range can be selected for umbrella sampling of the chemical potentials. The $\beta\mu$–μ and

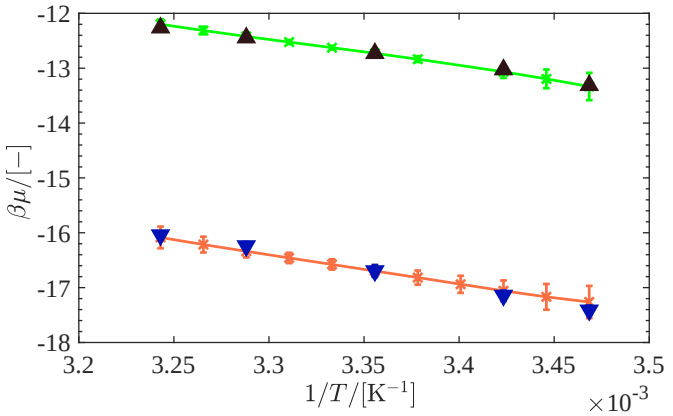


Figure 7.2: Plot showing $\beta\mu$–β for a water-methanol mixture, $x_{\mathrm{MeOH}} = 0.8$. To compute the chemical potentials, independent simulations are performed at temperatures between $T = 288$ K and $T = 308$ K. Downward-pointing triangles and upward-pointing triangles denote data for water and methanol from independent CFC*NPT* simulations. Asterisks connected by line, indicate data obtained using umbrella sampling from a single simulation at $T = 298$ K and $P = 1$ bar. Error bars are smaller than the symbol sizes. Raw data are listed in Table 7.3.

μ–P plots for water-methanol mixture at $x_{\mathrm{MeOH}} = 0.8$ are shown in Fig. 7.2. The lines indicate the data points are obtained from a single simulation. To compute the partial molar properties of water and methanol at $T = 298$ K, a central difference scheme with high order approximation $\mathcal{O}(h^4)$ is used to numerically evaluate $(\partial(\beta\mu)/\partial\beta)_P$ and $(\partial\mu/\partial P)_T$. For each component type, the partial molar excess enthalpy of water and methanol is then computed by subtracting the ideal gas part. For different mixture compositions, partial molar enthalpies of methanol and water are computed using umbrella sampling and used as reference. The results are compared to those obtained from the CFCMC method. The results are shown in Table 7.4 for the mixture with mole fractions of methanol between $x_{\mathrm{MeOH}} = 0.2$ and $x_{\mathrm{MeOH}} = 1$. Since the density of methanol mixtures at $T = 298$ K are high, the WTPI method is not used here. Excellent agreement is found between both methods. It is observed both for the LJ system and water-methanol mixture that the accuracy and uncertainty of the partial molar properties

Table 7.4: Chemical potentials $\left[\text{kJ.mol}^{-1}\right]$, partial molar excess enthalpies $\left[\text{kJ.mol}^{-1}\right]$ and partial molar volumes $\left[\text{cm}^3\text{.mol}^{-1}\right]$ for water-methanol mixtures at different compositions at $T = 298$ K and $P = 1$ bar, using simulations in the CFC*NPT* ensemble (Eqs. 5.5 and 5.6) and umbrella sampling (Eqs. 7.2 and 7.4). Umbrella sampling is performed at $T = 298$ K and $P = 1$ bar for all mixture compositions. x_i denotes the mole fraction of component i. The chemical potentials are calculated from Eq. 2.9 using independent simulations. Numbers in brackets indicate uncertainties.

				CFC*NPT*		Umbrella sampling	
i	x_i	ρ_i	μ_i^{tot}	$\bar{h}_i^{\text{ex}}$	$\bar{v}_i$	$\bar{h}_i^{\text{ex}}$	$\bar{v}_i$
MeOH	0.2	289.4(3)	-34.0(1)	-40(10)	40(10)	-40(10)	40(10)
H_2O	0.8	650.9(3)	-38.9(1)	-47(7)	17(14)	-47(7)	17(14)
MeOH	0.4	485.7(6)	-32.7(2)	-42(9)	39(9)	-42(8)	39(9)
H_2O	0.6	409.6(5)	-39.3(1)	-47(7)	16(5)	-47(7)	16(5)
MeOH	0.6	620.7(4)	-32.2(1)	-41(5)	40(5)	-41(5)	40(5)
water	0.4	233.0(2)	-39.7(1)	-49(9)	11(7)	-49(9)	11(7)
MeOH	0.8	713.7(6)	-31.5(1)	-41(6)	45(5)	-41(6)	45(5)
H_2O	0.2	100.3(1)	-41.4(3)	-47(12)	21(16)	-47(12)	21(16)
MeOH	1	777.0(9)	-31.1(1)	-42(3)	37(5)	-42(3)	37(5)

from umbrella sampling and the CFCMC method are similar. It is also observed that the uncertainties associated with partial molar properties are an order of magnitude larger compared to the uncertainties of the chemical potentials. This is due to the potential drawback of Eqs. 5.5 and 5.6 when subtracting two large numbers with a (relatively) small difference [20]. This may induce larger error bars compared to the chemical potential. A similar potential drawback is observed when using umbrella sampling to compute partial molar properties. Numerically computing the derivatives $(\partial(\beta\mu)/\partial\beta)_P$ and $(\partial\mu/\partial P)_T$ also involves subtracting two numbers (or several numbers) over a relatively small temperature or pressure difference. Therefore, accurate estimates of the chemical potential are needed to obtain accurate values for the partial molar properties. Based on the results, it is clear that the overall performance of both methods is very similar in terms of accuracy and precision.

To investigate the limitations of umbrella sampling in free energy calculations, a wide temperature and pressure range is selected for estimating the chemical potential of pure methanol from a single CFCNPT simulation. The chemical potentials of pure methanol in the temperature range of $T = 266$ K and $T = 340$ K, and pressure range of $P = 1$ bar to $P = 1001$ bar were computed using Eqs. 7.2 to 7.5 from a single simulation at $T = 298$ K and $P = 1$ bar. It should be noted that a wide temperature and pressure range is only selected to test the limitations of the method. The temperature and pressure range should be selected to ensure sufficient overlap in configuration space and energy between the systems. As a reference, independent CFCNPT simulations of methanol were performed at each temperature and pressure to compute the corresponding chemical potentials. The results are shown in Table 7.5. It can be seen that the estimated chemical potentials for $\Delta T = \pm 15$ K, relative to $T = 298$ K, are in excellent agreement with the chemical potentials obtained from independent simulations. It is clear that the sampling becomes more difficult when the temperature difference increases. The uncertainties of the estimated chemical potentials is an order of magnitude higher for $|\Delta T| > 15$ K. This sampling difficulty can be explained based on the overlap between the energy or configuration space of a system at two different temperatures, or pressures. In Fig. 7.3(a), the probability distribution of enthalpy per molecule $p\,(h)$ for pure methanol at different temperatures is shown ($h = H/N$, $N = 410$). For the distribution $p\,(h)$ at $T = 266$ K, no overlap is observed with the distribution $p\,(h)$ at $T = 298$ K, and the method fails to estimate the chemical potential at $T = 266$ K, as can be seen in Table 7.5(a). At $T = 283$ K, sufficient overlap is observed with the distribution $p\,(h)$ at $T = 298$ K, and the estimated chemical potential is in excellent agreement with the reference value. It should be noted that the uncertainties of the chemical potentials from independent simulations are always smaller. However, for $\Delta T = \pm 15$ K, the uncertainties of the estimated chemical potentials is small. In Fig. 7.3(a), it is also shown that the overlaps between distributions $p\,(h)$ for $T = 298$ K and $T = 320$ K become very small. This leads to large uncertainties in the estimated values for the chemical potentials. It is shown in Fig. 7.3(b) that for $P < 100$ bar the estimated chemical potentials of methanol are in excellent agreement

Table 7.5: Chemical potentials $\left[\text{kJ.mol}^{-1}\right]$ of pure methanol obtained from umbrella sampling (Eqs. 7.2 to 7.5) and independent simulations in the CFCNPT ensemble (Eq. 2.9). Umbrella sampling is performed in the CFCNPT ensemble of methanol at $T = 298$ K and $P = 1$ bar. Boltzmann averages obtained from umbrella sampling are highlighted in gray. Numbers in brackets indicate uncertainties.

(a)		Umbrella sampling			Independent CFCNPT simulations		
T/K	P/bar	μ_A^{Ig}	μ_A^{ex}	μ_A	μ_A^{Ig}	μ_A^{ex}	μ_A
340	1	N/A	N/A	N/A	-12.115(1)	-18.24(3)	-30.36(3)
320	1	-11.31(3)	-20(3)	-31(3)	-11.322(2)	-19.35(7)	-30.67(7)
314	1	-11.10(2)	-20(1)	-31(1)	-	-	-
308	1	-10.880(7)	-20.1(2)	-31.0(2)	-10.854(2)	-20.01(9)	-30.86(9)
298	1	-10.46(2)	-20.63(8)	-31.10(8)	-	-	-
288	1	-10.087(4)	-21.3(4)	-31.4(4)	-	-	-
283	1	-9.906(8)	-21(1)	-31(1)	-9.895(1)	-21.44(9)	-31.33(9)
279	1	-9.74(1)	-21(2)	-31(2)	-	-	-
274	1	-9.58(1)	-21(4)	-31(4)	-9.555(1)	-22.1(1)	-31.6(1)
266	1	N/A	N/A	N/A	-9.254(2)	-22.5(1)	-31.8(1)

(b)		Umbrella sampling			Independent CFCNPT simulations		
T/K	P/bar	μ_A^{Ig}	μ_A^{ex}	μ_A	μ_A^{Ig}	μ_A^{ex}	μ_A
298	1	-10.468(2)	-20.63(8)	-31.10(8)	-	-	-
298	11	-10.465(2)	-20.59(8)	-31.06(8)	-10.464(2)	-20.57(8)	-31.04(8)
298	21	-10.462(2)	-20.55(9)	-31.02(9)	-	-	-
298	31	-10.459(2)	-20.52(9)	-30.98(9)	-10.459(2)	-20.45(3)	-30.91(3)
298	41	-10.456(2)	-20.48(9)	-30.94(9)	-	-	-
298	51	-10.453(2)	-20.44(9)	-30.89(9)	-	-	-
298	61	-10.450(2)	-20.40(9)	-30.85(9)	-10.450(1)	-20.32(9)	-30.78(9)
298	101	-10.438(2)	-20.3(1)	-30.7(1)	-10.439(2)	-20.16(8)	-30.61(8)
298	201	-10.410(2)	-19.8(2)	-30.2(2)	-	-	-
298	501	-10.361(7)	-18(2)	-28(2)	-10.343(2)	-18.75(8)	-29.09(8)
298	801	-10.342(9)	-15(4)	-25(4)	-	-	-
298	1001	-10.33(2)	-12(6)	-22(6)	-10.469(2)	-20.63(8)	-31.10(8)

with the chemical potentials obtained from independent simulations. For pressures ranging between 100 and 500 bar, the uncertainties of the results from umbrella sampling are up to two orders of magnitude larger than those obtained from independent simulations. The estimated values of the chemical potentials are however within chemical accuracy (1 kcal/mol) [361].

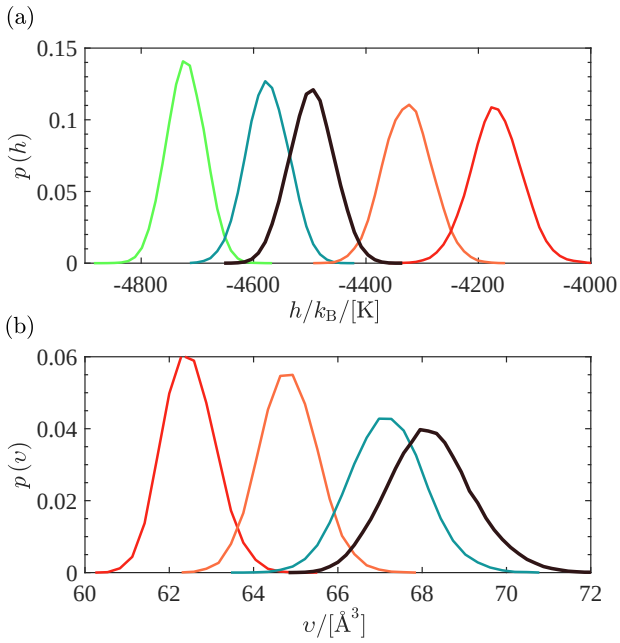


Figure 7.3: (a) Probability distribution of the enthalpy per molecule of methanol; $h = H/N$ at $P = 1$ bar and different temperatures: $T = 266$ K (green), $T = 285$ K (teal), $T = 298$ K (black), $T = 320$ K (orange), $T = 340$ K (red). (b) Probability distribution of the volume per molecule of methanol; $v = V/N$ at $T = 298$ K and different pressures: $P = 1$ bar (black), $P = 100$ bar (teal), $P = 500$ bar (orange), $P = 1000$ bar (red). The number of molecules for pure methanol in the liquid phase is $N = 410$.

In Fig. 7.3(b), the probability distributions of volume per molecule $p(v)$ for pure methanol at different pressures are shown ($v = V/N$, $N = 410$). Sufficient overlap is observed in Fig. 7.3(b) for the distributions $p(v)$ at $P = 1$ bar and $P = 100$ bar. This is expected since the compressibility of liquid methanol is very low at room temperature. Therefore, it is possible to estimate the chemical potentials of methanol accurately for any pressure between $P = 1$ bar and $P = 100$ bar, see Table 7.5(b) and Fig. 7.3(b). For $P > 500$ bar, the uncertainty of the estimated chemical potentials increases significantly (three orders of magnitude larger compared to independent

simulations) and the method starts to break down. This can be explained by examining the overlap between the distributions $p(\upsilon)$ at high pressures and $P = 1$ bar. In Fig. 7.3(b), no significant overlap is observed between the distribution $p(\upsilon)$ at $P = 1$ and the distributions at $P > 500$. At these conditions, the excess chemical potentials of methanol computed from umbrella sampling deviate significantly from the excess chemical potentials computed from independent simulations. Here, we illustrate the sampling issue of the excess chemical potential at high pressures by plotting $p(\lambda)$ as a function of pressure, computed from umbrella sampling at $T = 298$ K and $P = 1$ bar. For pure methanol ($N = 410$), it is clearly observed in Fig. 7.4 that the computed $p(\lambda)$ shows scatter for pressures significantly different than that of the simulation. This clearly illustrates that the method breaks down when the pressure difference becomes large, and the statistics for the excess chemical potential become very poor. We observed that for liquid methanol ($N = 410$), the computed chemical potentials from umbrella sampling are in excellent agreement with the reference values obtained from independent simulations, for $\Delta T = \pm 15$ K and $\Delta P = 100$ bar. The pressure and temperature range for accurate estimation of the chemical potentials, or other thermodynamic properties may differ from one system to another. However, plotting distributions $p(h)$ and $p(\upsilon)$ for a wide pressure and temperature range can readily visualize at what range umbrella sampling can be applied to obtain accurate results. As shown in Fig. 7.3, an investigation of the overlap between energies, volumes at different temperatures and pressures can easily indicate the boundaries at which the method starts to fail. For systems that lack sufficient overlap (i.e. $\Delta\beta$ or ΔP is too large), it is expected that the performance of the method is poor. If one is interested in nonlinear variations of μ (e.g. higher order derivatives of μ), our method may not work well due to a lack of overlap. To compute partial molar properties, selecting a narrow region (a small temperature or pressure range) still allows for a numerical evaluation of $\bar{h} = (\partial(\beta\mu)/\partial\beta)_P$ and $\bar{\upsilon} = (\partial\mu/\partial P)_T$. Therefore, partial molar properties can be computed from umbrella sampling without selecting a wide temperature or pressure range.

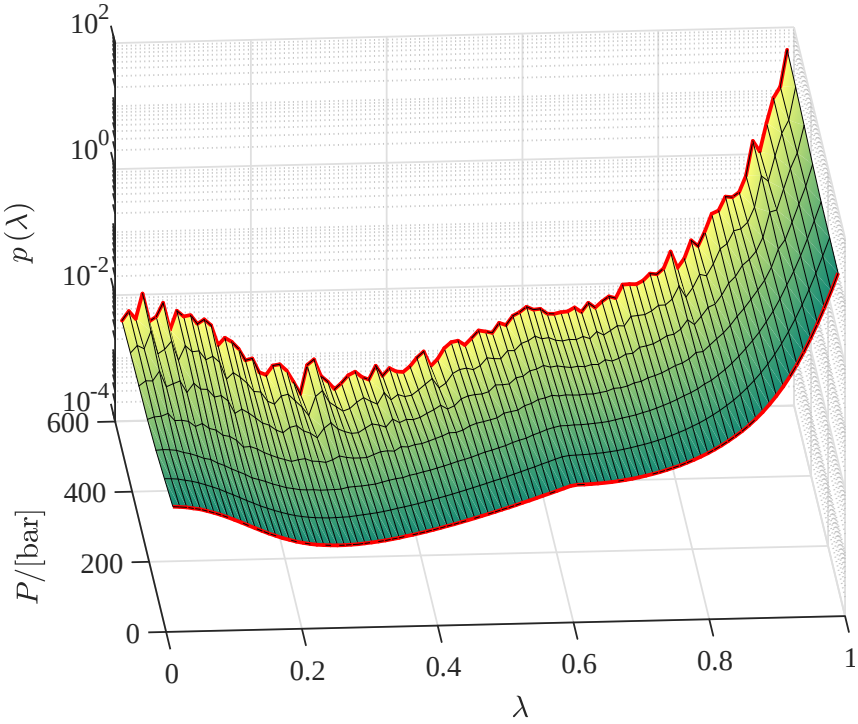


Figure 7.4: Probability distribution of λ for pure methanol obtained from a CFC*NPT* simulation at $T = 298$ K and $P = 1$ bar. Eq. 7.4 is used to compute $p(\lambda)$ for pressures up to $P = 600$ bar from a single simulation. The red lines indicate the Boltzmann distribution $p(\lambda)$ at $P = 1$ bar, and the distribution $p(\lambda)$ computed for $P = 600$ bar.

As an example of a strong hydrogen bond forming system, we considered the Haber-Bosch process ($N_2 + 3\,H_2 \rightleftharpoons 2\,NH_3$). Umbrella sampling was used to compute partial molar properties of the mixture at $T = 573$ K and $P = 800$ bar, at chemical equilibrium. The same equilibrium composition of the mixture is used as in chapter 5. Independent simulations are performed at $T = 573$ K and $P = 800$ bar in the CFC*NPT* ensemble. In each simulation, the fractional molecule of only one component is present. Partial molar properties obtained from umbrella sampling (Eqs. 7.2 and 7.4) and

Table 7.6: Partial molar enthalpies $\left[\text{kJ.mol}^{-1}\right]$ and partial molar volumes $\left[\text{cm}^3\text{.mol}^{-1}\right]$ of ammonia, nitrogen and hydrogen at $T = 573$ K and $P = 800$ bar. The composition of the mixture is obtained from simulations of ammonia synthesis in the reaction ensemble, as explained in our earlier work [142]. Partial molar properties are computed in the CFC*NPT* ensemble using Eqs. 5.5 and 5.6, and umbrella sampling using Eqs. 7.2 and 7.4. Numbers in brackets indicate uncertainties.

	CFC*NPT*		Umbrella Sampling	
	$\bar{h}$	$\bar{v}$	$\bar{h}$	$\bar{v}$
NH_3	-10(1)	50(7)	-10(1)	50(7)
N_2	11(1)	108(6)	11(1)	108(6)
H_2	12(1)	94(6)	12(1)	94(6)

the CFCMC method (Eqs. 5.5 and 5.6) are compared in Table 7.6. Excellent agreement is observed between the results obtained from umbrella sampling and the CFCMC method.

7.5 Conclusions

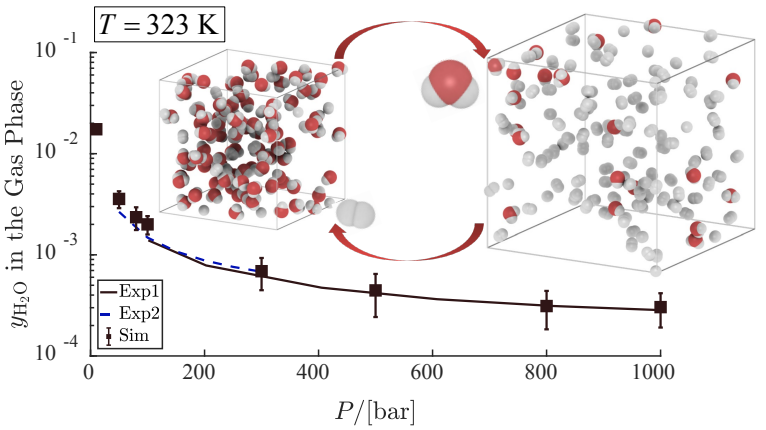
We introduced an alternative method to obtain accurate estimates of the excess chemical potential of a component for a wide temperature and pressure range from a single simulation. This method combines umbrella sampling and the CFCMC technique. Using the values of the estimated chemical potentials, the partial molar enthalpies and volumes of a component are obtained by numerically evaluating the derivatives $(\partial(\beta\mu)/\partial\beta)_P$ and $(\partial\mu/\partial P)_V$, respectively. This method does not have the disadvantages of the WTPI method. As a proof of concept, the values of the chemical potential for a binary LJ mixture were estimated using umbrella sampling from a single simulation in the CFC*NPT* ensemble at $T^* = 2$ and $P^* = 6$. For a temperature range between $T^* = 1.82$ and $T^* = 2.22$ and pressure range between $P^* = 5.95$ and $P^* = 6.05$, excellent agreement was observed between the estimated chemical potentials and those obtained from independent CFC*NPT* simulations of the LJ system, at each temperature and

pressure. Partial molar properties obtained from umbrella sampling were in excellent agreement with the CFCMC method introduced in chapter 5, and the original method of Frenkel et al. [150, 153]. We observed that the accuracy and precision of the averages obtained from the CFCMC method in chapter 5 and umbrella sampling are very similar. To apply our method to a system with complex intermolecular interactions, we considered aqueous mixtures of methanol with different compositions. We applied our method to estimate the chemical potentials of methanol and water for a temperature range between $T = 288$ K and $T = 308$ K from a single simulation at $T = 298$ K and $P = 1$ bar. Excellent agreement was found between the results obtained from umbrella sampling and those obtained from independent simulations in the CFCNPT ensemble, with relative differences well below 1%. As an example of a strong hydrogen bond forming system, our method was applied to a mixture of ammonia, nitrogen, and hydrogen at chemical equilibrium at $T = 573$ K and $P = 800$ bar. It was observed that partial molar properties of ammonia, nitrogen and hydrogen obtained from umbrella sampling and the CFCMC method are in excellent agreement. We investigated the limitations of our method for liquid methanol ($N = 410$ molecules). It was observed that for the temperature difference $\Delta T\pm = 15$ K, very accurate estimates of chemical potential at different temperatures were obtained from umbrella sampling from a single CFCMC simulation. In addition, it was found that for a pressure difference $\Delta P = 100$ bar, accurate estimates of chemical potential at different pressures were obtained from umbrella sampling from a single simulation. Lack of sufficient overlap (ΔP or $\Delta\beta$) between different states may result in a poor performance of the method. Based on the results it can be concluded that combining umbrella sampling with CFCMC provides a powerful tool for accurate estimation of chemical potentials for an appreciable temperature and pressure range and computation of partial molar properties. This can potentially reduce the number of simulations of a system for free energy calculations for a specific temperature and/or pressure range.

Chapter 8

Solubility of Water in High Pressure Hydrogen

This chapter is based on the following paper: Rahbari, A.; Brenkman, J.; Hense, R.; Ramdin, M.; van den Broeke , L. J. P.; Schoon, R.; Henkes, R.; Moultos, O. A.; Vlugt, T. J. H.; Solubility of Water in Hydrogen at High Pressure: A Molecular Simulation Study, Journal of Chemical & Engineering Data, 2019, 64, 4103-4115, Ref. [32]



8.1 Introduction

The world population is expected to grow rapidly, from 7.6 billion currently, to about 9.8 billion in 2050 [362]. Due to increasing prosperity, the worldwide consumption of energy per individual will also increase. Even in the current modern world, several billion people still do not have access to basic needs, such as clean water, sanitation, nutrition, health care, and education [363]. These are all examples of the Sustainable Development Goals, adopted by all United Nations Member States in 2015 [363]. Access to energy is a key enabler to reach these basic needs. The worldwide energy demand is therefore expected to increase by 40% by 2040 [364]. At the same time, CO_2 emissions need to be reduced to reach the goals of the Paris agreement [365]. 80% of the total primary energy supply is currently produced by fossil fuels, such as coal, oil, and natural gas [364]. To reach the goals of the Paris agreement, attempts are made to replace fossil fuels with renewable alternatives such as wind and solar (PV) energy. Current expectations are that by 2040, 40% of the total generated electricity will be from renewable energy sources [364].

Unlike fossil fuels, energy production from intermittent renewable sources, including wind power and solar energy, critically depend on the availability of these sources leading to an uncontrollable energy output [366]. For direct integration to the power grid, uncontrollable availability of intermittent renewable energy sources within 10% of the installed capacity is acceptable without major technical problems [366]. However, large scale integration of intermittent energy sources above this limit is expected to cause frequent mismatches between the supply and demand of energy. To avoid this, integration of energy storage technologies is proposed as one of the promising solutions for stable and flexible supply of electricity [367, 368]. Different types of technologies have been developed for electrical energy storage including: hydrostorage, flywheels, batteries, and hydrogen produced by electrolysis etc [366, 367, 369, 370]. One of the most popular alternatives for energy storage is hydrogen [369]. Hydrogen has the advantage that it can be stored for long periods and converted to electricity without pollution [371]. Hydrogen has a broad span of applications, like

fuel cells, fuel for heating, transportation, or even as a raw material for the chemical industry [366, 371, 372]. Since hydrogen has a very low density at standard conditions, it has a very low volumetric energy density. For practical storage and transportation purposes, the density of hydrogen must be increased significantly [373]. The density of hydrogen can be increased by compression, cooling, or a combination of both, depending on the scale and application [373]. One of the emerging applications for hydrogen is found in sustainable transportation [371]. In Fuel Cell Electric Vehicles (FCEV), hydrogen is stored in compressed form in pressurized cylinders at $P = 350$ bar or $P = 700$ bar [373]. In practice, a passenger car needs a tank capacity of around 100 to 150 litres to store 4 to 6 kg of hydrogen, which provides a range of approximately 500 km [373]. High pressure storage tanks with pressures of at least $P = 875$ bar [154, 373] are installed at refuelling stations, to fuel a vehicle within the target time of three to five minutes [373]. Conventional compressor types that are currently used are piston, compressed air, diaphragm, or ionic compressors, depending mainly on the capacity of the refuelling station [373]. The conventional compressor requires on average 6 kWh/kg of energy to compress hydrogen from 10 to 400 bar [154].

An alternative compressor is the Electrochemical Hydrogen Compressor (EHC). HyET BV [60] has developed an EHC that works with pressures up to 1000 bar, and has the potential of bringing compression costs down to 3 kWh/kg [154]. The working principle of an EHC operation is similar to a Proton Exchange Membrane (PEM) fuel cell [374]. A single EHC stack consists of a low pressure and a high pressure side, separated by a membrane that is only permeable for hydrogen protons, and not for molecules. The membrane is positioned between two platinum catalysts containing electrodes. Once a potential difference is applied over the electrodes, a hydrogen molecule splits into two protons. The protons then travel through the membrane where conversion to hydrogen molecules takes place at elevated pressure [374]. This is shown schematically in Fig. 8.1. In the EHC, the proton transfer through the membrane is enabled by water. The EHC has several advantages compared to traditional technologies [154, 375–377]: *(1)* the EHC has a higher efficiency, especially at high compression ratios [378].

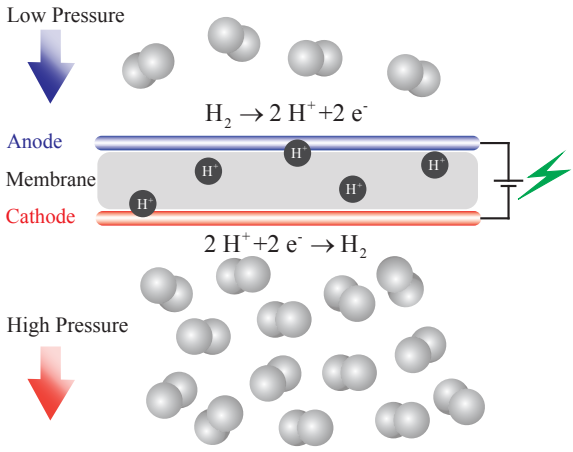


Figure 8.1: Schematic representation of electrochemical compression of hydrogen [154]. Using platinum-alloy catalysts, hydrogen molecules are split on the low pressure side of the membrane. Protons are forced through the membrane and form hydrogen molecules on the high pressure side.

In theory, the compression ratio using EHC can go to infinity, from an electrochemical perspective. The mechanical strength and back diffusion losses are the main limitations for higher pressure ratios for the EHC [378]; *(2)* due to the highly selective membrane that only allows the permeation of protons, contaminants are prevented from passing the membrane [378]. This means that the EHC performs both as a compressor and a purifier of hydrogen gas [378]; *(3)* the compressor has no moving parts, resulting in lower maintenance costs and making lubricants, which may contaminate the compressed hydrogen, redundant; *(4)* the EHC operates silently, since it has no rotating parts. This makes the EHC suitable for locations such as refuelling stations, where acoustical emission is a constraint; *(5)* the EHC is a compact device that is well suited to scale up [378]. Disadvantages of the EHC are similar to those of fuel cells, mainly high material costs. For instance, the platinum catalyst which is required to resist the corrosive environments in the compressor, is very expensive [378]. Another disadvantage is related to the proton transport through the membrane. Water enables

the proton transport through the membrane and therefore the membrane always needs to be hydrated [379]. Therefore, the resulting hydrogen gas is saturated with water which can be an issue depending on the application. The International Organization for Standardization (ISO) stated that water provides a transport mechanism for water-soluble contaminants such as K^+ and Na^+ when present as an aerosol [155]. Both K^+ and Na^+ can affect the fuel cell and are not recommended to exceed 0.05 μmol K^+ or Na^+ per mol hydrogen fuel mixture [155]. Another reason is that upon expansion, the temperature of high-pressure hydrogen gas increases. This is in sharp contract to CO_2 at room temperature, which has a Joule-Thomson coefficient of opposite sign [297]. To reduce the temperature of hydrogen when fuelling the tank, deep cooling is used to avoid ice formation. To avoid potential issues, the ISO has directed the maximum allowed concentration of impurities for gaseous hydrogen, including water, see Table 1 of the ISO 14687-2:2012 [155]. Note that having to add water at the inlet and to remove water at the outlet of the compressor is an important source of inefficieny for this EHC. The maximum concentration of water in the gaseous hydrogen, used for PEM fuel cells in road vehicles is limited to 5 μmol water per mol hydrogen fuel mixture [155]. This poses two important questions: *(1)* what is the solubility of water in hydrogen at high pressures? *(2)* if this solubility is too large, what is the best method to reduce the water content? To answer these questions, an accurate description of Vapor Liquid Equilibrium (VLE) of the $H_2O - H_2$ system at high pressures is required. Published experimental data that describe these systems is scarce. To the best of our knowledge, the only experimental data describing phase coexistencce of $H_2O - H_2$ for pressures exceeding 300 bar are from 1927 (limited to $T = 323$ K [61]). Wiebe and Gaddy studied also the solubility of hydrogen gas in liquid water at high pressures up to $P = 1013.25$ bar [380]. Therefore, molecular simulation and thermodynamic modelling are needed to determine the water content in the compressed hydrogen. In industrial applications, cubic type Equations of State (EoS) are one the most commonly used methods to study VLE, because of their simplicity [6, 13, 327, 381, 382]. In this chapter, the Peng-Robinson (PR) EoS and the Soave Redlich-Kwong (SRK) EoS with van der Waals mixing rules are used to predict the phase coexitence

of $H_2O - H_2$ at elevated pressures. However, molar volumes of the liquid phase and fugacity coefficients at high pressures obtained from PR-EoS and SRK-EoS modelling (with conventional mixing rules) are known to deviate significantly from experiments [9–12]. In this chapter, we show that both the PR-EoS and SRK-EoS fail to describe the liquid phase and the gas phase compositions, with or without fitted binary interaction parameters (k_{ij}'s). Since water is a highly polar molecule, either modifications of the conventional mixing rules are required [13], or more physically based models (*i.e.* SAFT types EoS [18] or molecular simulations [14]) should be used to describe the phase behavior of the $H_2O - H_2$ system [18]. It was found that a temperature dependent parameter k_{ij} is still required for SAFT type EoS modelling [383]. Therefore, force field based molecular simulation could be considered as a natural tool to study the phase coexistence of the $H_2O - H_2$ system. In this chapter, different molecular force fields for water and hydrogen are considered for describing the phase coexistence compositions of the liquid and gas phase of the $H_2O - H_2$ system, especially at high pressures. To evaluate the accuracy of the results from molecular simulations, we have performed an extensive literature survey on the VLE of $H_2O - H_2$ mixtures, at high pressures [61, 380, 384–391]. We show in this chapter that the best predictions of the VLE of the $H_2O - H_2$ system at high pressures (in both phases) are obtained using molecular simulations. No adjustable k_{ij}'s were used for the molecular simulations in this study.

This chapter is organized as follows. In Section 8.2, the molecular simulation techniques used in this study are explained and simulation details (molecular simulations and EoS modelling) and force field details for water and hydrogen are provided. Our results obtained from molecular simulations and EoS modelling are presented and compared with experimental data in Section 8.3. Our conclusions are summarized in Section 8.4. In appendix A.11, we present a detailed overview of available experimental VLE and solubility data for the $H_2O - H_2$ system at high pressures.

8.2 Modelling and Methodology

8.2.1 Simulation Technique

The natural choice for VLE phase equilibrium calculations is the GE method introduced by Panagiotopoulos [42–44], which is used extensively in molecular simulation studies [14]. As shown in chapter 2, in the GE, the vapor and liquid phase are simulated in two simulation boxes, which can exchange molecules, volume and energy. At coexistence, the pressures, temperatures and chemical potentials of each component are equal in both boxes [14]. In chapter 2, we combined the CFCMC method the GE with to obtain the chemical potentials and densities of coexisting gas and liquid phases for water, methanol, hydrogen sulfide and carbon dioxide. Since molecule exchanges in the CFCMC GE are performed using fractional molecules with scaled interactions, molecule transfers between coexisting phases are facilitated leading to a more efficient sampling of coexistence densities. It was shown in chapter 2 that the use of fractional molecules significantly improves the efficiency of the VLE calculations and the calculations of chemical potentials at coexistence. In simulations in the CFCMC GE, the chemical potential of a component is obtained using Eq. 2.9. The chemical potential is directly related to the fugacity. We show in appendix A.9 that the fugacity coefficient of component type i in phase j follows from:

$$\phi_{ij} = \frac{1}{Z_{\text{mix}}} \times \frac{p\left(\lambda_{ij} = 0\right)}{p\left(\lambda_{ij} = 1\right)} \tag{8.1}$$

where Z_{mix} is the compressibility factor of the mixture.

Based on the limited experimental solubility data available in literature at $T = 323$ K and pressures above $P = 300$ bar [61], we know that the solubility of water in the gas phase at high pressures ($P = 100$ bar to $P = 1000$ bar) is about a couple of hundred PPMs (molar), or less. At lower temperatures, due to the low the solubility of water in hydrogen, a very large number of hydrogen molecules (up to a million) in the gas phase would be required in the simulations to have on average a single water

molecule in the gas phase. The solubility of hydrogen in the liquid phase is also very low, *e.g.* mole fractions ranging from between 0.003 to 0.115 at $T = 323$ K and pressures between of $P = 25$ bar and $P = 1000$ bar. This makes most simulations of the $H_2O - H_2$ system in the CFCMC GE at low temperatures and high pressures impractical, as a very large system is needed to have at least a single component of each type in each box. One could in principle simulate the VLE of $H_2O - H_2$ in the CFCMC GE using a smaller system size. This would lead to poor statistics for the average number of H_2 molecules in the liquid phase, and H_2O molecules in the gas phase. Therefore, very long simulations are required. To circumvent these issues, both the gas and liquid phases (almost pure hydrogen gas and pure liquid water, respectively) are simulated independently in CFCNPT ensemble. For details on the CFCNPT ensemble [20], the reader is referred to chapters 3 and 5. By varying the mixture composition in the gas and liquid phases around the equilibrium state, the coexistence compositions are obtained by imposing equal chemical potentials for both phases.

At high pressures we know that the solute is almost pure in both phases, *i.e.* hydrogen in the gas phase and water in the liquid phase. For a solution close to infinite dilution, one can express the variation of the excess chemical potential of the solute, *i.e.* hydrogen in the liquid phase and water in the gas phase as a function of the number density of the solute:

$$\mu_{ij}^{\text{ex}}(\rho_{ij}) = A_{ij} + B_{ij}\rho_{ij} + C_{ij}\rho_{ij}^2 + \cdots \tag{8.2}$$

To obtain the terms A_{ij}, B_{ij}, $\cdots$, multiple simulations are performed at constant temperature and pressure, for different concentrations of the solute. In the region of interest (very dilute solutions) $\mu_{ij}^{\text{ex}}(\rho_{ij})$ depends linearly on the number density. As the solvent in both phases is almost a pure component, one can assume that the excess chemical potential of the solvent is independent of the number of few solute molecules in that phase. The coexistence densities are then obtained by imposing equal chemical potentials of each component using Eq. 8.2. Note that at conditions where both methods are applicable to obtain phase coexistence (*i.e.* simulations in the CFCMC GE and the CFCNPT ensemble), we have verified that both

methods (*i.e.* CFCMC GE and imposing equal chemical potentials) yield the same results.

8.2.2 Simulation Details

Depending on the temperature and pressure, molecular simulations are performed in the CFCMC GE (chapter 2) or in the CFCNPT ensemble (chapters 3 and 5). All simulations were performed using our in-house code. All molecules are rigid and the interactions between the molecules only consist of LJ and Coulombic interactions (see Eq. 4.1). To treat the electrostatic interactions, the Ewald summation was used with a relative precision of 1×10^{-6}. In CFCMC GE simulations of $H_2O - H_2$ mixtures, fractional molecules of water and hydrogen are present which are used to facilitate molecule exchanges between the phases.

Simulations in the CFCMC GE ensemble were started with 730 molecules of water and 600 molecules of hydrogen. For all temperatures and pressures, 10^5 equilibration cycles were carried out followed by $4 \cdot 10^6$ production cycles. For the rest of the simulation details in the CFCMC GE ensemble, the reader is referred to chapter 2. Independent CFCNPT simulations of the liquid phase, close to infinite dilution of hydrogen, were performed with 730 water molecules with $N_{H_2} \in \langle 0, 10 \rangle$ hydrogen molecules. Similarly, independent CFCNPT simulations of the gas phase, close to infinite dilution of water, were performed with 600 hydrogen molecules with $N_{H_2O} \in \langle 0, 7 \rangle$ water molecules. For the rest of the simulation details in the CFCNPT ensemble, the reader is referred to chapters 3 and 5. Details about the force field parameters for different water and hydrogen models and cutoff radii for LJ interactions are provided in the Supporting Information of Ref. [32].

8.2.3 Force Fields

To model the VLE of $H_2O - H_2$ mixtures, molecular force fields are considered to predict the density and composition of the gas and liquid phases. As the most commonly used force fields are developed based on

single-phase coexistence data [241, 256], we have screened these force fields using single-phase hydrogen (gas phase) and single-phase water (liquid phase) simulations. Force fields for water and hydrogen are selected based on predicting bulk properties of pure phases such as densities, chemical potentials and fugacity coefficients.

The densities and fugacity coefficients of molecular hydrogen in the gas phase are computed at different pressures using several force fields from the literature. The results are compared with REFPROP [144, 145]. Common force fields for molecular hydrogen in literature include single site [392–394], two-site [395] and multi-site potentials with (permanent) charge interactions [396–398]. Single-site hydrogen models are capable of predicting bulk thermodynamic properties of hydrogen accurately. The single-site hydrogen model by Buch [392] reproduces the bulk properties of hydrogen accurately up to high pressures. Multi-site hydrogen potentials that consider charge-quadrupoles and polarizability are more relevant for modeling hydrogen sorption in highly heterogeneous systems [395, 396, 398–401]. The densities and the excess chemical potentials predicted by different force fields of water in the liquid phase are computed as a function of pressure. The results are compared to those obtained from REFPROP [143, 145]. Even though water is a flexible and polarizable molecule, to date most molecular simulations studies consider rigid molecular potentials of water with constant point-charges [77, 80, 241, 255, 402]. It is computationally advantageous to use these simplified water potentials, which can predict thermodynamic and transport properties of water in good agreement with experiments. To obtain a more physical description of water, polarizable force fields have been developed to account for polarization effects [218, 241, 243–251]. Compared to the fixed-charge water potentials, thermodynamic properties of polarizable force fields are not fully known [243]. Commonly used fixed-charge force fields for water are three-site potentials: TIP3P [258], SPC [229, 259] and SPC/E [235], four-site potentials: TIP4P/2005 [255], TIP4P/Ew [204], OPC [256], and a five-site potential: TIP5P/Ew [205]. In chapters 2 and 4, we have shown that the computed excess chemical potentials of water for the three-site potentials TIP3P and SPC are in good agreement with values obtained from an empirical Helmholtz equation of

state [143] based on experimental data [144]. It is well-known that the TIP4P/2005 water outperforms the three-site models for predicting bulk properties of water such as the density [255]. In chapters 2 and 4, we have also shown that the computed excess chemical potentials of water obtained from four-site and five-site potentials show larger deviations from experimental data compared to three-site potentials [59, 222] .

8.2.4 Equation of State Modelling

The PR-EoS [151] and SRK-EoS [403] with the conventional van der Waals mixing rules are used to predict the $H_2O - H_2$ VLE. These equations of state are the most widely used in industry and perform best for describing the VLE of non-polar mixtures [345]. It is well-known that the molar volume of the liquid phase predicted by cubic equations of state is inaccurate [330, 404]. Since the solubility of small gas nonpolar molecules in the liquid phase are dominated by entropic effects (*i.e.*, molar volume), the solubility of H_2 in H_2O is predicted poorly. We have used both zero k_{ij}'s and k_{ij}'s fitted on high pressure experimental data. Parameters used for EoS modelling are provided in Table A7.

8.3 Results and Discussion

8.3.1 Molecular Simulations

The densities and the fugacity coefficients of pure hydrogen between $P = 100$ bar and $P = 1000$ bar obtained from CFCNPT simulations and EoS modelling are compared to those obtained from REFPROP [405], see Fig. 8.2. Since the differences between the results obtained for $P < 400$ bar is very small, only the results between $P = 400$ bar and $P = 1000$ bar are shown, and the raw data are provided in the Supporting Information of Ref. [32]. Hydrogen models used for this study include single-site models: *i.e.* Hirschfelder [393], Vrabec [394], Buch [392], two-site model: *i.e.* Cracknell [395] and the multi-site model of Marx [398]. It is clear that the densities obtained using the Buch [392] and Marx [398] force fields are in ex-

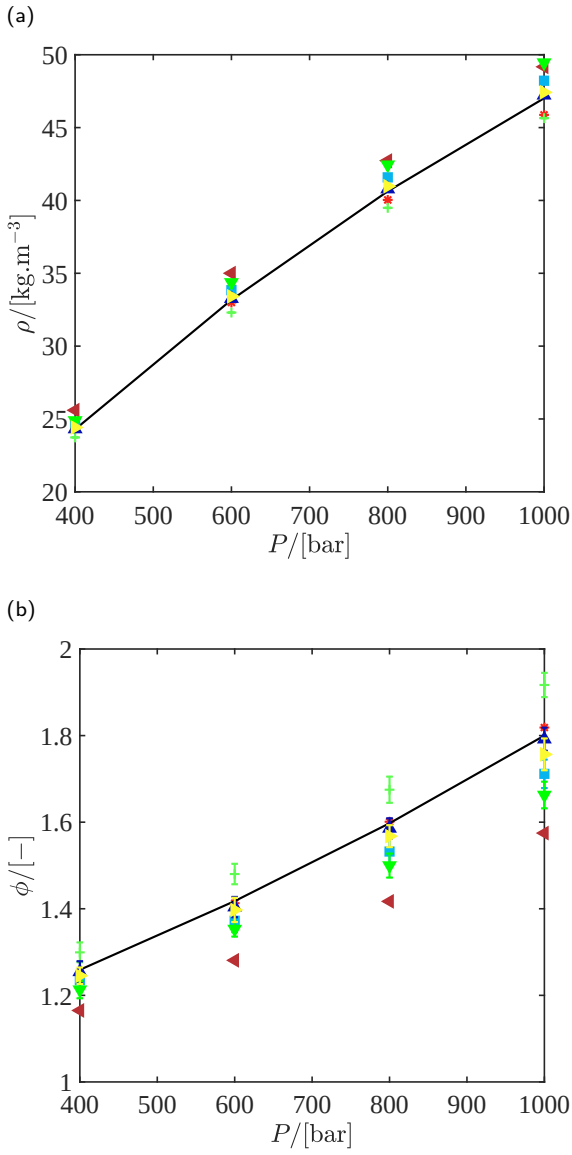


Figure 8.2: Comparison of different models to predict (a): the density and (b): the fugacity coefficient of pure hydrogen in the gas phase at $T = 323$ K and pressures ranging between $P = 10$ and $P = 1000$ bar. PR-EoS (left-pointing triangle), SRK-EoS (asterisk), experimental data from REFPROP [144, 145] (lines), molecular force fileds: Hirschfelder [393] (squares), Vrabec [394] (Plus signs), Buch [392] (upward-pointing triangles), Cracknell [395] (downward-pointing triangles) and Marx [398] (right-pointing triangles). Parameters used for EoS modelling are provided in Table A7. Raw simulation data are provided in the Supporting Information of Ref. [32].

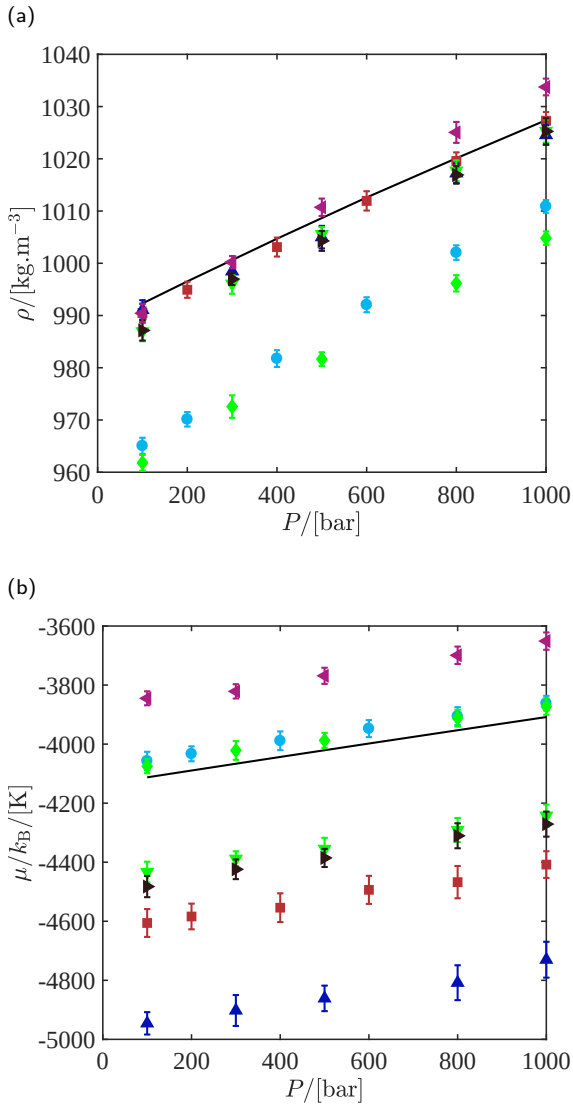


Figure 8.3: Comparison of different force fields of water to predict (a): the density and (b): the chemical potential in the liquid phase at $T = 323$ K and pressures ranging between $P = 10$ and $P = 1000$ bar: TIP3P [258] (diamonds), SPC [259] (circles), SPC/E [235] (right-pointing triangles), TIP4P/2005 [255] (squares), TIP4P/Ew [261] (downward-pointing triangles), OPC [256] (upward-pointing triangles), TIP5P/Ew [205] (left-pointing triangles). The reference state for the chemical potential is the ideal gas. In both subfigures, the lines are obtained from REFPROP [144, 145]. Raw data are in the Supporting Information of Ref. [32].

cellent agreement with experimental data up to $P = 1000$ bar. The results obtained from the PR-EoS and SRK-EoS deviate from experimental data for $P > 400$ bar. The calculated fugacity coefficients of pure hydrogen in the gas phase are best predicted using the Buch [392] and Marx [398] force fields. The calculated fugacity coefficients from the SRK-EoS are in excellent agreement with experiments. The simulation results show that both the Buch and Marx force fields outperform the other molecular models in predicting bulk densities and fugacity coefficients of hydrogen at high pressures. This means that considering a quadrupole moment for hydrogen does not strictly improve the bulk properties of hydrogen in the gas phase. Including the quadrupole moment may improve the prediction of phase coexistence in the liquid phase, as observed by Sun et al. [383]. Therefore, the Marx force field is considered further for VLE simulations of $H_2O - H_2$ mixtures. The densities and chemical potentials of TIP3P [258], SPC [259], SPC/E [235], TIP4P/2005 [255], TIP4P/Ew [261], OPC [256] and TIP5P/Ew [205] force fields between $P = 100$ bar and $P = 1000$ bar obtained from CFCNPT simulations are compared to the IAPWS empirical EoS [143, 145], see Fig. 8.3. Raw data are provided in the Supporting Information of Ref. [32]. It is shown in Fig. 8.3(a) that the force fields TIP5P/Ew and TIP4P/2005 clearly outperform the TIP3P and SPC force fields in predicting the density of liquid water (on average around 2%) over the whole pressure range. The TIP4P/2005 water is parameterized based on temperature of maximum density of liquid water, the stability of several ice polymorphs etc [255]. The TIP5P/Ew model is obtained from reparametrization of the TIP5P model [406] which is also a very accurate model capable of predicting maximum density of liquid water around 4° C [205]. Note that the deviations of the densities obtained from the TIP3P and SPC models decrease with increasing pressure. As shown in Fig. 8.3, the chemical potential of water is best predicted using the TIP3P and SPC force fields over the whole temperature range. This observation is also in agreement with previous works [59, 222]. The performance of the TIP3P and the SPC force fields are very similar in calculating the densities and chemical potentials of water. The TIP3P force field has been parameterized to the vaporization energy and density of liquid water [258]. This is consistent with the fact that the

computed chemical potential of TIP3P water is in better agreement with IAPWS empirical EoS, compared to TIP4P/2005 or TIP5P/Ew models. As shown in Fig. 8.3, the average deviation of the chemical potential of TIP3P force field from the IAPWS empirical EoS [143, 145] is about +50 K (in units of energy/k_B) for the whole pressure range. The average deviations of the chemical potentials for the TIP4P/2005 and TIP5P/Ew force fields from IAPWS empirical EoS are ca. −500 K and +250 K, respectively. The performance of the SPC/E force field is very similar to the TIP4P/2005 force field for predicting the densities and chemical potentials of water. For the 4-site water force fields, the densities and chemical potentials of the TIP4P/2005 force field show the best agreement with the experiments. Due to overall difference between the predicted densities and chemical potentials of these water models, it is not a priori clear which water model is best fitted for predicting the VLE of $H_2O - H_2$ mixtures. Therefore, three water models are considered (TIP3P,TIP4P/2005,TIP5P/Ew) in combination with the Marx force field (for hydrogen) for phase coexistence calculations of $H_2O - H_2$ mixtures, using molecular simulations.

