

Molecular Simulation of Binary and Ternary Vapour-Liquid Equilibria

An Advanced Gibbs-Duhem Integration Study

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr ir J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 31 januari 2005 om 13.00 uur

door

Aleidus VAN 'T HOF

scheikundig ingenieur

geboren te Leiden

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr ir J. de Swaan Arons

Prof. dr S.W. de Leeuw

Toegevoegd promotor: Dr ir C.J. Peters

Samenstelling promotiecommissie:

Rector Magnificus,

Prof. dr ir J. de Swaan Arons,

Prof. dr S.W. de Leeuw,

Dr ir C.J. Peters,

Prof. dr ir M.-O. Coppens,

Prof. dr G.J. Witkamp,

Prof. dr J.M.J. van Leeuwen,

Prof. dr W.J. Briels,

Voorzitter

Technische Universiteit Delft, promotor

Technische Universiteit Delft, promotor

Technische Universiteit Delft, toegevoegd promotor

Technische Universiteit Delft

Technische Universiteit Delft

Universiteit Leiden

Universiteit Twente

ISBN: 90-9019032-5

Copyright © 2005 by Aleidus van 't Hof

Printed by Wöhrmann Print Service

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the publisher.

Printed in the Netherlands

.....de wijsheid – waar moet je haar zoeken,
en het inzicht – waar is het te vinden?
Geen sterveling kent de weg erheen,
de wijsheid is niet in het land der levenden.

.....

Maar God kent haar wegen
en weet waar ze verblijft.

.....

En Hij sprak tot de mens:
‘Ontzag voor de Heer – dat is wijsheid;
het kwaad mijden – dat is inzicht.’

Job 28 (De Nieuwe Bijbelvertaling)

Table of Contents

Chapter 1.	Introduction	1
1.1.	The need for accurate prediction of thermophysical properties	1
1.2.	Molecular simulation	2
1.3.	Problem definition	4
1.4.	Outline of this thesis	5
1.5.	References	7
Chapter 2.	Gibbs-Duhem integration and the semigrand-canonical ensemble	9
2.1.	Gibbs-Duhem integration among other simulation methods	9
2.2.	Legendre transforms and the Gibbs-Duhem equation	11
2.3.	Semigrand variables	14
2.4.	Clapeyron equations	15
2.5.	Description of several ensembles	16
2.6.	Numerical integration of the Clapeyron equation	23
2.7.	Conclusions	26
2.8.	References	27
Chapter 3.	Principles of Monte Carlo importance sampling	30
3.1.	Computation of multidimensional integrals; the problem	30
3.2.	Numerical integration and Monte Carlo importance sampling	31
3.3.	Computation of statistical-mechanical ensemble averages	33
3.4.	Markov chains	34
3.5.	Biasing the underlying matrix of the Markov chain	36
3.6.	Conclusions	42
3.7.	References	42
Chapter 4.	Computing the initial coexistence point and the initial integrand	44
4.1.	Introduction	44
4.2.	Predicting the initial slope	44
4.3.	Predicting the initial point	61
4.4.	Conclusions	75
4.5.	References	76
Chapter 5.	Simulation results for binary vapour-liquid equilibria	81
5.1.	Force field	81
5.2.	Combining rules	91
5.3.	Simulation details	94
5.4.	Simulation results	95
5.5.	Error analysis of Gibbs-Duhem integration	107
5.6.	Conclusions	111
5.7.	References	112

Chapter 6.	Advanced Gibbs-Duhem integration	117
6.1.	Introduction	117
6.2.	Histogram-reweighting technique	118
6.3.	Multiple-histogram reweighting in the semigrand-canonical ensemble	122
6.4.	Advanced Gibbs-Duhem integration	125
6.5.	Parallel tempering in the semigrand-canonical ensemble	130
6.6.	Simulation results for binary vapour-liquid equilibria	131
6.7.	Error analysis of advanced Gibbs-Duhem integration	149
6.8.	Conclusions	152
6.9.	References	153
Chapter 7.	Extension to ternary mixtures	157
7.1.	Introduction	157
7.2.	Clapeyron equation, initial point, and initial slope	158
7.3.	Advanced Gibbs-Duhem integration in ternary mixtures	161
7.4.	Simulation details and results	163
7.5.	Conclusions	168
7.6.	References	169
Appendix A.	Clapeyron equations	171
Appendix B.	Derivation of a third-order variable-step-size predictor equation and a fourth-order variable-step-size corrector equation	174
B.1.	Third-order predictor equation with variable step size	174
B.2.	Fourth-order corrector equation with variable step size	175
B.3.	Reference	176
Appendix C.	Derivation of the Coulombic contribution to the virial	177
Appendix D.	Simulation results	182
	Samenvatting en aanbevelingen	192
	Summary and recommendations	197
	Dankwoord	201
	Curriculum Vitae and Publications	204

Chapter 1. Introduction

1.1. The need for accurate prediction of thermophysical properties

In everyday life, each person encounters, either noticed or without ever being observed, the competition between vapour and liquid phases. Consider, for example, the reservoir on the radiator which is filled with water and is meant to maintain a specific level of the humidity in the air. Since the reservoir is empty after some time, it does not really establish a vapour-liquid equilibrium situation; it rather searches for an equilibrium situation but never finds it due to ever changing ambient conditions. In case the room that contains the radiator is a closed system, a real vapor-liquid equilibrium situation would occur, provided the reservoir contains enough water. In fact there are no practical examples of true vapour-liquid equilibrium situations in the real world. However, many continuous flow processes in the chemical and process industry approach the true equilibrium situation.

The importance of knowledge of vapour-liquid equilibria and other thermodynamic properties to the process industry will be illustrated with an example.

Carrying out a certain reaction usually requires subsequent purification of the product by removal of by-products and/or a solvent. Sometimes separation itself is the main function of an entire process. Separation processes usually account for the larger part of the operational and capital investment costs [1]. Clearly, there is a large incentive for research directed towards cost reducing improvements of separation processes and/or the development of new ones. The selection of a proper separation process and the subsequent design of the separation equipment require knowledge of the thermodynamic data behind the process [1-3].

Consider, for example, a distillation column in which two components are separated based on a difference in volatility. The stages in such a column operate at a different temperature and a slightly different pressure. The vapour is moving to the top of the column while the liquid is moving downward. Simply stated, when the vapour and the liquid arrive on a particular stage, they exchange matter; part of the more volatile component moves to the vapour phase and a fraction of the less volatile component moves to the liquid phase, so as to establish vapour-liquid coexistence at the temperature and pressure of interest.

However, within the finite residence time on that stage, the true coexistence situation is never reached. The relative difference between the actual situation and true coexistence is a measure of the stage efficiency.

Obviously, knowledge of vapour-liquid equilibria is needed to select the right working pressure and number of stages of the column that are required to meet the imposed purity of the distillate.

Other thermodynamic properties like heat capacities, Joule-Thomson coefficients, isothermal compressibilities, and isobaric expansivities are of interest to the process industry as well. Many of these properties are tabulated and can be found in literature. Unfortunately, the larger part of these data consists of pure-component properties. Since chemical reactions, separations and many other processes involve mixtures, these data are usually not sufficient for design of process equipment. Furthermore, due to high costs of special equipment, if multicomponent data are available the larger part of it covers only a range of moderate temperatures and/or pressures.

Experimental investigation of thermophysical properties is always limited by the ability to prepare a representative sample and/or to devise a proper experimental setup. Furthermore, the performance of experiments is often expensive and time-consuming. Therefore, it is often

worthwhile to maximize the amount of information extracted from scarce experimental data. This can, for example, be accomplished by optimizing a thermodynamic model to those data. Subsequently, the optimized model can be used to predict data at other thermodynamic states.

Thermodynamic models that are commonly used to describe phase equilibria in chemical-engineering applications are equations of state (EOS) and activity-coefficient models [4]. The adjustable parameters in these models need to be optimized to experimental data. Unfortunately, the physical significance of the parameters is often lost as they include any incapability of the model to describe the experimental data.

In case multicomponent vapour-liquid equilibria are modelled, one often uses binary interaction parameters to extend the applicability of the model. Moreover, optimization of EOS-parameters to data sets at very different temperatures is often not possible without introducing temperature-dependent binary interaction parameters.

One can imagine that tabulation of the EOS-parameters and binary interaction parameters for only a limited number of components over a small range of state conditions would already require a large database. Nevertheless, the easy implementation and mathematical simplicity of these thermodynamic models explain their widespread use in engineering applications.

1.2. Molecular simulation

In Section 1.1, the use of thermodynamic models to increase the amount of information from scarce experimental data was discussed. On the other hand, one can develop *force fields* for the components of interest. These force fields describe how molecules feel about each other. The development of force fields can be accomplished with quantum-chemical methods or by optimizing the parameters to experimental data. Having found a proper force field, the properties of matter can be described on a molecular level. Statistical mechanics can be used to relate microscopic behaviour to bulk thermodynamic properties.

Statistical-mechanical descriptions of a system typically involve multidimensional integrals that cannot be computed analytically. However, *molecular simulation* may provide the numerical solution to the full statistical-mechanical problem. The results, are exact within the numerical accuracy of the computer provided an accurate simulation method is used and a very long simulation is performed. *Monte Carlo* (MC) and *Molecular Dynamics* (MD) are the proper techniques to perform molecular simulation.

In MD simulations [5, 6], Newton's equations of motion are solved numerically using discrete time steps. As the system is followed in time, MD provides dynamic properties as well as static properties.

In Monte Carlo simulations [7, 8], the momenta of the molecules are integrated out, as the dynamics of the system is unimportant. The name *Monte Carlo* refers to the famous gambling city and arises from the abundant use of random numbers. A random walk through the molecular configurations that correspond to the equilibrium state of the system is performed. The transition probabilities, that represent the probability to perform a move from a present configuration to a new configuration, are defined so as to sample configurations with a relative frequency that is proportional to the relevant statistical-mechanical probability distribution. Since there is no explicit time scale, the order in which the configurations are generated is unimportant. A direct consequence is that MC simulation provides only static properties.

An advantage of MC simulation over MD simulation is the ability to perform unphysical moves that may enhance the sampling of the configurations accessible to the simulated

system. Years of research on MC simulation techniques have provided a whole arsenal of special tricks [9-11].

A characteristic that is simultaneously a weak and a strong feature of molecular simulation is the need to choose a proper force field. The advantage of such a choice is that one can make the force field as simple as possible, thereby reducing the computational costs of the simulation. Moreover, one can study the influence of different contributions to the intermolecular potential on the simulation results. The weak point is that it is hard to find out whether a potential is really representative for the component of interest. In order to reduce the number of force-field parameters, one usually relates the interactions between different components to those between like components. However, there is little information in pure-component interactions on how this relation should look like.

Obviously, molecular simulation is not a complete substitute for real experiments, as some experimental data are still needed to check the performance of a force field. Nevertheless, if a proper force field has been found, molecular simulation can be used to provide thermophysical data, possibly at state conditions in the region where experiments can hardly be performed and/or are too expensive due to costly equipment. The computer is insensitive to the physical conditions of the simulation. Thus, simulation at vacuum, at high pressures or at cryogenic temperatures is no problem. An additional advantage is that the input parameters of a simulation like a temperature, pressure or simulation volume are exact; they do not have an uncertainty like experimental input parameters.

Furthermore, it is worthwhile to mention a number of other advantages of molecular simulation over real experiments that are certainly not trivial. As both MC and MD are used to generate molecular configurations and/or motions, information is generated that is not directly or hardly available from *real* experiments. Molecular simulations provide insight regarding structure of phases, vaporization or condensation of phases, density fluctuations near the critical point, and many other phenomena.

Moreover, molecular simulations may be less labour-intensive than real experiments. For example, it is easier to quit a simulation and delete the input file than to clean experimental equipment!

Last but not least, molecular simulation may provide a number of thermodynamic variables from a single simulation, whereas the experimental measurement of these properties often requires different experimental equipment.

Notwithstanding the number of advantages of molecular simulation over real experiments, the use of molecular simulation methods in engineering applications is limited. The reason is probably the rather hard implementation of Monte Carlo and Molecular Dynamics simulation methods. Furthermore, molecular simulations are computationally expensive compared to the application of thermodynamic models. However, the progress that has been made in the development of new simulation methods over the last twenty years, as well as the ever-increasing computer speed, will enhance the applicability for engineering purposes.

1.3. Problem definition

Among the simulation methods that are applicable to predict phase equilibria is the *Gibbs-Duhem integration* method [12, 13]. Gibbs-Duhem integration implies the numerical integration of a Clapeyron differential equation. Clapeyron equations describe changes in the field variables that maintain phase coexistence. Field-type variables necessarily take on identical values in coexisting phases. The numerical integration traces a coexisting curve from a predetermined starting coexistence point and a predetermined value of the right-hand-side of the differential equation. The Clapeyron equation describes mono-variant phase coexistence; it applies to any type of coexistence situation as long as the number of degrees of freedom is one.

The Gibbs-Duhem integration method was originally developed to predict pure-component phase equilibria, but later on, it was extended to multicomponent mixtures [14]. It is suitable for the simulation of mixtures and/or for simulation of phase equilibria including a solid phase [15-20]. The method has been shown to be capable of predicting the binary phase behaviour, including vapour liquid and solid phases, of simple model force fields [17, 18].

Initially, it was the purpose of this Ph.D. project to see whether the Gibbs-Duhem integration method, combined with realistic force fields, would be capable of predicting the total phase behaviour, including vapour liquid and solid phases, of real systems in close agreement with experimental data. However, during the course of the project, a number of problems was encountered. Finding the solution to these problems became a project in itself. For that reason, this work is restricted to the prediction of vapour-liquid equilibria.

The method that is in the centre of interest is Gibbs-Duhem integration combined with Monte Carlo simulation. This work roughly focuses on two topics:

- 1) The ins and outs of Gibbs-Duhem integration are investigated and discussed.
- 2) The Gibbs-Duhem integration method is used to simulate vapour-liquid equilibria of binary and ternary mixtures.

Ad 1) From a scientific point of view, the simulation method in itself is interesting. Initially, it may seem superfluous to investigate the capabilities of the method as Gibbs-Duhem integration is a numerical integration technique that should bring us from coexistence point *A* to coexistence point *B*; if it works for simple model force fields, it should work for more complex force fields as well. Although this statement is formally true, the practical application of the method is often more difficult when complex force fields are used

In order to correctly implement Gibbs-Duhem integration, a *good estimate for the initial coexistence point* and a corresponding *good estimate for the right-hand side of the Clapeyron equation* are needed. The right-hand side will be called the *initial slope*. The computation of the initial coexistence point and the corresponding initial slope often require additional advanced simulation techniques if realistic force fields are used instead of simple force fields in which molecules are assumed to be spherical. The prediction of the initial coexistence point and initial slope with molecular simulation will be discussed in detail.

The convenience of the Gibbs-Duhem integration will be investigated by comparing the method to other simulation methods. As will be pointed out in this thesis, the conventional Gibbs-Duhem integration method is an *inefficient method*. Attempts have been made to improve the applicability and efficiency of the method by combining the Gibbs-Duhem integration method with *more recent simulation and analysis techniques*.

Ad 2) From an application-oriented point of view, it is not only important that the simulation method works properly and easily, but it is also interesting whether the simulation method, combined with realistic force fields, is capable of predicting thermodynamic properties in close agreement with experimental data. Provided the simulation method works fine, the agreement between simulation results and experimental data as well as predictive power mainly depend on the physical significance of the force fields that are used; the simulation method is just a tool to arrive at the results that correspond to the force field and the thermodynamic conditions.

Force fields are available for a number of industrially important components. The force fields that are used in this work are chosen so as to make sure that at least the simulated pure-component phase behaviour matches the experimental data. Thus, a thorough analysis of available force fields has to be made. For the method to be as predictive as possible, cross-component interactions are computed via so-called combining rules; *i.e.* interactions between unlike components/interaction groups are expressed in the interactions between like components/interaction groups. As pure-component simulated phase behaviour will match experimental data, the simulation studies performed on mixtures are helpful in *judging the appropriateness of different combining rules* for the interactions between molecules of unlike components.

1.4. Outline of this thesis

In Chapter 2, the Gibbs-Duhem integration technique is commented upon and its position among other simulation techniques that are commonly used to predict phase equilibria is clarified. The Gibbs-Duhem integration technique is a combination of a molecular simulation method and a mathematical scheme. Simulations are performed in a so-called *statistical-mechanical ensemble*. Short descriptions of the statistical-mechanical ensembles that are used in this work are given. The Gibbs-Duhem equation, expressed in appropriate independent variables, is derived from the fundamental thermodynamic equation. Furthermore, the mathematical scheme that is used to numerically integrate the Gibbs-Duhem equation is described.

Statistical mechanics provides the ensemble probability to observe a system in a particular configuration. The configuration of the system is given by the positions and orientations of all molecules. The generation of these configurations via Metropolis Monte Carlo sampling is commented upon in Chapter 3.

In a naïve approach, the configurations of the system are generated totally randomly. Thermodynamic properties are then obtained by averaging the instantaneous thermodynamic properties, corresponding to the generated configurations, weighted by their corresponding Boltzmann factor. However, the larger part of these generated configurations will have a negligible Boltzmann weight. Therefore, it is better to generate configurations in accordance with their corresponding Boltzmann weights. This method is called *Monte Carlo importance sampling*. As configurations are generated according to their probability of occurrence, thermodynamic properties can be obtained by simply averaging the instantaneous values of those properties.

In Chapter 3, the basic principles of Monte Carlo importance sampling and its applications in this work are described. Subsequently, the Metropolis Monte Carlo scheme [7, 8], which is used to efficiently generate configurations according to a particular probability distribution, is discussed. Chapter 3 is concluded with the description of several biased sampling schemes that are developed to enhance the sampling of configurations.

Gibbs-Duhem integration implies the numerical integration of a Clapeyron differential equation. It requires an initial coexistence point and a corresponding initial slope. Chapter 4 contains a discussion on the advantages and limitations of several simulation methods that can be used to predict the initial coexistence point.

In several publications preceding this work, and in one of the author's publications that emerged somewhere halfway his Ph.D. project, the initial slope was computed from the so-called *difference method* [14, 21-23]. Unfortunately, the difference method only provides accurate results in particular cases. For that reason, two other methods that are capable of providing accurate results are discussed as well. Moreover, the failure of the difference method is clarified.

In Chapter 5, the importance of a careful selection of a proper force field is stressed. Since it is not the objective of this work to optimize force-field parameters for the components of interest, existing force fields are selected based on a number of criteria. No mixture information is included in the force fields. In order to check whether use of pure-component force fields is sufficient to simulate mixture vapour-liquid equilibria in close agreement with experimental data, the prediction of pure-component phase equilibria should be close to the experimental pure-component phase behaviour in any case.

The conventional Gibbs-Duhem integration method [14] is tested on the vapour-liquid equilibria of a number of binary systems. The cross-component force-field parameters mainly determine whether the prediction is close to the experimental data or not. In Chapter 5, the shortcomings of the commonly used Lorentz-Berthelot combining rules are commented upon and the advantages and limitations of exploiting other combining rules are discussed.

In Chapter 6, an improved version of the conventional Gibbs-Duhem integration method is introduced. The conventional Gibbs-Duhem integration technique has a number of inefficiencies and shortcomings. For example, it employs only a small fraction of the generated configurations to predict the composition and densities at vapour-liquid phase coexistence. Moreover, the length of the simulations is not known beforehand. In order to improve the method, conventional Gibbs-Duhem integration is combined with multiple-histogram reweighting. The resulting method, which is used to simulate the vapour-liquid equilibria of a number of binary systems, is called *advanced Gibbs-Duhem integration*. The advantages of the improved method over the conventional Gibbs-Duhem integration method are discussed. Moreover, simulation results obtained with different combining rules are commented upon.

In Chapter 7, the extension of the advanced Gibbs-Duhem integration method to ternary mixtures is described. The application of the method to the vapour-liquid equilibria of a number of ternary systems is discussed. The extension to prediction of vapour-liquid equilibria of multicomponent mixtures is commented upon.

The Chapter Summary and Recommendations summarizes the objectives, results, major conclusions, and ideas about logical next steps in the improvement of the Gibbs-Duhem integration method.

1.5. References

1. King, C. J., *Separation processes*, Chemical Engineering Series, 2nd Ed., McGraw Hill, 1980
2. Sinnott, R. K., *Chemical Engineering Design*, Coulson & Richardson's Chemical Engineering, Volume 6, 2nd Ed., Pergamon Press, Oxford, Great Britain, 1993
3. Backhurst, J. R. and Harker, J. H., *Particle Technology and Separation Processes*, Coulson & Richardson's Chemical Engineering, Volume 2, 4th Ed., Pergamon Press, Oxford, Great Britain, 1991
4. Smith, J. M., Van Ness, H. C., and Abbott, M. M., *Introduction to chemical engineering thermodynamics*, Chemical Engineering Series, McGraw-Hill, Boston, 2001
5. Alder, B. J. and Wainwright, T. E., Studies in molecular dynamics. I. General Method, *J. Chem. Phys.*, 31(2) **1959** 459
6. Alder, B. J. and Wainwright, T. E., Studies in molecular dynamics. II. Behavior of a small number of elastic spheres, *J. Chem. Phys.*, 33(5) **1960** 1439
7. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.*, 21(6) **1953** 1087
8. Metropolis, N. and Ulam, S., The Monte Carlo Method, *J. Am. Stat. Assoc.*, 44(247) **1949** 335
9. Frenkel, D. and Smit, B., *Understanding molecular simulation; from algorithms to applications*, Computational Science Series, Volume 1, 2nd Ed., Academic press, London, 2002
10. Landau, D. P. and Binder, K., *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, Cambridge, 2000
11. Allen, M. P. and Tildesley, D. J., *Computer simulation of liquids*, Clarendon Press, Oxford, 1987
12. Kofke, D. A., Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line, *J. Chem. Phys.*, 98(5) **1993** 4149
13. Kofke, D. A., Gibbs-Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation, *Mol. Phys.*, 78(6) **1993** 1331
14. Mehta, M. and Kofke, D. A., Coexistence Diagrams of Mixtures by Molecular Simulation, *Chem. Eng. Sci.*, 49(16) **1994** 2633
15. Hitchcock, M. R. and Hall, C. K., Solid-liquid phase equilibrium for binary Lennard-Jones mixtures, *J. Chem. Phys.*, 110(23) **1999** 11433
16. Agrawal, R. and Kofke, D. A., Solid-Fluid Coexistence for Inverse-Power Potentials, *Phys. Rev. Lett.*, 74(1) **1995** 122
17. Lamm, M. H. and Hall, C. K., Equilibria between solid, liquid, and vapor phases in binary Lennard-Jones mixtures, *Fluid Phase Equil.*, 194-197 **2002** 197
18. Lamm, M. H. and Hall, C. K., Monte Carlo simulations of complete phase diagrams for binary Lennard-Jones mixtures, *Fluid Phase Equil.*, 182(1-2) **2001** 37
19. Agrawal, R., Thermodynamic and structural properties of model systems at solid-fluid coexistence. II. Melting and sublimation of the Lennard-Jones system, *Mol. Phys.*, 85(1) **1995** 43
20. Agrawal, R. and Kofke, D. A., Thermodynamic and structural properties of model systems at solid-fluid coexistence. I. FCC and BCC soft spheres, *Mol. Phys.*, 85(1) **1995** 23

-
21. Sindzingre, P., Ciccotti, G., Massobrio, C., and Frenkel, D., Partial enthalpies and related quantities in mixtures from computer simulation, *Chem. Phys. Lett.*, 136(1) **1987** 35
 22. Van 't Hof, A., de Leeuw, S. W., Hall, C. K., and Peters, C. J., Molecular simulation of binary vapour-liquid equilibria with components differing largely in volatility, *Mol. Phys.*, 102(3) **2004** 301
 23. Sindzingre, P., Massobrio, C., Ciccotti, G., and Frenkel, D., Calculation of partial enthalpies of an argon-krypton mixture by NpT molecular dynamics, *Chem. Phys.*, 129 **1989** 213

Chapter 2. Gibbs-Duhem integration and the semigrand-canonical ensemble

Several molecular simulation techniques that can be used to predict vapour-liquid equilibria are described in Section 2.1. In Section 2.2, the Gibbs-Duhem equation is derived from the most fundamental equation of thermodynamics by Legendre transformation. A Legendre transform that is of particular interest to this work is the semigrand form. Its derivation and the Gibbs-Duhem equation, expressed in semigrand variables, are given in Section 2.3. The key equation in this work is the Clapeyron differential equation. Its most general form, which is derived from the Gibbs-Duhem equations of coexisting phases, is given in Section 2.4. The ensembles of interest to this work are described in Section 2.5. Special attention is given to the semigrand-canonical ensemble. The numerical integration procedure of the Clapeyron equation is extensively described in Section 2.6. Conclusions are summarized in Section 2.7.

2.1. Gibbs-Duhem integration among other simulation methods

Simulation techniques that have been developed to predict phase equilibria can be roughly divided into *direct* methods, in which the properties of the coexisting phases are computed directly, and *indirect* methods in which the free energies or chemical potentials of the phases are computed in order to determine the thermodynamic state at which the phases are coexisting.

The most obvious way to simulate vapour-liquid equilibria is a direct approach that mimics experiment; phase coexistence is determined by studying the phase separation in a single simulation box [1, 2]. Such a simulation requires the formation of a vapour-liquid interface in the simulation box. Applications of direct interfacial simulations are given by Rowlinson and Widom [1] and by Gubbins [2]. As for a small system a large fraction of the molecules resides near the interface [3], energy management is governed by the interfacial energy. In practice this slows down the formation of two phases, provided that it will occur [2, 3]. An advantage of a direct interfacial simulation is that it may provide information on the structure of the interface and on the surface tension between the coexisting phases. However, such a simulation requires a large number of molecules and gives unreliable results near the vapour-liquid critical point [4].

On the other hand, indirect methods use independent simulation boxes for the phases that are to coexist. Clearly, these methods do not suffer from interfacial problems. In order to simulate phases that are at coexistence, one must provide the state point at which coexistence occurs. As this state point is not known from the beginning of the simulation, one has to search for it either by trial and error or with a semi-analytical scheme [5-14]. Thus, indirect methods require multiple simulations at sometimes uninteresting thermodynamic states. Several indirect methods require the computation of free energies or chemical potentials. Unfortunately, Monte Carlo importance sampling [See Chapter 3], applied to the computation of statistical-mechanical ensemble averages, does not directly provide thermal properties like free energies, but only mechanical properties like energies and pressures. However, much effort has been put in the development of techniques that provide thermal properties [3, 15]. Nevertheless, these techniques remain computationally demanding.

In 1987, Panagiotopoulos [16] developed the *quasi-direct* Gibbs ensemble method, which enables the computation of phase coexistence from a single simulation. The method has been further investigated and reviewed by Panagiotopoulos *et al.* [17-19] and by Smit *et al.* [20, 21]. Application of the Gibbs ensemble method implies simulation in two simulation boxes without interface. The two simulation boxes, that represent the coexisting phases, are coupled

mechanically and chemically by fixing the total volume and total number of molecules of the overall system. The absence of an interface shortens the needed simulation length and decreases the necessary system size compared to direct simulation methods. Extension of the Gibbs ensemble method to mixtures is rather straightforward [17]. A clear advantage of the method is the implicit coexistence condition; no computationally expensive special techniques to determine free energies or chemical potentials are needed. Although computationally cheap for systems comprising simple molecules at moderate densities, the method breaks down at higher densities, as the probability to insert a molecule into a dense phase is very small, especially for large molecules. Near vapour-liquid critical points, the Gibbs ensemble method breaks down because phase separation within one simulation box occurs due to a decreasing interfacial energy. Moreover, simulations in a small simulation volume with periodic boundaries cannot provide correct results when large density and/or composition fluctuations occur, as is the case near a vapour-liquid critical point.

Inspired by the Gibbs ensemble method, Kofke [22, 23] devised the *indirect* Gibbs-Duhem integration method. The method has the advantage of not being dependent of insertions and/or removals of molecules. This independence makes the method especially suitable for the computation of phase equilibria in which at least one of the phases is a solid one [24-32]. In Gibbs-Duhem integration, the phases are simulated in independent simulation boxes at identical thermodynamic states. The absence of an interface shortens the needed simulation length and decreases the necessary system size compared to direct simulation methods. Kofke originally developed the Gibbs-Duhem integration method for predicting pure-component phase equilibria [22, 23].

Gibbs-Duhem integration implies the numerical integration of a Clapeyron equation. A Clapeyron equation is derived from the Gibbs-Duhem equations of the coexisting phases. Thus, Gibbs-Duhem integration is not a proper name for the method. The original Clapeyron equation describes how the pressure changes with temperature in order for two phases to remain at coexistence; *i.e.* in order to maintain zero chemical-potential difference between the phases. Phase coexistence at a number of temperatures is determined by numerically integrating the Clapeyron equation. As the Clapeyron equation is a differential equation, an initial condition is necessary to predict a coexistence point at an incremented value of the independent integration variable, which equals temperature in the original Clapeyron equation.

Gibbs-Duhem integration is easily extended to binary or multicomponent mixtures [33] and multiphase systems [34] by deriving the appropriate Clapeyron equation. When multicomponent mixtures are simulated, the Gibbs-Duhem integration method can be conveniently combined with semigrand-canonical ensemble simulations [6].

Unfortunately, the Gibbs-Duhem integration method has no implicit coexistence condition like the Gibbs ensemble has. The coexistence condition is formed by the Gibbs-Duhem equations of the phases that are coupled through an initial coexistence point. Furthermore, if the initial point is no true coexistence point, Gibbs-Duhem integration will predict a series of quasi-coexistence points that deviate from true coexistence. However, in case a good estimate of the initial coexistence point is known, the method is appropriate for the prediction of complete phase diagrams [24, 28, 29].

Like the Gibbs ensemble method, the Gibbs-Duhem integration method breaks down in the vicinity of a critical point because the finite size and periodicity of the simulation box cannot describe long-range fluctuations. Moreover, the closer to the vapour-liquid critical point the more the physical properties of the liquid and vapour phase resemble each other, which often results in condensation or vaporization of one of the phases.

A few years ago, Kofke [35] reviewed the applications of the Gibbs-Duhem integration method. Although the most obvious applications of the method imply the numerical

integration of a conventional Clapeyron equation, it is equally well possible to integrate over a different type of field variable. Agrawal *et al.* [34], for example, computed three-phase equilibria as a function of the force-field parameters. Agrawal and Kofke [26, 27, 36] numerically integrated a Clapeyron equation that describes how the pressure of a soft-sphere system must change with the softness of the intermolecular potential at constant temperature for a solid and fluid phase to remain at the freezing point. Camp *et al.* [37] used Gibbs-Duhem integration to determine the isothermal isotropic-nematic liquid-crystal coexistence line for a system composed of hard prolate ellipsoids of revolution as a function of the ellipsoid elongation. Pandit and Kofke [38] studied how temperature, pressure and force-field parameters influence the ‘position’ of vapour-liquid azeotropic points.

In this work, the application of the Gibbs-Duhem integration method to vapour-liquid equilibria of binary and ternary systems is investigated. The simulations that are needed to apply Gibbs-Duhem integration are performed in the semigrand ensemble.

2.2. Legendre transforms and the Gibbs-Duhem equation

The equilibrium state of a single-phase, c – component system can be expressed as a function of $(c + 2)$ independent variables. The most fundamental form of such an expression is the *fundamental thermodynamic equation in its energy representation* [39]:

$$nU = f_U(nS, nV, \{N_i; i = 1 \dots c\}) \quad (2.1)$$

where U is the molar internal energy, V is the molar volume, S is the molar entropy, n is the total number of moles, and N_i is the number of moles of component i . The symbol f_U indicates that the internal energy is a function of the variables inside the parentheses. The differential form of the fundamental thermodynamic equation in its energy representation is given by:

$$d(nU) = Td(nS) - pd(nV) + \sum_{i=1}^c \mu_i dN_i \quad (2.2)$$

where p is the pressure and μ_i is the chemical potential of component i . Integration of Eq. (2.2) using Euler’s theorem [39] gives:

$$nU = TnS - pnV + \sum_{i=1}^c \mu_i N_i \quad (2.3)$$

Although Eqs. (2.2) and (2.3) contain all thermodynamic information of a single-phase c – component system in equilibrium, their independent variables are not always appropriate. However, the system can be expressed in other variables without loss of information by Legendre transformation of the fundamental thermodynamic equation.

If one of the independent intensive variables takes on only positive values, one can replace that variable by its inverse [40]. For convenience the temperature, which is always positive, is replaced by a reciprocal temperature, $\beta = 1/k_B T$, k_B being Boltzmann’s constant. The product of β and an energy is a dimensionless quantity. The use of dimensionless energies is

very useful as energies of a few hundreds of molecules are very small quantities. The product of energy and β is usually much larger.

The first Legendre transform of the molar internal energy, U , gives the molar Helmholtz energy, A :

$$\beta(nA) \equiv \beta(nU) - \frac{nS}{k_B} = f_A(\beta, nV, \{N_i; i=1 \dots c\}) \quad (2.4)$$

The corresponding differential form of the Helmholtz energy is:

$$d(\beta nA) = (nU)d\beta - \beta p d(nV) + \sum_{i=1}^c \beta \mu_i dN_i \quad (2.5)$$

Variables that are readily accessible in an experiment are pressure, temperature and composition. Solution thermodynamics is built upon the fundamental thermodynamic equation expressed in these independent variables. The thermodynamic potential corresponding to pressure, temperature and composition is the second Legendre transform of the molar internal energy, also called molar Gibbs energy, G :

$$\beta(nG) \equiv \beta(nU) - \frac{nS}{k_B} + \beta p(nV) = f_G(\beta, p, \{N_i; i=1 \dots c\}) \quad (2.6)$$

The differential form of the Gibbs energy is:

$$d(\beta nG) = (nH)d\beta + \beta(nV)dp + \sum_i \beta \mu_i dN_i \quad (2.7)$$

where $H \equiv U + pV$ is the molar enthalpy. The Gibbs energy, expressed as Eq. (2.7), plays the role of a generating function from which all other thermodynamic properties can be derived by simple mathematical operations [41]. Consider a generalized molar mixture property, M , expressed in the natural variables of the Gibbs energy:

$$\beta(nM) = f_M(\beta, p, \{N_i; i=1 \dots c\}) \quad (2.8)$$

The differential form of the generalized mixture property $\beta(nM)$ is given by:

$$d(\beta nM) = \left(\frac{\partial \beta nM}{\partial \beta} \right)_{p, N_i} d\beta + \left(\frac{\partial \beta nM}{\partial p} \right)_{\beta, N_i} dp + \sum_{i=1}^c \left(\frac{\partial \beta nM}{\partial N_i} \right)_{p, \beta, N_{j \neq i}} dN_i \quad (2.9)$$

The most right-hand partial derivative in Eq. (2.9) is called a partial molar property, which will be indicated with \bar{M}_i .

The total Legendre transform of the internal energy is expressed in intensive variables only. According to Eq. (2.3), these intensive variables are dependent. Thus there is no new type of energy related to the total Legendre transform:

$$\beta(nU) - \frac{nS}{k_B} + \beta p(nV) - \sum_{i=1}^c \beta \mu_i N_i = 0 \quad (2.10)$$

The corresponding differential form is the best-known form of the Gibbs-Duhem equation:

$$(nH)d\beta + \beta(nV)dp + \sum_{i=1}^c N_i d(\beta \mu_i) = 0 \quad (2.11)$$

The formalism of the Gibbs-Duhem equation is much more general than suggested by Eq. (2.11) and can be applied to a generalized molar mixture property, M , as well. In the special case of a constant number of moles, the Gibbs-Duhem equation for molar property M is given by:

$$\left(\frac{\partial \beta M}{\partial \beta} \right)_{p, x_i} d\beta + \left(\frac{\partial \beta M}{\partial p} \right)_{\beta, x_i} dp + \sum_{i=1}^c x_i d(\beta \bar{M}_i) = 0 \quad (2.12)$$

In order to simplify the notation, Eq. (2.12) is rewritten into a form introduced by Griffiths and Wheeler [40]. Griffiths and Wheeler divide variables into *fields* and their conjugate *densities*. Fields, in contrast with densities, take equal values in coexisting phases. The field-type intensive state of a thermodynamic phase with c components can be expressed in $(c+2)$ field variables, ϕ_i , with one dependent field, ϕ_0 , expressed as a function of the rest [40]. The dependent field is called the potential. Densities are defined by first order partial derivatives of the potential:

$$\rho'_j \equiv - \left(\frac{\partial \phi_0}{\partial \phi_j} \right)_{\phi_{i \neq j}} = \frac{\rho_j}{\rho_0} \quad (2.13)$$

where ϕ_j is the field conjugate to density ρ'_j . In this notation, the Gibbs-Duhem equation (2.12) can be written as:

$$d\phi_0 = - \sum_{j=1}^{c+1} \frac{\rho_j}{\rho_0} d\phi_j \quad (2.14)$$

Equation (2.14) will be used in Section 2.4 to derive a general form of the Clapeyron equation.

2.3. Semigrand variables

A Legendre transform that is of particular interest to this work is the semigrand transform. In order to clarify its position among other well-known Legendre transforms, the derivation of the semigrand Legendre transform is started from the fundamental thermodynamic equation of a c -component mixture expressed in the natural variables of the Helmholtz energy [See Eq. (2.5)]. Legendre transformation of all the $(\beta\mu_i \leftrightarrow N_i)$ terms in the expression for the Helmholtz energy leads to the Hill energy, $\beta(nL) \equiv -\beta p(nV)$, whereas Legendre transformation of only the $(\beta p \leftrightarrow nV)$ term results in the Gibbs energy, Eq. (2.6).

Transformation of only a subset of the $(\beta\mu_i \leftrightarrow N_i)$ terms gives a hybrid thermodynamic potential that is somewhere between the Helmholtz energy and the Hill energy. The semigrand transform is obtained from Legendre transformation of a different form of Eq. (2.5) in which some of the $(\beta\mu_i \leftrightarrow N_i)$ terms are grouped together:

$$d(\beta nA) = (nU)d\beta - \beta p d(nV) + \beta\mu_1 dn + \beta \sum_{i=2}^c \Delta\mu_{i1} dN_i \quad (2.15)$$

where $\Delta\mu_{i1} \equiv \mu_i - \mu_1$. Legendre transformation of the $(\beta\Delta\mu_{i1} \leftrightarrow N_i)$ terms and the $(\beta p \leftrightarrow nV)$ term results in the molar isobaric semigrand energy, Y :

$$d(\beta nY) = (nH)d\beta + \beta(nV)dp + \beta\mu_1 dn - \sum_{i=2}^c N_i d(\beta\Delta\mu_{i1}) \quad (2.16)$$

This choice of independent variables was introduced by Griffiths and Wheeler [40]. From now on, the total number of moles is fixed. Furthermore, only binary mixtures are considered in this chapter. The differential form of the isobaric semigrand energy can be rewritten in terms of fugacity fractions that are, in contrast to chemical potentials, bounded between zero and one [6, 33, 42]:

$$nN_{\text{Av}} d \ln \left\{ \frac{\hat{f}_1 + \hat{f}_2}{p^{\text{ref}}} \right\} = (nH')d\beta + \beta(nV)dp - N_{\text{Av}} \frac{N_2 - n\xi_2}{\xi_2(1 - \xi_2)} d\xi_2 \quad (2.17)$$

where H' is the configurational molar enthalpy (molar enthalpy minus the purely temperature-dependent ideal gas contribution), \hat{f}_i is the fugacity of component i in the mixture, $\xi_2 = \hat{f}_2 / (\hat{f}_1 + \hat{f}_2)$ is the fugacity fraction of component two, N_{Av} is Avogadro's number, and p^{ref} is the reference pressure at which the purely temperature-dependent ideal-gas contributions are evaluated. It is convenient to remove the system size from Eq. (2.17). After division of both sides of Eq. (2.17) by nN_{Av} one gets:

$$d \ln \left\{ \frac{\hat{f}_1 + \hat{f}_2}{p^{\text{ref}}} \right\} = h' d\beta + \beta v dp - \frac{x_2 - \xi_2}{\xi_2(1 - \xi_2)} d\xi_2 \quad (2.18)$$

The variables h' and v are now the configurational enthalpy *per molecule* and the volume *per molecule* respectively. In what follows, N_i represents the number of molecules of component i .

As the conjugate density of ξ_2 in Eq. (2.18) contains an additive constant, the following form of Eq. (2.18) is possible as well:

$$d \ln \left\{ \frac{\hat{f}_1}{p^{\text{ref}}} \right\} = h' d\beta + \beta v d p - \frac{x_2}{\xi_2 (1 - \xi_2)} d\xi_2 \quad (2.19)$$

Notice that Eqs. (2.18) and (2.19) relate only field variables. They are Gibbs-Duhem equations expressed in semigrand variables.

2.4. Clapeyron equations

Clapeyron equations describe changes in the field variables that maintain phase coexistence. They are derived from the Gibbs-Duhem equations of the coexisting phases. Below, a derivation is given for the most general form of the Clapeyron equation, starting from Eq. (2.14). The number of components and the number of coexisting phases are c and π respectively. In order to describe the change of field variables at phase coexistence, one needs to specify one field variable that is kept constant between the coexisting phases. This so-called *hidden field variable* is given a subscript '0'. At constant ϕ_0 , Eq. (2.14) can be rewritten into the following form:

$$\frac{\rho_1}{\rho_0} + \sum_{j=2}^{c+1} \frac{\rho_j}{\rho_0} \left(\frac{\partial \phi_j}{\partial \phi_1} \right)_{\phi_0} = 0 \quad (2.20)$$

Subtraction of Eq. (2.20) for phase δ , from an identical equation for phase γ gives:

$$\left(\frac{\rho_1^\gamma}{\rho_0^\gamma} - \frac{\rho_1^\delta}{\rho_0^\delta} \right) + \sum_{j=2}^{c+1} \left(\frac{\rho_j^\gamma}{\rho_0^\gamma} - \frac{\rho_j^\delta}{\rho_0^\delta} \right) \left(\frac{\partial \phi_j}{\partial \phi_1} \right)_{\phi_0} = 0 \quad (2.21)$$

Relation (2.21) can be presented for every pair of the π coexisting phases. A shorthand notation for Eq. (2.21) is:

$$\Delta_1^{\gamma\delta} + \sum_{j=2}^{c+1} \Delta_j^{\gamma\delta} \left(\frac{\partial \phi_j}{\partial \phi_1} \right)_{\phi_0} = 0 \quad (2.22)$$

The complete system of $(\pi - 1)$ independent equations is given by:

$$\begin{bmatrix} -\Delta_1^{12} \\ \vdots \\ -\Delta_1^{1(\pi-1)} \end{bmatrix} = \begin{bmatrix} \sum_{j=2}^{c+1} \Delta_j^{12} \left(\frac{\partial \phi_j}{\partial \phi_1} \right)_{\phi_0} \\ \vdots \\ \sum_{j=2}^{c+1} \Delta_j^{1(\pi-1)} \left(\frac{\partial \phi_j}{\partial \phi_1} \right)_{\phi_0} \end{bmatrix} = \begin{bmatrix} \Delta_2^{12} & \dots & \Delta_{c+1}^{12} \\ \vdots & \ddots & \vdots \\ \Delta_2^{1(\pi-1)} & \dots & \Delta_{c+1}^{1(\pi-1)} \end{bmatrix} \begin{bmatrix} \left(\frac{\partial \phi_2}{\partial \phi_1} \right)_{\phi_0} \\ \vdots \\ \left(\frac{\partial \phi_{c+1}}{\partial \phi_1} \right)_{\phi_0} \end{bmatrix} \quad (2.23)$$

The solution to this set of equations is the most general form of the Clapeyron equation:

$$\begin{bmatrix} \left(\frac{\partial \phi_2}{\partial \phi_1} \right)_{\phi_0} \\ \vdots \\ \left(\frac{\partial \phi_{c+1}}{\partial \phi_1} \right)_{\phi_0} \end{bmatrix} = \begin{bmatrix} \Delta_2^{12} & \dots & \Delta_{c+1}^{12} \\ \vdots & \ddots & \vdots \\ \Delta_2^{1(\pi-1)} & \dots & \Delta_{c+1}^{1(\pi-1)} \end{bmatrix}^{-1} \begin{bmatrix} -\Delta_1^{12} \\ \vdots \\ -\Delta_1^{1(\pi-1)} \end{bmatrix} \quad (2.24)$$

Equation (2.24), along with the semigrand Gibbs-Duhem equation (2.19), has been used to generate a number of Clapeyron equations, which are presented in Appendix A.

All Gibbs-Duhem integrations in this work are performed at constant temperature. The reason is the presence of a large amount of isothermal $p - x - y$ data in literature that can be used to compare simulation results with. The Clapeyron equation, expressed in semigrand variables, for isothermal vapour (V) - liquid (L) equilibria in a binary mixture is given by:

$$\left(\frac{\partial p}{\partial \xi_2} \right)_{\beta, \sigma} = \frac{\Delta x_2^{\text{LV}}}{\xi_2 (1 - \xi_2) \beta \Delta v^{\text{LV}}} \quad (2.25)$$

where the subscript σ indicates that the derivative is taken along the saturation line.

2.5. Description of several ensembles

In this section, the ensembles relevant to this work are discussed: the isobaric-isothermal ensemble, the Gibbs ensemble, and the semigrand-canonical ensemble.

Most semi-classical statistical-mechanical partition functions that are presented in literature are derived for systems consisting of spherical molecules. However, the molecules used in this work have spatial configurations and, thus, consist of a number of interaction sites. Several types of molecules that are considered are fully rigid while other types have partial intramolecular flexibility. In order to keep the statistical-mechanical relations general, derivations are presented for molecules with partial intramolecular flexibility. Throughout this thesis, the molecules have fixed bond lengths. The internal degrees of freedom consist of rotation about torsion angles and bond-angle bending. Branched or ring structures are not considered.

The interaction energy is separated into three contributions; $\mathcal{U}^{\text{intra}}$ comprises the contributions from bond-angle bending and rotations about torsion angles, \mathcal{U}^{ext} is the sum of the intermolecular Lennard-Jones (LJ) and Coulombic (C) interaction energy, and $\mathcal{U}^{\text{LJint+Cint}}$ comprises intramolecular Lennard-Jones and Coulombic interactions. The intramolecular Lennard-Jones and Coulombic interactions are only non-zero for interactions between sites on partially flexible molecules that are more than three chemical bonds apart. The sum of the different contributions to the interaction energy is indicated with $\mathcal{U}^{\text{total}}$.

In case molecules are rigid, $\mathcal{U}^{\text{intra}}$ and $\mathcal{U}^{\text{LJint+Cint}}$ can be put to zero. Although the intramolecular Lennard-Jones and Coulombic interactions are zero in this work, they are nevertheless considered in order to keep the resulting relations general.

A fully flexible molecule of type j has $3n_j$ degrees of freedom, where n_j is the number of interaction sites in the molecule. These degrees of freedom comprise the Cartesian coordinates of all interaction sites. However, the configurations of molecules with intramolecular flexibility and/or intramolecular constraints are most easily described in generalized coordinates. Transformation of Cartesian to generalized coordinates introduces a Jacobian in the partition function. All $n_j - 1$ bond lengths are fixed. Consequently, component j has $2n_j + 1$ internal degrees of freedom. It has $2n_j - 2$ angle-dependent generalized coordinates and 3 Cartesian coordinates defining the overall translation of the molecule.

2.5.1. Isobaric-isothermal ensemble

The isobaric-isothermal ensemble ($Np\beta$ -ensemble) has the number of molecules, N , pressure, p , and reciprocal temperature, β , as its independent variables. The ensemble was introduced by Wood [43] and later extended by McDonald [44]. The variables conjugate to the independent ones are allowed to fluctuate. Their values can be computed from ensemble averages. It is obvious from the choice of independent variables that the Gibbs energy is the thermodynamic potential corresponding to the $Np\beta$ -ensemble [See Eq. (2.7)].

One is referred to the books of Frenkel and Smit [3] and of Allen and Tildesley [15] for an extensive explanation of this ensemble, including the acceptance criteria of Monte Carlo trial moves. The semi-classical isothermal-isobaric partition function for a binary mixture is given by:

$$\Delta_{N_1 N_2 p \beta}^* = \frac{q_1^{N_1} q_2^{N_2} N}{N_1! N_2!} \int dV \frac{V^N}{V} \exp\{-\beta p V\} \cdot \int \prod_{j=1}^2 \prod_{i=1}^{N_j} J_{ij} \left(\mathbf{q}_i^{2n_j-2} \right) d\mathbf{q}_i^{2n_j-2} d\mathbf{s}^{3N} \exp\left\{-\beta \mathcal{U}^{\text{total}} \left(\mathbf{s}^{3N}, \mathbf{q}^{2(N_1 n_1 + N_2 n_2 - N)} \right)\right\} \quad (2.26)$$

where \mathbf{s}^{3N} represents the vector of overall translational coordinates of all N molecules scaled by the length of the simulation box, $\mathbf{q}_i^{2n_j-2}$ represents the configuration vector of molecule i which is of type j , J_{ij} is the Jacobian of transformation of molecule i which is of type j , and q_i is the kinetic contribution of a molecule of type i to the partition function.

There is one subtle difference between the partition function used in this work and the one given in abovementioned books; equation (2.26) contains a factor NV^{-1} , instead of an undetermined volume scale that keeps the partition function dimensionless [45]. The isobaric-isothermal partition function (2.26) can be derived from the more general isobaric-isothermal

partition function as derived by Corti [46, 47] and Corti and Soto-Campos [48] who use a *shell molecule* to explicitly define the volume. Without this shell molecule, the partition function redundantly counts volume states. The volume scale in the more general isobaric-isothermal partition function of Corti [46, 47] and Corti and Soto-Campos [48] depends on the interactions of the system with the surroundings as well as on the shape of the simulation box. In case the system is homogeneous, interactions with the surroundings are negligible, and periodic boundary conditions are used, the general partition function of Corti [46, 47] and Corti and Soto-Campos [48] reduces to Eq. (2.26) [45, 49]. The asterisk in Eq. (2.26) is meant to distinguish the corrected partition function from the conventional isothermal-isobaric partition function.

The conventional isobaric-isothermal partition function and the more general isobaric-isothermal partition function are identical in the thermodynamic limit. Simulations performed in this work showed that for a system size of 300 molecules, both partition functions provide results that match within the statistical uncertainties.

2.5.2. Gibbs ensemble

The Gibbs ensemble was introduced by Panagiotopoulos in 1987 [16]. The method has been investigated and refined by Panagiotopoulos *et al.* and Smit *et al.* [17, 20, 21]. The Gibbs ensemble and its applications have been reviewed in refs. [18, 19, 50, 51].

In the $NV\beta$ –Gibbs ensemble, two phases without physical contact are allowed to exchange volume and matter under the constraint of constant total volume and constant total number of molecules. Mechanical and chemical equilibrium are maintained by three types of Monte Carlo moves: translations/rotations of molecules, exchanges of volume between the phases, and exchanges of molecules between the phases. For further details concerning the Gibbs Ensemble, including the acceptance criteria of the Monte Carlo trial moves, one is referred to the original articles by Panagiotopoulos [16, 17, 19, 52], Smit *et al.* [20, 21], and the book of Frenkel and Smit [3].

2.5.3. Semigrand-canonical ensemble

The semigrand-canonical ensemble is a hybrid ensemble somewhere in between the canonical ensemble (constant N_1, N_2, V, β) and grand-canonical ensemble (constant μ_1, μ_2, V, β). The independent variables of the isobaric semigrand-canonical ensemble are total number of molecules, N , reciprocal temperature, β , pressure, p , and fugacity fractions, $\{\xi_i; i = 2 \dots c\}$. The densities that are conjugate to these imposed field variables fluctuate during the course of the simulation. In order to perform a semigrand-canonical simulation, it is necessary to sample from all configurations and compositions at constant number of molecules. Thus, the different Monte Carlo trial moves are translations and rotations of molecules, identity changes of molecules, and volume changes of the simulation box.

The constant number of molecules is both an advantage and a disadvantage; on the one hand there is no need for insertions/removals of molecules of different components, but on the other hand, identity changes have to be performed on molecules in order to satisfy the imposed fugacity fractions. These identity changes can be very efficient when molecules of different components are similar [6]. However, in case the components of interest have very different size, shape and/or charge distributions the acceptance probability of the identity changes can be dramatically low [See Chapter 3, Section 3.5.2]. As already stated in Section 2.1, the absence of insertions/removals of molecules makes the ensemble suitable for studying solid phases [24–26, 28–32, 53].

Following Frenkel and Smit [3], the isobaric semi-classical semigrand-canonical partition function for a binary mixture can be written as follows:

$$\mathcal{Y}'_{Np\beta\Delta\mu_{21}} = \int dV \frac{(Vq_1)^N \exp\{-\beta pV\}}{V(N-1)!} \sum_{\text{iden}} \left(\frac{q_2}{q_1}\right)^{N_2} \exp[\beta N_2 \Delta\mu_{21}] \cdot \int \prod_{j=1}^2 \prod_{i=1}^{N_j} J_{ij}(\mathbf{q}_i^{2n_j-2}) d\mathbf{q}_i^{2n_j-2} ds^{3N} \exp\left\{-\beta \mathcal{U}^{\text{total}}(\mathbf{s}^{3N}, \mathbf{q}^{2(N_1 n_1 + N_2 n_2 - N)})\right\} \quad (2.27)$$

where $\mathcal{Y}'_{Np\beta\Delta\mu_{21}}$ is the thermodynamic potential belonging to the isobaric semigrand-canonical ensemble. The summation indicated by *iden* represents a sum over all possible identities of all molecules [3, 6]. This partition function differs from the ones given by Frenkel and Smit [3] and Kofke and Glandt [6] by a factor NV^{-1} for reasons pointed out in Section 2.5.1. The partition functions with or without the factor NV^{-1} become equal in the thermodynamic limit [46-48]. The fugacity fraction can be introduced in Eq. (2.27) by separating the chemical potentials into a purely temperature-dependent part and a mixture fugacity:

$$\beta\mu_i = \beta\mu_i^{\text{ig}}(\beta; p^{\text{ref}}, x_i^{\text{ref}} = 1) + \ln\left(\frac{\hat{f}_i}{p^{\text{ref}}}\right) \quad (2.28)$$

where \hat{f}_i is the fugacity of component i in the mixture and p^{ref} is a reference pressure. The purely temperature-dependent contribution to the chemical potential can be worked out as follows:

$$\begin{aligned} \beta\mu_i^{\text{ig}}(\beta; p^{\text{ref}}, x_i^{\text{ref}} = 1) &= -\ln\left(\frac{q_i}{\beta p^{\text{ref}}} \int J_i(\mathbf{q}^{2n_i-2}) d\mathbf{q}^{2n_i-2} ds^3 \exp\left[-\beta \mathcal{U}^{\text{intra}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}\right]\right) \\ &= -\ln\left(\frac{q_i C_i(\beta)}{\beta p^{\text{ref}}}\right) \end{aligned} \quad (2.29)$$

where the symbol C_i is introduced to simplify the notation. Using Eq. (2.29), the chemical-potential difference in equation (2.27) can be expressed as:

$$\begin{aligned} \beta N_2 \Delta\mu_{21} &= \beta N_2 \left\{ \mu_2^{\text{ig}}(\beta; p^{\text{ref}}, x_2^{\text{ref}} = 1) - \mu_1^{\text{ig}}(\beta; p^{\text{ref}}, x_1^{\text{ref}} = 1) \right\} + N_2 \ln\left\{ \frac{\hat{f}_2}{\hat{f}_1} \right\} \\ &= \ln\left(\frac{q_1 C_1 \xi_2}{q_2 C_2 \xi_1}\right)^{N_2} \end{aligned} \quad (2.30)$$

Substitution of Eq. (2.30) into Eq. (2.27) gives:

$$\mathcal{Y}'_{Np\beta\xi_2}^* = \int dV \frac{(Vq_1)^N \exp\{-\beta pV\}}{V(N-1)!} \sum_{\text{idem}} \left(\frac{C_1 \xi_2}{C_2 \xi_1} \right)^{N_2}. \quad (2.31)$$

$$\int \prod_{j=1}^2 \prod_{i=1}^{N_j} J_{ij} \left(\mathbf{q}_i^{2n_j-2} \right) d\mathbf{q}_i^{2n_j-2} ds^{3N} \exp\left\{-\beta\mathcal{U}^{\text{total}}\left(\mathbf{s}^{3N}, \mathbf{q}^{2(N_1n_1+N_2n_2-N)}\right)\right\}$$

The bridge equation that links macroscopic thermodynamics with this partition function is [3, 6]:

$$\beta\mu_1 N = -\ln \mathcal{Y}'_{Np\beta\xi_2}^* \quad (2.32)$$

Instead of Eq. (2.31), in this work a partition function is used that is related to the configurational semigrand energy. This difference does not influence the simulation results; it only simplifies the notation and makes the semigrand-canonical partition function consistent with Eq. (2.19). The shift between semigrand and configurational-semigrand energies equals the purely temperature-dependent ideal-gas chemical potential of component one, evaluated at the reference pressure p^{ref} . Subtraction of this shift from Eq. (2.32) gives [See Eq. (2.29)]:

$$N \ln \left(\frac{\hat{f}_1}{p^{\text{ref}}} \right) = \beta\mu_1 N - \beta\mu_1^{\text{ig}}(\beta; p^{\text{ref}}) N$$

$$= -\ln \mathcal{Y}'_{Np\beta\xi_2}^* + N \ln \left(\frac{q_1 C_1}{\beta p^{\text{ref}}} \right) \quad (2.33)$$

$$= -\ln \mathcal{Y}_{Np\beta\xi_2}^*$$

The resulting configurational semi-classical isobaric semigrand partition function is given by:

$$\mathcal{Y}_{Np\beta\xi_2}^* = \int dV \left(\frac{\beta V p^{\text{ref}}}{C_1} \right)^N \frac{\exp\{-\beta pV\}}{V(N-1)!} \sum_{\text{idem}} \left(\frac{C_1 \xi_2}{C_2 \xi_1} \right)^{N_2}. \quad (2.34)$$

$$\int \prod_{j=1}^2 \prod_{i=1}^{N_j} J_{ij} \left(\mathbf{q}_i^{2n_j-2} \right) d\mathbf{q}_i^{2n_j-2} ds^{3N} \exp\left\{-\beta\mathcal{U}^{\text{total}}\left(\mathbf{s}^{3N}, \mathbf{q}^{2(N_1n_1+N_2n_2-N)}\right)\right\}$$

2.5.4. Sampling intramolecular configurations

The mathematical principle behind the sampling method that is presented in this section will be discussed in Chapter 3. Nevertheless, sampling of intramolecular configurations is described here as it simplifies the notation of the partition functions considerably.

Generating intramolecular configurations in a totally random fashion is far from efficient. The larger part of the generated configurations will have a high intramolecular energy with corresponding negligible Boltzmann weight. It is much more efficient to generate intramolecular configurations $\{\mathbf{q}_j^{2n_j-2}, \mathbf{s}_j^3\}$ according to the intramolecular Boltzmann

distribution Ψ_j . The probability to generate an intramolecular configuration between Ψ_j and $\Psi_j + d\Psi_j$ is:

$$\begin{aligned} d\Psi_j &= \frac{\exp\left\{-\beta\mathcal{U}^{\text{intra}}\left(\mathbf{q}^{2n_j-2}\right)\right\} J_j d\mathbf{q}^{2n_j-2} ds^3}{\int J_j d\mathbf{q}^{2n_j-2} ds^3 \exp\left\{-\beta\mathcal{U}^{\text{intra}}\left(\mathbf{q}^{2n_j-2}\right)\right\}} \\ &= \frac{\exp\left\{-\beta\mathcal{U}^{\text{intra}}\left(\mathbf{q}^{2n_j-2}\right)\right\} J_j d\mathbf{q}^{2n_j-2} ds^3}{z_j} \end{aligned} \quad (2.35)$$

The generation of new configurations according to a specific probability distribution can be performed using a rejection method [3, 15]. Notice that the intramolecular Boltzmann distribution does not depend on the intramolecular Lennard-Jones and Coulombic energies. The reason is that it is computationally very expensive to include the intramolecular Lennard-Jones and Coulombic energies.

Insertion of Eq. (2.35) into Eq. (2.34) gives the following semi-classical isobaric semigrand-canonical partition function:

$$\begin{aligned} \mathcal{Y}_{Np\beta\xi_2} &= \int dV \left(\frac{z_1 V \beta p^{\text{ref}}}{C_1} \right)^N \frac{\exp\{-\beta p V\}}{V(N-1)!} \sum_{\text{idem}} \left(\frac{z_2 C_1 \xi_2}{z_1 C_2 \xi_1} \right)^{N_2} \\ &\quad \int d\Psi^{N_1} d\Psi^{N_2} \exp\left\{-\beta\mathcal{U}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^{N_1}, \Psi^{N_2})\right\} \\ &= \int dV \left(\frac{V \beta p^{\text{ref}}}{\langle \mathcal{W}_{\text{IG},1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle} \right)^N \frac{\exp\{-\beta p V\}}{V(N-1)!} \sum_{\text{idem}} \left(\frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle \xi_2}{\langle \mathcal{W}_{\text{IG},2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle \xi_1} \right)^{N_2} \\ &\quad \int d\Psi^{N_1} d\Psi^{N_2} \exp\left\{-\beta\mathcal{U}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^{N_1}, \Psi^{N_2})\right\} \end{aligned} \quad (2.36)$$

where $\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle$ is the normalized Rosenbluth factor of an isolated chain of component i [3, 54]. This normalized Rosenbluth factor equals the Boltzmann factor, with contributions from the intramolecular Lennard-Jones and Coulombic energies, averaged over the generated ideal-chain conformations:

$$\begin{aligned} \langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle &= \frac{\int J_i(\mathbf{q}^{2n_i-2}) d\mathbf{q}^{2n_i-2} ds^3 \exp\left[-\beta\mathcal{U}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}}\right] \exp\left[-\beta\mathcal{U}^{\text{intra}}\right]}{\int J_i(\mathbf{q}^{2n_i-2}) d\mathbf{q}^{2n_i-2} ds^3 \exp\left[-\beta\mathcal{U}^{\text{intra}}\right]} \\ &= \int \exp\left[-\beta\mathcal{U}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi_i)\right] d\Psi_i \end{aligned} \quad (2.37)$$

2.5.5. Acceptance criteria for semigrand-canonical MC trial moves

The acceptance criteria for translations/rotations of molecules and volume changes in the semigrand ensemble are identical to their counterparts in the isobaric-isothermal ensemble [3]. A trial move that is typical for the semigrand-canonical ensemble is the identity change. Having found a useful version of the semigrand-canonical partition function, the acceptance criterion for this trial move can be derived. The probability to select a random molecule whose identity is to be changed is given by $1/N$. Assume that the selected molecule is of type i and that its identity is changed from type i to type j . The probabilities to find the system in the ‘old’ and the ‘new’ configuration are derived from the semi-classical partition function (2.36):

$$\pi^{\text{old}} = \frac{\left(\frac{\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_j}}{\langle \mathcal{W}_{\text{IG},j}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_i}} \right)^{N_j} \exp\{-\beta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}\} d\Psi^{N_i} d\Psi^{N_j}}{\text{Constant}} \quad (2.38)$$

$$\pi^{\text{new}} = \frac{\left(\frac{\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_j}}{\langle \mathcal{W}_{\text{IG},j}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_i}} \right)^{N_j+1} \exp\{-\beta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}\} d\Psi^{N_i-1} d\Psi^{N_j+1}}{\text{Constant}}$$

The probability to create a configuration for a molecule of type j is $d\Psi_j$. The principle of detailed balance gives [15]:

$$\left(\frac{\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_j}}{\langle \mathcal{W}_{\text{IG},j}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_i}} \right)^{N_j} \exp\{-\beta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}\} d\Psi^{N_i} d\Psi^{N_j} \frac{d\Psi_j}{N} \min(1, \chi) =$$

$$\left(\frac{\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_j}}{\langle \mathcal{W}_{\text{IG},j}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_i}} \right)^{N_j+1} \exp\{-\beta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}\} d\Psi^{N_i-1} d\Psi^{N_j+1} \frac{d\Psi_i}{N} \min\left(1, \frac{1}{\chi}\right) \quad (2.39)$$

Rearrangement gives the acceptance criterion:

$$\min \left(1, \exp \left[-\beta\Delta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}} + m_{ij} \ln \left\{ \frac{\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_j}}{\langle \mathcal{W}_{\text{IG},j}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle_{\xi_i}} \right\} \right] \right) \quad (2.40)$$

where $\Delta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}$ is the difference in the interaction energy resulting from the identity change. The parameter m_{ij} equals $+1$ for identity changes $i \rightarrow j$ and -1 for identity changes $j \rightarrow i$. A trial move is accepted when the pseudo-Boltzmann factor in Eq. (2.40) is larger than a random number, sampled uniformly between zero and one.

2.6. Numerical integration of the Clapeyron equation

2.6.1. Predictor–Corrector schemes

The Clapeyron equations used in this work can be classified as first-order non-linear differential equations. The integrand of a Clapeyron equation, F , depends on both the dependent and independent variables:

$$\left(\frac{\partial p}{\partial \xi_2}\right)_{\beta,\sigma} = \frac{(x_2^\gamma - x_2^\delta)}{\xi_2(1-\xi_2)\beta(v^\gamma - v^\delta)} = F(p, \xi_2) \quad (2.41)$$

Numerical integration combined with molecular simulation is a rather unusual. In conventional numerical integration of differential equations, the computation of the integrand would yield the exact value of $F(p, \xi_2)$ at the imposed pressure and fugacity fraction.

In Gibbs-Duhem integrations, the integrand is determined with molecular simulation in the semigrand-canonical ensemble, and the longer the simulation proceeds, the better the estimate for the integrand gets:

$$\left(\frac{\partial p}{\partial \xi_2}\right)_{\beta,\sigma} = \frac{1}{\xi_2(1-\xi_2)\beta} \frac{\langle x_2 \rangle_{Np\beta\xi_2}^\gamma - \langle x_2 \rangle_{Np\beta\xi_2}^\delta}{\langle v \rangle_{Np\beta\xi_2}^\gamma - \langle v \rangle_{Np\beta\xi_2}^\delta} \quad (2.42)$$

where $\langle \dots \rangle_{Np\beta\xi_2}^\gamma$ represents an ensemble average in phase γ , obtained from an isobaric semigrand-canonical simulation.

Since molecular simulation is computationally expensive, it is important to choose a numerical integration technique that requires *as little integrand evaluations as possible*. Kofke [22, 23] and Mehta and Kofke [33] suggested the use of *predictor-corrector equations*. Predictor-corrector equations are derived by approximating the integrand of the differential equation by an interpolating polynomial P^m of a specific order m [55]:

$$\begin{aligned} p([\xi_2]_i) &= p([\xi_2]_{i-1}) + \int_{[\xi_2]_{i-1}}^{[\xi_2]_i} d\xi_2 F[p(\xi_2), \xi_2] \\ &\approx p^{\text{est}}([\xi_2]_{i-1}) + \int_{[\xi_2]_{i-1}}^{[\xi_2]_i} d\xi_2 P^m(\xi_2) \end{aligned} \quad (2.43)$$

where p^{est} is a previously obtained saturation pressure. Inclusion of the integrand $F\{p([\xi_2]_i), [\xi_2]_i\}$ at the currently estimated coexistence point in the polynomial gives an implicit or *corrector* equation while incorporation of only previously determined integrands gives an explicit or *predictor* equation. Predictor-corrector equations may differ in their order and step-size adjustment. In the original implementation of the Gibbs-Duhem integration method [22, 23, 33], a fixed step-size predictor-corrector scheme was used. The order of the predictor-corrector equations increased in successive integration steps up to order three. The predictor-corrector formulae used by Kofke [22] and by Mehta and Kofke [33] are given in Table 2.1 and Table 2.2.

Escobedo and de Pablo [56] used a lower-order predictor-corrector scheme in their work. In general lower-order schemes are more stable and are easier to implement with a variable step-size. They used two versions of second-order variable-step-size corrector equations. Their predictor equation was a simple Euler equation in all cases. For their applications, this set of predictor-corrector equations proved to be successful. However, if the computation of the integrand is hard and consequently takes a long simulation, it is worthwhile to use higher-order predictor-corrector equations in order to reduce the number of corrector iterations [35]. Moreover, there is another reason to use higher-order predictor equations; the stability of the phases close to a critical point heavily depends on a reliable prediction of the new coexistence point; in the vicinity of a critical point, the binodal and spinodal curves are very close. Higher-order predictor equations generally provide more accurate predictions, provided that the integration step-size is small.

In this work, higher-order *variable step-size predictor-corrector equations* are used. The disadvantage of using different equations for successive steps [56] is of little relevance as the relations are far from complex and need to be programmed only once. The derivation of these equations is given in Appendix B. The higher-order variable step-size predictor-corrector equations are given in Table 2.3 and Table 2.4.

To start the numerical integration of a first-order differential equation, an initial point and an initial value for the integrand are needed. Different simulation techniques that can be used to predict the initial point and the initial integrand are described in Chapter 4. The numerical integration of Clapeyron equation (2.41) consists of the following steps:

1. Compute the saturation pressure, p_0 , at the imposed fugacity fraction $[\xi_2]_0$ (initial point) along with the value of the integrand $F\{p_0, [\xi_2]_0\}$ at the conditions of interest.
2. Increment the fugacity fraction $[\xi_2]_i = [\xi_2]_{i-1} + [\Delta\xi_2]_i$, and estimate the corresponding saturation pressure p_i with the predictor equation belonging to this integration step.
3. Collect samples of the relevant variables in two isobaric semigrand-canonical simulations (liquid and vapour) and compute the integrand $F\{p_i, [\xi_2]_i\}$ at the conditions of interest from ensemble averages.
4. Apply the corrector equation to update the value of p_i , using the integrand computed in step 3.
5. Repeat steps 2. to 4. until the specified convergence criterion is met.
6. Perform a long simulation at the estimated state point (coexistence point) so as to obtain the final value of the integrand and other relevant thermodynamic properties.
7. Go to step 2.

From this scheme it is clear that the pressure converges in an analytical sense while the other thermodynamic properties converge in a statistical sense [22].

Table 2.1: Predictor equations at different integration steps, used in refs. [22, 23, 33]

Integr. step	Predictor equation	Name
1	$p_1 = p_0 + \Delta\xi_2 F_0$	Euler's method
2	$p_2 = p_0 + 2\Delta\xi_2 F_1$	Midpoint method
3	$p_3 = p_1 + 2\Delta\xi_2 F_2$	Midpoint method
i	$p_i = p_{i-1} + \frac{\Delta\xi_2}{24} (55F_{i-1} - 59F_{i-2} + 37F_{i-3} - 9F_{i-4})$	Fourth-order explicit Adams-Bashforth

Table 2.2: Corrector equations at different integration steps, used in refs. [22, 23, 33]

Integr. step	Corrector equation	Name
1	$p_1 = p_0 + \frac{\Delta\xi_2}{2} (F_1 + F_0)$	Implicit trapezoidal rule
2	$p_2 = p_0 + \frac{\Delta\xi_2}{3} (F_2 + 4F_1 + F_0)$	Implicit Simpson's rule
3	$p_3 = p_2 + \frac{\Delta\xi_2}{24} (9F_3 - 19F_2 + 5F_1 - F_0)$	Third-order implicit Adams-Moulton
i	$p_i = p_{i-1} + \frac{\Delta\xi_2}{24} (9F_i + 19F_{i-1} - 5F_{i-2} + F_{i-3})$	Fourth-order implicit Adams-Moulton

Table 2.3: Higher-order variable step-size predictor-corrector equations

Step	Predictor equation	Corrector equation
1	$p_1 = p_0 + \Delta\xi_2 F_0$	$p_1 = p_0 + \frac{\Delta\xi_2}{2} (F_0 + F_1)$
2	$p_2 = p_1 + \frac{1}{2} \frac{\Delta\xi_2}{r_1} \{(2r_1 + 1)F_1 - F_0\}$	$p_2 = p_1 + D \sum_{j=-1}^1 C_j F_{1+j}$
$i \geq 3$	$p_i = p_{i-1} + G \sum_{j=-2}^1 H_j F_{i+j}$	$p_i = p_{i-1} + I \sum_{j=-2}^1 J_j F_{i+j}$

Table 2.4: Coefficient used in Table 2.3

Step	Coefficients
All steps	$r_i = [\xi_2]_i - [\xi_2]_{i-1}$
2	$C_{-1} = -\frac{1}{r_i} \quad C_0 = \frac{1}{r_i} + 4 + 3r_i \quad C_1 = 2 + 3r_i \quad D = \frac{\Delta\xi_i}{6(1+r_i)}$
$i \geq 3$	$H_{-2} = 3r_i + 2 \quad H_{-1} = -(3r_i r_{i-1} + 3r_i + 2)(r_{i-1} + 1) \quad I = \frac{\Delta\xi_i}{12}$ $H_0 = r_{i-1} \left\{ 3(r_i + 1)(r_i r_{i-1} + r_i + 1) + 3r_i(r_i r_{i-1} + r_i) - 1 \right\} \quad G = \frac{1}{6} \frac{\Delta\xi_i}{r_{i-1} r_i^2 (r_{i-1} + 1)}$ $J_1 = \frac{\left\{ 2r_i(r_i r_{i-1} + r_i) + 4(r_i r_{i-1} + r_i + 1)(r_i + 1) - 1 \right\}}{(r_i r_{i-1} + r_i + 1)(r_i + 1)} \quad J_{-2} = \frac{(2r_i + 1)}{r_i^2 r_{i-1} (r_{i-1} + 1)(r_i r_{i-1} + r_i + 1)}$ $J_0 = -\frac{\left\{ 4r_i(r_i r_{i-1} + r_i) + 2(r_i r_{i-1} + r_i + 1)(r_i + 1) - 1 \right\}}{r_i^2 (r_{i-1} + 1)} \quad J_{-1} = \frac{(2r_i r_{i-1} + 2r_i + 1)}{r_i^2 r_{i-1} (r_i + 1)}$

2.7. Conclusions

In this chapter, several simulation techniques that can be used to predict vapour-liquid equilibria were discussed. Among these methods is Gibbs-Duhem integration. The method has the advantage of being independent of insertions and/or removals of molecules. This independence makes the method especially suitable for the computation of phase equilibria in which at least one of the phases is a solid one.

When mixtures are simulated, the Gibbs-Duhem integration method can be conveniently combined with semigrand-canonical ensemble simulations.

Derivations for the most general form of the Gibbs-Duhem equation and the Clapeyron equation were given.

The isothermal-isobaric ensemble, Gibbs ensemble, and semigrand-canonical ensemble were briefly discussed. Conventional isobaric partition functions redundantly count volume states. In this thesis, corrected partition functions are used that count volume states in the right way. Simulations performed with both types of partition functions and moderate system sizes provided results that matched within the statistical uncertainties.

Higher-order predictor-corrector equations, that are used to numerically integrate the Clapeyron equation, were derived. Higher-order equations generally provide more accurate predictions/corrections, thereby decreasing the number of necessary computations of the integrand. A disadvantage of higher-order equations is the small integration step-size that is needed to preserve stability of the method.

This chapter was concluded with a description of the numerical-integration scheme that is used in the Gibbs-Duhem integration method.

2.8. References

1. Rowlinson, J. S. and Widom, B., *Molecular theory of capilarity*, International series of monographs on chemistry, Volume 8, Oxford University Press, New York, 1982
2. Gubbins, K. E., The role of computer simulation in studying phase equilibria, *Mol. Sim.*, 2 **1989** 223
3. Frenkel, D. and Smit, B., *Understanding molecular simulation; from algorithms to applications*, Computational Science Series, Volume 1, 2nd Ed., Academic press, London, 2002
4. Panagiotopoulos, A. Z., Monte Carlo methods for phase equilibria of fluids, *J. Phys.: Condens. Matter*, 12 **2000** R25
5. Adams, D. J., Calculating the low-temperature vapor line by Monte Carlo, *Mol. Phys.*, 32(3) **1976** 647
6. Kofke, D. A. and Glandt, E. D., Monte Carlo simulation of multicomponent equilibria in a semigrand canonical ensemble, *Mol. Phys.*, 64(6) **1988** 1105
7. Möller, D., Correction, *Mol. Phys.*, 75(6) **1992** 1461
8. Möller, D. and Fischer, J., Vapor liquid equilibrium of a pure fluid from test particle method in combination with NpT molecular dynamics simulations, *Mol. Phys.*, 69(3) **1990** 463
9. Lofti, A., Vrabec, J., and Fischer, J., Vapour liquid equilibria of the Lennard-Jones fluid from the NpT plus test particle method, *Mol. Phys.*, 76(6) **1992** 1319
10. Vrabec, J., Lofti, A., and Fischer, J., Vapour liquid equilibria of Lennard-Jones model mixtures from the NpT plus test particle method, *Fluid Phase Equil.*, 112 **1995** 173
11. Vrabec, J. and Fischer, J., Vapour liquid equilibria of mixtures from the NpT plus test particle method, *Mol. Phys.*, 85(4) **1995** 781
12. Fincham, D., Quirke, N., and Tildesley, D. J., Computer simulation of molecular liquid mixtures. I. A diatomic Lennard-Jones model mixture for CO₂/C₂H₆, *J. Chem. Phys.*, 84(8) **1986** 4535
13. Panagiotopoulos, A. Z., Suter, U. W., and Reid, R. C., Phase diagrams of non-ideal fluid mixtures from Monte Carlo simulation, *Ind. Engng. Chem. Fundam.*, 25 **1986** 525
14. Adams, D. J., Calculating the high-temperature vapor line by Monte Carlo, *Mol. Phys.*, 37(1) **1979** 211
15. Allen, M. P. and Tildesley, D. J., *Computer simulation of liquids*, Clarendon Press, Oxford, 1987
16. Panagiotopoulos, A. Z., Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.*, 61(4) **1987** 813
17. Panagiotopoulos, A. Z., Quirke, N., Stapleton, M. R., and Tildesley, D. J., Phase equilibria by simulation in the Gibbs ensemble: Alternative derivation, generalization and application to mixture and membrane equilibria, *Mol. Phys.*, 63(4) **1988** 527
18. Panagiotopoulos, A. Z., *Supercritical Fluids, Fundamentals for Applications*. In: Kiran, E. and Levelt Sengers, J. M. H., editors, *NATO ASI Series E, volume 273*, Kluwer, Dordrecht, 1994
19. Panagiotopoulos, A. Z., Direct determination of fluid phase equilibria by simulation in the Gibbs ensemble: A review, *Mol. Sim.*, 9 **1992** 1
20. Smit, B., de Smedt, P., and Frenkel, D., Computer simulations in the Gibbs ensemble, *Mol. Phys.*, 68(4) **1989** 931
21. Smit, B., Simulation of phase coexistence: from atoms to surfactants, *PhD thesis, University of Utrecht, The Netherlands*, **1990**

-
22. Kofke, D. A., Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line, *J. Chem. Phys.*, 98(5) **1993** 4149
 23. Kofke, D. A., Gibbs-Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation, *Mol. Phys.*, 78(6) **1993** 1331
 24. Hitchcock, M. R. and Hall, C. K., Solid-liquid phase equilibrium for binary Lennard-Jones mixtures, *J. Chem. Phys.*, 110(23) **1999** 11433
 25. Silva Fernandes, F. M. S., Fartaria, R. P. S., and Freitas, F. F. M., The starting state in simulations of the fluid-solid coexistence by Gibbs-Duhem integration, *Comp. Phys. Comm.*, 141 **2001** 403
 26. Agrawal, R., Thermodynamic and structural properties of model systems at solid-fluid coexistence. II. Melting and sublimation of the Lennard-Jones system, *Mol. Phys.*, 85(1) **1995** 43
 27. Agrawal, R. and Kofke, D. A., Thermodynamic and structural properties of model systems at solid-fluid coexistence. I. FCC and BCC soft spheres, *Mol. Phys.*, 85(1) **1995** 23
 28. Lamm, M. H. and Hall, C. K., Equilibria between solid, liquid, and vapor phases in binary Lennard-Jones mixtures, *Fluid Phase Equil.*, 194-197 **2002** 197
 29. Lamm, M. H. and Hall, C. K., Monte Carlo simulations of complete phase diagrams for binary Lennard-Jones mixtures, *Fluid Phase Equil.*, 182(1-2) **2001** 37
 30. Bolhuis, P. G. and Kofke, D. A., Numerical study of freezing in polydisperse colloidal suspensions, *J. Phys. : Condens. Matter*, 8 **1996** 9627
 31. Polson, J. M. and Frenkel, D., Calculation of solid-fluid phase equilibria for systems of chain molecules, *J. Chem. Phys.*, 109(1) **1998** 318
 32. Polson, J. M. and Frenkel, D., Numerical prediction of the melting curve of *n*-octane, *J. Chem. Phys.*, 111(4) **1999** 1501
 33. Mehta, M. and Kofke, D. A., Coexistence Diagrams of Mixtures by Molecular Simulation, *Chem. Eng. Sci.*, 49(16) **1994** 2633
 34. Agrawal, R., Mehta, M., and Kofke, D. A., Efficient Evaluation of Three-Phase Coexistence Lines, *Int. J. Thermophys.*, 15(6) **1994** 1073
 35. Kofke, D. A., Semigrand Canonical Monte Carlo Simulation; Integration Along Coexistence Lines. In: Ferguson, D. M., Siepmann, J. I., and Truhlar, D. G., editors, *Advances in Chemical Physics, Volume 105, Monte Carlo Methods in Chemistry*, Wiley, 1999
 36. Agrawal, R. and Kofke, D. A., Solid-Fluid Coexistence for Inverse-Power Potentials, *Phys. Rev. Lett.*, 74(1) **1995** 122
 37. Camp, P. J., Mason, C. P., Allen, M. P., Khare, A. A., and Kofke, D. A., The isotropic-nematic phase transition in uniaxial hard ellipsoid fluids: Coexistence data and the approach to the Onsager limit, *J. Chem. Phys.*, 105(7) **1996** 2837
 38. Pandit, S. P. and Kofke, D. A., Evaluation of a locus of azeotropes by molecular simulation, *AIChE J.*, 45(10) **1999** 2237
 39. Tester, J. W. and Modell, M., *Thermodynamics and Its Applications*, Prentice Hall International Series in the Physical and Chemical Engineering Sciences, 3rd Ed., Prentice Hall PTR, Upper Saddle River, NJ, 1997
 40. Griffiths, R. B. and Wheeler, J. C., Critical Points in Multicomponent Systems, *Phys. Rev. A*, 2(3) **1970** 1047
 41. Smith, J. M., Van Ness, H. C., and Abbott, M. M., *Introduction to chemical engineering thermodynamics*, Chemical Engineering Series, McGraw-Hill, Boston, 2001
 42. Leung, S. S. and Griffiths, R. B., Thermodynamic Properties near the Liquid-Vapor Critical Line in Mixtures of He³ and He⁴, *Phys. Rev. A*, 8(5) **1973** 2670

-
43. Wood, W. W., Monte Carlo calculations for hard disks in the isothermal-isobaric ensemble, *J. Chem. Phys.*, 48(1) **1968** 415
 44. McDonald, I. R., *NpT*-ensemble Monte Carlo calculations for binary liquid mixtures, *Mol. Phys.*, 23(1) **1972** 41
 45. Attard, P., On the density of volume states in the isobaric ensemble, *J. Chem. Phys.*, 103(22) **1995** 9884
 46. Corti, D. S., Isothermal-isobaric ensemble for small systems, *Phys. Rev. E*, 64(1) **2001** 016128
 47. Corti, D. S., Monte Carlo simulations in the isothermal-isobaric ensemble: the requirement of a 'shell' molecule and simulations of small systems, *Mol. Phys.*, 100(12) **2002** 1887
 48. Corti, D. S. and Soto-Campos, G., Deriving the isothermal-isobaric ensemble: The requirement of a "shell" molecule and applicability to small systems, *J. Chem. Phys.*, 108(19) **1998** 7959
 49. Han, K.-K. and Son, H. S., On the isothermal-isobaric ensemble partition function, *J. Chem. Phys.*, 115(16) **2001** 7793
 50. Smit, B., Computer Simulations in Chemical Physics. In: Allen, M. P. and Tildesley, D. J., editors, *NATO ASI Series C, volume 397*, Kluwer, Dordrecht, 1993
 51. Panagiotopoulos, A. Z., Observation, Prediction and Simulation of Phase Transitions in Complex Fluids. In: Baus, M., Rull, L. F., and Ryckaert, J.-P., editors, *NATO ASI Series C, volume 460*, Kluwer, Dordrecht, 1995
 52. Panagiotopoulos, A. Z., Exact Calculations of Fluid-Phase Equilibria by Monte Carlo Simulation in a New Statistical Ensemble, *Int. J. Thermophys.*, 10(2) **1989** 447
 53. Agrawal, R. and Kofke, D. A., *Mol. Phys.*, 85 **1995** 23
 54. Smit, B., Grand canonical Monte Carlo simulations of chain molecules: adsorption isotherms of alkanes in zeolites, *Mol. Phys.*, 85(1) **1995** 153
 55. Burden, R. L. and Douglas Faires, J., *Numerical Analysis*, 7th Ed., Brooks/Cole, Pacific Grove, 2001
 56. Escobedo, F. A. and de Pablo, J. J., Pseudo-ensemble simulations and Gibbs-Duhem integrations for polymers, *J. Chem. Phys.*, 106(7) **1997** 2911

Chapter 3. Principles of Monte Carlo importance sampling

This chapter starts with a short discussion on the mathematical problems that are encountered in statistical mechanics and thus in molecular simulations as well. In Section 3.2, several numerical methods that are commonly used to compute integrals are discussed. The most appropriate method to compute multidimensional integrals that cannot be solved analytically is the Monte Carlo method. A naïve implementation of the Monte Carlo method can be very inefficient. On the other hand, a combination of importance sampling and Monte Carlo integration can be conveniently used to compute the average of a function over a difficult multidimensional weighting function. Applications of Monte Carlo importance sampling that are typical for this work are described in Section 3.3. Generation of random variates in accordance with a particular weighting function is performed by construction of a Markov chain. The construction of a Markov chain of configurations with a limiting distribution that equals the relevant statistical-mechanical probability distribution via Metropolis Monte Carlo sampling is discussed in Section 3.4. Bias methods can be used to enhance sampling of configuration space. Several existing bias method and a new bias method are described in Section 3.5. Finally, conclusions are summarized in Section 3.6.

