

# **MOLECULAR SIMULATION OF VAPOUR-LIQUID-LIQUID EQUILIBRIUM**

by

**Suren Moodley**

**[BSc.(Eng)]**

University of KwaZulu-Natal

Submitted in fulfillment of the academic requirements for  
the degree of Master of Science in Engineering at the School  
of Chemical Engineering, University of KwaZulu-Natal

Durban  
2008

## Abstract

Phase equilibrium data is vital for designing chemical separation equipment. Traditionally, such data is obtained through laboratory experiments by sampling and analysing each phase of an equilibrated chemical mixture. An alternative means of generating such data is via molecular simulations, which also gives insight into the microscopic structure of the phases. This project was undertaken due to the lack of work on molecular simulations in predicting vapour-liquid-liquid equilibrium (VLLE).

Gibbs Ensemble Monte Carlo molecular simulations were performed in the isochoric-isothermal (NVT) and isobaric-isothermal (NPT) ensembles to determine the ability and limitations of the Transferable Potentials for Phase Equilibria (United-Atom) and Extended Simple Point Charge (SPC-E) force fields in predicting three-phase fluid equilibrium for two binary and three ternary industrially relevant mixtures: n-hexane/water (1), ethane/ethanol (2), methane/n-heptane/water (3), n-butane/1-butene/water (4) and n-hexane/ethanol/water (5).

The NPT ensemble proved inadequate for predicting VLLE for binary mixtures, as for both binary mixtures (1 and 2), the simulations reverted to two phases. This was due in part to the unlike-pair interactions between pseudoatoms in different molecules not being accurately predicted at the specified simulation conditions to reproduce experimental mixture densities and vapour pressures. It was also due to the sensitivity of the NPT ensemble to perturbations which probably removed the system from its three-phase trajectory in Gibbs phase space, since specifying even the correct pressure corresponding to the potential models was unsuccessful in obtaining stable VLLE. Furthermore, ternary VLLE could not be obtained for a mixture exhibiting an extremely narrow three-phase region (4) and simulations for a miscible, non-ideal mixture (5) gave mole fractions that were in poor agreement with experiment. Good results were obtained for mixture 3 which exhibits limited mutual solubilities and a large three phase region.

The NVT ensemble overcame the shortcomings of the NPT ensemble by producing three stable phases for the binary mixtures, revealing that the three-phase pressures were shifted by as much as 12%. Also, the narrow three-phase region of mixture 4 was overcome by adjusting the total system volume, producing three stable phases.

These were also the first successful binary VLLE simulations involving complex polyatomic molecules.

## **Preface**

The work presented in this dissertation was performed at the University of KwaZulu-Natal and Hogskolan i Borås. The work was supervised by Professor D. Ramjugernath and Professor K. Bolton. This dissertation is submitted as the full requirement for the degree M.Sc. in Chemical Engineering. All the work presented in this dissertation is original, unless otherwise stated. It has not (in whole or in part) been previously submitted to any tertiary institute as part of a degree.

---

S. Moodley

## **Acknowledgements**

I am indebted to the following individuals and organizations that have assisted me in this work:

- My supervisors, Professor Deresh Ramjugernath and Professor Kim Bolton, for their guidance, support and knowledge of thermodynamics.
- The National Research Foundation (NRF) and Swedish International Development Agency (SIDA) for financial assistance.
- Doctor Erik Johansson, for many discussions especially on the simulation of water.
- My family, for their patience during my years of study.

*Pigmaei gigantum humeris impositi plusquam ipsi gigantes vident.*

Dwarves placed upon the shoulders of giants see further than the giants themselves.

— Bernard of Chartres, 12<sup>th</sup> century philosopher

## Table of contents

1. INTRODUCTION .....	1
2. LITERATURE REVIEW .....	3
3. STATISTICAL MECHANICS.....	7
3.1 BASIC POSTULATES OF STATISTICAL MECHANICS .....	7
3.2 ENSEMBLE AVERAGES AND ENSEMBLES USED IN MOLECULAR SIMULATIONS.....	10
3.2.1 <i>Internal Energy</i> .....	11
3.2.2 <i>Pressure</i> .....	11
3.2.3 <i>Chemical potential</i> .....	13
3.3 MOLECULAR SIMULATION AND THE MONTE CARLO METHOD.....	14
3.3.1 RANDOM SAMPLING.....	16
3.3.2 IMPORTANCE SAMPLING .....	16
3.3.3 THE METROPOLIS METHOD .....	17
4. ENSEMBLES .....	21
4.1 CANONICAL (NVT) ENSEMBLE.....	21
4.2 ISOBARIC–ISOTHERMAL (NPT) ENSEMBLE .....	23
4.3 GRAND–CANONICAL (MVT) ENSEMBLE .....	25
4.3.1 <i>Histogram–reweighting grand canonical Monte Carlo simulations</i> .....	27
4.4 TRACING COEXISTENCE CURVES – GIBBS–DUHEM INTEGRATION.....	27
4.5 SEMI-GRAND ENSEMBLE .....	29
5. POTENTIAL ENERGY METHODS .....	31
5.1 INTRAMOLECULAR INTERACTIONS.....	31
5.1.1 <i>Bond Stretching</i> .....	31
5.1.2 <i>Bond Bending</i> .....	32
5.1.3 <i>Torsion Energy</i> .....	33
5.2 INTERMOLECULAR INTERACTIONS .....	34
5.2.1 <i>van der Waals interactions</i> .....	35
5.2.2 <i>Coulombic Interactions</i> .....	37

5.2.3 MIXING RULES.....	37
5.3 POTENTIAL MODELS .....	37
5.3.1 <i>Transferable Potentials for Phase Equilibria – United Atom description (TraPPE-UA)</i> .....	38
5.3.2 <i>Simple Point Charge – Extended model</i> .....	39
6. THE GIBBS ENSEMBLE.....	41
7. STANDARD MOLECULAR SIMULATION TECHNIQUES.....	50
7.1 RANDOM NUMBER GENERATION .....	50
7.2 PERIODIC BOUNDARY CONDITIONS AND THE MINIMUM IMAGE CONVENTION.....	50
7.3 ANALYTICAL TAIL CORRECTIONS.....	52
7.4 HARD–INNER CUTOFF RADIUS.....	53
7.5 LONG–RANGE INTERACTIONS – THE EWALD SUMMATION FOR POINT CHARGES.....	53
7.6 CONFIGURATIONAL BIAS MONTE CARLO METHODS.....	57
7.6.1 <i>Standard CBMC</i> .....	58
7.6.2 <i>Dual–cutoff CBMC</i> .....	60
7.6.3 <i>Coupled–Decoupled CBMC</i> .....	61
7.6.4 <i>Arbitrary Trial Distributions</i> .....	63
8. COMPUTATIONAL ASPECTS.....	66
8.1 HARDWARE.....	67
8.1.1 <i>Nodes</i> .....	67
8.1.2 <i>Network Subcomponents</i> .....	67
8.2 SOFTWARE .....	68
8.2.1 <i>Operating system</i> .....	68
8.2.2 <i>Parallel processing software</i> .....	69
8.2.3 <i>Resource Management Software</i> .....	69
8.3 ROCKS CLUSTER SUITE.....	70
8.3.1 <i>Installation</i> .....	73
8.3.1.1 <i>Frontend Node Installation</i> .....	74
8.3.1.2 <i>Slave Node Installation</i> .....	74

9. SIMULATION DETAILS AND RESULTS .....	75
9.1 SIMULATION DETAILS .....	75
<i>Coupled-decoupled CBMC settings</i> .....	76
9.1.1 Binary Simulations.....	77
9.1.1.1 <i>n-hexane/water</i> .....	77
9.1.1.2 <i>ethane/ethanol</i> .....	79
9.1.2 Ternary Simulations.....	79
9.1.2.1 <i>methane/water/n-heptane</i> .....	79
9.1.2.2 <i>n-butane/1-butene/water</i> .....	80
9.1.2.3 <i>water/ethanol/n-hexane</i> .....	80
9.2 NUMERICAL RESULTS.....	81
9.2.1 <i>n-hexane/water</i> .....	81
9.2.2 <i>ethane/ethanol</i> .....	84
9.2.3 <i>methane/water/n-heptane</i> .....	87
9.2.4 <i>n-butane/1-butene/water</i> .....	88
9.2.5 <i>water/ethanol/n-hexane</i> .....	89
9.3 GRAPHICAL RESULTS.....	90
10. DISCUSSION AND ANALYSIS OF RESULTS .....	106
10.1 BINARY SIMULATIONS .....	106
10.1.1 <i>n-hexane/water</i> .....	106
10.1.2 <i>ethane/ethanol</i> .....	110
10.2 TERNARY SIMULATIONS.....	113
10.2.1 <i>methane/water/n-heptane</i> .....	113
10.2.2 <i>n-butane/1-butene/water</i> .....	115
10.2.3 <i>water/ethanol/n-hexane</i> .....	117
10.3 SIMULATION TIMES.....	119
11. CONCLUSIONS AND RECOMMENDATIONS .....	121
12. REFERENCES .....	123
APPENDIX A – POTENTIAL MODEL PARAMETERS .....	132

A.1 TRAPPE - UA .....	132
A.2 SPC-E.....	133
A.3 NERDV3.....	134
APPENDIX B – EXPERIMENTAL DATA .....	136
B.1 WATER (1) + ETHANOL (2) + N-HEXANE.....	136
B.2 METHANE (1) + WATER (2) + N-HEPTANE (3).....	137
B.3 N-BUTANE (1) + 1-BUTENE (2) + WATER (3).....	137
B.4 ETHANE (1) + ETHANOL (2).....	137
B.5 N-HEXANE (1) + WATER (2).....	138
B.6 LIQUID DENSITY CORRELATIONS .....	138
APPENDIX C – SAMPLE INPUT AND OUTPUT FILES .....	140
C.1 PRINTOUT OF SIMULATION SETTINGS AND PARAMETERS .....	140
C.2 RUNTIME PRINTOUTS .....	147
C.3 FINAL PRINTOUTS .....	151

## List of Tables

Table 5-1 - List of pseudoatoms implemented in the TraPPE-UA force field that were relevant to this study. ....	39
Table 8-1 - Master and Slave Node Hardware Specifications.....	68
Table 8-2 - Network and Display Hardware. ....	68
Table 9-1 - Final simulation results for n-hexane/water mixture at 482 K in the NVT-Gibbs ensemble. ....	81
Table 9-2 - Simulation results for n-hexane/water mixture at 482 K and 4116.41 kPa using 213 water and 369 hexane molecules. ....	81
Table 9-3 - Simulation results for n-hexane/water mixture at 492 K and 48826.55 kPa using 213 water and 369 hexane molecules. ....	82
Table 9-4 - Preliminary LLE simulation results for 250 water and 200 n-hexane molecules at 3516 kPa and 473.15 K. ....	82
Table 9-5 - Example of a simulation where the systems enters a metastable single phase state at 3516 kPa and 473.15 K. ....	82
Table 9-6 – Results for n-hexane/water simulation in the NPT ensemble using the ‘shifted’ pressure (3663 kPa) obtained from a successful NVT ensemble simulation. ....	83
Table 9-7 – Results for 3-box NVT simulation at 311.15 K using 400 ethane and 100 ethanol molecules. ....	84
Table 9-8 - Results for 3-box NVT simulation at 311.15 K using 800 ethane and 200 ethanol molecules. ....	84
Table 9-9 – Results for 3-box NVT simulation at 311.15 K using 800 ethane and 200 ethanol molecules and the NERD force field. ....	84
Table 9-10 – Results for a 3-box NPT simulation using 1000 molecules, which reverted to two phases. ....	85
Table 9-11 – VLE results for ethane/ethane using 500 molecules at 311.15K in the NVT ensemble. ....	85
Table 9-12 - VLE results for ethane/ethane using 500 molecules at 311.15K in the NVT ensemble. Same as Table 9-11, except that a slightly larger total volume was used in the simulation. ....	85

Table 9-13 – VLE results for ethane/ethanol using 500 molecules at 311.15 K in the NVT ensemble. Same as Table 9-12, except that an even larger total volume was used in the simulation. ....	86
Table 9-14 – LLE results for ethane/ethanol using 500 molecules at 311.15 K in the NVT ensemble. ....	86
Table 9-15 - LLE results for ethane/ethanol using 500 molecules at 311.15 K in the NVT ensemble. Same as Table 9-14, except that a much smaller total volume was used in the simulation. ....	86
Table 9-16 – Results for 3-box NPT simulation at 275.5 K for methane/n-heptane/water at 120 kPa. ....	87
Table 9-17 - Results for 3-box NPT simulation at 275.5 K for methane/n-heptane/water at 2000 kPa. ....	87
Table 9-18 – Results for a successful 3-box NVT ensemble simulation for n-butane/1-butene/water at 310.93 K. ....	88
Table 9-19 – Results for NPT ensemble simulation for n-butane/1-butene/water at 310.93 K and 404.72 kPa. ....	88
Table 9-20 - Results for NPT ensemble simulation for n-butane/1-butene/water at 310.93 K and 101.33 kPa. This simulation reverted to two phases. ....	88
Table 9-22 – NPT simulation results for water/ethanol/n-hexane at 101.33 kPa at four different mixture compositions. ....	89
Table 10-3 – Comparison of the simulation and experimental results for phase compositions for n-butane/1-butene/water at 310.93 K.....	117
Table A-1 - TraPPE-UA Non-bonded parameters for pseudoatoms used in this work.....	132
Table A-2 - TraPPE-UA bond lengths for bonded-pairs used in this work.....	133
Table A-3 - TraPPE-UA torsion constants for dihedral angles. ....	133
Table A-4 - TraPPE-UA equilibrium angles and bending constants used in this work. ....	133
Table A-5 - SPC-E parameters used to represent Water in this work. ....	134
Table A-6 - NERDv3 Non-bonded parameters for pseudoatoms used in this work. ....	134
Table A-7 - NERDv3 bond-stretching parameters for bonded-pairs used in this work. ....	134
Table A-8 - NERDv3 torsion constants for dihedrals.....	134
Table A-9 - NERDv3 equilibrium angles and bending constants used in this work. ....	135

Table B-1 - Experimental VLLE data for Water (1) + Ethanol (2) + n-Hexane (3) at 101.3 kPa (Gomis <i>et al.</i> , 2007).	136
Table B- 2 - Experimental VLLE data for Methane (1) + Water (2) + n-Heptane (3) at 277.5 K (Susilo <i>et al.</i> , 2005).	137
Table B-3 - Experimental VLLE data for n-butane (1) + 1-butene (2) + water (3) at 310.93 K (Wehe and McKetta, 1961).	137
Table B-4 - Experimental VLLE data point for ethane (1) + ethanol (2) at 5.169 MPa (Kato <i>et al.</i> , 1999).	137
Table B-5 - Constants used in the in the liquid density correlations for pure components (Perry and Green, 2007).	139

## List of Figures

Figure 4-1 - The Canonical Ensemble. ....	21
Figure 4-2 - The Isobaric–Isothermal Ensemble. ....	23
Figure 4-3 - The Grand–Canonical Ensemble. ....	26
Figure 4-4 - The Semigrand Ensemble. ....	29
Figure 5-1 - The various intramolecular interactions that contribute towards the potential energy of a given molecular configuration. ....	33
Figure 5-2 - Representation of the variation of the torsion energy for the C—C—C—C bonded group in perfluoroalkanes as parameterized by Watkins and Jorgensen (2001) .....	34
Figure 5-3 - Illustration of the potential energy well of a Lennard–Jones fluid. ....	36
Figure 5-4 – An illustrative example using n-propane, showing the philosophy of united-atom force fields. ....	38
Figure 5-5 – Schematic representation of the SPC-E water model. ....	40
Figure 6-1 - Schematic of the Gibbs ensemble technique for a single component, two–phase system. ....	42
Figure 6-2 - Gibbs ensemble Monte Carlo trial moves. ....	44
Figure 6-3 - Schematic of the Gibbs ensemble simulation technique for a three-phase (vapour-liquid-liquid) system. ....	45
Figure 7-1 - Illustration of the method of periodic boundary conditions and minimum image convention for a 2–D system. ....	51
Figure 7-2 - Schematic of a polar molecule. ....	54
Figure 7-3 - Ewald summation splitting of a charge. ....	56
Figure 7-4 - The configurational–bias Monte Carlo method. ....	58
Figure 8-1 – Schematic of a typical Beowulf cluster setup .....	66
Figure 8-2 – Screenshot of the Ganglia cluster monitoring tool. ....	72
Figure 8-3 – Snapshot of the Rocks SQL database. ....	73
Figure 9-1 - Plot of the total potential energy in each simulation box during the pre-equilibration period for n-hexane/water in the NVT ensemble. ....	90

Figure 9-2 - Plot of the variation in the liquid-phase box volumes versus the number of Monte Carlo cycles for n-hexane/water.....	91
Figure 9-3 - Plot of the total potential energy in each simulation box during a production run for n-hexane/water in the NVT ensemble. ....	92
Figure 9-4 - Plot of the variation of the number of n-hexane molecules in each phase during a production run in the NVT ensemble for n-hexane/water. ....	93
Figure 9-5 - Plot of the pressure within each simulation box versus the number of Monte Carlo cycles during a production run.....	94
Figure 9-6 - Plot of the variation of individual box volumes during a production run. ....	95
Figure 9-7 - Comparison of box energies for ethane/ethanol in the NVT ensemble before and after box identity swaps. ....	96
Figure 9-8 - Comparison of number of ethane molecules in each simulation box for ethane/ethanol in the NVT during an identity swap between the vapour and ethane-rich liquid phases. ....	97
Figure 9-9 - Comparison of number of volume changes in each simulation box for ethane/ethanol in the NVT during an identity swap between the vapour and ethane-rich liquid phases. ....	98
Figure 9-10 - Phase diagram for ethane/ethanol at 311.15 K. ....	99
Figure 9-11 - VLLE ternary composition diagram for methane/water/n-heptane at 275.5 K and 120 kPa.....	100
Figure 9-12 - VLLE ternary composition diagram for methane/water/n-heptane at 275.5 K and 2000 kPa.....	101
Figure 9-13 - VLLE ternary composition diagram for n-butane/1-butene/water at 310.93 K and 404.72 kPa.....	102
Figure 9-14 - VLLE ternary composition diagram for n-butane/1-butene/water at 310.93 K and 101.33 kPa.....	103
Figure 9-15 - VLLE ternary composition diagram for n-butane/1-butene/water at 310.93 K in the NVT ensemble. ....	104
Figure 9-16 - VLLE diagrams for water/ethanol/n-hexane at 101.33 kPa obtained from NPT-Gibbs ensemble simulations using a total of 750 molecules. ....	105

## Nomenclature

$acc(o \rightarrow n)$	probability of accepting a move from old state $o$ to new state $n$
$\mathbf{b}$	vector co-ordinates of a periodic image or a bead's centre of mass, $\mathbf{b} = (b_x, b_y, b_z)$
$b_x, b_y, b_z$	$x$ -, $y$ - and $z$ -components of a periodic image
$C^{\text{Gauss}}$	integration constant for Gaussian arbitrary trial probability distribution
$\{\mathbf{c}\}_\kappa$	set of $\kappa$ trial orientations for Configurational-Bias Monte Carlo moves
$c_{0,1,2,\dots}$	cosine series torsional constants
$C$	number of components
$d_{\text{COM}}(\text{A})$	largest distance between the interaction site furthest from the centre of mass of molecule $\text{A}$
$F$	number of phases
$\mathbf{F}_{ij}$	pair force exerted on molecule $i$ by molecule $j$
$g_2(r_{ij})$	pair distribution function
$G$	Gibbs free energy $G = U - TS + PV$
$h$	Planck's constant
$H$	Hamiltonian, or enthalpy, $H = U + PV$
$\Delta H_{\text{vap}}$	latent heat of vaporization
I,II,III	index to denote distinct phases

### Configurational-Bias Monte Carlo move

$f$	number of trial insertions for the first atom in CBMC method
$k$	number of trial directions for segments after the first segment is placed used during a Configurational-Bias Monte Carlo move
$k_B$	Boltzmann's constant
$k_{\text{str}}$	bond stretching constant
$k_{\text{str}}$	bond angle bending constant
$K$	kinetic (ideal) contribution to the energy
$K_0$	overall constant used to non-dimensionalize the partition function

$K$	arbitrary constant
$l$	bond length, or number of segments to be grown during a Configurational-Bias Monte Carlo move
$L$	length of a cubic simulation box
$m$	molecule mass
$n_i$	fluctuating number of molecules in phase I or of type $i$
$n_{\text{chLJ}}$	number of non-bonded trials moves for Coupled-Decoupled CBMC
$n_{\text{ctor}}$	number of torsional trials moves for Coupled-Decoupled CBMC
$n_{\text{cbend}}$	number of angle-bending trials move for Coupled-Decoupled CBMC
$N$	number of molecules
$N_{\text{seg}}$	is the number of atomic groups/segments in the molecule
$\mathbf{N}$	denotes the set of all number of molecules in each phase, $\mathbf{N} = \{n_1, n_{\text{II}}, \dots, n_F\}$
$\mathbf{p}$	molecular momentum vector
$P$	pressure
$p_{\text{trial}}^{\text{arb}}$	arbitrary trial distribution
$\Omega$	denotes a probability density function
$\Omega(\mathbf{r}^N)$	probability of observing the system in a configuration around $\mathbf{r}^N$ for an ensemble
$q$	generalized coordinate
$Q$	ensemble partition function
$\mathbf{r}_i$	molecular or atomic position vector for interaction site $i$ , $\mathbf{r}_i = (r_x, r_y, r_z)_i$
$\mathbf{r}_{ij}$	vector separation between sites interaction sites $i$ and $j$ , $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$
$\hat{\mathbf{r}}_x, \hat{\mathbf{r}}_y, \hat{\mathbf{r}}_z$	three orthogonal unit vectors in three-dimensional space
$r$	molecular or atomic position scalar
$r_C$	cut-off radius
$r_{DC}$	cut-off radius for the DC-CBMC algorithm
$r_{ij}$	scalar separation between interaction sites $i$ and $j$ , $r_{ij} =  \mathbf{r}_{ij} $
$r_x, r_y, r_z$	$x$ -, $y$ - and $z$ -components of molecular position
$R$	total number of points generated in an importance sampling algorithm, or universal gas constant

RND	random number from a uniform distribution in the interval [0,1]
$\mathbf{s}$	dimensionless molecular or atomic position vector, $\mathbf{s} = \mathbf{r}/L$
$t$	time
$T$	absolute temperature
$U$	potential (excess) contribution to the energy. Equivalent to internal energy
$U(\mathbf{r}_i)$	potential energy of a trial position in configurational-bias Monte Carlo
$\bar{U}^{\text{ext}}$	approximate intermolecular potential for DC-CBMC algorithm
$U_{1j}^{\text{ext}}$	non-bonded potential due to the insertion of this first atom, for the $j^{\text{th}}$ attempt
$\delta U^{\text{ext}}$	difference between the full and approximate potentials for the DC-CBMC algorithm
$\Delta U_{i,1}^+$	energy change of inserting a molecule of type $i$ into phase I for the evaluation of the chemical potential
$V$	volume of a simulation box
$\mathbf{V}$	denotes the set of volumes for each of the phases, $\mathbf{V} = \{V_I, V_{II}, \dots, V_F\}$
$V_C$	cut-off volume, $V_C = \frac{4}{3}\pi r_C^3$
$V_0$	reference volume to non-dimensionalize the partition function
$w(r_{ij})$	intermolecular pair virial function for pairwise additive functions
$W$	total molecular virial
$w(\mathbf{\Gamma})$	weight function
$w_m^{\text{new}}$	Rosenbluth factor of segment $m$ for the new chain in the CBMC algorithm
$W^{\text{new}}$	total Rosenbluth weight for the new chain
$\bar{W}^{\text{ext}}$	approximate Rosenbluth weight for the new chain for the DC-CBMC algorithm
$x^{\text{Gauss}}$	mean of the Gaussian distribution in Arbitrary Trial Distribution regrowth

#### Greek Symbols

$\alpha(o \rightarrow n)$	probability of generating a new state $n$ from an old state $o$
$\beta$	inverse temperature, $\beta = 1/k_B T$
$\varepsilon_i$	kinetic energy of molecule $i$
$\varepsilon_{ij}$	Lennard-Jones energy parameter between sites $i$ and $j$

$\kappa$	parameter that determines width of Gaussian distribution
$\theta$	bond angle
$\Lambda$	thermal de Broglie wavelength
$\mu$	chemical potential
$\xi_i$	fugacity fraction of component $i$
$\pi$	pi
$\pi(o \rightarrow n)$	transition probability of going from an old state $o$ to a new state $n$
$\rho_i$	density of component $i$ or the density of phase $i$
$\zeta_i$	number of points generated per unit volume around $\mathbf{r}^N$
$\sigma_{ij}$	Lennard-Jones size parameter between sites $i$ and $j$
$\sigma^{\text{Gauss}}$	standard deviation of Gaussian distribution in Arbitrary Trial Distribution regrowth
$\Gamma(t)$	point in phase space at time $t$
$\phi$	torsion angle

#### Brackets and Operators

$\langle \dots \rangle$	ensemble average
-------------------------	------------------

#### Subscripts

0	an equilibrium bond length $l_0$ , an equilibrium bond angle $\theta_0$ for a harmonic bond angle potential or an equilibrium torsional angle $\phi_0$ for a harmonic torsional potential
bend	contribution from bond angle bending
<i>ens</i>	ensemble average
<i>g</i>	gas or vapour phase
<i>h</i>	arbitrary indexing variable
<i>i,j</i>	species type, or pseudoatomic group interaction site
iso	isothermal ensemble
low	lower integration limit
LJ	Lennard-Jones potential
<i>m</i>	next segment being placed during a Configurational-Bias Monte Carlo move
<i>obs</i>	observed variable

stretch	contribution from bond stretching
$t$	trial direction when growing molecule during a Configurational-Bias Monte Carlo move
<i>therm</i>	thermodynamic definition of the pressure
<i>time</i>	time average
tors	contribution from torsion
up	upper integration limit
vir	virial definition of the pressure
$x, y, z$	three orthogonal directions in three-dimensional space
$\alpha, \beta$	interaction sites
$\phi$	torsional angle

#### Superscripts

$C$	short-range contribution to a calculated property below the cut-off radius, $r_c$
<i>id</i>	ideal property or ideal contribution
gen	generation of a configuration
sel	selecting a trial/new configuration
intra	energy from all the <u>intra</u> molecular interactions
inter	energy from all the <u>inter</u> molecular interactions
new	new state
old	old state
sat	property at saturated conditions

#### Abbreviations

COM	centre of mass
CBMC	Configurational-Bias Monte Carlo
DC-CBMC	Dual Cut-Off Configurational-Bias Monte Carlo
GEMC	Gibbs Ensemble Monte Carlo
LJ	Lennard-Jones potential
NPT	constant number of molecules, constant pressure and constant temperature ensemble
NVT	constant number of molecules, constant volume and constant temperature ensemble
VLE	vapour-liquid equilibrium
VLLE	vapour-liquid-liquid equilibrium

$\mu VT$  constant chemical potential, constant volume and constant temperature ensemble  
 $\mu PT$  hypothetical constant chemical potential, constant volume and constant temperature  
'ensemble'

## 1. Introduction

The acquisition of phase equilibrium data for chemical systems is of vital importance in the chemical industry for the various chemical separation processes that are used, for example, distillation, stripping and solvent extraction. The most common method for acquiring such data is through laboratory experimentation, in which the data is generated by sampling the different phases and then analyzing the composition of each phase, usually through an analytical technique such as gas chromatography. However, the use of mechanical equipment in the laboratory is limited, especially in situations where one wishes to obtain data at high pressures or high temperatures since the material of construction may not endure such pressures or temperatures. An attractive, alternative method in the form of molecular simulation has been making steady progress since 1953 when Metropolis *et al.* (1953) performed the first liquid Monte Carlo simulations at the Los Alamos Laboratory, using only computers. Since then, a large number of studies have been undertaken to develop the molecular simulation of phase equilibrium (and other molecular phenomena) so as to eventually drastically reduce the need for conventional laboratory experimentation. These studies have been greatly aided by the rapid development and advancement of cheaper yet faster computational power, and nowadays, a simple computer cluster can be built from off-the-shelf components in conjunction with free open-source software.

This work, titled “**Molecular Simulation of Vapour–Liquid–Liquid Equilibrium**”, was carried out at the University of KwaZulu-Natal (South Africa) and the University College of Borås (Sweden) under the supervision of Professor D. Ramjugernath and Professor K. Bolton, respectively. There is a wealth of literature on the molecular simulation of vapour–liquid equilibrium (VLE) and liquid–liquid equilibrium (LLE) of binary and ternary mixtures (see Panagiotopoulos (1989); Smit *et al.* (1995); Potter *et al.* (1997); Nath *et al.* (1998); Siepmann *et al.* (1997); Martin and Siepmann (1999); Potoff and Siepmann (2001); Chen *et al.* (2001); Stubbs *et al.* (2001); Nath (2003) and Khare *et al.* (2004), amongst others). There are, however, very few studies which have explicitly

examined the vapour–liquid–liquid equilibrium (VLLE) behaviour of chemical mixtures. Thus, it was the purpose of this work to exclusively examine the VLLE of binary and ternary mixtures, containing industrially relevant compounds – simple, unbranched alkanes, alkenes and alcohols were investigated, along with water, in a qualitative manner.

This dissertation is structured as follows: In Chapter 2, a brief literature review is presented detailing previous three-phase simulation studies. In Chapter 3, the basic link between the behaviour of matter at the microscopic level and the consequent phenomena observed in the macroscopic world of thermodynamics, namely, statistical mechanics is discussed. Thereafter, a discussion ensues in Chapter 4 on the various statistical ensembles that are available. Chapter 5 presents the various methods and force field models employed in modeling the interaction energies of the molecular systems of interest, while Chapter 6 gives insight into the Gibbs ensemble method, which formed the crux of the simulations that were carried out. Chapter 7 discusses some of the modern standard molecular simulation techniques that are applied to polyatomic molecular systems, while Chapter 8 reviews the computation resources that were used in this study. Chapter 9 provides the details of the simulations and the consequent results and Chapter 10 provides an analysis of the findings of this work. Finally, Chapter 11 presents the conclusions resulting from the study along with a few recommendations for related future work.

## 2. Literature Review

While there has been numerous simulation studies concerning the two-phase equilibrium of various chemical mixtures, there is a scarcity in the literature of three-phase fluid equilibrium work.

Lopes and Tildesley (1997) proposed a method that extended the Gibbs ensemble simulation technique to multiphase<sup>1</sup> equilibrium for simple model systems in the canonical ensemble. Using simple monatomic Lennard-Jones beads all having the same size parameter,  $\sigma$ , they demonstrated for several simple model systems that attaining three phases was indeed possible. Phase diagrams for two-component, three-phase and three-component, three- and four-phase mixtures were generated by varying only the compositions of each mixture and energy parameters ( $\epsilon$ ), for each different bead type.

Using the method proposed by Lopes and Tildesley (1997), Kristof *et al.* (2002) investigated the high-pressure phase equilibrium of a ternary mixture containing carbon dioxide, methanol and water in the isobaric-isothermal (NPT) ensemble. Site-site potential models of the Lennard-Jones 6-12 type to model intra- and intermolecular interactions along with Coulomb potentials to represent the electrostatic contribution to the total energy were used; the Elementary Physical Models 2 (EPM2) force field (Harris and Yung, 1995) was used to represent carbon dioxide, a model proposed by van Leeuwen and Smith (1995) to represent methanol, and the Transferable Intermolecular Potentials – 4 Point (TIP4P) force field along with the Simple Point Charge (SPC) potential model were used in separate simulations to model water interactions. The multiphase Gibbs ensemble Monte Carlo method was used to test regions where three phases were observed experimentally. It was observed that the simulation boxes approached three distinct phases at higher pressures than the experimental pressures.

Chen *et al.* (2002) investigated the adsorption, surface tension and molecular ordering at the vapour-liquid interfaces of mutually saturated 1-butanol/water mixtures in the NVT

---

<sup>1</sup> In this context, ‘multiphase’ means ‘greater than two’ since the original Gibbs ensemble method, proposed by Panagiotopoulos, used two simulation boxes.

ensemble; here, liquid boxes were elongated along one axis to create vapour-liquid interfaces. The TIP4P force field was used to represent water interactions and 1-butanol was represented by the united-atom variant of the Transferable Potentials for Phase Equilibria (TraPPE-UA) family of force fields. Strictly speaking, this was not a three-phase fluid equilibrium simulation per se, since the ‘vapour’ phase simulation box was merely used to facilitate the transfer of molecules between the two liquid phases since this **‘avoids concurrent energy penalties associated with removal and insertion of a molecule from/into liquid phases’** (Chen *et al.*, 2002).

Zhang and Siepmann (2004) used Configurational-Bias Monte Carlo simulations in the Gibbs ensemble to investigate the vapour-liquid-liquid equilibria and microscopic structures for two ternary perfluoroalkane/alkane/carbon dioxide mixtures using the TraPPE-UA force field; n-decane/n-perfluorohexane/carbon dioxide and n-hexane/n-perfluorodecane/carbon dioxide were studied. However, prior to the three-phase simulations, special mixing parameters for the unlike interactions between  $\text{CH}_x$  and  $\text{CF}_y$  pseudoatoms were obtained from two-phase simulations of perfluoromethane and methane to obtain significantly better predictions for the solution critical temperature. This was done by implementing two multiplying, or correction factors for the Lennard-Jones size and energy cross-terms which were adjusted until good agreement with experimental data was obtained:

$$\sigma_{ij} = \frac{1+a_{ij}}{2}(\sigma_{ii} + \sigma_{jj}), \quad (2.1)$$

and

$$\varepsilon_{ij} = (1-b_{ij})\sqrt{\varepsilon_{ii}\varepsilon_{jj}}. \quad (2.2)$$

The calculated upper critical solution temperatures for the three-phase mixtures were in excellent agreement with experimental data.

Wick *et al.* (2004) produced molecular-level information on the retention mechanism in reversed-phase liquid chromatography using the Gibbs ensemble method for a three-phase system consisting of helium vapour as the reference state, n-hexadecane as the retentive phase and a mobile phase consisting of water and methanol, for which the composition was allowed to vary for four different methanol/water ratios. To decrease computation time, n-hexadecane was not allowed to swap into the mobile or vapour phases due to its aqueous solubility being extremely low and its saturated vapour pressure being low as well. The TIP4P force field was used to model water interactions whilst the TraPPE-UA force field was used to model alkane, alcohol and helium molecules. For other three-phase chromatography work, see Rafferty *et al.* (2007), Sun *et al.* (2007a), Sun *et al.* (2007b), Zhang *et al.* (2006) and Sun *et al.* (2006).

Finally, Chen and Siepmann (2005) studied octanol/water partition coefficients for eight solute molecules (methane, ethane, propane, n-butane, methanol, ethanol, 1-propanol and 1-butanol), as well as the structure of, and solvation in the two liquid phases. Again, simulations were performed in the NPT variant of the Gibbs ensemble using the TIP4P and TraPPE-UA force fields at 1 atmosphere and 298 K. Initially, all solute molecules were placed in the vapour phase simulation box and pure octanol and water were placed in separate simulation boxes as well. The simulation results reconciled the structure of 1-octanol which was initially deduced from spectroscopic measurements and diffraction experiments.

It is now apparent that there are a limited number of three-phase fluid equilibrium studies. Thus, it was the purpose of this work to assess the ability and limitations of certain force fields in, first and foremost, essentially producing three phases in molecular simulations and consequently predicting the VLLE of several industrially-relevant binary and ternary mixtures. To this end, the ubiquitous TraPPE-UA force field was used to model the alkane, alkene and alcohol molecules, while the Simple Point Charge - Extended (SPC-E) potential model was used to model water. The SPC-E model was selected because there have already been several studies in which the TIP4P water model was used in conjunction with

TraPPE-UA for three phases yielding acceptable results, as discussed in the preceding paragraphs. Furthermore, simulations involving water and alkanes may be compared to a recent extensive biphasic study on the solubility of water in alkanes and the vapour phase clustering of water (Johansson, 2007). It is interesting to note, considering the extensive survey of the literature, that there have been no studies based on *binary* three-phase fluid equilibrium for complex polyatomic molecules.

### 3. Statistical Mechanics

Real chemical systems contain an extremely large number of molecules, of the order of Avogadro's number ( $\sim 10^{23}$ ). With current computational power, to simulate even one mole of any substance would take hundreds, if not thousands, of years to obtain results. Clearly, a means to obtain credible results from a drastically smaller number of molecules is needed. Quantum mechanics enables the calculation of molecular properties at the most basic level (e.g. equilibrium geometries and potential energy surfaces); more importantly, the mechanical description of a system of molecules fully specifies its microscopic behaviour, while thermodynamics is used to calculate the macroscopic (or bulk) 'observable' properties of a collection of these molecules. Intuitively, there must be a link between these two worlds of vastly different sizes. This link is statistical mechanics.

Statistical mechanics is a formalism which aims at explaining the physical properties of matter *in bulk* on the basis of the dynamical behavior of its microscopic constituents (Pathria, 1972).

#### 3.1 Basic postulates of statistical mechanics

Central to the study of statistical mechanics is the notion of an **ensemble**. An ensemble is a hypothetical collection of an infinite number of non-interacting systems, each of which is in the same thermodynamic state as the system of interest (Levine, 2003). Although the members of the ensemble are thermodynamically identical, they exhibit a vast number of microstates<sup>2</sup>, since many different microstates are compatible with a given macrostate (or thermodynamic state). This may be explained simply as follows: consider a collection of non-interacting particles confined to an isolated 3-dimensional space of arbitrary shape. Each particle possesses a definite, though not unique, kinetic energy. Being monatomic, this kinetic energy will be manifested as translational energy for each particle in each of the three spatial dimensions. Evidently, the total energy of this system is constant. However, there are numerous ways of partitioning or distributing this energy amongst all the particles

---

<sup>2</sup> The term 'microstate' is here understood to be the same as a 'quantum' state.

in the system, and for each particle, its own kinetic energy can be further distributed amongst the three Cartesian coordinates. Mathematically, this translates to

$$\sum_{i=1}^N \varepsilon_i = E \quad (3.1)$$

where  $N$  is the total number of particles in the system,  $\varepsilon_i$  is the total kinetic energy of particle  $i$  and  $E$  is the total energy of the system of particles. This idea can be extended further to a simple experiment. Consider an isolated water bath, of which we wish to determine the temperature. One method is to take temperature measurements at discrete time intervals using a thermometer and thereafter taking a time-weighted average to obtain the average water temperature, according to the formula

$$\langle T \rangle_{time} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t T(t') dt' \quad (3.2)$$

where the  $\langle \dots \rangle_{time}$  denotes the time-averaged value of a macroscopically observable property,  $t$  is the total observation time and the prime denotes the dummy variable of integration. At this point, it is instructive to introduce the phase space formalism of J. Willard Gibbs.

The microscopic state of a model system is uniquely determined by the specification of the complete set of microscopic variables (Vesely, 2007). The number of such variables is of the same order as the number of particles. In contrast, the thermodynamic state is specified by a small number of measurable quantities such as mass, pressure, temperature or volume. It was explained earlier in this chapter that a huge number of different microstates are compatible with a single macrostate. The microscopic variables may be viewed as coordinates in a high-dimensional space; a particular microstate may be represented as a vector in that space. This high-dimensional space is called Gibbs phase space and the state vector is symbolized as  $\vec{\Gamma}$ . Thus, for the ideal gas, the state vector is defined by all position and momentum coordinates (Vesely, 2007):

$$\vec{\Gamma} = \{\vec{r}_1, \dots, \vec{r}_N; \vec{v}_1, \dots, \vec{v}_N\} \quad \vec{r}_i \in V; v_{i,\alpha} \in (\pm\infty) \quad (3.3)$$

where  $\vec{r}_i$  and  $\vec{v}_i$  refer to the position and velocity vectors of the  $i$ -th particle, respectively, and  $\alpha$  refers to the three spatial dimensions.

Going back to determining the temperature of the water bath, the second method that may be used is one in which the time average is replaced by an ensemble average. Consider the same water bath (with the same thermodynamic state). In this method, there exists a large number of the same macroscopic system, i.e., an ensemble. Indeed, there are a huge number of microstates compatible with the given macrostate since water is a polyatomic molecule, and with the addition of bond-stretching and bond-bending energies, more degrees-of-freedom are introduced to the molecular configuration of the system. Noting that each of these microscopic states may be represented as a particular point in phase space, the ensemble average may be written as

$$\langle T \rangle_{ens} = \sum_{\vec{\Gamma}} p_{ens}(\vec{\Gamma}) T(\vec{\Gamma}) \quad (3.4)$$

where  $p_{ens}(\vec{\Gamma})$  is the probability density of observing a particular point in  $\vec{\Gamma}$  space in an ensemble. Obviously, one cannot create an infinite number of copies of the same macroscopic chemical system in something as finite as the memory of a computer, thus, for Equation 3.4 to be valid, a necessary condition is that of **ergodicity**. Instead of considering an ensemble of systems, consider just one single system as it evolves in time according to the laws of mechanics. The ergodic hypothesis states that “**in the course of such a ‘natural evolution’ of the system any permitted microstate will be reached (or closely approximated) with the same relative frequency**” (Vesely, 2007).

The ergodic hypothesis has an important consequence: for the calculation of average values over the microstates it does not matter if averages are taken over states randomly

picked from an ensemble, or over the successive states of one single, isolated system. The corollary of the ergodic hypothesis is succinctly stated as:

$$\text{ensemble average} = \text{time average} \quad (3.5)$$

Thus, one would expect that for the simple experiment described above, provided that sufficient sampling time and correct sampling of phase space occurred, respectively, in each method, the two average temperatures would be identical.

### 3.2 Ensemble averages and ensembles used in molecular simulations

The most common ensembles used in the molecular simulation of chemical thermodynamics will be briefly discussed. First, the nature of the probability of finding a chemical system in a particular microstate, namely  $p_{ens}(\vec{\Gamma})$ , shall be explained.