The water content in the gas phase and the solubility of hydrogen in the liquid phase for the mixture defined by the TIP3P-Marx force fields are obtained from phase coexistence equilibrium calculations, see Fig. 8.4. To check the consistency between the results with both methods, phase coexistence calculations at $T = 323$ K and $P > 100$ bar are performed for both (*i.e.* CFCMC GE and CFCNPT). It is shown that both methods yield the same results within the error bars. At $T = 283$ K, all simulations are performed only in the CFCNPT ensemble for the whole pressure range. At $T = 310$ K and $P > 100$ bar, phase coexistence calculations are also performed using simulations in the CFCNPT ensemble. At $T = 366$ K and $T = 423$ K, phase coexistence calculations are performed using simulations in the CFCMC GE. Raw data from experimental results are provided in appendix A.11. Based on available experimental data at pressures above $P = 300$ bar [61], it is clear that the predicted solubility of TIP3P water in the gas phase is in good agreement with experimental data. At $T = 283$ K, no experimental solubilities have been found, and therefore only the results obtained from molecular simulations are shown. For all isotherms of water

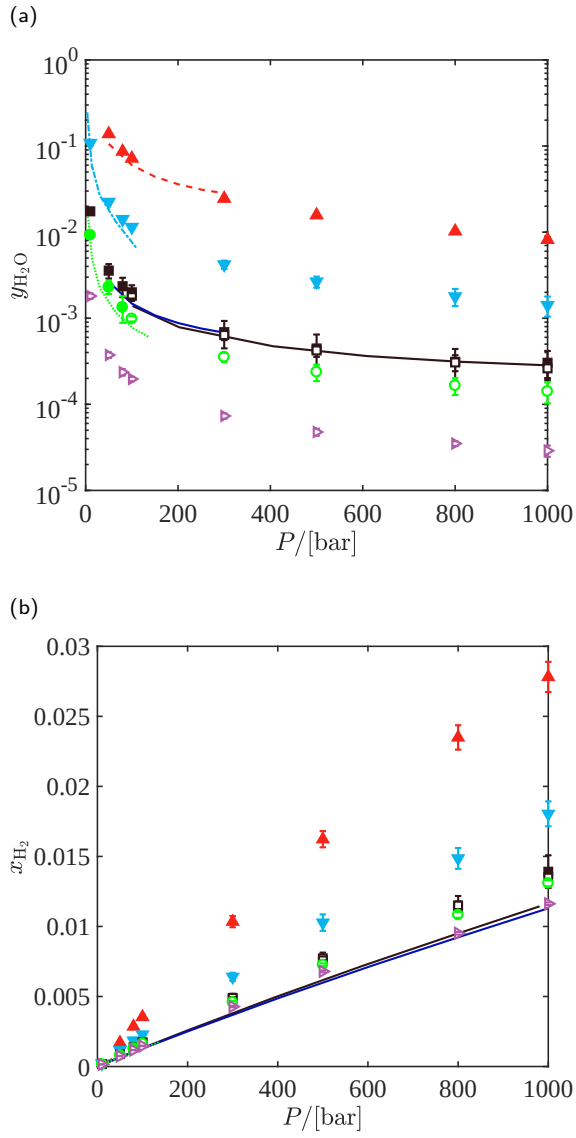


Figure 8.4: Vapor-Liquid equilibrium of $H_2O - H_2$ (TIP3P [258]-Marx [398]) at pressures ranging between $P = 10$ and $P = 1000$ bar. (a): y_{H_2O} in the gas phase and (b): x_{H_2} in the liquid phase. $T = 423$ K (upward-pointing triangles), $T = 366$ K (downward-pointing triangles), $T = 323$ K (squares), $T = 310$ K (circles), $T = 283$ (right-pointing triangles). Experimental data for $T = [423, 366, 323, 310]$ K are shown with dashed lines, dash-dot lines, solid lines and dotted lines, respectively. Published high pressure data are only available for $T = 323$ K [61]. Raw data are provided in the Supporting Information of Ref. [32].

vapor in the gas phase, it can be observed that the water content is slightly overpredicted at low pressures. At high pressures, the solubility of water in the gas phase is marginally underpredicted. From the condition of chemical equilibrium, we know that the chemical potential of water in the gas phase is equal to the chemical potential of water in the liquid phase. Therefore, it seems that good performance of the TIP3P force field to predict the isotherms of water in the gas phase is most likely related to how accurate it can predict μ_{H_2O} in the liquid phase. Based on the results shown in Figs. 8.3 and 8.4 it can be concluded that parametrization of the TIP3P force field based on the evaporation energy as one of the target quantities is essential for predicting the VLE of $H_2O - H_2$ mixtures. For all temperatures in this study (between $T = 283$ K and $T = 423$ K), it is observed that the solubility of water in the gas phase at coexistence is significantly higher than 5 μmol water per mol hydrogen (as allowed by the ISO standard [155]). Therefore, an additional step for removing water is needed.

The calculated isotherms for hydrogen in the liquid phase (TIP3P-Marx) are clearly overpredicted compared to experimental data as shown in Fig. 8.4(b). To the best of our knowledge, experimental solubility data for hydrogen isotherms in the liquid phase at pressures above ca. $P = 140$ bar are not available in the literature, except at $T = 323$ K [61]. The deviation from experimental solubilities of hydrogen at $T = 323$ K ranges from about 36% to 18% between $P = 50$ bar and $P = 1000$ bar, respectively. At $T = 366$ K, the deviation from experimental data is about 50% between $P = 50$ bar and $P = 100$ bar. At $T = 423$ K the deviation from experimental data is about 110% between $P = 50$ bar and $P = 80$ bar. Therefore, it can be concluded that the deviation of simulation results from experimental data increases with increasing temperature. Based on these results, it can also be concluded that the deviation from experimental solubilities decreases with increasing pressure. Similarly, better agreement is observed between experimental densities of water and those obtained based on TIP3P water at high pressures as also shown in Fig. 8.3. This suggests that predicting the density of the liquid phase (almost pure water) accurately may result in predicting the mixture compositions in better agreement with experiments.

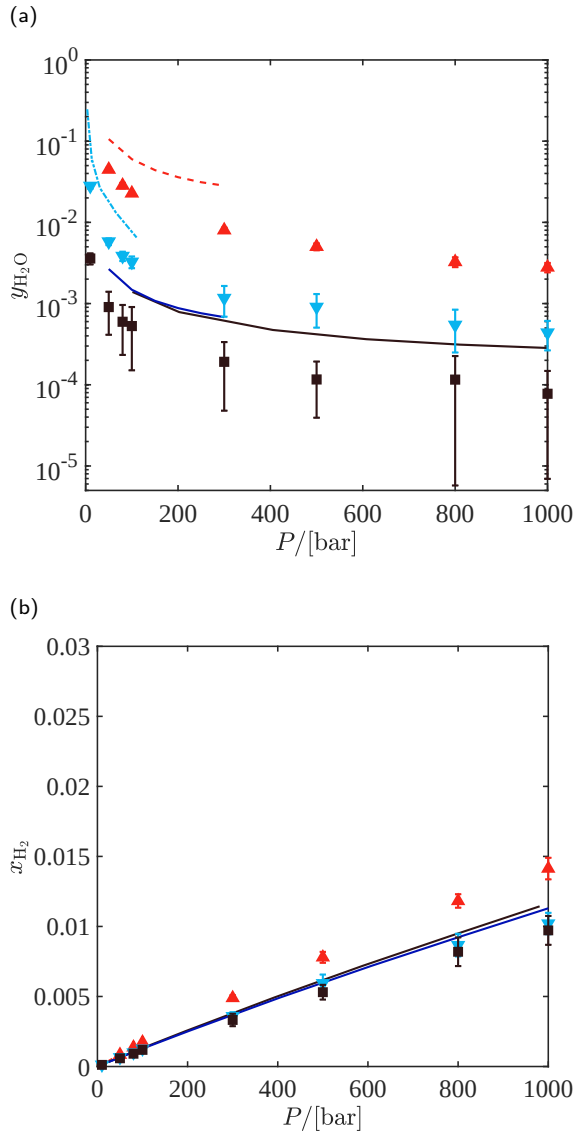


Figure 8.5: Vapor-Liquid equilibrium of $H_2O - H_2$ (TIP4P/2005 [255]-Marx [398]) at pressures ranging between $P = 10$ and $P = 1000$ bar. a: y_{H_2O} in the gas phase and b: x_{H_2} in the liquid phase. $T = 423$ K (upward-pointing triangles), $T = 366$ K (downward-pointing triangles), $T = 323$ K (squares). Experiential data for $T = [423, 366, 323, 310]$ K are shown with dashed lines, dash-dot lines, solid lines and dotted lines, respectively. Published high pressure data are only available for $T = 323$ K [61]. Raw data are provided in the Supporting Information of Ref. [32].

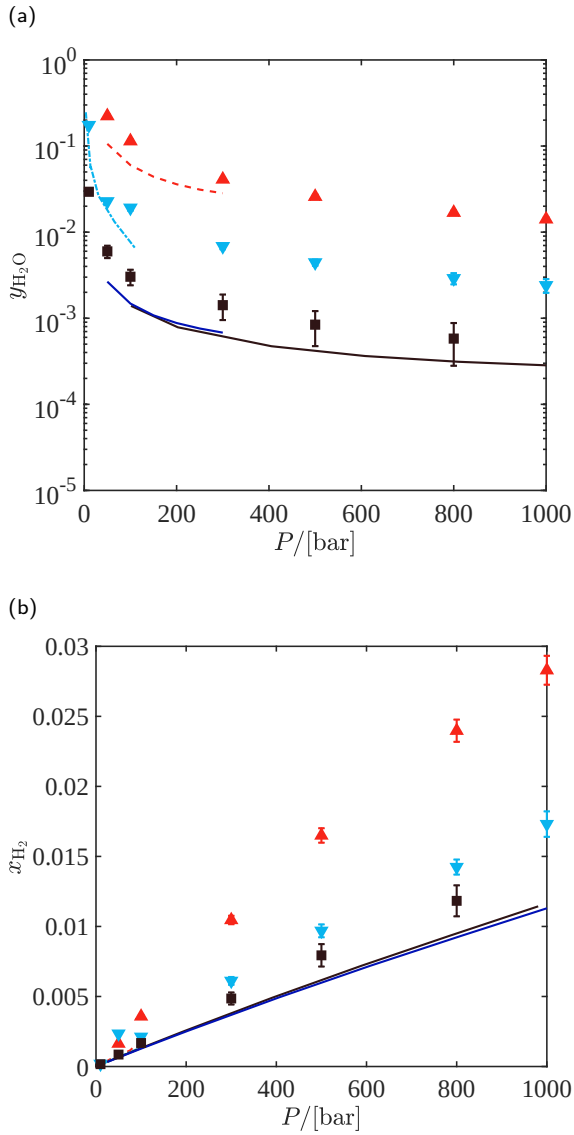


Figure 8.6: Vapor-Liquid equilibrium of $H_2O - H_2$ (TIP5P/Ew [205]-Marx [398]) at pressures ranging between $P = 10$ and $P = 1000$ bar. (a): y_{H_2O} in the gas phase and (b): x_{H_2} in the liquid phase. $T = 423$ K (upward-pointing triangles), $T = 366$ K (downward-pointing triangles), $T = 323$ K (squares). Experiential data for $T = [423, 366, 323, 310]$ K are shown with dashed lines, dash-dot lines, solid lines and dotted lines, respectively. Published high pressure data are only available for $T = 323$ K [61]. Raw data are provided in Supporting Information of Ref. [32].

The solubilities obtained from phase coexistence at equilibrium for the $H_2O - H_2$ mixture defined by the TIP4P/2005-Marx force fields are shown in Fig. 8.5. For this mixture, all simulations are performed in the CFCMC GE, at $T = 323$ K, $T = 366$ K and $T = 423$ K. It is clear from Fig. 8.5 that the solubilities of water in the gas phase are significantly underestimated for the whole pressure range. This is mainly due to the fact that the chemical potential of TIP4P/2005 water is significantly underpredicted, as shown in Fig. 8.3. Since the predicted water solubilities in the gas phase are systematically lower for the TIP4P/2005-Marx mixture (see Figs. 8.4a and 8.5a), the statistics for water solubilities obtained from CFCMC GE simulations are worse. This sampling issue is explained in Section 8.2.1. Similarly, the computed isotherms of hydrogen in the liquid phase are slightly underpredicted. For the mixture defined by TIP4P2005-Marx force fields, better agreement with experiments is observed for solubilities in the liquid phase for all temperatures. At $T = 366$ K, the deviation from experimental data is about 14% between $P = 50$ bar and $P = 100$ bar. At $T = 423$ K the deviation from experimental data is about 5% between $P = 50$ bar and $P = 80$ bar.

The solubilities obtained from phase coexistence at equilibrium for $H_2O - H_2$ mixture defined by TIP5P/Ew-Marx force fields are shown in Fig. 8.6. For this mixture, all simulations are performed in the CFCMC GE, at $T = 323$ K, $T = 366$ K and $T = 423$ K. In sharp contrast to the TIP4P/2005-Marx mixture, both calculated solubilities in the liquid and gas using the TIP5P/Ew-Marx mixture are overpredicted. The solubilities of hydrogen in the liquid phase are very similar to those obtained form the TIP3P-Marx force fields. To explain the results in a coherent way, it is important to consider the predicted water isotherms in the gas phase in Figs. 8.4 to 8.6 and the calculated chemical potentials of pure water in Fig. 8.3(b) simultaneously. From these figures, it can be concluded that underpredicting the solubilties in the gas phase is directly related to underpredcting the chemical potential of water (TIP4P/2005). Similarly, overpredicting the solubilties of water in the gas phase is directly related to overpredcting the chemical potential of water (TIP5P/Ew).

8.3.2 Equation of State Modelling

The water content in the gas phase and the solubility of hydrogen in the liquid phase are also calculated using the PR-EoS and SRK-EoS. High pressure experimental solubilites at $T = 323$ K were used to obtain the Binary Interaction Parameters (k_{ij}'s) for the PR-EoS and SRK-EoS. For $T = 323$ K, the isotherms of water and hydrogen in the gas and liquid phase are shown in Fig. 8.7 using both zero k_{ij}'s and non-zero k_{ij}'s. In Fig. 8.7, it is shown that the predicted solubilities in the liquid phase are significantly lower compared to experiments, using zero k_{ij}'s. The solubility of (nonpolar) gases is dominated by entropic effects which are related to the molar volume [407]. It is well-known that the predicted volumes of the liquid phase from PR-EoS or SRK-EoS, using conventional mixing rules, have significant errors [13, 330, 381]. Since 2017, more than 220 modifications of mixing rules for pure components and extensions to mixtures with the PR-EoS have been reported in literature [13]. This clearly indicates the need for more physically based models for thermodynamic modelling. In addition, the $H_2O - H_2$ system is highly polar in the liquid phase, and the performance of the conventional mixing rules for PR-EoS and SRK-EoS for polar mixture are known to be poor [13]. Therefore, it is expected that PR-EoS or SRK-EoS are not able to predict solubilities of hydrogen in liquid water accurately. With the fitted k_{ij}'s, the obtained solubilities of hydrogen in the liquid phase are in excellent agreement with experimental data for $p < 400$ bar. However, the solubilities in the gas phase deviate significantly using the fitted k_{ij}'s. Therefore, calculations of VLE of $H_2O - H_2$ mixtures using PR-EoS and SRK-EoS do not yield satisfactory results for both phases simultaneously, with or without adjusted k_{ij}'s.

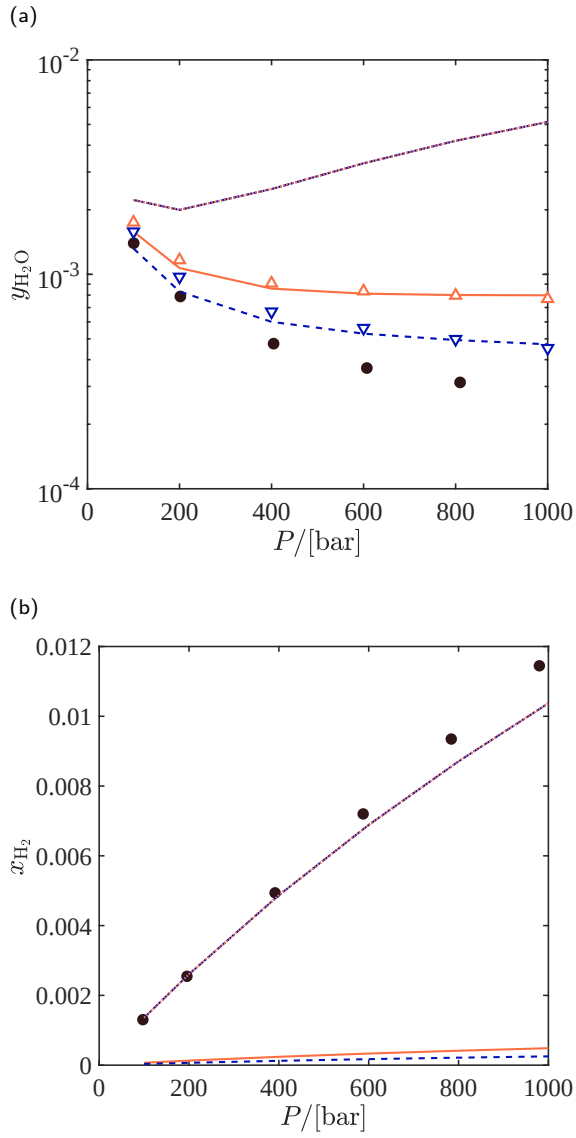


Figure 8.7: VLE of $H_2O - H_2$ at $T = 323$ K and pressures ranging between $P = 100$ and $P = 1000$ bar, obtained from EoS modelling. a: mole fraction of water in the gas phase, b: mole fraction of hydrogen in the liquid phase. Experimental solubilities are shown with circles (see appendix A.11). In both subfigures, the results are shown for $k_{ij} = 0$: PR-EoS [151] (lines), SRK-EoS [403] (dashed lines). The results from the γ-ϕ method are shown with open symbols: PR-EoS (upward-pointing triangles) and the SRK-EoS (downward-pointing triangles). The results for the fitted BIP for the PR-EoS ($k_{ij} = -0.89$) are shown with dash-dot lines. The results for the fitted BIP for the SRK-EoS ($k_{ij} = -1.51$) are shown with dotted lines.

8.4 Conclusions

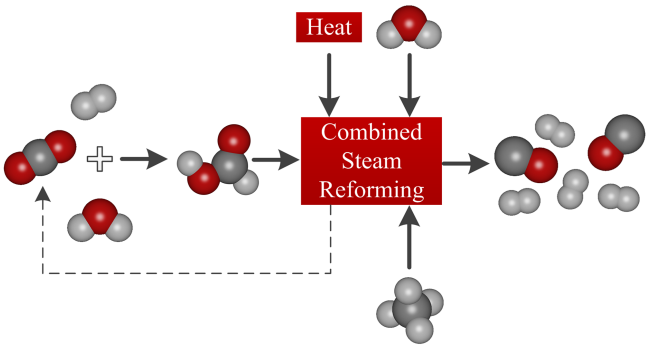
Molecular simulations are used to model the VLE behavior of $H_2O - H_2$ mixtures for pressures between $P = 10$ bar and $P = 1000$ bar. In appendix A.11, a detailed overview of available experimental data has been provided for this system. It is shown that commonly used cubic EoS, with conventional mixing rules fail to predict the composition of the gas and the liquid phases accurately. For the different molecular models for hydrogen, the Buch force field [392] (single-site model) and the Marx force field (including quadrupole moment) predict the density and fugacity coefficient of hydrogen in good agreement with experiments up to $P = 1000$ bar. In this study, no force field for rigid water with fixed point charges could accurately predict both the chemical potential and the density of water. The computed chemical potentials of TIP3P water [258] have the best agreement with experimental data from REFPROP [145] with a deviation of about +50 K (μ/k_B) for pressures between $P = 100$ bar and $P = 1000$ bar. This may be partly due to the fact that one of the target fitting parameters for the TIP3P force field is the heat of vaporization, unlike the TIP4P/2005 and TIP5P/Ew force fields. The computed chemical potentials (μ/k_B) of the TIP4P/2005 and TIP5P/Ew deviate on average by −500 K and +250 K from experimental data in this pressure range, respectively. Both the TIP4P/2005 and TIP5P/Ew force fields can predict the density of liquid water in good agreement with the experiments for the whole pressure range. From the simulations result, it is observed that solubilities of water in the gas phase are systematically underpredicted when using the TIP4P/2005 force field. This force field also underpredicts the chemical potential of liquid water compared to experiments. The highest solubilities in the gas phase are predicted using the TIP5P/Ew force field with the largest values for the calculated chemical potential of water. The best agreement between the predicted gas phase compositions and experiments for the whole pressure range are observed for the TIP3P fore field. This suggests that a suitable water force field for studying the VLE of $H_2O - H_2$ mixtures can be screened based on the chemical potential of the water model in the liquid phase. Based on

the screening of seven water force fields in this study, it turns out that the TIP3P and SPC force fields (with very similar values for chemical potential of liquid water) can best predict the equilibrium vapor phase coexistence composition of the $H_2O - H_2$ system. For all temperatures in this study, we observed that the solubility of water in the gas phase at coexistence is significantly higher than 5 μmol water per mol hydrogen (as allowed by the ISO standard). Therefore, an additional step for removing extra water from the gas phase is required. Despite the fact that the molecular simulations significantly outperform cubic EoS modelling for the VLE of $H_2O - H_2$ mixtures, the predicted liquid phase compositions need further improvements. The solubilities of hydrogen in the liquid phase are overpredicted using the TIP3P-Marx and TIP5P/Ew-Marx force fields. The best agreement between the calculated liquid phase composition and experiments is observed for the TIP4P/2005-Marx system (although the predicted solubilities are slightly lower). Further improvements in simulations of $H_2O - H_2$ systems may be realized taking polarizability of water molecules into account. Therefore, further molecular simulations of the $H_2O - H_2$ are recommended using polarizable force fields for water, especially to improve the predictions for the liquid phase composition.

Chapter 9

Combined Steam Reforming of Methane and Formic Acid

This chapter is based on the following maniuscript: Rahbari, A.; Ramdin, M.; van den Broeke, L. J. P.; Vlugt, T. J. H.; Combined Steam Reforming of Methane and Formic Acid to Produce Syngas with an Adjustable H_2:CO Ratio, Industrial & Engineering Chemistry Research 57, 10663-10674, Ref. [372]



9.1 Introduction

One of the consequences of the energy transition is that fossil fuel based production of chemicals will be replaced with renewable energy based processes [408–410]. The current infrastructure for producing chemicals is predominantly based on hydrogen and carbon. This means that to support the energy transition, a widely available and sustainable C_1 source is required. Therefore, the re-use of carbon dioxide will be an essential part of future chemical production processes [137, 411–413]. A range of efforts is underway to use carbon dioxide as a sustainable and economical source of C_1 to produce value-added chemicals [118, 411–413]. There are basically two pathways for the conversion of carbon dioxide: either by conventional hydrogenation or by electrochemical conversion.

Formic acid (FA) is one of the simplest products that can be made from carbon dioxide [118]. Recently, FA production by electrochemical reduction of CO_2 has gained significant interest [413–418]. In this process, the overall reaction in the electrochemical cell is the conversion of carbon dioxide with water to FA, according to:

$$CO_2 + H_2O \longrightarrow HCOOH + \frac{1}{2}\,O_2 \tag{R2}$$

The main advantage of the electrochemical conversion of carbon dioxide is that in the reaction water can be used as the hydrogen source. The cathodic half-cell reduction of carbon dioxide is described by the following reaction [416]:

$$CO_2 + 2\,H^+ + 2\,e^- \longrightarrow HCOOH \tag{R3}$$

The formation of FA is a two electron reaction, and the electric power to convert 1 kg of carbon dioxide to FA follows from [419]:

$$P = \frac{I \times U}{M_{CO_2}} \tag{9.1}$$

$$= \epsilon \, \frac{\lambda \times F \times Q \times U}{t \times M_{CO_2}} \tag{9.2}$$

with P the power input in kWh per kg carbon dioxide, I (A) the electric current, U the electrical potential which is in the order of 2.2 to 2.5 (V), λ is the number of electrons, $\lambda = 2$ (-), F is the Faraday coefficient which is equal to 96485 ($\mathrm{C \cdot mol^{-1}_{electron}}$), Q (C) is the total electric charge provided to the reactor, t (s) is the time, and M_{CO_2} ($\mathrm{g \cdot mol^{-1}}$) is the molecular mass of carbon dioxide. For an overall energy efficiency, ϵ, of around 70%, the energy required to convert 1 kg of carbon dioxide into 1 kg of FA is ca. 4 (kWh) [420]. A simple gross profit analysis using 5 dollar cents per kg carbon dioxide and an electricity price of 5 dollar cents per kWh leads to a cost price of around 25 dollar cents per kg FA [421].

For the hydrogenation reactions, the most sustainable approach to produce the required hydrogen is by water electrolysis, while traditional hydrogen production methods are based on fossil fuels [411, 422]. The conventional catalytic hydrogenation of carbon dioxide to FA proceeds according to [137, 423]:

$$CO_2 + H_2 \longrightarrow HCOOH \tag{R4}$$

FA is the simplest C_1 carboxylic acid, it is a non-toxic liquid between 281.55 K and 373.15 K, and it can be safely stored in aqueous solutions [138, 424]. In addition, hydrogenation of biomass derived feedstocks has been suggested as potential sustainable pathways to formate/formic acid production [412, 425–430]. Alternatively, value-added chemicals such as methanol, dimethyl ether and formate/formic acid can be produced by hydrogenation of carbon dioxide [137, 423]. To date, FA is mainly considered as a hydrogen storage material via its decomposition to hydrogen and carbon dioxide [119–137]. One of the key observations is that FA can be considered as a car-

bon monoxide carrier as well, via its decomposition to water and carbon monoxide [133, 134]. Basically, by combining the two main decomposition pathways towards hydrogen and carbon monoxide, and additional products such as water and carbon dioxide, FA can therefore be considered as a source for syngas, Yoshida et al. have reported the presence of FA as an intermediate in the Water Gas Shift Reaction (WGS) reaction reaction [138–140]:

$$H_2 + CO_2 \rightleftharpoons HCOOH \rightleftharpoons H_2O + CO \tag{R5}$$

On a molecular weight basis, FA contains 4.3 wt% hydrogen and 60.9 wt% carbon monoxide. Using a FA density of 1.22 $kg.L^{-1}$ at standard conditions leads to 53 grams H_2 per liter FA and 744 grams of carbon monoxide per liter FA. Based on the amount of 4.3 wt% or 53 grams of hydrogen, FA is identified as one of the most promising candidates for hydrogen storage [130, 431–433]. Considering the high carbon monoxide fraction in FA, it is interesting to explore the potential of FA as carbon monoxide carrier.

Typically, the WGS reaction is used together with Steam Reforming of Methane (SRM) to adjust the composition of the synthesis gas (syngas). This is one of the most common and oldest methods for syngas production [411, 434–443]. The reaction pathways for the SRM and WGS are:

$$CH_4 + H_2O \rightleftharpoons CO + 3\,H_2 \tag{R6}$$

$$CO + H_2O \rightleftharpoons CO_2 + H_2 \tag{R7}$$

Comparing Reactions R5 and R7 shows that by co-feeding FA to the SRM process, the WGS and the SRM reactions can be carried out in a single step. In this chapter, we show that using thermochemical equilibrium calculations the syngas composition (the H_2:CO molar ratio) can be adjusted to any required value between one and three by co-feeding FA to the SRM reaction. FA in the reactant feed decomposes to water and carbon monoxide, and/or to hydrogen and carbon dioxide which are all involved in the WGS reaction at high temperature. This can potentially change the conventional SRM

and WGS reactions (R7 and R6) from a two-step process into a single-step process.

This chapter is organized as follows. In Section 9.3, thermodynamic modelling of Reactions R5 and R6 is explained in detail. The Gibbs free energies of each component is calculated at standard pressure and temperatures between 400 K to 1400 K based on the partition function of isolated molecules. The Gibbs minimization method is used to calculate the composition of the product syngas at chemical equilibrium. Our results are summarized in Section 9.4. It is shown that the H_2:CO molar ratio can be adjusted to any value between one to three based on the initial concentration of the FA in the feed. Our conclusions are summarized in Section 9.5.

9.2 Applications of Formic Acid

9.2.1 Formic acid decomposition

The decomposition of FA can proceed according to two different pathways: decarbonylation (or dehydration) into carbon monoxide and water, or decarboxylation (dehydrogenation) into hydrogen and carbon dioxide:

$$\mathrm{HCOOH} \rightleftharpoons \mathrm{H_2O} + \mathrm{CO} \tag{R8}$$

$$\mathrm{HCOOH} \rightleftharpoons \mathrm{H_2} + \mathrm{CO_2} \tag{R9}$$

The selectivity towards FA dehydration or dehydrogenation depends on the temperature, pressure and the type of catalyst. For the heterogeneous FA decomposition, the dehydration/dehydrogenation selectivity of different solid catalysts has been studied [131, 444]. Metals and zinc oxide are predominantly active for Reaction R9, while other oxides are predominantly active for Reaction R8 [444]. Lopez et al. reported the results for different catalysts used for the heterogeneous FA decomposition reactions in the temperature range of $T = 573$ K to $T = 673$ K [131].

Blake and Hinshelwood investigated the homogeneous decomposition of FA acid in the gas phase for temperatures between $T = 709$ K and $T = 805$

K and concluded that catalytic effects become negligible at temperatures above $T = 773$ K [133]. Therefore, Reactions R8 and R9 are assumed to be in equilibrium at high temperatures which is a reasonable assumption since kinetics are fast and of minor importance [133]. In the temperature range of $T = 709$ K and $T = 805$ K, it was observed that Reaction R9 is of first order while Reaction R8 is second order. The reaction rates for packed and unpacked reactors were essentially the same for Reactions R8 and R9. In the beginning of 70's of the previous century, Blake et al. extended the experiment to the temperature range of $T = 820$ K to $T = 1053$ K [134]. In this temperature range, Reaction R9 was also observed to be a minor process, with typical CO:CO_2 = 10:1 molar ratios. Reaction R8 is second order for temperatures below $T = 943$ K and has an order of 1.5 for higher temperatures. The difference in yield of CO and CO_2 was attributed to the water-gas-shift reaction.

9.2.2 Synthesis of formic acid

Current industrial synthesis of FA is mainly based on fossil feedstocks using methanol carbonylation/methyl formate hydrolysis, and naphtha partial oxidation [431]. On a large scale, FA is produced in a two-step process of methanol carbonylation followed by methyl formate hydrolysis. In 2014, this two-step process was used to produce 81% of FA acid worldwide [445]. In the first step, carbon monoxide reacts with methanol at pressures around $P = 4$ MPa and temperatures around $T = 353$ K to produce methyl formate. FA and methanol are produced in the second step by methyl formate hydrolysis. The produced methanol is recycled back to the first step [429, 445]:

$$\mathrm{CO + CH_3OH \longrightarrow HCOOCH_3} \tag{R10}$$

$$\mathrm{HCOOCH_3 + H_2O \longrightarrow HCOOH + CH_3OH} \tag{R11}$$

The sum of Reactions R10 and R11 reduces to the direct reaction of carbon monoxide with water. FA synthesis based on methanol is a reliable and established process, however, it has some drawbacks [429, 445]. The process

uses carbon monoxide produced at high temperatures which is very energy intensive [429]. Also, a large excess of water is required to decompose methyl formate to FA (R11) [429]. The main application of FA is for the production of preservatives and as antibacterial agent [446], it is also used for dyeing in the leather industry. FA has received increased attention as a suitable material for controlled hydrogen storage and release [424, 431, 446–449]. A relatively new application is the use of FA in Direct Formic Acid Fuel Cells (DFACF) [450–453]. It has also been proposed to use FA for storage and transportation of carbon monoxide [454] or carbon dioxide [446, 455].

9.2.3 Established syngas technologies

Syngas refers to gas phase mixtures of hydrogen and carbon monoxide with various H_2:CO ratios [411, 436, 437]. Syngas can be produced by reforming almost any hydrocarbon source, such as naphtha, heavy oil, natural gas, biomass or coal [436, 440]. Currently, steam reforming of light hydrocarbons (e.g. methane, ethane, methanol, ethanol) is the most commonly used method for syngas production [411, 434–443]. An alternative source for syngas production are coal reserves, however, the investment costs associated with a coal-based syngas plant is approximately three times higher as compared to a natural gas-based plant [440]. Therefore, natural gas remains the major source for syngas production [411, 440]. Syngas is an intermediate in many industrial applications, and depending on the downstream process, the optimal H_2:CO molar ratio required in the syngas typically lies between one and three [353, 435]. The most common syngas applications in the chemical process industry are: methanol synthesis (H_2:CO = 2:1) [435, 456], Fischer-Tropsch (FT) synthesis (H_2:CO = 2:1) [457–459], oxo-synthesis or hydroformylation (H_2:CO = 1:1) [460–464], and acetic acid synthesis (H_2:CO = 1:1) [465]. As an illustrative example, Fig. 9.1 shows different reaction pathways leading to various syngas compositions by partial oxidation, steam reforming, carbon dioxide reforming, and the combined FA and steam option, as outlined in this work.

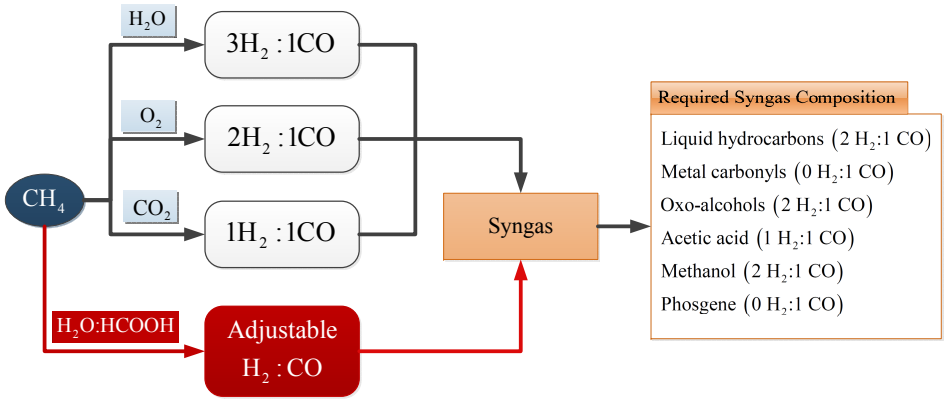


Figure 9.1: Different reaction pathways to reduce methane to syngas using oxygen (R13), steam (R6), carbon dioxide (R12) and an aqueous mixture of FA (R5) (proposed in this work). Synthesis of different products require favorable syngas H_2:CO ratios [466]. Synthesis of liquid hydrocarbons using the FT reaction (H_2:CO = 2:1) [435], metal carbonyls, oxo-alcohols (H_2:CO = 1:1) [464], acetic acid (H_2:CO = 1:1) [465], methanol synthesis (H_2:CO = 2:1) [456], and phosgene (H_2:CO = 0:1) [467].

To produce syngas from methane, various technologies have been developed, such as: SRM [411, 442] and WGS [411, 442], Carbon dioxide Reforming of Methane (CRM) [411, 442], catalytic Partial Oxidation of Methane (POM) [457, 458, 464, 468], combined partial oxidation and carbon dioxide Reforming of Methane or Autothermal Reforming of Methane (ARM) [469, 470], Combined Steam Reforming, and Carbon dioxide Reforming of Methane (CSRCRM) [437, 471].

The first industrial SRM plant was commissioned in the early 30's of the previous century [472, 473]. Methane is a very stable molecule and only at relatively high temperatures a high conversion rate to syngas is obtained [440, 474]. Syngas production from methane is divided into two steps. First, at high temperatures ranging from $T = 1073$ K to $T = 1273$ K and pressures ranging from $P = 20$ bar to $P = 40$ bar, the SRM reaction takes place. Second, the WGS is carried out after the SRM reaction to adjust the H_2:CO molar ratio [435, 442, 475]. SRM is typically performed

using Ni based catalysts [411]. This is related to the low cost and favorable activity of the Ni based catalysts, as compared to noble metals [443, 476]. Although noble metals are more coke resistant [440], the high cost and the limited availability make Ni catalysts a more practical choice in commercial applications [477]. SRM has two major drawbacks.

In particular, the Ni based processes suffer from coke formation which leads to deactivation of the catalyst. To avoid coke formation on the catalyst surface, excess steam is added which results in H_2 enriched syngas [478], and this will lead to a syngas composition with a H_2:CO molar ratio larger than three [438, 439, 469, 479]. The syngas compositions with high H_2:CO molar ratios do not meet the requirements for many downstream petrochemical processes e.g., FT synthesis [440, 457, 458, 473], acetic acid synthesis [473] or methanol synthesis [325, 326, 353, 479, 480]. The other disadvantage is that the SRM reaction is highly endothermic and subsequently highly energy-intensive [435, 474, 479, 481, 482].

Alternatively, in CRM (dry reforming), steam is replaced by carbon dioxide [442]:

$$CH_4 + CO_2 \longrightarrow 2\,H_2 + 2\,CO \tag{R12}$$

CRM is a CO_2-consuming reaction at temperatures between $T = 1073$ K to $T = 1273$ K resulting in syngas with H_2:CO = 1:1 molar ratio [411, 440, 483–485]. This syngas composition is more suitable for some downstream processes [440, 457, 458, 483, 484]. To lower the H_2:CO molar ratio of the syngas, CRM is widely used as a secondary reforming reaction after the SRM reaction [486]. CRM synthesis using Ni based catalysts, Co based catalysts, and noble metal based catalysts are reported extensively in the literature [468, 483, 485]. The main drawback of the CRM reaction is the rapid coke deposition, which can be explained by the Boudouard reaction [440, 487] ($2\,CO \longrightarrow C + CO_2$) taking place on the catalyst surface. Another disadvantage is the direct decomposition of methane [440, 487] ($CH_4 \longrightarrow C + 2\,H_2$) at high concentrations of CO_2 in the feed [411, 468, 483, 486, 488].

Catalytic Partial Oxidation of Methane (POM), also known as oxyreforming was introduced as an alternative to obtain syngas with a H_2:CO = 2:1 molar ratio, suitable for producing long chain hydrocarbons [411, 457, 458], and as a feed for methanol synthesis [325, 326, 479].

$$CH_4 + \frac{1}{2}O_2 \longrightarrow CO + 2\,H_2 \tag{R13}$$

POM is favorable for a wide range of temperatures allowing close to 100% methane conversion to syngas [474, 489]. The advantages include a short residence time and mild exothermicity [437, 440, 459, 473, 479, 490]. The main drawback is the presence of hot spots as a result of the high conversion rates of methane [437, 459, 477, 478, 491, 492]. Removing the heat produced in the reactor is difficult for large-scale operations, making the process difficult to control. From experiments by Prettre et al. it was shown that the catalytic oxidation of methane, with reactant feed composition CH_4:O_2 = 0.5:1, is not accurately represented by R13 [489, 493]. It seems that the POM reaction proceeds in two steps. The first step is exothermic which involves deep oxidation (combustion) of a part of the methane (approximately 25% of the starting moles) to carbon dioxide and steam. All oxygen is consumed during this process. In the second step, the residual methane reduces steam and carbon dioxide to syngas [474, 479, 489, 490, 493]. This is an endothermic process. The POM reaction mechanism can be described by the following three reactions [489, 493]:

$$CH_4 + 2\,O_2 \longrightarrow CO_2 + 2\,H_2O \tag{R14}$$

$$CH_4 + H_2O \rightleftharpoons CO + 3\,H_2 \tag{R15}$$

$$CH_4 + CO_2 \rightleftharpoons 2\,CO + 2\,H_2 \tag{R16}$$

The overall sequence of reactions, Reactions R14 to R16, using a Ni/Al_2O_3 catalyst results in syngas with a H_2:CO = 2:1 ratio as reported by Dissanayake et al. [489]. Yamamoto et al. have proposed the same reaction mechanism for partial oxidation of C_6^+ hydrocarbons using supported Ni catalysts [489, 494]. Different combinations of feedstock and catalysts can provide a specific H_2:CO molar ratio [442, 492].

Autothermal Reforming of Methane (ARM) is a combination of the POM and SRM-CRM process [495–498]. ARM is performed either in one or two separate reactors to reduce the energy consumption [411, 442]. The combination of the exothermic POM and endothermic SRM is energetically favorable [470]. ARM was originally designed for syngas production in ammonia and methanol plants in the 50s of the previous century [441]. The oxygen-steam flow is mixed with methane typically at around $T = 2200$ K [440], and methane is oxidized in a sub-stoichiometric flame. Combustion products enter the catalyst bed reactor, with high thermal stability and with the temperature in the range of $T = 1200$ K to $T = 1400$ K [440, 441]:

$$\mathrm{CH_4 + H_2O \rightleftharpoons CO + 3\,H_2} \tag{R17}$$

$$\mathrm{CH_4 + \frac{3}{2}\,O_2 \longrightarrow CO + 2\,H_2O} \tag{R18}$$

Adding steam is crucial for the ARM process as it prevents explosion hazards and suppresses coke formation [411, 470]. Equilibration of the syngas is further governed by the SRM and WGS reactions [441]. The H_2:CO molar ratio in the syngas can be precisely controlled by adjusting the H_2O:CH_4 and O_2:CH_4 molar ratios in the feed [440].

Combined Steam and Carbon Reforming of Methane (CSCRM) was proposed as an alternative to directly control the syngas composition [437, 471]. In this process, the H_2:CO molar ratio is adjusted by partially co-feeding carbon dioxide and steam with the reaction feed. Adding steam to CRM process drastically reduces the coke deposition on the catalyst [468, 499]. By changing the H_2O:CO_2:CH_4 ratio in the reaction feed, a H_2:CO ratio in the syngas between 1.5 to 2.5 is obtained [437, 441, 459, 468, 478, 479, 499–501].

9.3 Modelling and Methodology

For a single chemical reaction, the composition of the reaction product at chemical equilibrium is calculated from the method of equilibrium constants [2, 312, 502, 503]. In this approach, mole fractions are expressed as

functions of a single variable called the reaction coordinate (ϵ). The equilibrium constant is related to the individual mole fractions of the components and the stoichiometric coefficients. Therefore, ϵ is calculated for a single reaction [2, 312]. The method of equilibrium constants becomes numerically more difficult as the number of chemical species and reactions increases [2, 37, 502].

A necessary condition for chemical equilibrium is that the total Gibbs energy of the mixture reaches a minimum value at a given temperature and pressure. Based on this principle, the Gibbs minimization method [2, 117, 312] is used as a robust method to compute the composition of the reaction product at chemical equilibrium for multicomponent systems with simultaneous reactions [2, 117, 312, 504]. The solution obtained based on this method is less sensitive to the initial guess as compared to other methods [2, 117, 312]. The composition of the reaction product at chemical equilibrium is obtained by changing the initial composition such that the Gibbs energy of the mixture is minimized. The total number of atoms of each type should remain constant during this minimization process. The Gibbs free energy, or the chemical potential, of each component at the standard reference pressure, $P^\circ = 1$ bar, can be evaluated from the isolated molecule partition function [4, 142, 173, 505]:

$$\mu^\circ(T) = -RT \ln \left[\left(\frac{q(V,T)}{V} \right) \frac{k_\mathrm{B} T}{P^\circ} \right] \tag{9.3}$$

with $q(V,T)/V$ the temperature dependent part of the ideal gas partition function, k_B is the Boltzmann constant, P° is the standard reference pressure (1 bar), T is the temperature, and the volume $V = k_\mathrm{B}T/P^\circ$. Details on Eq. 9.3, zero of energy and reference state can be found in appendix A.2.1. Gibbs free energies of carbon monoxide, water, carbon dioxide, hydrogen, formic acid and methane at $P^\circ = 1$ bar, are computed using Eq. 9.3, and the results are provided in Table A6 in appendix A.4. The total Gibbs

energy of a multicomponent mixture equals [2, 3, 5]:

$$G^{\mathrm{t}} = \sum_{i=1}^{S} n_i \mu_i \tag{9.4}$$

with G^{t} the total Gibbs energy of the mixture, S is the number of components in the mixture, n_i is the number of moles of component i, μ_i is the chemical potential of component i in the mixture, and S is the total number of components in the mixture. Considering the standard state as an ideal gas, the chemical potential at any temperature and pressure is obtained from [20, 142]:

$$\mu_i = {\mu_i}^{\circ} + RT \ln \frac{y_i \varphi_i P}{P^{\circ}} \tag{9.5}$$

with R the universal gas constant, y_i is the mole fraction of component i, and φ_i is the fugacity coefficient of component i. The fugacity coefficient can be obtained from experimental volumetric data or an Equation of State [8]. Using the PR-EoS, φ_i is obtained from Eq. A90. Combining Eqs. 9.4 and 9.5 yields:

$$G^{\mathrm{t}} = \sum_{i=1}^{S} n_i {\mu_i}^{\circ} + RT \sum_{i=1}^{S} n_i \ln \frac{y_i \varphi_i P}{P^{\circ}} \tag{9.6}$$

At chemical equilibrium, the function G^t reaches a minimum. In a closed system, the minimization of Eq. 9.6 is subject to the constraints of the material balance [2, 117, 312]. In other words, the number of moles of each atom type remains constant during the reaction. For k types of atoms in the mixture, k independent mass balance equations are applied as constraints:

$$\sum_{i=1}^{S} n_i \alpha_{ik} = A_k \tag{9.7}$$

with A_k the number of atoms of type k, and α_{ik} is the number of atoms of type k present in molecule type i. Therefore, calculating the mixture compo-

sition at chemical equilibrium is reduced to minimizing Eq. 9.6 subject to the constraint of Eq. 9.7. The objective function, Eq. 9.6, is minimized using the function *fmincon* implemented in the MATLAB Optimization Toolbox [506]. In every iteration, the Peng-Robinson Equation of State (PR-EoS) [7, 8, 151] is used to evaluate the fugacity coefficients φ_i (Eq. A90). The mixture parameters are based on pure component parameters and van der Waals mixing rules [330, 382]. Parameters used for PR-EoS are provided in Table A7. The effect of the Binary Interaction Parameters (BIPs) are negligible for gaseous mixtures at high temperatures [20]. Therefore, the BIPs are set to zero in this work. Further details of the PR-EoS modeling are provided in appendix A.5. The standard Gibbs energies of Reactions R6 to R9 at P° are obtained based on the computed chemical potentials of individual components, Eq. 9.3, and the corresponding stoichiometric coefficients of the reaction:

$$\Delta G_r^\circ = \sum_{i=1}^{S} \nu_i \mu_i^\circ (T) \tag{9.8}$$

with ν_i the stoichiometric coefficient of component i. The standard reaction enthalpy ΔH_r° is directly computed using the Gibbs-Helmholtz equation [4]:

$$\left(\frac{\partial \Delta G_r^\circ / T}{\partial T} \right)_P = -\frac{\Delta H_r^\circ}{T^2} \tag{9.9}$$

9.4 Results and Discussion

9.4.1 Main reactions

In Fig. 9.2, the values for ΔG_r° and ΔH_r° (Eqs. 9.8 and 9.9) are plotted as a function of the temperature. The data in Fig. 9.2 are obtained for Reactions R6 to R9. The Gibbs energies are provided in Table A6. For more details on the computing of ΔG_r° and ΔH_r°, the reader is referred to appendix A.2.2 and A.2.1. The SRM reaction, R6, is endergonic, $\Delta G_r^\circ > 0$, at temperatures below $T = 880$ K [474] and exergonic, $\Delta G_r^\circ < 0$, at temperatures above $T = 880$ K. This indicates that the syngas production in the SRM reaction is favorable at high temperatures. The FA decomposition reactions, R8 and R9, are also endergonic for the temperature range of $T = 400$ K to $T = 1400$ K. Therefore, thermodynamic equilibrium favors high conversion of FA to water, hydrogen, carbon dioxide and carbon monoxide at high temperatures [133, 134]. The WGS reaction is endergonic at temperatures above $T = 1100$ K. At high enough temperatures, higher conversion of carbon dioxide and hydrogen to carbon monoxide and water is favored [507, 508]. The reaction enthalpies are calculated directly from the Gibbs-Helmholtz equation, Eq. 9.9. From the reaction enthalpies, ΔH, it is clear that Reactions R6 and R8 are endothermic and Reactions R7 and R9 are exothermic.

The Gibbs minimization method is used to compute the syngas equilibrium composition for the SRM and WGS reactions, R6 and R7. The reaction is studied with an equimolar feed mixture of water and methane, H_2O:CH_4 = 1:1, in the temperature range of $T = 800$ K to $T = 1500$ K, at $P = 1$ bar and $P = 25$ bar. The results are shown in Fig. 9.3. As expected, the H_2:CO molar ratios in the syngas are larger than three for the two pressures. It follows from Fig. 9.3 that full conversion of methane is achieved at $T = 1200$ K at $P = 1$ bar, while nearly full conversion of methane at $P = 25$ bar is not achieved until temperatures above $T = 1500$ K. For both pressures, low concentrations of carbon dioxide are observed in the syngas mixture at high temperatures. This is because the WGS equilibrium shifts towards carbon monoxide and water at high temperatures [435, 507–509].