3.1. Computation of multidimensional integrals; the problem

The relations that arise in statistical mechanics often contain multidimensional integrals of a Boltzmann factor or a pseudo-Boltzmann factor over *phase space*. These multidimensional integrals cannot be solved analytically. Phase space corresponds to the fictitious $6N$ -dimensional space formed by the 3 Cartesian coordinates and the 3 momentum coordinates of N interaction sites. As Monte Carlo simulations do not follow the system in time, the momenta are not relevant and can be integrated out of the multidimensional relations [1]. The reduced phase space that is set up by the $3N$ Cartesian coordinates of the interaction sites is called *configuration space*. For convenience, configuration space, scaled by the box length of the cubic simulation box, is indicated by Γ^{3N} in this chapter.

An example of a multidimensional statistical-mechanical integral is the semi-classical canonical partition function of a pure-component system consisting of N molecules in a volume V at reciprocal temperature β :

$$Q_{NV\beta} = \frac{(Vq)^N}{N!} \int \exp\{-\beta U(\Gamma^{3N})\} d\Gamma^{3N} \quad (3.1)$$

where q is the kinetic contribution of one molecule to the partition function. The absolute value of the multidimensional integral in Eq. (3.1) is of interest since it is related to the Helmholtz energy [See Chapter 2, Eq. (2.4)]. As will be pointed out in the next sections, the computation of the absolute value of the partition function is a hard task. However, this work mainly focuses on the computation of ensemble averages of configuration-dependent properties $h(\Gamma^{3N})$:

$$\langle h \rangle_{NV\beta} = \frac{(Vq)^N}{N!} \frac{\int h(\Gamma^{3N}) \exp\{-\beta U(\Gamma^{3N})\} d\Gamma^{3N}}{Q_{NV\beta}} = \int h(\Gamma^{3N}) \rho_{NV\beta}(\Gamma^{3N}) d\Gamma^{3N} \quad (3.2)$$

The differential quantity $\rho_{NV\beta}(\Gamma^{3N})d\Gamma^{3N}$ in Eq. (3.2) is the canonical probability to observe the system with a configuration between Γ^{3N} and $\Gamma^{3N}+d\Gamma^{3N}$. The brackets $\langle \rangle_{NV\beta}$ indicate a canonical ensemble average. In contrast to the absolute value of the partition function, the ensemble average of a property that depends on configuration space can be computed rather easily from Monte Carlo importance sampling, a technique that is explained in Section 3.2.

3.2. Numerical integration and Monte Carlo importance sampling

The principle of Monte Carlo importance sampling is illustrated with a simple example. Suppose one would like to know the value of the following, let's say hard to determine, integral of the function $f(x)$ between a lower bound x^{start} and an upper bound x^{end} :

$$I = \int_{x^{\text{start}}}^{x^{\text{end}}} f(x)dx \quad (3.3)$$

Furthermore, assume that the integral cannot be solved analytically. In the simplest numerical integration method, the average function value between the integration bounds is computed by averaging function values computed at the middle of regular subintervals. Subsequently, the average function value is multiplied by the integration domain:

$$I = \lim_{N \rightarrow \infty} \frac{x^{\text{end}} - x^{\text{start}}}{N} \sum_{i=1}^N f(x_i) \quad x_i = x^{\text{start}} + \left(i - \frac{1}{2}\right) \frac{x^{\text{end}} - x^{\text{start}}}{N} \quad (3.4)$$

Among the more sophisticated numerical methods are quadrature formulae's like Simpson's rule and the trapezoidal rule. In these methods, the function $f(x)$ is approximated by a closely fitting polynomial on regular subintervals.

It is also possible to compute the integral (3.3) by a process in which the function value is computed at random values that are uniformly sampled within the integration domain:

$$I = \underbrace{\lim_{N \rightarrow \infty} \frac{x^{\text{end}} - x^{\text{start}}}{N} \sum_{i=1}^N f(x_i)}_{N \text{ random variates } x_i, \text{ uniformly distributed between } x^{\text{start}} \text{ and } x^{\text{end}}} \quad (3.5)$$

This is a *Monte Carlo integration* method [2], so-called because the abundant use of random numbers probably made the inventor of the method think of the famous gambling resort.

When the function $f(x)$ has only large contributions from a tiny part of the integration domain, as is the case in many statistical-mechanical integrals, Monte Carlo integration with uniformly distributed random sampling is far from efficient. The larger part of the generated x_i -values has a negligible function value $f(x_i)$. Moreover, the x_i -values that have non-negligible function values are sampled infrequently, resulting in poor statistics. Fortunately, this sampling problem can be partially overcome by combining Monte Carlo integration with a technique called *importance sampling*.

Importance sampling, or sample-mean integration, introduces a weight function under the integral that gives priority to those x -values in the integration domain that contribute significantly to the integral. Instead of generating uniformly distributed random values of x , uniformly distributed random values of y are generated. The corresponding x -values are computed from the following cumulative probability distribution:

$$y = \int_{x^{\text{start}}}^x \frac{p(x')}{C} dx' \quad (3.6)$$

The weight function $p(x)$ resembles $f(x)$ and C is the normalization constant of $p(x)$. Introduction of the weight function into integral (3.3) results in:

$$I = C \int_{x^{\text{start}}}^{x^{\text{end}}} \frac{f(x) p(x)}{p(x) C} dx = C \int_0^1 \frac{f[x(y)]}{p[x(y)]} dy \quad (3.7)$$

Fortunately, there are methods to generate random variates in accordance with a particular probability distribution without using the transformation $x(y)$ in Eq. (3.6) and without computing the normalization constant C [See Section 3.4]. Thus, Eq. (3.7) can be simplified as follows:

$$I = C \int_{x^{\text{start}}}^{x^{\text{end}}} \frac{f(x')}{p(x')} dx' = \lim_{N \rightarrow \infty} \underbrace{\frac{C}{N} \sum_{i=1}^N \frac{f(x'_i)}{p(x'_i)}}_{N \text{ random variates } x'_i, \text{ from probability distribution } p(x)/C dx} \quad (3.8)$$

The problem with Eq. (3.8) is that the normalization constant, C , still needs to be computed. Notice that choosing $p(x) = f(x)$, as a weighting function, is no appropriate option since the normalization constant of $f(x)$ is the solution to the difficult integral itself. The approach outlined above works well in case the normalization constant of $p(x)$ can be computed rather easily. However, in statistical-mechanical integrals, $f(x)$ is a very complex function of x such that any $p(x)$ resembling $f(x)$ has a normalization constant that is hard to compute as well. Thus, the absolute value of a partition function and the corresponding thermodynamic potential are generally hard to determine via Monte Carlo importance sampling.

Now, the computation of a weighted average is discussed. Suppose one needs to compute the average value of the function $h(x)$, which is weighted by a very difficult function $f(x)$. Introduction of an importance-sampling weighting function $p(x)/C$ in the relation for the weighted average gives:

$$\langle h \rangle = \frac{\int_{x^{\text{start}}}^{x^{\text{end}}} h(x) f(x) dx}{\int_{x^{\text{start}}}^{x^{\text{end}}} f(x) dx} = \frac{C \int_{x^{\text{start}}}^{x^{\text{end}}} h(x) \frac{f(x) p(x)}{p(x) C} dx}{C \int_{x^{\text{start}}}^{x^{\text{end}}} \frac{f(x) p(x)}{p(x) C} dx} = \frac{\int_{x^{\text{start}}}^{x^{\text{end}}} h(x') \frac{f(x')}{p(x')} dx'}{\int_{x^{\text{start}}}^{x^{\text{end}}} \frac{f(x')}{p(x')} dx'} \quad (3.9)$$

Thus, the weighted average of the function $h(x)$ can be computed by generating random variates x' from probability distribution $p(x)/Cdx$. Obviously, the normalization constants in the numerator and the denominator of Eq. (3.9) cancel one another! Now, it is possible to make the ultimate choice for the weighting function, $p(x) = f(x)$. This choice gives the following simple relation for the weighted average:

$$\langle h \rangle = \frac{\int h(x') dx'}{\int dx'} = \lim_{N \rightarrow \infty} \frac{1}{N} \underbrace{\sum_{i=1}^N h(x'_i)}_{\substack{N \text{ random variates } x'_i \text{ from} \\ \text{prob. distribution } p(x)/Cdx}} \quad (3.10)$$

The Monte Carlo importance sampling scheme gains significant advantage over methodical approaches when applied to multidimensional integrals, as in contrast to methodical approaches, the integration error is independent of the dimension of the integral [3].

3.3. Computation of statistical-mechanical ensemble averages

Application of Monte Carlo importance sampling to statistical-mechanical ensemble averages is straightforward. In what follows, *intramolecular configurations* are distinguished from *intermolecular configurations*. The intramolecular configuration is the spatial configuration of an individual rigid or partially flexible molecule, which is defined with respect to a body-fixed coordinate system. The intermolecular configuration of a system of molecules defines the position of the body-fixed coordinate systems of all molecules with respect to a space-fixed coordinate system.

An example of Monte Carlo importance sampling is the generation of intramolecular configurations [See Chapter 2, Section 2.5.4]. Instead of generating totally random intramolecular configurations, intramolecular configurations Ψ^N are generated from the intramolecular Boltzmann probability distribution via a rejection method [1, 4]. For computational efficiency, this intramolecular Boltzmann distribution has no intramolecular Lennard-Jones and Coulombic contribution.

The intramolecular interaction energy is considered to be independent of the positions and configurations of other molecules. Thus, intramolecular configurations of individual molecules can be generated independently. When importance sampling is applied to the intramolecular configurations, the ensemble average of the configuration-dependent property h can be written as follows:

$$\langle h \rangle = \frac{\int h(\Psi^N, \mathbf{s}^N) \exp\{-\beta U^{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}(\Psi^N, \mathbf{s}^N)\} d\Psi^N d\mathbf{s}^N}{\int \exp\{-\beta U^{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}(\Psi^N, \mathbf{s}^N)\} d\Psi^N d\mathbf{s}^N} \quad (3.11)$$

As opposed to Eq. (2.35) in Chapter 2, the centre-of-mass coordinates \mathbf{s}^N in Eq. (3.11) are separated from the intramolecular configurations Ψ^N .

Intermolecular configurations are also generated with importance sampling. Provided that the intramolecular configurations of the molecules have already been generated, the probability to generate an intermolecular configuration between Ψ'^N and $\Psi'^N + d\Psi'^N$ is given by:

$$d\Psi'^N = \frac{\exp\left\{-\beta\mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}(\Psi'^N, \mathbf{s}^N)\right\} d\mathbf{s}^N}{C} \quad (3.12)$$

Generation of configurations in accordance with the total Boltzmann probability allows the computation of an ensemble average of a configuration-dependent property $h(\Gamma^{3N})$ from:

$$\langle h \rangle = \frac{\int h(\Psi'^N) d\Psi'^N}{\int d\Psi'^N} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N h[\Psi'^N] \quad (3.13)$$

The only problem that remains is the actual generation of *intermolecular configurations* Ψ'^N in accordance with a given probability distribution, provided *intramolecular* Ψ^N configurations of the molecules have already been generated. These intermolecular configurations can be efficiently generated by perturbing an existing intermolecular configuration. As will be described in the next section, this can be accomplished by constructing a Markov chain of successive intermolecular configurations with Metropolis Monte Carlo sampling.

3.4. Markov chains

Sampling of phase space in accordance with a specific probability density ρ^{limit} , for example $\rho^{\text{limit}} = \rho_{NV\beta}$, can be performed by constructing a Markov chain of configurations. The Markov chain used in this work is a stochastic sequence of trial configurations that satisfies the following conditions:

1. The Markov chain needs to be *ergodic*, *i.e.* all possible configurations must be within reach from any other configuration in a finite number of steps.
2. The outcome of each configuration depends only on the outcome of the immediately preceding configuration; *i.e.* the process has no memory.

The Markov chain is fully specified by the transition-probability matrix Π , the elements of which represent the transition probabilities π_{ij} that the system moves from a present configuration i into a configuration j . To simplify the notation in this section, discrete states are considered although configurations are by no means discrete. The transition-probability matrix Π for a system that can be in any of n discrete states is given by:

$$\Pi \equiv \begin{pmatrix} \pi_{11} & \cdots & \pi_{1n} \\ \vdots & \ddots & \vdots \\ \pi_{n1} & \cdots & \pi_{nn} \end{pmatrix} \quad (3.14)$$

During the construction of the Markov chain, the system moves from one state to another, selecting the new state according to the transition probabilities. The long-time average of the probability to visit the different states is called the limiting probability distribution, ρ^{limit} , of the Markov chain. As the states are discrete, the limiting probability distribution is a vector of length n . The limiting probability distribution is an eigenvector of the transition probability matrix Π [5]:

$$\rho^{\text{limit}} = \lim_{p \rightarrow \infty} [\Pi^p \rho^0] = \Pi \left(\lim_{p \rightarrow \infty} [\Pi^{p-1} \rho^0] \right) = \Pi \rho^{\text{limit}} \quad (3.15)$$

where Π^p is the p -step transition-probability matrix and ρ^0 is an initial guess for the limiting probability distribution. The eigenvector equation (3.15) comprises the following set of equalities:

$$\rho_i^{\text{limit}} = \sum_{j=1}^n \pi_{ji} \rho_j^{\text{limit}} \quad \{i = 1 \dots n\} \quad (3.16)$$

If the limiting probabilities ρ_j^{limit} and the transition probabilities π_{ij} all satisfy the principle of *microscopic reversibility*, which is also known as *detailed balance* [1],

$$\rho_i^{\text{limit}} \pi_{ij} = \rho_j^{\text{limit}} \pi_{ji}, \quad (3.17)$$

then they also satisfy the eigenvector equation. Detailed balance is a sufficient condition for the transition probabilities to satisfy Eq. (3.16), although it is unnecessarily strong [5].

In this work, the limiting distribution is given by the statistical-mechanical probability distribution of the ensemble of interest. A method to construct a Markov chain using the condition of detailed balance has been developed by Metropolis *et al.* in 1953 [6]. Suppose a trial move is performed from configuration i to configuration j . The probability to choose this move is given by the relevant element of the stochastic matrix α , which is also called the *underlying matrix of the Markov chain*. The total Metropolis Monte Carlo scheme is given by:

1. When in configuration i , choose with probability α_{ij} a new trial configuration j .
2. If $\rho_j^{\text{limit}} \geq \rho_i^{\text{limit}}$, accept j as the new state, otherwise accept state j with probability:

$$\min \left\{ 1, \frac{\alpha_{ji} \rho_j^{\text{limit}}}{\alpha_{ij} \rho_i^{\text{limit}}} \right\} \quad (3.18)$$

3. If trial configuration j is rejected, the system remains in configuration i .

As Eq. (3.18) contains a ratio of limiting probabilities, the normalization constants are not important.

3.5. Biasing the underlying matrix of the Markov chain

The original formulation of the Metropolis scheme, as described in Section 3.4, is based on a symmetric underlying stochastic matrix, such that $\alpha_{ij} = \alpha_{ji}$. This restriction is not necessary and in many cases inefficient. In the Metropolis Monte Carlo scheme, it often happens that α_{ij} is small while ρ_j^{limit} is much larger than ρ_i^{limit} . This indicates that, although the new state is favourable, the system can hardly find it. In this case, the system is subject to *entropic sampling problems*. On the other hand, α_{ij} can be large while ρ_j^{limit} is much smaller than ρ_i^{limit} . In this case, transitions that are selected very often are rarely accepted. These *sampling problems* are *energetic* in nature.

Biasing the underlying matrix of the Markov chain can be of help in these situations. The underlying matrix, as described in Section 3.4, is unbiased in the sense that no information on the present configuration is used to select new configurations. The bias methods used in this work include information on the present configuration in order to increase the diffusion of the system through configuration space without changing the limiting distribution. The bias should make the ratio $\alpha_{ji}\rho_j^{\text{limit}}/\alpha_{ij}\rho_i^{\text{limit}}$ close to one in order to increase the number of accepted Monte Carlo trial moves. This can be accomplished by choosing an asymmetric stochastic matrix α .

Many bias methods have been developed over the years. Examples are the force-bias method [7, 8], the virial-bias method [9, 10], the rotational-insertion-bias method [11], the association-bias method [12-19], the cavity-biased insertion method [20-23], and the configurational-bias method. In Section 3.5.1, the torque-bias and force-bias methods are described. Although the configurational-bias method is extensively used in this work, one is referred to the book of Frenkel and Smit [4] for a thorough discussion of the method. In Section 3.5.2, a new bias method is presented, which is developed in order to improve composition sampling in the semigrand-canonical ensemble.

3.5.1. Torque-biased rotations and force-biased translations

Moving molecules preferentially into the direction of the force and/or rotating molecules preferentially around the torque acting on the molecule, may increase the translational and rotational diffusion per unit computer time by a factor of 2 to 3 compared to conventional Metropolis sampling [24]. This is particularly advantageous in high-density systems and in systems with strong orientation-dependent interactions, in which diffusion through configuration space via conventional Metropolis sampling is poor. A force/torque-biased sampling scheme has been developed by Pangali *et al.* [7]. The method was further investigated by Rao *et al.* [8]. The probability to select a centre-of-mass translation in the interval $\{r_i, r_i + dr_i; i = x, y, z\}$, biased into the direction of the force-component f_i acting on the molecule is [7, 8]:

$$p(r_i)dr_i = \frac{\exp\{+\lambda\beta f_i r_i\} dr_i}{C(f_i, r_{\text{max}})} \quad (3.19)$$

where $C(f_i, r_{\text{max}})$ is a normalization constant, r_{max} is the maximum translation, and λ is a constant that determines the strength of the bias.

The normalization constant is given by:

$$C(f_i, r_{\max}) = \frac{2 \sinh\{\lambda \beta f_i r_{\max}\}}{\lambda \beta f_i} \quad (3.20)$$

The components of the translation vector \mathbf{r} are sampled in accordance with probability distribution (3.19) from [8]:

$$r_i = \frac{1}{\lambda \beta f_i} \left[\ln \left\{ \xi \left[\exp(+\lambda \beta f_i r_{\max}) - \exp(-\lambda \beta f_i r_{\max}) \right] + \exp(-\lambda \beta f_i r_{\max}) \right\} \right] \quad (3.21)$$

where ξ is a random number, uniformly sampled in the interval $\{0,1\}$. Pangali *et al.* [7] and Rao *et al.* [8] respectively used torque-biased Barker-and-Watts [25] rotations and torque-biased Euler-angle [1] rotations. Application of a torque-bias to Euler-angle rotations should lead to improved sampling compared to torque-biased Barker-and-Watts rotations. The reason is that for torque-biased Euler-angle rotations, both the rotation and the selection of the rotation axis are biased in the direction of the torque. However, Mehrotra *et al.* [26] showed with their simulation results for liquid water that the Euler-angle treatment did not lead to significant improvements compared to the simpler torque-biased Barker-and-Watts rotation scheme. Although the validity of this conclusion depends on the system that is simulated, in this work the computationally cheap torque-biased Barker and Watts rotations are used. The probability to select a rotation angle θ in the interval $\{\theta, \theta + d\theta\}$, around one of the space-fixed axes $\hat{\mathbf{n}}$, biased into the direction of the torque \mathbf{T} acting on the molecule is:

$$p(\theta) d\theta = \frac{\exp\{+\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}} \theta\} d\theta}{C(\mathbf{T}, \hat{\mathbf{n}}, \theta_{\max})} \quad (3.22)$$

where $C(\mathbf{T}, \hat{\mathbf{n}}, \theta_{\max})$ is a normalization constant, θ_{\max} is the maximum rotation, and ζ is a constant that determines the strength of the bias. The normalization constant is given by:

$$C(\mathbf{T}, \hat{\mathbf{n}}, \theta_{\max}) = \frac{2 \sinh\{\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}} \theta_{\max}\}}{\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}}} \quad (3.23)$$

The angle θ is sampled in accordance with probability distribution (3.22) from:

$$\theta = \frac{\left[\ln \left\{ \xi \left[\exp(\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}} \theta_{\max}) - \exp(-\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}} \theta_{\max}) \right] + \exp(-\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}} \theta_{\max}) \right\} \right]}{\zeta \beta \mathbf{T} \cdot \hat{\mathbf{n}}} \quad (3.24)$$

where ξ is a random number, uniformly sampled in the interval $\{0,1\}$.

The bias in the Markov chain is removed in the formulation of the acceptance criterion for the Monte Carlo trial moves.

The general acceptance criterion for a biased Monte Carlo trial move is written as:

$$\min \left\{ 1, \mathcal{A} \exp \left[-\beta (\Delta \mathcal{U} + \mathcal{B}) \right] \right\} \quad (3.25)$$

where $\Delta \mathcal{U}$ is the difference in interaction energy resulting from the trial move. The trial move is accepted in case the product of \mathcal{A} and the pseudo-Boltzmann factor in Eq. (3.25) exceeds a random number, sampled uniformly in the interval $\{0,1\}$. If the move is not accepted, the old situation is restored. The variables \mathcal{A} and \mathcal{B} , for the different biased trial moves, are given in Table 3.1.

According to Rao *et al.* [8], it is not clear which values of λ and ζ lead to the optimum diffusion through configuration space; the diffusion depends on the bias strength as well as on the maximum step sizes $\{r_{\max}, \theta_{\max}\}$. However, without optimization, a suitable choice for λ and ζ is 0.5 [8]. Maximum translations/rotations are often optimized during the equilibration phase of the simulation in order to get an overall acceptance of about 50% of the Monte Carlo trial moves.

Table 3.1: \mathcal{A} and \mathcal{B} in acceptance criterion (3.25)

Bias method	\mathcal{A}	\mathcal{B}
Force bias	$\prod_{i=1}^3 \frac{f_i^{\text{new}} \sinh \left\{ \lambda \beta f_i^{\text{old}} \delta r_{\max} \right\}}{f_i^{\text{old}} \sinh \left\{ \lambda \beta f_i^{\text{new}} \delta r_{\max} \right\}}$	$\lambda \left(\mathbf{f}^{\text{old}} + \mathbf{f}^{\text{new}} \right) \cdot \delta \mathbf{r}$
Torque bias	$\left(\frac{\mathbf{T}^{\text{new}} \cdot \hat{\mathbf{n}}}{\mathbf{T}^{\text{old}} \cdot \hat{\mathbf{n}}} \right) \frac{\sinh \left\{ \zeta \beta \mathbf{T}^{\text{old}} \cdot \hat{\mathbf{n}} \delta \theta_{\max} \right\}}{\sinh \left\{ \zeta \beta \mathbf{T}^{\text{new}} \cdot \hat{\mathbf{n}} \delta \theta_{\max} \right\}}$	$\zeta \left(\mathbf{T}^{\text{old}} + \mathbf{T}^{\text{new}} \right) \cdot \hat{\mathbf{n}} \delta \theta$

3.5.2. Energy-biased identity change in the semigrand-canonical ensemble

In this section, binary mixtures are considered. For a simple fluid comprising N single-site Lennard-Jones molecules, the number of intermolecular distances that needs to be computed in the evaluation of the intermolecular energy is $N(N-1)/2$. This number is already huge for a simulation system of moderate size. Thus, limited computer power forces us to choose *relatively small simulation systems*. Clearly, in a simulation of a diluted mixture, on average only a few molecules take the other identity.

It is important to have a reasonable acceptance of the identity-change trial move in order to enhance *diffusion through composition space*. This is particularly important in case only a few molecules takes the other identity. The average number of molecules that take the other identity in a semigrand-canonical simulation depends on the imposed fugacity fraction, on the system size, and on the differences in geometry and interaction potential of the molecules.

At first consider the ideal case, in which both components in the binary mixture are modelled as single-site Lennard-Jones molecules with identical force-field parameters. Thus, only the label of the molecules is different.

A molecule is picked randomly and an identity-change trial move is performed. In this ideal case, acceptance criterion (2.40) reduces to:

$$\min \left[1, \left(\frac{\xi_2}{1-\xi_2} \right)^{m_{12}} \right] \quad (3.26)$$

where $m_{12} = -1$ for an identity change $2 \rightarrow 1$ and $m_{12} = +1$ for the reverse change. The naïve implementation of the identity-change trial move described above picks a random molecule and changes its identity. This implementation is far from efficient for realistic molecules. For real molecules, *differences in size, shape and force-field parameters* are generally nonzero. These differences often have a dramatic effect on the fraction of accepted identity changes. The acceptance criterion for the identity-change trial move for the semi-flexible molecules described in Chapter 2, Section 2.5, was given by Eq. (2.40):

$$\min \left[1, \exp \left\{ -\beta \left(\mathcal{U}_{\text{new}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} - \mathcal{U}_{\text{old}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right) + m_{12} \ln \left(\frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle_{\xi_2}}{\langle \mathcal{W}_{\text{IG},2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle_{\xi_1}} \right) \right\} \right] \quad (3.27)$$

where $\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle$ is the normalized Rosenbluth factor of an isolated chain of component i with intramolecular Lennard-Jones and Coulombic interactions [4, 27] [See Chapter 2, Section 2.5.4].

Biasing the choice of the molecule whose identity is to be changed, instead of choosing a random molecule, may increase the overall fraction of accepted identity changes and the diffusion through *composition space*. The energy of a settled *old* molecule, $\mathcal{U}_{\text{old}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$, is generally low. Given the energy of the *old* molecule, the acceptance probability for the trial identity change is highest for *new* molecules of the other identity with low values of $\mathcal{U}_{\text{new}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$. One may try to find preferential orientations of a *new* molecule in an orientational-bias fashion. However, it is much easier to choose *old* settled molecules with a high energy $\mathcal{U}_{\text{old}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$. Although $\mathcal{U}_{\text{old}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$ is generally low, the higher its value, the larger the acceptance probability of the trial identity change. The molecule whose identity is to be changed is selected based on its energy and on the fugacity fraction that corresponds to its identity:

$$\pi_i = \frac{\exp \left\{ +\gamma \beta \mathcal{U}_i^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}, \text{old}} - \omega m_i \ln \left(\frac{\xi_2}{\xi_1} \right) \right\}}{\sum_{j=1}^N \exp \left\{ +\gamma \beta \mathcal{U}_j^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}, \text{old}} - \omega m_j \ln \left(\frac{\xi_2}{\xi_1} \right) \right\}} \quad (3.28)$$

where $m_i = -1$ when molecule i is of identity one and $m_i = +1$ when molecule i is of identity two. The parameters γ and ω determine the strength of the bias. The exponents in equation (3.28) have bound values as the energies of settled molecules are finite negative up to slightly positive. To obey the criterion of *detailed balance* [1, 4], the acceptance criterion (3.27) has to be changed as well.

The acceptance criterion for the *energy-biased identity change* is:

$$\min \left[1, \frac{\sum_{j=1}^N \exp \left\{ +\gamma \beta \mathcal{U}_j^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}, \text{old}}} - \omega m_j \ln \left(\frac{\xi_2}{\xi_1} \right) \right\}}{\sum_{j=1}^N \exp \left\{ +\gamma \beta \mathcal{U}_j^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}, \text{new}}} - \omega m_j \ln \left(\frac{\xi_2}{\xi_1} \right) \right\}} \left(\frac{\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle}{\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle} \right)^{m_{12}} \right] \cdot \exp \left\{ -(1-\gamma) \beta \Delta \mathcal{U}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + (1+2\omega) m_{12} \ln \left(\frac{\xi_2}{\xi_1} \right) \right\} \quad (3.29)$$

Logically, one would think that $\gamma = 1$ and $\omega = -0.5$ are suitable choices for the bias strengths as they remove the strongly fluctuating most right-hand exponent in Eq. (3.29). However, the *optimal* parameters depend on the system under study.

The different contributions to the bias are considered separately. In case the fugacity fraction of component two is small, the corresponding ensemble average of the mole fraction of component two will be small as well. This means that the trial identity change $1 \rightarrow 2$ is performed much more often than the reverse change. In such a case, it would be worthwhile to choose the molecules with identity two more often in order to increase the overall fraction of accepted identity changes. This increase in overall acceptance will only occur if the identity changes $1 \rightarrow 2$ do not suffer from serious energetic sampling problems.

As is clear from Eq. (3.29), ω should be assigned a negative value. The fugacity-fraction contribution to the bias increases the probability to select a molecule with identity two for $\xi_2 < 0.5$ and decreases the same probability for $\xi_2 > 0.5$. However, the differences between the imposed fugacity fraction and the ensemble average of the mole fraction of component two can be quite large. In case $\xi_2 > 0.5$ and $\langle x_2 \rangle \ll 0.5$, a positive value of ω is needed in order to select molecules with identity two more often. Thus, the sign and magnitude of ω should be selected very carefully.

The energy contribution to the bias increases the probability of changing the identity of a settled molecule with a high energy. This contribution to the bias has a more pronounced effect on the fraction accepted identity changes. There are two things about this contribution to the bias that should be noted. At first, the energies $\mathcal{U}_{\text{old}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$ of settled molecules are usually lower than those of the new molecules that did not find a favourite orientation/position. Selection criterion (3.28) should not choose the molecule whose identity changed in the immediately preceding Monte Carlo trial move too frequently! Secondly, the average energies of settled molecules with identity one are different from those of molecules with identity two. Selection probability (3.28) should not choose only molecules of one component. These phenomena are a threat to ergodicity and can be weakened by choosing a moderate bias strength.

The fractions of accepted identity changes in the liquid phase of two different systems are presented in Table 3.2 and Table 3.3 for different bias strengths. The probability $\pi_{1 \leftrightarrow 2}$ to select a molecule, given its identity changed in the immediately preceding Monte Carlo trial move, is also given in the tables. Reasonableness of bias strength is assessed based on the overall fraction of accepted changes as well as on the probability $\pi_{1 \leftrightarrow 2}$.

Methane and ethane are modelled with the TraPPE-EH potential [28], carbon dioxide is modelled with the TraPPE-AA force field, and difluoromethane is modelled with the potential of Higashi and Takada [29]. These force fields and other relevant simulation details are described in Chapter 5.

In case the electrostatic energy of the molecules is computed with the Ewald summation method [1], it makes sense to exclude the reciprocal-space contribution from selection probability (3.28). If this is not performed, the computationally expensive \mathbf{k} -space contributions per molecule will have to be computed after every Monte Carlo trial move. The fractions of accepted identity changes in Table 3.3 are obtained by excluding the \mathbf{k} -space contribution to the Ewald sum from the selection criterion.

Obviously, the energy-bias can increase the overall fraction of accepted identity changes in the liquid phase considerably. However, when the bias strength increases, the probability $\pi_{1\leftrightarrow 2}$ increases as well. From Table 3.3, it is seen that for approximately $\omega < -0.3$ the fugacity-fraction contribution to the bias has an adverse affect on the overall fraction of accepted identity changes. As the effect of the fugacity-fraction contribution to the bias is rather unpredictable, it is better to set ω to zero. A reasonable value for the bias strength of the energy contribution is $\gamma = 0.5$. However, optimal values of the bias strengths need to be optimized for every individual system.

Table 3.2: Simulation results for the binary system ethane(1)/methane(2) at $T = 250$ K, $p = 1.29$ MPa, and $\xi_2 = 0.10$; percentage of accepted trial identity changes

Bias strength	1 \rightarrow 2	2 \rightarrow 1	Overall Acc. , %	$\pi_{1\leftrightarrow 2}$
$\omega = 0, \gamma = 0$	0.64	44.	1.3	6.3e-3
$\omega = -0.5, \gamma = 0$	3.8	34.	6.8	1.0e-2
$\omega = -1, \gamma = 0$	18.	22.	20.	1.1e-1
$\omega = -0.5, \gamma = 0.25$	9.7	35.	15.	2.8e-2
$\omega = -0.5, \gamma = 0.5$	20.	37.	26.	7.5e-2
$\omega = -0.5, \gamma = 1$	43.	45.	44.	2.5e-1

Table 3.3: Simulation results for the binary system carbon dioxide(1)/difluoromethane(2) at $T = 244.3$ K, $p = 1.61$ MPa, and $\xi_2 = 0.05$; percentage of accepted trial identity changes

Bias strength	1 \rightarrow 2	2 \rightarrow 1	Overall Acc., %	$\pi_{1\leftrightarrow 2}$
$\omega = 0, \gamma = 0$	1.3	6.2	2.1	6.1e-3
$\omega = -0.1, \gamma = 0$	1.8	4.6	2.6	6.3e-3
$\omega = -0.2, \gamma = 0$	2.3	3.3	2.7	7.7e-3
$\omega = -0.3, \gamma = 0$	2.9	2.4	2.7	1.1e-2
$\omega = -0.5, \gamma = 0$	4.6	1.2	2.0	2.2e-2
$\omega = -1, \gamma = 0$	9.0	0.12	0.20	4.5e-2
$\omega = -0.5, \gamma = 0.25$	6.0	4.0	4.6	2.2e-2
$\omega = -0.5, \gamma = 0.5$	10.	11.	10.	4.4e-2
$\omega = -0.5, \gamma = 1$	11.	41.	20.	1.7e-1

3.6. Conclusions

In this chapter, the application of Monte Carlo importance sampling to the computation of statistical-mechanical integrals is described. Monte Carlo importance sampling combined with the Metropolis scheme can be conveniently used to compute statistical-mechanical ensemble averages. Computation of the absolute value of partition functions requires additional techniques.

In case the system that is simulated suffers from entropic or energetic sampling problems, it may be worthwhile to bias the underlying matrix of the Markov chain. A bias method chooses new favourable configurations using information about the present configuration.

The force-biased and torque-biased sampling schemes, that are used to enhance the diffusion of the system through configuration space, were discussed. A new bias method was developed in order to enhance composition sampling in the semigrand-canonical ensemble. It turned out that the increase in the overall fraction of accepted identity changes can be considerable. The values for the bias strengths of the different contributions to the bias are most easily set to $\gamma = 0.5$ and $\omega = 0$. However, optimal values of the bias strengths need to be optimized for every individual system.

3.7. References

1. Allen, M. P. and Tildesley, D. J., *Computer simulation of liquids*, Clarendon Press, Oxford, 1987
2. Metropolis, N. and Ulam, S., The Monte Carlo Method, *J. Am. Stat. Assoc.*, 44(247) **1949** 335
3. Landau, D. P. and Binder, K., *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, Cambridge, 2000
4. Frenkel, D. and Smit, B., *Understanding molecular simulation; from algorithms to applications*, Computational Science Series, Volume 1, 2nd Ed., Academic press, London, 2002
5. Van Kampen, N. G., *Stochastic processes in physics and chemistry*, North-Holland, Amsterdam, 1981
6. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.*, 21(6) **1953** 1087
7. Pangali, C., Rao, M., and Berne, B. J., On a novel Monte Carlo scheme for simulating water and aqueous solutions, *Chem. Phys. Lett.*, 55(3) **1978** 413
8. Rao, M., Pangali, C., and Berne, B. J., On the force bias Monte Carlo simulation of water: methodology, optimization and comparison with molecular dynamics, *Mol. Phys.*, 37(6) **1979** 1773
9. Jedlovsky, P. and Mezei, M., The anisotropic virial-biased sampling for Monte Carlo simulations in the isothermal-isobaric ensemble, *Mol. Phys.*, 96(2) **1999** 293
10. Mezei, M., Virial-bias Monte Carlo methods: efficient sampling in the (T, P, N) ensemble, *Mol. Phys.*, 48(5) **1983** 1075
11. Cracknell, R. F., Nicholson, D., Parsonage, N. G., and Evans, H., Rotational insertion bias: a novel method for simulating dense phases of structured particles, with particular application to water, *Mol. Phys.*, 71(5) **1990** 931
12. Tsangaris, D. M. and de Pablo, J. J., Bond-bias simulation of phase equilibria for strongly associating fluids, *J. Chem. Phys.*, 101(2) **1994** 1477

-
13. Chen, B. and Siepmann, J. I., Improving the Efficiency of the Aggregation-Volume-Bias Monte Carlo Algorithm, *J. Phys. Chem. B*, 105 **2001** 11275
 14. Chen, B. and Siepmann, J. I., A Novel Monte Carlo Algorithm for Simulating Strongly Associating Fluids: Applications to Water, Hydrogen Fluoride, and Acetic Acid, *J. Phys. Chem. B*, 104 **2000** 8725
 15. Wierzychowski, S. and Kofke, D. A., A general-purpose biasing scheme for Monte Carlo simulation of associating fluids, *J. Chem. Phys.*, 114(20) **2001** 8752
 16. Visco, D. P. and Kofke, D. A., Vapor-liquid equilibria and heat effects of hydrogen fluoride from molecular simulation, *J. Chem. Phys.*, 109(10) **1998** 4015
 17. Visco, D. P. and Kofke, D. A., Modeling the Monte Carlo simulation of associating fluids, *J. Chem. Phys.*, 110(12) **1999** 5493
 18. Busch, N. A., Wertheim, M. S., Chiew, Y. C., and Yarmush, M. L., A Monte Carlo method for simulating associating fluids, *J. Chem. Phys.*, 101(4) **1994** 3147
 19. Busch, N. A., Wertheim, M. S., and Yarmush, M. L., Monte Carlo simulation of n -member associating fluids: Application to antigen-antibody systems, *J. Chem. Phys.*, 104(11) **1996** 3962
 20. Mezei, M., Theoretical calculation of the liquid-vapor coexistence curve of water, chloroform and methanol with the cavity-biased Monte Carlo method in the Gibbs ensemble, *Mol. Sim.*, 9 **1992** 257
 21. Jedlovsky, P. and Mezei, M., Calculation of the Free Energy Profile of H₂O, O₂, CO, CO₂, NO, and CHCl₃ in a Lipid Bilayer with a Cavity Insertion Variant of the Widom Method, *J. Am. Chem. Soc.*, 122(21) **2000** 5125
 22. Mezei, M., A cavity-biased (μVT) Monte Carlo method for the computer simulation of fluids, *Mol. Phys.*, 40(4) **1980** 901
 23. Mezei, M., Grand-canonical ensemble Monte Carlo study of dense liquid Lennard-Jones soft spheres and water, *Mol. Phys.*, 61(3) **1987** 565
 24. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., *J. Chem. Phys.*, 21 **1953** 1087
 25. Barker, J. A. and Watts, R. O., Structure of water; A Monte Carlo Calculation, *Chem. Phys. Lett.*, 3(3) **1969** 144
 26. Mehrotra, P. K., Mezei, M., and Beveridge, D. L., Convergence acceleration in Monte Carlo computer simulation on water and aqueous solutions, *J. Chem. Phys.*, 78(6) **1983** 3156
 27. Smit, B., Grand canonical Monte Carlo simulations of chain molecules: adsorption isotherms of alkanes in zeolites, *Mol. Phys.*, 85(1) **1995** 153
 28. Chen, B. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, *J. Phys. Chem. B*, 103 **1999** 5370
 29. Higashi, S. and Takada, A., Molecular dynamics study of liquid CH₂F₂ (HFC-32), *Mol. Phys.*, 92(4) **1997** 641

Chapter 4. Computing the initial coexistence point and the initial integrand

Numerical integration of a Clapeyron differential equation requires an initial coexistence point and a corresponding initial value for the integrand. In Section 4.2, the advantages and disadvantages of several methods that may be used to compute this integrand are discussed. Simulation results for the initial value of the integrand are given in Section 4.2.5. Section 4.3 describes different methods that are used to compute the initial coexistence point. Simulation results for pure-component coexistence points are presented and discussed in Section 4.3.6. The conclusions that can be drawn from this chapter are summarized in Section 4.4.

4.1. Introduction

To start numerical integration of a Clapeyron differential equation, an initial coexistence point and an initial value for the integrand at the initial coexistence point are needed. In this chapter, the usefulness and limitations of different methods that are used to predict the initial coexistence point and the initial integrand are described.

The starting point can be any coexistence point, either taken from literature or computed from a suitable method. In case simulations are performed for systems containing single-site Lennard-Jones molecules, one can retrieve the coexistence point of interest from literature. However, for more realistic force fields, the coexistence points are generally not known.

All Gibbs-Duhem integrations in this work are performed isothermally. In a simulation of a binary mixture, the initial point is one of the pure-component coexistence points at the temperature of interest. In a ternary simulation, the numerical integration is started from the binary coexistence point at the fugacity fraction, pressure, and temperature of interest. From now on, the initial integrand of the Clapeyron equation will be indicated by *initial slope* and the initial coexistence point by *initial point*.

4.2. Predicting the initial slope

Logically, one would start to describe the techniques that can be used to predict the initial point since the initial slope is computed at the conditions of the initial point. However, a number of techniques that is needed to predict the initial slope can also be used to compute the initial point. For that reason, this section starts with a description of the methods that are used to compute the initial slope. In this chapter, binary mixtures are considered. An extension to ternary mixtures is given in Chapter 7. Further extension to multicomponent mixtures is straightforward.

As already mentioned, the initial point that is used to start a binary Gibbs-Duhem integration is one of the pure-component coexistence points at the temperature of interest. However, the Clapeyron equation [Eq. (2.25)] is not defined for a pure component. Fortunately, it is possible to rewrite the Clapeyron equation in the limit of infinite dilution of a solute (2) into a solvent (1).

The slope of the Clapeyron equation of a binary mixture in the limit of infinite dilution has been derived by Mehta and Kofke [1]:

$$\lim_{\xi_2 \rightarrow 0} \left(\frac{\partial p}{\partial \xi_2} \right)_{\beta, \sigma} = \frac{1}{\beta(v^\gamma - v^\delta)} \left\{ \left(\frac{f_1}{H_2} \right)^\gamma - \left(\frac{f_1}{H_2} \right)^\delta \right\} \quad (4.1)$$

where f_1 is the fugacity of the pure solvent, H_2 is Henry's constant of the infinitely-diluted solute in the solvent, and v^γ is the volume per molecule in phase γ . The subscript σ indicates that the derivative is taken along the saturation line. Limiting slopes for other Clapeyron equations are given in Appendix A, Table A.2.

The ratios f_1/H_2 in equation (4.1) must be expressed into thermodynamic properties that are readily accessible in a simulation. As derived below, the ratio f_1/H_2 can be expressed in the residual chemical potentials of the solvent and the solute. The chemical potential of component i in phase δ in a binary mixture is a function of pressure, temperature, and composition:

$$\beta\mu_i^\delta(p, \beta, x_i) = \beta\mu_i^\delta(p', \beta, x_i') + \ln \left\{ \frac{\hat{f}_i^\delta(p, \beta, x_i)}{\hat{f}_i^\delta(p', \beta, x_i')} \right\} \quad (4.2)$$

where \hat{f}_i^δ is the mixture fugacity of component i in phase δ . Primed properties indicate the reference state. The reference state in this work is the ideal-gas (ig) evaluated at the thermodynamic conditions of the mixture:

$$\beta\mu_i^\delta(p, \beta, x_i) = \beta\mu_i^{\text{ig}}(p, \beta, x_i) + \ln \left(\frac{\hat{f}_i^\delta(p, \beta, x_i)}{x_i p} \right)^\delta \quad (4.3)$$

The residual chemical potential is given by the difference between the total and the ideal-gas chemical potentials at the same state conditions. The residual chemical potentials of the solute (2) and the solvent (1) can be written as follows:

$$\beta\mu_1^{\text{res}, \delta}(p, \beta, x_1) = \ln \left(\frac{x_1 \gamma_1 f_1}{x_1 p} \right)^\delta \quad (4.4)$$

$$\beta\mu_2^{\text{res}, \delta}(p, \beta, x_2) = \ln \left(\frac{x_2 \gamma_2^{\text{HL}} H_2}{x_2 p} \right)^\delta \quad (4.5)$$

The activity coefficients γ_1 and γ_2^{HL} approach one in the limit of infinite dilution. The mixture fugacity of component two is described by Henry's law (HL) [2] as the fugacity of pure component two in phase δ does not always exist.

Subtraction of Eq. (4.4) from (4.5) gives the ratio of the solvent's fugacity over Henry's constant, expressed in terms of the residual chemical-potential difference $\Delta\mu_{21}^{\text{res}, \delta} = \mu_2^{\text{res}, \delta} - \mu_1^{\text{res}, \delta}$. In the limit of $\xi_2, x_2 \rightarrow 0$, the ratio f_1/H_2 is given by:

$$\left(\frac{f_1}{H_2}\right)^\delta = \lim_{x_2 \rightarrow 0} \exp\{-\beta\Delta\mu_{21}^{\text{res}, \delta}\} \quad (4.6)$$

Obviously, the residual chemical-potential difference is the key quantity that determines f_1/H_2 . In case no former simulation data are available, the ratios f_1/H_2 can be computed as an ensemble average by perturbing the $N_1N_2p\beta$ -ensemble.

There are several perturbation methods that can be used to compute the residual chemical-potential difference at infinite dilution, differing in complexity, accuracy, and CPU demands. As will be pointed out in the next sections, the perturbation method should be selected carefully, since not every perturbation method is capable of providing the correct chemical-potential difference.

The simplest method is to separately determine μ_2^{res} and μ_1^{res} by Widom trial insertions [3, 4] in the isobaric-isothermal ensemble. This method is described in Section 4.2.1. Mehta and Kofke [1] compute $\Delta\mu_{21}^{\text{res}}$ by performing trial identity changes in a pure solvent simulation. This perturbation method, which is called *difference method*, is discussed in Section 4.2.2. In Section 4.2.3, the *overlapping-distributions difference method* is used. This method reveals the shortcomings of the other methods. The different perturbation methods will be compared and assessed in Section 4.2.5.

In Sections 4.2.1 to 4.2.3, binary systems at constant temperature and constant pressure are considered. The semi-classical isobaric-isothermal partition function is given by:

$$\Delta_{N_1N_2p\beta}^* = \frac{(z_1q_1)^{N_1} (z_2q_2)^{N_2} N}{N_1!N_2!} \int dV \frac{V^N}{V} \exp\{-\beta pV\} \cdot \int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}\left(\Psi_1^{N_1}, \Psi_2^{N_2}\right)\right\} \quad (4.7)$$

where z_i is a normalization constant of the intramolecular Boltzmann distribution without intramolecular Lennard-Jones and Coulombic contributions [See Chapter 2, Eq. (2.35)], N_i is the number of molecules of type i , and Ψ_i is a random conformation generated from the intramolecular Boltzmann distribution. The situation of infinite dilution is approximated by the smallest possible mole fraction in a binary isothermal-isobaric simulation, *i.e.* $x_2 = 1/N$.

4.2.1. Residual chemical-potential difference in the isothermal-isobaric ensemble from Widom test-molecule insertions

The chemical potential has always been in the middle of interest since it is the key variable in the computation of phase equilibria. The best-known method to compute the chemical potential in a simulation was developed by Widom in 1963 [3]. In Widom's test-molecule method, *test molecules* are inserted into the simulation box. These test molecules do not influence the evolution of the system. The chemical potential is computed by ensemble-averaging the Boltzmann factors corresponding to the interaction energies of the test molecules. The method performs well in case sufficient trial insertions of molecules find no

overlap with the existing molecules in the system. Clearly, this condition can hardly be met in high-density systems with large molecules.

Many modifications of the original Widom method and other indirect methods have been developed over the years to overcome the sampling problems at high density. Among these methods are Kirkwood's coupling-parameter method [5-9], expanded-ensemble and gradual-insertion methods [6, 10-21] the chain-increment method [22-25], the acceptance-ratio method [26], the umbrella-sampling method [27-29], the $f - g$ sampling method [30-33], and methods that enhance the probability to find a preferential insertion position for the test molecule [34-38]. A method that is used to find a preferential insertion conformation/position is Rosenbluth sampling [39] with multiple-first-bead insertions [40]. This method is appropriate for the computation of the chemical potential of chain molecules [32, 41-44].

This section focuses on the computation of individual residual chemical potentials by Widom test-molecule insertions. The residual chemical potentials are computed by performing test insertions of solvent and solute molecules into a pure solvent [3]. The difference $\Delta\mu_{21}^{\text{res}}$ is computed by subtracting the solvent's residual chemical potential from the solute's residual chemical potential. In order to enhance the sampling of the test-molecule's Boltzmann factor at high densities, the multiple-first-bead Rosenbluth sampling scheme is used. Below, the derivation of the residual chemical potential in a binary mixture is given. The extension to Rosenbluth sampling is given in literature [45].

In order to facilitate the computation, the residual chemical-potential difference is divided into two contributions. The first contribution is a quasi-residual chemical potential which is referenced to the ideal-chain gas (superscript ic) and the second contribution is the ideal-gas (superscript ig) chemical potential which is referenced to the ideal-chain gas [46]:

$$\Delta\mu_{21}^{\text{res, Widom}} = \left(\Delta\mu_{21}^{\text{Widom}} - \Delta\mu_{21}^{\text{ic}} \right) - \left(\Delta\mu_{21}^{\text{ig}} - \Delta\mu_{21}^{\text{ic}} \right) \quad (4.8)$$

An ideal gas has contributions from all intramolecular interactions (superscript intra + LJ_{int} + C_{int}). On the other hand, the ideal-chain gas lacks intramolecular Lennard-Jones and Coulombic interactions. The partition functions of the ideal gas, Z_i^{ig} , and of the ideal-chain gas, Z_i^{ic} , at constant temperature and pressure are given by:

$$Z_i^{\text{ig}} = \frac{q_i}{\beta p} \int J_i d\mathbf{q}^{2n_i-2} ds^3 \exp \left\{ -\beta \mathcal{U}_i^{\text{intra} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\mathbf{q}^{2n_i-2}) \right\} \quad (4.9)$$

$$Z_i^{\text{ic}} = \frac{q_i}{\beta p} \int J_i d\mathbf{q}^{2n_i-2} ds^3 \exp \left\{ -\beta \mathcal{U}_i^{\text{intra}}(\mathbf{q}^{2n_i-2}) \right\} \quad (4.10)$$

Insertion of Eq. (2.35) into Eqs. (4.9) and (4.10) gives:

$$Z_i^{\text{ic}} = \frac{q_i z_i}{\beta p} \int d\Psi_i = \frac{q_i z_i}{\beta p} \quad (4.11)$$

$$Z_i^{\text{ig}} = \frac{q_i z_i}{\beta p} \int d\Psi_i \exp \left\{ -\beta \mathcal{U}_i^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\} \quad (4.12)$$

Having found the ideal-gas and ideal-chain-gas partition functions, one is able to give a relation for the second contribution on the right-hand side in Eq. (4.8):

$$\begin{aligned}
\beta(\Delta\mu_{21}^{\text{ig}} - \Delta\mu_{21}^{\text{ic}}) &= -\ln\left(\frac{Z_1^{\text{ic}}Z_2^{\text{ig}}}{Z_2^{\text{ic}}Z_1^{\text{ig}}}\right) \\
&= -\ln\left(\frac{\int d\Psi_2 \exp\{-\beta\Delta\mathcal{U}_2^{\text{LJint} + \text{Cint}}\}}{\int d\Psi_2} \frac{\int d\Psi_1}{\int d\Psi_1 \exp\{-\beta\Delta\mathcal{U}_1^{\text{LJint} + \text{Cint}}\}}\right) \quad (4.13) \\
&= -\ln\left(\frac{\langle \mathcal{W}_{\text{IG},2}^{\text{LJint} + \text{Cint}} \rangle}{\langle \mathcal{W}_{\text{IG},1}^{\text{LJint} + \text{Cint}} \rangle}\right)
\end{aligned}$$

where $\langle \mathcal{W}_{\text{IG},i}^{\text{LJint} + \text{Cint}} \rangle$ is the normalized Rosenbluth factor of an isolated chain of component i with intramolecular Lennard-Jones and Coulombic interactions [See Chapter 2, Section 2.5.4]. The first contribution to the residual chemical potential-difference in Eq. (4.8), in the limit $x_2 \rightarrow 0$, is given by:

$$\begin{aligned}
\lim_{x_2 \rightarrow 0} \beta(\Delta\mu_{21}^{\text{Widom}} - \Delta\mu_{21}^{\text{ic}}) &\approx -\ln\left\{\frac{\beta p}{N_1} \int d\Psi_2^{\text{test}} \left\langle V \exp\{-\beta\mathcal{U}_{\text{test},2}^{\text{ext} + \text{LJint} + \text{Cint}}\} \right\rangle_{N_1 p \beta}\right\} \\
&\quad + \ln\left\{\frac{\beta p}{N_1 + 1} \int d\Psi_1^{\text{test}} \left\langle V \exp\{-\beta\mathcal{U}_{\text{test},1}^{\text{ext} + \text{LJint} + \text{Cint}}\} \right\rangle_{N_1 p \beta}\right\} \quad (4.14) \\
&= -\ln\left\{\frac{N_1 + 1}{N_1} \frac{\int d\Psi_2^{\text{test}} \left\langle V \exp\{-\beta\mathcal{U}_{\text{test},2}^{\text{ext} + \text{LJint} + \text{Cint}}\} \right\rangle_{N_1 p \beta}}{\int d\Psi_1^{\text{test}} \left\langle V \exp\{-\beta\mathcal{U}_{\text{test},1}^{\text{ext} + \text{LJint} + \text{Cint}}\} \right\rangle_{N_1 p \beta}}\right\}
\end{aligned}$$

where $\mathcal{U}_{\text{test},i}^{\text{ext} + \text{LJint} + \text{Cint}}$ is the sum of the intermolecular and intramolecular Lennard-Jones and Coulombic energies of an inserted *test* molecule of identity i that does not take part in the evolution of the system of *real* molecules. The ensemble averages in Eq. (4.14) are obtained from a single pure-component isobaric-isothermal simulation. The ratio f_1/H_2 , computed from Widom's test-molecule method, can be related to ensemble averages as follows:

$$\left(\frac{f_1}{H_2}\right)^{\text{Widom}} \approx \frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJint} + \text{Cint}} \rangle}{\langle \mathcal{W}_{\text{IG},2}^{\text{LJint} + \text{Cint}} \rangle} \left(\frac{N_1 + 1}{N_1}\right) \frac{\int d\Psi_2 \left\langle V \exp\{-\beta\mathcal{U}_{\text{test},2}^{\text{ext} + \text{LJint} + \text{Cint}}\} \right\rangle_{N_1 p \beta}}{\int d\Psi_1 \left\langle V \exp\{-\beta\mathcal{U}_{\text{test},1}^{\text{ext} + \text{LJint} + \text{Cint}}\} \right\rangle_{N_1 p \beta}} \quad (4.15)$$

4.2.2. Residual chemical-potential difference in the isobaric-isothermal ensemble from the difference method

One of the perturbation methods that can be employed to compute the residual chemical-potential difference is the *difference* or *swap method*. It is more easily implemented than Widom's test-molecule method. The difference method was introduced by Shing [47, 48] in

order to compute activity coefficients and residual chemical-potential differences at infinite dilution. The method can be extended to non-diluted mixtures and is also suitable for the computation of other partial molar properties [49-57]. The chemical-potential difference at constant temperature and constant pressure can be approximated by:

$$\beta\Delta\mu_{21} \approx -\ln\left(\frac{\Delta_{N_1-1, N_2+1, p\beta}}{\Delta_{N_1 N_2 p\beta}}\right) \quad (4.16)$$

The derivations given by Shing [47, 48] and by Sindzingre *et al.* [49, 57] are applicable to systems with spherical or rigid molecules. Here, a straightforward derivation is given for multi-site molecules with intramolecular Lennard-Jones and Coulombic interactions. Employment of the same division as in Eq. (4.8) gives the following result for the residual chemical-potential difference:

$$\begin{aligned} \exp\left(-\beta\Delta\mu_{21}^{\text{res, Diff}}\right) &\approx \\ &\left\{ \frac{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJint+Cint}} \right\rangle}{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJint+Cint}} \right\rangle} \int \frac{N}{V} V^N \exp\{-\beta pV\} dV \int d\Psi_1^{N_1} d\Psi_2^{N_2} d\Psi_2^{\text{test}} \cdot \right. \\ &\left. \frac{\left\langle \exp\left\{-\beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext+LJint+Cint}}\left(\Psi_1^{N_1}, \Psi_2^{N_2}, \Psi_2^{\text{test}}\right)\right\} \exp\left\{-\beta\mathcal{U}^{\text{ext+LJint+Cint}}\left(\Psi_1^{N_1}, \Psi_2^{N_2}\right)\right\}}{\int \frac{N}{V} V^N \exp\{-\beta pV\} dV \int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}^{\text{ext+LJint+Cint}}\left(\Psi_1^{N_1}, \Psi_2^{N_2}\right)\right\}} \right. \quad (4.17) \\ &= \frac{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJint+Cint}} \right\rangle}{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJint+Cint}} \right\rangle} \int d\Psi_2^{\text{test}} \left\langle \exp\left\{-\beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext+LJint+Cint}}\left(\Psi_1^{N_1}, \Psi_2^{N_2}, \Psi_2^{\text{test}}\right)\right\} \right\rangle_{N_1 N_2 p\beta} \end{aligned}$$

where Ψ_2^{test} is a random test-molecule conformation, generated from the intramolecular Boltzmann probability of component two without intramolecular Lennard-Jones and Coulombic contributions. Clearly, the difference method is a Widom-like free-energy perturbation method, which gives the free energy between a reference system and a system that has a composition perturbation as an ensemble average taken over the reference system.

The simulation procedure is as follows. A simulation at constant $N_1 N_2 p\beta$ is performed. At regular intervals, one of the solvent (1) molecules is converted into a solute (2) molecule and the exchange energy $\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext+LJint+Cint}}$ is noted. The new molecule with identity two and the old molecule with identity one have the same overall centre-of-mass coordinates. The molecule is converted back before the simulation proceeds.

In the limit of pure component one, one gets the following relation for the residual chemical-potential difference:

$$\begin{aligned} \lim_{x_2 \rightarrow 0} \exp\left(-\beta\Delta\mu_{21}^{\text{res, Diff}}\right) &\approx \\ &\frac{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJint+Cint}} \right\rangle}{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJint+Cint}} \right\rangle} \int d\Psi_2^{\text{test}} \left\langle \exp\left\{-\beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext+LJint+Cint}}\left(\Psi_1^{N_1}, \Psi_2^{\text{test}}\right)\right\} \right\rangle_{N_1 p\beta} \quad (4.18) \end{aligned}$$

The result presented here, and a similar relation presented by Mehta and Kofke [1], are only valid in the canonical ensemble and the isobaric-isothermal ensemble. In Section 4.2.6, a relation will be presented that is valid in the Gibbs ensemble.

The ratio f_1/H_2 , computed from the difference method, can be related to an ensemble average obtained from a pure-component isobaric-isothermal simulation as follows:

$$\left(\frac{f_1}{H_2}\right)^{\text{Diff}} \approx \frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle}{\langle \mathcal{W}_{\text{IG},2}^{\text{LJ}_{\text{int}}+\text{C}_{\text{int}}} \rangle} \int d\Psi_2^{\text{test}} \left\langle \exp \left\{ -\beta \Delta \mathcal{U}_{1 \rightarrow 2}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \left(\Psi_1^{N_1}, \Psi_2^{\text{test}} \right) \right\} \right\rangle_{N_1 p \beta} \quad (4.19)$$

4.2.3. Residual chemical-potential difference in the isothermal-isobaric ensemble from the overlapping-distributions difference method

Shing and Gubbins [30, 33] and Powles *et al.* [31] showed how to combine test-molecule insertions and deletions for computing the chemical potential in a Monte Carlo simulation. The Shing-Gubbins method, which is also called $f-g$ method, is a special case of the Curve-Fitting method which was developed by Bennett [26] for computing the free-energy difference between two systems. In the original Widom test-molecule method [3, 4], a test molecule (perturbation) is inserted in the (unperturbed) reference system. The Boltzmann factor of the perturbation energy is then expressed as an ensemble average in the reference system. Thus, the reference system is simulated with regular trial insertions of test molecules. In contrast with the Widom method, the overlapping-distributions method uses two independent simulations, one for the unperturbed system (\otimes) and one for the perturbed system (\odot). During the course of the simulation that is performed in the unperturbed system, the energy change that results from perturbing it by adding a test molecule is sampled. In the perturbed system, the energy change that is needed to obtain the unperturbed system is noticed. Both simulations provide an energy distribution. The unperturbed system gives an energy distribution for a test molecule, while the perturbed system provides an energy distribution for a real molecule. The computation of the residual chemical potential is performed by a clever comparison of the two energy distributions.

Sindzingre *et al.* [49, 57] suggested the combination of the difference method and the $f-g$ method in order to get more accurate results for residual chemical-potential differences and other partial molar properties. Liu [50] and Liu and Govind [51, 52] actually performed computations with this combination of methods. This extended method will be called *overlapping-distributions difference method* in what follows.

Below, a derivation is given for the overlapping distributions of the difference method. A similar derivation for the residual chemical potential is given by Frenkel and Smit [45]. The unperturbed system has N_1 molecules with identity one, N_2 molecules with identity two and an ideal-chain molecule with identity two. The perturbed system comprises $N_1 - 1$ molecules with identity one, $N_2 + 1$ molecules with identity two and an ideal-chain molecule with identity one. As the residual chemical-potential difference in the $Np\beta$ -ensemble does not depend on the fluctuating volume, the volume dependence can be omitted in the derivation given below.

The probability density to find an energy difference $\mathcal{X} = \beta(\mathcal{U}_\ominus^{\text{ext}+\text{LJint}+\text{Cint}} - \mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}})$ when simulating the unperturbed system is:

$$\frac{\pi_\otimes(\mathcal{X})}{Z_2^{\text{ic}}} = \frac{\int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2})\right\} \delta\left(\mathcal{X} - \beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext}+\text{LJint}+\text{Cint}}\right)}{q_\otimes} \quad (4.20)$$

where q_\otimes is a normalization constant. The counterpart of Eq. (4.20) in the perturbed system is given by:

$$\pi_\ominus(\mathcal{X}) = \frac{\int d\Psi_2^{N_2+1} d\Psi_1^{N_1-1} \exp\left\{-\beta\mathcal{U}_\ominus^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1-1}, \Psi_2^{N_2+1})\right\} \delta\left(\mathcal{X} - \beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext}+\text{LJint}+\text{Cint}}\right)}{q_\ominus} Z_1^{\text{ic}} \quad (4.21)$$

The δ -function in Eq. (4.21) allows substituting $\mathcal{X} + \beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}$ for $\beta\mathcal{U}_\ominus^{\text{ext}+\text{LJint}+\text{Cint}}$:

$$\begin{aligned} \frac{q_\ominus}{q_\otimes} \pi_\ominus(\mathcal{X}) &= \frac{\int d\Psi_2^{N_2+1} d\Psi_1^{N_1-1} \exp\left\{-\beta\mathcal{U}_\ominus^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1-1}, \Psi_2^{N_2+1})\right\} \delta\left(\mathcal{X} - \beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext}+\text{LJint}+\text{Cint}}\right)}{\int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2})\right\} \int d\Psi_2 / \int d\Psi_1} \\ &= \frac{\int d\Psi_2^{N_2+1} d\Psi_1^{N_1-1} \exp\left\{-\beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1-1}, \Psi_2^{N_2+1})\right\} \delta\left(\mathcal{X} - \beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext}+\text{LJint}+\text{Cint}}\right) \int d\Psi_2}{\exp(+\mathcal{X}) \int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2})\right\} \int d\Psi_2} \\ &= \exp(-\mathcal{X}) \pi_\otimes(\mathcal{X}) \end{aligned} \quad (4.22)$$

where

$$\begin{aligned} \frac{q_\ominus}{q_\otimes} &= \frac{\int d\Psi_1^{N_1-1} d\Psi_2^{N_2+1} \exp\left\{-\beta\mathcal{U}_\ominus^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1-1}, \Psi_2^{N_2+1})\right\} \int d\Psi_1}{\int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2})\right\} \int d\Psi_2} \\ &= \frac{\int d\Psi_1^{N_1} d\Psi_2^{N_2+1} \exp\left\{-\beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2+1}) - \beta\mathcal{U}_\ominus^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2})\right\}}{\int d\Psi_1^{N_1} d\Psi_2^{N_2} \exp\left\{-\beta\mathcal{U}_\otimes^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2})\right\}} \\ &= \int d\Psi_2^{\text{test}} \left\langle \exp\left\{-\beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext}+\text{LJint}+\text{Cint}}(\Psi_1^{N_1}, \Psi_2^{N_2}, \Psi_2^{\text{test}})\right\} \right\rangle_{N_1 N_2 p \beta} \\ &= \exp\left(-\beta \left[\Delta\mu_{21} - \Delta\mu_{21}^{\text{ic}} \right]\right) \end{aligned} \quad (4.23)$$

The limit of infinite dilution is taken by setting N_2 in the reference system to zero and to one in the perturbed system. Combination of Eqs. (4.8), (4.13), (4.22), and (4.23) gives the following relation between the energy distributions:

$$\begin{aligned}\ln \pi_{\odot}(\mathcal{X}) &= \beta \left[\Delta\mu_{21} - \Delta\mu_{21}^{\text{ic}} \right] - \mathcal{X} + \ln \pi_{\otimes}(\mathcal{X}) \\ &= \beta \Delta\mu_{21}^{\text{res, OLD}} - \ln \left(\frac{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle}{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle} \right) - \mathcal{X} + \ln \pi_{\otimes}(\mathcal{X})\end{aligned}\quad (4.24)$$

The residual chemical-potential difference can now be obtained from a graphical analysis [45]. In order to do so, it is convenient to define two new distributions $f_{\otimes}(\mathcal{X})$ and $f_{\odot}(\mathcal{X})$:

$$\begin{aligned}f_{\otimes}(\mathcal{X}) &= \ln \pi_{\otimes}(\mathcal{X}) - \frac{1}{2} \left\{ \ln \frac{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle}{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle} + \mathcal{X} \right\} \\ f_{\odot}(\mathcal{X}) &= \ln \pi_{\odot}(\mathcal{X}) + \frac{1}{2} \left\{ \ln \frac{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle}{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle} + \mathcal{X} \right\}\end{aligned}\quad (4.25)$$

The possibility to compute the residual-chemical potential difference from the overlapping-distributions difference method depends on two factors. At first, as the name of the method indicates, the two distributions should have a region of overlap. The second condition is that subtraction of $f_{\otimes}(\mathcal{X})$ from $f_{\odot}(\mathcal{X})$ should give a constant value of $\beta \Delta\mu_{21}^{\text{res, OLD}}$ over a range of \mathcal{X} -values. The ratio $(f_1/H_2)^{\text{OLD}}$, computed from the overlapping-distributions difference method, is obtained from Eq. (4.6) and from $\beta \Delta\mu_{21}^{\text{res, OLD}}$ computed in the limit of infinite dilution.

4.2.4. Simulation details

In all $Np\beta$ -simulations, molecular translations and rotations were respectively force-biased and torque-biased. The parameters λ and ζ were both fixed at 0.5. The usual cubic periodic-boundary conditions with the minimum-image convention were employed [58]. The choice of intramolecular and intermolecular potential is extensively described in Chapter 5. The usual long-range corrections were applied [58] with a potential cutoff of at least three times the largest Lennard-Jones collision diameter σ in the simulated system. The cutoff was checked against half the box length after every successful volume change. The Ewald screening parameter and the upper bound in the reciprocal space were fixed at $\alpha = 5.6$ and $k_{\text{max}} = 6$ [58]. Maximum displacements used for translations, rotations, and volume changes were adjusted during the equilibration period of the simulation to yield a fraction of accepted MC trial moves of about 50%. Different maximum displacements were used for the liquid and the vapour phase. Liquid phases contained $N^{\text{L}} = 300$ molecules and vapour phases $N^{\text{V}} = 100$ molecules. After 5000 equilibration MC cycles, 100000 production MC cycles were performed. Every MC cycle consisted of $0.7N^i$ translations, $0.3N^i$ rotations, and 1 volume change, where i represents the liquid or the vapour phase. In addition to these moves, the Widom method [Section 4.2.1] included $0.5N^i$ trial insertions for molecules of both

components, the difference method [Section 4.2.2] included N^i trial-identity changes, and the overlapping-distributions difference method [Section 4.2.3] included $0.5N^i$ trial removals of molecules and $0.5N^i$ trial insertions of molecules. The distributions f_{\odot} and f_{\otimes} of the overlapping-distributions difference method contained 201 bins. The bin width was determined during the equilibration phase of the simulation.

Initially, the statistical uncertainty of the ensemble averages was computed with the block-averaging method of Flyvbjerg and Petersen [59]. This analysis provided an indication of the minimum block length needed to obtain uncorrelated blocks of samples. It turned out that blocks of a few thousands MC cycles were uncorrelated. The statistical errors were computed from application of the bootstrap method [60] to regularly saved blocks of 2500 MC cycles.

4.2.5. Comparison of the different methods

A comparison of the Widom method, difference method and overlapping-distributions difference method has already been performed by Liu [50] and by Liu and Govind [51, 52]. Their analyses were restricted to mixtures of similar single-site Lennard-Jones components. They only applied the difference method in the direction of increasing molecular size. It turned out that the three methods provided comparable results. The difference method had one order higher precision than the Widom method. For simple mixtures, the overlapping-distributions difference method had no advantage over the simple difference method. However, the overlapping-distributions difference method provided accurate results at higher densities than the other methods did.

Here, the analyses performed by Liu [50] and by Liu and Govind [51, 52] are extended to molecules with spatial configurations. Subsequently, the failure of the difference method when a large molecule is swapped with a smaller one is explained.

Numerical values of the liquid-phase residual chemical-potential differences and the accompanying ratios f_1/H_2 for a couple of binary systems at vapour-liquid coexistence are given in Table 4.1 and Table 4.2. The methods with which the initial points were computed are described in Section 4.3. The residual chemical-potential difference from the overlapping-distributions difference method, along with the statistical error, has been determined by averaging the values in the region of overlap of the energy distributions.

Table 4.1: Liquid-phase residual chemical-potential differences computed with different methods at vapour-liquid coexistence

System	Condition	$\beta\Delta\mu_{21}^{\text{res}}$ from different methods		
		Widom	Difference method	Overlapping distributions
C ₂ H ₆ (1)/CH ₄ (2)	250 K, 1.29 MPa	2.24(2)	2.54(11)	2.26(5)
CH ₄ (1)/CO ₂ (2)	150 K, 1.13 MPa	-1.89(2)	-1.70(3)	-1.90(4)
C ₃ H ₈ (1)/CH ₄ (2)	270 K, 0.433 MPa	3.57(5)	5.67(7)	3.6(1)
CHF ₃ (1)/CO ₂ (2)	254 K, 1.42 MPa	0.72(2)	1.31(8)	0.69(4)
CO ₂ (1)/CH ₂ F ₂ (2)	244.3 K, 1.61 MPa	-0.95(7)	-0.5(2)	-0.96(6)

Liquid-phase densities at the conditions given in Table 4.1 are moderate. Accordingly, the Widom test-molecule method can be efficiently used to compute the residual chemical-potential difference with high precision within reasonable time. The simulation results for the overlapping-distributions difference method agree within their statistical errors with the results obtained from the Widom method. In contrast to the simulation results of Liu [50], the

standard deviation of the results computed with the Widom method is smaller than the standard deviation computed from the other methods. The reason for the lower standard deviation is the use of the Rosenbluth sampling scheme. As the methods are quite different, a direct comparison based on the statistical error of both methods is hardly possible. One could compare the statistical errors of both methods for a given amount of CPU time. However, this has not been performed in this work.

Table 4.2: Liquid-phase ratios of solvent’s fugacity over solute’s Henry’s constant computed with different methods at vapour-liquid coexistence

System	Condition	f_1/H_2 from different methods		
		Widom	Difference method	Overlapping distributions
$C_2H_6(1)/CH_4(2)$	250 K, 1.29 MPa	0.107(2)	0.08(10)	0.105(6)
$CH_4(1)/CO_2(2)$	150 K, 1.13 MPa	6.6(1)	5.4(1)	6.7(2)
$C_3H_8(1)/CH_4(2)$	270 K, 0.433 MPa	0.028(1)	0.0034(2)	0.028(3)
$CHF_3(1)/CO_2(2)$	254 K, 1.42 MPa	0.48(1)	0.27(2)	0.50(2)
$CO_2(1)/CH_2F_2(2)$	244.3 K, 1.61 MPa	2.6(2)	1.6(3)	2.6(2)

Special care is needed when Rosenbluth sampling is implemented in the computation of the residual chemical potential [45, 61, 62]. The results of Rosenbluth sampling are only identical to those of non-biased Boltzmann sampling in the limit of an infinitely long simulation. Instead of generating a molecular conformation with a probability proportional to its Boltzmann weight, the Rosenbluth scheme accepts the conformation based on its Rosenbluth weight. Especially for large molecules, the elongated conformations are sampled poorly. However, as this work involves only small molecules, this limitation is less relevant.

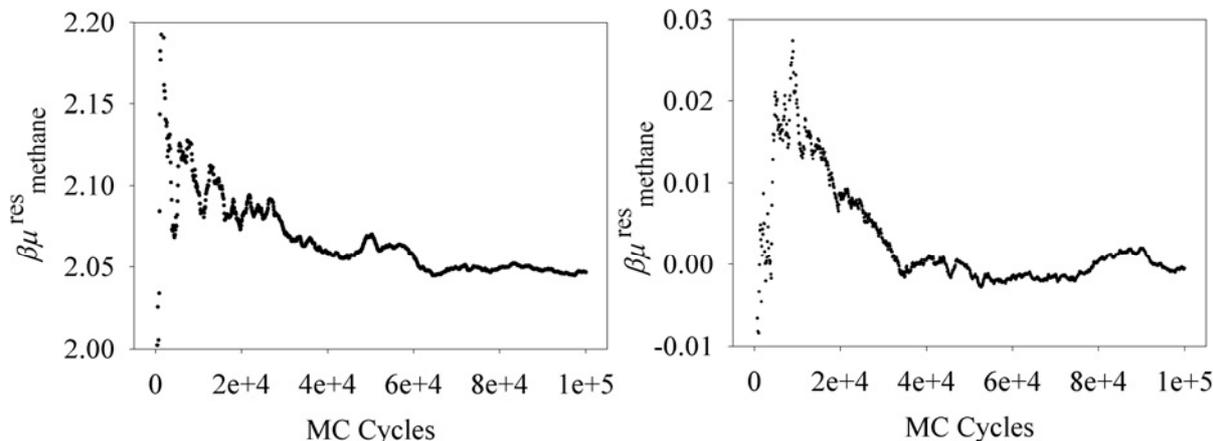


Figure 4.1: Moving average of the effective residual chemical potential of methane, infinitely diluted in ethane, at $T = 250$ K and $p = 1.29$ MPa . Left-hand picture: liquid phase. Right-hand picture: vapour phase.

An advantage of the Widom test-molecule method is its simplicity. Its application needs only minor modifications in a simulation program. However, its simplicity is cancelled out by the convergence characteristics of the method in systems at moderate to high densities or in systems with large molecules. At high density, most test-molecule insertions result in overlap with a corresponding negligible contribution to the ensemble average of the chemical

potential. Trial insertions with a high Boltzmann factor are rare, and once the system encounters such an insertion, the moving average of the chemical potential shows a jump. This jumpy behaviour is shown graphically in Figure 4.1.

A second drawback of the Widom approach is that the ensemble averages of both residual chemical potentials are computed individually; they both have a statistical uncertainty. Furthermore, several studies of the chemical potential of Lennard-Jones spheres in infinite periodic systems indicated that this quantity is rather system-size dependent [63-66]. Thus, one should always check whether the simulated system is large enough in order to avoid system-size dependence or one should apply a finite-size correction to the simulated chemical potential [63, 67].

As stated in Section 4.2.3, the $f-g$ sampling method of Shing and Gubbins [30, 33] has been developed to compute the residual chemical potential. As opposed to the Widom approach, the $f-g$ sampling method, or overlapping-distributions method, has built-in diagnostics; one can immediately see from the energy distributions if there is enough overlap to compute the residual chemical potential and if computation via Widom test-molecule insertions is possible. However, the computation of the residual chemical potential of the solvent and the solute via the $f-g$ sampling method requires four simulations per phase, which is computationally quite expensive.

The main advantage of the difference method is its easy applicability in a simulation. The computer code needs hardly to be changed. Since the virtual swap between an existing molecule and a molecule of the other component takes place at the same centre-of-mass coordinates, there is no serious insertion problem like in the Widom test-molecule method. Therefore, the method is applicable at higher densities than the Widom approach is. However, this advantage is less apparent when Rosenbluth sampling is used in the computation of the residual chemical potential.

The difference method determines the differences between residual chemical potentials as a single average. For mixtures consisting of not too dissimilar components, the statistical uncertainty of $\Delta\mu_{21}^{\text{res}} = \mu_2^{\text{res}} - \mu_1^{\text{res}}$, estimated with the difference method, can be much smaller than the sum of the statistical uncertainties of the individual values of μ_1^{res} and μ_2^{res} [48, 49, 56].

Another advantage of the difference method is that the simulated residual chemical-potential differences seem to be less size-dependent than their counterparts obtained from the Widom test-molecule method [48, 57].

As stated in refs. [57, 68], the difference method does not work when a large molecule is swapped with a smaller one. However, the method works quite well the other way around for systems comprising single-site Lennard-Jones molecules. This asymmetry of perturbation methods is well known. For example, when computing the chemical potential by perturbing the ensemble of interest, insertion of test molecules usually provides the right answer while removal of real molecules does not. To explain the asymmetry of the perturbation in the difference method, a *perturbed system* (subscript ‘per’) and an *unperturbed reference system* (subscript ‘ref’) are considered. The reference system comprises only molecules of type one, while the perturbed system has a molecule of type one replaced by a molecule of type two. Molecules of type one are larger than molecules of type two. Pictures of both systems are shown in Figure 4.2.

The relation between the perturbed system and the reference system is best explained by an analysis of phase-space overlap [69, 70]. The phase space accessible to the perturbed and the reference systems are given the symbols Γ_{per} and Γ_{ref} respectively. The region of phase

space that is important to the reference system is a subset of its total phase space and is indicated with Γ_{ref}^* . Configurations that contribute to Γ_{ref}^* have either a low energy or have a moderate energy and a high degeneracy.

Suppose one is interested in the ensemble average of the exponent $\exp\{-\beta[\mathcal{U}_{\text{per}} - \mathcal{U}_{\text{ref}}]\}$. Computation of the correct ensemble average of $\exp\{-\beta[\mathcal{U}_{\text{per}} - \mathcal{U}_{\text{ref}}]\}$ from a simulation in the reference system is only possible if Γ_{per}^* is a subset of Γ_{ref}^* or if the configurations important to both merely overlap. In these cases, a simulation in the reference system covers all configurations that contribute mostly to the ensemble average.

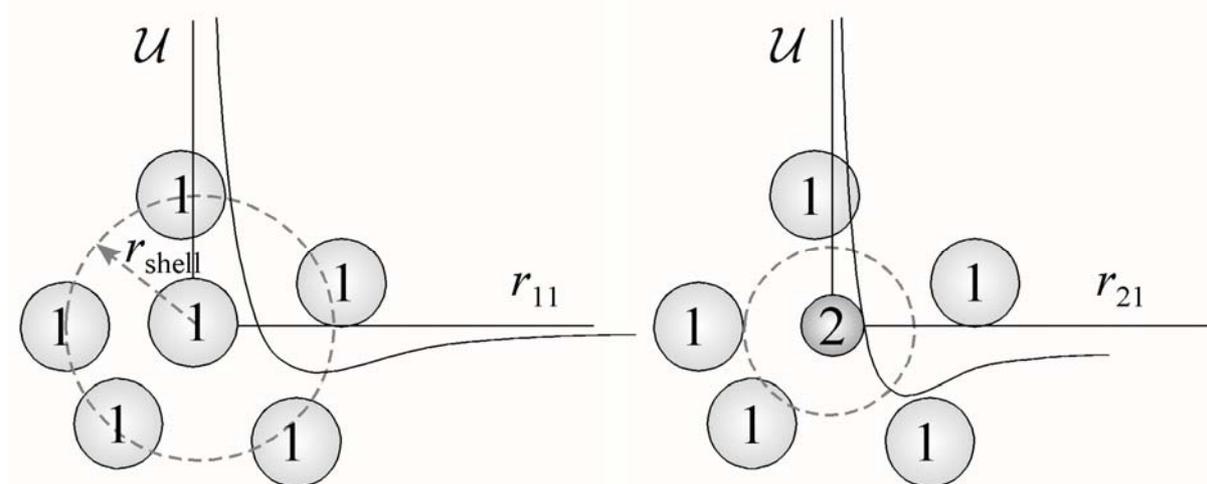


Figure 4.2: *Left-hand picture: favourable configuration of a reference system consisting of molecules with identity one. Right-hand picture: configuration of a perturbed system of molecules with identity one where the central molecule is replaced by a molecule with identity two. The curves represent the intermolecular potential.*

The asymmetry of the perturbation method is clear because if Γ_{per}^* is a subset of Γ_{ref}^* , the opposite cannot be true at the same time. Obviously, when the roles of the perturbed and the reference system are switched, the reference system never samples the larger part of the configurations that are important to the perturbed system. The configurations that are never sampled by the reference system may have large contributions to the ensemble average as \mathcal{U}_{per} is probably smaller than \mathcal{U}_{ref} for that part of Γ_{per}^* being outside of Γ_{ref}^* . The resulting *sampling problem is energetic in nature*; a part of the lower energies \mathcal{U}_{per} of the perturbed system is rarely sampled as \mathcal{U}_{ref} is high for those configurations.

The reason why a decrease of the molecular size in the difference method does not provide the correct ensemble average of $\exp\{-\beta[\mathcal{U}_{\text{per}} - \mathcal{U}_{\text{ref}}]\}$ has to do with *energetic sampling problems* as well. Favourable configurations in the reference system, Γ_{ref}^* , look like the one given in the left-hand picture of Figure 4.2; the first ‘shell’ of molecules is on average a distance $\langle r_{11} \rangle \approx r_{\text{shell}}$ apart from the central molecule. This distance roughly corresponds to the intermolecular distance at which the interaction potential for molecules of type one reaches its minimum value. Configurations with $\langle r_{11} \rangle \ll r_{\text{shell}}$ are highly improbable as they result in overlap of molecules with a corresponding high energy.