Allen and Tildesley (1987) defined a ‘weight’ function  $w_{ens}(\vec{\Gamma})$  in place of  $p_{ens}(\vec{\Gamma})$  which satisfies the following equations:

$$p_{ens}(\vec{\Gamma}) = \frac{w_{ens}(\vec{\Gamma})}{\sum_{\vec{\Gamma}} w_{ens}(\vec{\Gamma})} \quad (3.6)$$

$$Q_{ens} = \sum_{\vec{\Gamma}} w_{ens}(\vec{\Gamma}) \quad (3.7)$$

where  $Q_{ens}$  is the ensemble partition function which is a sum over states, and is unique for each ensemble (discussed later in this chapter). It may also be thought of as a normalizing factor for the probability  $p_{ens}$ . Thus, any thermodynamic property may be calculated by using the following equation:

$$\langle A \rangle_{ens} = \frac{\sum_{\vec{\Gamma}} w_{ens}(\vec{\Gamma}) A(\vec{\Gamma})}{Q_{ens}}. \quad (3.8)$$

### 3.2.1 Internal Energy

In general, the internal energy of a molecular configuration may be expressed in terms of the kinetic and potential energies. Using the Hamiltonian formalism, this translates to:

$$H(\mathbf{p}, \mathbf{q}) = E = K(\mathbf{p}) + U(\mathbf{q}), \quad (3.9)$$

where  $K$  is the total kinetic energy,  $U$  is the potential energy,  $\mathbf{p}$  refers to the momenta of all molecules and  $\mathbf{q}$  is the generalized coordinate system that is conjugate to the momenta. This splitting allows ease of calculation of thermodynamic averages, since the kinetic energy may be factored out and integrated analytically in isothermal systems as it is a quadratic function of momentum (du Preez, 2005). In fact, intuition dictates that since a Monte Carlo molecular simulation can be crudely seen as a series of snapshots where unphysical moves are permissible, the kinetic energy is of no consequence. Accordingly, it is the correct calculation of the potential energy that forms the backbone of an accurate and reliable Monte Carlo simulation since it is a function of the molecules' spatial coordinates.

### 3.2.2 Pressure

Molecular simulations are performed using periodic boundary conditions, wherein no conceptual boundaries exist through which molecules can pass and be detected. As such, it is not possible to measure the momentum flux against the boundaries in the simulation box (du Preez, 2005). There are two methods available for calculating the pressure during a

simulation: the first method uses the thermodynamic definition of pressure from a Maxwell relationship whilst the second method is derived from the virial theorem. It is important to note that pressure calculations are important even if the pressure has been specified prior to the simulation (for example, the NPT ensemble), since one of the conditions for thermodynamic equilibrium is equality of pressure across all phases.

### Thermodynamic pressure

Starting with the thermodynamic definition of pressure, where  $A$  is the Helmholtz free energy,

$$P_{\text{therm}} = -\left(\frac{\partial A}{\partial V}\right)_{N,T}, \quad (3.10)$$

it can be shown that the pressure may be split into an ideal component and a component that is due to intermolecular forces (Hummer and Grønbech–Jensen, 1998) as follows:

$$P_{\text{therm}} = \rho k_B T - \left\langle \frac{\partial U}{\partial V} \right\rangle, \quad (3.11)$$

If the potential energy does not depend explicitly on  $V$ , then the thermodynamic pressure is rewritten as:

$$P_{\text{therm}} = \rho k_B T - \frac{1}{3V} \sum_i \frac{\partial U}{\partial \mathbf{r}_i} \cdot \mathbf{r}_i. \quad (3.12)$$

### Virial pressure

Allen and Tildesley (1987) showed that starting with the virial theorem,

$$\left\langle \mathbf{q}_k \frac{\partial H}{\partial \mathbf{q}_k} \right\rangle = k_B T, \quad (3.13)$$

one may derive an expression for the system pressure that is split into ideal and intermolecular–force components:

$$P_{\text{virial}} = \rho k_B T + \frac{W_{\text{int}}}{V}, \quad (3.14)$$

where  $P_{\text{virial}}$  is the instantaneous pressure and  $W_{\text{int}}$  is the internal virial. For a three–dimensional system, the internal virial is defined as:

$$W_{\text{int}} = \frac{1}{3} \sum_i^{N-1} \sum_{j>i}^N \mathbf{r}_{ij} \cdot \mathbf{f}_{ij}, \quad (3.15)$$

where  $\mathbf{r}_{ij}$  is the distance vector between molecular centres and  $\mathbf{f}_{ij}$  is the force that molecule  $j$  exerts on molecule  $i$ . It is convenient to define the intermolecular pair virial function as:

$$w(r_{ij}) = -r_{ij} \frac{dU(r_{ij})}{dr_{ij}}, \quad (3.16)$$

In which case, the final expression for the virial pressure is:

$$P_{\text{virial}} = \left\langle \rho k_B T - \frac{1}{3V} \sum_{i=1}^{N-1} \sum_{j>i}^N r_{ij} \frac{dU(r_{ij})}{dr_{ij}} \right\rangle. \quad (3.17)$$

Since  $V$  has been factored out of the summation, this expression is to be used when the potential energy does not depend explicitly on the volume. If there is indeed volume dependence, then the thermodynamic definition of pressure must be used (du Preez, 2005).

### 3.2.3 Chemical potential

Evaluating the chemical potential is another important calculation since the condition for thermodynamic equilibrium requires the chemical potential for a species to be the same in all coexisting phases. Calculation of the chemical potential requires a special ‘particle–insertion’ method, developed by Widom (1963). The chemical potential of species  $i$  in phase I in a molecular system in the NVT–Gibbs Ensemble, in which the probability that either of the boxes contains zero molecules is small in addition to the boxes not changing identities, is given by:

$$\mu_{i,i} = -k_B T \ln \left[ \frac{1}{\Lambda^3} \left\langle \frac{V_1}{n_{i,i} + 1} \exp \left( -\frac{\Delta U_{i,i}^+}{k_B T} \right) \right\rangle \right] \quad (3.18)$$

where  $\Delta U_{i,i}^+$  is the potential energy change due to an insertion of particle type  $i$  in phase I. Strictly speaking, this method is not necessary in the Gibbs ensemble, since a particle transfer trial move is inherent to the Gibbs ensemble technique (McKnight, 2005). The remainder of this chapter is based largely on Chapter 3 of Frenkel and Smith (2002).

## 3.3 Molecular Simulation and the Monte Carlo Method

Thus far, the basic concepts of statistical mechanics have been discussed. This section discusses the role of the Monte Carlo scheme as used in molecular simulation. Specifically, emphasis is placed on the canonical (constant NVT) ensemble; however, these ideas are easily extended to other ensembles (see Chapter 4).

In the classical limit, the partition function for the NVT ensemble becomes an integral:

$$Q = c \int \exp[-E / k_B T] d\mathbf{p}^N d\mathbf{r}^N \quad (3.19)$$

where  $\mathbf{r}^N$  and  $\mathbf{p}^N$  stand for the coordinates and corresponding momenta, respectively, of all  $N$  particles.  $E$  is the total energy of the isolated system as a function of the coordinates and momenta of the particles, while  $c$  is a constant of proportionality that renders the partition function dimensionless.

Thus, using the definition of an ensemble average, the average of any observable  $A$  may be written as:

$$\langle A \rangle = \frac{\int A(\mathbf{p}^N, \mathbf{r}^N) \exp[-E/k_B T] d\mathbf{p}^N d\mathbf{r}^N}{\int \exp[-E/k_B T] d\mathbf{p}^N d\mathbf{r}^N} \quad (3.20)$$

In this equation, the observable  $A$  is expressed as a function of momenta and coordinates. It was discussed early in the chapter that the kinetic energy is a quadratic function of the momenta, so that integration can be carried out analytically. The problem arises when computing averages of functions  $A(\mathbf{r}^N)$ . Only in simple cases can this multidimensional integral be solved analytically; in most cases, numerical techniques are required.

At first glance, it would appear that a numerical quadrature technique such as Simpson's Rule be used to evaluate  $\langle A \rangle$ . However, such a method is useless even if the number of independent coordinates  $D \times N$  is very small. For example, in a 3-dimensional system consisting of 100 particles with 5 equidistant points along each axis, the integrand would have to have been evaluated at  $10^{210}$  points. Such computations cannot be performed with current computational power; furthermore, the answer would be subject to large statistical error (Frenkel and Smith, 2002). This is due to numerical quadratures working best on functions that are smooth over distances corresponding to the specified mesh size. For most intermolecular potentials, the term  $\exp(-E/k_B T)$  (the Boltzmann factor) is a rapidly varying function of a system's molecular configuration. Thus, accurate quadrature requires extremely small mesh spacing as far as molecular simulation is concerned. It is interesting

to note that for the evaluation of the integrand for a dense liquid the majority of points the Boltzmann factor is extremely small.

Indeed, better numerical techniques are required to compute statistical thermodynamic averages.

### 3.3.1 Random Sampling

The simplest Monte Carlo method used in the evaluation of integrals is random sampling. Consider the one-dimensional integral I:

$$I = \int_a^b f(x) dx \quad (3.21)$$

which can be rewritten as

$$I = (b - a) \langle f(x) \rangle \quad (3.22)$$

where  $\langle f(x) \rangle$  is the unweighted average of  $f(x)$  over the interval  $[a, b]$ . Thus, the average can be determined by evaluating  $f(x)$  at a large number of  $x$  values randomly distributed over  $[a, b]$ . As explained earlier in the chapter, for the integrals we wish to evaluate in determining thermodynamic averages, a large amount of computational time would be spent in those areas of phase space where the Boltzmann factor is extremely small. It would be more efficient to sample those points in phase space where the Boltzmann factor contributes significantly to the integral. This is the basic idea of importance sampling.

### 3.3.2 Importance Sampling

The problem that arises is how the sampling should be distributed in phase space. One method would be to sample from a non-uniform distribution over the range of integration and then correct for it. Frenkel and Smith (2002) showed that using a probability density  $\zeta(x)$  one may obtain an improved prediction of the integral. Thus, using  $\zeta(x)$ , the integral I becomes:

$$I = \int_a^b \frac{f(x)}{\zeta(x)} \zeta(x) dx \quad (3.23)$$

$$= \int_a^b \frac{f[x(u)]}{\zeta[x(u)]} du \quad (3.24)$$

$$\approx \frac{(a-b)}{\tau} \sum_{i=1}^{\tau} \frac{f[x(u_i)]}{\zeta[x(u_i)]} \quad (3.25)$$

where  $\tau$  is the number of random sample points taken from the distribution  $\zeta(x)$ . Unfortunately, the probability distributions and the partition functions are not known *a priori*, and thus the simple importance sampling scheme described above cannot be used to sample the multidimensional integrals of interest.

### 3.3.3 The Metropolis Method

In general it is not possible to evaluate integrals of the form  $\int \exp[-\beta U(\mathbf{r}^N)] d\mathbf{r}^N$  when using direct Monte Carlo sampling. The Metropolis method of sampling involves the construction of a random walk through phase space where the probability distribution is non-negligible. Frenkel and Smit (2002) use the analogy that this method is akin to determining the average depth of the river Nile by taking measurements within the Nile *only*, whereas the method of random sampling would sample the entire African region to determine the same average depth.

In order to generate points in phase space that contribute significantly to the integrals of interest, a relative probability proportional to the Boltzmann factor is used. The general approach is to prepare the system in a certain configuration which is denoted  $o$  (old). A new trial configuration is generated which is denoted as  $n$  (new). A decision must now be made whether to accept or reject the new configuration, based on the Boltzmann factor (which is a function of the molecular configuration).

Consider the integral for computing the thermal average of a macroscopic property  $A$ :

$$\langle A \rangle = \frac{\int A(\mathbf{r}^N) \exp[-\beta U(\mathbf{r}^N)] d\mathbf{r}^N}{\int \exp[-\beta U(\mathbf{r}^N)] d\mathbf{r}^N} \quad (3.26)$$

The probability density of finding the system in a configuration  $\mathbf{r}^N$  shall be denoted as

$$\Omega(\mathbf{r}^N) \equiv \frac{\exp[-\beta U(\mathbf{r}^N)]}{Q}, \quad (3.27)$$

while the probability of a transition from the old configuration ( $o$ ) to the new configuration ( $n$ ) is the transition probability matrix. The matrix elements  $\pi(o \rightarrow n)$  must satisfy the condition of not destroying an equilibrium distribution once it is reached (Frenkel and Smit, 2002)<sup>3</sup>. A much stronger condition known as the condition of detailed balance states that in equilibrium the average number of accepted trial moves from  $o$  to any  $n$  is exactly cancelled by the number of reverse moves. This implies:

$$\Omega(o)\pi(o \rightarrow n) = \Omega(n)\pi(n \rightarrow o), \quad (3.28)$$

---

<sup>3</sup> The average number of accepted trial moves that result in the transition  $o \rightarrow n$  must be identical to the average number of accepted trial moves from  $n \rightarrow o$  (Frenkel and Smith, 2002).

A Monte Carlo move consists of two stages: a trial move from state  $o$  to state  $n$  is first performed, and the probability of performing this trial move is denoted by the matrix  $\alpha(o \rightarrow n)$ ; <sup>4</sup> the next stage is to decide whether to accept or reject this trial move, and the probability of accepting this trial move from  $o$  to  $n$  is  $acc(o \rightarrow n)$ ,

Thus, one may write

$$\pi(o \rightarrow n) = \alpha(o \rightarrow n) \times acc(o \rightarrow n). \quad (3.29)$$

An important concept, namely, that of symmetry, is necessary for further development. It shall be assumed that the probability of performing a trial move from  $o$  to  $n$  is equal to the probability of the reverse move, i.e.

$$\alpha(o \rightarrow n) = \alpha(n \rightarrow o). \quad (3.30)$$

Therefore, Equation (3.28) may be written as

$$\Omega(o) acc(o \rightarrow n) = \Omega(n) acc(n \rightarrow o). \quad (3.31)$$

Recalling the definition of  $\Omega(\mathbf{r}^N)$ , it follows that

$$\frac{acc(o \rightarrow n)}{acc(n \rightarrow o)} = \frac{\Omega(n)}{\Omega(o)} = \exp\{-\beta[U(n) - U(o)]\}. \quad (3.32)$$

Metropolis *et al.* devised an efficient strategy for sampling of phase space using the following acceptance criteria:

---

<sup>4</sup>  $\alpha$  is referred to as the underlying matrix of the Markov chain.

$$\begin{aligned} \text{acc}(o \rightarrow n) &= \Omega(n)/\Omega(o) & \text{if } \Omega(n) < \Omega(o) \\ &= 1 & \text{if } \Omega(n) \geq \Omega(o) \end{aligned} \quad (3.33)$$

Thus, if the new molecular configuration has a lower energy than the old configuration, the move is accepted with probability  $\Omega(n)/\Omega(o)$ , otherwise, the move is accepted with a probability of one. In practice, however, a random number<sup>5</sup> is generated in the interval [0, 1]. The trial move is accepted if the random number is less than  $\text{acc}(o \rightarrow n)$ , and rejected otherwise. This ensures an equilibrium state is continuously approached in the chemical system, since the potential energy of each new state progressively decreases (provided that the trial move is accepted). It is important to note that if a trial move is not accepted, then the original state must be accounted for again, since

$$\pi(o \rightarrow o) = 1 - \sum_{n \neq o} \pi(o \rightarrow n). \quad (3.34)$$

Clearly, this is a positive number meaning that the original state needs to be recounted.

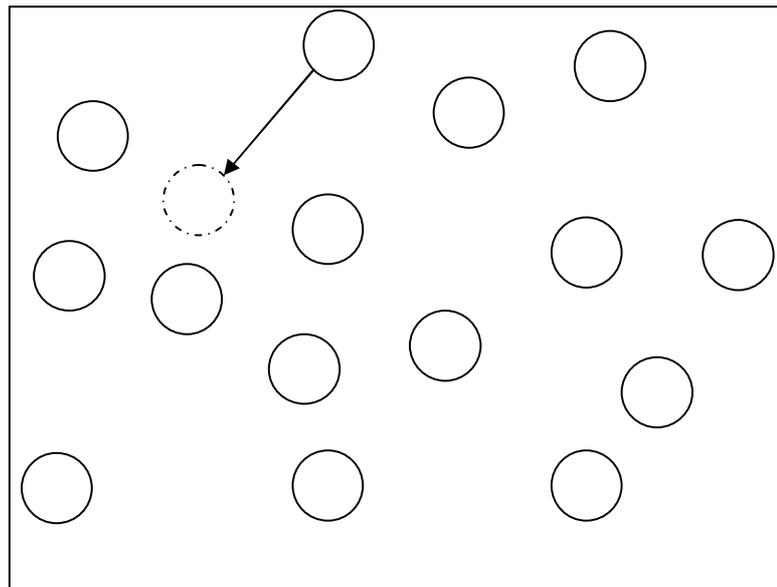
---

<sup>5</sup> Random numbers shall be discussed in the Chapter 7.

## 4. Ensembles

### 4.1 Canonical (NVT) Ensemble

In the context of statistical mechanics, ‘canonical’ means standard. This is the ensemble that was originally used by Metropolis *et al.* (1953) in the first Monte Carlo simulations. The total number of molecules (N), system temperature (T) and system volume (V) are fixed, and as such, only particle displacement moves (translational and orientation) are permitted (for the sake of simplicity, only monatomic molecules shall be used in the illustrations in the remainder of this chapter).



**Figure 4-1 - The Canonical Ensemble.** The number of molecules, system volume and temperature are constant. Only particle displacements are allowed (translational or orientation, in the case of polyatomic molecules).

The partition function for the NVT ensemble is

$$Q_{NVT} = \frac{1}{\Lambda^{3N} N!} \int \exp[-\beta U(\mathbf{r}^N)] d\mathbf{r}^N, \quad (4.1)$$

where  $\Lambda = \sqrt{h^2 / 2\pi m k_B T}$  is the *thermal* de Broglie wavelength, roughly the average de Broglie wavelength of the gas particles in an ideal gas at the temperature  $T$ ; when it is much smaller than the interparticle distance, the gas can be considered to be a classical gas. Otherwise, quantum effects will dominate and the gas must be treated as a Fermi gas or a Bose gas. From Equation 4.1 it is clear that the probability of finding the system of interest in a configuration  $\mathbf{r}^N$ , in terms of the potential energy is

$$\Omega(\mathbf{r}^N) \propto \exp[-\beta U(\mathbf{r}^N)], \quad (4.2)$$

which is the distribution that must be sampled during a simulation. The following procedure is used when carrying out simulations in the NVT ensemble:

1. Select a particle at random and calculate the potential energy of this configuration,  $U(o)$ .
2. Displace the particle from its current position, i.e.  $\mathbf{r}(o)$ , using

$$\mathbf{r}(n) = \mathbf{r}(o) + \Delta(\text{RND} - 0.5), \quad (4.3)$$

where  $\Delta/2$  is the maximum displacement, RND is a random number in the interval  $[0,1]$  and  $\mathbf{r}(n)$  is the new trial position. Then calculate the potential energy of this new configuration,  $U(n)$ .

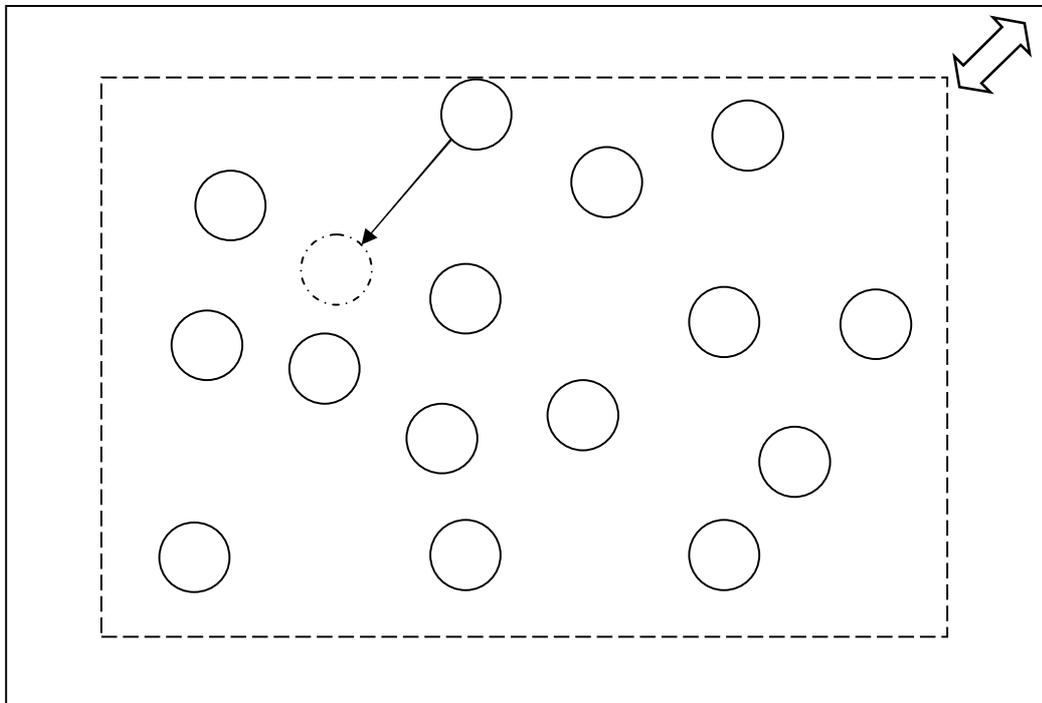
3. The acceptance probability is

$$\text{acc}(o \rightarrow n) = \min(1, \exp\{-\beta[U(n) - U(o)]\}). \quad (4.4)$$

If the move is rejected, the old configuration is maintained, and recounted.

## 4.2 Isobaric–Isothermal (NPT) Ensemble

The isobaric–isothermal ensemble has become very popular in molecular simulations due to real experiments being performed under constant pressure and temperature conditions, and importantly, in generating phase diagrams for multicomponent chemical systems. This ensemble offers added flexibility due to system volume changes also being possible, aside from the displacement moves used in the NPT ensemble.



**Figure 4-2 - The Isobaric–Isothermal Ensemble.** The number of molecules, system pressure and temperature are constant. In addition to particle displacements being allowed (translational or orientation, in the case of polyatomic molecules) volume changes of the simulation box are also permitted.

The volume is allowed to fluctuate so as to keep the system pressure constant. It is useful to define scaled coordinates for each particle:

$$\mathbf{r}_i = L\mathbf{s}_i \quad \text{for } i = \{1, 2, \dots, N\}, \quad (4.5)$$

where  $L = V^{1/3}$  is the length of the cubic box.<sup>6</sup> For a system containing  $N$  identical atoms, constant pressure  $P$ , with a total volume  $V$  and an imaginary piston that has a variable volume  $V_0$ , the partition function is

$$Q_{NPT} \equiv \frac{\beta P}{\Lambda^{3N} N!} \int V^N \exp(-\beta P V) dV \int \exp[-\beta U(\mathbf{s}^N; L)] d\mathbf{s}^N, \quad (4.6)$$

where  $U(\mathbf{s}^N; L)$  has been written to indicate that  $U$  depends on the *real* distances between the atoms. Now, the probability that the  $N$ -particle piston subsystem has a volume  $V$  is given by

$$\Omega_{NPT}(V) = \frac{V^N \exp(-\beta P V) \int \exp[-\beta U(\mathbf{s}^N; L)] d\mathbf{s}^N}{\int_0^V V'^N \exp(-\beta P V') dV' \int \exp[-\beta U(\mathbf{s}^N; L')] d\mathbf{s}^N}, \quad (4.7)$$

while the probability density of finding the subsystem at the given volume  $V$  in a particular molecular configuration, using scaled coordinates is:

$$\begin{aligned} \Omega_{NPT}(V; \mathbf{s}^N) &\propto V^N \exp(-\beta P V) \exp[-\beta U(\mathbf{s}^N; L)] \\ &= V^N \exp\{-\beta[U(\mathbf{s}^N; V) + P V - N\beta^{-1} \ln V]\}. \end{aligned} \quad (4.8)$$

In the NVT ensemble, it was shown that the particle coordinates were the only variables taken into account for the trial moves. In the NPT ensemble,  $V$  is treated as an additional coordinate since it is allowed to vary; concordantly, there are acceptance rules for such a change which must satisfy the underlying Markov chain. If the new trial move consists of a

---

<sup>6</sup> In this work, all simulations have been performed in cubic boxes.

volume change  $V(n) = V(o) + \Delta V$ , where  $\Delta V = \{-\Delta V_{\max}, \Delta V_{\max}\}$ , then this move will be accepted with the following probability:

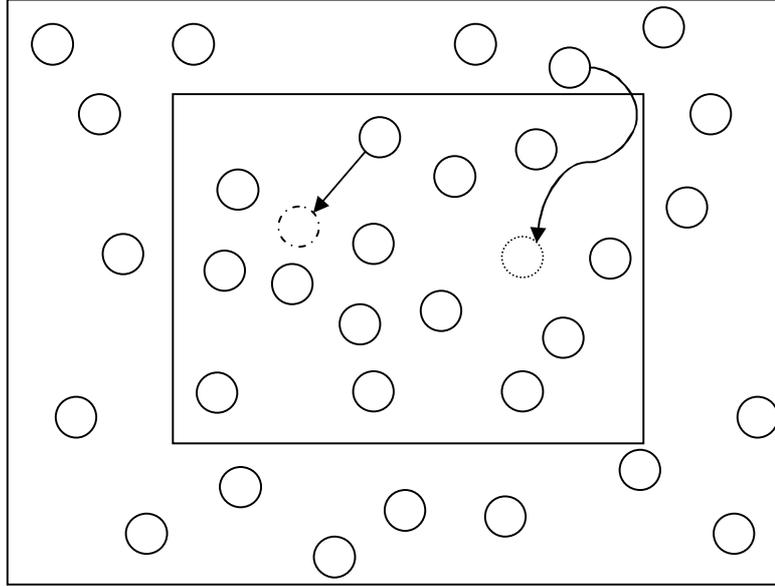
$$acc(o \rightarrow n) = \min \left( 1, \exp \left\{ -\beta \left[ U(\mathbf{s}^N; V(n)) - U(\mathbf{s}^N; V) + P(V(n) - V) - N\beta^{-1} \ln \frac{V(n)}{V} \right] \right\} \right). \quad (4.9)$$

It is now clear why scaled coordinates are used: each time a volume change is attempted, the potential energy for the system must be recalculated, which is a computationally expensive calculation. However, in this work, the potential energy function used to calculate the non-bonded intermolecular is the Lennard–Jones 12–6 model for which the potential energy of the new state is easily determined when scaled coordinates are used:

$$U_{\text{inter}}(L(n)) = \left( \frac{L}{L(n)} \right)^n U_{\text{inter}}(L). \quad (4.10)$$

### 4.3 Grand–Canonical ( $\mu VT$ ) Ensemble

Thus far, the ensembles that have been presented have had the total number of particles held constant. In contrast, the grand-canonical ensemble maintains a constant chemical potential whilst the number of particles is allowed to vary.



**Figure 4-3 - The Grand-Canonical Ensemble.** The system (enclosed by the inner rectangle) has a constant chemical potential, volume and temperature. In addition to displacement type moves, particle insertions are also permitted.

Particle insertions and deletions are allowed so as maintain a constant chemical potential within the volume  $V$ . Once again, it is useful to use scaled coordinates. The partition function may be written as:

$$Q_{\mu VT} = \sum_{N=0}^{\infty} \frac{\exp[\beta\mu N] V^N}{\Lambda^{3N} N!} \int \exp[-\beta U(\mathbf{s}^N)] d\mathbf{s}^N, \quad (4.11)$$

while the probability density is:

$$\Omega_{\mu VT}(N, \mathbf{s}^N) \propto \frac{\exp[\beta\mu N] V^N}{\Lambda^{3N} N!} \exp[-\beta U(\mathbf{s}^N)]. \quad (4.12)$$

The acceptance criteria for a displacement trial move in the grand-canonical ensemble are identical to those proposed for the canonical ensemble. Additionally, the acceptance criteria when particles are inserted or removed, respectively, are:

$$acc(N \rightarrow N+1) = \min \left[ 1, \frac{V}{\Lambda^3(N+1)} \exp\{\beta[\mu - U(N+1) + U(N)]\} \right] \quad (4.13)$$

and

$$acc(N \rightarrow N-1) = \min \left[ 1, \frac{\Lambda^3 N}{V} \exp\{\beta[-\mu - U(N-1) - U(N)]\} \right] \quad (4.14)$$

The grand-canonical ensemble is seldom used in the study of phase equilibrium and is more suited to adsorption studies (Frenkel and Smit, 2002).

#### 4.3.1 Histogram–reweighting grand canonical Monte Carlo simulations

At the outset of the development of molecular simulation it was theorized that a single calculation could be used to obtain information on the properties of a system for a range of state conditions, though at the time the lack of adequate computational power deterred further investigation into this concept. With the rapid development in computing technologies in recent times, several researchers have proven the usefulness of histogram–reweighting. The method was originally used to perform simulations near the critical point, after which new states near the critical point were calculated using reweighting.

#### 4.4 Tracing Coexistence Curves – Gibbs–Duhem Integration<sup>7</sup>

Much like the histogram–reweighting method, the Gibbs–Duhem integration method of studying phase coexistence relies on knowing a single point on the coexistence curve *a priori*. Kofke proposed a method that is the same as numerically integrating the Clausius–Clapeyron equation:

---

<sup>7</sup> This method of determining phase diagrams relies on simulation performed in the ensembles that have been discussed, and deserves to be mentioned in this chapter.

$$\frac{dP}{dT} = \frac{\Delta H}{P \Delta V / T} \quad (4.15)$$

where  $\Delta H$  is the difference in enthalpy of two phases. All quantities on the right hand side of the Clausius–Clapeyron equation can be computed during a simulation; thus,  $dP/dT$  may also be calculated. Gibbs–Duhem integration is potentially a highly efficient method for tracing coexistence curves (Frenkel and Smith, 2002), but errors in the integration of the Clausius–Clapeyron equation can result in significant errors when comparing the simulated coexistence curve with the true coexistence curve. Bolhuis and Frenkel (1997) assumed that the coexistence curve could be fitted to a polynomial in  $T$ :

$$\frac{dP}{dT} = \sum_{j=0}^3 \alpha_j T^j . \quad (4.16)$$

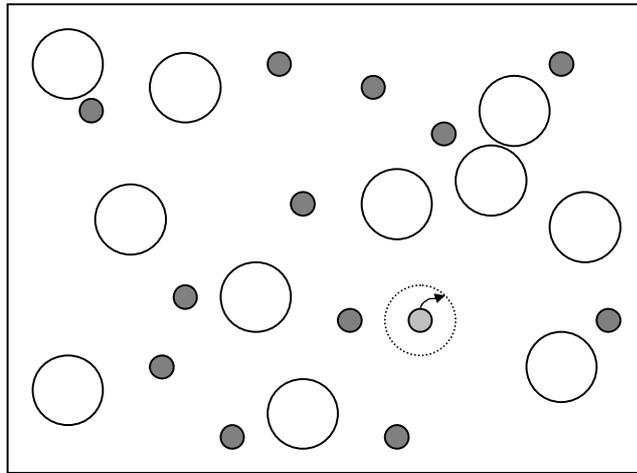
The integration scheme is performed as follows:

1. The original Gibbs–Duhem integration is carried out to obtain an initial estimate for the coexistence curve; at every point, the right–hand side of the Clausius–Clapeyron equation is calculated.
2. The numerical data is then fitted to the coefficients of the polynomial, Equation (4.16). A new estimate of the of the coexistence pressures is thus obtained.
3. The newly calculated pressures from Step 2 and the initial pressures are combined to improve stability.

The aforesaid scheme is iterated until convergence is achieved.

## 4.5 Semi-grand Ensemble

This method, yet another method for studying the phase behaviour of chemical mixtures uses a new type of ‘move’ called an identity change. The semi–grand ensemble is a variant of the grand–canonical ensemble that uses a fixed chemical potential on one particle type, thus imposing the chemical potential of other types of particles.



**Figure 4-4 - The Semigrand Ensemble. In this ensemble, molecular identity changes are allowed. Shown in the schematic is an attempt to transform a small particle into the larger type.**

Frenkel and Smith (2002) provide an elegant derivation for the semi–grand partition function starting with the grand–canonical particle function for an  $n$ –component mixture:

$$\mathcal{Q}_{\mu VT} = \sum_{N_1, N_2, \dots, N_n} \prod_{i=1}^n \frac{q_i^{N_i} \exp[\beta \mu_i N_i] V^N}{N_i!} \int \exp[-\beta U(\mathbf{s}^N)] d\mathbf{s}^N, \quad (4.17)$$

where  $N \equiv \sum_i N_i$  and  $q_i$  is the kinetic contribution to the partition function from species

*i*. Finally, the probability of accepting an identity change is:

$$acc(\xi_i \rightarrow \xi'_i) = \min \left\{ 1, \frac{\xi'_i}{\xi_i} \exp[-\beta \Delta U(\mathbf{s}^N)] \right\}, \quad (4.)$$

where  $\xi_i$  is the fugacity fraction of species  $i$ .

## 5. Potential Energy Models

Thus far, methods have been presented on determining what moves may be accepted in order to ensure that the points being sampled are in equilibrium phase space (typically, these would be points on an energy hypersurface). However, methods for calculating the intra- and intermolecular potential energies for a given configuration of molecules in a chemical system need to be discussed.

The total potential energy of the molecular system may be written as a sum of intramolecular and intermolecular energies:

$$U_{\text{total}} = U_{\text{intra}} + U_{\text{inter}}, \quad (5.1)$$

where  $U_{\text{intra}}$  refers to the intramolecular interactions and  $U_{\text{inter}}$  refers to the intermolecular interactions.

### 5.1 Intramolecular Interactions

#### 5.1.1 Bond Stretching

The stretching or distortion of bonds between atoms leads to a certain deformation energy, which is typically modelled as a Taylor series expansion around a natural bond length  $l_0$ . This energy is described as

$$U_{\text{str}} = U_0 + \frac{dU}{dl}(l-l_0) + \frac{1}{2!} \frac{d^2U}{dl^2}(l-l_0)^2 + \text{higher order terms}. \quad (5.2)$$

The term  $U_0$  is usually set to zero as a reference point, whilst the second term is zero since the gradient  $dU/dl$  at  $l = l_0$  is zero. It is common practice to exclude terms higher than the quadratic term in the expression given above, since they contribute very little to the bond-stretching energy. Thus, the stretching energy is described by a harmonic potential:

$$U_{\text{str}} = \frac{1}{2}k_{\text{str}}(l - l_0)^2, \quad (5.3)$$

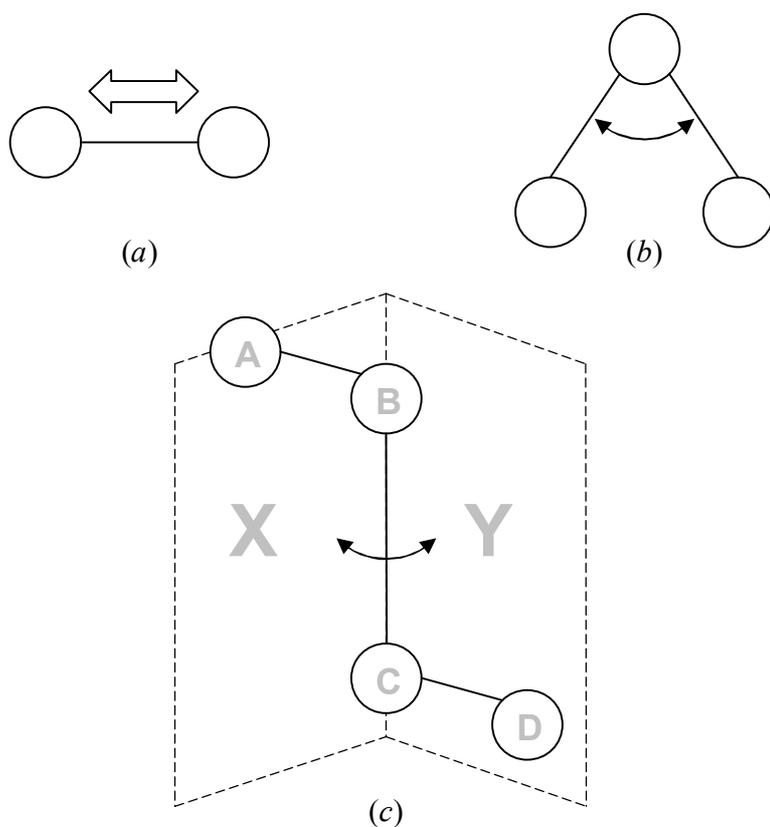
where  $k_{\text{str}}$  is the stretching constant.

### 5.1.2 Bond Bending

The bond bending component of the intramolecular interactions is only valid when a molecule has three or more atomic groups present. The bond angle is defined as the acute angle formed by two bonds that are connected to a common atomic group (see Figure 5-1). The functional form of bond bending term, like the bond stretching term, is harmonic:

$$U_{\text{bend}} = \frac{1}{2}k_{\text{bend}}(\theta - \theta_0)^2, \quad (5.4)$$

where  $k_{\text{bend}}$  is the bending constant and  $\theta_0$  is the equilibrium bending angle.



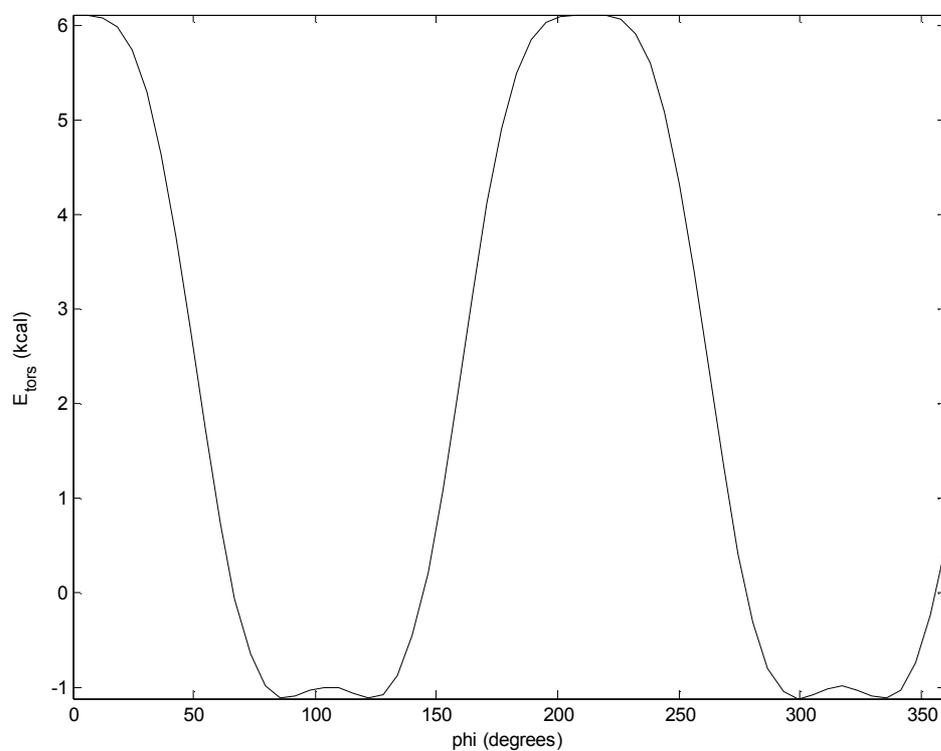
**Figure 5-1 - The various intramolecular interactions that contribute towards the potential energy of a given molecular configuration. From top left (clockwise): (a) Bond Stretching (b) Bond Bending (c) Torsion Energy.**

### 5.1.3 Torsion Energy

In order for torsion or ‘twisting’ energy to be present, there must be at least four bonded atomic groups in the molecule (see Figure 5-1). The angle formed by the planes A and B in the figure is referred to as the dihedral angle  $\phi$ , i.e. the angle formed by atoms A-B-C and atoms B-C-D. Since the orientation of the molecule does not matter when measuring the dihedral angle, it is clear that the energy interactions may be defined by a periodic function, specifically, a cosine series of the following standard form:

$$U_{\text{tors}} = c_0 + c_1[1 + \cos \phi] + c_2[1 - \cos 2\phi] + c_3[1 + \cos 3\phi], \quad (5.5)$$

where the  $c_i$  are constants that are determined (usually) by *ab initio* methods. This form of the torsion energy is implemented by many recently parameterized potential models, including the Transferable Potentials for Phase Equilibria – United-Atom (TraPPE-UA) force field (to be discussed later in this chapter).



**Figure 5-2 - Representation of the variation of the torsion energy for the C—C—C—C bonded group in perfluoroalkanes as parameterized by Watkins and Jorgensen (2001)**

## 5.2 Intermolecular Interactions

The intermolecular interactions are split into two parts – a van der Waals component that describes the non-polar interactions, and a Coulombic component that describes the charged (or polar) interactions. Thus, the contribution of each both parts may be summed to obtain the total intermolecular potential:

$$U_{\text{inter}} = U_{\text{vdW}} + U_{\text{Coul}} . \quad (5.6)$$

### 5.2.1 van der Waals interactions

All molecules are subject to two distinct forces, namely, an attractive force at long distances (termed the van der Waals or dispersion force) and a repulsive force at short distances. The latter force is as a result of overlapping electron orbitals, referred to as Pauli repulsion<sup>8</sup>. The Lennard-Jones 12–6 potential (Lennard-Jones, 1931) is a simple mathematical formula that models this behavior:

$$U_{\text{LJ}}^{ij}(r_{ij}) = 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (5.7)$$

where  $\varepsilon_{ij}$ ,  $\sigma_{ij}$  and  $r_{ij}$  are the depth of the potential well, the distance at which the force is zero, and the separation between two atomic sites  $i$  and  $j$ , respectively.

---

<sup>8</sup> From the Pauli Exclusion Principle.

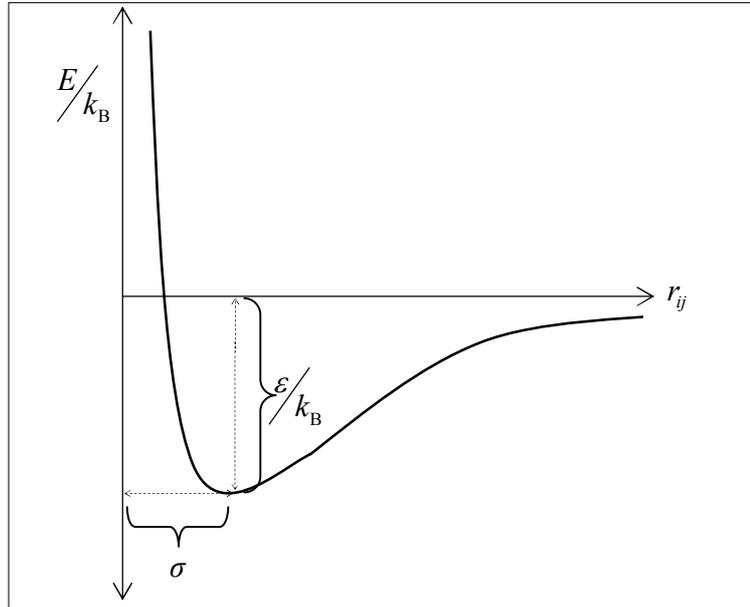


Figure 5-3 - Illustration of the potential energy well of a Lennard–Jones fluid.