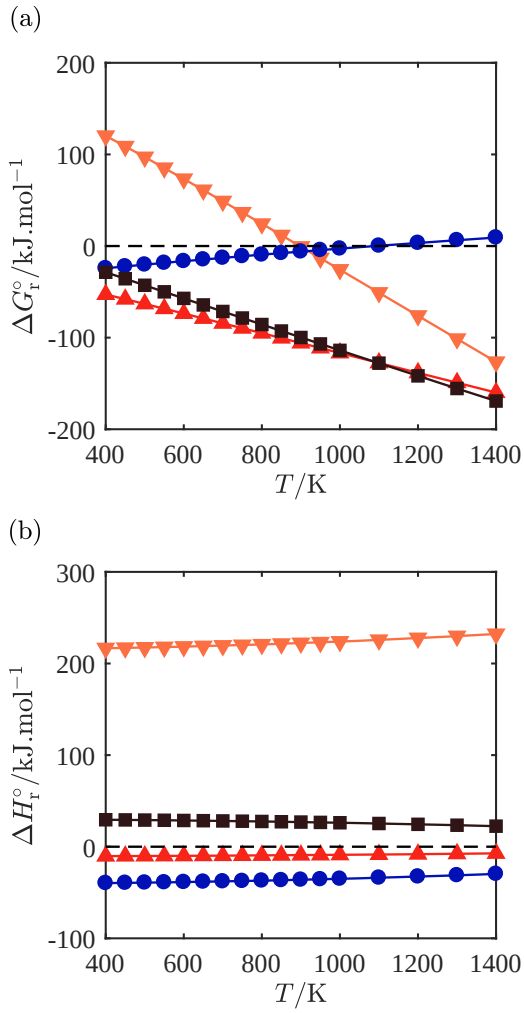


Figure 9.2: (a) Standard Gibbs energies of reaction and (b) reaction enthalpies for Reactions R8 and R9 (per mole of FA), R6 (per mole of methane) and R7 (per mole of water) as a function of temperature at $P^\circ = 1$ bar. The equilibrium constant is related to the Gibbs free energy change of the reaction [2, 4]. The symbols indicate: SRM (downward-pointing triangles), WGS (circles), Dehydration of FA (squares), Dehydrogenation of FA (upward-pointing triangles). A dashed line is used as a reference line at zero. Standard Gibbs energies of carbon monoxide, water, carbon dioxide, hydrogen, FA are provided in Table A6.

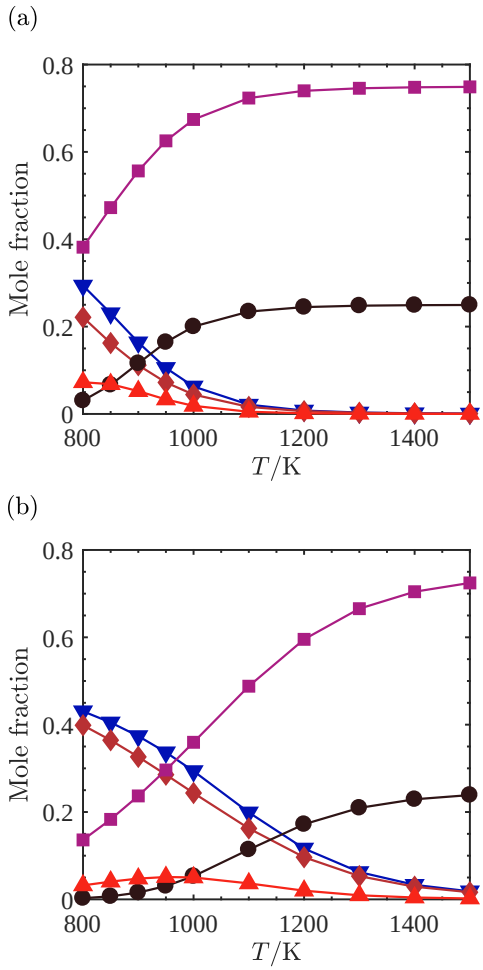


Figure 9.3: Equilibrium composition of syngas as a function of temperature computed using the Gibbs minimization method (Reactions R6 and R7). (a) at pressure of 1 bar and H_2O:CH_4 = 1:1 and (b) at pressure of 25 bar and H_2O:CH_4 =1:1. In both subfigures: mole fractions of hydrogen (squares), mole fractions of carbon monoxide (circles), mole fractions of methane (downward-pointing triangle), mole fractions of water (diamonds), and mole fractions of carbon dioxide (upward-pointing triangles).

9.4.2 FA combined with the SRM process: SRM-FA

To reduce the carbon-footprint of hydrogen and syngas production, alternative process schemes need to be developed. In Fig. 9.4, we propose a process scheme in which FA is combined with the SMR process to provide a wide range of H_2 to CO ratios. In this way both the CH_4:H_2O and the HCOOH:H_2O molar ratios can be varied. By using essentially both CH_4 and CO_2 as the C_1 feedstock the overall consumption of methane will be reduced.

For existing hydrogen and syngas production processes, there are two sources of carbon dioxide. To obtain the required product specifications for the hydrogen or the syngas, pressurized carbon dioxide is removed from the SRM and the WGS processes. Additionally, carbon dioxide is produced during heat generation and is present in the flue gas stream. The pressurized carbon dioxide stream from the existing hydrogen or syngas production units can be used as feedstock for the synthesis of FA, for both the electrochemical conversion or for the hydrogenation of carbon dioxide. It should be noted that large scale conversion of carbon dioxide to FA is not yet available. The advantage of the electrochemical route is that the product will be an aqueous FA stream. Various aqueous FA solutions, with different FA wt%, can be fed to the SRM-FA process, where the final syngas composition can be adjusted by the operating conditions for the pressure and temperature.

From the SRM process, syngas with a molar ratio of H_2:CO =3:1 is generally obtained. However, for most applications a lower H_2:CO molar ratio is required, see Fig. 9.1. To assess the potential of FA as a carbon monoxide carrier, the thermodynamic equilibrium of combining the FA decomposition reactions and the SRM reaction was evaluated. The composition of the feed mixture was defined by the molar ratio between water and methane, H_2O:CH_4, and varying the molar ratio between FA and water, HCOOH:H_2O. Two cases for the H_2O:CH_4 molar ratio are considered, H_2O:CH_4 = 1:1 and H_2O:CH_4 = 2:1. For the FA, a HCOOH:H_2O molar ratio in the range from 0.49 to 5.66 has been used, see Table 9.1. The equilibrium composition of

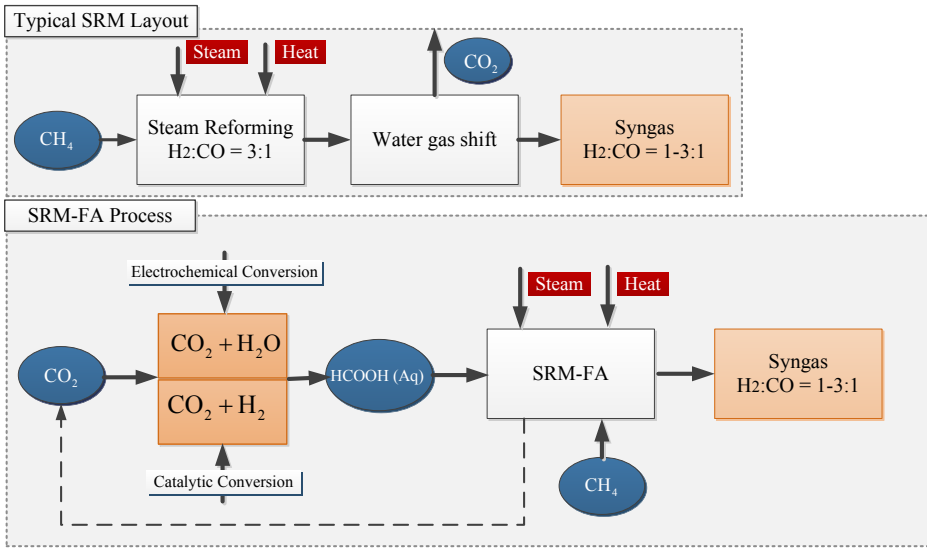


Figure 9.4: Comparison between a typical SRM layout and the layout for the proposed combined SRM-FA process. In the exiting SRM process, steam reforming is followed by the WGS process to adjust the H_2:CO ratio. In the alternative process, first FA is synthesized, and second the FA is added to the SMR to adjust the H_2:CO ratio. FA can be synthesized either by electrochemical conversion of CO_2 [414–418] or by conventional catalytic hydrogenation of CO_2 [137, 423].

Table 9.1: Different molar ratios of FA in FA-water mixtures used in the reactant feed. The corresponding mole percentage and weight percentage of FA (wt%) in the mixture is calculated based on the molar ratio between FA and water. The molar ratios between water and methane used in the simulations are H_2O:CH_4 = 1:1 and H_2O:CH_4 = 2:1.

FA:H_2O	FA:(FA + H_2O) %	FA wt%
0.11	10	22
0.49	33	56
1.00	50	72
1.50	60	79
5.66	85	94

the syngas is calculated using the Gibbs minimization method based on Reactions R5 and R6.

The results for the equilibrium syngas composition for the temperature range of $T = 900$ K to $T = 1500$ K at $P = 1$ bar are shown in Fig. 9.5, and the results for $P = 25$ bar are shown in Fig. 9.6. At $P = 1$ bar, full conversion of methane is achieved at temperatures up to $T = 1100$ K. By increasing the temperature further, the equilibrium favors conversion of hydrogen and carbon dioxide to water and carbon monoxide. This is in agreement with the equilibrium of the WGS reaction at high temperatures [128, 508, 510]. In addition, thermodynamic equilibrium favors complete FA decomposition (R5) in this temperature range. This leads to an increase in the mole fractions of water, carbon dioxide and carbon monoxide compared to the SRM-WGS process. Since the mole fraction of hydrogen is decreasing with the increase in temperature, contrary to the mole fraction of carbon monoxide, different H_2:CO molar ratios are obtained at different temperatures. The results show that by feeding FA to the SRM process, the equilibrium composition of the product syngas can be adjusted by changing the concentration of FA in the reactant feed. Future studies should investigate the effect of different types of catalyst for the combined SRM-FA process at different temperatures. The proposed method for adjusting the H_2:CO ratio by using FA is not limited to the methane steam reforming process.

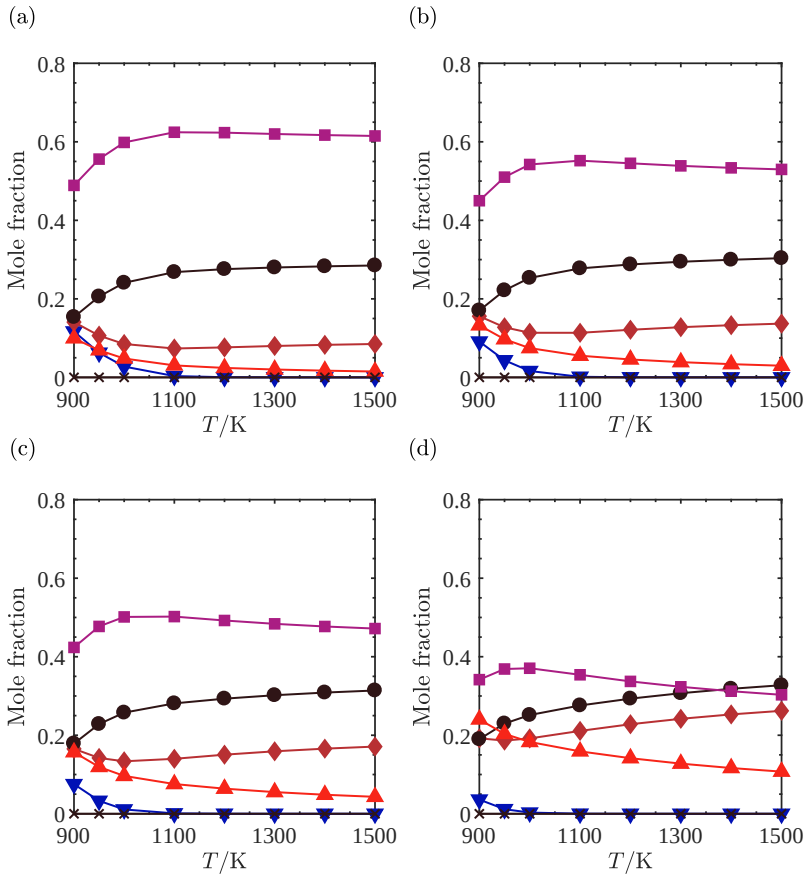


Figure 9.5: Equilibrium composition of syngas as a function of temperature obtained by co-feeding FA to the SRM reaction at 1 bar and $H_2O{:}CH_4 = 1$. The Gibbs minimization method is used to obtain the syngas equilibrium composition using on Eqs. R5 and R6. Initial mole fraction of FA relative to the mole fraction of water: (a) $HCOOH{:}H_2O = 0.49$, (b) $HCOOH{:}H_2O = 1.00$, (c) $HCOOH{:}H_2O = 1.50$, (d) $HCOOH{:}H_2O = 5.66$. In all subfigures: mole fractions of hydrogen (squares), mole fractions of carbon monoxide (circles), mole fractions of methane (downward-pointing triangle), mole fractions of water (diamonds), mole fractions of carbon dioxide (upward-pointing triangles), and mole fractions of FA (crosses).

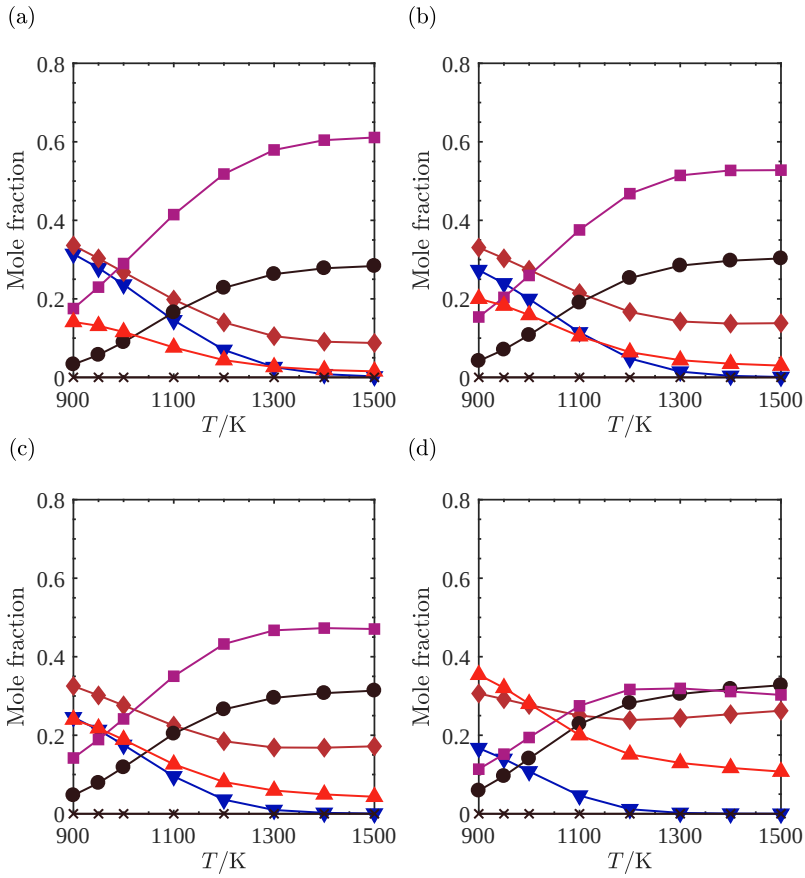


Figure 9.6: Equilibrium composition of syngas as a function of temperature obtained by co-feeding FA to the SRM reaction at 25 bar and $H_2O{:}CH_4 = 1$. The Gibbs minimization method is used to obtain the syngas equilibrium composition based on Reactions R5 and R6. Initial mole fraction of FA relative to Reactions mole fraction of water: (a) $HCOOH{:}H_2O = 0.49$, (b) $HCOOH{:}H_2O = 1.00$, (c) $HCOOH{:}H_2O = 1.50$, (d) $HCOOH{:}H_2O = 5.66$. In all subfigures: mole fractions of hydrogen (squares), mole fractions of carbon monoxide (circles), mole fractions of methane (downward-pointing triangle), mole fractions of water (diamonds), mole fractions of carbon dioxide (upward-pointing triangles), and mole fractions of FA (crosses).

First, it can be used in any process where adjustment of the H_2:CO ratio is required, see Fig. 9.1. Examples of this include: autothermal reforming, partial oxidation, gas-to-liquid technologies, naphtha reforming, biomass gasification etc. Second, formic acid can be used for energy storage by the use of fuel cells and formic acid reformers, to generate hydrogen, heat and electricity.

Carrying out the SRM-FA process at $P = 25$ bar changes the equilibrium composition of the reacting system, such that higher temperatures are required to fully reform methane and to reduce the carbon dioxide content in the syngas. This is in agreement with the Le Chatelier's principle [511, 512] which states that an increase in the pressure leads to a change in equilibrium composition to a new state in which fewer molecules/moles are present. Here, the thermodynamic equilibrium is shifted towards water, carbon dioxide and methane (R6) at low temperatures. Therefore, higher temperatures are required to reduce the methane and carbon dioxide concentrations in the syngas.

Based on the results shown in Figs. 9.5 and 9.6 it is clear that the concentrations of hydrogen and carbon monoxide can be adjusted by changing the FA concentration in the reactant feed. To have a clear overview of this principle in Fig. 9.7, the H_2:CO molar ratios in the syngas are plotted as a function of the composition of the reactant feed in the temperature range of $T = 800$ K to $T = 1500$ K at $P = 1$ bar and $P = 25$ bar. The composition of the reactant feed was obtained by adjusting the HCOOH:H_2O molar ratios between 0.11 and 5.66. Results shown in Fig. 9.7a and Fig. 9.7b correspond to H_2O:CH_4 = 1:1 molar ratio in the reactant feed at $P = 1$ bar and $P = 25$ bar, respectively. The results shown in Fig. 9.7c and Fig. 9.7d correspond to H_2O:CH_4 = 2:1 molar ratio in the reactant feed at $P = 1$ bar and $P = 25$ bar, respectively.

Thermochemical equilibrium calculations clearly show that Reactions R5 and R6 can be combined to produce syngas with an adjustable H_2:CO molar ratio ranging from one to three. The H_2:CO molar ratio can be adjusted by changing the HCOOH:H_2O:CH_4 ratio in the reactant feed at different temperatures. At high pressures, higher temperatures are required to reduce the concentration of methane and carbon dioxide in the product syngas, as

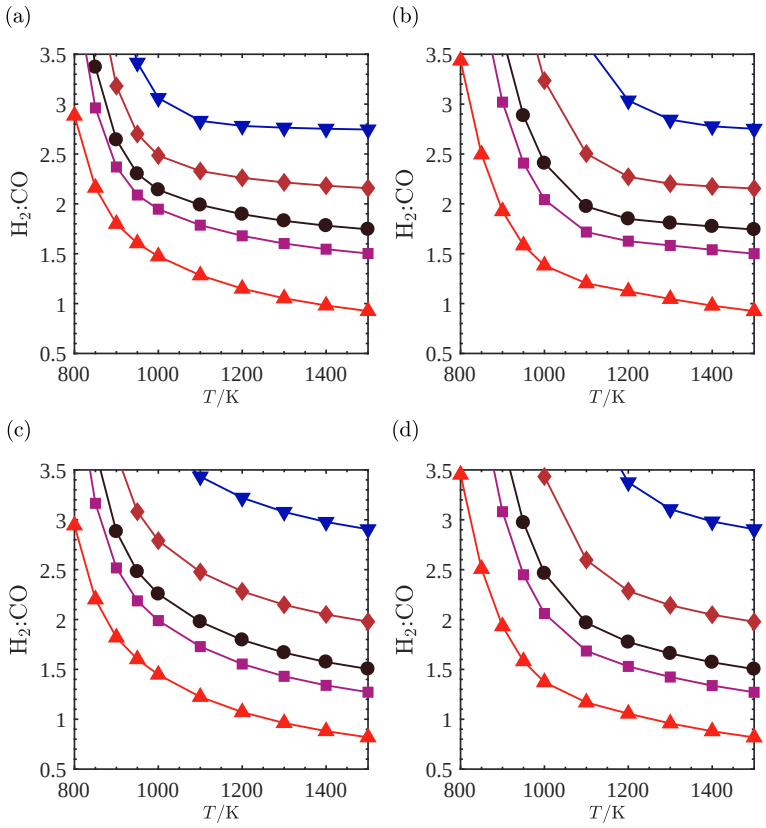


Figure 9.7: Different equilibrium syngas compositions (H_2:CO ratios) are obtained by co-feeding FA to the SRM reaction at different pressures and temperatures. (a) $P = 1$ bar, $H_2O{:}CH_4 = 1$ (b) $P = 25$ bar, $H_2O{:}CH_4 = 1$, (c) $P = 1$ bar, $H_2O{:}CH_4 = 2$, (d) $P = 25$ bar, $H_2O{:}CH_4 = 2$. In all subfigures, the initial mole fraction of FA relative to the mole fraction of water: $HCOOH{:}H_2O = 0.11$ (downward-pointing triangles), $HCOOH{:}H_2O = 0.49$ (diamonds), $HCOOH{:}H_2O = 1.00$ (circles), $HCOOH{:}H_2O = 1.50$ (squares), and $HCOOH{:}H_2O = 5.66$ (upward-pointing triangles).

shown in Fig. 9.6. However, adjusting the H_2:CO molar ratio in the syngas can be achieved at any temperature and pressure.

9.5 Conclusions

To adjust the H_2:CO molar ratio during syngas production, FA decomposition can be combined with the steam reforming of methane. The option to use FA as a syngas source is exploited by combining the two FA decomposition reactions at high temperatures. Essentially, FA can be considered as a combined hydrogen and carbon monoxide carrier. Thermodynamic equilibrium calculations show that the syngas composition can be controlled by adjusting the HCOOH:H_2O:CH_4 ratio in the reactant feed. It is possible to obtain different H_2:CO molar ratios between 1 to 3 in the product syngas. At higher pressures, higher temperatures are required for complete methane conversion and reducing carbon dioxide content in the syngas. Based on our results, it can be concluded that co-feeding FA to the SRM reaction can potentially reduce the traditional SRM and WGS processes from a two-step process to a single-step process able to produce syngas with adjustable H_2:CO . The proposed SMR-FA process based on CO_2 re-use may open up a range of new applications for formic acid.

Appendix

A.1 Partial Molar Properties in the CFC*NPT* Ensemble

A.1.1 Partial Molar Enthalpy in the *NPT* Ensemble

The partition function of a mixture of S distinguishable types of monoatomic components equals [14, 15]

$$Q_{N_i,P,T} = \beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \quad \text{(A1)}$$

This equation can be extended to mixtures of polyatomic molecules by simply multiplying by the ideal gas partition function of each polyatomic molecule (excluding the translational part) [4, 173]. As this only changes the reference state of the ideal gas contribution of partial molar properties, all derivations in this document are based on the formulation of Eq. A1. In this equation, U is the potential energy of the system, $\beta = 1/(k_\mathrm{B}T)$, k_B is the Boltzmann constant, s are the reduced coordinates of molecules in the system, V is the volume of the system, Λ_i is the thermal wavelength of a molecule of type i, N_i is the number of molecules of type i, P is the imposed pressure, and N is the total number of molecules present in the system:

$$N = \sum_{i=1}^{S} N_i \quad \text{(A2)}$$

Following Frenkel, Ciccotti and co-workers [150, 153], we can compute partial molar properties by calculating the ratio between partition functions with different number of molecules. We can write the partition function of

Eq. A1 when one additional molecule of type A is added to the system:

$$
\begin{aligned}
&Q_{N_\mathrm{A}+1,N_{i\neq A},P,T} = \\
&\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \cdot \frac{1}{\Lambda_\mathrm{A}^3 (N_\mathrm{A}+1)} \int \mathrm{d}V V^{N+1} \exp\left[-\beta P V\right] \\
&\times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]
\end{aligned}
\tag{A3}
$$

$s^{\mathrm{A}+}$ is the reduced coordinate of the additional molecule in the system, and $\Delta U^{\mathrm{A}+}$ is the interaction potential of this molecule with the rest of the system. The ratio between partition functions in Eq. A1 and Eq. A3 equals

$$
\begin{aligned}
&\frac{Q_{N_\mathrm{A}+1,N_{i\neq A},P,T}}{Q_{N_i,P,T}} = \\
&\frac{\left(\begin{aligned}
&\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \cdot \frac{1/\Lambda_\mathrm{A}^3}{N_\mathrm{A}+1} \int \mathrm{d}V V^N V \exp\left[-\beta P V\right] \\
&\times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]
\end{aligned}\right)}{\beta P \prod\limits_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)]}
\end{aligned}
\tag{A4}
$$

$$
\begin{aligned}
&= \frac{1/\Lambda_\mathrm{A}^3}{N_\mathrm{A}+1} \times \\
&\frac{\left(\begin{aligned}
&\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \cdot \int \mathrm{d}V V^N V \exp\left[-\beta P V\right] \\
&\times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]
\end{aligned}\right)}{\beta P \prod\limits_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)]}
\end{aligned}
\tag{A5}
$$

In terms of an ensemble average, we have

$$\frac{Q_{N_\mathrm{A}+1,N_{i\neq A},P,T}}{Q_{N_i,P,T}} = \frac{1/\Lambda_\mathrm{A}^3}{N_\mathrm{A}+1}\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T} \tag{A6}$$

We know that in the NPT ensemble, the chemical potential of species A equals [14]

$$\mu_\mathrm{A} = -\frac{1}{\beta}\ln\left(\frac{Q_{N_\mathrm{A}+1,N_{i\neq A},P,T}}{Q_{N_i,P,T}}\right) \tag{A7}$$

Combining Eq. A7 and Eq. A6 leads to

$$\mu_\mathrm{A} = -\frac{1}{\beta}\left[\ln\frac{\left(\frac{N}{\beta P}\right)/\Lambda_\mathrm{A}^3}{N_\mathrm{A}+1}\right] - \frac{1}{\beta}\left[\ln\left\langle\frac{\beta PV}{N}\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N,P,T}\right] \tag{A8}$$

The ideal part of the chemical potential equals [14]

$$\mu_\mathrm{A}^\mathrm{id} = -\frac{1}{\beta}\left[\ln\frac{\left(\frac{N}{\beta P}\right)/\Lambda_\mathrm{A}^3}{N_\mathrm{A}+1}\right] \tag{A9}$$

The excess part of the chemical potential is [14, 150]

$$\mu_\mathrm{A}^\mathrm{ex} = -\frac{1}{\beta}\left[\ln\left\langle\frac{\beta PV}{N}\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N,P,T}\right] \tag{A10}$$

These expressions for the chemical potential are in agreement with Ref. [14]. In principle, this expression can be split into an ideal gas part and an excess part, see Eq. 2.9. The finite-size effect of the ideal gas part can be corrected using the procedure outlined in [164]. The partial molar enthalpy $\bar{h}_\mathrm{A}$ and chemical potential of species A are related by

$$\bar{h}_\mathrm{A} = \left(\frac{\partial H}{\partial N_\mathrm{A}}\right)_{T,P,N_{i\neq\mathrm{A}}} = \left(\frac{\partial\beta\mu_\mathrm{A}}{\partial\beta}\right)_{P,N_i} \tag{A11}$$

For convenience, $\bar{h}_{\mathrm{A}}$ is expressed per molecule instead of per mole. In Eq. A11, H is the enthalpy of the system, μ_{A} is the chemical potential of component A, P is the imposed pressure, and T is the temperature of the mixture. It is important to note that the thermal wavelength (which appears in the ideal gas part of the chemical potential) is a function of β as well [15]:

$$\Lambda = \frac{\beta^{1/2}\, h}{\sqrt{2\pi m}} \tag{A12}$$

Combining Eqs. A9 and A12 leads to

$$\begin{aligned} \bar{h}_{\mathrm{A}}^{\mathrm{id}} &= -\frac{\partial}{\partial\beta}\left[\ln\left(\frac{\left(\frac{N}{\beta P}\right)}{N_{\mathrm{A}+1}} \Big/ \frac{\beta^{3/2}h^3}{\sqrt{8\pi^3 m^3}}\right)\right] \\ &= -\left[-\frac{1}{\beta} - \frac{3/2\ \beta^{1/2}}{\beta^{3/2}}\right] \\ &= -\left[-\frac{1}{\beta} - \frac{3}{2\beta}\right] \\ &= \frac{5}{2\beta} \end{aligned} \tag{A13}$$

in which contribution of the thermal wavelength is $3/(2\beta)$. For the excess part of the partial molar enthalpy, we combine Eqs. A10 and A11 and take the derivative with respect to β

$$\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = \left(\frac{\partial\beta\mu_{\mathrm{A}}^{\mathrm{ex}}}{\partial\beta}\right)_{P,N_i} = -\frac{\frac{\partial}{\partial\beta}\left\langle \frac{\beta PV}{N}\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}{\left\langle \frac{\beta PV}{N}\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}} \tag{A14}$$

Writing out the ensemble average leads to

$$\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = \frac{\frac{\partial}{\partial\beta}\left(\begin{array}{l}\int \mathrm{d}VV^N \frac{\beta PV}{N} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N,V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]\end{array}\right)}{\int \mathrm{d}VV^N \exp\left[-\beta PV\right] \int \mathrm{d}s^N \exp[-\beta U(s^N,V)]} - \frac{}{\left\langle \frac{\beta PV}{N} \exp\left[-\beta \Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}} \tag{A15}$$

The configurational part of the partition function of Eq. A1 is defined as

$$q_{N_i,P,T} = \int \mathrm{d}VV^N \exp\left[-\beta PV\right] \int \mathrm{d}s^N \exp[-\beta U(s^N,V)] \tag{A16}$$

Taking the derivative with respect to β, Eq. A15 leads to

$$\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = -\frac{1}{\left\langle \frac{\beta PV}{N} \exp[-\beta \Delta U^{\mathrm{A}+}]\right\rangle_{N_i,P,T}} \times \left[\frac{\frac{\partial}{\partial\beta}\left(\begin{array}{l}\int \mathrm{d}VV^N \frac{\beta PV}{N} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N,V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]\end{array}\right) \times q_{N_i,P,T}}{\left(q_{N_i,P,T}\right)^2} - \frac{\frac{\partial q_{N_i,P,T}}{\partial\beta} \times \left(\begin{array}{l}\int \mathrm{d}VV^N \frac{\beta PV}{N} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N,V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]\end{array}\right)}{\left(q_{N_i,P,T}\right)^2}\right] \tag{A17}$$

$$
\begin{aligned}
&= -\frac{1}{\left\langle \frac{\beta PV}{N} \exp[-\beta \Delta U^{\mathrm{A}+}] \right\rangle_{N_i,P,T}} \times \\
&\left[\frac{\int \mathrm{d}V V^N \frac{\beta PV}{N} \cdot \frac{\partial}{\partial \beta} \left(\begin{array}{l} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{q_{N_i,P,T}} \right. \\
&\quad + \frac{\left(\begin{array}{l} \int \mathrm{d}V V^N \frac{\partial \left(\frac{\beta PV}{N} \right)}{\partial \beta} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{q_{N_i,P,T}} \\
&\quad \left. - \frac{\frac{\partial q_{N_i,P,T}}{\partial \beta}}{q_{N_i,P,T}} \times \frac{\left(\begin{array}{l} \int \mathrm{d}V V^N \frac{\beta PV}{N} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{q_{N_i,P,T}} \right]
\end{aligned}
\tag{A18}
$$

$$
\begin{aligned}
&= -\frac{1}{\left\langle \frac{\beta PV}{N} \exp[-\beta \Delta U^{\mathrm{A}+}] \right\rangle_{N_i,P,T}} \times \frac{1}{q_{N_i,P,T}} \times \\
&\left[\begin{array}{l}
\int \mathrm{d}V V^N \frac{\beta PV}{N} \cdot \left(-\Delta U^{\mathrm{A}+} - PV - U(s^N, V)\right) \\
\times \left(\begin{array}{l} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \times \\ \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right) \\
+ \int \mathrm{d}V V^N \frac{1}{\beta} \times \frac{\beta PV}{N} \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\
\times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \\
- \int \mathrm{d}V V^N (-U(s^N, V) - PV) \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \times \\
\frac{1}{q_{N_i,P,T}} \times \left(\begin{array}{l} \int \mathrm{d}V V^N \frac{\beta PV}{N} \exp(-\beta PV) \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \times \\ \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)
\end{array} \right] \qquad \text{(A19)} \\
&= -\frac{\left[\begin{array}{l}
\left\langle \frac{\beta PV}{N} [-\Delta U^{\mathrm{A}+} - U(s^N, V) - PV] \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T} \\
+ \frac{1}{\beta} \left\langle \frac{\beta PV}{N} \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T} \\
- \left\langle -U(s^N, V) - PV \right\rangle \times \left\langle \frac{\beta PV}{N} \exp[-\beta \Delta U^{\mathrm{A}+}] \right\rangle_{N_i,P,T}
\end{array} \right]}{\left\langle \frac{\beta PV}{N} \exp[-\beta \Delta U^{\mathrm{A}+}] \right\rangle_{N_i,P,T}} \qquad \text{(A20)}
\end{aligned}
$$

The partial molar excess enthalpy of component A in the NPT ensemble of a multicomponent mixture equals

$$\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = -\frac{1}{\beta} + \frac{\left\langle (\Delta U^{\mathrm{A+}} + U(s^N, V) + PV) V \exp\left[-\beta \Delta U^{\mathrm{A+}}\right] \right\rangle_{N_i,P,T}}{\left\langle V \exp[-\beta \Delta U^{\mathrm{A+}}] \right\rangle_{N_i,P,T}} - \left\langle U(s^N, V) + PV \right\rangle_{N_i,P,T} \tag{A21}$$

The energy of the system with one additional molecule of component A can be expressed as

$$U(s^{N+1}, V) = \Delta U^{\mathrm{A+}} + U(s^N, V) \tag{A22}$$

We can write Eq. A21 as

$$\bar{h}_{\mathrm{A}}^{\mathrm{ex}} = -\frac{1}{\beta} + \frac{\left\langle (U(s^{N+1}, V) + PV) V \exp[-\beta \Delta U^{\mathrm{A+}}] \right\rangle_{N_i,P,T}}{\left\langle V \exp[-\beta \Delta U^{\mathrm{A+}}] \right\rangle_{N_i,P,T}} - \left\langle U(s^N, V) + PV \right\rangle_{N_i,P,T} \tag{A23}$$

This is the result that Frenkel, Ciccotti and co-workers previously found [150, 153]. It is instructive to show that $\bar{h}_{\mathrm{A}}^{\mathrm{ex}}$ equals zero for an ideal gas. Since intermolecular interactions for an ideal gas are zero, for an ideal gas Eq. A23 reduces to

$$\begin{aligned} \bar{h}_{\mathrm{A}}^{\mathrm{ex}} &= -\frac{1}{\beta} + \frac{\left\langle PV^2 \right\rangle_{N_i,P,T}}{\left\langle V \right\rangle_{N_i,P,T}} - \left\langle PV \right\rangle_{N_i,P,T} \\ &= -\frac{1}{\beta} + P \left[\frac{\left\langle V^2 \right\rangle_{N_i,P,T}}{\left\langle V \right\rangle_{N_i,P,T}} - \left\langle V \right\rangle_{N_i,P,T} \right] \end{aligned} \tag{A24}$$

For the average volume $\left\langle V \right\rangle_{N_i,P,T}$ we can write

$$\left\langle V \right\rangle_{N_i,P,T} = \frac{\int_0^\infty V^N V \exp[-\beta PV] \mathrm{d}V}{\int_0^\infty V^N \exp[-\beta PV] \mathrm{d}V} \tag{A25}$$

Both integrals in the nominator and denominator can be solved analytically according to

$$\int_0^\infty x^N \exp[-ax]\mathrm{d}x = \frac{N!}{a^{N+1}} \tag{A26}$$

Therefore, we have

$$\langle V\rangle_{N_i,P,T} = \left(\tfrac{(N+1)!}{(\beta P)^{N+1}}\right)/\left(\tfrac{N!}{(\beta P)^N}\right) = \frac{N+1}{\beta P} \tag{A27}$$

Similarly, for the term $\langle V^2\rangle_{N_i,P,T}$ we have

$$\begin{aligned}\langle V^2\rangle_{N_i,P,T} &= \frac{\int_0^\infty V^N V^2 \exp[-\beta PV]\mathrm{d}V}{\int_0^\infty V^N \exp[-\beta PV]\mathrm{d}V} \\ &= \left(\tfrac{(N+2)!}{(\beta P)^{N+2}}\right)/\left(\tfrac{N!}{(\beta P)^N}\right) \\ &= \tfrac{(N+2)(N+1)}{(\beta P)^2}\end{aligned} \tag{A28}$$

Combining Eqs. A24, A27 and A28 leads to

$$\begin{aligned}\bar{h}_{\mathrm{A}}^{\mathrm{ex}} &= -\frac{1}{\beta} + P\left[\frac{(N+2)(N+1)/(\beta P)^2}{N+1/\beta P} - \frac{N+1}{\beta P}\right] \\ &= -\frac{1}{\beta} + P\left[\frac{N+2}{\beta P} - \frac{N+1}{\beta P}\right] \\ &= 0\end{aligned} \tag{A29}$$

This is the expected result.

A.1.2 Partial Molar Volume in the *NPT* Ensemble

We know from thermodynamic relations that the partial molar volume of component A in a mixture of S components equals [8, 150, 153]

$$\bar{v}_{\mathrm{A}} = \left(\frac{\partial V}{\partial N_{\mathrm{A}}}\right)_{T,P,N_{i\neq \mathrm{A}}} = \left(\frac{\partial \mu_{\mathrm{A}}}{\partial P}\right)_{T,N_i} \tag{A30}$$

For convenience, partial molar properties are considered per molecule instead of per mole. In Eq. A30, V is the volume of the system, N_i denotes the number of molecules of component i, μ_{A} is the chemical potential of component A, P is the imposed pressure, T is the temperature of the mixture, $\beta = 1/(k_{\mathrm{B}}T)$, and k_{B} is the Boltzmann constant. Following Frenkel, Ciccotti, and co-workers [150, 153], combining Eqs. A8 and A30 leads to

$$\begin{aligned}
\bar{v}_{\mathrm{A}} &= -\frac{1}{\beta}\frac{\partial}{\partial P}\left[\ln \frac{\left(\frac{N}{\beta P}\right)/\Lambda_{\mathrm{A}}^3}{N_{\mathrm{A}}+1}\right] - \frac{1}{\beta}\frac{\partial}{\partial P}\left[\ln\left\langle \frac{\beta PV}{N}\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}\right] \\
&= \frac{1}{\beta P} - \frac{1}{\beta}\left[\frac{\frac{\partial}{\partial P}\left\langle PV\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}{\left\langle PV\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}\right] \\
&= \frac{1}{\beta P} - \frac{1}{\beta}\left[\frac{\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T} + P\times\frac{\partial}{\partial P}\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}{P\times\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}\right] \\
&= -\frac{1}{\beta}\times\frac{\frac{\partial}{\partial P}\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}{\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}}
\end{aligned} \tag{A31}$$

Starting from $\frac{\partial}{\partial P}\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T}$, we write out the ensemble average

$$\begin{aligned}
&\frac{\partial}{\partial P}\left\langle V\exp\left[-\beta\Delta U^{\mathrm{A}+}\right]\right\rangle_{N_i,P,T} = \\
&\frac{\partial}{\partial P}\left(\frac{\left(\begin{array}{l}\int \mathrm{d}V V^N V\exp[-\beta PV]\int \mathrm{d}s^N\exp[-\beta U(s^N,V)] \\ \times\int \mathrm{d}s^{\mathrm{A}+}\exp[-\beta\Delta U^{\mathrm{A}+}(s^{\mathrm{A}+},s^N,V)]\end{array}\right)}{\int \mathrm{d}V V^N\exp\left[-\beta PV\right]\int \mathrm{d}s^N\exp[-\beta U(s^N,V)]}\right)
\end{aligned} \tag{A32}$$

$$= \left[\begin{array}{l} \dfrac{\frac{\partial}{\partial P} \left(\begin{array}{l} \int \mathrm{d}V V^N V \exp[-\beta P V] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{\left(q_{N_i,P,T}\right)^2} \times q_{N_i,P,T} \\ \\ - \dfrac{\frac{\partial q_{N_i,P,T}}{\partial P} \times \left(\begin{array}{l} \int \mathrm{d}V V^N V \exp[-\beta P V] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{\left(q_{N_i,P,T}\right)^2} \end{array} \right] \tag{A33}$$

$$= \left[\begin{array}{l} \dfrac{\frac{\partial}{\partial P} \left(\begin{array}{l} \int \mathrm{d}V V^N V \exp[-\beta P V] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{q_{N_i,P,T}} \\ \\ - \dfrac{\frac{\partial q_{N_i,P,T}}{\partial P}}{q_{N_i,P,T}} \times \dfrac{\left(\begin{array}{l} \int \mathrm{d}V V^N V \exp[-\beta P V] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \end{array} \right)}{q_{N_i,P,T}} \end{array} \right] \tag{A34}$$

$$
\begin{aligned}
&= \frac{1}{q_{N_i,P,T}} \times \\
&\left[\begin{array}{l}
\displaystyle\int \mathrm{d}V V^N \left(-\beta V^2\right) \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \times \\
\displaystyle\int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)] \\
\displaystyle- \int \mathrm{d}V V^N \left(-\beta V\right) \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \times \frac{1}{q_{N_i,P,T}} \\
\displaystyle\times \int \mathrm{d}V V^N V \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\
\displaystyle\times \int \mathrm{d}s^{\mathrm{A}+} \exp[-\beta \Delta U^{\mathrm{A}+}(s^{\mathrm{A}+}, s^N, V)]
\end{array} \right]
\end{aligned}
\tag{A35}
$$

In terms of ensemble averages, we can write

$$
\begin{aligned}
&\frac{\partial}{\partial P} \left\langle V \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T} = \\
&\left\langle -\beta V^2 \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T} \\
&- \left\langle -\beta V \right\rangle_{N_i,P,T} \times \left\langle V \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T}
\end{aligned}
\tag{A36}
$$

Combining Eq. A36 and Eq. A31 leads to

$$
\begin{aligned}
&\bar{v}_{\mathrm{A}} = -\frac{1}{\beta} \times \\
&\left[\frac{\left\langle -\beta V^2 \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T} + \left\langle \beta V \right\rangle_{N_i,P,T} \times \left\langle V \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T}}{\left\langle V \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T}} \right]
\end{aligned}
\tag{A37}
$$

The partial molar volume of component A is [150, 153]

$$
\bar{v}_{\mathrm{A}} = \frac{\left\langle V^2 \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T}}{\left\langle V \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i,P,T}} - \left\langle V \right\rangle_{N_i,P,T} \tag{A38}
$$

This expression is identical to the one derived by Frenkel, Ciccotti, and co-workers [150, 153]. For an ideal gas, we can calculate the partial molar

volume analytically. Since there are no intermolecular interactions between ideal gas molecules, we can write

$$\bar{v}_{\mathrm{A}} = \frac{\left\langle V^2 \right\rangle_{N_i,P,T}}{\left\langle V \right\rangle_{N_i,P,T}} - \left\langle V \right\rangle_{N_i,P,T} \tag{A39}$$

Combining Eqs. A27, A28 and A39 leads to

$$\begin{aligned} \bar{v}_{\mathrm{A}} &= \frac{(N+2)(N+1)/(\beta P)^2}{(N+1)/\beta P} - \frac{N+1}{\beta P} \\ &= \frac{1}{\beta P} \\ &= \frac{V}{N} \end{aligned} \tag{A40}$$

This expression is the well-known result for the partial molar volume of an ideal gas.