The central molecule in the perturbed system is replaced by a molecule of type two. The interaction potential between a molecule of type two and a molecule of type one has a minimum at a distance that is smaller than r_{shell} . This is a good approximation when molecule two is smaller than molecule one. The central molecule in the perturbed system (right-hand

picture in Figure 4.2) prefers the molecules of type one to be closer than r_{shell} . So, part of Γ_{per}^* is never sampled in the reference system. The part of Γ_{per}^* being never sampled has large contributions to the ensemble average. Thus, the ensemble average of $\Delta\mu_{21}^{\text{res, Diff}}$ will be *systematically higher* than the true value and the ensemble average of $(f_1/H_2)^{\text{Diff}}$ will be *systematically lower* than the ensemble average of $\Delta\mu_{21}^{\text{res, Diff}}$ when a large molecule is swapped with a smaller molecule.

Another drawback of the difference method is the jumpy convergence of the ensemble averages of the ratio $(f_1/H_2)^{\text{Diff}}$. This behaviour is shown graphically in Figure 4.3. The jumps in the moving ensemble averages are caused by the unlikely situation of finding an existing settled molecule with relatively high energy that is replaced by a molecule of the other identity with low energy.

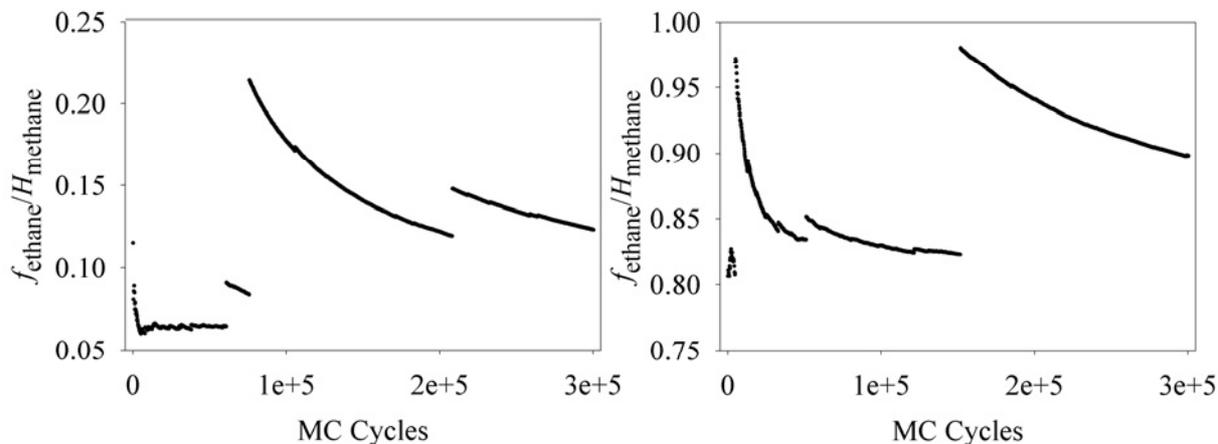


Figure 4.3: Convergence of the moving average of the ratio $f_{\text{ethane}}/H_{\text{methane}}$ at $T = 250\text{ K}$ and $p = 1.29\text{ MPa}$ in the liquid (left-hand picture) and the vapour phase (right-hand picture) versus number of Monte Carlo cycles. Notice the difference in scales on the vertical axis.

The discussion of the difference method in refs. [49, 51, 52, 57] is restricted to mixtures of single-site Lennard-Jones molecules. Liu [55] applied the method to rigid Lennard-Jones dumbbells. It is clear that the unilateral difference method should not be used when a large spherical molecule is swapped with a smaller one. Furthermore, notice that *larger* or *smaller* can hardly be defined for molecules with spatial configurations and intramolecular flexibility, as distances between interaction sites are now orientation-dependent. The systematic error may be lower in those cases.

Now, consider the simulation results in Table 4.1 and Table 4.2. Propane is a larger molecule than methane, although this difference is less clear than for purely spherical molecules. Theory predicts that the difference method is not able to predict the correct value of the residual chemical-potential difference when a propane molecule is swapped with a methane molecule. Indeed, as is clear from Table 4.1 and Table 4.2, $\Delta\mu_{21}^{\text{res, Diff}}$ is much larger than $\Delta\mu_{21}^{\text{res, Widom}}$ and $\Delta\mu_{21}^{\text{res, OI D}}$.

For the other couples of molecules in Table 4.1, it is not always clear which one is the larger or the smaller one. Moreover, the molecules in Table 4.1 have rather different force-field parameters. However, all results for $\Delta\mu_{21}^{\text{res, Diff}}$ are significantly larger than those obtained from the other methods. Thus, it is better not to use the difference method at all.

The overlapping-distributions difference method is applicable at higher densities than the conventional difference method is. As the overlapping-distributions difference method is a bilateral perturbation method, two simulations are needed [45]. Implementation in the simulation code is somewhat more complicated than it is for the Widom test-molecule method or the difference method.

An advantage of the overlapping-distributions difference method is the possibility to judge the reliability of the results. The energy distributions of the overlapping-distribution difference method immediately show if there is enough overlap to compute a reliable result from the difference between the distributions. One can also see from a separate energy distribution whether computation of the residual chemical-potential difference from the Widom test-molecule insertion method is possible or not.

The chemical potential computed from the Widom test-molecule method has considerable contributions from the poorly sampled low-energy tail of the energy distribution corresponding to the reference system. Since these low-energy configurations are rarely encountered in high-density systems, the moving average of the chemical potential shows the jumpy behaviour shown in Figure 4.1.

A nice feature of the overlapping-distributions difference method, compared to the Widom test-molecule method, is that the computed residual chemical-potential difference is computed by analysing the energy distributions once the simulation stopped. In case the region of overlap of the energy distributions f_{\otimes} and f_{\odot} is large enough, the residual chemical-potential difference can be computed from those parts of the distributions that are frequently sampled by both the perturbed and the reference system, resulting in a more accurate estimate.

Simulation results for the binary system propane/methane, obtained with the overlapping-distributions difference method, are shown in graphical form in Figure 4.4 and Figure 4.5. The reference system contains N propane molecules. The energy distribution f_{\otimes} is obtained by virtually changing the identity of a propane molecule to a methane molecule. On the other hand, the perturbed system contains $(N-1)$ propane molecules and one methane molecule. The energy distribution f_{\odot} is obtained by virtually changing the identity of the *real* methane molecule in the simulated system. The energy distribution f_{\odot} explores more of the attractive configurations of methane, while f_{\otimes} explores more of the less attractive and the repulsive configurations of methane.

From Figure 4.4, it is clear that there is considerable overlap between the energy distributions. This means that the perturbed and the unperturbed system sample to a large extent configurations that are representative for each other. The difference between the two distributions shows an acceptable range of constant $\Delta\mu_{21}^{\text{res,OID}}$ with small standard deviations. The corresponding ratio $(f_1/H_2)^{\text{OID}}$, along with the standard deviations, as a function of the perturbation energy is given in Figure 4.5. The actual value of $(f_1/H_2)^{\text{OID}}$ and its standard deviation are obtained by computing the average values in the region of constant $\Delta\mu_{21}^{\text{res,OID}}$. The overlapping-distributions for the binary system trifluoromethane/carbon dioxide are shown in graphical form in Figure 4.6.

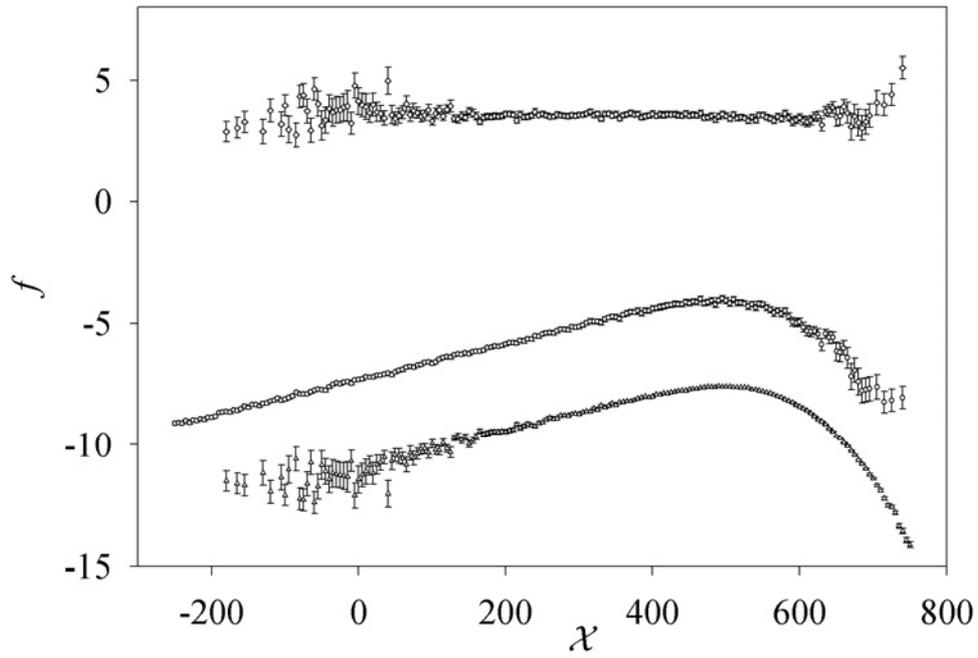


Figure 4.4: Overlapping distributions in the liquid phase of the system propane(1)/methane(2) at $T = 270 \text{ K}$ and $p = 0.4333 \text{ MPa}$. Upward triangles (lowest data series) and circles (data series in the middle) represent f_{\ominus} and f_{\odot} respectively. Diamonds (uppermost data series) indicate $\beta\Delta\mu_{21}^{\text{res,OID}}$.

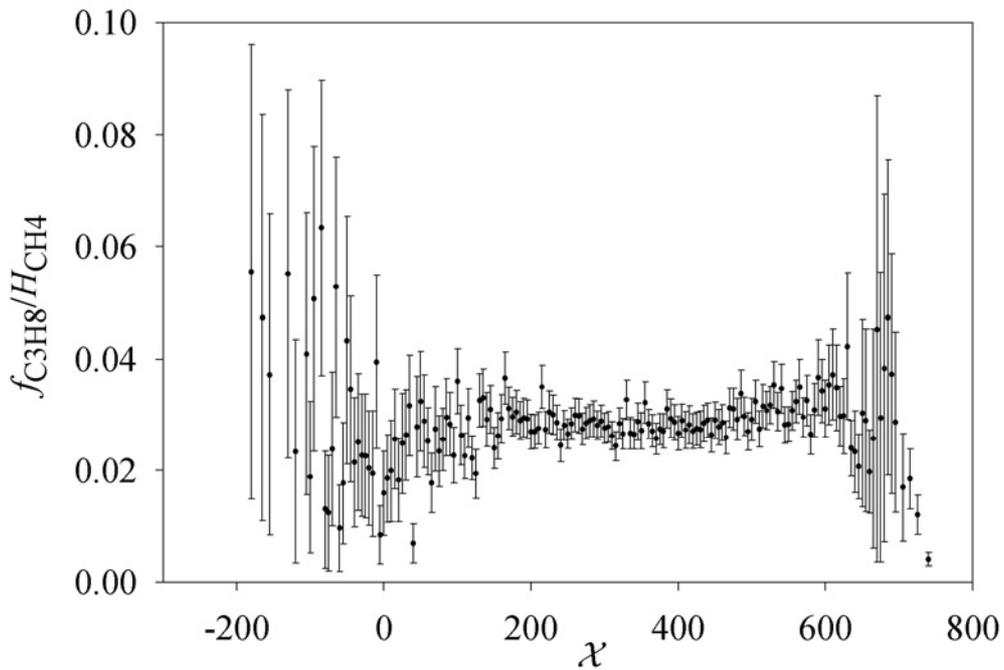


Figure 4.5: $f_{\text{propane}}/H_{\text{methane}}$ computed from the overlapping distributions in Figure 4.4.

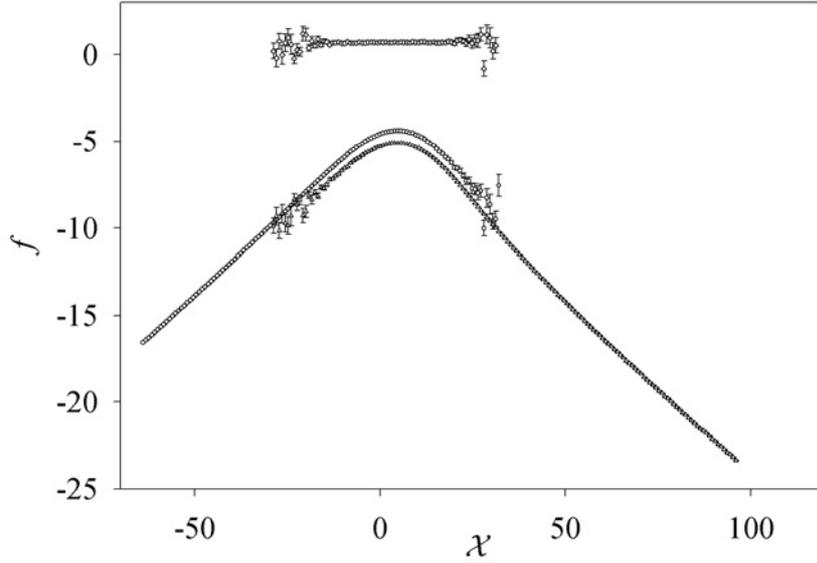


Figure 4.6: Overlapping distributions in the liquid phase of the system trifluoromethane(1)/carbon dioxide(2) at $T = 254 \text{ K}$ and $p = 1.423 \text{ MPa}$. Upward triangles (lowest data series) and circles (data series in the middle) represent f_{\otimes} and f_{\odot} respectively. Diamonds (uppermost data series) represent $\beta\Delta\mu_{21}^{\text{res, OLD}}$.

Finally, it is concluded that the overlapping-distributions difference method is the safest perturbation technique to compute the residual chemical-potential difference. The difference method can only be used when a small spherical molecule is swapped with a larger one. Computation of the residual chemical-potential difference by the Widom method with Rosenbluth sampling is a good alternative. However, the Widom method has no built-in diagnostics and extension with the overlapping-distributions method requires four computationally expensive independent simulations.

4.2.6. Residual chemical-potential difference in the Gibbs ensemble

Although it was concluded in Section 4.2.5 that the difference method has limited applicability in systems of molecules with spatial configurations, the author published simulation results for f_1/H_2 , obtained with the difference method in the Gibbs ensemble [71]. Computation of the ratio f_1/H_2 in the Gibbs ensemble is very attractive as a single simulation provides the initial slope and the initial point. Due to the fluctuating numbers of molecules in the coupled phases, the relation for the residual chemical-potential difference given in Eq. (4.17) is not valid in the Gibbs ensemble. The residual chemical-potential difference in phase γ , in the limit of infinite dilution is approximated by:

$$\lim_{x_2 \rightarrow 0} \exp\left(-\beta\Delta\mu_{21}^{\text{res, Diff}, \gamma}\right) \approx \frac{\left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJint} + \text{Cint}} \right\rangle \int d\Psi_2^{\text{test}} \left\langle N_1 \exp\left\{-\beta\Delta\mathcal{U}_{1 \rightarrow 2}^{\text{ext} + \text{LJint} + \text{Cint}}\left(\Psi_1^{N_1}, \Psi_2^{\text{test}}\right)\right\} \right\rangle_{\text{Gibbs}, \gamma}}{\left\langle \mathcal{W}_{\text{IG}, 2}^{\text{LJint} + \text{Cint}} \right\rangle \left\langle N_1 \right\rangle_{\text{Gibbs}, \gamma}} \quad (4.26)$$

Thus, the ratio f_1/H_2 , computed from the difference method in the Gibbs ensemble, is given by:

$$\left(\frac{f_1}{H_2}\right)^{\gamma, \text{Diff}} \approx \frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle \int d\Psi_2^{\text{test}} \langle N_1 \exp\left\{-\beta\Delta\mathcal{U}_{1\rightarrow 2}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi_1^{N_1}, \Psi_2^{N_2}, \Psi_2^{\text{test}})\right\} \rangle_{\text{Gibbs}, \gamma}}{\langle \mathcal{W}_{\text{IG},2}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle \langle N_1 \rangle_{\text{Gibbs}, \gamma}} \quad (4.27)$$

4.3. Predicting the initial point

4.3.1. Introduction

In Section 4.3, two methods are described that can be used to predict the initial point that is needed to start a binary Gibbs-Duhem integration: the $NV\beta$ -Gibbs ensemble method and the $Np\beta$ +test molecule method. All Gibbs-Duhem integrations in this work are performed isothermally. The initial point is a pure-component coexistence point. It is completely specified by the liquid and vapour densities at coexistence and by the vapour-liquid saturation pressure at the temperature of interest. The choice of the most appropriate method is mainly dictated by the liquid-phase density and by the density difference between the liquid and the vapour phase.

The $NV\beta$ -Gibbs ensemble has been described very shortly in Chapter 2. The simulation method is very convenient as it directly provides the coexisting densities and the saturation pressure at the temperature of interest. The method requires little expertise. However, its applications are limited to systems with liquid phases of moderate density. At high density, successful exchanges of molecules between the phases are rare; the majority of the insertions of molecules into a high-density phase result in overlap with other molecules.

Furthermore, the density difference between the phases limits the applicability of the Gibbs ensemble method. The $NV\beta$ -Gibbs ensemble cannot cope with a large density difference as the total volume and the total number of molecules of both phases are constant; a large density difference can only be obtained by putting (almost) all molecules in the liquid phase and by sending (almost) all the volume to the vapour phase. This problem can be overcome by taking a very large number of molecules. However, this would increase the length of the simulation with orders of magnitude.

In contrast with the quasi-direct Gibbs ensemble method, the $Np\beta$ +test molecule method is an indirect method. The basic idea of the $Np\beta$ +test molecule method and its extensions is the construction of the chemical potential as a function of pressure, temperature, and composition, around a reference point, and the subsequent calculation of the vapour-liquid saturation pressure from the criterion of equality of chemical potentials. Individual isothermal-isobaric simulations are performed for a liquid and a vapour phase at the thermodynamic conditions of the reference point.

As the simulations are performed at constant number of molecules, a large density difference between the liquid and the vapour phase does not limit the application of the method.

The Taylor series expansion for the chemical potential of component j , truncated after the first term, is given by:

$$\beta\mu_j(\beta, p, x_{1\dots c}) \approx \beta_0 [\mu_j]_0(\beta_0, p_0, [x_{1\dots c}]_0) + \left\{ \frac{\partial(\beta\mu_j)}{\partial p} \right\}_{\beta_0 p_0 [x_{1\dots c}]_0} (p - p_0) + \left\{ \frac{\partial(\beta\mu_j)}{\partial \beta} \right\}_{\beta_0 p_0 [x_{1\dots c}]_0} (\beta - \beta_0) + \sum_{i=1}^c \left\{ \frac{\partial(\beta\mu_j)}{\partial x_i} \right\}_{\beta_0 p_0 [x_{1\dots c}]_0} (x_i - [x_i]_0) \quad (4.28)$$

where the subscript 0 indicates the reference point about which the function is expanded. In the first applications of the method by Möller and Fischer [72] and by Lofti *et al.* [73], a Taylor series expansion in the pressure was used in order to compute vapour-liquid equilibria of pure components at a prescribed temperature.

The different modifications of the $Np\beta$ + test molecule method vary in the number of terms in the Taylor series expansion. For example, Boda *et al.* [74] used a Taylor series expansion in both temperature and pressure. Vrabec *et al.* [75-77] and Vrabec and Fischer [78] respectively extended the method to binary and ternary mixtures by including the composition dependence in the Taylor series expansion.

The test-molecule method has also been applied to the $NV\beta$ – ensemble [79, 80]. The $NV\beta$ + test molecule method does not use computationally expensive volume fluctuations. However, in order to compute a pure-component vapour-liquid coexistence point, the pressure needs to be computed from a canonical ensemble average.

In Section 4.3.2, the computation of the saturation pressure in the Gibbs ensemble is described. Two modifications of the pure-component $Np\beta$ + test molecule method are introduced in Sections 4.3.3 and 4.3.4. Simulation details and a comparison of the different methods are respectively given in Sections 4.3.5 and 4.3.6.

4.3.2. Pressure computation in the Gibbs ensemble

Ensemble averages of the densities at vapour-liquid coexistence in either phase are obtained in the usual way by sampling the density at regular intervals. The procedure to obtain the saturation pressure is a little bit more complicated.

The pressure is usually computed from the virial for pairwise-additive potentials [58]. The volume derivative of the Helmholtz energy is related to the pressure. The partition function of the $NV\beta$ – ensemble or the $NV\beta$ – Gibbs ensemble is linked with the Helmholtz energy. Thus, pressure can be computed by taking the volume derivative of the partition function. Scaling the coordinates of the molecules with the box-length of the cubic simulation box reveals the volume-dependence of the partition function.

The result for the canonical ensemble is [58]:

$$\begin{aligned}
\beta p &= -\left(\frac{\partial \beta A}{\partial V}\right)_{\beta, N} = \rho + \frac{\beta}{3V} \left\langle \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\alpha=1}^{n_i} \sum_{\gamma=1}^{n_j} \mathbf{f}_{i\alpha j\gamma} \cdot \mathbf{r}_{i\alpha j\gamma} \right\rangle_{NV\beta} \\
&= \rho - \frac{\beta}{V} \left\langle \underbrace{\frac{1}{3} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\alpha=1}^{n_i} \sum_{\gamma=1}^{n_j} \frac{d\mathcal{U}(r_{i\alpha j\gamma})}{dr_{i\alpha j\gamma}} \mathbf{r}_{i\alpha j\gamma} \cdot \mathbf{r}_{ij}}_{\text{Virial}} \right\rangle_{NV\beta}
\end{aligned} \tag{4.29}$$

where ρ is a molecular density, N is the number of molecules, n_i is the number of interaction sites on molecule i , \mathbf{r}_{ij} is the intermolecular-distance vector between the centres of mass of molecule i and j , and $\mathcal{U}(r)$ is the intermolecular potential. The variables $\mathbf{f}_{i\alpha j\gamma}$, $\mathbf{r}_{i\alpha j\gamma}$ and $r_{i\alpha j\gamma}$ respectively are the intermolecular-force vector, distance vector and absolute distance between sites $i\alpha$ and $j\gamma$. The use of the virial expression for the pressure in the $NV\beta$ -Gibbs ensemble requires the fluctuating volume to be incorporated in the ensemble average:

$$\beta p = \left\langle \rho - \frac{\beta}{3V} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\alpha=1}^{n_i} \sum_{\gamma=1}^{n_j} \frac{d\mathcal{U}(r_{i\alpha j\gamma})}{dr_{i\alpha j\gamma}} \frac{\mathbf{r}_{i\alpha j\gamma} \cdot \mathbf{r}_{ij}}{r_{i\alpha j\gamma}} \right\rangle_{\text{Gibbs}} \tag{4.30}$$

Usually, the liquid-phase virial is strongly fluctuating during the course of a simulation. For this reason, it is often impossible to get an ensemble average of the pressure in the liquid phase with small statistical uncertainty within reasonable time. On the other hand, the vapour-phase pressure at low density is dominated by the ideal-gas contribution and converges much faster. Therefore, one often takes the vapour-phase pressure as the true saturation pressure.

The derivative in relation (4.30) is easily derived for the Lennard-Jones force field. The derivative of the Coulombic contribution, which is computed with the Ewald summation method, is more complicated. Its derivation is given in Appendix C.

It is clear from Eqs. (4.29) and (4.30) that the virial pressure is computed from the forces between the molecules. As the computation of the forces takes place in the most ‘inner’ part of the computer code, computation of the pressure from the virial is computationally expensive. Fortunately, there is an alternative way to obtain the pressure in the $NV\beta$ -Gibbs ensemble without computing the virial. As mentioned above, the pressure is related to the derivative of the Helmholtz energy with respect to volume. Thus, the difference in Helmholtz energy resulting from a tiny volume perturbation can be related to the pressure:

$$\beta p = -\left(\frac{\partial \beta A}{\partial V}\right)_{\beta, N} \approx -\frac{\beta \{A(V + \Delta V) - A(V)\}}{\Delta V} \tag{4.31}$$

The working relation for the pressure, computed from volume perturbations, is given by [81]:

$$p \approx \frac{1}{\beta \Delta V} \ln \left\langle \left(\frac{V + \Delta V}{V} \right)^N \exp(-\beta \Delta \mathcal{U}) \right\rangle_{\text{Gibbs}, \gamma} \quad (4.32)$$

where ΔV is the volume change and $\Delta \mathcal{U}$ is the associated energy change. The volume change ΔV needs to be constant throughout the simulation [81]. The pressure can be computed by ensemble averaging the contributions for negative and positive volume changes separately. Application of Eq. (4.32) in the $NV\beta$ -Gibbs ensemble requires no extra computations as volume moves are already performed to equilibrate the pressure. From now on, this method will be called the *volume-perturbation method*.

Singh *et al.* [82] combined canonical-ensemble simulation results obtained from positive and negative volume changes with a Bennett-like formula without weighting function in order to improve the accuracy of the pressure estimate. Actually, this approach requires two independent simulations, one of a system with volume V perturbed to volume $V + \Delta V$, and one of a system with volume $V + \Delta V$ perturbed to volume V . Singh *et al.* [82] approximated the latter perturbation by $V \rightarrow V - \Delta V$. A nice feature of this approximation is that only one simulation is needed to obtain the pressure.

In this work, negative and positive volume changes in the Gibbs ensemble are combined with a Shing-Gubbins-type equation. Like Singh *et al.* [82], the perturbation $V + \Delta V \rightarrow V$ is approximated by $V \rightarrow V - \Delta V$ in this work. The perturbations $V \rightarrow V + \Delta V$ belong to an unperturbed system (\otimes), while the perturbations $V \rightarrow V - \Delta V$ belong to a perturbed system (\odot), although both perturbations actually take place in the same phase. The working equation for the computation of the pressure from the *volume-perturbation overlapping-distributions method* in the Gibbs ensemble is presented here without proof:

$$\ln \pi_{\odot}(\mathcal{X}) = -\beta p \Delta V - \mathcal{X} + \ln \pi_{\otimes}(\mathcal{X}) \quad (4.33)$$

where \mathcal{X} is given by:

$$\mathcal{X} = \beta \Delta \mathcal{U}^{\text{ext}} - N \ln \left(\frac{V + \Delta V}{V} \right) \quad (4.34)$$

As volume changes are computationally expensive, especially in the Gibbs ensemble, their number is minimized. This limits the ability to compute the pressure from Eq. (4.33) as it is hardly possible to obtain reliable distributions $\ln \pi_{\odot}(\mathcal{X})$ and $\ln \pi_{\otimes}(\mathcal{X})$ from a small number of samples. However, the overlap between the distributions can serve as a check of whether the use of the volume-perturbation method [Eq. (4.32)] is allowed or not.

4.3.3. Modified $Np\beta$ + test molecule method

In this work, a slightly modified version of the pure-component $Np\beta$ + test molecule method [72, 73] is used. The liquid-phase pressure dependence of the chemical potential is approximated by a Taylor series expansion truncated after the second term:

$$\beta_0\mu'^L(p; \beta_0) \approx \beta_0\mu'^L(p_0^L; \beta_0) + \beta_0v^L(p - p_0^L) - \frac{N^L}{2}\beta_0^2 \left\{ \langle (v^L)^2 \rangle - \langle v^L \rangle^2 \right\} (p - p_0^L)^2 \quad (4.35)$$

where v^L is the liquid volume per molecule, N^L is the number of molecules in the liquid phase, β_0 is the reciprocal reference temperature, and p_0^L is the reference pressure of the liquid phase. The prime in Eq. (4.35) indicates that the purely temperature-dependent part has been excluded; *i.e.* μ' is a configurational chemical potential. The purely temperature-dependent contribution to the chemical potential is equal in coexisting phases and is therefore not relevant in computations of phase coexistence. For the vapour-phase chemical potential, it is convenient to separate the residual chemical potential from μ' :

$$\beta_0\mu'^V(p; \beta_0) = -\ln\left(\frac{1}{\beta_0 p}\right) + \beta_0\mu^{\text{res}, V}(p; \beta_0) \quad (4.36)$$

According to Vrabec and Fischer [75, 76, 78], it is convenient to treat the ideal-gas part in Eq. (4.36) analytically and to use a truncated Taylor series expansion for the residual part. This approximation reveals the logarithmic dependence of the vapour-phase configurational chemical potential on the pressure:

$$\beta_0\mu'^V(p; \beta_0) \approx \beta_0\mu'^V(p_0^V; \beta_0) + \left(\beta_0v^V - \frac{1}{p_0^V}\right)(p - p_0^V) + \ln\left(\frac{p}{p_0^V}\right) \quad (4.37)$$

The term in Eq. (4.37) that is linear in pressure differs from its counterpart in ref. [78], Eq. 8, which is probably in error. Equating the liquid-phase and vapour-phase chemical potentials enables the estimation of the chemical potential at vapour-liquid coexistence and the corresponding vapour-liquid saturation pressure.

In principle, it should be possible to compute the saturation pressure from the $Np\beta$ + test molecule method by performing a single simulation. This is only possible in case a good estimate of the true saturation pressure is known. However, the method is more predictive when it also works in case the reference pressure is a bad guess. When the reference pressure lies far away from the true saturation pressure, one cannot expect the truncated Taylor series expansions (4.35) and (4.37) to predict the right saturation pressure. Therefore, the choice of the reference pressure needs to be refined during a series of simulations.

Unfortunately, every refinement that brings the reference pressure closer to the true saturation pressure requires a long simulation for obtaining a reliable estimate for the liquid-phase chemical potential. In general, a liquid-phase chemical potential is a very weak function of pressure. The typical pressure dependence of a liquid-phase chemical potential is given in Figure 4.12. One can take advantage of this phenomenon by fixing the liquid-phase reference pressure p_0^L in successive simulations. The values of the vapour-phase reference pressure

p_0^V are updated after every simulation. The simulation scheme of the modified $Np\beta + \text{test}$ molecule method is outlined below:

1. Choose a starting value for the pressure. This pressure is the initial reference pressure for both the liquid and the vapour phase, p_0^L and p_0^V .
2. Perform isobaric-isothermal simulations for a liquid and a vapour phase at p_0^L and p_0^V and determine the relevant ensemble averages.
3. Compute the estimates for the configurational chemical potential at vapour-liquid equilibrium, μ'^{est} , and the saturation pressure, p^{est} , from the intersection of Eqs. (4.35) and (4.37).
4. Refine the reference pressure for the vapour phase. In this work, a damping factor is used to avoid overshoot of the estimated saturation pressure:

$$p_{i+1}^V = p_i^V + \frac{1}{2}(p^{\text{est}} - p_i^V) \quad (4.38)$$

The reference pressure for the liquid phase remains unchanged.

5. Stop the sequence of simulations when the prescribed convergence criterion is met. Otherwise, go to step 2. using the updated value of the vapour-phase reference pressure.

As became clear in Sections 4.2.1 and 4.2.5, the computation of the chemical potential in a dense liquid phase is a hard task. Just like was performed in Section 4.2.1, the chemical potential is computed from the Rosenbluth factor.

4.3.4. Modified $Np\beta + \text{test}$ molecule method with overlapping distributions

In Section 4.3.3, a modified $Np\beta + \text{test}$ molecule method was introduced that should be capable of finding a pure-component saturation pressure from a number of successive simulations. The method computes the chemical potential from the Rosenbluth factor. In this section, the method is further improved by computing the chemical potential from the overlapping-distributions method of Shing and Gubbins [30, 33].

Below, a derivation of the distributions that are needed to compute the configurational chemical potential from the Shing-Gubbins overlapping-distributions method in the $Np\beta$ -ensemble is given. The derivation starts by defining a perturbed and an unperturbed system. The unperturbed system (\otimes) has N real molecules and an ideal-chain molecule while the perturbed system (\odot) comprises $(N+1)$ real molecules. An ideal-chain molecule is an ideal-gas molecule without intramolecular Lennard-Jones and Coulombic interactions. The variable that needs to be sampled is given by:

$$\mathcal{X} = \beta \Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} - \ln \left(\frac{\beta p V}{N+1} \right) \quad (4.39)$$

where $\Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$ is the energy change that occurs when the ideal-chain molecule is changed to a real molecule.

The probability density to find the unperturbed system in ‘state’ \mathcal{X} is:

$$\pi_{\otimes}(\mathcal{X}) = \frac{\left\{ \frac{q_{\otimes}^{N+1} z^{N+1}}{N! \beta p} \int dV \frac{\exp(-\beta pV) NV^N}{V} \int d\Psi^{N+1} \exp\left\{-\beta \mathcal{U}_{\otimes}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^N)\right\} \delta\left(\mathcal{X} - \beta \Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + \ln\left\{\frac{\beta pV}{N+1}\right\}\right) \right\}}{q_{\otimes}} \quad (4.40)$$

where q_{\otimes} is a normalization constant. The factor $N!$ instead of $(N+1)!$ in Eq. (4.40) is caused by the fact that the ideal-chain molecule is distinguishable from the real molecules. The counterpart of Eq. (4.40) in the perturbed system is given by:

$$\pi_{\ominus}(\mathcal{X}) = \frac{\left\{ \frac{q_{\ominus}^{N+1} z^{N+1}}{(N+1)!} \int dV \frac{\exp(-\beta pV) NV^{N+1}}{V} \int d\Psi^{N+1} \exp\left\{-\beta \mathcal{U}_{\ominus}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^{N+1})\right\} \delta\left(\mathcal{X} - \beta \Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + \ln\left\{\frac{\beta pV}{N+1}\right\}\right) \right\}}{q_{\ominus}} \quad (4.41)$$

For convenience the factor $q_{\ominus}^{N+1} z^{N+1}$ is omitted in what follows. The δ -function in Eq. (4.41) allows substituting $\mathcal{X} + \beta \mathcal{U}_{\otimes}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + \ln\{\beta pV/(N+1)\}$ for $\beta \mathcal{U}_{\ominus}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$:

$$\begin{aligned} \pi_{\ominus}(\mathcal{X}) &= \frac{\left\{ \frac{q_{\otimes}}{q_{\ominus}} \int dV \frac{\exp(-\beta pV) NV^{N+1}}{(N+1)! V} \int d\Psi^{N+1} \exp\left\{-\beta \mathcal{U}_{\ominus}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^{N+1})\right\} \delta\left(\mathcal{X} - \beta \Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + \ln\left\{\frac{\beta pV}{N+1}\right\}\right) \right\}}{\frac{1}{N! \beta p} \int dV \exp(-\beta pV) \frac{NV^N}{V} \int d\Psi^{N+1} \exp\left\{-\beta \mathcal{U}_{\otimes}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^N)\right\}} \\ &= \frac{\left\{ \frac{q_{\otimes}}{q_{\ominus}} \exp(-\mathcal{X}) \int dV \frac{\exp(-\beta pV) NV^N}{N! \beta p V} \int d\Psi^{N+1} \exp\left\{-\beta \mathcal{U}_{\otimes}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^N)\right\} \delta\left(\mathcal{X} - \beta \Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + \ln\left\{\frac{\beta pV}{N+1}\right\}\right) \right\}}{\frac{1}{N! \beta p} \int dV \exp(-\beta pV) \frac{NV^N}{V} \int d\Psi^{N+1} \exp\left\{-\beta \mathcal{U}_{\otimes}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^N)\right\} d\Psi^{N+1}} \\ &= \frac{q_{\otimes}}{q_{\ominus}} \exp\{-\mathcal{X}\} \pi_{\otimes}(\mathcal{X}) \end{aligned} \quad (4.42)$$

The ratio of normalization constants q_{\odot}/q_{\otimes} can be related to the chemical potential referenced to its ideal-chain value:

$$\begin{aligned}
\frac{q_{\odot}}{q_{\otimes}} &= \frac{\frac{1}{(N+1)!} \int dV \exp(-\beta pV) \frac{NV^{N+1}}{V} \int d\Psi^{N+1} \exp\{-\beta \mathcal{U}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^{N+1})\}}{\frac{1}{N! \beta p} \int dV \exp(-\beta pV) \frac{NV^N}{V} \int d\Psi^N \exp\{-\beta \mathcal{U}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^N)\}} \\
&= \frac{\beta p}{N+1} \int d\Psi^{\text{test}} \left\langle V \exp\left\{-\beta \Delta \mathcal{U}_{\text{test}}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}(\Psi^N, \Psi^{\text{test}})\right\} \right\rangle_{Np\beta} \\
&= \exp\left(-\beta \left[\mu - \mu^{\text{ic}}\right]\right)
\end{aligned} \tag{4.43}$$

where μ^{ic} is the ideal-chain chemical potential. Insertion of Eq. (4.43) into Eq. (4.42) gives:

$$\ln \pi_{\odot}(\mathcal{X}) = \beta \left(\mu - \mu^{\text{ic}}\right) - \mathcal{X} + \ln \pi_{\otimes}(\mathcal{X}) \tag{4.44}$$

The configurational chemical potential is given by:

$$\beta \mu' = \beta \left(\mu - \mu^{\text{ic}}\right) + \ln \left(\frac{p}{p_{\text{ref}}} \left\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \right\rangle \right) \tag{4.45}$$

Notice that the purely temperature-dependent Rosenbluth factor $\langle \mathcal{W}_{\text{IG}, 1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle$ and the reference pressure are not needed in the computation of phase equilibria, as they are equal in coexisting phases. The actual value of $\mu - \mu^{\text{ic}}$ is obtained from a graphical analysis of the difference between the following distributions:

$$f_{\odot} = \ln \pi_{\odot}(\mathcal{X}) + \frac{1}{2} \mathcal{X} \quad f_{\otimes} = \ln \pi_{\otimes}(\mathcal{X}) - \frac{1}{2} \mathcal{X} \tag{4.46}$$

Equation (4.44) is similar to the relation given by Sindzingre *et al.* [49], who derived it for single-site Lennard-Jones molecules. Kofke and Glandt [83] also used a similar result in order to compute pure-component fugacities of Lennard-Jones systems. They used a joint energy-volume distribution, which gives the same result but requires longer simulations.

In order to extend the applicability of the $Np\beta + \text{test}$ molecule method to systems at high density, the overlapping-distributions method combined with Rosenbluth sampling is used. This combination of methods has been worked out by Mooij and Frenkel [61]. The working equation that is used to obtain $\mu - \mu^{\text{ic}}$ is identical to Eq. (4.44). However, the variable that is sampled is now:

$$\mathcal{X} = -\ln \mathcal{W}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} - \ln \left\{ \frac{\beta pV}{N+1} \right\} \tag{4.47}$$

where $\mathcal{W}^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$ is the Rosenbluth factor of a grown test-molecule [45].

4.3.5. Simulation details

4.3.5.1. General simulation details

Here, only those simulation details are described that are different from the details given in Section 4.2.4. In all simulations, except for the applications of the Gibbs ensemble without computation of the virial, translations and rotations of molecules were respectively force-biased and torque-biased. The statistical errors have been computed with the bootstrap method [60] applied to regularly saved blocks of 2500 MC cycles. The force fields used to obtain the simulation results were extensively described in Chapter 5.

4.3.5.2. Gibbs ensemble

Gibbs ensemble simulations were performed with $N = 400$ molecules. The initial densities of the liquid and the vapour phase were chosen such that the vapour phase contained on average approximately 50 molecules. This can be accomplished by changing the ratio of the initial number of molecules in the liquid and the vapour phase and/or by changing the ratio of the initial densities of the liquid and the vapour phase.

The distributions f_{\odot} and f_{\otimes} contained 201 bins. The bin width was determined during the equilibration phase of the simulation. After 5000 equilibration MC cycles, 100000 production MC cycles were performed. Samples were taken during the production phase. Every MC cycle consisted of $0.7N$ trial translations, $0.3N$ trial rotations, and 2 trial changes of the coupled volumes (6 in case the pressure was computed from volume perturbations).

To achieve an acceptable percentage of successful exchanges of molecules between the coexisting phases, the configurational-bias method [45] with multiple-first-bead insertions [40] was used. The number of attempted exchanges of molecules and the number of trials in the growth process of a molecule were adjusted such that the percentage of accepted exchanges was approximately between 2 and 5 percent.

In the original configurational-bias method, molecular conformations are generated from the intramolecular Boltzmann distribution without intramolecular Lennard-Jones and Coulombic contributions. For rigid molecules, the generation of a new trial position for an interaction site is not based on the intramolecular energy but on the fixed geometry of the molecule only. In order to increase simulation speed for the exchanges of molecules between the phases, the intermolecular Coulombic contribution (superscript C_{ext}) was left out of the Rosenbluth factor. This approximation is often sufficient because it is the repulsive Lennard-Jones contribution that mainly determines whether a trial exchange of a molecule is rejected or not. The acceptance criterion for the exchange of a molecule, performed in this way, is [45, 84]:

$$\min \left\{ 1, \frac{\mathcal{W}_{\text{new}}^{\text{LJ}_{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}}}{\mathcal{W}_{\text{old}}^{\text{LJ}_{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}}} \exp \left\{ -\beta \Delta \mathcal{U}^{C_{\text{ext}}} \right\} \right\} \quad (4.48)$$

where $\mathcal{W}_{\text{old}}^{\text{LJ}_{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}}$ is the Rosenbluth factor of a regrown ‘old’ molecule, $\mathcal{W}_{\text{new}}^{\text{LJ}_{\text{ext}} + \text{LJ}_{\text{int}} + C_{\text{int}}}$ is the Rosenbluth factor of a grown ‘new’ molecule, and $\Delta \mathcal{U}^{C_{\text{ext}}}$ is the difference between the new and the old Coulombic contributions to the intermolecular energy.

4.3.5.3. $Np\beta$ + test molecule method

The $Np\beta$ + test molecule simulations were performed with 300 molecules in the liquid phase and 100 molecules in the vapour phase. Every simulation in the sequence of simulations performed at different pressures consisted of 2500 equilibration cycles and 20000 production

cycles. Chemical potentials were obtained either directly from the Rosenbluth factor or from the overlapping-distributions method and the Rosenbluth factor.

The distributions f_{\odot} and f_{\otimes} contained 201 bins. The bin width was determined during the equilibration phase of the simulation. The distribution f_{\odot} depends on the volume and on the Rosenbluth factors of an existing regrown molecule. This distribution converges much faster in the region of interest than f_{\otimes} , which depends on the energies of inserted test molecules. For that reason, $0.2N$ trial removals per MC cycle were performed in the perturbed phase (\odot), and N trial insertions per MC cycle in the unperturbed phase (\otimes). In addition to the trial insertions and removals, every production cycle included $0.3N$ trial translations, $0.1N$ trial rotations, and 2 trial changes of the volumes. The intersection point of the chemical-potential Taylor series expansions was computed with the bisection method.

When Rosenbluth sampling is used to compute the chemical potential or when the configurational-bias method is exploited to enhance the exchanges of molecules between phases, a molecule is grown site by site. Details of the growing process are extensively described in the book of Frenkel and Smit [45].

In particular at high densities, it is important to find the optimal way to insert a molecule. The optimal insertion procedure depends on the geometry of the molecule and on the intermolecular potential. Especially the Lennard-Jones collision diameters σ , which are a measure of the excluded volume of an interaction site, influence the optimal choice. Insertions of test-molecule or insertions of real molecules with high energy contribute negligibly to the ensemble average of the chemical potential or have a low probability of acceptance. It is advantageous to reject a trial insertion or test-molecule insertion with high energy, and thus a low Rosenbluth factor, as early during the growth process as possible in order to increase to the simulation speed.

A good and a bad example of the growing process of a molecule are given below. A simple example is test-molecule insertion of dimethyl sulfoxide (DMSO) molecules into a dense liquid phase. One of the force fields that were reviewed by Kalugin *et al.* [85] is taken as an example. The DMSO molecule consists of four interaction sites with partial charges to account for the electrostatic charge distribution. The interaction sites comprise two methyl groups, a sulphur atom, and an oxygen atom. The methyl groups have large Lennard-Jones size and energy parameters. As the force fields in ref. [85] do not have any intramolecular degrees of freedom, trial orientations solely depend on the fixed geometry of the molecule.

The most naïve way to grow the molecule is to simply follow the natural structure of the molecule. k_1 trial positions are generated [40, 45] for the first methyl group. Out of these positions, one trial position is chosen, based on its intermolecular Lennard-Jones energy. The k_2 trial positions for the sulphur interaction site are put on a sphere, a bond length $l_{\text{CH}_3\text{-S}}$ apart from the chosen position of the first site. Again, out of these trials one configuration is selected based on its intermolecular Lennard-Jones energy. As the geometry of the molecule is fixed, the other two interaction sites cannot be placed independently. The trial positions for one of the remaining two sites, say the second methyl group, are placed on a cone with bond length $l_{\text{CH}_3\text{-S}}$ and bond angle $\varphi_{\text{CH}_3\text{-S-CH}_3}$. Again, out of these trials one configuration is selected based on its intermolecular Lennard-Jones energy. These steps in the growing process are presented graphically in Figure 4.7.

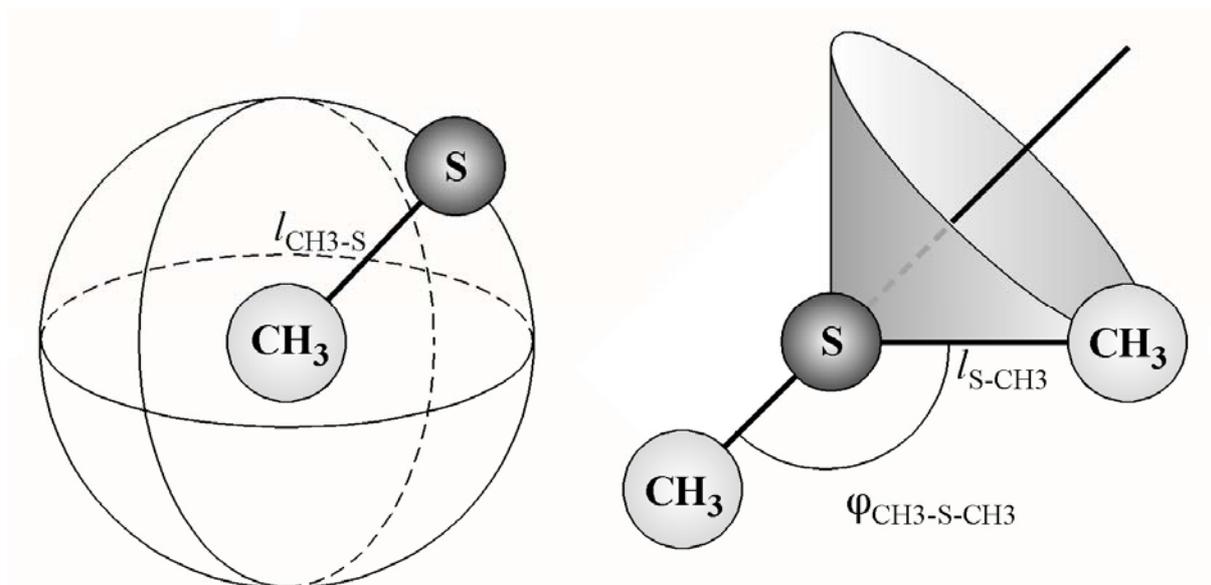


Figure 4.7: Steps in the growing process of a DMSO molecule. Left-hand picture: sulphur atom on a sphere. Right-hand picture: methyl group on a cone

A better growing scheme for the DMSO molecule is the following one. The interaction sites with the largest collision diameters/well depths have the highest probability to result in overlap of molecules. It is a waste of time when all but the last beads are grown and the last one cannot find a favourable position. Therefore, it is computationally advantageous to start searching for preferential positions of the largest groups. When overlap is noticed, the whole trial move can be rejected immediately.

The starting point for the growing process is again the generation of the trial positions for the first methyl group. Subsequently, the trial positions for the second methyl group are determined. In order to do so, it is needed to define a dummy chemical bond $l_{\text{CH}_3\text{-CH}_3}$, the length of which is computed easily as the geometry of the molecule is fixed. Out of the generated trial positions one configuration is selected based on its intermolecular Lennard-Jones energy. The trial configurations for the sulphur interaction site are placed on a cone with bond length $l_{\text{CH}_3\text{-S}}$ and a dummy bond angle $\varphi_{\text{CH}_3\text{-S-CH}_3}$. The fixed geometry of the molecule automatically dictates the position of the oxygen site. The steps in this growing process are presented graphically in Figure 4.8.

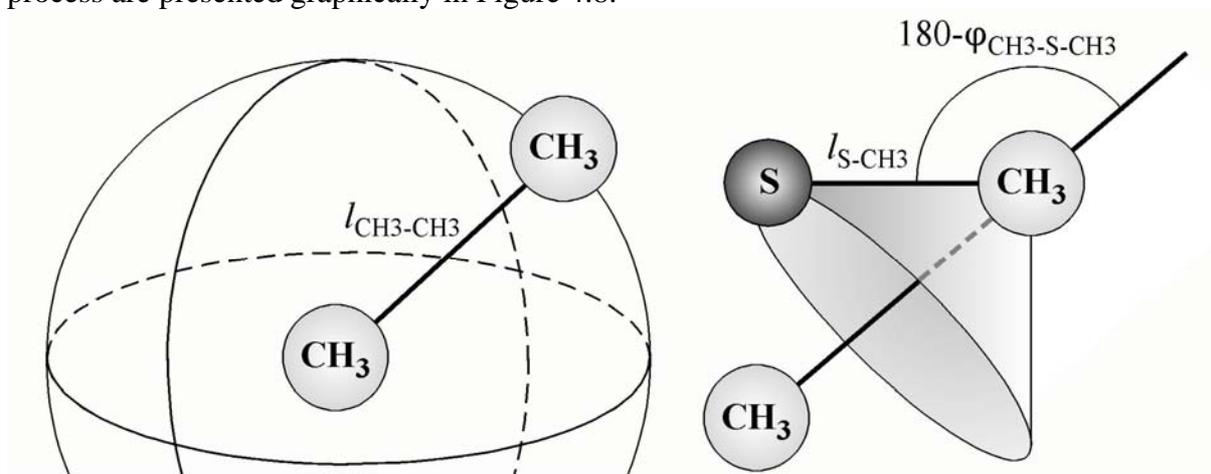


Figure 4.8: Steps in the growing process of a DMSO molecule. Left-hand picture: methyl group on a sphere. Right-hand picture: sulphur atom on a cone

4.3.6. Comparison of the different methods

The choice of the most suitable method to compute the initial point depends on a number of factors. The most important factor is the capability of the method to predict an accurate answer. Other factors are the length of the simulation needed to obtain a reliable result and the complexity of the simulation code.

Simulation results for several pure-component vapour-liquid coexistence points are given in Table 4.3. The $NV\beta$ – Gibbs ensemble directly provides the liquid and vapour densities at coexistence and the saturation pressure, all at the prescribed temperature. Table 4.3 presents both simulated saturation pressures obtained from the virial and from the volume-perturbation method. Remember that the overlapping distributions are only used to check the validity of the volume-perturbation method.

The $Np\beta$ + test molecule method provides the saturation pressure at the temperature of interest. Although, it is also possible to predict the coexisting densities at the predicted saturation pressure from a Taylor series expansion [86], in this work they are computed from a new simulation. Anyway, this new simulation is needed to compute the initial slope of the Clapeyron equation. For this reason, only the computed saturation pressures are presented in Table 4.3.

As the $Np\beta$ + test molecule method and the Gibbs ensemble method are completely different, simulation results obtained with both methods will not be compared. Instead of that, the applicability of the methods at different conditions is discussed. On the other hand, the different methods to compute the saturation pressure in the Gibbs ensemble can be compared rather easily.

Table 4.3: Simulation results for pure-component saturation pressures. Pressure [MPa] and/or configurational chemical potential $\beta\mu'$ along with statistical uncertainties at vapour-liquid coexistence conditions

System/condition	GE/virial [Eq. (4.30)]		GE/volume pert. [Eq.(4.32)]		$Np\beta$ + TM/overl. distr. [Eq. (4.44)]	
	p^L	p^V	p^L	p^V	p	$\beta\mu'$
$C_2H_6/250$ K	1.38(7)	1.323(5)	1.37(5)	1.312(7)		
$CO_2/280$ K	3.84(9)	4.09(5)	4.08(6)	4.03(2)		
$C_2H_6/144.26$ K					0.0076(7)	-8.9(9)
DMSO/278.5 K					0.0017(1)	-10.8(9)
DMSO/303.15 K					0.0063(5)	-9.57(8)
DMSO/328.94 K					0.019(1)	-8.55(7)

As discussed in Section 4.3.1, the Gibbs ensemble is a suitable method when the liquid density and the density-difference between the liquid and the vapour phase are moderate. In case the liquid density is moderate, it is worthwhile to use force-biased translations and torque-biased rotations. These bias methods require the computation of intermolecular forces. When the force between the molecules is computed, it requires little effort to compute the virial as well. Thus, at moderate densities, pressure evaluation via the virial is a suitable method.

On the other hand, when the liquid phase is far from dense, force and torque-biased methods are less relevant and computing the pressure from volume perturbations is an attractive alternative. Getting rid of the virial is particularly advantageous when the Ewald-summation method is used to compute the electrostatic interactions; the evaluation of electrostatic forces

is computationally very expensive. Obviously, the volume-perturbation method requires more volume-change trial moves than the virial method.

An advantage of computing the saturation pressure from volume perturbations is that it may provide reliable results from the liquid phase as well as from the vapour phase. When computing the saturation pressure from the virial, this is often not possible within reasonable time. The moving average of the saturation pressure, computed with both methods, in pure ethane [87] at $T = 250$ K is given in Figure 4.9. Obviously, both methods converge to the same saturation pressure. The moving average obtained from the volume-perturbation method converges much faster (both expressed in CPU demands and MC cycles). The overlapping distributions belonging to the volume-perturbation results in Figure 4.9, along with the computed saturation pressure, are presented in graphical form in Figure 4.10 (liquid phase) and in Figure 4.11 (vapour phase).

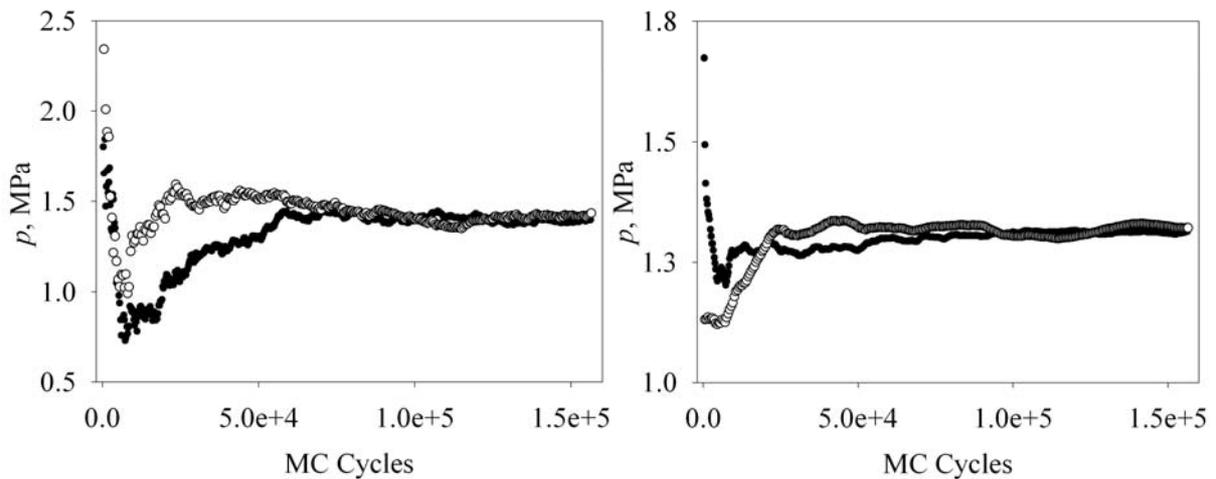


Figure 4.9: Moving average of the saturation pressure in pure ethane at $T = 250$ K. Left-hand picture: moving average of the liquid-phase pressure. Open symbols represent simulation results obtained from volume perturbations while solid symbols originate in the computation of the virial. Right-hand picture: same as left-hand picture but now for the vapour phase.

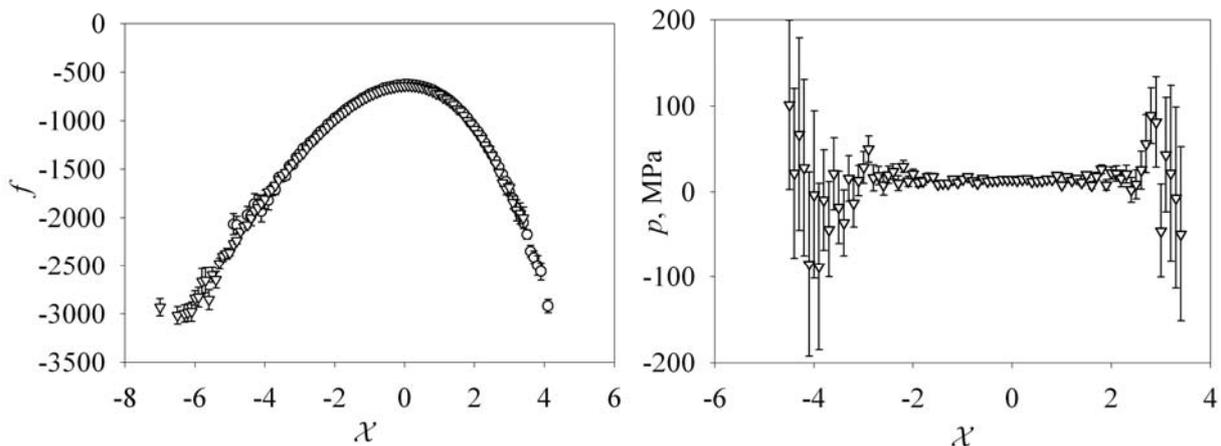


Figure 4.10: Simulation results for the volume-perturbation method. Left-hand picture: Overlapping distributions f_{\otimes} and f_{\odot} in the liquid phase of pure ethane at $T = 250$ K. Right-hand picture: saturation pressure obtained from the overlapping distributions.

As the distributions f_{\odot} and f_{\otimes} can hardly be distinguished, the regions of phase space that are important to the perturbed system and the unperturbed merely overlap. This indicates that the use of the volume-perturbation method was justified. It is expected that the distributions will move farther apart for larger values of the volume perturbation ΔV .

The modified $Np\beta$ + test molecule method required on average 5 updates of the vapour-phase reference pressure in order to converge to the vapour-liquid saturation pressure. Obviously, the number of updates depends on the initial estimate of the saturation pressure. It is not totally clear whether the number of updates needed purely originated from a poor estimate of the liquid-phase chemical potential, obtained within relatively short subsimulations, or from the finite range of the truncated Taylor series expansion of the vapour phase.

Simulation results for DMSO at $T = 303.15$ K are given in Figure 4.12. The distributions f_{\odot} and f_{\otimes} have a region of overlap that enables the computation of the configurational chemical potential μ' . The Taylor series expansions of the liquid-phase and vapour-phase configurational chemical potentials (obtained after the last update of the vapour-phase reference pressure) are also presented in Figure 4.12. The statistical uncertainty of the computed saturation pressure is governed by the statistical uncertainty of the configurational chemical potential at the reference pressure in the liquid phase.

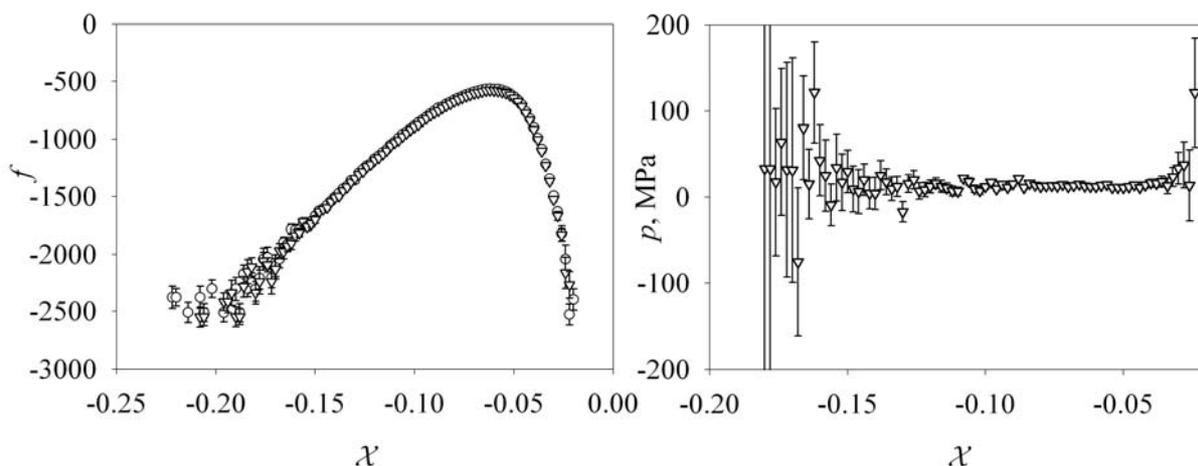


Figure 4.11: Simulation results for the volume-perturbation method. Left-hand picture: Overlapping distributions f_{\otimes} and f_{\odot} in the vapour phase of pure ethane at $T = 250$ K. Right-hand picture: pressure obtained from $f_{\odot} - f_{\otimes}$.

Predicted saturation pressures of DMSO along with the corresponding configurational chemical potentials are given in Table 4.3. The statistical uncertainty of the results is relatively high. Better results may be obtained when longer simulations are performed.

The prediction of one coexistence point with the modified $Np\beta$ + test molecule method is a time-consuming task. It requires four individual simulations per update of the vapour-phase reference pressure. Furthermore, the computation of the Rosenbluth factor is computationally expensive. Nevertheless, the combination of the $Np\beta$ + test molecule method with overlapping distributions and Rosenbluth sampling is very well capable of predicting vapour-liquid coexistence points at conditions where the Gibbs ensemble fails.

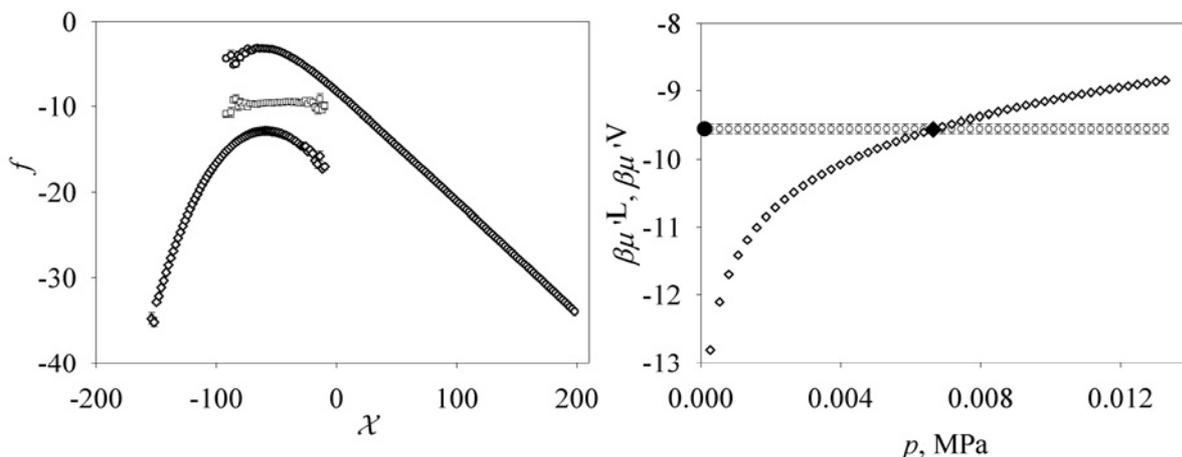


Figure 4.12: Simulation results for pure DMSO at $T = 303.15$ K computed from the overlapping-distributions $Np\beta +$ test molecule method. Left-hand picture: overlapping distributions from which the configurational chemical potential is computed. The lower curve represents the distribution obtained from a simulation in the perturbed phase (\odot), the uppermost curve represents the distribution obtained from a simulation in the reference phase (\otimes), and the data series between these distributions represents the liquid-phase configurational chemical potential. Right-hand picture: more or less horizontal data series represents the liquid-phase configurational chemical potential, logarithmic data series represents the vapour-phase configurational chemical potential, and the solid black symbols represent the configurational chemical potentials at the liquid and vapour-phase reference pressures.

4.4. Conclusions

In this chapter, the computation of the initial slope of the Clapeyron equation and the initial point are discussed. The initial point and initial slope are needed to start the numerical integration of the Clapeyron differential equation.

The initial slope at infinite dilution can be related to the difference between the residual chemical potentials of a solvent and an infinitely diluted solute. Three perturbation methods that can be used to predict the residual chemical-potential difference at infinite dilution were discussed: the Widom test-molecule method, the difference method, and the overlapping-distributions difference method. The Widom test-molecule method and the overlapping-distributions difference method provided comparable results.

The Widom test-molecule method, combined with Rosenbluth sampling, is applicable at relatively high densities. The overlapping-distributions difference method is also applicable at high densities. Furthermore, it has built-in diagnostics; one can easily see whether it is possible to compute a reliable answer.

The difference method provides only correct answers when a small single-site molecule is perturbed to a larger one. For non-spherical molecules, the difference method may work in neither perturbation direction, as the concepts *smaller* or *larger* are less clear. Thus, it is better not to use the difference method at all.

Two methods to predict the initial point were proposed: the $NV\beta -$ Gibbs ensemble method and a modified version of the $Np\beta +$ test molecule method.

The Gibbs ensemble method easily provides vapour-liquid coexistence points at moderate densities and for moderate density differences between the liquid and the vapour phase. The saturation pressure can be computed either from the virial or from volume perturbations. In case force-biased translations and torque-biased rotations are performed, for example at

moderate densities, it is convenient to compute the pressure from the virial. In other cases, one could use volume perturbations to compute the pressure. In general, the saturation pressure obtained from the volume-perturbation method converges faster than its counterpart that is computed from the virial.

The modification of the $Np\beta$ + test molecule method that is introduced in this work comprises a simulation scheme that should be able to predict the pure-component saturation pressure from a bad initial guess of the true saturation pressure. The configurational chemical potential is computed either directly from the Rosenbluth factor or indirectly from the overlapping-distributions method and the Rosenbluth factor, which makes the method applicable at high densities. The method is insensitive to large density differences between the phases as liquid and vapour-phase simulations are performed independently. Although the modification of the $Np\beta$ + test molecule method is computationally very demanding, it is capable of predicting initial points at conditions where the $NV\beta$ – Gibbs ensemble fails.

4.5. References

1. Mehta, M. and Kofke, D. A., Coexistence Diagrams of Mixtures by Molecular Simulation, *Chem. Eng. Sci.*, 49(16) **1994** 2633
2. Smith, J. M., Van Ness, H. C., and Abbott, M. M., *Introduction to chemical engineering thermodynamics*, Chemical Engineering Series, McGraw-Hill, Boston, 2001
3. Widom, B., Some Topics in the Theory of Fluids, *J. Chem. Phys.*, 39(11) **1963** 2808
4. Widom, B., Potential-Distribution Theory and the Statistical Mechanics of Fluids, *J. Phys. Chem.*, 86 **1982** 869
5. Kirkwood, J. G., Statistical Mechanics of Fluid Mixtures, *J. Chem. Phys.*, 3 **1935** 300
6. Smith, P. E. and van Gunsteren, W. F., Predictions of free-energy differences from a single simulation of the initial state, *J. Chem. Phys.*, 100(1) **1994** 577
7. Knopp, B., Suter, U. W., and Gusev, A. A., Atomistically Modeling of the Chemical Potential of Small Molecules in Dense Polymer Microstructures. 1. Method, *Macromolecules*, 30 **1997** 6107
8. McQuarrie, D. A., *Statistical Mechanics*, Harper's Chemistry Series, Harper and Row Publishers, New York, 1973
9. Shing, K. S. and Chung, S. T., Computer Simulation Methods for the Calculation of Solubility in Supercritical Extraction Systems, *J. Chem. Phys.*, 91 **1987** 1674
10. Kaminsky, R. D., Monte Carlo evaluation of ensemble averages involving particle number variations in dense fluid systems, *J. Chem. Phys.*, 101(6) **1994** 4986
11. Çagin, T. and Pettitt, B. M., Grand molecular dynamics: A method for open systems, *Mol. Sim.*, 6 **1991** 5
12. Attard, P., Simulation of the chemical potential and the cavity free energy of dense hard-sphere fluids, *J. Chem. Phys.*, 98(3) **1993** 2225
13. Lyubartsev, A. P., Martsinovski, A. A., Shevkunov, S. V., and Vorontsov-Velyaminov, P. N., New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles, *J. Chem. Phys.*, 96(3) **1992** 1776
14. Nezbeda, I. and Kolafa, J., A new version of the insertion particle method for determining the chemical potential by Monte Carlo simulation, *Mol. Sim.*, 5 **1991** 391
15. Lyubartsev, A. P., Laaksonen, A., and Vorontsov-Velyaminov, P. N., Free-energy calculations for Lennard-Jones systems and water using the expanded

-
- ensemble method - A Monte-Carlo and Molecular-Dynamics simulation study, *Mol. Phys.*, 82(3) **1994** 455
16. Vrabec, J., Kettler, M., and Hasse, H., Chemical potential of quadrupolar two-centre Lennard-Jones fluids by gradual insertion, *Chem. Phys. Lett.*, 356 **2002** 431
 17. Mon, K. K. and Griffiths, R. B., Chemical potential by gradual insertion of a particle in Monte Carlo simulation, *Phys. Rev. A*, 31(2) **1985** 956
 18. Khare, A. A. and Rutledge, G. C., Chemical potential of model benzene fluids using expanded ensemble Monte Carlo simulations, *J. Chem. Phys.*, 110(6) **1999** 3063
 19. Escobedo, F. A. and de Pablo, J. J., Monte Carlo simulation of the chemical potential of polymers in an expanded ensemble, *J. Chem. Phys.*, 103(7) **1995** 2703
 20. Wilding, N. B. and Müller, M., Accurate measurements of the chemical potential of polymeric systems by Monte Carlo simulation, *J. Chem. Phys.*, 101(5) **1994** 4324
 21. Müller, M. and Paul, W., Measuring the chemical potential of polymer solutions and melts in computer simulations, *J. Chem. Phys.*, 100(1) **1994** 719
 22. Kumar, S. K., Szleifer, I., and Panagiotopoulos, A. Z., Determination of the chemical potentials of polymeric systems from Monte Carlo simulations, *Phys. Rev. Lett.*, 66(22) **1991** 2935
 23. Sheng, Y.-J., Panagiotopoulos, A. Z., and Kumar, S. K., Effect of chain stiffness on polymer phase behavior, *Macromolecules*, 29(12) **1996** 4444
 24. Spyriouni, T., Economou, I. G., and Theodorou, D. N., Thermodynamics of Chain Fluids from Atomistic Simulation: A Test of the Chain Increment Method for Chemical Potential, *Macromolecules*, 30 **1997** 4744
 25. Kumar, S. K., Szleifer, I., and Panagiotopoulos, A. Z., Determination of the chemical potential of polymeric systems from Monte Carlo simulations. Reply, *Phys. Rev. Lett.*, 68(24) **1992** 3658
 26. Bennett, C. H., Efficient Estimation of Free Energy Differences from Monte Carlo Data, *J. Comput. Phys.*, 22 **1976** 245
 27. Torrie, G. M. and Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J. Comput. Phys.*, 23 **1977** 187
 28. Valleau, J. P. and Card, D. N., Monte Carlo Estimation of the Free Energy by Multistage Sampling, *J. Chem. Phys.*, 57(12) **1972** 5457
 29. Shing, K. S. and Gubbins, K. E., The chemical potential from computer simulation. Test particle method with umbrella sampling, *Mol. Phys.*, 43(3) **1981** 717
 30. Shing, K. S. and Gubbins, K. E., The chemical potential in dense fluids and fluid mixtures via computer simulation, *Mol. Phys.*, 46(5) **1982** 1109
 31. Powles, J. G., Evans, W. A. B., and Quirke, N., Non-destructive molecular-dynamics simulation of the chemical potential of a fluid, *Mol. Phys.*, 46(6) **1982** 1347
 32. de Pablo, J. J., Laso, M., and Suter, U. W., Estimation of the chemical potential of chain molecules by simulation, *J. Chem. Phys.*, 96(8) **1992** 6157
 33. Shing, K. S. and Gubbins, K. E., The chemical potential in non-ideal liquid mixtures: Computer simulation and theory, *Mol. Phys.*, 49(5) **1983** 1121
 34. Mezei, M., Theoretical calculation of the liquid-vapor coexistence curve of water, chloroform and methanol with the cavity-biased Monte Carlo method in the Gibbs ensemble, *Mol. Sim.*, 9 **1992** 257
 35. Deitrick, G. L., Scriven, L. E., and Davis, H. T., Efficient molecular simulation of chemical potentials, *J. Chem. Phys.*, 90(4) **1989** 2370

-
36. Stapleton, M. R. and Panagiotopoulos, A. Z., Application of excluded volume map sampling to phase equilibrium calculations in the Gibbs ensemble, *J. Chem. Phys.*, 92(2) **1990** 1285
 37. Mezei, M., A cavity-biased (μVT) Monte Carlo method for the computer simulation of fluids, *Mol. Phys.*, 40(4) **1980** 901
 38. Mezei, M., Grand-canonical ensemble Monte Carlo study of dense liquid Lennard-Jones soft spheres and water, *Mol. Phys.*, 61(3) **1987** 565
 39. Rosenbluth, M. N. and Rosenbluth, A. W., Monte Carlo calculation of the average extension of molecular chains, *J. Chem. Phys.*, 23(2) **1955** 356
 40. Esselink, K., Loyens, L. D. J. C., and Smit, B., Parallel Monte Carlo simulations, *Phys. Rev. E*, 51(2) **1995** 1560
 41. Harris, J. and Rice, S. A., A lattice model of a supported monolayer of amphiphile molecules: Monte Carlo simulations, *J. Chem. Phys.*, 88(2) **1988** 1298
 42. Frenkel, D. and Smit, B., Unexpected length dependence of the solubility of chain molecules, *Mol. Phys.*, 75(5) **1992** 983
 43. Siepmann, J. I., A method for the direct calculation of chemical potentials for dense chain systems, *Mol. Phys.*, 70 **1990** 1145
 44. Frenkel, D., Mooij, G. C. A. M., and Smit, B., Novel scheme to study structural and thermal properties of continuously deformable molecules, *J. Phys. : Condens. Matter*, 3 **1991** 3053
 45. Frenkel, D. and Smit, B., *Understanding molecular simulation; from algorithms to applications*, Computational Science Series, Volume 1, 2nd Ed., Academic press, London, 2002
 46. Smit, B., Grand canonical Monte Carlo simulations of chain molecules: adsorption isotherms of alkanes in zeolites, *Mol. Phys.*, 85(1) **1995** 153
 47. Shing, K. S., Infinite-dilution activity coefficients from computer simulation, *Chem. Phys. Lett.*, 119(2-3) **1985** 149
 48. Shing, K. S., Infinite-dilution activity coefficients of quadrupolar Lennard-Jones mixtures from computer simulation, *J. Chem. Phys.*, 85(8) **1986** 4633
 49. Sindzingre, P., Ciccotti, G., Massobrio, C., and Frenkel, D., Partial enthalpies and related quantities in mixtures from computer simulation, *Chem. Phys. Lett.*, 136(1) **1987** 35
 50. Liu, A., Calculation of chemical potentials of mixtures from computer simulations, *Mol. Phys.*, 89(6) **1996** 1651
 51. Liu, A. and Govind, R., Molecular dynamics simulation of the chemical potentials of mixtures, *Mol. Phys.*, 83(4) **1994** 779
 52. Liu, A. and Govind, R., Determination of thermodynamic excess properties of mixtures from computer simulation, *Mol. Sim.*, 15 **1995** 47
 53. Liu, A. and Beck, Th. L., Prediction of liquid-liquid phase equilibria of He+H₂ mixtures by NpT molecular dynamics simulations, *J. Chem. Phys.*, 105(6) **1996** 2424
 54. Liu, A. and Beck, Th. L., Methods of single-step and multistep particle switching in simulations of mixtures, *J. Phys. Chem.*, 100 **1996** 16002
 55. Liu, A., Excess thermodynamic properties of CO₂ + C₂H₆ mixtures from MD simulations, *Mol. Sim.*, 17 **1996** 75
 56. Heyes, D. M., Chemical potential, partial enthalpy and partial volume of mixtures by NpT molecular dynamics, *Chem. Phys.*, 159 **1992** 149
 57. Sindzingre, P., Massobrio, C., Ciccotti, G., and Frenkel, D., Calculation of partial enthalpies of an argon-krypton mixture by NpT molecular dynamics, *Chem. Phys.*, 129 **1989** 213

-
58. Allen, M. P. and Tildesley, D. J., *Computer simulation of liquids*, Clarendon Press, Oxford, 1987
 59. Flyvbjerg, H. and Petersen, H. G., Error estimates on averages of correlated data, *J. Chem. Phys.*, 91(1) **1989** 461
 60. Efron, B., *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982
 61. Mooij, G. C. A. M. and Frenkel, D., A systematic optimization scheme for configurational bias Monte Carlo, *Mol. Sim.*, 17 **1996** 41
 62. Batoulis, J. and Kremer, K., Statistical properties of biased sampling methods for long polymer chains, *J. Phys. A: Math. Gen.*, 21 **1988** 127
 63. Smit, B., Simulation of phase coexistence: from atoms to surfactants, *PhD thesis, University of Utrecht, The Netherlands*, **1990**
 64. Adams, D. J., Chemical potential of hard-sphere fluids by Monte Carlo methods, *Mol. Phys.*, 28(5) **1974** 1241
 65. Heinbuch, U. and Fischer, J., On the application of Widom's test particle method to homogeneous and inhomogeneous fluids, *Mol. Sim.*, 1 **1987** 109
 66. Guillot, B. and Guissani, Y., Investigation of the chemical potential by molecular dynamics simulation, *Mol. Phys.*, 54(2) **1985** 455
 67. Siepmann, J. I., McDonald, I. R., and Frenkel, D., Finite-size corrections to the chemical potential, *J. Phys. : Condens. Matter*, 4 **1992** 679
 68. Malanoski, A. P. and Monson, P. A., The phase behavior of a hard sphere chain model of a binary *n*-alkane mixture, *J. Chem. Phys.*, 112(6) **2000** 2870
 69. Kofke, D. A. and Cummings, P. T., Precision and accuracy of staged free-energy perturbation methods for computing the chemical potential by molecular simulation, *Fluid Phase Equil.*, 150-151 **1998** 41
 70. Kofke, D. A. and Cummings, P. T., Quantitative comparison and optimization of methods for evaluating the chemical potential by molecular simulation, *Mol. Phys.*, 92(6) **1997** 973
 71. Van 't Hof, A., de Leeuw, S. W., Hall, C. K., and Peters, C. J., Molecular simulation of binary vapour-liquid equilibria with components differing largely in volatility, *Mol. Phys.*, 102(3) **2004** 301
 72. Möller, D. and Fischer, J., Vapour liquid equilibrium of a pure fluid from test particle method in combination with *NpT* molecular dynamics simulations, *Mol. Phys.*, 69(3) **1990** 463
 73. Lofti, A., Vrabec, J., and Fischer, J., Vapour liquid equilibria of the Lennard-Jones fluid from the *NpT* plus test particle method, *Mol. Phys.*, 76(6) **1992** 1319
 74. Boda, D., Liszi, J., and Szalai, I., An extension of the *NpT* plus test particle method for the determination of the vapour-liquid equilibria of pure fluids, *Chem. Phys. Lett.*, 235 **1995** 140
 75. Vrabec, J., Lofti, A., and Fischer, J., Vapour liquid equilibria of Lennard-Jones model mixtures from the *NpT* plus test particle method, *Fluid Phase Equil.*, 112 **1995** 173
 76. Vrabec, J. and Fischer, J., Vapour liquid equilibria of mixtures from the *NpT* plus test particle method, *Mol. Phys.*, 85(4) **1995** 781
 77. Vrabec, J. and Fischer, J., Vapor-Liquid Equilibria of Binary Mixtures Containing Methane, Ethane, and Carbon Dioxide from Molecular Simulation, *Int. J. Thermophys.*, 17(4) **1996** 889
 78. Vrabec, J. and Fischer, J., Vapor-Liquid Equilibria of the Ternary Mixture CH₄+C₂H₆+CO₂ from Molecular Simulation, *AIChE J.*, 43(1) **1997** 212

-
79. Okumura, H. and Yonezawa, F., Method for liquid-vapor coexistence curves by test-particle insertions in the canonical ensemble, *J. Non-Cryst. Solids*, 293-295 **2001** 715
 80. Okumura, H. and Yonezawa, F., Reliable Determination of the Liquid-Vapor Critical Point by the *NVT* Plus Test Particle Method, *J. Phys. Soc. Jap.*, 70(7) **2001** 1990
 81. Harismiadis, V. I., Vorholz, J., and Panagiotopoulos, A. Z., Efficient pressure estimation in molecular simulations without evaluation the virial, *J. Chem. Phys.*, 105(18) **1996** 8469
 82. Singh, J. K., Kofke, D. A., and Errington, J. R., Surface tension and vapor-liquid phase coexistence of the square-well fluid, *J. Chem. Phys.*, 119(6) **2003** 3405
 83. Kofke, D. A. and Glandt, E. D., Monte Carlo simulation of multicomponent equilibria in a semigrand canonical ensemble, *Mol. Phys.*, 64(6) **1988** 1105
 84. Vlugt, T. J. H., Martin, M. G., Smit, B., Siepmann, J. I., and Krsihna, R., Improving the efficiency of the coniguration-bias Monte Carlo algorithm, *Mol. Phys.*, 94(4) **1998** 727
 85. Kalugin, O. N., Volobuev, M. N., Ishchenko, A. V., and Adya, A. K., Structure and dynamics of Na⁺ and Cl⁻ solvation shells in liquid DMSO: molecular dynamics simulations, *J. Mol. Liquids*, 91 **2001** 135
 86. Möller, D. and Fischer, J., Determination of an effective intermolecular potential for carbon dioxide using vapour-liquid phase equilibria from *NpT* + test particle simulations, *Fluid Phase Equil.*, 100 **1994** 35
 87. Chen, B. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, *J. Phys. Chem. B*, 103 **1999** 5370

Chapter 5. Simulation results for binary vapour-liquid equilibria

The theory of statistical mechanics forms a solid basis for the prediction of thermophysical properties from molecular simulation. In order to actually perform molecular simulation, one has to specify how molecules feel about each other; this feeling is expressed in intermolecular forces that can be derived from an intermolecular potential or force field. The nature of the intermolecular interactions and the origin and shortcomings of the well-known Lennard-Jones potential are commented upon in Section 5.1. A summary of available force fields for the components that are of interest to this work along with the selection criteria for a proper force field are also extensively described in Section 5.1. The rest of this chapter is organized as follows. In Section 5.2, the choice of combining rules that are used for interactions between dissimilar interaction sites is described. Simulation details of Gibbs-Duhem integrations are given in Section 5.3. Simulation results for pure-component vapour-liquid equilibria and for vapour-liquid equilibria of binary mixtures are given in Section 5.4. In Section 5.5, the major sources of error in the Gibbs-Duhem integration method are described. Finally, the conclusions concerning this chapter are summarized in Section 5.6.

5.1. Force field

5.1.1. Importance of a proper force field

The theory of statistical mechanics is well established and forms a solid basis for molecular simulation. In addition to the mathematical framework, to actually perform molecular simulation, knowledge about the way molecules feel about each other is crucial. This *feeling* is expressed in intermolecular forces. These forces can be derived from an *intermolecular potential*. The combination of intramolecular and intermolecular potentials is often called a *force field*. In what follows, the name force field will be used for both the potential and the corresponding parameters.