The  $r^{-12}$  term represents the repulsion between atoms when they are brought close to each other. The exponent 12 was chosen entirely for practical purposes and has no physical basis, though, theoretically, exponential behavior would be more appropriate. The long-range attractive forces are represented by the  $r^{-6}$  term. The  $\epsilon_{ij}$  and  $\sigma_{ij}$  terms are obtained from quantum chemistry calculations for each molecule type, and are thus unique.

Recently, more accurate methods of calculating the van der Waals interaction energy have been proposed, of which the most notable is the Buckingham exponential–6 model:

$$U_{\text{exp-6}}^{ij}(r_{ij}) = \begin{cases} \frac{\epsilon_{ij}^{\text{exp-6}}}{1-6/\omega} \left[ \frac{6}{\omega} \exp\left(\omega \left[1 - \frac{r_{ij}}{r_{\text{max}}}\right]\right) - C \left(\frac{r_{\text{max}}}{r_{ij}}\right)^6 \right] \\ \infty \end{cases} \quad (5.8)$$

where the cut-off distance  $r_{\text{max}}$  is the smallest positive value for which  $dU_{ij}^{\text{exp-6}}(r)/dr = 0$  (obtained by iterative solution).

### 5.2.2 Coulombic Interactions

The Coulombic interactions, relevant to systems of molecules that contain charged or partially charged species, are calculated using Coulomb's law of electric interaction between two charged bodies:

$$U_{\text{Coul}}^{ij} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \quad (5.9)$$

where  $q_i$  and  $q_j$  are the charges on interacting sites  $i$  and  $j$ , while  $\epsilon_0 = 8.854 \times 10^{-12} \text{ C}^2 \text{ N}^{-1} \text{ m}^{-2}$  (also  $\text{F m}^{-1}$ ) is the permittivity of free space.

### 5.2.3 Mixing Rules

To account for the heterogeneity between the atomic interaction sites of different molecules, the well-established Lorentz–Berthelot mixing rules provide a convenient means for calculating the cross-term ( $ij$ ) parameters:

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}}, \text{ and} \quad (5.10)$$

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}). \quad (5.11)$$

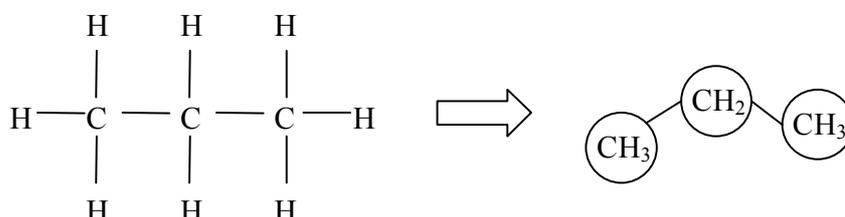
The energy cross-term is based on a geometric average (with no physical significance) whilst the size parameter cross-term is based on an average diameter of the interacting sites.

## 5.3 Potential Models

The two force fields used in this work are the united atom description of the TraPPE model developed by Martin and Siepmann (1998) and the extended SPC model, developed by Berendsen *et al.* (1987).

### 5.3.1 Transferable Potentials for Phase Equilibria – United Atom description (TraPPE-UA)

TraPPE-UA, developed by Martin and Siepmann (1998), models linear alkanes by representing each functional group i.e. methyl ( $\text{—CH}_3$ ) and ethyl ( $\text{—CH}_2$ ) groups, as single Lennard-Jones interactions sites, although separate parameters were developed for methane and ethane. The Lennard-Jones size and energy parameters were fitted to experimental critical temperatures and saturated liquid densities in the canonical ensemble; simulations revealed that the vapour pressures and liquid densities for the pure components were in good agreement with experimental data, showing small, systematic deviations.



**Figure 5-4 – An illustrative example using n-propane, showing the philosophy of united-atom force fields. The hydrogen atoms are ‘lumped’ onto the carbon atoms; thus, hydrogen interactions are implicitly accounted for in this force field methodology.**

TraPPE-UA utilizes rigid bonds; as such, there are no vibrations between the pseudoatoms in any molecule that is described by this force field. Thus, all inter- and intramolecular interactions that were described in the previous subchapter, with the exception of the vibration term, are applicable to this potential model.

In further studies, the model was extended to implement a variety of functional groups and also bonds types, including alcohols (Chen *et al.*, 2001) and alkenes (Wick *et al.*, 2000), respectively.

A list of TraPPE-UA pseudoatoms used in this work is shown in the table below<sup>9</sup>:

Pseudoatom	Description
<b>CH<sub>4</sub></b>	Methane
<b>CH<sub>3</sub>(sp<sup>3</sup>)</b>	methyl group; one sp <sup>3</sup> carbon bonded to three hydrogen atoms and one non-hydrogen atom.
<b>CH<sub>2</sub>(sp<sup>3</sup>)</b>	ethyl group; one sp <sup>3</sup> carbon bonded to two hydrogen atoms and two other non-hydrogen atoms.
<b>CH<sub>2</sub>(sp<sup>2</sup>)</b>	sp <sup>2</sup> carbon atom that is bonded to two hydrogen atoms and double-bonded to a non-hydrogen atom.
<b>O (sp<sup>3</sup>)</b>	sp <sup>3</sup> oxygen atom that is single bonded to one carbon atom and single bonded to one hydrogen atom, used in alcohols.
<b>H</b>	Hydrogen atom bonded to one oxygen atom.

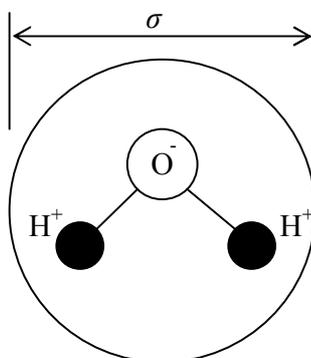
**Table 5-1 - List of pseudoatoms implemented in the TraPPE-UA force field that relevant to this study.**

### 5.3.2 Simple Point Charge – Extended, model of water molecule

The SPC-E force field (Berendsen *et al.*, 1987) is a reparameterized form of the SPC force field (Berendsen *et al.*, 1981). In this model, the water molecule is represented as single interaction site, that is to say, the Lennard-Jones size and energy parameters describe water as one pseudoatom. However, three charge sites are used to characterize the electrostatic interactions. These interaction sites model the molecules' Coulombic interactions and as

<sup>9</sup> A comprehensive list of pseudoatoms implemented in TraPPE-UA is available at <http://towhee.sourceforge.net/forcefields/trappeua.html>

such are key in predicting such phenomena as water clustering and association. For details on the values of this potential model's parameters, refer to Appendix A.



**Figure 5-5 – Schematic representation of the SPC-E water model.**

The SPC model has been found to give better predictions than SPC-E for the saturated vapour pressure of water, but SPC-E was found to better-predict the saturated liquid densities over the temperature range of interest (Boulougouris *et al.*, 1998). The justification of this model's use in this study is briefly discussed in Chapter 10.

## 6. The Gibbs Ensemble

The condition for the coexistence of two or more phases I, II, ... is that the pressure and temperature for all coexisting phases must be equal, in addition to the chemical potential of each species  $i$  being the same for that species in all phases (Frenkel and Smith, 2002):

$$(P_1=P_{II} = \dots = P), (T_1=T_{II} = \dots = T) \text{ and } (\mu_{1,i}=\mu_{II,i} = \dots = \mu_i). \quad (6.1)$$

It would seem appropriate to devise a constant  $\mu$ PT ensemble to simulate phase equilibrium but such an ensemble is not possible. The reason is twofold: firstly, if only intensive parameters are specified, then there is no physical limit on the size of the system (Frenkel and Smit, 2002)<sup>10</sup>, and secondly, one cannot stipulate all intensive variables *a priori* as it would correspond to an over-specification of the state of the system (McKnight, 2005). The Gibbs ensemble method has emerged as the technique of choice for determining the phase coexistence curves of fluid mixtures (Panagiotopoulos, 2000). Prior to the introduction of the Gibbs ensemble, molecular simulations had to locate phase coexistence indirectly. By performing several simulations and subsequently evaluating the macroscopic properties of each phase, a point where the temperature, pressure and the chemical potentials of all species are equal in each phase would be located (Widom, 1963, Lofti *et al.*, 1992, Kofke, 1993a and Kofke, 1993b).

A Gibbs ensemble simulation is performed in microscopic regions within each bulk phase, using periodic boundary conditions. Figure 6-1 is a schematic of a one-component two-phase system at constant  $N$ ,  $V$  and  $T$ , for which the trial move acceptance rules were originally derived (Panagiotopoulos, 1987).

---

<sup>10</sup> The set of parameters  $\mu$ ,  $P$  and  $T$  is said to be linearly dependent.

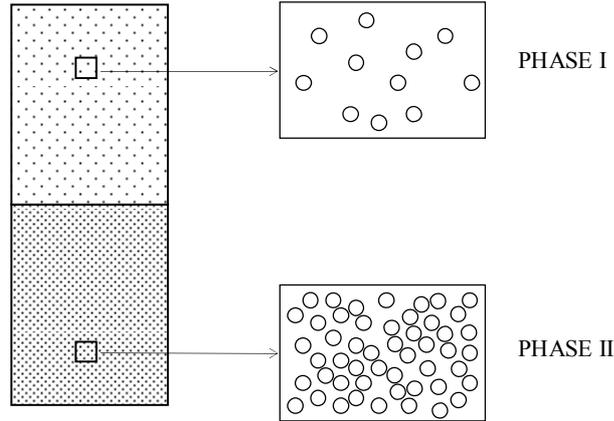


Figure 6-1 - Schematic of the Gibbs ensemble technique for a single component, two-phase system.

The partition function for this system is:

$$Q_{NVT}^{\text{GE}} = \frac{1}{\Lambda^{3N} N!} \sum_{n_1=0}^N \binom{N}{n_1} \int V_1^{n_1} V_2^{n_2} dV_1 \int \exp[-\beta U(n_1)] ds_1^{n_1} \times \int \exp[-\beta U(n_2)] ds_2^{n_2} \quad (6.2)$$

where  $N = n_1 + n_2$ . The three possible trial moves in the Gibbs ensemble are:

### 1. Particle displacements

The acceptance rule can be derived by imposing the condition of detailed balance,  $K(o \rightarrow n) = K(n \rightarrow o)$ . Assume that the new state  $n$  is obtained from the original state  $o$  by displacing a randomly selected particle within a single simulation box. The acceptance rule for a particle displacement is:

$$\text{acc}(o \rightarrow n) = \min\left(1, \exp\left\{-\beta[U(\mathbf{s}_{\text{new}}^{n_1}) - U(\mathbf{s}_{\text{old}}^{n_1})]\right\}\right), \quad (6.3)$$

where  $\mathbf{s}_{\text{new}}$  and  $\mathbf{s}_{\text{old}}$  refer, respectively, to the particle coordinates of the new and old configurations using the scaled coordinates formalism and the  $n_1$  superscript

denotes that the randomly displaced particle resides in box I. This acceptance rule is identical to the conventional NVT ensemble acceptance criterion.

## 2. Volume changes

For a mutual exchange of volume of two randomly selected boxes, chosen here as boxes I and II, the acceptance rule is:

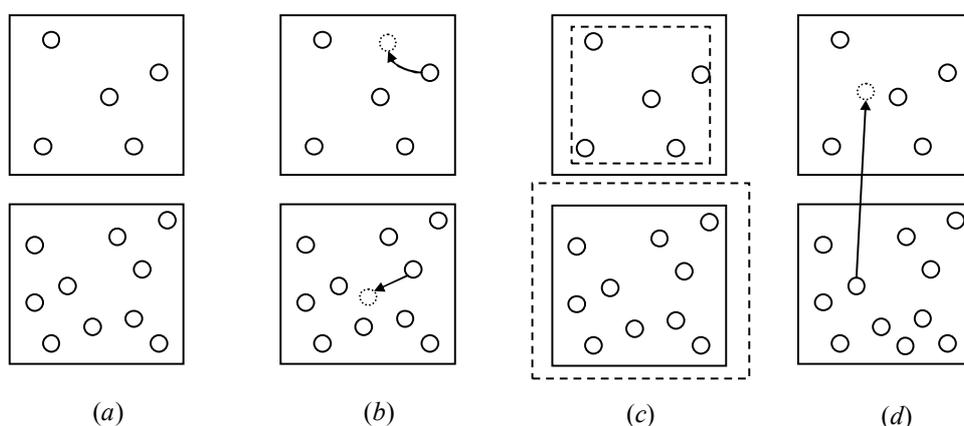
$$acc(o \rightarrow n) = \min \left\{ 1, \frac{(V_{\text{new}}^I)^{n_I} (V - V_{\text{new}}^{\text{II}})^{n_{\text{II}}}}{(V_{\text{old}}^I)^{n_I} (V - V_{\text{old}}^I)^{n_{\text{II}}}} \times \exp[-\beta U(\mathbf{s}_n^{n_I+n_{\text{II}}}) + \beta U(\mathbf{s}_o^{n_I+n_{\text{II}}})] \right\}, \quad (6.4)$$

where  $V_{\text{new}}^I = V_{\text{old}}^I + \Delta V$  is the new volume of box I,  $V_{\text{old}}^I$  is its original volume and  $V = V_{\text{old}}^I + V_{\text{old}}^{\text{II}}$ . This is the original method of changing the volume as derived by Panagiotopoulos *et al.* (1987 and 1988).

## 3. Particle insertions

The acceptance rule for removing a particle of species  $i$  from box I and inserting it in box II is written for a multi-component mixture as (Frenkel and Smit, 2002):

$$acc(o \rightarrow n) = \min \left\{ 1, \frac{n_{I,i} (V - V^I)^{n_{\text{II}}}}{(n_{\text{II}} + 1) V^I} \times \exp[-\beta U(\mathbf{s}_{\text{new}}^{n_I+n_{\text{II}}}) - \beta U(\mathbf{s}_{\text{new}}^{n_I+n_{\text{II}}})] \right\}. \quad (6.5)$$

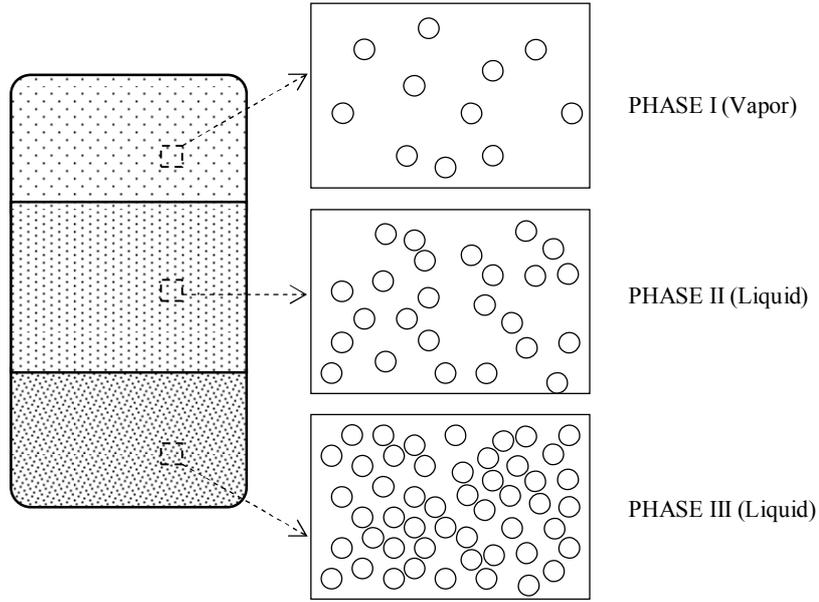


**Figure 6-2 - Gibbs ensemble Monte Carlo trial moves. (a) the original configuration. (b) particle displacements (c) volume changes (d) particle swaps.**

For the purposes of this work, both NVT and NPT variants of the Gibbs ensemble are used. Usually the NPT–Gibbs ensemble is implemented in phase equilibrium simulations since this is analogous to carrying out a real experiment under constant temperature, pressure and compositions conditions. The NVT analogue may also be used in multi-component, multi-phase simulations, but requires judicious selection of the total system volume. An important point to note is that the constant–P method is best applied to chemical systems that contain two or more components, since in a one–component system, the two–phase region is a line in the P–T plane, whilst in a two–component system, this region constitutes a plane (or a hypersurface, when more than two components are present).

In the original article whence the Gibbs ensemble was introduced, the partition function and corresponding acceptance rules were derived for the two–phase coexistence of a pure component. The partition function for the NVT variant of three– and multi–phase systems were derived by Canongia Lopes and Tildesley (1997), and are presented here, followed by the acceptance rules for the isobaric–isothermal ensemble for binary and ternary mixtures.

Consider a 3–phase chemical system at equilibrium, enclosed in a volume  $V$  and consisting of  $N$  identical particles, with temperature  $T$ :



**Figure 6-3 - Schematic of the Gibbs ensemble simulation technique for a three-phase (vapour-liquid-liquid) system.**

The NVT–Gibbs ensemble partition function can be written as a combination of the canonical partition functions of each phase. For **two** phases, the Gibbs ensemble partition function is:

$$Q_{NVT} = \sum_{n_1=0}^N \int_0^V Q_{n_1 V_1 T} Q_{n_1 V_1 T} dV_1, \quad (6.6)$$

where  $n_1$  is the number of particles in phase I and  $V_1$  denotes the volume of phase I. Note that the partition functions for each subsystem ( $Q_{n_1 V_1 T}$  and  $Q_{n_1 V_1 T}$ ) refer to the canonical ensemble.

For  $s$  subsystems, with  $V = \sum_{j=1}^s V_j$  and  $N = \sum_{j=1}^s N_j$ , a recursive argument is used to develop the Gibbs partition function. The Gibbs ensemble partition function is first written for the last two simulation boxes with total number of particles  $N_2 = n_{s-1} + n_s$  and total volume  $V_2 = V_{s-1} + V_s$ , which is equivalent to a single  $N_2 V_2 T$  canonical ensemble:

$$\bar{Q}_2 = \sum_{n_{s-1}=0}^{N_2} \int_0^{V_2} Q_{n_{s-1}V_{s-1}T} Q_{n_s V_s T} dV_{s-1} \quad (6.7)$$

A third simulation box may now be added with the total number of particles becoming  $N_3 = n_s + n_{s-1} + n_{s-2}$ , so that the new partition function is:

$$\begin{aligned} \bar{Q}_3 &= \sum_{n_{s-2}=0}^{N_3} \int_0^{V_3} Q_{n_{s-2}V_{s-2}T} \bar{Q}_2 dV_{s-2} \\ &= \sum_{n_{s-2}=0}^{N_3} \sum_{n_{s-1}=0}^{N_2} \int_0^{V_3} dV_{s-2} \int_0^{V_2} Q_{n_{s-2}V_{s-2}T} Q_{n_{s-1}V_{s-1}T} Q_{n_s V_s T} dV_{s-1} \end{aligned} \quad (6.8)$$

In general, for  $s$  phases, the final expression, after substituting the expression for the canonical ensemble (Equation 3.1) is:

$$\begin{aligned} \bar{Q}_s^{\text{NVT-GE}} &= \frac{1}{\Lambda^{3N} V^{s-1} N!} \sum_{n_1=0}^N \sum_{n_2=0}^{N-n_1} \cdots \sum_{n_{s-1}=0}^{N-(n_1+\cdots+n_{s-2})} \binom{N}{n_1} \times \binom{N-n_1}{n_2} \cdots \binom{N-n_1+\cdots+n_{s-2}}{n_{s-1}} \\ &\times \int_0^V dV_1 \int_0^{V-V_1} dV_2 \cdots \int_0^{V-(V_1+\cdots+V_{s-2})} dV_{s-1} \cdots \prod_{i=1}^s V_i^{n_i} \int \exp[-\beta U_i(\mathbf{r}^{n_i})] d\mathbf{r}^{n_i} \end{aligned} \quad (6.10)$$

where  $\mathbf{r}^{n_i}$  refers to all particle coordinates in box  $i$ ,  $U_i(\mathbf{r}^{n_i})$  is the total potential energy of box  $i$  and  $\binom{N}{n_1} = \frac{N!}{(N-n_1)!n_1!}$  counts the number of ways of choosing  $n_1$  molecules from  $N$  total molecules. Clearly, three phases will be present when, during the course of a simulation, each box has a different density. Obviously, when two boxes have the same density a first-order phase transition will have occurred. From Equation 6.10, the

statistical weight of a particular molecular configuration with  $n_1$ ,  $n_2$  and  $n_3$  molecules in each box  $V_1$ ,  $V_2$  and  $V_3$  will occur with the following probability distribution:

$$\Omega(\mathbf{r}^N, V) \propto \frac{V_1^{n_1} V_2^{n_2} V_3^{n_3}}{n_1! n_2! n_3!} \times \exp\{-\beta[U_1(\mathbf{r}_1) + U_2(\mathbf{r}_2) + U_3(\mathbf{r}_3)]\}. \quad (6.11)$$

Canongia Lopes and Tildesley (1997) showed that because all trial moves performed during a Gibbs ensemble Monte Carlo simulation occur between a pair of simulation boxes or within a single simulation box, the acceptance criteria for these trial moves will remain unchanged however many phases are present.

Thus, the acceptance criteria for producing a Markov chain of configurations with a probability distribution equivalent to Equation 6.11 and for the NPT version are the same as the rules that were originally derived by Panagiotopoulos *et al.* (1988), whether two or more simulation boxes are present (Canongia Lopes and Tildesley, 1998), since any trial move involves, at most, two simulation boxes.

The particle insertion and particle displacement acceptance criteria for the NPT–Gibbs ensemble are identical to the acceptance criteria for its NVT counterpart; however, for mutual exchange of volume between any two phases (here denoted as phases I and II), the acceptance criterion is different. For an increase in the volume of phase I of  $\Delta V_I$ ,  $V_I^{\text{new}} = V_I^{\text{old}} + \Delta V_I$  (Panagiotopoulos *et al.*, 1988):

$$\text{acc}(o \rightarrow n) = \min \left[ 1, \exp \left[ \begin{array}{l} -\beta\Delta U_I - \beta\Delta U_{\text{II}} + N_I \ln \frac{V_I + \Delta V_I}{V_I} \\ + N_{\text{II}} \ln \frac{V_{\text{II}} + \Delta V_{\text{II}}}{V_{\text{II}}} - \beta P(\Delta V_I + \Delta V_{\text{II}}) \end{array} \right] \right]. \quad (6.16)$$

It is important to note that the volume of any simulation box may now also vary independently, since a constant, specific pressure needs to be maintained across all phases

to fulfill the mechanical–equilibrium criterion. Note that for both the NVT and NPT versions of the Gibbs ensemble, that the probability distributions for the binary and ternary three-phase mixtures are not required. Rather, the acceptance rules form the backbone of the simulations.

## 7. Standard Molecular Simulation Techniques

This focus of this chapter is on the standard methods used in the algorithms that are implemented in the simulation of phase equilibrium. Since the theory that has been presented thus far has focused on monatomic systems, the techniques that are used in the simulation of polyatomic molecules will be explained here as well.

### 7.1 Random Number Generation

A random number generator is a computational device designed to generate a sequence of numbers that lack any pattern. RANLUX is an advanced pseudo-random number generator that is implemented by the Towhee Monte Carlo program, and was proposed by Luescher (1994) and is based on the RCARRY algorithm. The latter used a subtract-and-borrow algorithm with a period on the order of  $10^{171}$  (more than adequate for molecular simulations, considering the number of Monte Carlo cycles typically used) but still had detectable correlations between numbers. RANLUX generates pseudo-random numbers using RCARRY but throws away numbers to destroy correlations. In doing so, there is a slight decrease in the speed of generation of random numbers but a higher quality is obtained.

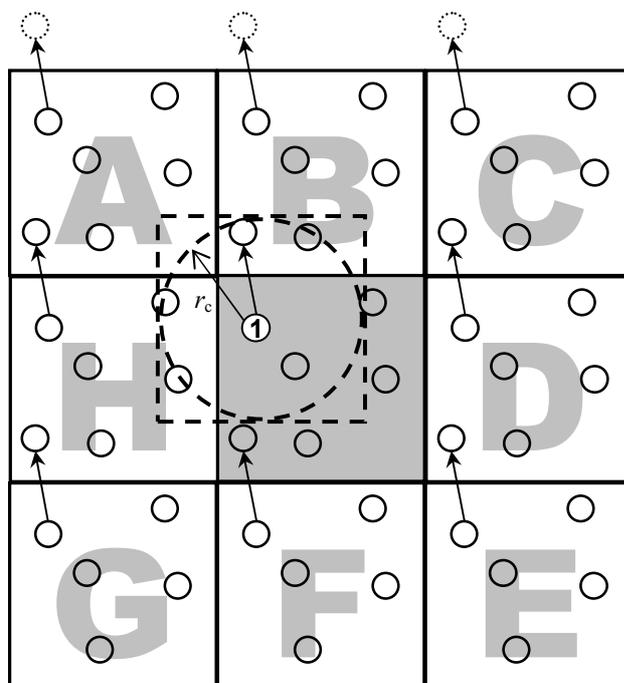
### 7.2 Periodic Boundary Conditions and the Minimum Image Convention

Typically, a molecular simulation uses less than 10000 molecules (Allen and Tildesley, 1987) and the time required for a simulation is generally dependent on the square of the number of molecules, since, for a system of  $N$  molecules, there will be  $\frac{1}{2}N(N-1)$  intermolecular interactions<sup>11</sup>. For a three-dimensional simulation box with free boundaries the number of molecules at the surface of this box is proportional to  $N^{\frac{1}{3}}$ . Thus, a

---

<sup>11</sup> This is true if all interactions are truncated at two-body interactions.

significant number of molecules reside at the simulation box surface. These molecules will experience different forces compared with molecules in the bulk phase, making it is necessary to choose boundaries which mimic the behaviour of an infinite bulk fluid.



**Figure 7-1 Illustration of the method of periodic boundary conditions and minimum image convention for a 2-D system (Allen and Tildesley, 1987).**

Surface effects are usually overcome by implementing periodic boundary conditions. A two-dimensional periodic system is shown in Figure 7-1. The centre cell is called the primitive cell, and an exact copy of it is tiled infinitely in 2-D space. When a molecule moves through a boundary in the primitive cell, its corresponding images in the remaining cells move across their corresponding boundaries (McKnight, 2005). Therefore, only the coordinates of molecules in the central box need to be stored in computer memory. Assuming pair-wise additivity of molecular interactions, the expression for the total potential energy of  $N$  molecules for a three-dimensional system is written as (Frenkel and Smit, 2002):

$$U^{\text{total}} = \frac{1}{2} \sum'_{i,j,\mathbf{n}} u(\mathbf{r}_{ij} + \mathbf{n}L), \quad (7.1)$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ ,  $\mathbf{n} = (n_x, n_y, n_z)$  is an arbitrary three-dimensional vector,  $L$  is the length of the cubic periodic box and the prime on the  $\mathbf{n}$ -sum indicates that when  $\mathbf{n} = 0$ , the  $i = j$ -terms are not to be counted. This form of calculating the potential energy is not very useful as it is not a finite sum. To overcome this problem the molecular interactions are split in two types: long- and short-range interactions. Above a certain cut-off radius, denoted  $r_c$ , all intermolecular interactions are truncated. Methods describing the truncation of short-range interactions are discussed in the next section.

The minimum image convention requires that an interacting molecule must not interact with molecules outside of a box that has the same size as the simulation box centred on the interacting molecule of interest. A molecule should not interact with a periodic image of itself, that is to say, molecule 1 in the primitive cell in Figure 7-1 should not interact with its corresponding molecules ( $1_A \rightarrow 1_H$ ), and neither should it interact with two periodic images of the same molecule as this introduces an artificial periodicity to the simulation (McKnight, 2005).

### 7.3 Analytical Tail Corrections

Before discussing the method of analytical tail corrections, it is instructive to introduce the radial distribution function. The radial distribution function describes the radial packing in an atomic or molecular system and provides a suitable means to correctly estimate the contribution to the potential energy from the long-ranged interactions when  $\int dr_{ij} r_{ij}^{-2} U(r_{ij}) < 1$  (Allen and Tildesley, 1987), as well as to calculate the ensemble average of an observable. The general form of the RDF is:

$$g(r) = \frac{n(r)}{4\pi\rho r^2\Delta r} = \frac{\rho(r)}{\rho}, \quad (7.2)$$

where  $n(r)$  is the mean number of atoms with a shell of width  $\Delta r$  (at a distance  $r$ ) and  $\rho$  is the bulk system density. An ensemble average in terms of the RDF is written as:

$$\langle A \rangle = \frac{1}{2} N \rho \int_0^{\infty} A(r) g(r) 4\pi r^2 dr. \quad (7.3)$$

Thus, using the Lennard–Jones potential, one may obtain the contribution to the system's total potential energy beyond a cutoff radius  $r_c$  as follows:

$$U_{\text{LJ}}^{\text{tail}} = \frac{1}{2} 4\pi\rho \int_{r_c}^{\infty} 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] r^2 dr \quad (7.4)$$

$$\therefore U_{\text{LJ}}^{\text{tail}} = \frac{8}{3} \pi \rho \epsilon \sigma^3 \left[ \frac{1}{3} \left( \frac{\sigma}{r_c} \right)^9 - \left( \frac{\sigma}{r_c} \right)^3 \right]. \quad (7.5)$$

The RDF is also used in thermodynamic modeling in order to take into account the density of molecules around one molecule.

#### 7.4 Hard–inner cutoff radius

A hard–inner cutoff radius is used to ensure that any attempt to either insert an atomic group at a point in space that is already occupied, or to move a molecule too close to such a point, will be rejected. Such insertions or moves would automatically be rejected upon calculation of the new potential energy of the molecular configuration; thus the hard–inner cutoff radius assists in the reduction of simulation times. The radius is centered at the middle of an atomic group, and ranges typically from 0Å to 2Å.

## 7.5 Long-Range Interactions – The Ewald summation for point charges

Compared with the nature of the van der Waals interaction energy for which analytical tail corrections are added after truncation, the Coulombic interaction cannot be treated in a similar manner. Consider the tail contribution of the potential  $U^*(r)$  in three dimensions:

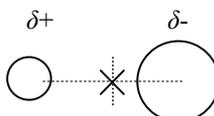
$$U_{\text{tail}} = \frac{N\rho}{2} \int_{r_c}^{\infty} dr U^*(r) 4\pi r^2, \quad (7.6)$$

which, for the Coulomb potential becomes:

$$U_{\text{tail}}^{\text{Coul}} = \frac{N\rho}{2} \int_{r_c}^{\infty} \frac{1}{r} 4\pi r^2 dr = \infty. \quad (7.7)$$

Clearly, tail corrections cannot be used, for Equation (7.7) confirms that Coulomb forces are long-ranged. Several techniques exist for the calculation of long-range interactions, of which the Ewald summation method is the most widely used (Frenkel and Smith, 2002). A rigorous discussion is presented by de Leeuw *et al.* (1980*a*, 1980*b* and 1983).

The force fields used in this work model the electrostatic interactions based on point charges. These charges are not necessarily placed at the center of a pseudo – atomic group (for example, a methyl —(CH<sub>3</sub>)— group), since for highly polar species, the central point charge of the entire pseudoatomic group will be located off-center (see Figure 7-2).



**Figure 7-2 - Schematic of a polar molecule. For the two atoms shown, which constitute a pseudoatomic group, the point charge is located somewhere between the two atoms at point 'X'.**

Consider an electrically neutral ( $\sum_i q_i = 0$ ) system of  $N$  charged molecules in a cubic box of volume  $V (= L^3)$ . It is assumed that periodic boundary conditions are applicable to this system. The total contribution of the Coulomb potential to the total potential energy of the system is:

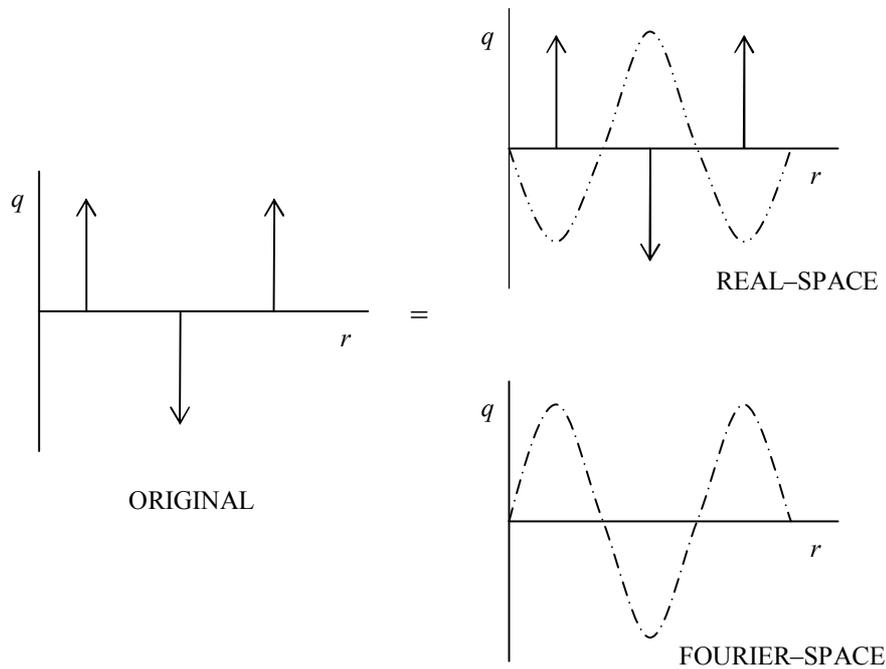
$$U^{\text{Coul}} = \frac{1}{8\pi\epsilon_0} \sum_{\mathbf{n}=0}^{\infty} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|}. \quad (7.8)$$

The prime on the summation indicates that  $i = j$ -terms are not to be counted when  $\mathbf{n} = [0, 0, 0]$ . As it stands, Equation 7.8 cannot be used since it is conditionally convergent. The Ewald summation method splits this equation into a real-space part and a Fourier-space part. Every point charge  $q_i$  is assumed to be surrounded by a diffuse charge distribution of opposite sign but of an equal magnitude. A Gaussian distribution (Allen and Tildesley, 1987) is used as a screening distribution for each charge:

$$\rho^{\text{Gauss}}(r_d) = q_i \kappa^3 \exp\left(\frac{-\kappa^2 r_d^2}{\pi^{3/2}}\right), \quad (7.9)$$

where  $r_d$  is the distance to the point charge  $q_i$  and  $\kappa$  is a parameter that determines the width of the Gaussian distribution.

The contribution to the electrostatic potential due to screened charges can be computed by direct summation; however, the contribution of interest is that due to point charges, specifically point  $i$ . Thus, a correction is needed. Figure 7-3 is a schematic of the screened point charges:



**Figure 7-3 - Ewald summation splitting of a charge.**

The compensating charges vary smoothly in real space, and are exactly cancelled by their Fourier-spaced counterpart. There are now three contributions to the electrostatic potential:

1. The potential due to point charge  $q_i$ ,
2. The potential due to Gaussian screening charge  $-q_i$ , and
3. The compensation charge with charge  $q_i$  (Fourier – space component)

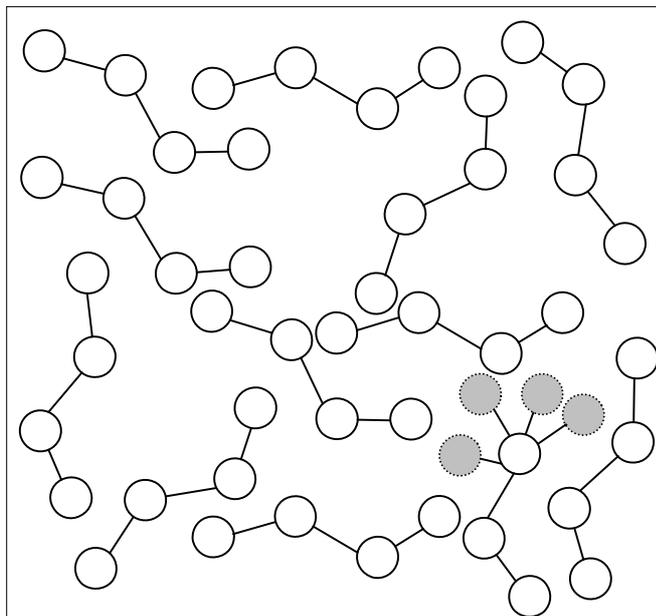
Frenkel and Smit (2002) provide a detailed derivation of all the terms described above, and a correction term to exclude Coulomb self-interactions. The final expression for the total Coulombic potential contribution is:

$$U^{\text{Coul}} = \frac{1}{2V} \sum_{\mathbf{k} \neq 0}^{\infty} \frac{4\pi}{k^2} |\rho(\mathbf{k})|^2 \exp(-k^2/4\alpha) - (\alpha/\pi)^2 \sum_{i=1}^N q_i^2 + \frac{1}{2} \sum_{i \neq j}^N \frac{q_i q_j \operatorname{erfc}(\sqrt{\alpha} r_{ij})}{r_{ij}}, \quad (7.10)$$

where  $\text{erfc}$  is the complementary error function and  $k = 2\pi\mathbf{n}/L$ . Clearly, the larger the  $\kappa$  term is made the faster the convergence of the real-space summation term. However, this results in a sharper distribution that requires more Fourier-space terms, which in turn necessitates larger computation time.

## 7.6 Configurational Bias Monte Carlo Methods

The theory presented so far has focused on simple monatomic molecules. The acceptance rules for monatomic molecules are equally valid for complex polyatomic molecules, but there will be a large number of rejections before a large molecule's insertion or displacement moves are accepted. To overcome this computationally expensive scenario, the configurational-bias Monte Carlo method was introduced by Siepmann (1990), and was derived from a lattice-based method developed by Rosenbluth and Rosenbluth (1955). Essentially, a molecule is grown atom-by-atom into those areas of a dense fluid that have a lower energy position, and this 'bias' is then corrected for afterwards effectively leading to a large increase in acceptance rates for polyatomic molecule insertions. This method also prevents overlapping with other molecules in space (Siepmann *et al.*, 1991), (de Pablo *et al.*, 1992), (Frenkel and Smith, 1992).



**Figure 7-4 - The configurational-bias Monte Carlo method. To complete the growth of the molecule, the next atomic group is placed at the most energy-favorable position in space.**

### 7.6.1 Standard CBMC<sup>12</sup>

In general, the calculation of the non-bonded interactions for a collection of molecules is time consuming. It was shown that the potential energy may be split into an intramolecular component and an intermolecular component.

Suppose that a molecule is to be grown atom-by-atom (the word ‘atom’ here refers to a pseudoatomic group) in a dense liquid. The Rosenbluth weight of the first segment is:

$$w_1(n) = \sum_{j=1}^{j=f} \exp[-\beta U_{1j}^{\text{ext}}], \quad (7.11)$$

where  $f$  is the total number of trial insertions for the first atom at random positions in the simulation box and  $U_{1j}^{\text{ext}}$  is the non-bonded potential due to the insertion of this first atom,

<sup>12</sup> This sub-section is based mostly on Siepmann & Frenkel (1992).

for the  $j^{\text{th}}$  attempt. Thus, using the definition of a probability, a particular position where the atom will be placed is selected using:

$$P_{li}^{\text{select}}(\mathbf{b}_i) = \frac{\exp[-\beta U_{1j}^{\text{ext}}]}{w_1(n)}, \quad (7.12)$$

where  $\mathbf{b}_i$  is the position of the center of mass of the atomic group. For the  $l$  remaining groups,  $k$  trial orientations are generated according to the Boltzmann weight of the internal potential of that atomic group:

$$P_{li}^{\text{gen}}(\mathbf{b}_i) d\mathbf{b} = \frac{\exp[-\beta U_{li}^{\text{int}}] d\mathbf{b}}{\int \exp[-\beta U_{li}^{\text{int}}] d\mathbf{b}}. \quad (7.13)$$

The trial orientation that is eventually selected is:

$$P_{li}^{\text{gen}}(\mathbf{b}_i) = \frac{\exp[-\beta U_{li}^{\text{ext}}]}{\sum_{j=1}^{j=k} \exp[-\beta U_{li}^{\text{ext}}]}, \quad (7.14)$$

according to the Boltzmann weight of its external potential. The procedure outlined above is continued until the entire molecule has been grown. Thus, the Rosenbluth weight of the **new** configuration is written as:

$$W(n) = \frac{w_1(n) \prod_{l=1}^{N_{\text{seg}}} \left[ \sum_{j=1}^{j=k} \exp[-\beta U_{li}^{\text{ext}}] \right]}{f \times k^{N-1}}, \quad (7.15)$$

where  $N_{\text{seg}}$  is the number of atomic groups/segments in the molecule. The Rosenbluth weight of the old configuration,  $W(o)$  is similarly defined, but the  $k^{\text{th}}$  or  $f^{\text{th}}$  trial orientation is the old configuration:

$$W(o) = \frac{w_1(o) \prod_{l=1}^{N_{\text{seg}}} \left[ \sum_{j=1}^{j=k} \exp[-\beta U_{li}^{\text{ext}}] \right]}{f \times k^{N-1}}. \quad (7.16)$$

From the original acceptance criterion for a particle displacement, namely,

$$\text{acc}(o \rightarrow n) = \min(1, \exp\{-\beta[U(n) - U(o)]\}),$$

the probability of accepting the configurational-bias regrowth is:

$$\text{acc}(o \rightarrow n) = \min\left(1, \frac{W(n)}{W(o)}\right). \quad (7.17)$$

## 7.6.2 Dual-cutoff CBMC

Vlugt *et al.* (1998) introduced an improved CBMC algorithm called dual-cutoff configurational-bias Monte Carlo (DC-CBMC). These algorithms sped up the calculations for systems of 144 octane and 3,4-dimethylhexane molecules by factors of 3 and 4 respectively, resulting in a decreased difference in simulation time between a linear molecule and a branched-isomer. It was observed that during a molecule re-growth, the CBMC selection process is hard-core in nature, that is to say, the selection of a trial orientation is strongly affected by molecules in close proximity. Therefore, one could consider only the nearby interactions during the growth phase and correct for the bias afterwards.