A.1.3 Partial Molar Enthalpy in the Continuous Fractional Component *NPT* Ensemble

The partition function of the NPT ensemble of a mixture of S components, expanded with a fractional molecule of component A equals [38, 42, 43, 142, 178]

$$Q_{\mathrm{CFC}NPT} = \beta P \left[\prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!}\right] \times \frac{1}{\Lambda_{\mathrm{A}}^3} \int_0^1 d\lambda \int \mathrm{d}V V^{N+1} \exp\left[-\beta P V\right] \times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda, V)] \tag{A41}$$

s indicates the scaled coordinates of molecules in the system, N is the total number of whole molecules, and the fractional molecule is distinguishable from whole molecules of the same type. $U_{\mathrm{frac}}^{\mathrm{A}}$ is the interaction potential of the fractional molecule with the rest of the system, and λ is a scaling factor. The value $\lambda = 0$ means that the fractional molecule of type A has no interactions with the surrounding molecules, and at $\lambda = 1$ the fractional molecule has full interactions with other molecules in the system *i.e.* the fractional molecule behaves as a whole molecule. Before deriving expressions for partial molar enthalpy and partial molar volume, we show that the chemical potential of component A corresponding to the conventional NPT ensemble (Eq. A7) can be computed in the CFCNPT ensemble and the result is identical. Following Refs. [38, 142], we can write

$$\begin{aligned} &\left\langle \frac{\delta_{\lambda=0}}{V/\Lambda_{\mathrm{A}}^3} \right\rangle_{\mathrm{CFC}NPT} = \\ &\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \left(\frac{1}{V/\Lambda_{\mathrm{A}}^3}\right) \exp\left[-\beta P V\right] \times \\ &\int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, \lambda \downarrow 0, s^N, V)] = \\ &\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \end{aligned} \tag{A42}$$

and

$$\left\langle \frac{\delta_{\lambda=1}}{N_{\mathrm{A}}+1} \right\rangle_{\mathrm{CFC}NPT} = \frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \frac{1}{N_{\mathrm{A}}+1} \exp\left[-\beta P V\right] \times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \uparrow 1, V)] \tag{A43}$$

$\delta_{\lambda=1}$ denotes the probability of λ approaching one, and $\delta_{\lambda=0}$ the probability of λ approaching zero. Combining the ensemble averages from Eqs. A42 and A43, we have

$$\frac{\left\langle \frac{\delta_{\lambda=1}}{N_{\mathrm{A}}+1} \right\rangle_{\mathrm{CFC}NPT}}{\left\langle \frac{\delta_{\lambda=0}}{V/\Lambda_{\mathrm{A}}^3} \right\rangle_{\mathrm{CFC}NPT}} = \frac{\prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \times \frac{1}{\Lambda_{\mathrm{A}}^3 (N_{\mathrm{A}}+1)} \int \mathrm{d}V V^{N+1} \exp\left[-\beta P V\right] \int \mathrm{d}s^{N+1} \exp[-\beta U(s^{N+1}, V)]}{\prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N} \exp\left[-\beta P V\right] \int \mathrm{d}s^{N} \exp[-\beta U(s^{N}, V)]} \tag{A44}$$

The ratio between these ensemble averages equals the ratio between the partition functions of the conventional NPT ensemble in which one system has an additional molecule of type A

$$\frac{\left\langle \frac{\delta_{\lambda=1}}{N_{\mathrm{A}}+1} \right\rangle_{\mathrm{CFC}NPT}}{\left\langle \frac{\delta_{\lambda=0}}{V/\Lambda_{\mathrm{A}}^3} \right\rangle_{\mathrm{CFC}NPT}} = \frac{Q_{N_A+1, N_{i\neq A}, P, T}}{Q_{N_i, P, T}} \tag{A45}$$

Therefore, with Eq. A45, we have shown that the chemical potential in the conventional NPT ensemble and the CFCNPT ensemble are the same by definition, so we have

$$\mu_{\mathrm{A}} = -\frac{1}{\beta} \ln \left(\frac{\left\langle \frac{\delta_{\lambda=1}}{N_A+1} \right\rangle_{\mathrm{CFC}NPT}}{\left\langle \frac{\delta_{\lambda=0}}{V/\Lambda_{\mathrm{A}}^3} \right\rangle_{\mathrm{CFC}NPT}} \right) \tag{A46}$$

For sufficiently large systems, λ_{A}, V, and N_{A} are uncorrelated [38], so therefore

$$\begin{aligned} \mu_{\mathrm{A}} &\approx -\frac{1}{\beta} \ln \left(\frac{\frac{1}{N_{\mathrm{A}}+1}}{\frac{1}{\left\langle V/\Lambda_{\mathrm{A}}^3 \right\rangle_{\mathrm{CFC}NPT}}} \times \frac{\langle \delta_{\lambda=1} \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0} \rangle_{\mathrm{CFC}NPT}} \right) \\ &\approx -\frac{1}{\beta} \ln \left(\frac{\langle V \rangle_{\mathrm{CFC}NPT}}{\Lambda_{\mathrm{A}}^3 \left(N_{\mathrm{A}} + 1 \right)} \right) - \frac{1}{\beta} \left(\frac{p(\lambda_{\mathrm{A}} \uparrow 1)}{p(\lambda_{\mathrm{A}} \downarrow 0)} \right) \end{aligned} \tag{A47}$$

In the above equations, $p(\lambda_{\mathrm{A}} \uparrow 1)$ denotes the probability of λ_{A} approaching one, and $p(\lambda_{\mathrm{A}} \downarrow 0)$ denotes the probability of λ_{A} approaching zero. The first term on the right hand side of Eq. A47 is the ideal gas part of the chemical potential of component A, and the second term is the excess chemical potential of component A. Therefore, the excess chemical potential is directly related to the probabilities of λ_{A} approaching one and zero [38, 142]. Similarly, we can write the partial molar enthalpy and partial molar volume in the conventional NPT ensemble as averages in the CFCNPT ensemble. Here, we derive expressions for the ensemble averages of Eq. A21

in terms of averages in the expanded ensemble. Starting from the term $\left\langle V \exp(-\beta \Delta U^{\mathrm{A}+}) \right\rangle_{N_i,P,T}$, we can write

$$
\begin{aligned}
&\frac{\langle \delta_{\lambda=1} \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0} / V \rangle_{\mathrm{CFC}NPT}} = \\
&\frac{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \times \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \uparrow 1, V)] \end{array} \right)}{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} V^{-1} \times \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N \lambda \downarrow 0, V)] \end{array} \right)} = \\
&\frac{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N V \times \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \uparrow 1, V)] \end{array} \right)}{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)]} \\
&= \left\langle V \exp(-\beta \Delta U^{\mathrm{A}+}) \right\rangle_{N_i,P,T}
\end{aligned}
\tag{A48}
$$

For the second ensemble average $\left\langle U(s^N,V)+PV\right\rangle_{N_i,P,T}$ in Eq. A21, we can start from

$$
\begin{aligned}
&\frac{\langle \delta_{\lambda=0}\,(U/V\ +P)\rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V\ \rangle_{\mathrm{CFC}NPT}} = \\
&\frac{\frac{\beta P}{Q_{\mathrm{CFC}NPT}}\prod_{i=1}^{S}\frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i}N_i!}\int \mathrm{d}V V^{N+1}\left(\begin{array}{l}\left(U(s^N,V)/V\ +P\right)\exp[-\beta PV]\\ \times\int \mathrm{d}s^N \exp[-\beta U(s^N,V)]\\ \times\int \mathrm{d}s^{\mathrm{A}}_{\mathrm{frac}} \exp[-\beta U^{\mathrm{A}}_{\mathrm{frac}}(s^{\mathrm{A}}_{\mathrm{frac}},s^N,\lambda\downarrow 0,V)]\end{array}\right)}{\frac{\beta P}{Q_{\mathrm{CFC}NPT}}\prod_{i=1}^{S}\frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i}N_i!}\int \mathrm{d}V V^{N+1}\left(\begin{array}{l}V^{-1}\exp[-\beta PV]\\ \times\int \mathrm{d}s^N \exp[-\beta U(s^N,V)]\\ \times\int \mathrm{d}s^{\mathrm{A}}_{\mathrm{frac}} \exp[-\beta U^{\mathrm{A}}_{\mathrm{frac}}(s^{\mathrm{A}}_{\mathrm{frac}},s^N,\lambda\downarrow 0,V)]\end{array}\right)} \\
&= \frac{\beta P\prod_{i=1}^{S}\frac{1}{\Lambda_i^{3N_i}N_i!}\int \mathrm{d}V V^{N}\left(\begin{array}{l}\left(U(s^N,V)\ +PV\right)\exp[-\beta PV]\\ \times\int \mathrm{d}s^N \exp[-\beta U(s^N,V)]\end{array}\right)}{\beta P\prod_{i=1}^{S}\frac{1}{\Lambda_i^{3N_i}N_i!}\int \mathrm{d}V V^{N}\exp[-\beta PV]\int \mathrm{d}s^N \exp[-\beta U(s^N,V)]} \\
&= \left\langle U(s^N,V)+PV\right\rangle_{N_i,P,T}
\end{aligned}
\tag{A49}
$$

For the third ensemble average $\left\langle (\Delta U^{\mathrm{A}+} + U(s^N, V) + PV)V \exp[-\beta \Delta U^{\mathrm{A}+}] \right\rangle_{N_i,P,T}$ in Eq. A21, we have

$$\frac{\left\langle \delta_{\lambda=1} \left(U^{\mathrm{A}}_{\mathrm{frac}} + U + PV \right) \right\rangle_{\mathrm{CFC}NPT}}{\left\langle \delta_{\lambda=0} / V \right\rangle_{\mathrm{CFC}NPT}} = \tag{A50}$$

$$\frac{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \times \left(\begin{array}{l} \left(U^{\mathrm{A}}_{\mathrm{frac}} + U(s^N, V) + PV \right) \exp[-\beta PV] \\ \times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}}_{\mathrm{frac}} \exp[-\beta U^{\mathrm{A}}_{\mathrm{frac}}(s^{\mathrm{A}}_{\mathrm{frac}}, s^N, \lambda \uparrow 1, V)] \end{array} \right)}{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \left(\begin{array}{l} V^{-1} \exp[-\beta PV] \\ \times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}}_{\mathrm{frac}} \exp[-\beta U^{\mathrm{A}}_{\mathrm{frac}}(s^{\mathrm{A}}_{\mathrm{frac}}, s^N, \lambda \downarrow 0, V)] \end{array} \right)} \tag{A51}$$

$$= \frac{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \left(\begin{array}{l} \left(U^{\mathrm{A}}_{\mathrm{frac}} + U(s^N, V) + PV \right) V \exp[-\beta PV] \\ \times \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s^{\mathrm{A}}_{\mathrm{frac}} \exp[-\beta U^{\mathrm{A}}_{\mathrm{frac}}(s^{\mathrm{A}}_{\mathrm{frac}}, s^N, \lambda \uparrow 1, V)] \end{array} \right)}{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \exp[-\beta PV] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)]} \tag{A52}$$

Similar to Eq. A22, we can define the total energy of the system at $\lambda = 1$

$$U(s^{N+1}, V) = U^{\mathrm{A}}_{\mathrm{frac}} + U(s^N, V) \tag{A53}$$

$U^{\mathrm{A}}_{\mathrm{frac}}$ is the interaction potential of the fractional molecule with the surrounding molecules. U is the total energy of the system including the fractional molecule which has full interactions with the rest of the system at $\lambda = 1$. The ratio in Eq. A52 equals the ensemble average in the conventional NPT ensemble as shown in Eq. A21. Combining Eqs. A50 and A52 leads to

$$\frac{\left\langle \delta_{\lambda=1} (U + PV) \right\rangle_{\mathrm{CFC}NPT}}{\left\langle \delta_{\lambda=0} / V \right\rangle_{\mathrm{CFC}NPT}} = \left\langle (U(s^{N+1}, V) + PV)V \exp[-\beta \Delta U^{\mathrm{A}+}] \right\rangle_{N_i,P,T} \tag{A54}$$

Therefore, the two ensemble averages yield identical results. Combining Eqs. A48, A49 and A54 with Eq. A21 yields

$$\begin{aligned}\bar{h}_{\mathrm{A}}^{\mathrm{ex}} &= -\frac{1}{\beta} + \frac{\frac{\langle \delta_{\lambda=1}(U+PV)\rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V\ \rangle_{\mathrm{CFC}NPT}}}{\frac{\langle \delta_{\lambda=1}\rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V\ \rangle_{\mathrm{CFC}NPT}}} - \frac{\langle \delta_{\lambda=0}\,(U/V\ +P)\rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V\ \rangle_{\mathrm{CFC}NPT}} \\ &= -\frac{1}{\beta} + \frac{\langle \delta_{\lambda=1}\,(U+PV)\rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=1}\rangle_{\mathrm{CFC}NPT}} - \frac{\langle \delta_{\lambda=0}\,(U/V\ +P)\rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V\ \rangle_{\mathrm{CFC}NPT}} \\ &= -\frac{1}{\beta} + \langle H\,(\lambda_{\mathrm{A}} \uparrow 1)\rangle_{\mathrm{CFC}NPT} - \frac{\langle H/V\,(\lambda_{\mathrm{A}} \downarrow 0)\ \rangle_{\mathrm{CFC}NPT}}{\langle 1/V\,(\lambda_{\mathrm{A}} \downarrow 0)\ \rangle_{\mathrm{CFC}NPT}}\end{aligned} \tag{A55}$$

$\langle H(\lambda_{\mathrm{A}} \uparrow 1)\rangle_{\mathrm{CFC}NPT}$ is ensemble average enthalpy of the system in the limit at which λ_{A} approaches one. $\langle H/V\ (\lambda_{\mathrm{A}} \downarrow 0)\rangle_{\mathrm{CFC}NPT}$ is ensemble average of the ratio between the total enthalpy and the volume of the system in the limit at which λ_{A} approaches zero. It is important to note that the assumption $\langle H\,(\lambda_{\mathrm{A}} \downarrow 0)\rangle \approx \langle H/V(\lambda_{\mathrm{A}} \downarrow 0)\rangle / \langle 1/V(\lambda_{\mathrm{A}} \downarrow 0)\rangle$ leads to an incorrect answer for an ideal gas as can be seen from Eqs. A24 and A29. A typical plot for ensemble averages in Eq. A55 is shown for LJ system in Fig. A1

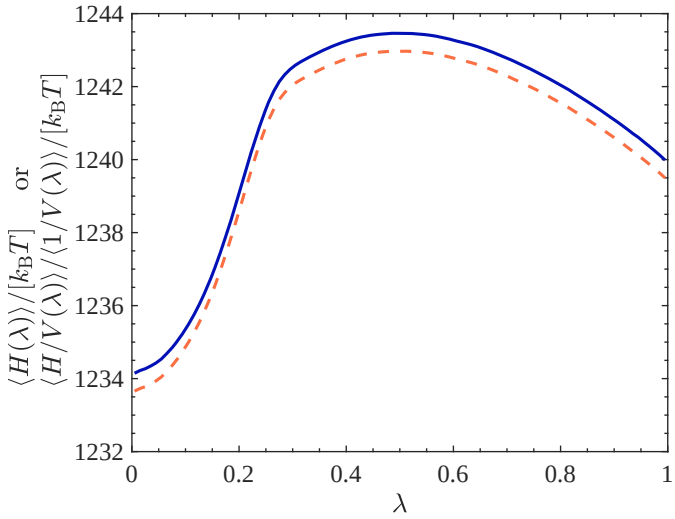


Figure A1: Ensemble average $\langle H(\lambda)\rangle$ (line) and ensemble average $\langle H/V\rangle(\lambda)/\langle 1/V\rangle(\lambda)$ (dashed line) in the CFC*NPT* ensemble simulation of a binary color mixture (50%-50%) consisting of 200 molecules. Temperature, reduced pressure and reduced density of the mixture are $T^* = 2$, $P^* = 9$ and $\langle \rho^* \rangle = 0.880$, respectively.

A.1.4 Partial Molar Volume in the Continuous Fractional Component *NPT* Ensemble

Similar to the partial molar enthalpy, we can write the averages of Eq. A38 in terms of averages in the expanded NPT ensemble. We already know that the denominator of Eq. A38 can be expressed as an ensemble average in the expanded NPT ensemble, as shown in Eq. A48. For the ensemble average in the nominator of equation Eq. A38, we can start from

$$
\begin{aligned}
&\frac{\langle \delta_{\lambda=1} V \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0} / V \, \rangle_{\mathrm{CFC}NPT}} = \\
&\frac{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} V \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \uparrow 1, V)] \end{array} \right)}{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} V^{-1} \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \downarrow 0, V)] \end{array} \right)} = \\
&\frac{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N V^2 \times \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \uparrow 1, V)] \end{array} \right)}{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \times \left(\exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \right)} = \\
&\left\langle V^2 \exp\left[-\beta \Delta U^{\mathrm{A}+}\right] \right\rangle_{N_i, P, T}
\end{aligned}
\tag{A56}
$$

For the second term on the right hand side of Eq. A38, we can start from

$$\frac{\langle \delta_{\lambda=0} \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V \rangle_{\mathrm{CFC}NPT}} = \frac{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \downarrow 0, V)] \end{array} \right)}{\frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda_{\mathrm{A}}^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} V^{-1} \left(\begin{array}{l} \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)] \\ \times \int \mathrm{d}s_{\mathrm{frac}}^{\mathrm{A}} \exp[-\beta U_{\mathrm{frac}}^{\mathrm{A}}(s_{\mathrm{frac}}^{\mathrm{A}}, s^N, \lambda \downarrow 0, V)] \end{array} \right)}$$

$$= \frac{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N V \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)]}{\beta P \prod_{i=1}^{S} \frac{1}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^N \exp\left[-\beta P V\right] \int \mathrm{d}s^N \exp[-\beta U(s^N, V)]}$$

$$= \langle V \rangle_{N_i, P, T} \tag{A57}$$

Substituting Eqs. A48, A56 and A57 in Eq. A38 leads to

$$\begin{aligned} \bar{v}_{\mathrm{A}} &= \frac{\frac{\langle \delta_{\lambda=1} V \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V \rangle_{\mathrm{CFC}NPT}}}{\frac{\langle \delta_{\lambda=1} \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V \rangle_{\mathrm{CFC}NPT}}} - \frac{\langle \delta_{\lambda=0} \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V \rangle_{\mathrm{CFC}NPT}} \\ &= \frac{\langle \delta_{\lambda=1} V \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=1} \rangle_{\mathrm{CFC}NPT}} - \frac{\langle \delta_{\lambda=0} \rangle_{\mathrm{CFC}NPT}}{\langle \delta_{\lambda=0}/V \rangle_{\mathrm{CFC}NPT}} \end{aligned} \tag{A58}$$

We can write

$$\bar{v}_{\mathrm{A}} = \langle V(\lambda_{\mathrm{A}} \uparrow 1) \rangle_{\mathrm{CFC}NPT} - \langle 1/V \ (\lambda_{\mathrm{A}} \downarrow 0) \rangle_{\mathrm{CFC}NPT}^{-1} \tag{A59}$$

$\langle V(\lambda_{\mathrm{A}} \uparrow 1) \rangle_{\mathrm{CFC}NPT}$ is the ensemble average of the volume when λ_{A} approaches one, and $\langle 1/V \ (\lambda_{\mathrm{A}} \downarrow 0) \rangle_{\mathrm{CFC}NPT}$ is the ensemble average of the inverse of the volume when λ_{A} approaches zero. In Eq. A59, one may be tempted to assume that $\langle 1/V \ (\lambda_{\mathrm{A}} \downarrow 0) \rangle_{\mathrm{CFC}NPT}^{-1} \approx \langle V(\lambda_{\mathrm{A}} \downarrow 0) \rangle_{\mathrm{CFC}NPT}$. However, this assumption leads to an incorrect answer for an ideal gas as can be seen from Eqs. A38 and A39. A typical plot for ensemble averages in Eq. A59 is shown for LJ system in Fig. A2

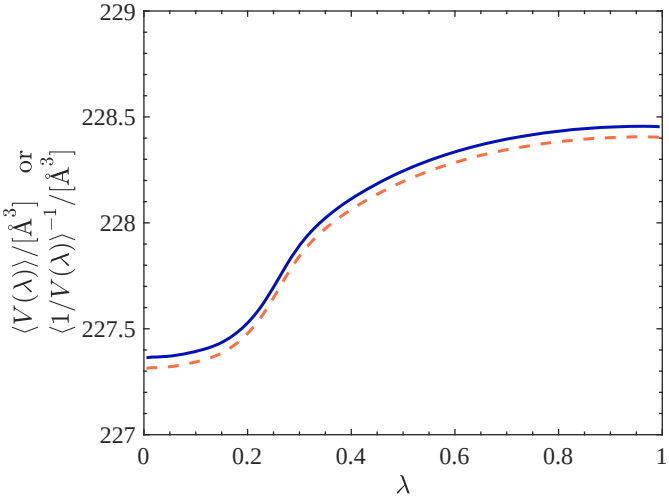


Figure A2: Ensemble average $\langle V(\lambda)\rangle$ (line) and ensemble average $\langle 1/V\rangle^{-1}(\lambda)$ (dashed line) in the CFC*NPT* ensemble simulation of a binary color mixture (50%-50%) consisting of 200 molecules. Temperature, reduced pressure and reduced density of the mixture are $T^* = 2$, $P^* = 9$ and $\langle \rho^* \rangle = 0.880$, respectively.

A.2 Partition Functions of Isolated Molecules

Molecular partition functions are used to calculate thermochemical properties such as the internal energy, entropy, chemical potential, heat capacity, etc. [4, 513]. These quantities can be obtained from different sources including available thermochemistry data or quantum calculations, each with a well-defined choice of reference state for zero of energy. The main purpose of this section is to explain how reference states for energy calculations can be chosen consistently using different data sets or computer programs. It is assumed that the reader is partially familiar with statistical mechanics and basic concepts in computational chemistry, and this section should not be blindly used as a "cook book" for computation chemistry problems. The theoretical part in this section is based on the Physical Chemistry book by McQuarrie [4] and Essential Statistical Thermodynamics from Computational Chemistry Comparison and Benchmark Data Base (CCCBDB) [340]. In this section, commonly used thermodynamic data sets (JANAF tables [334, 514]) and the Gaussian09 software [513] are used to provide examples on how to calculate molecular partition functions for nitrogen, hydrogen and ammonia. The results are used in chapter 5 to solve the chemical equilibrium in MC simulations of the Haber-Bosch process in the reaction ensemble [20]. To compute the molecular partition function, a complete set of molecular energy levels is required, which is almost never available [340, 515]. As an approximation, the energy of a molecule can be estimated by decoupling translational, vibrational, rotational and electronic contributions, which means that different energy contributions are unaffected by each other. Using this approximation, the molecular partition function of an isolated molecule can be written as [4]

$$q(V,T) = q_{\text{trans}}(V,T)q_{\text{rot}}(T)q_{\text{vib}}(T)q_{\text{elec}}(T) \tag{A60}$$

in which the terms on the right hand side of Eq. A60 denote translational, rotational, vibrational and electronic partition functions. Here, it is explained briefly how different contributions in Eq. A60 are calculated to obtain $q(V,T)$.

Translational Partition Function: the translational contribution is obtained from [4]

$$q_{\text{trans}}(V,T) = \frac{V}{\Lambda^3} \qquad \Lambda = \frac{h}{\sqrt{2\pi M k_{\text{B}} T}} \tag{A61}$$

in which Λ is the thermal de Broglie wavelength of the molecule, k_{B} is the Boltzmann constant, h is the plank constant and T is the temperature, and the mass M is the sum of all atomic masses within the molecule. The choice of volume in Eq. A61 may be defined differently depending on the data set or software. In Gaussian09 [513], the ideal gas law is used to calculate the volume at atmospheric pressure: $V = k_{\text{B}}T/P$ [513]. It is shown later in this section that only the temperature dependent part of the partition function is used to obtain the chemical potential. Therefore, the choice of volume in Eq. A61 does not affect the chemical potential.

Electronic Partition Function: For a monotonic ideal gas, the electronic partition function is obtained using

$$q_{\text{elec}}(T) = \sum_i g_{\text{e}i} \exp[-\beta \varepsilon_{\text{e}i}] \tag{A62}$$

in which $g_{\text{e}i}$ and $\varepsilon_{\text{e}i}$ are the degeneracy and the energy of the i^{th} electronic level, respectively. The degeneracy of electronic levels is determined by the spin multiplicity [513], $2S + 1$, in which $S =$ is the net electron spin or total spin quantum number [340]. For a monoatomic ideal gas, the zero of electronic energy is fixed at the ground state ($\varepsilon_{\text{e}0} = 0$). Normally, only the first and the second term in the summation in Eq. A62 are considered for the electronic contribution to the partition function [4]. This is because the excited electronic energy levels at typical temperatures are around tens of thousands of wave numbers [4], which means that the excited energy levels for most substances are nearly inaccessible even at temperatures up to $T = 1000$ K. As an illustrative example, hydrogen atom has a first excited state $^2P_{1/2}$ with the energy 82258 in units of cm^{-1} [4]. At $T = 1000$ K, the contribution of the second term in Eq. A62 is in order of 10^{-52}. When considering excited electronic states, the translation, vibrational and rotational

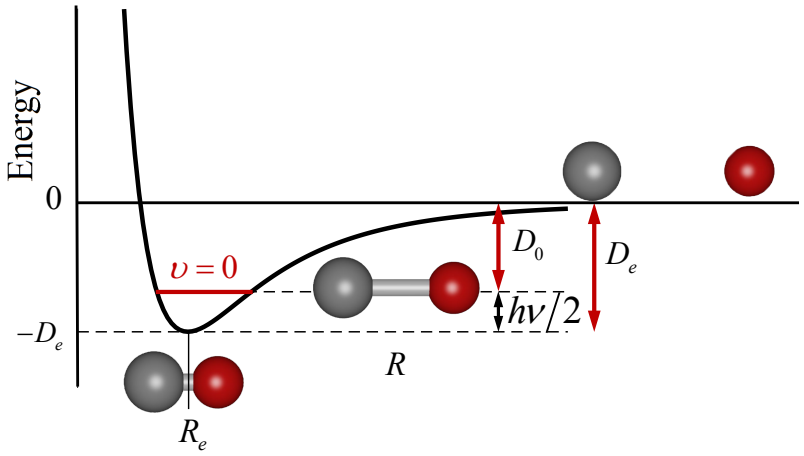


Figure A3: Electronic potential curve of a diatomic molecule as a function of internuclear distance R. $-D_\mathrm{e}$ is the energy of the molecule in the ground state (lowest value), relative to infinitely dissociated atoms. The vibrational energy of the ground state ($\upsilon = 0$) is called zero-point energy which equals $h\nu/2$. $-D_0$ is the corresponding dissociation (atomization) energy when the zero-point energy is selected as zero of energy.

contirbutions can be approximated as those in the electronic ground state if no other data is available [515]. For diatomic or polyatomic ideal gas, the arbitrary zero of electronic energy is taken to be the infinitely dissociated atoms at rest in their ground electronic state [4]. The ground electronic level for a diatomic ideal gas is shown Fig. A3. It is shown in Fig. A3 that the energy difference between the minimum of the internuclear potential well fully dissociated limit (bare nuclea and free electrons) is denoted with D_e, which is the dissociation energy of the molecule. The energy of the ground electronic state is $\varepsilon_{\mathrm{e}1} = -D_\mathrm{e}$. Note the same definition of the molecular ground state holds for a polyatomic molecule. The electronic partition function for a diatomic or polyatomic ideal gas molecule is

$$q_\mathrm{elec}(T) = g_{\mathrm{e}1} \exp[\beta D_\mathrm{e}] + \cdots \tag{A63}$$

in which the contributions from exited electronic levels are not considered at ordinary temperatures. It should be noted that in Gaussian09, the

electronic contribution only contains the degeneracy of the electronic ground state [513], which means

$$q_{\text{elec}}(T) = g_{\text{e1}} \tag{A64}$$

The distinction between Eqs. A63 and A64 is very important when obtaining the molecular partition function from Gaussian09, as the energy reference for the electronic contribution is different by D_{e}.

Vibrational partition function: Under the harmonic-oscillator approximation, the accessible vibrational energy levels of a diatomic molecule, relative to the bottom of the internuclear potential well, as shown in Fig. A3 are obtained from [4]

$$\varepsilon_{\upsilon} = (\upsilon + 1/2\ )\, h\nu \qquad \upsilon = 0, 1, 2, \cdots \tag{A65}$$

in which $\nu = (k/\mu^{1/2}\ )/2\pi$ is the frequency of vibration. k is the force constant of the molecule, υ is the quantum number and μ is the reduced mass. A zero of energy is also required for the vibrational energy levels. Two choices are common for zero of vibrational energy: *(1)* the minimum of the internuclear potential energy curve as shown in Fig. A3, which means $\varepsilon_0 = h\nu/2$. *(2)* The energy of the ground vibrational state is set to zero, which means $\varepsilon_0 = 0$. As shown in Fig. A3, the corresponding dissociation energy is $D_0 = D_{\text{e}} - h\nu/2$ which means that the zero-point energy is taken to be the ground vibrational state. For a polyatomic molecule, the vibrational motion is described using an independent harmonic oscillator, in terms of normal coordinates. The vibrational energies are written as

$$\varepsilon_{\upsilon} = \sum_{j=1}^{\alpha} (\upsilon_j + 1/2\ )\, h\nu_j \qquad \upsilon_j = 0, 1, 2, \cdots \tag{A66}$$

in which j denotes the j^{th} normal mode, and α is the vibrational degree of freedom. For a linear molecule with n atoms, $\alpha = 3n - 5$, and a nonlinear molecule, $\alpha = 3n - 6$. Similar to a diatomic molecule. Considering the first

choice for the zero of vibrational energy, the vibrational partition function of a diatomic molecule equals [4]:

$$q_{\mathrm{vib}}^{\mathrm{BOT}}(T) = \prod_{j=1}^{\alpha} \frac{\exp\left[-\Theta_{\mathrm{vib},j}/2T\right]}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\right]} \tag{A67}$$

in which $\Theta_j = h\nu_j/k_{\mathrm{B}}$ is the characteristic vibrational temperature corresponding to the j^{th} normal mode. The notation "BOT" refers to the zero of energy at the bottom of the internuclear potential well ($-D_{\mathrm{e}}$). This notation is used in Gaussian09 to report the vibrational partition function using Eq. A67. The exponential term in the nominator on the right hand side of Eq. A67, is the contribution of the ground vibrational state [4]. Note that the vibrational energy levels are non-degenerate. Considering the ground vibrational state as zero of vibrational energy, the vibrational partition function is obtained using [513]

$$q_{\mathrm{vib}}^{\mathrm{V=0}}(T) = \prod_{j=1}^{\alpha} \frac{1}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\right]} \tag{A68}$$

The notation "V=0" refers to the vibrational ground state is taken to be the arbitrary zero-point energy. This notation is also used in Gaussian09 to report the vibrational partition function using Eq. A68. When the vibrational ground state is taken to be zero of energy, the corresponding dissociation energy, for a general case of a polyatomic molecule relative to the is obtained from, see Fig. A3:

$$D_0 = D_{\mathrm{e}} - \sum_{j=1}^{\alpha} h\nu_j/2 \tag{A69}$$

The second term on the right hand side of Eq. A69 is the zero-point vibrational energy at zero Kelvin, usually abbreviated as ZPE or ZPVE [340, 513]. Normally, empirical scaling factors are used to scale the vibrational frequencies obtained from ab inito calculations to correct for vibrational

anharmonicity and incomplete electron correlations [340, 516]. The empirical scaling factors are found in literature and reference data bases [340, 516].

Rotational partition function: Under the rigid-rotator approximation, the rotational energy levels of a diatomic molecule are written as [4]

$$\varepsilon_J = \frac{\hbar^2 J(J+1)}{2I} \qquad J = 0, 1, 2, \cdots \tag{A70}$$

in which I is the moment of inertia of the rotor. The degeneracy of each rotational level equals $g_J = 2J + 1$. As shown in Eq. A70, rotation of a rigid molecule is quantized and only certain rotational energy levels can be occupied [515]. Based on Eq. A70, a convenient choice for the zero of rotational energy is the level $J = 0$. The energies and degeneracies of a linear polyatomic molecule are the same as a diatomic molecule [4]. The expression for the rotational partition function of a linear polyatomic molecule or a diatomic molecule is [4]

$$q_{\text{rot}}(T) = \frac{T}{\sigma \Theta_{\text{rot}}} \tag{A71}$$

in which $\Theta_{\text{rot}} = h^2/8\pi^2 I k_{\text{B}}$ is the characteristic rotational temperature, and σ is the symmetry number or rotational symmetry number of the molecule. The symmetry number is the number of unique orientations of the rigid molecule obtained by interchanging identical atoms [4, 340, 513, 515]. For a heteronuclear diatomic or unsymmetrical molecule, $\sigma = 1$, and for a homonuclear diatomic or symmetrical molecule $\sigma = 2$. The rotational symmetry number can be identified using group theory if the point group of the molecule is known. The symmetry number corresponding to different point groups is provided in Table A1 [340]. As a reference, a list of molecules and corresponding symmetry numbers is found on the website of CCCBDB [340]. It is also possible to obtain σ from counting manually the number of unique orientations of molecule. It is highly recommended to double-check the point group of the molecule from Gaussian09 output. For

Table A1: Point groups and the corresponding symmetry number. The values in this table are obtained from the Computational Chemistry Comparison and Benchmark DataBase (CCCBDB) website [340]

Group	σ
$C_1, C_i, C_s, C_{\infty\nu}$	1
$C_n, C_{n\nu}, C_{nh}$	n
$D_{\infty h}$	2
D_n, D_{nh}, D_{nd}	$2n$
T, T_d	12
S_n	$n/2$
O_h	24
I_h	60

non-linear, polyatomic molecules, the partition function is obtained from [4]

$$q_{\mathrm{rot}}(T) = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right) \tag{A72}$$

in which $\Theta_{\mathrm{rot,j}}$ is the characteristic rotational temperature corresponding to the three principal moments of inertia; A,B,C [4]. If $\Theta_{\mathrm{rot,A}} = \Theta_{\mathrm{rot,B}} = \Theta_{\mathrm{rot,C}}$ the molecule is a symmetric top, and if $\Theta_{\mathrm{rot,A}} \neq \Theta_{\mathrm{rot,B}} \neq \Theta_{\mathrm{rot,C}}$ the molecule is a asymmetric top. The molecule is called symmetric top if $\Theta_{\mathrm{rot,A}} = \Theta_{\mathrm{rot,B}} \neq \Theta_{\mathrm{rot,C}}$ [4].

Molecular partition function: Combining Eqs. A60, A61, A63, A67 and A72 , the molecular partition function for a polyatomic ideal gas equals

$$\begin{aligned} q(V,T) = \frac{V}{\Lambda^3}\cdot\frac{\pi^{1/2}}{\sigma}&\left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}}\right) \\ &\times\prod_{j=1}^{\alpha}\frac{\exp\left[-\Theta_{\mathrm{vib},j}/2T\right]}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\right]}\cdot g_{\mathrm{e1}}\exp[D_{\mathrm{e}}/k_{\mathrm{B}}T] \end{aligned} \tag{A73}$$

in which the zero-point energy for the vibrational ground state is the bottom of the internuclear potetnial well, and the zero of electronic energy is the

dissociated atoms in the ground state. This means that the ground state electronic energy equals $-D_\mathrm{e}$ as shown in Fig. A3. In the thermochemistry output of Gaussian09, the energy of the electronic ground state is taken to be zero, and the corresponding partition function denoted by "Q Total Bot" is the same as Eq. A73 except for the factor of $\exp[\beta D_\mathrm{e}]$ [513]. It is important to consider this factor when evaluating equilibrium constant of reaction from molecular partition functions. This will be highlighted later in this section. The molecular partition function is also obtained by combining Eqs. A60, A61, A63, A68, A69 and A72

$$\begin{aligned} q(V,T) = \frac{V}{\Lambda^3} \cdot \frac{\pi^{1/2}}{\sigma} \left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}} \right) \\ \times \prod_{j=1}^{\alpha} \frac{1}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\right]} \cdot g_{\mathrm{e1}} \exp[D_0/k_\mathrm{B}T\,] \end{aligned} \tag{A74}$$

in which the zero-point for the vibrational energy is the ground state energy, and the corresponding the ground state electronic energy is $-D_0$ as shown in Fig. A3. In thermodynamic tables, it is however common to take the ground state of molecule (vibrational and electronic) as zero of energy, instead of the dissociated atoms. This is performed by factoring out the contribution of the ground state energy of the molecule from the partition function:

$$\begin{aligned} q(V,T) &= \sum_j \exp\left[-\varepsilon_j/k_\mathrm{B}T\,\right] \\ &= \exp\left[-\varepsilon_0/k_\mathrm{B}T\,\right] \sum_j \exp\left[-(\varepsilon_j-\varepsilon_0)/k_\mathrm{B}T\,\right] \\ &= \exp\left[-\varepsilon_0/k_\mathrm{B}T\,\right] q^0(V,T) \end{aligned} \tag{A75}$$

The notation $q^0(V,T)$ for the molecular partition functions highlights the fact that the ground state energy of the molecule is zero (instead of $-D_0$). By comparing Eqs. A74 and A75, it is clear that $\varepsilon_0 = -D_0$. For a general

case of a polyatomic molecule, we have

$$q^0(V,T) = \frac{V}{\Lambda^3} \cdot \frac{\pi^{1/2}}{\sigma} \left(\frac{T^3}{\Theta_{\mathrm{rot,A}}\Theta_{\mathrm{rot,B}}\Theta_{\mathrm{rot,C}}} \right) \times \prod_{j=1}^{\alpha} \frac{1}{1-\exp\left[-\Theta_{\mathrm{vib},j}/T\right]} \cdot g_{\mathrm{e1}} \tag{A76}$$

$q^0(V,T)$ is reported in the thermochemistry output of Gaussian09 as "Q Total V=0" [513].

Atomization energy (D_0) : The experimental values for the atomization energy of several molecules are reported in the NIST database [340] and other thermodynamic references [4]. The atomization energy can also be obtained using ab inito calculation of molecular energies. As shown in Fig. A3, for a diatomic molecule, D_0 is obtained from the difference between the molecular energy (vibrational and electronic) in the ground state and the dissociated atoms in their respective ground state. The same approach holds for a polyatomic molecule. Using the notation as in the Gaussian09 manual [513], the atomization energy of a polyatomic molecule is obtained from

$$D_0 = \sum_{i=1}^{N} y_i(\varepsilon_{\mathrm{e},i}) - (\varepsilon_{\mathrm{e}} + \varepsilon_{\mathrm{ZPE}}) \tag{A77}$$

in which y_i indicates the number of atoms of kind i in the molecule. $\varepsilon_{\mathrm{e},i}$ is the electronic energy of the i^{th} atom (dissociated). The second term on the right hand side is the sum of electronic and ZPE of the molecule (vibrational energy in the ground state). It should be noted that computing the atomization energy to chemical accuracy (usually defined as 1 kcal/mol) is not trivial. Advanced methods (*e.g.* Gaussian-n composite methods [517, 518] are recommended for accurate calculation of atomization energies. This is beyond the scope of this thesis. From JANAF tables, the atomization energy is obtained from the difference between the enthalpy of formation of

the molecule and the dissociated atoms:

$$D_0 = \sum_{i=1}^{N} y_i \Delta_f H_i^\circ(0\text{ K}) - \Delta_f H^\circ(0\text{ K}) \tag{A78}$$

in which $\Delta_f H_i^\circ(0\text{ K})$ is the enthalpy of formation of the i^{th} atom (dissociated) and $\Delta_f H^\circ(0\text{ K})$ is the enthalpy of formation of the molecule.

A.2.1 Chemical Potentials

In Ref. [4], it is shown that the standard chemical potential and the molecular partition function are related:

$$\mu^\circ(T) = -RT \ln \left[\left(\frac{q(V,T)}{V} \right) \frac{k_\text{B} T}{P^\circ} \right] \tag{A79}$$

$\mu^\circ(T)$ is the standard chemical potential of an ideal gas molecule at standard reference pressure ($P^\circ = 1$) bar. To tabulate the chemical potential in Eq. A79, a zero of energy is required. In thermodynamic tables, it is common to take the ground state energy of the molecule to be zero. This leads to

$$\mu^\circ(T) - E_0 = -RT \ln \left[\left(\frac{q^0(V,T)}{V} \right) \frac{k_\text{B} T}{P^\circ} \right] \tag{A80}$$

in which $E_0 = N_\text{A} \varepsilon_0$. Choosing this energy reference, we show that the E_0 is the standard molar enthalpy of molecule at $T = 0$ K ($E_0 = -D_0$). For an ideal gas, the enthalpy can be written in terms of the partition function:

$$\begin{aligned} H^\circ(T) &= U + RT \\ &= N k_\text{B} T^2 \left(\frac{\partial \ln q(V,T)}{\partial T} \right)_V + RT \\ &= N\varepsilon_0 + N k_\text{B} T^2 \left(\frac{\partial \ln q^0(V,T)}{\partial T} \right)_V + RT \end{aligned} \tag{A81}$$

Table A2: An example of a JANAF table entry [334, 514]. In the JANAF tables, the standard Gibbs energy (chemical potential) is referenced to the enthalpy at $T_\mathrm{r} = 298.15$ K. To obtain the chemical potential of a molecule in which the ground state energy is taken to be zero, the values for the Gibbs energy are shifted relative to the enthalpy at $T_\mathrm{r} = 0$ K, as described in Eq. A82.

Nitrogen Oxide	(NO_2)						N_1O_2 (g)
	$\mathbf{J.K^{-1}.mol^{-1}}$			$\mathbf{kJ.mol^{-1}}$			
$\boldsymbol{T}$/K	$\boldsymbol{C_P^\circ}$	$\boldsymbol{S^\circ}$	$-[\boldsymbol{G^\circ} - \boldsymbol{H^\circ}(\boldsymbol{T}_\mathrm{r})]/\boldsymbol{T}$	$\boldsymbol{H} - \boldsymbol{H^\circ}(\boldsymbol{T}_\mathrm{r})$	$\boldsymbol{\Delta_f H^\circ}$	$\boldsymbol{\Delta_f G^\circ}$	$\mathbf{\log K_f}$
0	0	0	INFINITE	-10.186	35.927	35.927	INFINITE
100	33.276	202.563	271.168	-6.861	34.898	39.963	-20.874
200	34.385	225.852	243.325	-3.495	33.897	45.422	-11.863
250	35.593	233.649	240.634	-1.746	33.46	48.355	-10.103
298.15	36.974	240.034	240.034	0	33.095	51.258	-8.98

Evaluating the enthalpy in Eq. A81 at $T = 0$ shows that for one mole of an ideal gas $H^\circ(0\ \mathrm{K}) = E_0 = N_\mathrm{A}\varepsilon_0$, in which N_A is the Avogadro number. The ideal gas partition function $q^0(V,T)$ in Eq. A81 is the same as in Eq. A76 which can be obtained by rearranging Eq. A79 or Eq. A80. $q^0(V,T)$ obtained using rigid rotator-harmonic oscillator approximation which agrees reasonably well with experiments [4]. To improve the accuracy, experimental data may be used to complement the theoretical calculations [4]. The combination of experimentally determined thermodynamic properties and theoretical calculations can be found in the Joint, Army, Navy Air Force (JANAF) tables [334, 514]. Thermodynamic functions and parameters including the Gibbs free energy, enthalpy and heat capacity are extensively tabulated in JANAF tables [334, 514]. A JANAF table entry for NO_2 is provided as an example in Table A2. Using JANAF tables, it is possible to calculate $q^0(V,T)$ in Eq. A80 without performing direct quantum mechanical calculations.

As we know for a pure component $\mu^\circ = G^\circ$. For a pure substance, $\mu^\circ(T) - E_0^\circ =$ in Eq. A80 is obtained from JANAF tables (denoted with $G^\circ(T) - H_0^\circ$). To obtain the chemical potential of a molecule in which the ground state energy is taken to be zero (see Eq. A80), the values for the Gibbs energy are shifted relative to the enthalpy at $T_\mathrm{r} = 0$ K. By rearranging

the data in the fourth and fifth columns (see Table A2):

$$\begin{aligned} -\frac{(G^\circ(T) - H^\circ(0\text{ K}))}{T} &= R\ln\left[\left(\frac{q^0(V,T)}{V}\right)\frac{k_\mathrm{B}T}{P^\circ}\right] \\ &= -\frac{(G^\circ(T) - H^\circ(298.15\text{ K}))}{T} + \frac{(H^\circ(0) - H^\circ(298.15\text{ K}))}{T} \end{aligned} \tag{A82}$$

The term on the left hand side of Eq. A82 is the chemical potential as shown in Eq. A80 divided by the temperature.

A.2.2 Chemical Equilibrium

For a general case of homogeneous gas phase chemical reaction, it is well-known that the Gibbs energy of reaction and the equilibrium constant are related to chemical potentials of reactants and products (at P°). For a multicomponent reacting mixture of S distinguishable components we have [4, 117]

$$\begin{aligned} \Delta G_r^\circ(T) &= \sum_{i=1}^{S} \nu_i \mu_i^\circ(T) \\ &= -RT\ln K(T) \\ &= -RT\ln\left[\prod_{i=1}^{S}\left(\frac{q(V,T)}{V}\right)\frac{k_\mathrm{B}T}{P^\circ}\right]^{\nu_i} \\ &= -RT\sum_{i=1}^{S}\nu_i\ln\left[\left(\frac{q(V,T)}{V}\right)\frac{k_\mathrm{B}T}{P^\circ}\right] \end{aligned} \tag{A83}$$

ν_i is the stoichiometric coefficient of component i, and $K(T)$ is the equilibrium constant of the reaction [4]. The reaction enthalpy at standard pressure is calculated directly from the Gibbs-Helmholtz equation [2, 5]

$$\left(\frac{\partial \Delta G_r^\circ/T}{\partial T}\right)_P = -\frac{\Delta H_r^\circ}{T^2} \tag{A84}$$

A.3 Ammonia Synthesis Reaction

The ideal gas partition functions are calculated for nitrogen, hydrogen, and ammonia using experimental thermochemistry data [4], JANAF tables [514] and quantum calculations using Gaussian09 [513]. The frequency analysis is performed using two different basis sets: B3LYP level of theory with a 6-31G** basis set, and MP2 level of theory with a 6-311G** basis set [513]. It should be noted that other software packages are also available for performing frequency analysis, such as ADF [519], Spartan [520] etc. For details about Gaussian09 input files, the reader is referred to the manual [513]. The characteristic vibrational and rotational temperatures of nitrogen, hydrogen, and ammonia obtained from experimental data and ab intio calculations are provided in Table A3. For atomization energies of components, only the experimental data provided by McQuarrie [4] are used and not the ones obtained from Gaussian09 [513]. Atomization energies can also be obtained from JANAF tables using Eq. A78. The results are shown in Table A4. One can easily see that computation of the atomization energies using a single basis set results in energy differences well above chemical accuracy [361]. Advanced methods such as (*e.g.* Gaussian-n composite methods [517, 518] are recommended for ab initio calculation of atomization energies which is beyond the scope of this thesis. The thermochemical data in Tables A3 and A4 are used to compute the partition functions of nitrogen, hydrogen and ammonia at temperatures between $T = 573$ K and $T = 873$ K. The results are presented in Table A5. Note that the corresponding vibrational and rotational partition functions for diatomic molecules were used for hydrogen and nitrogen, and for ammonia the corresponding vibrational and rotational partition functions were used.

Table A3: Characteristic vibrational and rotational temperatures of nitrogen, hydrogen and ammonia obtained from experimental data [4] and ab initio calculations using Gaussian09 [513]. In the table, Gaussian (1) denotes frequency calculations using B3LYP level of theory with a 6-31G** basis set, and Gaussian (2) denotes frequency calculations using MP2 level of theory with a 311G**basis set. To correct for vibrational anharmonicity, the vibrational frequencies obtained from Gaussian09 are scaled by empirical factors, 0.96 for the B3LYP/6-31G** and 0.95 [340, 516]. The numbers in parentheses indicate the degenacy of the modes.

	McQuarrie		Gaussian (1)		Gaussian (2)	
	Θ_{vib}/[K]	Θ_{rot}/[K]	Θ_{vib}/[K]	Θ_{rot}/[K]	Θ_{vib}/[K]	Θ_{rot}/[K]
N_2	3374	2.88	3393	2.83	2979	2.76
H_2	6332	85.2	6161	87.16	6196	88.29
NH_3	1360	13.60	1509	14.10	1542	14.24
	4800	13.60	4781	14.10	4819	14.24
	2330(2)	8.92	2340(2)	9.14	2292(2)	9.18
	4880(2)		4958(2)		5015(2)	

Table A4: Experimental atomization energies [4] and atomization energies computed for nitrogen, hydrogen, ammonia using Gaussian09 using the B3LYP level of theory with a 6-31G** basis set. [513]

Component	D_0 / [kJ/mol]	
	McQuarrie	Gaussian
N_2	941.6	917.6
H_2	432.1	432.1
NH_3	1158	1149.8

Table A5: Computed ideal gas partition functions of nitrogen, hydrogen, ammonia obtained as defined in this. The reported values are based on experimental data (McQuarrie and JANAF tables) [4, 514] and quantum computations using Gaussian09 [513]. In the table, Gaussian (1) denotes frequency calculations using B3LYP level of theory with a 6-31G** basis set, and Gaussian (2) denotes frequency calculations using MP2 level of theory with a 311G** basis set. It is important to note that for the values obtained from Gaussian09, the experimental atomization energies provided by McQuarrie [4] are used (see Table A4). One can easily see that using the atomization energies from Gaussian09, listed in Table A4, results in large deviations.

	q/V $[N_2]/[Å^{-3}]$		q/V $[H_2]/[Å^{-3}]$		q/V $[NH_3]/[Å^{-3}]$	
$T/[K]$	Gaussian (1)	Gaussian (2)	Gaussian (1)	Gaussian(2)	Gaussian (1)	Gaussian (2)
573	$2.65 \cdot 10^{90}$	$2.72 \cdot 10^{90}$	$5.94 \cdot 10^{40}$	$5.87 \cdot 10^{40}$	$1.40 \cdot 10^{110}$	$1.38 \cdot 10^{110}$
673	$7.00 \cdot 10^{77}$	$7.22 \cdot 10^{77}$	$1.25 \cdot 10^{35}$	$1.23 \cdot 10^{35}$	$5.01 \cdot 10^{94}$	$4.94 \cdot 10^{94}$
773	$3.50 \cdot 10^{68}$	$3.62 \cdot 10^{68}$	$8.09 \cdot 10^{30}$	$7.99 \cdot 10^{30}$	$1.94 \cdot 10^{83}$	$1.92 \cdot 10^{83}$
873	$2.46 \cdot 10^{61}$	$2.56 \cdot 10^{61}$	$4.96 \cdot 10^{27}$	$4.90 \cdot 10^{27}$	$3.33 \cdot 10^{74}$	$3.29 \cdot 10^{74}$
$T/[K]$	McQuarrie	JANAF	McQuarrie	JANAF	McQuarrie	JANAF
573	$2.60 \cdot 10^{90}$	$2.67 \cdot 10^{90}$	$6.08 \cdot 10^{40}$	$6.53 \cdot 10^{40}$	$1.50 \cdot 10^{110}$	$1.46 \cdot 10^{110}$
673	$6.89 \cdot 10^{77}$	$7.04 \cdot 10^{77}$	$1.27 \cdot 10^{35}$	$1.36 \cdot 10^{35}$	$5.42 \cdot 10^{94}$	$5.26 \cdot 10^{94}$
773	$3.44 \cdot 10^{68}$	$3.52 \cdot 10^{68}$	$8.27 \cdot 10^{30}$	$8.79 \cdot 10^{30}$	$2.12 \cdot 10^{83}$	$2.06 \cdot 10^{83}$
873	$2.42 \cdot 10^{61}$	$2.48 \cdot 10^{61}$	$5.07 \cdot 10^{27}$	$5.38 \cdot 10^{27}$	$3.65 \cdot 10^{74}$	$3.58 \cdot 10^{74}$

A.4 Combined SRM with FA decomposition

The Gibbs free energy (or chemical potential) of each component at standard pressure ($P^\circ = 1$ bar) can be evaluated from its ideal gas partition as shown in Eq. A79. The temperature dependent part of the ideal gas partition function, $q(V,T)/V$, can be computed using Eq. A74. In this work, we have used experimental data from the NIST database [340] to compute the ideal gas partition function. Alternatively, quantum mechanical ab initio packages e.g. Gaussian [513] can be used to evaluate the terms in Eq. A74 [4]. Chemical potentials, μ° , of methane, water, carbon monoxide, hydrogen, carbon dioxide and Formic Acid (FA) are listed in table Table A6.

Table A6: Gibbs free energy or μ°, in kJ $\cdot$ mol^{-1}, of carbon monoxide, water, carbon dioxide, hydrogen, formic acid and methane at $P^\circ = 1$ bar, based on Eqs. A74 and A79.