It is the force field that mainly determines the simulation results and their physical significance. Unfortunately, knowledge of intermolecular forces is still limited. However, there are nevertheless advanced force fields that describe the different contributions to the intermolecular energy in detail. Obviously, the use of these force fields for molecular simulation of large systems of molecules is computationally very expensive. For that reason, many researchers developed much simpler force fields that account *effectively* for the different contributions to the intermolecular energy. Other common simplifications are the united-atom approach, in which different atoms have been united into one interaction site, and the use of rigid molecules instead of fully or partially flexible molecules. Furthermore, it is common use to simplify the force field by expressing the interactions between molecules of different components in terms of the interactions between equal molecules with so-called *combining rules*. The advantage of this approach is the possibility to predict mixture properties from pure-component force fields.

These simplified, computationally efficient, effective force fields usually predict the physical properties of interest rather well. However, the simplicity goes at the cost of physical significance; the simple force fields cannot be expected to provide acceptable predictions of physical properties that heavily depend on the intramolecular geometry of the molecule or on the right balance between the different contributions to the force field. Moreover, simple force fields are often only applicable at state conditions close to the ones they are optimized to.

Clearly, there is some freedom in choosing the extent of complexity and/or the physical significance of the force field. This freedom has led to a variety of force fields for individual components.

It is not the purpose of this work to develop new force fields. One of the aims is to study whether the Gibbs-Duhem integration method is capable of, and suitable for, the prediction of vapour-liquid equilibria of mixtures. For this reason, it makes sense to choose existing force fields that accurately predict pure-component vapour-liquid equilibria. From a theoretical point of view, it does not matter whether the pure-component saturation pressures and vapour and liquid densities at coexistence are close to the experimental values or not; the numerical integration of the Clapeyron equation should trace the coexistence curve from the saturation pressure of the first component to the saturation pressure of the other component. However, from an application-oriented point of view, it is nonsense to study phase equilibria of mixtures when the pure-component phase behaviour is not well predicted.

A second aim of this work is the investigation of the influence of the Lorentz-Berthelot combining rules and other combining rules on the predicted vapour-liquid equilibria of mixtures.

5.1.2. Origin of the Lennard-Jones force field

Intermolecular forces are strongly repulsive at short range and attractive at long range. This is evident from the properties of condensed phases; attractive interactions keep the molecules closer than in a gas phase, however, the low compressibility of condensed phases indicates that there is a minimum favourable intermolecular distance.

Intermolecular forces are Coulombic in nature and have their origin in the interactions between the electrons and protons of the molecules. A description of the repulsive forces at short range, which are also called *overlap forces*, is given by Stone [1]. The long-range attractive forces can be divided into three contributions: (1) *attractive electrostatic*, (2) *induction*, and (3) *dispersion* or *Van der Waals* forces.

Ad 1) Electrostatic interactions are either attractive or repulsive and are pairwise additive. The attractive electrostatic contributions arise from the classical interaction between the ‘static’, and thus ‘permanent’, charge distributions on molecules. Sometimes, a distinction is made between electrostatic interactions involving molecules with net charges, and polar interactions involving neutral polar molecules.

The electrostatic interactions do not distort the electron clouds of the molecules and are therefore called *first order*. In simulations, the electron clouds are often modelled with sets of point charges that mimic the multipole moments on the molecules.

The electrostatic energy of N molecules, resulting from the interaction between permanent multipoles, can be computed by applying Coulomb’s law to sets of point charges:

$$\mathcal{U}^{\text{el}} = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N-1} \sum_{\gamma=1}^{n_i} \sum_{j=i+1}^N \sum_{\delta=1}^{n_j} \frac{q_i^\gamma q_j^\delta}{r_{i\gamma j\delta}} \quad (5.1)$$

where ϵ_0 is the vacuum permittivity, q_i^γ is partial charge γ on molecule i , n_i is the number of partial charges on molecule i , and $r_{i\gamma j\delta}$ is the distance between partial charges $i\gamma$ and $j\delta$.

Ad 2) The electron cloud of a molecule with a permanent multipole generates an electric field that may distort the electron cloud of a nearby molecule without permanent multipole. The

induced multipole then interacts with the permanent multipole to produce an attractive force. The strength of this force depends on the polarizabilities of the molecules.

The distortion of the electron clouds makes this a *second-order* interaction, which is *not pairwise additive*. For neutral molecules, the largest contribution to the induction energy comes from dipole-induced dipole interactions and can be shown to be proportional to r^{-6} , where r is the distance between interaction sites [1, 2].

Ad 3) The continuous motion of the electrons in the electron clouds of molecules without static multipole moments generates a fluctuating instantaneous dipole and other multipoles. The rapidly changing dipole induces a dipole and other multipoles on a nearby molecule. The correlation between instantaneous multipoles on the molecules in the system results in a lowering of the energy that is called *dispersion energy*. This type of interaction is clearly *second order* and *not pairwise additive*.

As there is no permanent multipole involved in the interaction, dispersion interactions are always present. For non-polar neutral molecules, this is the only contribution to the long-range attractive energy. Taking into account contributions between instantaneous dipole-dipole, dipole-quadrupole, quadrupole-quadrupole, and between other higher multipoles results in the following expression for the dispersion interaction:

$$\mathcal{U}^{\text{disp}}(r) = \frac{C_6}{r^6} + \frac{C_8}{r^8} + \frac{C_{10}}{r^{10}} + \dots \quad (5.2)$$

See refs. [1, 2] for a description of the coefficients C_i in Eq. (5.2).

Consider a system comprising N molecules that consist of a single interaction site. It is convenient to write the potential energy of the system into the following form, which is exact when written out to the final term:

$$\mathcal{U}^{\text{tot}} = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \mathcal{U}_{ij} + \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^N \Delta\mathcal{U}_{ijk} + \dots \quad (5.3)$$

where \mathcal{U}_{ij} is the pair potential energy between molecules i and j , which represents the interaction energy between two isolated molecules. The contribution $\Delta\mathcal{U}_{ijk}$ is a first correction to the pair potential which is generally non-zero as the presence of a third molecule distorts the electron clouds of the isolated pair of molecules.

A general approximation, which simplifies the computation of the interaction energy of an assembly of molecules considerably, is that of *pairwise additivity*. In this approximation it is assumed that the multibody correction terms in Eq. (5.3) are small compared to the pair-potential energy and that the intermolecular energy can be computed from interactions between pairs of molecules only:

$$\mathcal{U}^{\text{tot}} \approx \sum_{i=1}^{N-1} \sum_{j=i+1}^N \mathcal{U}_{ij} \quad (5.4)$$

The leading terms in the induction and the dispersion contributions fall off as r^{-6} . Although one often assumes that the repulsive contribution falls off as r^{-12} , there is evidence that an exponential approximation is better [1].

Lennard-Jones [3] adopted the following pairwise-additive form of the intermolecular potential:

$$\mathcal{U}^{\text{LJ}}(r) = 4\varepsilon \left\{ \left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right\} \quad (5.5)$$

Clearly, this potential has effective contributions for both the attractive and repulsive interactions. The parameter ε is called the *well-depth* of the potential. The parameter σ , sometimes called *collision diameter*, equals the intermolecular separation at which the potential is zero.

Equation (5.5) is used both for interactions between atoms and for interactions between ‘sites’ in polyatomic molecules. The Lennard-Jones intermolecular potential is generally adopted for computation of thermophysical properties with molecular simulation. Intermolecular potentials that are used less often are the exp-6 potential [1, 2] and the Kihara potential [4].

The partial charges, needed to compute the electrostatic intermolecular energy, can be obtained by optimizing them to experimentally determined multipole moments or to electron-density distributions computed from *ab initio* methods.

Given the partial charges, the Lennard-Jones parameters are usually optimized to *ab initio* results or to experimental data like densities, vapour pressures, and enthalpies of vaporization. Obviously, the Lennard-Jones parameters are optimized to the interactions that are not accounted for by the partial charges. The rest of the interactions include dispersion, repulsion, and the part of induction that is not accounted for in the partial charges. As these contributions are not pairwise additive, the multibody contributions in Eq. (5.3) are effectively incorporated in the pair potential energy, Eq. (5.4). This makes the Lennard-Jones potential an *effective pairwise-additive potential*. Clearly, the use of such an intermolecular potential has an adverse effect on the predictive power.

5.1.3. Force fields used in this work

Non-bonded interaction sites interact via an effective pairwise-additive Lennard-Jones 12-6 potential (external Lennard-Jones contribution). The Lennard-Jones interactions are truncated at a cutoff distance r_c that is at least three times larger than the largest collision diameter, σ , used in the simulation. Long-range corrections are applied beyond the cutoff [5]. Interaction sites on semi-flexible molecules more than three chemical bonds apart interact via the Lennard-Jones potential (internal Lennard-Jones contribution). The electrostatic contribution to the intermolecular potential is computed with the Ewald summation method [5].

The resulting intermolecular potential, including Lennard-Jones long-range corrections, is:

$$\begin{aligned}
\tilde{\mathcal{U}} = & 4 \sum_i^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \tilde{\varepsilon}_{\gamma\delta} \left\{ \left(\frac{\tilde{\sigma}_{\gamma\delta}}{\tilde{r}_{i\gamma j\delta}} \right)^{12} - \left(\frac{\tilde{\sigma}_{\gamma\delta}}{\tilde{r}_{i\gamma j\delta}} \right)^6 \right\} + 4 \sum_i^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \tilde{\varepsilon}_{\gamma\delta} \left\{ \left(\frac{\tilde{\sigma}_{\gamma\delta}}{\tilde{r}_{i\gamma i\delta}} \right)^{12} - \left(\frac{\tilde{\sigma}_{\gamma\delta}}{\tilde{r}_{i\gamma i\delta}} \right)^6 \right\} \\
& \text{external Lennard-Jones term} \qquad \qquad \qquad \text{internal Lennard-Jones term} \\
& + \frac{8\pi\tilde{\rho}(N-1)}{9} \sum_{i=1}^N \sum_{j=1}^N x_i x_j \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \tilde{\varepsilon}_{\gamma\delta} \tilde{\sigma}_{\gamma\delta}^3 \left\{ \left(\frac{\tilde{\sigma}_{\gamma\delta}}{\tilde{r}_c} \right)^9 - 3 \left(\frac{\tilde{\sigma}_{\gamma\delta}}{\tilde{r}_c} \right)^3 \right\} - \frac{\alpha}{\tilde{L}\sqrt{\pi}} \sum_{i=1}^N \sum_{\gamma=1}^{n_i} (\tilde{q}_i^\gamma)^2 \\
& \text{Lennard-Jones long-range correction} \qquad \qquad \qquad \text{point-self energy term} \\
& + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{\tilde{q}_i^\gamma \tilde{q}_j^\delta \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L})}{\tilde{r}_{i\gamma j\delta}} - \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \frac{\tilde{q}_i^\gamma \tilde{q}_i^\delta \operatorname{erf}(\alpha \tilde{r}_{i\gamma i\delta} / \tilde{L})}{\tilde{r}_{i\gamma i\delta}} \\
& \text{real-space term} \\
& + \sum_{\mathbf{n} \neq \mathbf{0}} \frac{\exp\left(-\frac{\pi^2 \|\mathbf{n}\|^2}{\alpha^2}\right)}{2\pi\tilde{L}\|\mathbf{n}\|^2} \left\{ \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \cos\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right]^2 + \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \sin\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right]^2 \right\} \\
& \text{reciprocal-space term}
\end{aligned} \tag{5.6}$$

The variable \mathbf{n} in Eq. (5.6) is a reciprocal lattice vector of the periodic lattice [5], $r_{i\gamma j\delta}$ is a site-site intermolecular distance, and α is an Ewald-sum screening parameter [5]. The variables in Eq. (5.6) are made dimensionless using the Lennard-Jones parameters σ_{11} and ε_{11} :

$$\tilde{\sigma}_{\gamma\delta} = \frac{\sigma_{\gamma\delta}}{\sigma_{11}} \quad \tilde{\varepsilon}_{\gamma\delta} = \frac{\varepsilon_{\gamma\delta}}{\varepsilon_{11}} \quad \tilde{q}_i^\gamma = \frac{q_i^\gamma}{\sqrt{4\pi\varepsilon_0\varepsilon_{11}}} \quad \tilde{L} = \frac{L}{\sigma_{11}} \quad \tilde{\rho} = \rho\sigma_{11}^3 \quad \tilde{\mathcal{U}} = \frac{\mathcal{U}}{\varepsilon_{11}} \quad r_{i\gamma} = \frac{\tilde{r}_{i\gamma}}{\sigma_{11}} \tag{5.7}$$

where $\sigma_{\gamma\delta}$ and $\varepsilon_{\alpha\gamma}$ respectively are the Lennard-Jones collision-diameter and well-depth parameters for the interaction between site γ and site δ , L is the length of the cubic simulation box, q_i^γ is the partial charge of interaction site γ on molecule i , and ρ is a molecular density.

Intramolecular flexibility, if present, comprises bond-angle bending and rotation about torsion angles. Bond lengths are fixed throughout this work. The only semi-flexible molecules that are considered are linear alkanes that are modelled with the TraPPE-EH force field [6]. The TraPPE-EH force field uses a harmonic bond-bending potential for the angle θ_{CCC} between three carbon atoms:

$$\mathcal{U}^{\text{bond bending}} = \frac{k_\theta}{2} (\theta_{\text{CCC}} - \theta_0)^2 \tag{5.8}$$

where θ_0 is the expectation value of θ_{CCC} and k_θ is a force constant. All θ_{CCH} angles regarding hydrogen atoms in a methyl group as well as all angles θ_{HCH} are fixed.

The torsional motions of the dihedral angles are governed by the following potentials:

$$\mathcal{U}^{\text{torsion}}(\phi_{\text{CCCC}}) = c_1 \{1 + \cos(\phi)\} + c_2 \{1 - \cos(2\phi)\} + c_3 \{1 + \cos(3\phi)\} \quad (5.9)$$

$$\mathcal{U}^{\text{torsion}}(\phi_{\text{XCCH}}) = c_X \{1 + \cos(3\phi)\} \quad (5.10)$$

where X represents a methyl hydrogen atom or a carbon atom, and ϕ is the torsion angle. Equation (5.10) differs from Eq. (3) in ref. [6]; a plus sign is used in Eq. (5.10) to make the definition of the torsion angle consistent with Eq. (5.9). The force constants used in Eqs. (5.8)–(5.10) are given in Table 5.2. The TraPPE-EH force field does not take into account (1-5) intramolecular Lennard-Jones interactions involving one hydrogen atom and (1-6) interactions involving two hydrogen atoms.

5.1.4. Choice of force-field parameters

5.1.4.1. Introduction

As stated in Section 5.1.1, optimization of force-field parameters is not the purpose of this work. Based on published simulation results for vapour-liquid equilibria, suitable force fields are selected for the components of interest. This selection should be performed very carefully as many force fields are optimized to liquid-phase data only and/or to physical properties over a small range of state conditions. Moreover, in Section 5.1.1 it was pointed out that the simulated saturation pressures should be close to their experimental counterparts. Unfortunately, most simulation studies do not focus on the prediction of pressures.

There is a large number of force fields that predict vapour-liquid densities at coexistence and saturation pressures very well. However, a large part of these force fields has exclusively been optimized to thermodynamic properties of one type of molecule. Moreover, different researchers used different levels of approximation and different methods to optimize the force-field parameters. Thus, there is often a lack of consistency between force fields and force-field parameters that are optimized by different people.

Consistency between force-field parameters is strongly related to *transferability*. Force-field parameters are transferable when the parameters of a site of type α in a molecule of component i equal those of a site of type α in a molecule of component j .

A typical example of a molecule that has been modelled with a variety of force fields is ethane. Ethane can be modelled as a two-centre rigid Lennard-Jones (2CLJ) molecule, in which a methyl group is united into one interaction site [7]. Coexisting liquid and vapour densities and saturation pressures predicted with this force field show excellent agreement with experimental data [See Section 5.4.2]. However, the bond length is far from realistic and the force field is not transferable.

On the other hand, the TraPPE-EH force field treats all eight atoms of ethane explicitly and includes rotation about the torsion angle [6]. This model is more complex, physically realistic, and transferable. However, although the TraPPE-EH force is applicable to the whole series of linear alkanes, the constraint of consistency between force-field parameters of equal interaction sites in dissimilar molecules has a slightly adverse effect on the coexistence properties of individual components.

Whenever possible, force fields for the components of interest to this work are chosen based on their extent of transferability, predictive power of pure-component vapour-liquid equilibria, and a good balance between simplicity and physical significance.

5.1.4.2. Linear alkanes

There are a number of force fields available for the homologous series of linear alkanes. Some of them have been extended to branched alkanes. The larger part of these force fields, for example the TraPPE-UA force field [8, 9], the NERD force field [10, 11], and the OPLS-UA force field [12] approximate CH_x – groups by united atoms for computational efficiency. Above-mentioned force fields have partial intramolecular flexibility and transferable Lennard-Jones groups. The OPLS-UA force field [12] uses geometric combining rules for both well-depth and collision-diameter parameters, while the other force fields employ the Lorentz-Berthelot combining rules for the interactions between dissimilar interaction sites. See Section 5.2 for a discussion of combining rules.

A result of the transferability of these united-atom models is that the predicted thermophysical properties for individual linear alkanes are relatively far off from the experimental values. The TraPPE-UA, NERD, and OPLS force fields predict vapour-liquid saturation pressures that are systematically too low or too high [6, 8, 11].

The force field developed by Errington and Panagiotopoulos [13] is similar to the ones described above, however, it uses the exp-6 potential instead of the Lennard-Jones potential. The vapour-liquid equilibria predicted with this force field are much closer to experimental data than those simulated with the TraPPE-UA and NERD force fields. However, this is not surprising as the exp-6 potential uses three parameters per interaction site opposed to the two parameters for a Lennard-Jones model.

The use of all-atom force fields is physically more realistic but is also computationally expensive. In addition to increased physical significance, an all-atom model has more degrees of freedom (parameters) than its united-atom counterpart and, therefore, should be able to predict vapour-liquid equilibria in closer agreement with experiment.

Chen and Siepmann tested a number of transferable all-atom force fields for alkanes and subsequently developed their own TraPPE-EH force field [6, 14]. The TraPPE-EH force field treats the hydrogen atoms explicitly, has partial intramolecular flexibility, and employs the Lorentz-Berthelot combining rules. The vapour-liquid densities at coexistence and the saturation pressures, predicted with the TraPPE-EH potential, for a series of n -alkanes were in close agreement with experimental data [6, 15].

In addition to the transferable force fields mentioned above, a number of force fields has been developed for individual linear alkanes. As these force fields have exclusively been optimized to thermodynamic properties of one type of molecule, they often predict results in very close agreement with experimental data. Examples are united-atom force fields for methane (1CLJ) [16-18], ethane (2CLJ) [7, 16, 19, 20], and a united-atom model with point quadrupole for ethane (2CLJQ) [21].

In this chapter, the non-transferable united-atom force fields for methane and ethane published by Lofti *et al.* [17, 18] and by Vrabec and Fischer [7] are used in order to show how well these very simple force fields perform. On the other hand, in Chapter 6, the more detailed and physically realistic TraPPE-EH force field [6] will be used. The corresponding force-field parameters for different alkanes are given in Table 5.1.

5.1.4.3. Carbon dioxide

A number of force fields has been developed for the prediction of vapour-liquid equilibria of carbon dioxide. Among them are the EPM and EPM2 models of Harris and Young [22], and the all-atom TraPPE-AA force field [23]. These force fields use a Lennard-Jones potential for the repulsive and dispersive interactions, while point charges are used for the electrostatic and induction interactions. Carbon dioxide was modelled as a three-site rigid linear molecule. Interactions between dissimilar interaction sites were expressed in those between similar sites using the Lorentz-Berthelot combining rules.

Potoff *et al.* [24] also modelled carbon dioxide as a three-site rigid linear molecule with point charges. However, their force field employed an exp-6 potential instead of a Lennard-Jones potential. Potoff *et al.* [24] performed two series of simulations. In the first series, the Lorentz-Berthelot combining rules were used for all interactions, while the second series employed the Lorentz-Berthelot combining rules for interactions between similar molecules and the Kong combining rules for interactions between dissimilar molecules. Simulation results for vapour-liquid equilibria computed with this force field were not available.

There is also a number of force fields that model carbon dioxide as a two-centre Lennard-Jones molecule, with or without quadrupole moment (2CLJQ or 2CLJ) [19, 21, 25-28].

In order to acquire agreement between experimental and simulated phase behaviour of alkane/carbon dioxide mixtures, one often needs special combining rules [23, 24]. However, Potoff and Siepmann [23] state that the simple Lorentz-Berthelot combining rules are sufficient in case the TraPPE-AA force field for carbon dioxide is combined with the TraPPE-EH force field for alkanes. In order to obtain consistency between these force fields, the Lennard-Jones parameters of carbon dioxide had to be optimized to vapour-liquid phase-equilibrium data of *n*-alkane/carbon dioxide mixtures. In practice, the force-field parameters were only optimized to vapour-liquid equilibria of pure carbon dioxide and the mixture propane/carbon dioxide at $T = 294.15$ K.

According to Potoff and Siepmann [23], the EPM2 and TraPPE-AA force fields perform equally well. Nevertheless, the transferability of the TraPPE-AA force field makes this potential a more suitable choice. Force-field parameters for carbon dioxide are given in Table 5.1.

5.1.4.4. Dimethyl sulfoxide

An overview of different force fields for dimethyl sulfoxide (DMSO) was given by Kalugin *et al.* [29]. The force fields given in the overview comprise four interaction sites: an oxygen atom, a sulphur atom and two united-atom methyl groups.

More recent force fields were published by Cordeiro [30], Vishnyakov *et al.* [31], and Geerke *et al.* [32]. The latter force fields are similar to the ones given by Kalugin *et al.* [29]. There is also a fully flexible all-atom force field that treats the hydrogen atoms explicitly [33]. Except for the force fields developed by Geerke *et al.* [32] and Cordeiro [30], all force fields are to be used in combination with the Lorentz-Berthelot combining rules.

Unfortunately, the available force fields for DMSO were developed for studying structural and thermodynamic properties of the liquid phase. Their capabilities to predict vapour-liquid equilibria have not been studied before. Thus, the choice of a proper force field for DMSO is based on a criterion that differs from the ones described in Section 5.1.4.1.

The difference between the liquid-phase and the vapour-phase density of DMSO at moderate temperatures is too large to employ the Gibbs ensemble [See Chapter 4, Section 4.3]. Thus, the $Np\beta$ + test molecule method is used to compute vapour-liquid coexistence. A successful application of this method depends on the ability to insert test molecules with a nonzero Boltzmann factor into a simulation box. The larger the Lennard-Jones interaction beads, the harder they find a favourable insertion position in a dense phase. In this work, the force field developed by Rao and Singh [34] is chosen, as its collision diameters and well-depths are of moderate size/strength. Force-field parameters for DMSO are given in Table 5.1.

5.1.4.5. Difluoromethane

The number of force fields that is available for difluoromethane is increasing as the molecule is considered to be an alternative refrigerant (R32).

Gao *et al.* [35] used an effective Stockmayer potential to simulate vapour-liquid equilibria of difluoromethane. They modelled difluoromethane as a 1CLJ molecule with an orientation-dependent dipole. This simple force field predicted vapour-liquid phase equilibria in quite good agreement with experimental data over a broad range of temperatures.

Lago *et al.* [20] also modelled difluoromethane as one united-atom group with a point dipole. The Kihara potential [4] was used for the dispersive and repulsive interactions. The difference between simulated and experimental saturation pressures and liquid and vapour densities at coexistence, predicted with this force field, increased with decreasing temperature.

Higashi and Takada [36] developed a rigid all-atom force field with Lennard-Jones interaction sites. A partial charge was put on every interaction site to model the first-order electrostatic and induction interactions. This potential is to be used with the Lorentz-Berthelot combining rules. As the Lennard-Jones parameters were optimized to vapour-liquid equilibria, the force field predicts coexisting vapour and liquid densities in close agreement with experimental data. However, the saturation pressure deviates from the experimental values at low temperatures, as is clear from simulation results presented in Section 5.4.5 and from simulation results published by Lísal *et al.* [37].

The force field developed by Potter *et al.* [38] is similar to the one developed by Higashi and Takada [36] except for the presence of partial intramolecular flexibility and the use of a different combining rule for the H...F interaction. The force field consists of transferable groups and has been optimized for all fluoromethanes. The predicted coexistence densities and saturation pressures were relatively far off from the experimental ones.

Jedlovsky and Mezei [39] improved the all-atom force field of Potter *et al.* [38] in order to obtain better structural and thermodynamic properties of the liquid phase. They optimized the parameters of the force field to liquid-phase data at low temperatures. It is not surprising that simulated vapour-liquid equilibria at low temperatures were in close agreement with experimental data. However, the agreement at higher temperatures became progressively worse. Simulation results for the saturation pressures were not given in the article.

Fermeglia *et al.* [40] optimized force-field parameters for a couple of alternative refrigerants. The alternative refrigerants, among which is difluoromethane, were modelled as effective all-atom potentials with intramolecular flexibility. The overlap and dispersion forces were modelled with a Lennard-Jones (9-6) potential. Partial charges were used to model the first-order electrostatic and induction interactions. The force-field parameters are to be used with sixth-order combining rules [40]. Lennard-Jones parameters and partial charges for a particular type of interaction site differ from component to component. This force field predicts vapour and liquid densities at coexistence that are in close agreement with experimental data. Saturation pressures were not given in the article published by Fermeglia *et al.* [40].

In addition to the force fields mentioned above, there are two all-atom force fields the parameters of which were optimized for studying liquid-phase properties [41, 42]. Based on the considerations above, the force field developed by Higashi and Takada is chosen for further research [36]. Force-field parameters for difluoromethane are given in Table 5.1.

Table 5.1: Lennard-Jones parameters, partial charges, bond lengths, angles, and ionization potentials for several components

Comp. Group	ϵ/k_B K	σ Å	q e	l Å	\angle °	I eV
methane (1CLJ) [7]						
CH ₄	148.99	3.7275				12.51 [43]
ethane (2CLJ) [7]						
CH ₃	135.57	3.500		$l_{\text{CH}_3\text{-CH}_3}=2.345$		11.52 [43]
carbon dioxide [23]						
C	27.0	2.80	+0.70	$l_{\text{C-O}}=1.16$		13.773 [43]
O	79.0	3.05	-0.35			
dimethyl sulfoxide [34]						
CH ₃	33.12	3.66	+0.160	$l_{\text{CH}_3\text{-S}}=1.80$	$\angle_{\text{CH}_3\text{SO}}=106.75$	9.01 [43]
S	101.65	3.56	+0.139	$l_{\text{S=O}}=1.53$	$\angle_{\text{CH}_3\text{-S-CH}_3}=97.4$	
O	80.58	2.94	-0.299			
TraPPE-EH potential for linear alkanes [6]						
C _{CH4}	0.01	3.31		$l_{\text{C-C}}=1.535$ [44]	$\angle_{\text{C-C-H}}=111.^b$	12.51 [43] CH ₄
C _{CH2}	5.0	3.65		$l_{\text{C-H}}=0.55$	$\angle_{\text{H-C-H}}=107.9^b$	11.52 [43] C ₂ H ₆
C _{CH3}	4.0	3.30				10.95 [43] C ₃ H ₈
H	15.3	3.31				
difluoromethane [36]						
C	42.0	3.46	+0.3850	$l_{\text{C-F}}=1.37$	$\angle_{\text{FCF}}=108.6$	12.71 [43]
F	37.0	2.95	-0.2415 ^a	$l_{\text{C-H}}=1.09$	$\angle_{\text{HCH}}=113.8$	
H	29.0	2.20	+0.0490			
trifluoromethane [45]						
CH	85.48	3.50	+0.275	$l_{\text{CH-F}_3}=1.67$		13.86 [43]
F ₃	85.48	3.50	-0.275			

^a charge of F-atom has been changed slightly in order to obey the condition of electroneutrality

^b angles have been changed slightly as structure is overspecified

5.1.4.6. Trifluoromethane

Like difluoromethane, trifluoromethane is considered to be an alternative refrigerant (R23). The larger part of the all-atom force fields for trifluoromethane [38, 41, 42, 45-48] has been developed to study liquid-phase data. Only a few of them were intended to predict vapour-liquid equilibria. Song *et al.* [45] developed a two-site model with point charges on both interaction sites. The Lennard-Jones parameters of the two sites were equal. This force field is physically not very realistic but the agreement between experimental and simulated vapour-liquid equilibria was surprisingly good. Its simplicity is attractive but its predictive power in mixtures is still to be investigated.

The all-atom force field of Potter *et al.* [38] employs a Lennard-Jones potential with additional partial charges [See Section 5.1.4.5]. The force field has partial intramolecular flexibility. Its parameters have been optimized for all three fluoromethanes and are transferable. The predicted vapour and liquid densities at coexistence were close to their experimental counterparts, while the predicted saturation pressures deviated substantially.

As the agreement between simulated and experimental results is much better for the force field of Song *et al.* [45], this potential will be used for trifluoromethane in what follows. Force-field parameters for trifluoromethane are given in Table 5.1.

Table 5.2: TraPPE-EH intramolecular parameters (k_B is Boltzmann's constant)

Eq. (5.10)	Eq. (5.9)	Eq. (5.8)
$c_C/k_B = 854$ K	$c_1/k_B = 355.03$ K	$k_\theta/k_B = 58765$ K/rad ²
$c_H/k_B = 717$ K	$c_2/k_B = -68.19$ K	
	$c_3/k_B = 791.32$ K	

5.2. Combining rules

Simulation of the thermophysical properties of mixtures requires knowledge of the interactions between dissimilar components. It is possible to optimize pure-component force-field parameters to mixture data or to develop new force fields for interactions between dissimilar components. However, these approaches reduce the predictive power of molecular simulation since experimental mixture data are needed to predict other mixture data.

For this reason, many researchers developed empirical expressions in which the force-field parameters of interactions between dissimilar sites are expressed into the parameters of interactions between similar sites.

From now on, the interactions between different sites will be called *unlike interactions* and the ones between equal sites *like interactions*. The empirical expressions for unlike interactions are called *combining* or *combination rules*. Although the use of combining rules has a number of advantages, there is little information in effective pairwise-additive pure-component Lennard-Jones potentials that predicts how unlike molecules will interact.

Assume that molecules of type i and j are modelled as single interaction sites. The most simple and most widely used combining rules for Lennard-Jones parameters are the Lorentz rule for the unlike-components collision-diameter parameters [49]:

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}, \quad (5.11)$$

and the Berthelot rule for the unlike-components well-depth parameters [50, 51]:

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}. \quad (5.12)$$

The Lorentz rule is exact for mixtures of hard spheres. The Berthelot combining rule is a simplification of the combining rule that can be derived from the London theory of dispersion forces [2, 52]. It is known to overestimate the unlike interactions [1, 2]. Several researchers tried to correct for this systematic deviation by inclusion of binary-interaction parameters in the Lorentz-Berthelot (LB) combining rules [7, 53-58]:

$$\sigma_{ij} = (1 - l_{ij}) \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad \varepsilon_{ij} = (1 - k_{ij}) \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (5.13)$$

The optimized value for l_{ij} is often close to zero [7, 16, 53, 57]. Over the years, a large variety of combining rules has been developed, differing in complexity and theoretical background. An overview of a large number of combining rules was given by Diaz Peña *et al.* [59, 60] and by Maitland *et al.* [2].

The derivation of a combining rule usually starts from assumptions about the nature of the dispersion interactions, the repulsive interactions, and about the form of the intermolecular potential. An example is the derivation of the Hudson-McCoubrey (HMCC) combining rule for ε_{ij} . The London theory of dispersion forces gives the following approximation for the dispersion potential between unlike molecules [1]:

$$\mathcal{U}_{ij}^{\text{disp}} \approx -\frac{3}{2} \left(\frac{\langle u_i \rangle \langle u_j \rangle}{\langle u_i \rangle + \langle u_j \rangle} \right) \frac{\alpha_i \alpha_j}{(4\pi\varepsilon_0)^2 r_{ij}^6} \quad (5.14)$$

where $\langle u_i \rangle$ is the average excitation energy of molecule i , α_i is the static polarizability of atom i , ε_0 is the vacuum permittivity, and r_{ij} is the distance between site i and site j . Though only a rough approximation, usually the average excitation energies are put equal to the ionization potentials [1, 61]:

$$\mathcal{U}_{ij}^{\text{disp}} \approx -\frac{3}{2} \left(\frac{I_i I_j}{I_i + I_j} \right) \frac{\alpha_i \alpha_j}{(4\pi\varepsilon_0)^2 r_{ij}^6} \quad (5.15)$$

where I_i is the ionization potential of molecule i . Equating relation (5.15) with the Lennard-Jones (6-12) potential gives the following form for the unlike-components well depth ε_{ij} :

$$\varepsilon_{ij}^{\text{HMCC}} \approx 2 \frac{\sqrt{I_i I_j}}{I_i + I_j} \left(\frac{2\sqrt{\sigma_{ii}\sigma_{jj}}}{\sigma_{ii} + \sigma_{jj}} \right)^6 \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (5.16)$$

Equation (5.16) is the HMCC combining rule [52]. In this work, the larger part of the molecules consists of a number of interaction sites. In case Eq. (5.16) is used for interactions between interaction sites on multi-site molecules, the ionization potential of the molecule belonging to the site of interest is used. Division of Eq. (5.16) by the Berthelot combining rule gives:

$$\frac{\varepsilon_{ij}^{\text{HMCC}}}{\varepsilon_{ij}^{\text{Berthelot}}} = 2 \frac{\sqrt{I_i I_j}}{I_i + I_j} \left(\frac{2\sqrt{\sigma_{ii}\sigma_{jj}}}{\sigma_{ii} + \sigma_{jj}} \right)^6 \quad (5.17)$$

The right-hand side of Eq. (5.17) is plotted in Figure 5.1 versus ratios of ionization potentials and collision diameters. From Figure 5.1, it is clear that the HMCC combining rule predicts lower values of the unlike-interaction well depths than the Berthelot combining rule does.

Although superior to the Berthelot combining rule, the HMCC combining rule still tends to overestimate the dispersion energy belonging to interactions between unlike molecules [2]. For molecules with equal collision diameters and ionization potentials, Eq. (5.16) reduces to the Berthelot combining rule.

Khalaf Al-Matar and Rockstraw [62] state that combining rules are meant to reduce the number of parameters in a force field. In their opinion, incorporation of ‘additional’ parameters like ionization potentials should be avoided whenever possible.

Of course, it makes sense to exclude parameters derived from experimental mixture data. However, incorporation of ionization potentials or polarizabilities, which are pure-component properties that are available for a lot of molecules, is not against the purpose of combining rules.

Potoff *et al.* [24] showed that a combination of Kong combining rules [63] for interactions between molecules of dissimilar components and LB combining rules for interactions between molecules of similar components improves the agreement between experimental and simulated vapour-liquid equilibria of binary systems *n*-alkane/carbon dioxide considerably.

The different treatment of interactions between similar and dissimilar molecules was needed since the force-field parameters for the pure components were optimized using the LB combining rules. It is questionable whether above-mentioned improvement is general or not; the effect of using different combining rules on the simulation results depends on the force field that is used. Apart from that, it makes more sense to compare the performance of different sets of combining rules when all interactions are computed with that set. However, in order to do so, one needs to optimize pure-component force-field parameters with the Kong combining rules or with any other set of combining rules.

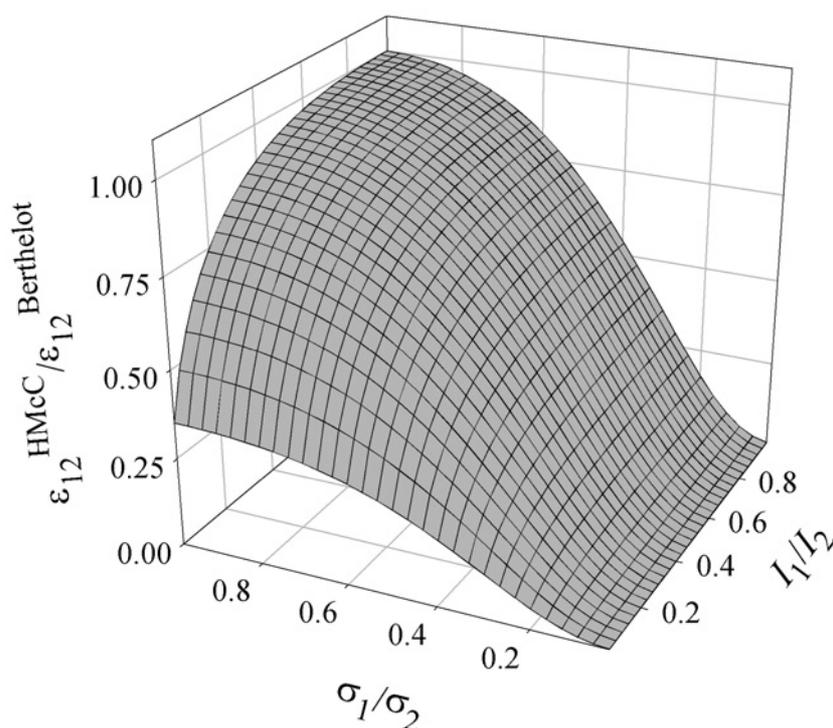


Figure 5.1: Right-hand side of Eq. (5.17) versus ratios of ionization potentials and collision diameters.

Potoff and Siepmann [23] showed that the simple LB combining rules can do a good job when predicting vapour-liquid equilibria of binary and ternary mixtures of *n*-alkanes, carbon dioxide, and nitrogen. In order to reach this result, pure-component force-field parameters of carbon dioxide were optimized to vapour-liquid equilibria of the binary mixture propane/carbon dioxide so as to find a good balance between the Lennard-Jones (dispersion and repulsion) interactions and the Coulombic (first-order electrostatic and induction) interactions. However, as will be shown in Chapter 6, optimization of the force-field

parameters to vapour-liquid equilibria of one binary mixture at a single temperature is probably to strong a confinement to obtain this balance.

The simulations in this chapter are performed with the LB combining rules for all interactions. As the existing force fields in this work are to be used in combination with a particular combining rule, that combining rule has to be used for the interactions between similar molecules.

In case simulated vapour-liquid equilibria of binary mixtures are far off from experiment, the only thing that can be done to improve the prediction is changing the combining rule for interactions between molecules of different components. In Chapter 6, it will be shown that the use of different combining rules within a single simulation can improve the agreement between simulation results and experimental data considerably. In these cases, the LB combining rule will be used for all interactions between molecules of the same component and the LHMcC combining rules for all interactions between dissimilar components.

5.3. Simulation details

Details of pure-component Gibbs-ensemble simulations, $Np\beta$ + test molecule simulations, and simulations that are performed in order to compute the initial slope of the Clapeyron equation were given in Chapter 4. Simulation details that do not differ from those given in Chapter 4, Section 4.3.5.1, will not be mentioned here.

The Gibbs-Duhem integration method has been described in Chapter 2. Below, the simulation details for the Gibbs-Duhem integration method are described. The statistical uncertainties of densities and mole fractions, obtained from a simulation at the estimated saturation pressure, were computed with the block-averaging method of Flyvbjerg and Petersen [64]. A brief error analysis of the Gibbs-Duhem integration method is given in Section 5.5.

The computation of one coexistence point contains a *predictor* phase, a *corrector* phase and a *production* phase. The predictor phase takes 2500 MC cycles to equilibrate the system and 5000 production cycles in which the relevant ensemble averages are computed. The value of the integrand of the Clapeyron equation at the imposed pressure and fugacity fraction is computed from these ensemble averages. The integrand is used in the corrector equation to update the estimate of the saturation pressure. A corrector simulation is performed at the updated estimate of the saturation pressure. This simulation takes 1000 equilibration MC cycles and again 5000 MC production cycles.

It is possible to compute the integrand from samples obtained from a single corrector simulation. However, it takes a long simulation to get a converging value of the ensemble average of the integrand. Therefore, it is convenient to accumulate samples taken during successive corrector simulations in order to get a converging value of the integrand as soon as possible. Adding samples obtained at successive corrector steps is often a fair approximation because the consecutive updated estimates of the saturation pressures will not differ very much.

When the estimated saturation pressure converged to a stable value, the production phase is started. The production phase takes 1000 equilibration MC cycles and 10000 production MC cycles.

A disadvantage of the Gibbs-Duhem integration method is that the number of corrector iterations, and thus the total simulation length (expressed in MC cycles), is not known beforehand. Furthermore, addition of samples taken during successive corrector simulations, which are performed at slightly different pressures, may destabilize the simulation, especially in the vicinity of the vapour-liquid critical point. Another disadvantage is that the samples taken during the corrector simulations do not contribute to the ensemble averages of the

densities and mole fractions at the predicted coexistence point. An improved version of the Gibbs-Duhem integration method that does not suffer from these drawbacks is presented in Chapter 6.

The simulations in the semigrand-canonical ensemble were performed with 300 molecules in the liquid phase and 100 molecules in the vapour phase. The composition sampling was performed without bias. Every MC cycle consisted of $2N + 2$ trial moves: $0.5N$ translations, $0.5N$ rotations, N identity changes, and one change of the liquid-phase and the vapour-phase volumes.

The variable step-size predictor-corrector equations, as described by Escobedo and De Pablo [See Chapter 2, Section 2.6], were used [65]. It might have been advantageous to adjust the integration step-size on the fly. Such an integration scheme increases the stability of the method by choosing a small integration step in case $|\partial^2 p / \partial \xi_2^2|$ is large. Nevertheless, in this work, predetermined integration steps are used.

5.4. Simulation results

5.4.1. Methane

In this section, it is shown how well a very simple force field can perform. Methane is modelled as a single-site Lennard-Jones molecule (1CLJ). The force field parameters are given in Table 5.1. Simulated liquid and vapour densities at coexistence and saturation pressures for a number of temperatures, along with the statistical uncertainties, are given in Table 5.3 and in Figure 5.2.

As indicated in Table 5.3, the results at the lowest temperature have been computed with the $Np\beta$ -test molecule method ($Np\beta$ +TM) without the use of overlapping distributions. The other results were computed in the Gibbs ensemble (GE).

The good agreement between simulated and experimental data is evident from Figure 5.2. Of course, this result is not unexpected, as the force-field parameters were optimized to vapour-liquid equilibria. However, the force field has only two parameters (degrees of freedom) and nevertheless describes the equilibria very well over a broad range of temperatures.

The deviation of the simulation results from the experimental data increases in the vicinity of the critical point. It has not been investigated whether this effect is caused by incapacibilities of the force field or originates in finite-size effects.

Table 5.3: Simulated vapour-liquid equilibria of methane. The numbers in parentheses denote the statistical uncertainties in the last digit.

T K	p^{sim} MPa	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	Method
100	0.03(1)	0.532(3)	440.3(9)	$Np\beta$ +TM
120	0.187(4)	3.15(7)	410.3(4)	GE Virial
130	0.37(1)	6.0(2)	394.7(4)	GE Virial
140	0.65(1)	10.1(2)	377.3(4)	GE Virial
150	1.05(2)	16.0(4)	359.5(4)	GE Virial
160	1.63(3)	25.4(5)	339.2(5)	GE Virial
170	2.22(4)	34.7(9)	315.(1.)	GE Virial
180	3.2(1)	55.(3.)	282.(3.)	GE Virial

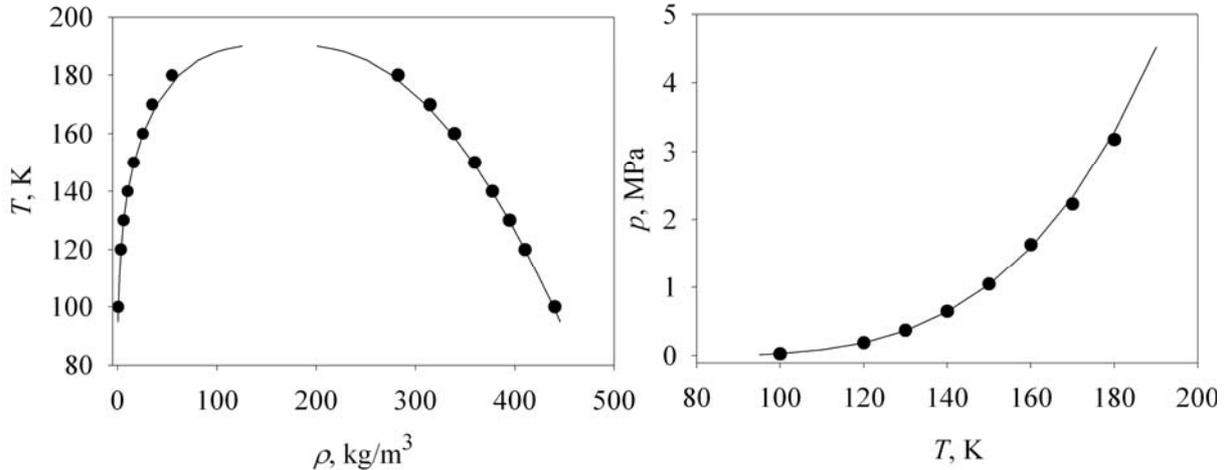


Figure 5.2: Vapour-liquid equilibria of methane. Left-hand picture: liquid and vapour densities at coexistence. Right-hand picture: saturation pressures. The solid circles represent simulated densities and pressures, and the solid lines have been obtained from an equation of state published on the web by NIST [66].

5.4.2. Ethane

In this section, ethane is modelled as a rigid dumbbell with two united-atom Lennard-Jones groups (2CLJ). The force-field parameters, which were optimized to vapour-liquid equilibria [7], are given in Table 5.1. Simulation results for vapour-liquid equilibria, over a broad range of temperatures, along with the statistical uncertainties, are given in Table 5.4.

In a prior publication, simulation results at $T = 144.26$ K and $T = 160$ K were presented. These coexistence points were computed with the $Np\beta$ +test molecule method without exploiting overlapping distributions [67]. Table 5.4 also contains results obtained with the $Np\beta$ +test molecule method with overlapping distributions ($Np\beta$ +TM OID). The results obtained with and without the use of overlapping distributions (approximately) match within their statistical uncertainties. Data at higher temperatures were obtained from Gibbs ensemble simulations. A comparison between experimental and simulated data is given in Figure 5.3.

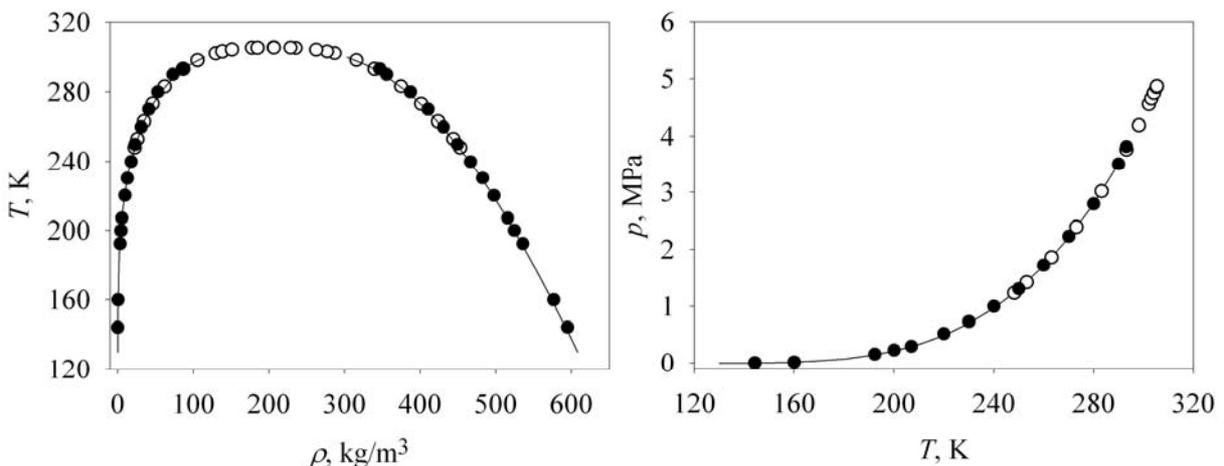


Figure 5.3: Vapour-liquid equilibria of ethane. Left-hand picture: liquid and vapour densities at coexistence. Right-hand picture: saturation pressures. Solid circles represent simulated saturation pressures and liquid and vapour densities at coexistence, open circles represent experimental data published by Douslin et al. [68], and solid lines have been obtained from an equation of state published on the web by NIST [66].

Table 5.4: Simulated vapour-liquid equilibria of ethane. The numbers in parentheses denote the statistical uncertainties in the last digit.

T	p^{sim}	$\rho_{\text{V}}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	Method
K	MPa	kg/m ³	kg/m ³	
144.26	0.006(2)	0.15(9)	594.7(5)	$Np\beta + \text{TM}$
160.00	0.020(2)	0.5(1)	576.5(5)	$Np\beta + \text{TM}$
192.37	0.162(9)	3.2(2)	535.8(5)	$Np\beta + \text{TM}$
144.26	0.0056(5)	0.145(4)	594.(1.)	$Np\beta + \text{TM OID}$
160.00	0.023(1)	0.530(2)	575.6(4)	$Np\beta + \text{TM OID}$
200.00	0.234(9)	4.5(2)	524.6(4)	GE Virial
207.00	0.301(7)	5.6(1)	515.6(4)	GE Virial
220.00	0.52(2)	9.5(5)	497.6(5)	GE Virial
230.00	0.73(2)	13.1(5)	482.7(5)	GE Virial
240.00	1.00(2)	17.8(5)	466.6(6)	GE Virial
250.00	1.30(4)	22.9(8)	450.(2.)	GE Virial
260.00	1.72(4)	31.(1.)	430.6(8)	GE Virial
270.00	2.22(4)	41.1(9)	410.(1.)	GE Virial
280.00	2.79(3)	53.(1.)	387.(1.)	GE Virial
290.00	3.51(5)	73.(3.)	356.(2.)	GE Virial
293.15	3.82(5)	86.(5.)	346.(3.)	GE Virial

As the force-field parameters were optimized to vapour-liquid equilibria, it is not surprising that simulation results and experimental data perfectly agree. Even at (very) low temperature, the experimental and simulation results agree very well. This is a very good performance for such a simple force field.

5.4.3. Carbon dioxide

The vapour-liquid equilibria of carbon dioxide, which are simulated with the TraPPE-AA force field, are presented in Table 5.5 and in Figure 5.4. The statistical uncertainties of the results are also given in Table 5.5.

Potoff and Siepmann [23] published their results in graphical form only. However, from a graphical comparison, it is known that their predictions are comparable with those presented in this work; the match between experimental and simulated data is quite good over a broad range of temperatures.

Table 5.5: Simulated vapour-liquid equilibria of carbon dioxide. The numbers in parentheses denote the statistical uncertainties in the last digit.

T	p^{sim}	$\rho_{\text{V}}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	Method
K	MPa	kg/m ³	kg/m ³	
220	0.64(3)	16.7(5)	1160.(2)	GE Virial
230	0.87(3)	22.2(8)	1123.(3)	GE Virial
240	1.25(5)	32.(2.)	1085.(1)	GE Virial
250	1.77(5)	46.(1.)	1046.(3)	GE Virial
260	2.42(8)	63.(3.)	997.(5.)	GE Virial
270	3.16(4)	82.(1.)	950.(3.)	GE Virial
280	4.26(2)	117.8(6)	895.(1.)	GE Virial

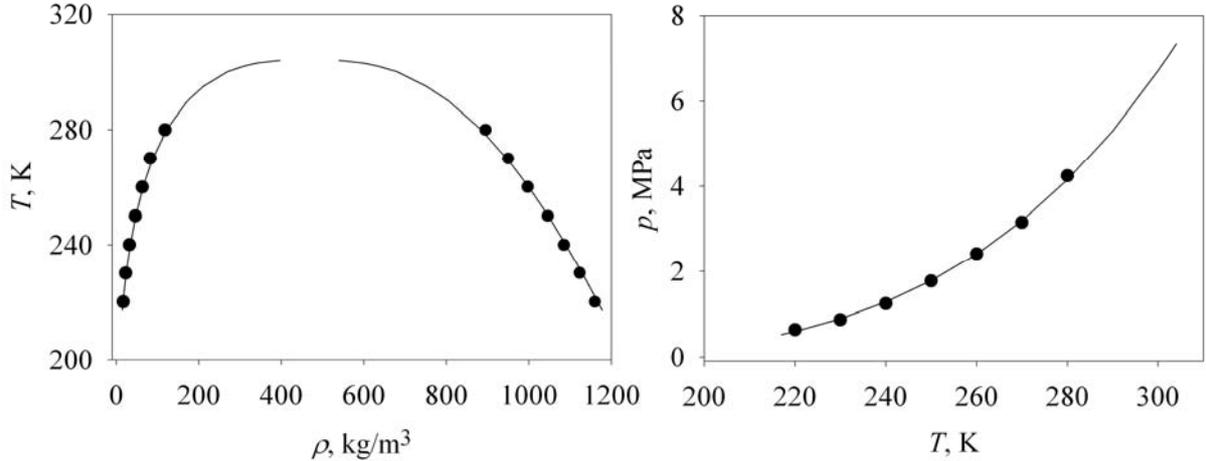


Figure 5.4: Vapour-liquid equilibria of carbon dioxide. Left-hand picture: liquid and vapour densities at coexistence. Right-hand picture: saturation pressures. The solid circles represent simulated densities and pressures, and the solid lines have been obtained from an equation of state published on the web by NIST [66].

5.4.4. DMSO

The force-field parameters for DMSO, as given by Rao *et al.* [34], have not been optimized to vapour-liquid equilibria. Moreover, there are no experimental vapour densities to compare the simulation results obtained in this work with. Therefore, the capabilities of the force field to predict phase behaviour of pure DMSO over a broad range of temperatures have not been checked.

Simulation results, along with the corresponding statistical uncertainties, are given in Table 5.6. and in Figure 5.5. The simulations for DMSO give reasonable liquid densities, as was to be expected since the force-field parameters have been optimized to liquid-phase data. Unfortunately, the predicted vapour-liquid saturation pressures are much too high.

In one of the author's prior publications [67], simulated vapour-liquid equilibria were presented that had been computed with the $Np\beta$ +test molecule method. From Chapter 4, it is known that the $Np\beta$ +test molecule method combined with overlapping distributions is applicable at higher densities and has built-in diagnostics. Simulation results, obtained with the $Np\beta$ +test molecule method with and without exploiting overlapping distributions, are given in Table 5.6.

Table 5.6: Simulated vapour-liquid equilibria of DMSO. The numbers in parentheses denote the uncertainties in the last digit(s).

T K	p^{sim} MPa	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	Method
303.15	0.0092(4)	0.3(1)	1091.(2.)	$Np\beta$ + TM
328.94	0.016(8)	0.5(1)	1049.(3.)	$Np\beta$ + TM
278.5	0.0017(1)	0.0572	1122.82	$Np\beta$ + TM OID
303.15	0.0063(2)	0.1967	1087.023	$Np\beta$ + TM OID
328.94	0.019(1)	0.5685	1051.97	$Np\beta$ + TM OID

The saturation pressures computed with both methods at $T = 328.94$ K match within their statistical uncertainties. This is not the case for the results obtained at $T = 303.15$ K. In this

chapter, the saturation pressure computed at $T = 303.15$ K with the simple $Np\beta$ +test molecule method is used as a starting point for Gibbs-Duhem integration [See Section 5.4.8]. As the results obtained from the overlapping-distributions $Np\beta$ +test molecule method are generally more accurate, those values will be used in Chapter 6.

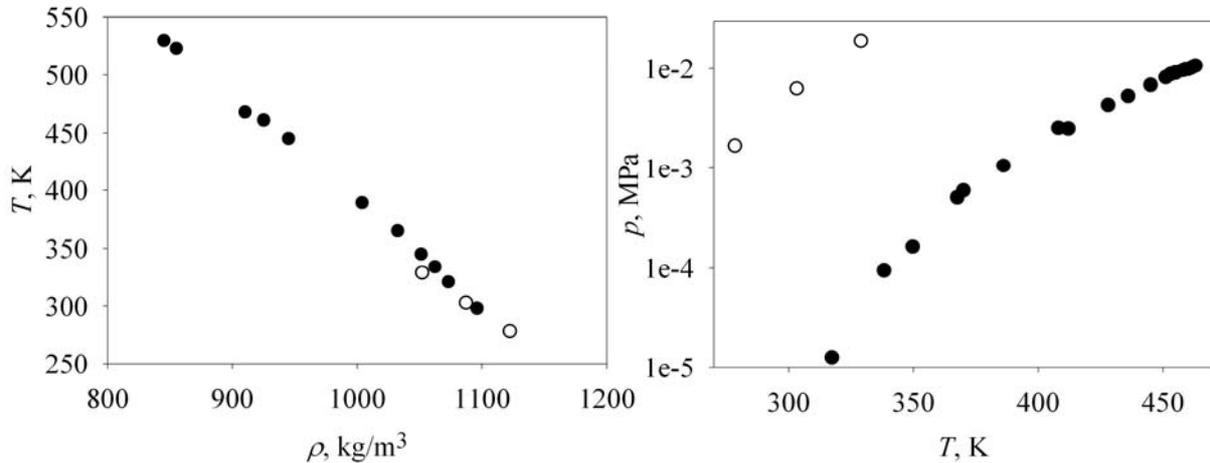


Figure 5.5: Vapour-liquid equilibria of DMSO. Left-hand picture: liquid densities at coexistence. Right-hand picture: saturation pressures. Solid circles are experimental data published by Campbell [69]. Open circles represent simulation results.

5.4.5. Difluoromethane (R32)

Experimental and simulated liquid and vapour densities at coexistence, saturation pressures, and corresponding statistical uncertainties are given in Table 5.7 and in Figure 5.6. Results were computed in the Gibbs ensemble. The pressure was obtained from volume perturbations (VP).

The simulation results are close to the experimental data. Like the results of Lísal *et al.* [37] and of Higashi and Takada [36], the predicted saturation pressures presented in this work are slightly lower than the experimental ones.

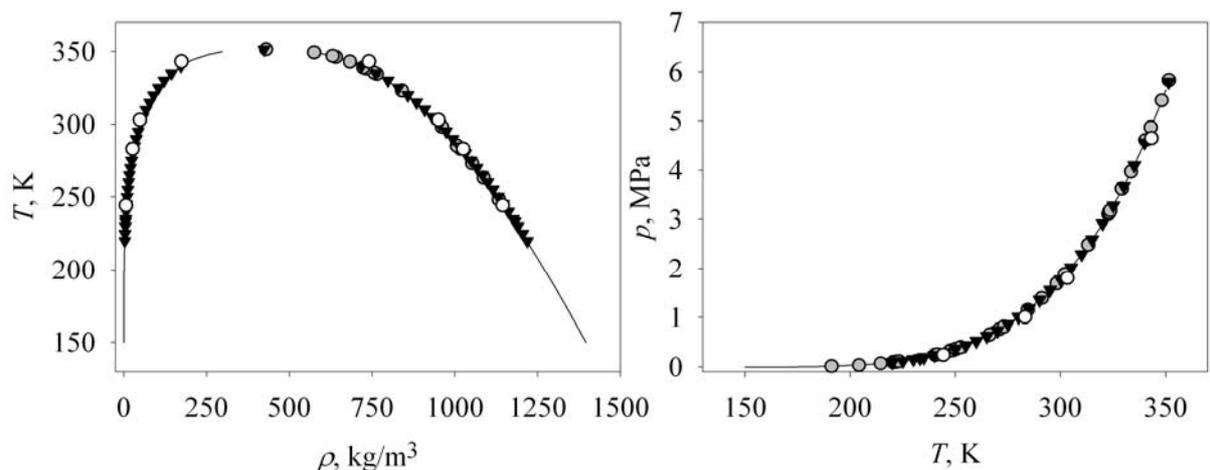


Figure 5.6: Vapour-liquid equilibria of difluoromethane. Left-hand picture: liquid and vapour densities at coexistence. Right-hand picture: saturation pressures. Solid grey circles, solid black downward triangles and solid lines respectively represent experimental liquid densities at coexistence published by Malbrunot *et al.* [70], saturation pressures and liquid and vapour densities at coexistence published by Defibaugh *et al.* [71], and correlated data published on the web by NIST [66]. Open circles are simulation results obtained in this work.

Table 5.7: Simulated vapour-liquid equilibria of difluoromethane. The numbers in parentheses denote the uncertainties in the last digit.

T K	p^{sim} MPa	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	Method
244.26	0.241(2)	6.64(6)	1144.3(9)	GE VP
283.15	1.008(6)	26.6(2)	1025.(1.)	GE VP
303.15	1.80(1)	48.9(3)	950.(1.)	GE VP
343.26	4.65(5)	174.(3.)	740.(4.)	GE VP

5.4.6. Pure trifluoromethane (R23)

Experimental and simulated liquid and vapour densities at coexistence, saturation pressures, and the corresponding statistical uncertainties are given in Table 5.8 and in Figure 5.7. As indicated in Table 5.8, the result at the lowest temperature has been computed from the $Np\beta$ + test molecule method combined with overlapping distributions. The other results were obtained from Gibbs ensemble simulations. The pressure in the Gibbs ensemble was computed from volume perturbations.

Table 5.8: Simulated vapour-liquid equilibria of trifluoromethane. The numbers in parentheses denote the uncertainties in the last digit(s).

T K	p^{sim} MPa	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	Method
180.	0.062(2)	2.99(5)	1500.(1.)	$Np\beta$ + TM OID
254.	1.37(1)	60.4(5)	1179.(1.)	GE VP
263.15	1.89(1)	87.1(9)	1133.(2.)	GE VP
273.15	2.39(1)	113.1(9)	1063.(2.)	GE VP
283.25	3.08(2)	158.(2.)	984.(2.)	GE VP
293.15	3.87(2)	300.(10.)	783.(11.)	GE VP

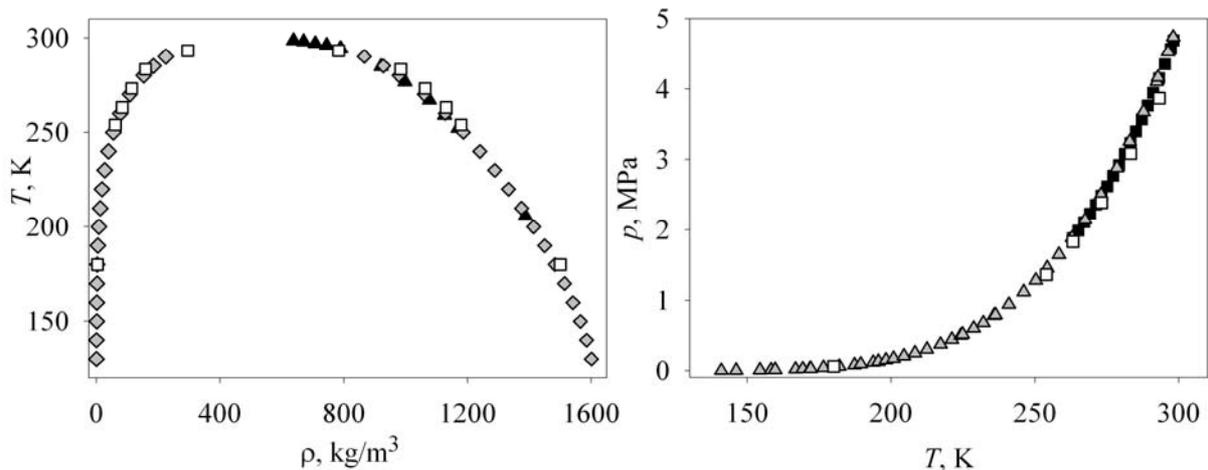


Figure 5.7: Vapour-liquid equilibria of trifluoromethane. Left-hand picture: liquid and vapour densities at coexistence. Right-hand picture: saturation pressures. Upward grey triangles represent experimental data published by Hou and Martin [72]. Solid grey diamonds originate in the equation of state of Rubio et al. [73]. Solid black squares are unpublished data from Mooijer et al. [74]. Open squares are simulation results obtained in this work.

The good agreement of simulated and experimental data is surprising as the employed force field is far from realistic. Like the results published by Song *et al.* [45], the saturation pressures presented in this section get progressively worse at higher temperatures. Notice the large statistical uncertainties of the results obtained at $T = 293.15$ K, which indicate the limitations of the Gibbs ensemble near the vapour-liquid critical point.

5.4.7. Binary mixture methane-ethane

This binary system has been chosen for several reasons, the main of which is its simplicity and the availability of simple force fields. Sources of experimental data for vapour-liquid equilibria of the binary system methane/ethane are given in Table 5.9. An overview of data sets published before 1976 was given by Miller *et al.* [75]. Table 5.9 mainly includes recent data sets.

Table 5.9: Sources of experimentally determined vapour-liquid equilibria of the system methane/ethane

Reference	Temp. range, K	Press. range, MPa	Number of sets
Davalos <i>et al.</i> [76, 77]	250	1.3-8	1
Wichterle and Kobayashi [78]	130.37-199.92	0.001-5	12
Wei <i>et al.</i> [79]	210-270	1.8-6.5	4
Gupta <i>et al.</i> [80]	260, 270, 280	1.7-7.0	3
Price [81]	144.26-283.15	0.05-5.5	6
Miller <i>et al.</i> [75]	160, 180	0.02-3.3	2

The vapour-liquid equilibria of this mixture have been simulated with a variety of methods, force fields, and combining rules. Vrabec and Fischer simulated the system with the mixture extension of the $Np\beta$ + test molecule method at $T = 160$ K, $T = 199.93$ K, and $T = 250$ K [7]. They used a 1CLJ model for methane [17, 18] and a 2CLJ model for ethane [7]. Binary interaction parameters were included in the LB combining rules. The agreement of experimental and simulated data was very good.

Liu and Beck [57] studied the system at $T = 250$ K with the intermolecular potentials and binary interaction parameters that were also used by Vrabec and Fischer [7]. However, Liu and Beck performed their simulations in the $Np\beta$ – Gibbs ensemble.

Zhang and Duan [16] used temperature-scaling Gibbs ensemble simulations to predict vapour-liquid equilibria of methane/ethane at $T = 180$ K, $T = 210$ K, $T = 250$ K, and $T = 280$ K. They used slightly modified versions of the OPLS force field for methane (1CLJ) [12] and the TraPPE-UA force field [8] for ethane (2CLJ), along with the commonly used LB combining rules. The agreement of their simulation results for pure-components and binary mixtures with experimental data was quite good.

De Pablo and Prausnitz [82] performed simulations at $T = 213.7$ K in the $Np\beta$ – Gibbs ensemble. They used a 1CLJ model for methane and a 2CLJ model for ethane. The force-field parameters were taken from the OPLS force field [12]. The OPLS force field exploits geometric combining rules for both well-depth and collision-diameter parameters. However, de Pablo and Prausnitz also performed simulations with an arithmetic combining rule for the collision-diameters (LB), resulting in better agreement with experimental data.

In this section, the 1CLJ force field for methane as given by Lofti [17, 18], and the 2CLJ united-atom force field for ethane published by Vrabec and Fischer [7] are used. These force fields are by no means transferable but their simplicity and known power to predict vapour-liquid equilibria in close agreement with experimental data are attractive. The well-predicted pure-component saturation pressures make this binary mixture a good test case for the LB

combining rules; as the points of suspension of the phase envelope as well as the vapour-liquid pure-component phase behaviour are well predicted, the agreement between simulated and experimental mixture phase behaviour mainly depends on the interactions between dissimilar components .

In order to start numerical integration of the Clapeyron equation over a range of fugacity fractions, ξ_{CH_4} , one needs the liquid and vapour densities at coexistence and the saturation pressure of pure ethane along with the ratios $f_{\text{C}_2\text{H}_6}/H_{\text{CH}_4}$ at the temperature of interest.

In this section, simulation results are presented for the vapour-liquid equilibria of the binary system methane/ethane at three temperatures, $T = 144.26$ K, $T = 160.00$ K, and $T = 192.37$ K. The simulated saturation pressures and liquid and vapour densities at coexistence of pure ethane at these temperatures were taken from Table 5.4. The ratios $f_{\text{C}_2\text{H}_6}/H_{\text{CH}_4}$ are given in Table 5.10. In a prior publication [67], the author of this work presented simulation results for the ratios $f_{\text{C}_2\text{H}_6}/H_{\text{CH}_4}$ that were computed from the difference method (DM). Table 5.10, also contains simulation results obtained from the overlapping-distributions difference method (DM+OID). This method is known to be more accurate [see Chapter 4]. For the latter simulation results, the statistical uncertainty has been computed as well.

Table 5.10: Simulated ratios of solvent's (ethane) fugacity over solute's (methane) Henry's constant.

T K	p MPa	$\left(\frac{f_{\text{C}_2\text{H}_6}}{H_{\text{CH}_4}}\right)_{\text{L}}^{\text{sim}}$	$\left(\frac{f_{\text{C}_2\text{H}_6}}{H_{\text{CH}_4}}\right)_{\text{V}}^{\text{sim}}$	Method
144.26	0.0059	0.00212	0.995	DM in $Np\beta$ ensemble
160.00	0.0196	0.00178	0.990	DM in $Np\beta$ ensemble
192.37	0.162	0.0111	0.951	DM in GE
144.26	0.0056	0.0067(3)	1.02(5)	DM + OID
160.00	0.023	0.0136(4)	1.00(3)	DM + OID
192.37	0.162	0.0406(9)	0.97(2)	DM + OID

Simulation results for binary vapour-liquid equilibria are presented in Table 5.11 and in Figure 5.8 to Figure 5.10. Error bars for the mole fractions are given in the figures. Figure 5.9 also contains simulation results published by Vrabec and Fischer [7]. The force fields that were used in this work are identical to those used by Vrabec and Fischer. However, Vrabec and Fischer used binary interaction parameters in the LB combining rules.

The good agreement of simulation results and experimental data is obvious. The Gibbs-Duhem integration method traces the whole coexistence curve, starting from the low saturation pressure of pure ethane, without any difficulties. The accuracy of the initial coexistence points is evident in the tendency towards the correct saturation pressure of pure methane.

The statistical uncertainties of the mole fractions at $T = 160.00$ K are smaller than those given by Vrabec and Fischer [7]. However, the Gibbs-Duhem integration method suffers from other sources of errors as well, and a direct comparison based on statistical errors is therefore not fair. A brief error analysis of the Gibbs-Duhem integration method is given in Section 5.5. A comparison between simulation results and experimental data shows that the LB combining rules do a perfect job and that no binary interaction parameters are needed. Apart from that, it is questionable whether the effect of incorporation of small binary interaction parameters is larger than the statistical uncertainties of the simulations.

At $T = 192.37$ K, methane is supercritical, which prevented accurate simulations closer to the vapour-liquid critical point of the mixture than the results given in Table 5.11 and in Figure 5.10. The Gibbs-Duhem integration method fails close to the critical point because the liquid and vapour probability distributions that describe the probability to observe a simulating box with a given composition, volume and energy, have a region of overlap. The closer the system comes to the critical point, the easier unilateral phase changes occur.

Another source of error in the vicinity of the critical point is a result of finite-size effects; the small simulation box with periodic boundaries cannot handle long-range fluctuations. No attempts have been made to simulate closer to the critical point.

Table 5.11: Simulated vapour-liquid equilibria of the binary system ethane/methane at different temperatures

$T = 144.26$ K									
p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.00590	0.000	0.000	595.	0.152	0.302	0.328	0.983	539.	4.03
0.00985	0.00460	0.399	594.	0.210	0.393	0.443	0.989	518.	5.68
0.0148	0.00990	0.598	592.	0.271	0.445	0.520	0.990	501.	6.42
0.0298	0.0255	0.797	590.	0.471	0.509	0.610	0.993	480.	7.49
0.0594	0.0559	0.898	586.	0.870	0.547	0.664	0.994	468.	8.05
0.0732	0.0690	0.918	583.	1.06	0.589	0.726	0.995	452.	8.78
0.113	0.119	0.948	576.	1.61	0.637	0.792	0.996	433.	9.58
0.176	0.187	0.968	566.	2.53	0.688	0.863	0.997	414.	10.4
0.244	0.252	0.979	553.	3.46	0.743	0.934	0.999	392.	11.3
$T = 160.00$ K									
p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.0230	0.00000	0.000	576.	0.53	0.328	0.179	0.935	546.	4.35
0.0287	0.00290	0.197	575.	0.59	0.381	0.221	0.948	539.	5.03
0.0380	0.0104	0.398	573.	0.71	0.454	0.258	0.956	533.	6.05
0.0565	0.0188	0.597	572.	0.93	0.563	0.329	0.967	519.	7.50
0.0745	0.0291	0.694	571.	1.15	0.736	0.450	0.977	496.	10.0
0.0886	0.0364	0.748	569.	1.32	0.863	0.536	0.982	476.	11.9
0.109	0.0475	0.797	570.	1.58	1.04	0.661	0.988	446.	14.6
0.210	0.105	0.895	560.	2.85	1.27	0.830	0.994	398.	18.6
0.256	0.137	0.915	554.	3.45	1.49	0.969	0.999	349.	22.1
0.288	0.163	0.927	550.	3.84					
$T = 192.37$ K									
p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.162	0.00000	0.000	536.	3.19	1.56	0.344	0.901	471.	19.8
0.202	0.00850	0.194	534.	3.59	1.72	0.379	0.911	464.	22.0
0.266	0.0263	0.386	529.	4.29	1.91	0.430	0.920	452.	24.5
0.389	0.0546	0.584	526.	5.60	2.14	0.491	0.931	438.	28.0
0.507	0.0853	0.682	521.	6.93	2.44	0.564	0.941	420.	32.9
0.598	0.110	0.734	517.	7.92	2.82	0.654	0.953	393.	34.0
0.729	0.137	0.779	512.	9.44	3.31	0.775	0.964	352.	49.0
1.32	0.289	0.881	481.	16.5	3.59	0.834	0.971	334.	56.2

In Section 4.2, the computation of the initial coexistence point that is needed to start the numerical integration is discussed. The choice of the method to compute this point depends on the density of the liquid phase and on the density-difference between the liquid and the vapour phase. As shown in Table 5.4, the starting points at $T = 144.26$ K and $T = 160.00$ K have been computed from the $Np\beta$ -test molecule method. However, at these temperatures there was no need to start the numerical integration from the higher boiling component (ethane); the Gibbs-Duhem integration procedure could have been started from pure methane, which has a much higher saturation pressure. At this pressure, the density difference between the liquid and the vapour phase of pure methane is small enough to perform a Gibbs ensemble simulation. Nevertheless, this section shows the capability of the combination of the $Np\beta$ -test molecule method and Gibbs-Duhem integration to predict the phase envelope from the (very) low-pressure side. Moreover, when the second component is supercritical, as is the case for methane at $T = 192.37$ K, one simply has to start from the saturation pressure of the pure high-boiling component.

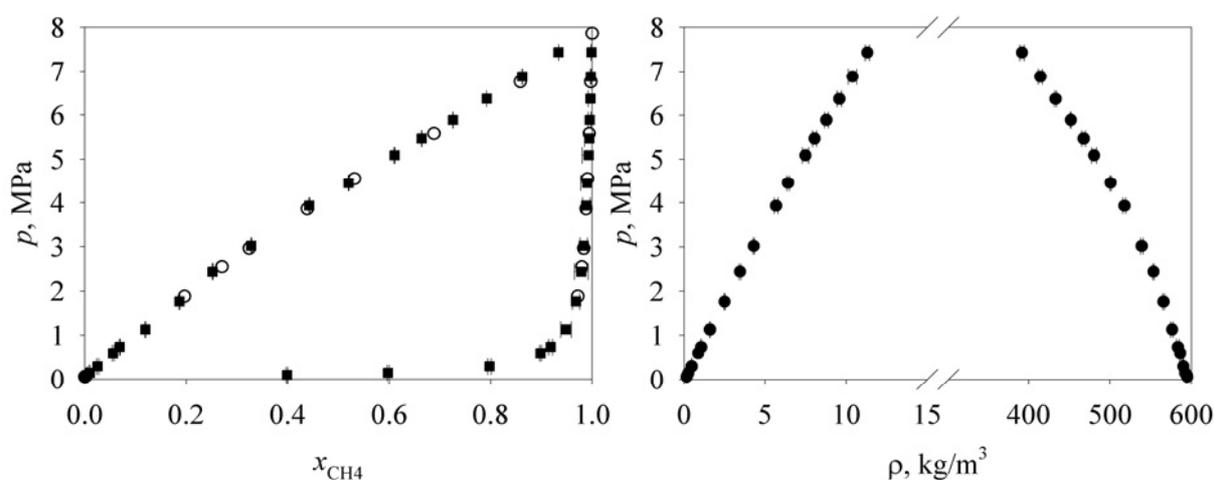


Figure 5.8: Experimental and simulated vapour-liquid equilibria of the binary mixture ethane/methane at $T = 144.26$ K. Open and solid symbols represent experimental [78] and simulated coexistence data respectively. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour densities at coexistence.

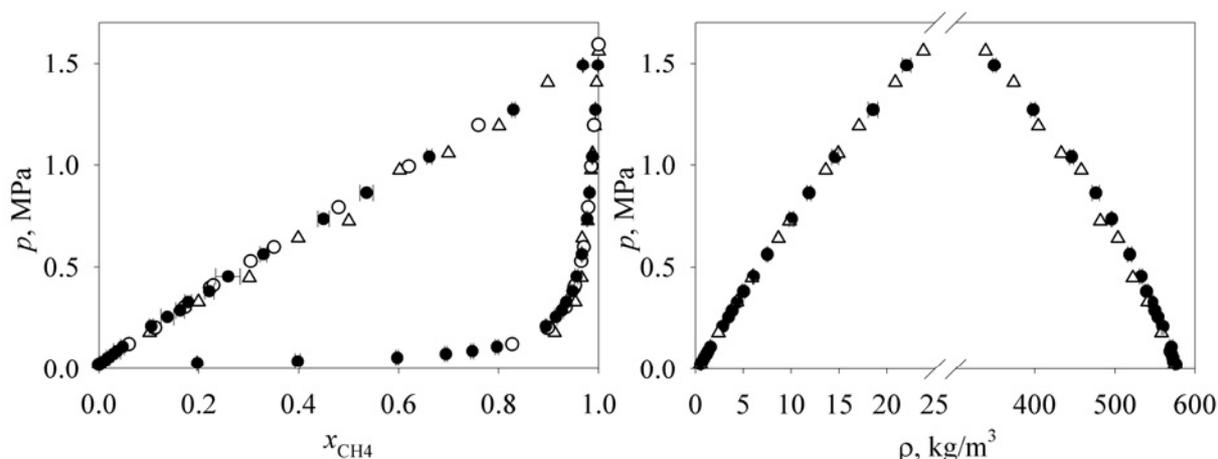


Figure 5.9: Experimental and simulated vapour-liquid equilibria of the binary mixture ethane/methane at $T = 160.0$ K. Open and solid circles respectively represent experimental [75] and simulated coexistence data. Open upward triangles represent simulated vapour-liquid equilibria by Vrabec and Fischer [7]. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour densities at coexistence.

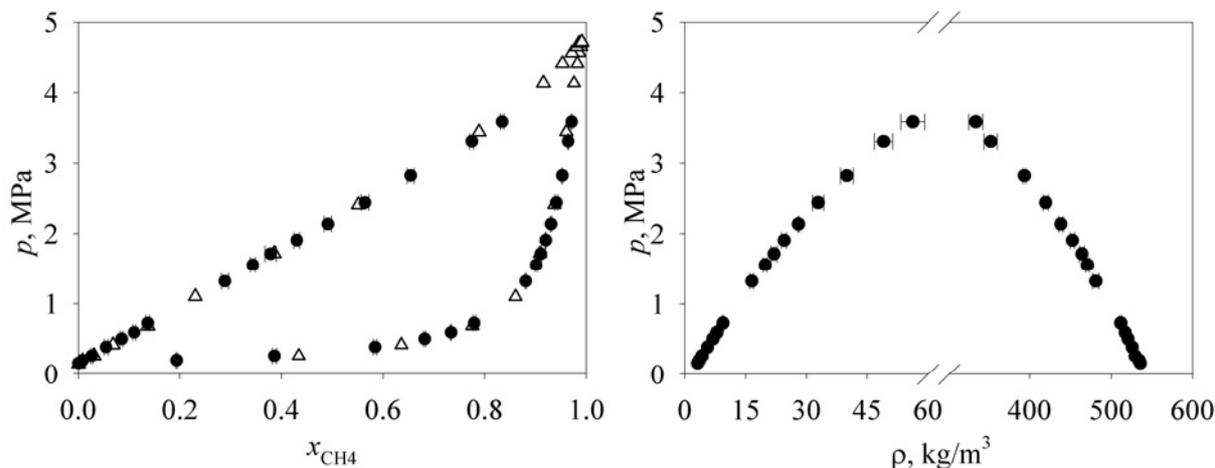


Figure 5.10: Experimental and predicted vapour-liquid equilibria of the binary mixture ethane/methane at $T = 192.37$ K. Open upward triangles and solid circles represent experimental [78] and simulated coexistence data respectively. Left-hand picture: p - x_{CH_4} -data. Right-hand picture: liquid and vapour densities at coexistence.

5.4.8. Binary system DMSO/carbon dioxide

Sources of experimental data for vapour-liquid equilibria of the binary system DMSO/carbon dioxide are given in Table 5.12. This mixture is used in the gas-antisolvent process to recrystallize particles (explosives, polymers, pharmaceuticals, etc.) [83, 84]. It has not been studied before with molecular simulation.

Table 5.12: Sources of experimentally determined vapour-liquid equilibria of the system DMSO/carbon dioxide

Reference	Temp. range, K	Pressure range, MPa	Number of sets
Vega Gonzalez <i>et al.</i> [85]	309.44-328.94	4.1-12.6	4
Kordikowski <i>et al.</i> [86]	298.15, 303.15	0.2-6.6	2
Rajasingam <i>et al.</i> [87]	298.2, 398.2, 318.2	1.0-9.0	3

The initial point that is used to start the numerical integration is the saturation pressure of pure DMSO at the temperature of interest. Furthermore, the liquid and vapour densities at coexistence, along with the ratios $f_{\text{DMSO}}/H_{\text{CO}_2}$, are needed to determine the starting value of the integrand of the Clapeyron equation.

In this section, simulation results for the vapour-liquid equilibria of the binary system DMSO/carbon dioxide at $T = 303.15$ K are presented. The saturation pressure and liquid and vapour densities at coexistence of DMSO at this temperature are taken from Table 5.6. The ratios $f_{\text{DMSO}}/H_{\text{CO}_2}$ are given in Table 5.13. DMSO and carbon dioxide were respectively modelled with the force field of Rao and Singh [34] and the TraPPE-AA force field [23]. The interactions between unlike interaction sites are modelled with the LB combining rules. In one of the author's publications [67], simulation results for the ratios $f_{\text{DMSO}}/H_{\text{CO}_2}$ computed from the difference method (DM) were presented. Table 5.13 also contains simulation results obtained with the overlapping-distributions difference method (DM + OID), which is known to be more accurate at high densities [See Chapter 4]. Nevertheless, the Gibbs-Duhem-integrations that are presented in this section were started from the initial point and initial slope computed without exploiting overlapping distributions.

The simulation results for the vapour-liquid equilibria of the binary mixture can be found in Table 5.14. These results are also presented in graphical form in Figure 5.11, together with

bubble-point data published by Kordikowski *et al.* [86]. Error bars for the mole fractions are given in the plots.

The combination of the $Np\beta$ + test molecule method and Gibbs-Duhem integration provides us with predictions of vapour-liquid equilibria where experiments can hardly be performed. The bubble-point data of Kordikowski *et al.* [86] are reproduced. Notice that no vapour-phase data were available to compare the simulation results with.

The convergence of the corrector iterations/simulations and the composition sampling for this binary mixture are poor, as can be seen from the scattering and the error bars of the liquid-phase mole fractions. The predictor-corrector method does not provide accurate predictions in the limit of pure carbon dioxide where the integrand of the Clapeyron equation rapidly changes.

In the diluted regions, only a few molecules take the other identity. Composition sampling in the liquid phase is further deteriorated by the high density of the liquid phase, and by the differences in size, shape, and polarity between DMSO and carbon dioxide. However, the latter two problems may be solved by increasing the system size and/or by exploiting energy-biased identity changes.

The sampling problems described above prevented accurate predictions at higher pressures than those shown in Figure 5.11. In Chapter 6, it will be shown that the use of the energy bias in the identity-change trial move improves the sampling of the composition considerably.

