UNIVERSITY OF KWAZULU-NATAL

Molecular Simulation of Vapour-Liquid Equilibrium Using Beowulf Clusters

p 191 missing

T.J. McKnight

[B.Sc. (Eng)]

March, 2006

Molecular Simulation of Vapour-Liquid Equilibrium Using Beowulf Clusters

Tyrone J. McKnight

A thesis submitted in the
School of Chemical Engineering
University of KwaZulu-Natal
Durban

(All work performed in the
School of Chemical Engineering
University of Natal, Durban
South Africa
and
Högskolan i Borås,
Sweden)

In fulfillment of the requirements for the degree of
Doctor of Philosophy

March, 2006

DECLARATION

The work presented in this thesis was performed at the University of KwaZulu-Natal, Durban and at the Högskolan i Borås, Sweden from January 2001 to December 2005. The work was supervised by Professor D. Ramjugernath, and Associate Professor K Bolton

This thesis is presented as the full requirement for the degree of Ph.D. in Chemical Engineering. All the work presented in this thesis is original unless otherwise stated and has not (in whole or part) been submitted previously to any tertiary institute as part of a degree.

T.J. McKnight

I hereby certify that I find this work to be suitable for submission for the degree of Ph.D. in Chemical Engineering.

D. Ramjugernath

ACKNOWLEDGEMENTS

I would like to acknowledge the following people for their contribution to this work:

- Firstly, my direct supervisors Deresh Ramjugernath and Kim Bolton for their advice and assistance during this project. I would especially like to thank both Deresh and Kim for the work they put in behind the scenes which ensured that the project collaboration was successful. I would also like acknowledge my co-supervisor Mathew Starzak for several very useful discussions particularly at the beginning of the project which I felt guided the work towards a more fundamental approach.

- I must extend a massive debt of gratitude to my friends Anders Askåsen and Daniel Johansson for their outstanding IT support, guidance and genuine friendship. I learnt so much from you guys and had an absolute blast the whole time. I look back with fondness and absolute respect for two outstanding friends whom I feel thoroughly privileged and lucky to have worked with. I think that you two are large part of what I am handing in today and I will always appreciate your kind encouragements and, most of all, your patience.

- On a personal note I would also like to thank Kim Bolton and Peter Ahlström for their friendship and support particularly when I visited Sweden in September 2001. Kim, thank you for all the effort you put in behind the scenes to make that trip and the ones which followed fruitful and enjoyable. Peter, thank you for inviting me along to the Beowulf Conference in Linköping in 2001 as well. In the months which followed I was able to draw extensively on what I had learnt. I also can't thank you enough for having been the one who contacted Thijs and got permission to use the BIGMAC program code. I don't think that the work relationship would have been anything like what we enjoyed had it not been for you inputs into the project from Sweden.

- I would like to extend a special word of appreciation to Thijs Vlugt for allowing the project to make use of his serial and parallel versions of the BIGMAC program code. During the course of the work, several revisions were made to the basic program code but the BIGMAC code served as the reliable foundation upon which essentially all of the following work could be built. Not only was the code a pleasure to work with but Thijs's advice in the second year of the project provided invaluable insights as to where the project should go.

- On several levels I want to thank Roger and Michela Harris for the many wonderful hours we spent as friends and colleagues together. Michela, thank you for all the kind words and support which you gave while I was studying. Roger, I really appreciated your friendship when we started and the awesome advice and guidance which you so richly gave me. Post-grad would have been a lot harder and a lot more boring without you. I missed our late night conversations most of all after you left. I think back with amazement at the kindness both of you extended so regularly and so selflessly. It made both my university experience and my work more enjoyable and more productive.

- Thank you Jenny and Sarushen for the many hours of counsel and advice that you gave me. Jen, you and I have been in this boat for a good many years now and, looking back, I wouldn't have wanted to have done it with anyone else. Thank you for the discussions, thank you for the bananas, thank you for the lifts and thank you most of all for a being a genuinely outstanding friend from whom I could find a rich oasis of wisdom.

- Thank you Nick for the in-depth discussions which we had in 2004. I don't know if I said anything useful, probably didn't, but I was so grateful to be able to bounce those crazy ideas off a sharp wit and keen mind over those filter coffees late at night. I have missed our conversations but I look back with sincere appreciation on every minute.

- Financially, I would like to make a sincere point of thanking the SASOL, the NRF and SIDA for Financial support during my degree.

- My office colleagues, Kasuren, Prathika, Scott, Minal, John-Roy, Warren , Myriam, Alex, Viran and Yash in the Thermodynamics Research Unit for their ideas, our laughs and most of all your support and superb friendship.

- On a most personal note my parents. Mom and dad, thank you for all the times you politely listened when I spoke so vaguely about what I thought I knew. Thank you for your patience, but most of all, thank you for always being the encouragement that I needed in the dark times when nothing seemed to be going right. In large measure, everything that might be worthwhile in this work I owe to you both.

ABSTRACT

This work describes the installation of a Beowulf cluster at the University of KwaZulu-Natal and the subsequent Monte Carlo molecular simulation study which was undertaken to investigate the pure component and binary mixture phase equilibrium properties for systems of industrial relevance. The work was undertaken both at the University of KwaZulu-Natal in South Africa and at the Högskolan i Borås in Sweden.

To conduct the molecular simulation work, a 20-node Beowulf cluster called yoda was assembled and installed at the University of KwaZulu-Natal. The project saw the installation of a Linux operating system as well as the necessary PBS queuing software, the SSH remote login software and the BIGMAC molecular simulation program code.

The first objective of this work was to use the BIGMAC program code to study hydrocarbon mixtures for industrially relevant systems. To this end, five different *n*-alkane and 1-alkene mixtures were simulated using the Spyriouni *et al.* (1999) olefin force field as well as the NERD and TraPPE force fields. These mixtures included an ethane + propene binary at 277.6K, a 1-hexene + n-octane binary at 328.15K, an n-dodecane + 1-octadecene binary at 760mmHg, a propene + 1-butene binary at 294.3K and a 1-butene + 1-hexene mixture at 373K. It was shown that the NERD and TraPPE force fields yield P-x-y data sets that are, in general, shifted to higher pressures due to their overestimation of the pure component vapour pressures but that the x-y data is accurately reproduced for all three force fields used.

Building on the results from the binary simulations, the NERD and TraPPE force field parameters were extended to the simulation of conjugated alkenes (1,3-butadiene; 1,3-pentadiene and 1,3,5-hexatriene) to examine their ability to simulate molecules for which they were not explicitly parameterized. In this part of the work, it was shown that the effects of intramolecular equilibrium geometries can have a significant impact on the predicted thermophysical properties. While the results for 1,3-butadiene were very successful, the results for the remaining conjugated alkenes and the results from the TraPPE force field simulations were more disappointing. This work suggested that a reparameterization of the united-atom force field parameters was necessary.

Another issue which emerged from the simulation of the five initial binary mixtures was the difficulty which was experienced when trying to use the N_1N_2PT Gibbs Ensemble to simulate

narrow phase envelope regions. As a result, a new methodology was developed in terms of which new variants of isothermal Gibbs Ensembles can be developed. It was shown how novel Gibbs Ensemble partition functions and their corresponding thermodynamic potentials can all emerge from a single governing equation which is a special case of the generalized multiphase μPT ensemble. In terms of this analysis, the novel μ_1N_2VT and ξ_1NVT Gibbs Ensembles were developed, characterised and tested. It is shown how these two new ensembles reliably overcome the mass balance constraints inherent in the traditional N_1N_2PT Gibbs Ensemble. These novel Gibbs Ensembles were used to study two close-boiling mixtures and, to the knowledge of this work, this represents the first time that such close-boiling mixtures have been directly studied using a Gibbs Ensemble.

Finally, a new force field for *n*-alkanes and alkenes called the TBORG (Transferability Based On Relaxing Geometries) force field was developed. This new force field represents a large regression effort which sought to improve the intermolecular parameters for *n*-alkanes and alkenes by allowing for a larger and more versatile intramolecular parameter set while still retaining the conceptual appeal of transferable intermolecular parameters. The *n*-alkane parameters were highly successful, although the parameters for alkenes gave less satisfactory results. Both sets of parameters, however, achieved particular success in improving the estimation of pure component vapour pressures and vapour densities for carbons in the range of C_2 to C_8 without sacrificing the accurate prediction of liquid densities, critical temperatures and critical densities. The regressed parameters point towards the need for a wider intramolecular parameter set to be used in conjunction with anisotropic parameters in order to improve the predictions of transferable united-atom force field models.

TABLE OF CONTENTS

"I don't understand why people are frightened of new ideas. I am afraid of the old ones." – John Cage

LIST OF FIGURES

"I have not failed. I've just found 10,000 ways that won't work." – Thomas Alva Edison

LIST OF TABLES

"Whenever you find yourself on the side of the majority it is time to pause and reflect." – Mark Twain

NOMENCLATURE

"The devil is in the details" – German proverb

$acc(o \to n)$	probability of accepting a move from old state o to new state n
A	constant for the law of rectilinear diameters
$\mathcal{A}$	macroscopic variable
A	Helmholtz free energy, $A = U - TS$
b	vector co-ordinates of a periodic image, $\mathbf{b} = (b_x, b_y, b_z)$
b_x, b_y, b_z	x-, y- and z-components of a periodic image
B	constant for the density scaling law used to regress the critical temperature
$\{\mathbf{c}\}_k$	set of k trial orientations for Configurational-Bias Monte Carlo moves
$c_{0,1,2,...}$	cosine series torsional constants
C	number of components
DOF	number of intensive variable degrees of freedom which need to be specified
$d_{COM}(\mathrm{A})$	largest distance between the interaction site furthest from the centre of mass of molecule A
f_i	fugacity of species i
$f(\mathbf{N},\mathbf{V})$	compounded function of **N** and **V** for the generalized $Q_{\mu PT}$ ensemble
F	number of phases
$\mathbf{F}_{ij}$	pair force exerted on molecule i by molecule j
F_{obj}	objective function
$\mathcal{F}_{\tilde{A}}$	"Massieu" function in the entropy representation
$g_2(r_{ij})$	pair distribution function
$g(\mathbf{N})$	compounded function of **N** for the generalized $Q_{\mu PT}$ ensemble
G	Gibbs free energy $G = U - TS + PV$
h	Planck's constant
$\mathcal{H}$	Hamiltonian
H	enthalpy, $H = U + PV$
ΔH_{vap}	latent heat of vaporization

I,II,III	index to denote distinct phases
J	simulated property used for regression of intermolecular parameters
k_0	number of trial positions for the first segment placed used during a Configurational-Bias Monte Carlo move
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_s	bond stretching constant
k_θ	bond angle bending constant
k_ϕ	harmonic torsional constant
$\mathcal{K}$	kinetic (ideal) contribution to the energy
K_0	overall constant used to non-dimensionalize the partition function
K	arbitrary constant
l	bond length
ℓ	number of segments to be grown during a Configurational-Bias Monte Carlo move
L	Hill energy ($L = -PV$)
L_b	length of a cubic simulation box
m	molecule mass
$M_{\bar{x}}$	number of extensive parameter specifications replaced by intensive parameter specifications in the Legendre transform
$M_{\bar{y}}$	number of intensive parameter specifications replacing extensive parameter specifications in the Legendre transform
n_i	fluctuating number of molecules in phase I or of type i
N	number of molecules
$N_{\{int\}}$	number of interaction sites
N	denotes the set of all number of molecules in each phase, $\mathbf{N} = \{n_I, n_{II}, ..., n_F\}$
p	molecular momentum vector
P	pressure
$\mathcal{P}$	denotes a probability density function
$\mathcal{P}(\mathbf{r}^N)_{ens}$	probability of observing the system in a configuration around $\mathbf{r}^N$ for an ensemble

$\wp(\Gamma)$	phase space probability density
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 , $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$
$\hat{\mathbf{r}}_x, \hat{\mathbf{r}}_z, \hat{\mathbf{r}}_z$	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 , $r_{ij} = \lvert\mathbf{r}_{ij}\rvert$
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
R_G	Universal Gas Constant
$\mathcal{R}_{[0,1]}$	random number from a uniform distribution in the interval [0,1]
s	standard deviation
$\mathbf{s}$	dimensionless molecular or atomic position vector, $\mathbf{s} = \mathbf{r}/L$
S	entropy
$\hat{S}$	instantaneous value of the entropy
t	time
T	absolute temperature
T_r	reduced temperature, $T_r = T/T_{cr}$
U	potential (excess) contribution to the energy. Equivalent to internal energy
$\bar{U}^{ext}$	approximate intermolecular potential for DC-CBMC algorithm
δU^{ext}	difference between the full and approximate potentials for the DC-CBMC algorithm
$\Delta U^{+}_{I,i}$	energy change of inserting a molecule of type i into phase I for the evaluation of the chemical potential
v_j	regression parameter "j"
$\mathbf{v}_j$	set of all regression parameters
V	volume of a simulation box
$\mathbf{V}$	denotes the set of volumes for each of the phases, $\mathbf{V} = \{V_I, V_{II}, ..., V_F\}$

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 viral function for pairwise additive functions
$\mathcal{W}$	total molecular viral
$w(\Gamma)$	weight function
w_m^{new}	Rosenbluth factor of segment m for the new chain
W^{new}	total Rosenbluth weight for the new chain
$\bar{W}^{ext}$	approximate Rosenbluth weight for the new chain for the DC-CBMC algorithm
x_i	mole fraction of component i in the liquid
$\mathbf{X}$	set of extensive quantities in the entropy representation of the fundamental relation
$\bar{\mathbf{X}}$	set of fixed extensive quantities for a simulation
$\tilde{\mathbf{X}}$	set of variable extensive quantities for a simulation
X_{iI}	extensive variable i in phase I of the entropy representation of the fundamental relation
$\mathbf{X}'$	set of extensive variables in the energy representation of the fundamental relation
$\hat{\mathbf{X}}$	set of the instantaneous values of the set of variable extensive quantities $\tilde{\mathbf{X}}$
y_i	mole fraction of component i in the vapour
$\mathbf{Y}$	set of intensive quantities in the energy representation of the fundamental relation
$\bar{\mathbf{Y}}$	set of intensive quantities which have a conjugate extensive quantity which is fixed for the simulation
$\tilde{\mathbf{Y}}$	set of intensive quantities which have a conjugate extensive quantity which is variable for the simulation
Y_{iI}	intensive variable i in phase I
z_i	overall mole fraction of component i across all phases
Z_{RA}	modified compressibility fact or for the modified Rackett equation