T [K]	μ°_{CO}	$\mu^\circ_{H_2O}$	$\mu^\circ_{CO_2}$	$\mu^\circ_{H_2}$	μ°_{HCOOH}	$\mu^\circ_{CH_4}$
800	-1229.8	-1068.9	-1771.9	-536.1	-2212.9	-1793.5
825	-1235.5	-1074.5	-1778.4	-540.1	-2220.7	-1799.3
850	-1241.2	-1080.1	-1784.9	-544.1	-2228.5	-1805.2
875	-1246.9	-1085.8	-1791.4	-548.1	-2236.4	-1811.2
900	-1252.7	-1091.5	-1798.0	-552.2	-2244.3	-1817.2
925	-1258.5	-1097.2	-1804.6	-556.3	-2252.3	-1823.2
950	-1264.3	-1102.9	-1811.3	-560.4	-2260.3	-1829.3
975	-1270.1	-1108.7	-1818.0	-564.5	-2268.4	-1835.4
1000	-1276.0	-1114.5	-1824.7	-568.6	-2276.6	-1841.6
1050	-1287.7	-1126.2	-1838.2	-576.9	-2293.0	-1854.1
1100	-1299.6	-1138.0	-1851.9	-585.3	-2309.7	-1866.7
1150	-1311.5	-1149.8	-1865.7	-593.8	-2326.5	-1879.6
1200	-1323.5	-1161.8	-1879.6	-602.3	-2343.5	-1892.6
1250	-1335.5	-1173.9	-1893.6	-610.9	-2360.8	-1905.8
1300	-1347.7	-1186.0	-1907.8	-619.6	-2378.2	-1919.1
1350	-1359.9	-1198.3	-1922.0	-628.3	-2395.7	-1932.6
1400	-1372.1	-1210.6	-1936.4	-637.0	-2413.5	-1946.2
1450	-1384.5	-1223.0	-1950.9	-645.9	-2431.4	-1960.0
1500	-1396.8	-1235.5	-1965.4	-654.7	-2449.5	-1973.9

A.5 Equation of State Modeling

Peng-Robinson Equation of State

The Peng-Robinson Equation of State (PR-EoS) [151] is used to compute partial molar properties:

$$P = \frac{RT}{\upsilon_m - b_m} - \frac{a_m}{\upsilon_m(\upsilon_m + b_m) + b_m(\upsilon_m - b_m)} \tag{A85}$$

υ_{m} is the molar volume of the mixture per mole of the mixture. a_{m} and b_{m} are defined based on pure component parameters (a_i and b_i) and van der Waals mixing rules [330, 382]:

$$a_{\mathrm{m}} = \sum_{i=1}^{S} \sum_{j=1}^{S} x_i x_j a_{ij} \tag{A86}$$

$$b_{\mathrm{m}} = \sum_{j=1}^{S} x_i b_i \tag{A87}$$

$$a_{ij} = (1 - k_{ij}) \left(a_i a_j\right)^{1/2} \tag{A88}$$

x_i is the mole fraction of component i, k_{ij} is a Binary Interaction Parameter (BIP) between components i and j. Pure component parameters a_i and b_i are defined by

$$\begin{aligned} a_i &= 0.45724 \frac{R^2 T_{c,i}^2}{P_{c,i}} \left[1 + \left(0.37464 + 1.54226\omega_i - 0.26992\omega_i^2\right) \left(1 - T_{r,i}^{1/2}\right)\right]^2 \\ b_i &= \frac{0.0778 R T_{c,i}}{P_{c,i}} \end{aligned} \tag{A89}$$

ω_i is the acentric factor of a pure component which is usually available for common components in literature [11, 521], $T_{r,i} = T/T_{c,i}$ is the reduced temperature of component i at temperature T, $T_{c,i}$ and $P_{c,i}$ are critical temperature and critical pressure of component i which are available in literature [297]. The Fugacity coefficient of component i in a mixture is obtained from [312, 330, 382]

$$\begin{aligned} \ln \varphi_i = {} & \frac{b_i}{b_{\mathrm{m}}} \left(Z_{\mathrm{m}} - 1\right) - \ln \left(Z_{\mathrm{m}} - B_{\mathrm{m}}\right) \\ & - \frac{A_{\mathrm{m}}}{2\sqrt{2} B_m} \left(\frac{2 \sum_{k=1}^{S} y_k a_{ik}}{a_{\mathrm{m}}} - \frac{b_i}{b_{\mathrm{m}}} \right) \ln \left(\frac{Z_{\mathrm{m}} + 2.414 B_{\mathrm{m}}}{Z_{\mathrm{m}} - 0.414 B_{\mathrm{m}}} \right) \end{aligned} \tag{A90}$$

with φ_i the fugacity coefficient of component i, Z_m the compressibility factor of the mixture, and A_m and B_m are defined as

$$A_\mathrm{m} = a_\mathrm{m}\,(T)\,P/R^2T^2\ ,\ B_\mathrm{m} = b_\mathrm{m}P/RT\ ,\ Z_\mathrm{m} = PV/RT \tag{A91}$$

To obtain an expression for the partial molar volume of a certain component, we start from the minus 1 rule [5]

$$\left(\frac{\partial V_\mathrm{m}}{\partial N_k}\right)_{T,P,N_{i\neq k}} \left(\frac{\partial N_k}{\partial P}\right)_{T,V_\mathrm{m},N_{i\neq k}} \left(\frac{\partial P}{\partial V_\mathrm{m}}\right)_{T,N_i} = -1 \tag{A92}$$

V_m is the volume of the mixture, N_k is the number of moles of component k, and the notation $N_{i\neq k}$ means that number of molecules of all components except component k are kept constant during differentiation. The term $\left(\frac{\partial V_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}}$ is the partial molar volume of component k which equals

$$\bar{v}_k = -\frac{\left(\frac{\partial P}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}}}{\left(\frac{\partial P}{\partial V_\mathrm{m}}\right)_{T,N_i}} \tag{A93}$$

An analytic expression for the partial molar volume is obtained by applying the derivatives of Eq. A93 to the PR-EoS (Eq. A85). We can write Eq. A85 as

$$\begin{aligned} P &= \frac{RT}{\frac{V_\mathrm{m}}{N} - b_\mathrm{m}} - \frac{a_\mathrm{m}}{\frac{V_\mathrm{m}}{N^2}(V_\mathrm{m} + Nb_\mathrm{m}) + \frac{b_\mathrm{m}}{N}(V_\mathrm{m} - Nb_\mathrm{m})} \\ &= \frac{NRT}{V_\mathrm{m} - Nb_\mathrm{m}} - \frac{N^2 a_\mathrm{m}}{V_\mathrm{m}(V_\mathrm{m} + Nb_\mathrm{m}) + Nb_\mathrm{m}(V_\mathrm{m} - Nb_\mathrm{m})} \end{aligned} \tag{A94}$$

N is the number of moles of the mixture. Starting from the denominator Eq. A93 and taking the derivative of P (Eq. A94) with respect to V_m leads to

$$\begin{aligned} &\left(\frac{\partial P}{\partial V_\mathrm{m}}\right)_{T,N_i} = \left(\frac{\partial}{\partial V_\mathrm{m}}\right)_{T,N_i} \left(\frac{NRT}{V_\mathrm{m} - Nb_\mathrm{m}} - \frac{N^2 a_\mathrm{m}}{V_\mathrm{m}(V_\mathrm{m} + Nb_\mathrm{m}) + Nb_\mathrm{m}(V_\mathrm{m} - Nb_\mathrm{m})}\right) \\ &= -\frac{NRT}{(V_\mathrm{m} - Nb_\mathrm{m})^2} + \frac{2N^2 a_\mathrm{m}(V_\mathrm{m} + Nb_\mathrm{m})}{[V_\mathrm{m}(V_\mathrm{m} + Nb_\mathrm{m}) + Nb_\mathrm{m}(V_\mathrm{m} - Nb_\mathrm{m})]^2} \end{aligned} \tag{A95}$$

The derivative in the nominator of Eq. A93 equals

$$\left(\frac{\partial P}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} = \frac{RT(V_\mathrm{m}-Nb_\mathrm{m})+NRT\left(\frac{\partial Nb_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}}}{(V_\mathrm{m}-Nb_\mathrm{m})^2} - \frac{\left(\frac{\partial N^2 a_\mathrm{m}}{\partial N_k}\right)_{T,V_m,N_{i,j\neq k}} - \frac{N^2 a_\mathrm{m}\left(\frac{\partial}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}}[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]}}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]^2} \tag{A96}$$

Taking the derivative on the right hand side and rearranging leads to

$$\begin{aligned}\left(\frac{\partial P}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} &= \frac{RT(V_\mathrm{m}-Nb_\mathrm{m})+NRT\left(\frac{\partial Nb_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}}}{(V_\mathrm{m}-Nb_\mathrm{m})^2} \\ &- \frac{Na_\mathrm{m}\left(\frac{\partial N}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} + N\left(\frac{\partial Na_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i,j\neq k}}}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]} \\ &- \frac{-\frac{N^2 a_\mathrm{m}\left[2v_\mathrm{m}\left(\frac{\partial Nb_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} - 2Nb_\mathrm{m}\left(\frac{\partial Nb_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}}\right]}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]}}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]}\end{aligned} \tag{A97}$$

Here, two analytic derivatives of a_m and b_m are required which are listed below

$$\begin{aligned}\left(\frac{\partial N b_\mathrm{m}}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} &= \left(\frac{\partial}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} \left(N \sum_{i=1}^{S} x_i b_i\right) \\ &= \left(\frac{\partial}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} \left(N \sum_{i=1}^{S} \frac{N_i}{N} b_i\right) \\ &= \sum_{i=1}^{S} \left(\frac{\partial N_i b_i}{\partial N_k}\right)_{T,V_\mathrm{m},N_{i\neq k}} \\ &= b_k\end{aligned} \tag{A98}$$

$$
\begin{aligned}
\left(\frac{\partial N a_{\mathrm{m}}}{\partial N_k}\right)_{T,V_m,N_{i,j\neq k}} &= \left(\frac{\partial}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i,j\neq k}} \left(N\sum_{i=1}^{S}\sum_{j=1}^{S} x_i x_j a_{ij}\right) \\
&= \left(\frac{\partial}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i,j\neq k}} \left(N\sum_{i=1}^{S}\sum_{j=1}^{S} \frac{N_i}{N^2}\cdot N_j a_{ij}\right) \\
&= \left(\frac{\partial}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i,j\neq k}} \left(\sum_{i=1}^{S}\sum_{j=1}^{S} \frac{N_i}{N} N_j a_{ij}\right) \\
&= \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij} N_j \left(\frac{\partial \frac{N_i}{N}}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i\neq k}} + \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij}\frac{N_i}{N}\left(\frac{\partial N_j}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i\neq k}} \\
&= \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij} N_j \frac{1}{N}\left(\frac{\partial N_i}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i\neq k}} + \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij} N_i N_j \left(\frac{\partial \frac{1}{N}}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i\neq k}} \\
&\quad + \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij}\frac{N_i}{N}\left(\frac{\partial N_j}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i\neq k}} \\
&= \sum_{j=1}^{S} a_{kj}\frac{N_j}{N} - \sum_{i=1}^{S}\sum_{j=1}^{S} a_{ij}\frac{N_i}{N^2} N_j + \sum_{i=1}^{S} a_{ik}\frac{N_i}{N} \\
&= 2\sum_{i=1}^{S} x_i a_{ki} - a_{\mathrm{m}}
\end{aligned}
\tag{A99}
$$

Replacing Eqs. A98 and A99 in Eq. A97 leads to

$$
\begin{aligned}
\left(\frac{\partial P}{\partial N_k}\right)_{T,V_{\mathrm{m}},N_{i\neq k}} &= \frac{RT(V_{\mathrm{m}} - N b_{\mathrm{m}}) + NRTb_k}{(V_{\mathrm{m}} - N b_{\mathrm{m}})^2} \\
&- \frac{2N\sum_{i=1}^{S} x_i a_{i\neq k} - \frac{2a_{\mathrm{m}} b_k N^2 (V_{\mathrm{m}} - N b_{\mathrm{m}})}{[V_{\mathrm{m}}(V_{\mathrm{m}} + N b_{\mathrm{m}}) + N b_{\mathrm{m}}(V_{\mathrm{m}} - N b_{\mathrm{m}})]}}{[V_{\mathrm{m}}(V_{\mathrm{m}} + N b_{\mathrm{m}}) + N b_{\mathrm{m}}(V_{\mathrm{m}} - N b_{\mathrm{m}})]}
\end{aligned}
\tag{A100}
$$

Replacing Eqs. A95 and A100 in Eq. A93, the partial molar volume of component k becomes

$$\bar{v}_k = \frac{\frac{RT(V_\mathrm{m}-Nb_\mathrm{m})+NRTb_k}{(V_\mathrm{m}-Nb_\mathrm{m})^2} - \frac{2N\sum_{i=1}^{S} x_i a_{ki} - \frac{2a_\mathrm{m}b_k N^2(V_\mathrm{m}-Nb_\mathrm{m})}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]}}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]}}{\left[\frac{NRT}{(V_\mathrm{m}-Nb_\mathrm{m})^2} - \frac{2N^2a_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})}{[V_\mathrm{m}(V_\mathrm{m}+Nb_\mathrm{m})+Nb_\mathrm{m}(V_\mathrm{m}-Nb_\mathrm{m})]^2}\right]} \tag{A101}$$

Factorizing N leads to

$$\begin{aligned}\bar{v}_k = & \frac{\frac{NRT(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})+NRTb_k}{N^2(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})^2} - \frac{2N\sum_{i=1}^{S} x_i a_{ki} - \frac{2a_\mathrm{m}b_k N^3(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})}{[N^2\frac{V_\mathrm{m}}{N}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})+N^2b_\mathrm{m}(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})]}}{[N^2\frac{V_\mathrm{m}}{N}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})+N^2b_\mathrm{m}(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})]}}{\left[\frac{NRT}{N^2(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})^2} - \frac{2a_\mathrm{m}N^3(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})}{[N^2\frac{V_\mathrm{m}}{N}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})+N^2b_\mathrm{m}(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})]^2}\right]} \\ = & \frac{\frac{1}{N}\left[\frac{RT(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})+RTb_k}{(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})^2}\right] - \frac{1}{N}\left[\frac{2\sum_{i=1}^{S} x_i a_{ki} - \frac{2a_\mathrm{m}b_k(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})}{[\frac{V_\mathrm{m}}{N}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})+b_\mathrm{m}(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})]}}{[\frac{V_\mathrm{m}}{N}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})+b_\mathrm{m}(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})]}\right]}{\frac{1}{N}\left[\frac{RT}{(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})^2} - \frac{2a_\mathrm{m}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})}{[\frac{V_\mathrm{m}}{N}(\frac{V_\mathrm{m}}{N}+b_\mathrm{m})+b_\mathrm{m}(\frac{V_\mathrm{m}}{N}-b_\mathrm{m})]^2}\right]}\end{aligned} \tag{A102}$$

$v_m = \frac{V_\mathrm{m}}{N}$ is the molar volume of the mixture which is solved directly from Eq. A85. Therefore, partial molar volume of component k in the mixture

equals

$$\bar{\upsilon}_k = \frac{\dfrac{RT(\upsilon_\mathrm{m}-b_\mathrm{m})+RTb_k}{(\upsilon_\mathrm{m}-b_\mathrm{m})^2} - \dfrac{2\sum_{i=1}^{S} x_i a_{ki} - \frac{2a_\mathrm{m}b_k(\upsilon_\mathrm{m}-b_\mathrm{m})}{[\upsilon_\mathrm{m}(\upsilon_\mathrm{m}+b_\mathrm{m})+b_\mathrm{m}(\upsilon_\mathrm{m}-b_\mathrm{m})]}}{[\upsilon_\mathrm{m}(\upsilon_\mathrm{m}+b_\mathrm{m})+b_\mathrm{m}(\upsilon_\mathrm{m}-b_\mathrm{m})]}}{\left[\dfrac{RT}{(\upsilon_\mathrm{m}-b_\mathrm{m})^2} - \dfrac{2a_\mathrm{m}(\upsilon_\mathrm{m}+b_\mathrm{m})}{[\upsilon_\mathrm{m}(\upsilon_\mathrm{m}+b_\mathrm{m})+b_\mathrm{m}(\upsilon_\mathrm{m}-b_\mathrm{m})]^2}\right]} \tag{A103}$$

The partial molar enthalpy of component k is related to its partial molar volume [3]

$$\left(\frac{\partial \bar{h}_k}{\partial P}\right)_{T,N_i} = \bar{\upsilon}_k - T\left(\frac{\partial \bar{\upsilon}_k}{\partial T}\right)_{P,N_i} \tag{A104}$$

The right hand side of Eq. A104 is only related to excess partial molar volume as the ideal gas term drops out:

$$\begin{aligned}
&\bar{\upsilon}_k - T\left(\frac{\partial \bar{\upsilon}_k}{\partial T}\right)_{P,N_i} = \\
&\left[\bar{\upsilon}_k^\mathrm{id} - T\left(\frac{\partial \bar{\upsilon}_k^\mathrm{id}}{\partial T}\right)_{P,N_i}\right] + \bar{\upsilon}_k^\mathrm{ex} - T\left(\frac{\partial \bar{\upsilon}_k^\mathrm{ex}}{\partial T}\right)_{P,N_i} = \\
&\left[\frac{RT}{P} - T\left(\frac{R}{P}\right)\right] + \bar{\upsilon}_k^\mathrm{ex} - T\left(\frac{\partial \bar{\upsilon}_k^\mathrm{ex}}{\partial T}\right)_{P,N_i} = \\
&\bar{\upsilon}_k^\mathrm{ex} - T\left(\frac{\partial \bar{\upsilon}_k^\mathrm{ex}}{\partial T}\right)_{P,N_i}
\end{aligned} \tag{A105}$$

Given that the ideal part of the enthalpy is only a function of temperature, we can write

$$\left(\frac{\partial \bar{h}_k^\mathrm{ex}}{\partial P}\right)_{P,N_i} = \bar{\upsilon}_k - T\left(\frac{\partial \bar{\upsilon}_k}{\partial T}\right)_{P,N_i} \tag{A106}$$

The partial molar excess enthalpy can be computed numerically from

$$\bar{h}_k^\mathrm{ex} = \int_0^P \mathrm{d}P \left[\bar{\upsilon}_k - T\left(\frac{\partial \bar{\upsilon}_k}{\partial T}\right)_{P,N_i}\right] \tag{A107}$$

Table A7: Critical temperatures (T_c), pressures (P_c), acentric factors (ω) used for PR-EoS modelling, and enthalpies of formation ($\bar{h}_f^\circ$) of the components at the standard reference state ($P^\circ = 1$ bar, $T^\circ = 298$ K) [334, 514, 521].

Component	T_c /[K]	P_c /[Pa]	ω	$\bar{h}_f^\circ$/[kJ.mol^{-1}]
N_2	126.19	3395800	0.0372	0
H_2	33.14	1296400	-0.219	0
NH_3	405.4	11333000	0.25601	-45.94 ± 0.35
H_2O	647.1	22064000	0.3443	-241.826 ± 0.040
CO	132	3400000	0.066	-110.53 ± 0.17
CO_2	304	7300000	0.228	-393.51 ± 0.13
HCOOH	577	7500000	0.445	-425.5 ± 0.3
CH_4	190	4500000	0.012	-74.6 ± 0.3

Eqs. A103 and A107 were numerically verified with the expressions derived by Michelsen and Mollerup [8]. Exact agreement was found between the results from both methods.

PC-SAFT Equation of State

The Perturbed Chain - Statistical Association Fluid Theory (PC-SAFT) EoS [18, 19, 331, 332, 347] is a theoretically derived model, based on statistical mechanics principles. The basis of the model relies on applying rigorous perturbation theory [331] for systems which are comprised of a repulsive core and multiple attractive sites, resulting in an expression for the Helmholtz energy. In this way, the Helmholtz energy of a molecular fluid can be obtained as the sum of the Helmholtz energies of a simple reference fluid (which is known accurately) and various perturbation terms. For details and the exact mathematical relations, the reader is referred to Refs. [18, 19, 331, 332, 347]. The calculation of mixture properties requires appropriate mixing and combining rules. In this work, the van der Waals mixing rules, as proposed by Gross and Sadowski [18], were used for the dispersion term, while the Lorentz-Berthelot mixing rules were applied to

calculate the segment energy and diameter parameters [15]. The PC-SAFT parameters for the mixtures are therefore

$$\overline{m^2 \frac{\varepsilon}{k_\mathrm{B}T}\sigma^3} = \sum_{i=1}^{\mathrm{S}}\sum_{j=1}^{\mathrm{S}} x_i x_j m_i m_j \left[\frac{\varepsilon_{ij}}{k_\mathrm{B}T}\right] \sigma_{ij}^3 \tag{A108}$$

$$\overline{m^2 \left[\frac{\varepsilon}{k_\mathrm{B}T}\right]^2 \sigma^3} = \sum_{i=1}^{\mathrm{S}}\sum_{j=1}^{\mathrm{S}} x_i x_j m_i m_j \left[\frac{\varepsilon_{ij}}{k_\mathrm{B}T}\right]^2 \sigma_{ij}^3 \tag{A109}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}(1 - k_{ij}) \tag{A110}$$

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \tag{A111}$$

In these equations, m_i is the number of spherical segments in component i, ϵ_i is the dispersion energy between spherical segments of component i, σ_i is the temperature-independent diameter of each spherical segment in component i, and k_{ij} is the BIP between components i and j. In this work, the binary interactions parameters k_{ij} are set to zero. For associating mixtures, two more combining rules have to be applied for the cross-association energy and volume [332]:

$$\varepsilon^{A_i B_j} = \frac{1}{2}(\varepsilon^{A_i B_i} + \varepsilon^{A_j B_j}) \tag{A112}$$

$$\kappa^{A_i B_j} = \sqrt{\kappa^{A_i B_i} \kappa^{A_j B_j}} \left(\frac{\sqrt{\sigma_i \sigma_j}}{\frac{1}{2}(\sigma_i + \sigma_j)}\right)^3 \tag{A113}$$

where $\epsilon^{A_i B_i}$ is the association energy and $\kappa^{A_i B_i}$ is the association volume of component i. In this work, ammonia was treated as an associating molecule with 4 association sites, as proposed by Mejbri and Bellagi [333, 343]. The PC-SAFT EoS parameters that were used in our calculations are summarized in Table A8.

Table A8: PC-SAFT EoS parameters for the components studied in this work.

Component	m	$\sigma/[\text{Å}]$	$\epsilon/k_\text{B}/[\text{K}]$	$\epsilon^{AB}/k_\text{B}/[\text{K}]$	κ^{AB}	Ref.
H_2	0.8285	2.973	12.53	-	-	[343]
N_2	1.2053	3.313	90.96	-	-	[19]
NH_3	2.5785	2.2677	75.092	1041.5	0.37213	[333]

A.6 Evaluating Enthalpies

The reaction enthalpy is defined as the difference between the enthalpy of the reaction products and enthalpy of the reactants, with respect to their stoichiometric coefficients [5]:

$$\Delta\bar{h} = \sum_{\text{P}} \nu_i \bar{h}_i(T,P) - \sum_{\text{A}} \nu_j \bar{h}_j(T,P) \tag{A114}$$

The partial molar enthalpy of a component at temperature T and pressure P equals

$$\begin{aligned} \bar{h}_i(T,P) &= \bar{h}^\circ_{f,i} + \left[\bar{h}_i(T,P_\text{ref}) - \bar{h}_i(T_\text{ref},P_\text{ref})\right] + \left[\bar{h}^\text{ex}_i(T,P) - \bar{h}^\text{ex}_i(T,P_\text{ref})\right] \\ &= \bar{h}^\circ_{f,i} + \left[\bar{h}_i(T,P_\text{ref}) - \bar{h}_i(T_\text{ref},P_\text{ref})\right] + \bar{h}^\text{ex}_i(T,P) \end{aligned} \tag{A115}$$

T_ref and P_ref are the reference temperature and pressure at 298 K and 1 bar, respectively. $\bar{h}^\circ_{f,i}$ is the formation enthalpy of component i at (T_ref,P_ref) and its value can be found in literature [3, 334, 514]. The second term on the right hand side of Eq. A118 is associated with enthalpy difference at (T, P_ref) relative to the reference state at (T_ref, P_ref) and at constant composition. This is often expressed as follows (the Shomate equation [514]):

$$\bar{h}_i(T,P_\text{ref}) - \bar{h}_i(T_\text{ref},P_\text{ref}) = AT + \frac{BT^2}{2} + \frac{CT^3}{3} + \frac{DT^4}{4} - \frac{E}{T} + F - H \tag{A116}$$

The temperature T in Eq. A119 is in units of $\frac{\mathrm{K}}{1000}$. The coefficients A to H for NH_3, N_2 and H_2 are taken from NIST-JANAF Thermochemical Tables [334, 514]. The coefficients A to H and the left hand side of Eq. A119 evaluated at $T = 573$ K are listed in Table A9 for all components. The third term on the right hand side of Eq. A118 is associated with the enthalpy difference between states (T, P) and (T, P_{ref}) which accounts for deviation from ideal gas behavior relative to the standard reference pressure [3]. The term $\bar{h}_i^{\mathrm{ex}}(T, P_{\mathrm{ref}})$ in Eq. A118 can be considered zero at high temperatures. As explained in the previous sectons, $\bar{h}_i^{\mathrm{ex}}(T, P_{\mathrm{ref}})$ can be obtained either from the PR-EoS (Eq. A107), the PC-SAFT EoS, or MC simulations (Eq. A55).

A.7 Evaluating the Reaction Enthalpy at High Pressure

The reaction enthalpy is defined as the difference between the enthalpy of the reaction products and enthalpy of the reactants, with respect to their stoichiometric coefficients [5]:

$$\Delta\bar{h} = \sum_{\mathrm{P}} \nu_i \bar{h}_i(T,P) - \sum_{\mathrm{A}} \nu_j \bar{h}_j(T,P) \tag{A117}$$

The partial molar enthalpy of a component at temperature T and pressure P equals

$$\begin{aligned} &\bar{h}_i(T,P) = \\ &\bar{h}^{\circ}_{f,i} + \left[\bar{h}_i(T,P_{\mathrm{ref}}) - \bar{h}_i(T_{\mathrm{ref}},P_{\mathrm{ref}})\right] + \left[\bar{h}^{\mathrm{ex}}_i(T,P) - \bar{h}^{\mathrm{ex}}_i(T,P_{\mathrm{ref}})\right] = \\ &\bar{h}^{\circ}_{f,i} + \left[\bar{h}_i(T,P_{\mathrm{ref}}) - \bar{h}_i(T_{\mathrm{ref}},P_{\mathrm{ref}})\right] + \bar{h}^{\mathrm{ex}}_i(T,P) \end{aligned} \tag{A118}$$

T_{ref} and P_{ref} are the reference temperature and pressure at 298 K and 1 bar, respectively. $\bar{h}^{\circ}_{f,i}$ is the formation enthalpy of component i at (T_{ref},P_{ref}) and its value can be found in literature [3, 334, 514]. The second term on the right hand side of Eq. A118 is associated with enthalpy difference at (T, P_{ref}) relative to the reference state at (T_{ref}, P_{ref}) and at constant composition. This is often expressed as follows (the Shomate equation [514]):

$$\bar{h}_i(T,P_{\mathrm{ref}}) - \bar{h}_i(T_{\mathrm{ref}},P_{\mathrm{ref}}) = AT + \frac{BT^2}{2} + \frac{CT^3}{3} + \frac{DT^4}{4} - \frac{E}{T} + F - H \tag{A119}$$

The temperature T in Eq. A119 is in units of $\frac{\mathrm{K}}{1000}$. The coefficients A to H for NH_3, N_2 and H_2 are taken from NIST-JANAF Thermochemical Tables [334, 514]. The coefficients A to H and the left hand side of Eq. A119 evaluated at $T = 573$ K are listed in Table A9 for all components. The third

term on the right hand side of Eq. A118 is associated with the enthalpy difference between states (T, P) and (T, P_{ref}) which accounts for deviation from ideal gas behavior relative to the standard reference pressure [3]. The term $\bar{h}_i^{\text{ex}}(T, P_{\text{ref}})$ in Eq. A118 can be considered zero at high temperatures. As explained in the previous sectons, $\bar{h}_i^{\text{ex}}(T, P_{\text{ref}})$ can be obtained either from the PR-EoS (Eq. A107), the PC-SAFT EoS, or MC simulations (Eq. A55).

Table A9: Calculation of the enthalpy differences for each component in the ammonia synthesis reaction at temperature of 573 K relative to the reference temperature ($T_{\mathrm{ref}} = 298.15\mathrm{K}$) at $P_{\mathrm{ref}} = 0.1\mathrm{MPa}$, using the Shomate equation (Eq. A119 [514]). The parameters A to H were taken from NIST thermochemistry database [334, 514]. The parameters are applicable to the temperature range of 298 K to 1400 K for $\mathrm{NH_3}$, 500 K to 2000 K for $\mathrm{N_2}$ and 298 K to 1000 K for $\mathrm{H_2}$. Enthalpies are in units of $\mathrm{kJ.mol^{-1}}$.

Component	$\mathrm{NH_3}$	$\mathrm{N_2}$	$\mathrm{H_2}$
A	19.99563	19.50583	33.066178
B	49.77119	19.88705	-11.363417
C	-15.37599	-8.598535	11.432816
D	1.921168	1.369784	-2.772874
E	0.189174	0.527601	-0.158558
F	-53.30667	-4.935202	-9.980797
G	203.8591	212.3900	172.707974
H	-45.89806	-	-

$\bar{h}_i(T, P_{\mathrm{ref}}) - \bar{h}_i(T_{\mathrm{ref}}, P_{\mathrm{ref}})$			
$T/[\mathrm{K}]$	$\mathrm{NH_3}$	$\mathrm{N_2}$	$\mathrm{H_2}$
573	10.97	8.08	8.01

A.8 Umbrella Sampling

Considering the partition function in the NPT ensemble expanded with a fractional molecule, as in Eq. A41, the probability of $\lambda' = \lambda$ is written as

$$\begin{aligned}\left\langle \delta\left(\lambda-\lambda'\right)\right\rangle_{\beta} = \frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \exp[-\beta P V] \\ \times \int_0^1 \mathrm{d}\lambda' \int \mathrm{d}s^N \exp[-\beta U_{\mathrm{total}}\left(s^N, s_{\mathrm{frac}}, \lambda', V\right)] \delta\left(\lambda-\lambda'\right)\end{aligned} \tag{A120}$$

where $U_{\mathrm{total}}\left(s^N, s_{\mathrm{frac}}, \lambda', V\right)$ is the total interaction potential between the molecules including the fractional molecule. Multiplying and dividing the integrand on the right hand side of Eq. A120 by a biasing factor proportional to the Boltzmann factor of total enthalpy of the system at $T^\star$, $\exp[-\beta^\star\left(U_{\mathrm{total}} + PV\right)]$, leads to

$$\begin{aligned}\left\langle \delta\left(\lambda-\lambda'\right)\right\rangle_{\beta} = \frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \exp[\Delta\beta P V] \\ \times \int_0^1 \mathrm{d}\lambda' \exp[-\beta^\star P V] \int \mathrm{d}s^N \exp[\Delta\beta U_{\mathrm{total}}\left(s^N, s_{\mathrm{frac}}, \lambda', V\right)] \\ \times \exp[-\beta U_{\mathrm{total}}\left(s^N, s_{\mathrm{frac}}, \lambda', V\right)] \delta\left(\lambda-\lambda'\right)\end{aligned} \tag{A121}$$

where $\Delta\beta = \beta^\star - \beta$. Rearranging Eq. A121 leads to

$$\begin{aligned}\left\langle \delta\left(\lambda-\lambda'\right)\right\rangle_{\beta} = \frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda^3}{\Lambda_i^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \\ \times \int_0^1 \mathrm{d}\lambda' \left(\delta\left(\lambda-\lambda'\right) \exp\left[\Delta\beta H_{\mathrm{total}}\left(s^N, s_{\mathrm{frac}}, \lambda', V\right)\right]\right) \\ \times \exp[-\beta^\star P V] \int \mathrm{d}s^N \exp[-\beta^\star U_{\mathrm{total}}\left(s^N, s_{\mathrm{frac}}, \lambda', V\right)]\end{aligned} \tag{A122}$$

which means that the distribution $p(\lambda)$ in the CFCNPT ensemble can be sampled by performing a simulation in the CFC$NPT^{\star}$ ensemble. Eq. A122 can be written as

$$p\left(\lambda\right)|_{\beta} = c \cdot \left\langle \delta(\lambda^{'} - \lambda) \exp\left[\left(\beta^{*} - \beta\right) H\right] \right\rangle_{\beta^{\star}} \tag{A123}$$

where c is a normalization constant. In a similar manner, one can calculate other ensemble averages, such as the density, in the CFCNPT ensemble by performing a simulation in the CFC$NPT^{\star}$ ensemble. To compute the distribution $p(\lambda)$ at a different pressure, one can simply multiply and divide the right hand side of equation Eq. A120 by $\exp[-\beta P^{\star} V]$ leading to

$$\begin{aligned} \left\langle \delta\left(\lambda - \lambda'\right) \right\rangle_{P} = \frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda^{3}}{\Lambda_{i}^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \exp[\beta V \Delta P] \\ \times \int_{0}^{1} \mathrm{d}\lambda' \exp[-\beta P^{\star} V] \int \mathrm{d}s^{N} \exp[-\beta U_{\mathrm{total}}\left(s^{N}, s_{\mathrm{frac}}, \lambda', V\right)] \\ \times \delta\left(\lambda - \lambda'\right) \end{aligned} \tag{A124}$$

where $\Delta P = P^{\star} - P$. Rearranding Eq. A124 leads to

$$\begin{aligned} \left\langle \delta\left(\lambda - \lambda'\right) \right\rangle_{P} = \frac{\beta P}{Q_{\mathrm{CFC}NPT}} \prod_{i=1}^{S} \frac{1/\Lambda^{3}}{\Lambda_{i}^{3N_i} N_i!} \int \mathrm{d}V V^{N+1} \\ \times \int_{0}^{1} \mathrm{d}\lambda' \left(\delta\left(\lambda - \lambda'\right) \exp\left[\beta V \Delta P\right]\right) \exp[\beta P^{\star} V] \\ \times \int \mathrm{d}s^{N} \exp[-\beta U_{\mathrm{total}}\left(s^{N}, s_{\mathrm{frac}}, \lambda', V\right)] \end{aligned} \tag{A125}$$

which means that the distribution $p(\lambda)$ in the CFCNPT ensemble can be sampled by performing a simulation in the CFC$NP^{\star}T$ ensemble. Eq. A125 can be written as

$$p\left(\lambda\right)|_{P} = c \cdot \left\langle \delta(\lambda^{'} - \lambda) \exp\left[\beta V \left(P^{\star} - P\right)\right] \right\rangle_{P^{\star}} \tag{A126}$$

In a similar, manner, one can calculate other ensemble averages, such as the density, in the CFCNPT ensemble by running a simulation in the CFC$NP^{\star}T$ ensemble.

A.9 Fugacity Coefficients from CFCMC Simulations

Consider a multicomponent system that is simulated in an ensemble that is either open (e.g. grand-canonical ensemble) or closed (NPT ensemble). In this system, we would like to calculate the fugacity coefficient ϕ_i of component i. We assume that a fractional molecule of component i is present. The chemical potential of component i equals (Eq. 2.9)

$$\mu_i = \mu_i^0 + RT \ln \frac{\langle \rho_i \rangle}{\rho_0} + \mu_i^{\text{ex}} = \mu_i^0 + RT \ln \frac{\langle \rho_i \rangle}{\rho_0} - RT \ln \frac{p(\lambda_i = 1)}{p(\lambda_i = 0)} \tag{A127}$$

in which μ_i^0 is the reference state of the chemical potential which depends on the temperature but not on the pressure, $\langle \rho_i \rangle$ is the average number density of i, μ_{ex}^i is the excess chemical potential of i, ρ_0 is an arbitrary reference density (to make the argument of the logarithm dimensionless), and $p(\lambda_i)$ is the probability distribution of the coupling parameter of the fractional molecule of i. We are considering a large system so $\rho_i = N_i/V \approx (N_i + 1)/V$ [164]. In classical thermodynamics, the chemical potential of i is usually expressed as [4, 312]

$$\mu_i = \mu_i^\star + RT \ln \left(\frac{y_i P \phi_i}{P_0} \right) \tag{A128}$$

in which $\mu_i^\star$ is a reference chemical potential (which is different from μ_i^0), y_i is the mole fraction of i, P is the pressure, and P_0 is a reference pressure (usually 1 bar). The reference chemical potentials μ_i^0 and $\mu_i^\star$ only depend on the temperature and not on the pressure or composition of the system. To find an expression for the fugacity coefficient ϕ_i, consider a system in which the pressure P is approaching zero while the composition of the mixture is constant. In this limit, $\phi_i = 1$ and $\mu_{\text{ex}}^i = 0$. We have

$$\mu_i^0 + RT \ln \frac{\langle \rho_i \rangle}{\rho_0} = \mu_i^\star + RT \ln \left(\frac{y_i P}{P_0} \right) \tag{A129}$$

In this limit, the ideal gas law can also be used to calculate the average number density of i,

$$\langle \rho_i \rangle = \frac{y_i P}{RT} \tag{A130}$$

This leads to

$$\mu_i^0 - \mu_i^\star = RT \ln \left(\frac{\rho_0 RT}{P_0} \right) \tag{A131}$$

This equation can be used to eliminate the reference state in Eq. A128 leading to

$$\mu_i^{\text{ex}} = RT \ln \left(\frac{y_i P \phi_i}{RT \langle \rho_i \rangle} \right) \tag{A132}$$

so

$$\phi_i = \frac{RT \langle \rho_i \rangle}{y_i P} \exp \left[\mu_{\text{ex}}^i / (RT) \right] \tag{A133}$$

If the system consists of N_t molecules in total (including component i, and not counting fractional molecules), we have $\langle \rho_i \rangle \approx N_i / \langle V \rangle$ ($\langle V \rangle$ being the average volume) and $y_i = N_i / N_t$. We finally have

$$\phi_i = \frac{N_t RT}{P \langle V \rangle} \exp \left[\mu_i^{\text{ex}} / (RT) \right] = \frac{\exp \left[\mu_i^{\text{ex}} / (RT) \right]}{Z_m} \tag{A134}$$

in which $Z_m = \frac{P \langle V \rangle}{N_t RT}$ is the compressibility of the mixture. The fugacity coefficient ϕ_i thus depends on both the excess chemical potential of i and the overall deviation from ideal gas behavior of the mixture.

A.10 Activity Coefficients from Molecular Simulation

Eq. A127 described the chemical potential of component i in a mixture, with respect to the ideal gas. For a liquid, the chemical potential of Eq. A127 in can also be written in terms of the activity coefficient of component i in the mixture:

$$\mu_i = \mu_i^\star + k_\mathrm{B}T \ln (\gamma_i x_i) \tag{A135}$$

where $\mu_i^\star$ is the reference chemical potential of component i which depends only on temperature and pressure. x_i is the mole fraction of component i, and γ_i is the activity coefficient of component i which depends on temperature, pressure and composition. For a pure component ($\gamma_i x_i = 1$), the reference chemical potential $\mu_i^\star$ is obtained by combining Eqs. A127 and A135

$$\mu_i^\star = \mu_i^\circ + k_\mathrm{B}T \ln \frac{\langle \rho_i \rangle}{\rho_0} + \mu_{0i}^\mathrm{ex} \tag{A136}$$

in which $\langle \rho_{0i} \rangle$ is the ensemble average number density of pure i and μ_{0i}^ex is the excess chemical potential of pure i with respect to the ideal gas. By combining Eqs. A135 and A136, the chemical potential of component i in a mixture can be written as

$$\mu_i = \mu_i^\circ + k_\mathrm{B}T \ln \frac{\langle \rho_i \rangle}{\rho_0} + \mu_{0i}^\mathrm{ex} + k_\mathrm{B}T \ln (\gamma_i x_i) \tag{A137}$$

The activity coefficient can be obtained by combining Eqs. A127 and A137. This leads to

$$k_\mathrm{B}T \ln \frac{\langle \rho_i \rangle}{\langle \rho_{0i} \rangle} + \mu_i^\mathrm{ex} - \mu_{0i}^\mathrm{ex} = k_\mathrm{B}T \ln (\gamma_i x_i) \tag{A138}$$

By rearranging this equation, we obtain

$$\gamma_i = \frac{\langle \rho_i \rangle}{x_i \langle \rho_{0i} \rangle} \exp \left[\frac{\mu_i^\mathrm{ex} - \mu_{0i}^\mathrm{ex}}{k_\mathrm{B}T} \right] \tag{A139}$$

This is the same result as obtained by Sadowski and co-workers [294].

A.11 Experimental Solubilities

A.11.1 $H_2O - H_2$ mixtures - Liquid phase

Table A10: Experimental solubilities of hydrogen in $H_2O - H_2$ mixtures (liquid phase) at coexistence. Experimental data are converted to mole fractions for different temperatures and pressures. For the original units for each data set, see the indicated references below. For conversion to mole fractions, standard conditions at $T = 273.15$ K and $P = 1.01325$ atm are considered, unless otherwise mentioned in the reference.

T/[K]	P/[bar]	x_{H_2}	Ref
273.15	25	4.31×10^{-4}	[380]
273.15	51	8.63×10^{-4}	[380]
273.15	101	1.71×10^{-3}	[380]
273.15	203	3.35×10^{-3}	[380]
273.15	405	6.40×10^{-3}	[380]
273.15	608	9.25×10^{-3}	[380]
273.15	811	1.19×10^{-2}	[380]
273.15	1013	1.42×10^{-2}	[380]
298.15	25	3.50×10^{-4}	[380]
298.15	51	6.94×10^{-4}	[380]
298.15	101	1.39×10^{-3}	[380]
298.15	203	2.71×10^{-3}	[380]
298.15	405	5.24×10^{-3}	[380]
298.15	608	7.64×10^{-3}	[380]
298.15	811	9.90×10^{-3}	[380]
298.15	1013	1.20×10^{-2}	[380]
310.15	1.013	1.33×10^{-5}	[384]
310.93	3.4	4.50×10^{-5}	[385]
310.93	13.8	1.81×10^{-4}	[385]
310.93	31.0	4.10×10^{-4}	[385]
310.93	65.5	8.62×10^{-4}	[385]
310.93	103.4	1.33×10^{-3}	[385]
310.93	137.9	1.76×10^{-3}	[385]
323.15	25	3.26×10^{-4}	[380]
323.15	31.8	4.02×10^{-4}	[386]
323.15	51	6.49×10^{-4}	[380]
323.15	60.3	7.63×10^{-4}	[386]

Continued on next page

T/[K]	P/[bar]	x_{H_2}	Ref
323.15	101	1.29×10^{-3}	[380]
323.15	119.3	1.51×10^{-3}	[386]
323.15	203	2.54×10^{-3}	[380]
323.15	405	4.92×10^{-3}	[380]
323.15	608	7.18×10^{-3}	[380]
323.15	811	9.34×10^{-3}	[380]
323.15	1013	1.14×10^{-2}	[380]
348.15	25	3.32×10^{-4}	[380]
348.15	51	6.63×10^{-4}	[380]
348.15	101	1.32×10^{-3}	[380]
348.15	203	2.60×10^{-3}	[380]
348.15	405	5.04×10^{-3}	[380]
348.15	608	7.35×10^{-3}	[380]
348.15	811	9.54×10^{-3}	[380]
348.15	1013	1.17×10^{-2}	[380]
366.48	3.4	3.74×10^{-5}	[385]
366.48	13.8	2.0×10^{-4}	[387]
366.48	13.8	1.80×10^{-4}	[385]
366.48	27.6	3.7×10^{-4}	[387]
366.48	31.0	4.26×10^{-4}	[385]
366.48	55.2	7.5×10^{-4}	[387]
366.48	65.5	8.93×10^{-4}	[385]
366.48	110.3	1.50×10^{-3}	[387]
366.48	137.9	1.840×10^{-3}	[385]
373.15	21	3.01×10^{-4}	[388]
373.15	25	3.70×10^{-4}	[380]
373.15	31	4.48×10^{-4}	[388]
373.15	42	5.94×10^{-4}	[388]
373.15	42	5.17×10^{-4}	[389]
373.15	51	7.30×10^{-4}	[380]
373.15	57.0	8.20×10^{-4}	[386]
373.15	62	8.77×10^{-4}	[388]
373.15	62	7.61×10^{-4}	[389]
373.15	82	1.15×10^{-3}	[388]
373.15	82	1.02×10^{-3}	[389]
373.15	101	1.45×10^{-3}	[380]
373.15	102	1.32×10^{-3}	[389]

Continued on next page

$T/[\mathrm{K}]$	$P/[\mathrm{bar}]$	$x_{\mathrm{H_2}}$	Ref
373.15	120.9	1.75×10^{-3}	[386]
373.15	153.7	2.23×10^{-3}	[386]
373.15	203	2.85×10^{-3}	[380]
373.15	405	5.46×10^{-3}	[380]
373.15	608	7.96×10^{-3}	[380]
373.15	811	1.03×10^{-2}	[380]
373.15	1013	1.25×10^{-2}	[380]
398.15	23	3.44×10^{-4}	[388]
398.15	33	5.04×10^{-4}	[388]
398.15	43	6.54×10^{-4}	[388]
398.15	52	7.18×10^{-4}	[389]
398.15	63	9.32×10^{-4}	[388]
398.15	82	1.13×10^{-3}	[389]
398.15	83	1.21×10^{-3}	[388]
398.15	87	1.21×10^{-3}	[389]
398.15	102	1.43×10^{-3}	[389]
422.04	31.0	5.00×10^{-4}	[385]
422.04	65.5	1.16×10^{-3}	[385]
422.04	103.4	1.88×10^{-3}	[385]
423.15	21	3.26×10^{-4}	[388]
423.15	30	4.95×10^{-4}	[388]
423.15	31	4.370×10^{-4}	[389]
423.15	40	6.55×10^{-4}	[388]
423.15	42	5.98×10^{-4}	[389]
423.15	52	7.79×10^{-4}	[389]
423.15	51.8	9.00×10^{-4}	[386]
423.15	54.3	9.79×10^{-4}	[386]
423.15	55	8.75×10^{-4}	[388]
423.15	62	9.43×10^{-4}	[389]
423.15	76	1.15×10^{-3}	[388]
423.15	75.9	1.35×10^{-3}	[386]
423.15	82	1.34×10^{-3}	[389]
423.15	87.1	1.60×10^{-3}	[386]
448.15	22	3.28×10^{-4}	[388]
448.15	29	4.83×10^{-4}	[388]
448.15	39	6.59×10^{-4}	[388]
448.15	52	7.46×10^{-4}	[389]

Continued on next page

T/[K]	P/[bar]	x_{H_2}	Ref
448.15	60	9.56×10^{-4}	[388]
448.15	62	9.43×10^{-4}	[389]
448.15	80	1.24×10^{-3}	[388]
473.15	26	3.30×10^{-4}	[388]
473.15	31	1.00×10^{-3}	[388]
473.15	36	5.74×10^{-4}	[388]
473.15	42	5.52×10^{-4}	[389]
473.15	46	8.44×10^{-4}	[388]
473.15	52	7.51×10^{-4}	[389]
473.15	76	1.22×10^{-3}	[388]
473.15	82	1.24×10^{-3}	[389]
473.15	102	1.40×10^{-3}	[389]
473.15	118	1.62×10^{-3}	[389]
477.59	27.6	3.5×10^{-4}	[387]
477.59	31.0	4.29×10^{-4}	[385]
477.59	55.2	1.05×10^{-3}	[387]
477.59	65.5	1.46×10^{-3}	[385]
477.59	103.4	2.57×10^{-3}	[385]
477.59	110.3	2.71×10^{-3}	[387]
498.15	33	3.22×10^{-4}	[388]
498.15	36	4.47×10^{-4}	[388]
498.15	41	5.91×10^{-4}	[388]
498.15	51	8.10×10^{-4}	[388]
498.15	62	9.93×10^{-4}	[389]
498.15	71	1.21×10^{-3}	[388]
498.15	92	1.45×10^{-3}	[389]
523.15	44	2.70×10^{-4}	[388]
523.15	47	4.45×10^{-4}	[388]
523.15	52	6.50×10^{-4}	[388]
523.15	60	8.93×10^{-4}	[388]
523.15	70	1.18×10^{-3}	[388]
548.15	63	3.01×10^{-4}	[388]
548.15	65	4.70×10^{-4}	[388]
548.15	68	6.54×10^{-4}	[388]
548.15	72	8.48×10^{-4}	[388]
548.15	80	1.19×10^{-3}	[388]
573.15	89	2.86×10^{-4}	[388]

Continued on next page

$T/[\mathrm{K}]$	$P/[\mathrm{bar}]$	$x_{\mathrm{H_2}}$	Ref
573.15	91	5.42×10^{-4}	[388]
573.15	93	7.24×10^{-4}	[388]
573.15	96	8.80×10^{-4}	[388]
573.15	100	1.25×10^{-3}	[388]
574.81	110.3	1.41×10^{-3}	[387]
588.70	110.3	2.26×10^{-3}	[387]
588.71	137.9	2.98×10^{-3}	[385]

A.11.2 $H_2O - H_2$ mixtures - Gas phase

Table A11: Experimental solubilities of water in $H_2O - H_2$ mixtures (gas phase) at coexistence. Experimental data are converted to mole fractions for different temperatures and pressures. For the original units for each data set, see references below. For conversion to mole fractions, standard conditions at $T = 273.15$ K and $P = 1.01325$ atm are considered, unless otherwise mentioned in the reference.

$T/$[K]	$P/$[bar]	x_{H_2}	Ref
310.93	3.4	1.96×10^{-2}	[385]
310.93	13.8	4.88×10^{-3}	[385]
310.93	31.0	2.22×10^{-3}	[385]
310.93	65.5	1.16×10^{-3}	[385]
310.93	103.4	7.6×10^{-4}	[385]
310.93	137.9	6.0×10^{-4}	[385]
310.95	13.8	4.88×10^{-3}	[390]
310.95	31.0	2.22×10^{-3}	[390]
310.95	65.5	1.16×10^{-3}	[390]
310.95	103.4	7.60×10^{-4}	[390]
323.15	50	2.66×10^{-3}	[391]
323.15	100	1.48×10^{-3}	[391]
323.15	101.3	1.38×10^{-3}	[61]
323.15	150	1.08×10^{-3}	[391]
323.15	200	8.8×10^{-4}	[391]
323.15	202.7	7.84×10^{-4}	[61]
323.15	250	7.6×10^{-4}	[391]
323.15	300	6.8×10^{-4}	[391]
323.15	405.3	4.64×10^{-4}	[61]
323.15	608.0	3.61×10^{-4}	[61]
323.15	1013.3	2.80×10^{-4}	[61]
343.15	50	6.60×10^{-3}	[391]
343.15	100	3.58×10^{-3}	[391]
343.15	150	2.57×10^{-3}	[391]
343.15	200	2.07×10^{-3}	[391]
343.15	250	1.77×10^{-3}	[391]
343.15	300	1.57×10^{-3}	[391]
366.45	13.8	3.86×10^{-3}	[390]
366.45	31.0	2.64×10^{-3}	[390]

Continued on next page

$T/[\mathrm{K}]$	$P/[\mathrm{bar}]$	$y_{\mathrm{H_2O}}$	Ref
366.45	65.5	1.32×10^{-3}	[390]
366.45	137.9	6.68×10^{-3}	[390]
366.48	3.4	2.47×10^{-1}	[385]
366.48	13.8	5.54×10^{-2}	[387]
366.48	13.8	5.86×10^{-2}	[385]
366.48	27.6	3.08×10^{-2}	[387]
366.48	31.0	2.64×10^{-2}	[385]
366.48	55.2	1.63×10^{-2}	[387]
366.48	65.5	1.32×10^{-2}	[385]
366.48	110.3	6.51×10^{-3}	[387]
366.48	137.9	6.68×10^{-3}	[385]
422.04	31.0	1.570×10^{-1}	[385]
422.04	65.5	7.53×10^{-2}	[385]
422.04	103.4	5.13×10^{-2}	[385]
423.15	50	1.06×10^{-1}	[391]
423.15	100	6.00×10^{-2}	[391]
423.15	150	4.40×10^{-2}	[391]
423.15	200	3.60×10^{-2}	[391]
423.15	250	3.12×10^{-2}	[391]
423.15	300	2.83×10^{-2}	[391]
448.15	50	2.06×10^{-1}	[391]
448.15	100	1.21×10^{-1}	[391]
448.15	150	9.27×10^{-2}	[391]
448.15	200	7.80×10^{-2}	[391]
448.15	250	6.88×10^{-2}	[391]
448.15	300	6.23×10^{-2}	[391]
473.15	50	2.66×10^{-1}	[391]
473.15	100	1.48×10^{-1}	[391]
473.15	150	1.08×10^{-1}	[391]
473.15	200	8.8×10^{-2}	[391]
473.15	250	7.6×10^{-2}	[391]
473.15	300	6.9×10^{-2}	[391]
477.59	27.6	6.34×10^{-1}	[387]
477.59	31.0	5.55×10^{-1}	[385]
477.59	55.2	3.26×10^{-1}	[387]
477.59	65.5	2.85×10^{-1}	[385]
477.59	103.4	1.88×10^{-1}	[385]

Continued on next page

$T/$[K]	$P/$[bar]	y_{H_2O}	Ref
477.59	110.3	1.72×10^{-1}	[387]
498.15	50	5.54×10^{-1}	[391]
498.15	100	3.22×10^{-1}	[391]
498.15	150	2.44×10^{-1}	[391]
498.15	200	2.08×10^{-1}	[391]
498.15	250	1.83×10^{-1}	[391]
498.15	300	1.67×10^{-1}	[391]
523.15	50	8.26×10^{-1}	[391]
523.15	100	4.96×10^{-1}	[391]
523.15	150	3.83×10^{-1}	[391]
523.15	200	3.27×10^{-1}	[391]
523.15	250	2.94×10^{-1}	[391]
523.15	300	2.71×10^{-1}	[391]
548.15	100	6.86×10^{-1}	[391]
548.15	150	5.36×10^{-1}	[391]
548.15	200	4.61×10^{-1}	[391]
548.15	250	4.16×10^{-1}	[391]
548.15	300	3.86×10^{-1}	[391]
573.15	100	9.00×10^{-1}	[391]
573.15	150	7.22×10^{-1}	[391]
573.15	200	6.33×10^{-1}	[391]
573.15	250	5.82×10^{-1}	[391]
573.15	300	5.51×10^{-1}	[391]
574.81	110.3	8.45×10^{-1}	[387]
588.7	110.3	9.67×10^{-1}	[387]
588.71	137.9	8.10×10^{-1}	[385]

References

(1) Çengel, Y.; A., B. M., *Thermodynamics: an engineering approach*, 8th ed.; McGraw-Hill: New York, USA, 2015.

(2) Smith, J. M.; Van Ness, H. C.; Abbott, M. M., *Introduction to chemical engineering thermodynamics*, 7th ed.; McGraw-Hill: New York, USA, 2005.

(3) Moran, M. J.; Shapiro, H. N., *Fundamentals of engineering thermodynamics*, 5th ed.; John Wiley & Sons: West Sussex, England, 2006.

(4) McQuarrie, D. A.; Simon, J. D., *Physical chemistry: a molecular approach*, 1st ed.; University Science Books: Sausalito, California, 1997.