Let  $r_{\text{cut}^*}$  be the cutoff radius beyond which potential contributions to the DC-CBMC re-growth are truncated. The external potential may then be split into two components<sup>13</sup>:

---

<sup>13</sup> This is an arbitrary separation; the potential can, in general, be split in any manner as long as the interactions are independent (du Preez, 2005).

$$U^{\text{ext}} = \bar{U}^{\text{ext}}(r < r_{\text{cut}^*}) + \delta U^{\text{ext}}(r_{\text{cut}^*} < r < r_{\text{cut}}), \quad (7.18)$$

where  $\bar{U}^{\text{ext}}$  is the computationally less-expensive potential,  $\delta U^{\text{ext}}$  is the difference between the two potential energies and  $r_{\text{cut}}$  is the cutoff radius for non-bonded interactions, as described in Chapter 4. Since  $\bar{U}^{\text{ext}}$  is a shortened potential, the Rosenbluth weights are calculated faster and consequently, molecule growths occur faster. However, this leads to a bias that must be corrected. The acceptance rule is derived as such:

$$\frac{\text{acc}(o \rightarrow n)}{\text{acc}(n \rightarrow o)} = \frac{\exp[-\beta U^{\text{ext}}(n)] \exp[-\beta U^{\text{ext}}(o)]}{\exp[-\beta U^{\text{ext}}(o)]} \frac{\bar{W}(n)}{\bar{W}(o) \exp[-\beta U^{\text{ext}}(o)]}, \quad (7.19)$$

$$= \frac{\bar{W}(n)}{\bar{W}(o)} \exp[-\beta \{\delta U^{\text{ext}}(n) - \delta U^{\text{ext}}(o)\}] \quad (7.20)$$

where  $\bar{W}(o)$  and  $\bar{W}(n)$  are the Rosenbluth weights that are calculated using  $\bar{U}^{\text{ext}}$ . Thus, the DC-CBMC acceptance criterion is:

$$\text{acc}(o \rightarrow n) = \min\left(1, \frac{\bar{W}(n)}{\bar{W}(o)} \exp[-\beta \{\delta U^{\text{ext}}(n) - \delta U^{\text{ext}}(o)\}]\right). \quad (7.21)$$

### 7.6.3 Coupled-Decoupled CBMC

In 1999 Vlught *et al.* revealed a flaw in the Boltzmann rejection technique if a molecule contains any atom that is bonded to three or more other atoms. To this end, the coupled-decoupled CBMC algorithm, which is a generalization of the standard CBMC method, was developed by Martin and Siepmann (1999) which also avoid problems for sequentially generating dihedral and bond bending angles. At the time, bond lengths were still rigid, but the algorithm later included decoupled flexible bond lengths (Martin and Thompson, 2004) for force fields such as NERD (Nath *et al.*, 1998). It is currently implemented in the MCCCIS Towhee code, and it is important to remember that this method of particle re-

growth is only effective when the appropriate parameters are chosen carefully, since simulation times can increase without judicious setting of the parameters. This is due to more calculations being performed per trial compared with standard CBMC, because dihedral and bond bending angles are also generated (du Preez, 2005).

For the biased regrowth of a molecule, Martin and Siepmann (1999) initially proposed the inclusion of bond bending and dihedral energies with the Lennard–Jones energy. However, it was proven to be computationally inefficient since it takes a large number of trials to find reasonable bond angles. Similar inefficiencies were reported for molecular conformations when dealing with dihedral energies, since once the dihedral angles were chosen in a biased approach, only one possible trial site could be used for the Lennard–Jones selection. Eventually, the approach used was to couple the biased selections so that each biased selection sent several possible conformations to the subsequent selection step. Thus, the probability of generating a conformation is:

$$P^{\text{gen}} = \prod_{n=1}^{n_{\text{step}}} \left[ \frac{\exp(-\beta U^{\text{LJ}}(i)) W_{\text{T}}(i)}{W_{\text{L}}(n)} \right] \left[ \frac{\exp(-\beta U^{\text{tors}}(j)) W_{\text{B}}(j)}{W_{\text{T}}(i)} \right] \times \dots \times \left[ \frac{\exp(-\beta U^{\text{bend}}(k))}{W_{\text{B}}(j)} \right], \quad (7.22)$$

where

$$W_{\text{L}}(n) = \sum_{i=1}^{n_{\text{chLJ}}} \exp(-\beta U^{\text{LJ}}(i)) W_{\text{T}}(i), \quad (7.23)$$

$$W_{\text{T}}(i) = \sum_{j=1}^{n_{\text{chTors}}} \exp(-\beta U^{\text{tors}}(j)) W_{\text{B}}(j), \text{ and} \quad (7.24)$$

$$W_{\text{B}}(j) = \sum_{k=1}^{n_{\text{chbend}}} \exp(-\beta U^{\text{bend}}(k)).$$

(7.25)

In Equations 7.13 to 7.25,  $n_{\text{chtor}}$ ,  $n_{\text{chLJ}}$  and  $n_{\text{chbend}}$  are the number of trial sites for the torsional, Lennard-Jones and bond-bending interactions, respectively, and  $W_L(n)$ ,  $W_T(i)$  and  $W_B(j)$  are the corresponding Rosenbluth weights. The move is then accepted with the following probability:

$$P^{\text{acc}} = \min \left[ 1, \frac{\prod_{n=1}^{n_{\text{step}}} W_L(n)_{\text{new}}}{\prod_{n=1}^{n_{\text{step}}} W_L(n)_{\text{old}}} \right]. \quad (7.26)$$

One disadvantage of coupling the different energy types is that  $n_{\text{chLJ}} \times n_{\text{chtor}} \times n_{\text{chbend}}$  trial vectors need to be generated for the bond angle bias selection, compared with  $(f + k)$  vectors for the standard CBMC method (Martin and Siepmann, 1999).

#### 7.6.4 Arbitrary Trial Distributions

Yet another formulation of the original CBMC method, this technique uses arbitrary distributions, different from the standard Boltzmann distribution, to generate trial bond lengths, angles and dihedral angles (Martin and Frischknecht, 2006). Arbitrary trial distributions have been shown to provide acceptance rates similar to the standard CBMC method, but are computationally less expensive and have subsequently been implemented in the coupled–decoupled algorithm.

The concept of arbitrary trial distribution CBMC is based upon the energy bias schemes developed by Snurr *et al.* (1993) in adsorption studies. Martin and Frischknecht (2006) showed that that scheme is applicable to any trial generation and selection steps used in a CBMC move. A particular trial position ( $\mathbf{r}_i$ ) is generated according to a positive arbitrary

trial distribution ( $p_{\text{trial}}^{\text{arb}}(\mathbf{r}_i)$ ) that is normalized over the range of trial positions. The selection probability for one of the generated trials being used for further growth during step  $k$  is:

$$P^{\text{select}}(\mathbf{r}_i) = \frac{[p_{\text{trial}}^{\text{ideal}}(\mathbf{r}_i)/p_{\text{trial}}^{\text{arb}}(\mathbf{r}_i)]\exp[-\beta U(\mathbf{r}_i)]}{W_k}, \quad (7.27)$$

where  $p_{\text{trial}}^{\text{ideal}}(\mathbf{r}_i)$  is the distribution of trials in an ideal system where no intermolecular forces are present,  $U(\mathbf{r}_i)$  is the potential energy of the trial position and  $W_k$  is a modified Rosenbluth weight which contains the ratio of trial distribution probabilities:

$$W_k = \sum_{j=1}^{n_{\text{trial}}^k} \frac{p_{\text{trial}}^{\text{ideal}}(\mathbf{r}_i)}{p_{\text{trial}}^{\text{arb}}(\mathbf{r}_i)} \exp[-\beta U(\mathbf{r}_i)], \quad (7.28)$$

where  $n_{\text{trial}}^k$  is the number of trials generated during step  $k$  of the CBMC growth. A CBMC growth consisting of  $n_{\text{step}}$  growth steps obeys the following acceptance probability:

$$P^{\text{accept}} = \min \left[ 1, \frac{\prod_{k=1}^{n_{\text{step}}} W_k^{\text{new}}}{\prod_{k=1}^{n_{\text{step}}} W_k^{\text{old}}} \right]. \quad (7.29)$$

Arbitrary trial distribution CBMC uses a truncated Gaussian trial probability distribution:

$$p_{\text{trial}}^{\text{Gauss}} = \frac{1}{C^{\text{Gauss}} \sigma^{\text{Gauss}} (2\pi)^{1/2}} \exp \left[ -\frac{1}{2} \left( \frac{m - \mathbf{x}^{\text{Gauss}}}{\sigma^{\text{Gauss}}} \right)^2 \right] \quad (7.30)$$

on the interval  $(m_{\min}, m_{\max})$ .  $x^{\text{Gauss}}$  is the mean,  $\sigma^{\text{Gauss}}$  is the standard deviation and  $C^{\text{Gauss}}$  is an integration constant defined as:

$$C^{\text{Gauss}} = 0.5[\text{erf}(m_{\max}) - \text{erf}(m_{\min})], \quad (7.31)$$

all referring to either a bond length, bond angle or dihedral angle. This distribution is now used in the generation of trial bond lengths, bond angles and dihedral angles.

## 8. Computational Aspects

Molecular simulations are computationally expensive and thus require intensive computing capability. To this end, simulations were performed using the MCCC (Monte Carlo for Complex Chemical Systems) Towhee code (Martin, 2007), which is written in the FORTRAN language and compiled with the C programming language, on Beowulf cluster Ruby (University College of Borås). In general, a computer cluster refers to a group of computers that work quite intimately to the extent that the entire group may be viewed as a single entity. A Beowulf cluster is a type of high-performance computing cluster that has become popular in scientific computing and was originally developed by Thomas Sterling and Donald Becker at NASA. One of the main attractions of a Beowulf cluster is that it is significantly cheaper than conventional supercomputers, since a Beowulf cluster uses free and open source software (e.g. Linux) and off-the-shelf (commodity) hardware components. Figure 8-1 illustrates the basic set-up of a Beowulf computer cluster.

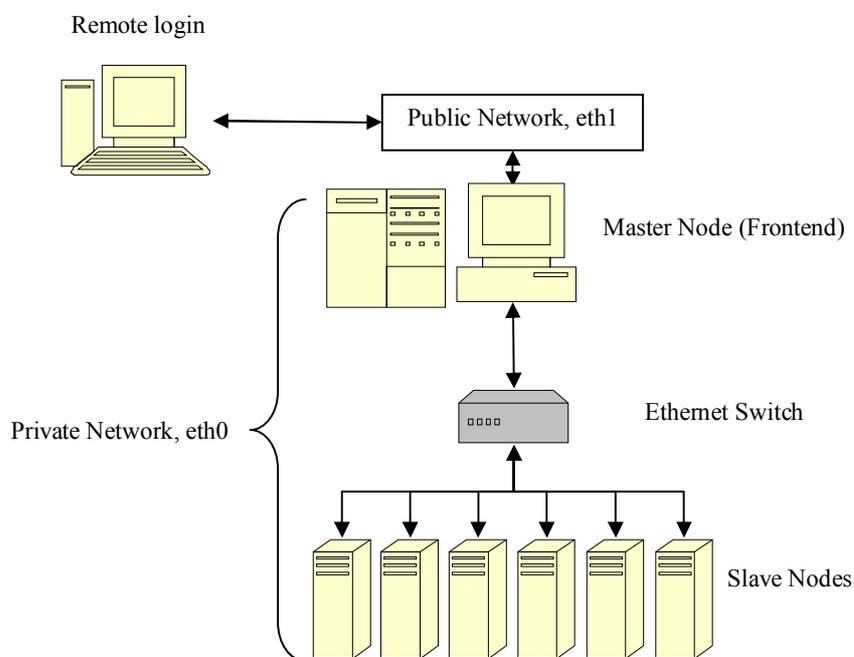


Figure 8-1 – Schematic of a typical Beowulf cluster setup

## **8.1 Hardware**

### **8.1.1 Nodes**

In general computer network terminology, a node refers to a device that is connected as part of a computer network; in a Beowulf cluster, these devices are computers and the various network subcomponents (to be discussed later in this chapter). Two types of computer nodes exist in a Beowulf cluster – these are the master node and the slave nodes. The master node's key function is to direct different computational tasks to the slave nodes, and then retrieve the results from the slave nodes and direct it to the end-user. It must be borne in mind that the master node may also function as an additional computational node, but this tends to decrease the calculation speed of the entire cluster. Thus, the sole purpose of the master node in this work is to oversee the correct distribution of computational tasks to the slave nodes.

The sole purpose of a slave node is to perform the calculation tasks that have been assigned to it, and then to pass the results back to the master node. It is here that one of the cost-saving aspects of a Beowulf cluster is most evident; since the slave nodes are dedicated computational devices, it is not necessary for these nodes to have their own display devices except during the software installation process.

### **8.1.2 Network Subcomponents**

The network subcomponents along with the correct software and protocols (or means of communication), allow for the exchange of data between each slave node and the master node. These subcomponents constitute the private network, and include the Ethernet switch, network interface cards (NIC) on each node and the links (standard moulded copper cabling). Each network interface card has a unique MAC address, which allows a particular node to be identified on the private network. Ruby utilises a standard 100Megbit Ethernet connection and consists of one master and six slave nodes.

	Master Node	Slave Nodes
CPU	AMD Opteron 156 single core with 1MiB L2 Cache	AMD Opteron 146 single core with 1MiB L2 Cache
RAM	1GiB DIMM RAM	512MiB PC 133 SDRAM
Hard Disk	80GiB SATA HDD	40GiB Hard Drive
CD-ROM	12x DVD-ROM Drive	-
Graphics Card	32MiB Onboard SiS Chipset	8MiB Graphics Card
Network Card	Onboard LAN connection	Onboard LAN connection

**Table 8-1 – Master and Slave Node Hardware Specifications.**

Network Switch	3Com Superstack 3 3250 48-PORT
Keyboard and Mouse	One set each for the master node and for slave node administration
Monitor	One for the master node and one for slave node administration

**Table 8-2 – Network and Display Hardware**

## 8.2 Software

For any Beowulf cluster to be fully functional, the following software is required:

1. Operating system
2. Parallel processing software (if parallel computation is enabled)
3. Resource management software

### 8.2.1 Operating system

The operating system of choice for most Beowulf clusters is the Unix-based Linux operating system. Linux is an open-source operating system that comes in different variants or distributions, each suited to a particular application. The Linux source code for the kernel is free to modify by developers thus providing for a continual improvement of the software. Furthermore, some Linux distributions are available free on the Internet and this increases the cost:performance ratio for Beowulf clusters.

### 8.2.2 Parallel processing software

The purpose of parallel processing is to split a computational task into multiple processes or threads, and execute each thread on its own processor and in doing so, obtain results faster. It is instructive to note that not every computational algorithm may be parallelized, since the different portions of a program must be independent to be executed separately on its own processor (although some serial algorithms may be redesigned to run in parallel). This may be summed up as:

“One woman can have a baby in nine months, but nine women can't have a baby in one month.” (Brooks, 2005)

Parallel computing is achieved through communication amongst the nodes of the cluster setup, using a Message Passing application programmer Interface<sup>14</sup> (MPI), designed to provide access to parallel hardware. Two popular MPI implementations are LamMPI<sup>15</sup> and MPICH<sup>16</sup>.

### 8.2.3 Resource Management Software

The processing power that is afforded by a Beowulf cluster is a resource that needs to be efficiently managed and controlled. Several users may have access to the cluster and will require computational power to perform calculations; thus, a queuing system is used to store user jobs until processing power is available. Scheduling is used to balance job priorities and to have as many jobs running at any given time; it is also important that each job is distributed to the correct node. A user should also be able to terminate or suspend a job if required (McKnight, 2005). To this end, the Sun Grid Engine (SGE) which is distributed under the open source license is used on Ruby as a resource management tool

---

<sup>14</sup> <http://www.mpi-forum.org>

<sup>15</sup> [www.osc.edu/lam.html](http://www.osc.edu/lam.html)

<sup>16</sup> [www.mcs.anl.gov/mpi/mpich](http://www.mcs.anl.gov/mpi/mpich)

that requires a user to only know how to submit jobs and get results. SGE is resource management software that:

1. Accepts jobs submitted by users
2. Schedules them for execution on appropriate systems based on resource management policies
3. Can submit 100s of jobs without worrying where or when it will run ('Inside SGE', <http://www.rocksclusters.org/>)

Two versions of SGE exist; they are Sun Grid Engine, which is distributed under the open source license (SISSL license) and Sun N1 Grid Engine.

### **8.3 Rocks Cluster Suite**

The software components that have been discussed in the preceding sections can be installed independently; however, for organizations that do not have access to a full-time dedicated cluster administrator, the installation process can be difficult; moreover, maintaining the cluster presents an even greater problem. High-performance computer clusters are now the computing tool of choice for a myriad of scientific disciplines (Sacerdoti *et al.*, 2003), so one can imagine the lack of adequately skilled cluster experts.

The Rocks cluster suite (<http://www.rocksclusters.org/>), comprises all of the necessary software to install and maintain a cluster and supports a variety of commodity architectures, including the ubiquitous x86 architectures that are implemented by the Intel Pentium 4 and AMD Athlon chipsets. Rocks Version 3.3.0 (Fuji distribution) is used on Ruby. Rocks v3.3.0 is built on the Linux distribution based on the Red Hat Enterprise line, and allows users to add additional software modules called 'rolls' that plug into the base program. Thus, aside from the operating system, all other software utilities that were mentioned in the preceding sections would be classified as rolls. Rocks provides a good solution for computer clusters in that the same software configuration is installed and maintained across the entire network on all computers.

In order to install a barebones cluster, the Rocks Base CD, the HPC Roll CD and the Kernel Roll CD are required. Additionally, for Ruby, the Ganglia Cluster Toolkit, SGE, Intel and Java rolls were installed.

The XML-based cluster monitoring tool named Ganglia functions as the ‘nervous’ system in the cluster that alerts the user to components that might be damaged. It also enables the user to monitor loads on each slave node CPU, thus providing a valuable tool in tuning the performance of parallel programs (Sacerdoti *et al.*, 2003).

Figure 8-2 is a snapshot of the Ganglia cluster monitoring tool. This snapshot presents to the user an overall picture of the ‘health’ of the cluster, with information such as the average workload on each CPU and the current status of each CPU (i.e. whether it idle, active or not working). Additionally, detailed information on each node may be viewed.

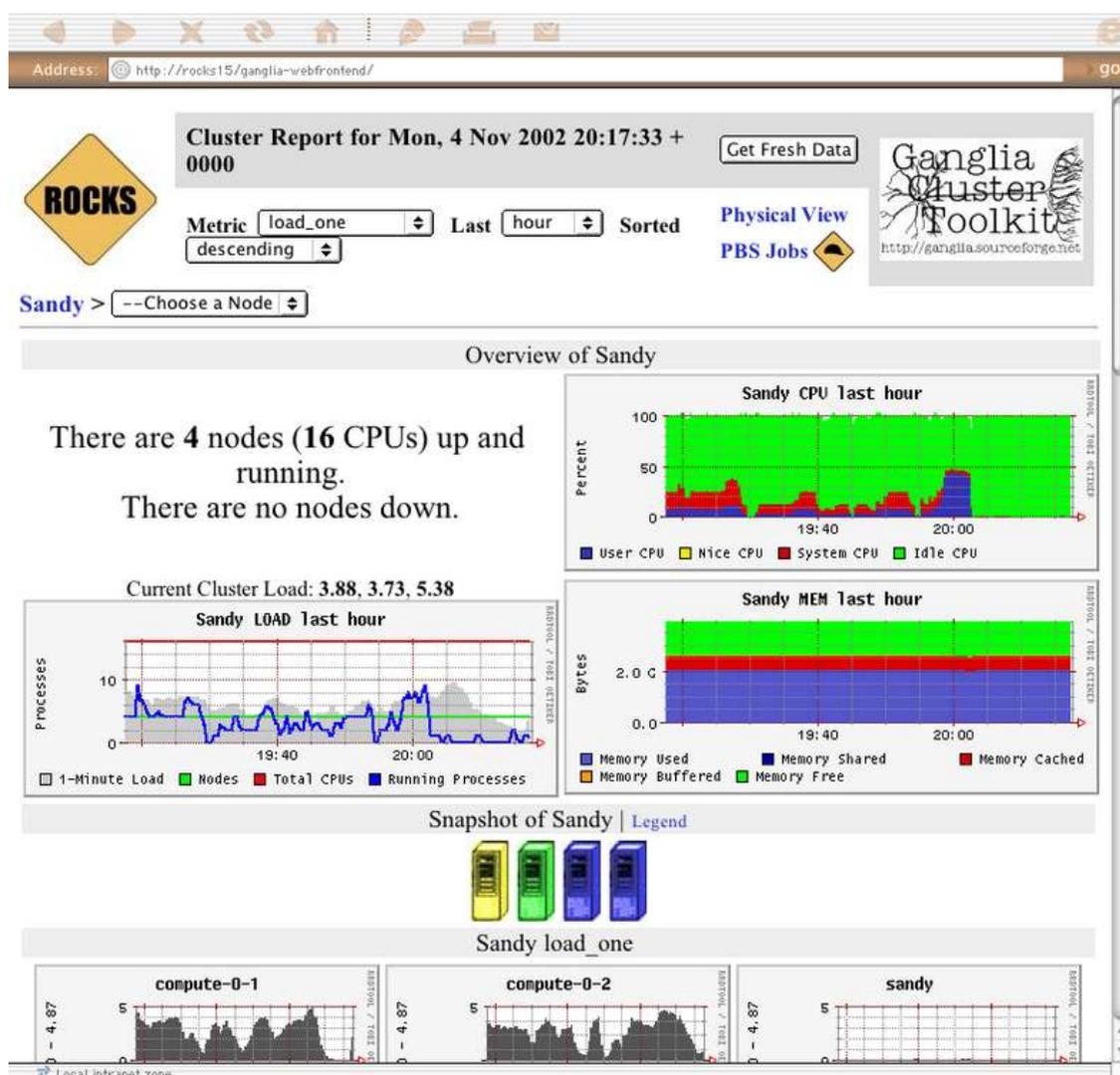
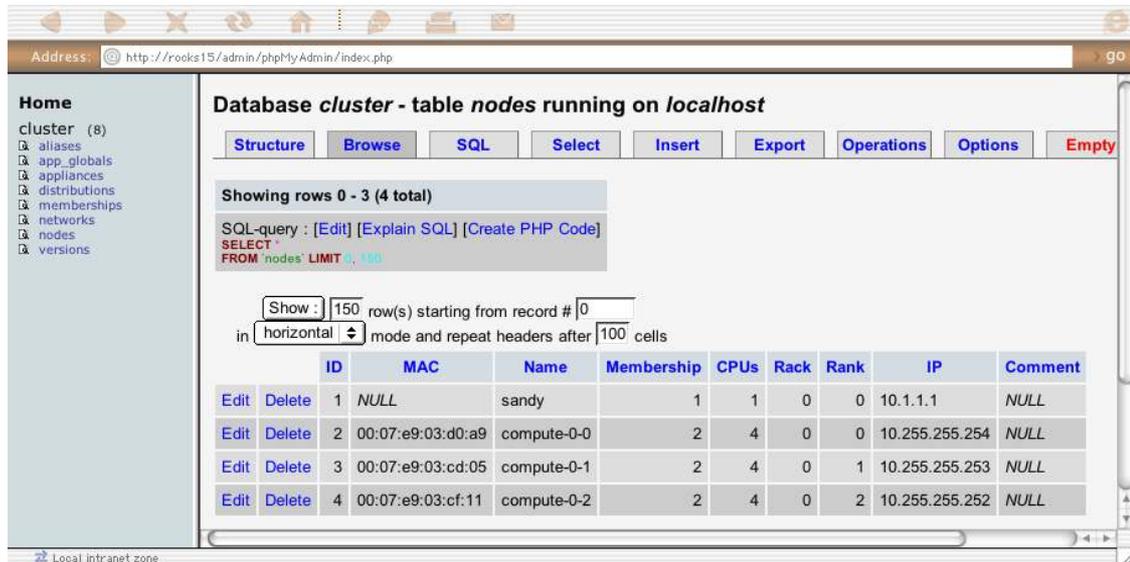


Figure 8-2 – Screenshot of the Ganglia cluster monitoring tool.

Ganglia was designed at Berkeley by Matt Massie (University of California) in 2000, and is currently developed by an open source partnership between Berkeley, the San Diego Supercomputer Centre and others. It is distributed through Sourceforge.net (<http://sourceforge.net>) under the GPL software license (<http://www.gnu.org/copyleft/gpl.html>).

Rocks also has a web database application that lists all active computers that are installed in the cluster. Based on SQL, the database allows Queries, Inserts, Updates, and Deletes; any changes made within the web application immediately take effect on services that rely on

the database. Thus, only hosts on the internal network are permitted to access to this page (see Figure 8-3).



The screenshot shows a web browser window displaying the Rocks SQL database interface. The address bar shows `http://rooks15/admin/phpMyAdmin/index.php`. The main content area is titled "Database cluster - table nodes running on localhost". It features a navigation menu with options: Structure, Browse, SQL, Select, Insert, Export, Operations, Options, and Empty. Below the menu, it indicates "Showing rows 0 - 3 (4 total)" and displays the SQL query: `SELECT * FROM nodes LIMIT 0, (4)`. The interface also shows settings for "Show: 150 row(s) starting from record # 0" and "in horizontal mode and repeat headers after 100 cells". A table with 10 columns (ID, MAC, Name, Membership, CPUs, Rack, Rank, IP, Comment) and 4 rows is displayed. Each row has "Edit" and "Delete" links.

	ID	MAC	Name	Membership	CPUs	Rack	Rank	IP	Comment
Edit Delete	1	NULL	sandy	1	1	0	0	10.1.1.1	NULL
Edit Delete	2	00:07:e9:03:d0:a9	compute-0-0	2	4	0	0	10.255.255.254	NULL
Edit Delete	3	00:07:e9:03:cd:05	compute-0-1	2	4	0	1	10.255.255.253	NULL
Edit Delete	4	00:07:e9:03:cf:11	compute-0-2	2	4	0	2	10.255.255.252	NULL

Figure 8-3 – Snapshot of the Rocks SQL database.

### 8.3.1 Installation

Prior to commencing the installation, mandatory checks were made on the cluster hardware setup. The power cables to each computer's power supply were checked, along with the LAN cables to each computer's network interface card to ensure that they were correctly plugged-in.

Furthermore, since all slave (or in Rocks terminology, 'compute') nodes did not have CD-ROM drives, installing the software on these nodes required the installation to be performed via a network boot, using the software images on the CD-ROM in the master (or 'frontend') node. To this end, the default boot option on each node was set to PXE boot (Preboot EXecution Environment). Naturally, once the installation of the software was successful on each node, the default boot option was reverted to the hard drive disk.

#### 8.3.1.1 Frontend Node Installation

The frontend installation was initiated by first selecting the CD-ROM as the default boot device. The base CD, along with all the necessary rolls was inserted at the onscreen prompts. The user only truly interacts with the computer during the installation process when prompted to enter the hard disk drive partitioning options (default values were used) and to enter the cluster information at the 'Cluster Information Screen'. During this part of the process, the default values for the public network (eth0) interface were used, whilst the values for the public network (eth1) variables, i.e. the IP Address, Netmask, Gateway and Primary DNS were changed.

The IP address is a unique address that is assigned to the computer's network card's MAC address (which in turn is a hexadecimal number unique to every network interface card that is manufactured). This is important for the master node as it allows the user to access the cluster from a remote location that has internet connectivity, but is not part of the cluster network. The IP addresses for the slave nodes however, are automatically assigned via DHCP within the cluster setup. Finally, a root password was chosen for the root user who has authority over the entire cluster (e.g. to create and delete user accounts and to edit system files).

### **8.3.1.2 Slave Node Installation**

Once the frontend installation was successful, the compute nodes were installed. All compute nodes were initially switched off, after changing the default boot option to PXE boot). The # `insert-ethers` command was executed – this invoked a program that captured compute node DHCP requests and added their information to the Rocks MySQL database (<http://www.rocksclusters.org>). During this process, the monitor was connected to each node in succession; each node was switched on and detected by the frontend node and the clustering software was installed on all nodes. This can be a highly efficient process if there are no hardware or boot configuration problems, since multiple nodes may have Rocks software installed simultaneously. On average, this process lasts for approximately fifteen minutes per node.

## 9. Simulation Details and Results

Two binary mixtures and three ternary mixtures were investigated; these were: n-hexane/water, ethane/ethanol, methane/water/n-heptane, n-butane/1-butene/water and water/ethanol/n-hexane. All organic molecules were modeled by the TraPPE-UA force field, while the SPC-E potential model was used to represent water; the justification for their use was given in Chapter 2. However, a very brief discussion on the use of the SPC-E force field is also given in Chapter 10. Literature data for the articulated mixtures is listed in Appendix B. This section presents the simulation settings used in each Monte Carlo simulation along with reasons as to why they were chosen.

### 9.1 Simulation Details

The MCCCSTowhee code was used in all simulation work in this study, utilizing the coupled-decoupled CBMC algorithm. Estimates for the initial simulation box volumes for liquid phases were obtained by using a temperature-dependent density correlation (Perry and Green, 2007), while vapour phase box volumes were estimated using the ideal gas equation of state. It must be stressed that for NPT simulations, the initial estimates for the box volumes are important only from a computational time perspective, since this ensemble spontaneously finds the equilibrium values of the box volumes by constantly varying their sizes independently, for each phase. However, for NVT simulations, more accurate methods are required. Fortunately, due to the nature of this study being a qualitative one, experimental densities, where available, were used as a starting point in constant volume simulations. Each simulation was performed in duplicate using different integer seeds to initialize the random number generator. Molecules were initially arranged on a cubic lattice. For all simulations, block averages were calculated every  $N_{MC}/5$  cycles, where  $N_{MC}$  is the total number of Monte Carlo cycles used in a simulation. In all simulations, the criteria used in deciding the mixtures to be equilibrated were the densities, compositions and pressures not showing increasing or decreasing trends over the block averages obtained

from the simulations; in other words, the statistical uncertainties (standard deviations) for the final values of the properties of interest were not large. It will be seen in Chapter 10 that lengthy equilibration times were required especially for the ternary simulations involving water and ethanol, due to there being many charged sites present. Unless otherwise stated, the same simulation parameters, as discussed for the n-hexane/water mixture, including move probabilities were used in all simulations.

### **Coupled-decoupled CBMC settings**

The default values suggested by Martin (2008) were used in all simulations. This caused no problems since the coupled-decoupled algorithm was initially designed to address problems regarding the sequential generation of torsion and bond bending angles in branched alkanes (Martin and Siepmann, 1999), while this study dealt with straight chain molecules. For clarity, the main parameters of this molecule re-growth algorithm will be summarized.

Trial bond lengths were generated according to a bounded  $r^2$  probability distribution within a range of 85% to 115% of the equilibrium bond length between two pseudoatoms. Ten trial positions were sampled for the first pseudoatom inserted in a simulation box during CBMC swap moves and ten trial positions each for the remaining pseudoatom(s). Dihedral angle trials were generated uniformly on  $(-\pi, \pi)$ , with 360 trial dihedral angles sampled. Since all molecules used in the simulations were unbranched and short-chained, no further dihedral angles needed to be sampled once each molecule was fully grown (except for n-hexane and n-heptane). To sample bending angles, one thousand trial angles were generated on  $(-\pi, \pi)$  for each pseudoatomic triplet.

To calculate the contribution of electrostatically charged sites to the potential energy of each box, Ewald summations with tin-foil boundary conditions were used with  $\kappa \times L = 5$  and  $k_{\max} = 5$ . The tin-foil boundary condition essentially amounts to a neutralizing system boundary (du Preez, 2005). Values of the parameters used in this study were taken directly from work done by Zhang and Siepmann (2004). These are typical values used in most

Monte Carlo molecular simulations. As discussed in Chapter 7, the Ewald summation technique is used to determine the contribution of the Coulomb potential energy which is a long-ranged force and thus should not be truncated like the Lennard-Jones non-bonded interactions.

## 9.1.1 Binary Simulations

### 9.1.1.1 n-hexane/water

NPT-Gibbs ensemble Monte Carlo simulations were used to study the three-phase coexistence of n-hexane and water at two different state points, viz.,  $P = 4116.41$  kPa with  $T = 482$  K and  $P = 4882.55$  kPa with  $T = 492$  K, using an overall n-hexane composition  $z_{\text{n-hexane}} = 0.614$ . A total of  $N = 600$  molecules was used in each simulation, comprising 369 n-hexane and 231 water molecules. This is a reasonable system size since in a study of the mutual solubilities of long-chain alkanes and water, Johansson (2007) used no less than 200 water and 100 alkane molecules in each simulation. Additionally, the pressures and corresponding temperatures and compositions of interest were selected based on isochoric heat capacity measurements done by Kamilov *et al.* (2001).

These simulations required the use of both TraPPE-UA and SPC-E force fields; for model parameters, refer to Appendix A. The minimum allowable size of a simulation box was set to  $17.0$  Å; that is to say, all Lennard-Jones non-bonded interactions were truncated beyond  $8.5$  Å, in keeping with the minimum image convention. This value was chosen due to the small volume occupied by the water molecules. Even doubling the number of water molecules does not significantly change the box volume and would thus serve to only increase the simulation times without affecting the thermodynamic averages. A hard-inner cutoff radius of  $0.8$  Å was used as well so that any attempts to insert or move a molecule within this spherical region were automatically rejected, thereby improving the simulation efficiency. Care was taken to ensure that the volume of any box did not drop below that of the corresponding minimum box length, thus, the total number of molecules had to be

judiciously selected to ensure that the simulation was simultaneously efficient (in terms of computation time) as well as producing reliable results.

During the pre-equilibration runs, the pure phases were allowed to equilibrate in separate simulation boxes – all water molecules were placed in one box, while 309 n-hexane molecules were placed in a second simulation box and the remaining 60 n-hexane molecules were placed in a large third simulation box. For this part of the simulation, only volume, translation and rotation moves were allowed, along with configurational-bias re-growths with the following probabilities: volume:translation:rotation:configurational-bias re-growth = 0.01:0.33:0.33:0.33. It was discussed in Chapter 4 that the volume move is computationally expensive, thus, this move accounted for only 1% of all moves used in the simulations. The configurational-bias regrowth move effectively amounts to being a conformational move for chain-like molecules. Values for the other move types were chosen based on simulations done by previous workers (Martin and Siepmann, 1998). It was found that equilibrating pure phases in their respective simulation boxes sped up the ensuing equilibration period, especially for mixtures exhibiting a high degree of mutual insolubility, since, when swap moves were finally enabled in the equilibration period, not that many large molecules would be transferred to denser phases.

The pre-equilibration runs consisted of at least  $2.5 \times 10^4$  Monte Carlo cycles, where one Monte Carlo cycle consists of  $N$  moves, where  $N$  is the total number of molecules used in a simulation. Thereafter, swap moves were allowed with equal probabilities between each pair of simulation boxes with the total probability set to 10% and the remaining probabilities were equally distributed among translation, rotation and CBMC re-growths (the volume-change probability remained at 1%).

Equilibration runs consisted of at least  $1.5 \times 10^5$  MC cycles during which the densities, potential energies, pressures and compositions were monitored for convergence. It is worth mentioning that during the pre-equilibration runs, the maximum allowable centre-of-mass translations, rotations and volume displacements were updated every ten cycles for the first two thousand cycles, to roughly yield acceptance rates of 50% for each move type. This preserved symmetry in the underlying Markov chain, that is to say, once equilibrium was

realized in each simulation, the average number of accepted trial moves that resulted in the system leaving an old state were identical to the number of trial moves that would result from all other new states going to the ‘old’ state.

### **9.1.1.2 ethane/ethanol**

Simulations for this mixture were performed in the NPT-Gibbs ensemble for 400 ethane and 100 ethanol molecules at  $P = 5.169$  MPa and  $T = 311.15$  K, as well as two other simulations using 800 ethane and 200 ethanol molecules. These temperature and pressure values were obtained from a high pressure phase equilibrium study by Kato *et al.* (1999). Move probabilities were the same as those used for the n-hexane/water mixture.

## **9.1.2 Ternary Simulations**

### **9.1.2.1 methane/water/n-heptane**

Simulations at  $T = 275$  K for two pressures, 120 kPa and 2000 kPa, were performed in the NPT ensemble for a mixture consisting of 400 methane, 350 water and 250 n-heptane molecules. A pre-equilibration run of  $5 \times 10^3$  cycles was used to obtain a 50% acceptance rate for molecular translations, box volume changes and molecular rotations (except for methane, which is monatomic in the TraPPE-UA potential model). No swaps were allowed during this period; thereafter,  $2.5 \times 10^4$  cycles were used to further equilibrate the pure components before swap moves were allowed for equilibration runs of at least  $1 \times 10^5$  cycles. Move probabilities were the same as those used in the binary simulations. Aside from using a hard inner cut-off radius of 0.7 Å and truncating non-bonded Lennard-Jones beyond 10 Å, all other simulation parameters remained unchanged.

### 9.1.2.2 n-butane/1-butene/water

Simulations were performed at  $T = 310.93$  K and  $P = 404.72$  kPa for a mixture of 300-butane, 450 1-butene and 250 water molecules in the NPT-Gibbs ensemble as well as the NVT-Gibbs ensemble.

### 9.1.2.3 water/ethanol/n-hexane

Four state points at atmospheric pressure at temperatures of 329.45 K, 329.51 K, 329.77 K and 330.54 K were investigated for mixtures of compositions  $\{z_{\text{water}}; z_{\text{ethanol}}; z_{\text{n-hexane}}\} = \{0.2; 0.31; 0.49\}$ ,  $\{0.2; 0.245; 0.555\}$ ,  $\{0.2; 0.145; 0.655\}$  and  $\{0.2; 0.06; 0.74\}$ , respectively, for a total of 750 molecules in each NPT-Gibbs ensemble simulation. Equilibration periods consisted of at least  $1 \times 10^5$  Monte Carlo cycles, followed by production runs of at least  $5 \times 10^4$  cycles.

It shall be clear in the next Chapter that not all simulations in the NPT variant of the Gibbs ensemble were successful in achieving three distinct coexisting phases. To this end, the unsuccessful simulations were then attempted in the NVT-Gibbs ensemble. Compared to its NPT counterpart, the only difference here is that the total volume of the system is now a constant and must be distributed accordingly amongst the three simulation boxes in order to achieve mechanical equilibrium.

The simulation parameters remained exactly the same as those for the constant pressure simulations. However, better estimates of the box volumes were required; this is discussed in the next chapter. The ternary simulations used no less than 750 molecules in each simulation, while the binary simulations used at least 500 molecules; this ensured that converged statistics were obtained during a run.

## 9.2 Numerical results

This section lists the numerical results obtained from all simulation work that was undertaken in this study. The uncertainties (standard deviations) in these values are listed below each thermodynamic average.