Greek Symbols

$\alpha(o \to n)$	probability of generating a new state n from an old state o
β	inverse temperature, $\beta = 1/k_B T$

β	universal scaling exponent for the density scaling law to regress the critical temperature
γ_i	activity coefficient of component i in the liquid
δ_{ij}	anisotropic parameter for AUA force fields
ε_{ij}	Lennard-Jones energy parameter between sites i and j
η	total number of measured simulation properties used for parameter regression
θ	bond angle
λ	number of force field parameters to be regressed
Λ	thermal de Broglie wavelength
μ	chemical potential
$\mu_i^0(T)$	reference chemical potential of species i in the ideal gas state
ξ_i	fugacity fraction of component i
π	pi
$\pi(o \to n)$	transition probability of going from an old state o to a new state n
ρ_i	density of component i or the density of phase i
ς_i	number of points generated per unit volume around r^N
σ_{ij}	Lennard-Jones size parameter between sites i and j
$\Upsilon(t)$	point in phase space at time t
Υ_o	old point in phase space
Υ_n	new point in phase space
ϕ	torsional angle
$\hat{\phi}_i$	fugacity coefficient of component i in solution in the vapour phase
Φ_i	measures deviations from ideal behaviour in the vapour, $\left(\Phi_i = \hat{\phi}_i P_i^{sat} / f_i\right)$
Ψ_{ens}	ensemble thermodynamic potential
Ω	variable equivalent to the entropy, S/k_B

Brackets and Operators

$\langle ... \rangle$	ensemble average

∇_r	gradient vector operator

Subscripts

0	an equilibrium bond length l_0, an equilibrium bond angle θ_0 for a harmonic bond angle potential or an equilibrium torsional angle ϕ_0 for a harmonic torsional potential
a,b	interaction sites
bend	contribution from bond angle bending
Be6	Buckingham exponential-6 potential
cr	critical property for the density ρ, temperature T or pressure P
ens	ensemble average
exp	value measured experimentally
g	gas or vapour phase
h	arbitrary indexing variable
i,j	species type
iso	isothermal ensemble
I,II,III	phase label
l	liquid phase
low	lower integration limit
LJ	Lennard-Jones potential
m	next segment being placed during a Configurational-Bias Monte Carlo move
obs	observed variable
stretch	contribution from bond stretching
t	trial direction when growing molecule during a Configurational-Bias Monte Carlo move
thm	thermodynamic definition of the pressure
time	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
α, β	interaction sites

ϕ	torsional angle

Superscripts

0	property at standard / reference conditions
C	short-range contribution to a calculated property below the cut-off radius, r_C
ex	excess property or excess contribution
ext	external energy from all the intermolecular interactions
GE	partition function, probability density or thermodynamic potential for a "Gibbs-type" of ensemble
id	ideal property or ideal contribution
int	internal energy from all the intramolecular interactions
LR	long-range contribution to a calculated property
new	new state
old	old state
RG	partition function, probability density or thermodynamic potential for the Reaction Gibbs Ensemble
sat	property at saturated conditions
SG	partition function, probability density or thermodynamic potential for the Semi-Grand Ensemble
T	total value of an extensive property for all components in all phases

Abbreviations

AUA	anisotropic united-atom approximation
AUA4	anisotropic united-atom force fields of Ungerer *et al.* (2000) and Bourasseau *et al.* (2003)
Be6	Buckingham exponential-6 potential
COM	centre of mass
CBMC	Configurational-Bias Monte Carlo
DC-CBMC	Dual Cut-Off Configurational-Bias Monte Carlo
EH	Explicit Hydrogen approximation
GEMC	Gibbs Ensemble Monte Carlo
LJ	Lennard-Jones potential

nAEP	Buckingham exponential-6 *n*-alkane potential of Errington and Panagiotopoulos (1999a)
NPT	constant number of molecules, constant pressure and constant temperature ensemble
N_1N_2PT	constant number of molecules of species 1, constant number of molecules of species 2, constant pressure and constant temperature Gibbs Ensemble
NVT	constant number of molecules, constant volume and constant temperature ensemble
RGEMC	Reaction Gibbs Ensemble Monte Carlo
SA1	Spyriouni *et al.* olefin force field (Spyriouni *et al.* (1999))
SG	Semi-Grand Ensemble
TBORG	Transferability Based On Relaxing Geometries
UA	(isotropic) united-atom approximation
VLE	vapour-liquid equilibrium
μVT	constant chemical potential, constant volume and constant temperature ensemble
μPT	hypothetical constant chemical potential, constant volume and constant temperature 'ensemble'
μ_1N_2VT	constant chemical potential of species 1, constant number of molecules of species 2, constant volume and constant temperature Gibbs Ensemble
ξ_1NVT	constant fugacity fraction of species 1, constant total number of molecules, constant volume and constant temperature Gibbs Ensemble

CHAPTER 1:
INTRODUCTION

"If one doesn't know to which port one is sailing, no wind is favourable" – Lucius Annaeus Seneca

Separation operations are ubiquitous in the chemical industry. Separation operations of one form or another are found in the purification of chemicals or in the removal of wastes from process streams. Most of these separation processes involve the formation of a new phase or the addition of a new phase. Consequently, knowledge of phase equilibria is crucial to the design of efficient separation strategies. The number of known chemical compounds exceeds six million and certain estimates project that this number may increase at a rate of 100,000 new chemical compounds per year. Clearly this creates a data gap between the synthesised chemical of the organic chemist and the measured pure component property data of the physical chemist. The situation becomes even worse for binary mixtures where the number of unique binary combinations roughly increases proportional to the square of the number of pure components and where nearly all of the mixtures of interest are usually complicated functions of temperature (and pressure).

With the rapid advances in computer hardware and clustering technologies in recent years, however, molecular modelling and molecular simulation techniques have emerged as a powerful means to predict phase equilibria and to explore the link between microscopic chemical behaviour and macroscopic properties. The ability to simulate different phase diagrams could in the near future, therefore, become viable for the design of separation processes and crucial to uncovering the underlying phenomenological factors at work.

Ungerer (2003) has suggested that the greatest current challenges facing the molecular simulation community may be grouped into the following three categories:

- Development of accurate force fields to correctly describe intermolecular interactions.
- Development of reliable mixing rules to correctly model cross-interactions between different functional groups or atomic species.
- Development of efficient algorithms to improve the sampling efficiency of computer simulations and to further reduce the time for simulating physical properties.

This work addresses the first and third points above.

In Chapter Two, the installation of a Linux Beowulf cluster called yoda is described. The installation procedure, the general operating principles and the various softwares required are discussed.

As a necessary first step, Chapter Three outlines the basic theory of the thermodynamics of Monte Carlo simulations. In particular, the concept of ensembles is fleshed out and the notion of importance sampling is introduced. Several thermodynamic partition functions are given to establish the link between macroscopic thermodynamics and statistical thermodynamics. It is also explained how thermodynamic degrees of freedom can be understood in a simulation sense.

The primary simulation method used throughout this work is the Gibbs Ensemble Monte Carlo method and Chapter Four details how this algorithm is implemented. The strengths and weaknesses of alternative simulation strategies are also discussed. A new unification formalism is also developed in terms of which it is shown how the partition function of novel isothermal Gibbs Ensembles can be developed from a single governing equation which also directly provides the thermodynamic potential. The Semi-Grand Ensemble is also introduced as a prelude to the development of two new isothermal Gibbs Ensembles in Chapter Eight.

Chapter Five describes the auxiliary algorithms, the intra- and intermolecular potentials used as well as some of the calculation procedures for the chemical potential and the pressure. The essential Configurational-Bias Monte Carlo algorithm is also explained as this is required for all of the subsequent simulation work.

The simulation results of five binary *n*-alkane and 1-alkene mixtures using three different force fields are presented in Chapter Six. The performance of the NERD, TraPPE and Spyriouni *et al.* (1999) olefin force fields are discussed and their strengths and weaknesses elucidated for these relatively ideal mixtures. As justification for the development of new Gibbs Ensembles, the limitations of current N_1N_2PT simulations are explained.

Chapter Seven describes the extension of the NERD and TraPPE force fields to the simulation of conjugated alkene molecules. The objective was to test the versatility of transferable force fields to new homologous series for which the existing parameters sets should be sufficient without adjusting the intermolecular parameters. It is shown that the magnitude of the intramolecular equilibrium

geometries play a major role in the accurate prediction of the pure component phase equilibrium properties.

Building on the foundation laid in Chapter Five, Chapter Eight introduces the novel μ_1N_2VT and ξ_1NVT Gibbs Ensembles. Using these two new algorithms, the simulations of two close-boiling mixtures are presented. It is shown how the μ_1N_2VT and ξ_1NVT Gibbs Ensembles have, for the first time, permitted the simulation of a binary mixture with such a narrow phase envelope. The new Gibbs Ensembles are also shown to provide phase compositions with considerably lower statistical uncertainties than conventional N_1N_2PT Gibbs Ensembles.

Chapters Six and Seven clearly established several deficiencies with respect to current transferable united-atom force fields. Chapter Nine builds on the results of these two Chapters and presents the new TBORG *n*-alkane and alkene force fields. It is shown how the incorporation of a larger intramolecular parameter set is able to provide a force field which, for the first time, is able to provide an accurate prediction of the vapour pressures and vapour densities while not sacrificing accuracy with respect to the prediction of the liquid densities, the critical temperatures and critical densities for molecules in the range of C_2 to C_8.

CHAPTER 2:
BEOWULF CLUSTER YODA

"No one understands the value of the unforgiving minute" – General George S. Patton

2.1 Introduction

In 1965, the co-founder of Intel Corporation, Gordon Moore, made his famous observation that the number of transistors on a silicon chip roughly doubles every 2 years (Moore (1965)). Since 1965, however, developments in microprocessor technology have seen "Moore's Law" come to predict an average doubling of PC speeds every 18 months. Indeed, even for this work over the three year period beginning in December 2001 and lasting until January 2005, entry level computer speeds from South African retailers moved from approximately 1.2GHz at the end of 2001 up to over 4GHz at the beginning of 2005. With this increase in computer processing power expected to continue beyond 2015, clustered computer technologies offer an exciting new tool which can be explored to meet the computer-intensive requirements of molecular simulation of phase equilibrium.

2.2 The Beowulf Cluster as a Computational Tool

Until the mid 1990's supercomputing (or high performance) computing was essentially reserved for those who had access to budgets in the millions of dollars. The Beowulf evolution has completely changed this scenario. Using relatively cheap, off-the-shelf components from a local computer hardware vendor, it is now possible for anyone in the research community to have access to a dedicated computational resource by building a Beowulf cluster.

In most general terms, computer clusters are ensembles of independently operational computers integrated by means of an interconnection network which supports user accessible software for organizing and controlling concurrent computing tasks that may co-operate on a common application program or workload (Sterling (2001)).

2.2.1 A Brief History

In 1993, Thomas Sterling and Donald Becker led the so-called "Beowulf project" which aimed to build a gigaflops workstation for under $50,000. Several developments in previous years

had made this project possible. Indeed, by 1993, the new Intel 80386 processor represented a major performance advance over its predecessor the 80286, RAM memory had decreased to comparatively low cost levels and 10 MBps Ethernet was commonly available for standard PCs as a local area networking technology. In addition, the Linux operating system (first released by Linus Torvalds in 1991) had come to the point where it could be used in a cluster environment, and the Parallel Virtual Machine (PVM, http://www.csm.ornl.gov/pvm/, a library of linkable functions that could allow routines running on separate but networked machines to exchange data and co-ordinate their operation) had managed to become the first major cross-platform parallel programming message passing model to achieve wide acceptance. The project ran essentially outside the main parallel processing community and was based on commodity software and publicly available technology. The Beowulf project succeeded and a 16-node cluster of Intel 486 computers costing $40,000 ran in 1994. In the process, the Beowulf project developed all the necessary Ethernet driver software for Linux and additional low-level cluster management tools and demonstrated the cost effectiveness of Beowulf class systems for real-world applications.

In 1997, a Beowulf cluster costing under $50,000 won the Gordon Bell Prize for performance / unit price. This trend has continued and Beowulf clusters have increased their share of the "Top 500" supercomputer list (http://www.top500.org). In June 2001, 33 Beowulfs were in the Top 500, while in June 2005, this number had risen to 304. The major advantages of a Beowulf cluster include:

- Scalability: new nodes may readily be added to an existing cluster to increase its size,
- Availability of new software: the extensive community of software developers on the internet have ensured the ease of Beowulf installation and support, and
- Performance / unit price: the Beowulf cluster display the lowest $ / computational power statistics.