(5) Sandler, S. I., *Chemical, biochemical, and engineering thermodynamics*, 4th ed.; John Wiley & Sons: Hoboken, N.J., USA, 2006.

(6) Hendriks, E.; Kontogeorgis, G. M.; Dohrn, R.; de Hemptinne, J.-C.; Economou, I. G.; Žilnik, L. F.; Vesovic, V. Industrial requirements for thermodynamics and transport properties. *Ind. Eng. Chem. Res.* **2010**, *49*, 11131–11141.

(7) Kontogeorgis, G. M.; Folas, G. K., *Thermodynamic models for industrial applications: from classical and advanced mixing rules to association theories*, 1st ed.; John Wiley & Sons: Wiltshire, Great Britain, 2009.

(8) Michelsen, M.; Mollerup, J. M., *Thermodynamic models: fundamental & computational aspects*, 2nd ed.; Tie-Line Publications: Holte, Denmark, 2007.

(9) Harstad, K. G.; Miller, R. S.; Bellan, J. Efficient high-pressure state equations. *AIChE J.* **1997**, *43*, 1605–1610.

(10) Jhaveri, B. S.; Youngren, G. K. Three-parameter modification of the Peng-Robinson equation of state to improve volumetric predictions. *SPE reservoir engineering* **1988**, *3*, 1033–1040.

(11) Poling, B. E.; Prausnitz, J. M.; O'Connell, J. P., *The properties of gases and liquids*, 5th ed.; McGraw-Hill New York: New York, USA, 2001.

(12) Valderrama, J. O. The state of the cubic equations of state. *Ind. Eng. Chem. Res.* **2003**, *42*, 1603–1618.

(13) Lopez-Echeverry, J. S.; Reif-Acherman, S.; Araujo-Lopez, E. Peng-Robinson equation of state: 40 years through cubics. *Fluid Phase Equilib.* **2017**, *447*, 39 –71.

(14) Frenkel, D.; Smit, B., *Understanding molecular simulation: from algorithms to applications*, 2nd ed.; Academic Press: San Diego, California, 2002.

(15) Allen, M. P.; Tildesley, D. J., *Computer simulation of liquids*, 2nd ed.; Oxford university press: Oxford, United Kingdom, 2017.

(16) Onsager, L. Crystal statistics. I. a two-dimensional model with an order-disorder transition. *Phys. Rev.* **1944**, *65*, 117–149.

(17) Vlugt, T. J. H.; van der Eerden, J. P. J. M.; Dijkstra, M.; Smit, B.; and Frenkel, D. Introduction to molecular simulation and statistical thermodynamics. available from: `http://homepage.tudelft.nl/v9k6y/imsst/index.html`, 2009.

(18) Gross, J.; Sadowski, G. Perturbed-Chain SAFT: an equation of state based on a perturbation theory for chain molecules. *Ind. Eng. Chem. Res.* **2001**, *40*, 1244–1260.

(19) Diamantonis, N. I.; Economou, I. G. Evaluation of Statistical Associating Fluid Theory (SAFT) and perturbed chain-SAFT equations of state for the calculation of thermodynamic derivative properties of fluids related to carbon capture and sequestration. *Energy Fuels* **2011**, *25*, 3334–3343.

(20) Rahbari, A.; Hens, R.; Nikolaidis, I. K.; Poursaeidesfahani, A.; Ramdin, M.; Economou, I. G.; Moultos, O. A.; Dubbeldam, D.; Vlugt, T. J. H. Computation of partial molar properties using Continuous Fractional Component Monte Carlo. *Mol. Phys.* **2018**, *116*, 3331–3344.

(21) Frenkel, D. In *Computer modelling of fluids polymers and solids*, Catlow, C. R. A., Parker, S. C., Allen, M. P., Eds.; Springer Netherlands: Dordrecht, 1990, pp 83–123.

(22) Frenkel, D. Introduction to Monte Carlo methods. *NIC series* **2004**, *23*, 2–60.

(23) Visualizing the trillion-fold Increase in computing power. available from: `https://www.visualcapitalist.com/visualizing-trillion-fold-increase-computing-power`.

(24) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.

(25) Rosenbluth, M. N.; Rosenbluth, A. W. Monte Carlo calculation of the average extension of molecular chains. *J. Comput. Phys.* **1955**, *23*, 356–359.

(26) Frenkel, D.; Smit, B. Simulation of phase coexistence for complex molecules. *Computers in physics* **1997**, *11*, 246–255.

(27) Dougan, L.; Bates, S. P.; Hargreaves, R.; Fox, J. P.; Crain, J.; Finney, J. L.; Réat, V.; Soper, A. K. Methanol-water solutions: a bi-percolating liquid mixture. *J. Chem. Phys.* **2004**, *121*, 6456–6462.

(28) Dixit, S; Crain, J; Poon, W.; Finney, J.; Soper, A. Molecular segregation observed in a concentrated alcohol–water solution. *Nature* **2002**, *416*, 829.

(29) Dixit, S.; Soper, A. K.; Finney, J. L.; Crain, J. Water structure and solute association in dilute aqueous methanol. *Europhys. Lett.* **2002**, *59*, 377.

(30) Shiflett, M. B.; Maginn, E. J. The solubility of gases in ionic liquids. *AIChE J.* **2017**, *63*, 4722–4737.

(31) Ramdin, M.; de Loos, T. W.; Vlugt, T. J. H. State-of-the-art of CO_2 capture with ionic liquids. *Ind. Eng. Chem. Res.* **2012**, *51*, 8149–8177.

(32) Rahbari, A.; Brenkman, J.; Hens, R.; Ramdin, M.; van den Broeke, L. J. P.; Schoon, R.; Henkes, R.; Moultos, O. A.; Vlugt, T. J. H. Solubility of water in hydrogen at high Pressures: a molecular simulation study. *J. Chem. Eng. Data* **2019**, *64*, 4103–4115.

(33) Zou, R.; Zhong, R.; Han, S.; Xu, H.; Burrell, A. K.; Henson, N.; Cape, J. L.; Hickmott, D. D.; Timofeeva, T. V.; Larson, T. E.; Zhao, Y. A Porous metal-organic replica of α-PbO_2 for capture of nerve agent surrogate. *J. Amer. Chem. Soc.* **2010**, *132*, 17996–17999.

(34) Vieira Soares, C.; Maurin, G.; Leitão, A. A. Computational exploration of the catalytic degradation of sarin and its simulants by a titanium metal-organic framework. *J. Phys. Chem. C* **2019**, *123*, 19077–19086.

(35) Jamali, S. H.; Wolff, L.; Becker, T. M.; de Groen, M.; Ramdin, M.; Hartkamp, R.; Bardow, A.; Vlugt, T. J. H.; Moultos, O. A. OCTP: a tool for on-the-fly calculation of transport properties of fluids with the order-n algorithm in LAMMPS. *Journal of Chemical Information and Modeling* **2019**, *59*, PMID: 30742429, 1290–1294.

(36) Model 3 owner's manual., `https://www.tesla.com/sites/default/files/model_3_owners_manual_north_america_en.pdf`, Palo Alto, United States.

(37) Smith, W. R. The computation of chemical equilibria in complex systems. *Ind. Eng. Chem. Fundam.* **1980**, *19*, 1–10.

(38) Poursaeidesfahani, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H. Direct free energy calculation in the Continuous Fractional Component Gibbs ensemble. *J. Chem. Theory Comput.* **2016**, *12*, 1481–1490.

(39) Smit, B.; Frenkel, D. Calculation of the chemical potential in the Gibbs ensemble. *Mol. Phys.* **1989**, *68*, 951–958.

(40) Widom, B. Some topics in the theory of fluids. *J. Chem. Phys.* **1963**, *39*, 2808–2812.

(41) Rahbari, A.; Hens, R.; Dubbeldam, D.; Vlugt, T. J. H. Improving the accuracy of computing chemical potentials in CFCMC simulations. *Mol. Phys.* **2019**, *117*, 3493–3508.

(42) Panagiotopoulos, A. Z. Direct determination of fluid-phase equilibria by simulation in the Gibbs ensemble - a review. *Mol. Simul.* **1992**, *9*, 1–23.

(43) Panagiotopoulos, A. Z. Molecular simulation of phase equilibria: simple, ionic and polymeric fluids. *Fluid Phase Equilib.* **1992**, *76*, 97–112.

(44) Panagiotopoulos, A. Z. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Mol. Phys.* **1987**, *61*, 813–826.

(45) Shi, W.; Maginn, E. J. Continuous Fractional Component Monte Carlo: an adaptive biasing method for open system atomistic simulations. *J. Chem. Theory Comput.* **2007**, *3*, 1451–1463.

(46) Siepmann, J. I.; Frenkel, D. Configurational bias Monte Carlo: a new sampling scheme for flexible chains. *Mol. Phys.* **1992**, *75*, 59–70.

(47) Laso, M.; de Pablo, J. J.; Suter, U. W. Simulation of phase equilibria for chain molecules. *J. Chem. Phys.* **1992**, *97*, 2817–2819.

(48) Vlugt, T. J. H.; Martin, M. G.; Smit, B; Siepmann, J. I.; Krishna, R Improving the efficiency of the configurational-bias Monte Carlo algorithm. *Mol. Phys.* **1998**, *94*, 727–733.

(49) Vlugt, T. J. H. Efficiency of parallel CBMC simulations. *Mol. Simul.* **1999**, *23*, 63–78.

(50) Escobedo, F. A.; de Pablo, J. J. Extended continuum configurational bias Monte Carlo methods for simulation of flexible molecules. *J. Chem. Phys.* **1995**, *102*, 2636–2652.

(51) Frenkel, D; Mooij, G. C. A. M.; Smit, B Novel scheme to study structural and thermal properties of continuously deformable molecules. *J. Phys.: Condens. Matt.* **1992**, *4*, 3053–3076.

(52) De Pablo, J. J.; Laso, M.; Suter, U. W. Estimation of the chemical potential of chain molecules by simulation. *J. Chem. Phys.* **1992**, *96*, 6157–6162.

(53) Consta, S.; Vlugt, T. J. H.; Hoeth, J. W.; Smit, B.; Frenkel, D. Recoil growth algorithm for chain molecules with continuous interactions. *Mol. Phys.* **1999**, *97*, 1243–1254.

(54) Combe, N.; Vlugt, T. J. H.; Wolde, P. R. T.; Frenkel, D. Dynamic pruned-enriched Rosenbluth method. *Mol. Phys.* **2003**, *101*, 1675–1682.

(55) Houdayer, J. The wormhole move: a new algorithm for polymer simulations. *J. Chem. Phys.* **2002**, *116*, 1783–1787.

(56) Boon, N. Efficient configurational-bias Monte-Carlo simulations of chain molecules with "swarms" of trial configurations. *J. Chem. Phys.* **2018**, *149*, 064109.

(57) Martin, M. G.; Siepmann, J. I. Novel Configurational-Bias Monte Carlo method for branched molecules. transferable potentials for phase equilibria. 2. united-atom description of branched alkanes. *J. Phys. Chem. B* **1999**, *103*, 4508–4517.

(58) Shah, J. K.; Maginn, E. J. A general and efficient Monte Carlo method for sampling intramolecular degrees of freedom of branched and cyclic molecules. *J. Chem. Phys.* **2011**, *135*, 134121.

(59) Rahbari, A.; Poursaeidesfahani, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H. Chemical potentials of water, methanol, carbon dioxide and hydrogen sulphide at low temperatures using continuous fractional component Gibbs ensemble Monte Carlo. *Mol. Simul.* **2018**, *44*, 405–414.

(60) HyET Group — High yield Energy Technologies., `https://www.hyetgroup.com`, Arnhem, the Netherlands.

(61) Bartlett, E. P. The concentration of water vapor in compressed hydrogen, nitrogen and a mixture of these gases in the presence of condensed water. *J. Amer. Chem. Soc.* **1927**, *49*, 65–78.

(62) Winchester, L. J.; Dodge, B. F. The chemical equilibrium of the ammonia synthesis reaction at high temperatures and extreme pressures. *AIChE J.* **1956**, *2*, 431–436.

(63) Lísal, M.; Bendová, M.; Smith, W. R. Monte Carlo adiabatic simulation of equilibrium reacting systems: The ammonia synthesis reaction. *Fluid Phase Equilib.* **2005**, *235*, 50–57.

(64) Peng, X.; Wang, W.; Huang, S. Monte Carlo simulation for chemical reaction equilibrium of ammonia synthesis in MCM-41 pores and pillared clays. *Fluid Phase Equilib.* **2005**, *231*, 138–149.

(65) Gillespie, L. J.; Beattie, J. A. The thermodynamic treatment of chemical equilibria in systems composed of real gases. I. an approximate equation for the mass action function applied to the existing data on the Haber equilibrium. *Phys. Rev.* **1930**, *36*, 743–753.

(66) Turner, C. H.; Johnson, J. K.; Gubbins, K. E. Effect of confinement on chemical reaction equilibria: The reactions $2NO \Leftrightarrow (NO)_2$ and $N_2 + 3H_2 \Leftrightarrow 2NH_3$ in carbon micropores. *J. Comput. Phys.* **2001**, *114*, 1851–1859.

(67) Jennings, J. R., *Catalytic ammonia synthesis: fundamentals and practice*, 1st ed.; Plenum Press: New York, USA, 1991.

(68) Xiao, H.; Valera-Medina, A.; Marsh, R.; Bowen, P. J. Numerical study assessing various ammonia/methane reaction models for use under gas turbine conditions. *Fuel* **2017**, *196*, 344–351.

(69) Nozari, H.; Karabeyoğlu, A. Numerical study of combustion characteristics of ammonia as a renewable fuel and establishment of reduced reaction mechanisms. *Fuel* **2015**, *159*, 223–233.

(70) Xiao, H.; Howard, M.; Valera-Medina, A.; Dooley, S.; Bowen, P. J. Study on reduced chemical mechanisms of ammonia/methane combustion under gas turbine conditions. *Energy Fuels* **2016**, *30*, 8701–8710.

(71) Zamfirescu, C.; Dincer, I. Using ammonia as a sustainable fuel. *J. Power Sources* **2008**, *185*, 459–465.

(72) Zamfirescu, C.; Dincer, I. Ammonia as a green fuel and hydrogen source for vehicular applications. *Fuel Process. Technol.* **2009**, *90*, 729–737.

(73) Kitano, M.; Inoue, Y.; Yamazaki, Y.; Hayashi, F.; Kanbara, S.; Matsuishi, S.; Yokoyama, T.; Kim, S.-W.; Hara, M.; Hosono, H. Ammonia synthesis using a stable electride as an electron donor and reversible hydrogen store. *Nature Chem.* **2012**, *4*, 934–940.

(74) van der Ham, C. J.; Koper, M. T.; Hetterscheid, D. G. Challenges in reduction of dinitrogen by proton and electron transfer. *Chem. Soc. Rev.* **2014**, *43*, 5183–5191.

(75) Guo, X.; Zhu, Y.; Ma, T. Lowering reaction temperature: Electrochemical ammonia synthesis by coupling various electrolytes and catalysts. *J. Energy Chem.* **2017**, *26*, 1107–1116.

(76) Onufriev, A. V.; Izadi, S. Water models for biomolecular simulations. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1347.

(77) Vega, C.; Abascal, J. L. F. Simulating water with rigid non-polarizable models: a general perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663–19688.

(78) Guillot, B. A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.* **2002**, *101*, 219 –260.

(79) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(80) Tsimpanogiannis, I. N.; Moultos, O. A.; Franco, L. F. M.; de M. Spera, M. B.; Erdős, M.; Economou, I. G. Self-diffusion coefficient of bulk and confined water: a critical review of classical molecular simulation studies. *Mol. Simul.* **2019**, *45*, 425–453.

(81) Tröster, P.; Lorenzen, K.; Tavan, P. Polarizable six-point water models from computational and empirical optimization. *J. Phys. Chem. B* **2014**, *118*, 1589–1602.

(82) Finney, J. L. The water molecule and its interactions: the interaction between theory, modelling, and experiment. *J. Mol. Liq.* **2001**, *90*, 303 –312.

(83) Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **2008**, *108*, 74–108.

(84) Zhong, Y.; Warren, G. L.; Patel, S. Thermodynamic and structural properties of methanol-water solutions using nonadditive interaction models. *J. Comput. Chem.* **2008**, *29*, 1142–1152.

(85) Martínez-Jiménez, M.; Saint-Martin, H. A Four-Site Molecular Model for Simulations of Liquid Methanol and Water-Methanol Mixtures: MeOH-4P. *J. Chem. Theory Comput.* **2018**, *14*, 2526–2537.

(86) Dopazo-Paz, A.; Gómez-Álvarez, P.; González-Salgado, D. Thermodynamics and structure of the {water+methanol} system viewed from three simple additive pair-wise intermolecular potentials based on the rigid molecule approximation. *Collect. Czech. Chem. Commun.* **2010**, *75*, 617–635.

(87) Moučka, F.; Nezbeda, I. Water-methanol mixtures with non-Lorentz-Berthelot combining rules: a feasibility study. *J. Mol. Liq.* **2011**, *159*, 47–51.

(88) Ferrario, M.; Haughney, M.; McDonald, I. R.; Klein, M. L. Molecular-dynamics simulation of aqueous mixtures: Methanol, acetone, and ammonia. *J. Chem. Phys.* **1990**, *93*, 5156–5166.

(89) Mallamace, F.; Corsaro, C.; Mallamace, D.; Vasi, C.; Vasi, S.; Stanley, H. E. Dynamical properties of water-methanol solutions. *J. Chem. Phys.* **2016**, *144*, 064506.

(90) Matsumoto, M.; Takaoka, Y.; Kataoka, Y. Liquid-vapor interface of water-methanol mixture. I. Computer simulation. *J. Chem. Phys.* **1993**, *98*, 1464–1472.

(91) Wensink, E. J. W.; Hoffmann, A. C.; van Maaren, P. J.; van der Spoel, D. Dynamic properties of water/alcohol mixtures studied by computer simulation. *J. Chem. Phys.* **2003**, *119*, 7308–7317.

(92) González-Salgado, D.; Nezbeda, I. Excess properties of aqueous mixtures of methanol: simulation versus experiment. *Fluid Phase Equilib.* **2006**, *240*, 161–166.

(93) Wormald, C.; Badock, L.; Lloyd, M. Excess enthalpies for (water + methanol) at $T = 423$ K to $T = 523$ and pressures up to 20 MPa. a new flow mixing calorimeter. *J. Chem. Thermodyn.* **1996**, *28*, 603 –613.

(94) Wormald, C.; Yerlett, T. Molar enthalpy increments for (0.5 H_2O + 0.5 CH_3OH) at temperatures up to 573.2 K and pressures up to 13.0 MPa. *J. Chem. Thermodyn.* **2000**, *32*, 97–105.

(95) Simonson, J.; Bradley, D.; Busey, R. Excess molar enthalpies and the thermodynamics of (methanol + water) to 573 K and 40 MPa. *J. Chem. Thermodyn.* **1987**, *19*, 479–492.

(96) Kuroki, T.; Kagawa, N.; Endo, H.; Tsuruno, S.; Magee, J. W. Specific heat capacity at constant volume for water, methanol, and their mixtures at temperatures from 300 K to 400 K and pressures to 20 MPa. *J. Chem. Eng. Data* **2001**, *46*, 1101–1106.

(97) Xiao, C.; Bianchi, H.; Tremaine, P. R. Excess molar volumes and densities of (methanol+water) at temperatures between 323 K and 573 K and pressures of 7.0 MPa and 13.5 MPa. *J. Chem. Thermodyn.* **1997**, *29*, 261–286.

(98) Lama, R. F.; Lu, B. C.-Y. Excess thermodynamic properties of aqueous alcohol solutions. *J. Chem. Eng. Data* **1965**, *10*, 216–219.

(99) Vlček, L; Nezbeda, I Excess properties of aqueous mixtures of methanol: simple models versus experiment. *J. Mol. Liq.* **2007**, *131-132*, 158–162.

(100) Tomaszkiewicz, I.; Randzio, S.; Gierycz, P. Excess enthalpy in the methanol-water system at 278.15, 298.15 and 323.15 K under pressures of 0.1, 20 and 39 MPa: II. experimental results and their analytical presentation. *Thermochimica Acta* **1986**, *103*, 281–289.

(101) Gmehling, J; Onken, U; Arlt, W; Grenzheuser, P; Weidlich, U; Kolbe, B; Rarey, J, *Chemistry data series, vapor-liquid equilibrium data collection, Vol. 1, Part 1c*, 1st ed.; DECHEMA: Frankfurt am Main, Germany, 1986.

(102) Mikhail, S. Z.; Kimel, W. R. Densities and viscosities of methanol-water mixtures. *J. Chem. Eng. Data* **1961**, *6*, 533–537.

(103) Tanaka, H.; Gubbins, K. E. Structure and thermodynamic properties of water-methanol mixtures: role of the water-water interaction. *J. Chem. Phys.* **1992**, *97*, 2626–2634.

(104) Schrader, J.; Schilling, M.; Holtmann, D.; Sell, D.; Villela Filho, M.; Marx, A.; Vorholt, J. A. Methanol-based industrial biotechnology: current status and future perspectives of methylotrophic bacteria. *Trends. Biotechnol.* **2009**, *27*, 107–115.

(105) Li, X.; Faghri, A. Review and advances of direct methanol fuel cells (DMFCs) part I: design, fabrication, and testing with high concentration methanol solutions. *J. Power Sources* **2013**, *226*, 223 –240.

(106) Olah, G. A. Beyond oil and gas: the methanol economy. *Angew. Chem. Int. Ed.* **2005**, *44*, 2636–2639.

(107) Hogarth, M. P.; Hards, G. A. Direct methanol fuel cells. *Platin. Met. Rev.* **1996**, *40*, 150–159.

(108) Heinzel, A.; Barragán, V. A review of the state-of-the-art of the methanol crossover in direct methanol fuel cells. *J. Power Sources* **1999**, *84*, 70–74.

(109) Aricò, A. S.; Srinivasan, S.; Antonucci, V. DMFCs: from fundamental aspects to technology development. *Fuel Cells* **2001**, *1*, 133–161.

(110) Mahmood, A.; Bano, S.; Kim, S.-G.; Lee, K.-H. Water-methanol separation characteristics of annealed SA/PVA complex membranes. *J. Memb. Sci.* **2012**, *415-416*, 360 –367.

(111) Shah, D; Kissick, K; Ghorpade, A; Hannah, R; Bhattacharyya, D Pervaporation of alcohol-water and dimethylformamide-water mixtures using hydrophilic zeolite NaA membranes: mechanisms and experimental results. *J. Memb. Sci.* **2000**, *179*, 185–205.

(112) Winarto; Takaiwa, D.; Yamamoto, E.; Yasuoka, K. Water-methanol separation with carbon nanotubes and electric fields. *Nanoscale* **2015**, *7*, 12659–12665.

(113) Bergmann, D. L.; Eckert, C. A. Measurement of limiting activity coefficients for aqueous systems by differential ebulliometry. *Fluid Phase Equilib.* **1991**, *63*, 141 –150.

(114) Slusher, J. T. Infinite dilution activity coefficients in hydrogen-bonded mixtures via molecular dynamics: the water/methanol system. *Fluid Phase Equilibria* **1999**, *154*, 181 –192.

(115) Sherman, S. R.; Trampe, D. B.; Bush, D. M.; Schiller, M.; Eckert, C. A.; Dallas, A. J.; Li, J.; Carr, P. W. Compilation and correlation of limiting activity coefficients of nonelectrolytes in water. *Ind. Eng. Chem. Res.* **1996**, *35*, 1044–1058.

(116) Eckert, C. A.; Sherman, S. R. Measurement and prediction of limiting activity coefficients. *Fluid Phase Equilib.* **1996**, *116*, Proceedings of the Seventh International Conference on Fluid Properties and Phase Equilibria for Chemical Process Design, 333 –342.

(117) Gmehling, J.; Kolbe, B.; Kleiber, M.; Rarey, J., *Chemical thermodynamics for process simulation*, 1st ed.; Wiley-VCH Verlag & GmbH Co. KGaA.: Weinheim, Germany, 2012.

(118) Walter, L. Carbon dioxide as a raw material: the synthesis of formic acid and its derivatives from CO_2. *Angew. Chem. Int. Ed.* **1995**, *34*, 2207–2221.

(119) Saito, K.; Shiose, T.; Takahashi, O.; Hidaka, Y.; Aiba, F.; Tabayashi, K. Unimolecular decomposition of formic acid in the gas phase on the ratio of the competing reaction channels. *J. Phys. Chem. A* **2005**, *109*, 5352–5357.

(120) Saito, K.; Kakumoto, T.; Kuroda, H.; Torii, S.; Imamura, A. Thermal unimolecular decomposition of formic acid. *J. Chem. Phys.* **1984**, *80*, 4989–4996.

(121) Columbia, M.; Thiel, P. The interaction of formic acid with transition metal surfaces, studied in ultrahigh vacuum. *J. Electroanal Chem.* **1994**, *369*, 1–14.

(122) Iglesia, E.; Boudart, M. Decomposition of formic acid on copper, nickel, and copper-nickel alloys: III. Catalytic decomposition on nickel and copper-nickel alloys. *J. Catal.* **1983**, *81*, 224–238.

(123) Iglesia, E.; Boudart, M. Decomposition of formic acid on copper, nickel, and copper-nickel alloys: I. Preparation and characterization of catalysts. *J. Catal.* **1983**, *81*, 204–213.

(124) Iglesia, E.; Boudart, M. Decomposition of formic acid on copper, nickel, and copper-nickel alloys: II. Catalytic and temperature-programmed decomposition of formic acid on $CuSiO_2$, $CuAl_2O_3$, and Cu powder. *J. Catal.* **1983**, *81*, 214–223.

(125) Iglesia, E. Unimolecular and bimolecular formic acid decomposition on copper. *J. Phys. Chem.* **1986**, *90*, 5272–5274.

(126) Herron, J. A.; Scaranto, J.; Ferrin, P.; Li, S.; Mavrikakis, M. Trends in formic acid decomposition on model transition metal surfaces: a density functional theory study. *ACS Catal.* **2014**, *4*, 4434–4445.

(127) Lee, H. J.; Kang, D.-C.; Pyen, S. H.; Shin, M.; Suh, Y.-W.; Han, H.; Shin, C.-H. Production of H_2-free CO by decomposition of formic acid over ZrO_2 catalysts. *Appl. Catal., A* **2017**, *531*, 13–20.

(128) Yu, J.; Savage, P. E. Decomposition of formic acid under hydrothermal conditions. *Ind. Eng. Chem. Res.* **1998**, *37*, 2–10.

(129) Bjerre, A. B.; Soerensen, E. Thermal decomposition of dilute aqueous formic acid solutions. *Ind. Eng. Chem. Res.* **1992**, *31*, 1574–1577.

(130) Müller, K.; Brooks, K.; Autrey, T. Hydrogen storage in formic acid: a comparison of process options. *Energy & Fuels* **2017**, *31*, 12603–12611.

(131) López, F.; Stone, F. Formic acid decomposition over α-chromia-alumina solid solution catalysts. *Zeitschrift für Phys. Chem.* **1978**, *111*, 247–256.

(132) Wescott, B. B.; Engelder, C. J. The catalytic decomposition of formic acid. *J. Phys. Chem.* **1926**, *30*, 476–479.

(133) Blake, P. G.; Hinshelwood Cyril Norman, P. The homogeneous decomposition reactions of gaseous formic acid. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **1960**, *255*, 444–455.

(134) Blake, P. G.; Davies, H. H.; Jackson, G. E. Dehydration mechanisms in the thermal decomposition of gaseous formic acid. *J. Chem. Soc. B* **1971**, 1923–1925.

(135) Hinshelwood, C. N.; Topley, B. CXV.-The energy of activation in heterogeneous gas reactions with relation to the thermal decomposition of formic acid vapour. *J. Chem. Soc., Trans.* **1923**, *123*, 1014–1025.

(136) Hinshelwood, C. N.; Topley, B. The influence of temperature on two alternative modes of decomposition of formic acid. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **1922**, *100*, 575–581.

(137) Alvarez, A.; Bansode, A.; Urakawa, A.; Bavykina, A. V.; Wezendonk, T. A.; Makkee, M.; Gascon, J.; Kapteijn, F. Challenges in the greener production of formates/formic acid, methanol, and DME by heterogeneously catalyzed CO_2 hydrogenation Processes. *Chem. Rev.* **2017**, *117*, 9804–9838.

(138) Yoshida, K.; Wakai, C.; Matubayasi, N.; Nakahara, M. NMR spectroscopic evidence for an intermediate of formic acid in the water-gas-shift Reaction. *J. Phys. Chem. A* **2004**, *108*, 7479–7482.

(139) Chihiro, W.; Ken, Y.; Yasuo, T.; Nobuyuki, M.; Masaru, N. Effect of concentration, acid, temperature, and metal on competitive reaction pathways for decarbonylation and decarboxylation of formic acid in hot water. *Chem. Lett.* **2004**, *33*, 572–573.

(140) Bröll, D.; Kaul, C.; Krämer, A.; Krammer, P.; Richter, T.; Jung, M.; Vogel, H.; Zehner, P. Chemistry in supercritical water. *Angew. Chem. Int. Ed.* **1999**, *38*, 2998–3014.

(141) Shi, W.; 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.* **2008**, *29*, 2520–2530.

(142) Poursaeidesfahani, A.; Hens, R.; Rahbari, A.; Ramdin, M.; Dubbeldam, D.; Vlugt, T. J. H. Efficient application of Continuous Fractional Component Monte Carlo in the reaction ensemble. *J. Chem. Theory Comput.* **2017**, *13*, 4452–4466.

(143) Wagner, W.; Pruß, A. The IAPWS formulation 1995 for the thermodynamic properties of ordinary water substance for general and scientific use. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387–535.

(144) Lemmon, E. W.; Huber, M. L.; McLinden, M. O. NIST reference fluid thermodynamic and transport properties–REFPROP. *NIST standard reference database* **2002**, *23*, v7.

(145) Lemmon, E. W.; Span, R. Short fundamental equations of state for 20 industrial fluids. *J. Chem. Eng. Data* **2006**, *51*, 785–850.

(146) Ewald, P. P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik* **1921**, *369*, 253–287.

(147) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. Classical electrostatics for biomolecular simulations. *Chemical Reviews* **2014**, *114*, 779–814.

(148) Toukmaji, A. Y.; Board Jr, J. A. Ewald summation techniques in perspective: a survey. *Comput. Phys. Commun.* **1996**, *95*, 73–92.

(149) Fennell, C. J.; Gezelter, J. D. Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. *J. Chem. Phys.* **2006**, *124*, 234104.

(150) Sindzingre, P.; Ciccotti, G.; Massobrio, C.; Frenkel, D. Partial enthalpies and related quantities in mixtures from computer simulation. *Chem. Phys. Lett.* **1987**, *136*, 35–41.

(151) Peng, D.-Y.; Robinson, D. B. A new two-constant equation of state. *Ind. Eng. Chem. Fundam.* **1976**, *15*, 59–64.

(152) Josephson, T. R.; Singh, R.; Minkara, M. S.; Fetisov, E. O.; Siepmann, J. I. Partial molar properties from molecular simulation using multiple linear regression. *Mol. Phys.* **2019**, *117*, 3589–3602.

(153) Sindzingre, P.; Massobrio, C.; Ciccotti, G.; Frenkel, D. Calculation of partial enthalpies of an argon-krypton mixture by NPT molecular dynamics. *Chem. Phys.* **1989**, *129*, 213–224.

(154) Bouwman, P. Electrochemical Hydrogen Compression (EHC) solutions for hydrogen infrastructure. *Fuel Cells Bulletin* **2014**, *2014*, 12–16.

(155) *Hydrogen fuel - Product specification - Part 2: Proton exchange membrane (PEM) fuel cell applications for road vehicles*; International Standard, `www.iso.org/committee/54560/x/catalogue/`; Geneva, CH: International Organization for Standardization, 2012.

(156) Raabe, G.; Maginn, E. J. Molecular modeling of the vapor-liquid equilibrium properties of the alternative refrigerant 2,3,3,3-tetrafluoro-1-propene (HFO-1234yf). *J. Phys. Chem. Lett.* **2010**, *1*, 93–96.

(157) Maginn, E. J. From discovery to data: What must happen for molecular simulation to become a mainstream chemical engineering tool. *AIChE Journal* **2009**, *55*, 1304–1310.

(158) Sheridan, Q. R.; Mullen, R. G.; Lee, T. B.; Maginn, E. J.; Schneider, W. F. Hybrid computational strategy for predicting CO_2 solubilities in reactive ionic liquids. *J. Phys. Chem. C* **2018**, *122*, 14213–14221.

(159) Shiflett, M. B.; Maginn, E. J. The solubility of gases in ionic liquids. *AIChE Journal* **2017**, *63*, 4722–4737.

(160) Paluch, A. S.; Maginn, E. J. Predicting the solubility of solid phenanthrene: a combined molecular simulation and group contribution approach. *AIChE Journal* **2013**, *59*, 2647–2661.

(161) Fetisov, E. O.; Siepmann, J. I. Prediction of vapor–liquid coexistence properties and critical points of polychlorinated biphenyls from Monte Carlo simulations with the TraPPE–EH force field. *J. Chem. Eng. Data* **2014**, *59*, 3301–3306.

(162) Becker, T. M.; Luna-Triguero, A.; Vicent-Luna, J. M.; Lin, L.-C.; Dubbeldam, D.; Calero, S.; Vlugt, T. J. H. Potential of polarizable force fields for predicting the separation performance of small hydrocarbons in M-MOF-74. *Phys. Chem. Chem. Phys.* **2018**, *20*, 28848–28859.

(163) Hens, R.; Vlugt, T. J. H. Molecular simulation of vapor-liquid equilibria using the Wolf method for electrostatic interactions. *J. Chem. Eng. Data* **2018**, *63*, 1096–1102.

(164) Martin, M. G.; Siepmann, J. I. Calculating Gibbs free energies of transfer from Gibbs ensemble Monte Carlo simulations. *Theor. Chem. Acc.* **1998**, *99*, 347–350.

(165) Chen, B.; Siepmann, J. I.; Klein, M. L. Direct Gibbs ensemble Monte Carlo simulations for solid-vapor phase equilibria: applications to Lennard-Jonesium and carbon dioxide. *J. Phys. Chem. B* **2001**, *105*, 9840–9848.

(166) Bai, P.; 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.* **2017**, *13*, 431–440.

(167) Dinpajooh, M.; Bai, P.; Allan, D. A.; Siepmann, J. I. Accurate and precise determination of critical properties from Gibbs ensemble Monte Carlo simulations. *J. Chem. Phys.* **2015**, *143*, 114113.

(168) Bai, P.; Siepmann, J. I. Selective adsorption from dilute solutions: Gibbs ensemble Monte Carlo simulations. *Fluid Phase Equilib.* **2013**, *351*.

(169) Poursaeidesfahani, A.; de Lange, M. F.; Khodadadian, F.; Dubbeldam, D.; Rigutto, M.; Nair, N.; Vlugt, T. J. H. Product shape selectivity of MFI-type, MEL-type, and BEA-type zeolites in the catalytic hydroconversion of heptane. *J. Catal.* **2017**, *353*, 54 –62.

(170) Castillo, J. M.; Vlugt, T. J. H.; Dubbeldam, D.; Hamad, S.; Calero, S. Performance of chiral zeolites for enantiomeric separation revealed by molecular simulation. *J. Phys. Chem. C* **2010**, *114*, 22207–22213.

(171) Matito-Martos, I.; Rahbari, A.; Martin-Calvo, A.; Dubbeldam, D.; Vlugt, T. J. H.; Calero, S. Adsorption equilibrium of nitrogen dioxide in porous materials. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4189–4199.

(172) Fetisov, E. O.; Kuo, I.-F. W.; Knight, C.; VandeVondele, J.; Van Voorhis, T.; Siepmann, J. I. First-Principles Monte Carlo Simulations of Reaction Equilibria in Compressed Vapors. *ACS. Cent. Sci.* **2016**, *2*, 409–415.

(173) Balaji, S. P.; Gangarapu, S.; Ramdin, M.; Torres-Knoop, A.; Zuilhof, H.; Goetheer, E. L. V.; Dubbeldam, D.; Vlugt, T. J. H. Simulating the reactions of CO_2 in aqueous monoethanolamine solution by reaction ensemble Monte Carlo using the Continuous Fractional Component Method. *J. Chem. Theory Comput.* **2015**, *11*, 2661–2669.

(174) Fetisov, E. O.; Shah, M. S.; Knight, C.; Tsapatsis, M.; Siepmann, J. I. Understanding the reactive adsorption of H_2S and CO_2 in sodium-exchanged zeolites. *ChemPhysChem* **2018**, *19*, 512–518.

(175) Mullen, R. G.; Corcelli, S. A.; Maginn, E. J. Reaction ensemble Monte Carlo simulations of CO_2 absorption in the reactive ionic liquid triethyl(octyl)phosphonium 2-cyanopyrrolide. *J. Phys. Chem. Lett.* **2018**, *9*, 5213–5218.

(176) Mullen, R. G.; Maginn, E. J. Reaction ensemble Monte Carlo simulation of xylene isomerization in bulk phases and under confinement. *J. Chem. Theory Comput.* **2017**, *13*, 4054–4062.

(177) Rosch, T. W.; Maginn, E. J. Reaction ensemble Monte Carlo simulation of complex molecular systems. *J. Chem. Theory Comput.* **2011**, *7*, 269–279.

(178) Neubauer, B; Boutin, A; Tavitian, B; Fuchs, A. Gibbs ensemble simulations of vapour-liquid phase equilibria of cyclic alkanes. *Mol. Phys.* **1999**, *97*, 769–776.

(179) Siepmann, J. I.; McDonald, I. R.; Frenkel, D Finite-size corrections to the chemical potential. *J. Phys.: Condens. Matt.* **1992**, *4*, 679.

(180) Recht, J.; Panagiotopoulos, A. Z. Finite-size effects and approach to criticality in Gibbs ensemble simulations. *Mol. Phys.* **1993**, *80*, 843–852.

(181) Torres-Knoop, A.; Burtch, N. C.; Poursaeidesfahani, A.; Balaji, S. P.; Kools, R.; Smit, F. X.; Walton, K. S.; Vlugt, T. J. H.; 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* **2016**, *120*, 9148–9159.

(182) Boulougouris, G. C.; Economou, I. G.; Theodorou, D. N. On the calculation of the chemical potential using the particle deletion scheme. *Mol. Phys.* **1999**, *96*, 905–913.

(183) Boulougouris, G. C.; Economou, I. G.; Theodorou, D. N. Calculation of the chemical potential of chain molecules using the staged particle deletion scheme. *J. Comput. Phys.* **2001**, *115*, 8231–8237.

(184) Widom, B. Potential-distribution theory and the statistical mechanics of fluids. *J. Phys. Chem.* **1982**, *86*, 869–872.

(185) Kofke, D. A. Free energy methods in molecular simulation. *Fluid Phase Equilib.* **2005**, *228-229*, 41–48.

(186) Torres-Knoop, A.; Balaji, S. P.; Vlugt, T. J. H.; Dubbeldam, D. A comparison of advanced Monte Carlo methods for open systems: CFCMC vs CBMC. *J. Chem. Theory Comput.* **2014**, *10*, 942–952.

(187) Mooij, G. C. A. M.; Frenkel, D. The overlapping distribution method to compute chemical potentials of chain molecules. *J. Phys.: Condens. Matt.* **1994**, *6*, 3879–3888.

(188) Coskuner, O.; Deiters, U. K. Hydrophobic interactions by Monte Carlo simulations. *Zeitschrift für Phys. Chem.* **2006**, *220*, 349–369.

(189) Shing, K.; Gubbins, K. The chemical potential in dense fluids and fluid mixtures via computer simulation. *Mol. Phys.* **1982**, *46*, 1109–1128.

(190) Shing, K.; Gubbins, K. The chemical potential in non-ideal liquid mixtures. *Mol. Phys.* **1983**, *49*, 1121–1138.

(191) Daly, K. B.; Benziger, J. B.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Massively parallel chemical potential calculation on graphics processing units. *Comput. Phys. Commun.* **2012**, *183*, 2054–2062.

(192) Shing, K.; Gubbins, K. Free energy and vapour-liquid equilibria for a quadrupolar Lennard-Jones fluid. *Mol. Phys.* **1982**, *45*, 129–139.

(193) Mon, K. K.; Griffiths, R. B. Chemical potential by gradual insertion of a particle in Monte Carlo simulation. *Phys. Rev. A* **1985**, *31*, 956–959.

(194) Squire, D. R.; Hoover, W. G. Monte Carlo simulation of vacancies in rare-gas crystals. *J. Chem. Phys.* **1969**, *50*, 701–706.

(195) Mruzik, M. R.; Abraham, F. F.; Schreiber, D. E.; Pound, G. M. A Monte Carlo study of ion–water clusters. *J. Chem. Phys.* **1976**, *64*, 481–491.

(196) Escobedo, F. A.; de Pablo, J. J. Expanded grand canonical and Gibbs ensemble Monte Carlo simulation of polymers. *J. Chem. Phys.* **1996**, *105*, 4391–4394.

(197) Yoo, B.; Marin-Rimoldi, E.; Mullen, R. G.; Jusufi, A.; Maginn, E. J. Discrete Fractional Component Monte Carlo simulation study of dilute nonionic surfactants at the air–water interface. *Langmuir* **2017**, *33*, 9793–9802.

(198) Paluch, A. S.; Shah, J. K.; Maginn, E. J. Efficient solvation free energy calculations of amino acid analogs by expanded ensemble molecular simulation. *J. Chem. Theory Comput.* **2011**, *7*, 1394–1403.

(199) Wang, F.; Landau, D. P. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **2001**, *86*, 2050–2053.

(200) Poulain, P.; Calvo, F.; Antoine, R.; Broyer, M.; Dugourd, P. Performances of Wang-Landau algorithms for continuous systems. *Phys. Rev. E* **2006**, *73*, 056704.

(201) Sikora, B. J.; Colón, Y. J.; Snurr, R. Q. Continuous fractional component Monte Carlo simulations of high-density adsorption in metal–organic frameworks. *Mol. Simul.* **2015**, *41*, 1339–1347.

(202) Robinson, G. W.; Singh, S; Zhu, S.-B.; Evans, M. W., *Water in biology, chemistry and physics: experimental overviews and computational methodologies*, 1st ed.; World Scientific: Singapore, 1996; Vol. 9.

(203) Price, D. J.; Brooks III, C. L. A modified TIP3P water potential for simulation with Ewald summation. *J. Comput. Phys.* **2004**, *121*, 10096–10103.

(204) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Comput. Phys.* **2004**, *120*, 9665–9678.

(205) Rick, S. W. A reoptimization of the five-site water potential (TIP5P) for use with Ewald sums. *J. Chem. Phys.* **2004**, *120*, 6085–6093.

(206) Jorgensen, W. L Optimized intermolecular potential functions for liquid alcohols. *J. Phys. Chem.* **1986**, *90*, 1276–1284.

(207) Chen, B.; Potoff, J. J.; Siepmann, J. I. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. united-atom description of primary, secondary, and tertiary alcohols. *J. Phys. Chem. B* **2001**, *105*, 3093–3104.

(208) Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. Development of the transferable potentials for phase equilibria model for hydrogen sulfide. *J. Phys. Chem. B* **2015**, *119*, 7041–7052.

(209) Potoff, J. J.; Siepmann, J. I. Vapor–liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AIChE J.* **2001**, *47*, 1676–1682.

(210) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.

(211) Widom, B. Structure of interfaces from uniformity of the chemical potential. *J. Stat. Phys.* **1978**, *19*, 563–574.

(212) Willemsen, S.; Vlugt, T. J. H.; Hoefsloot, H.; Smit, B Combining dissipative particle dynamics and Monte Carlo techniques. *J. Comput. Phys.* **1998**, *147*, 507–517.

(213) Ferrario, M.; Haughney, M.; McDonald, I. R.; Klein, M. L. Molecular-dynamics simulation of aqueous mixtures: Methanol, acetone, and ammonia. *J. Comput. Phys.* **1990**, *93*, 5156–5166.

(214) Luzar, A.; Chandler, D. Structure and hydrogen bond dynamics of water–dimethyl sulfoxide mixtures by computer simulations. *J. Comput. Phys.* **1993**, *98*, 8160–8173.

(215) Luzar, A.; Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **1996**, *379*, 55.

(216) Jedlovszky, P; Brodholt, J. P.; Bruni, F.; Ricci, M. A.; Soper, A. K.; Vallauri, R. Analysis of the hydrogen-bonded structure of water from ambient to supercritical conditions. *J. Comput. Phys.* **1998**, *108*, 8528–8540.

(217) Luzar, A. Resolving the hydrogen bond dynamics conundrum. *J. Comput. Phys.* **2000**, *113*, 10663–10675.

(218) Bauer, B. A.; Patel, S. Properties of water along the liquid-vapor coexistence curve via molecular dynamics simulations using the polarizable TIP4P-QDP-LJ water model. *J. Comput. Phys.* **2009**, *131*, 084709.

(219) Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Mol. Simul.* **2015**, *42*, 81–101.

(220) Dubbeldam, D.; Torres-Knoop, A.; Walton, K. S. On the inner workings of Monte Carlo codes. *Mol. Simul.* **2013**, *39*, 1253–1292.

(221) Harris, J.; Rice, S. A. A lattice model of a supported monolayer of amphiphile molecules: Monte Carlo simulations. *J. Comput. Phys.* **1988**, *88*, 1298–1306.

(222) Rahbari, A.; Hens, R.; Jamali, S. H.; Ramdin, M.; Dubbeldam, D.; Vlugt, T. J. H. Effect of truncating electrostatic interactions on predicting thermodynamic properties of water–methanol systems. *Mol. Simul.* **2019**, *45*, 336–350.

(223) Shirts, M. R.; Pande, V. S. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* **2005**, *122*, 134508.

(224) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely precise free energy calculations of amino acid side chain analogs: comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.* **2003**, *119*, 5740–5761.

(225) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. In *Annual Reports in Computational Chemistry*, Spellmeyer, D., Wheeler, R., Eds., Chapter 4 Alchemical Free Energy Calculations: Ready for Prime Time?; Elsevier: United States, 2007, pp 41–59.

(226) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the analysis of free energy calculations. *J Comput. Aided. Mol. Des.* **2015**, *29*, 397–411.

(227) Naden, L. N.; Pham, T. T.; Shirts, M. R. Linear basis function approach to efficient alchemical free energy calculations. 1. removal of uncharged atomic sites. *J. Chem. Theory Comput.* **2014**, *10*, 1128–1149.

(228) Naden, L. N.; Shirts, M. R. Linear basis function approach to efficient alchemical free energy calculations. 2. Inserting and deleting particles with Coulombic interactions. *J. Chem. Theory Comput.* **2015**, *11*, 2536–2549.

(229) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Comput. Phys.* **1983**, *79*, 926–935.

(230) Wang, F.; Landau, D. Determining the density of states for classical statistical models: a random walk algorithm to produce a flat histogram. *Phys. Rev. E* **2001**, *64*, 056101.

(231) De Reuck, K.; Craven, R. Methanol, international thermodynamic tables of the fluid state, vol. 12. *IUPAC, Blackwell Scientific Publications, London* **1993**.

(232) Span, R.; Wagner, W. A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa. *J. Phys. Chem. Ref. Data* **1996**, *25*, 1509–1596.

(233) Poursaeidesfahani, A.; Rahbari, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H. Computation of thermodynamic properties in the Continuous Fractional Component Monte Carlo Gibbs ensemble. *Mol. Simul.* **2017**, *43*, 189–195.

(234) Adler, J.; Parmryd, I. Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytom. A* **2010**, *77A*, 733–742.

(235) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(236) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r^{-1} summation. *J. Chem. Phys.* **1999**, *110*, 8254–8282.

(237) Wolf, D. Reconstruction of NaCl surfaces from a dipolar solution to the Madelung problem. *Phys. Rev. Lett.* **1992**, *68*, 3315–3318.

(238) Waibel, C.; Gross, J. Modification of the Wolf method and evaluation for molecular simulation of vapor-liquid equilibria. *J. Chem. Theory Comput.* **2018**, *14*, 2198–2206.

(239) Waibel, C.; Feinler, M. S.; Gross, J. A modified shifted force approach to the Wolf summation. *J. Chem. Theory Comput.* **2019**, *15*, 572–583.

(240) Shi, W.; Maginn, E. J. Atomistic simulation of the absorption of carbon dioxide and water in the ionic liquid 1-n-Hexyl-3-methylimidazolium Bis(trifluoromethylsulfonyl)imide ([hmim][Tf_2N]. *J. Phys. Chem. B* **2008**, *112*, 2045–2055.

(241) Vega, C.; Abascal, J. L. F.; Conde, M. M.; Aragones, J. L. What ice can teach us about water interactions: a critical comparison of the performance of different water models. *Faraday Discuss.* **2009**, *141*, 251–276.

(242) Halgren, T. A.; Damm, W. Polarizable force fields. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236 –242.

(243) Jiang, H.; Moultos, O. A.; Economou, I. G.; Panagiotopoulos, A. Z. Hydrogen-bonding polarizable intermolecular potential model for water. *J. Phys. Chem. B* **2016**, *120*, 12358–12370.

(244) Gladich, I.; Roeselová, M. Comparison of selected polarizable and nonpolarizable water models in molecular dynamics simulations of ice I_h. *Phys. Chem. Chem. Phys.* **2012**, *14*, 11371–11385.

(245) Chen, B.; Xing, J.; Siepmann, J. I. Development of polarizable water force fields for phase equilibrium calculations. *J. Phys. Chem. B* **2000**, *104*, 2391–2401.

(246) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, 1–17.

(247) Kunz, A.-P. E.; van Gunsteren, W. F. Development of a nonlinear classical polarization model for liquid water and aqueous solutions: COS/D. *J. Phys. Chem. A* **2009**, *113*, 11570–11579.

(248) Lamoureux, G.; MacKerell, A. D.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

(249) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245 –249.

(250) Ren, P.; Ponder, J. W. Polarizable atomic multipole water model for molecular mechanics simulation. *J. Phys. Chem. B* **2003**, *107*, 5933–5947.

(251) Laury, M. L.; Wang, L.-P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised parameters for the AMOEBA polarizable atomic multipole water model. *J. Phys. Chem. B* **2015**, *119*, 9423–9437.

(252) Becker, T. M.; Dubbeldam, D.; Lin, L.-C.; Vlugt, T. J. H. Investigating polarization effects of CO_2 adsorption in MgMOF-74. *J. Comput. Sci.* **2016**, *15*, 86–94.

(253) Antila, H. S.; Salonen, E. In *Biomolecular simulations: methods and protocols*, Monticelli, L., Salonen, E., Eds.; Humana Press: Totowa, NJ, 2013; Chapter 9, pp 215–241.