### 9.2.1 n-hexane/water

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-hexane	Water		
Vapor	0.76 <sub>1</sub>	0.24 <sub>1</sub>	97 <sub>4</sub>	3663 <sub>80</sub>
n-hexane-rich liquid	0.958 <sub>1</sub>	0.042 <sub>1</sub>	413 <sub>4</sub>	3613 <sub>73</sub>
water-rich liquid	0.00012 <sub>4</sub>	0.99988 <sub>4</sub>	821 <sub>3</sub>	2237 <sub>390</sub>

**Table 9-1 - Final simulation results for n-hexane/water mixture at 482 K in the NVT-Gibbs ensemble. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-hexane	Water		
Vapor	0.934 <sub>19</sub>	0.066 <sub>19</sub>	377 <sub>26</sub>	4531 <sub>367</sub>
n-hexane-rich liquid	0.9589 <sub>18</sub>	0.0411 <sub>18</sub>	424 <sub>3</sub>	4071 <sub>124</sub>
water-rich liquid	0.00012 <sub>6</sub>	0.99988 <sub>6</sub>	826 <sub>6</sub>	2152 <sub>2007</sub>

**Table 9-2 - Simulation results for n-hexane/water mixture at 482 K and 4116.41 kPa using 213 water and 369 hexane molecules. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-hexane	water		
Vapor	0.940 <sub>5</sub>	0.060 <sub>5</sub>	397 <sub>5</sub>	4882 <sub>334</sub>
n-hexane-rich liquid	0.940 <sub>6</sub>	0.060 <sub>6</sub>	394 <sub>11</sub>	4892 <sub>152</sub>
water-rich liquid	0.00012 <sub>2</sub>	0.99988 <sub>2</sub>	815 <sub>7</sub>	3921 <sub>1790</sub>

**Table 9-3 - Simulation results for n-hexane/water mixture at 492 K and 48826.55 kPa using 213 water and 369 hexane molecules. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-hexane	water		
n-hexane-rich liquid	0.937 <sub>65</sub>	0.063 <sub>65</sub>	401 <sub>9</sub>	10651 <sub>17737</sub>
water-rich liquid	0.000071 <sub>19</sub>	0.999929 <sub>19</sub>	838 <sub>3</sub>	50797 <sub>76010</sub>

**Table 9-4 - Preliminary LLE simulation results for 250 water and 200 n-hexane molecules at 3516 kPa and 473.15 K. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-hexane	water		
Vapor	0.449 <sub>21</sub>	0.551 <sub>21</sub>	68.3 <sub>3</sub>	3794 <sub>1325</sub>
n-hexane-rich liquid	0.460 <sub>12</sub>	0.540 <sub>12</sub>	67.1 <sub>3</sub>	2499 <sub>460</sub>
water-rich liquid	0.438 <sub>12</sub>	0.562 <sub>12</sub>	66.8 <sub>1</sub>	2453 <sub>2329</sub>

**Table 9-5 - Example of a simulation where the systems enters a metastable single phase state at 3516 kPa and 473.15 K. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-hexane	water		
Vapor	0.880 <sub>67</sub>	0.120 <sub>67</sub>	295 <sub>106</sub>	3691 <sub>95</sub>
n-hexane-rich liquid	0.957 <sub>4</sub>	0.043 <sub>4</sub>	419 <sub>4</sub>	3671 <sub>57</sub>
water-rich liquid	0.0001 <sub>8</sub>	0.9999 <sub>8</sub>	826 <sub>2</sub>	2142 <sub>826</sub>

**Table 9-6 – Results for n-hexane/water simulation in the NPT ensemble using the ‘shifted’ pressure (3663 kPa) obtained from a successful NVT ensemble simulation. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

## 9.2.2 ethane/ethanol

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.9962 <sub>3</sub>	0.0038 <sub>3</sub>	140 <sub>18</sub>	5600 <sub>190</sub>
ethane-rich liquid	0.945 <sub>5</sub>	0.055 <sub>5</sub>	292 <sub>10</sub>	5756 <sub>213</sub>
ethanol-rich liquid	0.376 <sub>55</sub>	0.624 <sub>55</sub>	625 <sub>24</sub>	6085 <sub>407</sub>

**Table 9-7 – Results for 3-box NVT simulation at 311.15 K using 400 ethane and 100 ethanol molecules. Box identity swaps occurred periodically in this simulation, hence the large uncertainties in the densities. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.9903 <sub>52</sub>	0.0097 <sub>52</sub>	168 <sub>2</sub>	5721 <sub>165</sub>
ethane-rich liquid	0.963 <sub>15</sub>	0.037 <sub>15</sub>	273 <sub>2</sub>	5696 <sub>149</sub>
ethanol-rich liquid	0.609 <sub>54</sub>	0.391 <sub>54</sub>	513 <sub>19</sub>	5815 <sub>70</sub>

**Table 9-8 - Results for 3-box NVT simulation at 311.15 K using 800 ethane and 200 ethanol molecules. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.9972 <sub>2</sub>	0.0028 <sub>2</sub>	108 <sub>5</sub>	4665 <sub>81</sub>
ethane-rich liquid	0.928 <sub>21</sub>	0.072 <sub>21</sub>	349 <sub>13</sub>	4664 <sub>101</sub>
ethanol-rich liquid	0.518 <sub>53</sub>	0.482 <sub>53</sub>	553 <sub>24</sub>	4584 <sub>180</sub>

**Table 9-9 – Results for 3-box NVT simulation at 311.15 K using 800 ethane and 200 ethanol molecules and the NERD force field. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.938 <sub>48</sub>	0.062 <sub>48</sub>	62.8 <sub>9</sub>	5122 <sub>32</sub>
ethane-rich liquid	0.940 <sub>4</sub>	0.060 <sub>4</sub>	61.3 <sub>7</sub>	5264 <sub>71</sub>
ethanol-rich liquid	0.163 <sub>36</sub>	0.837 <sub>36</sub>	714 <sub>18</sub>	7984 <sub>1880</sub>

**Table 9-10 – Results for a 3-box NPT simulation using 1000 molecules, which reverted to two phases. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.9952 <sub>8</sub>	0.0048 <sub>8</sub>	103.2 <sub>1</sub>	5181 <sub>49</sub>
Liquid	0.346 <sub>89</sub>	0.654 <sub>89</sub>	577 <sub>1</sub>	4887 <sub>397</sub>

**Table 9-11 – VLE results for ethane/ethanol using 500 molecules at 311.15K in the NVT ensemble. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.995 <sub>1</sub>	0.005 <sub>1</sub>	70.94 <sub>89</sub>	4229 <sub>46</sub>
Liquid	0.286 <sub>37</sub>	0.714 <sub>37</sub>	658 <sub>18</sub>	4533 <sub>498</sub>

**Table 9-12 - VLE results for ethane/ethanol using 500 molecules at 311.15K in the NVT ensemble. Same as Table 11, except that a slightly larger total volume was used in the simulation. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.99261 <sub>97</sub>	0.00739 <sub>97</sub>	48.12 <sub>44</sub>	3230 <sub>30</sub>
Liquid	0.183 <sub>19</sub>	0.817 <sub>19</sub>	701.4 <sub>9.7</sub>	3928 <sub>489</sub>

**Table 9-13 – VLE results for ethane/ethanol using 500 molecules at 311.15 K in the NVT ensemble. Same as Table 12, except that an even larger total volume was used in the simulation. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Liquid	0.9955 <sub>20</sub>	0.0045 <sub>20</sub>	265 <sub>36</sub>	6225 <sub>224</sub>
Liquid	0.601 <sub>55</sub>	0.399 <sub>55</sub>	523 <sub>25</sub>	6841 <sub>340</sub>

**Table 9-14 – LLE results for ethane/ethanol using 500 molecules at 311.15 K in the NVT ensemble. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions		Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	ethane	ethanol		
Vapor	0.969 <sub>6</sub>	0.031 <sub>6</sub>	368.3 <sub>7.7</sub>	10399 <sub>581</sub>
Liquid	0.603 <sub>73</sub>	0.397 <sub>73</sub>	535 <sub>34</sub>	10680 <sub>611</sub>

**Table 9-15 - LLE results for ethane/ethanol using 500 molecules at 311.15 K in the NVT ensemble. Same as Table 14, except that a much smaller total volume was used in the simulation. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

### 9.2.3 methane/water/n-heptane

Phase	Mole fractions			Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	methane	n-heptane	water		
Vapor	0.9839 <sub>32</sub>	0.0154 <sub>5</sub>	0.0006 <sub>6</sub>	0.911 <sub>16</sub>	119.97 <sub>19</sub>
Hydrocarbon-rich liquid	0.008336 <sub>701</sub>	0.991663 <sub>701</sub>	0.000001 <sub>1</sub>	702.5 <sub>1</sub>	92.8 <sub>16.7</sub>
water-rich liquid	0.00002 <sub>1</sub>	0	0.99998 <sub>1</sub>	1008 <sub>63</sub>	-781.1 <sub>226</sub>

**Table 9-16 – Results for 3-box NPT simulation at 275.5 K for methane/n-heptane/water at 120 kPa. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions			Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	methane	n-heptane	water		
Vapor	0.99822 <sub>57</sub>	0.00172 <sub>58</sub>	0.000059 <sub>75</sub>	14.78 <sub>59</sub>	1994 <sub>16</sub>
Hydrocarbon-rich liquid	0.125 <sub>2</sub>	0.8746 <sub>25</sub>	0.000002 <sub>2</sub>	685.2 <sub>9</sub>	5169 <sub>2474</sub>
water-rich liquid	0.00007 <sub>4</sub>	0	0.99993 <sub>4</sub>	1016 <sub>23</sub>	13345 <sub>11081</sub>

**Table 9-17 - Results for 3-box NPT simulation at 275.5 K for methane/n-heptane/water at 2000 kPa. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

### 9.2.4 n-butane/1-butene/water

Phase	Mole fractions			Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-butane	1-butene	water		
Vapor	0.372 <sub>5</sub>	0.624 <sub>5</sub>	0.0040 <sub>8</sub>	13.56 <sub>59</sub>	548 <sub>22</sub>
Hydrocarbon-rich liquid	0.419 <sub>4</sub>	0.581 <sub>4</sub>	0.00004 <sub>1</sub>	563.2 <sub>9</sub>	790 <sub>133</sub>
water-rich liquid	0.0000051 <sub>102</sub>	0.0000019 <sub>38</sub>	0.99999 <sub>1</sub>	993 <sub>7</sub>	533.3 <sub>2055</sub>

**Table 9-18 – Results for a successful 3-box NVT ensemble simulation for n-butane/1-butene/water at 310.93 K. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions			Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-butane	1-butene	water		
Vapor	0.397 <sub>3</sub>	0.5984 <sub>38</sub>	0.0044 <sub>7</sub>	9.57 <sub>7</sub>	419 <sub>5</sub>
Hydrocarbon-rich liquid	0.402 <sub>8</sub>	0.587 <sub>16</sub>	0.0037 <sub>18</sub>	9.53 <sub>38</sub>	265 <sub>310</sub>
water-rich liquid	0.00001 <sub>1</sub>	0.00002 <sub>3</sub>	0.99997 <sub>3</sub>	0.991 <sub>6</sub>	-5413 <sub>16356</sub>

**Table 9-19 – Results for NPT ensemble simulation for n-butane/1-butene/water at 310.93 K and 404.72 kPa. This simulation reverted to two phases. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

Phase	Mole fractions			Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	n-butane	1-butene	water		
Vapor	0.597 <sub>13</sub>	0.390 <sub>7</sub>	0.013 <sub>6</sub>	2.23 <sub>3,05</sub>	177 <sub>130</sub>
Hydrocarbon-rich liquid	0.5799 <sub>399</sub>	0.4077 <sub>324</sub>	0.0124 <sub>74</sub>	2.24 <sub>7,21</sub>	211 <sub>801</sub>
water-rich liquid	0.00008 <sub>15</sub>	0.00004 <sub>7</sub>	0.99988 <sub>22</sub>	983 <sub>9</sub>	-3928 <sub>15599</sub>

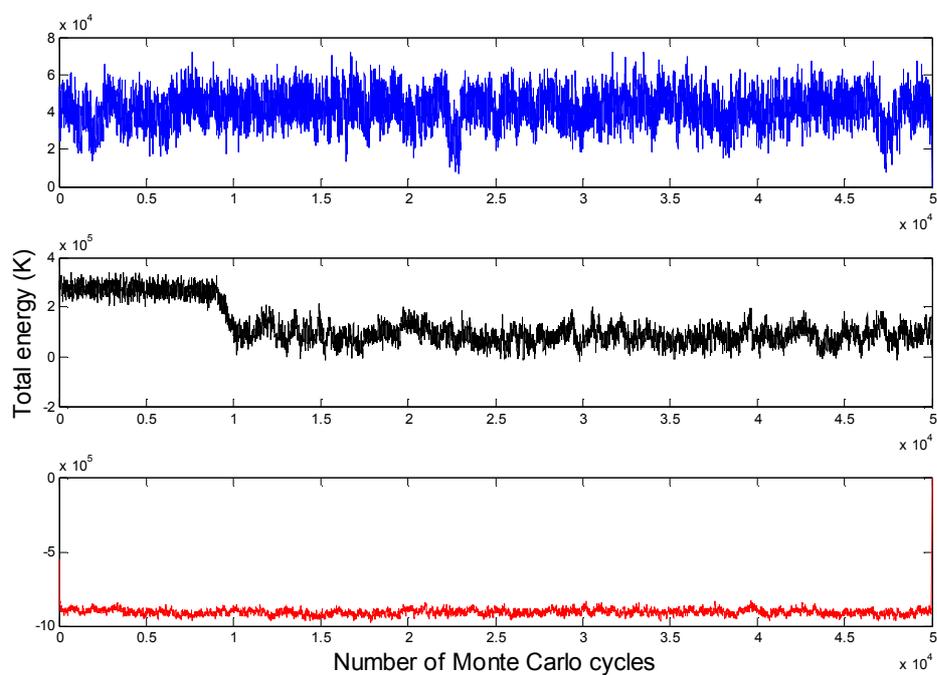
**Table 9-20 - Results for NPT ensemble simulation for n-butane/1-butene/water at 310.93 K and 101.33 kPa. This simulation reverted to two phases. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

## 9.2.5 water/ethanol/n-hexane

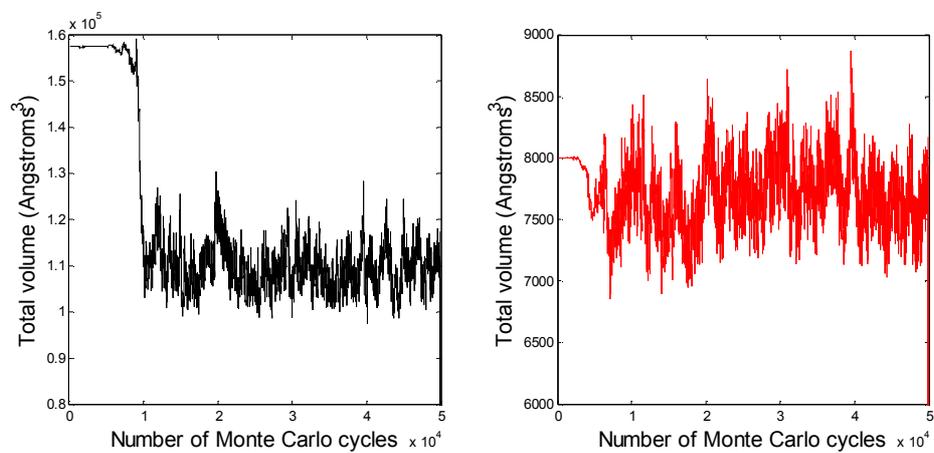
Phase	Mole fractions			Specific Density [kg.m <sup>-3</sup> ]	Virial Pressure [kPa]
	water	ethanol	n-hexane		
T = 330.540 K					
Vapor	0.0637 <sub>39</sub>	0.0664 <sub>12</sub>	0.8699 <sub>28</sub>	3.0152 <sub>56</sub>	101.66 <sub>7</sub>
Hydrocarbon-rich liquid	0.00021 <sub>6</sub>	0.00305 <sub>25</sub>	0.99674 <sub>3</sub>	621 <sub>6</sub>	9655 <sub>2457</sub>
water-rich liquid	0.9642 <sub>59</sub>	0.0358 <sub>58</sub>	4 <sub>9</sub> × 10 <sup>-7</sup>	965 <sub>6</sub>	-21832 <sub>14326</sub>
T = 329.770 K					
Vapor	0.0535 <sub>33</sub>	0.1696 <sub>35</sub>	0.7769 <sub>38</sub>	3.0152 <sub>563</sub>	101.68 <sub>5</sub>
Hydrocarbon-rich liquid	0.0005 <sub>4</sub>	0.0131 <sub>40</sub>	0.9864 <sub>44</sub>	621 <sub>6</sub>	6456 <sub>1383</sub>
water-rich liquid	0.950 <sub>16</sub>	0.0496 <sub>16</sub>	0 <sub>0</sub>	965 <sub>6</sub>	-1323 <sub>2633</sub>
T = 329.510 K					
Vapor	0.0470 <sub>18</sub>	0.2283 <sub>47</sub>	0.7247 <sub>36</sub>	2.825 <sub>16</sub>	101.72 <sub>16</sub>
Hydrocarbon-rich liquid	0.2089 <sub>108</sub>	0.6953 <sub>156</sub>	0.0958 <sub>220</sub>	730 <sub>9</sub>	-1149 <sub>1345</sub>
water-rich liquid	0.9517 <sub>164</sub>	0.0483 <sub>164</sub>	0 <sub>0</sub>	961 <sub>8</sub>	895 <sub>1790</sub>
T = 329.450 K					
Vapor	0.0570 <sub>44</sub>	0.2310 <sub>59</sub>	0.7120 <sub>27</sub>	2.80 <sub>1</sub>	101.66 <sub>9</sub>
Hydrocarbon-rich liquid	0.1332 <sub>207</sub>	0.7084 <sub>242</sub>	0.1584 <sub>437</sub>	714 <sub>8</sub>	-869 <sub>725</sub>
water-rich liquid	0.9129 <sub>132</sub>	0.0871 <sub>132</sub>	0 <sub>0</sub>	938 <sub>8</sub>	-45460 <sub>24036</sub>

**Table 9-21 – NPT simulation results for water/ethanol/n-hexane at 101.33 kPa at four different mixture compositions. The statistical uncertainty in the final digit(s) of each value are shown as subscripts.**

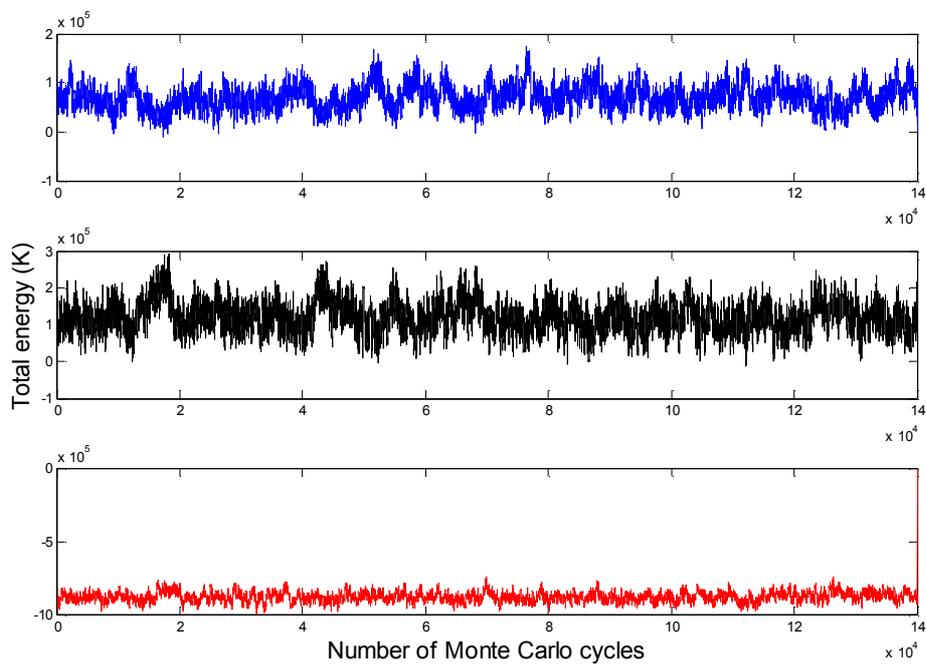
### 9.3 Graphical Results



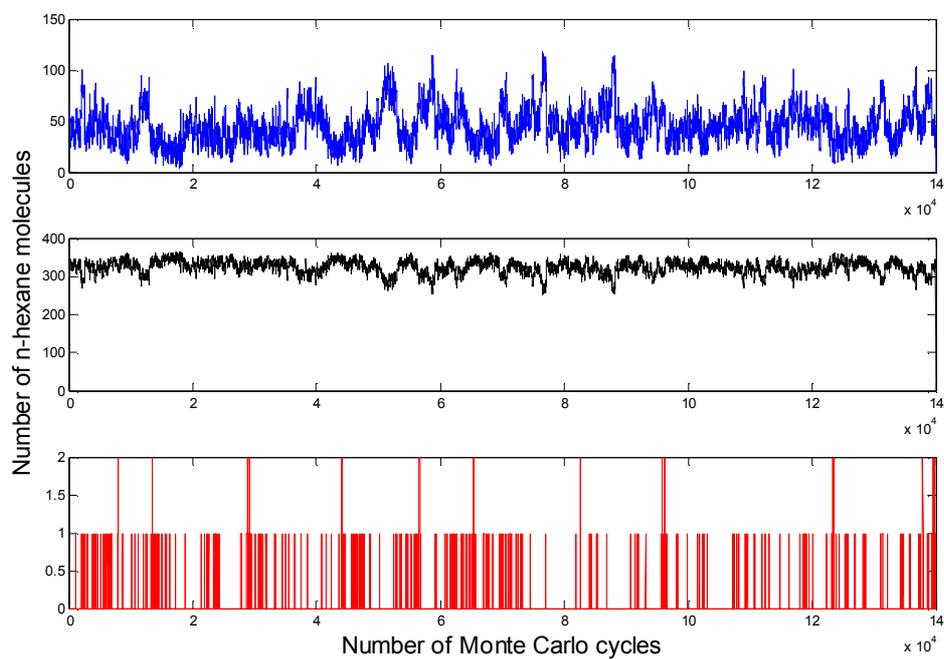
**Figure 9-1 - Plot of the total potential energy in each simulation box during the pre-equilibration period for n-hexane/water in the NVT ensemble. The vapour phase is shown as the upper-most graph, n-hexane phase as the middle graph and water phase as the bottom graph.**



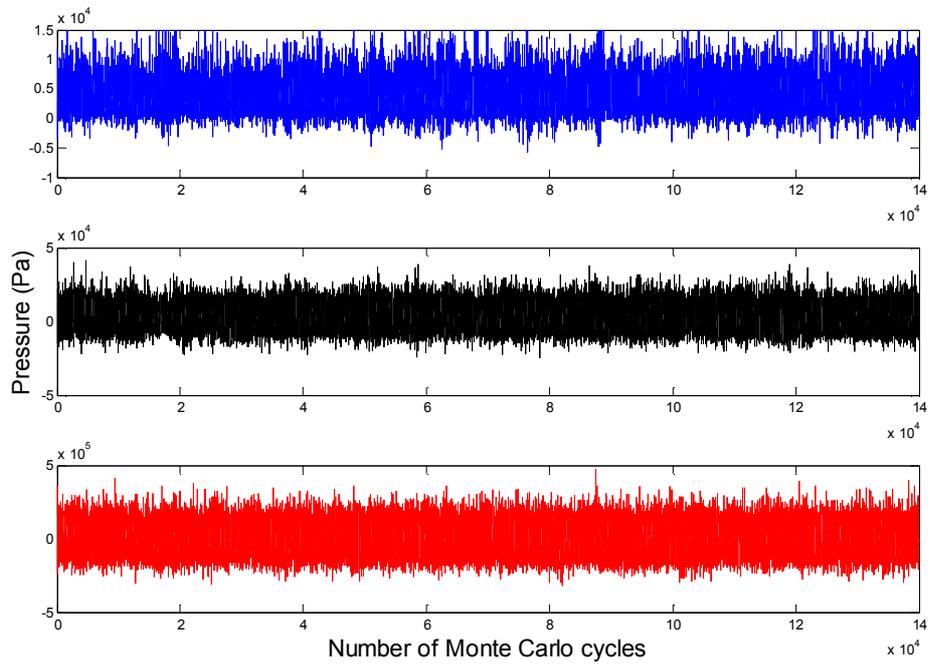
**Figure 9-2 - Plot of the variation in the liquid-phase box volumes versus the number of Monte Carlo cycles for n-hexane/water. The n-hexane box is shown on the left, while the water box is shown on the right. The variation has been shown here for the first 50 000 cycles.**



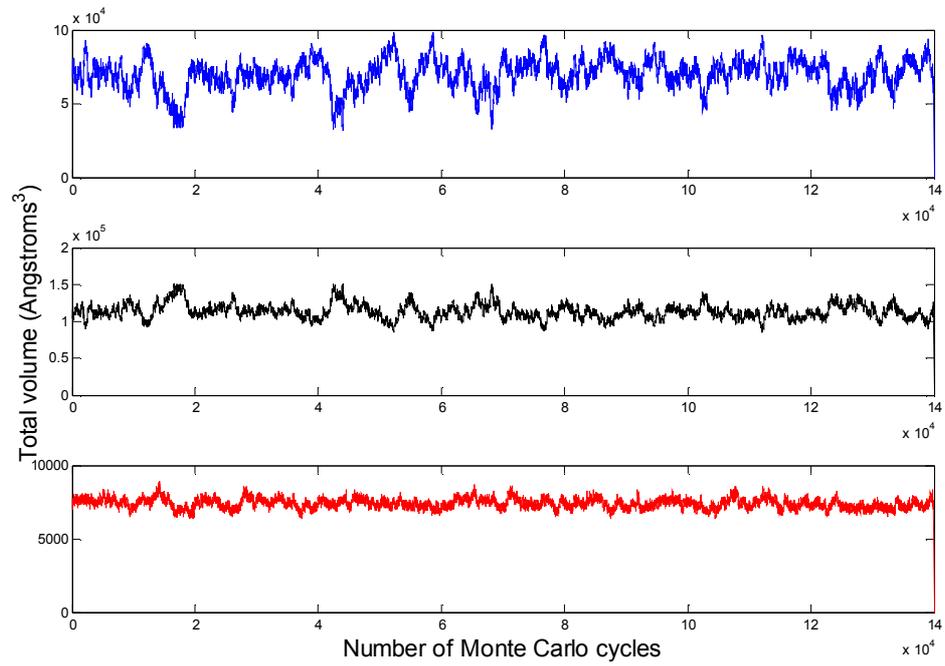
**Figure 9-3 - Plot of the total potential energy in each simulation box during a production run for n-hexane/water in the NVT ensemble. Vapour: top graph; n-hexane phase: middle; water phase: bottom.**



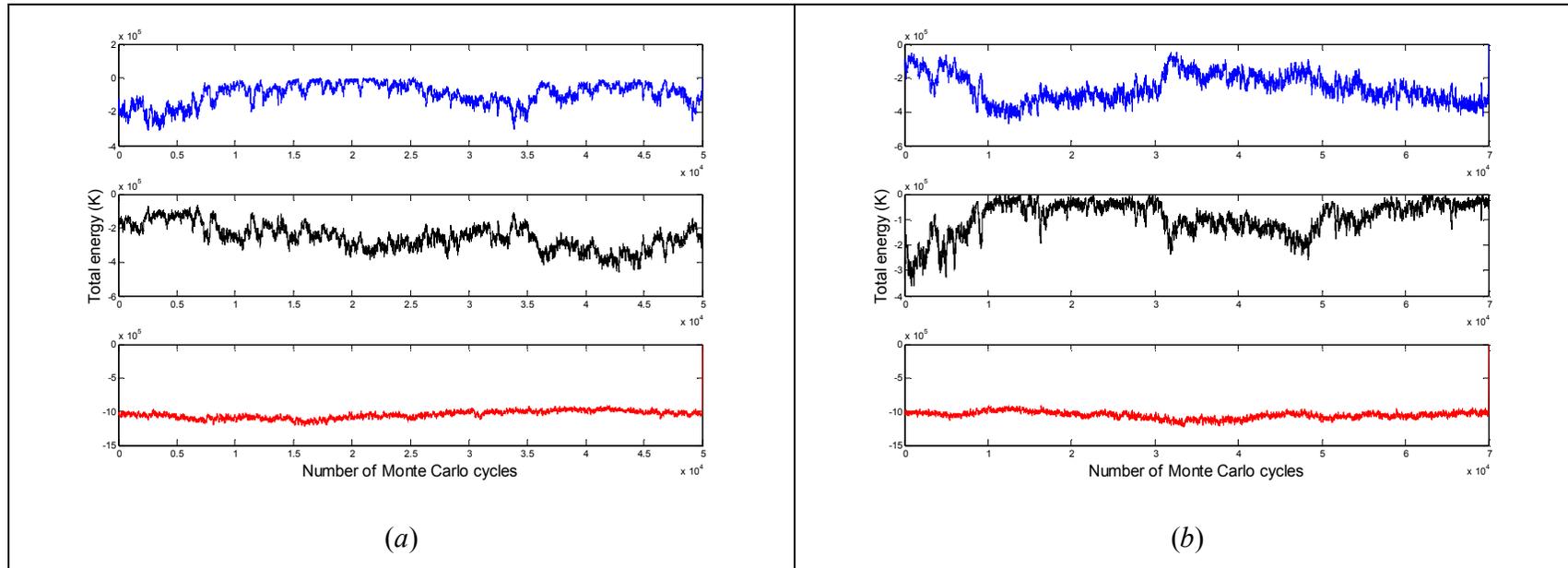
**Figure 9-4 - Plot of the variation of the number of n-hexane molecules in each phase during a production run in the NVT ensemble for n-hexane/water. Vapour: top graph; n-hexane phase: middle; water phase: bottom.**



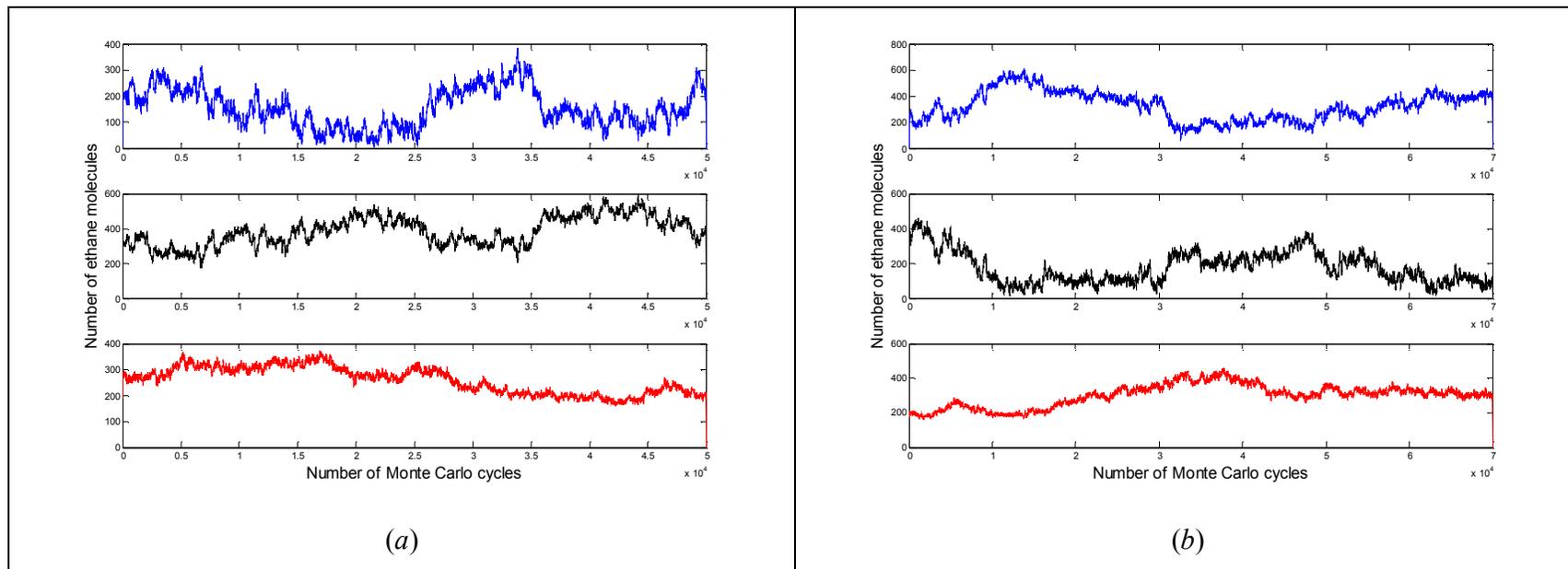
**Figure 9-5 - Plot of the pressure within each simulation box versus the number of Monte Carlo cycles during a production run. Vapour: top graph; n-hexane phase: middle; water phase: bottom.**



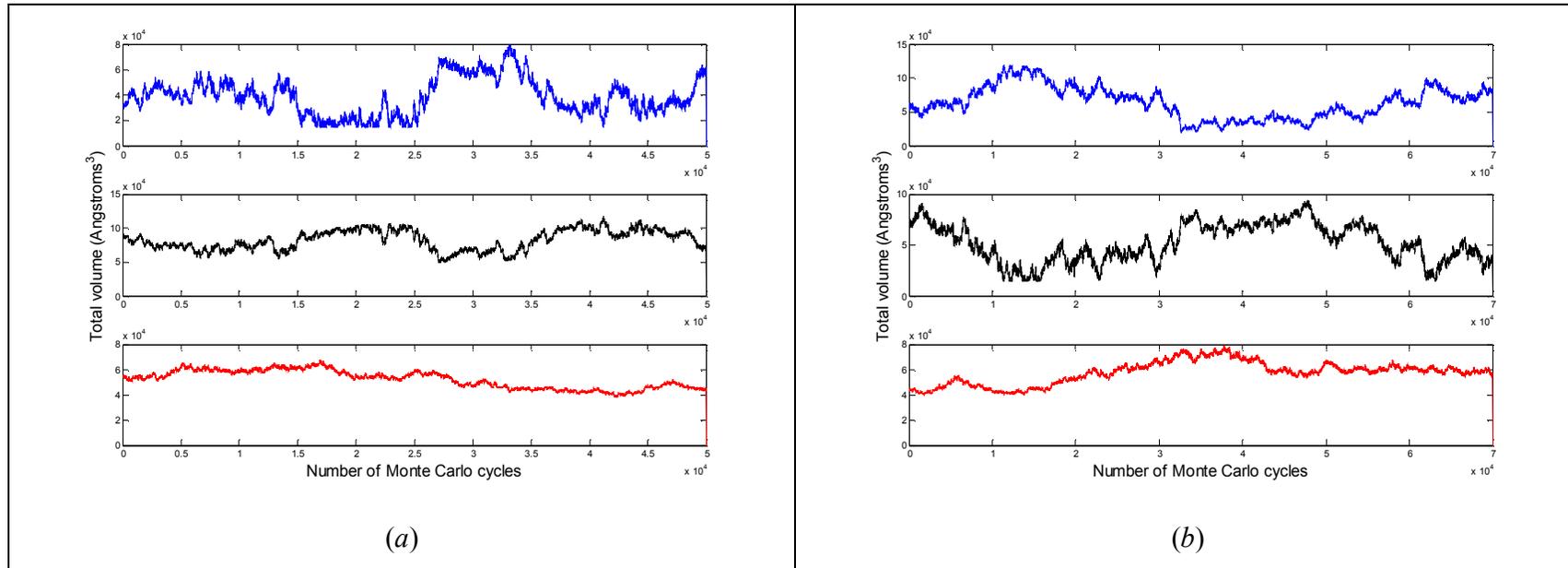
**Figure 9-6 - Plot of the variation of individual box volumes during a production run. Vapour: top graph; n-hexane phase: middle; water phase: bottom.**



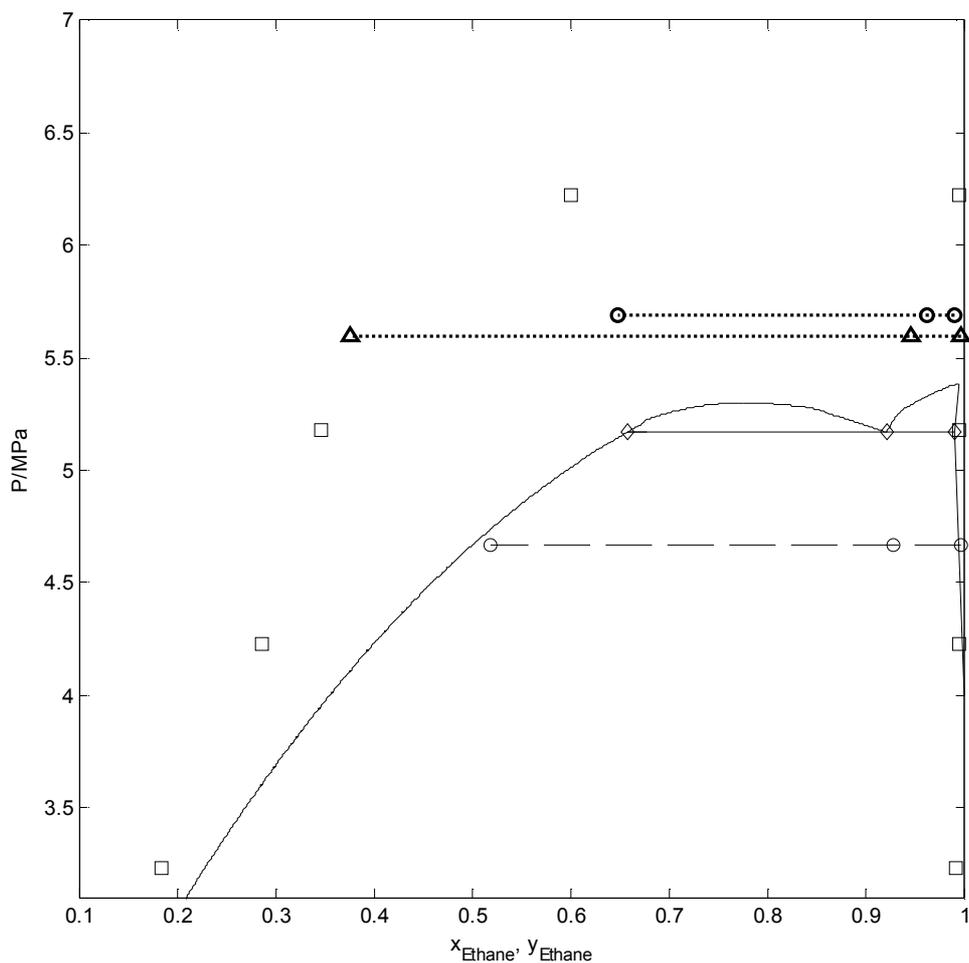
**Figure 9-7 - Comparison of box energies for ethane/ethanol in the NVT ensemble before and after box identity swaps. (a) Just before the identity swap (b) After the identity swap.**



**Figure 9-8 - Comparison of number of ethane molecules in each simulation box for ethane/ethanol in the NVT during an identity swap between the vapour and ethane-rich liquid phases. (a) Just before the identity swap (b) After the identity swap.**



**Figure 9-9 - Comparison of number of volume changes in each simulation box for ethane/ethanol in the NVT during an identity swap between the vapour and ethane-rich liquid phases. (a) Just before the identity swap (b) After the identity swap.**



**Figure 9-10 - Phase diagram for ethane/ethanol at 311.15 K. The experimental phase envelope is shown as a solid line (Kato *et al.*, 1999), with the experimental VLLE line joined by hollow diamonds ( $\diamond$ ). Simulations using the TraPPE-UA force field in the two phase regions as well as those which reverted to two phases are shown as squares ( $\square$ ). The simulated VLLE region for 500 molecules is marked by circles ( $\circ$ ), while for 1000 molecules, the phase compositions are marked by triangles, with points being joined by dotted lines. The VLLE line predicted by the NERDv3 force field is shown by a dashed line, joining circles marking the three phases.**

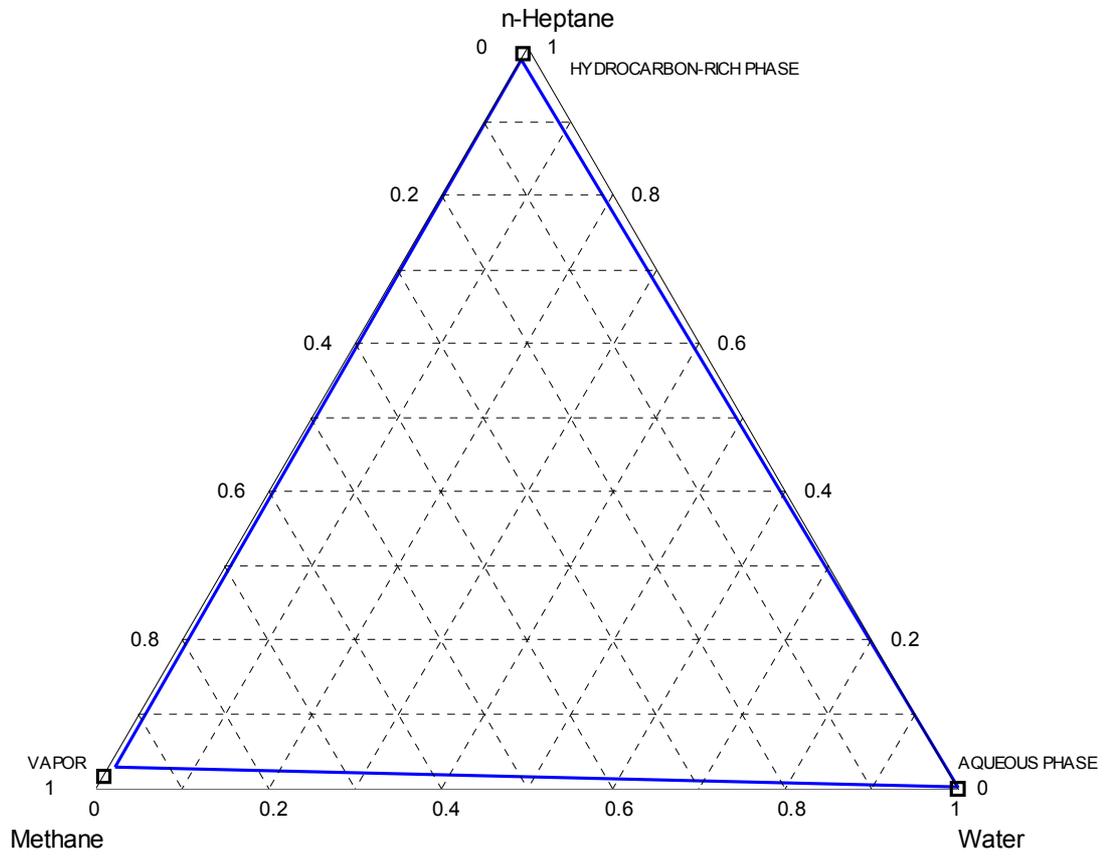
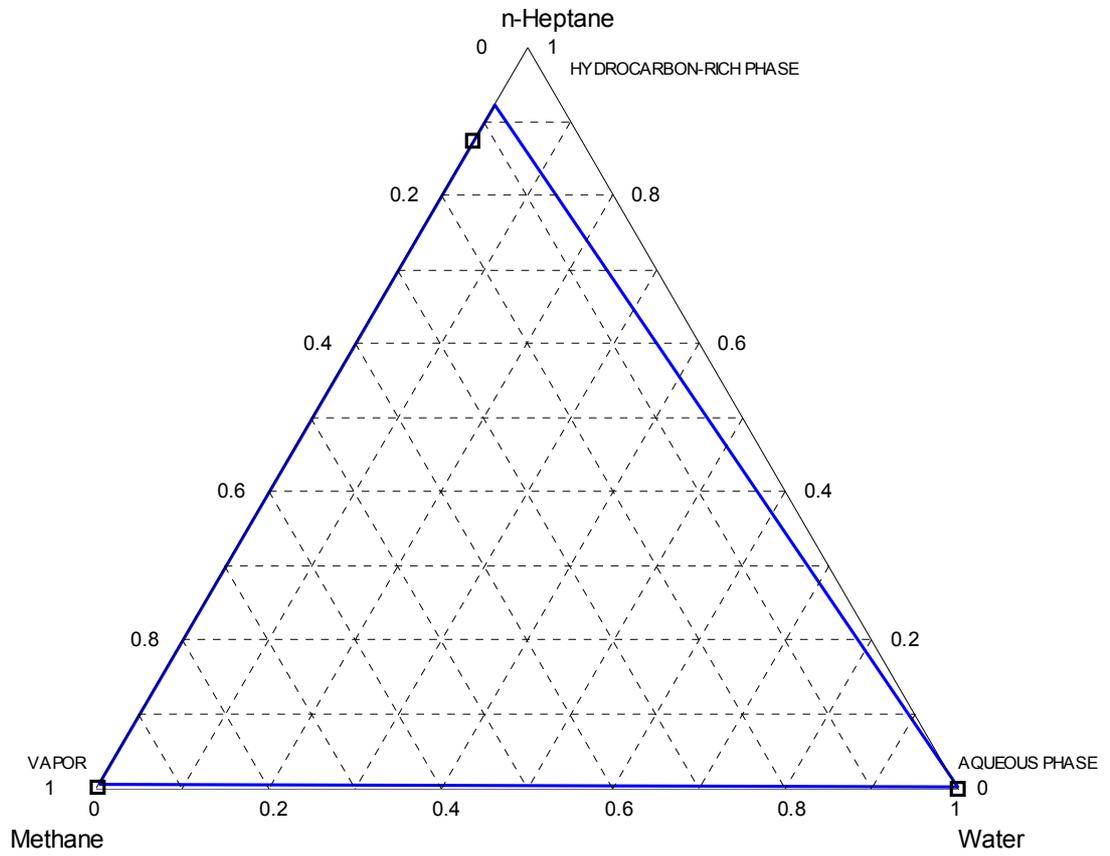
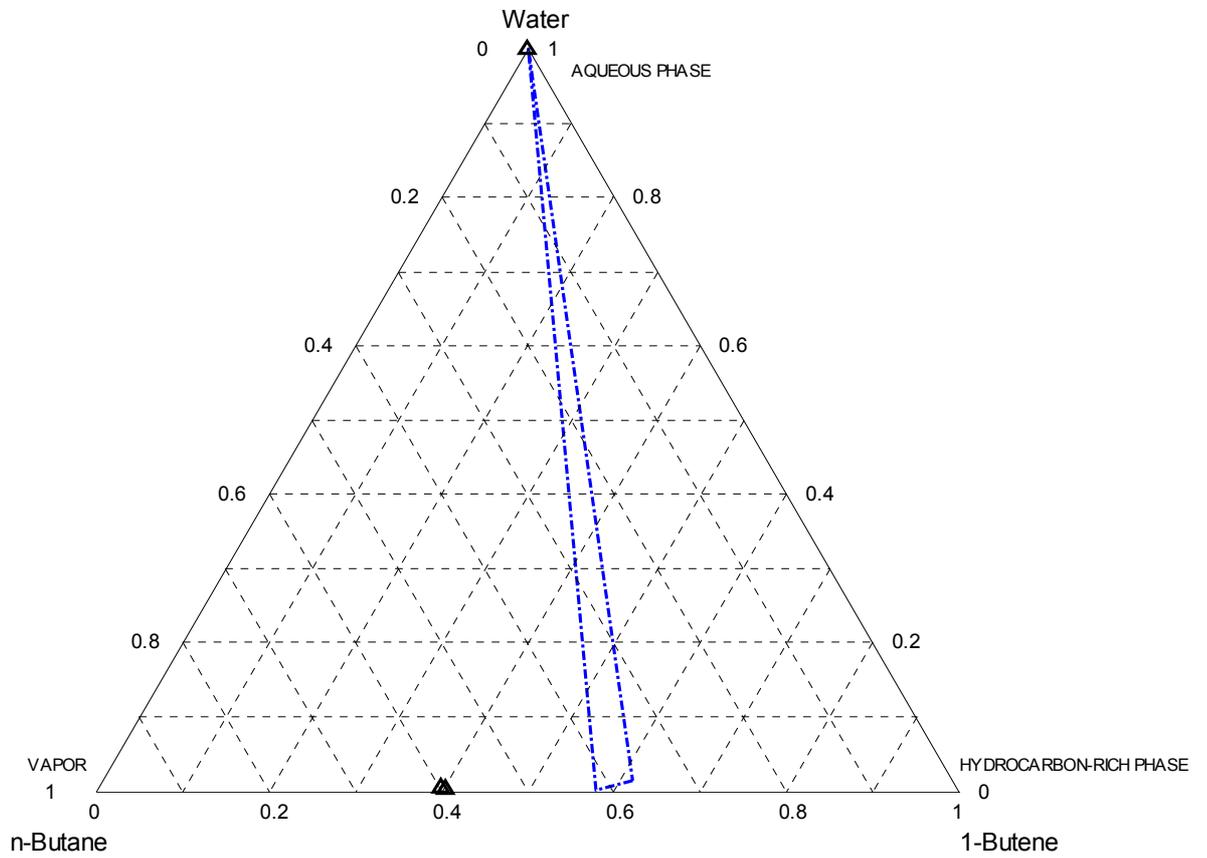


Figure 9-11 - VLLE ternary composition diagram for methane/water/n-heptane at 275.5 K and 120 kPa. The experimental three phase region is shown as a solid triangle (Susilo *et al.*, 2005), while the region predicted by NPT simulations is plotted using squares (□).