A Beowulf cluster may be conveniently thought of in terms of a layered structure of interlocking hardwares and softwares that facilitate the interface between the computational resource and the user. This is illustrated in Figure 2-1:

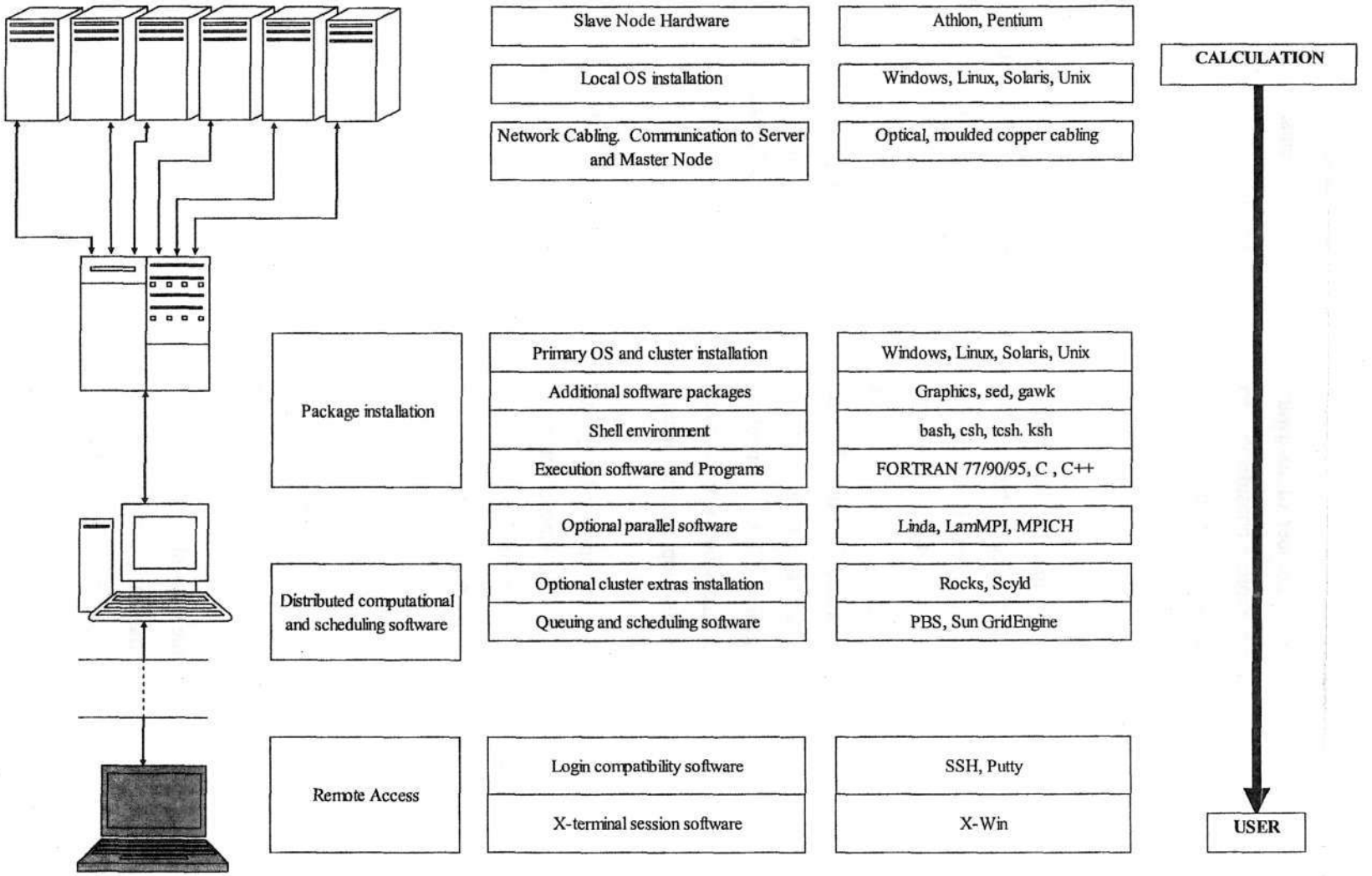


Figure 2-1: Schematic representation of the hardware and software architecture of a Beowulf cluster.

2.2.2 Elements of a Cluster

A Beowulf cluster comprises a number of hardware and software components. Unlike traditional mainframes, the user can express considerable freedom in the selection of the system architecture and the range of software which is to be installed.

2.2.2.1 The Hardware nodes

The hardware node is the principal building block of a Beowulf cluster. Normally a distinction is made between the controlling or "master node" and the computational or "slave nodes". In certain installations it is possible to install the master node as a computational node but this adds to the load on the master node with the benefit of only one additional computational PC. For this project, the master node was therefore installed as a dedicated administration node to maximise its administration performance. The Intel Pentium and Athlon AMD represent the two broad families from which slave and master nodes tend to be drawn. Ideally, all nodes within a cluster should be of similar specification. This is particular true if massively parallel jobs are to be run in order to minimize load imbalances between the computational nodes. This is, however, no longer strictly necessary with recent developments in queuing technology

At the time of the purchase of the hardware for the yoda cluster in November 2001, the Athlon AMD 1.2 GHz architecture was noted to have a computational performance advantage over the corresponding Pentium II machines in a similar price bracket. For this reason, it was decided to proceed with the Athlon AMD's as the nodes of choice.

2.2.2.2 The Network

The network provides the means of exchanging data between the clustered nodes and the means for co-ordinating their actions. The subcomponents of the network are the network interface controllers, the network switches and the network channels and links. In late 2001, considerable interest was being expressed in the Beowulf community in the use of Gigabit Ethernet connections and optical fibre cabling. Since massively parallel calculations were not likely to be undertaken on yoda, it was decided to utilise the standard 100Megabit Ethernet connection and the standard moulded copper cabling.

2.2.2.3 Operating System

Linux has emerged as the dominant Unix-like operating system. There is some debate within the Beowulf community surrounding the use of Windows-based Beowulf clusters (Sterling (2001)). The greater stability of the Linux operating platform; the large global community of Linux kernel and application developers who readily provide support, bug-fixes and software enhancements; as well the much lower cost of the Linux operating system made possible by the open source software initiative (http://www.opensource.org/) and GNU (Gnu's Not Unix), meant that the installation of a Linux operating system on the yoda cluster had a distinct cost advantage over the installation of a Windows-based system.

Aside from direct costs, the essentially freely available software and data manipulation applications on a Linux machine are more suited to dealing with the large quantities of data generated in the output file from a typical simulation. These include the various shell programming environments like bash, tcsh and csh for the manipulation of files as well as the column and row scripting and manipulation packages such as sed (http://www.gnu.org/software/sed/sed.html) and gawk (http://www.gnu.org/software/gawk/gawk.html) (Siever *et al.* (2000) and Welsh *et al.* (1999)). In addition, graphical packages like xmgrace (http://plasma-gate.weizmann.ac.il/Grace/) provide graphical analysis tools equal in utility to Microsoft Excel and which are perhaps more ideally suited to the representation of large vector arrays of data.

2.2.2.4 Resource Management

The distributed resources of a Beowulf cluster require the services of a software system for resource management. Several such software systems exist to provide administrators with the ability to balance contending demands and to apply policies to govern the use of the cluster. The tasks that need to be performed include:

- Queuing: Different people submit different jobs to a given Beowulf cluster. A queuing system is necessary to buffer job requests until resources are available to process them.
- Scheduling: This is necessary to balance the priorities of jobs with the existing cluster governance policies to optimally fit as many jobs onto a cluster at any given time or load.
- Resource control: This becomes important to put jobs on the correct nodes, to start jobs, to terminate jobs and to suspend jobs as required.
- Monitoring: Software needs to be in place in order to continuously track the status of the cluster and its resource utilization. This is needed for proper cluster management.

- Accounting: A billing system is necessary to determine the remaining time due to a user as well as to assess overall system availability.

The two most prominent packages used for this type of application are the Portable Batch System (PBS, http://www.openpbs.org/) and the Sun GridEngine (http://gridengine.sunsource.net/) softwares. An additional package for resource management is the Maui Scheduler (http://www.clusterresources.com/products/maui/). The Maui scheduler in fact sits on top of other lower resource managers and offers additional resource management options. It was not installed for this project due to the limited demands that were placed on the Beowulf cluster yoda.

2.2.2.5 Program Software

Common programming languages used for molecular simulation include C and FORTRAN 77. Although there are now more sophisticated variants of these two programming languages (C++; ; FORTRAN 90; FORTRAN 95), their portability, ease of integration into a myriad of installation environments, ease of compilation and low processing overhead, all ensure that they still find appeal in the programming community. For this project, a serial and a parallel version of the BIGMAC programming code as provided by Thijs Vlugt were modified and utilized for the simulation of phase equilibria.

2.2.2.6 Parallel Environment

Most simply, parallel computation on a Beowulf is achieved by dividing a computation into parts and making use of multiple processes on separate processors to execute each of these parts. The most simple approach to achieve this is to have each of the processors send and receive messages. This is made possible by a message-passing library specification, MPI (Message Passing Interface, http:///www.mpi-forum.org). MPICH (http://www-unix.mcs.anl.gov/mpi/mpich/) and LamMPI (http://fink.sourceforge.net/pdb/package.php/lammpi) are the two most popular implementations of MPI.

In late 2001, the BIGMAC code was used to simulate the vapour-liquid equilibrium of a system of 300 molecules of (propane + *n*-pentane) at 64°C on the Beowulf clusters yoda (http://yoda.nu.ac.za) and Ingvar (www.nsc.liu.se/systems/ingvar/) at mole fractions of 0.1 and 0.9 propane. At the time of the test, both the Ingvar and yoda clusters had similar slave node specifications. Typically, a

Monte Carlo simulation will only achieve a linear speed increase with increasing numbers of nodes running a job in parallel when the number of interaction sites in a computation exceeds 1,000 sites (Vlugt (2002)). As may be seen from Figure 2-2, the speed of execution of the BIGMAC job when run on 8 nodes is not significantly higher than that achieved when the job is only run on 2 nodes. This is due the bottleneck of processes (and algorithms) requiring network communication (message-passing) and hence the use of any more than 2 nodes per job represented a waste of system resources for the yoda cluster. For 8 nodes, this network communication overhead in fact leads to the increase in the CPU time as shown in Figure 2-2 for the 90mol% simulations on yoda and Ingvar. The typical system sizes used in this work were consistently less than this "threshold" value of about 1000 interaction sites and hence the serial BIGMAC code was modified and used for the results presented in this work.

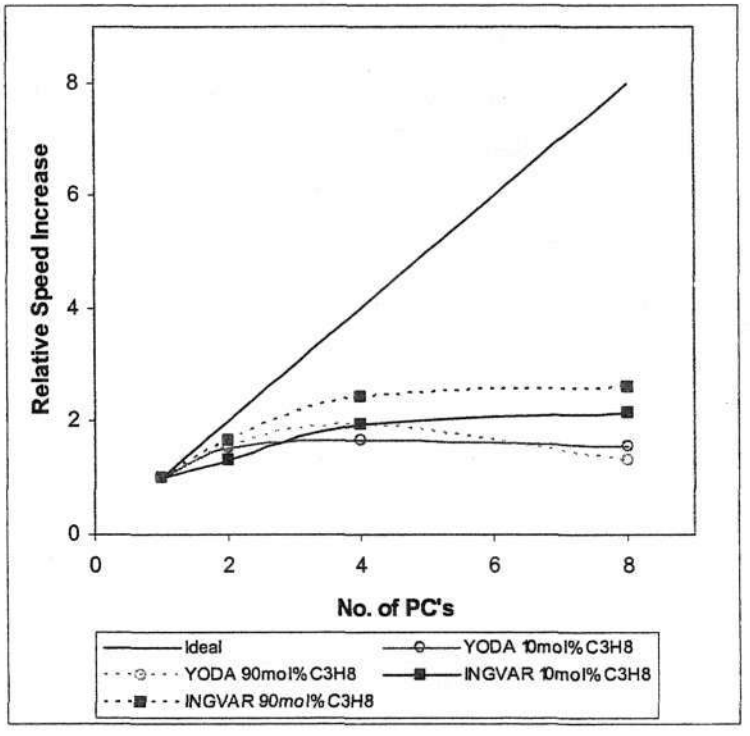


Figure 2-2: Performance of the BIGMAC parallel program code for the 300 molecule binary system of propane + *n*-pentane at 64°C on the yoda and Ingvar Beowulf clusters.

2.2.2.7 Remote Access Software

An important advantage of the Beowulf cluster is that the access of the user may be remote with respect to the physical location of the Beowulf cluster itself. The Beowulf cluster may thus

exist as a dedicated computational resource. This is achieved by means of a remote login software. In terms of this work, the Open SSH (Secure Shell) software was used (http://ww.openssh.com). This software is freely available for academic use and hence represented no additional financial overhead for the project.

In this way, using relatively cheap off-the-shelf components and hardware, a 20-node Beowulf cluster called "yoda" was assembled for this work. The following hardware specifications describe the equipment used to assemble the Beowulf cluster yoda:

Table 2-1: Specifications of the master and slave nodes for the Beowulf cluster yoda.

	Master Node	Slave Node
CPU	1.2GHz AMD Thuderbird	1.2GHz AMD Thuderbird
Motherboard	Asus Socket A7VLT33-VM	Asus Socket A7VLT33-VM
Slots	Five PCI and 4 Integrated AGP	Five PCI and 4 Integrated AGP
RAM	256Mb PC 133 SDRAM	256Mb PC 133 SDRAM
Hard Disk	80Gb Hard Drive	20Gb Hard Drive
Floppy Drive	Samsung 3.5" 1.44Mb	Samsung 3.5" 1.44Mb
CD-ROM	52X CD-ROM Drive	-
Graphics Card	8Mb Graphics Card	8Mb Graphics Card
Network Card	DFE-570TX, 4x100MB, TP, PCI	3Com FastEtherlink XL-3C905C-TX-M TP PCI

Table 2-2: Additional Hardware for the Beowulf cluster yoda.

Network Switch	One 3COM 24-port 10/100Mbit Fast Ethernet Switch
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

Three important cost-saving features included:

- Only the master node and the first slave node had CD-ROM drives. This is because once the master node had been installed, all remaining installations could be achieved by boot disk.
- Only 1 master node monitor and one mobile slave node monitor were purchased

- The slave node hard-drives could be limited to 20Gb capacity since this greatly exceeded the required installation space.

2.2.3 Installation of the Beowulf cluster yoda

The installation of yoda and was completed at the end of 2001. The Red Hat 7.1 provided a 2.4.2 kernel and almost all of the necessary software that was required for the installation and administration of the cluster. Version 7.1 was the most recent version of Red Hat Linux available at the time of the installation.

2.2.3.1 Partitioning the Hard Drive

A total of 9 partitions were created on the hard drive of the master node:

- A boot partition, "/dev/hda1".
- Another primary partition, "/dev/hda2".
- A third primary partition created as the system's swap drive.
- A fourth partition, "/dev/hda4", was created as an extended partition to include the rest of the disk partitions. These remaining partitions included "/usr", "/usr/local", "/var", "/tmp" and "/home".

2.2.3.2 Local Area Network Settings and Running Services

The master node is part of two different networks, an internal network with a range of private IP addresses (10.0.0.0) and an external network with a public IP address. The eth0 device was used for the host for the public IP address and the eth1 device was used for the host for the private IP address. In addition, the following services were selected to be run on the machine: crond, dhcpd, gpm, ipchains, iptables, keytable, netfs, network, nfs, nfslock, ntpd, portmap, random, rawdevices, syslog, tftp, xfs, xinetd and sshd2.