(254) Rick, S. W.; Stuart, S. J. In *Reviews in Computational Chemistry*, Lipkowitz, K. B., Salonen Boyd, D. B., Eds.; Wiley-VCH: Hoboken, New Jersey ,United States, 2002; Vol. 18; Chapter 3, pp 89–146.

(255) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(256) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: a different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.

(257) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic improvement of a classical molecular model of water. *J. Phys. Chem. B* **2013**, *117*, 9956–9972.

(258) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(259) Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.

(260) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(261) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-EW. *J. Chem. Phys.* **2004**, *120*, 9665–9678.

(262) Gonzalez-Salgado, D.; Vega, C. A new intermolecular potential for simulations of methanol: The OPLS/2016 model. *J. Chem. Phys.* **2016**, *145*, 034508.

(263) Fanourgakis, G. S. An extension of Wolf's method for the treatment of electrostatic interactions: application to liquid water and aqueous solutions. *J. Phys. Chem. B* **2015**, *119*, 1974–1985.

(264) Fukuda, I.; Nakamura, H. Non-Ewald methods: theory and applications to molecular systems. *Biophys. Rev.* **2012**, *4*, 161–170.

(265) Vlugt, T. J. H.; García-Pérez, E.; Dubbeldam, D.; Ban, S.; Calero, S. Computing the heat of adsorption using molecular simulations: the effect of strong Coulombic interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1107–1118.

(266) Mendoza, F. N.; López-Lemus, J.; Chapela, G. A.; Alejandre, J. The Wolf method applied to the liquid-vapor interface of water. *J. Chem. Phys.* **2008**, *129*, 024706.

(267) McCann, B. W.; Acevedo, O. Pairwise alternatives to Ewald summation for calculating long-range electrostatics in ionic liquids. *J. Chem. Theory Comput.* **2013**, *9*, 944–950.

(268) Cerutti, D. S.; Duke, R. E.; Darden, T. A.; Lybrand, T. P. Staggered Mesh Ewald: an extension of the Smooth Particle-Mesh Ewald method adding great versatility. *J. Chem. Theory Comput.* **2009**, *5*, 2322–2338.

(269) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(270) Neelov, A.; Holm, C. Interlaced P3M algorithm with analytical and ik-differentiation. *J. Chem. Phys.* **2010**, *132*, 234103.

(271) Isele-Holder, R. E.; Mitchell, W.; Hammond, J. R.; Kohlmeyer, A.; Ismail, A. E. Reconsidering dispersion potentials: reduced cutoffs in mesh-based Ewald solvers can be faster than truncation. *J. Chem. Theory Comput.* **2013**, *9*, 5412–5420.

(272) Dubbeldam, D.; Walton, K. S.; Vlugt, T. J. H.; Calero, S. Design, parameterization, and implementation of atomic force fields for adsorption in nanoporous materials. *Adv. Theory Simul.* **2019**, *2*, 1900135.

(273) Saito, M. Molecular dynamics simulations of proteins in solution: artifacts caused by the cutoff approximation. *J. Chem. Phys.* **1994**, *101*, 4055–4061.

(274) Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.* **1994**, *15*, 667–683.

(275) Hansen, J. S.; Schrøder, T. B.; Dyre, J. C. Simplistic coulomb forces in molecular dynamics: comparing the Wolf and Shifted-Force approximations. *J. Phys. Chem. B* **2012**, *116*, 5738–5743.

(276) Levitt, M.; Hirshberg, M.; Sharon, R.; Daggett, V. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Computer Physics Communications* **1995**, *91*, 215 –231.

(277) Evjen, H. M. On the stability of certain heteropolar crystals. *Phys. Rev.* **1932**, *39*, 675–687.

(278) Pekka, M.; Lennart, N. Structure and dynamics of liquid water with different long-range interaction truncation and temperature control methods in molecular dynamics simulations. *J. Comput. Chem.* **2002**, *23*, 1211–1219.

(279) Patra, M.; Karttunen, M.; Hyvönen, M.; Falck, E.; Lindqvist, P.; Vattulainen, I. Molecular dynamics simulations of lipid bilayers: major artifacts due to truncating electrostatic interactions. *Biophys. J.* **2003**, *84*, 3636 –3645.

(280) van der Spoel, D.; van Maaren, P. J. The origin of layer structure artifacts in simulations of liquid water. *J. Chem. Theory Comput.* **2006**, *2*, 1–11.

(281) van der Spoel, D.; van Maaren, P. J.; Berendsen, H. J. C. A systematic study of water models for molecular simulation: derivation of water models optimized for use with a reaction field. *J. Chem. Phys.* **1998**, *108*, 10220–10230.

(282) Ghobadi, A. F.; Taghikhani, V.; Elliott, J. R. Investigation on the solubility of SO_2 and CO_2 in imidazolium-based ionic liquids using NPT Monte Carlo simulation. *J. Phys. Chem. B* **2011**, *115*, 13599–13607.

(283) Mahadevan, T. S.; Garofalini, S. H. Dissociative water potential for molecular dynamics simulations. *J. Phys. Chem. B* **2007**, *111*, 8919–8927.

(284) Carlos, A.; Alejandro, G.-V. Monte Carlo simulations of primitive models for ionic systems using the Wolf method. *Mol. Phys.* **2006**, *104*, 1475–1486.

(285) Smith, W; Forester, T.; Todorov, I.; Leslie, M The DLpoly 2 user manual. *CCLRC, Daresbury Laboratory, Daresbury, Warrington WA4 4AD, England* **2001**, *2*, Available at: `ftp://ftp.dl.ac.uk/ccp5/DL_POLY/DL_POLY_CLASSIC/DOCUMENTS/USRMAN2.19.pdf`.

(286) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Amer. Chem. Soc.* **1995**, *117*, 5179–5197.

(287) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(288) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(289) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Amer. Chem. Soc.* **1996**, *118*, 11225–11236.

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

(291) Shi, W.; Myers, C. R.; Luebke, D. R.; Steckel, J. A.; Sorescu, D. C. Theoretical and experimental studies of CO_2 and H_2 separation

using the 1-ethyl-3-methylimidazolium acetate ([emim][CH_3COO]) ionic liquid. *J. Phys. Chem. B* **2012**, *116*, 283–295.

(292) Shi, W.; Sorescu, D. C. Molecular simulations of CO_2 and H_2 sorption into ionic liquid 1-n-hexyl-3-methylimidazolium bis(trifluoromethylsulfonyl)amide ([hmim][Tf_2N]) confined in carbon nanotubes. *J. Phys. Chem. B* **2010**, *114*, 15029–15041.

(293) Balaji, S. P.; Schnell, S. K.; McGarrity, E. S.; Vlugt, T. J. H. A direct method for calculating thermodynamic factors for liquid mixtures using the Permuted Widom test particle insertion method. *Mol. Phys.* **2013**, *111*, 287–296.

(294) Hempel, S.; Fischer, J.; Paschek, D.; Sadowski, G. Activity coefficients of complex molecules by molecular simulation and Gibbs-Duhem integration. *Soft Mater.* **2012**, *10*, 26–41.

(295) Goodwin, R. D. Methanol thermodynamic properties from 176 to 673 K at pressures to 700 bar. *J. Phys. Chem. Ref. Data* **1987**, *16*, 799–892.

(296) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(297) Lemmon, E. W.; Huber, M. L.; McLinden, M. O. NIST reference fluid thermodynamic and transport properties–REFPROP. *NIST standard reference database* **2002**, *23*, `https://www.nist.gov/srd/refprop`, v7.

(298) Schnell, S. K.; Englebienne, P.; Simon, J.-M.; Krüger, P.; Balaji, S. P.; Kjelstrup, S.; Bedeaux, D.; Bardow, A.; Vlugt, T. J. H. How to apply the Kirkwood–Buff theory to individual species in salt solutions. *Chem. Phys. Lett.* **2013**, *582*, 154–157.

(299) Alberty, R. A., *Thermodynamics of biochemical reactions*, 1st ed.; John Wiley & Sons: New Jersey, USA, 2005.

(300) Engel, T.; Reid, P. J., *Thermodynamics, statistical thermodynamics, & kinetics*, 3rd ed.; Prentice Hall Upper Saddle River: Upper Saddle River, NJ, USA, 2010.

(301) Podgoršek, A; Jacquemin, J; Pádua, A.; Costa Gomes, M. Mixing enthalpy for binary mixtures containing ionic liquids. *Chem. Rev.* **2016**, *116*, 6075–6106.

(302) Almantariotis, D.; Fandiño, O.; Coxam, J.-Y.; Gomes, M. C. Direct measurement of the heat of solution and solubility of carbon dioxide in 1-hexyl-3-methylimidazolium bis[trifluoromethylsulfonyl]amide and 1-octyl-3-methylimidazolium bis[trifluoromethylsulfonyl]amide. *Int. J. Greenhouse Gas Control* **2012**, *10*, 329–340.

(303) Shiflett, M. B.; Elliott, B. A.; Lustig, S. R.; Sabesan, S.; Kelkar, M. S.; Yokozeki, A. Phase behavior of CO_2 in room-temperature ionic liquid 1-ethyl-3-ethylimidazolium acetate. *ChemPhysChem* **2012**, *13*, 1806–1817.

(304) Firaha, D. S.; Hollóczki, O.; Kirchner, B. Computer-aided design of ionic liquids as CO_2 Absorbents. *Angew. Chem. Int. Ed.* **2015**, *54*, 7805–7809.

(305) Zhang, B.; van Duin, A. C. T.; Johnson, J. K. Development of a ReaxFF reactive force field for tetrabutylphosphonium glycinate/CO_2 Mixtures. *J. Phys. Chem. B* **2014**, *118*, 12008–12016.

(306) Šerbanović, S. P.; Djordjević, B. D.; Grozdanić, D. K. Excess molar volume prediction for some hydrocarbons and related mixtures by means of simple cubic equations of state. *Fluid Phase Equilib.* **1990**, *57*, 47–65.

(307) Abdulagatov, I. M.; Bazaev, A. R.; Bazaev, E. A.; Saidakhmedova, M. B.; Ramazanova, A. E. Volumetric properties of near-critical and supercritical water + pentane mixtures: molar, excess, partial, and apparent volumes. *J. Chem. Eng. Data* **1998**, *43*, 451–458.

(308) Abdulagatov, I. M.; Bazaev, A. R.; Bazaev, E. A.; Khokhlachev, S. P.; Saidakhmedova, M. B.; Ramazanova, A. E. Excess, partial, and molar volumes of n-alkanes in near-critical and supercritical water. *J. Solution Chem.* **1998**, *27*, 731–753.

(309) Liu, H.; O'Connell, J. P. On the measurement of solute partial molar volumes in near-critical fluids with supercritical fluid chromatography. *Ind. Eng. Chem. Res.* **1998**, *37*, 3323–3330.

(310) Chang, R.; Sengers, J. L. Behavior of dilute mixtures near the solvent's critical point. *J. Phys. Chem.* **1986**, *90*, 5921–5927.

(311) Schnell, S. K.; Skorpa, R.; Bedeaux, D.; Kjelstrup, S.; Vlugt, T. J. H.; Simon, J.-M. Partial molar enthalpies and reaction enthalpies from equilibrium molecular dynamics simulation. *J. Chem. Phys.* **2014**, *141*, 144501.

(312) Walas, S. M., *Phase Equilibria in Chemical Engineering*; Butterworth-Heinemann: USA, 1985.

(313) Kirkwood, J. G.; Buff, F. P. The statistical mechanical theory of solutions. I. *J. Chem. Phys.* **1951**, *19*, 774–777.

(314) Krüger, P.; Schnell, S. K.; Bedeaux, D.; Kjelstrup, S.; Vlugt, T. J. H.; Simon, J.-M. Kirkwood–Buff integrals for finite volumes. *J. Phys. Chem. Lett.* **2012**, *4*, 235–238.

(315) Dawass, N.; Krüger, P.; Simon, J.-M.; Vlugt, T. J. H. Kirkwood–Buff integrals of finite systems: shape effects. *Mol. Phys.* **2018**, *116*, 1573–1580.

(316) Dawass, N.; Krüger, P.; Schnell, S. K.; Simon, J.-M.; Vlugt, T. J. H. Kirkwood-Buff integrals from molecular simulation. *Fluid Phase Equilib.* **2019**, *486*, 21 –36.

(317) Krüger, P.; Vlugt, T. J. H. Size and shape dependence of finite-volume Kirkwood-Buff integrals. *Phys. Rev. E* **2018**, *97*, 051301–5.

(318) Ben-Naim, A. Y., *Statistical thermodynamics for chemists and biochemists*, 1st ed.; Springer Science & Business Media: New York, USA, 1992.

(319) Kofke, D. A.; Cummings, P. T. Quantitative comparison and optimization of methods for evaluating the chemical potential by molecular simulation. *Mol. Phys.* **1997**, *92*, 973–996.

(320) Nezbeda, I.; Kolafa, J. A new version of the insertion particle method for determining the chemical Potential by Monte Carlo simulation. *Mol. Simul.* **1991**, *5*, 391–403.

(321) Johnson, J. K.; Panagiotopoulos, A. Z.; Gubbins, K. E. Reactive canonical Monte Carlo: a new simulation technique for reacting or associating fluids. *Mol. Phys.* **1994**, *81*, 717–733.

(322) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

(323) Rahbari, A.; Hens, R.; Moultos, O. A.; Dubbeldam, D.; Vlugt, T. J. H. Multiple free energy calculations from single state point Continuous Fractional Component Monte Carlo simulation using umbrella sampling. *J. Chem. Theory Comput.* in press, DOI: 10.1021/acs.jctc.9b01097.

(324) Erisman, J. W.; Sutton, M. A.; Galloway, J.; Klimont, Z.; Winiwarter, W. How a century of ammonia synthesis changed the world. *Nature Geosci.* **2008**, *1*, 636–639.

(325) Graaf, G. H.; Winkelman, J. G. Chemical equilibria in methanol synthesis including the water–gas shift reaction: a critical reassessment. *Ind. Eng. Chem. Res.* **2016**, *55*, 5854–5864.

(326) Graaf, G.; Sijtsema, P.; Stamhuis, E.; Joosten, G. Chemical equilibria in methanol synthesis. *Chem. Eng. Sci.* **1986**, *41*, 2883–2890.

(327) Diamantonis, N. I.; Boulougouris, G. C.; Mansoor, E.; Tsangaris, D. M.; Economou, I. G. Evaluation of cubic, SAFT, and PC-SAFT equations of state for the vapor-liquid equilibrium modeling of CO_2 mixtures with other gases. *Ind. Eng. Chem. Res.* **2013**, *52*, 3933–3942.

(328) Turner, H. C.; Brennan, J. K.; Lisal, M.; Smith, W. R.; Johnson, K. J.; Gubbins, K. E. Simulation of chemical reaction equilibria by the reaction ensemble Monte Carlo method: a review. *Mol. Simul.* **2008**, *34*, 119–146.

(329) Smith, W. R.; Lísal, M. Direct Monte Carlo simulation methods for nonreacting and reacting systems at fixed total internal energy or enthalpy. *Phys. Rev. E* **2002**, *66*, 011104.

(330) Lin, C.-T.; Daubert, T. E. Estimation of partial molar volume and fugacity coefficient of components in mixtures from the soave and Peng-Robinson equations of state. *Ind. Eng. Chem. Process Des. Dev.* **1980**, *19*, 51–59.

(331) Barker, J. A.; Henderson, D. Perturbation theory and equation of state for fluids: the square-well potential. *J. Chem. Phys.* **1967**, *47*, 2856–2861.

(332) Gross, J.; Sadowski, G. Application of the perturbed-chain SAFT equation of state to associating systems. *Ind. Eng. Chem. Res.* **2002**, *41*, 5510–5515.

(333) Mejbri, K.; Bellagi, A. Modelling of the thermodynamic properties of the water–ammonia mixture by three different approaches. *Int. J. Ref.* **2006**, *29*, 211–218.

(334) Chase, M. W.; Curnutt, J.; Prophet, H; McDonald, R.; Syverud, A. JANAF thermochemical tables, 1975 supplement. *J. Phys. Chem. Ref. Data.* **1975**, *4*, 1–176.

(335) Glotzer, S. C.; Stauffer, D.; Jan, N. Monte Carlo simulations of phase separation in chemically reactive binary mixtures. *Phys. Rev. Lett.* **1994**, *72*, 4109.

(336) Smith, W. R.; Triska, B The reaction ensemble method for the computer simulation of chemical and phase equilibria. I. Theory and basic examples. *J. Comput. Phys.* **1994**, *100*, 3019–3027.

(337) Hansen, N.; Jakobtorweihen, S.; Keil, F. J. Reactive Monte Carlo and grand-canonical Monte Carlo simulations of the propene metathesis reaction system. *J. Comput. Phys.* **2005**, *122*, 164705.

(338) Ungerer, P.; Collell, J.; Yiannourakou, M. Molecular modeling of the volumetric and thermodynamic properties of kerogen: influence of organic type and maturity. *Energy & Fuels* **2014**, *29*, 91–105.

(339) Ghahremanpour, M. M.; van Maaren, P. J.; Ditz, J. C.; Lindh, R.; van der Spoel, D. Large-scale calculations of gas phase thermochemistry: enthalpy of formation, standard entropy, and heat capacity. *J. Comput. Phys.* **2016**, *145*, 114305–12.

(340) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 20, August 2019, Editor: Russell D. Johnson III, `http://cccbdb.nist.gov`.

(341) Gillespie, L. J.; Beattie, J. A. The thermodynamic treatment of chemical equilibria in systems composed of real gases. I. An approximate equation for the mass action function applied to the existing data on the Haber equilibrium. *Phys. Rev.* **1930**, *36*, 743.

(342) Abudour, A. M.; Mohammad, S. A.; Jr., R. L. R.; Gasem, K. A. Generalized binary interaction parameters for the Peng-Robinson equation of state. *Fluid Phase Equilib.* **2014**, *383*, 156–173.

(343) Aspen Plus V8.8 Database, Aspen Technology Inc, Bedford, United States, `https://www.aspentech.com`.

(344) Abudour, A. M.; Mohammad, S. A.; Jr., R. L. R.; Gasem, K. A. Predicting PR EOS binary interaction parameter using readily available molecular properties. *Fluid Phase Equilib.* **2017**, *434*, 130–140.

(345) Twu, C. H.; Coon, J. E.; Bluck, D. Comparison of the Peng-Robinson and Soave-Redlich-Kwong equations of state using a new zero-pressure-based mixing rule for the prediction of high-pressure and high-temperature phase equilibria. *Ind. Eng. Chem. Res.* **1998**, *37*, 1580–1585.

(346) Wei, Y. S.; Sadus, R. J. Equations of state for the calculation of fluid-phase equilibria. *AIChE J.* **2000**, *46*, 169–196.

(347) Meragawi, S. E.; Diamantonis, N. I.; Tsimpanogiannis, I. N.; Economou, I. G. Hydrate - fluid phase equilibria modeling using PC-SAFT and Peng-Robinson equations of state. *Fluid Phase Equilib.* **2016**, *413*, 209–219.

(348) Walpole, R. E.; Myers, R. H.; Myers, S. L.; Ye, K., *Probability & statistics for engineers & scientists*, 9th ed.; Prentice Hall: Boston, USA, 2012.

(349) Torres-Knoop, A.; Poursaeidesfahani, A.; Vlugt, T. J. H.; Dubbeldam, D. Behavior of the enthalpy of adsorption in nanoporous materials close to saturation conditions. *J. Chem. Theory Comput.* **2017**, *13*, 3326–3339.

(350) Karavias, F.; Myers, A. L. Isosteric heats of multicomponent adsorption: thermodynamics and computer simulations. *Langmuir* **1991**, *7*, 3118–3126.

(351) Van den Bergh, J.; Ban, S.; Vlugt, T. J. H.; Kapteijn, F. Modeling the loading dependency of diffusion in zeolites: the relevant site model extended to mixtures in DDR-type zeolite. *J. Phys. Chem. C*, *113*, 21856–21865.

(352) Walpole, R. E.; Myers, R. H.; Myers, S. L.; Ye, K., *Probability & statistics for engineers & scientists*, 9th ed.; Prentice Hall: Boston, USA, 2012.

(353) Chang, T.; Rousseau, R. W.; Kilpatrick, P. K. Methanol synthesis reactions: calculations of equilibrium conversions using equations of state. *Ind. Eng. Chem. Process Des. Dev.* **1986**, *25*, 477–481.

(354) Poursaeidesfahani, A.; Torres-Knoop, A.; Rigutto, M.; Nair, N.; Dubbeldam, D.; Vlugt, T. J. H. Computation of the Heat and Entropy of Adsorption in Proximity of Inflection Points. *J. Phys. Chem. C* **2016**, *120*, 1727–1738.

(355) Vlugt, T. J. H.; Krishna, R.; Smit, B. Molecular Simulations of Adsorption Isotherms for Linear and Branched Alkanes and Their Mixtures in Silicalite. *J. Phys. Chem. B* **1999**, *103*, 1102–1118.

(356) Vlugt, T. J. H.; Zhu, W; Kapteijn, F; Moulijn, J. A.; Smit, B; Krishna, R Adsorption of linear and branched alkanes in the zeolite silicalite-1. *J. Amer. Chem. Soc.* **1998**, *120*, 5599–5600.

(357) Krishna, R.; Smit, B.; Vlugt, T. J. H. Sorption-Induced Diffusion-Selective Separation of Hydrocarbon Isomers Using Silicalite. *J. Phys. Chem. A* **1998**, *102*, 7727–7730.

(358) Dubbeldam, D.; Calero, S.; Vlugt, T. J. H.; Krishna, R.; Maesen, T. L. M.; Smit, B. United atom force field for alkanes in nanoporous materials. *J. Phys. Chem. B* **2004**, *108*, 12301–12313.

(359) Panagiotopoulos, A. Z.; Reid, R. C. On the relationship between pairwise fluctuations and thermodynamic derivatives. *J. Chem. Phys.* **1986**, *85*, 4650–4653.

(360) Ray, J. R. Ensembles and Computer Simulation Calculation of Response Functions. *Handbook of Materials Modeling* **2005**, 729–743.

(361) Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges. *Theor. Chem. Acc.* **2012**, *131*, 1079.

(362) World population prospects: the 2017 revision, key findings and advance tables, ESA/P/WP/248., `https://population.un.org/wpp/Publications/Files/WPP2017_KeyFindings.pdf`, New York, 2017.

(363) *United Nations. (2015) General Assembly Resolution A/RES/70/1. Transforming our world: the 2030 agenda for sustainable development*; tech. rep. A/RES/70/1, `http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E`; New York: United Nations, General Assembly, 2015.

(364) *World Energy Outlook 2018*; tech. rep., `www.iea.org/weo`; International Energy Agency, 2018.

(365) *Adoption of the Paris Agreement,* FCCC/CP/2015/L.9/Rev.1; tech. rep. FCCC/CP/2015/L.9/Rev.1, `https://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf`; Paris: United Nations, Framework Convention on Climate Change, 2015.

(366) Gallo, P.; Amann-Winkel, K.; Angell, C. A.; Anisimov, M. A.; Caupin, F.; Chakravarty, C.; Lascaris, E.; Loerting, T.; Panagiotopoulos, A. Z.; Russo, J.; Sellberg, J. A.; Stanley, H. E.; Tanaka, H.; Vega, C.; Xu, L.; Pettersson, L. G. M. Water: a tale of two liquids. *Chem. Rev.* **2016**, *116*, 7463–7500.

(367) Zakeri, B.; Syri, S. Electrical energy storage systems: a comparative life cycle cost analysis. *Renew. Sust. Energy Rev.* **2015**, *42*, 569 –596.

(368) Evans, A.; Strezov, V.; Evans, T. J. Assessment of utility energy storage options for increased renewable energy penetration. *Renew. Sust. Energy Rev.* **2012**, *16*, 4141 –4147.

(369) Mahlia, T.; Saktisahdan, T.; Jannifar, A.; Hasan, M.; Matseelar, H. A review of available methods and development on energy storage; technology update. *Renew. Sust. Energy Rev.* **2014**, *33*, 532 –545.

(370) Chen, H.; Cong, T. N.; Yang, W.; Tan, C.; Li, Y.; Ding, Y. Progress in electrical energy storage system: a critical review. *Progress in Natural Science* **2009**, *19*, 291–312.

(371) Johnston, B.; Mayo, M. C.; Khare, A. Hydrogen: the energy source for the 21st century. *Technovation* **2005**, *25*, 569 –585.

(372) Rahbari, A.; Ramdin, M.; van den Broeke, L. J. P.; Vlugt, T. J. H. Combined steam reforming of methane and formic acid to produce syngas with an adjustable H_2:CO ratio. *Ind. Eng. Chem. Res.* **2018**, *57*, 10663–10674.

(373) Adolf, J.; Balzer, C.; Louis, J. *Energy of the future? sustainable mobility through fuel cells and H_2*; tech. rep.; Hamburg: Shell Deutschland Oil GmbH, Wuppertal Institut, 2017.

(374) Bampaou, M.; Panopoulos, K. D.; Papadopoulos, A. I.; Seferlis, P.; Voutetakis, S. An electrochemical hydrogen compression model. *Chem. Eng. Trans.* **2018**, *70*, 1213–1218.

(375) Ströbel, R.; Oszcipok, M.; Fasil, M.; Rohland, B.; Jörissen, L.; Garche, J. The compression of hydrogen in an electrochemical cell based on a PE fuel cell design. *J. Power Sources* **2002**, *105*, 208 –215.

(376) Suermann, M.; Kiupel, T.; Schmidt, T. J.; Büchi, F. N. Electrochemical hydrogen compression: efficient pressurization concept derived from an energetic evaluation. *J. Electrochem. Soc.* **2017**, *164*, F1187–F1195.

(377) Rohland, B.; Eberle, K.; Ströbel, R.; Scholta, J.; Garche, J. Electrochemical hydrogen compressor. *Electrochim. Acta.* **1998**, *43*, 3841 –3846.

(378) Nordio, M.; Rizzi, F.; Manzolini, G.; Mulder, M.; Raymakers, L.; Van Sint Annaland, M.; Gallucci, F. Experimental and modelling study of an electrochemical hydrogen compressor. *Chem. Eng. J.* **2019**, *369*, 432–442.

(379) Casati, C.; Longhi, P.; Zanderighi, L.; Bianchi, F. Some fundamental aspects in electrochemical hydrogen purification/compression. *J. Power Sources* **2008**, *180*, 103–113.

(380) Wiebe, R.; Gaddy, V. L. The solubility of hydrogen in water at 0, 50, 75 and 100 C° from 25 to 1000 atmospheres. **2005**, *56*, 76–79.

(381) Iwai, Y.; Margerum, M. R.; Lu, B. C.-Y. A new three-parameter cubic equation of state for polar fluids and fluid mixtures. *Fluid Phase Equilib.* **1988**, *42*, 21 –41.

(382) Kwak, T.; Mansoori, G. A. van der Waals mixing rules for cubic equations of state. applications for supercritical fluid extraction modelling. *Chem. Eng. Sci.* **1986**, *41*, 1303–1309.

(383) Sun, R.; Lai, S.; Dubessy, J. Calculations of vapor-liquid equilibria of the H_2O-N_2 and H_2O-H_2 systems with improved SAFT-LJ EOS. *Fluid Phase Equilib.* **2015**, *390*, 23 –33.

(384) Meyer, M.; Tebbe, U.; Piiper, J. Solubility of inert gases in dog blood and skeletal muscle. *Pflügers Archiv European Journal of Physiology* **1980**, *384*, 131–134.

(385) Gillespie, P. C.; Wilson, G. M. *Vapor-liquid equilibrium data on water-substitue gas components*; tech. rep.; Provo, Utah, 1980, RR_41, 1–34.

(386) Kling, G; Maurer, G The solubility of hydrogen in water and in 2-aminoethanol at temperatures between 323 K and 423 K and pressures up to 16 M Pa. *Chem. Thermodynamics* **1991**, *23*, 531–541.

(387) Devaney, W; Berryman, J. M.; Kao, P.-L.; Eakin, B *High temperature V-L-E measurements for substitute gas components*; tech. rep.; Houston, 1978, pp 1–27.

(388) Jung, J. Loslichkeit von Kohlenmonoxid und Wasserstoff in Wasser zwisen 0 C und 300 C., Ph.D. Thesis, RWTH Aachen, 1962.

(389) Ipatev, V.; Teodorovich, V. Equilibrium compositions of vapor-gas mixtures over solutions. *Zh.Obshch.Khim.* **1934**, *4*, 395–399.

(390) Ugrozov, V. V. Equilibrium compositions of vapor-gas mixtures over solutions. **1996**, *70*, 1240–1241.

(391) Maslennikova, V. Y.; Goryunova, N.; Subbotina, L.; Tsiklis, D. The solubility of water in compressed hydrogen. *Russian Journal of Physical Chemistry* **1976**, *50*, 240–243.

(392) Buch, V. Path integral simulations of mixed para-D_2 and ortho-D_2 clusters: the orientational effects. *J. Chem. Phys.* **1994**, *100*, 7610–7629.

(393) Hirschfelder, C.; Curtiss, F.; Bird, R. B., *Molecular theory of gases and liquids*; Wiley: New York, 1954.

(394) Köster, A.; Thol, M.; Vrabec, J. Molecular models for the hydrogen age: hydrogen, nitrogen, oxygen, argon, and water. *J. Chem. Eng. Data* **2018**, *63*, 305–320.

(395) Cracknell, R. F. Molecular simulation of hydrogen adsorption in graphitic nanofibres. *Phys. Chem. Chem. Phys.* **2001**, *3*, 2091–2097.

(396) Belof, J. L.; Stern, A. C.; Space, B. An accurate and transferable intermolecular diatomic hydrogen potential for condensed phase simulation. *J. Chem. Theory Comput.* **2008**, *4*, 1332–1337.

(397) Forrest, K. A.; Pham, T.; McLaughlin, K.; Belof, J. L.; Stern, A. C.; Zaworotko, M. J.; Space, B. Simulation of the mechanism of gas sorption in a metal-organic framework with open metal sites: molecular hydrogen in PCN-61. *J. Phys. Chem. C* **2012**, *116*, 15538–15549.

(398) Marx, D.; Nielaba, P. Path-integral Monte Carlo techniques for rotational motion in two dimensions: quenched, annealed, and no-spin quantum-statistical averages. *Phys. Rev. A* **1992**, *45*, 8968–8971.

(399) Camp, J.; Stavila, V.; Allendorf, M. D.; Prendergast, D.; Haranczyk, M. Critical factors in computational characterization of hydrogen storage in metal-organic frameworks. *J. Phys. Chem. C* **2018**, *122*, 18957–18967.

(400) Yang, Q.; Zhong, C. Molecular simulation of carbon dioxide/methane/hydrogen mixture adsorption in metal-organic frameworks. *J. Phys. Chem. B* **2006**, *110*, 17776–17783.

(401) Darkrim, F.; Levesque, D. Monte Carlo simulations of hydrogen adsorption in single-walled carbon nanotubes. *J. Chem. Phys.* **1998**, *109*, 4981–4984.

(402) Vega, C. Water: one molecule, two surfaces, one mistake. *Mol. Phys.* **2015**, *113*, 1145–1163.

(403) Soave, G. Equilibrium constants from a modified Redlich-Kwong equation of state. *Chem. Eng. Sci.* **1972**, *27*, 1197 –1203.

(404) Péneloux, A.; Rauzy, E.; Fréze, R. A consistent correction for Redlich-Kwong-Soave volumes. *Fluid Phase Equilib.* **1982**, *8*, 7 –23.

(405) Leachman, J. W.; Jacobsen, R. T.; Penoncello, S.; Lemmon, E. W. Fundamental equations of state for parahydrogen, normal hydrogen, and orthohydrogen. *J. Phys.: Chem. Ref. Data.* **2009**, *38*, 721–748.

(406) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(407) Wilhelm, E.; Waghorne, E.; Hefter, G.; Hummel, W.; Maurer, G; Rebelo, L. P. N.; da Ponte, M. N.; Battino, R.; Clever, L.; van Hook, A.; Domanska-Zelazna, U; Tomkins, R. P. T.; Richon, D.; de Stafani, V.; Coquelet, C; Costa Gomes, M.; Siepmann, J. I.; Anderson, K. E.; Eckert, F.; Grolier, J.-P.; Boyer, S; Salminen, J.; Dohrn, R.; Leiberich, R.; Fele Zilnik, L.; Fages, J.; Macedo, M. E.; Gmehling, J.; Brennecke, J.; Cordes, W.; Prausnitz, J.; Goodwin, A. R.; Marsh, K.; Peters, C. J.; Voigt, W.; Koenigsberger, E.; May, P.; Kamps, A. P.-S.; Shariati, A; Raeissi, S; Padua, A. A. H.; Sauceau, M; Pinho, S. P. P.; Kaskiala, T.; Kobylin, P., *Developments and applications in solubility*, 1st ed.; Letcher, T. M., Ed.; The Royal Society of Chemistry: 2007.

(408) Buttler, A.; Spliethoff, H. Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: a review. *Renew. Sust. Energy Rev.* **2018**, *82*, 2440 –2454.

(409) Riese, J.; Grünewald, M.; Lier, S. Utilization of renewably generated power in the chemical process industry. *Energy Sust. Soc.* **2014**, *4*, 18.

(410) Schiffer, Z. J.; Manthiram, K. Electrification and decarbonization of the chemical industry. *Joule* **2017**, *1*, 10 –14.

(411) Goeppert, A.; Czaun, M.; Jones, J.-P.; Surya Prakash, G. K.; Olah, G. A. Recycling of carbon dioxide to methanol and derived products - closing the loop. *Chem. Soc. Rev.* **2014**, *43*, 7995–8048.

(412) Centi, G.; Perathoner, S. Opportunities and prospects in the chemical recycling of carbon dioxide to fuels. *Catal. Today* **2009**, *148*, 191–205.

(413) Centi, G.; Quadrelli, E. A.; Perathoner, S. Catalysis for CO_2 conversion: a key technology for rapid introduction of renewable energy in the value chain of chemical industries. *Energy Environ. Sci.* **2013**, *6*, 1711–1731.

(414) Kortlever, R.; Peters, I.; Koper, S.; Koper, M. T. M. Electrochemical CO_2 reduction to formic acid at low overpotential and with high faradaic efficiency on carbon-supported bimetallic Pd-Pt nanoparticles. *ACS Catal.* **2015**, *5*, 3916–3923.

(415) Natsui, K.; Iwakawa, H.; Ikemiya, N.; Nakata, K.; Einaga, Y. Stable and highly efficient electrochemical production of formic acid from carbon dioxide using diamond electrodes. *Angew. Chem. Int. Ed.* **2018**, *57*, 2639–2643.

(416) Lee, S.; Ju, H.; Machunda, R.; Uhm, S.; Lee, J. K.; Lee, H. J.; Lee, J. Sustainable production of formic acid by electrolytic reduction of gaseous carbon dioxide. *J. Mater. Chem. A* **2015**, *3*, 3029–3034.

(417) Chaplin, R.; Wragg, A. Effects of process conditions and electrode material on reaction pathways for carbon dioxide electroreduction with particular reference to formate formation. *J. Appl. Electrochem.* **2003**, *33*, 1107–1123.

(418) Xu, L.; Leung Dennis, Y. C.; Huizhi, W.; Leung Michael, K. H.; Jin, X. Electrochemical reduction of carbon dioxide to formic acid. *ChemElectroChem* **2014**, *1*, 836–849.

(419) Chen, A.; Lin, B.-L. A simple framework for quantifying electrochemical CO_2 fixation. *Joule* **2018**, *2*, 594–606.

(420) Thorson, M. R.; Siil, K. I.; Kenis, P. J. Effect of cations on the electrochemical conversion of CO_2 to CO. *Journal of the Electrochemical Society* **2013**, *160*, F69–F74.

(421) Jouny, M.; Luc, W.; Jiao, F. General techno-economic analysis of CO_2 electrolysis systems. *Ind. Eng. Chem. Res.* **2018**, *57*, 2165–2177.

(422) Steinberg, M.; Cheng, H. C. Modern and prospective technologies for hydrogen production from fossil fuels. *Int. J. Hydrogen Energy* **1989**, *14*, 797 –820.

(423) Moret, S.; Dyson, P. J.; Laurenczy, G. Direct synthesis of formic acid from carbon dioxide by hydrogenation in acidic media. *Nature Commun.* **2014**, *5*, 4017.

(424) Tedsree, K.; Li, T.; Jones, S.; Chan, C. W. A.; Yu, K. M. K.; Bagot, P. A.; Marquis, E. A.; Smith, G. D.; Tsang, S. C. E. Hydrogen production from formic acid decomposition at room temperature using a Ag-Pd core-shell nanocatalyst. *Nature Nanotech.* **2011**, *6*, 302.

(425) Xing, R.; Qi, W.; Huber, G. W. Production of furfural and carboxylic acids from waste aqueous hemicellulose solutions from the pulp and paper and cellulosic ethanol industries. *Energy Environ. Sci.* **2011**, *4*, 2193–2205.

(426) Jin, F.; Enomoto, H. Rapid and highly selective conversion of biomass into value-added products in hydrothermal conditions: chemistry of acid/base-catalysed and oxidation reactions. *Energy Environ. Sci.* **2011**, *4*, 382–397.

(427) Kamm, B.; Gruber, P. R.; Kamm, M., *Biorefineries-industrial processes and products*; Wiley Online Library: WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2007.

(428) Thomas, S.; A., P. R. A process for the synthesis of formic acid by CO_2 hydrogenation: thermodynamic aspects and the role of CO. *Angew. Chem. Int. Ed.* **2011**, *50*, 7278–7282.

(429) Bulushev, D. A.; Ross, J. R. H. Towards sustainable production of formic acid. *ChemSusChem* **2018**, *11*, 821–836.

(430) Bernskoetter, W. H.; Hazari, N. Reversible hydrogenation of carbon dioxide to formic acid and methanol: Lewis acid enhancement of base metal catalysts. *Acc. Chem. Res.* **2017**, *50*, 1049–1058.

(431) Grasemann, M.; Laurenczy, G. Formic acid as a hydrogen source - recent developments and future trends. *Energy Environ. Sci.* **2012**, *5*, 8171–8181.

(432) Zhu, Q.-L.; Tsumori, N.; Xu, Q. Immobilizing extremely catalytically active palladium nanoparticles to carbon nanospheres: a weakly-capping growth approach. *J. Amer. Chem. Soc.* **2015**, *137*, 11743–11748.

(433) Henricks, V.; Yuranov, I.; Autissier, N.; Laurenczy, G. Dehydrogenation of formic acid over a homogeneous Ru-TPPTS catalyst: unwanted CO production and its successful removal by PROX. *Catalysts* **2017**, *7*, 348.

(434) Newsome, D. S. The water-gas shift reaction. *Catal. Rev.* **1980**, *21*, 275–318.

(435) Moulijn, J. A.; Makkee, M.; Van Diepen, A. E., *Chemical process technology*, 2nd ed.; John Wiley & Sons: West Sussex, PO19 8SQ, United Kingdom, 2013.

(436) Rostrup-Nielsen, J. R.; Sehested, J.; Norskov, J. K. Hydrogen and synthesis gas by steam- and CO_2 reforming. *Adv. Catal.* **2002**, *47*, 65–139.

(437) Roh, H. S.; Koo, K. Y.; Jeong, J. H.; Seo, Y. T.; Seo, D. J.; Seo, Y. S.; Yoon, W. L.; Park, S. B. Combined reforming of methane over supported Ni catalysts. *Catal. Lett.* **2007**, *117*, 85–90.

(438) Rostrup-Nielsen, J. R. In *Catalysis: science and technology*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1984, pp 1–117.

(439) Roh, H. S.; Jun, K. W.; Dong, W. S.; Chang, J. S.; Park, S. E.; Joe, Y. I. Highly active and stable Ni/Ce-ZrO_2 catalyst for H_2 production from methane. *J. Mol. Catal. A* **2002**, *181*, 137–142.

(440) Pena, M.; Gomez, J.; Fierro, J. New catalytic routes for syngas and hydrogen production. *Appl. Catal., A* **1996**, *144*, 7–57.

(441) Song, X.; Guo, Z. Technologies for direct production of flexible H_2/CO synthesis gas. *Energy. Convers. Manag.* **2006**, *47*, 560–569.

(442) LeValley, T. L.; Richard, A. R.; Fan, M. The progress in water gas shift and steam reforming hydrogen production technologies - a review. *Int. J. Hydrogen Energy* **2014**, *39*, 16983–17000.

(443) Sehested, J. Four challenges for nickel steam-reforming catalysts. *Catal. Today* **2006**, *111*, 103–110.

(444) Schwab, G.-M.; Schwab-Agallidis, E. On selective catalysis. *J. Amer. Chem. Soc.* **1949**, *71*, 1806–1816.

(445) Hietala, J.; Vuori, A.; Johnsson, P.; Pollari, I.; Reutemann, W.; Kieczka, H. In *Ullmann's encyclopedia of industrial chemistry*; American Cancer Society: 2016, pp 1–22.

(446) Zhao, Z.; Chen, Z.; Lu, G. Computational discovery of Nickel-based catalysts for CO_2 reduction to formic acid. *J. Phys. Chem. C* **2017**, *121*, 20865–20870.

(447) Yoo, J. S.; Abild-Pedersen, F.; Norskov, J. K.; Studt, F. Theoretical analysis of transition-metal catalysts for formic acid decomposition. *ACS Catal.* **2014**, *4*, 1226–1233.

(448) Trillo, J. M.; Munuera, G.; Criado, J. M. Catalytic decomposition of formic acid on metal oxides. *Catal. Rev.* **1972**, *7*, 51–86.

(449) Boddien, A.; Gärtner, F.; Federsel, C.; Sponholz, P.; Mellmann, D.; Jackstell, R.; Junge, H.; Beller, M. CO_2-"neutral" hydrogen storage based on bicarbonates and formates. *Angew. Chem. Int. Ed.* **2011**, *50*, 6411–6414.

(450) Yu, X.; Pickup, P. G. Recent advances in direct formic acid fuel cells (DFAFC). *J. Power Sources* **2008**, *182*, 124–132.

(451) Liu, Z.; Hong, L.; Tham, M. P.; Lim, T. H.; Jiang, H. Nanostructured Pt/C and Pd/C catalysts for direct formic acid fuel cells. *J. Power Sources* **2006**, *161*, 831 –835.

(452) Kundu, A.; Jang, J. H.; Gil, J. H.; Jung, C.; Lee, H.; Kim, S.-H.; Ku, B.; Oh, Y. Micro-fuel cells–current development and applications. *J. Power Sources* **2007**, *170*, 67 –78.

(453) Piola, L.; Fernandez-Salas, J. A.; Nahra, F.; Poater, A.; Cavallo, L.; Nolan, S. P. Ruthenium-catalysed decomposition of formic acid: fuel cell and catalytic applications. *Mol. Catal.* **2017**, *440*, 184–189.

(454) Schnabel, T.; Cortada, M.; Vrabec, J.; Lago, S.; Hasse, H. Molecular model for formic acid adjusted to vapor-liquid equilibria. *Chem. Phys. Lett.* **2007**, *435*, 268–272.

(455) Mura, M. G.; Luca, L. D.; Giacomelli, G.; Porcheddu, A. Formic acid: a promising bio-renewable feedstock for fine chemicals. *Adv. Synth. Catal.* **2012**, *354*, 3180–3186.

(456) Hindermann, J. P.; Hutchings, G. J.; Kiennemann, A. Mechanistic aspects of the formation of hydrocarbons and alcohols from CO hydrogenation. *Catal. Rev.* **1993**, *35*, 1–127.

(457) Dry, M. E. The Fischer-Tropsch process: 1950-2000. *Catal. Today* **2002**, *71*, 227–241.

(458) Schulz, H. Short history and present trends of Fischer-Tropsch synthesis. *Appl. Catal., A* **1999**, *186*, 3 –12.

(459) Roh, H. S.; Koo, K. Y.; Joshi, U. D.; Yoon, W. L. Combined H_2O and CO_2 reforming of methane over Ni-Ce-ZrO_2 catalysts for gas to liquids (GTL). *Catal. Lett.* **2008**, *125*, 283–288.

(460) Jang, W. J.; Jeong, D. W.; Shim, J. O.; Roh, H. S.; Son, I. H.; Lee, S. J. H_2 and CO production over a stable Ni-MgO-$Ce_{0.8}Zr_{0.2}O_2$ catalyst from CO_2 reforming of CH_4. *Int. J. Hydrogen Energy* **2013**, *38*, 4508 –4512.

(461) Roh, H. S.; Jun, K. W. Carbon dioxide reforming of methane over Ni catalysts supported on Al_2O_3 modified with, La_2O_3, MgO, and CaO. *Catal. Surv. Asia* **2008**, *12*, 239–252.

(462) Choudhary, V. R.; Mammon, A. S.; Sansare, S. D. Selective oxidation of methane to CO and H_2 over Ni/MgO at low temperatures. *Angew. Chem. Int. Ed.* **1992**, *31*, 1189–1190.

(463) Dong, W.-S.; Roh, H.-S.; Jun, K.-W.; Park, S.-E.; Oh, Y.-S. Methane reforming over Ni/Ce-ZrO_2 catalysts: effect of nickel content. *Appl. Catal., A* **2002**, *226*, 63–72.

(464) Hyun-Seog, R.; Ki-Won, J.; Wen-Sheng, D.; Sang-Eon, P.; Yung-Il, J. Partial oxidation of methane over Ni/θ-Al_2O_3 Catalysts. *Chem. Lett.* **2001**, *30*, 666–667.

(465) Knifton, J. F. Syngas reactions: IX. Acetic acid from synthesis gas. *J. Catal.* **1985**, *96*, 439 –453.

(466) Wender, I. Reactions of synthesis gas. *Fuel Process. Technol.* **1996**, *48*, 189 –297.

(467) Ajmera, S. K.; Losey, M. W.; Jensen, K. F.; Schmidt, M. A. Microfabricated packed-bed reactor for phosgene synthesis. *AIChE Journal* **2001**, *47*, 1639–1647.

(468) Choudhary, V. R.; Mondal, K. C. CO_2 reforming of methane combined with steam reforming or partial oxidation of methane to syngas over $NdCoO_3$ perovskite-type mixed metal-oxide catalyst. *Applied Energy* **2006**, *83*, 1024–1032.

(469) Cai, X.; Cai, Y.; Lin, W. Autothermal reforming of methane over Ni catalysts supported over ZrO_2-CeO_2-Al_2O_3. *J. Nat. Gas Chem.* **2008**, *17*, 201–207.

(470) Cai, X.; Dong, X.; Lin, W. Autothermal reforming of methane over Ni catalysts supported on CuO-ZrO_2-CeO_2-Al_2O_3. *J. Nat. Gas Chem.* **2006**, *15*, 122–126.

(471) Olah, G. A.; Goeppert, A.; Prakash, G. K. S. Chemical recycling of carbon dioxide to methanol and dimethyl ether: from greenhouse gas to renewable, environmentally carbon neutral fuels and synthetic hydrocarbons. *J. Org. Chem.* **2009**, *74*, 487–498.

(472) Byrne Jr, P.; Gohr, E.; Haslam, R. Recent progress in hydrogenation of petroleum. *Ind. Eng. Chem.* **1932**, *24*, 1129–1135.

(473) Tsang, S.; Claridge, J.; Green, M. Recent advances in the conversion of methane to synthesis gas. *Catal. Today* **1995**, *23*, 3–15.

(474) Vermeiren, W.; Blomsma, E.; Jacobs, P. Catalytic and thermodynamic approach of the oxyreforming reaction of methane. *Catal. Today* **1992**, *13*, 427 –436.

(475) van Beurden, P. On the catalytic aspects of steam-methane reforming. *Energy Research Centre of the Netherlands (ECN), technical report I-04-003* **2004**.

(476) Jeong, J. H.; Lee, J. W.; Seo, D. J.; Seo, Y.; Yoon, W. L.; Lee, D. K.; Kim, D. H. Ru-doped Ni catalysts effective for the steam reforming of methane without the pre-reduction treatment with H_2. *Appl. Catal., A* **2006**, *302*, 151–156.

(477) Wang, S.; Lu, G. Q. M.; Millar, G. J. Carbon dioxide reforming of methane to produce synthesis gas over metal-supported catalysts: state of the art. *Energy & Fuels* **1996**, *10*, 896–904.

(478) Koo, K. Y.; Roh, H.-S.; Seo, Y. T.; Seo, D. J.; Yoon, W. L.; Park, S. B. A highly effective and stable nano-sized Ni/MgO-Al_2O_3 catalyst for gas to liquids (GTL) process. *Int. J. Hydrogen Energy* **2008**, *33*, 2036–2043.

(479) Qin, D.; Lapszewicz, J.; Jiang, X. Comparison of partial oxidation and steam-CO_2 mixed reformingof CH_4 to syngas on MgO-supported metals. *J. Catal.* **1996**, *159*, 140–149.

(480) Wang, S.; Lu, G. Q. M. Catalytic activities and coking characteristics of oxides-supported Ni catalysts for CH_4 reforming with carbon dioxide. *Energy & Fuels* **1998**, *12*, 248–256.

(481) Oh, Y. S.; Roh, H. S.; Jun, K. W.; Baek, Y. S. A highly active catalyst, Ni/Ce-ZrO_2/θ-Al_2O_3, for on-site H_2 generation by steam methane reforming: pretreatment effect. *Int. J. Hydrogen Energy* **2003**, *28*, 1387–1392.

(482) Hou, Z.; Chen, P.; Fang, H.; Zheng, X.; Yashima, T. Production of synthesis gas via methane reforming with CO_2 on noble metals and small amount of noble-(Rh-) promoted Ni catalysts. *Int. J. Hydrogen Energy* **2006**, *31*, 555–561.

(483) Rostrupnielsen, J.; Hansen, J. H. B. CO_2-reforming of methane over transition metals. *J. Catal.* **1993**, *144*, 38–49.

(484) Ross, J.; van Keulen, A.; Hegarty, M.; Seshan, K. The catalytic conversion of natural gas to useful products. *Catal. Today* **1996**, *30*, 193–199.

(485) Bradford, M. C. J.; Vannice, M. A. CO_2 reforming of CH_4. *Catal. Rev.* **1999**, *41*, 1–42.

(486) Ashcroft, A.; Cheetham, A.; Green, M. Partial oxidation of methane to synthesis gas using carbon dioxide. *Nature* **1991**, *352*, 225–226.

(487) Claridge, J. B.; Green, M. L. H.; Tsang, S. C.; York, A. P. E.; Ashcroft, A. T.; Battle, P. D. A study of carbon deposition on catalysts during the partial oxidation of methane to synthesis gas. *Catal. Lett.* **1993**, *22*, 299–305.