Table 5.13: Simulated ratios of solvent's (DMSO) fugacity over solute's (carbon dioxide) Henry's constant

T K	p MPa	$\left(\frac{f_{\text{DMSO}}}{H_{\text{CO}_2}}\right)_{\text{L}}^{\text{sim}}$	$\left(\frac{f_{\text{DMSO}}}{H_{\text{CO}_2}}\right)_{\text{V}}^{\text{sim}}$	Method
303.15	0.0092(4)	0.000	0.957	DM in $Np\beta$ ensemble
303.15	0.0063(2)	0.0008(1)	1.00(6)	DM + OI in $Np\beta$ ensemble

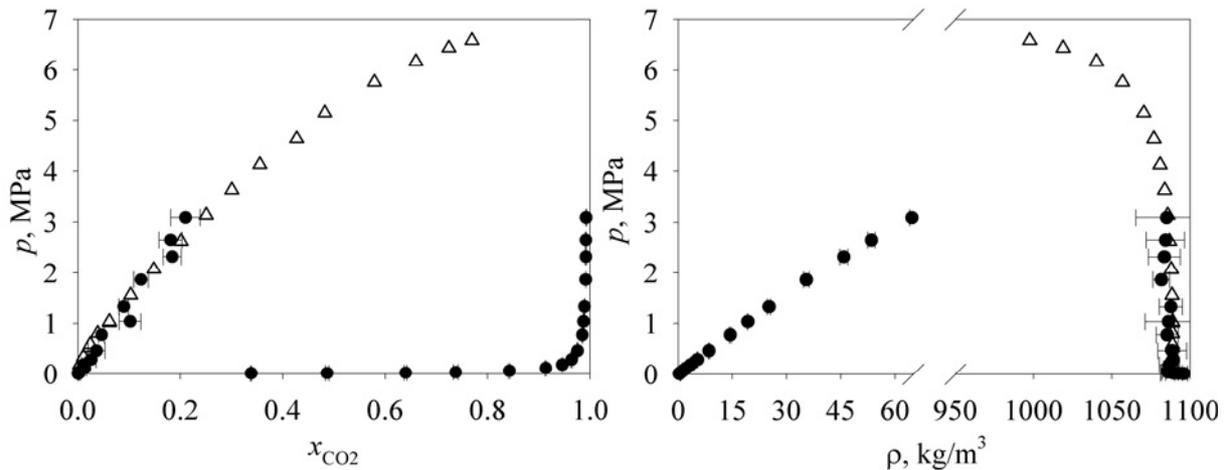


Figure 5.11: Vapour-liquid equilibria of the binary mixture DMSO/carbon dioxide at $T = 303.15$ K. Open upward triangles represent experimental bubble-point data [86], while solid circles are simulated coexistence data. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour-phase densities at coexistence.

Table 5.14: Simulated vapour-liquid equilibria of the binary system DMSO/carbon dioxide at $T = 303.15\text{ K}$

p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.00917	0.0000	0.000	1096.	0.312	0.466	0.0362	0.976	1089.	8.52
0.0143	0.0002	0.338	1090.	0.383	0.779	0.0462	0.985	1085.	14.4
0.0185	0.0005	0.487	1093.	0.461	1.04	0.1017	0.988	1086.	19.2
0.0263	0.0029	0.640	1088.	0.590	1.33	0.0892	0.989	1088.	25.1
0.0367	0.0012	0.738	1086.	0.781	1.86	0.1232	0.992	1082.	35.4
0.0613	0.0047	0.843	1086.	1.22	2.32	0.1841	0.992	1083.	45.8
0.117	0.0134	0.914	1087.	2.18	2.65	0.1808	0.992	1084.	53.4
0.187	0.0100	0.946	1087.	3.43	3.09	0.2099	0.993	1085.	64.6
0.291	0.0258	0.964	1089.	5.31					

5.5. Error analysis of Gibbs-Duhem integration

5.5.1. Propagation of statistical errors

A thorough error analysis of Gibbs-Duhem integration is hard. For that reason, only the major sources of error are considered in this section. The errors arising from the use of finite-order corrector equations will be neglected.

The integrands of the Clapeyron equation at successive state points are computed with Monte Carlo simulation. The ensemble averages of the integrands have a statistical uncertainty. As the estimates of subsequent saturation pressures are coupled via the corrector equations, the variance of the saturation pressure at a certain state point depends on the variances of predetermined saturation pressures/integrands. The set of corrector equations (every integration step is assigned a corrector equation, see Chapter 2) forms a linear system that can be written in the following form:

$$\mathbf{p} = [\boldsymbol{\beta}]\mathbf{F} + \mathbf{p}' \quad (5.18)$$

The matrix $[\boldsymbol{\beta}]$ contains either zeros or values that depend solely on the step sizes of the integration process. The vectors \mathbf{p} , \mathbf{p}' , and \mathbf{F} contain the saturation pressures and integrands corresponding to successive integration steps. After successive substitution of corrector equations into one another, one obtains the influence of the saturation pressure p_n at integration step n , on the predetermined integrands $\{F_i, i = 0 \dots n\}$.

The result for the constant-step-size corrector equations used in refs. [88, 89] is:

$$\mathbf{p} = \begin{bmatrix} \sum_{i=0}^1 \beta_{1i} F_i \\ \sum_{i=0}^2 \beta_{2i} F_i \\ \sum_{i=0}^3 \left(\sum_{j=2}^3 \beta_{ji} \right) F_i \\ \vdots \\ \sum_{i=0}^n \left(\sum_{j=2}^n \beta_{ji} \right) F_i \end{bmatrix} + p_0 \quad (5.19)$$

where p_0 is the saturation pressure corresponding to the initial coexistence point. The result for the variable-step-size corrector equations used in refs. [65, 67] and in this chapter is:

$$\mathbf{p} = \begin{bmatrix} \sum_{i=0}^1 \beta_{1i} F_i \\ \sum_{i=0}^2 \left(\sum_{j=1}^2 \beta_{ji} \right) F_i \\ \vdots \\ \sum_{i=0}^n \left(\sum_{j=1}^n \beta_{ji} \right) F_i \end{bmatrix} + p_0 \quad (5.20)$$

It is convenient to replace the summations over different elements of the matrix $[\beta]$ by elements of a new matrix $[\Omega]$:

$$\mathbf{p} = [\Omega] \mathbf{F} + p_0 \quad (5.21)$$

The variance of the pressure at state point n is given by:

$$\sigma_{p_n}^2 = \Omega_n \Sigma(\mathbf{F}) \Omega_n^T \quad (5.22)$$

where $\Sigma(\mathbf{F})$ is the variance-covariance matrix of the integrands, and Ω_n is the n^{th} row vector of the matrix $[\Omega]$. As the integrands at successive integration points are statistically independent, $\Sigma(\mathbf{F})$ is a diagonal matrix with the variances of the elements of \mathbf{F} on the diagonal.

The version of the Clapeyron equation that is numerically integrated in this work is:

$$\left(\frac{\partial \ln p}{\partial \xi_2} \right)_{\beta, \sigma} = F(\xi_2, p) \quad (5.23)$$

This means that the saturation pressure in Eqs. (5.18) to (5.22) is to be substituted by $\ln p$. The variance in the saturation pressure at integration point n can be computed from the error propagation law:

$$\sigma_{p_n}^2 = \sigma_{\ln p_n}^2 \left(\frac{\partial p_n}{\partial \ln p_n} \right)^2 = p_n^2 \sigma_{\ln p_n}^2 \quad (5.24)$$

The propagation of the statistical uncertainties via the corrector equations in the binary mixture ethane/methane at $T = 193.27$ K is given in Table 5.15.

Table 5.15: Propagation of statistical errors; standard deviation of predicted saturation pressures at successive integration steps in the binary mixture methane/ethane at $T = 193.27$ K

p^{sim} MPa	$\sigma_{p^{\text{sim}}}$ MPa								
0.180	0.09	0.348	0.09	0.837	0.0901	1.68	0.0906	2.78	0.0939
0.201	0.09	0.388	0.09	0.984	0.0901	1.87	0.0908	3.00	0.0958
0.229	0.09	0.438	0.09	1.12	0.0902	1.98	0.091	3.26	0.0999
0.246	0.09	0.506	0.09	1.20	0.0902	2.11	0.0912	3.49	0.1043
0.265	0.09	0.576	0.09	1.29	0.0903	2.25	0.0916	3.72	0.1164
0.288	0.09	0.642	0.09	1.40	0.0903	2.40	0.092	3.84	0.1201
0.315	0.09	0.727	0.0901	1.53	0.0904	2.58	0.0927	4.03	0.1304

As is clear from Table 5.15, it is the statistical uncertainty of the saturation pressure corresponding to the initial coexistence point, p_0 , that mainly determines the statistical error of the saturation pressures at successive integration points. Thus, in case the statistical error of the initial saturation pressure is small, the statistical error of the estimated saturation pressure after a number of integrations steps is probably small as well for simulations of sufficient length.

5.5.2. Influence of the error in the initial coexistence point on the deviation from the true coexistence line

The numerical integration commences from an initial coexistence point. The estimated saturation pressure at the initial coexistence point has a statistical uncertainty and will deviate from the *true* saturation pressure at the temperature of interest. As a consequence, the liquid-phase value of $\ln([\hat{f}_1 + \hat{f}_2]/p^{\text{ref}})$ [see Eq. (2.18)] will be different from the vapour-phase value. The Gibbs-Duhem integration method is designed to keep the difference $\Delta^{\text{LV}} \ln([\hat{f}_1 + \hat{f}_2]/p^{\text{ref}})$ between the phases constant.

Kofke [90] analyzed the influence of the error in the initial coexistence point on the difference between the true outcome and the result obtained from Gibbs-Duhem integration. Such a comparison is only possible when the true coexistence line is known. Unfortunately, this is

not the case in this work. The analysis goes as follows. The variable that is fixed throughout the series of integration steps is the temperature. During numerical integration, the fugacity fraction is the independent variable and the pressure the dependent one.

The Taylor series expansion of $\ln([\hat{f}_1 + \hat{f}_2]/p^{\text{ref}})$, up to first order in the pressure, is given by:

$$\begin{aligned} \ln \left\{ \frac{\hat{f}_1 + \hat{f}_2}{p^{\text{ref}}} \right\} (p; [\xi_2]_0, \beta) &\approx \ln \left\{ \frac{\hat{f}_1 + \hat{f}_2}{p^{\text{ref}}} \right\} (p^{\text{eq}}; [\xi_2]_0, \beta) + \\ &\quad \left(\frac{\partial \ln \left\{ (\hat{f}_1 + \hat{f}_2) / p^{\text{ref}} \right\}}{\partial p} \right)_{[\xi_2]_0, p^{\text{eq}}, \beta} (p - p^{\text{eq}}) \quad (5.25) \\ &= \ln \left\{ \frac{\hat{f}_1 + \hat{f}_2}{p^{\text{ref}}} \right\} (p^{\text{eq}}; [\xi_2]_0, \beta) + \beta \langle V \rangle_{[\xi_2]_0, p^{\text{eq}}, \beta} (p - p^{\text{eq}}) \end{aligned}$$

Subtraction of the truncated Taylor series expansion for the liquid phase from the expansion for the vapour phase gives:

$$\Delta^{\text{LV}} \ln \left\{ \frac{\hat{f}_1 + \hat{f}_2}{p^{\text{ref}}} \right\} (p; [\xi_2]_0, \beta) \approx \beta \Delta^{\text{LV}} \langle V \rangle_{[\xi_2]_0, p^{\text{eq}}, \beta} (p - p^{\text{eq}}) \quad (5.26)$$

Assume that the Gibbs-Duhem integration method keeps the difference $\Delta^{\text{LV}} \ln([\hat{f}_1 + \hat{f}_2]/p^{\text{ref}})$ unchanged. This assumption results in the following relation for the deviation of the predicted saturation pressure from the true saturation pressure at coexistence point i :

$$(p_i - p_i^{\text{eq}}) = \frac{\Delta^{\text{LV}} \langle V \rangle_{[\xi_2]_0, p_0^{\text{eq}}, \beta}}{\Delta^{\text{LV}} \langle V \rangle_{[\xi_2]_i, p_i^{\text{eq}}, \beta}} (p_0 - p_0^{\text{eq}}) \quad (5.27)$$

Thus, it makes sense to perform the integration into the direction of increasing $\Delta^{\text{LV}}V$ in order to weaken the influence of the error in the estimate of the saturation pressure corresponding to the initial coexistence point [90]. Of course, this effect is only important if the difference $\Delta^{\text{LV}}V$ changes significantly over the range of fugacity fractions.

Suppose the absolute deviation from the true initial saturation pressure equals the standard deviation of the estimate of the initial saturation pressure, σ_{p0} . It is assumed that the difference $\Delta^{\text{LV}}V$ at the true saturation pressure can be approximated by its counterpart obtained from ensemble averages at the actual estimate of the saturation pressure.

The influence of the error in the estimate of the initial saturation pressure in the binary mixture ethane/methane at $T = 193.27$ K is given in Figure 5.12. As methane is supercritical at $T = 193.27$ K, the difference $\Delta^{\text{LV}}V$ changes significantly over the integration range. Thus, the error in the estimate of the initial saturation pressure is magnified when the vapour-liquid critical point of the mixture is approached.

From a theoretical point of view, the numerical integration should have been started from the methane side as the difference Δ^{LV} increases when integrating from the methane-side to the ethane-side. However, it is questionable whether the result would have been different when the integration had started from the supercritical side, as the computation of the saturation pressure of a mixture close to the critical point involves a high statistical uncertainty or completely fails.

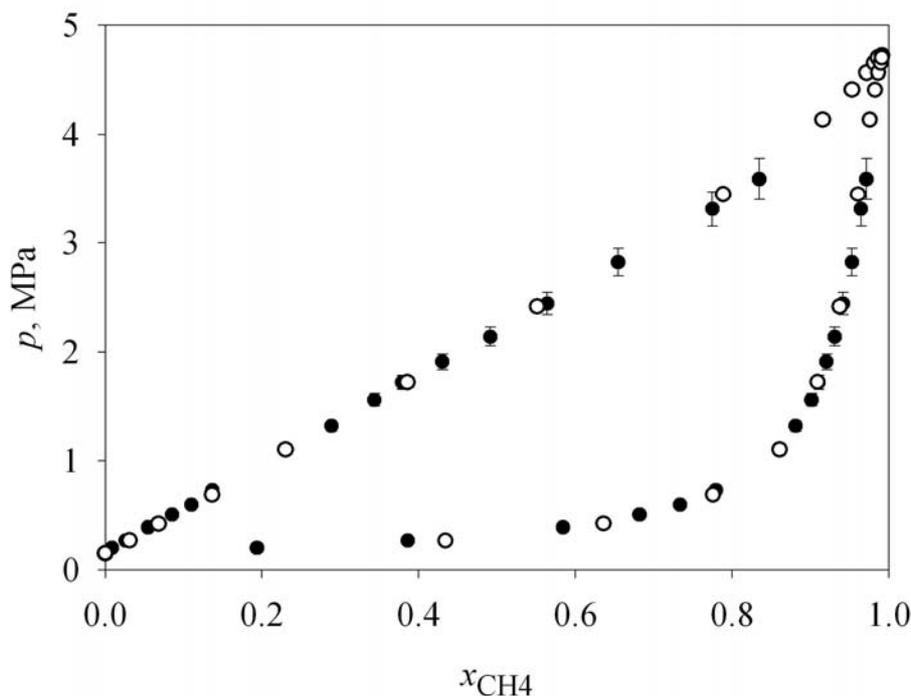


Figure 5.12: Influence of the error in the estimate of the saturation pressure corresponding to the initial coexistence point on the error in the estimated saturation pressures at successive integration points

5.6. Conclusions

This chapter started with a description of the origin and shortcomings of the well-known Lennard-Jones force field. The choice of a proper force field is crucial in molecular simulation of thermophysical properties. It is the force field that mainly determines the simulation results and their physical significance.

Existing force fields for the components of interest were selected, based on former simulation results for pure-component vapour-liquid equilibria, extent of transferability, and a right balance between physical significance and simplicity. The selection of a proper force field should be performed very carefully as there are many force fields for individual components that have limited physical significance and/or predictive power.

In this chapter, pure-component simulations were performed in order to obtain vapour-liquid equilibria of methane, ethane, carbon dioxide, difluoromethane, trifluoromethane, and DMSO. Simulations of binary mixtures were performed for the systems methane/ethane and DMSO/carbon dioxide.

The importance of combining rules for the prediction of phase equilibria of binary mixtures has been stressed. It is concluded that the force fields used in this work predict pure-component vapour-liquid equilibria in close agreement with experimental data. Thus, any deviation of the predicted mixture vapour-liquid equilibria from the true equilibria mainly

originates from a bad combining rule. Although the Berthelot combining rule generally overpredicts the interactions between unlike molecules, the simple Lorentz-Berthelot combining rules were appropriate for the simulations performed in this chapter; the results for binary mixtures were in close agreement with experimental data.

In general, the Gibbs-Duhem integration method easily traces the coexistence curve from the initial coexistence point to the coexistence point of component two. The combination of the $Np\beta$ + test molecule method and the Gibbs-Duhem integration method is capable of tracing the vapour-liquid coexistence curve from a coexistence point with a very low saturation pressure.

Like other simulation methods, the Gibbs-Duhem integration method fails close to the critical point. Furthermore, the method becomes unstable when the integrand of the Clapeyron equation changes very rapidly. Efficient sampling of compositions in diluted regions requires a large number of molecules and a bias method.

A serious disadvantage of the Gibbs-Duhem integration method is that the number of corrector iterations/simulations is not known beforehand, and that the larger part of the information generated during the corrector simulations does not contribute to the ensemble averages of the liquid and vapour densities and mole fractions at coexistence.

The major sources of error that arise from the Gibbs-Duhem integration method have been investigated. The error in the saturation pressure corresponding to the initial coexistence point seems to have the largest influence on the results.

5.7. References

1. Stone, A. J., *The theory of intermolecular forces*, International series of monographs on chemistry, Volume 32, Reprint with corrections, Oxford University Press, New York, 2000
2. Maitland, G. C., Rigby, M., Smith, E. B., and Wakeham, W. A., *Intermolecular Forces, Their Origin and Determination*, International series of monographs on chemistry, Volume 3, Reprint with corrections, Oxford University Press, New York, 1987
3. Lennard-Jones, J. E., *Proc. Roy. Soc. A*, 106 **1924** 463
4. Kihara, T., Virial coefficients and models of molecules in gases, *Rev. Mod. Phys.*, 25(4) **1953** 831
5. Allen, M. P. and Tildesley, D. J., *Computer simulation of liquids*, Clarendon Press, Oxford, 1987
6. Chen, B. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, *J. Phys. Chem. B*, 103 **1999** 5370
7. Vrabc, J. and Fischer, J., Vapor-Liquid Equilibria of Binary Mixtures Containing Methane, Ethane, and Carbon Dioxide from Molecular Simulation, *Int. J. Thermophys.*, 17(4) **1996** 889
8. Martin, M. G. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 1. United-Atom Description of *n*-Alkanes, *J. Phys. Chem. B*, 102 **1998** 2569
9. Martin, M. G. and Siepmann, J. I., Novel Configurational-Bias Monte Carlo Method for Branched Molecules. Transferable Potentials for Phase Equilibria. 2. United-Atom Description of Branched Alkanes, *J. Phys. Chem. B*, 103 **1999** 4508
10. Nath, S. K. and de Pablo, J. J., Simulation of vapour-liquid equilibria for branched alkanes, *Mol. Phys.*, 98(4) **2000** 231

-
11. Nath, S. K., Escobedo, F. A., and de Pablo, J. J., On the simulation of vapor-liquid equilibria for alkanes, *J. Chem. Phys.*, 108(23) **1998** 9905
 12. Jorgensen, W. L., Madura, J. D., and Swenson, C. J., Optimized Intermolecular Potential Functions for Liquid Hydrocarbons, *J. Am. Chem. Soc.*, 106 **1984** 6638
 13. Errington, J. R. and Panagiotopoulos, A. Z., A New Intermolecular Potential Model for the *n*-Alkane Homologous Series, *J. Phys. Chem. B*, 103 **1999** 6314
 14. Chen, B., Martin, M. G., and Siepmann, J. I., Thermodynamic Properties of the Williams, OPLS-AA, and MMFF94 All-Atom Force Fields for Normal Alkanes, *J. Phys. Chem. B*, 102 **1998** 2578
 15. Martin, M. G., Chen, B., and Siepmann, J. I., Molecular Structure and Phase Diagram of the Binary Mixture of *n*-Heptane and Supercritical Ethane: A Gibbs Ensemble Monte Carlo Study, *J. Phys. Chem. B*, 104 **2000** 2415
 16. Zhang, Z. and Duan, Z., Phase equilibria of the system methane-ethane from temperature scaling Gibbs Ensemble Monte Carlo simulation, *Geochim. Cosmochim. Acta*, 66(19) **2002** 3431
 17. Lofti, A., Vrabec, J., and Fischer, J., Vapour liquid equilibria of the Lennard-Jones fluid from the *NpT* plus test particle method, *Mol. Phys.*, 76(6) **1992** 1319
 18. Lofti, A., *PhD thesis, Ruhr-Universität Bochum, Germany*, **1993**
 19. Fischer, J., Lustig, R., Breitenfelder-Manske, H., and Lemming, W., Influence of Intermolecular Potential Parameters on Orthobaric Properties of Fluids Consisting of Spherical and Linear Molecules, *Mol. Phys.*, 52(2) **1984** 485
 20. Lago, S., Garzón, B., Calero, S., and Vega, C., Accurate Simulations of the Vapor-Liquid Equilibrium of Important Organic Solvents and Other Diatomics, *J. Phys. Chem. B*, 101 **1997** 6763
 21. Vrabec, J., Stoll, J., and Hasse, H., A Set of Molecular Models for Symmetric Quadrupolar Fluids, *J. Phys. Chem. B*, 105 **2001** 12126
 22. Harris, J. G. and Yung, K. H., Carbon Dioxide's Liquid-Vapor Coexistence Curve and Critical Properties As Predicted by a Simple Molecular Model, *J. Phys. Chem.*, 99 **1995** 12021
 23. Potoff, J. J. and Siepmann, J. I., Vapor-Liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen, *AIChE J.*, 47(7) **2001** 1676
 24. Potoff, J. J., Errington, J. R., and Panagiotopoulos, A. Z., Molecular Simulation of Phase Equilibria for Mixtures of Polar and Non-Polar Components, *Mol. Phys.*, 97(10) **1999** 1073
 25. Murthy, C. S., O'Shea, S. F., and McDonald, I. R., Electrostatic Interactions in Molecular Crystals. Lattice Dynamics of Solid Nitrogen and Carbon Dioxide, *Mol. Phys.*, 50(3) **1983** 531
 26. Murthy, C. S. and Singer, K., *Mol. Phys.*, 44 **1981** 135
 27. Möller, D. and Fischer, J., Determination of an effective intermolecular potential for carbon dioxide using vapour-liquid phase equilibria from *NpT* + test particle simulations, *Fluid Phase Equil.*, 100 **1994** 35
 28. Garzón, B., Lago, S., Vega, C., De Miguel, E., and Rull, L. F., Computer Simulation of Vapor-Liquid Equilibria of Linear Quadrupolar Fluids. Departures from the Principle of Corresponding States, *J. Chem. Phys.*, 101(5) **1994** 4166
 29. Kalugin, O. N., Volobuev, M. N., Ishchenko, A. V., and Adya, A. K., Structure and dynamics of Na⁺ and Cl⁻ solvation shells in liquid DMSO: molecular dynamics simulations, *J. Mol. Liquids*, 91 **2001** 135
 30. Cordeiro, J. M. M., Study of Liquid Dimethyl Sulfoxide by Computer Simulation, *Mol. Eng.*, 8 **1999** 303

-
31. Vishnyakov, A., Lyubartsev, A. P., and Laaksonen, A., Molecular Dynamics Simulations of Dimethyl Sulfoxide and Dimethyl Sulfoxide-Water Mixture, *J. Phys. Chem. A*, 105 **2001** 1702
 32. Geerke, D. P., Oostenbrink, C., Van der Vegt, N. F. A., and van Gunsteren, W. F., An Effective Force Field For Molecular Dynamics Simulations of Dimethyl Sulfoxide and Dimethyl Sulfoxide-Water Mixtures, *J. Phys. Chem. B*, 108 **2004** 1436
 33. Strader, M. L. and Feller, S. E., A Flexible All-Atom Model of Dimethyl Sulfoxide for Molecular Dynamics Simulations, *J. Phys. Chem. A*, 106 **2002** 1074
 34. Rao, B. G. and Singh, U. C., A Free Energy Perturbation Study of Solvation in Methanol and Dimethyl Sulfoxide, *J. Am. Chem. Soc.*, 112 **1990** 3803
 35. Gao, G. T., Wang, W. C., and Zeng, X. C., Vapor-liquid equilibria for pure HCFC/HFC substances by Gibbs ensemble simulation of Stockmayer potential molecules, *Fluid Phase Equil.*, 137 **1997** 87
 36. Higashi, S. and Takada, A., Molecular dynamics study of liquid CH₂F₂ (HFC-32), *Mol. Phys.*, 92(4) **1997** 641
 37. Lísal, M., Smith, W. R., and Aim, K., Direct molecular-level Monte Carlo simulation of Joule-Thomson processes, *Mol. Phys.*, 101(18) **2003** 2875
 38. Potter, S. C., Tildesley, D. J., Burgess, A. N., and Rogers, S. C., A transferable potential model for the liquid-vapour equilibria of fluoromethanes, *Mol. Phys.*, 92(5) **1997** 825
 39. Jedlovsky, P. and Mezei, M., Computer simulation study of liquid CH₂F₂ with a new effective pair potential model, *J. Chem. Phys.*, 110(6) **1999** 2991
 40. Fermeglia, M., Ferrone, M., and Priol, S., Development of an all-atoms force field from ab initio calculations for alternative refrigerants, *Fluid Phase Equil.*, 210 **2003** 105
 41. Lísal, M. and Vacek, V., Effective potentials for liquid simulation of the alternative refrigerants HFC-32: CH₂F₂ and HFC-23: CHF₃, *Fluid Phase Equil.*, 118 **1996** 61
 42. Palmer, B. J. and Anchell, J. L., Molecular Mechanics Parameters for Fluorine-Substituted Methanes from *ab initio* Quantum Calculations, *J. Phys. Chem.*, 99 **1995** 12239
 43. Lide, D. R., *CRC Handbook of chemistry and physics*, CRC Press, 73rd edition, Boca Raton, Florida, 1992
 44. Siepmann, J. I., 2003, Personal communication
 45. Song, W., Patel, N., and Maroncelli, M., A 2-Site Model for Simulating Supercritical Fluoroform, *J. Phys. Chem. B*, 106 **2002** 8783
 46. Hloucha, M. and Deiters, U. K., Monte Carlo study of the thermodynamic properties and the static dielectric constant of liquid trifluoromethane, *Fluid Phase Equil.*, 149 **1998** 41
 47. Gough, C. A., DeBolt, S. E., and Kollman, P. A., Derivation of Fluorine and Hydrogen Atom Parameters Using Liquid Simulations, *J. Comput. Chem.*, 13(8) **1992** 963
 48. Böhm, H. J., Ahlrichs, R., Scharf, P., and Schiffer, H., Intermolecular potentials for CH₄, CH₃F, CHF₃, CH₃Cl, CH₂Cl₂, CH₃CN and CO₂, *J. Chem. Phys.*, 81(3) **1984** 1389
 49. Lorentz, H. A., *Ann. Phys.*, 12 **1881** 127
 50. Berthelot, D., *Compt. Rend.*, 126 **1898** 1857
 51. Berthelot, D. C., *R. Hebd. Séanc. Acad. Sci. , Paris*, 126 **1898** 1703

-
52. Hudson, G. H. and McCoubrey, J. C., Intermolecular forces between unlike molecules, *Trans. Faraday Soc.*, 56 **1960** 761
 53. Vrabec, J. and Fischer, J., Vapor-Liquid Equilibria of the Ternary Mixture CH₄+C₂H₆+CO₂ from Molecular Simulation, *AIChE J.*, 43(1) **1997** 212
 54. Song, W., Rossky, P. J., and Maroncelli, M., Modeling alkane + perfluoroalkane interactions using all-atom potentials: Failure of the usual combining rules, *J. Chem. Phys.*, 119(17) **2003** 9145
 55. Möller, D., Óprzynski, J., Müller, A., and Fischer, J., Prediction of thermodynamic properties of fluid mixtures by molecular-dynamics simulations - methane/ethane, *Mol. Phys.*, 75(2) **1992** 363
 56. Vrabec, J. and Fischer, J., Vapour liquid equilibria of mixtures from the *NpT* plus test particle method, *Mol. Phys.*, 85(4) **1995** 781
 57. Liu, A. and Beck, Th. L., Vapor-Liquid Equilibria of Binary and Ternary Mixtures Containing Methane, Ethane, and Carbon Dioxide from Gibbs Ensemble Simulations, *J. Phys. Chem. B*, 102 **1998** 7627
 58. Kronome, G., Szalai, I., Wendland, M., and Fischer, J., Extension of the *NpT* + test particle method for the calculation of phase equilibria of nitrogen + ethane, *J. Mol. Liquids*, 85 **2000** 237
 59. Diaz Peña, M., Pando, C., and Renuncio, J. A. R., Combination rules for intermolecular potential parameters. I. Rules based on approximations for the long-range dispersion energy, *J. Chem. Phys.*, 76(1) **1982** 325
 60. Diaz Peña, M., Pando, C., and Renuncio, J. A. R., Combination rules for intermolecular potential parameters. II. Rules based on approximations for the long-range dispersion energy and an atomic distortion model for the repulsive interactions, *J. Chem. Phys.*, 76(1) **1982** 333
 61. London, F., *Z. Physik*, 63 **1930** 245
 62. Khalaf Al-Matar, A. and Rockstraw, D. A., A Generating Equation for Mixing Rules and Two New Mixing Rules for Interatomic Potential Energy Parameters, *J. Comput. Chem.*, 25 **2004** 660
 63. Kong, C. L. and Chakrabarty, M. R., Combining Rules for Intermolecular Potential Parameters. III. Application to the Exp 6 Potential, *J. Phys. Chem.*, 77(22) **1973** 2668
 64. Flyvbjerg, H. and Petersen, H. G., Error estimates on averages of correlated data, *J. Chem. Phys.*, 91(1) **1989** 461
 65. Escobedo, F. A. and de Pablo, J. J., Pseudo-ensemble simulations and Gibbs-Duhem integrations for polymers, *J. Chem. Phys.*, 106(7) **1997** 2911
 66. NIST Chemistry WebBook available on the web, <http://webbook.nist.gov/chemistry/>, 2004
 67. Van 't Hof, A., de Leeuw, S. W., Hall, C. K., and Peters, C. J., Molecular simulation of binary vapour-liquid equilibria with components differing largely in volatility, *Mol. Phys.*, 102(3) **2004** 301
 68. Douslin, D. R. and Harrison, R. H., Pressure, volume, temperature relations of ethane, *J. Chem. Thermodyn.*, 5(4) **1973** 491
 69. Campbell, A. N., The density and vapour pressure of dimethylsulfoxide at various temperatures and the (hypothetical) critical density, *Can. J. Chem.*, 57 **1979** 705
 70. Malbrunot, P. F., Meunier, P. A., and Scatena, G. M., Pressure-Volume-Temperature Behavior of Difluoromethane, *J. Chem. Eng. Data*, 13(1) **1968** 16
 71. Defibaugh, D. R., Morrison, G., and Weber, L. A., Thermodynamic Properties of Difluoromethane, *J. Chem. Eng. Data*, 39(2) **1994** 333

-
72. Hou, Y. C. and Martin, J. J., Physical and Thermodynamic Properties of Trifluoromethane, *AIChE J.*, 5(1) **1959** 125
 73. Rubio, R. G., Zollweg, J. A., Palanco, J. M. G., Calado, J. C. G., Miller, J., and Streett, W. B., Thermodynamic Properties of Simple Molecular Fluids: Tetrafluoromethane and Trifluoromethane, *J. Chem. Eng. Data*, 36 **1991** 171
 74. Mooijer-Van den Heuvel, M. M., Nasrifar, Kh., and Peters, C. J., *Unpublished data*, **2002**
 75. Miller, R. C., Kidnay, A. J., and Hiza, M. J., Liquid + vapor equilibria in methane + ethene and in methane + ethane from 150.00 to 190.00 K, *J. Chem. Thermodyn.*, 9 **1977** 167
 76. Davalos, J., Anderson, W. R., Phelps, R. E., and Kidnay, A. J., Liquid-Vapor Equilibria at 250.00 K for Systems Containing Methane, Ethane, and Carbon Dioxide, *J. Chem. Eng. Data*, 21(1) **1976** 81
 77. Davalos, J., Anderson, W. R., Phelps, R. E., and Kidnay, A. J., Correction, *J. Chem. Eng. Data*, 21(4) **1976** 510
 78. Wichterle, I. and Kobayashi, R., Vapor-Liquid Equilibrium of Methane-Ethane System at Low Temperatures and High Pressures, *J. Chem. Eng. Data*, 17(1) **1972** 9
 79. Wei, M. S. W., Brown, T. S., Kidnay, A. J., and Sloan, D. E., Vapor + liquid equilibria for the ternary system methane + ethane + carbon dioxide at 230 K and its constituent binaries at temperatures from 207 to 270 K, *J. Chem. Eng. Data*, 40 **1995** 726
 80. Gupta, M. K., Gardner, G. C., Hegarty, M. J., and Kidnay, A. J., Liquid-Vapor Equilibria for the N₂ + CH₄ + C₂H₆ System from 260 to 280 K, *J. Chem. Eng. Data*, 25 **1980** 313
 81. Price, A. R., PhD Thesis, *Rice Inst. Houston, Texas* **1957**
 82. de Pablo, J. J. and Prausnitz, J. M., Phase equilibria for fluid mixtures from Monte-Carlo simulation, *Fluid Phase Equil.*, 53 **1989** 177
 83. Lim, G. B., Lee, S. Y., Koo, K. K., Park, B. S., and Kim, H. S., *Proceedings of the 5th meeting on supercritical fluids, Tome 1, Perrut, M.; Subra, P., Eds., Nice, 1998*
 84. Reverchon, E., Della Porta, G., De Rosa, I., Subra, P., and Letourneur, D., *5th Conference on Supercritical Fluids and their Applications*, Garda, 1999
 85. Vega Gonzalez, A., Tufeu, R., and Subra, P., High-Pressure Vapor-Liquid Equilibrium for the Binary Systems Carbon Dioxide + Dimethyl Sulfoxide and Carbon Dioxide + Dichloromethane, *J. Chem. Eng. Data*, 47 **2002** 492
 86. Kordikowski, A., Schenk, A. P., van Nielen, R. M., and Peters, C. J., Volume Expansions and Vapor-Liquid Equilibria of Binary Mixtures of a Variety of Polar Solvents and Certain Near-Critical Solvents, *J. Supercrit. Fluids*, 8 **1995** 205
 87. Rajasingam, R., Lioe, L., Pham, Q. T., and Lucien, F. P., Solubility of carbon dioxide in dimethylsulfoxide and *N*-methyl-2-pyrrolidone at elevated pressure, *J. Supercrit. Fluids*, *In Press*, **2004**
 88. Kofke, D. A., Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line, *J. Chem. Phys.*, 98(5) **1993** 4149
 89. Mehta, M. and Kofke, D. A., Coexistence Diagrams of Mixtures by Molecular Simulation, *Chem. Eng. Sci.*, 49(16) **1994** 2633
 90. Kofke, D. A., Semigrand Canonical Monte Carlo Simulation; Integration Along Coexistence Lines. In: Ferguson, D. M., Siepmann, J. I., and Truhlar, D. G., editors, *Advances in Chemical Physics, Volume 105, Monte Carlo Methods in Chemistry*, Wiley, 1999

Chapter 6. Advanced Gibbs-Duhem integration

In this chapter, an advanced version of the Gibbs-Duhem integration method is discussed, which is more efficient than the conventional method. The improved method is a combination of Gibbs-Duhem integration and the multiple-histogram reweighting technique. The multiple-histogram reweighting technique is used to increase the amount of information extracted from simulation results, as will be described in Section 6.2. In Section 6.3, the relations that are needed to apply multiple-histogram reweighting in the semigrand ensemble are derived. The details of the implementation of multiple-histogram reweighting in the Gibbs-Duhem integration method are described in Section 6.4. The advanced Gibbs-Duhem integration method is tested on the vapour-liquid equilibria of a number of interesting binary mixtures with different combining rules for interactions between similar and dissimilar molecules. The details of the simulations, along with the actual simulation results, are presented in Section 6.6. A brief error analysis of the advanced Gibbs-Duhem integration method is given in Section 6.7. Finally, the conclusions concerning this chapter are summarized in Section 6.8.

6.1. Introduction

From Chapter 5, it became clear that the conventional Gibbs-Duhem integration method is far from efficient. The large number of samples generated during the corrector simulations is merely used to predict the saturation pressure. The liquid and vapour densities and mole fractions at coexistence are computed during a production simulation at the predicted saturation pressure. Instead of this approach, it would be worthwhile to extract as much information from the generated samples as possible.

In a conventional simulation, the samples that are collected at a particular thermodynamic state point are only used to compute ensemble averages at those state points. However, the samples can also be used to predict ensemble averages at nearby state points by implementing reweighting or single-histogram reweighting. Furthermore, samples obtained at *different* thermodynamic state points can be used to improve the estimate of an ensemble average over a whole range of thermodynamic state points.

The corrector simulations in the conventional Gibbs-Duhem integration method were intended to improve the estimate of the saturation pressure. The number of required corrector simulations was not known beforehand. On average, the number of corrector iterations was about ten, but the number can be much larger. As will be pointed out in Section 6.4, the use of the histogram-reweighting technique can be conveniently used to replace the prediction, correction, and production simulations by a fixed number of simulations. Moreover, current simulations may improve the estimates of formerly computed coexistence points.

In the next section, it is pointed out how to obtain ensemble averages at different state points from a single simulation by reweighting or single-histogram reweighting. The extension to multiple-histogram reweighting is given in Section 6.2.2. The techniques are most easily described in the canonical ensemble; an extension to the semigrand-canonical ensemble is given in Section 6.3.

6.2. Histogram-reweighting technique

6.2.1. Single-histogram reweighting

The configurations generated during a Monte Carlo simulation contain a large amount of information. In a conventional simulation, only a small fraction of these data is used. It is a waste of data (and time!) not to try to extract as much information as possible from the configurations. A simulation at a single thermodynamic state point can provide much more information than only a few ensemble averages.

Reweighting methods and histogram-reweighting methods are examples of simulation analysis tools that increase the amount of information extracted from the simulations. Reweighting provides an estimate of the probability distribution of the ensemble of interest at thermodynamic state variables close to the ones the simulation is performed at [1-6]. From the reweighted probability distribution, one can compute ensemble averages at nearby state variables. Obviously, reweighting of data obtained from a single simulation is an extrapolation method. Suppose the instantaneous property \mathcal{O} is sampled during the course of a simulation in the canonical ensemble performed at reciprocal temperature β_0 . A Monte Carlo (MC) estimate of the canonical probability of observing a configuration i with corresponding energy \mathcal{U}_i , based on n samples, is given by:

$$\pi^{\text{est}}(i; \beta_0) = \frac{\exp(-\beta_0 \mathcal{U}_i)}{\sum_{j=1}^n \exp(-\beta_0 \mathcal{U}_j)} = \frac{\exp(-\beta_0 \mathcal{U}_i)}{Q_{NV\beta_0}^{\text{est}}} \quad (6.1)$$

where $Q_{NV\beta_0}^{\text{est}}$ is an estimate of the canonical partition function at temperature β_0 , based on n samples. This MC estimate of the probability can be reweighted to a different temperature, β , as follows:

$$\begin{aligned} \pi^{\text{est}}(i; \beta) &= \frac{\exp\{-\beta \mathcal{U}_i\}}{Q_{NV\beta}^{\text{est}}} \frac{Q_{NV\beta_0}^{\text{est}}}{Q_{NV\beta_0}^{\text{est}}} \\ &= \frac{\exp\{-(\beta - \beta_0) \mathcal{U}_i\} \exp\{-\beta_0 \mathcal{U}_i\} \sum_{j=1}^n \exp\{-\beta_0 \mathcal{U}_j\}}{Q_{NV\beta_0}^{\text{est}} \sum_{j=1}^n \exp\{-(\beta - \beta_0) \mathcal{U}_j\} \exp\{-\beta_0 \mathcal{U}_j\}} \\ &= \pi^{\text{est}}(i; \beta_0) \frac{\exp\{-(\beta - \beta_0) \mathcal{U}_i\}}{\langle \exp\{-(\beta - \beta_0) \mathcal{U}\} \rangle_{\beta_0, n}} \end{aligned} \quad (6.2)$$

where $\langle \rangle_{\beta_0, n}$ is a canonical ensemble average computed from n samples at temperature β_0 . The ensemble average of \mathcal{O} at a different temperature β is given by:

$$\langle \mathcal{O} \rangle_{\beta, n} = \frac{\langle \mathcal{O}(\mathcal{U}) \exp\{-(\beta - \beta_0) \mathcal{U}\} \rangle_{\beta_0, n}}{\langle \exp\{-(\beta - \beta_0) \mathcal{U}\} \rangle_{\beta_0, n}} \quad (6.3)$$

It is often convenient to sum over energy levels instead of over configurations (individual samples). As Monte Carlo simulations provide *continuous* energies, the energies need to be discretized. Energy level j includes all energies \mathcal{U} within the following interval:

$$\mathcal{U}_j - \frac{1}{2}\Delta\mathcal{U} \leq \mathcal{U} < \mathcal{U}_j + \frac{1}{2}\Delta\mathcal{U} \quad (6.4)$$

where $\Delta\mathcal{U}$ is the bin width of the energy levels. From now on, the number of samples is dropped. The canonical probability to observe the system with energy \mathcal{U} at temperature β_0 , written in terms of energy levels, is given by:

$$\pi(\mathcal{U}; \beta_0) = \frac{\Omega(N, V, \mathcal{U}) \exp\{-\beta_0 \mathcal{U}\}}{\sum_{\mathcal{U}'} \Omega(N, V, \mathcal{U}') \exp\{-\beta_0 \mathcal{U}'\}} = \frac{\Omega(N, V, \mathcal{U}) \exp\{-\beta_0 \mathcal{U}\}}{\mathcal{Q}_{NV\beta_0}} \quad (6.5)$$

where $\Omega(N, V, \mathcal{U})$ is the *degeneracy* or microcanonical partition function (often called density of states). The degeneracy is the number of configurations with energy \mathcal{U} at constant number of molecules and at constant volume.

All samples of the energy that are obtained during the course of a simulation, are collected into a histogram $\mathcal{H}_{\beta_0}(\mathcal{U})$. The subscript β_0 indicates the temperature at which the histogram is collected. The energy levels, as given in Eq. (6.4), correspond to the histogram bins. As the configurations are generated in accordance with their Boltzmann weights, the histogram can serve as an estimate of the true canonical probability distribution (6.5) at temperature β_0 :

$$\pi^{\text{est}}(\mathcal{U}; \beta_0) = \frac{\mathcal{H}_{\beta_0}(\mathcal{U})}{\mathcal{N}_{\beta_0}} \quad (6.6)$$

where \mathcal{N}_{β_0} is the sum of the entries of histogram $\mathcal{H}_{\beta_0}(\mathcal{U})$. Combination of Eqs. (6.5) and (6.6) enables one to estimate the degeneracy:

$$\begin{aligned} \Omega^{\text{est}}(N, V, \mathcal{U}) &= \frac{\mathcal{H}_{\beta_0}(\mathcal{U})}{\mathcal{N}_{\beta_0}} \exp\{+\beta_0 \mathcal{U}\} \mathcal{Q}_{NV\beta_0} \\ &= \frac{\mathcal{H}_{\beta_0}(\mathcal{U})}{\mathcal{N}_{\beta_0}} \exp\{+\beta_0(\mathcal{U} - A_{NV\beta_0})\} \end{aligned} \quad (6.7)$$

where $A_{NV\beta_0}$ is the Helmholtz energy at temperature β_0 . From now on, the subscript β_0 is replaced by a subscript 0.

Insertion of Eq. (6.7) into Eq. (6.5) shows how the canonical probability distribution at any reciprocal temperature β can be obtained by reweighting the histogram collected at reciprocal temperature β_0 :

$$\pi^{\text{est}}(\mathcal{U}; \beta) = \frac{\frac{\mathcal{H}_0(\mathcal{U})}{\mathcal{N}_0} \exp\{-(\beta - \beta_0)\mathcal{U}\}}{\sum_{\mathcal{U}'} \frac{\mathcal{H}_0(\mathcal{U}')}{\mathcal{N}_0} \exp\{-(\beta - \beta_0)\mathcal{U}'\}} \quad (6.8)$$

The Helmholtz energies in the numerator and denominator of probability distribution (6.8) cancel one another. The ensemble average of the property \mathcal{O} at temperature β is now computed from:

$$\langle \mathcal{O} \rangle_\beta = \frac{\sum_{\mathcal{U}} \mathcal{O}(\mathcal{U}) \mathcal{H}_0(\mathcal{U}) \exp\{-(\beta - \beta_0)\mathcal{U}\}}{\sum_{\mathcal{U}} \mathcal{H}_0(\mathcal{U}) \exp\{-(\beta - \beta_0)\mathcal{U}\}} \quad (6.9)$$

In order to show the limitations of the single-histogram reweighting method, a schematic picture of the density of states and several histograms collected at different temperatures is presented in Figure 6.1.

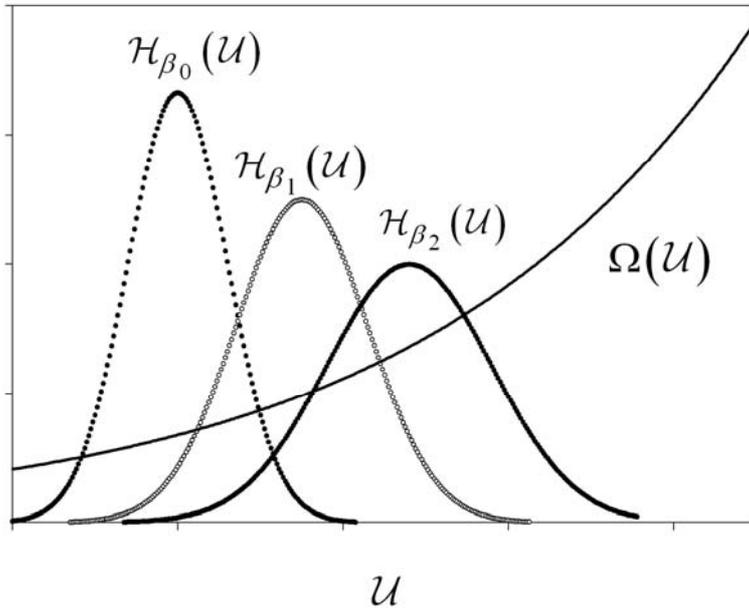


Figure 6.1: Schematic picture of the density of states and several histograms collected at different temperatures, as a function of the energy \mathcal{U} .

Every histogram obtained from a finite simulation is subject to statistical errors. Furthermore, the range of energies sampled during a finite simulation at a single temperature β_i is relatively small. Thus, a single simulation provides a reliable estimate of the degeneracy $\Omega^{\text{est}}(\mathcal{U})$ only over a limited range of energies; only there where the histograms in Figure 6.1

peak. This means that the extrapolative power of Eqs. (6.8) and (6.9) is limited to a relatively narrow range of temperatures around β_i [7-11].

If the reweighting temperature, β , differs too much from β_i , the peak of the reweighted histogram will shift to the tails of the measured histogram where the statistical uncertainty is high. Nevertheless, there are methods that can be used to obtain a degeneracy that is valid over a broad range of energies, as will be described in the next section.

6.2.2. Multiple-histogram reweighting

The degeneracy is independent of state variables and can in principle be estimated from every histogram [See Eq. (6.7)]. From Figure 6.1, it is clear that a simulation at a single temperature only estimates the degeneracy accurately where the histogram peaks. It is also clear from Figure 6.1 that the histograms collected at nearby temperatures have a region of overlap. Thus, a series of histograms collected at nearby temperatures cover a large part of the energies available to the system. A clever combination of those histograms may be used to estimate the degeneracy over a broad range of energies. Such a combination can be performed with the multiple-histogram reweighting (MHR) method of Ferrenberg and Swendsen [1, 2, 10, 12]. In this method, every individual histogram bin is assigned a normalized weight $r_i(\mathcal{U})$. The degeneracy, computed from a linear combination of the histogram estimates collected at R state points, is given by:

$$\Omega^{\text{est}}(N, V, \mathcal{U}) = \sum_{i=1}^R r_i(\mathcal{U}) \frac{\mathcal{H}_i(\mathcal{U})}{\mathcal{N}_i} \exp\left(+\beta_i(\mathcal{U} - A_{NV\beta_i})\right) \quad (6.10)$$

The weights for each energy level are found by minimizing the variance in the estimated degeneracy [2, 10, 12]. The resulting equation for the estimated degeneracy is given by:

$$\Omega^{\text{est}}(N, V, \mathcal{U}) = \frac{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(\mathcal{U})}{\sum_{j=1}^R \mathcal{N}_j g_j^{-1} \exp\left\{-\beta_j(\mathcal{U} - A_{NV\beta_j})\right\}} \quad (6.11)$$

where the parameters g_i depend on the correlation time of successive configurations collected at state point i [2, 10, 12]. The Helmholtz energies in Eq. (6.11) are estimated from:

$$\exp\left\{-\beta_i A_{NV\beta_i}\right\} = \sum_{\mathcal{U}} \Omega^{\text{est}}(N, V, \mathcal{U}) \exp(-\beta_i \mathcal{U}) \quad (6.12)$$

Equation (6.11) and the Helmholtz energies (6.12) at the R state points form a dependent set of equations which can be solved self-consistently with a Newton-Raphson method [12, 13]. The optimization of the Helmholtz energies is only possible when histograms obtained from different simulations overlap sufficiently. As Eqs. (6.11) and (6.12) determine the Helmholtz energies only within an additive constant, it is convenient to set one of the Helmholtz energies to zero and compute the other values with respect to this reference state. In case the histograms are extended to a thermodynamic state point with known Helmholtz energy, all absolute Helmholtz energies at other state points can be estimated.

Finally, the probability to observe the system with energy, \mathcal{U} , at reciprocal temperature, β , valid over a broad range of temperatures and energies is given by:

$$\pi^{\text{est}}(\mathcal{U};\beta) = \frac{\Omega^{\text{est}}(N,V,\mathcal{U})\exp(-\beta\mathcal{U})}{\sum_{\mathcal{U}'} \Omega^{\text{est}}(N,V,\mathcal{U}')\exp(-\beta\mathcal{U}')} \quad (6.13)$$

6.3. Multiple-histogram reweighting in the semigrand-canonical ensemble

As the simulations in this work are performed in the semigrand-canonical ensemble, the multiple-histogram equations need to be extended to this ensemble. Fortunately, such an extension is straightforward. In the binary semigrand-canonical ensemble, the total energy, $\mathcal{U}^{\text{tot}} \equiv \mathcal{U}^{\text{ext}} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}} + \text{intra}$, volume, V , and number of molecules of component two, N_2 , fluctuate while the reciprocal temperature, β , pressure, p , fugacity fraction, ξ_2 , and total number of molecules, N , are fixed. The semigrand-canonical semi-classical partition function was given in Chapter 2, Eq. (2.31):

$$\mathcal{Y}'_{Np\beta\xi_2}^* = \int dV \frac{[Vq_1(\beta)]^N \exp\{-\beta pV\}}{V(N-1)!} \sum_{\text{idn}} \left(\frac{C_1(\beta)\xi_2}{C_2(\beta)\xi_1} \right)^{N_2} \cdot \int \prod_{j=1}^2 \prod_{i=1}^{N_j} J_{ij}(\mathbf{q}_i^{2n_j-2}) d\mathbf{q}_i^{2n_j-2} d\mathbf{s}^{3N} \exp\left\{-\beta\mathcal{U}^{\text{tot}}(\mathbf{s}^{3N}, \mathbf{q}^{2(N_1n_1+N_2n_2-N)})\right\} \quad (6.14)$$

where the asterisk indicates that the partition function does not count volume states redundantly [14-16]. The prime in equation (6.14) indicates that the thermodynamic potential belonging to this semi-classical partition function is the full semigrand energy.

In order to simplify the notation, the factor $J_{ij}(\mathbf{q}_i^{2n_j-2})d\mathbf{q}_i^{2n_j-2}d\mathbf{s}^{3N}$ will be indicated by $d\mathbf{q}'_{ij}$ in what follows. So as to be able to use MHR in the semigrand-canonical ensemble, a discrete partition function like $Q_{NV\beta_0}$ in Eq. (6.5) must be derived. This can be performed by choosing a total energy $\mathcal{U}_0^{\text{tot}}$, volume V_0 , and number of molecules of component two N_2^0 , and by introducing Dirac δ -functions and a Kronecker delta in Eq. (6.14):

$$\mathcal{Y}'_{Np\beta\xi_2}^* = \sum_{N_2^0} \int dV_0 \int d\mathcal{U}_0^{\text{tot}} \int dV \delta(V-V_0) \frac{[Vq_1(\beta)]^N \exp\{-\beta pV\}}{V(N-1)!} \cdot \sum_{\text{idn}} \delta_{N_2, N_2^0} \left(\frac{C_1(\beta)\xi_2}{C_2(\beta)\xi_1} \right)^{N_2} \int \prod_{j=1}^2 \prod_{i=1}^{N_j} d\mathbf{q}'_{ij} \delta(\mathcal{U}^{\text{tot}} - \mathcal{U}_0^{\text{tot}}) \exp\{-\beta\mathcal{U}^{\text{tot}}\} \quad (6.15)$$

Now, the integrand in Eq. (6.15) is written as a product of degeneracy and pseudo-Boltzmann factor.

As was mentioned already, the degeneracy is independent of state variables. Keeping all parts that depend on temperature, pressure and fugacity fraction out of the degeneracy results in:

$$\mathcal{Y}'_{Np\beta\xi_2} = \sum_{N_2^0} \int dV_0 \int d\mathcal{U}_0^{\text{tot}} \Omega^* \left(N_2^0, V_0, \mathcal{U}_0^{\text{tot}} \right) \cdot \exp \left\{ -\beta \mathcal{U}_0^{\text{tot}} - \beta p V_0 + N \ln \left(q_1(\beta) V_0 \int d\mathbf{q}'_1 \right) + N_2^0 \ln \left(\frac{\xi_2 C_1(\beta) \int d\mathbf{q}'_2}{\xi_1 C_2(\beta) \int d\mathbf{q}'_1} \right) \right\} \quad (6.16)$$

The degeneracy Ω^* in Eq. (6.16) is given by:

$$\Omega^* \left(N_2^0, V_0, \mathcal{U}_0^{\text{tot}} \right) = \int dV \frac{\delta(V - V_0)}{V(N-1)!} \sum_{\text{idem}} \delta_{N_2, N_2^0} \frac{\prod_{j=1}^2 \prod_{i=1}^{N_j} \int d\mathbf{q}'_{ij} \delta(\mathcal{U}^{\text{tot}} - \mathcal{U}_0^{\text{tot}})}{\prod_{j=1}^2 \prod_{i=1}^{N_j} \int d\mathbf{q}'_{ij}} \quad (6.17)$$

The asterisk in Eq. (6.17) again indicates that the degeneracy belongs to an ensemble that correctly counts volume states. As the degeneracy is estimated from histograms, the discretized version of partition function (6.16) is needed:

$$\mathcal{Y}'_{Np\beta\xi_2} = \sum_{N_2^0} \sum_{V_0} \sum_{\mathcal{U}_0^{\text{tot}}} \Omega^* \left(N_2^0, V_0, \mathcal{U}_0^{\text{tot}} \right) \cdot \exp \left\{ -\beta \mathcal{U}_0^{\text{tot}} - \beta p V_0 + N \ln \left(q_1(\beta) V_0 \int d\mathbf{q}'_1 \right) + N_2^0 \ln \left(\frac{\xi_2 C_1(\beta) \int d\mathbf{q}'_2}{\xi_1 C_2(\beta) \int d\mathbf{q}'_1} \right) \right\} \quad (6.18)$$

where the degeneracy Ω^* is now given by:

$$\Omega^* \left(N_2^0, V_0, \mathcal{U}_0^{\text{tot}} \right) = \frac{\Delta V_0 \Delta \mathcal{U}_0^{\text{tot}}}{(N-1)!} \int dV \frac{\delta(V - V_0)}{V} \sum_{\text{idem}} \delta_{N_2, N_2^0} \frac{\prod_{j=1}^2 \prod_{i=1}^{N_j} \int d\mathbf{q}'_{ij} \delta(\mathcal{U}^{\text{tot}} - \mathcal{U}_0^{\text{tot}})}{\prod_{j=1}^2 \prod_{i=1}^{N_j} \int d\mathbf{q}'_{ij}} \quad (6.19)$$

The probability of observing the system with energy, \mathcal{U}^{tot} , volume, V , and number of molecules of component two, N_2 , at thermodynamic state point $\{\xi_2, p, \beta\}$ is:

$$\pi \left(N_2, V, \mathcal{U}^{\text{tot}}; \xi_2, p, \beta \right) = \frac{\Omega^* \left(N_2, V, \mathcal{U}^{\text{tot}} \right) \exp \left\{ -\beta \mathcal{U}^{\text{tot}} - \beta p V + N \ln \left(q_1(\beta) V \int d\mathbf{q}'_1 \right) + N_2 \ln \left(\frac{\xi_2 C_1(\beta) \int d\mathbf{q}'_2}{\xi_1 C_2(\beta) \int d\mathbf{q}'_1} \right) \right\}}{\mathcal{Y}'_{Np\beta\xi_2}} \quad (6.20)$$

Over the course of a simulation performed at thermodynamic state point i , a histogram containing all energy-volume-composition combinations the system encounters is collected:

$$\mathcal{H}_i(N_2, V, \mathcal{U}^{\text{tot}}) \equiv \mathcal{H}(N_2, V, \mathcal{U}^{\text{tot}}; [\xi_2]_i, p_i, \beta_i) \quad (6.21)$$

Simulations performed at R different thermodynamic state points can be combined with the MHR method in order to obtain an improved estimate of the degeneracy $\Omega^*(N_2, V, \mathcal{U}^{\text{tot}})$ that is valid over a large range of temperatures, pressures and fugacity fractions. The estimate for the degeneracy is given by:

$$\Omega^{*, \text{est}}(N_2, V, \mathcal{U}^{\text{tot}}) = \frac{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(N_2, V, \mathcal{U}^{\text{tot}})}{\sum_{j=1}^R \frac{\mathcal{N}_j}{g_j} \exp \left\{ \begin{aligned} & -\beta_j \mathcal{U}^{\text{tot}} - \beta_j p_j V + N \ln \left\{ q_1 V \int d\mathbf{q}'_1 \right\} + \\ & N_2 \ln \left(\frac{C_1 \int d\mathbf{q}'_2 [\xi_2]_j}{C_2 \int d\mathbf{q}'_1 [\xi_1]_j} \right) + \beta_j [\mu_1]_j N \end{aligned} \right\}} \quad (6.22)$$

where \mathcal{N}_j is the sum of the entries in histogram $\mathcal{H}_j(N_2, V, \mathcal{U}^{\text{tot}})$ and $[\mu_1]_j N$ is the total semigrand energy at state point j . As the kinetic contribution to the pseudo-Boltzmann factor at a particular state point is a constant, it can be combined with the semigrand energy. In the same way, the semigrand energy is replaced by the configurational semigrand energy [See Chapter 2, Eq. (2.33)]:

$$N \ln(\hat{f}_1 / p^{\text{ref}}) = \beta \mu_1 N + N \ln \left\{ \frac{q_1(\beta) C_1(\beta)}{\beta p^{\text{ref}}} \right\} \quad (6.23)$$

where p^{ref} is a reference pressure, the value of which is trivial in the computations performed in this work. The estimate of the degeneracy and semigrand configurational energies are found self-consistently by iterating Eq. (6.24) and Eq. (6.26), which are given below:

$$\Omega^{*, \text{est}}(N_2, V, \mathcal{U}^{\text{tot}}) = \frac{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(N_2, V, \mathcal{U}^{\text{tot}})}{\sum_{j=1}^R \frac{\mathcal{N}_j}{g_j} \exp \left\{ -\beta_j \mathcal{U}^{\text{tot}} - \beta_j p_j V + \kappa(\beta_j, N_2, V) + N_2 \ln \left(\frac{[\xi_2]_j}{[\xi_1]_j} \right) + N \ln \left([\hat{f}_1]_j / p^{\text{ref}} \right) \right\}} \quad (6.24)$$

The variable $\kappa(\beta_j, N_2, V)$ in Eq. (6.24) is defined by:

$$\kappa(\beta_j, N_2, V) \equiv N \ln(V \beta_j p^{\text{ref}}) + N_1 \ln \left(\frac{\int d\mathbf{q}'_1}{C_1(\beta_j)} \right) + N_2 \ln \left(\frac{\int d\mathbf{q}'_2}{C_2(\beta_j)} \right) \quad (6.25)$$

The configurational semigrand energy at thermodynamic state point j is given by:

$$\frac{N}{\beta} \ln \left(\left[\hat{f}_1 \right]_j / p^{\text{ref}} \right) = -\frac{1}{\beta} \ln \left\{ \sum_{N_2} \sum_V \sum_{\mathcal{U}^{\text{tot}}} \Omega^* \left(N_2, V, \mathcal{U}^{\text{tot}} \right) \right. \\ \left. \exp \left[-\beta_j \mathcal{U}^{\text{tot}} - \beta_j p_j V + \kappa \left(\beta_j, N_2, V \right) + N_2 \ln \left(\frac{\left[\xi_2 \right]_j}{\left[\xi_1 \right]_j} \right) \right] \right\} \quad (6.26)$$

The probability to observe the system with volume, V , energy, \mathcal{U}^{tot} , and number of molecules of component two, N_2 , at reciprocal temperature, β , fugacity fraction, ξ_2 , and pressure, p , is:

$$\pi \left(N_2, V, \mathcal{U}^{\text{tot}}; \beta, \xi_2, p \right) = \\ \frac{\Omega^* \left(N_2, V, \mathcal{U}^{\text{tot}} \right) \exp \left\{ -\beta \mathcal{U}^{\text{tot}} - \beta p V + \kappa \left(\beta, N_2, V \right) + N_2 \ln \left(\frac{\xi_2}{\xi_1} \right) \right\}}{\sum_{N'_2} \sum_{V'} \sum_{\mathcal{U}'^{\text{tot}}} \Omega^* \left(N'_2, V', \mathcal{U}'^{\text{tot}} \right) \exp \left\{ -\beta \mathcal{U}'^{\text{tot}} - \beta p V' + \kappa \left(\beta, N'_2, V' \right) + N'_2 \ln \left(\frac{\xi_2}{\xi_1} \right) \right\}} \quad (6.27)$$

In Section 6.4, a simplified version of Eq. (6.27) will be combined with predictor-corrector methods in order to predict phase equilibria of binary mixtures.

6.4. Advanced Gibbs-Duhem integration

The search for phase equilibrium requires the frequent adjustment of state variables, so as to fulfil the coexistence conditions, and the subsequent computation of the corresponding relevant ensemble averages. As the multiple-histogram reweighting method is a perfect means to obtain ensemble averages over a range of state variables from only a few simulations, it is very convenient in phase-equilibrium computations.

To actually compute phase coexistence from the multiple-histogram reweighting method, there must be a connection (overlap of histograms) between the liquid and the vapour phase. If this connection can be realized, the optimized probability distribution of the *order parameter* close to vapour-liquid phase-coexistence conditions is double-peaked (*bimodal*). Examples of the order parameter are the density in a pure-component simulation and the composition in a binary simulation. Each of the two peaks of the bimodal distribution corresponds to a phase. The coexistence point is given by those thermodynamic state variables for which the areas under the two peaks are equal (*equal peak weight criterion*) [17, 18].

In case the bimodal probability distribution of the order parameter can be measured by one histogram, collected during a *single* simulation in a single simulation box, phase coexistence is obtained by tuning the thermodynamic state variables so as to obtain equal areas under the liquid and the vapour peak [17, 18]. However, accurate prediction of the relative peak areas in a single simulation is often difficult as the free-energy barrier for condensation or

vaporization, which is high at low temperatures, decreases the frequency of sampling both phases.

A possible solution to this problem is to lower the free-energy barrier by exploiting multicanonical sampling methods [18-22]. In case the above procedure does not work, or is computationally too expensive, another connection between the liquid and the vapour phase must be sought for.

The connection may be realized by performing a simulation at a state point in the vicinity of the vapour-liquid critical point. Close to the critical point, the free-energy barrier for condensation or vaporization is so low that a single simulation will frequently sample both liquid-like and vapour-like configurations. Additional histograms must be collected at subcritical liquid and vapour-phase conditions.

Application of the MHR method of Ferrenberg and Swendsen to all liquid and vapour-phase histograms simultaneously results in an estimated degeneracy. An estimate of the bimodal probability distribution of the order parameter, that is valid over a range of state variables close to coexistence, can be extracted from this degeneracy. The criterion of equal peak weight can be easily applied to this probability distribution. Clearly, application of the above-mentioned methods requires some knowledge about the position of the vapour-liquid critical point and the coexistence line before the series of simulation is started.

Applications of the MHR method to computation of vapour-liquid equilibria in the grand-canonical ensemble are given in refs. [23-30]. Applications in the isothermal-isobaric ensemble are given by Conrad and de Pablo [31] and by Chang and Sandler [32].

Meijer and El Azhar [33] proposed an improved version of the Gibbs-Duhem integration technique (GDI), which they called *coexistence-line free-energy difference integration* (CFDI). They applied their method to pure-component phase equilibria. The method is a combination of Gibbs-Duhem integration and Bennett's method [34]. Bennett's method uses energy distributions in order to compute the free-energy difference between two systems and is so to say a *double-histogram method*.

Like GDI, CFDI is developed to keep the free-energy difference between two phases constant. Even if the initial point has a non-zero free-energy difference, *i.e.* it is not really at coexistence, the method is capable of predicting the deviation from the true coexistence line for a series of estimated coexistence points near the true coexistence line.

For a given free-energy difference, $\Delta^{LV}G_i$, between pure-component liquid and vapour phases, the true saturation pressure at the temperature of interest can be estimated from first-order Taylor series expansions of the Gibbs energy, G , expanded about a reference pressure. The saturation pressure at an incremented value of the temperature is estimated from first-order Taylor series expansions of the Gibbs energy, G , expanded about a reference pressure and temperature. This estimate can be refined by computing the free-energy difference $\Delta^{LV}G_j$ at the new state point j . This free-energy difference is computed from the free-energy difference $\Delta^{LV}G_i$ of the former integration point and from the liquid and vapour-phase free-energy differences between the old state point and the new state point:

$$\Delta^{LV}G_j = \Delta^{LV}G_i + (G_j^L - G_i^L) - (G_j^V - G_i^V) \quad (6.28)$$

An advantage of the CFDI method over the conventional GDI method is that every individual point is checked against the criterion of phase coexistence. For a pure-component system at given pressure and temperature, this means zero Gibbs-energy difference between the vapour and the liquid phase.

The conventional GDI method traces the coexistence curve from an estimate of the initial coexistence point via predictor-corrector equations. If the initial point is not a true coexistence point, the GDI method traces a curve of non-zero free-energy difference between the phases. Thus, the difference between the CFDI method and the GDI method lies in the knowledge of the free energy at the initial point and in the way phase coexistence is computed. In the author's opinion, the advantage of the CFDI method over the GDI method is limited as knowledge of the free-energy difference at the initial point enables one to find the true initial coexistence point.

Numerical integration via CFDI suffers from accumulated statistical errors that arise from the successive computation of free-energy differences. On the other hand, GDI suffers from accumulating statistical errors due to successive use of simulated integrands in the corrector equations. The error caused by the use of finite-order corrector equations is assumed to be negligible for small integration steps. Thus, in case the initial point is a true coexistence point, both methods have a similar performance.

A further improvement of the CFDI method would imply the prediction of a new coexistence point from the Clapeyron equation, the computation of the bimodal probability distribution of the order parameter (possibly using multicanonical methods), and the computation of phase coexistence from the criterion of equal peak weight. This improved method reduces the contribution of the Clapeyron equation to the first estimation of the coexistence point; the actual coexistence point is computed from histogram reweighting.

Escobedo [35-37] developed novel formulations of Gibbs-Duhem integrations in a pseudo-ensemble framework. He extensively described the analogies and differences between conventional Gibbs-Duhem integration and multiple-histogram reweighting. A coexistence point computed with the conventional GDI method is not improved by data collected at subsequent integration steps. The MHR method uses histograms collected near the coexistence line. The actual coexistence points are computed when all histograms have been collected and reweighted. As histograms collected at nearby state points need some overlap in order to combine them, different histograms improve the estimate of a single coexistence point. Both methods use an approximate analytical scheme to combine data from simulations at subsequent state points. In the MHR method, the method of Ferrenberg and Swendsen is used, while the GDI method uses predictor-corrector equations. For sharply peaked histograms, the methods become alike [36].

In this work, the Clapeyron equation is still used to compute phase coexistence. In a conventional Gibbs-Duhem integration scheme, a 'new' coexistence point at an incremented value of the independent integration variable is estimated from the predictor equation. Actual liquid and vapour-phase simulations performed at the predicted coexistence point, along with applications of the relevant corrector equation, enable one to compute an improved estimate of the coexistence point. The dependent variable is the saturation pressure and the integration variable is the fugacity fraction. Thus, after every corrector simulation, the estimate of the saturation pressure is updated. This search for the 'true' saturation pressure, which is close to the *predicted* saturation pressure in case small integration steps are used, can also be performed with histogram reweighting.

Suppose a liquid-phase and a vapour-phase histogram are collected at the estimated coexistence point that has been obtained from the predictor equation. From Section 6.2, it is known that histograms and ensemble averages can be reweighted to nearby thermodynamic state points. The condition for phase coexistence used in this work is implied in the corrector equation and in formerly determined coexistence points. Thus, the liquid and vapour-phase histograms need not overlap (neither direct nor indirect), which is clearly an advantage over other methods. Application of the relevant corrector equation and histogram reweighting provides an improved estimate of the coexistence point. This procedure requires *only one*

liquid and vapour-phase simulation per coexistence point. However, as pointed out in Section 6.2, the extrapolative power of a single histogram is limited to a narrow range of state variables. This is in particular true for large systems at low temperature. Therefore, the single-histogram reweighting approach is only appropriate in case the integration step is very small. In order to make the method applicable for larger integration steps, a number of simulations is performed at state variables that cover the region nearby the coexistence point that has been estimated from the predictor equation. This enables the computation of phase coexistence from multiple-histogram reweighting by *interpolation* instead of *extrapolation* to nearby state conditions.

The total procedure of the *advanced Gibbs-Duhem integration method* is as follows. Equations (6.24) and (6.27) can be simplified considerably by remembering that all simulations in this work are performed isothermally. Thus, reweighting to other temperatures is not needed to compute phase coexistence. The conjugate density of the reciprocal temperature, β , is the energy, \mathcal{U} . It is possible to perform the summation over the energy beforehand. This is convenient as it reduces the histograms by one dimension. The pseudo-Boltzmann factor in Eq. (6.24) is referred to a certain reference state, say $(p_0, [\xi_2]_0, \beta)$. Division of the denominator and numerator of Eq. (6.24) by the pseudo-Boltzmann factor at the reference state point gives:

$$\Omega^{*\text{ est}}(N_2, V, \mathcal{U}^{\text{tot}}) = \frac{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(N_2, V, \mathcal{U}^{\text{tot}}) \exp\left(+\beta \mathcal{U}^{\text{tot}} + \beta p_0 V - \kappa(\beta, N_2, V) - N_2 \ln \left\{ \frac{[\xi_2]_0}{[\xi_1]_0} \right\}\right)}{\sum_{j=1}^R \mathcal{N}_j g_j^{-1} \exp\left\{-\beta(p_j - p_0)V + N_2 \ln \left(\frac{[\xi_2]_j [\xi_1]_0}{[\xi_1]_j [\xi_2]_0} \right) + N \ln \left([\hat{f}_1]_j / p^{\text{ref}} \right)\right\}} \quad (6.29)$$

It is convenient to define a new *temperature-dependent pseudo-degeneracy*:

$$\Theta^*(N_2, V) \equiv \frac{\sum_{\mathcal{U}^{\text{tot}}} \Omega^{*\text{ est}}(N_2, V, \mathcal{U}^{\text{tot}}) \exp\left(-\beta \mathcal{U}^{\text{tot}} - \beta p_0 V + \kappa + N_2 \ln \left\{ \frac{[\xi_2]_0}{[\xi_1]_0} \right\}\right)}{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(N_2, V)} \quad (6.30)$$

$$= \frac{\sum_{j=1}^R \mathcal{N}_j g_j^{-1} \exp\left\{-\beta(p_j - p_0)V + N_2 \ln \left(\frac{[\xi_2]_j [\xi_1]_0}{[\xi_1]_j [\xi_2]_0} \right) + N \ln \left([\hat{f}_1]_j / p^{\text{ref}} \right)\right\}}{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(N_2, V)}$$

Using Eq. (6.30), probability distribution (6.27) simplifies to:

$$\pi(N_2, V; \xi_2, p) = \frac{\Theta^*(N_2, V) \exp\left(-\beta(p - p_0)V + N_2 \ln \left\{ \frac{\xi_2 [\xi_1]_0}{\xi_1 [\xi_2]_0} \right\}\right)}{\sum_{N'_2} \sum_{V'} \Theta^*(N'_2, V') \exp\left(-\beta(p - p_0)V' + N'_2 \ln \left\{ \frac{\xi_2 [\xi_1]_0}{\xi_1 [\xi_2]_0} \right\}\right)} \quad (6.31)$$

The starting point for the numerical integration of the Clapeyron equation is the saturation pressure of one of the pure components, along with the liquid and vapour densities at coexistence, and the ratios f_1/H_2 .

The predictor equation is used to make a first estimate of the saturation pressure at an incremented value of the fugacity fraction. In order to cover the region of interest, $m-1$ thermodynamic state points $\{p_i, [\xi_2]_i; i=1 \dots m-1\}$ are chosen with values nearby the estimated saturation pressure at the imposed fugacity fraction. Thus, m simulations per coexistence point are performed. This *fixed number of simulations per coexistence point* is a clear advantage of the GDI method combined with MHR over the conventional GDI method and the CFDI method.

During the course of the simulations, m histograms $\mathcal{H}_i(N_2, V)$ are collected, containing combinations of the encountered volumes and compositions. These histograms are collected for both liquid and vapour phases. The histograms of the liquid phase are combined with the method of Ferrenberg and Swendsen to obtain an estimate of the probability distribution $\pi^L(N_2, V; \xi_2, p)$ that is valid over the range of pressures and fugacity fractions covered by the simulations. The same procedure applies to the vapour phase.

The estimated integrand of the Clapeyron equation at any state point $\{p, \xi_2; \beta\}$ is obtained from:

$$F_{Np\beta\xi_2} = \frac{1}{\xi_1 \xi_2} \frac{1}{\beta p} \frac{\langle N_2/N \rangle_{Np\beta\xi_2}^L - \langle N_2/N \rangle_{Np\beta\xi_2}^V}{\langle V/N \rangle_{Np\beta\xi_2}^L - \langle V/N \rangle_{Np\beta\xi_2}^V} \quad (6.32)$$

The properties within brackets are computed from:

$$\langle \mathcal{O} \rangle_{Np\beta\xi_2}^\alpha = \sum_{N_2^\alpha} \sum_{V^\alpha} \mathcal{O}(N_2, V) \pi^\alpha(N_2, V; p, \xi_2) \quad (6.33)$$

In the procedure described above, the probability distribution $\pi^\alpha(N_2, V; \xi_2, p)$ is estimated by combining m histograms at a fixed fugacity fraction (single coexistence point). However, in case the integration steps are not too large, histograms obtained at different fugacity fractions (different coexistence points) have considerable overlap as well. Combination of the histograms collected at multiple integration steps extends the region of validity of $\pi^\alpha(N_2, V; \xi_2, p)$.

Having found the liquid-phase and vapour-phase probability distributions, the corrector equation is used to iterate to the saturation pressure at the imposed fugacity fraction. One can also use the corrector equation to compute coexistence points at interpolated values of the fugacity fraction; *i.e.* a whole *coexistence line* can be computed instead of only coexistence points at predetermined fugacity fractions. The densities and mole fractions at the coexistence points are found from application of Eq. (6.33).

The advanced Gibbs-Duhem integration method enables one to search for the saturation pressure at a particular mole fraction or density of the liquid or the vapour phase. As histograms collected at the current integration step also improve the estimate of that part of the degeneracy that is important at prior integration steps, the current simulation improves the

predictions of formerly computed coexistence points as well. This *retroactive behaviour* is another advantage over the conventional GDI method and the CFDI method.

The multiple-histogram estimate for $\pi^\alpha(N_2, V; \xi_2, p)$ enables the computation of the right-hand side of Eq. (6.32) as a function of pressure and fugacity fraction. An example of this ‘surface’ in 3-dimensional space is given in Figure 6.2. The surface has been obtained by combining 90 liquid-phase and 90 vapour-phase histograms. Figure 6.2 also contains the computed coexistence line (Clapeyron equation).

Obviously, the advanced Gibbs-Duhem integration method provides a smooth surface for $F_{Np\beta\xi_2}$ in the range of pressures and fugacity fractions covered by the simulations. As pointed out in Section 6.2, outside this range, *i.e.* farther away from the coexistence line, the extrapolative power of the histogram-reweighting method is limited due to bad sampling of the ‘tails’ of the histograms.

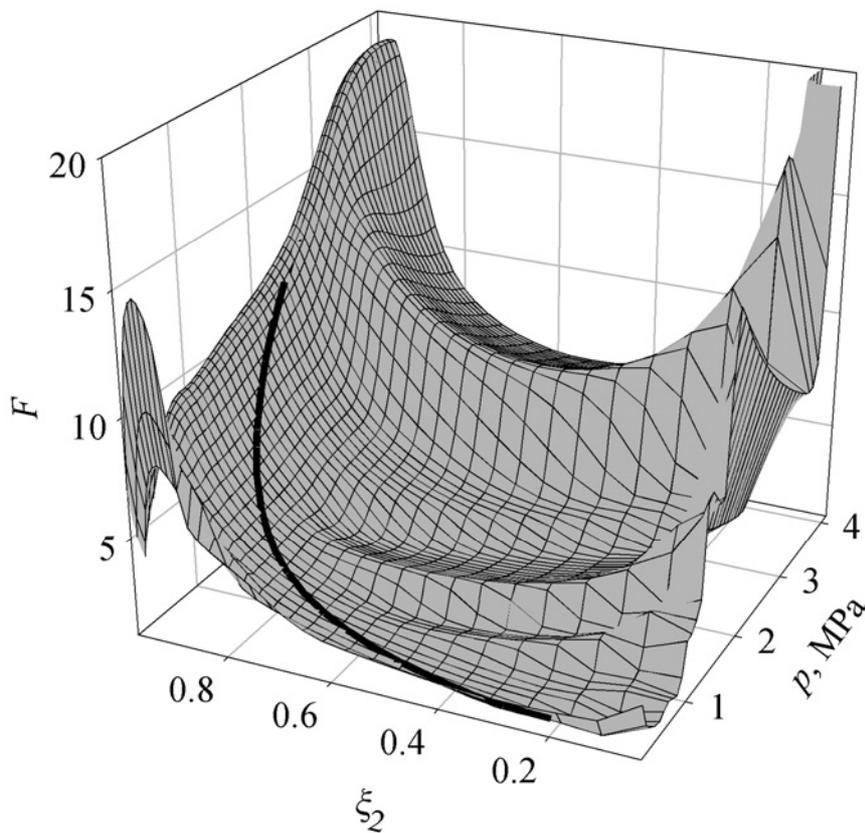


Figure 6.2: *Integrand of the Clapeyron differential equation as a function of pressure and fugacity fraction. The solid line represents the integrand at vapour-liquid equilibrium conditions. Simulations are performed with united-atom force fields for the binary mixture ethane/methane at $T = 192.37$ K.*

6.5. Parallel tempering in the semigrand-canonical ensemble

The parallel-tempering method has been developed to enhance sampling of systems that have an energy landscape with many local minima. Especially at low temperature, the energy barriers between the local minima prevent efficient sampling of the whole energy landscape; *i.e.* the system is trapped in a single local minimum. However, at higher temperatures the system crosses the energy barriers more easily. The first implementations of the parallel-

tempering method used a number of parallel simulations, say m , each at a different temperature.

In addition to conventional trial moves, the parallel-tempering scheme includes swaps of configurations belonging to systems with *neighbouring temperatures*. This implementation enables low-temperature systems that are trapped in a local energy minimum to cross the energy barriers at a higher temperature. Clearly, the neighbouring temperatures should be not too different in order to assure a reasonable percentage of accepted parallel-tempering configuration swaps. A description of the parallel-tempering method and an overview of related methods is given in the book of Frenkel and Smit [38].

Implementation of parallel tempering in the advanced Gibbs-Duhem integration method is very convenient. Per coexistence point, m simulations are performed simultaneously, each at a different thermodynamic state point $\{[\xi_2]_i, p_i; i=1 \dots m\}$. In order to apply multiple-histogram reweighting to the simulation results, the histograms at neighbouring thermodynamic state points should have a region of overlap. The parallel-tempering method is a perfect means to check whether there is sufficient overlap; without overlap, the fraction of accepted parallel-tempering configuration swaps approaches zero. Furthermore, parallel-tempering configuration swaps improve sampling of phase space.

The derivation of the acceptance criterion for the parallel-tempering configuration swaps starts by defining an extended semi-classical semigrand-canonical partition function. This extended partition function is the product of the m individual partition functions at different thermodynamic states [38]:

$$\begin{aligned} \mathcal{Y}^{\text{extended}} &\equiv \prod_{j=1}^n \mathcal{Y}_{N p_j \beta [\xi_2]_j} \\ &= \prod_{j=1}^n \int dV_j \left(\frac{V_j \beta p^{\text{ref}}}{\langle \mathcal{W}_{\text{IG},1}^{\text{LJint} + \text{Cint}} \rangle} \right)^N \frac{\exp\{-\beta p_j V_j\}}{V_j (N-1)!} \sum_{\text{idem}} \left(\frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJint} + \text{Cint}} \rangle_{[\xi_2]_j}^{[N_2]_j}}{\langle \mathcal{W}_{\text{IG},2}^{\text{LJint} + \text{Cint}} \rangle_{[\xi_1]_j}} \right) \cdot \\ &\quad \int d\Psi^{[N_1]_j} d\Psi^{[N_2]_j} \exp\left\{-\beta \mathcal{U}_j^{\text{ext} + \text{LJint} + \text{Cint}} \left(\Psi^{[N_1]_j}, \Psi^{[N_2]_j} \right)\right\} \end{aligned} \quad (6.34)$$

A configuration swap between neighbouring states i and j is accepted with probability:

$$\min \left\{ 1, \exp \left[\beta (p_j - p_i) (V_j - V_i) + \ln \left\{ \frac{[\xi_2]_i [\xi_1]_j}{[\xi_1]_i [\xi_2]_j} \right\} \left([N_2]_j - [N_2]_i \right) \right] \right\} \quad (6.35)$$

6.6. Simulation results for binary vapour-liquid equilibria

6.6.1. Simulation details

Simulation details for conventional Gibbs-Duhem integrations were given in Chapter 5. In this section, only the details are discussed that are typical for the advanced GDI method. A brief error analysis of the advanced Gibbs-Duhem integration method is given in Section 6.7. In the applications discussed in this chapter, the energy-biased trial identity change was used with bias strengths $\gamma = -0.5$ and $\omega = 0$.