2.2.3.3 LILO Configuration

The installation of LILO on the Master Boot Record was only necessary on the master node because the slave nodes were booted using dhcp and bootp over the network. In addition, a high

security firewall was configured to ensure network security. SSH traffic was explicitly allowed on port 22 and the firewall rules were configured only to apply on eth0.

2.2.3.4 Setting up the Hostnames of the Nodes Used

The hostnames of all the nodes in the cluster had to be defined and this was done in the "/etc/hosts" file which served as a naming service for the cluster's internal network. In addition, the "/etc/resolv.conf" file was configured to provide access to the Internet Domain Name System. This particular file is needed in a series of routines to resolve IP-address to domain names or vice versa.

2.2.3.5 Installation of the DHCPD and BOOTP

DHCP stands for Dynamic Host Control Protocol, and its server provides its clients with IP addresses according to those defined in its configuration file, "/etc/dhcpd.conf". The DHCPD that came with Red Hat Linux 7.1 also featured a BOOTP server and by defining each client separately, this feature was enabled and each node could be booted off the network. Each client had to be listed in the DHCP configuration file and given an IP-address based on its MAC address (a six byte unique number, identifying that particular Ethernet card.). The DHCPD only applies to eth1. In order to reflect this explicitly, a small alteration was made to the DHCPD SystemV control script found in "/etc/init.d/dhcpd".

2.2.3.6 Installation of Secure Shell (SSH)

At the time of installation, the OpenSSH version 3.1.0 was the latest version of SSH and it was installed on the cluster. After installation, a few changes had to be made to the file "/etc/ssh2/sshd2_config" for the server to work in the cluster environment. Most significantly, every user must be able to logon to every node without giving their password, as this is needed by MPI and "rdist". A method to solve this problem is to use host-based authentication within the cluster together with password authentication for access from the outside. This required:

- Changing "AllowedAuthentications" to "hostbased, password".
- Checking that "IgnoreRhosts" is set to "no"
- Changing "DefaultDomain" to the appropriate domain

- Creating a file named ".shosts" with permissions set to 0400. This file had to be located in the home directories of all users and had to contain all hostnames and a username for each host. The owner and the group of the file had to be explicitly set to be the user.

The directory named "knownhosts" in "/etc/ssh2/" must contain all public keys from every node. The SSH server was restarted by issuing the command "/etc/rc.d/init.d/sshd restart".

2.2.3.7 Installation of Slave Nodes

Next, it was possible to install a single slave node as an image which could be used for cloning additional nodes. For compatibility reasons, the same version of RedHat Linux 7.1 was used. The actual partitioning of each node's hard drive is done by BpBatch during the installation phase of the imagefiles.

The installation and configuration procedure of the slave nodes was the same as for the master node, except for the generation of keys. Since the same pair of keys had to be used on every node, the keys in "/etc/ssh2/" on the master node had to be copied to the slave node and then used to overwrite any created keys that were local to the slave node.

In order to add new nodes to the cluster, a few changes had to be made to the master node before the new nodes could be used:

- The new node with its network information (MAC address, IP-number) had to be added to "/etc/dhcpd.conf".
- The new node's hostname and IP-number had to be added to "/etc/hosts".
- The new node had to be added "/etc/distfile".
- A copy of "/etc/ssh2/knowhosts/node00.yoda.ssh-dss.pub" had to be made and renamed to "*newhostname*.yoda.ssh-dss.pub".
- The new node's hostname had to be added to "/var/spool/PBS/server_priv/nodes".

Importantly, however, whenever a new user was added or deleted, this information had to be manually updated on all nodes.

The administration of slave nodes can be a tedious task. The availability of cloning software, however, alleviates this burden. For this work, the freely downloadable remote-boot processor BpBatch (http://www.bpbatch.org) was used. BpBatch allows one to partition hard drives as well as to create images of entire hard disk partitions and file systems. It then provides a framework to

distribute and install these images on a cluster of PC's. The software takes control at an early stage of the boot process of the computer. Network cards supporting PXE booting were thus required. BpBatch's loader must be specified as the bootfilename in the DHCP information for the client computer. BpBatch's loader can then use TFTP to retrieve the main BpBatch program from the server. This also required the creation of a "/tftpboot" directory with the correct file permissions. Once BpBatch is started, it looks for a default script on its network server. Every command found in this script will be executed in a serial order, until the end of the script is reached.

The clone images were created with the help of BpBatch's MrZip. MrZip is a tool used for compressing a selected amount of a disk system. After completion the image files were moved to the "/tftpboot" directory, from where BpBatch can collect and use the images as required.

2.2.3.8 Installation of the Message Passing Interface (MPI)

MPICH is a portable, open-source implementation of the Message-Passing Interface standard and is suitable for clusters built on standard Ethernet network devices and wiring. At the time of installation, the latest MPICH version was MPICH-1.2.3 and this was selected for installation. In terms of the website documentation at http://www-unix.mcs.anl.gov/mpi/mpich, MPICH implements all features found in the MPI 1.2 standard. Every node in the cluster required that a local copy of MPICH be installed. This was achieved by putting the compiled version of MPICH in an nfs-mountable directory, mounting this on every node and then installing it on each node. The "/usr/lib/mpich/share/machines.LINUX" file was modified to reflect all nodes in the cluster while the "/usr/lib/mpich/bin/mpirun.args" file was modified to integrate MPICH with PBS.

2.2.3.9 Installation of the Portable Batch Scheduler (PBS)

At the time of installation, PBS had become the *de facto* standard for job scheduling on Linux Beowulf clusters. In order to perform the workload management functions of queuing, scheduling and monitoring, PBS consists of the following fundamental parts:

- The Job Server. This daemon is run continuously on the master node and its main function is to receive and create batch jobs, modify jobs if needed and protect the jobs in case of a system crash.
- The Job Executor (MOM). This daemon handles the actual execution of jobs. When it receives a copy of a job from the server, it creates a new session. Apart from the actual

execution, the executor is also responsible for returning the job output to the user. A job executor daemon runs on each computer node that is part of the cluster.

- The Job Scheduler. The Job Scheduler daemon deals with enforcing the resource policies defined by the system administrator. The Scheduler is in close contact with the MOM's running on each node in order to determine what resources are available and the status of each job.

The OPENPBS software was downloaded and installed in the "/usr/pbs/" directory while the default "/var/spool/PBS/" directory was used for the spooling directory. Both the Job Server and the Job Scheduler were configured by modifying configuration files located in the "/var/spool/PBS/" directory on the master node. In particular:

- The "/var/spool/PBS/server_privs/nodes" file was edited to include the names of all 19 available computational nodes. These names are read by the server during bootup.
- The "/var/spool/PBS/server_privs/serv_conf" file was modified to define 4 default queues ("qmini", "qsmall", "qmedium", and "qlarge") and one routing queue. The "/var/spool/PBS/server_priv/qmgr" command was used to apply the queue configuration.

Each slave computational node also required the installation of a server to handle the execution of jobs sent from the master node server. This required:

- Setting "PBS_SCP=/usr/local/bin/scp" in the "/etc/pbs.conf" file in order to specify the application to use during the copying of output files between nodes.
- Specifying the execution server to use in the "/var/spool/PBS/mom_priv/config" file on each node.

2.3 Cost Considerations

Although a Beowulf cluster exhibits very favourable price per unit performance statistics, the useful lifespan of a Beowulf cluster is typically less than 3 years (Vlugt (2002)). This is because the relative speed of all of the computational nodes is then ¼ that of entry level computer in accordance with Moore's Law. In addition, manufacturer guarantees typically do not extend beyond 2 years for the most expensive components such as the motherboards and hard-drives.

The greatest hazard to the yoda installation proved to be the semi-regular electrical power outages (usually around the holiday season in South Africa) which ultimately further reduced the lifespan of

the hard-drives. Multiple UPS units for both the master nodes and slave nodes with automatic shutdown softwares thus proved essential to ensure stable operation and cluster protection.

Network and communication latency is a crucial issue which must be adequately addressed in any cluster installation. Although this issue is most serious for parallel computing applications where the transfer of large quantities of data between the slave nodes and master node must be achieved efficiently, even the serial calculations of the yoda cluster suffered from severe packet loss depending on how heavily the cluster was loaded with computational jobs and on how many jobs were queued. The replacement of the handmade cabling with moulded Belkin "FastCat" Cat5e cabling essentially eliminated this problem at the end of 2003. Data packet losses were dramatically reduced and the network latency decreased by two orders of magnitude in "ping" tests down to approximately 0.2 milliseconds.

CHAPTER 3:

THERMODYNAMICS OF THE MONTE CARLO SIMULATION OF PHASE EQUILIBRIA

"To see the world in a grain of sand;
and heaven in a wildflower
Hold infinity in the palm of your hand;
And eternity in an hour" – William Blake

3.1 Ensemble Averages

Legend holds that in 1872 a $25,000 wager was made between California Governor Leland Stanford and two other men, James R Keene and Fredrick MacCrellish, concerning whether all four of a horse's legs left the ground in a full gallop. It was a commonly held Victorian belief that when a horse galloped it always maintained one hoof in contact with the ground. Stanford was convinced otherwise, but since a horse's trot or canter was too fast for the human eye to see, the point seemed impossible to prove. He thought that photography would decisively settle the matter and hired Edweard James Muybridge who at the time held the position of 'Official Photographer for the US Government for the Pacific Coast'.

Muybridge's subject was Stanford's horse, and he calculated that he needed 12 cameras 22 inches apart to record the separate parts of the horse's stride. As it ran across the trip wires, the horse pulled out a pin allowing the shutter to move and take a picture.

Muybridge was successful, Stanford won the bet and the work culminated in a lecturing circuit for Muybridge concerning animal locomotion and development of the zoopraxiscope which was the precursor to Edison's phonograph.

The natural question of course is: What exactly is the link between the molecular simulation of phase equilibria and an obscure 19th century bet? Essentially, the genius of Muybridge was in realising that a sufficient number of still photos taken at regular intervals could be used to decompose an apparently continuous variable (a horse's gallop) into discrete, manageable images to observe the horse's gallop. This analogy is, therefore, both convenient and appropriate for understanding the essence of a Monte Carlo simulation. Basically, the idea behind a Monte Carlo

simulation is to calculate a macroscopic variable by substituting our almost intuitive understanding of a macroscopic variable being continuous with discrete microscopic 'pictures' of the system.

This is consistent with the way in which one measures macroscopic variables. Consider a pressure gauge. When one observes the reading on the gauge, one is not actually observing the instantaneous pressure, but rather the reading that one takes is an average of the fluctuating pressure observed on the gauge. These different fluctuating values can be thought of as co-ordinates in multidimensional space, namely phase space. A macroscopically observed variable may thus be written as a time average as follows:

$$\mathcal{A}_{obs} = \langle \mathcal{A} \rangle_{time} = \left\langle \mathcal{A}\left(\Upsilon(t)\right) \right\rangle_{time} = \lim_{t_{obs}\to\infty} \left\{ \frac{1}{t_{obs}} \int_0^{t_{obs}} dt\, \mathcal{A}\left(\Upsilon(t)\right) \right\} \qquad (3\text{–}1)$$

where $\mathcal{A}$ is the macroscopic variable, $\langle\ \rangle_{time}$ denotes an average over time, $\Upsilon(t)$ denotes a particular point in phase space at time t and subscript t_{obs} denotes the observation time. In practice it is not feasible to calculate $\mathcal{A}_{obs}$ over an infinite time and hence the link with the horseracing analogy. The idea, therefore, is to replace the time average by an ensemble average where an ensemble is a large collection of points Υ in phase space with the same macroscopic parameter values, e.g. constant number of molecules, constant volume and constant temperature (NVT) or constant number of molecules, constant pressure and constant temperature (NPT). One may therefore replace the time average of Equation (3–1) with an average over all the members of the ensemble 'frozen' at a particular time:

$$\mathcal{A}_{obs} = \langle \mathcal{A} \rangle_{ens} = \sum_{\Upsilon} \mathcal{A}(\Upsilon) \mathcal{P}_{ens}(\Upsilon) \qquad (3\text{–}2)$$

where $\langle\ \rangle_{ens}$ denotes an ensemble average and $\mathcal{P}_{ens}(\Upsilon)$ represents the probability density of observing a particular point in phase space of a given ensemble. It should be stressed that a necessary requirement for Equation (3–2) to hold valid is that the system be 'ergodic', that is that all state points eventually access all other state points and that there are no inaccessible points in phase space that should in fact be accessible.

It is more convenient to define a 'weight' function $w_{ens}(\Upsilon)$ in place of $\mathcal{P}_{ens}(\Upsilon)$ which satisfies the following equations (Allen and Tildesley (1987)):

$$\mathcal{P}_{ens}(\Upsilon) = Q_{ens}^{-1} w_{ens}(\Upsilon) \qquad (3\text{–}3)$$

$$Q_{ens} = \sum_{\Upsilon} w_{ens}(\Upsilon) \qquad (3\text{–}4)$$

$$\langle \mathcal{A} \rangle_{ens} = \left\{ \sum_{\Upsilon} w_{ens}(\Upsilon) \mathcal{A}(\Upsilon) \right\} \Big/ \sum_{\Upsilon} w_{ens}(\Upsilon) \qquad (3\text{–}5)$$

where Q_{ens} is the ensemble partition function (also called the sum over states).