(488) Jiang, H.; Li, H.; Zhang, Y. Tri-reforming of methane to syngas over Ni/Al_2O_3-thermal distribution in the catalyst bed. *J. Fuel Chem. Tech.* **2007**, *35*, 72–78.

(489) Dissanayake, D.; Rosynek, M. P.; Kharas, K. C.; Lunsford, J. H. Partial oxidation of methane to carbon monoxide and hydrogen over a Ni/Al_2O_3 catalyst. *J. Catal.* **1991**, *132*, 117–127.

(490) Ashcroft, A.; Cheetham, A.; Foord, J. a.; Green, M.; Grey, C.; Murrell, A. Selective oxidation of methane to synthesis gas using transition metal catalysts. *Nature* **1990**, *344*, 319–321.

(491) Koo, K. Y.; Roh, H.-S.; Jung, U. H.; Yoon, W. L. CeO_2 Promoted Ni/Al_2O_3 catalyst in combined steam and carbon dioxide reforming of methane for gas to liquid (GTL) process. *Catal. Lett.* **2009**, *130*, 217–221.

(492) Jing, Q.; Lou, H.; Fei, J.; Hou, Z.; Zheng, X. Syngas production from reforming of methane with CO_2 and O_2 over Ni/SrO-SiO_2 catalysts in a fluidized bed reactor. *Int. J. Hydrogen Energy* **2004**, *29*, 1245–1251.

(493) Prettre, M; Eichner, C.; Perrin, M The catalytic oxidation of methane to carbon monoxide and hydrogen. *Trans. Faraday Soc.* **1946**, *42*, 335b–339.

(494) Yoshitomi, S.; Morita, Y.; ichi Yamamoto, K. Catalytic partial oxidation of hydrocarbons at high temperature. *Bulletin of The Japan Petroleum Institute* **1962**, *4*, 15–27.

(495) Song, C.; Pan, W. Tri-reforming of methane: a novel concept for catalytic production of industrially useful synthesis gas with desired H_2/CO ratios. *Catal. Today* **2004**, *98*, 463 –484.

(496) Yoo, J.; Bang, Y.; Han, S. J.; Park, S.; Song, J. H.; Song, I. K. Hydrogen production by tri-reforming of methane over nickel-alumina aerogel catalyst. *J. Mol. Catal. A* **2015**, *410*, 74 –80.

(497) Halmann, M.; Steinfeld, A. Thermoneutral tri-reforming of flue gases from coal- and gas-fired power stations. *Catal. Today* **2006**, *115*, 170 –178.

(498) Lee, S. H.; Cho, W.; Ju, W.-S.; Cho, B. H.; Lee, Y. C.; Baek, Y. S. Tri-reforming of CH_4 using CO_2 for production of synthesis gas to dimethyl ether. *Catal. Today* **2003**, *87*, 133 –137.

(499) Choudhary, V. R.; Rajput, A. M. Simultaneous carbon dioxide and steam reforming of methane to syngas over NiO-CaO catalyst. *Ind. Eng. Chem. Res.* **1996**, *35*, 3934–3939.

(500) Zhang, Q.-H.; Li, Y.; Xu, B.-Q. Reforming of methane and coalbed methane over nanocomposite Ni/ZrO_2 catalyst. *Catal. Today* **2004**, *98*, 601–605.

(501) Pompeo, F.; Nichio, N. N.; Ferretti, O. A.; Resasco, D. Study of Ni catalysts on different supports to obtain synthesis gas. *Int. J. Hydrogen Energy* **2005**, *30*, 1399 –1405.

(502) Krieger, F. J.; White, W. B. A Simplified method for computing the equilibrium composition of gaseous systems. *J. Chem. Phys.* **1948**, *16*, 358–360.

(503) Brinkley Jr., S. R. Calculation of the equilibrium composition of systems of many constituents. *J. Chem. Phys.* **1947**, *15*, 107–110.

(504) Stone, E. Complex chemical equilibria: application of Newton-Raphson method to solve non-linear equations. *J. Chem. Educ.* **1966**, *43*, 241–244.

(505) Hill, T. L., *An introduction to statistical thermodynamics*, 1st ed.; Dover Publications Inc.: New York, USA, 2015.

(506) *Optimization Toolbox User's Guide*; MathWorks, Inc.: The MathWorks, USA, 2016.

(507) Bustamante, F.; Enick, R. M.; Cugini, A.; Killmeyer, R. P.; Howard, B. H.; Rothenberger, K. S.; Ciocco, M. V.; Morreale, B. D.; Chattopadhyay, S.; Shi, S. High-temperature kinetics of the homogeneous reverse water-gas shift reaction. *AIChE Journal* **2004**, *50*, 1028–1041.

(508) Joo, O.-S.; Jung, K.-D.; Moon, I.; Rozovskii, A. Y.; Lin, G. I.; Han, S.-H.; Uhm, S.-J. Carbon dioxide hydrogenation to form methanol via a reverse-water-gas-shift reaction (the CAMERE process). *Ind. Eng. Chem. Res.* **1999**, *38*, 1808–1812.

(509) Callaghan, C. A. Kinetics and catalysis of the water-gas-shift reaction: a microkinetic and graph theoretic approach., Ph.D. Thesis, Worcester Polytechnic Institute Worcester, MA, 2006.

(510) Tingey, G. Kinetics of the water-gas equilibrium reaction. I. the reaction of carbon dioxide with hydrogen. *J. Phys. Chem.* **1966**, *70*, 1406–1412.

(511) De Heer, J. The principle of Le Châtelier and Braun. *J. Chem. Educ.* **1957**, *34*, 375.

(512) Campbell, J. A. Le Châtelier's principle, temperature effects, and entropy. *J. Chem. Educ.* **1985**, *62*, 231.

(513) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Egidi, F. L. F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 09, Revision C.02., Gaussian Inc. Wallingford CT, 2016.

(514) Chase, M. W. NIST-JANAF Themochemical Tables, Fourth Edition. *J. Phys. Chem. Ref. Data* **1998**, *4*, 1–1951.

(515) Irikura, K. K. In *Computational thermochemistry*, 1998; Chapter 22, pp 402–418.

(516) Irikura, K. K.; Johnson, R. D.; Kacker, R. N. Uncertainties in scaling factors for ab Initio vibrational frequencies. *J. Phys. Chem. A* **2005**, *109*, PMID: 16834237, 8430–8437.

(517) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 theory. *J. Chem. Phys.* **2007**, *126*, 084108.

(518) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. Gaussian-3 (G3) theory for molecules containing first and second-row atoms. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(519) Te Velde, G.; Bickelhaupt, F.; Baerends, E.; Fonseca Guerra, C.; Dr. van Gisbergen, S.; Snijders, J.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

(520) Wavefunction Inc., Irvine, CA. `www.wavefun.com`, United States.

(521) Yaws, C. L., *Thermophysical properties of chemicals and hydrocarbons*, 2nd ed.; Gulf Professional Publishing: Oxford, UK, 2014.

Summary

Improving and developing simulation techniques are key to obtaining higher efficiency and accuracy in molecular simulations of dense liquid systems. The methodology development introduced in this thesis is relevant both for academia and industrial applications. In this thesis, the methods developments/improvements for molecular simulations are introduced followed by applications for realistic systems and systems of industrial relevance. The Continuous Fractional Component Monte Carlo (CFCMC) simulation methodology considers an expanded ensemble to solve the problem of low insertion/deletion acceptance probabilities in open ensembles. It allows for a direct calculation of the chemical potential by binning of the coupling parameter λ and using the probabilities $p(\lambda = 0)$ and $p(\lambda = 1)$. In chapter 2, the application of the CFCMC method in the GE is introduced. Chemical potentials of coexisting gas and liquid phases for water, methanol, hydrogen sulfide and carbon dioxide for the temperature range T=220 K to T=375 K are computed using two different methodologies: *(1)* the WTPI method in the conventional GE, and *(2)* the CFCGE MC method. It is shown that the WTPI method fails to accurately compute the chemical potentials of water and methanol in the liquid phase at low temperatures, while accurate chemical potentials in the liquid phase are computed using the CFCGE MC method. For the CFCGE MC method, the statistical uncertainty for computed chemical potentials of water and methanol in the liquid phase are considerably smaller compared to the WTPI method. For the water models considered in this study (SPC, TIP3P/EW, TIP4P/EW,

TIP5P/EW), computed excess chemical potentials based on three-site models are in better agreement with chemical potentials computed from an empirical equation of state from the NIST database. For water, orientational biasing is applied during test particle insertions to check whether certain orientations of test particles are energetically unfavorable. A two-dimensional Overlapping Distribution Method (ODM) in the NVT ensemble is derived for this purpose. It is shown that the failure of the WTPI method for systems with a strong hydrogen bonding network does not depend on orientation of the test molecule in that system. For all systems in this study, the WTPI method breaks down when the void fraction of the system drops below approximately 0.50.

In chapter 3, we show that extrapolating $p(\lambda)$ to compute the chemical potential leads to systematic errors when the distribution is steep. We propose an alternative binning scheme which improves the accuracy of computed chemical potentials. We also investigate the use of multiple fractional molecules needed in simulations of multiple components, and show that these fractional molecules are very weakly correlated and that calculations of chemical potentials are hardly affected. The statistics of Boltzmann averages in systems with multiple fractional molecules is shown to be poor. Good agreement is found between CFCMC averages (uncorrected for the bias) and Boltzmann averages when the number of fractional molecules is less than 1% of the total number of all molecules. We found that, in dense systems, biased averages have a much smaller uncertainty compared to Boltzmann averages.

In chapter 4, the performance of the spherical cutoff methods in MC and MD simulations was compared to the Ewald summation. The use of a spherical cutoff method for handling long range electrostatic interactions is computationally advantageous compared to the Ewald summation, especially for dense systems. Cutoff based methods in MC simulations are most attractive for dense systems as the effect of the long range interactions beyond the cutoff radius becomes negligible due to the efficient screening of charges. The radial distribution functions obtained from the Ewald summation and the Damped-Shifted Force (DSF) method were in excellent agreement. Numerical artifacts appeared at the cutoff radius when the

original Wolf method was used to calculate the electrostatic interactions. The calculated excess mixing enthalpies, excess chemical potentials, and activity coefficients of water and methanol obtained from the Wolf method were in good agreement with the DSF method. Our simulation results show that the numerical artifacts of the original Wolf method have little effect on energy calculations in aqueous methanol mixtures. The combination of the TraPPE and OPLS/2016 force fields with five water models, TIP3P, SPC/E, OPC, TIP4P/2005 and TIP4P/EW was used to compute mixing enthalpies, excess chemical potentials, and activity coefficients of water and methanol. Excess chemical potentials and activity coefficients were computed in an expanded version of the NPT ensemble. We found the best agreement between experimental data for all the computed properties of water-methanol mixtures for the TIP4P/2005-TraPPE force fields.

In chapter 5, an alternative method for calculating partial molar excess enthalpies and partial molar volumes of components in Monte Carlo (MC) simulations is developed. This method combines the original idea of Frenkel, Ciccotti, and co-workers [Chem. Phys. Lett., 1987, 136 pp 35-41] with the CFCMC technique. The method is tested for a system of Lennard-Jones particles at different densities. As an example of a realistic system, partial molar properties of a [NH_3, N_2, H_2] mixture at chemical equilibrium are computed at different pressures ranging from $P = 10$ MPa to $P = 80$ MPa. Results obtained from MC simulations are compared to those obtained from the PC-SAFT Equation of State (EoS) and the Peng-Robinson EoS. Excellent agreement is found between the results obtained from MC simulations and PC-SAFT EoS, and significant differences were found for PR EoS modeling. This is expected as it is well-known that the PR EoS fails at high pressures. We find that the reaction is much more exothermic at higher pressures. Moreover, the reaction enthalpy of the ammonia synthesis reaction is be computed directly by simple linear regression of the enthalpy as a function of the number of reactant molecules. The results is in excellent agreement with that of the CFCMC method.

In chapter 6, It is shown that in the grand-canonical ensemble, expressions for thermodynamic derivatives obtained from least squares multiple linear regression are identical to the expressions obtained from fluctuations.

This provides a conceptually simple and computationally efficient approach to obtain thermodynamic properties from fluctuations in multicomponent systems. Identical expressions are only obtained if higher order terms are not included in the linear regression model. Multiple linear regression is thermodynamically consistent with fluctuations both in constant-volume and constant-pressure ensembles. In the grand-canonical ensemble, multiple linear regression can be used to obtain the heat of adsorption even around sharp inflection points where the fluctuation approach is known to fail.

In chapter 7, an alternative method is introduced for free energy calculations at multiple temperatures and pressures from a single simulation. This method combines umbrella sampling with CFCMC simulations to obtain accurate estimates of chemical potentials for systems with appreciable overlapping density of states at (T^*, P) or (T, P^*) with the system of interest (T, P). The chemical potentials estimated at different temperatures and pressures are used to obtain partial molar properties. The partial molar excess enthalpy $\bar{h}^{\text{ex}}$ is obtained by numerically evaluating $(\partial(\beta\mu^{\text{ex}})/\partial\beta)_P$, and the partial molar volume υ is obtained by numerically evaluating $(\partial\mu/\partial P)_T$. This method is tested for a system of LJ at different densities. The results are compared to those obtained from independent simulations. The partial molar properties for LJ systems are compared to those obtained from the method in chapter 5, and the WTPI method. We found excellent agreement between all methods. As an example of a realistic system, the chemical potentials and partial molar properties of different mixtures of water-methanol are computed using umbrella sampling and the method in chapter 5. To test the limits of the method, we considered estimating properties of pure methanol at temperatures and pressures far from the system of interest ($T = 298$ K and $P = 1$ bar). The results show that for temperatures ± 15 K, accurate estimates of the chemical potential of methanol are obtained from a single simulation of 410 molecules. The method breaks down gradually by increasing the temperature difference. The chemical potentials of methanol ($N = 410$ molecules) are estimated accurately for pressures up to $P = 500$ bar. The method breaks down for pressures between $P = 500$ bar and $P = 1000$ bar. For larger systems this range becomes smaller since the relative fluctuations of energy and volume become smaller. For larger

systems this range becomes smaller since the relative fluctuations of energy and volume become smaller. We find that this method is a powerful tool for free energy calculations, especially for dense systems.

In the last two chapters of thesis, the focus is on industrially relevant applications. In chapter 8, simulation techniques developed in previous chapters are applied to simulate and verify the main thermophysical properties of $H_2O - H_2$ mixtures in the gas phase at elevated pressures (up to $P = 1000$ bar). The results are used to optimize the electrochemical compression of hydrogen using a recently developed compressor by HyET BV. Hydrogen is one of the most popular alternatives for energy storage. Due to its low volumetric energy density, hydrogen should be compressed for practical storage and transportation purposes. Recently, Electrochemical Hydrogen Compressors (EHC) have been developed that are capable of compressing hydrogen up to $P = 1000$ bar, and have the potential of reducing compression costs to 3 kWh/kg. As EHC compressed hydrogen is saturated with water, the maximum water content in gaseous hydrogen should meet the fuel requirements issued by the International Organization for Standardization (ISO) when refuelling Fuel Cell Electric Vehicles (FCEV). The ISO 14687-2:2012 standard has limited the water concentration in hydrogen gas to 5 µmol water per mol hydrogen fuel mixture. Knowledge on the vapor liquid equilibrium of $H_2O - H_2$ mixtures is crucial for designing a method to remove H_2O from compressed H_2. To the best of our knowledge, the only experimental high pressure data ($P > 300$ bar) for $H_2O - H_2$ phase coexistence is from 1927 [J. Am. Chem. Soc., 1927, 49, pp 65-78]. In this paper, we have used molecular simulation and thermodynamic modelling to study the phase coexistence of the $H_2O - H_2$ system for temperatures between $T = 283$ K to $T = 423$ K and pressures between $P = 10$ bar and $P = 1000$ bar. It is shown that the PR EoS and SRK EoS with van der Waals mixing rules fail to accurately predict the equilibrium coexistence compositions of the liquid and gas phase, with or without fitted binary interaction parameters. We have shown that the solubility of water in compressed hydrogen is adequately predicted using force field based molecular simulations. The modelling of phase coexistence of $H_2O - H_2$ mixtures may be improved by using polarizable models for water. In the appendix, we

present a detailed overview of available experimental solubility data for the $H_2O - H_2$ system at high pressures.

In chapter 9, we have considered formic acid as a source for both hydrogen and carbon monoxide to adjust the H_2:CO ratio in the syngas. Syngas is an important intermediate in the chemical process industry. It is used for the production of hydrocarbons, acetic acid, oxo-alcohols and other chemicals. Depending on the target product and stoichiometry of the reaction, an optimum (molar) ratio between hydrogen and carbon monoxide (H_2:CO) in the syngas is required. Different technologies are available to control the H_2:CO molar ratio in the syngas. Using thermochemical equilibrium calculations, we show that the syngas composition can be controlled by co-feeding formic acid (FA) into the SRM process. The H_2:CO molar ratio can be adjusted to a value between one and three by adjusting the concentration of FA in the reaction feed. At steam reforming conditions, typically above 900 K, FA can decompose to water and carbon monoxide, and/or to hydrogen and carbon dioxide. Our results show that co-feeding FA into the SRM process can adjust the H_2:CO molar ratio in a single step. This can potentially be an alternative to the WGS process.

Samenvatting

Om moleculaire simulaties uit te kunnen voeren in systemen bij hoge dichtheden is het van belang om simulatietechnieken te ontwikkelen en te verbeteren. De ontwikkeling van deze methoden is relevant voor zowel wetenschappelijke als industriële toepassingen. In dit proefschrift bekijken we eerst de ontwikkelingen op het gebied van moleculaire simulatietechnieken en daarna hun toepassing in realistische systemen die interessant zijn voor de industrie. Om het probleem op te lossen van de kleine acceptatiekans van het toevoegen danwel verwijderen van deeltjes in open ensembles is de *Continuous Fractional Component Monte Carlo* (CFCMC) methode ontwikkeld. Met behulp van deze methode, die gebruik maakt van de koppelingsparameter λ, is het mogelijk om in een simulatie direct de chemische potentiaal te berekenen door middel van de kansen $p(\lambda = 0)$ en $p(\lambda = 1)$. In Hoofdstuk 2 wordt uitgelegd hoe de CFCMC methode gebruikt kan worden in het Gibbs Ensemble (GE). Voor water, methanol, waterstofsulfide en koolstofdioxide wordt de chemische potentiaal uitgerekend in zowel de gas- als vloeistoffase bij temperaturen van T=220 K tot T=375 K en wordt gebruik gemaakt van (1) de WPTI methode in het conventionele GE en (2) de CFCMC methode in het GE. Hier wordt aangetoond dat, bij lage temperaturen, de chemische potentiaal van vloeibaar water en methanol nauwkeurig kan worden berekend met de CFCMC methode en dat de WPTI methode het daarentegen laat afweten. Bovendien zijn de foutmarges voor de CFCMC methode een stuk kleiner dan die voor de WPTI methode. Van de berekende excess chemische potentialen voor verschillende watermodellen (SPC,

TIP3P/EW, TIP4P/EW en TIP5P/EW) komen de drie-*site*-modellen het beste overeen met de chemische potentialen uit empirische toestandfuncties in de NIST database. Voor simulaties met water is gebruik gemaakt van *orientational biasing* voor het testdeeltje om te zien welke oriëntaties energetisch onwaarschijnlijk zijn. In het NVT ensemble is hiervoor een Overlapping Distribution Method (ODM) ontwikkeld. We laten zien dat de slechte werking van de WPTI methode in systemen met een netwerk van sterke waterstofbruggen niet afhangt van de oriëntatie van het testdeeltje. Uit dit onderzoek blijkt dat de WPTI methode niet werkt wanneer de zogenaamde *void fraction* van het systeem lager is dan ongeveer 0.5.

In Hoofdstuk 3 laten we zien dat extrapolatie van een sterk stijgende kansdichtheidsfunctie $p(\lambda)$ leidt tot een systematische fout in de berekening van de exces chemische potentiaal. Daarom introduceren we een alternatief *binning scheme* dat de berekening van de chemische potentiaal nauwkeuriger maakt. Verder onderzoeken we eventuele effecten van meerdere *fractional* moleculen in een systeem en zien we dat deze vrijwel niet gecorreleerd zijn aan elkaar en nauwelijks invloed hebben op de berekende chemische potentialen. In een systeem met meerdere *fractional* moleculen zijn Boltzmannstatistieken onnauwkeurig. Als het aantal *fractional* moleculen kleiner is dan 1% van het totaal aantal moleculen vinden we vrijwel gelijke Boltzmann en *biased* gemiddelden. Bovendien zien we dat de *biased* gemiddelden een veel kleinere foutmarge hebben dan de Boltzmann gemiddelden.

In Hoofdstuk 4 bestuderen we verschillende methoden om interacties (sferisch) af te breken in MC en MD simulaties. Dit wordt vergeleken met de Ewald methode. Het gebruik van deze methoden is rekentechnisch efficiënter, vooral voor systemen met een hoge ladingsdichtheid omdat *screening* van de ladingen de interacties over langere afstanden vrijwel verwaarloosbaar maakt. We zien in dit hoofdstuk dat de radiale distrubutiefuncties (RDF) in simulaties met de Ewald en *Damped-Shifted Force* (DSF) methoden goed overeenkomen. Er worden kleine afwijkingen waargenomen in de RDF wanneer de Wolf methode wordt gebruikt. We zien geen verschil tussen de Wolf en DSF methode in de berekening van mengenthalpie, chemische potentiaal en activiteitscoëfficiënten voor water en methanol. Onze simulaties laten zien dat er geen verschillen optreden in de energieberekening van

water-methanolmengsels. Voor het berekenen van mengenthalpieën, excess chemische potentialen en activiteitscoëfficiënten van water en methanol zijn de combinaties gebruikt van de TraPPE en OPLS/2016 krachtvelden met vijf watermodellen: TIP3P, SPC/E, OPC, TIP4P/2005 en TIP4P/EW. Om excess chemische potentialen en activiteitscoëfficiënten te berekenen zijn simulaties uitgevoerd in het extended NPT ensemble. De beste resultaten (in vergelijking met experimentele data) worden behaald voor de combinatie TIP4P/2005 met TraPPE.

In Hoofdstuk 5 besteden we aandacht aan de ontwikkeling van een methode voor het berekenen van partieel molaire excess enthalpieëen en partieel molaire volumes in MC simulaties. Deze methode combineert het werk van Frenkel, Ciccotti, en anderen [Chem. Phys. Lett., 1987, 136 pp 35-41] met de CFCMC methode. We testen deze methode eerst voor een simpel Lennard-Jones systeem voor verschillende dichtheden. Daarna berekenen we de partiële molaire grootheden van een [NH_3, N_2, H_2]-mengsel bij verschillende drukken van $P = 10$ MPa tot $P = 80$ MPa. De resultaten uit deze MC simulaties worden vergeleken met waarden uit de PC-SAFT toestandsfunctie (EOS) en de Peng-Robinson (PR) toestandsfunctie. De resulaten uit de simulaties komen goed overeen met die van PC-SAFT maar wijken af van die van PR-EOS. Dit kan worden verklaard uit het feit dat PR-EOS niet geschikt is voor systemen bij hoge druk. We concluderen dat de reactie, $N_2 + 3H_2 \rightleftharpoons 2NH_3$, sterker exothermisch wordt bij een toenemende druk. De reactie enthalpie voor de ammonia-reactia kan worden berekend door lineaire regressie van de enthalpie en het aantal reactanten. De resultaten zijn gelijk aan de resultaten van de CFCMC methode.

In Hoofdstuk 6 laten we zien dat de uitdrukking voor "thermodynamische afgeleiden"die wordt gevonden met meervoudige lineaire regressie gelijk is aan die in de fluctuatie-methode. Dit biedt een eenvoudige en snelle methode om thermodynamische eigenschappen te berekenen uit de fluctuaties in systemen met meerdere componenten. Dezelfde uitdrukkingen worden gevonden als hogere orde termen niet worden meegenomen in dit model. Meervoudige lineaire regressie is thermodynamisch consistent in ensembles met een constant volume en ensembles met een constante druk. In het groot-canoniek ensemble kan meervoudige lineaire regressie gebruikt

worden voor het berekenen van de adsorptie-warmte. Dit in tegenstelling tot de fluctuatie-methode waarvan bekend is dat deze niet werkt rond de buigpunten.

In Hoofdstuk 7 beschouwen we een methode voor het berekenen van de vrije energie uit één simulatie voor verschillende temperaturen en drukken. Dit is een combinatie van *umbrella sampling* en de CFCMC methode. Deze methode kan gebruikt worden om de chemische potentialen te bepalen voor systemen in de toestanden (T^*, P) of (T, P^*) indien hun toestandsruimte genoeg overlap heeft met die van het gesimuleerde systeem (T, P). Deze chemische potentialen kunnen op hun beurt weer gebruikt worden om andere partieel molaire grootheden te bepalen. De partiële molaire excess enthalpie $\bar{h}^{\text{ex}}$ kan berekend worden door numeriek de afgeleide $(\partial(\beta\mu^{\text{ex}})/\partial\beta)_P$ te bepalen. Op dezelfde manier kan het partieel molair volume v worden berekend door het bepalen van $(\partial\mu/\partial P)_T$. Deze methode is getest in een Lennard-Jones systeem en de partiële molaire grootheden zijn vergeleken met resultaten uit de methode beschreven in Hoofdstuk 5 en de WPTI methode. De resultaten van alle methoden zijn identiek. Voor een meer realistisch systeem bepalen we de partiële molaire grootheden van verschillende water-methanol mengsels met *umbrella sampling* en de methode uit Hoofdstuk 5. Door naar eigenschappen te kijken van systemen verder weg van het referentiepunt ($T = 298$ K en $P = 1$ bar) kunnen we de grenzen van de toepasbaarheid van onze methode bestuderen. De resultaten laten zien dat we voor temperaturen ± 15 K de chemische potentiaal nog steeds nauwkeurig kunnen bepalen uit één simulatie van 410 moleculen. Voor grotere temperatuursverschillen worden de resultaten onnauwkeuriger. Voor drukken tot 500 bar kunnen we nauwkeurig de chemische potentiaal bepalen maar tussen de $P = 500$ bar en $P = 1000$ bar neemt de onnauwkeurigheid toe. Voor systemen met een groter aantal deeltjes zullen deze grenzen kleiner worden omdat de fluctuaties in energie en volume dan relatief kleiner zijn. We concluderen dat deze methode geschikt is voor de bepaling van vrije energieën, in het bijzonder voor systemen met hoge dichtheden.

De laatste twee hoofdstukken van dit proefschrift richten zich op toepassingen die relevant zijn voor de industrie. In Hoofdstuk 8 worden de nieuwe simulatietechnieken gebruikt voor de berekening van thermofysische

eigenschappen van water-waterstofmengsels in de gasfase onder hoge druk (tot $P = 1000$ bar). De resultaten worden gebruikt voor het optimalizeren van de elektrochemische compressie van waterstof in een compressor, recent ontwikkeld door HyET BV. Waterstof is een van de populairste alternatieven voor energie-opslag. Omdat waterstof een lage energiedichtheid heeft is compressie vanuit praktisch oogpunt nodig om het op te slaan en te vervoeren. Er zijn elektrochemische waterstofcompressors (EHC) ontwikkeld die in staat zijn waterstof te comprimeren tot $P = 1000$ bar en de compressiekosten kunnen terugbrengen tot 3 kWh/kg. Het gecomprimeerde waterstof in een EHC is verzadig met water en hoort te voldoen aan de eisen van de Internationale Organisatie voor Standaardisatie (ISO) als het gebruikt wordt als brandstof voor waterstofauto's (FCEV). De ISO 14687-2:2012 standaard hanteert een limiet van 5 µmol water per mol waterstof als brandstofmengsel. Voor het verwijderen van water uit het waterstof-brandstofmengsel is inzicht in het gas-vloeistof evenwicht van groot belang. Voor zover wij weten is het artikel [J. Am. Chem. Soc., 1927, 49, pp 65-78] uit 1927 het enige met experimentele data over gas-vloeistofevenwichten van water-waterstofmengsels bij drukken hoger dan 300 bar. Wij gebruiken moleculaire simulatie en thermodynamische modellen voor de bepaling van het gas-vloeistofevenwicht van water-waterstofmengsels bij temperaturen van $T = 283$ K tot $T = 423$ K en drukken van $P = 10$ bar tot $P = 1000$ bar. Verder laten we zien dat de PR-EOS en SRK-EOS toestandsfuncties niet gebruikt kunnen worden om de evenwichtscomposities te bepalen in de gas- en vloeistoffase, in beide gevallen: met en zonder gefitte binaire interactie parameters. Moleculaire simulaties die gebruik maken van krachtvelden zijn wel geschikt hiervoor. Het modelleren van de evenwichten in water-waterstofmengsels kan wellicht worden verbeterd door gebruik te maken van polariseerbare modellen van water. In het Appendix presenteren we een uitgebreid overzicht van beschikbare experimentele data over gas-vloeistofevenwichten en oplosbaarheden van water-waterstofsystemen onder hoge druk.

In Hoofdstuk 9 gebruiken we mierenzuur om voor waterstof en koolstofmonoxide de verhouding H_2:CO in syngas aan te passen. Syngas is een belangrijk tussenproduct in de (chemische) procesindustrie en wordt

gebruikt voor de productie van koolwaterstoffen, azijnzuur, oxo-alcoholen en andere chemicaliën. Afhankelijk van het gewenste product en de stoichiometrie van de reactie is er een optimale verhouding tussen waterstof en koolstofmonoxide in het syngas. Er bestaan verschillende technieken om deze ratio aan te passen in syngas. We laten met thermochemische berekeningen zien dat de compositie van syngas kan worden aangepast door mierenzuur toe te voegen aan het SRM proces. De verhouding H_2:CO kan worden veranderd tussen 1 en 3 door de concentratie mierenzuur aan te passen. In het geval van stoomreforming (typisch bij temperaturen boven 900 K) kan mierenzuur ontbinden in water en koolstofmonoxide en/of waterstof en koolstofdioxide. Onze resultaten laten zien dat door het toevoegen van mierenzuur aan het SRM proces de verhouding H_2:CO in één enkele stap kan worden aangepast. Dit kan een interessant alternatief zijn voor het WGS proces.

Conclusions

Molecular simulation is a valuable tool to design and optimize industrial processes and can simultaneously provide better insight into the physics of these processes. By modelling the interactions between molecules, we can improve our understanding of how materials behave, especially at extreme conditions where conducting experiments are dangerous or expensive. Many important thermodynamic and transport properties are not easily obtainable using conventional simulation techniques, especially in open ensembles. This is mostly due to low probability of molecule exchanges in dense systems. For dense systems, computation of equilibrium compositions, chemical potentials and partial molar properties requires several (long) simulations. With the increase in density, computation of some properties become almost impossible using conventional methods. Improving simulation techniques and increased computational power make molecular simulations more attractive/relevant for many industrial applications. In chapter 3, we have improved the accuracy of computing chemical potentials in CFCMC simulations. Chemical potentials can be used to verify the condition of chemical equilibrium, compute activity coefficients in mixtures, and predict solubilities at extreme conditions. In addition to efficient free energy calculations, application of the CFCMC method is coupled with improved molecule exchange efficiency in phase equilibrium simulations. In chapter 5, we developed an alternative method, within the framework of the CFCMC method, to calculate partial molar properties from a single simulation. Partial molar properties are computationally difficult to calculate and

are experimentally difficult to measure at extreme conditions. Contribution of the partial molar enthalpies in calculating the reaction enthalpy can be significant at high pressures which can be calculated using our method without the drawbacks of the Widom's Test Particle Insertion (WTPI) method. We calculated partial molar enthalpies of nitrogen, hydrogen and ammonia to compute the reaction enthalpy of Haber-Bosch process at high pressures. We showed that the contribution of the partial molar excess enthalpies is not negligible for this process. It is expected that partial molar properties at high pressures are more accurately predicted using a physically based EoS such as PC-SAFT or advanced MC techniques, compared to a cubic EoS. However, cubic EoS are widely used to study other industrially important applications due to their simplicity. By developing faster and more efficient simulation techniques, we can encourage our collaborators from industry to use molecular simulations more often.

Based on the method proposed by Josephson and Siepmann, we computed the reaction enthalpy of Haber-Bosch process directly by simple linear regression of the enthalpy as a function of the number of reactant molecules. Obtaining partial molar properties from multiple linear regression, strongly suggests that fluctuations and multiple linear regression lead to identical values for thermodynamic derivatives. In chapter 6, we we rigorously proved the equivalence of the two approaches for computing thermodynamic derivatives in open ensembles of an n-component system. This means that in the grand-canonical ensemble, multiple linear regression can be used to obtain the heat of adsorption even around sharp inflection points where the fluctuation approach is known to fail. Using multiple linear regression for computing thermodynamic derivatives is simple and allows for using legacy data without any additional requirements.

In chapter 7, we developed a method to perform free energy calculations for mixtures at multiple temperatures and pressures from a single simulation, by combining umbrella sampling and the CFCMC technique. One can perform a simulation of a mixture at a certain pressure and temperature, and accurately compute the chemical potential at other pressures and temperatures close to the simulation conditions. Using the values of the estimated chemical potentials, the partial molar enthalpies and volumes of a compo-

nent are obtained by numerically evaluating the derivatives $(\partial(\beta\mu)/\partial\beta)_P$ and $(\partial\mu/\partial P)_V$, respectively. The method developed in this chapter also aims at obtaining more information about thermochemical properties of the system (including at near temperatures and pressures) to potentially save simulation time. We found excellent agreement with the results obtained in chapter 7.

In the last two chapters, we applied thermodynamic modelling to two different systems of industrial relevance. Hydrogen is an example of a renewable source for fuel and feedstock for production of chemicals. One of the emergent technologies in this area is the electrochemical compression of hydrogen to very high pressures. Due to the inner working of the electrochemical compressor, developed by HyET BV, water is dissolved in the compressed gaseous hydrogen in small, yet unknown, amounts. To develop and optimize of new drying technology, accurate molecular modelling is required to describe the phase behaviour of hydrogen at high pressures. Requirements for dry hydrogen for application in a hydrogen fuel cell car are found in the ISO standard. In this thesis, we applied advanced molecular simulation techniques to model the thermodynamics of $H_2O - H_2$ systems for a range of composition, temperature and pressure conditions. The main focus has been on high pressure, electrochemical hydrogen applications. The thermodynamic behaviour of hydrogen above $P = 500$ bar is not extensively studied. While conducting experiments at such high pressures are very expensive, molecular simulations can be used to provide detailed and comprehensive of the phase behaviour of the systems with hydrogen. Although the results for the high pressure VLE of $H_2O - H_2$ from molecular simulations were reassuring, we feel that further improvements may be realized taking polarizability of water molecules into account. Therefore, further molecular simulations of the $H_2O - H_2$ are recommended using polarizable force fields for water, especially to improve the predictions for the liquid phase composition.

Improved molecular models are necessary to describe physical/chemical processes accurately. One of the most studied molecules in the field of molecular simulation is water. Water is a flexible and polarizable molecule. To account for polarization, recently polarizable force fields for water have

been developed. Some properties of water, such as vapor pressure, critical properties, dielectric constant, and virial coefficient are most accurately predicated by considering polarization effects. However, the performance of different polarizable force fields to compute chemical potentials and activity coefficients of water is not fully investigated in literature, and more work needs to be done in this field. Further improvements in simulations of $H_2O - H_2$ systems may be realized taking polarizability of water into account.

Syngas is an important intermediate in the chemical process industry. It is used for the production of hydrocarbons, acetic acid, oxo-alcohols and other chemicals. Depending on the target product and stoichiometry of the reaction, an optimum (molar) ratio between hydrogen and carbon monoxide (H_2:CO) in the syngas is required. To adjust the H_2:CO ratio we have considered formic acid as a source for both hydrogen and carbon monoxide (chapter 9). Using thermochemical equilibrium calculations, we show that the syngas composition can be controlled by co-feeding formic acid into the SRM process. Due to low pressures, and temperatures, it was not needed to use molecular simulations for computing phase and reaction equilibria. However, with the increase in pressure, or decrease in temperature, it is strongly recommended to perform molecular simulations to describe the thermochemical properties of the system more accurately compared to conventional equation of state modelling.

Curriculum Vitae

Ahmadreza Rahbari was born on 21 September 1987 in Tehran, Iran. In 2005, he started a B.S.c degree in Mechanical Engineering at Azad University, South-Tehran Branch in Iran. Subsequently, he moved to Trondheim, Norway, in 2009 to transfer his studies and start a 5-year integrated M.S.c program at Norwegian University of Science and Technology (NTNU). After completing a full year intensive Norwegian language course at NTNU in 2009-2010, Ahmadreza started his Mechanical Engineering track. As part of his master's education, he participated in Erasmus Exchange program in 2012 and moved to the Netherlands to continue as a guest student at Delft University of Technology (TU Delft). In 2014, he completed his Master of Science (master i teknologi) in Mechanical Engineering (produktutvikling og produksjon). Shortly after his graduation, he joined the Graduate Engineering Program at FMC Technologies in Kongsberg, Norway. As a graduate engineering trainee, he had the opportunity to work in various departments at FMC Technologies and acquire on-the-job training, and cross-functional knowledge about the products and business processes withing the organization in a relatively short time. In the beginning of 2016, Ahmadreza returned back to Delft to join the Engineering Thermodynamic group of Prof. dr. ir. T. J. H. Vlugt at TU Delft as a PhD candidate in collaboration with Dr. D. Dubbeldam from the University of Amsterdam. His focus was studying thermodynamic equilibrium in industrially relevant systems or emerging applications relevant for industrial purposes. A part of his PhD studies was spent on improving and extending simulation techniques.

Another part was focused to do research related to hydrogen technologies and energy transition. During his PhD, he enjoyed collaborations with innovative companies and different research groups and companies such as HyET BV, Tielo-Tech BV, the University of Amsterdam (The Netherlands), Universidad Pablo de Olavide (Spain), Texas A&M University at Qatar (Qatar), the Institute of Nanoscience and Nanotechnology (Greece), and the University of Minnesota (USA).

List of Publications

Publications included in this thesis:

1. Rahbari, A.; Poursaeidesfahani, A.; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H.; Chemical Potentials of Water, Methanol, Carbon Dioxide, and Hydrogen Sulfide at Low Temperatures using Continuous Fractional Component Gibbs Ensemble Monte Carlo, Molecular Simulation, 2018, 44, 405-414

2. Rahbari, A.; Hens, R.; Nikolaidis, I. K. ; Poursaeidesfahani, A.; Ramdin, M.; Economou, I.G.; Moultos, O. A. ; Dubbeldam, D.; Vlugt, T. J. H., Computation of Partial Molar Properties using Continuous Fractional Component Monte Carlo, Molecular Physics, 2018, 116, 3331–3344

3. Rahbari, A.; Ramdin, M.; van den Broeke, L. J. P.; Vlugt, T. J. H.; Combined Steam Reforming of Methane and Formic Acid to Produce Syngas with an Adjustable H_2:CO Ratio, Industrial & Engineering Chemistry Research, 2018, 57, 10663-10674

4. Rahbari, A.; Hens, R.; Jamali, S. H.; Ramdin, M.; Dubbeldam, D.; Vlugt, T. J. H.; Effect of truncating electrostatic interactions on predicting thermodynamic properties of water-methanol systems, Molecular Simulation, 2019, 45, 336-350

5. Rahbari, A.; Brenkman, J.; Hense, R.; Ramdin, M.; van den Broeke, L. J. P.; Schoon, R.; Henkes, R.; Moultos, O. A.; Vlugt, T. J. H.; Solubility of Water in Hydrogen at High Pressure: a Molecular Simulation Study, Journal of Chemical & Engineering Data, 2019, 64, 4103-4115

6. Rahbari, A.; Hens, R.; Dubbeldam , D.; Vlugt, T. J. H., Improving the accuracy of computing chemical potentials in CFCMC simulations, Molecular Simulations, 2019, 117, 3493-3508

7. Rahbari, A.; Hens, R.; Moultos, O.A.; Dubbeldam, D.; Vlugt, T. J. H., Multiple free energy calculations from single state point Continuous Fractional Component Monte Carlo simulation using umbrella sampling, Journal of Computational Theory and Computation, 2020, in press. DOI:10.1021/acs.jctc.9b01097

Publications not included in this thesis:

1. Hens R.; Rahbari A.; Caro-Ortiz S.; Dawass N.; Erdős M.; Poursaeidesfahani, A.; Salehi H. S.; Celebi A. T. ; Ramdin M.; Moultos O. A.; Dubbeldam, D.; Vlugt T. J. H.; Brick-CFCMC: open source software for Monte Carlo simulations of phase and reaction equilibria using the Continuous Fractional Component technique, Manuscript in preparation.

2. Poursaeidesfahani, A.; Rahbari, A,; Torres-Knoop, A.; Dubbeldam, D.; Vlugt, T. J. H.; Computation of thermodynamic properties in the Continuous Fractional Component Monte Carlo Gibbs ensemble: Molecular Simulation, 2016, 43, 189-195

3. Poursaeidesfahani, A.; Hens, R.; Rahbari, A,; Ramdin, M.; Dubbeldam , D.; Vlugt, T. J. H.; Efficient application of Continuous Fractional Component Monte Carlo in the Reaction Ensemble: Journal of Chemical Theory and Computation, 2017, 13, 4452-4466

4. Matito-Martos, I.; Rahbari, A.; Martin-Calvo, A.; Dubbeldam, D.; Vlugt, T. J. H.; Calero, S.; Adsorption equilibrium of nitrogen dioxide in porous materials, Physical Chemistry Chemical Physics, 2018, 20, 4189-4199

Acknowledgements

The time has come to say goodbye to one of the most unique experiences of my life, my PhD journey. It is rightly called "the PhD rollercoaster ride" filled with successful moments, typical academic setbacks, and occasionally unknown turnarounds. It was a combination of all these circumstances which has ultimately given long-term benefits to my personal development. In hindsight, I can see how my PhD journey has changed me, my social life, and my relationships for the better. During this journey, I was lucky enough to be able to work alongside a very resourceful and knowledgeable group of scientists, researchers and friends. I would like to thank my supervisors, colleagues, friends and family who in their own way helped me overcome the challenges. First of all, I would like to thank my supervisors Thijs J. H. Vlugt and David Dubbeldam:

Thijs, ik herinner me nog steeds ons allereerste gesprek op Skype vanuit Noorwegen in 2015, toen ik met mijn zeer beperkte woordenschat probeerde jou in het Nederlands te begroeten. Ik wil jou daarom graag in het Nederlands bedanken. Van harte bedankt! Het was een privilege om onderzoek te kunnen doen onder jouw begeleiding. Tijdens mijn promotie heb jij mij goede steun geboden. Zonder jouw begeleiding, zou het onmogelijk zijn geweest om tot dit proefschrift te komen. Onze samenwerking kan het best worden gekenmerkt door boeiende, wetenschappelijke discussies, en oprechte en snelle communicatie. Het meest interessante was het ontwikkelen van methodes, het gebruik van statistisch mechanische vergelijkingen en de wetenschappelijke discussies die daarbij kwamen kijken. Uiteindelijk

hebben we samen een groot aantal artikelen geschreven. Dit zal lang een goede herinnering blijven aan de tijd van mijn promotie. Ik wil ook mijn andere promotor, David, hartelijk bedanken. Bedankt voor alle feedback die jij gegeven hebt op mijn werk. Ik heb altijd genoten van onze vergaderingen bij jou aan de universiteit van Amsterdam, onze samenwerking aan de RASPA-workshops en de MolSim cursussen. Deze gesprekken hadden altijd een belangrijke toegevoegde waarde aan onze artikelen.

Special thanks to other members of my committee for reviewing and accepting my dissertation: Prof.dr. N. van der Vegt, Prof.dr.-ing.habil. J. Vrabec, Prof.dr. J.H. van Esch, Prof.dr.ir. W. de Jong, and Dr. Othon Moultos. Many thanks to my mentor Dr. HosseinNia.

Collaborations with different scientists and fellow researchers over the past four years have contributed significantly to the quality of this thesis. Hereby, I would like to thank Dr. Albert Bos, Prof.dr.ir. R.A.W.M. Henkes, Rogier Schoon, Dr. Ioannis Economou, Dr. Tyler Josephson. Many thanks to Prof. S. Calero, and Prof. Kjelstrup for providing us the opportunity to visit their research groups in Seville and Trondheim.

I am particularly thankful for having such an amazing friend and colleague, Seyed Hossein. We met in 2012 in Delft when I was participating in the Erasmus exchange program from Norway, and our friendship has been growing ever since. In 2015, you let me know about a PhD position at the Process and Energy Department which ultimately lead to my decision to move back to the Netherlands. You are one of the most resourceful and selfless people I have ever known. You never hesitated to help me with your scientific insight even at your busiest moments. I cannot thank you enough.

I have had the pleasure of working with great colleagues and fellow researchers from the Process and Energy department and other research institutes. Thank you, Ali, for all the collaborations over different scientific topics. I am grateful that our friendship has grown since we met in 2012. You have been very helpful to me also outside of work. I would like to thank Remco Hens. Our PhD projects have been progressing simultaneously, and I am happy to see that the development of Brick-CFCMC, the specialized and efficient Monte Carlo code for phase equilibrium calculations, is in its final stages. I am thankful for our efficient collaboration which was based

on mutual respect and confidence. Many thanks to Mahinder and Peter for our collaborations on different topics. Your approach to science has always offered an all-round perspective, namely, the chemical engineering, economics and modelling perspective. My gratitude also goes to Ariana-Torres Knoop, Jeroen Brenkman, Ilias Nikolaidis, and Ismael Matito for our collaborations.

I would also like to thank my great colleagues and friends at TU Delft. Nikos, you are a great scientist and musician, and I am grateful that we could play some Iranian classical music together, and I enjoyed playing the Bouzouki very much. Elyas, I always enjoyed our conversations about Persian literature, especially Rumi's poems. Vilborg, you are one of the most positive and high-spirited people I have ever met. I would also like to thank, Brian, Remco Hartkamp, Hirad, Rishab, Karsten, Noura, Mate, Max, Alper, Sebastian, Fatma, Stephan, Selvaraj, Jie, Weiwei, Rohit, Marloes, Jurriaan, Farnaz, Maryam, Mengmeng, Christos, Luis, Hakan, Carla, Gustavo, Sebastian, Sathish, Fatima, Erfan, Sergio and Tim.

Many thanks to my closest friends: Arash, you are one of the kindest and most resourceful and helpful people I have ever met, and certainly one of the best cooks. I am grateful that our friendship has grown since we met in 2012. Many thanks to Saman, Shima, Arghavan, Soheil, Marzie, Antje, Amir, Sanaz, Reza Farzad, Hossein Norouzi and Alireza Ghasemi for all the nice memories during my PhD.

Jeg vil gjerne benytte anledningen til å anerkjenne en person som har spilt en meget viktig rolle i hele mitt eventyr i Europa. Odette, tusen hjertelig takk. Jeg kan godt si at det hele eventyret begynte da du spurte meg, i telefonen: "hvorfor kommer du ikke for å studere i Norge?" I den tiden kunne jeg veldig lite om Norge, men du ga meg all informasjon jeg trengte for å kunne gjøre et klokt valg. Jeg var overbevisst da om at det kom til å gå bra. Hadde jeg ikke begynt på NTNU i 2009, hadde jeg mest sannsynlig ikke fått muligheten til å ta en doktorgradsutdanning på universitetet i Delft, tror jeg. Du hjalp meg masse med søknaden min til NTNU. Jeg vil gjerne uttrykke min takknemlighet en gang til, til deg og Mansour, og si at jeg kommer til å huske denne tjenesten ganske lenge. Dere lærte meg om den norske kulturen og tok vare på meg frem til jeg kunne

fungere selvstendig. I tillegg har dere alltid vært til støtte for meg. Jeg vil også takke Hilda og Bjørn for støtten dere har gitt meg. Dere tok vare på meg de aller første dagene i Trondheim da hybelen min ikke var klar ennå. Dere har kjørt til meg på kalde og mørke vinterdager med nydelig hjemmelaga mat. Det er jeg takknemlig for. Bjørn, du har vært utrolig snill mot meg og hjulpet meg et par ganger med å flytte i Trondheim. Tusen takk for det. Jeg kommer til å huske deres hjelp veldig lenge. I løpet av oppholdet mitt i Trondheim ble jeg kjent med andre fantastiske mennesker. Armin, Karen, Ida, Inger-Marie. Dere har alltid vært åpne for interessante samtaler og diskusjoner på tvers av de norske, armenske og iranske kulturene. Ingvill, tusen takk for at du inspirerte meg til å dra på utveksling til Nederland. Som du vet, har min nåværende PhD-utdannelse en sterk sammenheng med min utvekslingsperiode.

I tillegg vil jeg gjerne takke min norsklærer, eller min norske mor, Sissel. Du har vært utrolig snill og behjelpelig mot meg. Etter norskkurset fortsatte du å gi meg råd og støtte på mange områder for at jeg kunne fungere bedre i det norske samfunnet. Jeg har alltid vært velkommen hos deg. Det virker som om du fremdeles føler deg ansvarlig for dine studenter selv om vi har vært ferdige med norskkurset i mange år. Da jeg søkte jobb, leste og korrigerte du alle mine følgebrev, og alltid på kort varsel. Siden 2009 har vi fremdeles kontakt og vi skypes av og til. Jeg er takknemlig for din støtte.

I løpet av min PhD-utdannelse har jeg vært på besøk til Trondheim, og vil gjerne takke Lauren, Arun, Tatiana, Egil, Fahd, Nicoleta og Elin. I tillegg har jeg vært i kontakt med mine gamle venner fra FMC Technologies i Kongsberg. Tusen takk Thomas, Arya, Abdi, Martin og Ole for turen til Barcelona og Amsterdam. Thomas, tusen takk for at du inviterte meg for å feire jul sammen med familien din. Det betyr veldig mye for meg. Jeg vil også takke mine tidligere kolleger fra FMC Technologies som jeg fremdeles har kontakt med: Gustavo, Mari, Simon og Rina.

I would like to thank my parents, Firouzeh and Alireza for their unconditional love and support. Mom, Dad, without your support, I would not have been able to come so far. You made sacrifices so that I could have the education which was best for my future. I think the biggest sacrifice for us all is living apart in different countries, which has not been easy. I am

grateful for all your help. I would not have succeeded without your support. I dedicate this PhD degree to you, although it seems like such a small token of appreciation compared to your unconditional love and support. I would like to thank my dear sister, Azade, for always making sure I am doing fine, and for listening to me talk about my challenges. You have always called me even at your busy moments, and I thank you for always believing in me. Thank you for all the nice gifts you have sent all these years. I am grateful. Eventhough we are living far away from each other, I hope I can spend more time with you, Mehran, and dear Kaveh.