**Figure 9-12 - VLLE ternary composition diagram for methane/water/n-heptane at 275.5 K and 2000 kPa. The experimental three phase region is shown as a solid triangle (Susilo *et al.*, 2005), while the region predicted by NPT simulations is plotted using squares (□).**



**Figure 9-13 - VLE ternary composition diagram for n-butane/1-butene/water at 310.93 K and 404.72 kPa. The experimental three phase region is shown as a dashed triangle (Wehe and McKetta, 1961), while the region predicted by NPT simulations is plotted using triangles. This simulation reverted to two phases, hence the close proximity of the bottom two triangle symbols.**

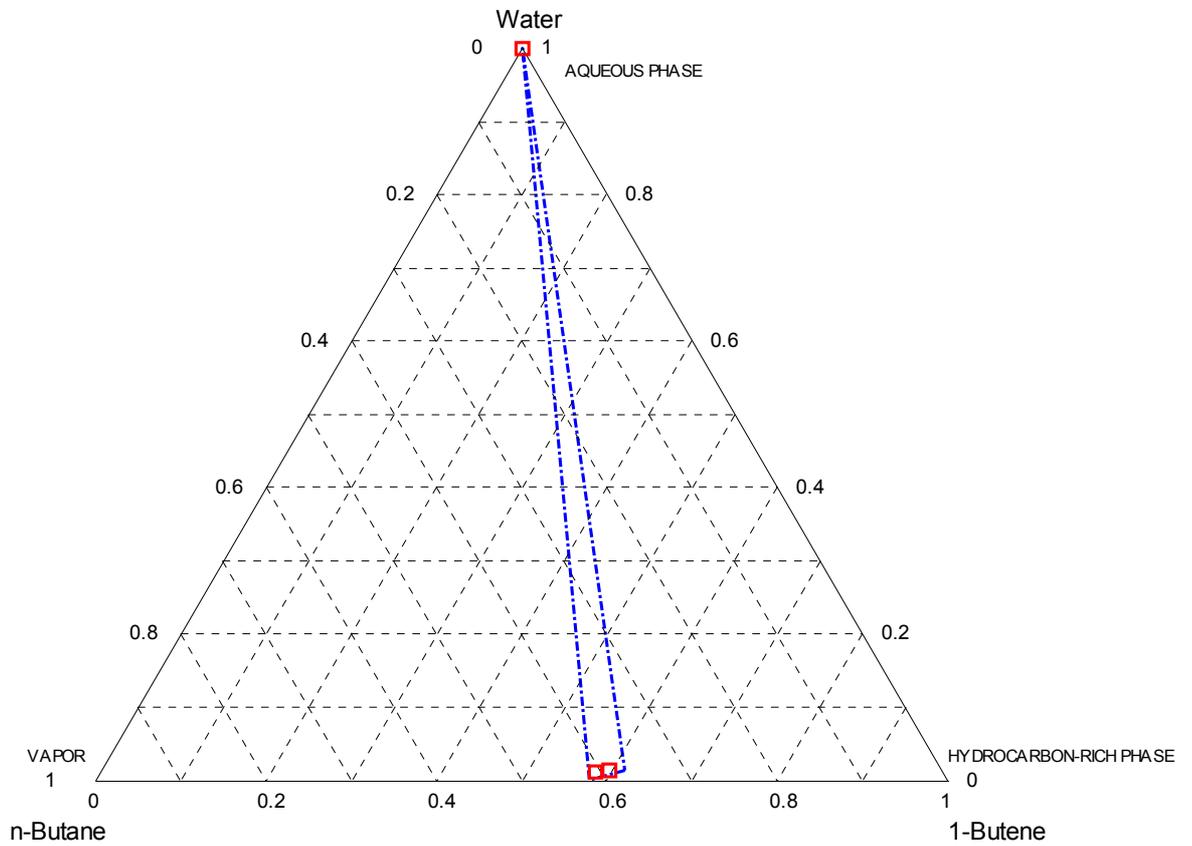
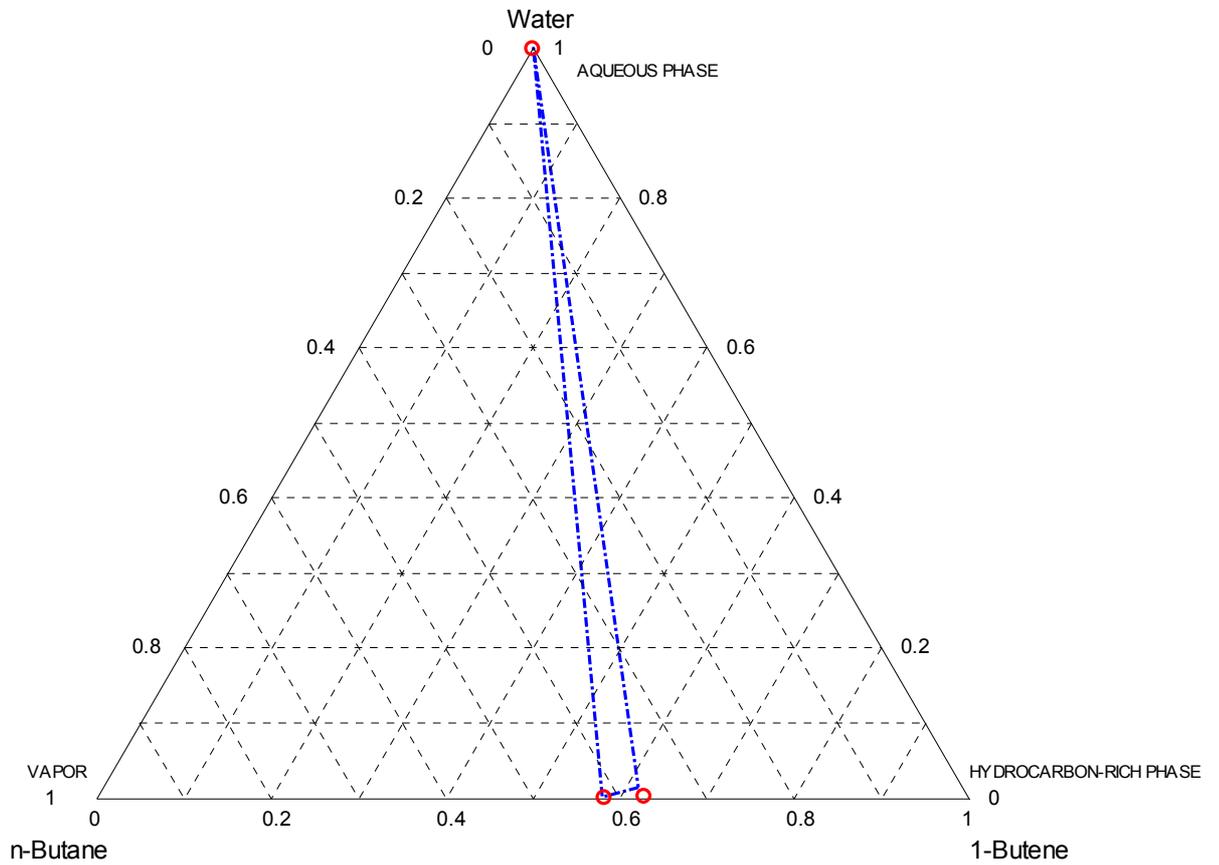


Figure 9-14 - VLLE ternary composition diagram for n-butane/1-butene/water at 310.93 K and 101.33 kPa. The experimental three phase region is shown as a dashed triangle (Wehe and McKetta, 1961), while the region predicted by NPT simulations is plotted using squares ( $\square$ ). This simulation also reverted to two phases, hence the proximity of the bottom two square symbols.



**Figure 9-15 - VLE ternary composition diagram for n-butane/1-butene/water at 310.93 K in the NVT ensemble. The experimental three phase region is shown as a dashed-triangle (Wehe and McKetta, 1961), while the region predicted by simulation is plotted using circles (○).**

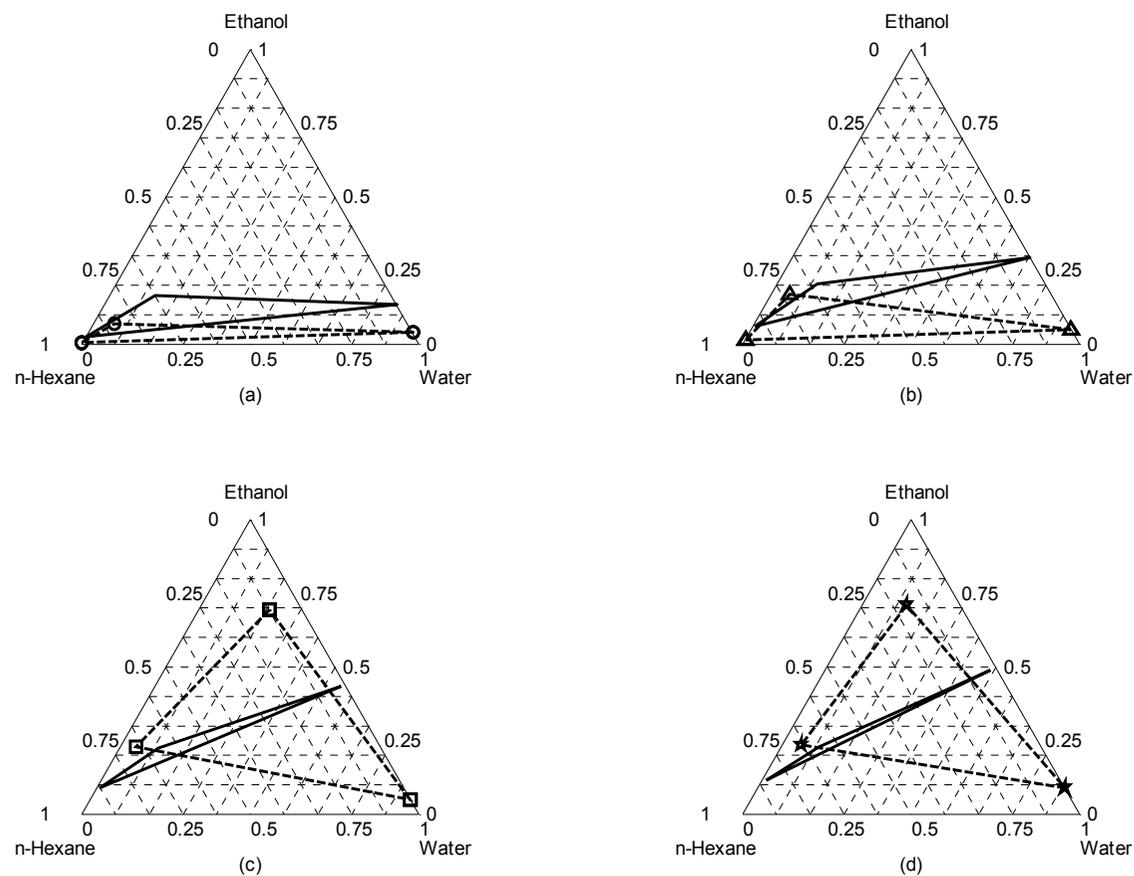


Figure 9-16 - VLE diagrams for water/ethanol/n-hexane at 101.33 kPa obtained from NPT-Gibbs ensemble simulations using a total 750 molecules. (a) 329.45 K using 232 ethanol molecules. (b) 329.51 K using 184 ethanol molecules. (c) 329.77 K using 108 ethanol molecules (d) 330.54 K using 45 ethanol molecules. Experimental three-phase envelopes are shown as solid lines (Gomis *et al.*, 2007), while simulation envelopes are shown as dashed lines joining the symbols which indicate phase compositions.

## 10. Discussion and Analysis of Results

The chosen mixtures were investigated due to the availability of reliable experimental data (see Appendix B). However, for the n-hexane/water mixture, no literature solubility data was available at the temperatures of interest; nevertheless, comparisons for the coexisting liquid phases were possible using temperature-dependent solubility correlations reported by Tsonopoulos and Wilson (1983).

### 10.1 Binary Simulations

#### 10.1.1 n-hexane/water

It is worth mentioning that prior to conducting the three phase simulations, two phase simulations in the NPT-Gibbs ensemble were undertaken to reproduce liquid-liquid equilibrium data reported by Johansson (2007), in order to validate the TraPPE-UA and SPC-E force fields as they are implemented in Towhee (see Table 9-4). The solubilities and densities obtained from these two phase simulations were in excellent agreement with previously reported simulations at similar conditions and consequently this model was used for the three systems studies here. However, two-phase formation was observed, with liquid-liquid equilibrium at slightly lower pressures and vapour-liquid equilibrium at much lower pressures and some simulations produced metastable states with three identical phases (Table 9-5). Nevertheless, this was promising as it suggested that an intermediate pressure existed between the two VL and LL state points which should produce three phases.

A study by Kamilov *et al.* (2001) on the isochoric heat capacities during liquid-liquid-vapour to liquid-vapour and liquid-liquid transitions of aqueous n-hexane/water mixtures provided a better starting point for the three-box NPT simulations. At the temperatures of

interest (that is to say, those which yielded three phases during experiments; see Chapter 9), the corresponding three-phase pressures were obtained from a temperature-dependent correlation reported by Tsonopoulos and Wilson (1983).

Using 369 n-hexane and 231 water molecules, the first seventy thousand cycles suggested three-phase formation, but it was observed that the density in the vapour box was steadily increasing along with a decrease in the water mole fraction in the same box. After an initial  $1 \times 10^5$  cycles, the mixture was left to equilibrate for a further  $5 \times 10^4$  cycles. Analysis of the simulation results revealed that these simulations reverted to two liquid phases i.e. a water-rich liquid phase in one box and n-hexane rich liquid phases in the other two boxes. The n-hexane/water simulations in the NPT-Gibbs ensemble both reverted to two phases in four independent simulations (each simulation was performed in duplicate). Although no emptying of any of the simulation boxes occurred, two boxes having identical densities and compositions, rich in n-hexane, were obtained. The density values of these two boxes were very close to that of pure n-hexane liquid, hence the assertion that the simulation reverted to two liquid phases (see Tables 9-2 and 9-3). This suggested that the specified pressure was too high to produce an additional vapour phase; in other words, the input pressure should have been lower.

Clearly, using the TraPPE-UA and SPC-E force fields in the NPT-Gibbs ensemble would not produce vapour-liquid-liquid equilibrium for this mixture, even when specifying the exact three-phase *experimental* conditions (pressures, temperatures and compositions) during a simulation. This is probably due to the simplicity of the Lorentz-Berthelot combining rules for unlike-pseudoatoms, which in this case underestimates the interactions between n-hexane and water atomic groups at the simulation conditions. Another possibility for not being able to obtain three phases in the NPT ensemble lies in the ensemble itself – the ensemble may be highly sensitive to large perturbations during the course of a simulation. These perturbations would result in sampling of points in Gibbs phase space where two phases exist, as opposed to sampling points where three phases would have been realized.

As outlined earlier in the chapter, progressively decreasing the simulation pressure until three phases were obtained would have been a cumbersome task. To this end, the overall

composition used in the constant pressure simulations was used in an NVT-Gibbs ensemble simulation at 482 K. The total system volume, now constant, was estimated by using the liquid density of the experimental mixture at 482 K ( $\rho_{\text{exp}} = 312.50 \text{ kg.m}^{-3}$  and  $z_1 = 0.614$ ), along with temperature dependent liquid density correlations (Perry & Green, 2007). An equilibration run consisting of  $7 \times 10^4$  cycles suggested the formation of three distinct phases – a n-hexane-rich vapour phase ( $\sim 75$  mole % n-hexane) and n-hexane-rich and water-rich liquid phases. At this stage of the simulation, the standard deviations of the densities of each simulation box between the calculation ‘blocks’ were promising. An unsigned deviation of  $7.85 \text{ kg.m}^{-3}$  for the vapour phase ( $\rho = 103.6 \text{ kg.m}^{-3}$ ) was obtained. A further  $7 \times 10^4$  cycles yielded thermodynamic averages which confirmed stability in the three distinct coexisting phases, followed by a production run of  $1.4 \times 10^5$  cycles in which the final equilibrium averages were obtained (Table 9-1). Figures 9-1 and 9-2 illustrate the fluctuations of the potential energy and liquid box volumes, respectively, versus the number of Monte Carlo cycles during the pre-equilibration period. A significant decrease in the potential energy (by approximately 50%) occurred in the n-hexane-rich box and a similar trend is seen in the box volume. The stability of the simulation during the equilibration period, i.e. when points in equilibrium phase space are sampled, is apparent in Figures 9-3 to 9-6, which show changes for the potential energies; number of n-hexane molecules; pressures and box volumes in the simulation boxes.

Thus, the NVT-Gibbs ensemble proved to be effective in overcoming the shortfalls of its constant pressure counterpart in binary three-phase simulations. The vapour phase mole composition ( $y_1 = 0.75_1$ ) could not be compared with any literature data, since no vapour phase experimental mole fractions for n-hexane/water mixtures have been reported. However, comparisons were possible for the liquid phases. At 482 K, a correlation of the form  $\ln x_{\text{hc}} = A + B/T + C \ln(T)$  (Tsonopoulos and Wilson, 1983) for the solubility of n-hexane (expressed as a mole fraction) in the water-rich phase predicts  $x_{\text{hc}}$  to be  $2.749 \times 10^{-4}$ . This is in very good agreement with the simulation value,  $x_{\text{hc}}^{\text{sim}} = 1.194 \times 10^{-4}$ , with a deviation of 56.6%. In the context of these simulations, ‘good agreement’ is understood to mean that the simulated mole fractions are of the same order of magnitude as those stated in the literature. A similar correlation for the solubility of water in the n-

hexane-rich liquid gives  $x_w = 0.1334$ , while the simulation solubility is  $x_w^{\text{sim}} = 0.0417$  giving a 68.7% deviation. In both cases, the simulations in this work slightly under-predicted the mutual solubilities. At 450 K, Johansson (2007) reported the solubility of water in n-hexane as  $x_w = 0.0132$ , also using the TraPPE-UA and SPC-E force fields – as expected, the mutual solubilities follow the trend of being greater at higher temperatures. Johansson (2007) obtained densities of  $488.9 \text{ kg.m}^{-3}$  and  $877.9 \text{ kg.m}^{-3}$  for the n-hexane- and water-rich phases respectively, while the corresponding simulation densities in this work were  $412.9 \text{ kg.m}^{-3}$  and  $821.2 \text{ kg.m}^{-3}$ . This is to be expected, since the NVT simulations were performed at a higher temperature, thus lower densities were obtained. The NVT simulation, besides producing three distinct phases, was also able to qualitatively agree with correlation-based solubility data, where available.

The average virial pressure during the NVT simulation was  $3663 \text{ kPa} \pm 80 \text{ kPa}$ ; as predicted earlier in this discussion, the actual simulation pressure that would have predicted three-phase formation was much lower than the experimental pressure,  $4116.41 \text{ kPa}$ , at 482 K by approximately 12%. Hence, only liquid phases were obtained from the NPT simulations.

For completeness a simulation in the NPT ensemble using the new ‘shifted’ pressure of  $3663 \text{ kPa}$ , obtained from the NVT simulations, was performed (again at 482 K; see Table 9-6). An equilibration run consisting of 120 000 Monte Carlo cycles was performed, followed by a 150 000-cycle production run. It was found that specifying the correct three-phase pressure (along with temperature and composition) did not produce a stable vapour phase – the density and composition of this phase fluctuated between those of the vapour and liquid phases corresponding to the NVT simulations, even when using vastly different box sizes for the ‘vapour’ and hexane-rich liquid boxes. Thus, the NVT ensemble is evidently more reliable than the NPT ensemble in predicting the VLLE of binary mixtures, provided that the potential models used in the simulations can adequately model the unlike-pair interactions especially in regions of high mutual solubility. This reconciles with an earlier discussion that the NPT ensemble is most probably very sensitive to perturbations in phase space even when the exact three-phase simulation pressure which corresponds to the articulated force fields (obtained from NVT simulations) is used.

### 10.1.2 ethane/ethanol

NPT-Gibbs ensemble simulations for mixtures of ethane/ethanol also produced two phases: an ethanol-rich liquid phase and a vapour phase, present in two simulation boxes, which was rich in ethane (Table 9-10). Using densities reported by Kato *et al.* (1999), simulations were performed using the experimental densities from liquid-liquid and vapour-liquid regions to get estimates for the total volume to use in the NVT ensemble; the total volumes used in these simulations were the sums of the individual experimental phase volumes within each liquid-liquid and vapour-liquid region. The results of these simulations indicated that the calculated densities were in agreement with the corresponding densities in the LL and VL experimental regions. Thus, there was a total system volume between the simulated LL and VL regions which would result in three-phase formation.

Using the experimental densities of the vapour and ethane- and ethanol-rich liquid phases, the total volume that was used in an NVT ensemble simulation was calculated using methods described earlier in this chapter. Using 400 ethane and 100 ethanol molecules, a simulation at 311.15 K was done in the NVT-Gibbs ensemble. Three distinct phases were obtained, although there was periodic swapping of identities between the vapour and ethane-rich liquid boxes. Hence, the final block averages obtained from this simulation were not very reliable since knowing when the start of an identity switch occurred was not possible; even if it was possible, there would not be a sufficient number of Monte Carlo cycles between each switch for calculating reliable thermodynamic averages (see Table 9-7). Figures 9-7 to 9-9 show comparisons of the total energies, number of ethane molecules and volumes, respectively, for the vapour and ethane-rich liquid phases. These fluctuations have a direct effect on the densities of the phases of concern and thus affect three-phase stability. The swapping phenomenon was initially attributed to being a system size effect; to investigate this, another simulation was performed using 800 ethane and 200 ethanol molecules, scaling the other simulation parameters and total system volume accordingly. Once again, an identity swap was observed between the vapour and ethane-rich liquid boxes. Upon analyzing the volumes of

these two boxes, it became apparent that the two volumes were nearly identical. Thus, yet another simulation was attempted, this time using vastly different volumes for the aforesaid boxes. This overcame the identity-swapping behaviour and vapour-liquid-liquid equilibrium was obtained in this simulation. The simulation results (Table 9-8) indicate reasonable agreement with experimental data (Kato *et al.*, 1999) for compositions as well as densities; Table 10-1 lists the experimental and simulation compositions and densities of each phase, as well as unsigned percent deviations.

	Mole fractions			Densities		
	Experiment	Simulation	Deviation	Experiment	Simulation	Deviation
<b>Phase</b>						
Vapour	0.990	0.9903	0.030 %	155.8	168	7.83 %
Ethane-rich liquid	0.921	0.963	4.56 %	355.1	273	23.1 %
Ethanol-rich liquid	0.677	0.609	4.28 %	491.3	513	4.42 %

**Table 10-1 – Comparison of the simulation and experimental results for phase compositions and densities for ethane/ethanol at 311.15 K for a successful NVT simulation using 1000 molecules. Mole fractions refer to those of ethane. Experimental data was obtained from a study by Kato *et al.* (1999).**

As can be seen, the differences between the experimental and simulations values for the compositions are in fair agreement, although for the densities, slightly larger deviations for the vapour and ethane-rich liquid phases are observed. This is due to a fair amount of swapping of ethane molecules between these two phases (although box identity swaps were not observed).

Again, there is a significant difference between the calculated virial pressure and the experimental pressure which produced three phases, with the three-phase simulation pressure of 5.889 MPa being 13.93% higher than the experimental value of 5.169 MPa, hence only vapour-liquid equilibrium was observed in the NPT simulations. This agrees

with previous VLE simulation work done by Chen *et al.* (2001) in which azeotropic pressures for mixtures of n-hexane/methanol and n-hexane/ethanol were over-predicted by as much as 30%, due to unlike pair interactions between alkane and alcohol molecules being weak. Furthermore, pure component vapour pressures were over-predicted by as much as 10%.

For completeness sake, several other VLE and LLE points were calculated in the NVT ensemble; these results are listed in Tables 9-11 through to 9-15. Figure 9-10 shows that there is indeed a pressure shift of the entire phase diagram for the ethane/ethanol mixture.

As a matter of interest, the NERD Version 3 force field (Nath *et al.*, 1998) was used in a once-off NVT simulation for the same ethane/ethanol mixture. Although no standard deviations were reported in a VLE study using this force field (Khare *et al.*, 2004), there is qualitative agreement since for a mixture consisting of n-heptane and 1-pentanol, the mole fraction of n-heptane was over-predicted in both phases causing the simulated envelope to shift to lower pressures; in this study, the three-phase VLLE line was shifted to 4.665 MPa (see Table 9-9 and Figure 9-10) – this value is 9.75% lower than the experimental pressure (5.169 MPa).

Although the results of the NVT-Gibbs ensemble simulations revealed deviations from experiment, this ensemble nevertheless provided a step in the right direction for further binary three- and multi-phase work for fluid phase equilibria. The deviations from experiment for the successful simulations (those which produced three phases) arose due to the Lorentz-Berthelot combining rules for unlike-interactions for the molecules of interest not being adequate in predicting their corresponding mixture properties except in dilute composition regions. Considering the extensive literature survey carried out prior to and during this study, it is believed that these are the first successful binary vapour-liquid-liquid equilibrium atomistic simulations involving complex molecules (Lopes and Tildesley (1997) performed simulations for simple, model Lennard-Jones beads).

A simple trial-and-error scheme for attaining three-phase binary equilibrium in the NVT ensemble would be as follows:

- Step 1: Using estimates for phase volumes in the experimental liquid-liquid and vapour-liquid regions for a given composition, perform simulations to ascertain whether the predicted regions indeed correspond to the experimental regions, or previous simulation work.
- Step 2: If successful, progressively decrease the volume used in the VLE simulation from Step 1 and increase the volume used in LLE simulation, in subsequent simulations that implement three boxes, until three phase formation becomes apparent in either of the simulations.

If Step 1 is unsuccessful, the following should be done: if Step 1 produced VLE in both simulations, then the total volume used in the simulation which gave the highest pressure should be further decreased so that two liquid phases are obtained; the opposite is true if two liquid-liquid regions were obtained in Step 1. Then proceed to Step 2.

## **10.2 Ternary simulations**

### **10.2.1 methane/water/n-heptane**

Simulations in the NPT ensemble for this mixture were successful with very good quantitative agreement at both pressures (see Tables 9-16 and 9-17), although at 2000 kPa the mole fraction of n-heptane in the hydrocarbon liquid phase was significantly under-predicted. This may be attributed to the fact that as the system pressure increases, more methane molecules will be transferred to the hydrocarbon liquid phase; furthermore, due to the densities in the other two phases increasing, there would be very low acceptance rates for swap moves for a large molecule such as n-heptane. Thus, while the number of methane molecules increases in the hydrocarbon liquid phase, the number of n-heptane molecules would be nearly constant in this phase, hence the under-prediction of the mole

fraction of n-heptane. n-Heptane is virtually insoluble in water at the temperature of interest (275.5 K) and considering the lengthy simulation times, it is only reasonable that there were no configurational-bias regrowths of this large molecule in the aqueous phase. The experimental solubilities (expressed as mole fractions) at 275.5 K for n-heptane in water, at 120 kPa and 2000 kPa, are  $1 \times 10^{-6}$  and  $1.1 \times 10^{-5}$  respectively; thus, simulation times could have been drastically reduced if swap moves for n-heptane from the hydrocarbon liquid box to the aqueous phase box were disallowed, since the coupled-decoupled regrowth procedure uses multiple trial sites during the growing of chain molecules.

The success of these simulations is due to the high degree of mutual insolubility amongst the three components, especially at ambient conditions – simulations in dilute regions for highly non-ideal mixtures such as alkanes and water usually yield excellent liquid densities and compositions, due to the unlike-pair interactions being less prominent; this is why the SPC-E force field was used instead of the SPC force field, since SPC-E better-predicts water liquid densities than SPC, which better-predicts the vapour density (Boulougouris *et al.*, 1998). Furthermore, in all of the mixtures that were simulated, water was the least volatile component. Thus, better results were obtained at 120 kPa than at 2000 kPa. Although the deviations of the simulation mole fractions from experiment are significant especially in those phases which are rich in one component (mole fraction  $> 0.9$ ; see Table 10-2 for a list of percent deviations), these simulations are considered successful due to overall phase stability and mole fractions obtained from the simulations being of the same order of magnitude as those reported by experimental work. Furthermore, considering that the experimental mole fractions for those components that weren't in high concentrations in a given phase were very small, e.g. methane in the aqueous phase, one would expect a significant statistical uncertainty from the simulations. More so, since there were no regrowths of large n-heptane molecules in the dense aqueous phases, large deviations were obtained. Better comparisons could have been made if uncertainties in the experimental measurements were reported.

The methane/water/n-heptane mixture is a good example of a ternary mixture that exhibits highly non-ideal behaviour with a very large three-phase composition region. In such mixtures, one is free to choose from many combinations of the number of molecules of each type to use in a simulation, at the same time staying far from the phase boundaries. There is thus very little risk of a simulation reverting to two (or fewer) phases.

		<b>Mole fractions <math>\times 10^6</math></b>					
		Methane			n-Heptane		
<b>Pressure [kPa]</b>	<b>Phase</b>	Experiment	Simulation	Deviation	Experiment	Simulation	Deviation
		120	Vapour	963987	983955.5	-2.07 %	29214
	HC	15111	8335.5	44.84 %	984045	991663.4	-0.77 %
	Aqueous	47	19	59.57 %	1	0	100.00 %
2000	Vapour	994973	998221.4	-0.33 %	4291	1719.4	59.93 %
	HC	76446	125422.9	-64.07 %	922922	874575.1	5.24 %
	Aqueous	523	72.1	86.21 %	11	0	100.00 %

**Table 10-2 - Comparison of the simulation and experimental results for phase compositions for methane/water/n-heptane at 275.5 K. HC refers to the hydrocarbon-rich liquid phase. Experimental data was obtained from a study by Susilo *et al.* (2005).**

### 10.2.2 n-butane/1-butene/water

This mixture exhibits an extremely narrow three-phase region (see Figure 34) and proved challenging to simulate in the NPT ensemble. Two simulations at 404.72 kPa and 101.33 kPa (both at 310.93 K) yielded a water-rich liquid phase in one box and a hydrocarbon-rich liquid phase in the remaining two simulation boxes (each having almost identical densities and compositions; see Tables 9-19 and 9-20). It must be noted that the initial compositions were selected by locating the central point of the triangular three-phase region to stay as far as possible from the phase boundaries. The initial analyses of these simulations during the equilibration periods suggested three distinct phases had formed but the uncertainties in the densities were not promising.

The final equilibrium values of the volumes of each simulation box from the NPT simulation at 404.72 kPa were used as a starting point in a constant volume simulation. The total volume from the NPT simulation was decreased by 15% and used as the volume in the NVT ensemble, keeping all other simulation conditions, except pressure, the same as in the NPT simulations. This proved successful as three phases were obtained albeit at a much higher pressure (548 kPa) with satisfactory uncertainties in the coexisting phase densities (see Table 9-18). The predicted compositions were in excellent agreement with experimental data (see Figure 9-15), almost completely reproducing the experimental three-phase composition boundary. Thus, it was shown that the NVT-Gibbs ensemble may also be used to overcome extremely narrow three-phase composition regions for ternary mixtures.

The excellent agreement for the phase compositions was due to n-butane and 1-butene having low solubilities in water (the experimental mole fractions are  $2.25 \times 10^{-5}$  and  $12.69 \times 10^{-5}$ , respectively), so one would expect reasonable results in the dilute region, although the mole fraction of water in the vapour and hydrocarbon phases. The vapour and hydrocarbon-rich phases were also in excellent agreement with experimental data. This is due to the unlike-pair interactions between saturated alkane pseudoatoms and vinylic alkene pseudoatoms being well accounted for (Wick *et al.*, 2000), so that there is no need for special interaction parameters to simulate mixtures of alkanes and alkenes to good accuracy. Wick *et al.* (2000) obtained excellent results for a supercritical mixture of ethene and n-heptane.

<b>Mole fractions</b>						
	<b>n-Butane</b>			<b>1-Butene</b>		
	Experiment	Simulation	Deviation	Experiment	Simulation	Deviation
<b>Phase</b>						
Vapour	0.373	0.372	0.17 %	0.614	0.62397	-1.62 %
HC	0.42	0.419	0.32 %	0.579	0.581281	-0.39 %
Aqueous	$2.25 \times 10^{-5}$	$0.51 \times 10^{-5}$	77.33 %	$12.7 \times 10^{-5}$	$1.9 \times 10^{-6}$	98.50 %
	<b>Water</b>					
	Experiment	Simulation	Deviation			
<b>Phase</b>						
Vapour	0.0128	0.00368	71.25 %			
HC	0.0012	$4.43 \times 10^{-5}$	96.31 %			
Aqueous	0.9998506	0.999993	-0.01 %			

**Table 10-3 – Comparison of the simulation and experimental results for phase compositions for n-butane/1-butene/water at 310.93 K. HC refers to the hydrocarbon-rich liquid phase. Experimental data was obtained from a study by Wehe and McKetta (1961).**

### 10.2.3 water/ethanol/n-hexane

Due to the availability of multiple VLLE state points for this mixture, four simulations at the conditions stated in Chapter 9 were performed in the NPT-Gibbs ensemble. The sizes of the three-phase composition regions were reasonable at each state point of interest, thus the number of water molecules in each simulation was kept constant at 150 molecules to ensure reasonable statistics for the aqueous phase in each simulation.

The equilibration runs for these mixtures, especially at high ethanol concentrations, required at least  $1.5 \times 10^5$  Monte Carlo cycles in order to obtain reasonable statistical uncertainties in the mole fractions of each component in each phase, more so for the organic and aqueous phases. Figure 9-16 shows the ternary composition diagrams for the four NPT simulations that were performed, each time increasing the number of ethanol molecules (while decreasing the number of n-hexane molecules to preserve the total number of molecules used in each simulation).

As can be seen, at very low ethanol concentrations the shape of the three-phase region is predicted somewhat qualitatively. An increase in the number of ethanol molecules progressively decreases the experimental three-phase region. This was not true for the simulations, as the simulated three-phase regions became progressively larger. This was due to there being very few ethanol molecules being transferred to the aqueous phase, essentially rendering the composition of the aqueous phase constant over all simulations, as there were also n-hexane molecules transferred to the water phase only at the highest simulation temperature used here (330.54 K), with a mole fraction  $x_3 = 0.0000004$ . This is in poor agreement with the corresponding experimental solubility ( $x_3^{\text{exp}} = 0.001$ ). Due to there being virtually no n-hexane transferred to the aqueous phase, poor agreement with experiment in all simulations for the water-rich phase was obtained. Considering the reasonable statistical uncertainties in the densities and compositions of each phase, it was assumed that the simulations had converged, so no further equilibration cycles were necessary.

Thus, an increase in the number of alcohol molecules here diminished the ability of the force fields to even qualitatively predict the equilibrium solubilities and hence the three-phase regions. This may be attributed again to the simplicity of the Lorentz-Berthelot mixing rules and the unlike-pair interactions being strong enough. As mentioned earlier in this Chapter, previous two-phase studies show systematic deviations for alcohol/alkane mixtures by over-estimating the alkane solubility in both phases. This is evident in this particular ternary mixture in the vapour and ethanol rich phases. Clearly, good agreement was obtained only at low concentrations of ethanol. In a *quantitative* study, better results would have been obtained if the binary simulations were carried out to better-predict the vapour-liquid and liquid-liquid phase envelopes of mixtures of alkanes/alcohols, alkanes/water and alcohols/water over large composition ranges. This would have required successive modifications of the unlike-pair interactions for the size and energy parameters in many simulations for each mixture until satisfactory agreement with experimental data was obtained. This certainly is an arduous task and was not attempted in this work due to limited time and computational resources (discussed in the next sub-section).

### 10.3 Simulation Times

The purpose of this work was to determine whether vapour-liquid-liquid equilibrium could be predicted by two common force fields. For binary mixtures, this proved oftentimes impossible using the NPT ensemble, until the NVT variant of the Gibbs ensemble was used. In both cases, extremely long equilibration times were required to obtain acceptable statistical uncertainties for compositions, pressures and densities. The virial pressure calculation for pressures in liquid simulation boxes, especially those rich in water, would have required an extremely large number of Monte Carlo cycles in order to be similar to that of the vapour phase. Thus, in all simulations, the final simulation pressure was taken to be equivalent to the vapour phase virial pressure as this is usually a very reliable estimate of the equilibrium pressure for a mixture (Martin, 2008). As can be seen from the tabulated numerical results (Chapter 9), some simulations which reverted to two phases were terminated as soon as there were similarities in the phase compositions and densities of any two boxes; hence there were slightly larger uncertainties in these compositions compared to successful runs. Experience from previous three-box simulations for binary and ternary mixtures in this work showed that they would eventually converge upon two phases and this was necessary since only six computational nodes were available for simulation purposes.

Considering that each simulation was duplicated to ensure that metastable states were not obtained on six computational nodes (the figures listed in Table 10-4 above are not exaggerations!), one obtains a better appreciation of the qualitative nature of this study. Furthermore, to obtain accurate results for the virial pressures in the simulations, pressure calculations were performed after each Monte Carlo cycle.

Table 10-4 lists typical simulation times for the mixtures that were studied.

Mixture	Ensembl		$N_{MC}$	
	e	Time [Days]		
		Equilibratio n		Productio n
n-hexane/water	NPT	10.68	5.34	60 000
	NVT	13.79	13.79	140 000
ethane/ethanol	NPT	6.10	3.81	50000
	NVT	15.26	5.72	75 000
n-hexane/ethanol/water	NPT	13.85	5.77	50 000
methane/n-heptane/water	NPT	21.12	8.80	50 000
n-butane/1-butene/water	NPT	13.18	6.59	75 000
	NVT	20.73	8.29	80 000

**Table 10-4 - Typical simulation times for mixtures studied in this work.  $N_{MC}$  refers to the number of cycles used during a production run.**

An issue that arises when performing simulations of highly immiscible mixtures, where the experimental solubility of a chemical species can be of the order of magnitude  $-4$  (and lower), is the trade-off between the total number of molecules that are used in the simulation and the length of the simulation i.e. the total number of Monte Carlo cycles that are used. If one uses, say, 600 total molecules in a simulation, then it is difficult to imagine how solubilities of orders of magnitudes  $-4$  or lower in the dilute region will be calculated. However, the number of swap moves that are carried out in a simulation is directly proportional to the number of Monte Carlo cycles, which in turn is equal to the number of molecules multiplied by the number of MC cycles. Thus, if 100 000 MC cycles are used, then a total of  $600 \times 1 \times 10^5 = 6 \times 10^7$  moves will occur. If 10% of these moves are interphase swaps, then a total of  $6 \times 10^6$  swap moves will have occurred. It is now clear, depending on the rate of accepted swaps, that for the scenario presented here, mole fractions of the order of magnitude  $-6$  may be predicted by the simulation. Obviously, for

simulations in which much lower solubilities are expected then a greater number of MC cycles must be used.

## 11. Conclusions and Recommendations

The aim of this study was to investigate the ability of the Gibbs ensemble, using Monte Carlo molecular simulations, in predicting the three-phase fluid equilibrium for several industrially-relevant mixtures using two common potential models, TraPPE-UA and SPC-E. The simulation results for each mixture were then compared to the corresponding experimental data.