For the purposes of this thesis, it is necessary to initially consider the partition functions for three 'basic' ensembles. It is assumed throughout this thesis that the basic simulation box for all equations is a cube and that the dimensionality of all simulations boxes is thus 3. As a starting point for further discussion, it is instructive to write the quasi-classical partition function for the pure component canonical (NVT) ensemble (Allen and Tildesley (1987)):

$$\begin{aligned} Q_{NVT} &= \frac{1}{N! h^{3N}} \int d\mathbf{p}^N \exp\left[-\mathcal{K}(\mathbf{p}^N) \Big/ k_B T \right] \int d\mathbf{r}^N \exp\left[-U(\mathbf{r}^N) \Big/ k_B T \right] \\ &= Q_{NVT}^{id} Q_{NVT}^{ex} \end{aligned} \qquad (3\text{–}6)$$

where Q_{NVT}^{id} denotes the kinetic (ideal gas) NVT partition function, Q_{NVT}^{ex} denotes the potential (excess) NVT partition function, h is Planck's constant, k_B is Boltzmann's constant, T is the absolute temperature, $\mathcal{K}(\mathbf{p}^N) = \sum_{i=1}^{N} \mathbf{p}_i^2 / 2m$ is the kinetic energy, $U(\mathbf{r}^N)$ is the potential energy and the integration is performed over all molecular momenta vectors $\mathbf{p}^N$ and molecular position vectors $\mathbf{r}^N$ for the N molecules in the system. The exponent '3N' in the term h^{3N} arises because of the 3-dimensional nature of the simulation box that has been assumed to apply for N molecules. The $N!$ term accounts for the indistinguishability of the molecules. The final line in Equation (3–6) results because the energy is always expressible as the sum of the kinetic (**p**-dependent) and potential (**r**-dependent) contributions. In particular, the ideal contribution may be evaluated analytically because $\mathcal{K}$ is a quadratic function of the momenta to give:

$$Q_{NVT}^{id} = \frac{V^N}{N! \Lambda^{3N}} \qquad (3\text{–}7)$$

where V is the volume and Λ is the thermal de Broglie wavelength where

$$\Lambda = \left(h^2 / 2\pi m k_B T \right)^{1/2} \qquad (3\text{–}8)$$

where m is the molecule mass. The excess part is then:

$$Q_{NVT}^{ex} = V^{-N} \int d\mathbf{r}^N \exp\left[-U(\mathbf{r}^N) \Big/ k_B T \right] \qquad (3\text{–}9)$$

For Monte Carlo simulations, this analytical integration of the kinetic contribution to the partition function is usually possible. It is convenient to recognise that the partition function represents a sum over all possible molecular arrangements, i.e. a sum over states.

Monte Carlo simulations in the isothermal-isobaric ensemble or NPT ensemble were first described by Wood (1968), and a neat derivation of the partition function of the NPT ensemble is given by Frenkel and Smit (1996). In the NPT ensemble, the volume is allowed to fluctuate so as to keep the total system pressure constant. One must therefore account for these fluctuations in the volume in the excess part of the partition function of a pure component:

$$Q_{NPT} = \frac{1}{K_0 N! \Lambda^{3N}} \int d(\ln V) V^{N+1} \exp[-\beta P V] \left(\int ds^N \exp\left[-\beta U(\mathbf{s}^N)\right] \right) \qquad (3\text{–}10)$$

where $\beta = 1/k_B T$, P is the pressure and the integration is now performed over all dimensionless vector co-ordinates $\mathbf{s}^N$ where $\mathbf{s} = \mathbf{r}/L_b$ and L_b in the box length of a cubic simulation box. The reason for the separation of the term V^N from the integral over molecular co-ordinates is because the partition function must now be integrated over all possible volumes. It is important to note that the function of the term K_0 is to non-dimensionalize the partition function. In this instance, K_0 has units of volume. It is somewhat arbitrary since it cancels off in practice (Equation (3–16) reveals why). It should also be noted that in this formulation of the NPT ensemble partition function, the fluctuations in the volume are integrated over the logarithm of the volume (Eppinga and Frenkel (1984)).

The Grand Canonical (μVT) Ensemble was first implemented for classical fluids by Norman and Filinov (1969). An elegant derivation of the basic partition function is given in Frenkel and Smit (1996) which may be written as follows for a pure component:

$$Q_{\mu VT} = \sum_{N=0}^{\infty} \left(\frac{\exp[\beta \mu N] V^N}{N! \Lambda^{3N}} \int d\mathbf{s}^N \exp\left[-\beta U(\mathbf{s}^N)\right] \right) \qquad (3\text{–}11)$$

where μ is the chemical potential and the partition function now includes a summation over numbers of molecules since the number of molecules may fluctuate in this ensemble.

The connection from the partition function to classical thermodynamics is made by defining a thermodynamic potential Ψ_{ens} (McQuarrie (1976)):

$$\Psi_{ens} = -\ln[Q_{ens}] \qquad (3\text{–}12)$$

The thermodynamic potential for the NVT, NPT and μVT ensembles may then be written as (Allen and Tildesley (1987)):

$$A/k_BT = -\ln\left[Q_{NVT}\right] \tag{3–13}$$

$$G/k_BT = -\ln\left[Q_{NPT}\right] \tag{3–14}$$

$$L/k_BT = -PV = -\ln\left[Q_{\mu VT}\right] \tag{3–15}$$

where A is the Helmholtz free energy, G is the Gibbs free energy and L is the Hill energy.

3.2 The Metropolis Method

If one considers Equation (3–5) and, say, Equation (3–6) for the NVT ensemble, it would seem that the effort in evaluating an average of the form $\langle \mathcal{A} \rangle_{ens}$ typically rests in the evaluation of an expression of the form (Frenkel and Smit (1996)):

$$\begin{aligned} \langle \mathcal{A}_{\mathrm{NVT}} \rangle &= \frac{\int d\mathbf{r}^N \left(\mathcal{A}\left(\mathbf{r}^N\right) \exp\left[-\beta U\left(\mathbf{r}^N\right)\right]\right)}{\int d\mathbf{r}^N \exp\left[-\beta U\left(\mathbf{r}^N\right)\right]} \\ &= \frac{\int d\mathbf{r}^N \left(\mathcal{A}\left(\mathbf{r}^N\right) \exp\left[-\beta U\left(\mathbf{r}^N\right)\right]\right)}{Q_{NVT}} \end{aligned} \tag{3–16}$$

where the kinetic contributions cancel because of the ratio of the two integrals. Conventional numerical integration techniques (e.g. Simpson's rule) are in general not feasible to use to evaluate the above expression due to the intractably large number of functional evaluations that would be required to evaluate N molecules over 3-dimensional space for all possible arrangements of these molecules. Also, many of these configurations may also have a low probability density. A random Monte Carlo sampling integration scheme is also not a feasible strategy because one does not know the distribution $\mathcal{P}\left(\mathbf{r}^N\right)_{ens}$ *a priori*. Equation (3–16) is, however, a ratio of two integrals and Metropolis *et al.* (1953) were able to show that it is possible to devise an efficient Monte Carlo importance sampling scheme to evaluate such integrals.

This may be understood by recognising that the probability of observing the system in a configuration around $\mathbf{r}^N$ for the NVT ensemble, $\mathcal{P}\left(\mathbf{r}^N\right)_{NVT}$, is given by:

$$\mathcal{P}\left(\mathbf{r}^N\right)_{NVT} = \frac{\exp\left[-\beta U\left(\mathbf{r}^N\right)\right]}{\int d\mathbf{r}^N \exp\left[-\beta U\left(\mathbf{r}^N\right)\right]} = \frac{\exp\left[-\beta U\left(\mathbf{r}^N\right)\right]}{Q_{NVT}} \tag{3–17}$$

If one could then generate points in configuration space according to the probability distribution $\mathcal{P}\left(\mathbf{r}^N\right)_{NVT}$ around $\mathbf{r}^N$, then on average:

$$\left\langle \mathcal{A}_{NVT} \right\rangle \approx \frac{1}{R}\sum_{i=1}^{R} \varsigma_i \mathcal{A}\left(\mathbf{r}_i^N\right) \tag{3–18}$$

where ς_i is the number of points generated per unit volume around $\mathbf{r}^N$ and R are the total number of points generated. This is the basic idea behind importance sampling, i.e. one transforms the task from generating points in configuration space **randomly** and then giving them a weight given by $\mathcal{P}\left(\mathbf{r}^N\right)_{ens}$, to generating points in phase space according to the distribution given by $\mathcal{P}\left(\mathbf{r}^N\right)_{ens}$ and then giving each of these samples an **equal** weighting.

Consider the analogy of a student trying to determine the most popular beer in a well frequented student night club and restaurant street area on a Friday night. One possible approach would be to attempt to go **into every restaurant or night club on the street** without exception and accumulate statistics on how many beers of each type are consumed in each establishment. Of course, several of these haunts will be closed (most clubs usually have only one or two nights a week when they are designated as the place to be seen). Thus there would therefore be several wasted excursions to closed or not very busy bars. An alternative strategy would be to choose to only go into those night clubs which have the largest number of cars parked outside and to then question the clientele who seem to be drinking the most beers as to what their favourite brew is. The first strategy is analogous to the approach of conventional numerical integration using say Simpson's Rule: it is **very** time intensive. The second approach represents an implementation of the Metropolis importance sampling technique: so long as a sufficient number of people are questioned in a sufficient number of bars (there might be reduced prices in different bars on different beers and experience dictates that price invariably overrides student brand loyalty) then a good estimate of the most consumed beer will be obtained. In principle, the first approach would give us total statistics on the amount of beer consumed in the evening and so would be able to provide information on the total beer consumed for all types in all places on the Friday of interest. This is analogous to obtaining information regarding the quantity Q_{ens}. The importance sampling approach, however, cannot provide this information (i.e. cannot provide Q_{ens}) since the curious student only enters those night clubs and restaurants that have the largest crowds outside.

In terms of the Monte Carlo simulation of phase equilibria, the solution therefore becomes to set up a Markov chain of states which has a limiting distribution given by $\mathcal{P}\left(\mathbf{r}^N\right)_{ens}$. A Markov chain is a sequence of trials that satisfies two conditions (Allen and Tildesley (1987)):

- The outcome of each trial belongs to a finite set of outcomes $\{\Upsilon_1, \Upsilon_2, ..., \Upsilon_o, \Upsilon_n, ...\}$
- The outcome of each trial depends only on the trial that immediately precedes it

Two states Υ_o and Υ_n are linked by a transition probability $\pi(o \to n)$ which is the probability of going from an old state to a new state. The collection of all $\pi(o \to n)$ forms a transition probability matrix which is both stochastic (random) and ergodic and which must satisfy the condition that it must not destroy the limiting distribution $\mathcal{P}\left(\mathbf{r}^N\right)_{ens}$ for the ensemble once this distribution is reached (Frenkel and Smit (1996)). It is convenient to enforce the unnecessarily strong condition of microscopic reversibility which requires that the number of moves from an old state o to a new state n is exactly cancelled by the number of reverse moves at equilibrium. This detailed balance condition may be expressed as (Frenkel and Smit (1996)):

$$\mathcal{P}^{\text{old}}\pi(o \to n) = \mathcal{P}^{\text{new}}\pi(n \to o) \tag{3–19}$$

In practice a Monte Carlo algorithm consists of two basic steps. The first is the attempt to perform a trial move from state o to state n while the second stage involves the decision to accept or reject this trial move. This may be expressed as:

$$\pi(o \to n) = acc(o \to n)\alpha(o \to n) \tag{3–20}$$

where the probability of accepting a trial move is denoted by $acc(o \to n)$ and where $\alpha(o \to n)$ is the probability of performing a trial move from state o to state n (Allen and Tildesley (1987), Frenkel and Smit (1996)). In the original Metropolis *et al.* (1953) implementation, α was chosen to be a symmetrical matrix (i.e. $\alpha(o \to n) = \alpha(n \to o)$) which combined with Equations (3–19) and (3–20) leaves:

$$\frac{acc(o \to n)}{acc(n \to o)} = \frac{\mathcal{P}^{\text{new}}}{\mathcal{P}^{\text{old}}} \tag{3–21}$$

There exist different possible choices for selecting $acc(o \to n)$ and a discussion of this is contained in Allen and Tildesley (1987). The choice of Metropolis *et al.*, however, appears to be the most efficient, i.e.:

$$\begin{aligned} acc(o \to n) &= \mathcal{P}^{new}/\mathcal{P}^{old} \quad \text{if} \quad \mathcal{P}^{new} < \mathcal{P}^{old} \\ &= 1 \quad \text{if} \quad \mathcal{P}^{new} \geq \mathcal{P}^{old} \end{aligned} \tag{3–22}$$

The second line of Equation (3–22) indicates that the probability of accepting a move cannot exceed unity. The transition probability for going from state o to state n can then be written as:

$$\begin{aligned} \pi(o \to n) &= \alpha(o \to n) \quad \text{if} \quad \mathcal{P}^{new} \geq \mathcal{P}^{old} \\ &= \alpha(o \to n)\left[\mathcal{P}^{new}/\mathcal{P}^{old}\right] \quad \text{if} \quad \mathcal{P}^{new} < \mathcal{P}^{old} \\ \pi(o \to o) &= 1 - \sum_{n \neq o} \pi(o \to n) \end{aligned} \tag{3–23}$$

The procedure by which one accepts a new state is thus as follows. The first line of Equation (3–23) indicates that if the probability of observing the new state $\mathcal{P}^{new}$ is higher than the probability of observing the old state $\mathcal{P}^{old}$ then the transition probability is equal to the probability of generating the trial move, i.e. it is unconditionally accepted without applying a stochastic acceptance criterion. The second line of Equation (3–23) states that if $\mathcal{P}^{new}$ is lower than $\mathcal{P}^{old}$, then the trial move is only conditionally accepted. This is determined by generating a random number, $\mathcal{R}_{[0,1]}$, from a uniform distribution in the interval [0,1]. The trial move is accepted if $\mathcal{R}_{[0,1]} < acc(o \to n)$ and is rejected otherwise. The probability that $\mathcal{R}_{[0,1]}$ is less than $acc(o \to n)$ is clearly equal to $acc(o \to n)$ and so this acceptance rule guarantees that one generates the correct acceptance criterion for a trial move from state n to state o (Frenkel and Smit (1996)). The final line of Equation (3–23) indicates that the old state must be recounted if the trial move is rejected.

3.3 Degrees of Freedom

In many ways, molecular simulation can conveniently be thought of as a kind of 'computer experiment'. This is because molecular simulation effectively aims to describe physical systems such as those encountered in chemical thermodynamics. The quantities encountered in chemical thermodynamics can also be determined in a statistical sense by molecular simulation methods. Of particular relevance to the simulation of phase equilibria are the following variables:

Table 3-1: Intensive and extensive variable pairing and contribution to total thermodynamic equilibrium.