The number of state points per integration step was fixed at three. Thus, the computation of one coexistence point required simulation of three liquid phases and three vapour phases. Every simulation took 2500 MC cycles to equilibrate the system and 10000 production MC cycles to collect the histograms at the state points of interest.

The simulations in the semigrand-canonical ensemble were performed with 300 molecules in the liquid phase and 150 molecules in the vapour phase. Every MC cycle (per phase) consisted of $0.5N$ translations, $0.5N$ rotations, N identity changes, a change of the volume, and $0.05N$ parallel-tempering configuration switches. These switches were only performed between liquid phases with neighbouring state points or between vapour phases with neighbouring state points. The higher-order variable step-size predictor-corrector equations, as described in Chapter 2, were used.

The bin width for the volume dimension of the histograms was determined during the second half of the equilibration phase of the simulation at the very first integration step. The bin width was given such a value so as to produce approximately 40 volume bins. As a large number of histograms, obtained at different thermodynamic state points, is combined, the total volume and composition may be very different in successive simulations.

For the composition, this is no problem since the number of bins for N_2 cannot exceed the total number of molecules. However, the change of volume can be more problematic. Consider the vapour-liquid equilibria of a binary mixture in which component two is supercritical. A schematic picture of the vapour-liquid equilibria of such a system is given in Figure 6.3, left-hand picture. The numerical integration of the Clapeyron equation starts at pure component one. The volume bin width is determined during the equilibration period of the very first simulation. As this simulation takes place at low pressure, the difference between the liquid and vapour-phase molar volumes can be huge. Since the initial number of bins for the liquid and the vapour phase are equal, the vapour-phase volume bin width is much larger than the volume bin width of the liquid phase. Now, the integration proceeds to higher fugacity fractions and higher pressures. The mixture approaches the vapour-liquid critical point, resulting in a decreasing difference between the molar volumes of the vapour and the liquid phase. Thus, the number of volume bins in the liquid phase grows. On the other hand, the number of volume bins in the vapour phase decreases.

In order to avoid an unbridled increase of liquid-phase volume bins or a low resolution of the volume distribution in the vapour phase, a logarithmic volume bin width, $\Delta \ln V$, is used. The lower (–) and upper (+) limits of the volumes that contribute to bin i in phase α are computed from:

$$\begin{aligned} V_i^{\alpha,-} &= \exp\left\{\ln\left(V_{\text{ref}}^{\alpha}\right) + i\Delta \ln V^{\text{L}}\right\} \\ V_i^{\alpha,+} &= \exp\left\{\ln\left(V_{\text{ref}}^{\alpha}\right) + (i+1)\Delta \ln V^{\text{L}}\right\} \end{aligned} \quad (6.36)$$

where V_{ref}^{α} is the reference volume of phase α . The discretized volume that belongs to bin i is given by:

$$V_i^{\alpha} = \frac{V_i^{\alpha,-} + V_i^{\alpha,+}}{2} \quad (6.37)$$

If the shape of the volume distribution is approximately known from the outset of the simulation, this knowledge can be used to define an improved version of Eq. (6.37).

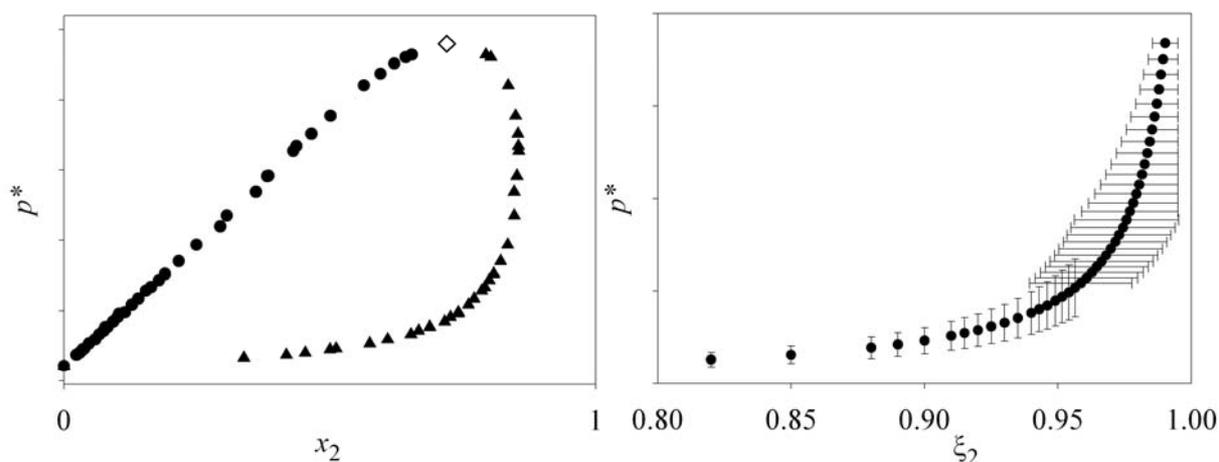


Figure 6.3: Schematic pictures of a system in which component two is supercritical. *Left-hand picture: vapour-liquid equilibria. Right-hand picture: saturation pressure versus fugacity fraction. Solid spheres between vertical bars represent first estimates of the saturation pressures obtained from the predictor equation at an imposed fugacity fraction. Solid spheres between horizontal bars represent first estimates of the fugacity fractions obtained from the predictor equation at imposed saturation pressures. The ends of the bars represent the state points in the neighbourhood of the estimated coexistence point.*

The right-hand picture in Figure 6.3 shows how the thermodynamic state points at successive integration points are chosen. In case the slope of the Clapeyron equation is moderate, the fugacity fraction is incremented with an amount that makes sure that the composition dimensions of the histograms at state point i have enough overlap.

As already mentioned, the saturation pressure at the incremented fugacity fraction is estimated from the predictor equation. Two other state points that only differ in pressure are chosen; one pressure about 10% lower and one about 10% higher than the predicted saturation pressure. As the pressure changes only slightly with increasing fugacity fraction (moderate slope), the volume dimensions of the histograms will have enough overlap.

From Figure 6.3, it is clear that the slope of the Clapeyron equation can reach very large values in case component two is supercritical. This means that a tiny increment in the fugacity fraction results in a large change of the predicted saturation pressure. It is questionable whether the volume dimension of the histograms collected at the initially estimated saturation pressure will have enough overlap with the volume dimensions of histograms collected at former thermodynamic state points. Therefore, it is convenient to increment the saturation pressure instead of the fugacity fraction. The fugacity fraction at the incremented saturation pressure is then estimated from a polynomial fit to former coexistence points. Two other state points that only differ in fugacity fraction are chosen, one fugacity fraction about 3% lower and one about 3% higher than the initially estimated fugacity fraction at the imposed saturation pressure.

A lack of overlap between histograms collected at nearby state points is easily detected from an increasing number of MHR optimization iterations, or earlier during the simulation, from a low percentage of accepted parallel-tempering MC moves.

Like other authors do either implicitly [31, 32, 39] or explicitly [26], the correlation between subsequent samples is neglected in the application of the MHR method of Ferrenberg and Swendsen to the collected histograms. This means that the g_j 's in Eq. (6.30) are put to one. The influence of this approximation has not been investigated but seems to be small [10].

6.6.2. Binary system methane/ethane

Sources of experimental data for vapour-liquid equilibria of the binary system methane/ethane were given in Table 5.9. Simulation results, obtained with very simple force fields, were presented in Chapter 5. An overview of simulations performed by other researchers was also given in Chapter 5. In this section, simulation results at $T = 250$ K are presented. Instead of the simple force fields used in Chapter 5, the advanced TraPPE-EH force field is now used [40] for both ethane and methane. The simulation results obtained at this temperature will be used in Chapter 7 as a starting point for a ternary Gibbs-Duhem integration.

Simulation results for the initial point (pure ethane) and the initial slope of the Clapeyron equation in the limit of pure ethane are given in Table 6.1.

Table 6.1: Simulated ratios of solvent's (ethane) fugacity over solute's (methane) Henry's constant.

T K	p^{sim} MPa	ρ_L^{sim} kg/m ³	ρ_V^{sim} kg/m ³	$\left(\frac{f_{\text{C}_2\text{H}_6}}{H_{\text{CH}_4}}\right)_L^{\text{sim}}$	$\left(\frac{f_{\text{C}_2\text{H}_6}}{H_{\text{CH}_4}}\right)_V^{\text{sim}}$	Comb. rule unlike molecules	Method f_1/H_2
250	1.313(7)	443.5(4)	23.3(2)	0.108(6)	0.83(2)	LHMcC	DM+OID

The saturation pressure for pure ethane at $T = 250$ K differs from the one presented in Chapter 4, Table 4.1 and Table 4.2. However, the results match within the statistical uncertainty. The saturation pressure given here (and in Table 4.3) was obtained from a longer simulation and has a lower statistical uncertainty. The Lorentz-Berthelot (LB) combining rules for the interactions between like molecules and the Lorentz-Hudson-McCoubrey (LHMcC) combining rules for interactions between unlike molecules were used. The difference between the Hudson-McCoubrey and Berthelot combining rules is negligible for this binary mixture as ethane and methane have similar ionization potentials and almost identical collision diameters (See Table 5.1).

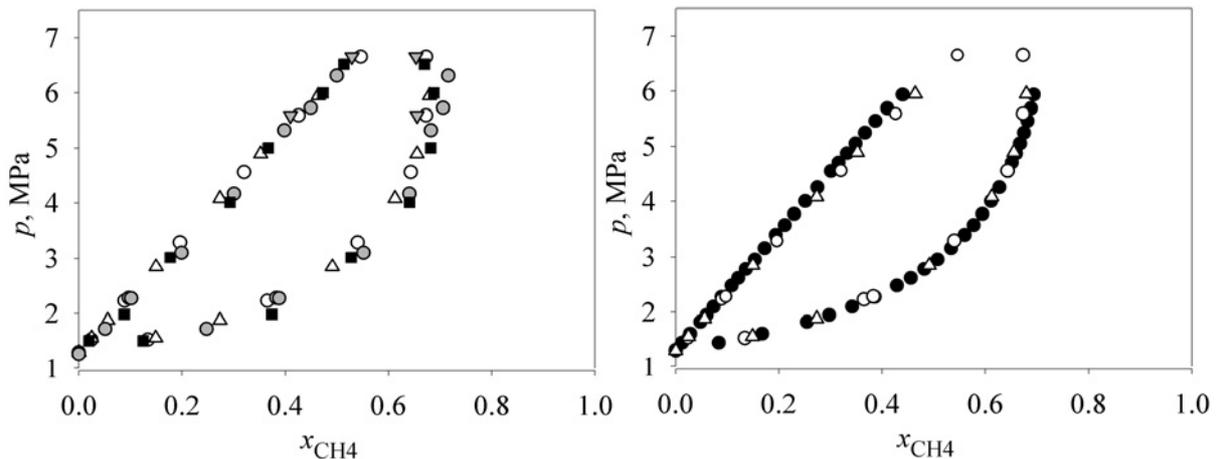


Figure 6.4: Vapour-liquid equilibria of the binary mixture ethane/methane at $T = 250$ K. Open circles and open upward triangles respectively represent experimental vapour-liquid equilibria published by Davalos et al. [41] and Wei et al. [42]. Left-hand picture: grey circles, black squares, and grey downward triangles respectively represent simulation results by Vrabec and Fischer [43], by Liu and Beck [44], and by Zhang and Duan [45]. Right-hand picture: black circles are simulation results obtained in this work.

Simulation results for the vapour-liquid equilibria of the binary mixture are given in Figure 6.4 and in Figure 6.5. These figures also include simulation results published by other authors. The simulation results are also presented in tabular form in Appendix D, Table D.1.

The simulation results presented in this section, which are performed with the advanced Gibbs-Duhem integration method, are much *smoother* than the simulation results published by other researchers. In general, they are also smoother than the results obtained from the conventional Gibbs-Duhem integration method. This phenomenon is caused by the retroactive nature of multiple-histogram reweighting; histograms collected during the course of a current simulation improve the estimate of the degeneracy in the range that is important to previously computed coexistence points.

The simulation results obtained in this work could even well have been presented by a line as the combination of the Clapeyron equation and the optimized degeneracies of the liquid and the vapour phase enable the computation of phase coexistence at any saturation pressure or at any fugacity fraction within the ranges covered by the simulations.

The simulation results obtained with advanced Gibbs-Duhem integration are in good agreement with experimental data. Unfortunately, there are no experimental density-data to compare the simulation results with. However, Vrabec and Fischer [43] compared their simulation results for liquid and vapour densities at coexistence with predictions from an equation of state (EOS). Their results were in good agreement with the EOS-predictions, and the results presented in this section are in good agreement with the results of Vrabec and Fischer, as is clear from Figure 6.5. Thus, it is stated that the simple force fields of Vrabec and Fischer perform equally well (at least at this temperature) as the advanced TraPPE-EH force field. Nevertheless, the complexity of the TraPPE-EH force field is justified by its transferability.

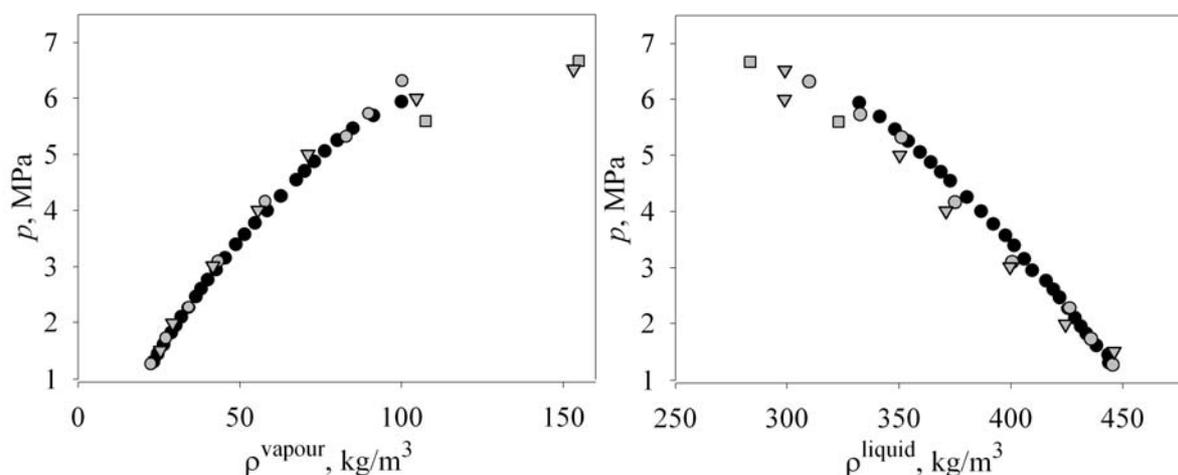


Figure 6.5: Liquid and vapour densities at coexistence of the binary mixture ethane/methane at $T = 250\text{ K}$. Downward triangles are simulation results by Liu and Beck [44]. Grey circles are simulation results published by Vrabec and Fischer [43]. Grey squares are simulation results taken from Zhang and Duan [45]. Black circles are simulation result obtained in this work. Left-hand picture: vapour density. Right-hand picture: liquid density.

6.6.3. Binary system propane/methane

The binary simulation results in this section will be used in Chapter 7 as a starting point for a ternary Gibbs-Duhem integration. Sources of experimental data for vapour-liquid equilibria of the binary system propane/methane were summarized by Webster and Kidnay [46].

The advanced TraPPE-EH force field [47] is used for both propane and methane. Simulation results for the initial point (pure propane) and the initial slope of the Clapeyron equation in the limit of pure propane are given in Table 6.2. The LB combining rules are used for the interactions between like molecules and the LHMcC combining rules for the interactions between unlike molecules. The difference between the Hudson-McCoubrey and the Berthelot combining rules is negligible for this mixture as propane and methane have similar ionization potentials and almost identical collision diameters (See Table 5.1).

Table 6.2: Simulated ratios of solvent's (propane) fugacity over solute's (methane) Henry's constant

T K	p^{sim} MPa	ρ_L^{sim} kg/m ³	ρ_V^{sim} kg/m ³	$\left(\frac{f_{\text{C}_3\text{H}_8}}{H_{\text{CH}_4}}\right)_L^{\text{sim}}$	$\left(\frac{f_{\text{C}_3\text{H}_8}}{H_{\text{CH}_4}}\right)_V^{\text{sim}}$	Comb. rule unlike molecules	Method f_1/H_2
270	0.432(9)	526.2(3)	9.3(2)	0.32(5)	0.89(2)	LHMcC	DM+OID

Simulation results for the vapour-liquid equilibria of the binary mixture are presented in graphical form in Figure 6.6. The results are also given in tabular form in Appendix D, Table D.2. The simulated vapour-liquid equilibria perfectly match the experimental data over a large range of fugacity fractions/pressures. The results close to the vapour-liquid critical point deviate from the experimental ones, probably due to finite size effects. The simulation results for ethane/methane [Section 6.6.2] and for propane/methane show how well the advanced TraPPE-EH force field performs.

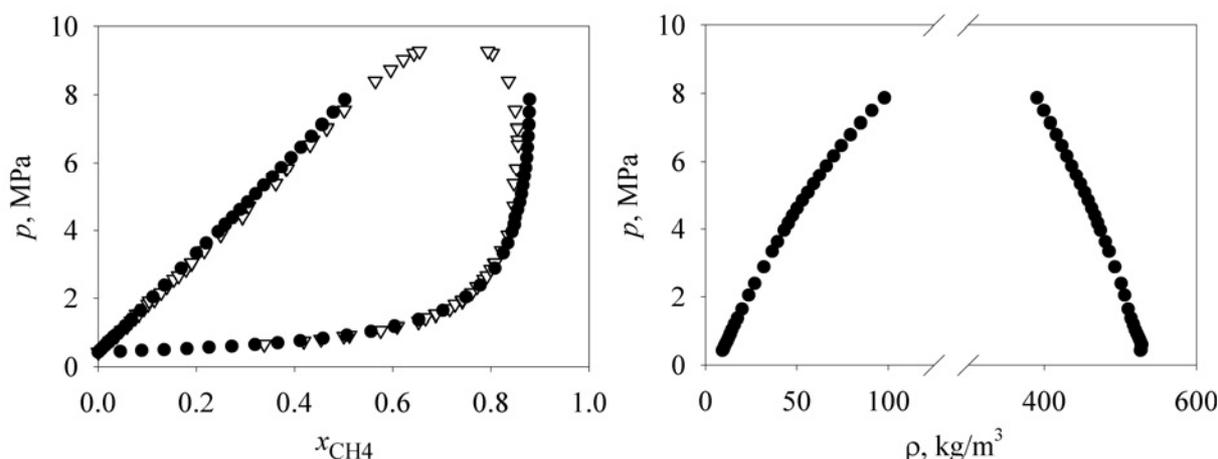


Figure 6.6: Vapour-liquid equilibria of the binary mixture propane/methane at $T = 270$ K . Open downward triangles represent experimental data from Webster and Kidnay [46]. Solid circles represent simulation results obtained in this work. Left-hand picture: pxy -data. Right-hand picture: liquid and vapour densities at coexistence.

6.6.4. Binary system ethane/carbon dioxide

The vapour-liquid equilibria of the binary mixture ethane/carbon dioxide have been simulated by a number of researchers with a variety of simulation techniques, force fields and combining rules. The reason is probably the occurrence of maximum-pressure azeotropy. Furthermore, the system exhibits critical azeotropy [48], however, this phenomenon has not been studied with molecular simulation.

Maximum-pressure azeotropy occurs when component one is a bad solvent for component two and vice versa. Although azeotropy is not a very exciting phenomenon from a thermodynamic point of view, it is a challenge to predict the right azeotropic pressure and composition with molecular simulation.

Several sources of experimental data for vapour-liquid equilibria of the binary system ethane/carbon dioxide are given in Table 6.3. Other sources are listed in ref. [49].

Table 6.3: Sources of experimentally determined vapour-liquid equilibria of the binary system ethane/carbon dioxide

Reference	Temp. range, K	Press. range, MPa	Number of sets
Davalos <i>et al.</i> [49]	250	1.3-2.1	1
Fredenslund and Mollerup [50]	223-293	0.5-5.6	5
Wei <i>et al.</i> [51]	207-270	0.3-3.6	6
Price [52]	144.26-283.15	0.7-5.5	6
Hamam and Lu [53]	222.04-288.71	0.7-5.7	4
Ohgaki and Katayama [54]	283.15-298.15	3.0-6.5	5
Nagahama <i>et al.</i> [55]	253	1.4-2.3	1
Gugnoni <i>et al.</i> [56]	241.45-283.15 ^a	1.0-4.9	4
Brown <i>et al.</i> [57]	207-270	0.3-3.6	10

^a Only liquid-phase compositions and densities

Vrabec and Fischer [43] simulated this system at $T = 223.15$ K, $T = 250$ K, and $T = 283.15$ K with the mixture extension of the $Np\beta$ + test molecule method. They used a 2CLJ model for ethane [58, 59] and a 2CLJQ model for carbon dioxide [43]. Binary interaction parameters with values close to zero were included in the LB combining rules [See Section 5.2, Eq. (5.13)]. Liu and Beck [44] used the same intermolecular potentials and binary interaction parameters to study the system at $T = 250$ K with the $Np\beta$ – Gibbs ensemble.

Stoll *et al.* [60] also used a 2CLJ model for ethane and a 2CLJQ model for carbon dioxide with LB combining rules. A binary interaction parameter for the unlike-components well-depths was used. They simulated the binary mixture at $T = 223.15$ K, $T = 263.15$ K, and $T = 283.15$ K. Their results can hardly be compared with experimental data as they were only given in graphical form. Furthermore, they were presented over a range of pressures that is at least three times larger than the difference between the pure-component saturation pressures at a single temperature.

Potoff *et al.* [25] used an exponential-6 potential with additional point charges to account for the electrostatic interactions between carbon dioxide molecules. Either the LB combining rules for all interactions or the LB combining rules for the interactions between like molecules and the Kong combining rules for the interactions between unlike molecules were used. Carbon dioxide was modelled as a rigid linear three-site molecule with point charges and ethane was modelled as a 2CLJ rigid united-atom molecule. Their simulations were performed in the grand-canonical ensemble. Phase equilibria were computed by applying the multiple-histogram reweighting method to the simulation results.

In this section, simulation results of vapour-liquid equilibria at $T = 207$ K, $T = 250$ K, and $T = 263.15$ K are presented. Every simulation was performed twice. The first simulation was performed with the LB combining rules for all interactions, while in the second case the LB combining rules for the interactions between like components and the LHMCC combining rules for the interactions between unlike components were used. Ethane and carbon dioxide were modelled with the TraPPE-EH [40] and TraPPE-AA [23] force fields respectively. Simulation results for the initial point (pure ethane) and the initial slope of the Clapeyron equation in the limit of pure ethane are given in Table 6.4. The saturation pressures of pure ethane were computed from volume perturbations in the Gibbs ensemble.

Table 6.4: Coexistence points of pure ethane and simulated ratios of solvent's (ethane) fugacity over solute's (carbon dioxide) Henry's constant.

T K	p^{sim} MPa	ρ_L^{sim} kg/m ³	ρ_V^{sim} kg/m ³	$\left(\frac{f_{\text{C}_2\text{H}_6}}{H_{\text{CO}_2}}\right)_L^{\text{sim}}$	$\left(\frac{f_{\text{C}_2\text{H}_6}}{H_{\text{CO}_2}}\right)_V^{\text{sim}}$	Comb. rule unlike molecules	Simulation Method
207	0.297(5)	506.6(4)	5.5(1)	0.20(1)	0.97(1)	LHMcC	DM+OID
250	1.313(7)	443.5(4)	23.3(2)	0.32(1)	0.89(1)	LHMcC	DM+OID
263.15	1.85(1)	419.8(5)	33.2(2)	0.34(2)	0.87(2)	LHMcC	DM+OID
207	0.297(5)	506.6(4)	5.5(1)	0.252(9)	0.961(8)	LB	DM+OID
250	1.313(7)	443.5(4)	23.3(2)	0.37(3)	0.89(2)	LB	DM+OID
263.15	1.85(1)	419.8(5)	33.2(2)	0.40(3)	0.88(3)	LB	DM+OID

Simulation results for the vapour-liquid equilibria of the binary mixture are presented in graphical form in Figure 6.7 to Figure 6.11. The results are also given in tabular form in Appendix D, Table D.3.

Figure 6.8 and Figure 6.9 include simulation results published by Liu and Beck [44] and by Vrabec and Fischer [43]. In addition to the simulation results obtained in this work, Figure 6.10 also contains simulation results published by Potoff *et al.* [25]. The simulation results at $T = 207$ K show that, though the prediction is qualitatively correct, even the combination of the advanced TraPPE-AA force field for carbon dioxide and the TraPPE-EH force field for ethane, do not predict the right azeotropic pressure. The azeotropic composition is predicted fairly well. Thus, one can state that either optimization of the TraPPE-AA force field for carbon dioxide to a single binary alkane/CO₂ mixture [See Section 5.1.4.3] is not enough or the LB combining rules are simply not good enough to capture the phase behaviour of ethane/carbon dioxide mixtures. It is also concluded from Figure 6.7 (and from the simulation results at other temperatures) that the use of the LHMcC combining rules, instead of the LB combining rules, for interactions between unlike molecules brings experimental data and simulation results much closer.

It is known from experiment that pure carbon dioxide forms a stable solid phase at $T = 207$ K [61]. However, the Gibbs-Duhem integration method traces the vapour-liquid coexistence line up to pure carbon dioxide.

The expected phase behaviour of the binary mixture at $T = 207$ K, based on experimental points from ref. [62], is indicated in Figure 6.7. The uppermost experimental vapour-liquid saturation pressures in Figure 6.7 are close to LVS- three-phase equilibrium. The solid phase is expected to consist of pure carbon dioxide [62]. As the formation of a liquid-solid interface in the liquid simulation has a high free-energy barrier, freezing of the liquid phase was not observed. Generally, simulation of a solid phase is only possible when the starting configuration is a crystal structure. Nevertheless, the end-point of the simulation, at $\xi_{\text{CO}_2} \approx 1$ is very close to the experimental sublimation pressure of pure carbon dioxide [61]!

The simulation results at $T = 250$ K show the same trends as those obtained at $T = 207$ K. The results obtained with the advanced Gibbs-Duhem integration method are much smoother than those published by Vrabec and Fischer [43] and by Liu and Beck [44]. However, the simulation results taken from Vrabec and Fischer are closer to experimental data. Although their force fields were very simple, this is not surprising as they used a binary interaction parameter that was optimized to mixture data.

Notice the rather odd simulation results of Liu and Beck [44] that are obtained by exploiting force fields identical to those of Vrabec and Fischer [43].

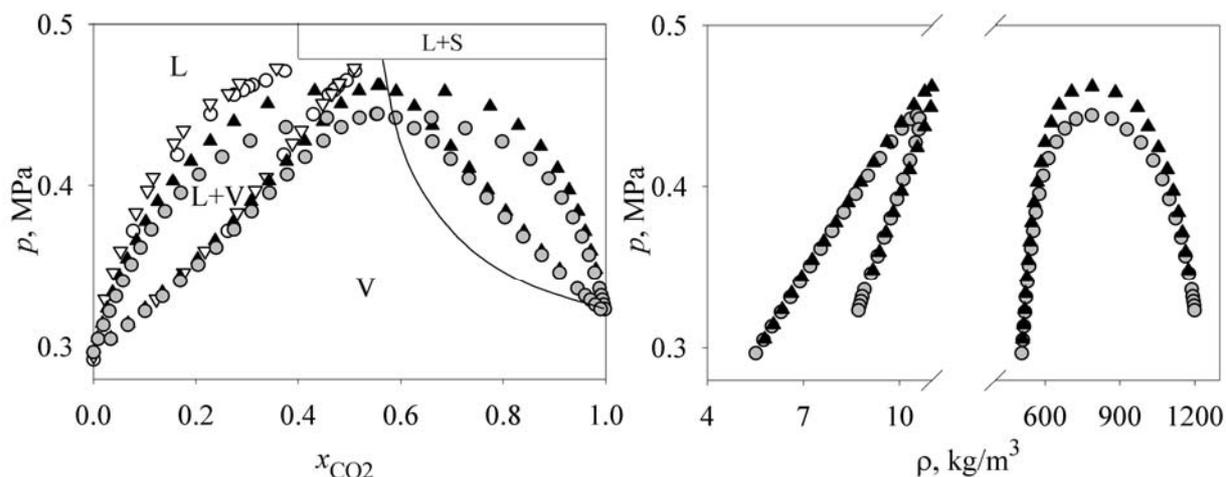


Figure 6.7: Vapour-liquid equilibria of the binary mixture ethane/carbon dioxide at $T = 207 \text{ K}$. Open circles and open downward triangles respectively represent experimental vapour-liquid equilibria published by Brown et al. [57] and Wei et al. [51]. Solid grey circles represent simulation results obtained with LB combining rules, while solid black upward triangles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Left-hand picture: p - x data. Solid lines indicate a qualitative prediction of the phase behaviour based on measurements published in ref. [62]. Right-hand picture: liquid and vapour densities at coexistence.

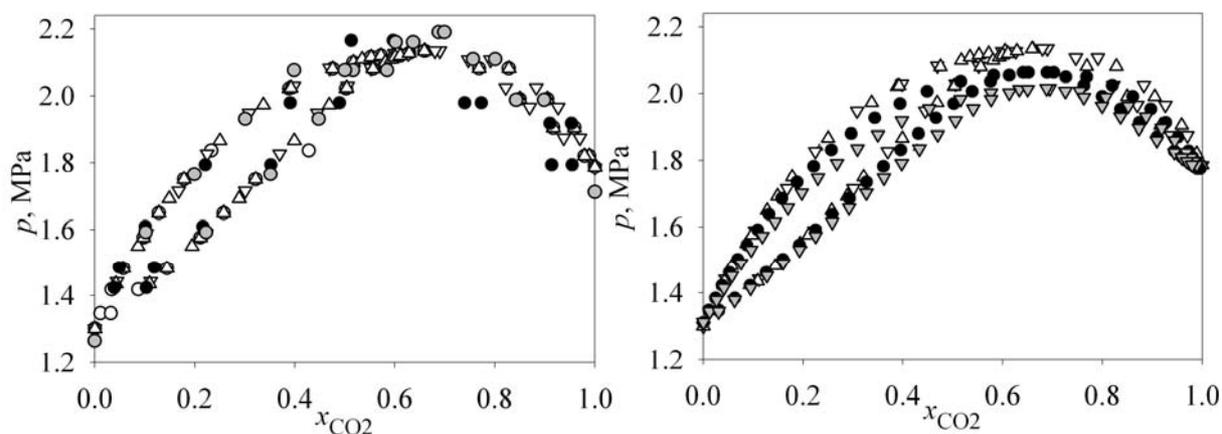


Figure 6.8: p - x -data of the binary mixture ethane/carbon dioxide at $T = 250 \text{ K}$. Open circles, open upward and downward triangles respectively represent experimental vapour-liquid equilibria published by Wei et al. [51], Brown et al. [57], and Davalos et al. [49]. Left-hand picture: solid black circles and solid grey circles are simulation results published by Liu and Beck [44] and Vrabcic and Fischer [43]. Right-hand picture: simulation results obtained in this work. Solid grey downward triangles represent simulation results obtained with LB combining rules, while solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules.

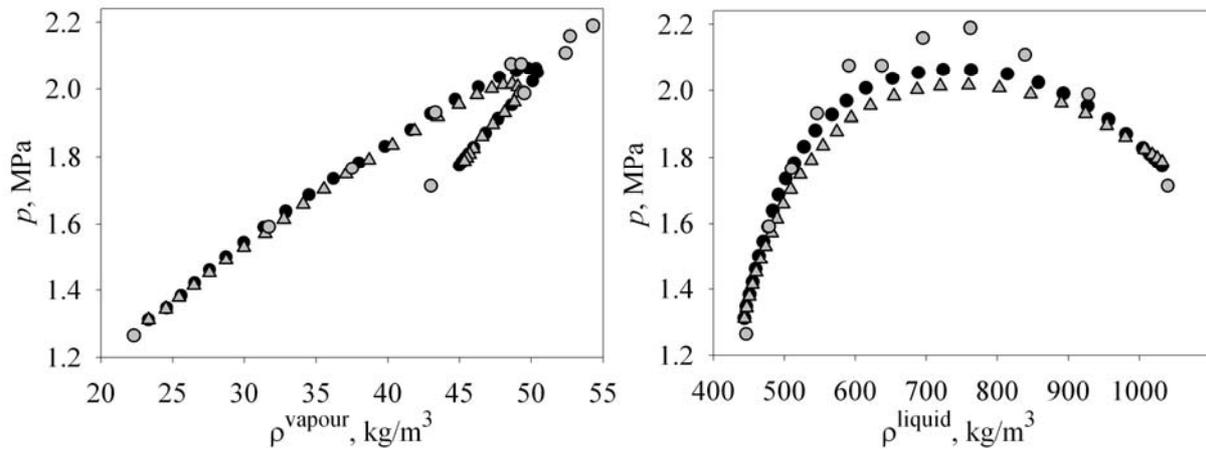


Figure 6.9: Liquid and vapour densities at coexistence of the binary mixture ethane/carbon dioxide at $T = 250$ K . Solid grey circles are simulation results published by Vrabc and Fischer [43]. Solid grey upward triangles represent simulation results obtained with LB combining rules, while solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Left-hand picture: vapour density. Right-hand picture: liquid density.

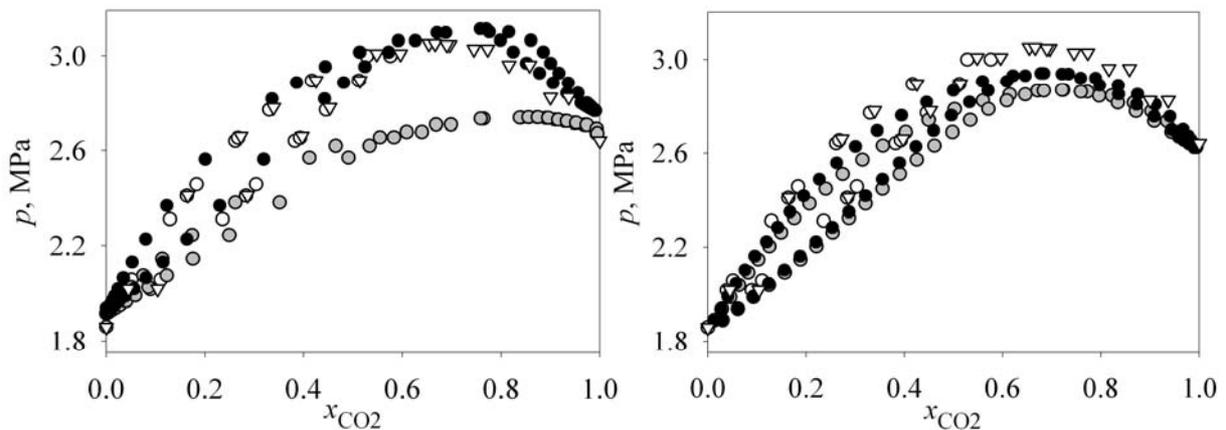


Figure 6.10: pxy-data of the binary mixture ethane/carbon dioxide at $T = 263.15$ K . Open circles, and open downward triangles respectively represent experimental vapour-liquid equilibria published by Brown et al. [57] and Fredenslund et al. [50]. Left-hand picture: solid black circles and solid grey circles are simulation results by Potoff et al. [25]. Solid grey circles are simulated with LB combining rules and solid black circles with LB combining rules for interaction between like molecules and Kong combining rules for interactions between unlike molecules. Right-hand picture: simulation results obtained in this work. Solid grey circles represent simulation results obtained with LB combining rules, while solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules.

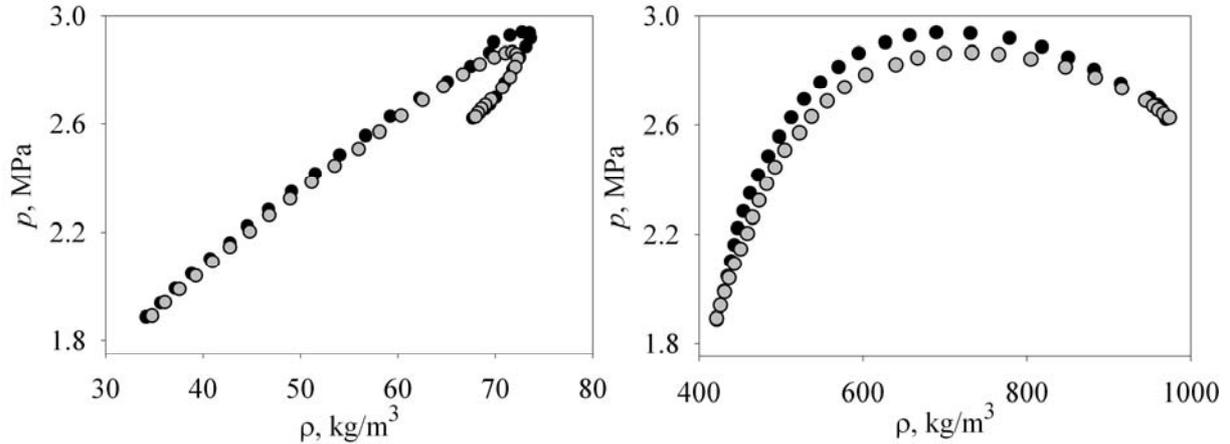


Figure 6.11: Liquid and vapour densities at coexistence of the binary mixture ethane/carbon dioxide at $T = 263.15$ K. Solid grey circles represent simulation results obtained with LB combining rules, while solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Left-hand picture: vapour density. Right-hand picture: liquid density.

The simulation results obtained at $T = 263.15$ K qualitatively agree with their experimental counterparts. Like the results obtained at $T = 207$ K and $T = 250$ K, the predicted saturation pressures at $T = 263.15$ K are systematically lower than the experimental values.

The results published by Potoff *et al.* [25], obtained with the LB combining rules for the interactions between like molecules and the Kong combining rules for the interactions between unlike molecules, have slightly higher saturation pressure than the experimental coexistence points have. On the other hand, their predictions with LB combining rules for all interactions have far too low saturation pressures. Notice that the predicted coexistence curves *with* and *without* Kong combining rules end up in rather different saturation pressures of pure carbon dioxide. This phenomenon is strange as the pure-component interactions for both series of simulations solely depend on the LB combining rules.

6.6.5. Binary system difluoromethane/carbon dioxide

Difluoromethane is considered to be an alternative refrigerant (R32). Thermophysical properties of alternative refrigerants, including mixtures of hydrofluorocarbons, are needed to develop refrigeration cycles. Simulation studies can be of help to predict those properties.

Sources of experimental data for vapour-liquid equilibria of the binary system difluoromethane/carbon dioxide are given in Table 6.5. This binary mixture has not been investigated before with molecular simulation.

Table 6.5: Sources of experimentally determined vapour-liquid equilibria of the binary system difluoromethane/carbon dioxide

Reference	Temp. range, K	Press. range, MPa	Number of sets
Rivollet <i>et al.</i> [63]	283.15-343.23	1.1-6.6	8
Diefenbacher and Türk [64]	280, 295, 310	1.7-6.4	3
Adams and Stein [65]	222-283	0.1-4.5	4

As pointed out in Chapter 5, the force field of Higashi and Takada [66] is used for difluoromethane and the TraPPE-AA potential [23] for carbon dioxide.

In this work, the vapour-liquid equilibria of the binary mixture at $T = 283.15$ K and $T = 303.15$ K were predicted. Every simulation was performed twice. The first simulation was performed with the LB combining rules for all interactions, while in the second case the LB combining rules for the interactions between like components and the LHMCC combining rules for the interactions between unlike components were used.

Simulation results for the initial slope of the Clapeyron equation in the limit of pure difluoromethane are given in Table 6.6. Simulation results for the initial points were taken from Chapter 5, Table 5.7.

Table 6.6: Simulated ratios of solvent's (difluoromethane) fugacity over solute's (carbon dioxide) Henry's constant

T K	p MPa	$\left(\frac{f_{\text{CH}_2\text{F}_2}}{H_{\text{CO}_2}}\right)_{\text{L}}^{\text{sim}}$	$\left(\frac{f_{\text{CH}_2\text{F}_2}}{H_{\text{CO}_2}}\right)_{\text{V}}^{\text{sim}}$	Comb. rule unlike molecules	Simulation Method
283.15	1.008	0.167(1)	0.836(1)	LB	DM + OID
303.15	1.80	0.28(2)	0.77(1)	LB	DM + OID
283.15	1.008	0.21(2)	0.83(2)	LHMCC	DM + OID
303.15	1.80	0.29(1)	0.77(2)	LHMCC	DM + OID

Simulation results for the vapour-liquid equilibria of the binary mixture are presented in graphical form in Figure 6.12 and in Figure 6.13. The results are also presented in tabular form in Appendix D, Table D.4.

The simulation results performed with both sets of combining rules qualitatively agree with the experimental data. Unfortunately, it is not clear which set of combining rules performs best. From Figure 6.12 and Figure 6.13, it is concluded that the results for different combining rules differ only slightly. The simulation results obtained with the LB combining rules for all interactions correctly predict the bubble-point pressures but underestimate the dew-point pressures. On the other hand, the results obtained with the LHMCC combining rules for the interactions between unlike molecules overestimate the bubble-point pressures and predict the dew-point pressures fairly well.

The simulations for this binary mixture were performed with a limited number of MC cycles per coexistence point in order to compensate the computational burden of the all-atom difluoromethane force field. A direct consequence of the short simulations is the scattering in the predicted liquid-phase densities.

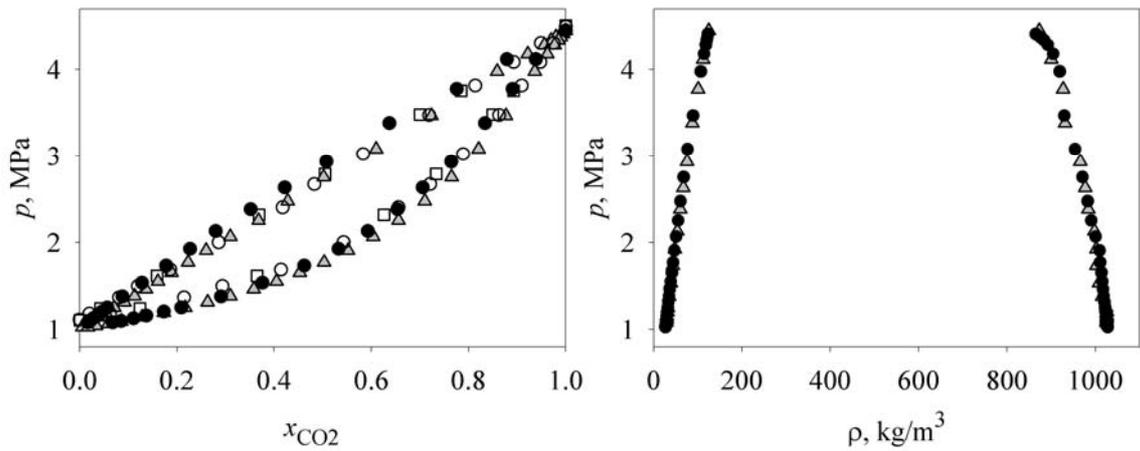


Figure 6.12: Vapour-liquid equilibria of the binary mixture difluoromethane/carbon dioxide at $T = 283.15 \text{ K}$. Open circles and open squares respectively represent experimental data published by Adams and Stein [65] and Rivollet et al. [63]. Solid grey upward triangles represent simulation results obtained with LB combining rules, while solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Left-hand picture: p - x_{CO_2} -data. Right-hand picture: liquid and vapour densities at coexistence.

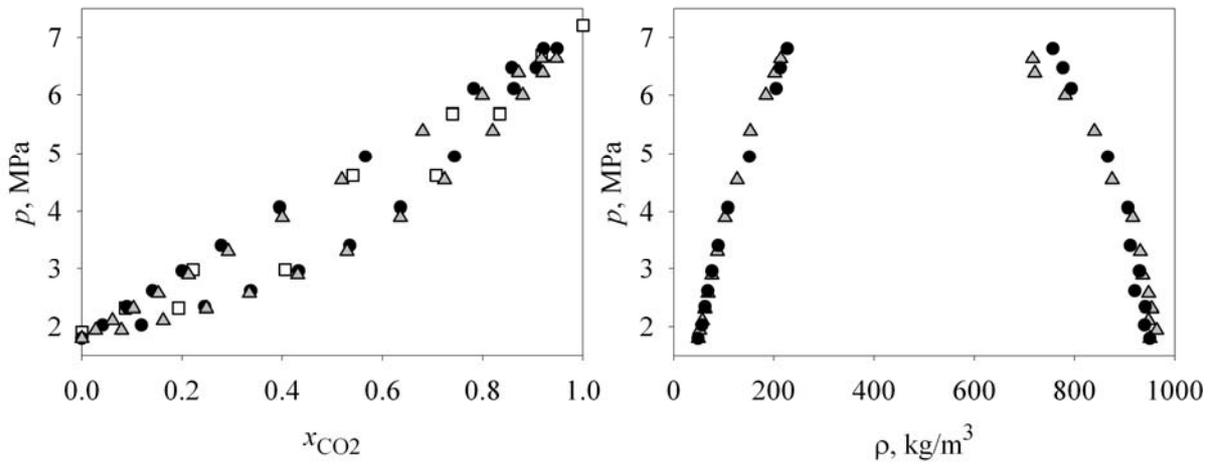


Figure 6.13: Vapour-liquid equilibria of the binary mixture difluoromethane/carbon dioxide at $T = 303.15 \text{ K}$. Open squares represent experimental data from Adams and Stein [65]. Solid grey upward triangles represent simulation results obtained with LB combining rules, while solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Left-hand picture: p - x_{CO_2} -data. Right-hand picture: liquid and vapour densities at coexistence.

6.6.6. Binary system trifluoromethane/carbon dioxide

Like difluoromethane, trifluoromethane is considered to be an alternative refrigerant (R23). Sources of experimental data for vapour-liquid equilibria of the binary system trifluoromethane /carbon dioxide are given in Table 6.7.

Table 6.7: Sources of experimentally determined vapour-liquid equilibria of the binary system trifluoromethane/carbon dioxide

Reference	Temp. range, K	Press. range, MPa	Numb. of sets
Roth et al. [67]	254.00-293.13	1.5-5.6	4
Mooijer-Van den Heuvel et al. [68]	263.15-302.61	2.0-7.0	7

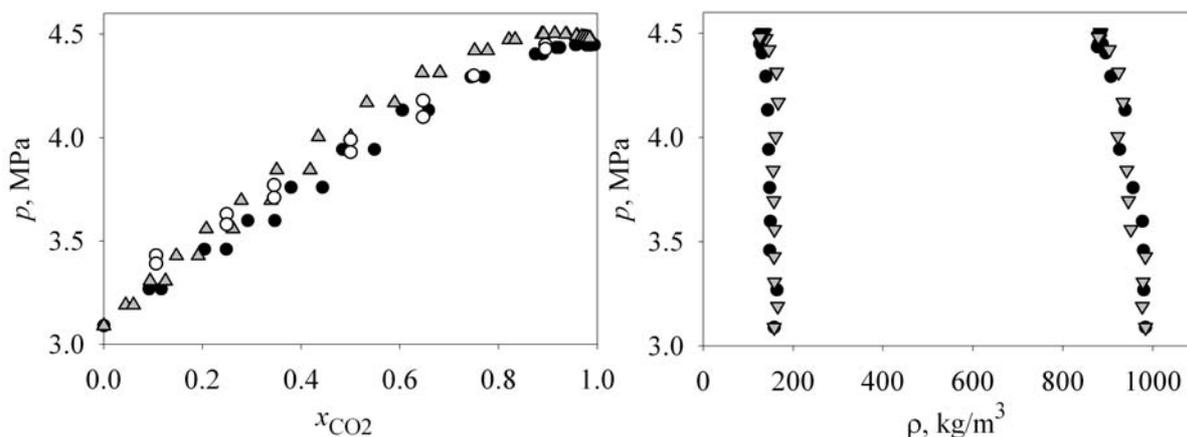


Figure 6.14: Vapour-liquid equilibria of the binary mixture trifluoromethane/carbon dioxide at $T = 283.25$ K. Open circles represent experimental data from Mooijer-Van den Heuvel *et al.* [68]. Solid black circles represent simulation results obtained with LB combining rules, while solid grey upward triangles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour densities at coexistence.

The vapour-liquid equilibria of this binary mixture have not been simulated before. In this work, simulations at $T = 283.25$ K were performed. As pointed out in Chapter 5, the force field of Song *et al.* [69] was used for trifluoromethane and the TraPPE-AA potential [23] for carbon dioxide. Simulation results for the vapour-liquid equilibria of the binary mixture are presented in graphical form in Figure 6.14. The results are also given in tabular form in Appendix D, Table D.5.

As the saturation pressures of the fluoroform force field of Song *et al.* [69] predict values that are low compared to experimental data, all dew and bubble point pressures on the fluoroform-side are systematically low. The results simulated with the LB combining rules for the interactions between like molecules and the LHMCC combining rules for the interactions between unlike molecules clearly overpredict the bubble and dew-point pressures. The results of the simulations performed with the LB combining rules for all interactions are closer to the experimental data.

It is expected that further optimization of the force-field parameters of fluoroform will improve the agreement between the binary simulation results and the experimental data, provided LB combining rules are used.

6.6.7. Binary system propane/carbon dioxide

Sources of experimentally determined vapour-liquid equilibria for this binary mixture are given in Table 6.8. Simulation results for this binary mixture are used in Chapter 7 as a starting point for the simulation of vapour-liquid equilibria in a ternary mixture.

The vapour-liquid equilibria of this system have been studied by Moon and Moon [70] at $T = 244.26$ K and $T = 266.48$ K with the $Np\beta$ -Gibbs ensemble. Carbon dioxide was modelled as a simple 2CLJ molecule [71] and propane as a rigid united-atom molecule [72]. The Lennard-Jones well-depth, ϵ_{ij} , for interactions between unlike molecules, included a large binary interaction parameter, $k_{\text{CO}_2\text{-C}_3\text{H}_8} = 0.1$. Their simulation results were in good agreement with experimental data. The large binary interaction parameter was needed to compensate for the simplicity of the exploited force field for carbon dioxide.

Table 6.8: Sources of experimentally determined vapour-liquid equilibria of the binary system propane/carbon dioxide

Reference	Temp. range, K	Press. range, MPa	Number of sets
Webster and Kidnay <i>et al.</i> [46]	230, 270	0.1-3.1	2
Akers <i>et al.</i> [73]	233.15-273.15	0.1-3.5	3
Hamam and Lu [53]	244.26, 266.48	0.5-2.6	2
Reamer <i>et al.</i> [74] ^a	277.59-344.26	0.5-6.9	5
Nagahama <i>et al.</i> [55]	250, 273.15	0.2-3.4	2
Poettmann and Katz [75] ^b	290.43-366.54	1.1-7.1	6

^a reference includes volumetric data, ^b pT measurements at fixed composition

Potoff *et al.* [25] used an exponential-6 potential for both carbon dioxide and propane. Either the LB combining rules for all interactions or the LB combining rules for the interactions between like molecules and the Kong combining rules for the interactions between unlike molecules were used. Their simulations were performed in the grand-canonical ensemble. Phase equilibria were computed from the multiple-histogram reweighting method.

Potoff and Siepmann [23] used experimental vapour-liquid equilibria of this binary mixture at $T = 294.26$ K to make the TraPPE-AA force field for carbon dioxide consistent with the TraPPE-EH force field for *n*-alkanes [See Chapter 5]. For this reason, it is not surprising that these force fields predict the vapour-liquid equilibria of this mixture, in particular at $T = 294.26$ K, very well.

In this work, simulation results at $T = 230$ K, $T = 270$ K and $T = 294.26$ K are presented. Like Potoff and Siepmann did, the TraPPE-AA force field [23] for carbon dioxide and the TraPPE-EH force field [40] for propane were used in this work. In contrast to the simulation study performed by Potoff and Siepmann, the LB combining rules for interactions between like molecules and the LHMcC combining rules for interactions between unlike molecules were exploited in this work.

Simulation results for the initial point (pure propane) and the initial slope of the Clapeyron equation in the limit of pure propane are given in Table 6.9. The saturation pressures of pure propane have been computed from volume perturbations in the Gibbs ensemble.

Table 6.9: Simulated ratios of solvent's (propane) fugacity over solute's (carbon dioxide) Henry's constant

T K	p^{sim} MPa	ρ_L^{sim} kg/m ³	ρ_V^{sim} kg/m ³	$\left(\frac{f_{\text{C3H8}}}{H_{\text{CO2}}}\right)_L^{\text{sim}}$	$\left(\frac{f_{\text{C3H8}}}{H_{\text{CO2}}}\right)_V^{\text{sim}}$	Comb. rule unlike molecules	Simulation Method
230	0.096(4)	576.1(4)	2.3(1)	0.029(4)	0.96(2)	LHMcC	DM + OID
270	0.432(9)	526.2(3)	9.3(2)	0.088(9)	0.90(1)	LHMcC	DM + OID
294.26	0.88(1)	493.7(8)	18.5(2)	0.093(1)	0.85(3)	LHMcC	DM + OID

Simulation results for the vapour-liquid equilibria of the binary mixture are presented in graphical form in Figure 6.15 to Figure 6.18. The results are also given in tabular form in Appendix D, Table D.6.

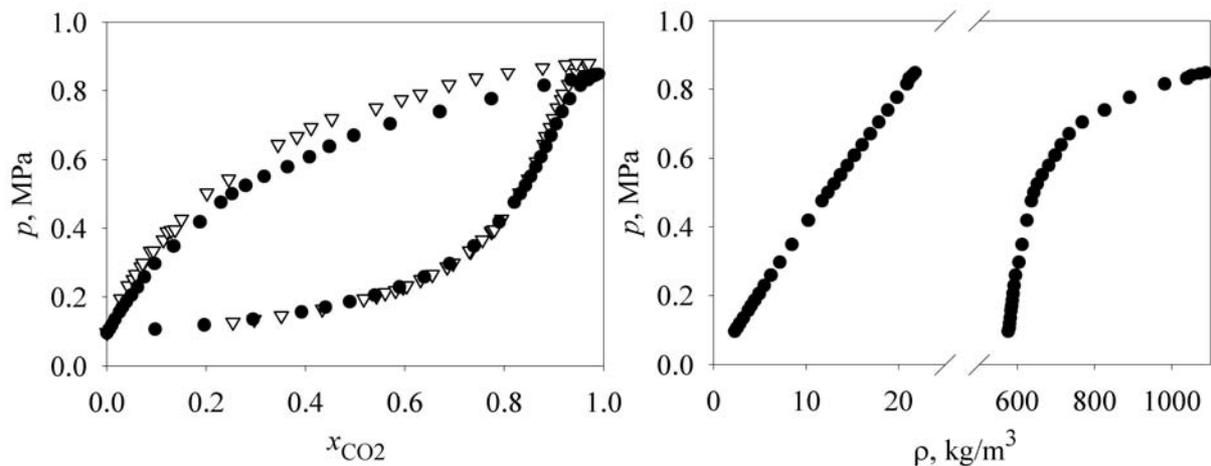


Figure 6.15: Vapour-liquid equilibria of the binary mixture propane/carbon dioxide at $T = 230\text{ K}$. Open downward triangles are experimental data from Webster and Kidnay [46]. Solid circles are simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour densities at coexistence.

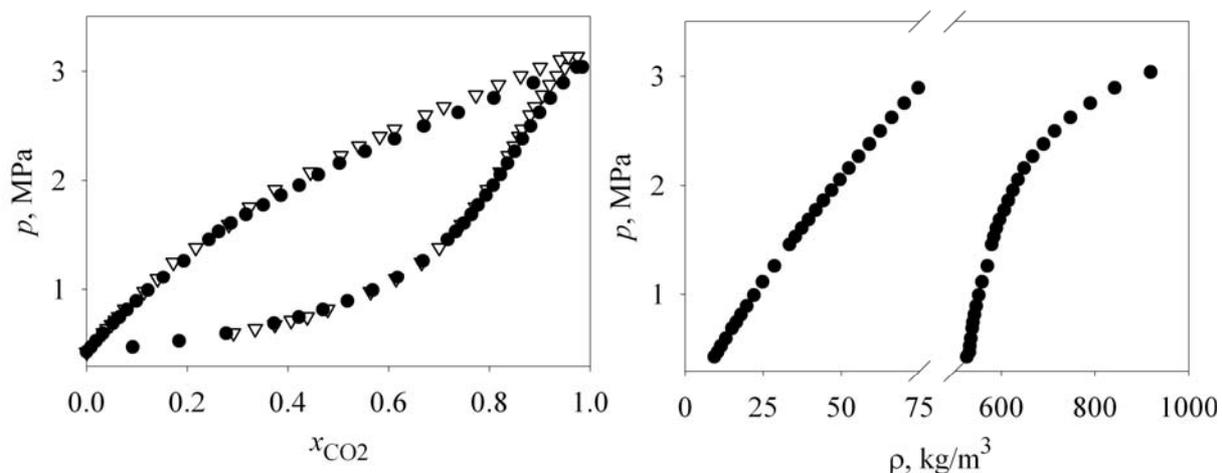


Figure 6.16: Vapour-liquid equilibria of the binary mixture propane/carbon dioxide at $T = 270\text{ K}$. Open downward triangles are experimental data from Webster and Kidnay [46]. Solid circles are simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour densities at coexistence.

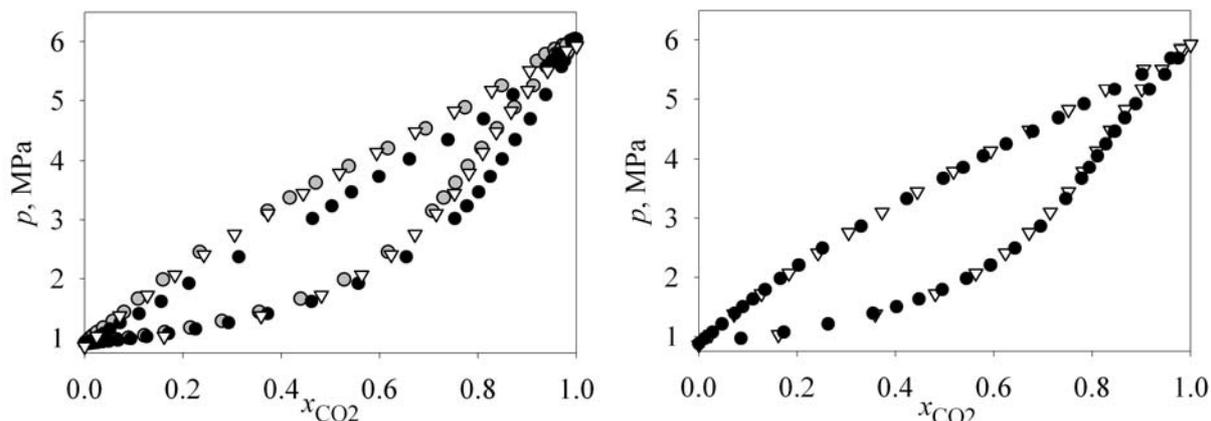


Figure 6.17: Vapour-liquid equilibria of the binary mixture propane/carbon dioxide at $T = 294.26$ K. Open downward triangles represent experimental vapour-liquid equilibria published by Reamer *et al.* [74]. Left-hand picture: solid black circles and solid grey circles are simulation results published by Potoff *et al.* [25]. Solid black circles were simulated with LB combining rules and solid grey circles with LB combining rules for interactions between like molecules and Kong combining rules for interactions between unlike molecules. Right-hand picture: solid circles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules.

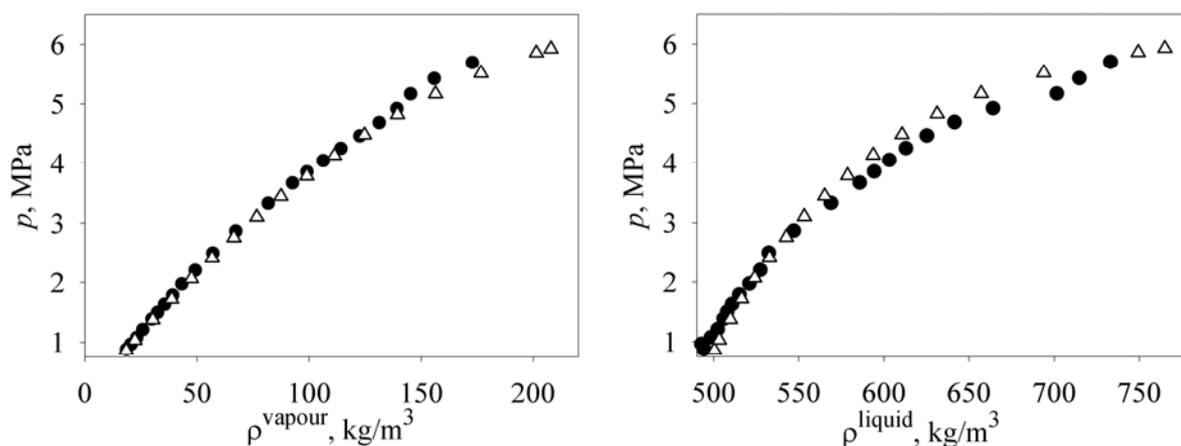


Figure 6.18: Liquid and vapour densities at coexistence of the binary mixture propane/carbon dioxide at $T = 294.26$ K. Open upward triangles represent experimental data published by Reamer *et al.* [74]. Solid circles represent simulation results obtained with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Left-hand picture: vapour density. Right-hand picture: liquid density.

From Figure 6.17, it is clear that the simulation results of Potoff *et al.* [25], performed with the LB combining rules for interactions between like molecules and the Kong combining rules for interactions between unlike molecules, agree very well with the experimental data. The simulation results of Potoff and Siepmann [23] at $T = 294.26$ K were not available in tabular form. However, from their simulation results in graphical form, it is obvious that the liquid and vapour-phase compositions were slightly too high [23].

The simulation results in this work at $T = 294.26$ K show even better agreement with experimental data than the results of Potoff *et al.* [25] and Potoff and Siepmann [23] do. From Figure 6.18, it is seen that the vapour and liquid densities at coexistence are also in good agreement with experimental values. Unfortunately, the simulated saturation pressures at $T = 230$ K and $T = 270$ K are systematically low at high mole fractions of carbon dioxide.

As the predictions of the binary vapour-liquid equilibria are qualitatively correct, further optimization of the force-field parameters of carbon dioxide to low-temperature data will improve the predictive power of the TraPPE force field.

6.6.8. Binary system DMSO/carbon dioxide

Details on this binary mixture have been given in Chapter 5. Here, simulation results obtained at $T = 278.5$ K and $T = 303.15$ K are presented. In contrast to the simulation results presented in Section 5.4.8, the LB combining rules for interactions between like molecules and the LHMCC combining rules for interactions between unlike molecules were used in this section. The ratios $f_{\text{DMSO}}/H_{\text{CO}_2}$, which are needed to start the numerical integration are given in Table 6.10. The saturation pressures of pure DMSO have been computed with the $Np\beta$ +TM method combined with overlapping distributions. DMSO is described by the force field of Rao and Singh [23], while carbon dioxide is modelled with the TraPPE-AA potential [23]. Simulation results for binary vapour-liquid equilibria at $T = 278.5$ K and $T = 303.15$ K are given in Appendix D, Table D.7. The results are also shown graphically in Figure 6.19 and in Figure 6.20.

Table 6.10: Simulated ratios of solvent's (DMSO) fugacity over solute's (carbon dioxide) Henry's constant

T K	p MPa	$\left(\frac{f_{\text{DMSO}}}{H_{\text{CO}_2}}\right)_L^{\text{sim}}$	$\left(\frac{f_{\text{DMSO}}}{H_{\text{CO}_2}}\right)_V^{\text{sim}}$	Comb. rule unlike molecules	Simulation Method
278.15	0.0017(1)	0.00047(17)	1.00(14)	LHMCC	DM + OID
303.15	0.0063(5)	0.00046(11)	0.95(8)	LHMCC	DM + OID

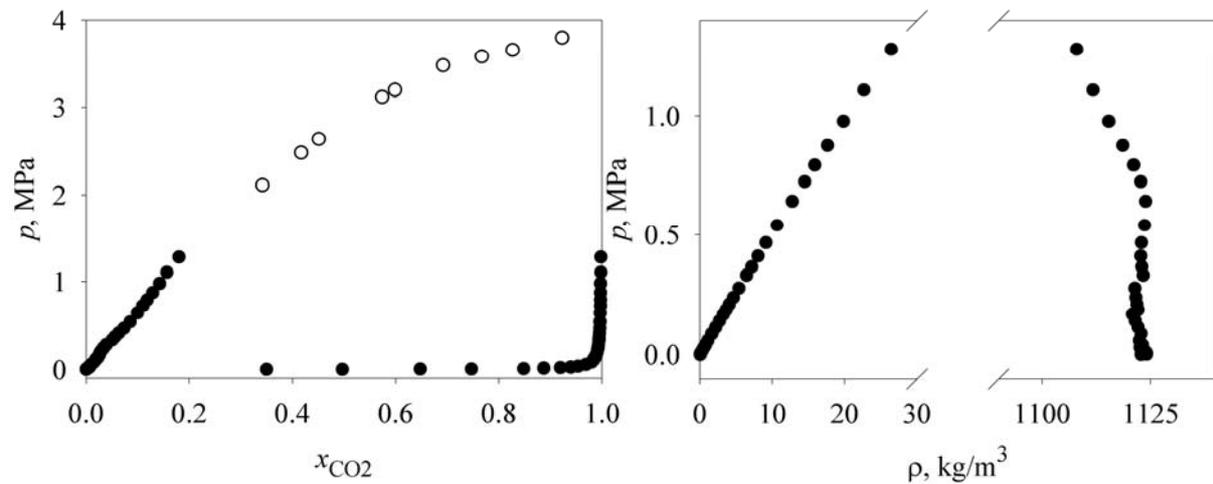


Figure 6.19: Vapour-liquid equilibria of the binary mixture DMSO/carbon dioxide at $T = 278.5$ K . Open circles represent experimental bubble-point data from Florusse and Peters [76]. Solid circles are simulation results computed with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Left-hand picture: p - x -data. Right-hand picture: liquid and vapour densities at coexistence.

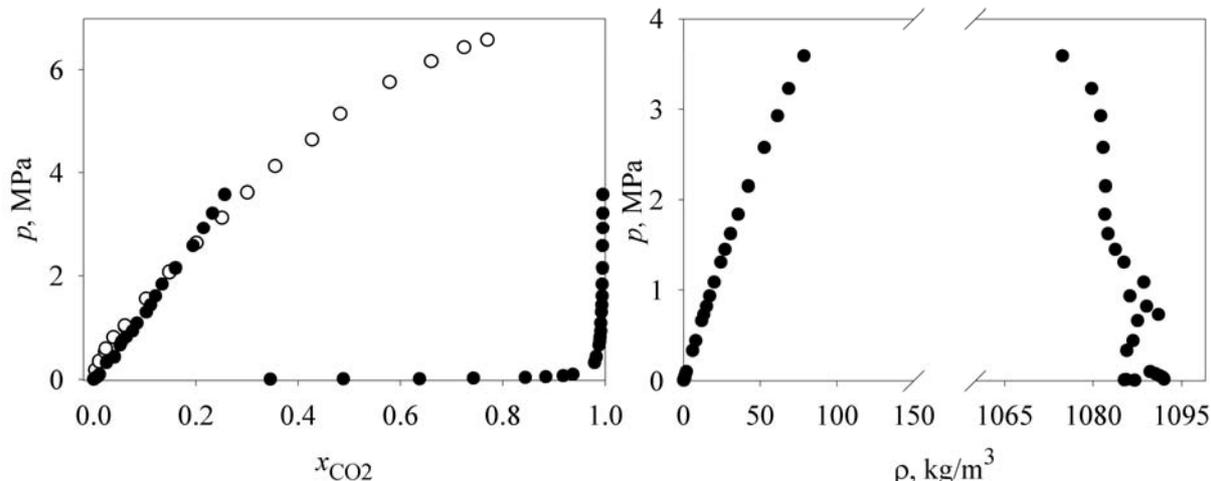


Figure 6.20: Vapour-liquid equilibria of the binary mixture DMSO/carbon dioxide at $T = 303.15$ K. Open circles represent experimental bubble-point data published by Kordikowski et al. [77]. Solid black circles are simulation results computed with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Left-hand picture: p - x_{CO_2} -data. Right-hand picture: liquid and vapour densities at coexistence.

In Chapter 5, Section 5.4.8, simulation results for this binary mixture at $T = 303.15$ K were presented. Those simulations were performed with the LB combining rules for all interactions. A comparison of the performance of different combining rules is difficult as the results given in Chapter 5 have large statistical uncertainties.

The simulation results presented in this section have been obtained by performing energy-biased trial identity changes. This bias method improved sampling of the composition considerably. The enhanced composition sampling, along with the retroactive nature of advanced Gibbs-Duhem integration, results in p - x_{CO_2} -predictions that are much smoother than those computed in Chapter 5. However, as can be seen from the odd course of the liquid-phase densities, sampling of volumes is not satisfactory. Another origin of the odd course might be a bad estimate for the initial coexistence point or for the initial slope.

Simulations at higher pressures/fugacity fractions than the ones given in Figure 6.19 and Figure 6.20 were not possible, as the slope of the Clapeyron equation changed very rapidly ($|\partial^2 p / \partial \xi_2^2|$ is large) at fugacity fractions close to one.

6.7. Error analysis of advanced Gibbs-Duhem integration

A thorough error analysis of the advanced Gibbs-Duhem integration method is difficult. A relatively short simulation and/or a narrow range of sampled fluctuating variables reduce the extrapolative power of histogram reweighting. However, in this work it is checked whether histograms at nearby thermodynamic state points overlap sufficiently. Furthermore, interpolation between histograms (state points) is performed instead of extrapolation to regions where no data have been collected. This means that the poorly sampled *tails of the histograms* are less important. For this reason, the errors that are specific for histogram reweighting [8, 9] are neglected.

Errors arising from the use of finite-order corrector equations are neglected as well. The influence of an incorrect initial coexistence point on the deviation of the simulated

coexistence line from the true coexistence line has already been given in Chapter 5 and will not be discussed here.

The integrands of the Clapeyron equation obtained at successive coexistence points are computed using Monte Carlo simulation. It is assumed that the integration variable is the fugacity fraction. The ensemble averages of the integrands have a statistical uncertainty. As the successive estimates of the saturation pressures are coupled via corrector equations, the variance of the saturation pressure at a particular thermodynamic state point depends on the variances of formerly determined saturation pressures/integrands. From Chapter 5, it is known that the pressure vector \mathbf{p} (element i is the estimated saturation pressure at thermodynamic state point i) can be written as [See Eq. 5.21]:

$$\mathbf{p} = [\mathbf{\Omega}]\mathbf{F} + p_0 \quad (6.38)$$

where the vector \mathbf{F} contains the integrands of the Clapeyron equation at successive integration steps, p_0 is the saturation pressure corresponding to the estimate of the initial coexistence point, and $[\mathbf{\Omega}]$ is a matrix that contains parameters from the corrector equations. The variance of the saturation pressure at state point n is given by:

$$\sigma_{p_n}^2 = \mathbf{\Omega}_n \mathcal{Z}(\mathbf{F}) \mathbf{\Omega}_n^T \quad (6.39)$$

where $\mathcal{Z}(\mathbf{F})$ is the variance-covariance matrix of the integrands, and $\mathbf{\Omega}_n$ is the n^{th} row vector of the matrix $[\mathbf{\Omega}]$. The integrands at successive integration points are statistically dependent as they are computed from a degeneracy that is optimized to samples collected at different state points. Thus, the covariances in $\mathcal{Z}(\mathbf{F})$ are generally non-zero. The version of the Clapeyron equation that is numerically integrated in this chapter is:

$$\left(\frac{\partial \ln p}{\partial \xi_2} \right)_{\beta, \sigma} = F(\xi_2, p) \quad (6.40)$$

This means that the saturation pressure in Eqs. (6.38) and (6.40) is to be substituted by $\ln p$. The variance in the saturation pressure can be computed from the error propagation law:

$$\sigma_{p_i}^2 = \sigma_{\ln p_i}^2 \left(\frac{\partial p_i}{\partial \ln p_i} \right)^2 = p_i^2 \sigma_{\ln p_i}^2 \quad (6.41)$$

The variances and covariances in the matrix $\mathcal{Z}(\mathbf{F})$ are computed as follows. Instead of collecting samples at state point i into a single histogram, samples are collected on a regular basis into M subhistograms. These subhistograms are more or less statistically independent. Ensemble averages are obtained from application of the multiple-histogram reweighting method to the sum of the subhistograms:

$$\mathcal{H}_i(N_2, V) = \sum_{j=1}^M \mathcal{H}_{ij}(N_2, V; \beta, p_i, [\xi_2]_i) \quad (6.42)$$

The variance and the covariance of the integrands and of other ensemble averages are computed as follows:

1. From the M subhistograms collected at thermodynamic state point i , M subhistograms are picked randomly, not trying to avoid multiple sampling. The sum of these selected subhistograms forms a bootstrap sample of the histogram at state point i , $\mathcal{H}_i^*(N_2, \mathcal{U}, V)$. Bootstrap-sample histograms are needed for both the liquid and the vapour phase.
2. Repeat step 1. for every state point that takes part in the multiple-histogram reweighting process.
3. Perform multiple-histogram reweighting with the bootstrap samples of the histograms.
4. Application of the corrector equations to the optimized bootstrap sample probabilities $\pi^*(N_2, V, \mathcal{U}; \xi_2, p)$ for the liquid and the vapour phase provides bootstrap samples of the saturation pressures p_i^* , densities $\{\rho_i^{*, \text{liq}}, \rho_i^{*, \text{vap}}\}$, mole fractions $\{[x_2^{*, \text{liq}}]_i, [x_2^{*, \text{vap}}]_i\}$, and integrands F_i^* at coexistence [See Section 6.4].
5. Repeat steps 1. to 4. a large number of times, say N_B times, and compute the bootstrap estimate of the variance of variable Y_i at thermodynamic state point i and the covariance between variables Y_i and Y_k at thermodynamic state points i and k from:

$$\sigma_{Y_i}^2 = \frac{1}{N_B - 1} \sum_{j=1}^{N_B} (Y_{ij}^* - \langle Y_i \rangle)^2 \quad (6.43)$$

$$\sigma_{Y_i Y_k}^2 = \frac{1}{N_B - 1} \sum_{j=1}^{N_B} (Y_{ij}^* - \langle Y_i \rangle)(Y_{kj}^* - \langle Y_k \rangle) \quad (6.44)$$

Although it is possible to compute asymmetric errors for the computed ensemble averages, this has not been performed in this work. The propagation of the statistical uncertainties via the corrector equations in the binary mixture ethane/methane at $T = 193.27$ K is given in Table 6.11.

Table 6.11: Propagation of statistical errors; standard deviation of predicted saturation pressures at successive integration steps in the binary mixture methane/ethane at $T = 193.27$ K

p^{sim} MPa	$\sigma_{p^{\text{sim}}}$ MPa								
0.18	0.09	0.39	0.09	1.13	0.09	2.12	0.11	3.67	0.18
0.20	0.09	0.44	0.09	1.21	0.10	2.25	0.11	3.78	0.21
0.23	0.09	0.51	0.09	1.30	0.10	2.41	0.11	3.94	0.28
0.25	0.09	0.58	0.09	1.41	0.10	2.58	0.11	4.11	0.38
0.26	0.09	0.64	0.09	1.54	0.10	2.78	0.11		
0.29	0.09	0.73	0.09	1.70	0.10	3.00	0.12		
0.31	0.09	0.84	0.09	1.88	0.11	3.25	0.12		
0.35	0.09	0.99	0.09	1.99	0.11	3.46	0.13		

As is clear from Table 6.11, it is the statistical uncertainty of the estimate of the initial saturation pressure that mainly determines the statistical error of the saturation pressures at successive coexistence points. Thus, in case the statistical error of the initial saturation pressure is small, the statistical error of the saturation pressure after a number of integration steps is probably small as well. The errors for the advanced GDI method are slightly larger than those for the conventional GDI method. However, the errors due to the use of a finite-order corrector equation (not investigated in this work) will be smaller for the advanced GDI method as subsequent simulations improve formerly predicted coexistence points.

6.8. Conclusions

In this chapter, the single and multiple-histogram reweighting methods were discussed. An advanced version of Gibbs-Duhem integration, which is a combination of conventional Gibbs-Duhem integration and multiple-histogram reweighting, was introduced. The application of the advanced Gibbs-Duhem integration method depends on the ability to compute the degeneracy and semigrand-canonical probability distribution from histograms collected at different state points. The relations for the degeneracy and semigrand-canonical probability distribution were derived.

The advantages of the advanced Gibbs-Duhem integration method over the conventional method are the fixed number of simulations and the retroactive improvement of formerly computed coexistence points. This retroactive correction results in smooth simulation results. The corrector equation can be used to compute a whole coexistence line from a limited number of simulations. Furthermore, it is possible to search for the coexistence point belonging to a particular liquid-phase or vapour-phase mole fraction or density.

The advanced Gibbs-Duhem integration method, combined with parallel-tempering configuration swaps, has been used to simulate the vapour-liquid equilibria of a number of binary mixtures: ethane/methane, propane/methane, ethane/carbon dioxide, difluoromethane/carbon dioxide, trifluoromethane/carbon dioxide, propane/carbon dioxide, and DMSO/carbon dioxide.

In general, the simulations with the advanced Gibbs-Duhem integration method performed much faster than those performed with the conventional Gibbs-Duhem integration method. This increase in simulation speed is due to the fixed number of simulations opposed to the free number of corrector simulations in the conventional Gibbs-Duhem integration method.

The use of different combining rules for interactions between like and unlike molecules considerably improved the match between experimental data and simulation results for several binary mixtures. However, in two cases, it was not clear which set of combining rules performed best. Nevertheless, on average, the combination of the Lorentz-Berthelot combining rules for the interactions between like molecules and the Lorentz-Hudson-McCoubrey combining rules for the interactions between unlike molecules, performed better than simply exploiting the Lorentz-Berthelot combining rules for all interactions.

Although the TraPPE force field had been optimized for interactions between carbon dioxide and linear alkanes, it did not reproduce the experimental saturation pressures of the binary system ethane/carbon dioxide. Thus, further optimization of force-field parameters to other *n*-alkane/carbon dioxide systems is needed. In the systems propane/carbon dioxide and trifluoromethane/carbon dioxide, it is not the combining rule that causes deviations from experimental vapour-liquid equilibria but it is the force field of one of the pure components; the agreement between the simulation results and experimental vapour-liquid equilibria can be improved by further optimization of pure-component force-field parameters of carbon dioxide and trifluoromethane.