The first successful binary liquid-liquid-vapour equilibrium simulations for complex molecules were performed for two mixtures in the NVT-Gibbs ensemble: n-hexane/water and ethane/ethanol. The simulations in the NPT ensemble were unsuccessful and reverted to two-phase liquid-liquid equilibrium for n-hexane/water and vapour-liquid equilibrium for ethane/ethanol due to there being significant under- and over- predictions, respectively, in the simulation pressures (as confirmed by the virial pressures obtained from the NVT simulations). This was due to the unlike-pair interactions for each different molecular pair not being well-accounted for. Even using the three-phase virial pressure obtained from the NVT ensemble simulations for n-hexane/water as the input pressure in the NPT ensemble did not produce three distinct phases since the density and composition of the vapour box periodically changed, making it impossible to obtain meaningful ensemble averages. This suggests, aside from limitations in the unlike-pair interactions, that the NPT ensemble is very sensitive to perturbations when sampling regions of phase space which should give three phases.

The NVT ensemble has also been shown to overcome an extremely narrow VLLE ternary composition envelope, in which the NPT ensemble produced only two distinct phases. Excellent agreement with experimental data was obtained for a ternary n-butane/1-butene/water mixture.

For the water/ethanol/n-hexane mixture, experimental data indicated a fair degree of solubility amongst the components across all phases, especially as the proportion of ethanol molecules increased. However, the simulations showed large deviations from experiment,

especially when the number of ethanol molecules in the mixtures became higher. There were hardly any accepted swap moves and configurational-bias regrowths for n-hexane into the dense water-rich phase, as well as the solubility of ethanol in water also being underestimated. Thus, while the NPT-Gibbs ensemble produced three distinct phases, the limitations in these simulations were the unlike-pair interactions.

The NPT and NVT variants of the Gibbs ensemble are thus adequate in predicting three-phase formation. However, the NVT variant is clearly the ensemble of choice when performing simulations for ternary mixtures that exhibit very narrow three-phase regions and also for simulations involving binary mixtures. Furthermore, one is more likely to obtain VLLE in an NVT simulation than in an NPT simulation as the total volume is easier to control than using a shifted pressure in the NPT ensemble; Gibbs phase rule states that for VLLE in binary mixtures, there is all but one degree of freedom. Thus, once the pressure is fixed, there are no additional degrees of freedom; hence, the simulation pressure specified in the NPT ensemble would have to be extremely close to the experimental three-phase pressure, assuming that the unlike-pair interactions are accurate enough.

The greatest limitations arose from the force fields and the NPT ensemble. In regard to the force fields, the combining rules that were used could not accurately predict the mixture behaviour. For ternary mixtures, provided that a set of force fields can adequately reproduce the liquid-liquid and vapour-liquid phase envelopes for each binary pairing, there would be better predictions which are in qualitative and quantitative agreement with experimental data. The way forward is to re-parameterize existing force fields, or develop improved mixing rules so that better quantitative agreement is obtained with the two-phase regions of binary mixtures. One would then expect more accurate results for the simulation of vapour-liquid-liquid equilibrium. Furthermore, more work needs to be done to ascertain whether the NPT ensemble is viable for predicting three-phase fluid equilibrium of binary mixtures.

## 12. References

1. Lennard-Jones, J. E (1931), 'Cohesion', *Proceedings of the Physical Society*, **43**, 461-482.
2. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M., Teller, A. H. and Teller, E. (1953), 'Equation of state calculations by fast computing machines', *Journal of Chemical Physics* **21**, 1087–1092.
3. Rosenbluth, M. N. and Rosenbluth, A. W. (1955), 'Monte Carlo Calculation of the Average Extension of Molecular Chains', *Journal of Chemical Physics* **23**(2), 356–359.
4. Wehe, A. H. and McKetta, J. J. (1961), 'n-Butane-1 -Butene-Water System in the Three-phase Region', *J. Chem. Eng. Data*, **6**(2), 167-172.
5. Widom, B., (1963), 'Some topics in the theory of fluids', *Journal of Chemical Physics* **39**, 2808–2812.
6. Pathria R. K., 1972, "Statistical Mechanics", 1<sup>st</sup> edition, Pergamon Press, U.K.
7. de Leeuw, S. W., Perram, J. W. and Smith, E. R. (1980a), 'Simulation of Electronic Systems in Periodic Boundary Conditions. I. Lattice Sums and Dielectric Constants', *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **373**(1752), 27–56.
8. de Leeuw, S. W., Perram, J. W. and Smith, E. R. (1980b), 'Simulation of Electronic Systems in Periodic Boundary Conditions. II. Equivalence of Boundary Conditions', *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **373**(1752), 57–66.

9. de Leeuw, S. W., Perram, J. W. and Smith, E. R. (1980c), 'Simulation of Electronic Systems in Periodic Boundary Conditions. III. Further Theory and Applications', *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **388**(1794), 173–193.
10. Rowlinson, J S, and Swinton, F L, 1982, 'Liquid and Liquid mixtures, 3<sup>rd</sup> Edition', Butterworth Scientific, London.
11. Tsonopoulos, C. and Wilson, G. M. (1983), 'High-Temperature Mutual Solubilities of Hydrocarbons and Water', *AIChE Journal*, **29**(6), 990-999.
12. Allen, M. P., and Tildesley, D. J., (1987), "Computer Simulation of Liquids", 2<sup>nd</sup> Edition, Clarendon Press, Oxford.
13. Berendsen, H. J. C., Grigera, J. R. and Straatsma, T. P. (1987), 'The Missing Term in Effective Pair Potentials', *J. Phys. Chem.*, **91**, 6269-6271.
14. Panagiotopoulos, A. Z. (1987), 'Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble', *Molecular Physics* **61**(4), 813–826.
15. Panagiotopoulos, A. Z., Quirke, N., Stapleton, M. and Tildesley, D. J., 1988, 'Phase equilibria by simulation in the Gibbs ensemble. Alternative derivation, generalization and application to mixture and membrane equilibria', *Molecular Physics* **63**(4), 527–545.
16. Rowlinson, J S, and Widom, B, 1989, 'Molecular Theory of Capillarity', Oxford University Press, New York.

17. Panagiotopoulos, A. Z. (1989), 'Exact Calculations of Fluid-Phase Equilibria by Monte Carlo Simulation in a New Statistical Ensemble', *Int. J. Thermophys.* **10** 447-457.
18. J. I. Siepmann (1990), "A method for the direct calculation of chemical potentials for dense chain systems", *Molecular Physics* **70** 1145–1158.
19. Laso, M., de Pablo, J. J., Suter, U. W. (1992), 'Simulation of phase equilibria for chain molecules', *Journal of Chemical Physics* **97** 2817-2819.
20. Frenkel, D., Siepmann, J. I. (1992), 'Configurational bias Monte Carlo: a new sampling scheme for flexible chains', *Molecular Physics* **75** 59-70.
21. Kofke, D. A. (1993*a*), 'Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line', *Journal of Chemical Physics* **98**(5), 4149–4162.
22. Kofke, D. A. (1993*b*), 'Gibbs–Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation', *Molecular Physics* **78**(6), 1331–1336.
23. Snurr, R. Q., Bell, A. T. and Theodorou, D. N. (1993); 'Prediction of Adsorption of Aromatic Hydrocarbons in Silicalite from Grand Canonical Monte Carlo Simulations with Biased Insertions', *Journal of Physical Chemistry* **97**, 13742.
24. M. Luscher (1994), 'A portable high-quality random number generator for lattice field theory calculations', *Computer Physics Communications* **79**, 100.
25. Smit, B., Karaborni, S. and Siepmann, J. I. (1995), 'Computer simulations of vapor-liquid phase equilibria of *n*-alkanes', *J. Chem. Phys.* **102**, 2126-2140.

26. Siepmann, J. I., Martin, M. G., Mundy, C. J. and Klein, M. L., 'Intermolecular potentials for branched alkanes and the vapour-liquid phase equilibria of n-heptane, 2-methylhexane, and 3-ethylpentane', *Mol. Phys.* **90**, 687-693 (1997).
27. Boulougouris, C., Economou, I. G. and Theodorou, D. N. (1998), 'Engineering a Molecular Model for Water Phase Equilibrium over a Wide Temperature Range', *J. Phys. Chem. B*, **102**, 1029-1035.
28. Canongia Lopes, J. N., Tildesley, D. J. (1997), 'Multiphase equilibria using the Gibbs ensemble Monte Carlo Method', *Molecular Physics* **92**(2), 187-195.
29. Hummer, G. and Grønbech-Jensen, N. (1998), 'Pressure calculation in polar and charged systems using Ewald summation: Results for the extended simple point charge model of water', *Journal of Chemical Physics* **109**(7), 2791-2797.
30. Martin, M.G. and Siepmann, J.I. (1998), 'Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes', *J. Phys. Chem. B*, **102**, 2569-2577.
31. Nath, S. K., Escobedo, F. A., de Pablo, J. J. (1998), 'On the simulation of vapor-liquid equilibria for alkanes', *Journal of Chemical Physics* **108** 9905-9911.
32. Potter, S. C., Tildesley, D. J., Burgess, A. N. and Rogers, S. C. (1998), 'A transferable potential model for the liquid-vapour equilibria of fluoromethanes', *Mol. Phys.* **92** 825-833.
33. Vlugt, T. J. H., Martin, M. G., Smit, B., Siepmann, J. I., Krishna, R. (1998), 'Improving the efficiency of the configurational-bias Monte Carlo algorithm', *Molecular Physics* **94** 727-733.

34. Kato, M., Tanaka, H. and Yoshikawa, H (1999), 'High-Pressure Phase Equilibrium for Ethane + Ethanol at 311.15 K', *J. Chem. Eng. Data*, **44**, 116-117.
35. Martin, M. G., Siepmann, J. I. (1999), 'Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. united-atom description of branched alkanes', *Journal of Physical Chemistry B* **103** 4508-4517.
36. Vlugt, T. J. H., Krishna R., Smit, B. (1999), 'Molecular Simulations of Adsorption Isotherms for Linear and Branched Alkanes and Their Mixtures in Silicalite' *Journal of Physical Chemistry B* **103** 1102-1118.
37. Panagiotopoulos, A. Z. (2000), 'Monte Carlo methods for phase equilibria of fluids', *Journal of Physics: Condensed Matter* **12**, R25–R52.
38. Wick, C.D., Martin, M.G. and Siepmann, J.I. (2000), 'Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and of alkylbenzenes', *J. Phys. Chem. B*, **104**, 8008-8016.
39. Chen, B., Potoff, J.J. and Siepmann, J.I. (2001), 'Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary and tertiary alcohols', *J. Phys. Chem. B*, **105**, 3093-3104.
40. Kamilov, I. K., Stepanov, G. V., Abdulagatov, I. M., Rasulov, A. R. and Milikhina, E. I. (2001), "Liquid-Liquid-Vapor, Liquid-Liquid, and Liquid-Vapor Phase Transitions in Aqueous n-Hexane Mixtures from Isochoric Heat Capacity Measurements", *J. Chem. Eng. Data*, **46**, 1556-1567.

41. Potoff, J. J. and Siepmann, J. I. (2001), 'Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide and nitrogen', *AICHE J.*, **47**, 1676-1682.
42. Stubbs, J. M., Chen, B., Potoff, J. J., and Siepmann, J. I. (2001), 'Monte Carlo calculations for the phase equilibria of alkanes, alcohols, water and their mixtures', *Fluid Phase Equil.*, **183/184**, 301-309.
43. Watkins, E. K., Jorgensen, W. L. (2001), 'Perfluoroalkanes: Conformational analysis and liquid-state properties from ab initio and Monte Carlo calculations', *Journal of Physical Chemistry A*, **105**, 4118.
44. Chen, B. and Siepmann, J. I. and Klein, M. L. (2002), 'Vapor-Liquid Interfacial Properties of Mutually Saturated Water/1-Butanol Solutions', *J. Am. Chem. Soc.*, **124**, 12232-12237.
45. Frenkel, D. and Smit, B. (2002), *Understanding Molecular Simulations: From Algorithms to Applications*, Vol. 1 of Computational Science Series, 2<sup>nd</sup> edition, Academic Press, San Diego.
46. Levine I. N., 2003, "Physical Chemistry", 5<sup>th</sup> edition, McGraw-Hill, New York.
47. Nath, S. K (2003), 'Molecular Simulation of Vapor-Liquid Phase Equilibria of Hydrogen Sulfide and Its Mixtures with Alkanes', *Journal of Physical Chemistry B* **107** 9498-9504.
48. Khare, R., Sum, A. K., Nath, S. K. and de Pablo, J. J. (2004), 'Simulation of Vapor-Liquid Phase Equilibria of Primary Alcohols and Alcohol-Alkane Mixtures', *Journal of Physical Chemistry B* **108** 10071-10076 (2004)
49. Martin, M. G., Thompson, A. P. (2004); 'Industrial property prediction using Towhee and LAMMPS' *Fluid Phase Equilibria* **217** 105-110.

50. Wick, C. D., Siepmann, J. I. and Schure, M. R. (2004), 'Simulation Studies on the Effects of Mobile-Phase Modification on Partitioning in Liquid Chromatography', *Anal. Chem.*, **76**, 2886-2892.
51. du Preez, N. B. (2005), Determination of Phase Equilibria for Long-Chain Linear Hydrocarbons by Monte Carlo Simulation, MScEng Thesis, University of KwaZulu-Natal.
52. McKnight, T. (2005), Molecular Simulation of Binary VLE using Beowulf Clusters, PhD Thesis, University of KwaZulu-Natal.
53. Susilo, R., Lee, J.D. and Englezos, P. (2005), 'Liquid-liquid equilibrium data of water with neohexane, methylcyclohexane, tert-butyl methyl ether, n-heptane and vapor-liquid-liquid equilibrium with methane', *Fluid Phase Equilibria*, **231**, 20-26.
54. Zhang, L. and Siepmann, J.I. (2005), 'Pressure Dependence of the Vapor-Liquid-Liquid Phase Behavior in Ternary Mixtures Consisting of n-Alkanes, n-Perfluoroalkanes, and Carbon Dioxide', *J. Phys. Chem. B*, **109**, 2911-2919.
55. Vesely, F. J., (2005), Franz Vesely Personal Home Page.  
[http://homepage.univie.ac.at/Franz.Vesely/sp\\_english/sp/node14.html](http://homepage.univie.ac.at/Franz.Vesely/sp_english/sp/node14.html).
56. Chen, B. and Siepmann, J. I. (2006), 'Microscopic Structure and Solvation in Dry and Wet Octanol', *J. Phys. Chem. B*, **110**, 3555-3563.
57. Martin, M. G., Frischknecht, A. L. (2006), 'Using arbitrary trial distributions to improve intramolecular sampling in configurational-bias Monte Carlo', *Molecular Physics* **104** 2439-2456.

58. Sun, L., Klotz, W.L., Siepmann, J.I. and Schure, M.R. (2006), 'Retention in gas-liquid chromatography with a polyethylene oxide stationary phase: Molecular simulation and experiment', *J. Chromatogr. A*, **1126**, 373-380.
59. Sun, L., Siepmann, J.I. and Schure, M.R. (2006), 'Conformation and solvation structure for an isolated n-octadecane chain in water, methanol, and their mixtures', *J. Phys. Chem. B*, **110**, 10519-10525.
60. Zhang, L., Rafferty, J.L., Siepmann, J.I., Chen, B. and Schure, M.R. (2006), 'Chain conformation and solvent partitioning in reversed-phase liquid chromatography: Monte Carlo simulations for various water/methanol concentrations', *J. Chromatogr. A*, **1126**, 219-231.
61. Gomis, V., Font, A., Pedraza, R. and Saquete, M.D. (2007), 'Isobaric vapor-liquid and vapor-liquid-liquid equilibrium data for the water-ethanol-hexane system', *Fluid Phase Equilib.*, **259**, 66-70.
62. Perry, R.H. and Green, D.W., 2007, "Perry's Chemical Engineers' Handbook", McGraw-Hill, New York.
63. Johansson, E., Bolton, K., Theodorou, D. N. and Ahlström, P. (2007), "Monte Carlo simulations of equilibrium solubilities and structure of water in n-alkanes and polyethylene", *J. Chem. Phys.*, **126**, 224902-1-224902-7.
64. Rafferty, J.L., Zhang, L., Siepmann, J.I. and Schure, M.R. (2007), 'Retention mechanism in reversed-phase liquid chromatography: A molecular perspective', *Anal. Chem.*, **79**, 6551-6558.
65. Sun, L., Siepmann, J.I. and Schure, M.R. (2007) 'Monte Carlo simulations of an isolated n-octadecane chain solvated in water-acetonitrile mixtures', *J. Chem. Theor. Comp.*, **3**, 350-357.

66. Martin, M. G., (2008), MCCCCS Towhee Home Page. <http://towhee.sourceforge.net>
67. Bolhuis, P. G. and Frenkel, D. (1996), 'Tracing the phase boundaries of hard spherocylinders', *J. Chem. Phys.*, **106**, 666-687.
68. Brooks, F. P., 1995, "The Mythical Man-Month: Essays on Software Engineering, Anniversary Edition", 2<sup>nd</sup> Edition, Addison-Wesley Professional.
69. Poncela, A., Rubio, A. M. and Juan J. Freire (2003), 'Gibbs ensemble simulation of symmetric mixtures composed by the homopolymers AA, BB and their common block copolymer AB', *J. Chem. Phys.*, **118**, 425-433.

## Appendix A

### Potential Model Parameters

This appendix lists the parameters for the force fields that were used in this work; for TraPPE-UA (Martin and Siepmann (1998), Wick *et al.* (2000) and Chen *et al.* (2001)) and NERDv3 (Nath *et al.* (1998) and Khare *et al.* (2004)), only values for molecules studied in this work are listed. For a more comprehensive list, refer to the relevant papers cited above.

#### A.1 TraPPE

Non-bonded parameters

United-Atom	$\epsilon/k_B$ [K]	$\sigma$ [Å]	Charge, $q$
CH <sub>4</sub>	147.9	3.73	-
CH <sub>3</sub> (Ethane)	104.1	3.775	-
CH <sub>3</sub>	98	3.75	-
CH <sub>2</sub> (sp <sup>3</sup> )	46	3.75	-
CH <sub>2</sub> (sp <sup>2</sup> )	85	3.95	-
CH (sp <sup>2</sup> )	47	3.73	-
$\alpha$ -CH <sub>2</sub> (sp <sup>3</sup> )	46	3.75	+0.265
O	93	3.02	-0.700
H	0	0	+4.35

**Table A-1 - TraPPE-UA Non-bonded parameters for pseudoatoms used in this work.**

Bond-length parameters

Bond	$l_0$
CH <sub>x</sub> -CH <sub>y</sub>	1.54
CH <sub>x</sub> -OH	1.43
O-H	0.945
CH <sub>x</sub> =CH <sub>y</sub>	1.33

**Table A-2 - TraPPE-UA bond lengths for bonded-pairs used in this work.**

Torsional parameters

Torsion	$c_0/k_B$ [K]	$c_1/k_B$ [K]	$c_2/k_B$ [K]	$c_3/k_B$ [K]
CH <sub>x</sub> -(CH <sub>2</sub> )-(CH <sub>2</sub> )-CH <sub>y</sub>	0.0	335.03	-68.19	791.32
CH <sub>x</sub> -(CH <sub>2</sub> )-(CH <sub>2</sub> )-OH	0.0	176.62	-53.34	769.93
CH <sub>x</sub> -(CH <sub>2</sub> )-(O)-H	0.0	209.82	-29.17	187.93
CH <sub>x</sub> =(CH)-(CH <sub>2</sub> )-CH <sub>y</sub>	688.5	86.36	-109.77	-282.24

**Table A-3 - TraPPE-UA torsion constants for dihedral angles.**

Bond-bending parameters

Bend	$\theta_0$ [deg.]	$k_\theta/k_B$
CH <sub>x</sub> -(CH <sub>2</sub> )-CH <sub>y</sub>	114	62500
CH <sub>x</sub> -(CH <sub>y</sub> )-O	109.47	50400
CH <sub>x</sub> -(O)-H	108.5	55400
CH <sub>x</sub> =(CH)-CH <sub>y</sub>	119.7	70420

**Table A-4 - TraPPE-UA equilibrium angles and bending constants used in this work.**

## A.2 SPC-E

The SPC-E force field models Water as a single site when calculating non-bonded interactions with other molecules (Berendsen, Grigera and Straatsma, 1987). As such, there are no bond-bending, stretching or dihedral terms.

## Non-bonded and parameters

$\epsilon$ [kcal/mol]	$\sigma$ [Å]	Charge, $q$		H-O-H angle (deg.)
0.155	3.166	H: +0.4238	O: -0.84760	109.47

**Table A-5 - SPC-E parameters used to represent Water in this work.****A.3 NERDv3**

## Non-bonded parameters

United-Atom	$\epsilon/k_B$		
	[K]	$\sigma$ [Å]	Charge, $q$
CH <sub>3</sub> (Ethane)	100.6	3.825	-
CH <sub>3</sub>	104	3.91	-
$\alpha$ -CH <sub>2</sub> (sp <sup>3</sup> )	45.8	3.93	+0.290
O	108	2.98	-0.710
H	3.89	0.98	+4.20

**Table A-6 - NERDv3 Non-bonded parameters for pseudoatoms used in this work.**

## Bond-stretching parameters

Bond	$l_0$	$k_r/k_B$ [K]
CH <sub>x</sub> -CH <sub>y</sub>	1.54	168380
CH <sub>x</sub> -OH	1.43	198448
O-H	0.961	312706
CH <sub>x</sub> =CH <sub>y</sub>	1.34	48250

**Table A-7 - NERDv3 bond-stretching parameters for bonded-pairs used in this work.**

## Torsional parameters

Torsion	$c_0/k_B$ [K]	$c_1/k_B$ [K]	$c_2/k_B$ [K]	$C_3/k_B$ [K]
CH <sub>x</sub> -(CH <sub>2</sub> )-(O)-H	0.0	359.25	59.053	220.82

**Table A-8 - NERDv3 torsion constants for dihedrals.**

## Bond-bending parameters

Bend	$\theta_0$ [deg.]	$k_\theta/k_B$
CH <sub>x</sub> -(CH <sub>y</sub> )-O	108	60136
CH <sub>x</sub> -(O)-H	107.5	27662

**Table A-9 - NERDv3 equilibrium angles and bending constants used in this work.**

## Appendix B

### Experimental Data

#### B.1 water (1) + ethanol (2) + n-hexane

$T_b$ [K]	Organic phase			Aqueous Phase			Vapour phase		
	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
334.440	0.002	0.000	0.998	1.000	0.000	0.000	0.214	0.000	0.786
332.570	0.001	0.004	0.995	0.965	0.035	0.000	0.159	0.076	0.765
330.800	0.002	0.015	0.983	0.893	0.106	0.001	0.139	0.150	0.711
330.540	0.002	0.021	0.976	0.869	0.131	0.001	0.137	0.162	0.702
330.050	0.003	0.028	0.969	0.791	0.208	0.001	0.128	0.185	0.688
329.920	0.003	0.039	0.958	0.754	0.244	0.002	0.124	0.195	0.681
329.820	0.004	0.049	0.947	0.732	0.266	0.002	0.123	0.197	0.680
329.770	0.005	0.060	0.935	0.704	0.293	0.003	0.122	0.202	0.676
329.710	0.006	0.065	0.930	0.674	0.322	0.004	0.120	0.207	0.673
329.510	0.008	0.087	0.905	0.555	0.433	0.012	0.114	0.221	0.665
329.480	0.010	0.101	0.889	0.524	0.461	0.014	0.112	0.225	0.663
329.450	0.012	0.113	0.875	0.489	0.492	0.019	0.110	0.229	0.661
329.420	0.015	0.133	0.852	0.431	0.541	0.029	0.107	0.235	0.658
329.280	0.022	0.172	0.806	0.353	0.593	0.053	0.103	0.242	0.656
329.500	0.037	0.235	0.728	0.296	0.624	0.079	0.100	0.248	0.653
329.530	0.047	0.272	0.680	0.248	0.620	0.132	0.096	0.253	0.652
329.550	0.055	0.297	0.648	0.234	0.609	0.157	0.096	0.254	0.650
329.560	0.061	0.316	0.622	0.219	0.604	0.178	0.095	0.256	0.649

Table B-1 - Experimental VLLE data for Water (1) + Ethanol (2) + n-Hexane (3) at 101.3 kPa (Gomis *et al.*, 2007).

**B.2 methane (1) + water (2) + n-heptane (3)**

$P$ [kPa]	Organic phase			Aqueous Phase			Vapour phase		
	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
120	15111	844	984045	47	999953	1	963987	6800	29214
2000	76446	632	922922	523	999466	11	994973	737	4291

Table B- 2 - Experimental VLLE data for Methane (1) + Water (2) + n-Heptane (3) at 277.5 K (Susilo *et al.*, 2005). Mole fractions have been multiplied by  $1 \times 10^6$ .

**B.3 n-butane (1) + 1-butene (2) + water (3)**

$P$ [Pa]	Organic phase			Aqueous Phase			Vapour phase		
	$x_1$	$x_2$	$x_3$	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$
404.72	0.4	0.57	0.001	2.2	12.6	0.999850	0.37	0.61	0.012
0	2	9	2	5	9	6	3	4	8

Table B-3 - Experimental VLLE data for n-Butane (1) + 1-Butene (2) + Water (3) at 310.93 K (Wehe and McKetta, 1961). The mole fractions of n-Butane and 1-Butene in the aqueous phase have been multiplied by  $1 \times 10^5$ .

**B.4 ethane (1) + ethanol (2)**

$T$ [K]	Ethanol-rich liquid		Ethane-rich liquid		Vapour	
	$x_1$	$x_2$	$x_1$	$x_2$	$y_1$	$y_2$
311.15	0.677	0.323	0.921	0.079	0.990	0.010

Table B-4 - Experimental VLLE data point for Ethane (1) + Ethanol (2) at 5.169 MPa (Kato, Tanaka and Yoshikawa, 1999).

### B.5 n-hexane (1) + water (2)

The predictive correlations that were used in to determine experimental solubilities of n-Hexane in Water and Water in n-Hexane were taken from Tsonopoulos and Wilson (1983):

The solubility of n-Hexane in Water (expressed as the logarithm of the mole fraction of n-Hexane is:

$$\ln x_{hc} = -367.98497 + 16128.646/T[\text{K}] + 52.820813 \ln(T[\text{K}]), \quad \text{B1}$$

while the solubility of Water in n-Hexane is expressed as:

$$\ln x_w = -45.1714 - 1635.73/T[\text{K}] + 7.53503 \ln(T[\text{K}]). \quad \text{B2}$$

Additionally, the three-phase pressures at the temperatures of interest were obtained using the following temperature dependent correlation (Tsonopoulos and Wilson, 1983):

$$\ln P_3 = 9.8127 - 4047.70/T[\text{K}]. \quad \text{B3}$$

### B.6 Liquid Density Correlations

Estimates for the pure-phase box volumes were obtained from the following predictive equation (Perry and Green, 1997):

$$\rho[\text{kmol.m}^{-3}] = C_1 / C_2^\psi, \quad \text{B4}$$

where  $\psi = 1 + (1 - \{T[\text{K}]/C_3\}^{C_4})$ . Values for the  $C$  constants are listed in Table B-5.

<u>Chemical Species</u>	<u>C<sub>1</sub></u>	<u>C<sub>2</sub></u>	<u>C<sub>3</sub></u>	<u>C<sub>4</sub></u>	<u>Molecular Weight</u>
Ethane	1.9122	0.27937	305.32	0.29187	30.07
Ethanol	1.2400	0.27342	508.30	0.23530	46.07
n-Hexane	0.7082	0.26411	507.60	0.27537	86.18
n-Heptane	0.6126	0.26211	540.2	0.2814	100.13
1-Butene	1.0972	0.26490	419.95	0.29043	87.80
1-Propanol	1.2350	0.27136	536.78	0.24000	60.10
n-Butane	1.0677	0.27188	425.12	0.28688	58.12
Methane	2.9214	0.28976	190.56	0.2888	16.043
Water	5.4590	0.30542	647.13	0.08100	18.02

**Table B-5 - Constants used in the in the liquid density correlations for pure components (Perry and Green, 2007).**

## Appendix C

### Sample Input and Output Files

This appendix lists an abridged output file for a n-Hexane/Water simulation in the NVT-Gibbs ensemble. Matlab® was used to extract information from the output files to generate the graphs presented in this study.

#### C.1 Printout of simulation settings and parameters

In this section, the first part of the output file is shown, listing the simulation settings as well as the interaction parameters for each molecules type.

```

Reading from towhee_input file: towhee_input
in directory: current directory
inputformat:Towhee
random_luxlevel:          3
random_allow_restart:    T
ensemble: nvt
temperature: 483.020
nmolty:                   2
nmolctyp:                 231          369
numboxes:                 3
stepstyle: cycles
nstep:                    140000
printfreq:                1
blocksize:                28000
moviefreq:                100000
backupfreq:               1000
runoutput: full
Full output of updates and block averages
pdb_output_freq:         70000
loutdft: F
loutlammps: F
pressurefreq:            1
trmaxdispfreq:          14000
volmaxdispfreq:          14000
chempotperstep:          0          0
potentialstyle: internal
ffnumber:                2
ff_filename:
/home/SUREN/towhee-5.2.3/ForceFields/towhee_ff_TraPPE-UA
/home/SUREN/towhee-5.2.3/ForceFields/towhee_ff_SPC-E
classical_potential: Lennard-Jones
READCLASSICAL: pot_num:    1 potential name: Lennard-Jones

```

```

classical_mixrule: Lorentz-Berthelot
lshift: F
ltailc: T
rmin: 1.00000
rcut: 9.00000
rcutin: 7.00000
electrostatic_form: coulomb
coulombstyle: ewald_fixed_kmax
kalp: 5.60000
kmax: 5
dielect: 1.00000
Setting up force field parameters from files
opening forcefield file: 1
opening forcefield file: 2
Lorentz-Berthelot Mixing rules
Arithmetic mean of sigma terms
Geometric mean of epsilon term
nfield: 0
solvation_style: none
No solvation model used
limit: F
initboxtype: dimensions
initstyle Box: 1
full cbmc full cbmc
initstyle Box: 2
full cbmc full cbmc
initstyle Box: 3
full cbmc full cbmc
Box: 1 initlattice: simple cubic simple cubic
Box: 2 initlattice: simple cubic simple cubic
Box: 3 initlattice: simple cubic simple cubic
Box: 1 initmol: 0 100
Box: 2 initmol: 0 269
Box: 3 initmol: 231 0
Box: 1 inix, iniy, iniz: 4 7 4
Box: 2 inix, iniy, iniz: 6 7 7
Box: 3 inix, iniy, iniz: 6 6 7
Box idim hmatrix: 1 1 44.85000 0.00000 0.00000
Box idim hmatrix: 1 2 0.00000 44.85000 0.00000
Box idim hmatrix: 1 3 0.00000 0.00000 44.85000
Box idim hmatrix: 2 1 45.36000 0.00000 0.00000
Box idim hmatrix: 2 2 0.00000 45.36000 0.00000
Box idim hmatrix: 2 3 0.00000 0.00000 45.36000
Box idim hmatrix: 3 1 19.52000 0.00000 0.00000
Box idim hmatrix: 3 2 0.00000 19.52000 0.00000
Box idim hmatrix: 3 3 0.00000 0.00000 19.52000
itest: 1 pairbox: 1 2
itest: 2 pairbox: 1 3
itest: 3 pairbox: 2 3
pmvol: 0.100000E-01
pmvlpr: 0.330000 0.667000 1.00000
rmvol: 0.100000
taval: 0.500000
pm2boxrbswap: 0.00000
pm2rbswmt: 0.400000 1.00000
pm2rbswpr: 0.400000 0.800000 1.00000
pm2boxcbswap: 0.105000

```

```
pm2cbswmt: 0.450000      1.00000
pm2cbswpr: 0.330000      0.670000      1.00000
pm1boxcbswap: 0.00000
pm1cbswmt: 1.00000      1.00000
pmavb1: 0.00000
pmavb1in: 0.500000
pmavb1mt: 0.500000      1.00000
moltyp: 1 pmavb1ct: 1.000000 1.000000
moltyp: 2 pmavb1ct: 1.000000 1.000000
avb1rad: 4.50000
pmavb2: 0.00000
pmavb2in: 0.500000
pmavb2mt: 1.00000      1.00000
moltyp: 1 pmavb2ct: 1.000000 1.000000
moltyp: 2 pmavb2ct: 1.000000 1.000000
avb2rad: 4.50000
pmavb3: 0.00000
pmavb3mt: 1.00000      1.00000
moltyp: 1 pmavb3ct: 1.000000 1.000000
moltyp: 2 pmavb3ct: 1.000000 1.000000
avb3rad: 4.50000
pmcb: 0.330000
pmcbmt: 0.500000      1.00000
pmall: 0.600000      1.00000
pmback: 0.00000
pmbkmt: 1.00000      1.00000
pmpivot: 0.00000
pmpivmt: 1.00000      1.00000
pmconrot: 0.00000
pmcrmt: 1.00000      1.00000
pmcrback: 0.00000
pmcrbmt: 1.00000      1.00000
pmplane: 0.00000
pmplanebox: 0.500000      0.700000      1.00000
planewidth: 3.00000
pmrow: 0.00000
pmrowbox: 0.500000      0.700000      1.00000
rowwidth: 3.00000
pmtraat: 0.00000
pmtamt: 0.500000      1.00000
rmtraa: 0.300000E-01
tatraa: 0.500000
pmtracm: 0.670000
pmtcmt: 0.580000      1.00000
rmtrac: 0.500000
tatrac: 0.650000
pmrotate: 1.00000
pmromt: 0.500000      1.00000
rmrot: 0.800000E-01
tarot: 0.800000
cbmc_style: coupled-decoupled
coupled_decoupled_form: Martin and Siepmann JPCB 1999
Coupled-decoupled form from M.G. Martin;
J.I. Siepmann; J. Phys. Chem. B 103 2977-2980 (1999)
cbmc_setting_style: default ideal
input_style: basic connectivity map
nunit: 3
```

```

nmaxcbmc:          3
lpdbnames:  F
  using the SPC-E      force field
charge_assignment: manual
  Building the input file for molecule type:      1
unit:   1 name:HW      charge:   0.42380
unit:   2 name:OW      charge:  -0.84760
unit:   3 name:HW      charge:   0.42380
input_style: basic connectivity map
nunit:          6
nmaxcbmc:          6
lpdbnames:  F
  using the TraPPE-UA force field
charge_assignment: bond increment
  Building the input file for molecule type:      2
unit:   1 name:CH3*(sp3)
unit:   2 name:CH2**(sp3)
unit:   3 name:CH2**(sp3)
unit:   4 name:CH2**(sp3)
unit:   5 name:CH2**(sp3)
unit:   6 name:CH3*(sp3)
Charges assigned for Molecule Type:      2
Unit:   1 nbname: CH3*(sp3) Charge:   0.00000
Unit:   2 nbname: CH2**(sp3) Charge:   0.00000
Unit:   3 nbname: CH2**(sp3) Charge:   0.00000
Unit:   4 nbname: CH2**(sp3) Charge:   0.00000
Unit:   5 nbname: CH2**(sp3) Charge:   0.00000
Unit:   6 nbname: CH3*(sp3) Charge:   0.00000
Total charge for Molecule Type:      2 is:   0.00000
Verifying input structures are consistent
Determining cyclic subunits for molecule type      1
Determining cyclic subunits for molecule type      2
Default total charge on molecule  1 is  0.00000
Default total charge on molecule  2 is  0.00000
Total charge in the simulation system:  0.00000
Bond Types
Type:   1 Style: Fixed Length: 1.5400
Type:  16 Style: Fixed Length: 1.0000
Angle Types
Type:   1 Style: Standard Harmonic Angle:   114.000 Constant:
31250.0
Type:  20 Style: Fixed Angle Angle:   109.470
Torsion Types
Type:   1 Style: Old UA OPLS Cosine Series
      k0:   0.0 k1:  355.0 k2:   -68.2 k3:   791.3
      with 1-4 vdw and scaled (0.500) 1-4 coulomb
Improper Torsion Types
  No Improper Types
  Canonical Gibbs ensemble
  3-dimensional periodic box
  Additional Center-of-Mass cutoff
  Dual Cutoff Configurational-bias Monte Carlo
  Coupled-decoupled Configurational-bias MC
  Coulombic inter- and intra-molecular interactions
    with an Ewald sum
    including the real-space terms up to half the shortest box length

```

```

Molecular mass for molecule type 1 is 18.0148 g/mol
Molecular mass for molecule type 2 is 86.1766 g/mol
Reading in initial conformation from towhee_initial
Initial version: 5
PROPER RESULTS ONLY WITH INITIALISATION FROM 25 INTEGERS
OBTAINED WITH RLUXUT
FULL INITIALIZATION OF RANLUX WITH 25 INTEGERS:
    6915574 15881122 13997077 9983268 4960010
    3644328 1571928 4909142 9882476 14341548
    2170679 12379367 9490550 13341825 11496213
    8390550 11244060 3121403 7265867 12619516
    980763 7280783 1877264 9421043 3100115
RANLUX LUXURY LEVEL SET BY RLUXIN TO: 3
new maximum displacements read from towhee_initial
box: 1
molecule type: 1
Max displacement for Atom translate: 0.030000
Max displacement for COM translate: 4.033061
Max displacement for rotation: 0.831536
molecule type: 2
Max displacement for Atom translate: 0.030000
Max displacement for COM translate: 2.430330
Max displacement for rotation: 0.355826
box: 2
molecule type: 1
Max displacement for Atom translate: 0.030000
Max displacement for COM translate: 1.480958
Max displacement for rotation: 0.953631
molecule type: 2
Max displacement for Atom translate: 0.030000
Max displacement for COM translate: 0.828284
Max displacement for rotation: 0.124619
box: 3
molecule type: 1
Max displacement for Atom translate: 0.030000
Max displacement for COM translate: 0.318994
Max displacement for rotation: 0.108494
molecule type: 2
Max displacement for Atom translate: 0.030000
Max displacement for COM translate: 0.589077
Max displacement for rotation: 0.085825
Max disp. for 3D Volume: 0.5957E-01 0.3457E-01 0.3267E-01
Max disp. for unit cell perturbation
Boxes 1 and 2 idim 1 rmcCell: 0.1000E+01 0.0000E+00 0.0000E+00
Boxes 1 and 2 idim 2 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 1 and 2 idim 3 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 1 and 3 idim 1 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 1 and 3 idim 2 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 1 and 3 idim 3 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 2 and 3 idim 1 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 2 and 3 idim 2 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
Boxes 2 and 3 idim 3 rmcCell: 0.0000E+00 0.0000E+00 0.0000E+00
new box dimensions read from towhee_initial
Box 1 hmatrix(1,x): 41.65034 0.00000 0.00000
Box 1 hmatrix(2,x): 0.00000 41.65034 0.00000
Box 1 hmatrix(3,x): 0.00000 0.00000 41.65034

```

Box	2	hmatrix(1,x):	48.09488	0.00000	0.00000
Box	2	hmatrix(2,x):	0.00000	48.09488	0.00000
Box	2	hmatrix(3,x):	0.00000	0.00000	48.09488
Box	3	hmatrix(1,x):	19.55866	0.00000	0.00000
Box	3	hmatrix(2,x):	0.00000	19.55866	0.00000
Box	3	hmatrix(3,x):	0.00000	0.00000	19.55866

Box:	1	Initial calp:	0.13445
Box:	1	Initial kmax:	5
Box:	2	Initial calp:	0.11644
Box:	2	Initial kmax:	5
Box:	3	Initial calp:	0.28632
Box:	3	Initial kmax:	5

Energies exclusively from internal potentials

Nonbonded Force Field

Lennard-Jones 12-6 potential

with tail corrections

$$u(r) = 4 * \epsilon [ (\sigma/r)^{12} - (\sigma/r)^6 ] - \text{shift}$$

Num. Atom(i)	Num. Atom(j)	sigma	epsilon	shift	1-4sig	1-4eps
3 CH3*(sp3)	3 CH3*(sp3)	3.7500	98.0000	0.0000	0.0000	0.0000
3 CH3*(sp3)	4 CH2**(sp3)	3.8500	67.1416	0.0000	0.0000	0.0000
3 CH3*(sp3)	30 OW	3.4578	87.5403	0.0000	0.0000	0.0000
3 CH3*(sp3)	31 HW	1.8750	0.0000	0.0000	0.0000	0.0000
4 CH2**(sp3)	4 CH2**(sp3)	3.9500	46.0000	0.0000	0.0000	0.0000
4 CH2**(sp3)	30 OW	3.5578	59.9755	0.0000	0.0000	0.0000
4 CH2**(sp3)	31 HW	1.9750	0.0000	0.0000	0.0000	0.0000
30 OW	30 OW	3.1656	78.1970	0.0000	0.0000	0.0000
30 OW	31 HW	1.5828	0.0000	0.0000	0.0000	0.0000
31 HW	31 HW	0.0000	0.0000	0.0000	0.0000	0.0000