Extensive property	Conjugate intensive property	Equilibrium component
Number of molecules, N	Chemical potential, μ	Chemical equilibrium
Volume, V	Pressure, P	Mechanical equilibrium
Internal energy, E	Temperature, T	Thermal equilibrium

Total thermodynamic equilibrium is achieved when the system under study achieves chemical equilibrium, mechanical equilibrium and thermal equilibrium. An understanding of the pairing of these variables is therefore significant for equilibrium chemical thermodynamics because the specification of either the extensive or conjugate intensive variable of a pair is required to determine each of the three equilibrium components of total thermodynamic equilibrium. The conjugate variable of the pairing which is not specified then becomes the dependent variable and must be determined from the equilibrium condition of the system. In this way, the specification of either the number of molecules or the chemical potential determines the chemical equilibrium, the specification of pressure or volume determines the mechanical equilibrium and the specification of temperature or internal energy determines the thermal equilibrium. This, in turn, will affect the specific ensemble that will be appropriate for a given simulation (e.g. grand-canonical (μVT) ensemble, isothermal-isobaric (NPT) ensemble, canonical (NVT) ensemble). The number of intensive and extensive parameters which need to be specified may be determined from the Gibb-Duhem phase rule (Smith *et al.* (1996)):

$$DOF = \mathrm{C} - \mathrm{F} + 2 \qquad (3\text{–}24)$$

where DOF is the number of intensive degrees of freedom which need to be specified, C is the number of components, and F is the number of phases present. Thus for a two-component and two phase system, two intensive parameters will need to be set in order to specify the equilibrium condition of the system under study. An important prescription is that it is necessary to stipulate at least one extensive property. In the simulation of a two-phase system, two extensive properties will need to be specified to place two physical constraints on the system. This has the effect of placing physical bounds on both of the simulation "boxes".

CHAPTER 4:

GIBBS ENSEMBLE MONTE CARLO SIMULATIONS

"Make everything as simple as possible but not simpler" – Albert Einstein

4.1 Phase Coexistence

The condition for the coexistence of two or more phases (I,II,...), is that the temperature for all coexisting phases must be equal, the pressure in all phases must be equal and the chemical potential of each species i must be the same for that species in all phases (Smith *et al.* (1996)):

$$\left(T_{\text{I}}=T_{\text{II}}=...=T\right),\ \left(P_{\text{I}}=P_{\text{II}}=...=P\right),\ \left(\mu_{\text{I},i}=\mu_{\text{II},i}=...=\mu_{i}\right) \qquad (4\text{–}1)$$

It might, therefore, seem appropriate to stipulate a constant μPT ensemble to simulate phase equilibrium. This approach has two inherent flaws. Any sensible simulation requires at least one extensive parameter specification in order to place a physical bound on the size of the system (Frenkel and Smit (1996)). In addition, this μPT ensemble would correspond to an over-specification of the state of the system in terms of Equation (3–24) because it is, in general, not possible to stipulate all intensive variables *a priori*. This would violate the Phase Rule.

In a physical experiment, a first-order phase transition is normally quite easy to locate. So long as one is at a temperature sufficiently far away from the critical point, one will observe that at a suitable temperature and density that an initially homogenous phase will divide into two phases separated by an interface. Before the advent of the Gibbs Ensemble, molecular simulations had to locate phase coexistence indirectly. This usually involved conducting several simulations, evaluating the macroscopic properties of the individual phases in each of these simulations and then finding the point where the temperature, pressure and the chemical potentials of all species were equal in all phases. This approach underpins the application of the NPT + test particle method (Lofti *et al.* (1992)) in conjunction with the Widom test particle insertion method (Widom (1963)), and the Gibbs-Duhem integration technique (Kofke (1993a) and Kofke (1993b)).

Recent reviews of different simulation methods include those by Panagiotopoulos (2000) and Ungerer (2003).

4.2 The Gibbs Ensemble

The direct simulation of phase coexistence is complicated by the presence of an interface between two coexisting phases. Typically in a simulation one tries to minimize the sample size to reduce computational cost (assuming only pair-wise additive interactions, the simulation time typically scales approximately proportional to N^2). For a two-phase simulation of limited size with an explicit interface, however, this would increase the relative effect of the interface since the relative percentage of molecules residing near the interface increases rapidly with decreasing system size (Frenkel and Smit (1996)).

The Gibbs Ensemble (Panagiotopoulos (1987), Panagiotopoulos *et al.* (1988), Smit *et al.* (1989), Smit and Frenkel (1989)) represents perhaps the most widely used simulation technique for the study of first order phase transitions. Gibbs Ensemble simulations are performed in two microscopic regions away from the interface. Each region is simulated within standard periodic boundary conditions using the minimum image convention. The Gibbs Ensemble thus enables the direct simulation of coexisting phases without the difficulty of simulating the interface (see Figure 4-1). This is achieved by a coupling of the phases by means of suitable acceptance / rejection criteria (discussed below).

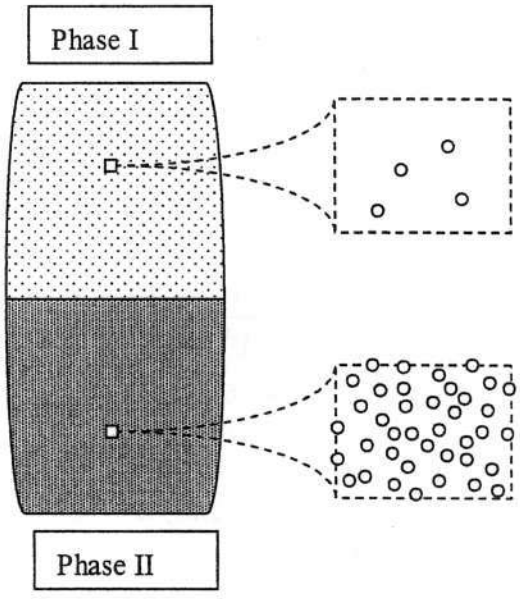


Figure 4-1: Schematic representation of the two coexisting phases in a Gibbs Ensemble simulation.

The Gibbs Ensemble comes in NVT and NPT (or N_1N_2PT) variants with the former variant being the one of choice for one-component simulations whereas the latter variant is the preferable ensemble for binary simulations (the stipulation of P and T providing the specification of the necessary two intensive degrees of freedom for a binary two-phase system). The partition function for the NVT version of the Gibbs Ensemble at temperature T for N total molecules distributed over two volumes V_I and $V = V - V_I$ is given by (Smit and Frenkel (1989), Panagiotopoulos (1992)):

$$Q_{NVT}^{GE} = \frac{1}{K_0 \Lambda^{3N}} \sum_{n_I=0}^{N} \frac{1}{n_I!(N-n_I)!} \int_0^V dV_I V_I^{n_I} (V - V_I)^{N-n_I} \times \int ds_{II}^{N-n_I} \exp\left[-\beta U_{II}\left(s_{II}^{N-n_I}\right)\right] \int ds_I^{n_I} \exp\left[-\beta U_I\left(s_I^{n_I}\right)\right] \tag{4-2}$$

where n_I is the number of molecules in phase I. It has, however, been shown to be advantageous to perform the volume integration over $\ln(V)$ (Eppinga and Frenkel (1984)) in which case the integral becomes (Frenkel and Smit (1996):

$$Q_{NVT}^{GE} = \frac{1}{\Lambda^{3N} K_0} \sum_{n_I=0}^{N} \frac{1}{n_I!(N-n_I)!} \times \int_0^V d\left(\ln\left[\frac{V_I}{V - V_I}\right]\right)\left(\frac{V_I(V - V_I)}{V} V_I^{n_I} (V - V_I)^{N-n_I}\right) \times \int ds_{II}^{N-n_I} \exp\left[-\beta U_{II}\left(s_{II}^{N-n_I}\right)\right] \int ds_I^{n_I} \exp\left[-\beta U_I\left(s_I^{n_I}\right)\right] \tag{4-3}$$

where $\frac{1}{n_I!(N-n_I)!}$ counts the number of ways of choosing n_I molecules from N total molecules. Thus it becomes clear that the probability of observing a configuration around $s_I^{n_I}$ and $s_{II}^{N-n_I}$ with n_I molecules in phase I with volume V_I is given by:

$$\mathcal{P}\left(n_I, V_I, s_I^{n_I}, s_{II}^{N-n_I}\right)_{NVT}^{GE} \propto \frac{V_I^{n_I+1}(V - V_I)^{N-n_I+1}}{V(N-n_I)!n_I!} \times \exp\left[-\beta\left(U_I\left(s_I^{n_I}\right) + U_{II}\left(s_{II}^{N-n_I}\right)\right)\right] \tag{4-4}$$

In order to achieve the above distribution, and to satisfy the requirements for phase coexistence (Section (4.1)) and thermodynamic equilibrium (Section (3.3)), three basic moves are required during simulations in the Gibbs Ensemble:

1) molecule translations to satisfy the requirement for thermal (i.e. temperature) equilibrium,
2) volume exchanges between the two boxes such that the total volume is conserved to satisfy the requirement for mechanical (i.e. pressure) equilibrium, and

3) molecule transfers between the two simulation boxes to satisfy the requirement for chemical (i.e. chemical potential) equilibrium

and are shown in Figure 4-2:

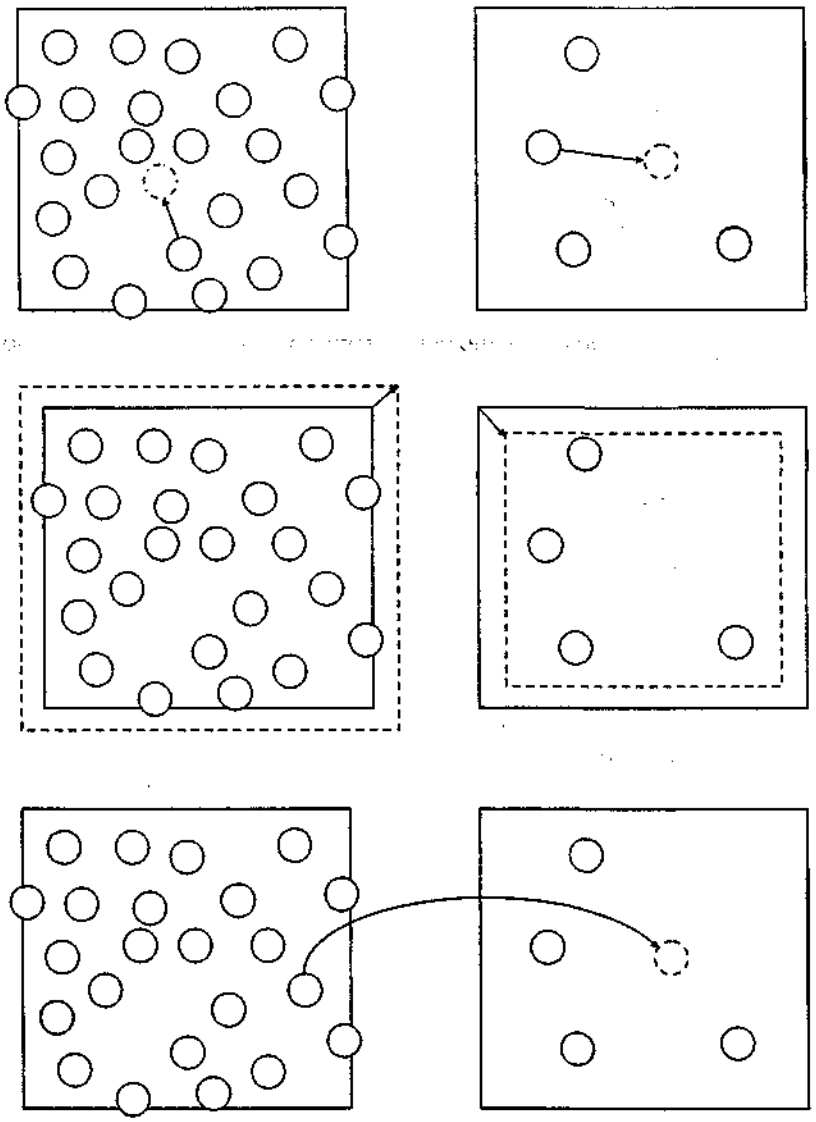


Figure 4-2: The three principle Monte Carlo moves in the Gibbs Ensemble to ensure total thermodynamic equilibrium of two coexisting phases: molecule displacements, volume changes and transfers of molecules.

Using Equations (3–21) and (4–4) it is possible to derive the acceptance criteria for performing:

1) a trial molecule displacement in phase I (Panagiotopoulos (1987)):

$$acc(o \to n) = \frac{\left(\mathcal{P}_{NVT}^{GE}\right)^{new}}{\left(\mathcal{P}_{NVT}^{GE}\right)^{old}} = \min\left\{1, \exp\left[-\beta\left(\Delta U_I\left(\mathbf{s}_I^{n_I}\right)\right)\right]\right\} \quad (4\text{–}5)$$

where the $\min\{\ \}$ function denotes that one accepts a move with a maximum probability of unity, and $\Delta U_I\left(\mathbf{s}_I^{n_I}\right) = U_I^{new}\left(\mathbf{s}_I^{n_I}\right) - U_I^{old}\left(\mathbf{s}_I^{n_I}\right)$,

2) a volume exchange in box I by an amount ΔV, $V_I^{new} = V_I^{old} + \Delta V$ (Panagiotopoulos (1987)):

$$acc(o \to n) = \frac{\left(\mathcal{P}_{NVT}^{GE}\right)^{new}}{\left(\mathcal{P}_{NVT}^{GE}\right)^{old}} = \min\left\{\begin{array}{l} 1, \left(\frac{V_I^{new}}{V_I^{old}}\right)^{n_I+1}\left(\frac{V - V_I^{new}}{V - V_I^{old}}\right)^{N-n_I+1} \\ \times \exp\left[-\beta\left(\Delta U_I\left(\mathbf{s}_I^{n_I}\right) + \Delta U_{II}\left(\mathbf{s}_{II}^{N-n_I}\right)\right)\right] \end{array}\right\} \quad (4\text{–}6)$$

3) a molecule transfer by removing a molecule from box I and inserting this molecule into box II (Panagiotopoulos (1987)):

$$acc(o \to n) = \frac{\left(\mathcal{P}_{NVT}^{GE}\right)^{new}}{\left(\mathcal{P}_{NVT}^{GE}\right)^{old}} = \min\left\{\begin{array}{l} 1, \frac{n_I(V - V_I)}{(N - n_I + 1)V_I} \\ \times \exp\left[-\beta\left(\Delta U_I\left(\mathbf{s}_I^{n_I}\right) + \Delta U_{II}\left(\mathbf{s}_{II}^{N-n_I}\right)\right)\right] \end{array}\right\} \quad (4\text{–}7)$$

Additional moves like molecule rotation (which has an identical acceptance criterion to equation (4–5)) are also permitted, but the above three moves are required to ensure phase equilibrium. The key phase-coupling moves are the molecule transfer move which enforces equal chemical potentials of each species in all phases $\left(\mu_{I,i} = \mu_{II,i} = \ldots = \mu_i\right)$ without requiring *a priori* knowledge of the equilibrium chemical potentials of each of the components, and the volume change move which ensures equality of pressure between the phases $\left(P_I = P_{II} = \ldots = P\right)$ without requiring *a priori* knowledge of the equilibrium pressure. In the original work, the Gibbs Ensemble Monte Carlo method (GEMC) was introduced as a combination of the μVT, NPT and NVT ensembles. It has, however, been shown that the NVT version of the Gibbs Ensemble is rigorously equivalent to the NVT ensemble in the thermodynamic limit of an infinite system size (Smit *et al.* (1989)).