6.9. References

1. Ferrenberg, A. M. and Swendsen, R. H., New Monte Carlo Technique for Studying Phase Transitions, *Phys. Rev. Lett.*, 61(23) **1988** 2635
2. Ferrenberg, A. M. and Swendsen, R. H., Optimized Monte Carlo Data Analysis, *Computers in Physics*, sep/oct **1989** 101
3. Valleau, J. P. and Card, D. N., Monte Carlo Estimation of the Free Energy by Multistage Sampling, *J. Chem. Phys.*, 57(12) **1972** 5457
4. McDonald, I. R. and Singer, K., Calculation of Thermodynamic Properties of Liquid Argon from Lennard-Jones Parameters by a Monte Carlo Method, *Discuss. Faraday Soc.*, 43 **1967** 40
5. Chesnut, D. A. and Salsburg, Z. W., Monte Carlo Procedure for Statistical Mechanical Calculations in a Grand Canonical Ensemble of Lattice Systems, *J. Chem. Phys.*, 38(12) **1963** 2861
6. Salsburg, Z. W., Jacobson, J. D., Fickett, W., and Wood, W. W., Application of the Monte Carlo method to the lattice gas model. 1. Two-dimensional triangular lattice, *J. Chem. Phys.*, 30(1) **1959** 65
7. Ferrenberg, A. M. and Landau, D. P., Critical behavior of the three-dimensional Ising model: A high-resolution Monte Carlo study, *Phys. Rev. B*, 44(10) **1991** 5081
8. Ferrenberg, A. M., Landau, D. P., and Swendsen, R. H., Statistical errors in histogram reweighting, *Phys. Rev. E*, 51(5) **1995** 5092
9. Newman, M. E. J. and Palmer, R. G., Error estimation in the histogram Monte Carlo method, *J. Stat. Phys.*, 97(5-6) **1999** 1011
10. Swendsen, R. H., Modern methods of analyzing Monte Carlo computer simulations, *Physica A*, 194 **1993** 53
11. McDonald, I. R., NpT -ensemble Monte Carlo calculations for binary liquid mixtures, *Mol. Phys.*, 23(1) **1972** 41
12. Ferrenberg, A. M. and Swendsen, R. H., Optimized Monte Carlo Data Analysis, *Phys. Rev. Lett.*, 63(12) **1989** 1195
13. Ferguson, D. M. and Garrett, D. G., Simulated annealing - Optimal histogram methods. In: Ferguson, D. M., Siepmann, J. I., and Truhlar, D. G., editors, *Advances in Chemical Physics, Volume 105, Monte Carlo Methods in Chemistry*, Wiley, 1999
14. Corti, D. S. and Soto-Campos, G., Deriving the isothermal-isobaric ensemble: The requirement of a "shell" molecule and applicability to small systems, *J. Chem. Phys.*, 108(19) **1998** 7959
15. Corti, D. S., Isothermal-isobaric ensemble for small systems, *Phys. Rev. E*, 64(1) **2001** 016128
16. Corti, D. S., Monte Carlo simulations in the isothermal-isobaric ensemble: the requirement of a 'shell' molecule and simulations of small systems, *Mol. Phys.*, 100(12) **2002** 1887
17. Borgs, C. and Kappler, S., Equal weight versus equal height: a numerical study of an asymmetric first-order transition, *Phys. Lett. A*, 171(1-2) **1992** 37
18. Wilding, N. B., Computer simulation of fluid phase transitions, *Am. J. Phys.*, 69(11) **2001** 1147
19. Berg, B. A. and Neuhaus, T., Multicanonical algorithms for first order phase transitions, *Phys. Lett. B*, 267 **1991** 249
20. Berg, B. A., Multicanonical simulations step by step, *Comp. Phys. Comm.*, 153 **2003** 397

-
21. Berg, B. A. and Neuhaus, T., Multicanonical ensemble - A new approach to simulate 1st-order phase transitions, *Phys. Rev. Lett.*, 68(1) **1992** 9
 22. Gelb, L. D. and Gubbins, K. E., Studies of binary liquid mixtures in cylindrical pores: phase separation, wetting and finite-size effects from Monte Carlo simulations, *Physica A*, 244 **1997** 112
 23. Potoff, J. J. and Siepmann, J. I., Vapor-Liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen, *AIChE J.*, 47(7) **2001** 1676
 24. Kim, S.-Y., Panagiotopoulos, A. Z., and Floriano, M. A., Ternary oil-water-amphiphile systems: self-assembly and phase equilibria, *Mol. Phys.*, 100(14) **2002** 2213
 25. Potoff, J. J., Errington, J. R., and Panagiotopoulos, A. Z., Molecular Simulation of Phase Equilibria for Mixtures of Polar and Non-Polar Components, *Mol. Phys.*, 97(10) **1999** 1073
 26. Shi, W. and Johnson, J. K., Histogram reweighting and finite-size scaling study of the Lennard-Jones fluids, *Fluid Phase Equil.*, 187-188 **2001** 171
 27. Kiyohara, K., Gubbins, K. E., and Panagiotopoulos, A. Z., Phase coexistence properties of polarizable water models, *Mol. Phys.*, 94(5) **1998** 803
 28. Errington, J. R., Kiyohara, K., Gubbins, K. E., and Panagiotopoulos, A. Z., Monte Carlo simulation of high-pressure phase equilibria in aqueous systems, *Fluid Phase Equil.*, 150-151 **1998** 33
 29. Kiyohara, K., Gubbins, K. E., and Panagiotopoulos, A. Z., Phase coexistence properties of polarizable Stockmayer fluids, *J. Chem. Phys.*, 106(8) **1997** 3338
 30. Shi, W., Zhao, X., and Johnson, J. K., Phase transitions of adsorbed fluids computed from multiple-histogram reweighting, *Mol. Phys.*, 100(13) **2002** 2139
 31. Conrad, P. B. and de Pablo, J. J., Comparison of histogram reweighting techniques for a flexible water model, *Fluid Phase Equil.*, 150-151 **1998** 51
 32. Chang, J. and Sandler, S. I., Determination of liquid-solid transition using histogram reweighting method and expanded ensemble simulations, *J. Chem. Phys.*, 118(18) **2003** 8390
 33. Meijer, E. J. and El Azhar, F., Novel procedure to determine coexistence lines by computer simulation. Application to hard-core Yakawa model for charge-stabilized colloids, *J. Chem. Phys.*, 106(11) **1997** 4678
 34. Bennett, C. H., Efficient Estimation of Free Energy Differences from Monte Carlo Data, *J. Comput. Phys.*, 22 **1976** 245
 35. Escobedo, F. A. and de Pablo, J. J., Pseudo-ensemble simulations and Gibbs-Duhem integrations for polymers, *J. Chem. Phys.*, 106(7) **1997** 2911
 36. Escobedo, F. A., Tracing coexistence lines in multicomponent fluid mixtures by molecular simulation, *J. Chem. Phys.*, 110(24) **1999** 11999
 37. Escobedo, F. A., Novel pseudoensembles for simulation of multicomponent phase equilibria, *J. Chem. Phys.*, 108(21) **1998** 8761
 38. Frenkel, D. and Smit, B., *Understanding molecular simulation; from algorithms to applications*, Computational Science Series, Volume 1, 2nd Ed., Academic press, London, 2002
 39. Potoff, J. J. and Panagiotopoulos, A. Z., Critical point and phase behavior of the pure fluid and a Lennard-Jones mixture, *J. Chem. Phys.*, 109(24) **1998** 10914
 40. Chen, B. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, *J. Phys. Chem. B*, 103 **1999** 5370
 41. Davalos, J., Anderson, W. R., Phelps, R. E., and Kidnay, A. J., Liquid-Vapor Equilibria at 250.0 K for Systems Containing Methane, Ethane, and Carbon Dioxide, *J. Chem. Eng. Data*, 21(1) **1976** 81

-
42. Wei, W., Brown, T. S., Kidnay, A. J., and Sloan, D. E., Vapor + liquid equilibria for the ternary system methane + ethane + carbon dioxide at 230 K and its constituent binaries at temperatures from 207 K to 270 K, *J. Chem. Eng. Data*, 40 **1995** 726
 43. Vrabec, J. and Fischer, J., Vapor-Liquid Equilibria of Binary Mixtures Containing Methane, Ethane, and Carbon Dioxide from Molecular Simulation, *Int. J. Thermophys.*, 17(4) **1996** 889
 44. Liu, A. and Beck, Th. L., Vapor-Liquid Equilibria of Binary and Ternary Mixtures Containing Methane, Ethane, and Carbon Dioxide from Gibbs Ensemble Simulations, *J. Phys. Chem. B*, 102 **1998** 7627
 45. Zhang, Z. and Duan, Z., Phase equilibria of the system methane-ethane from temperature scaling Gibbs Ensemble Monte Carlo simulation, *Geochim. Cosmochim. Acta*, 66(19) **2002** 3431
 46. Webster, L. A. and Kidnay, A. J., Vapor-Liquid Equilibria for the Methane-Propane-Carbon Dioxide Systems at 230 K and 270 K, *J. Chem. Eng. Data*, 46 **2001** 759
 47. Chen, B. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen description of normal alkanes, *J. Phys. Chem. B*, 103 **1999** 5370
 48. Rowlinson, J. S., *Liquids and liquid mixtures*, 2nd Ed., Butterworth & Co, Belfast, 1969
 49. Davalos, J., Anderson, W. R., Phelps, R. E., and Kidnay, A. J., Liquid-Vapor Equilibria at 250.00 K for Systems Containing Methane, Ethane, and Carbon Dioxide, *J. Chem. Eng. Data*, 21(1) **1976** 81
 50. Fredenslund, A. and Mollerup, J., Measurement and Prediction of Equilibrium Ratios for the C₂H₆+CO₂ System, *J. Chem. Soc. Far. Trans. I*, 70 **1974** 1653
 51. Wei, M. S. W., Brown, T. S., Kidnay, A. J., and Sloan, D. E., Vapor + liquid equilibria for the ternary system methane + ethane + carbon dioxide at 230 K and its constituent binaries at temperatures from 207 to 270 K, *J. Chem. Eng. Data*, 40 **1995** 726
 52. Price, A. R., PhD Thesis, *Rice Inst. Houston, Texas* **1957**
 53. Hamam, S. E. M. and Lu, B. C. Y., Vapor-Liquid Equilibrium in the Ethane-Carbon Dioxide System, *Can. J. Chem. Eng.*, 52 **1974** 283
 54. Ohgaki, K. and Katayama, T., Isothermal vapor-liquid equilibrium data for the ethane-carbon dioxide system at high pressures, *Fluid Phase Equil.*, 1 **1977** 27
 55. Nagahama, K., Konishi, H., Hoshino, D., and Hirata, M., Binary vapor-liquid equilibria of carbon dioxide-light hydrocarbons at low temperature, *J. Chem. Eng. Jap.*, 7(5) **1974** 323
 56. Gugnoni, R. J., Eldridge, J. W., Okay, V. C., and Lee, T. J., Carbon Dioxide-Ethane Phase Equilibrium and Densities from Experimental Measurements and the B-W-R Equation, *AIChE J.*, 20(2) **1974** 357
 57. Brown, T. S., Kidnay, A. J., and Sloan, D. E., Vapor-liquid equilibria in the carbon dioxide-ethane system, *Fluid Phase Equil.*, 40 **1988** 169
 58. Lofti, A., Vrabec, J., and Fischer, J., Vapour liquid equilibria of the Lennard-Jones fluid from the *NpT* plus test particle method, *Mol. Phys.*, 76(6) **1992** 1319
 59. Lofti, A., *PhD thesis, Ruhr-Universität Bochum, Germany*, **1993**
 60. Stoll, J., Vrabec, J., and Hasse, H., Vapor-Liquid Equilibria of Mixtures Containing Nitrogen, Oxygen, Carbon Dioxide, and Ethane, *AIChE J.*, 49(8) **2003** 2187
 61. Angus, S., Armstrong, B., and Reuck, K. M., *IUPAC, International Thermodynamic Tables of the Fluid State*, Volume 3, Carbon Dioxide, Pergamon Press, 1976

-
62. Jensen, R. H. and Kurata, F., Heterogeneous Phase Behavior of Solid Carbon Dioxide in Light Hydrocarbons at Cryogenic Temperatures, *AIChE J.*, 17(2) **1971** 357
 63. Rivollet, F., Chapoy, A., Coquelet, C., and Richon, D., Vapor-liquid equilibrium data for the carbon dioxide (CO₂) + difluoromethane (R32) system at temperatures from 283.12 to 343.25 K and pressures up to 7.46 MPa, *Fluid Phase Equil.*, 218 **2004** 95
 64. Diefenbacher, A. and Türk, M., (Vapour + liquid) Equilibria of binary mixtures of CO₂, CH₂F₂, CHF₃, and SF₆, *J. Chem. Thermodyn.*, 34 **2002** 1361
 65. Adams, R. A. and Stein, F. P., Vapor-Liquid Equilibria for Carbon Dioxide-Difluoromethane System, *J. Chem. Eng. Data*, 16(2) **1971** 146
 66. Higashi, S. and Takada, A., Molecular dynamics study of liquid CH₂F₂ (HFC-32), *Mol. Phys.*, 92(4) **1997** 641
 67. Roth, H., Peters-Gerth, P., and Lucas, K., Experimental vapor-liquid equilibria in the systems R22-R23, R22-CO₂, CS₂-R22, R23-CO₂, CS₂-R23 and their correlation by equations of state, *Fluid Phase Equil.*, 73 **1992** 147
 68. Mooijer-Van den Heuvel, M. M., Nasrifar, Kh., and Peters, C. J., *Unpublished data*, **2002**
 69. Song, W., Patel, N., and Maroncelli, M., A 2-Site Model for Simulating Supercritical Fluoroform, *J. Phys. Chem. B*, 106 **2002** 8783
 70. Moon, S.-D. and Moon, B. K., Gibbs Ensemble Monte Carlo Simulation for Vapor-Liquid Equilibrium of Binary Mixtures CO₂/C₃H₈, CO₂/CH₃OCH₃, and CO₂/CH₃COCH₃, *Bull. Korean Chem. Soc.*, 21(11) **2000** 1133
 71. Murthy, C. S., Singer, K., and McDonald, I. R., Interaction site models for carbon dioxide, *Mol. Phys.*, 44(1) **1981** 135
 72. Smit, B., Karaboni, S., and Siepmann, J. I., Computer simulations of vapor-liquid phase equilibria of *n*-alkanes, *J. Chem. Phys.*, 102(5) **1995** 2126
 73. Akers, W. W., Kelley, R. E., and Lipscomb, T. G., Low-temperature phase equilibria. Carbon dioxide-propane system, *Ind. Eng. Chem.*, 46(12) **1954** 2535
 74. Reamer, H. H., Sage, B. H., and Lacey, W. N., Phase Equilibria in Hydrocarbon Systems. Volumetric and phase behavior of the propane-carbon dioxide system, *Ind. Eng. Chem.*, 43(11) **1951** 2515
 75. Poettmann, F. H. and Katz, D. L., Phase Behavior of Binary Carbon Dioxide-Paraffin Systems, *Ind. Eng. Chem.*, 37(9) **1945** 847
 76. Florusse, L. J. and Peters, C. J., 2004, Personal communications
 77. Kordikowski, A., Schenk, A. P., van Nielen, R. M., and Peters, C. J., Volume Expansions and Vapor-Liquid Equilibria of Binary Mixtures of a Variety of Polar Solvents and Certain Near-Critical Solvents, *J. Supercrit. Fluids*, 8 **1995** 205

Chapter 7. Extension to ternary mixtures

In this chapter, simulation results for vapour-liquid equilibria in ternary mixtures are presented. Extension of the advanced Gibbs-Duhem integration method to ternary mixtures is straightforward. Some comments on the choice of the Clapeyron equation, the computation of the initial slope, and the initial point are given in Section 7.2. The relations needed to apply multiple-histogram reweighting to ternary mixtures are given in Section 7.3. Simulation results and simulation details are presented in Section 7.4. Finally, this chapter ends with several conclusions concerning this chapter.

7.1. Introduction

In Chapter 6, the advanced Gibbs-Duhem integration (GDI) method has been tested on the vapour-liquid equilibria of a number of binary systems. The advanced GDI method could be conveniently used to predict the vapour-liquid equilibria of those systems. Although binary mixtures are surely interesting, industry often works with multicomponent mixtures. Thus, prediction of vapour-liquid equilibria of multicomponent mixtures with molecular simulation would be worthwhile.

Like the conventional GDI method [1], the advanced GDI method is easily extended to the simulation of vapour-liquid equilibria in multicomponent mixtures. Although extension to multiphase equilibria in multicomponent systems is also possible [See for example ref. [2]], this chapter focuses on vapour-liquid equilibria in ternary mixtures.

The semigrand-canonical ensemble is particularly suitable for simulation of mixtures as the total number of molecules is fixed. The number of molecules of the individual components fluctuates by performing trial identity changes. Thus, the number of components does not influence the complexity of the simulation.

Other methods that can be used rather easily to simulate vapour-liquid equilibria in multicomponent mixtures are the mixture extension of the $Np\beta + \text{TM}$ method [3-6] and the $Np\beta - \text{Gibbs}$ ensemble [7, 8].

The mixture version of the $Np\beta + \text{TM}$ method relies on the accurate computation of a number of ensemble averages that occur in the Taylor series expansion of the chemical potentials. Although the capability of the $Np\beta - \text{Gibbs}$ ensemble to predict vapour-liquid equilibria still relies on the ability to successfully insert molecules into the liquid phase, the application to mixtures brings some ease; instead of performing insertions/removals of molecules for all components, it is convenient to only perform these moves for the smallest component and to carry out trial identity changes for the larger molecules [8].

Examples of simulation studies of vapour-liquid equilibria in ternary mixtures are given by a number of authors [9-16]. The Clapeyron equation presented in this chapter, along with the conventional GDI method, has also been used by Attwood [13] to predict solid-fluid equilibria in a ternary mixture.

7.2. Clapeyron equation, initial point, and initial slope

There are several choices for a Clapeyron equation describing vapour-liquid equilibria in a ternary mixture. For example, Escobedo [15] used a Clapeyron equation expressed in open-ensemble variables.

In this chapter, a Clapeyron equation at constant temperature and pressure is chosen. The availability of a large amount of experimental data at these conditions makes this choice of variables very convenient. The fugacity fractions act as dependent and independent integration parameters. A Clapeyron equation describing two-phase equilibria in a ternary mixture, which corresponds to abovementioned choice of variables, is given by [See Appendix A]:

$$\left(\frac{\partial \xi_1}{\partial \xi_3} \right)_{p, \beta, \sigma} = - \frac{\xi_1 \left(\xi_3 \Delta x_2^{\text{LV}} - \xi_2 \Delta x_3^{\text{LV}} \right)}{\xi_3 \left(\xi_1 \Delta x_2^{\text{LV}} - \xi_2 \Delta x_1^{\text{LV}} \right)} \quad (7.1)$$

where ξ_i is the fugacity fraction of component i , σ indicates that the derivative is taken along the saturation line, and $\Delta x_i^{\text{LV}} = x_i^{\text{L}} - x_i^{\text{V}}$. Other Clapeyron equations at constant temperature and pressure can be obtained from Eq. (7.1) by interchanging the indices.

The numerical integration of differential equation (7.1) commences at zero fugacity fraction of component three, $\xi_3 = 0$. Clearly, the corresponding initial point equals the vapour-liquid coexistence point of a binary mixture at the specified temperature, saturation pressure and fugacity fraction $\xi_2 = 1 - \xi_1$.

In this chapter, the initial point is the result of GDI in a binary mixture. The advanced GDI method, applied to a binary mixture, provides the vapour-liquid coexistence line expressed as pressure versus fugacity fraction at constant temperature. A polynomial fit to this coexistence line enables us to predict the fugacity fraction $\xi_2 = 1 - \xi_1$ at every saturation pressure in the range that is covered by the simulations. Once the thermodynamic state of the coexistence point of interest $(p, \xi_2; \beta)$ has been obtained, the mole fractions and densities at coexistence can be computed from the optimized semigrand-canonical probability distributions of the liquid and the vapour phase.

From the previous chapters, it became clear that a GDI series has no internal checkpoints. In case the initial point is not a *true* coexistence point, the GDI method traces a curve of non-zero free-energy difference between the vapour and the liquid phase. Moreover, the deviation of the simulation results from the true coexistence line grows when the difference between the molar volumes of the phases decreases during successive integration steps. In case the binary GDI results deviate substantially from true coexistence, it would be better to compute the initial point for the ternary integration from another method.

The $Np\beta$ -Gibbs ensemble [7, 8] is a suitable alternative, provided that the liquid-phase density is moderate and/or the densities between the liquid and the vapour phase are not too different. In other cases, one could use the mixture version of the $Np\beta + \text{TM}$ method [3-6]. In contrast with a Gibbs-Duhem integration series, these methods provide a quasi-direct estimate of the initial point.

Equation (7.1) is not defined at $\xi_3 = 0$. However, the right-hand side of Eq. (7.1) (slope or integrand) in the limit $\xi_3 \rightarrow 0$ can be obtained by rewriting the fugacity fractions in terms of fugacities. Equation (7.1), rewritten in the limit $\xi_3 \rightarrow 0$, is given in Appendix A. For completeness, the derivation of the limiting case is given here.

The fugacities of components two and three are given by:

$$\hat{f}_2 = x_2 \gamma_2^{(123)} f_2 \quad \hat{f}_3 = x_3 \gamma_{3, \text{HL}}^{(123)} H_3^{(123)} \quad (7.2)$$

where \hat{f}_i is the fugacity of component i in the ternary mixture, f_2 is the fugacity of pure component two, $\gamma_2^{(123)}$ is the activity coefficient of component two in the ternary mixture, $H_3^{(123)}$ is Henry's constant of component three in the ternary mixture, and $\gamma_{3, \text{HL}}^{(123)}$ is the activity coefficient according to Henry's law of component three in the ternary mixture. These variables are evaluated at the temperature and pressure of the system. Insertion of these fugacities into Eq. (7.1) gives:

$$\left(\frac{\partial \xi_1}{\partial \xi_3} \right)_{p, \beta, \sigma} = -\xi_1 \frac{\Delta x_2^{\text{LV}} - \left(\frac{x_2 \gamma_2^{(123)} f_2}{x_3 \gamma_{3, \text{HL}}^{(123)} H_3^{(123)}} \right)^\alpha \Delta x_3^{\text{LV}}}{\xi_1 \Delta x_2^{\text{LV}} - \xi_2 \Delta x_1^{\text{LV}}} \quad (7.3)$$

where α represents either the liquid or the vapour phase. Thus, Eq. (7.3) can also be written as:

$$\begin{aligned} \left(\frac{\partial \xi_1}{\partial \xi_3} \right)_{p, \beta, \sigma} &= -\xi_1 \frac{\Delta x_2^{\text{LV}} - \left(\frac{x_2 \gamma_2^{(123)} f_2}{x_3 \gamma_{3, \text{HL}}^{(123)} H_3^{(123)}} \right)^{\text{L}} x_3^{\text{L}} + \left(\frac{x_2 \gamma_2^{(123)} f_2}{x_3 \gamma_{3, \text{HL}}^{(123)} H_3^{(123)}} \right)^{\text{V}} x_3^{\text{V}}}{\xi_1 \Delta x_2^{\text{LV}} - \xi_2 \Delta x_1^{\text{LV}}} \\ &= -\xi_1 \frac{\Delta x_2^{\text{LV}} - \left(\frac{x_2 \gamma_2^{(123)} f_2}{\gamma_{3, \text{HL}}^{(123)} H_3^{(123)}} \right)^{\text{L}} + \left(\frac{x_2 \gamma_2^{(123)} f_2}{\gamma_{3, \text{HL}}^{(123)} H_3^{(123)}} \right)^{\text{V}}}{\xi_1 \Delta x_2^{\text{LV}} - \xi_2 \Delta x_1^{\text{LV}}} \end{aligned} \quad (7.4)$$

In the limit $x_3 \rightarrow 0$, the activity coefficient $\gamma_{3, \text{HL}}^{(123)}$ according to Henry's law approaches one:

$$\left(\frac{\partial \xi_1}{\partial \xi_3} \right)_{p, \beta, \sigma, \xi_3 \rightarrow 0} = -\xi_1 \frac{x_2^{\text{L}} \left\{ 1 - \left(\frac{\gamma_2^{(12)} f_2}{H_3^{(12)}} \right)^{\text{L}} \right\} - x_2^{\text{V}} \left\{ 1 - \left(\frac{\gamma_2^{(12)} f_2}{H_3^{(12)}} \right)^{\text{V}} \right\}}{\xi_1 \Delta x_2^{\text{LV}} - \xi_2 \Delta x_1^{\text{LV}}} \quad (7.5)$$

In Eq. (7.5), the superscripts (123) are substituted by (12) as, at infinite dilution of component three, the system can be considered as a binary mixture. In order to determine the initial slope of the Clapeyron equation, the ratios indicated in Eq. (7.5) are needed. These ratios can be computed analogously to those needed to determine the initial slope in a binary mixture.

The residual chemical potentials of components two and three in the ternary mixture are given by:

$$\beta\mu_2^{\text{res}}(p, \beta, \{x_i, i = 1 \dots 3\}) = \ln \left(\frac{\gamma_2^{(123)} f_2}{p} \right) \quad (7.6)$$

$$\beta\mu_3^{\text{res}}(p, \beta, \{x_i, i = 1 \dots 3\}) = \ln \left(\frac{\gamma_{3, \text{HL}}^{(123)} H_3^{(123)}}{p} \right) \quad (7.7)$$

In the limit $x_3 \rightarrow 0$, the residual chemical-potential difference $\beta\Delta\mu_{32}^{\text{res}} = \beta(\mu_3^{\text{res}} - \mu_2^{\text{res}})$ can be written as:

$$\lim_{x_3 \rightarrow 0} \beta\Delta\mu_{32}^{\text{res}} = -\ln \left(\frac{\gamma_2^{(12)} f_2}{H_3^{(12)}} \right) \quad (7.8)$$

or,

$$\frac{\gamma_2^{(12)} f_2}{H_3^{(12)}} = \lim_{x_3 \rightarrow 0} \exp \left\{ -\beta\Delta\mu_{32}^{\text{res}} \right\} \quad (7.9)$$

Thus, in order to obtain the slope of Clapeyron equation (7.5) in the limit $x_3 \rightarrow 0$, $\Delta\mu_{32}^{\text{res}}$ must be computed for both the liquid and the vapour phase. This can be conveniently performed by writing the right-hand side of Eq. (7.9) as an ensemble average in the binary $N_1 N_2 p \beta$ -ensemble. Just like in Chapter 4, Section 4.3.2, the residual chemical-potential difference is computed from the overlapping-distributions difference method. The limit of infinite dilution of component three in the binary mixture is again approximated by $x_3 = 1/N$. It is possible that component two and component three have very different size, shape, and force-field parameters. In that case, the distributions from which the residual-chemical potential difference is to be determined may have a limited region of overlap or no region of overlap at all. However, notice that it is also possible to compute the initial slope by changing the identity of a molecule ‘one’ to a molecule ‘three’ by interchanging the indices in Eqs. (7.5) and (7.9). This is convenient in case molecules of type one and type three are more alike than molecules of type two and type three.

Attwood [13] also derived a relation for the limiting slope of the Clapeyron equation (7.1) in the limit of infinite dilution of one of the components:

$$\left(\frac{\partial \xi_1}{\partial \xi_3} \right)_{p, \beta, \sigma, \xi_3 \rightarrow 0} = -\xi_1 \frac{\left\{ \left(\frac{\hat{f}_1}{H_3^{(12)}} + \frac{\hat{f}_2}{H_3^{(12)}} \right)^{\text{L}} - \left(\frac{\hat{f}_1}{H_3^{(12)}} + \frac{\hat{f}_2}{H_3^{(12)}} \right)^{\text{V}} \right\} (1 - \xi_1) + \Delta x_1^{\text{LV}}}{\Delta x_1^{\text{LV}}} \quad (7.10)$$

Relation (7.10) requires the computation of *four* different ratios of a mixture fugacity over Henry’s constant. Attwood [13] computes $\hat{f}_j/H_k^{(ij)}$ from the ensemble average of $\exp\{-\beta\Delta\mu_{kj}^{\text{res}}\}$ in the limit of infinite dilution of component k . However, from Eqs. (7.2)

and (7.9), it is clear that it equals $x_j \exp\{-\beta\Delta\mu_{kj}^{\text{res}}\}$ in the limit of infinite dilution of component k . Fortunately, equation (7.10) can be rewritten to the more convenient form, Eq. (7.5), which requires only two ratios $\gamma_j^{(ij)} f_j / H_k^{(ij)}$.

7.3. Advanced Gibbs-Duhem integration in ternary mixtures

7.3.1. Multiple-histogram reweighting in ternary mixtures

Formally, the application of multiple-histogram reweighting to simulation results obtained from ternary isobaric semigrand-canonical ensemble simulations would require four-dimensional histograms. To reduce the computational burden of multiple-histogram reweighting, the energy and volume dimensions are left out of the histograms.

The field variables conjugate to the volume and the energy are pressure and reciprocal temperature. As the ternary simulations are performed at constant pressure and temperature, reweighting to other pressures and/or temperatures is not needed to compute phase coexistence. Thus, the fluctuations in the volume and energy need not be collected into histograms. The pseudo-degeneracy of phase α , estimated from a number of histograms collected at R thermodynamic state points that differ in fugacity fractions only, is given by:

$$\Theta^{*\alpha}(N_2, N_3) = \frac{\sum_{i=1}^R g_i^{-1} \mathcal{H}_i(N_2, N_3)}{\sum_{j=1}^R \mathcal{N}_j g_j^{-1} \exp \left\{ N_2 \ln \left(\frac{[\xi_2]_j [\xi_1]_0}{[\xi_1]_j [\xi_2]_0} \right) + N_3 \ln \left(\frac{[\xi_3]_j [\xi_1]_0}{[\xi_1]_j [\xi_3]_0} \right) + N \ln \left([\hat{f}_1]_j / p^{\text{ref}} \right) \right\}} \quad (7.11)$$

where g_j depends on the correlation time of successive configurations collected at state point j [17-19], $N \ln([\hat{f}_1]_j / p^{\text{ref}}) / \beta$ is the total configurational semigrand energy at state point j , $\mathcal{H}_j(N_2, N_3)$ is a histogram collected at state point j , and \mathcal{N}_j is the sum of the entries in histogram $\mathcal{H}_j(N_2, N_3)$. The subscript '0' indicates the reference state point. The corresponding semigrand-canonical probability to observe phase α with a particular number of molecules of component two, N_2 , and a particular number of molecules of component three, N_3 , at a particular set of fugacity fractions is given by:

$$\pi^\alpha(N_2, N_3; \xi_2, \xi_3) = \frac{\Theta^{*\alpha}(N_2, N_3) \exp \left(N_2 \ln \left\{ \frac{\xi_2 [\xi_1]_0}{\xi_1 [\xi_2]_0} \right\} + N_3 \ln \left\{ \frac{\xi_3 [\xi_1]_0}{\xi_1 [\xi_3]_0} \right\} \right)}{\sum_{N'_2} \sum_{N'_3} \Theta^{*\alpha}(N'_2, N'_3) \exp \left(N'_2 \ln \left\{ \frac{\xi_2 [\xi_1]_0}{\xi_1 [\xi_2]_0} \right\} + N'_3 \ln \left\{ \frac{\xi_3 [\xi_1]_0}{\xi_1 [\xi_3]_0} \right\} \right)} \quad (7.12)$$

The procedure that must be followed to compute phase coexistence from Eq. (7.12) and the corrector equations was given in Chapter 6. Without volume and energy entries in the histograms, one cannot provide the liquid and vapour densities at coexistence and energy-related quantities at a predicted coexistence point.

A possible solution to this problem, which is otherwise not performed in this work, is the following. Suppose histograms are collected from all volume-energy-composition combinations encountered during the course of a simulation. The optimization of the free energies and the global liquid-phase and vapour-phase pseudo-degeneracies is performed on histograms that do not include the volume and energy dimensions. As all volume and energy bins corresponding to a certain composition have the same pseudo-Boltzmann weight, the liquid and vapour densities at coexistence can be computed from the full histograms.

Notice that only the fluctuations in the conjugate densities of the dependent and independent field-type integration variables show up in the semigrand-canonical probability that is needed to compute phase coexistence in a multicomponent mixture. Thus, even for the computation of vapour-liquid equilibria in a multicomponent mixture, only two-dimensional histograms are needed in the optimization of the global pseudo-degeneracies of the liquid and the vapour phase.

7.3.2. Acceptance criteria

Acceptance criteria for rotational, translational and volume MC trial moves are identical in pure-component and multicomponent mixtures. Here, the trial moves that include changes in the composition are presented.

The overall acceptance criterion for an *unbiased* identity-change trial move in a c -component semigrand-canonical ensemble is:

$$\min \left\{ 1, \exp \left[-\beta \Delta U^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}} + \sum_{i=2}^c m_{1i} \ln \left(\frac{\langle \mathcal{W}_{\text{IG},1}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle_{\xi_i}}{\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle_{\xi_1}} \right) \right] \right\} \quad (7.13)$$

where $\Delta U^{\text{ext} + \text{LJ}_{\text{int}} + \text{C}_{\text{int}}}$ is the energy change resulting from the identity change, and $\langle \mathcal{W}_{\text{IG},i}^{\text{LJ}_{\text{int}} + \text{C}_{\text{int}}} \rangle$ is the normalized Rosenbluth factor of an isolated chain of component i with intramolecular Lennard-Jones and Coulombic interactions [See Chapter 2, Section 2.5.4]. The parameters m_{1j} , corresponding to the different possibilities for the identity-change trial move in a ternary mixture, are given in Table 7.1.

Table 7.1: Values of the parameters m_{1j} in Eq. (7.13)

	1 → 2	2 → 1	2 → 3	3 → 2	1 → 3	3 → 1
m_{12}	1	-1	-1	1	1	-1
m_{13}	0	0	1	-1	0	0

The acceptance criterion for the parallel-tempering configuration switch between phase i and phase j in a c -component mixture is given by:

$$\min \left\{ 1, \exp \left[\sum_{k=2}^c \left([N_k]_j - [N_k]_i \right) \ln \left\{ \frac{[\xi_k]_i [\xi_1]_j}{[\xi_1]_i [\xi_k]_j} \right\} \right] \right\} \quad (7.14)$$

7.4. Simulation details and results

7.4.1. Simulation details

The simulation details for the ternary simulations do not differ substantially from the ones given in Chapter 6. Of course, the dependent and independent integration parameters differ from those in a binary simulation. However, the principle of the numerical integration of the Clapeyron equation is identical to that given in Chapter 6. Three liquid and three vapour phases were used per integration step (coexistence point). The thermodynamic state points at which the simulations are performed differ only in their fugacity fractions.

7.4.2. Ternary system ethane/methane/carbon dioxide

Sources of experimental data for this ternary system are given in Table 7.2. The vapour-liquid equilibria of this ternary system have been simulated by a number of authors. Vrabec and Fischer [3] used the mixture version of the $Np\beta + TM$ method to simulate the vapour-liquid equilibria at 233.15 K/2 MPa, 250.5 K/2 MPa, and 250.5 K/3.04 MPa. They used a 2CLJ model for ethane, a 1CLJ model for methane, and a 2CLJQ model for carbon dioxide. The Lorentz-Berthelot (LB) combining rules were used for all interactions. Binary interaction parameters with values close to zero were included in the LB combining rules [See Section 5.2, Eq. (5.13)].

Liu and Beck [12] used the same force fields and binary interaction parameters to study the system at 250 K/3.04 MPa with the $Np\beta - Gibbs$ ensemble.

Guo *et al.* [14] simulated the mixture at $T = 220$ K and $T = 233$ K at various pressures with the $Np\beta - Gibbs$ ensemble. Ethane, methane, and carbon dioxide were respectively modelled as 2CLJ, 1CLJ, and 2CLJ molecules. LB combining rules with a binary interaction parameter were used to compute all interactions.

Table 7.2: Sources of experimentally determined vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide

Reference	Temp. range, K	Press. range, MPa	Number of sets
Davalos <i>et al.</i> [20, 21]	250	2.16-3.04	3
Knapp <i>et al.</i> [22]	220.0-251.1	2.0-5.0	11
Wei <i>et al.</i> [23]	230	1.15-6.59	8

In this work, simulation results for the vapour-liquid equilibria of the ternary system at $T = 250$ K and at four different pressures are presented. The simulation results at $T = 250$ K/ $p = 1.9$ MPa are *pure predictions*; there are no experimental data at these conditions to compare the simulation results with.

The advanced TraPPE-EH [24] potential was used for methane and ethane and the TraPPE-AA [25] potential for carbon dioxide. The interactions between like molecules were described with the LB combining rules while the Lorentz-Hudson-McCoubrey (LHMcC) combining rules were used for the interactions between unlike molecules. Ratios of the fugacity over Henry's constant, computed from the overlapping-distributions difference method, are given in Table 7.3 and in Table 7.4.

The fugacity fraction at the saturation pressure of interest was computed from a polynomial fit to the binary coexistence data presented in Chapter 6. The liquid and vapour densities and mole fractions at coexistence in Table 7.3 and Table 7.4 were obtained from the optimized semigrand-canonical probability distribution [Chapter 6, Eq. (6.33)].

Simulation results are given in graphical form in Figure 7.1 to Figure 7.4. Results in tabular form are given in Appendix D, Table D.9. The results obtained at $T = 250 \text{ K} / p = 1.9 \text{ MPa}$ show two separate vapour-liquid coexistence regions. From the experimental data presented in Figure 6.9, Chapter 6, it is obvious that this behaviour should also occur at $T = 250 \text{ K} / p = 2.128 \text{ MPa}$. However, the combination of the TraPPE-EH potential for ethane and the TraPPE-AA potential for carbon dioxide, along with the abovementioned combining rules, predicts a one-phase region for the binary mixture ethane/carbon dioxide at this pressure. Thus, the simulation results for the ternary system at $T = 250 \text{ K} / p = 2.128 \text{ MPa}$ contain only one two-phase region.

The simulation results at the three highest pressures are close to the experimental data. The results obtained at $p = 2.128 \text{ MPa}$ show a small deviation from the experimental data close to zero mole fraction of ethane. The results obtained at $p = 3.040 \text{ MPa}$ are comparable to those obtained by Vrabec and Fischer [3] and by Liu and Beck [12].

It is concluded that the simple force fields used by Vrabec and Fischer perform equally well as the advanced TraPPE potentials. However the simple force fields of Vrabec and Fisher exploited a binary interaction parameter in the LB combining rules. Moreover, the complexity of the TraPPE force fields is justified by its transferability.

Table 7.3: Initial points, initial slopes, and saturation pressures at $T = 250 \text{ K}$ of the ternary mixture ethane(1)/ methane(2)/ carbon dioxide(3)

p MPa	ξ_2	x_2^L	x_2^V	ρ^L kg/m ³	ρ^V kg/m ³	$\left(\frac{f_2\gamma_2^{(12)}}{H_3^{(12)}}\right)_L^{\text{sim}}$	$\left(\frac{f_2\gamma_2^{(12)}}{H_3^{(12)}}\right)_V^{\text{sim}}$
1.900	0.3316	0.0553	0.2815	432	29.6	2.86(10)	1.10(4)
2.128	0.4098	0.0763	0.3505	428	32.3	2.87(9)	1.10(3)
2.533	0.5131	0.1138	0.4405	421	37.0	2.82(10)	1.13(3)
3.040	0.6027	0.1608	0.5195	410	43.8	2.73(7)	1.15(2)

Table 7.4: Initial point and initial slope at $T = 250 \text{ K} / 1.9 \text{ MPa}$ of the ternary mixture ethane(1)/carbon dioxide(2)/methane(3)

p MPa	ξ_2	x_2^L	x_2^V	ρ^L kg/m ³	ρ^V kg/m ³	$\left(\frac{f_2\gamma_2^{(12)}}{H_3^{(12)}}\right)_L^{\text{sim}}$	$\left(\frac{f_2\gamma_2^{(12)}}{H_3^{(12)}}\right)_V^{\text{sim}}$
1.900	0.8861	0.9352	0.8852	965	47.4	0.066(3)	0.84(2)

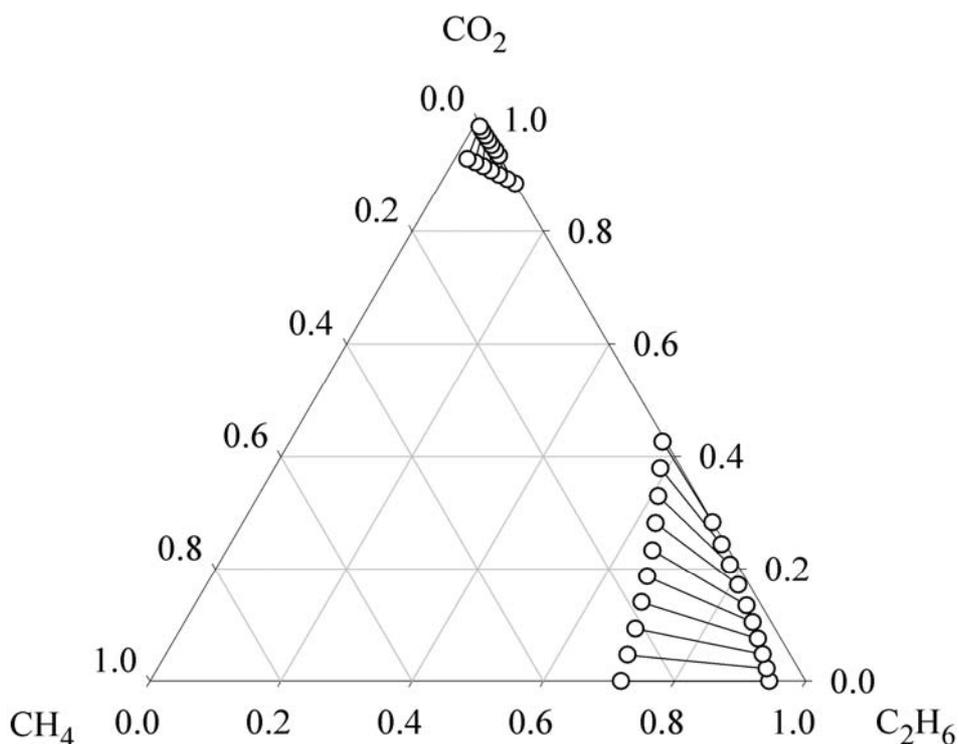


Figure 7.1: Vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide at $T = 250 \text{ K} / p = 1.9 \text{ MPa}$. Open circles are simulation results obtained in this work with LB combining rules for interactions between like molecules and LHM_cC combining rules for interactions between unlike molecules. No experimental data are available at this pressure. Solid lines represent tie lines between coexistence points.

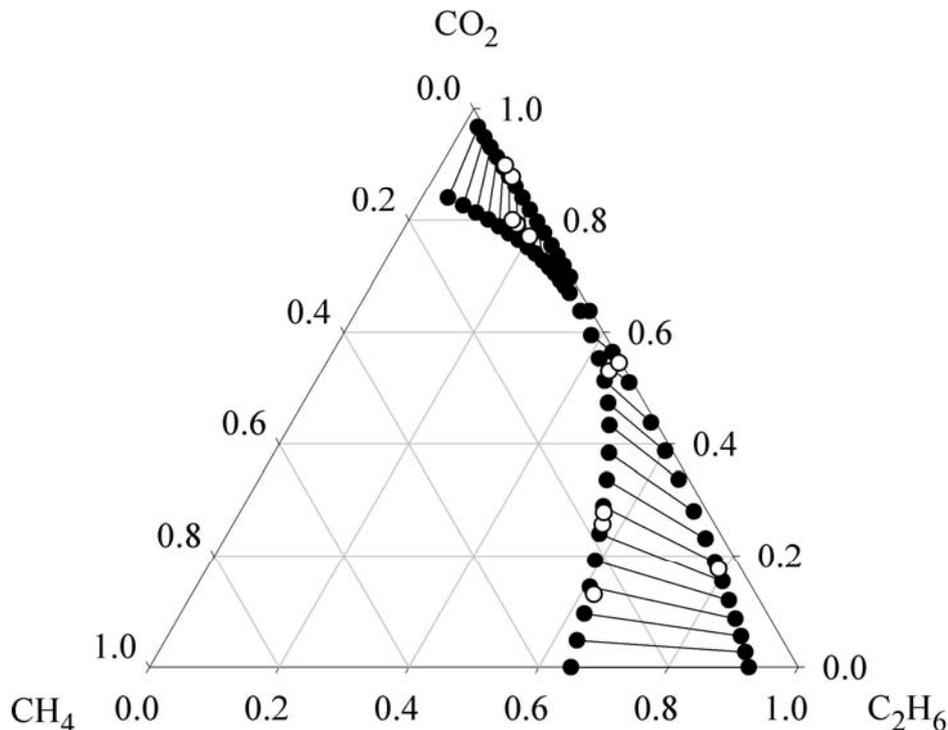


Figure 7.2: Vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide at $T = 250 \text{ K} / p = 2.128 \text{ MPa}$. Open circles represent experimental data published by Davalos et al. [20, 21]. Solid black circles are simulation results obtained in this work with LB combining rules for interactions between like molecules and LHM_cC combining rules for interactions between unlike molecules. Solid lines represent tie lines between coexistence points.

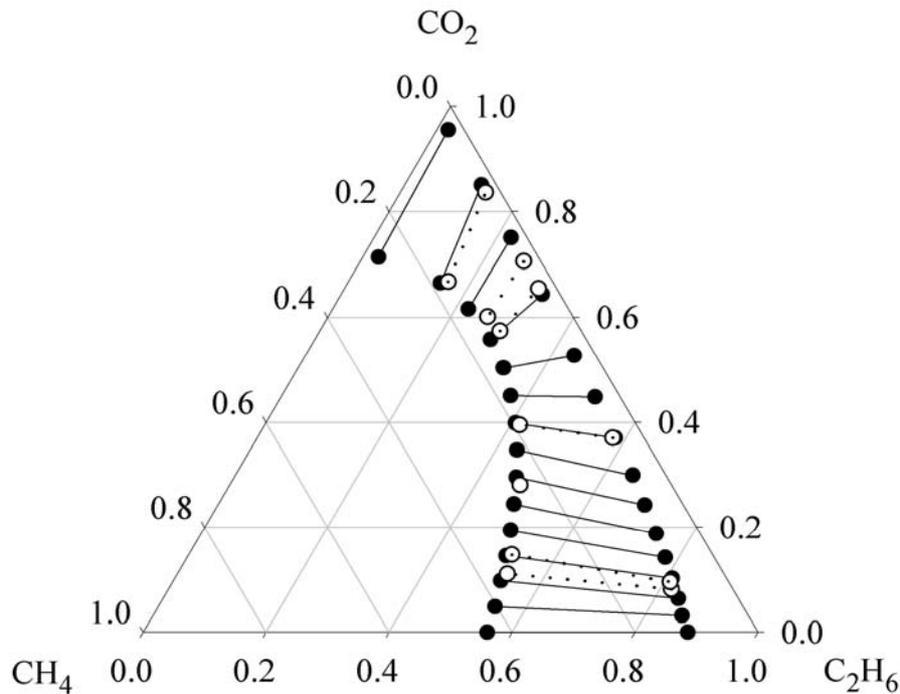


Figure 7.3: Vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide at $T = 250 \text{ K}$ / $p = 2.533 \text{ MPa}$. Open circles represent experimental data published by Davalos et al. [20, 21]. Dashed lines represent tie lines between experimental coexistence points. Solid black circles are simulation results obtained in this work with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Solid lines represent tie lines between coexistence points obtained in this work.

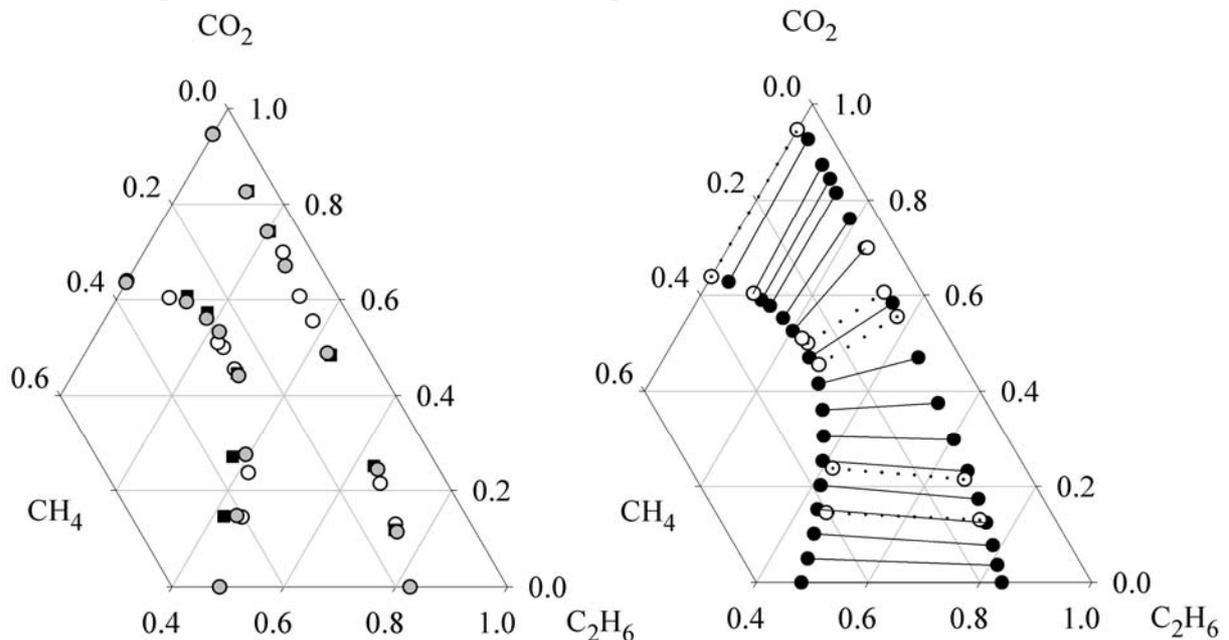


Figure 7.4: Vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide at $T = 250 \text{ K}$ / $p = 3.040 \text{ MPa}$. Open circles represent experimental data published by Davalos et al. [20, 21]. Dashed lines represent tie lines between experimental coexistence points. Solid grey circles and solid black squares represent simulation results published by Vrabec and Fischer [3] and by Liu and Beck [12] respectively. Solid black circles are simulation results obtained in this work with LB combining rules for interactions between like molecules and LHMcC combining rules for interactions between unlike molecules. Solid lines represent tie lines between coexistence points obtained in this work.

7.4.3. Ternary system propane/methane/carbon dioxide

Experimental data for this ternary system were published by Webster and Kidnay [26]. The vapour-liquid equilibria of this ternary mixture have not been simulated before.

In this work, simulation results for vapour-liquid equilibria at $T = 270 \text{ K} / p = 2.8 \text{ MPa}$ are presented. The advanced TraPPE-EH [24] potential is used for methane and propane and the TraPPE-AA [25] potential for carbon dioxide. The interactions between like molecules were described with the LB combining rules while the LHMCC combining rules were used for the interactions between unlike molecules.

The ratios of the fugacities over Henry's constant, computed from the overlapping-distributions difference method, are given in Table 7.5. Simulation results are given in graphical form in Figure 7.5. Results in tabular form are given in Appendix D, Table D.8. The initial point is obtained from the simulation results for the binary mixture propane/methane that were presented in Chapter 6.

Table 7.5: Initial point and initial slope at $T = 270 \text{ K} / 2.8 \text{ MPa}$ for the ternary mixture propane(1)/methane(2)/ carbon dioxide(3)

p MPa	ξ_2	x_2^L	x_2^V	ρ^L kg/m ³	ρ^V kg/m ³	$\left(\frac{f_2\gamma_2^{(12)}}{H_3^{(12)}}\right)_L^{\text{sim}}$	$\left(\frac{f_2\gamma_2^{(12)}}{H_3^{(12)}}\right)_V^{\text{sim}}$
2.8	0.8555	0.1631	0.8038	494	30.9	2.57(9)	1.09(4)

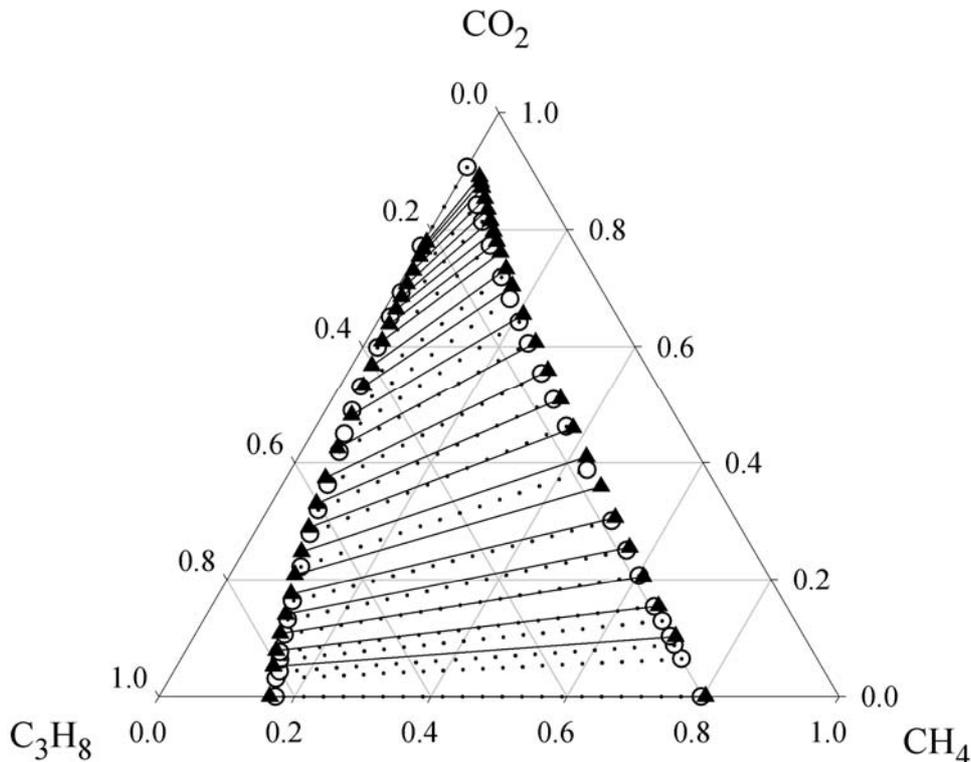


Figure 7.5: Vapour-liquid equilibria of the ternary system propane/methane/carbon dioxide at $T = 270 \text{ K} / p = 2.8 \text{ MPa}$. Open circles represent the experimental data published by Webster and Kidnay [26]. Dashed lines represent tie lines between experimental coexistence points. Solid black upward triangles are simulation results obtained in this work with LB combining rules for interactions between like molecules and LHMCC combining rules for interactions between unlike molecules. Solid lines represent tie lines between coexistence points obtained in this work.

Like the simulation results for the vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide, the simulation results for this ternary system are close to the experimental data.

The simulated liquid-phase and vapour-phase mole fractions of carbon dioxide near zero mole fraction of methane are slightly larger than the experimental values. This is in agreement with the simulation results for the binary mixture propane/carbon dioxide that were presented in Chapter 6, Figure 6.16. This agreement confirms the conclusion that the advanced GDI method easily traces the true coexistence curve in case the numerical integration is started from an initial point that is close to the true coexistence point of interest.

7.5. Conclusions

In this chapter, the extension of the advanced Gibbs-Duhem integration method to vapour-liquid equilibria of ternary mixtures is discussed. The advanced Gibbs-Duhem integration method was used to simulate the vapour-liquid equilibria of a number of ternary mixtures: ethane/methane/carbon dioxide at four pressures and propane/methane/carbon dioxide at a single pressure.

It was shown that the combination of the advanced TraPPE force field combined with the Lorentz-Berthelot combining rules for the interactions between like molecules and the Lorentz-Hudson-McCoubrey combining rules for the interactions between unlike molecules is very well capable of predicting results close to the experimental ternary vapour-liquid equilibria.

The advanced Gibbs-Duhem integration method is very convenient for the simulation of ternary vapour-liquid equilibria. As the volume and energy (continuous properties) are not needed in the computation of phase coexistence at constant temperature and pressure, the histograms contain only entries for (discrete) numbers of molecules. Such a reduction of the dimension of the histograms increases the optimization speed of the global pseudo-degeneracies of both phases.

However, without volume entries in the histograms, one cannot provide the liquid and vapour densities and energy-related quantities at a predicted coexistence point. A possible solution to this problem is the following. Histograms are collected from all volume, energy, and composition combinations encountered during the course of a simulation. The optimization of the free energies and the global liquid-phase and vapour-phase pseudo-degeneracies is performed using histograms that have no volume and energy dimensions. The liquid and vapour densities at coexistence are computed from the full histograms.

An identical procedure applies to the computation of vapour-liquid equilibria in multicomponent mixtures; only two-dimensional histograms are needed in the optimization of the pseudo-degeneracies of the liquid and the vapour phase.

The computation of the starting point from a previous binary Gibbs-Duhem integration series is not very convenient if one is only interested in the phase behaviour of the ternary system. Furthermore, in case the initial point that is used to start a binary Gibbs-Duhem integration is no true coexistence point, the initial point for the ternary simulation is probably even worse. Thus, if one is sure the binary Gibbs-Duhem-integration results deviate substantially from true coexistence, it is better to compute the starting point for the ternary integration from another method. Possible candidates are the mixture version of the $Np\beta + \text{TM}$ method [3-6] and the $Np\beta - \text{Gibbs}$ ensemble [7, 8].

7.6. References

1. Mehta, M. and Kofke, D. A., Coexistence Diagrams of Mixtures by Molecular Simulation, *Chem. Eng. Sci.*, 49(16) **1994** 2633
2. Agrawal, R., Mehta, M., and Kofke, D. A., Efficient Evaluation of Three-Phase Coexistence Lines, *Int. J. Thermophys.*, 15(6) **1994** 1073
3. Vrabec, J. and Fischer, J., Vapor-Liquid Equilibria of the Ternary Mixture CH₄+C₂H₆+CO₂ from Molecular Simulation, *AIChE J.*, 43(1) **1997** 212
4. Kronome, G., Szalai, I., Wendland, M., and Fischer, J., Extension of the NpT + test particle method for the calculation of phase equilibria of nitrogen + ethane, *J. Mol. Liquids*, 85 **2000** 237
5. Vrabec, J., Lofti, A., and Fischer, J., Vapour liquid equilibria of Lennard-Jones model mixtures from the NpT plus test particle method, *Fluid Phase Equil.*, 112 **1995** 173
6. Vrabec, J. and Fischer, J., Vapour liquid equilibria of mixtures from the NpT plus test particle method, *Mol. Phys.*, 85(4) **1995** 781
7. Panagiotopoulos, A. Z., Quirke, N., Stapleton, M. R., and Tildesley, D. J., Phase equilibria by simulation in the Gibbs ensemble: Alternative derivation, generalization and application to mixture and membrane equilibria, *Mol. Phys.*, 63(4) **1988** 527
8. Panagiotopoulos, A. Z., Exact Calculations of Fluid-Phase Equilibria by Monte Carlo Simulation in a New Statistical Ensemble, *Int. J. Thermophys.*, 10(2) **1989** 447
9. Carrero-Mantilla, J. and Llano-Restrepo, M., Vapor-liquid equilibria of the binary mixtures nitrogen + methane, nitrogen + ethane and nitrogen + carbon dioxide, and the ternary mixture nitrogen + methane + ethane from Gibbs-ensemble molecular simulation, *Fluid Phase Equil.*, 208(1-2) **2003** 155
10. Kim, S.-Y., Panagiotopoulos, A. Z., and Floriano, M. A., Ternary oil-water-amphiphile systems: self-assembly and phase equilibria, *Mol. Phys.*, 100(14) **2002** 2213
11. Vrabec, J. and Fischer, J., Vapor-Liquid Equilibria of the Ternary Mixture CH₄-C₂H₆-CO₂ from molecular simulation, *AIChE J.*, 43(1) **1997** 212
12. Liu, A. and Beck, Th. L., Vapor-Liquid Equilibria of Binary and Ternary Mixtures Containing Methane, Ethane, and Carbon Dioxide from Gibbs Ensemble Simulations, *J. Phys. Chem. B*, 102 **1998** 7627
13. Attwood, B. C., PhD Thesis, Monte Carlo simulations of solid-fluid phase equilibria in binary and ternary mixtures, *NCSU, Raleigh NC, USA* **2003**
14. Guo, M.-X., Zhao, E., and Lu, B. C. Y., Molecular simulations of enthalpies for CH₄-C₂H₆-CO₂ mixtures at saturation conditions, *Thermochim. Acta*, 254 **1995** 55
15. Escobedo, F. A., Tracing coexistence lines in multicomponent fluid mixtures by molecular simulation, *J. Chem. Phys.*, 110(24) **1999** 11999
16. Lísal, M., Smith, W. R., and Nezbeda, I., Molecular Simulation of Multicomponent Reaction and Phase Equilibria in MTBE Ternary System, *AIChE J.*, 46(4) **2000** 866
17. Ferrenberg, A. M. and Swendsen, R. H., Optimized Monte Carlo Data Analysis, *Phys. Rev. Lett.*, 63(12) **1989** 1195
18. Swendsen, R. H., Modern methods of analyzing Monte Carlo computer simulations, *Physica A*, 194 **1993** 53
19. Ferrenberg, A. M. and Swendsen, R. H., Optimized Monte Carlo Data Analysis, *Computers in Physics*, sep/oct **1989** 101

-
20. Davalos, J., Anderson, W. R., Phelps, R. E., and Kidnay, A. J., Liquid-Vapor Equilibria at 250.00 K for Systems Containing Methane, Ethane, and Carbon Dioxide, *J. Chem. Eng. Data*, 21(1) **1976** 81
 21. Davalos, J., Anderson, W. R., Phelps, R. E., and Kidnay, A. J., Correction, *J. Chem. Eng. Data*, 21(4) **1976** 510
 22. Knapp, H., Yang, X., and Zhang, Z., Vapor-liquid equilibria in ternary mixtures containing nitrogen, methane, ethane, and carbon dioxide at low temperatures and high pressures, *Fluid Phase Equil.*, 54 **1990** 1
 23. Wei, W., Brown, T. S., Kidnay, A. J., and Sloan, D. E., Vapor + liquid equilibria for the ternary system methane + ethane + carbon dioxide at 230 K and its constituent binaries at temperatures from 207 K to 270 K, *J. Chem. Eng. Data*, 40 **1995** 726
 24. Chen, B. and Siepmann, J. I., Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes, *J. Phys. Chem. B*, 103 **1999** 5370
 25. Potoff, J. J. and Siepmann, J. I., Vapor-Liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen, *AIChE J.*, 47(7) **2001** 1676
 26. Webster, L. A. and Kidnay, A. J., Vapor-Liquid Equilibria for the Methane-Propane-Carbon Dioxide Systems at 230 K and 270 K, *J. Chem. Eng. Data*, 46 **2001** 759

Appendix A. Clapeyron equations

In this appendix, the original Clapeyron equation and a number of other Clapeyron equations are presented, which are derived from the semigrand form of the Gibbs-Duhem equation [See Chapter 2, Section 2.3]:

$$d \ln \left\{ \sum_{i=1}^c \hat{f}_i / p^{\text{ref}} \right\} = h' d\beta + \beta v d p - \sum_{i=1}^c \frac{x_i}{\xi_i} d\xi_i \quad (\text{A.1})$$

where c is the number of components in the mixture, p^{ref} is a reference pressure, \hat{f}_i is the fugacity of component i in the mixture, h' is the configurational enthalpy per molecule, v is the volume per molecule, and ξ_i is the fugacity fraction of component i . To derive a Clapeyron equation, one degree of freedom is needed. The number of degrees of freedom, F , for a c -component mixture with π coexisting phases is given by the phase rule:

$$F = 2 - \pi + c \quad (\text{A.2})$$

Thus, one needs to fix $F - 1$ variables in order to derive a Clapeyron equation. Table A.1 lists a number of Clapeyron equations that describe two-phase equilibria in one-component systems up to three-phase equilibria in ternary mixtures. The symbol σ in Table A.1 indicates that the derivatives are taken along the saturation line.

Table A.1: Clapeyron equations

c	π	F	Equation
1	2	1	$\left(\frac{d p}{d \beta} \right)_{\sigma} = - \frac{\Delta h^{12}}{\beta \Delta v^{12}}$ (The original Clapeyron equation)
2	2	2	$\left(\frac{\partial p}{\partial \beta} \right)_{\xi_2, \sigma} = - \frac{\Delta h^{12}}{\beta \Delta v^{12}}$ $\left(\frac{\partial p}{\partial \xi_2} \right)_{\beta, \sigma} = \frac{\Delta x_2^{12}}{\xi_1 \xi_2 \beta \Delta v^{12}}$ $\left(\frac{\partial \beta}{\partial \xi_2} \right)_{p, \sigma} = \frac{\Delta x_2^{12}}{\xi_1 \xi_2 \Delta h'^{12}}$

Table A.1: Continued

2	3	1	$\begin{bmatrix} \left(\frac{d\beta}{d\xi_2}\right)_\sigma \\ \left(\frac{dp}{d\xi_2}\right)_\sigma \end{bmatrix} = \frac{1}{\xi_1\xi_2} \begin{bmatrix} \Delta h'^{12} & \beta\Delta v^{12} \\ \Delta h'^{13} & \beta\Delta v^{13} \end{bmatrix}^{-1} \begin{bmatrix} \Delta x_2^{12} \\ \Delta x_2^{13} \end{bmatrix}$
3	2	3	$\begin{aligned} \left(\frac{\partial\beta}{\partial\xi_3}\right)_{p,\xi_1,\sigma} &= \frac{\xi_2\Delta x_3^{12} - \xi_3\Delta x_2^{12}}{\xi_2\xi_3\Delta h'^{12}} \\ \left(\frac{\partial\xi_1}{\partial\xi_3}\right)_{p,\beta,\sigma} &= -\frac{\xi_1(\xi_3\Delta x_2^{12} - \xi_2\Delta x_3^{12})}{\xi_3(\xi_1\Delta x_2^{12} - \xi_2\Delta x_1^{12})} \\ \left(\frac{\partial p}{\partial\xi_3}\right)_{\beta,\xi_1,\sigma} &= \frac{\xi_2\Delta x_3^{12} - \xi_3\Delta x_2^{12}}{\xi_2\xi_3\beta\Delta v^{12}} \end{aligned}$
3	3	2	$\begin{aligned} \begin{bmatrix} \left(\frac{\partial\beta}{\partial\xi_3}\right)_{p,\sigma} \\ \left(\frac{\partial\xi_1}{\partial\xi_3}\right)_{p,\sigma} \end{bmatrix} &= \frac{1}{\xi_2\xi_3} \begin{bmatrix} \Delta h'^{12} & \frac{\xi_1\Delta x_2^{12} - \xi_2\Delta x_1^{12}}{\xi_1\xi_2} \\ \Delta h'^{13} & \frac{\xi_1\Delta x_2^{13} - \xi_2\Delta x_1^{13}}{\xi_1\xi_2} \end{bmatrix}^{-1} \begin{bmatrix} \xi_2\Delta x_3^{12} - \xi_3\Delta x_2^{12} \\ \xi_2\Delta x_3^{13} - \xi_3\Delta x_2^{13} \end{bmatrix} \\ \begin{bmatrix} \left(\frac{\partial\beta}{\partial\xi_3}\right)_{\xi_1,\sigma} \\ \left(\frac{\partial p}{\partial\xi_3}\right)_{\xi_1,\sigma} \end{bmatrix} &= \frac{1}{\xi_2\xi_3} \begin{bmatrix} \Delta h'^{12} & \beta\Delta v^{12} \\ \Delta h'^{13} & \beta\Delta v^{13} \end{bmatrix}^{-1} \begin{bmatrix} \xi_2\Delta x_3^{12} - \xi_3\Delta x_2^{12} \\ \xi_2\Delta x_3^{13} - \xi_3\Delta x_2^{13} \end{bmatrix} \\ \begin{bmatrix} \left(\frac{\partial p}{\partial\xi_3}\right)_{\beta,\sigma} \\ \left(\frac{\partial\xi_1}{\partial\xi_3}\right)_{\beta,\sigma} \end{bmatrix} &= \frac{1}{\xi_2\xi_3} \begin{bmatrix} \beta\Delta v^{12} & \frac{\xi_1\Delta x_2^{12} - \xi_2\Delta x_1^{12}}{\xi_1\xi_2} \\ \beta\Delta v^{13} & \frac{\xi_1\Delta x_2^{13} - \xi_2\Delta x_1^{13}}{\xi_1\xi_2} \end{bmatrix}^{-1} \begin{bmatrix} \xi_2\Delta x_3^{12} - \xi_3\Delta x_2^{12} \\ \xi_2\Delta x_3^{13} - \xi_3\Delta x_2^{13} \end{bmatrix} \end{aligned}$

The starting point for the numerical integration of the Clapeyron equations presented in Table A.1 is a pure-component coexistence point. As the Clapeyron equations for mixtures are not defined at zero fugacity fraction of one of the components, they need to be rewritten in the limit of infinite dilution. The results are shown in Table A.2. The variables f_1 and H_2 represent the fugacity of pure component one and Henry's constant of component two, evaluated at the temperature and pressure of the system. The variables f_2 , $H_3^{(12)}$, and $\gamma_2^{(12)}$ are the fugacity of pure component two, Henry's constant of infinitely diluted component three in a ternary mixture, and the activity coefficient of component two in the infinitely diluted ternary mixture, all evaluated at the temperature and pressure of the system.

Table A.2: Clapeyron equations in the limit of infinite dilution of one of the components

c	π	F	Equation
2	2	2	$\left(\frac{\partial p}{\partial \xi_2}\right)_{\beta, \sigma, \xi_2 \rightarrow 0} = \frac{1}{\beta \Delta v^{12}} \Delta \left(\frac{f_1}{H_2}\right)^{12} \quad \left(\frac{\partial \beta}{\partial \xi_2}\right)_{p, \sigma, \xi_2 \rightarrow 0} = \frac{1}{\Delta h'^{12}} \Delta \left(\frac{f_1}{H_2}\right)^{12}$
2	3	1	$\begin{bmatrix} \left(\frac{d\beta}{d\xi_2}\right)_{\sigma, x_2 \rightarrow 0} \\ \left(\frac{dp}{d\xi_2}\right)_{\sigma, x_2 \rightarrow 0} \end{bmatrix} = \begin{bmatrix} \Delta h'^{12} & \beta \Delta v^{12} \\ \Delta h'^{13} & \beta \Delta v^{13} \end{bmatrix}^{-1} \begin{bmatrix} \Delta \left(\frac{f_1}{H_2}\right)^{12} \\ \Delta \left(\frac{f_1}{H_2}\right)^{13} \end{bmatrix}$
3	2	3	$\left(\frac{\partial \beta}{\partial \xi_3}\right)_{p, \xi_1, \sigma, \xi_3 \rightarrow 0} = \frac{\Delta \left(\left[f_2 \gamma_2^{(12)} x_2 \right] / H_3^{(12)} \right)^{12} - \Delta x_2^{12}}{\xi_2 \Delta h'^{12}}$ $\left(\frac{\partial \xi_1}{\partial \xi_3}\right)_{p, \beta, \sigma, \xi_3 \rightarrow 0} = \xi_1 \frac{\Delta \left(\left[f_2 \gamma_2^{(12)} x_2 \right] / H_3^{(12)} \right)^{12} - \Delta x_2^{12}}{\xi_1 \Delta x_2^{12} - \xi_2 \Delta x_1^{12}}$ $\left(\frac{\partial p}{\partial \xi_3}\right)_{\beta, \xi_1, \sigma, \xi_3 \rightarrow 0} = \frac{\Delta \left(\left[f_2 \gamma_2^{(12)} x_2 \right] / H_3^{(12)} \right)^{12} - \Delta x_2^{12}}{\xi_2 \beta \Delta v^{12}}$
3	3	2	$\begin{bmatrix} \left(\frac{\partial \beta}{\partial \xi_3}\right)_{p, \sigma, \xi_3 \rightarrow 0} \\ \left(\frac{\partial \xi_1}{\partial \xi_3}\right)_{p, \sigma, \xi_3 \rightarrow 0} \end{bmatrix} = \frac{1}{\xi_2} \begin{bmatrix} \Delta h'^{12} & \frac{\xi_1 \Delta x_2^{12} - \xi_2 \Delta x_1^{12}}{\xi_1 \xi_2} \\ \Delta h'^{13} & \frac{\xi_1 \Delta x_2^{13} - \xi_2 \Delta x_1^{13}}{\xi_1 \xi_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta \left(\frac{f_2 \gamma_2^{(12)} x_2}{H_3^{(12)}} \right)^{12} - \Delta x_2^{12} \\ \Delta \left(\frac{f_2 \gamma_2^{(12)} x_2}{H_3^{(12)}} \right)^{13} - \Delta x_2^{13} \end{bmatrix}$ $\begin{bmatrix} \left(\frac{\partial \beta}{\partial \xi_3}\right)_{\xi_1, \sigma, \xi_3 \rightarrow 0} \\ \left(\frac{\partial p}{\partial \xi_3}\right)_{\xi_1, \sigma, \xi_3 \rightarrow 0} \end{bmatrix} = \frac{1}{\xi_2} \begin{bmatrix} \Delta h'^{12} & \beta \Delta v^{12} \\ \Delta h'^{13} & \beta \Delta v^{13} \end{bmatrix}^{-1} \begin{bmatrix} \Delta \left(\frac{f_2 \gamma_2^{(12)} x_2}{H_3^{(12)}} \right)^{12} - \Delta x_2^{12} \\ \Delta \left(\frac{f_2 \gamma_2^{(12)} x_2}{H_3^{(12)}} \right)^{13} - \Delta x_2^{13} \end{bmatrix}$ $\begin{bmatrix} \left(\frac{\partial p}{\partial \xi_3}\right)_{\beta, \sigma, \xi_3 \rightarrow 0} \\ \left(\frac{\partial \xi_1}{\partial \xi_3}\right)_{\beta, \sigma, \xi_3 \rightarrow 0} \end{bmatrix} = \frac{1}{\xi_2} \begin{bmatrix} \beta \Delta v^{12} & \frac{\xi_1 \Delta x_2^{12} - \xi_2 \Delta x_1^{12}}{\xi_1 \xi_2} \\ \beta \Delta v^{13} & \frac{\xi_1 \Delta x_2^{13} - \xi_2 \Delta x_1^{13}}{\xi_1 \xi_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta \left(\frac{f_2 \gamma_2^{(12)} x_2}{H_3^{(12)}} \right)^{12} - \Delta x_2^{12} \\ \Delta \left(\frac{f_2 \gamma_2^{(12)} x_2}{H_3^{(12)}} \right)^{13} - \Delta x_2^{13} \end{bmatrix}$

Appendix B. Derivation of a third-order variable-step-size predictor equation and a fourth-order variable-step-size corrector equation

B.1. Third-order predictor equation with variable step size

The integrand of the Clapeyron equation, F , is approximated by an interpolating polynomial of third order:

$$\begin{aligned}
 p(\xi_i) &= p(\xi_{i-1}) + \int_{\xi_{i-1}}^{\xi_i} F(\xi) d\xi \\
 &\approx p(\xi_{i-1}) + \int_{\xi_{i-1}}^{\xi_i} \sum_{k=i-3}^{i-1} F(\xi_k) L_k(\xi) d\xi
 \end{aligned} \tag{B.1}$$

where the L_k 's are given by:

$$\begin{aligned}
 L_{i-3}(\xi) &= \frac{(\xi - \xi_{i-2})(\xi - \xi_{i-1})}{(\xi_{i-3} - \xi_{i-2})(\xi_{i-3} - \xi_{i-1})} \\
 L_{i-2}(\xi) &= \frac{(\xi - \xi_{i-3})(\xi - \xi_{i-1})}{(\xi_{i-2} - \xi_{i-3})(\xi_{i-2} - \xi_{i-1})} \\
 L_{i-1}(\xi) &= \frac{(\xi - \xi_{i-3})(\xi - \xi_{i-2})}{(\xi_{i-1} - \xi_{i-3})(\xi_{i-1} - \xi_{i-2})}
 \end{aligned} \tag{B.2}$$

Actual integration over the polynomial results in:

$$\begin{aligned}
 p(\xi_i) &\approx p(\xi_{i-1}) + \frac{1}{6} \left[\frac{\Delta h_{i-1}^2 (3\Delta h_{i-2} + 2\Delta h_{i-1})}{\Delta h_{i-3} (\Delta h_{i-3} + \Delta h_{i-2})} F(\xi_{i-3}) \right. \\
 &\quad - \frac{\Delta h_{i-1}^2 (3\Delta h_{i-3} + 3\Delta h_{i-2} + 2\Delta h_{i-1})}{\Delta h_{i-3} \Delta h_{i-2}} F(\xi_{i-2}) + \\
 &\quad \left. \frac{\Delta h_{i-1} \left\{ 3 \left(1 + \frac{\Delta h_{i-1}}{\Delta h_{i-2}} \right) (\Delta h_{i-3} + \Delta h_{i-2} + \Delta h_{i-1}) + 3(\Delta h_{i-3} + \Delta h_{i-2}) - \frac{\Delta h_{i-1}^2}{\Delta h_{i-2}} \right\}}{(\Delta h_{i-3} + \Delta h_{i-2})} F(\xi_{i-1}) \right]
 \end{aligned} \tag{B.3}$$

where the variable Δh_{i-1} is defined by:

$$\Delta h_{i-1} \equiv \xi_i - \xi_{i-1} \tag{B.4}$$

Introduction of a new variable $r_i \equiv \Delta h_{i-1}/\Delta h_i$ simplifies expression (B.3):

$$p(\xi_i) \approx p(\xi_{i-1}) + \frac{1}{6} \frac{\Delta h_{i-1}}{r_{i-2} r_{i-1}^2 (r_{i-2} + 1)} \cdot \left\{ \begin{aligned} &(3r_{i-1} + 2)F(\xi_{i-3}) \\ &-(3r_{i-1}r_{i-2} + 3r_{i-1} + 2)(r_{i-2} + 1)F(\xi_{i-2}) \\ &+ r_{i-2} \{3(r_{i-1} + 1)(r_{i-1}r_{i-2} + r_{i-1} + 1) + 3r_{i-1}(r_{i-1}r_{i-2} + r_{i-1}) - 1\} F(\xi_{i-1}) \end{aligned} \right\} \quad (\text{B.5})$$

For a constant step size Δh , this equation reduces to the third-order Adams-Bashforth explicit method [1].

B.2. Fourth-order corrector equation with variable step size

The integrand of the Clapeyron equation is approximated by an interpolating polynomial of fourth order:

$$p(\xi_i) = p(\xi_{i-1}) + \int_{\xi_{i-1}}^{\xi_i} F(\xi) d\xi \quad (\text{B.6})$$

$$\approx p(\xi_{i-1}) + \int_{\xi_{i-1}}^{\xi_i} \sum_{k=i-3}^i F(\xi_k) L_k(\xi) d\xi$$

where the L_k 's are given by:

$$L_{i-3}(\xi) = \frac{(\xi - \xi_{i-2})(\xi - \xi_{i-1})(\xi - \xi_i)}{(\xi_{i-3} - \xi_{i-2})(\xi_{i-3} - \xi_{i-1})(\xi_{i-3} - \xi_i)}$$

$$L_{i-2}(\xi) = \frac{(\xi - \xi_{i-3})(\xi - \xi_{i-1})(\xi - \xi_i)}{(\xi_{i-2} - \xi_{i-3})(\xi_{i-2} - \xi_{i-1})(\xi_{i-2} - \xi_i)}$$

$$L_{i-1}(\xi) = \frac{(\xi - \xi_{i-3})(\xi - \xi_{i-2})(\xi - \xi_i)}{(\xi_{i-1} - \xi_{i-3})(\xi_{i-1} - \xi_{i-2})(\xi_{i-1} - \xi_i)}$$

$$L_i(\xi) = \frac{(\xi - \xi_{i-3})(\xi - \xi_{i-2})(\xi - \xi_{i-1})}{(\xi_i - \xi_{i-3})(\xi_i - \xi_{i-2})(\xi_i - \xi_{i-1})} \quad (\text{B.7})$$

Actual integration over the polynomial gives:

$$\begin{aligned}
p(\xi_i) \approx p(\xi_{i-1}) &+ \frac{1}{12} \left\{ \frac{\Delta h_{i-1}^3 (2\Delta h_{i-2} + \Delta h_{i-1})}{\Delta h_{i-3} (\Delta h_{i-3} + \Delta h_{i-2}) (\Delta h_{i-3} + \Delta h_{i-2} + \Delta h_{i-1})} F(\xi_{i-3}) \right. \\
&- \frac{\Delta h_{i-1}^3 (2\Delta h_{i-3} + 2\Delta h_{i-2} + \Delta h_{i-1})}{\Delta h_{i-3} \Delta h_{i-2} (\Delta h_{i-2} + \Delta h_{i-1})} F(\xi_{i-2}) \\
&- \frac{\Delta h_{i-1} \left\{ 4\Delta h_{i-2} + 2(\Delta h_{i-3} + \Delta h_{i-2} + \Delta h_{i-1}) \frac{(\Delta h_{i-2} + \Delta h_{i-1})}{(\Delta h_{i-3} + \Delta h_{i-2})} - \frac{\Delta h_{i-1}^2}{(\Delta h_{i-3} + \Delta h_{i-2})} \right\}}{\Delta h_{i-2}} F(\xi_{i-1}) \\
&\left. + \frac{\Delta h_{i-1} \left\{ 2\Delta h_{i-2} \frac{(\Delta h_{i-3} + \Delta h_{i-2})}{(\Delta h_{i-2} + \Delta h_{i-1})} + 4(\Delta h_{i-3} + \Delta h_{i-2} + \Delta h_{i-1}) - \frac{\Delta h_{i-1}^2}{(\Delta h_{i-2} + \Delta h_{i-1})} \right\}}{(\Delta h_{i-3} + \Delta h_{i-2} + \Delta h_{i-1})} F(\xi_i) \right\} \quad (B.8)
\end{aligned}$$

Introduction of the variable $r_i \equiv \Delta h_{i-1}/\Delta h_i$ simplifies expression (B.8):

$$\begin{aligned}
p(\xi_i) = p(\xi_{i-1}) &+ \frac{1}{12} \frac{\Delta h_{i-1} (2r_{i-1} + 1)}{r_{i-1}^2 r_{i-2} (r_{i-2} + 1) (r_{i-1} r_{i-2} + r_{i-1} + 1)} F(\xi_{i-3}) \\
&- \frac{1}{12} \frac{\Delta h_{i-1} (2r_{i-1} r_{i-2} + 2r_{i-1} + 1)}{r_{i-1}^2 r_{i-2} (r_{i-1} + 1)} F(\xi_{i-2}) \\
&- \frac{1}{12} \frac{\Delta h_{i-1} \{ 4r_{i-1} (r_{i-1} r_{i-2} + r_{i-1}) + 2(r_{i-1} r_{i-2} + r_{i-1} + 1) (r_{i-1} + 1) - 1 \}}{r_{i-1}^2 (r_{i-2} + 1)} F(\xi_{i-1}) \\
&+ \frac{1}{12} \frac{\Delta h_{i-1} \{ 2r_{i-1} (r_{i-1} r_{i-2} + r_{i-1}) + 4(r_{i-1} r_{i-2} + r_{i-1} + 1) (r_{i-1} + 1) - 1 \}}{(r_{i-1} r_{i-2} + r_{i-1} + 1) (r_{i-1} + 1)} F(\xi_i) \quad (B.9)
\end{aligned}$$

For a constant step size Δh , this equation reduces to the fourth-order Adams-Moulton implicit method [1].