Number of MC cycles:	140000
Number of molecules:	600
Temperature [K]:	483.02000

Initial Energies for Box 1

Total molecules in this box 69

Molecules of type 1 : 16

Molecules of type 2 : 53

total vibration	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
bond-bond(1-2)	0.000 [K]	0.00000 [kcal/mol]
total angle	46990.474 [K]	93.38021 [kcal/mol]
regular	46990.474 [K]	93.38021 [kcal/mol]
angle-angle	0.000 [K]	0.00000 [kcal/mol]
total torsion	75766.515 [K]	150.56442 [kcal/mol]
regular	75766.515 [K]	150.56442 [kcal/mol]
improper	0.000 [K]	0.00000 [kcal/mol]
total nonbond	-37428.472 [K]	-74.37845 [kcal/mol]
intramolecular	-6339.668 [K]	-12.59829 [kcal/mol]
2-body nonbond	-27581.904 [K]	-54.81120 [kcal/mol]
3-body nonbond	0.000 [K]	0.00000 [kcal/mol]
tail correct.	-3506.901 [K]	-6.96897 [kcal/mol]
total coulombic	-17506.969 [K]	-34.79013 [kcal/mol]
real space	-17334.254 [K]	-34.44690 [kcal/mol]
intramolec.	0.000 [K]	0.00000 [kcal/mol]
intermolec.	-17334.254 [K]	-34.44690 [kcal/mol]
self	-218556.590 [K]	-434.31912 [kcal/mol]
correction	217963.990 [K]	433.14149 [kcal/mol]
recip sum	419.885 [K]	0.83440 [kcal/mol]

external field	0.000 [K]	0.00000 [kcal/mol]
solvation	0.000 [K]	0.00000 [kcal/mol]
total classical	67821.54717498 [K]	134.7760530455 [kcal/mol]
Initial Energies for Box 2		
Total molecules in this box	332	
Molecules of type 1 :	16	
Molecules of type 2 :	316	
total vibration	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
bond-bond(1-2)	0.000 [K]	0.00000 [kcal/mol]
total angle	302605.503 [K]	601.34245 [kcal/mol]
regular	302605.503 [K]	601.34245 [kcal/mol]
angle-angle	0.000 [K]	0.00000 [kcal/mol]
total torsion	434499.189 [K]	863.44367 [kcal/mol]
regular	434499.189 [K]	863.44367 [kcal/mol]
improper	0.000 [K]	0.00000 [kcal/mol]
total nonbond	-593509.834 [K]	-1179.43215 [kcal/mol]
intramolecular	-32410.519 [K]	-64.40670 [kcal/mol]
2-body nonbond	-484272.758 [K]	-962.35450 [kcal/mol]
3-body nonbond	0.000 [K]	0.00000 [kcal/mol]
tail correct.	-76826.557 [K]	-152.67095 [kcal/mol]
total coulombic	-3578.049 [K]	-7.11035 [kcal/mol]
real space	-3502.923 [K]	-6.96106 [kcal/mol]
intramolec.	0.000 [K]	0.00000 [kcal/mol]
intermolec.	-3502.923 [K]	-6.96106 [kcal/mol]
self	-189270.811 [K]	-376.12195 [kcal/mol]
correction	188887.095 [K]	375.35943 [kcal/mol]
recip sum	308.590 [K]	0.61324 [kcal/mol]
external field	0.000 [K]	0.00000 [kcal/mol]
solvation	0.000 [K]	0.00000 [kcal/mol]
total classical	140016.80874348 [K]	278.2436206268 [kcal/mol]
Initial Energies for Box 3		
Total molecules in this box	199	
Molecules of type 1 :	199	
Molecules of type 2 :	0	
total vibration	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
bond-bond(1-2)	0.000 [K]	0.00000 [kcal/mol]
total angle	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
angle-angle	0.000 [K]	0.00000 [kcal/mol]
total torsion	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
improper	0.000 [K]	0.00000 [kcal/mol]
total nonbond	122286.953 [K]	243.01057 [kcal/mol]
intramolecular	0.000 [K]	0.00000 [kcal/mol]
2-body nonbond	127069.920 [K]	252.51536 [kcal/mol]
3-body nonbond	0.000 [K]	0.00000 [kcal/mol]
tail correct.	-4782.967 [K]	-9.50479 [kcal/mol]
total coulombic	-976478.272 [K]	-1940.47309 [kcal/mol]
real space	-911726.452 [K]	-1811.79725 [kcal/mol]
intramolec.	0.000 [K]	0.00000 [kcal/mol]
intermolec.	-911726.452 [K]	-1811.79725 [kcal/mol]
self	-5788638.952 [K]	-11503.27504 [kcal/mol]
correction	5714660.958 [K]	11356.26479 [kcal/mol]

```

    recip sum          9226.175 [K]          18.33440 [kcal/mol]
  external field          0.000 [K]          0.00000 [kcal/mol]
  solvation              0.000 [K]          0.00000 [kcal/mol]
  total classical      -854191.31893204 [K]  -1697.4625219678
[kcal/mol]
  initial virial pressure in box 1 =      4940.79
  initial virial pressure in box 2 =     -171.34
  initial virial pressure in box 3 =    -76109.48

```

## C.2 Runtime Printouts

The runtime printouts for the first 30 Monte Carlo cycles, along with the updating of maximum displacements for translation, rotation and volume moves after every 14 000 cycles is shown here; the total number of cycles used in this simulation was 140 000.

```
+++++ start of markov chain +++++
```

Cycle	Box	Energy [K]	Volume [A <sup>3</sup> ]	Press. [kPa]	Molecules
1	B: 1	0.6222E+05	0.7413E+05	2914.2	17 53
	B: 2	0.1286E+06	0.1094E+06	13406.4	15 316
	B: 3	-0.8503E+06	0.7482E+04	-13045.1	199 0
2	B: 1	0.6763E+05	0.7413E+05	2335.3	15 57
	B: 2	0.1381E+06	0.1095E+06	14218.6	17 312
	B: 3	-0.8513E+06	0.7314E+04	198810.7	199 0
3	B: 1	0.7891E+05	0.7413E+05	4056.0	15 58
	B: 2	0.1461E+06	0.1096E+06	13770.2	17 311
	B: 3	-0.8554E+06	0.7290E+04	209840.0	199 0
4	B: 1	0.8006E+05	0.7221E+05	5965.7	17 59
	B: 2	0.1593E+06	0.1113E+06	-2280.7	14 310
	B: 3	-0.8563E+06	0.7449E+04	29003.1	200 0
5	B: 1	0.7823E+05	0.7123E+05	6368.9	18 56
	B: 2	0.1550E+06	0.1124E+06	-6038.0	13 313
	B: 3	-0.8607E+06	0.7396E+04	43607.0	200 0
6	B: 1	0.8153E+05	0.7293E+05	5827.6	18 55
	B: 2	0.1450E+06	0.1107E+06	652.1	13 314
	B: 3	-0.8632E+06	0.7396E+04	61288.6	200 0
7	B: 1	0.7391E+05	0.7291E+05	7017.1	17 53
	B: 2	0.1397E+06	0.1106E+06	2211.3	14 316
	B: 3	-0.8625E+06	0.7445E+04	30345.8	200 0
8	B: 1	0.8261E+05	0.7288E+05	4466.5	14 54
	B: 2	0.1380E+06	0.1106E+06	7458.2	16 315
	B: 3	-0.8726E+06	0.7467E+04	-71597.1	201 0
9	B: 1	0.8511E+05	0.7208E+05	2355.1	15 57
	B: 2	0.1419E+06	0.1115E+06	351.5	15 312
	B: 3	-0.8733E+06	0.7390E+04	-10508.8	201 0
10	B: 1	0.7098E+05	0.7213E+05	1814.1	15 54
	B: 2	0.1359E+06	0.1115E+06	-1222.2	15 315
	B: 3	-0.8779E+06	0.7337E+04	41874.0	201 0
11	B: 1	0.6716E+05	0.7213E+05	6310.5	17 53
	B: 2	0.1431E+06	0.1115E+06	1225.0	13 316
	B: 3	-0.8742E+06	0.7337E+04	75107.0	201 0
12	B: 1	0.6882E+05	0.7180E+05	4705.6	14 53
	B: 2	0.1453E+06	0.1116E+06	1711.9	16 316
	B: 3	-0.8724E+06	0.7572E+04	-159109.0	201 0

13	B: 1	0.6382E+05	0.7459E+05	703.1	12	50
	B: 2	0.1319E+06	0.1090E+06	22132.7	18	319
	B: 3	-0.8709E+06	0.7447E+04	-15509.2	201	0
14	B: 1	0.6210E+05	0.7344E+05	1031.2	13	52
	B: 2	0.1300E+06	0.1101E+06	5703.2	18	317
	B: 3	-0.8718E+06	0.7464E+04	-89878.5	200	0
15	B: 1	0.6444E+05	0.7171E+05	3438.7	14	54
	B: 2	0.1333E+06	0.1119E+06	-6482.0	17	315
	B: 3	-0.8736E+06	0.7409E+04	-107979.6	200	0
16	B: 1	0.6657E+05	0.7171E+05	2621.2	16	54
	B: 2	0.1580E+06	0.1120E+06	-6999.8	15	315
	B: 3	-0.8796E+06	0.7256E+04	92115.8	200	0
17	B: 1	0.6792E+05	0.7159E+05	3043.0	18	53
	B: 2	0.1616E+06	0.1121E+06	-5935.4	13	316
	B: 3	-0.8772E+06	0.7269E+04	106093.4	200	0
18	B: 1	0.6309E+05	0.7159E+05	1898.6	17	52
	B: 2	0.1590E+06	0.1119E+06	-2369.7	14	317
	B: 3	-0.8768E+06	0.7473E+04	-122751.3	200	0
19	B: 1	0.6557E+05	0.7256E+05	539.8	19	52
	B: 2	0.1559E+06	0.1110E+06	4329.4	12	317
	B: 3	-0.8768E+06	0.7414E+04	-61155.3	200	0
20	B: 1	0.6448E+05	0.7278E+05	-336.5	18	52
	B: 2	0.1579E+06	0.1112E+06	3292.0	13	317
	B: 3	-0.8828E+06	0.7053E+04	362873.3	200	0
21	B: 1	0.6665E+05	0.7260E+05	420.3	18	53
	B: 2	0.1559E+06	0.1113E+06	-1742.9	13	316
	B: 3	-0.8855E+06	0.7091E+04	259264.1	200	0
22	B: 1	0.6335E+05	0.7188E+05	2370.0	19	52
	B: 2	0.1634E+06	0.1118E+06	3063.1	12	317
	B: 3	-0.8792E+06	0.7267E+04	22700.4	200	0
23	B: 1	0.6285E+05	0.7172E+05	3564.1	18	54
	B: 2	0.1619E+06	0.1121E+06	-1515.8	13	315
	B: 3	-0.8808E+06	0.7185E+04	114450.2	200	0
24	B: 1	0.6613E+05	0.7207E+05	3329.7	17	54
	B: 2	0.1540E+06	0.1115E+06	2351.5	13	315
	B: 3	-0.8812E+06	0.7407E+04	-133476.5	201	0
25	B: 1	0.6557E+05	0.7357E+05	-488.8	17	53
	B: 2	0.1428E+06	0.1100E+06	8750.8	12	316
	B: 3	-0.8854E+06	0.7424E+04	-186347.6	202	0
26	B: 1	0.6418E+05	0.7270E+05	1440.2	16	53
	B: 2	0.1384E+06	0.1109E+06	2077.7	14	316
	B: 3	-0.8796E+06	0.7380E+04	-114491.5	201	0
27	B: 1	0.6468E+05	0.7303E+05	1844.8	15	51
	B: 2	0.1369E+06	0.1107E+06	2320.7	15	318
	B: 3	-0.8778E+06	0.7229E+04	64429.7	201	0
28	B: 1	0.6419E+05	0.7340E+05	1990.8	15	51
	B: 2	0.1354E+06	0.1102E+06	12374.9	15	318
	B: 3	-0.8765E+06	0.7376E+04	-29473.0	201	0
29	B: 1	0.6304E+05	0.7328E+05	1980.5	16	52
	B: 2	0.1432E+06	0.1103E+06	17872.4	13	317
	B: 3	-0.8811E+06	0.7391E+04	-33633.2	202	0
30	B: 1	0.6614E+05	0.7172E+05	5599.3	16	53

.....

13999	B: 1	0.3067E+05	0.5270E+05	6028.4	6	17
	B: 2	0.1388E+06	0.1301E+06	-9417.0	8	352
	B: 3	-0.9412E+06	0.8182E+04	-71609.6	217	0

Updating maximum translational/rotational displacements

Box: 1 Molecule:	1	Attempts	Accepted	New Displacement
Translate COM		91746.	61251.	4.142357
Rotate		76933.	62457.	0.843839
Box: 1 Molecule:	2	Attempts	Accepted	New Displacement
Translate COM		144078.	96853.	2.513433
Rotate		165786.	134725.	0.361450
Box: 2 Molecule:	1	Attempts	Accepted	New Displacement
Translate COM		93252.	60833.	1.486314
Rotate		78191.	63006.	0.960540
Box: 2 Molecule:	2	Attempts	Accepted	New Displacement
Translate COM		1055527.	685446.	0.827504
Rotate		1219878.	976565.	0.124703
Box: 3 Molecule:	1	Attempts	Accepted	New Displacement
Translate COM		1469957.	956522.	0.319344
Rotate		1232269.	986113.	0.108527
Box: 3 Molecule:	2	Attempts	Accepted	New Displacement
Translate COM		106.	64.	0.547183
Rotate		106.	87.	0.088051

Updating 3D volume maximum displacements

Boxes 1 and 2 Tries:	27373	Accepted:	13863	Max Disp.:	0.603E-01
Boxes 1 and 3 Tries:	28429	Accepted:	14359	Max Disp.:	0.349E-01
Boxes 2 and 3 Tries:	28056	Accepted:	14214	Max Disp.:	0.331E-01

14000 B: 1	0.2959E+05	0.5279E+05	5430.2	7	16
B: 2	0.1404E+06	0.1300E+06	-10285.8	7	353
B: 3	-0.9405E+06	0.8182E+04	-102268.7	217	0

Block averages were taken calculated after every 28 000 cycles. The averages for the first block are shown below:

Block Averages (BA) for block 1

BA Box: 1 Volume [A^3]	0.66109288E+05
BA Box: 1 Specific density [g/ml]	0.89369983E-01
BA Box: 1 Virial Pressure [kPa]	0.35238882E+04
BA Box: 1 Total Classical	0.61942107E+05
BA Box: 1 Inter vdw	-.18895639E+05
BA Box: 1 Angle	0.37736613E+05
BA Box: 1 Torsion	0.53682946E+05
BA Box: 1 Intra vdw	-.47131615E+04
BA Box: 1 External Field	0.00000000E+00
BA Box: 1 Vibration	0.00000000E+00
BA Box: 1 Coulomb	-.58686522E+04
BA Box: 1 Tail vdw	-.22743594E+04
BA Box: 1 Solvation	0.00000000E+00
BA Box: 1 u (Gibbs Total) [K] Type 1	-.66615535E+04
BA Box: 1 u (Gibbs Total) [K] Type 2	-.89472252E+03
BA Box: 1 Number density [nm-3] Type 1	0.19410954E+00
BA Box: 1 Number density [nm-3] Type 2	0.58394230E+00
BA Box: 1 Mol Fraction Type 1	0.25525204E+00
BA Box: 1 Mol Fraction Type 2	0.74474796E+00
BA Box: 1 Stress Tensor Virial S <sub>xx</sub> [kPa]	0.44054323E+04
BA Box: 1 Stress Tensor Virial S <sub>yy</sub> [kPa]	0.44801484E+04
BA Box: 1 Stress Tensor Virial S <sub>zz</sub> [kPa]	0.44437142E+04
BA Box: 1 Stress Tensor Virial S <sub>xy</sub> [kPa]	0.52732231E+04
BA Box: 1 Stress Tensor Virial S <sub>xz</sub> [kPa]	0.53112387E+04
BA Box: 1 Stress Tensor Virial S <sub>yz</sub> [kPa]	0.52677490E+04
BA Box: 1 Stress Tensor Virial P <sub>tail</sub> [kPa]	-.91921008E+03
BA Box: 1 Radius of Gyration Type: 1	0.32821
BA Box: 1 Radius of Gyration Type: 2	2.09567

```

BA Box: 2 Volume [A^3] 0.11742170E+06
BA Box: 2 Specific density [g/ml] 0.40772154E+00
BA Box: 2 Virial Pressure [kPa] 0.35740909E+04
BA Box: 2 Total Classical 0.12933469E+06
BA Box: 2 Inter vdw -.58859883E+06
BA Box: 2 Angle 0.31661379E+06
BA Box: 2 Torsion 0.44764523E+06
BA Box: 2 Intra vdw -.39692029E+05
BA Box: 2 External Field 0.00000000E+00
BA Box: 2 Vibration 0.00000000E+00
BA Box: 2 Coulomb -.66334643E+04
BA Box: 2 Tail vdw -.79478619E+05
BA Box: 2 Solvation 0.00000000E+00
BA Box: 2 u (Gibbs Total) [K] Type 1 -.66843137E+04
BA Box: 2 u (Gibbs Total) [K] Type 2 -.88827452E+03
BA Box: 2 Number density [nm-3] Type 1 0.12559729E+00
BA Box: 2 Number density [nm-3] Type 2 0.28229141E+01
BA Box: 2 Mol Fraction Type 1 0.42946117E-01
BA Box: 2 Mol Fraction Type 2 0.95705388E+00
BA Box: 2 Stress Tensor Virial S_xx [kPa] 0.22539774E+05
BA Box: 2 Stress Tensor Virial S_yy [kPa] 0.22333695E+05
BA Box: 2 Stress Tensor Virial S_zz [kPa] 0.22386792E+05
BA Box: 2 Stress Tensor Virial S_xy [kPa] 0.19689935E+05
BA Box: 2 Stress Tensor Virial S_xz [kPa] 0.19932077E+05
BA Box: 2 Stress Tensor Virial S_yz [kPa] 0.19529815E+05
BA Box: 2 Stress Tensor Virial P_tail [kPa] -.18845996E+05
BA Box: 2 Radius of Gyration Type: 1 0.32821
BA Box: 2 Radius of Gyration Type: 2 2.09629
BA Box: 3 Volume [A^3] 0.74530525E+04
BA Box: 3 Specific density [g/ml] 0.81780645E+00
BA Box: 3 Virial Pressure [kPa] 0.20749198E+04
BA Box: 3 Total Classical -.88006299E+06
BA Box: 3 Inter vdw 0.13483533E+06
BA Box: 3 Angle 0.24371526E+02
BA Box: 3 Torsion 0.36185876E+02
BA Box: 3 Intra vdw -.29638195E+01
BA Box: 3 External Field 0.00000000E+00
BA Box: 3 Vibration 0.00000000E+00
BA Box: 3 Coulomb -.10149559E+07
BA Box: 3 Tail vdw -.50332405E+04
BA Box: 3 Solvation 0.00000000E+00
BA Box: 3 u (Gibbs Total) [K] Type 1 -.71797356E+04
BA Box: 3 u (Gibbs Total) [K] Type 2 0.92261749E+03
BA Box: 3 Number density [nm-3] Type 1 0.27321551E+02
BA Box: 3 Number density [nm-3] Type 2 0.34168876E-02
BA Box: 3 Mol Fraction Type 1 0.99987074E+00
BA Box: 3 Mol Fraction Type 2 0.12926346E-03
BA Box: 3 Stress Tensor Virial S_xx [kPa] 0.20962007E+05
BA Box: 3 Stress Tensor Virial S_yy [kPa] 0.20373219E+05
BA Box: 3 Stress Tensor Virial S_zz [kPa] 0.20944554E+05
BA Box: 3 Stress Tensor Virial S_xy [kPa] 0.19978573E+06
BA Box: 3 Stress Tensor Virial S_xz [kPa] 0.19975309E+06
BA Box: 3 Stress Tensor Virial S_yz [kPa] 0.20051925E+06
BA Box: 3 Stress Tensor Virial P_tail [kPa] -.18685007E+05
BA Box: 3 Radius of Gyration Type: 1 0.32821
BA Box: 3 Radius of Gyration Type: 2 2.08554

```

### C.3 Final printouts

+++++ end of markov chain +++++

Final hmatrix (general box dimensions)

```

Box: 1
  hmatrix(1,x) 40.49474 0.00000 0.00000
  hmatrix(2,x) 0.00000 40.49474 0.00000
  hmatrix(3,x) 0.00000 0.00000 40.49474
Box: 2
  hmatrix(1,x) 48.93848 0.00000 0.00000
  hmatrix(2,x) 0.00000 48.93848 0.00000
  hmatrix(3,x) 0.00000 0.00000 48.93848
Box: 3
  hmatrix(1,x) 19.46351 0.00000 0.00000
  hmatrix(2,x) 0.00000 19.46351 0.00000
  hmatrix(3,x) 0.00000 0.00000 19.46351

```

\* 3D Volume Change Moves \*

```

Box 1 and 2 Tries: 277340 Accepted: 138992 Acp. Ratio: 0.501 Max Disp.:
0.602E-01
Box 1 and 3 Tries: 284099 Accepted: 142304 Acp. Ratio: 0.501 Max Disp.:
0.352E-01
Box 2 and 3 Tries: 279181 Accepted: 140148 Acp. Ratio: 0.502 Max Disp.:
0.340E-01

```

\* Configurational-Bias SWAP Moves \*

```

Molecule type: 1
  From box 2 to box 1 Attempted: 593713 Grown: 593713 Accepted: 130275
  From box 3 to box 1 Attempted: 610616 Grown: 610616 Accepted: 8057
  From box 1 to box 2 Attempted: 593033 Grown: 593031 Accepted: 130344
  From box 3 to box 2 Attempted: 592842 Grown: 592840 Accepted: 7549
  From box 1 to box 3 Attempted: 610788 Grown: 610708 Accepted: 7984
  From box 2 to box 3 Attempted: 591718 Grown: 591638 Accepted: 7624
Molecule type: 2
  From box 2 to box 1 Attempted: 724365 Grown: 724365 Accepted: 191145
  From box 3 to box 1 Attempted: 745899 Grown: 745899 Accepted: 1284
  From box 1 to box 2 Attempted: 725587 Grown: 725587 Accepted: 191163
  From box 3 to box 2 Attempted: 724054 Grown: 724054 Accepted: 864
  From box 1 to box 3 Attempted: 745548 Grown: 742187 Accepted: 1274
  From box 2 to box 3 Attempted: 726149 Grown: 722921 Accepted: 874

```

\* Configurational-Bias REGROWTH Moves \*

```

Molecule type: 1 Box: 1
  Length Attempts Regrown Accepted %Regrown %Accep.
  1 75338 75338 61385 100.00 81.48
  2 491519 491519 309205 100.00 62.91
Molecule type: 2 Box: 1
  Length Attempts Regrown Accepted %Regrown %Accep.
  5 1172480 1172480 739593 100.00 63.08
Molecule type: 1 Box: 2
  Length Attempts Regrown Accepted %Regrown %Accep.
  1 76853 76853 64655 100.00 84.13
  2 502753 502753 315907 100.00 62.84
Molecule type: 2 Box: 2
  Length Attempts Regrown Accepted %Regrown %Accep.
  5 8282589 8282589 2615709 100.00 31.58
Molecule type: 1 Box: 3
  Length Attempts Regrown Accepted %Regrown %Accep.
  1 1107236 1107236 611899 100.00 55.26

```

2	7196944	7196934	784665	100.00	10.90
Molecule type:	2	Box:	3		
	Length	Attempts	Regrown	Accepted	%Regrown
	5	638	638	75	100.00
					%Accep.
					11.76

## \* COM Translation Moves \*

Molecule:	1	Box:	1	Attempts:	991680.	Accepted:	642872.	Accepted:
					64.827	%		
Molecule:	2	Box:	1	Attempts:	1483778.	Accepted:	966248.	Accepted:
					65.121	%		
Molecule:	1	Box:	2	Attempts:	1015301.	Accepted:	657620.	Accepted:
					64.771	%		
Molecule:	2	Box:	2	Attempts:	10506984.	Accepted:	6833143.	Accepted:
					65.034	%		
Molecule:	1	Box:	3	Attempts:	14556126.	Accepted:	9460839.	Accepted:
					64.996	%		
Molecule:	2	Box:	3	Attempts:	803.	Accepted:	502.	Accepted:
					62.516	%		

## \* Rotation Moves \*

Molecule:	1	Box:	1	Attempts:	831592.	Accepted:	664375.	Accepted:
					79.892	%		
Molecule:	2	Box:	1	Attempts:	1713047.	Accepted:	1372352.	Accepted:
					80.112	%		
Molecule:	1	Box:	2	Attempts:	851250.	Accepted:	678461.	Accepted:
					79.702	%		
Molecule:	2	Box:	2	Attempts:	12141464.	Accepted:	9719195.	Accepted:
					80.050	%		
Molecule:	1	Box:	3	Attempts:	12175831.	Accepted:	9742350.	Accepted:
					80.014	%		
Molecule:	2	Box:	3	Attempts:	862.	Accepted:	678.	Accepted:
					78.654	%		

## Final Energies for Box 1

Total molecules in this box		65	
Molecules of type 1 :		20	
Molecules of type 2 :		45	
total vibration	0.000 [K]	0.00000 [kcal/mol]	
regular	0.000 [K]	0.00000 [kcal/mol]	
bond-bond(1-2)	0.000 [K]	0.00000 [kcal/mol]	
total angle	36745.319 [K]	73.02088 [kcal/mol]	
regular	36745.319 [K]	73.02088 [kcal/mol]	
angle-angle	0.000 [K]	0.00000 [kcal/mol]	
total torsion	57383.269 [K]	114.03294 [kcal/mol]	
regular	57383.269 [K]	114.03294 [kcal/mol]	
improper	0.000 [K]	0.00000 [kcal/mol]	
total nonbond	-36899.516 [K]	-73.32730 [kcal/mol]	
intramolecular	-5802.707 [K]	-11.53123 [kcal/mol]	
2-body nonbond	-28264.675 [K]	-56.16801 [kcal/mol]	
3-body nonbond	0.000 [K]	0.00000 [kcal/mol]	
tail correct.	-2832.134 [K]	-5.62806 [kcal/mol]	
total coulombic	-4254.334 [K]	-8.45428 [kcal/mol]	
real space	-3980.130 [K]	-7.90938 [kcal/mol]	
intramolec.	0.000 [K]	0.00000 [kcal/mol]	
intermolec.	-3980.130 [K]	-7.90938 [kcal/mol]	
self	-280991.908 [K]	-558.39157 [kcal/mol]	
correction	280185.356 [K]	556.78878 [kcal/mol]	
recip sum	532.347 [K]	1.05789 [kcal/mol]	
external field	0.000 [K]	0.00000 [kcal/mol]	
solvation	0.000 [K]	0.00000 [kcal/mol]	

total classical	52974.73788527 [K]	105.2722384064
[kcal/mol]		
Final Energies for Box	2	
Total molecules in this box	334	
Molecules of type 1 :	10	
Molecules of type 2 :	324	
total vibration	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
bond-bond(1-2)	0.000 [K]	0.00000 [kcal/mol]
total angle	321534.762 [K]	638.95897 [kcal/mol]
regular	321534.762 [K]	638.95897 [kcal/mol]
angle-angle	0.000 [K]	0.00000 [kcal/mol]
total torsion	431934.113 [K]	858.34631 [kcal/mol]
regular	431934.113 [K]	858.34631 [kcal/mol]
improper	0.000 [K]	0.00000 [kcal/mol]
total nonbond	-602101.252 [K]	-1196.50515 [kcal/mol]
intramolecular	-38274.386 [K]	-76.05947 [kcal/mol]
2-body nonbond	-487486.965 [K]	-968.74182 [kcal/mol]
3-body nonbond	0.000 [K]	0.00000 [kcal/mol]
tail correct.	-76339.901 [K]	-151.70386 [kcal/mol]
total coulombic	-714.114 [K]	-1.41910 [kcal/mol]
real space	-746.377 [K]	-1.48321 [kcal/mol]
intramolec.	0.000 [K]	0.00000 [kcal/mol]
intermolec.	-746.377 [K]	-1.48321 [kcal/mol]
self	-116255.088 [K]	-231.02395 [kcal/mol]
correction	116027.526 [K]	230.57174 [kcal/mol]
recip sum	259.825 [K]	0.51633 [kcal/mol]
external field	0.000 [K]	0.00000 [kcal/mol]
solvation	0.000 [K]	0.00000 [kcal/mol]
total classical	150653.50958014 [K]	299.3810410470
[kcal/mol]		
Final Energies for Box	3	
Total molecules in this box	201	
Molecules of type 1 :	201	
Molecules of type 2 :	0	
total vibration	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
bond-bond(1-2)	0.000 [K]	0.00000 [kcal/mol]
total angle	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
angle-angle	0.000 [K]	0.00000 [kcal/mol]
total torsion	0.000 [K]	0.00000 [kcal/mol]
regular	0.000 [K]	0.00000 [kcal/mol]
improper	0.000 [K]	0.00000 [kcal/mol]
total nonbond	142490.978 [K]	283.16033 [kcal/mol]
intramolecular	0.000 [K]	0.00000 [kcal/mol]
2-body nonbond	147442.486 [K]	293.00004 [kcal/mol]
3-body nonbond	0.000 [K]	0.00000 [kcal/mol]
tail correct.	-4951.508 [K]	-9.83972 [kcal/mol]
total coulombic	-1024585.149 [K]	-2036.07184 [kcal/mol]
real space	-957871.661 [K]	-1903.49774 [kcal/mol]
intramolec.	0.000 [K]	0.00000 [kcal/mol]
intermolec.	-957871.661 [K]	-1903.49774 [kcal/mol]
self	-5875400.536 [K]	-11675.68903 [kcal/mol]
correction	5799544.344 [K]	11524.94640 [kcal/mol]
recip sum	9142.704 [K]	18.16853 [kcal/mol]
external field	0.000 [K]	0.00000 [kcal/mol]

solvation		0.000 [K]		0.00000 [kcal/mol]	
total classical		-882094.17083708 [K]		-1752.9115113394	
[kcal/mol]					
Averages	Units	Type	Box 1	Box 2	Box 3
Volume	nm <sup>3</sup>		0.70040E+02	0.11354E+03	0.74071E+01
Molecule Number		1	13.844	14.169	202.987
Molecule Number		2	45.675	323.301	0.024
Molar Volume	ml/mol		0.71924E+03	0.20183E+03	0.21946E+02
Specific Density	g/ml		0.09744633	0.41292673	0.82121825
Number Density	nm <sup>-3</sup>	1	0.19764	0.12411	27.43676
Number Density	nm <sup>-3</sup>	2	0.63964	2.85960	0.00317
Mole Fraction		1	0.2428917	0.0417471	0.9998806
Mole Fraction		2	0.7571083	0.9582529	0.0001194
Radius of Gyration	A	1	0.3282120	0.3282120	0.3282120
Radius of Gyration	A	2	2.0955485	2.0963818	2.0798137
Virial Pressure	kPa		0.36632E+04	0.36131E+04	0.22365E+04
Virial S <sub>xx</sub>	kPa		0.47447E+04	0.22853E+05	0.22127E+05
Virial S <sub>yy</sub>	kPa		0.47793E+04	0.23033E+05	0.21024E+05
Virial S <sub>zz</sub>	kPa		0.47581E+04	0.22920E+05	0.20074E+05
Virial S <sub>xy</sub>	kPa		0.56682E+04	0.20075E+05	0.19831E+06
Virial S <sub>xz</sub>	kPa		0.56915E+04	0.20013E+05	0.19927E+06
Virial S <sub>yz</sub>	kPa		0.56847E+04	0.19926E+05	0.19982E+06
Virial P <sub>tail</sub>	kPa		-.10975E+04	-.19322E+05	-.18838E+05
Virial p <sub>i</sub> <x <sub>i</sub> ><p <sub>v</sub> >	kPa	1	0.88976E+03	0.15084E+03	0.22363E+04
Virial p <sub>i</sub> <x <sub>i</sub> ><p <sub>v</sub> >	kPa	2	0.27734E+04	0.34623E+04	0.26699E+00
Ideal Pressure	kPa		0.55853E+04	0.19904E+05	0.18305E+06
Ideal p <sub>i</sub> <N/V>kT	kPa	1	0.13184E+04	0.82790E+03	0.18303E+06
Ideal p <sub>i</sub> <N/V>kT	kPa	2	0.42669E+04	0.19076E+05	0.21175E+02
-<dU/dV>	kPa		-.19581E+04	-.16102E+05	-.17655E+06
Thermodynamic Pressure	kPa		0.36272E+04	0.38014E+04	0.64940E+04
Thermo p <sub>i</sub> <x <sub>1</sub> ><p <sub>t</sub> >	kPa	1	0.88102E+03	0.15870E+03	0.64932E+04
Thermo p <sub>i</sub> <x <sub>1</sub> ><p <sub>t</sub> >	kPa	2	0.27462E+04	0.36427E+04	0.77522E+00
Total Classical	K		0.6978E+05	0.1201E+06	-0.8792E+06
Inter vdw	K		-0.2429E+05	-0.5841E+06	0.1345E+06
Angle	K		0.4381E+05	0.3106E+06	0.2320E+02
Torsion	K		0.6223E+05	0.4387E+06	0.3361E+02
Intra vdw	K		-0.5491E+04	-0.3891E+05	-0.2875E+01
External Field	K		0.0000E+00	0.0000E+00	0.0000E+00
Vibration	K		0.0000E+00	0.0000E+00	0.0000E+00
Coulomb	K		-0.6483E+04	-0.6227E+04	-0.1014E+07
Tail vdw	K		-0.2890E+04	-0.7897E+05	-0.5045E+04
Solvation	K		0.0000E+00	0.0000E+00	0.0000E+00
u (Density)	K	1	-6548.023	-6770.173	-4165.824
u (NVT Insertion)	K	1	-118.364	107.422	-2475.009
u (NpT Insertion)	K	1	-113.591	105.537	-2470.636
u (Den. + NVT Insert)	K	1	-6666.387	-6662.752	-6640.833
u (Den. + NpT Insert)	K	1	-6661.614	-6664.636	-6636.460
u (Gibbs Total)	K	1	-6665.683	-6668.134	-6639.089
u (Density)	K	2	-7105.499	-6393.538	-9664.765
u (NVT Insertion)	K	2	6231.890	5510.218	8794.210
u (NpT Insertion)	K	2	6224.994	5510.507	8782.623
u (Den. + NVT Insert)	K	2	-873.610	-883.320	-870.555
u (Den. + NpT Insert)	K	2	-880.506	-883.031	-882.142
u (Gibbs Total)	K	2	-882.114	-881.311	935.108
G: Sum{<u <sub>i</sub> ><N <sub>i</sub> >}	kJ/mol		-0.1102E+04	-0.3154E+04	-0.1120E+05
U	kJ/mol		0.5802E+03	0.9985E+03	-0.7310E+04
pV: <p><V>	kJ/mol		0.1545E+03	0.2470E+03	0.9973E+01

H: <U> + <p><V>	kJ/mol	0.7346E+03	0.1245E+04	-0.7300E+04
H: <U + pV>	kJ/mol	0.7355E+03	0.1241E+04	-0.7304E+04
S: (<H> - <G>)/T	kJ/K mol	0.3803E+01	0.9109E+01	0.8083E+01
Z: <p><V>/<N>RT		0.646203	0.182225	0.012233

Block Averages (5 blocks)	Units	Type	Box	Average	Standard Deviation
Specific Density	g/ml		1	0.97446E-01	0.47975E-02
Specific Density	g/ml		2	0.41293E+00	0.39006E-02
Specific Density	g/ml		3	0.82122E+00	0.33363E-02
Virial Pressure	kPa		1	0.36632E+04	0.80471E+02
Virial S_xx	kPa		1	0.47447E+04	0.17455E+03
Virial S_yy	kPa		1	0.47793E+04	0.17802E+03
Virial S_zz	kPa		1	0.47581E+04	0.17425E+03
Virial S_xy	kPa		1	0.56682E+04	0.23207E+03
Virial S_xz	kPa		1	0.56915E+04	0.22973E+03
Virial S_yz	kPa		1	0.56847E+04	0.25155E+03
Virial P_tail	kPa		1	-0.10975E+04	0.10917E+03
Virial Pressure	kPa		2	0.36131E+04	0.72650E+02
Virial S_xx	kPa		2	0.22853E+05	0.48465E+03
Virial S_yy	kPa		2	0.23033E+05	0.43225E+03
Virial S_zz	kPa		2	0.22920E+05	0.37106E+03
Virial S_xy	kPa		2	0.20075E+05	0.24854E+03
Virial S_xz	kPa		2	0.20013E+05	0.20965E+03
Virial S_yz	kPa		2	0.19926E+05	0.23951E+03
Virial P_tail	kPa		2	-0.19322E+05	0.35807E+03
Virial Pressure	kPa		3	0.22365E+04	0.39016E+03
Virial S_xx	kPa		3	0.22127E+05	0.24562E+04
Virial S_yy	kPa		3	0.21024E+05	0.22899E+04
Virial S_zz	kPa		3	0.20074E+05	0.19816E+04
Virial S_xy	kPa		3	0.19831E+06	0.19403E+04
Virial S_xz	kPa		3	0.19927E+06	0.11727E+04
Virial S_yz	kPa		3	0.19982E+06	0.15799E+04
Virial P_tail	kPa		3	-0.18838E+05	0.14714E+03
Thermodynamic Pressure	kPa		1	0.36273E+04	0.76944E+02
Thermodynamic Pressure	kPa		2	0.38013E+04	0.93927E+02
Thermodynamic Pressure	kPa		3	0.64933E+04	0.24528E+03
Total Classical	K		1	0.69783E+05	0.46579E+04
Inter vdw	K		1	-0.24288E+05	0.31376E+04
Angle	K		1	0.43810E+05	0.35191E+04
Torsion	K		1	0.62234E+05	0.49196E+04
Intra vdw	K		1	-0.54908E+04	0.44509E+03
External Field	K		1	0.00000E+00	0.00000E+00
Vibration	K		1	0.00000E+00	0.00000E+00
Coulomb	K		1	-0.64827E+04	0.47281E+03
Tail vdw	K		1	-0.28903E+04	0.36121E+03
Solvation	K		1	0.00000E+00	0.00000E+00
Total Classical	K		2	0.12010E+06	0.61269E+04
Inter vdw	K		2	-0.58409E+06	0.46148E+04
Angle	K		2	0.31061E+06	0.33981E+04
Torsion	K		2	0.43871E+06	0.51082E+04
Intra vdw	K		2	-0.38907E+05	0.44478E+03
External Field	K		2	0.00000E+00	0.00000E+00
Vibration	K		2	0.00000E+00	0.00000E+00
Coulomb	K		2	-0.62268E+04	0.26892E+03
Tail vdw	K		2	-0.78965E+05	0.63096E+03
Solvation	K		2	0.00000E+00	0.00000E+00
Total Classical	K		3	-0.87922E+06	0.45576E+04

Inter vdw	K	3	0.13446E+06	0.11350E+04
Angle	K	3	0.23197E+02	0.76853E+01
Torsion	K	3	0.33611E+02	0.11627E+02
Intra vdw	K	3	-0.28750E+01	0.10483E+01
External Field	K	3	0.00000E+00	0.00000E+00
Vibration	K	3	0.00000E+00	0.00000E+00
Coulomb	K	3	-0.10137E+07	0.56636E+04
Tail vdw	K	3	-0.50450E+04	0.17621E+02
Solvation	K	3	0.00000E+00	0.00000E+00
u (Gibbs Total)	K	1 1	-6664.228	36.759
u (Gibbs Total)	K	1 2	-6668.019	10.798
u (Gibbs Total)	K	1 3	-6443.531	398.312
u (Gibbs Total)	K	2 1	-882.066	6.673
u (Gibbs Total)	K	2 2	-881.297	3.610
u (Gibbs Total)	K	2 3	1006.109	304.285
u (NpT Insertion)	K	1 1	-112.505	34.917
u (NpT Insertion)	K	1 2	105.981	16.406
u (NpT Insertion)	K	1 3	-2277.931	395.064
u (NpT Insertion)	K	2 1	6225.868	18.839
u (NpT Insertion)	K	2 2	5510.485	3.456
u (NpT Insertion)	K	2 3	8855.070	307.128
u (NVT Insertion)	K	1 1	-116.737	39.085
u (NVT Insertion)	K	1 2	107.713	17.095
u (NVT Insertion)	K	1 3	-2277.694	399.782
u (NVT Insertion)	K	2 1	6232.258	18.819
u (NVT Insertion)	K	2 2	5510.231	3.523
u (NVT Insertion)	K	2 3	8867.741	310.693
Number Density	nm-3	1 1	0.19764E+00	0.80142E-02
Number Density	nm-3	1 2	0.12411E+00	0.39623E-02
Number Density	nm-3	1 3	0.27437E+02	0.11480E+00
Number Density	nm-3	2 1	0.63964E+00	0.33174E-01
Number Density	nm-3	2 2	0.28596E+01	0.27523E-01
Number Density	nm-3	2 3	0.31742E-02	0.10705E-02
Mole Fraction		1 1	0.2428917	0.0103509
Mole Fraction		1 2	0.0417471	0.0014701
Mole Fraction		1 3	0.9998806	0.0000403
Mole Fraction		2 1	0.7571083	0.0103509
Mole Fraction		2 2	0.9582529	0.0014701
Mole Fraction		2 3	0.0001194	0.0000403
Molarity	M	1 1	0.32830E+00	0.13313E-01
Molarity	M	1 2	0.20616E+00	0.65819E-02
Molarity	M	1 3	0.45576E+02	0.19071E+00
Molarity	M	2 1	0.10625E+01	0.55106E-01
Molarity	M	2 2	0.47502E+01	0.45718E-01
Molarity	M	2 3	0.52728E-02	0.17783E-02
Radius of Gyration	A	1 1	0.32821	0.00000
Radius of Gyration	A	1 2	0.32821	0.00000
Radius of Gyration	A	1 3	0.32821	0.00000
Radius of Gyration	A	2 1	2.09555	0.00021
Radius of Gyration	A	2 2	2.09638	0.00020
Radius of Gyration	A	2 3	2.08175	0.00653

-----block averages -----

Box: 1

Block Energy	Density	Virial Press.	Mol frags
1 0.61942107E+05	0.89369983E-01	0.35238882E+04	0.25525204 0.74474796
2 0.67507182E+05	0.95672216E-01	0.36528888E+04	0.23963390 0.76036610

```
3 0.75020024E+05 0.10377270E+00 0.36578969E+04 0.22773154 0.77226846
4 0.73340158E+05 0.98723117E-01 0.37199679E+04 0.23804220 0.76195780
5 0.71103183E+05 0.99693644E-01 0.37612480E+04 0.25379901 0.74620099
Box:      2
Block Energy      Density      Virial Press.  Mol fracs
1 0.12933469E+06 0.40772154E+00 0.35740909E+04 0.04294612 0.95705388
2 0.12385573E+06 0.41052175E+00 0.35391594E+04 0.04134731 0.95865269
3 0.11893838E+06 0.41253580E+00 0.35612898E+04 0.04125681 0.95874319
4 0.11126945E+06 0.41927620E+00 0.36561529E+04 0.03948080 0.96051920
5 0.11710338E+06 0.41457836E+00 0.37350143E+04 0.04370453 0.95629547
Box:      3
Block Energy      Density      Virial Press.  Mol fracs
1 -.88006299E+06 0.81780645E+00 0.20749198E+04 0.99987074 0.00012926
2 -.88072047E+06 0.81804136E+00 0.30122064E+04 0.99985110 0.00014890
3 -.88140730E+06 0.81981531E+00 0.19975450E+04 0.99985474 0.00014526
4 -.88349740E+06 0.82531069E+00 0.21033966E+04 0.99995992 0.00004008
5 -.87039878E+06 0.82511744E+00 0.19946574E+04 0.99986662 0.00013338

real 19851m4.599s
user 19808m31.349s
sys  0m7.414s
Source code : /home/erj/SUREN/towhee-5.2.3/Source
Wed Mar 26 04:07:23 CET 2008
```

The “real 19851m4.599s” string shown above refers to the total running time of this particular simulation, i.e. ~ 14 days.