The Gibbs Ensemble may also be written for the case where the total system pressure is kept constant (Panagiotopoulos *et al.* (1988)). This NPT version of the Gibbs Ensemble is ideally suited for the study of binary vapour-liquid equilibrium (VLE) and for the simulation of non-saturated phase behaviour of pure compounds because two intensive degrees of freedom are required for both of these cases. This may be seen from Equation (3–24). The essential difference between the NVT and NPT versions of the Gibbs Ensemble is that the box volumes are allowed to change volumes

independently in the NPT Gibbs Ensemble while in the NVT version of the Gibbs Ensemble the total system volume of both boxes is conserved. For a binary two-phase system, with N_1 molecules of type 1 where $N_1 = n_{I,1} + n_{II,1}$, and N_2 molecules of type 2 where $N_2 = n_{I,2} + n_{II,2}$ the partition function for the N_1N_2PT Gibbs Ensemble is:

$$\begin{aligned} Q^{GE}_{N_1N_2PT} = \frac{1}{\Lambda_1^{3N_1}\Lambda_2^{3N_2}K_0}\sum_{n_{I,1}=0}^{N_1}\sum_{n_{I,2}=0}^{N_2}\frac{1}{n_{I,1}!\left(N_1-n_{I,1}\right)!n_{I,2}!\left(N_2-n_{I,2}\right)!} \\ \times\int d\left(\ln V_I\right)\left(V_I^{n_1+1}\exp\left[-\beta PV_I\right]\right) \\ \times\int d\left(\ln V_{II}\right)\left(V_{II}^{N-n_1+1}\exp\left[-\beta PV_{II}\right]\right) \\ \times\left\{\int ds_I^{n_1}\exp\left[-\beta U\left(\mathbf{s}_I^{n_1}\right)\right]\int ds_{II}^{N-n_1}\exp\left[-\beta U\left(\mathbf{s}_{II}^{N-n_1}\right)\right]\right\} \end{aligned} \tag{4–8}$$

K_0 now has units proportional to the square of the volume. Thus:

$$\begin{aligned} \mathcal{P}\begin{pmatrix} n_{I,1}, n_{I,2}, V_I, \\ V_{II}, \mathbf{s}_I^{n_1}, \mathbf{s}_{II}^{N-n_1} \end{pmatrix}^{GE}_{N_1N_2PT} \propto \frac{V_I^{(n_1+1)}V_{II}^{(N-n_1+1)}}{n_{I,1}!\left(N_1-n_{I,1}\right)!n_{I,2}!\left(N_2-n_{I,2}\right)!} \\ \times\exp\left[-\beta\left(P\left(V_I+V_{II}\right)+U\left(\mathbf{s}_I^{n_1}\right)+U\left(\mathbf{s}_{II}^{N-n_1}\right)\right)\right] \end{aligned} \tag{4–9}$$

The molecule transfer and molecule translation acceptance criteria for this ensemble are identical to the acceptance criteria for the NVT Gibbs Ensemble but the volume move now has a different acceptance criterion. Thus for an increase in the volume of phase I of ΔV_I, $V_I^{new} = V_I^{old} + \Delta V_I$:

$$\begin{aligned} acc\left(o\to n\right) &= \frac{\left(\mathcal{P}^{GE}_{N_1N_2PT}\right)^{new}}{\left(\mathcal{P}^{GE}_{N_1N_2PT}\right)^{old}} \\ &= \min\left\{1,\left(\left(\frac{V_I^{new}}{V_I^{old}}\right)^{n_1+1}\exp\left[-\beta\left(P\left(\Delta V_I\right)+\Delta U_I\left(\mathbf{s}_I^{n_1}\right)\right)\right]\right)\right\} \end{aligned} \tag{4–10}$$

It is important to note that the total number of molecules is fixed in both versions of the Gibbs Ensemble.

The molecule transfer step is crucial for the algorithm to work properly. This makes the Gibbs Ensemble difficult to implement for dense phases since the probability of inserting a molecule successfully is then very low. This tends to be a weakness of all 'chemical potential based' algorithms.

4.3 The Semi-Grand Ensemble

Before the advent of Configurational-Bias techniques (discussed in Section (5.7)), the insertion or deletion of molecules in μVT ensemble simulations proved prohibitively difficult. The molecule transfer step is also crucial for the Gibbs Ensemble algorithm. The Gibbs and μVT ensembles thus tend to be awkward to implement successfully for dense phases since the probability of inserting a molecule is then very low.

As an alternative to the difficulty that will be encountered when trying to insert a molecule in a simulation box for an ensemble in which the total volume is fixed (e.g. the μVT or NVT version of the Gibbs Ensemble), one might consider trying to change molecule identities for the case of a multicomponent mixture. In this case, the total number of molecules will remain constant and thus each identity change corresponds to the destruction of a molecule of a certain species and the creation of a molecule of a different species type and is shown schematically in Figure 4-3. This coupling of a molecule destruction to the creation of a molecule of another type will in general be more successful than the separate creation and destruction of molecules when sampling configurations. This is precisely the reasoning behind the formulation of the Semi-Grand Ensemble (Kofke and Glandt (1988)).

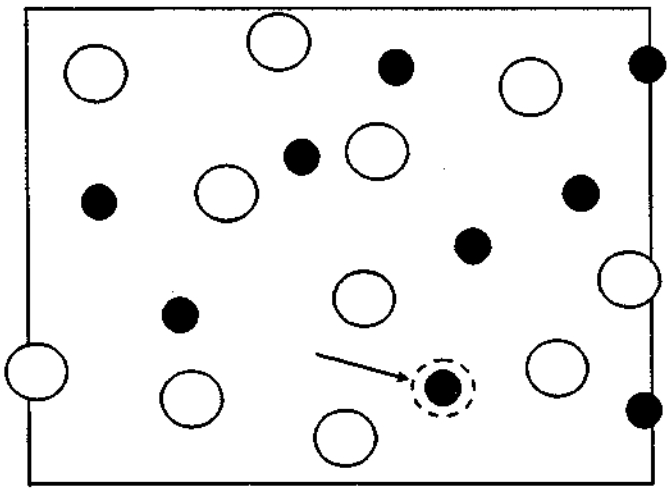


Figure 4-3: Molecule identity change move in the Semi-Grand Ensemble. The smaller molecule is being transformed into the same type as the larger molecule.

In the original formulation, the Semi-Grand Ensemble was derived for the case of constant simulation box volume. For a multicomponent mixture, however, it will in general be necessary to

specify at least two degrees of freedom (Equation 3–24). Consequently, it is most convenient to write the partition function for Semi-Grand Ensemble at constant pressure (Frenkel and Smit (1996)):

$$Q_{NPT}^{SG} = \frac{1}{K_0 \Lambda_1^{3N} N!} \int_0^V dV V^N \exp[-\beta PV] \times \left\{ \sum_{\text{identities}} \prod_{i=1}^{C} \left(\frac{\xi_i}{\xi_1} \right)^{n_i} \int ds^N \exp\left[-\beta U\left(s^N\right) \right] \right\} \tag{4–11}$$

where Q_{NPT}^{SG} is the isobaric Semi-Grand partition function, C is the number of components, n_i is the number of molecules of type i and ξ_i is the fugacity fraction of species i which varies between 0 and 1 and is defined as:

$$\xi_i \equiv \frac{f_i}{\sum_{j=1}^{m} f_j} \tag{4–12}$$

where f_i is the fugacity of species i and is defined in terms of the chemical potential of species i as (Frenkel and Smit (1996)):

$$\mu_i\left(P,T,\{x_i\}\right) \equiv \mu_i^0\left(T\right) + k_B T \ln\left[f_i\right] \tag{4–13}$$

where x_i is the mole fraction of component i, and where

$$\mu_i^0\left(T,\{P=1\}\right) = k_B T \ln\left[\beta \Lambda_i^3\right] = k_B T \ln\left[\beta \Lambda_i^3 P\right] = k_B T \ln\left[\rho_i \Lambda_i^3\right] \tag{4–14}$$

where $\mu_i^0\left(T\right)$ is the chemical potential in the ideal gas reference state ($P=1$) of species i (Frenkel and Smit (1996)). The summation over identities in Equation (4–11) indicates that the partition function must be evaluated over all possible identities for the N molecules in the mixture. It must be noted that at equilibrium, $\left(f_{I,i}=f_{II,i}=\ldots=f_i\right)$ because $\left(\mu_{I,i}=\mu_{II,i}=\ldots=\mu_i\right)$ at equilibrium (Smith *et al.* (1996)). The probability of observing a configuration around s^N of N molecules in a volume V with n_i molecules of type i in the Semi-Grand Ensemble is given by:

$$\wp\left(N,V,s^N,n_i\right)_{NPT}^{SG} \propto V^N \exp[-\beta PV] \left(\frac{\xi_i}{\xi_1} \right)^{n_i} \exp\left[-\beta U\left(s^N\right) \right] \tag{4–15}$$

Using Equation (3–21) and Equation (4–15), it is possible to show that the acceptance criterion for changing the identity of a molecule of species i into a molecule of species j is given by:

$$acc\left(o \to n\right) = \frac{\left(\wp^{SG}\right)^{\text{new}}}{\left(\wp^{SG}\right)^{\text{old}}} = \min\left\{1, \frac{\xi_j}{\xi_i} \exp\left[-\beta\left(\Delta U\left(s^N\right)\right) \right] \right\} \tag{4–16}$$

It may be appear that this ensemble is directly applicable to the simulation of phase equilibrium. The probability given in Equation (4–15) includes specification of two intensive variables, $\beta = 1/k_B T$ and P, as well as a ratio of two intensive variables, ξ_i/ξ_1. Frenkel and Smit (1996) outline a procedure to determine phase equilibrium at a given β and P. The procedure is, unfortunately, iterative in nature. For a binary mixture this method requires the gradual insertion of molecules of type 1 into a phase I composed of pure molecules of type 2. Thereafter one performs a similar task for a second phase II, but instead, molecules of type 2 are added to a phase initially composed of molecules of type 1 only. Equilibrium properties can, therefore, not be obtained from a direct simulation. The Semi-Grand Ensemble does, however, have considerable advantage for the case of multicomponent systems where the number of components present is greater than two. This is because the composition of a phase is then determined in an averaged statistical sense. In contrast, the Gibbs Ensemble requires that the overall composition for the components be specified *a priori*. The difficulty inherent in *a priori* specification of overall composition is explained with reference to the well known "lever-arm rule" in Section (8.1). In Sections (8.2) and (8.3) two methods that marry the advantages of the Gibbs Ensemble and the Semi-Grand Ensemble by the use of new ensemble formulations are described.

The characteristic thermodynamic potential for this isobaric Semi-Grand Ensemble is given by (Frenkel and Smit (1996)):

$$\beta\mu_1 N = -\ln\left[Q_{NPT}^{SG}\right] \tag{4–17}$$

The presence of the term $\mu_1 N$ in Equation (4–17) and the fugacity ratio ξ_i both originate from the specification that the total number of molecules is constant. It is important to note that in this formulation of the Semi-Grand Ensemble, the partition function of Equation (4–11) applies only to the simulation of one phase and not to the simulation of two-coexisting phases explicitly.

Panagiotopoulos (1989) has proposed an additional move for the Gibbs Ensemble which corresponds to an identity swap between two molecules in two different but co-existing phases. The move is in principle similar to the identity swap move of the Semi-Grand Ensemble but differs in the sense that the identity swap is performed in both phases of the Gibbs Ensemble in such a way that the total number of molecules of each species involved in the swap move are conserved, i.e. n_1 and n_2 remain unchanged. The acceptance criterion for this move is given by:

$$acc(o \to n) = \frac{\left(\mathcal{Q}^{GE}_{N_1 N_2 PT}\right)^{new}}{\left(\mathcal{Q}^{GE}_{N_1 N_2 PT}\right)^{old}}$$

$$= \min\left\{1, \left[\frac{(N_1 - n_{I,1})}{(n_{I,1}+1)} \frac{n_{I,2}}{(N_2 - n_{I,2}+1)} \times \exp\left[-\beta\left(\Delta U_I\left(\mathbf{s}_I^{n_1}\right) + \Delta U_{II}\left(\mathbf{s}_{II}^{N-n_1}\right)\right)\right]\right]\right\} \quad (4\text{–}18)$$

Although this move in the Gibbs Ensemble is in principle similar to the identity swap move of the Semi-Grand Ensemble, the inverse identity swap move which is performed simultaneously in the other simulation box has the net effect of cancelling off the fugacity coefficient ratio in the acceptance criterion for this move.

The identity swap move has the effect of ensuring that the chemical potential difference between two components in a mixture remains constant across the phases. The identity swap move is therefore of particular use when one of the components is considerably more difficult to swap successfully. For a binary and two-phase system, if one performs conventional molecule swaps for the more easily swapped component according to Equation (4–7) and also identity changes, the following conditions are therefore satisfied:

$$\begin{aligned} \mu_{I,1} &= \mu_{II,1} \\ \mu_{I,1} - \mu_{I,2} &= \mu_{II,1} - \mu_{II,2} \end{aligned} \quad (4\text{–}19)$$

The first line of Equation (4–19) originates from the transfer steps while the second line originates from the identity swaps. Equation (4–19) is sufficient to ensure equality of the chemical potentials of the species present. Martin *et al.* (2000) and Panagiotopoulos (1989) both report that the relative acceptance of the identity swap move is one to two orders of magnitude higher than that for the conventional molecule swap move of Equation (4–7). In general, therefore, for an n-component and two-phase mixture, then only one species swap move and (n-1) distinct identity swap moves are necessary to ensure chemical equilibrium. This move is therefore of considerable advantage for the case of simulating binary and multicomponent VLE.