B.3. Reference

1. Burden, R. L. and Douglas Faires, J., *Numerical Analysis*, 7th Ed., Brooks/Cole, Pacific Grove, 2001

Appendix C. Derivation of the Coulombic contribution to the virial

In case the Ewald-summation method is used to compute the electrostatic interaction energy, the reduced (See Chapter 5, Eqs. 5.6 and 5.7) Coulombic contribution can be written as:

$$\begin{aligned}
 \tilde{\mathcal{U}}^{\text{Coulomb}} = & \underbrace{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{\tilde{q}_i^\gamma \tilde{q}_j^\delta \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L})}{\tilde{r}_{i\gamma j\delta}} - \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \frac{\tilde{q}_i^\gamma \tilde{q}_i^\delta \operatorname{erf}(\alpha \tilde{r}_{i\gamma i\delta} / \tilde{L})}{\tilde{r}_{i\gamma i\delta}}}_{\text{real-space term}} + \\
 & \underbrace{\sum_{\mathbf{n} \neq \mathbf{0}} \frac{\exp\left(-\frac{\pi^2 \|\mathbf{n}\|^2}{\alpha^2}\right)}{2\pi \tilde{L} \|\mathbf{n}\|^2} \left\{ \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \cos\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right]^2 + \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \sin\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right]^2 \right\}}_{\text{reciprocal-space term}} + \\
 & \underbrace{-\frac{\alpha}{\tilde{L}\sqrt{\pi}} \sum_{i=1}^N \sum_{\gamma=1}^{n_i} (\tilde{q}_i^\gamma)^2}_{\text{point-self energy}}
 \end{aligned} \tag{C.1}$$

The Coulombic contribution to the virial \tilde{W} , in reduced form, is given by

$$\tilde{W}^{\text{Coulomb}} = -\tilde{V} \frac{\partial \tilde{\mathcal{U}}^{\text{Coulomb}}}{\partial \tilde{V}} \tag{C.2}$$

Thus, the volume derivative of Eq. (C.1) is needed. The position of interaction site γ on molecule i is given by the sum of the center-of-mass vector $\tilde{\mathbf{r}}_i$ of molecule i and the intramolecular vector $\tilde{\mathbf{d}}_{i\gamma}$ (See Figure C.1). The vector $\tilde{\mathbf{r}}_{i\gamma j\delta}$ between site γ on molecule i and site δ on molecule j can be written as follows (See Figure C.1):

$$\tilde{\mathbf{r}}_{i\gamma j\delta} = \tilde{\mathbf{r}}_{ij} + \tilde{\mathbf{d}}_{j\delta} - \tilde{\mathbf{d}}_{i\gamma} = \tilde{\mathbf{r}}_{ij} + \tilde{\mathbf{d}}_{i\gamma j\delta} \tag{C.3}$$

It is only the intermolecular separation vector $\tilde{\mathbf{r}}_{ij}$ that scales with the volume:

$$\tilde{\mathbf{r}}_{i\gamma j\delta} = \tilde{V}^{1/3} \tilde{\mathbf{s}}_{ij} + \tilde{\mathbf{d}}_{i\gamma j\delta} \tag{C.4}$$

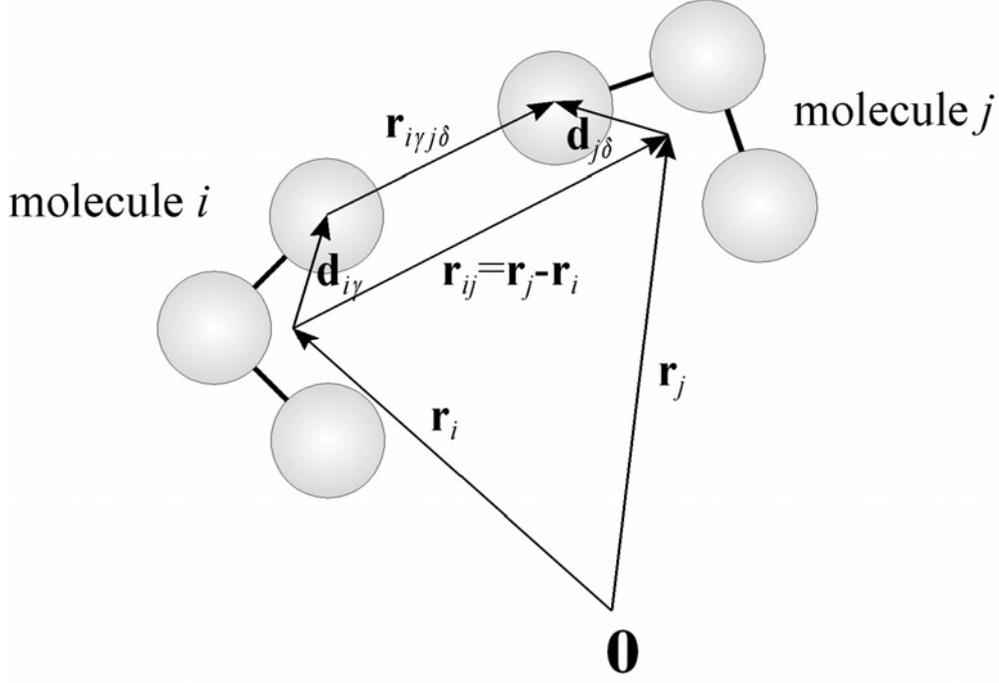


Figure C.1: Distance between interaction sites γ on molecule i and δ on molecule j

The volume derivative of the intermolecular real-space contribution can be worked out as follows:

$$\begin{aligned}
\frac{\partial \tilde{\mathcal{U}}^{\text{real space (inter)}}}{\partial \tilde{V}} &= \frac{\partial}{\partial \tilde{V}} \left\{ \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{\tilde{q}_i^\gamma \tilde{q}_j^\delta \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L})}{\tilde{r}_{i\gamma j\delta}} \right\} = \\
& \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{\partial}{\partial \tilde{r}_{i\gamma j\delta}} \frac{\tilde{q}_i^\gamma \tilde{q}_j^\delta \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L})}{\tilde{r}_{i\gamma j\delta}} \frac{\partial \tilde{r}_{i\gamma j\delta}}{\partial \tilde{\mathbf{r}}_{i\gamma j\delta}} \frac{d\tilde{\mathbf{r}}_{i\gamma j\delta}}{d\tilde{V}} = \\
& - \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{\tilde{q}_i^\gamma \tilde{q}_j^\delta \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L})}{\tilde{r}_{i\gamma j\delta}^2} \frac{\partial \tilde{r}_{i\gamma j\delta}}{\partial \tilde{\mathbf{r}}_{i\gamma j\delta}} \frac{d\tilde{\mathbf{r}}_{i\gamma j\delta}}{d\tilde{V}} \\
& + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \frac{1}{\tilde{r}_{i\gamma j\delta}} \frac{\partial \tilde{q}_i^\gamma \tilde{q}_j^\delta \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L})}{\partial \tilde{r}_{i\gamma j\delta} / \tilde{L}} \frac{\partial (\tilde{r}_{i\gamma j\delta} / \tilde{L})}{\partial (\tilde{\mathbf{r}}_{i\gamma j\delta} / \tilde{L})} \frac{d(\tilde{\mathbf{r}}_{i\gamma j\delta} / \tilde{L})}{d\tilde{V}} = \\
& - \frac{1}{3\tilde{V}} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \tilde{q}_i^\gamma \tilde{q}_j^\delta \left\{ \begin{array}{l} \operatorname{erfc}(\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L}) \frac{\tilde{\mathbf{r}}_{ij} \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}}{\tilde{r}_{i\gamma j\delta}^3} \\ - \frac{2\alpha}{\tilde{L}\sqrt{\pi}} \exp\left(-[\alpha \tilde{r}_{i\gamma j\delta} / \tilde{L}]^2\right) \left(1 - \frac{\tilde{\mathbf{r}}_{ij} \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}}{\tilde{r}_{i\gamma j\delta}^2}\right) \end{array} \right\} \quad (\text{C.5})
\end{aligned}$$

where Eqs. (C.6) and (C.7) have been used.

$$\frac{\partial \tilde{r}_{i\gamma j\delta}}{\partial \tilde{\mathbf{r}}_{i\gamma j\delta}} \frac{d\tilde{\mathbf{r}}_{i\gamma j\delta}}{d\tilde{V}} = \frac{\partial (\tilde{\mathbf{r}}_{i\gamma j\delta} \cdot \hat{\tilde{\mathbf{r}}}_{i\gamma j\delta})}{\partial \tilde{\mathbf{r}}_{i\gamma j\delta}} \frac{d(\tilde{V}^{1/3} \tilde{\mathbf{s}}_{ij} + \tilde{\mathbf{d}}_{i\gamma j\delta})}{d\tilde{V}} = \frac{1}{3\tilde{V}} \frac{\tilde{\mathbf{r}}_{ij} \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}}{\tilde{r}_{i\gamma j\delta}} \quad (\text{C.6})$$

$$\begin{aligned}
\frac{\partial(\tilde{\mathbf{r}}_{i\gamma j\delta}/\tilde{L})}{\partial(\tilde{\mathbf{r}}_{i\gamma j\delta}/\tilde{L})} \frac{d(\tilde{\mathbf{r}}_{i\gamma j\delta}/\tilde{L})}{d\tilde{V}} &= \frac{\partial(\tilde{\mathbf{r}}_{i\gamma j\delta} \cdot \hat{\mathbf{r}}_{i\gamma j\delta})}{\partial\tilde{\mathbf{r}}_{i\gamma j\delta}} \frac{d(\tilde{\mathbf{s}}_{ij} + \tilde{V}^{-1/3}\tilde{\mathbf{d}}_{i\gamma j\delta})}{d\tilde{V}} \\
&= -\frac{1}{3\tilde{V}} \tilde{V}^{-1/3} \tilde{\mathbf{d}}_{i\gamma j\delta} \cdot \hat{\mathbf{r}}_{i\gamma j\delta} \\
&= -\frac{1}{3\tilde{V}} \tilde{V}^{-1/3} \frac{(\tilde{\mathbf{r}}_{i\gamma j\delta} - \tilde{\mathbf{r}}_{ij}) \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}}{\tilde{r}_{i\gamma j\delta}} \\
&= -\frac{1}{3\tilde{V}} \tilde{V}^{-1/3} \left(\tilde{r}_{i\gamma j\delta} - \frac{\tilde{r}_{ij}}{\tilde{r}_{i\gamma j\delta}} \right) \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}
\end{aligned} \tag{C.7}$$

The vector $\hat{\mathbf{r}}_{i\gamma j\delta}$ in Eqs. (C.6) and (C.7) is of unit length. The volume derivative of the intramolecular real-space contribution is given by:

$$\begin{aligned}
\frac{\partial\tilde{\mathcal{U}}^{\text{real space (intra)}}}{\partial\tilde{V}} &= -\frac{\partial}{\partial\tilde{V}} \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \frac{\tilde{q}_i^\gamma \tilde{q}_i^\delta \text{erf}(\alpha\tilde{r}_{i\gamma i\delta}/\tilde{L})}{\tilde{r}_{i\gamma i\delta}} = \\
&= \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \frac{\tilde{q}_i^\gamma \tilde{q}_i^\delta \text{erf}(\alpha\tilde{r}_{i\gamma i\delta}/\tilde{L})}{\tilde{r}_{i\gamma i\delta}^2} \frac{\partial\tilde{r}_{i\gamma i\delta}}{\partial\tilde{\mathbf{r}}_{i\gamma i\delta}} \frac{d\tilde{\mathbf{r}}_{i\gamma i\delta}}{d\tilde{V}} + \\
&\quad - \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \frac{1}{\tilde{r}_{i\gamma i\delta}} \frac{\partial\tilde{q}_i^\gamma \tilde{q}_i^\delta \text{erf}(\alpha\tilde{r}_{i\gamma i\delta}/\tilde{L})}{\partial(\tilde{r}_{i\gamma i\delta}/\tilde{L})} \frac{\partial(\tilde{r}_{i\gamma i\delta}/\tilde{L})}{\partial(\tilde{\mathbf{r}}_{i\gamma i\delta}/\tilde{L})} \frac{d(\tilde{\mathbf{r}}_{i\gamma i\delta}/\tilde{L})}{d\tilde{V}} \\
&= \frac{1}{3\tilde{V}} \frac{2\alpha}{\tilde{L}\sqrt{\pi}} \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \tilde{q}_i^\gamma \tilde{q}_i^\delta \exp\left(-[\alpha\tilde{r}_{i\gamma i\delta}/\tilde{L}]^2\right)
\end{aligned} \tag{C.8}$$

where Eqs. (C.9) and (C.10) have been used:

$$\frac{\partial\tilde{r}_{i\gamma i\delta}}{\partial\tilde{\mathbf{r}}_{i\gamma i\delta}} \frac{d\tilde{\mathbf{r}}_{i\gamma i\delta}}{d\tilde{V}} = \frac{\partial(\tilde{\mathbf{d}}_{i\gamma i\delta} \cdot \hat{\mathbf{d}}_{i\gamma i\delta})}{\partial\tilde{\mathbf{d}}_{i\gamma i\delta}} \frac{d(\tilde{\mathbf{d}}_{i\gamma i\delta})}{d\tilde{V}} = 0 \tag{C.9}$$

$$\begin{aligned}
\frac{\partial(\tilde{r}_{i\gamma i\delta}/\tilde{L})}{\partial(\tilde{\mathbf{r}}_{i\gamma i\delta}/\tilde{L})} \frac{d(\tilde{\mathbf{r}}_{i\gamma i\delta}/\tilde{L})}{d\tilde{V}} &= \frac{\partial(\tilde{\mathbf{d}}_{i\gamma i\delta} \cdot \hat{\mathbf{d}}_{i\gamma i\delta})}{\partial\tilde{\mathbf{d}}_{i\gamma i\delta}} \frac{d(\tilde{V}^{-1/3}\tilde{\mathbf{d}}_{i\gamma i\delta})}{d\tilde{V}} \\
&= -\frac{1}{3\tilde{V}} \tilde{V}^{-1/3} \tilde{\mathbf{d}}_{i\gamma i\delta} \cdot \hat{\mathbf{d}}_{i\gamma i\delta} \\
&= -\frac{1}{3\tilde{V}\tilde{L}} \tilde{r}_{i\gamma i\delta}
\end{aligned} \tag{C.10}$$

In order to take the derivative of the reciprocal-space contribution, the volume dependence of this contribution is revealed:

$$\tilde{\mathcal{U}}^{\text{reciprocal space}} = \sum_{\mathbf{n} \neq \mathbf{0}}^{\infty} \frac{\tilde{V}^{-\frac{1}{3}} \exp\left(-\frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2}\right)}{2\pi \|\mathbf{n}\|^2} \left\{ \begin{array}{l} \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \cos\left(2\pi \mathbf{n} \cdot \left\{ \tilde{V}^{-\frac{1}{3}} \tilde{\mathbf{d}}_{i\alpha} + \tilde{\mathbf{s}}_i \right\}\right) \right]^2 + \\ \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \sin\left(2\pi \mathbf{n} \cdot \left\{ \tilde{V}^{-\frac{1}{3}} \tilde{\mathbf{d}}_{i\alpha} + \tilde{\mathbf{s}}_i \right\}\right) \right]^2 \end{array} \right\} \quad (\text{C.11})$$

The derivative to the reduced volume is given by

$$\begin{aligned} \frac{\partial \tilde{\mathcal{U}}^{\text{reciprocal space}}}{\partial \tilde{V}} &= -\frac{\tilde{V}^{-\frac{1}{3}}}{3\tilde{V}} \sum_{\mathbf{n} \neq \mathbf{0}}^{\infty} \frac{1}{2\pi \|\mathbf{n}\|^2} \exp\left(-\frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2}\right) \left\{ \text{COS}^2 + \text{SIN}^2 \right\} \\ &+ \frac{4\pi}{3\tilde{V}} \sum_{\mathbf{n} \neq \mathbf{0}}^{\infty} \frac{\tilde{V}^{-\frac{2}{3}} \exp\left(-\frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2}\right)}{2\pi \|\mathbf{n}\|^2} \cdot \left\{ \begin{array}{l} \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \mathbf{n} \cdot \tilde{\mathbf{d}}_{i\gamma} \sin\left(2\pi \mathbf{n} \cdot \left\{ \tilde{V}^{-\frac{1}{3}} \tilde{\mathbf{d}}_{i\gamma} + \tilde{\mathbf{s}}_i \right\}\right) \right] \text{COS} - \\ \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \mathbf{n} \cdot \tilde{\mathbf{d}}_{i\gamma} \cos\left(2\pi \mathbf{n} \cdot \left\{ \tilde{V}^{-\frac{1}{3}} \tilde{\mathbf{d}}_{i\gamma} + \tilde{\mathbf{s}}_i \right\}\right) \right] \text{SIN} \end{array} \right\} \end{aligned} \quad (\text{C.12})$$

where SIN and COS are defined by:

$$\text{SIN} \equiv \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \sin\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right] \quad \text{COS} \equiv \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \cos\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right] \quad (\text{C.13})$$

After several simplifications one gets:

$$\begin{aligned} \frac{\partial \tilde{\mathcal{U}}^{\text{reciprocal space}}}{\partial \tilde{V}} &= -\frac{1}{3\tilde{V}} \sum_{\mathbf{n} \neq \mathbf{0}}^{\infty} \frac{\exp\left(-\frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2}\right)}{2\pi \tilde{L} \|\mathbf{n}\|^2} \left\{ \text{COS}^2 + \text{SIN}^2 \right\} \\ &+ \frac{2}{3\tilde{V}} \sum_{\mathbf{n} \neq \mathbf{0}}^{\infty} \frac{\exp\left(-\frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2}\right)}{2\pi \tilde{L} \|\mathbf{n}\|^2} \left\{ \begin{array}{l} \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{d}}_{i\gamma} \sin\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right] \text{COS} - \\ \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{d}}_{i\gamma} \cos\left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma}\right) \right] \text{SIN} \end{array} \right\} \end{aligned} \quad (\text{C.14})$$

The total Coulombic contribution to the virial is given by:

$$\begin{aligned}
3\tilde{W}^{\text{Coulomb}} &= -3\tilde{V} \left\{ \frac{\partial \tilde{U}^{\text{Coulomb}}}{\partial \tilde{V}} \right\} = \\
&\sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{\gamma=1}^{n_i} \sum_{\delta=1}^{n_j} \tilde{q}_i^\gamma \tilde{q}_j^\delta \left\{ \operatorname{erfc} \left(\frac{\alpha \tilde{r}_{i\gamma j\delta}}{\tilde{L}} \right) \frac{\tilde{\mathbf{r}}_{ij} \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}}{\tilde{r}_{i\gamma j\delta}^3} - \frac{2\alpha \exp \left(- \left[\frac{\alpha \tilde{r}_{i\gamma j\delta}}{\tilde{L}} \right]^2 \right)}{\tilde{L}\sqrt{\pi}} \left(1 - \frac{\tilde{\mathbf{r}}_{ij} \cdot \tilde{\mathbf{r}}_{i\gamma j\delta}}{\tilde{r}_{i\gamma j\delta}^2} \right) \right\} \\
&- \frac{2\alpha}{\tilde{L}\sqrt{\pi}} \sum_{i=1}^N \sum_{\gamma=1}^{n_i-1} \sum_{\delta=\gamma+1}^{n_i} \tilde{q}_i^\gamma \tilde{q}_i^\delta \exp \left(- \left[\frac{\alpha \tilde{r}_{i\gamma i\delta}}{\tilde{L}} \right]^2 \right) + \sum_{\mathbf{n} \neq \mathbf{0}} \frac{\exp \left(- \frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2} \right)}{2\pi \tilde{L} \|\mathbf{n}\|^2} \{ \operatorname{COS}^2 + \operatorname{SIN}^2 \} \\
&- 2 \sum_{\mathbf{n} \neq \mathbf{0}} \frac{\exp \left(- \frac{\|\mathbf{n}\|^2 \pi^2}{\alpha^2} \right)}{2\pi \tilde{L} \|\mathbf{n}\|^2} \left\{ \begin{aligned} &\left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{d}}_{i\gamma} \sin \left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma} \right) \right] \operatorname{COS} \\ &- \left[\sum_{i=1}^N \sum_{\gamma=1}^{n_i} \tilde{q}_i^\gamma \frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{d}}_{i\gamma} \cos \left(\frac{2\pi}{\tilde{L}} \mathbf{n} \cdot \tilde{\mathbf{r}}_{i\gamma} \right) \right] \operatorname{SIN} \end{aligned} \right\} \frac{\alpha \sum_{i=1}^N \sum_{\gamma=1}^{n_i} (\tilde{q}_i^\gamma)^2}{\tilde{L}\sqrt{\pi}} \quad (\text{C.15})
\end{aligned}$$

Appendix D. Simulation results

LB: Lorentz/Berthelot combining rules
 LHMCC: Lorentz/Hudson-McCoubrey combining rules

Table D.1: Vapour-liquid equilibria of the binary system ethane/methane at $T = 250$ K

LB (interactions between like molecules), LHMCC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
1.313	0.000	0.000	444	23.3	3.578	0.211	0.578	398	51.5
1.448	0.013	0.084	444	24.6	3.779	0.230	0.594	392	54.6
1.612	0.028	0.168	438	26.4	4.004	0.251	0.611	387	58.4
1.821	0.048	0.255	434	28.8	4.259	0.275	0.627	380	62.7
1.949	0.060	0.297	431	30.2	4.548	0.301	0.644	373	67.4
2.096	0.073	0.342	429	31.9	4.707	0.316	0.652	369	70.0
2.269	0.089	0.386	425	34.1	4.876	0.332	0.660	364	73.0
2.474	0.108	0.429	422	36.4	5.058	0.349	0.667	359	76.3
2.616	0.122	0.455	419	38.0	5.253	0.367	0.675	354	80.1
2.774	0.136	0.482	416	40.1	5.463	0.387	0.682	348	84.9
2.954	0.154	0.508	410	42.6	5.690	0.410	0.688	341	91.2
3.160	0.172	0.534	406	45.4	5.938	0.440	0.694	332	99.9
3.398	0.194	0.560	401	48.8					

Table D.2: Vapour-liquid equilibria of the binary system propane/methane at $T = 270$ K

LB (interactions between like molecules), LHMCC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.432	0.000	0.000	526	9.3	2.886	0.169	0.808	493	31.8
0.454	0.002	0.045	527	9.5	3.344	0.201	0.825	485	36.4
0.478	0.003	0.089	527	9.7	3.631	0.220	0.834	480	39.4
0.505	0.005	0.135	527	10.0	3.971	0.244	0.843	474	43.1
0.536	0.007	0.181	527	10.2	4.173	0.258	0.848	470	45.3
0.570	0.010	0.226	527	10.4	4.382	0.274	0.850	467	47.7
0.609	0.012	0.272	528	10.7	4.601	0.289	0.855	462	50.2
0.654	0.015	0.320	527	11.2	4.831	0.304	0.859	458	53.0
0.707	0.019	0.366	526	11.7	5.073	0.321	0.862	453	56.0
0.769	0.023	0.412	525	12.2	5.327	0.337	0.865	448	59.1
0.843	0.029	0.458	524	12.8	5.593	0.355	0.868	442	62.5
0.933	0.035	0.507	522	13.6	5.872	0.373	0.871	436	66.1
1.045	0.044	0.556	520	14.5	6.166	0.393	0.873	430	70.0
1.188	0.054	0.604	517	15.7	6.474	0.413	0.875	423	74.3
1.379	0.068	0.653	514	17.5	6.798	0.434	0.876	416	79.2
1.646	0.087	0.703	510	20.0	7.138	0.456	0.877	408	84.8
2.046	0.112	0.749	506	23.6	7.495	0.479	0.878	400	91.0
2.393	0.135	0.778	501	27.0	7.870	0.502	0.878	390	97.8

Table D.3: Vapour-liquid equilibria of the binary system ethane/carbon dioxide at different temperatures

<i>T</i> = 207 K LB combining rules for all interactions									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.297	0.000	0.000	507	5.5	0.444	0.552	0.554	788	10.5
0.305	0.009	0.034	511	5.8	0.442	0.660	0.590	868	10.6
0.313	0.020	0.067	515	6.0	0.436	0.726	0.626	928	10.6
0.322	0.031	0.101	519	6.3	0.427	0.798	0.662	984	10.5
0.331	0.044	0.135	524	6.6	0.416	0.852	0.697	1033	10.3
0.341	0.058	0.170	529	6.9	0.405	0.889	0.733	1069	10.1
0.351	0.074	0.204	536	7.2	0.392	0.914	0.768	1098	9.9
0.362	0.092	0.239	544	7.5	0.380	0.935	0.803	1124	9.7
0.373	0.113	0.274	552	7.9	0.369	0.953	0.839	1146	9.5
0.384	0.140	0.309	563	8.3	0.357	0.967	0.875	1163	9.3
0.396	0.171	0.343	576	8.6	0.346	0.978	0.910	1177	9.1
0.407	0.206	0.378	592	9.0	0.336	0.988	0.945	1188	8.9
0.418	0.251	0.413	613	9.4	0.332	0.991	0.960	1192	8.8
0.428	0.306	0.449	645	9.7	0.329	0.994	0.970	1195	8.8
0.436	0.376	0.484	679	10.1	0.326	0.996	0.980	1197	8.7
0.442	0.456	0.519	724	10.3	0.323	0.998	0.990	1199	8.7

<i>T</i> = 207 K LB (interactions between like molecules), LHMCC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.297	0.000	0.000	507	5.5	0.450	0.340	0.483	656	10.4
0.305	0.008	0.034	509	5.8	0.459	0.432	0.519	707	10.8
0.314	0.017	0.068	513	6.1	0.462	0.559	0.555	788	11.0
0.324	0.027	0.101	517	6.3	0.458	0.686	0.591	882	11.0
0.334	0.037	0.135	521	6.6	0.449	0.774	0.626	970	11.0
0.344	0.050	0.169	526	6.9	0.437	0.829	0.662	1012	10.8
0.355	0.066	0.203	531	7.3	0.424	0.873	0.697	1054	10.5
0.366	0.084	0.238	539	7.6	0.411	0.906	0.733	1089	10.3
0.378	0.102	0.273	546	8.0	0.397	0.928	0.769	1114	10.1
0.390	0.125	0.308	555	8.4	0.384	0.946	0.805	1134	9.8
0.403	0.155	0.342	568	8.8	0.372	0.959	0.840	1150	9.6
0.415	0.190	0.377	583	9.2	0.360	0.971	0.875	1163	9.4
0.428	0.228	0.412	600	9.6	0.348	0.981	0.910	1174	9.2
0.440	0.274	0.448	624	10.0					

Table D.3: Continued

<i>T</i> = 250 K LB combining rules for all interactions									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
1.31	0.000	0.000	444	23.3	1.95	0.453	0.541	621	44.9
1.34	0.013	0.031	447	24.5	1.98	0.516	0.577	654	46.2
1.38	0.027	0.063	451	25.4	2.00	0.577	0.613	687	47.2
1.41	0.041	0.095	455	26.5	2.01	0.633	0.650	719	48.0
1.45	0.057	0.128	460	27.6	2.02	0.694	0.688	759	48.7
1.49	0.076	0.160	466	28.7	2.01	0.755	0.726	803	49.0
1.53	0.096	0.193	473	30.0	1.99	0.807	0.763	847	49.1
1.57	0.118	0.225	482	31.4	1.96	0.852	0.799	890	48.8
1.61	0.144	0.258	490	32.7	1.93	0.888	0.836	924	48.1
1.66	0.170	0.292	498	34.1	1.90	0.919	0.873	954	47.3
1.70	0.197	0.327	509	35.5	1.86	0.945	0.909	980	46.6
1.75	0.230	0.362	522	37.1	1.82	0.968	0.944	1007	45.9
1.79	0.268	0.398	538	38.7	1.81	0.978	0.959	1018	45.7
1.84	0.309	0.433	554	40.3	1.80	0.984	0.969	1025	45.5
1.88	0.350	0.470	574	41.9	1.79	0.989	0.979	1032	45.3
1.92	0.398	0.505	594	43.5					

<i>T</i> = 250 K LB (interactions between like molecules), LHMCC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
1.31	0.000	0.000	444	23.3	2.01	0.449	0.539	615	46.3
1.35	0.012	0.031	446	24.6	2.04	0.516	0.576	652	47.8
1.38	0.025	0.063	451	25.6	2.06	0.582	0.614	689	49.0
1.42	0.038	0.094	455	26.5	2.06	0.641	0.651	724	49.8
1.46	0.053	0.127	459	27.6	2.06	0.702	0.689	763	50.3
1.50	0.069	0.160	464	28.7	2.05	0.769	0.726	814	50.4
1.55	0.088	0.193	470	30.0	2.02	0.820	0.763	858	50.1
1.59	0.109	0.225	477	31.4	1.99	0.861	0.800	893	49.5
1.64	0.132	0.259	484	32.9	1.95	0.897	0.837	927	48.6
1.68	0.158	0.293	492	34.5	1.91	0.927	0.874	957	47.7
1.73	0.188	0.327	502	36.2	1.87	0.952	0.909	981	46.8
1.78	0.222	0.361	514	38.0	1.83	0.973	0.945	1005	46.0
1.83	0.257	0.396	527	39.8	1.81	0.981	0.960	1015	45.6
1.88	0.297	0.431	544	41.6	1.80	0.986	0.970	1021	45.4
1.93	0.344	0.467	567	43.0	1.79	0.991	0.980	1027	45.2
1.97	0.395	0.502	587	44.7					

Table D.3: Continued

<i>T</i> = 263.15 K LB combining rules for all interactions									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
1.85	0.000	0.000	420	33.2	2.74	0.560	0.450	578	64.7
1.89	0.035	0.013	421	34.8	2.79	0.595	0.502	603	66.7
1.94	0.070	0.029	426	36.1	2.82	0.630	0.560	640	68.4
1.99	0.105	0.045	431	37.6	2.85	0.665	0.614	667	69.9
2.04	0.140	0.063	436	39.2	2.86	0.700	0.671	699	71.0
2.09	0.175	0.082	443	41.0	2.87	0.735	0.723	733	71.7
2.15	0.210	0.103	451	42.8	2.86	0.770	0.772	766	72.1
2.20	0.245	0.125	459	44.8	2.84	0.805	0.822	805	72.3
2.26	0.280	0.149	465	46.8	2.81	0.840	0.867	847	72.1
2.32	0.315	0.176	473	48.9	2.78	0.875	0.903	883	71.5
2.39	0.350	0.207	482	51.1	2.74	0.910	0.935	916	70.7
2.45	0.385	0.240	493	53.5	2.69	0.945	0.963	945	69.6
2.51	0.420	0.275	505	56.0	2.67	0.960	0.974	955	69.0
2.57	0.455	0.314	523	58.1	2.66	0.970	0.981	961	68.6
2.63	0.490	0.356	537	60.3	2.64	0.980	0.987	967	68.2
2.69	0.525	0.402	556	62.5	2.63	0.990	0.994	974	68.0
<i>T</i> = 263.15 K									
LB (interactions between like molecules), LHMCC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
1.85	0.000	0.000	420	33.2	2.82	0.445	0.533	570	67.5
1.89	0.013	0.031	422	34.2	2.87	0.499	0.570	595	69.4
1.94	0.026	0.062	426	35.7	2.90	0.559	0.608	627	69.8
1.99	0.041	0.093	430	37.2	2.93	0.621	0.646	657	71.5
2.05	0.057	0.124	435	38.8	2.94	0.678	0.684	689	72.8
2.10	0.075	0.156	439	40.7	2.94	0.734	0.721	731	73.5
2.16	0.095	0.188	443	42.7	2.92	0.789	0.759	779	73.6
2.22	0.120	0.220	447	44.5	2.89	0.835	0.797	818	73.1
2.28	0.142	0.253	454	46.7	2.85	0.875	0.836	850	72.5
2.35	0.167	0.287	462	49.1	2.81	0.911	0.873	882	71.8
2.42	0.195	0.321	472	51.5	2.75	0.941	0.908	915	71.0
2.49	0.227	0.356	484	54.0	2.70	0.967	0.944	950	70.0
2.56	0.262	0.390	498	56.7	2.67	0.977	0.960	960	69.4
2.63	0.301	0.423	512	59.2	2.66	0.983	0.970	964	68.9
2.69	0.344	0.459	528	62.2	2.64	0.989	0.980	967	68.3
2.76	0.394	0.496	548	65.1	2.62	0.995	0.990	970	67.7

Table D.4: Vapour-liquid equilibria of the binary system difluoromethane/carbon dioxide at different temperatures

<i>T</i> = 283.15 K LB combining rules for all interactions									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
4.45	1.000	1.000	873	124.7	1.77	0.223	0.503	1011	44.0
4.41	0.990	0.995	865	122.9	1.65	0.189	0.453	1014	41.2
4.37	0.980	0.990	875	121.6	1.55	0.161	0.405	1016	38.7
4.34	0.970	0.985	883	120.3	1.46	0.136	0.358	1017	36.5
4.28	0.956	0.977	892	118.2	1.38	0.113	0.310	1019	34.8
4.18	0.922	0.962	905	114.1	1.31	0.092	0.263	1021	33.6
3.98	0.859	0.936	919	107.2	1.25	0.071	0.217	1023	32.2
3.47	0.724	0.877	930	89.4	1.19	0.053	0.173	1024	30.7
3.08	0.610	0.821	954	76.8	1.14	0.038	0.128	1024	29.7
2.75	0.502	0.765	971	67.9	1.09	0.025	0.086	1023	28.8
2.48	0.428	0.710	983	60.9	1.07	0.017	0.059	1024	28.2
2.25	0.368	0.656	991	55.7	1.04	0.010	0.034	1026	27.3
2.07	0.310	0.604	1001	51.0	1.03	0.005	0.017	1028	26.8
1.91	0.260	0.552	1010	46.9					

<i>T</i> = 283.15 K LB (interactions between like molecules), LHMmC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
4.45	1.000	1.000	873	124.7	1.73	0.178	0.462	1002	43.4
4.12	0.879	0.939	900	112.0	1.54	0.128	0.376	1008	39.1
3.78	0.776	0.891	927	100.6	1.38	0.088	0.291	1015	35.3
3.38	0.637	0.834	931	87.5	1.25	0.057	0.209	1024	32.2
2.93	0.508	0.765	966	75.7	1.20	0.046	0.173	1026	31.0
2.63	0.422	0.706	977	67.1	1.16	0.036	0.137	1026	30.0
2.38	0.352	0.654	984	60.4	1.13	0.028	0.111	1026	29.4
2.13	0.280	0.593	998	53.7	1.10	0.021	0.085	1025	28.8
1.93	0.227	0.533	1002	47.9	1.08	0.017	0.068	1025	28.4

<i>T</i> = 303.15 K LB combining rules for all interactions									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
1.80	0.000	0.000	950	48.9	3.90	0.400	0.636	916	102.4
1.94	0.028	0.079	964	51.7	4.54	0.519	0.724	875	126.6
2.11	0.061	0.162	948	57.2	5.39	0.680	0.820	840	153.1
2.31	0.103	0.249	954	62.3	6.00	0.799	0.880	782	183.9
2.57	0.153	0.334	948	68.6	6.38	0.871	0.920	721	201.8
2.89	0.213	0.431	936	76.4	6.63	0.917	0.947	716	213.4
3.32	0.292	0.529	931	87.3					

Table D.4: Continued

$T = 303.15 \text{ K}$
 LB (interactions between like molecules), LHMcC (interactions between unlike molecules)

p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³
1.80	0.000	0.000	950	48.9	4.06	0.395	0.636	907	108.0
2.03	0.042	0.119	940	56.3	4.94	0.566	0.744	866	151.0
2.34	0.090	0.245	941	62.3	6.11	0.782	0.862	794	204.4
2.61	0.141	0.337	920	67.9	6.47	0.858	0.907	777	213.1
2.95	0.200	0.433	930	76.5	6.81	0.922	0.949	757	226.9
3.42	0.278	0.535	912	89.0					

Table D.5: Vapour-liquid equilibria of the binary system trifluoromethane/carbon dioxide at $T = 283.25 \text{ K}$

LB combining rules for all interactions

p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³
3.090	0.000	0.000	984	157.9	4.405	0.875	0.889	895	131.1
3.268	0.092	0.116	980	164.1	4.435	0.917	0.924	877	129.2
3.459	0.204	0.248	980	147.7	4.449	0.956	0.957	880	126.7
3.599	0.292	0.346	977	149.5	4.449	0.979	0.979	888	125.5
3.760	0.380	0.444	957	147.9	4.448	0.984	0.984	888	125.4
3.944	0.484	0.548	926	145.9	4.448	0.987	0.987	888	125.4
4.132	0.605	0.658	939	143.4	4.449	0.990	0.990	887	125.4
4.293	0.744	0.770	907	139.3	4.450	0.993	0.994	886	125.5

LB (interactions between like molecules), LHMcC (interactions between unlike molecules)

p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³
3.090	0.000	0.000	984	157.9	4.422	0.752	0.778	903	145.9
3.191	0.044	0.060	977	165.8	4.475	0.821	0.834	882	138.8
3.306	0.094	0.125	978	158.0	4.501	0.888	0.891	884	136.1
3.428	0.147	0.191	984	157.5	4.504	0.914	0.914	883	134.0
3.559	0.208	0.262	951	158.1	4.500	0.938	0.936	884	131.3
3.697	0.278	0.339	946	157.0	4.493	0.960	0.958	883	128.4
3.843	0.350	0.418	942	155.5	4.489	0.970	0.968	882	127.2
4.005	0.435	0.500	923	161.0	4.487	0.975	0.974	881	126.6
4.169	0.533	0.589	934	166.9	4.485	0.980	0.979	880	126.0
4.314	0.645	0.681	924	162.7	4.482	0.986	0.984	879	125.5

Table D.6: Vapour-liquid equilibria of the binary system propane/carbon dioxide at different temperatures.

<i>T</i> = 230 K									
LB (interactions between like molecules), LHMcC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.096	0.000	0.000	576	2.3	0.476	0.229	0.820	636	11.7
0.106	0.005	0.097	579	2.5	0.500	0.251	0.831	643	12.3
0.119	0.010	0.196	580	2.8	0.525	0.279	0.842	651	13.0
0.135	0.016	0.294	581	3.2	0.552	0.316	0.853	664	13.7
0.157	0.025	0.391	584	3.7	0.579	0.363	0.863	681	14.4
0.170	0.031	0.440	585	4.1	0.608	0.408	0.873	697	15.2
0.186	0.039	0.489	587	4.4	0.639	0.448	0.883	713	16.0
0.205	0.049	0.539	589	4.9	0.671	0.497	0.894	734	16.9
0.229	0.061	0.589	592	5.5	0.704	0.570	0.904	768	17.8
0.258	0.076	0.639	595	6.2	0.739	0.670	0.916	826	18.8
0.297	0.096	0.690	604	7.2	0.776	0.774	0.931	891	19.8
0.348	0.134	0.739	612	8.4	0.815	0.880	0.953	981	20.9
0.419	0.187	0.789	624	10.2					
<i>T</i> = 270 K									
LB (interactions between like molecules), LHMcC (interactions between unlike molecules)									
p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.432	0.000	0.000	526	9.3	1.692	0.316	0.764	596	39.5
0.476	0.009	0.091	531	10.4	1.777	0.351	0.777	606	41.8
0.531	0.019	0.183	533	11.4	1.866	0.386	0.793	615	44.3
0.601	0.033	0.277	535	13.0	1.959	0.422	0.807	625	46.8
0.691	0.052	0.372	538	15.0	2.057	0.460	0.821	636	49.5
0.748	0.065	0.421	540	16.3	2.160	0.502	0.836	649	52.4
0.815	0.080	0.469	543	17.8	2.268	0.553	0.850	667	55.5
0.895	0.099	0.518	546	19.6	2.381	0.611	0.865	690	58.9
0.992	0.122	0.568	551	22.0	2.500	0.670	0.882	714	62.4
1.113	0.152	0.618	558	24.8	2.625	0.738	0.900	747	66.1
1.265	0.193	0.668	570	28.5	2.756	0.809	0.921	790	70.1
1.462	0.243	0.717	579	33.3	2.894	0.887	0.947	842	74.6
1.535	0.263	0.733	584	35.2	3.039	0.974	0.984	919	80.3
1.612	0.287	0.749	589	37.3					

Table D.6: Continued

$T = 294.26$ K

LB (interactions between like molecules), LHMcC (interactions between unlike molecules)

p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.88	0.000	0.000	494	18.5	3.33	0.423	0.746	569	81.8
0.97	0.012	0.086	493	20.5	3.67	0.497	0.778	586	92.6
1.08	0.028	0.173	498	23.1	3.86	0.538	0.795	594	99.1
1.21	0.047	0.263	502	25.8	4.05	0.579	0.811	603	106.2
1.39	0.073	0.354	506	29.8	4.25	0.625	0.828	613	114.1
1.50	0.090	0.402	508	32.5	4.47	0.679	0.846	625	122.6
1.63	0.110	0.448	511	35.5	4.69	0.732	0.866	641	131.2
1.79	0.135	0.495	515	39.1	4.92	0.784	0.889	664	139.0
1.98	0.166	0.544	521	43.2	5.17	0.846	0.916	701	145.2
2.21	0.204	0.593	527	49.2	5.43	0.902	0.948	715	155.8
2.50	0.251	0.643	532	57.1	5.70	0.960	0.975	733	172.7
2.87	0.330	0.695	547	67.2					

Table D.7: Vapour-liquid equilibria of the binary system DMSO/carbon dioxide at different temperatures

$T = 278.5$ K

LB (interactions between like molecules), LHMcC (interactions between unlike molecules)

p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$	p^{sim}	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$	$\rho_{\text{V}}^{\text{sim}}$
MPa			kg/m ³	kg/m ³	MPa			kg/m ³	kg/m ³
0.0017	0.0000	0.0000	1123	0.1	0.2091	0.0299	0.9914	1122	4.1
0.0026	0.0001	0.3500	1124	0.1	0.2383	0.0336	0.9923	1122	4.6
0.0034	0.0002	0.4965	1124	0.1	0.2771	0.0390	0.9934	1121	5.4
0.0049	0.0005	0.6475	1124	0.1	0.3310	0.0509	0.9942	1123	6.4
0.0068	0.0009	0.7470	1124	0.2	0.3665	0.0559	0.9948	1123	7.2
0.0114	0.0021	0.8486	1124	0.2	0.4107	0.0631	0.9953	1123	8.0
0.0156	0.0032	0.8872	1123	0.3	0.4668	0.0731	0.9958	1123	9.1
0.0213	0.0047	0.9191	1123	0.4	0.5409	0.0856	0.9962	1124	10.7
0.0284	0.0062	0.9393	1123	0.6	0.6433	0.0997	0.9968	1124	12.8
0.0379	0.0066	0.9536	1123	0.8	0.7260	0.1099	0.9972	1123	14.5
0.0568	0.0095	0.9690	1122	1.1	0.7942	0.1184	0.9973	1121	15.9
0.0850	0.0149	0.9798	1123	1.6	0.8767	0.1293	0.9974	1119	17.6
0.1128	0.0191	0.9842	1122	2.2	0.9785	0.1426	0.9975	1115	19.8
0.1404	0.0228	0.9873	1121	2.7	1.1079	0.1565	0.9977	1112	22.7
0.1680	0.0261	0.9891	1121	3.3	1.2786	0.1800	0.9978	1108	26.4
0.1863	0.0272	0.9905	1122	3.6					

Table D.7: Continued

$T = 303.15 \text{ K}$
 LB (interactions between like molecules), LHMCC (interactions between unlike molecules)

p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³	p^{sim} MPa	$x_{\text{CO}_2}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$\rho_{\text{L}}^{\text{sim}}$ kg/m ³	$\rho_{\text{V}}^{\text{sim}}$ kg/m ³
0.0063	0.0000	0.0000	1087	0.2	0.8179	0.0636	0.9900	1089	15.0
0.0097	0.0005	0.3455	1085	0.3	0.9327	0.0757	0.9912	1086	17.1
0.0126	0.0008	0.4882	1086	0.3	1.0847	0.0849	0.9916	1089	20.1
0.0179	0.0014	0.6372	1092	0.4	1.2972	0.1029	0.9925	1085	24.2
0.0249	0.0022	0.7426	1092	0.5	1.4378	0.1117	0.9931	1084	27.0
0.0415	0.0041	0.8439	1092	0.8	1.6135	0.1211	0.9937	1083	30.6
0.0563	0.0070	0.8838	1091	1.1	1.8410	0.1342	0.9941	1082	35.7
0.0769	0.0092	0.9177	1091	1.4	2.1468	0.1601	0.9945	1082	42.3
0.1022	0.0116	0.9366	1090	1.9	2.5735	0.1945	0.9949	1082	52.6
0.3283	0.0248	0.9791	1086	5.9	2.9268	0.2148	0.9952	1081	61.3
0.4384	0.0407	0.9826	1087	8.0	3.2241	0.2325	0.9953	1080	68.6
0.6564	0.0517	0.9877	1088	11.9	3.5900	0.2562	0.9952	1075	78.6
0.7281	0.0551	0.9890	1091	13.3					

Table D.8: Vapour-liquid equilibria of the ternary system propane/methane/carbon dioxide at $T = 270 \text{ K}$ and $p = 2.8 \text{ MPa}$

LB (interactions between like molecules), LHMCC (interactions between unlike molecules)

$x_{\text{CH}_4}^{\text{sim}}$	$x_{\text{CO}_2}^{\text{sim}}$	$x_{\text{C}_3\text{H}_8}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$y_{\text{C}_3\text{H}_8}^{\text{sim}}$	$x_{\text{CH}_4}^{\text{sim}}$	$x_{\text{CO}_2}^{\text{sim}}$	$x_{\text{C}_3\text{H}_8}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$y_{\text{C}_3\text{H}_8}^{\text{sim}}$
0.1631	0.0000	0.8369	0.8038	0.0000	0.1962	0.0426	0.4814	0.4760	0.2079	0.6551	0.1370
0.1426	0.0515	0.8059	0.7085	0.1023	0.1892	0.0351	0.5315	0.4334	0.1680	0.7036	0.1284
0.1334	0.0785	0.7881	0.6575	0.1543	0.1882	0.0293	0.5675	0.4032	0.1443	0.7328	0.1229
0.1246	0.1074	0.7680	0.6100	0.2049	0.1851	0.0231	0.6101	0.3668	0.1215	0.7612	0.1173
0.1161	0.1404	0.7435	0.5644	0.2550	0.1806	0.0194	0.6386	0.3420	0.1071	0.7795	0.1134
0.1069	0.1753	0.7178	0.5180	0.3060	0.1760	0.0167	0.6642	0.3191	0.0935	0.7972	0.1093
0.0959	0.2084	0.6957	0.4704	0.3593	0.1703	0.0140	0.6857	0.3003	0.0797	0.8156	0.1047
0.0858	0.2478	0.6664	0.4238	0.4098	0.1664	0.0112	0.7076	0.2812	0.0660	0.8342	0.0998
0.0760	0.2891	0.6349	0.3793	0.4592	0.1615	0.0087	0.7306	0.2607	0.0526	0.8529	0.0945
0.0667	0.3304	0.6029	0.3359	0.5091	0.1550	0.0064	0.7544	0.2392	0.0394	0.8715	0.0891
0.0582	0.3739	0.5679	0.2927	0.5589	0.1484	0.0053	0.7665	0.2282	0.0329	0.8807	0.0864
0.0503	0.4245	0.5252	0.2501	0.6075	0.1424	0.0043	0.7786	0.2171	0.0264	0.8899	0.0837

Table D.9: Vapour-liquid equilibria of the ternary system ethane/methane/carbon dioxide at $T = 250$ K and different pressures

LB (interactions between like molecules), LHMmC (interactions between unlike molecules)											
$p = 1.9$ MPa						$p = 2.128$ MPa					
$x_{\text{CH}_4}^{\text{sim}}$	$x_{\text{CO}_2}^{\text{sim}}$	$x_{\text{C}_2\text{H}_6}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$y_{\text{C}_2\text{H}_6}^{\text{sim}}$	$x_{\text{CH}_4}^{\text{sim}}$	$x_{\text{CO}_2}^{\text{sim}}$	$x_{\text{C}_2\text{H}_6}^{\text{sim}}$	$y_{\text{CH}_4}^{\text{sim}}$	$y_{\text{CO}_2}^{\text{sim}}$	$y_{\text{C}_2\text{H}_6}^{\text{sim}}$
0.0553	0.0000	0.9447	0.2815	0.0000	0.7185	0.0763	0.0000	0.9237	0.3505	0.0000	0.6495
0.0479	0.0223	0.9298	0.2481	0.0464	0.7055	0.0685	0.0264	0.9051	0.3170	0.0478	0.6352
0.0417	0.0477	0.9106	0.2129	0.0933	0.6938	0.0604	0.0551	0.8845	0.2821	0.0952	0.6227
0.0350	0.0754	0.8896	0.1801	0.1402	0.6797	0.0535	0.0861	0.8604	0.2495	0.1427	0.6078
0.0287	0.1042	0.8671	0.1483	0.1866	0.6651	0.0472	0.1190	0.8338	0.2175	0.1911	0.5914
0.0227	0.1348	0.8425	0.1171	0.2328	0.6501	0.0402	0.1532	0.8066	0.1863	0.2399	0.5738
0.0168	0.1713	0.8119	0.0881	0.2813	0.6306	0.0335	0.1888	0.7777	0.1554	0.2889	0.5557
0.0114	0.2080	0.7806	0.0602	0.3291	0.6107	0.0269	0.2311	0.7420	0.1269	0.3359	0.5372
0.0062	0.2435	0.7503	0.0326	0.3783	0.5891	0.0210	0.2797	0.6993	0.0997	0.3838	0.5165
0.0010	0.2827	0.7163	0.0054	0.4255	0.5691	0.0155	0.3368	0.6477	0.0741	0.4331	0.4928
0.0000	0.9352	0.0648	0.0000	0.8852	0.1148	0.0114	0.3874	0.6012	0.0564	0.4728	0.4708
0.0007	0.9453	0.0540	0.0085	0.8929	0.0986	0.0080	0.4379	0.5541	0.0414	0.5128	0.4458
0.0014	0.9544	0.0442	0.0170	0.9009	0.0821	0.0059	0.5091	0.4850	0.0298	0.5535	0.4167
0.0021	0.9626	0.0353	0.0254	0.9086	0.0660	0.0037	0.5648	0.4315	0.0212	0.5952	0.3836
0.0027	0.9707	0.0266	0.0339	0.9161	0.0500	0.0026	0.6380	0.3594	0.0163	0.6378	0.3459
0.0032	0.9790	0.0178	0.0425	0.9229	0.0346	0.0026	0.6983	0.2991	0.0171	0.6702	0.3127
0.0036	0.9871	0.0093	0.0513	0.9291	0.0196	0.0028	0.7181	0.2791	0.0184	0.6817	0.2999
						0.0030	0.7362	0.2608	0.0204	0.6928	0.2868
						0.0033	0.7541	0.2426	0.0228	0.7042	0.2730
						0.0034	0.7766	0.2200	0.0259	0.7149	0.2592
						0.0039	0.7966	0.1995	0.0300	0.7271	0.2429
						0.0044	0.8181	0.1775	0.0350	0.7394	0.2256
						0.0049	0.8400	0.1551	0.0411	0.7517	0.2072
						0.0055	0.8605	0.1340	0.0486	0.7640	0.1874
						0.0061	0.8789	0.1150	0.0573	0.7762	0.1665
						0.0068	0.8962	0.0970	0.0669	0.7882	0.1449
						0.0076	0.9135	0.0789	0.0776	0.8004	0.1220
						0.0084	0.9313	0.0603	0.0896	0.8129	0.0975
						0.0092	0.9495	0.0413	0.1034	0.8258	0.0708
						0.0101	0.9673	0.0226	0.1194	0.8400	0.0406
$p = 2.533$ MPa						$p = 3.040$ MPa					
0.1138	0.0000	0.8862	0.4405	0.0000	0.5595	0.1608	0.0000	0.8392	0.5195	0.0000	0.4805
0.1082	0.0310	0.8608	0.4034	0.0484	0.5482	0.1509	0.0358	0.8133	0.4832	0.0495	0.4673
0.0976	0.0634	0.8390	0.3695	0.0975	0.5330	0.1383	0.0767	0.7850	0.4464	0.1008	0.4528
0.0881	0.1024	0.8095	0.3367	0.1454	0.5179	0.1265	0.1242	0.7493	0.4149	0.1516	0.4335
0.0796	0.1419	0.7785	0.3054	0.1941	0.5005	0.1158	0.1734	0.7108	0.3842	0.2019	0.4139
0.0713	0.1879	0.7408	0.2746	0.2442	0.4812	0.1063	0.2315	0.6622	0.3548	0.2525	0.3927
0.0629	0.2423	0.6948	0.2451	0.2954	0.4595	0.0975	0.2979	0.6046	0.3270	0.3055	0.3675
0.0540	0.2990	0.6470	0.2182	0.3473	0.4345	0.0868	0.3755	0.5377	0.3013	0.3603	0.3384
0.0468	0.3702	0.5830	0.1954	0.3983	0.4063	0.0751	0.4697	0.4552	0.2812	0.4155	0.3033
0.0406	0.4481	0.5113	0.1772	0.4504	0.3724	0.0643	0.5839	0.3518	0.2702	0.4698	0.2600
0.0349	0.5277	0.4374	0.1624	0.5031	0.3345	0.0574	0.6975	0.2451	0.2718	0.5254	0.2028
0.0291	0.6438	0.3271	0.1556	0.5580	0.2864	0.0522	0.7606	0.1872	0.2759	0.5524	0.1717
0.0260	0.7510	0.2230	0.1629	0.6160	0.2211	0.0496	0.8149	0.1355	0.2868	0.5775	0.1357
0.0250	0.8497	0.1253	0.1839	0.6651	0.1510	0.0459	0.8445	0.1096	0.2953	0.5902	0.1145
0.0265	0.9549	0.0186	0.2606	0.7134	0.0260	0.0453	0.8735	0.0812	0.3042	0.6033	0.0925
						0.0442	0.9270	0.0288	0.3354	0.6279	0.0367

Samenvatting en aanbevelingen

Moleculaire simulatie van binaire en ternaire vloeistof-damp evenwichten

Voor procesontwerp en efficiëntieverbeteringen van bestaande processen is kennis van de fysische eigenschappen van zuivere stoffen en mengsels onontbeerlijk. Veel fysische eigenschappen voor zuivere stoffen zijn in de literatuur terug te vinden. Eigenschappen van mengsels zijn echter minder vaak gemeten en gedocumenteerd. Bovendien komt het vaak voor dat grootheden slechts gemeten zijn in een beperkt gebied van fysische condities. Experimentele apparatuur die bij extreme condities, zoals zeer hoge of zeer lage temperaturen en drukken, gebruikt kan worden, is zeer duur.

De laatste jaren is de rekenkracht van computers met sprongen vooruit gegaan. Het is dan ook mogelijk om rekenintensieve methoden als moleculaire simulatie in plaats van echte experimenten in te zetten om fysische grootheden te voorspellen, bijvoorbeeld bij condities waar experimenten simpelweg te duur of niet uitvoerbaar zijn.

In dit proefschrift wordt moleculaire simulatie gebruikt voor het voorspellen van vloeistof-damp evenwichten van binaire en ternaire mengsels. De nadruk ligt niet zo zeer op voorspelling bij extreme fysische condities als wel op het gemak van de simulatiemethode en op de mate waarin experimentele data kunnen worden gereproduceerd. De simulatiemethode die centraal staat is Gibbs-Duhem integratie.

In hoofdstuk twee werden enkele moleculaire simulatiemethoden, die gebruikt kunnen worden voor het voorspellen van vloeistof-damp evenwichten, besproken. Eén van die methoden is Gibbs-Duhem integratie. Een Gibbs-Duhem vergelijking geeft aan hoe toestandsgrootheden in een individuele fase van elkaar afhangen. Uit de Gibbs-Duhem vergelijkingen van fasen die met elkaar in evenwicht zijn kan een zogenaamde Clapeyronvergelijking worden afgeleid.

Gibbs-Duhem integratie komt neer op de numerieke integratie van de Clapeyronvergelijking vanaf een vooraf berekend coëxistentiepunt. De integrand van de Clapeyronvergelijking wordt berekend met behulp van moleculaire simulatie.

Voor het uitvoeren van moleculaire simulatie is een statistisch-mechanisch ensemble nodig. Dit ensemble geeft de waarschijnlijkheid om het te simuleren systeem in een bepaalde configuratie tegen te komen. Gibbs-Duhem integratie kan goed worden gecombineerd met het semigroot-canoniek ensemble dat bij uitstek geschikt is voor het simuleren van mengsels bij constante druk en temperatuur. Druk en temperatuur zijn experimenteel makkelijk toegankelijke variabelen zodat er voldoende experimentele gegevens zijn om de simulatieresultaten mee te vergelijken. De positie van het semigroot-canoniek ensemble onder de meer bekende ensembles werd uitgewerkt.

De meest algemene vorm van de Clapeyronvergelijking werd afgeleid. Via deze algemene vergelijking en de Gibbs-Duhem vergelijking, uitgedrukt in semigrote variabelen, werd een aantal, voor dit werk relevante, Clapeyronvergelijkingen afgeleid.

De statistische mechanica geeft de waarschijnlijkheid om een systeem in een bepaalde moleculaire configuratie aan te treffen. Het efficiënt genereren van moleculaire configuraties via de statistisch-mechanische waarschijnlijkheidsverdeling wordt uitgevoerd met behulp van Monte Carlo importance sampling en de Metropolis methode. De grondbeginselen van Monte Carlo importance sampling werden uitgelegd in hoofdstuk drie.

Het random kiezen van nieuwe configuraties is vaak verre van efficiënt. De efficiëntie kan worden verbeterd door toepassing van een zogenaamde biasmethode die preferente configuraties zoekt met behulp van informatie over de huidige configuratie. Bestaande biasmethoden die veelvuldig in dit proefschrift werden gebruikt zijn de configurationele biasmethode en de moment/kracht biasmethode. Er werd een nieuwe biasmethode ontwikkeld voor het efficiënt genereren van mengselsamenstellingen in het semigroot-canoniek ensemble. Deze biasmethode kan een aanzienlijke verbetering bewerkstelligen in vloeistoffen met hoge dichtheid en/of met moleculen die verschillen in grootte, vorm en krachtenveldparameters.

Numerieke integratie van de Clapeyronvergelijking vereist een initieel coëxistentiepunt en een initiële waarde voor de integrand. De berekening van deze randvoorwaarden werd beschreven in hoofdstuk vier. Geschikte methoden om het startpunt te berekenen zijn Gibbs ensemble simulaties en toepassing van de testmoleculenmethode in het $Np\beta$ -ensemble.

De Gibbs ensemble methode is verreweg de meest eenvoudige manier om fasenevenwicht te berekenen. De methode is echter niet toepasbaar wanneer het verschil tussen de dichtheid van de vloeistof- en de dampfase groot is. Ook bij hoge vloeistofdichtheid is de methode niet toepasbaar omdat succesvolle uitwisselingen van moleculen tussen de fasen schaars zijn.

In de meer rekenintensieve testmoleculenmethode in het $Np\beta$ -ensemble worden onafhankelijke simulaties voor de vloeistof- en de dampfase gebruikt zodat het dichtheidsverschil tussen de fasen geen limitatie is. De bestaande testmoleculenmethode in het $Np\beta$ -ensemble werd verbeterd. De verbeterde versie gaat efficiënt om met gegenereerde vloeistofconfiguraties, heeft niet noodzakelijkerwijs een goede initiële schatting nodig om de juiste verzadigingsdruk te voorspellen en is toepasbaar bij relatief hoge vloeistofdichtheden.

Er werden drie methoden onderzocht voor het berekenen van de initiële integrand in de limiet van een oneindig verdunde oplossing: de Widom testmoleculenmethode en de verschilmethode met of zonder gebruik van overlappende verdelingen. De Widom testmoleculenmethode is, met behulp van andere technieken, toepasbaar bij relatief hoge vloeistofdichtheden. Van de drie methoden is de verschilmethode het makkelijkst te implementeren in een simulatie. Theoretisch en via simulaties werd echter aangetoond dat de methode niet accuraat is wanneer moleculen van het oplosmiddel groter zijn dan die van de opgeloste stof.

De verschilmethode met overlappende verdelingen is rekenintensiever dan de normale verschilmethode. De methode is echter toepasbaar bij hogere dichtheden en geeft uitsluitel over de mogelijkheid om een betrouwbaar antwoord te berekenen. De verschilmethode met overlappende verdelingen is dus duidelijk de veiligste methode voor het berekenen van de initiële integrand.

Om moleculaire simulaties te kunnen uitvoeren moet de interactie tussen moleculen gekwantificeerd worden. De interactie wordt meestal uitgedrukt in een zogenaamd krachtenveld. In hoofdstuk vijf werd het belang van de keuze van een goed krachtenveld benadrukt.

Voor vele stoffen in een scala aan krachtenvelden beschikbaar, welke verschillen in wiskundige complexiteit, fysische relevantie en overdraagbaarheid van de bijbehorende parameters. Het voorspellend vermogen is echter vaak gelimiteerd omdat vele soorten interacties, waaronder de niet-additieve, effectief worden uitgedrukt in een paarsgewijs additief krachtenveld. Bovendien zijn veel krachtenveldparameters geoptimaliseerd aan slechts enkele fysische eigenschappen en in een klein bereik van toestandsgrootheden. In dit proefschrift was het primaire selectie criterium voor een geschikt krachtenveld een goede overeenkomst tussen simulatieresultaten en experimentele vloeistof-damp evenwichten van de zuivere stof.

De interacties tussen ongelijke moleculen worden uitgedrukt in de interacties tussen identieke moleculen via zogenaamde combinatieregels. Deze combinatieregels hebben een grote invloed op het fasengedrag van mengsels. Omdat de conventionele Berthelot combinatieregels over het algemeen een te sterke interactie tussen ongelijke moleculen voorspelt, werd in dit proefschrift een andere combinatieregels voorgesteld. De interacties tussen moleculen van dezelfde soort zijn echter vaak gekwantificeerd met behulp van de Lorentz-Berthelot combinatieregels, zodat verschillende combinatieregels voor verschillende soorten interacties gebruikt moeten worden.

Het potentieel van de Gibbs-Duhem integratiemethode werd getest door het voorspellen van de vloeistof-damp evenwichten van enkele binaire mengsels. De coëxistentiepunten van de zuivere stoffen, *i.e.* de ophangpunten van het binaire diagram, werden separaat berekend. In het geval beide componenten subkritisch waren, traceerde de Gibbs-Duhem integratiemethode de coëxistentielijn moeiteloos van het initiële punt naar het coëxistentiepunt van de tweede component.

Een deel van de simulatieresultaten werd verkregen met zeer simpele en fysisch weinig realistische krachtevelden en Lorentz-Berthelot combinatieregels. De overeenkomst met experimentele data was echter uitstekend.

De Gibbs-Duhem integratiemethode werd onbetrouwbaar wanneer de numerieke integratie een kritisch punt naderde en/of wanneer het initiële punt te veel afweek van het ware coëxistentiepunt.

In hoofdstuk zes werd een verbeterde Gibbs-Duhem integratiemethode geïntroduceerd. Verbeteringen waren nodig omdat de conventionele Gibbs-Duhem integratiemethode zeer inefficiënt omgaat met gegenereerde configuraties. Alle informatie die wordt verzameld tijdens het voorspellen van een verzadigingsdruk draagt niet bij tot de uiteindelijke waarden voor dichtheid en samenstelling bij fasenevenwicht. Bovendien staat het aantal benodigde simulaties voor het bepalen van die verzadigingsdruk niet van tevoren vast.

In dit proefschrift werd een combinatie van de multihistogrammethode en Gibbs-Duhem integratie, die de bovengenoemde nadelen ondervangt, gepresenteerd. In deze *geavanceerde Gibbs-Duhem integratiemethode* draagt elke gegenereerde configuratie bij aan het te voorspellen coëxistentiepunt. Een ander belangrijk voordeel is de retroactieve werking van de methode; *i.e.* een actuele simulatie verbetert ook de voorspelling van eerdere coëxistentiepunten.

Met de methode kan een complete coëxistentielijn berekend worden. Het is ook mogelijk om een coëxistentiepunt bij een bepaalde samenstelling of dichtheid te berekenen. Dit is een voordeel ten opzichte van de conventionele Gibbs-Duhem integratiemethode omdat, vanuit een experimentele invalshoek, een samenstelling of dichtheid een interessantere variabele is dan een fugaciteitsfractie.

De geavanceerde Gibbs-Duhem integratiemethode werd toegepast op een aantal binaire systemen. Over het algemeen kwamen de simulatieresultaten goed overeen met experimentele data. In de meeste gevallen verbeterde het gebruik van verschillende combinatieregels voor interacties tussen gelijke en ongelijke moleculen de voorspelling. De retroactieve werking van de geavanceerde Gibbs-Duhem integratiemethode resulteert in 'gladde' simulatieresultaten. Het vaste aantal simulaties maakt de geavanceerde methode veel sneller dan de conventionele methode.

In hoofdstuk zeven werd de uitbreiding van de geavanceerde Gibbs-Duhem integratiemethode naar ternaire mengsels besproken. Uitbreiding tot ternaire en multicomponentmengsels is interessant omdat industriële mengsels vaak uit meer dan twee componenten bestaan. Een probleem is dat multicomponentsimulaties ook multidimensionale histogrammen met zich

meebrengen. De multihistogrammethode kan dan een groot deel van de simulatietijd opslokken. Voor het bepalen van vloeistof-damp fasenevenwicht zijn echter slechts tweedimensionale histogrammen nodig. De rest van de dimensies hoeft geen deel uit te maken van de multihistogrammethode.

Toepassing van de geavanceerde Gibbs-Duhem integratiemethode op enkele ternaire systemen resulteerde in ‘gladde’ voorspellingen van vloeistof-damp evenwichten.

De startpunten voor ternaire simulaties waren simulatieresultaten van binaire Gibbs-Duhem integraties. Het kan echter zinvol zijn om het startpunt via een Gibbs ensemble simulatie of met de $Np\beta$ +testmoleculenmethode te berekenen omdat deze methoden het voorspelde coëxistentiepunt op een directere manier aan de voorwaarden voor fasenevenwicht toetsen. Bovendien vereisen deze methoden slechts één simulatie in plaats van een hele serie.

In dit proefschrift werden verschillende facetten van Gibbs-Duhem integratie uitgewerkt. Tijdens de toepassing van de methode kwam een aantal problemen aan het licht. Een deel van die problemen werd ondervangen door Gibbs-Duhem integratie te combineren met andere simulatietechnieken. De computercode voor Gibbs-Duhem integratie en voor de berekening van het initiële coëxistentiepunt en de initiële integrand zijn geschikt om door anderen gebruikt te worden voor het voorspellen van vloeistof-damp evenwichten. De onderzoeksresultaten bieden echter ook aangrijpingspunten voor verdere verbeteringen die direct of indirect met Gibbs-Duhem integratie te maken hebben. Uit het onderzoek kwamen de volgende aanbevelingen voort:

- 1) Uit foutenanalyses van beide Gibbs-Duhem integratietechnieken kwam naar voren dat de fout in het initiële coëxistentiepunt een negatieve weerslag heeft op de voorspelde coëxistentielijn. Beide Gibbs-Duhem integratiemethoden hebben geen interne controlepunten waarin de voorspelling direct aan de voorwaarden voor fasenevenwicht worden getest. Het verdient aanbeveling om de voorspelling om de zoveel integratiestappen te toetsen.
- 2) Uit dit onderzoek is gebleken dat de conventionele Berthelot combinatieregels een te sterke interactie tussen ongelijke moleculen voorspelt. In dit proefschrift werd aangetoond dat het gebruik van verschillende combinatieregels voor interacties tussen gelijke en ongelijke moleculen een aanzienlijke verbetering in de voorspelling kan bewerkstelligen. Het gebruik van verschillende combinatieregels is echter inconsequent. Het zou dan ook beter zijn om krachtenveldparameters van zuivere stoffen met betere combinatieregels te optimaliseren zodat slechts één set combinatieregels voor alle interacties kan worden gebruikt.
- 3) Er bestaat een aantal ver doorgevoerde overdraagbare krachtenvelden. Helaas gaat de overdraagbaarheid van de parameters vaak ten koste van de voorspelling van individuele fysische grootheden. Vooral de voorspelling van de verzadigde dampdruk wijkt vaak veel af van experimentele waarden. Daarom zijn deze modellen niet geschikt voor kwantitatieve voorspellingen van vloeistof-damp evenwichten. Niet-overdraagbare krachtenvelden zijn vaak zeer goed geoptimaliseerd aan fysische eigenschappen. Helaas zien verschillende krachtenvelden voor individuele componenten er soms zo verschillend uit dat men met recht kan twijfelen aan de fysische betekenis van die modellen. De vraag is of er ooit onderzoekers zijn die zo'n model willen gebruiken. Het is wellicht zinvol om tot een bepaalde consensus te komen over hoe een goed krachtenveld er zou moeten uitzien en volgens welke procedures de parameters geoptimaliseerd moeten worden. Het is wellicht ook zinvol om parameters aan afgeleide grootheden te optimaliseren omdat deze grootheden veel gevoeliger zijn voor de parameters.

4) De geavanceerde Gibbs-Duhem integratiemethode resulteert in een waarschijnlijkheidsverdeling die de kans aangeeft dat een systeem, gegeven de thermodynamische toestand, aangetroffen wordt in een bepaalde configuratie. Het moet dus mogelijk zijn om met behulp van deze waarschijnlijkheidsverdeling op een eenvoudige manier allerlei fysische eigenschappen, zoals warmtecapaciteiten, isotherme compressibilitieten, en isobare expansiviteiten op en in de buurt van de coëxistentielijn te berekenen. Dit vereist hoogstwaarschijnlijk veel langere simulaties dan degenen die in dit werk werden gepresenteerd. Verder zal de correlatie tussen verschillende configuraties mee moeten worden genomen in de multihistogrammethode.

Aleidus van 't Hof

Summary and recommendations

Molecular Simulation of Binary and Ternary Vapour-Liquid Equilibria

Knowledge of physical properties of pure components and mixtures is essential when designing new processes or improving the efficiency of existing processes. Many pure-component physical properties can be retrieved from literature. However, mixture properties are less often measured and documented. Furthermore, the mixture properties that are available are often measured only at moderate physical conditions. Experimental equipment that can be used at extreme conditions, like very high or very low temperatures and pressures, is very expensive.

Computer power has increased considerably over the last years. Thus, it is possible to exploit computationally demanding methods like molecular simulation to predict physical properties. This is particularly attractive at conditions where real experiments are expensive or impracticable. In this thesis, molecular simulation is used to predict vapour-liquid equilibria of binary and ternary mixtures. Instead of focusing on predictions at extreme conditions, this work rather concentrates on the convenience of the simulation method and the extent to what experimental data can be reproduced. The simulation method that is in the centre of interest is Gibbs-Duhem integration.

In chapter two, several molecular simulation methods that can be used to predict vapour-liquid equilibria are discussed. Among them is Gibbs-Duhem integration. A Gibbs-Duhem equation expresses the mutual dependence of state variables in a single phase. Combination of the Gibbs-Duhem equations of coexisting phases enables one to derive a so-called Clapeyron equation.

Gibbs-Duhem integration implies the numerical integration of the Clapeyron equation from a predetermined starting point. The integrand of the Clapeyron equation is computed with molecular simulation.

In order to perform molecular simulations, one needs a statistical-mechanical ensemble. This ensemble provides the probability to observe the system under study in a particular configuration. Gibbs-Duhem integration can be conveniently combined with the semigrand-canonical ensemble, which is exquisitely suitable for the simulation of mixtures at constant pressure and temperature. As temperature and pressure are variables that are readily accessible in experiments, there are sufficient experimental data to compare the simulation results with. The position of the semigrand-canonical ensemble among other well-known ensembles has been clarified in chapter two.

The most general form of the Clapeyron equation was derived. Combination of this general form with the Gibbs-Duhem equation, expressed in semigrand variables, resulted in a number of Clapeyron equations of which two have actually been used in this work.

Statistical mechanics provides the probability to observe a system in a particular molecular configuration. Efficient generation of molecular configurations in accordance with the statistical-mechanical probability distribution is performed with Monte Carlo importance sampling and the Metropolis method. The basic principles of Monte Carlo importance sampling were explained in chapter three.

Choosing new configurations totally randomly can be very inefficient. The efficiency can be improved by exploiting a bias method that chooses favourable configurations using information about the present configuration. Existing bias methods that were extensively used

in this work are the configurational-bias method and the torque/force-bias method. A new bias method was developed in order to improve composition sampling in the semigrand-canonical ensemble. The improvement of the composition sampling can be considerable in liquid phases with high density and/or with molecules that largely differ in size, shape, and force-field parameters.

Numerical integration of the Clapeyron equation requires an initial coexistence point and an initial value for the integrand. The computation of these boundary conditions was discussed in chapter 4. The Gibbs ensemble method and the $Np\beta$ +test molecule method are suitable methods to compute the initial point.

The Gibbs ensemble method is by far the most straightforward method to compute phase coexistence. However, the method cannot be used in case the density difference between the liquid and the vapour phase is large. Furthermore, the Gibbs ensemble is not applicable at high liquid-phase density as successful exchanges of molecules between the phases are rare. The more computationally expensive $Np\beta$ +test molecule method uses individual liquid and vapour-phase simulations and can be applied at any difference in density between the phases. The existing $Np\beta$ +test molecule method has been improved. The improved method efficiently uses generated liquid-phase configurations, does not necessarily need a good guess to predict the right saturation pressure, and is applicable at relatively high liquid-phase densities.

Three methods, that can be used to compute the initial integrand in the limit of an infinite dilution, were discussed: the Widom test-molecule method and the difference method with or without overlapping distributions. The Widom test-molecule method, combined with other techniques, is applicable at relatively high densities. Among the three methods, the difference method can be implemented most easily in a simulation. However, it was shown theoretically and with simulation results that the method is not accurate in case the solvent molecules are larger than the solute molecules.

The overlapping-distributions difference method is computationally demanding. However, the method is applicable at high densities and provides the possibility to check whether the answer is reliable. Obviously, the overlapping-distributions difference method is the safest method to compute the residual-chemical potential difference.

In order to perform molecular simulation, the interactions between molecules need to be quantified. The interactions are usually expressed in a so-called force field. The importance of the choice of a proper force field was stressed in chapter five.

There are many different force fields available for individual components. These force fields differ in mathematical complexity, physical significance, and transferability of the parameters. Their predictive power is often limited as different kinds of interactions, including the non-additive ones, are effectively optimized to a pairwise additive potential. Moreover, the parameters are often optimized to only a limited number of physical properties over a narrow range of state variables. In this work, the primary selection criterion for an appropriate force field was the capability to predict results in close agreement with pure-component phase equilibria.

The interactions between dissimilar molecules are expressed in the interactions between similar molecules via so-called combining rules. The influence of these combining rules on the phase behaviour of mixtures can be substantial. As the conventional Berthelot combining rule generally overpredicts the interactions between dissimilar interaction groups, a different combining rule has been proposed in this work. However, since the interactions between different interaction groups on similar molecules are usually optimized with Lorentz-Bethelot

combining rules, different combining rules must be used for interactions between like and unlike molecules.

The capabilities of the Gibbs-Duhem integration method were tested on the vapour-liquid equilibria of a number of binary systems. The coexistence points of the pure components (points of suspension of the binary diagram) were computed separately. Provided both components were subcritical at the conditions of interest, the Gibbs-Duhem integration method easily traced the coexistence curve from the initial point to the coexistence point of the other component.

Part of the simulation results was obtained with very simple, physically not very realistic, force fields. However, the agreement with experimental data was excellent. The Gibbs-Duhem integration method got unreliable when a critical point was approached and/or when the initial point deviated too much from the true coexistence point.

In chapter six, an improved Gibbs-Duhem integration method was introduced. Improvements were necessary since the conventional Gibbs-Duhem integration method inefficiently exploits the larger part of the generated configurations. The information that is gathered during the prediction of the saturation pressure does not contribute to the ensemble averages of the densities and mole fractions. Moreover, the number of corrector simulations is not fixed from the outset.

In this thesis, a combination of the multiple-histogram reweighting method and Gibbs-Duhem integration was presented, which overcomes above-mentioned disadvantages. In the *advanced Gibbs-Duhem integration method*, every generated configuration contributes to the coexistence point. Another important advantage is the retroactive nature of the method; *i.e.* a current simulation improves formerly predicted coexistence points.

The method enables one to predict a whole coexistence curve instead of only individual coexistence points at predetermined integration steps. Furthermore, the method provides the opportunity to predict coexistence points at fixed compositions or densities. This is a clear advantage over the conventional Gibbs-Duhem integration method as, from an experimental point of view, a composition or density is a more interesting variable than a fugacity fraction. The advanced Gibbs-Duhem integration method has been applied to a number of binary systems. In general, the agreement between simulation results and experimental data was good. In most cases, the use of different combining rules for interactions between like and unlike molecules improved this agreement. The retroactive nature of the advanced method resulted in smooth simulation results. Due to the fixed number of simulations, the advanced method performed much faster than the conventional Gibbs-Duhem integration method did.

The extension of the advanced Gibbs-Duhem integration method to ternary mixtures was discussed in chapter seven. Extension to ternary or multicomponent mixtures is interesting as industrial mixtures are usually composed of more than two components. A disadvantage of multicomponent simulations is that they are accompanied by multidimensional histograms. However, in order to determine phase coexistence, only two dimensions of a histogram are needed. The rest of the dimensions need not take part in multiple-histogram reweighting.

Application of the advanced Gibbs-Duhem integration method to several ternary systems resulted in smooth simulation results in close agreement with experimental data.

The initial points for the ternary Gibbs-Duhem integrations were simulation results taken from binary Gibbs-Duhem integrations. It might be worthwhile to compute the initial point from a $Np\beta$ – Gibbs ensemble simulation or from the $Np\beta +$ test molecule method as these methods compute the initial point in a more direct sense from the conditions of phase coexistence. Furthermore, they require a single simulation instead of a whole series.

In this thesis, the ins and outs of Gibbs-Duhem integration were discussed. During application of the method, a number of problems appeared. Part of the problems has been overcome by a clever combination of Gibbs-Duhem integration and other simulation techniques. The computer code for Gibbs-Duhem integration, computation of the initial point and initial integrand can be conveniently used for the prediction of vapour-liquid coexistence. However, this work also provides points of interest for further research. The following recommendations emerged from this work:

1) A result of the error analyses of both Gibbs-Duhem integration techniques revealed the adverse effect of the error in the initial coexistence point on the course of the predicted coexistence line. Both Gibbs-Duhem integration techniques have no internal checkpoints where the prediction is tested in a direct way against the criterion of phase coexistence. It should be worthwhile to perform checks on a regular basis in order to be sure to predict a curve of zero free-energy difference.

2) Simulation results obtained in this work indicated that the Berthelot combining rule overestimates the interactions between dissimilar molecules. In this thesis, it was shown that the use of different combining rules for interactions between similar and dissimilar molecules can improve the predictions considerably. However, the use of different combining rules is inconsistent. For that reason, it should be worthwhile to optimize new force-field parameters with improved combining rules such that a single set of combining rules can be used for all interactions.

3) There is a number of extended transferable force fields available. Unfortunately, the transferability of the parameters goes at the cost of the quantitative prediction of individual physical properties. In particular the prediction of the saturation pressure is usually far off from the experimental values. For that reason, these transferable force fields are not suitable for the quantitative prediction of vapour-liquid equilibria. On the other hand, there are highly optimized non-transferable force fields. Unfortunately, force fields for an individual type of molecule often look so different that there is reason to doubt the physical significance of these models. It is questionable whether researchers will ever apply such force fields. It might be worthwhile to reach a certain level of consensus about how a good force field should look like and according to what procedure the parameters should be optimized. Moreover, it may be sensible to also optimize force-field parameters to derived properties as these properties are very sensitive to the parameters.

4) The advanced Gibbs-Duhem integration method results in a probability distribution that provides the probability to observe the system in a certain configuration, at a particular thermodynamic state point. It should be possible to predict a variety of derived physical properties, like heat capacities, isothermal compressibilities, and isobaric expansivities close to and at the coexistence line from the probability distributions. Computation of derived properties will probably require much longer simulations than the ones performed in this work. Furthermore, it will be important to take into account the correlation between subsequent samples into the multiple histogram reweighting.

Aleidis van 't Hof

Dankwoord

Op het moment van schrijven van dit dankwoord zijn de wielrenners van de Tour de France enkele dagen onderweg. Ik moest bij het kijken naar een zware bergetappe onwillekeurig aan de afgelopen vier jaar van mijn promotie denken. Op het eerste gezicht hebben wetenschap en wielersport weinig met elkaar te maken en een àl te letterlijke vergelijking gaat in sommige opzichten dan ook mank. Er zijn echter vele analogieën waarvan een aantal in het onderstaande de revue zal passeren. Hierbij zullen mijn belangrijkste ‘helpers in de strijd’ aan bod komen.

*Net als in de Tour de France gaat het bij een promotie grotendeels om individuele prestaties; het zelfstandig verrichten van vernieuwend onderzoek binnen een bepaalde tijdsspanne. Elke wielrenner behoort tot een bepaalde ploeg met een **ploegleiding** die hem zo nu en dan corrigeert, aanzet tot prestaties en, indien nodig, op het juiste pad brengt. Mijn ploegleiding bestond uit mijn promotoren prof. J. de Swaan Arons en prof. S.W. de Leeuw en mijn dagelijkse begeleider en co-promotor dr C.J. Peters.*

Mijn co-promotor, dr C.J. Peters, wil ik bedanken voor zijn vertrouwen in het welslagen van mijn onderzoek en voor de mogelijkheden die hij me heeft gegeven om mezelf te ontplooiën. Hij heeft me in grote mate vrijgelaten in de keuze van het onderwerp van het promotieonderzoek en de daadwerkelijke invulling daarvan.

Daarnaast bood hij me de mogelijkheid om college te geven aan eerstejaarsstudenten van de opleidingen Scheikundige Technologie en Life Science and Technology. Het geven van college vereiste een gedegen voorbereiding; één gek kan namelijk meer vragen dan tien wijzen kunnen antwoorden! Dingen die vroeger zo logisch of aannemelijk leken, hadden toch diepere gronden dan ik me vooraf had gerealiseerd. Ik heb er veel van geleerd.

Mijn promotor, prof. J. de Swaan Arons, wil ik bedanken voor de interesse die hij in mijn project heeft getoond. Ook al besloot hij om slechts zijdelings betrokken te zijn bij mijn werk, toch heeft hij altijd de moeite genomen om de globale lijnen van mijn onderzoek te doorgronden. Zijn adviezen over het ‘verkopen van je werk’ en de opbouw van een proefschrift zullen me altijd bijblijven. Daarnaast waren de besprekingen aan de Koornmarkt altijd gezellig; onder het genot van een kopje Chinese thee bespraken we, naast mijn promotiewerk, onderwerpen die varieerden van opvoeding tot wereldproblematiek.

Ongeveer halverwege de promotieperiode vonden we prof. S.W. de Leeuw bereid om, naast prof. J. de Swaan Arons, als promotor op te treden. Zijn gedegen kennis van simulatietechnieken en van fysica in het algemeen was meer dan welkom. Ik wil hem bedanken voor zijn interesse in mijn werk en vooral ook voor de mogelijkheid die hij bood om zomaar even zijn kamer binnen te wippen om een actueel probleem te bespreken.

*Elke ploeg bestaat uit een aantal wielrenners met verschillende specialiteiten: de één is een sprinter, de ander is een klimmer en nog een ander probeert het eindklassement te winnen. Zo waren er in de groep ‘Physical Chemistry and Molecular Thermodynamics’ promovendi, afstudeerders, post-docs, gasten en wetenschappelijk personeel met verschillende onderzoeksonderwerpen en interesses. Deze **ploegmaten** hebben in sommige gevallen de sprint voor me aangezet en me in andere gevallen geholpen met een steile klim.*

I would like to thank my colleagues at the research group ‘Physical Chemistry and Molecular Thermodynamics’. Although I realize I haven’t been the most entertaining colleague, in my opinion we had a great time together! Miranda, Gerard, Diana, José, Ali, Sona, Wei, Peijun, Eliane, Louw, Eugène, Karin, Laura, Maaïke, and Vincent, thanks for your friendship and cooperation.

Ook wil ik Arjan Minnigh danken voor zijn bijdrage aan de totstandkoming van dit proefschrift. Arjan’s afstudeerwerk werd ingepast in mijn promotieonderzoek. We hebben samen heel wat triviale en minder triviale problemen opgelost door er met elkaar over te discussiëren.

*Net als in de Tour bestaat het gevaar dat je **internationale concurrenten** er met de buit vandoor gaan. Het is bovendien praktisch bijna onmogelijk om alle aan dit werk gelieerde literatuur te lezen. Pas als het werk gepubliceerd is weet je dat de buit binnen is. Het bediscussiëren van je eigen werk met anderen gebeurt dan ook in goed vertrouwen.*

A number of scientists was willing to discuss several topics concerning this work with me, which I really appreciated. I would like to thank prof. D.A. Kofke, prof. D. Corti, prof. B. Smit, prof. D. Frenkel, prof. J.I. Siepmann, and dr Monica Hitchcock-Lamb who shared part of her computer code with me.

*Een deel van het succes van het wielerspektakel hangt af van het aantal **toeschouwers** dat de deelnemers toejuicht. De één heeft meer stimulans nodig dan de ander om het tot de eindstreep vol te houden. Zelf heb ik ervaren dat het best stil kan zijn langs de weg naarmate het onderzoek vordert; er zijn er die het onderzoek niet begrijpen en weer anderen die moeite hebben om de maatschappelijke relevantie van onderzoek dat niet direct mens-gerelateerd is in te zien. Gelukkig waren er altijd mensen bij wie ik mijn ei kwijt kon.*

Mijn ouders wil ik bedanken voor de stabiele gezinssituatie waarin ik heb mogen opgroeien en voor de mogelijkheden die ze me hebben geboden om goed onderwijs te volgen. Ook al was het waarschijnlijk niet altijd even duidelijk waar ik precies mee bezig was, mijn ouders, schoonouders en verdere aanverwanten hebben altijd met interesse naar mijn werk gevraagd. Verder bewaar ik goede herinneringen aan de ‘broederlunches’ in de aula. Cornelis en Pieter, bedankt voor de gezellige tijd die we samen hadden tijdens onze studies en/of ons promotieonderzoek aan de TU Delft.

Ook wil ik enkele vrienden bedanken voor hun interesse in mijn werk en voor de welkome afleiding die ze vormden op momenten dat het promotiewerk zich andermaal in een lokaal minimum bevond. Arjan en Mirjam Grevengoed en Hendrik-Jan en Marliekse Talsma, bedankt voor de warme vriendschap die Gerine en ik met jullie mochten, en naar ik hoop nog mogen, beleven.

*Helaas weet de promovendus, in tegenstelling tot de wielrenner, niet van te voren hoe het parcours er uit zal zien. In mijn geval startte de etappe rustig, kwamen de hoge bergen ongeveer halverwege en was de finish vlak. Het is belangrijk dat je op grote hoogte en in diepe dalen een **mental coach** of **ploegarts** aan je zijde hebt...*

Lieve Gerine, waren de eerste twee jaar van mijn promotietijd redelijk constant, de laatste twee jaar hebben we geleefd onder continu veranderende omstandigheden. De combinatie van

een promotie en het krijgen en opvoeden van onze jongens Joost en Marnix ging ons niet in de koude kleren zitten!

Je had soms moeite me te overtuigen van de irrelevantie van het aanzetten van een nieuwe simulatie halverwege de nacht; wijze woorden die ik nu beter op hun waarde weet te schatten! Ik wil je bedanken voor jouw niet te onderschatten bijdrage aan dit resultaat. Het promotieonderzoek heeft jou ook heel wat energie en hoofdbrekens gekost.....wat mij betreft word jij ook doctor!

Joost en Marnix, bedankt voor het relativiserings- en incasseringsvermogen dat jullie me hebben geleerd door op jullie ontdekkingsstocht door de wereld op dat aantrekkelijke knopje van die computer te drukken, die al een week met een berekening bezig was

Curriculum Vitae

July 20 th , 1977	Born in Leiden, The Netherlands
1989-1990	Secondary School at Christelijk Lyceum, Alphen a/d Rijn, The Netherlands
1990-1995	Secondary School at Christelijk College Nassau-Veluwe, Harderwijk, The Netherlands
June 1995	Graduation Pre-University Education (VWO)
1995-2000	Delft University of Technology, Delft, The Netherlands, Chemical Technology (M.Sc.)
June 2000	Graduation 'cum laude'
November 2001	Winner of the 'Unilever Researchprijs'
2000-2004	Ph.D. position at the research group Physical Chemistry and Molecular Thermodynamics, Delft University of Technology, Delft, The Netherlands

Publications

A. van 't Hof, M.L. Japas, and C.J. Peters, Description of liquid-liquid equilibria including the critical region with the crossover-NRTL model, *Fluid Phase Equil.* 192(1-2) **2001** 27-48

A. van 't Hof, S.W. de Leeuw, C.K. Hall, and C.J. Peters, Molecular simulation of binary vapour-liquid equilibria of components differing largely in volatility, *Mol. Phys.* 102(3) **2004** 301

A. van 't Hof, S.W. de Leeuw, and C.J. Peters, Several issues concerning Gibbs-Duhem Integration, **2005**, *in preparation*

A. van 't Hof, S.W. de Leeuw, and C.J. Peters, Advanced Gibbs-Duhem Integration, **2005**, *in preparation*

A. van 't Hof, S.W. de Leeuw, and C.J. Peters, Extension of the Advanced Gibbs-Duhem Integration method to multicomponent systems, **2005**, *in preparation*