4.4 Pseudo-Ensemble Alternatives to the Gibbs Ensemble

Although the primary focus of this work is the use of the Gibbs Ensemble to simulate phase equilibrium, pseudo-ensembles represent an important new class of simulation technique that have emerged in recent years. As pointed out by Escobedo (1998), the advent of pseudo-ensembles has greatly enhanced the numerical tools available to molecular simulation. Mehta and Kofke (1995)

first proposed a scheme to mimic a traditional μVT simulation by means of an NPT type of ensemble in which molecule insertions and deletions were replaced by volume fluctuations. Attard (1997) has also proposed a similar approach. Both schemes rely on virtual or "faked" insertion moves to evaluate the instantaneous chemical potential of the system being simulated (*e.g.* by test-particle insertion methods such as the Widom test-particle insertion method of Widom (1963)).

Camp and Allen (1996) subsequently used the idea of replacing molecule insertions with volume moves to formulate a pseudo Gibbs Ensemble in which molecule transfers are replaced by concerted volume moves. Instantaneous evaluation of the chemical potential is again a requirement of this method. Escobedo and de Pablo (1997) then proposed a variant of this method which essentially allows for the substitution of volume moves with molecule transfers in another pseudo Gibbs Ensemble. This method avoids direct volume changes and is advantageous for polymeric systems.

Using a general formalism, Escobedo (1998) has been able to unify pseudo-ensembles with Histogram Reweighting techniques and NPT + test particle methods (Möller and Fischer (1990), Lofti *et al.* (1992), Boda *et al.* (1995)). In particular, it has been explicitly shown that NPT + test particle methods represent low order extrapolation schemes of pseudo-ensembles using Taylor series while Histogram Reweighting is effectively an infinite order pseudo-ensemble extrapolation scheme. The formalism also suggested the formulation of an interpolative approach for the direct simulation of dew- and bubble-points.

The idea of combining pseudo-ensembles with multiple phases has spawned several new molecular simulation approaches. Bode *et al.* (1996a) used separate μVT simulations to determine third order Taylor series of the pressure as a function of the imposed values of chemical potential and temperature around a neighbourhood of $\{T_0, \mu_0\}$. In a sense, this method can be thought of as the reverse approach to that of an NPT + test particle simulation. This is because an NPT + test particle simulation achieves phase coexistence by enforcing pressure and temperature and then determining the point at which the chemical potentials of two phases are equal, while the approach of Boda *et al.* (1996) first imposes the chemical potentials and then seeks the conditions at which the chemical potentials of the two phases are the same.

In a related approach, Vrabec and Hasse (2002) have formulated a Grand Equilibrium method for pure components and mixtures which also uses Taylor series expansions. The method rather uses

NPT simulations to evaluate the chemical potentials of all components and their derivatives at constant pressure to formulate first-order Taylor series expansions for the chemical potentials as functions of pressure at constant composition. Next, a pseudo-μVT simulation is performed in which the chemical potentials are set according to the instantaneous pressure (which must be calculated during this second simulation) according to the functions which have been determined for the chemical potentials. In this way, the outputted chemical potential function from the first simulation of the one phase is used as an input for the second simulation of the other phase. The decoupling of the phases, the expansions used for the chemical potentials and the multiple simulations required to evaluate the functions for the chemical potentials necessarily place this approach in the class of pseudo-ensembles multiphase pseudo-ensembles.

In order to compensate for the well-known inability of the Gibbs Ensemble to simulate near-critical data accurately, Miyano (1997) proposed an iterative equation of state algorithm-based Gibbs Ensemble approach. In this scheme, the pressure and vapour composition are systematically refined until equality of chemical potentials between the vapour and the liquid was obtained. This algorithm did allow for the simulation of binary data much closer to the critical point without the usually large fluctuations found in the Gibbs Ensemble.

Recently, the Bubble Point Ensemble algorithm has been developed and used for the simulation of mixtures of fluids with simple Lennard-Jones interactions (Ungerer *et al.* (1999)) as well as for alkane mixtures of chain molecules (Ungerer *et al.* (2001)). The idea behind this ensemble was also suggested by Escobedo (1999). This development has been significant because it has facilitated the simulation of bubble points at a fixed liquid composition by a direct algorithm. The method relies heavily on fake molecule insertions and deletions in the liquid phase and is shown schematically in Figure 4-4:

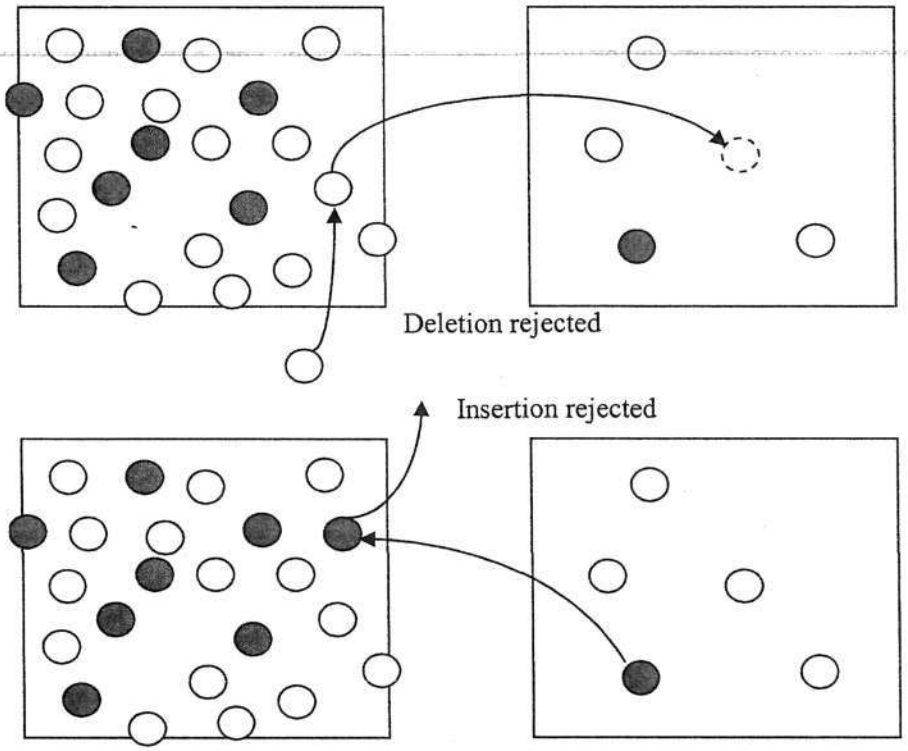


Figure 4-4: Diagram showing how insertions and deletions of molecules in the liquid phase are never actually carried out.

In this method, the composition of the liquid phase is kept constant. The usual Gibbs Ensemble moves are used except that transfers of molecules to or from the liquid phase are never actually performed. The liquid phase thus remains at constant composition while the vapour phase undergoes effective changes. These "fake" or virtual transfer moves require modifications to the usual transfer acceptance criteria (Ungerer *et al.* (1999)). The simulation proceeds by imposing the chemical potential calculated in the liquid phase by "fake" molecule insertions in the liquid phase on the vapour phase. This imposed chemical potential is progressively refined during the simulation. Unlike the methods of Boda *et al.* (1996) and Vrabec and Hasse (2002), the Bubble Point Ensemble thus has direct mechanical coupling of the coexisting phases which has been shown by Ungerer *et al.* (1999) to improve convergence. A hybrid method was proposed in which the average box sizes and average molecule numbers from the Bubble Point Ensemble are used to initialize an NVT Gibbs Ensemble simulation. This approach was selected because the Bubble Point Ensemble was found to have larger density fluctuations than the Gibbs Ensemble (Ungerer *et al.* (1999), Ungerer *et al.* (2001)). This might be attributable to slight errors which creep into the calculation of the chemical potential which is imposed on the vapour phase. Following the

discussion from Section (3.3), it is important to note that an NVT ensemble has too few intensive parameters specified for a two-phase and two-component system. Clearly, only T is specified for the final NVT Gibbs Ensemble when there are in fact two intensive degrees of freedom. It therefore becomes very important to sample the chemical potential in the liquid phase over a sufficiently long interval such that a good chemical potential is imposed on the vapour box.

The key difference between the pseudo-ensembles and pseudo Gibbs Ensembles discussed thus far and conventional Gibbs Ensemble simulations rests in the way in which two coexisting phases are coupled at the level of microscopic fluctuations in Gibbs Ensemble (e.g. molecule transfers which implicitly assume equality of chemical potentials of individual components in all phases), whereas these fluctuations are selectively decoupled for any pseudo-ensemble approach. The virtual molecule transfers to the liquid phase in the Bubble Point Ensemble are an example of this decoupling. The key issue, though, is that this convenient decoupling of certain types of moves in pseudo-ensembles allows one to avoid some of the mass balance and volume constraints which can hinder a standard Gibbs Ensemble simulation (Escobedo (1999)). An example of this is the way in which a constant liquid composition can be imposed in the Bubble Point Ensemble while still allowing molecule transfers to satisfy the requirement for chemical equilibrium.

Pseudo-ensembles can, however, suffer from the disadvantage that inaccuracies present in intermediate computed variables propagate as errors in imposed variables. This can cause errors of varying magnitude in the final computed results. The higher the desired accuracy or extrapolation order, the longer a simulation must be run. Escobedo (1998) alluded to the necessary compromise that must be struck between simulation length and the extrapolation order when considering using different pseudo-ensemble techniques. Also, Ungerer *et al.* (1999) have pointed out that the repeated evaluation of chemical potentials in methods like those mentioned by Escobedo (1998) and Miyano (1998) can be very expensive computationally. Systematic deviations from GEMC simulations errors have also been found in certain cases. Vrabec and Hasse (2002) found the isothermal dew line of a Lennard-Jones fluid mixture to be underestimated, while this result is also evident in the results from Miyano (1997) for points away from the critical point.

There has been some debate concerning the validity of microscopic reversibility as a necessary condition for a simulation to be valid. For example, during a simulation it is common practice to adjust the maximum rotation or translation step size so that a predetermined target acceptance of rotation and translation moves are achieved by the end of the simulation. Small translations or

rotations will have a high probability of being accepted but will sample configuration space less efficiently. Miller *et al.* (2000) have argued, however, that adjusting step size can bias the potential energy surface particularly for inhomogeneous systems. It was shown explicitly that this can lead to erroneous results since the system is no longer truly microscopically reversible. Manousiouthakis and Deem (1999), however, have argued that since most Monte Carlo simulation moves lead to ergodic sampling, the local balance condition together with the requirement for regular sampling is sufficient to ensure a correct simulation.

Pseudo-ensembles are, in general, non-ergodic due to the use of interpolation schemes and "fake" moves. "Fake" or virtual moves are inherently non-ergodic. For a simulation to be unconditionally ergodic, the reverse move must be possible to satisfy the condition of microscopic reversibility. This is not possible for virtual moves. This can lead to ergodicity errors. It can be shown that there are circumstances for which the condition of the microscopic reversibility can be relaxed (Miller *et al.* (2000), Rull *et al.* (1995)). The point, however, remains that while there may be circumstances for which microscopic reversibility is an unnecessarily strong requirement, the imposition of this condition will always guarantee that a simulation is ergodic (Frenkel and Smit (1996)). It has also been shown that "real" moves lead to a faster convergence (Escobedo (1999)). The use of pseudo-ensembles to take advantage of the array of simulation options provided by these simulation tools must therefore be done with careful consideration and on a case-by-case basis (Escobedo (1999)).

4.5 Histogram Reweighting

Histogram Reweighting represents an exciting new alternative to the simulation of phase equilibrium properties. Although the primary focus of this work is the simulation of phase equilibrium using the Gibbs Ensemble, some attention must be given to this new method because it represents a simulation methodology that is receiving renewed attention.

As pointed out by Panagiotopoulos (2000), it has been known for some time that a single calculation can, in principle, be used to obtain information for a system for a range of state conditions. Grand canonical Histogram Reweighting makes use of this idea. In general, multiple overlapping histograms at different temperatures and chemical potentials will need to be collected to cover the conditions of interest. Once this has been achieved, all the thermodynamic properties over the region of interest can be obtained.

Ferrenberg and Swendsen (1988) originally proposed a means by which to combine multiple histograms to obtain a global free energy function by minimizing the differences between predicted and observed histograms. Subsequently, Ferrenberg and Swendsen (1989) also proposed an optimized method to combine the collected histograms for a wide range of parameter values of different Hamiltonians. The implementation of Histogram Reweighting has additionally been described by Swendsen (1993), Potoff and Panagiotopoulos (1998), and Errington and Panagiotopoulos (1998a), Errington and Panagiotopoulos (1998b) and Errington and Panagiotopoulos (1999a) have also outlined a general procedure for generating pure component coexistence data effectively.

Most recently, the method has been used to parameterise intermolecular force fields for *n*-alkanes (Errington and Panagiotopoulos (1999a)), benzene and cyclohexane (Errington and Panagiotopoulos (1999b)), carboxylic acids (Kamath *et al.* (2004)), as well as glycols, ethers, ketones and aldehydes (Stubbs *et al.* (2004). Conrad and de Pablo (1998) have developed an alternative NPT ensemble Histogram Reweighting technique.

While Histogram Reweighting works well for pure components and while the method has been applied to the simulation of some alcohol mixtures (Chen *et al.* (2001)), carbon dioxide and nitrogen mixtures (Potoff and Siepmann (2001)), polar and non-polar mixtures (Potoff *et al.* (1999)) and water-alkane mixtures (Errington *et al.* (1998c) and Boulougouris *et al.* (2000)), the increasing number of histograms that are required to simulate multicomponent mixtures make this method computationally expensive for mixtures with two or more components (Ungerer *et al.* (1999). The method does, however, have the distinct advantage that the simulated data is not limited to discrete simulated points as for the Gibbs Ensemble and for certain cases of pseudo-ensembles. Also, Histogram Reweighting is ideally suited to the simulation of near-critical phase equilibrium data. In this way, it can be used to simulate mixtures over a wide temperature range even though the computational overload is increased (Potoff *et al.* (1999)).

4.6 The Reaction Gibbs Ensemble

Recently, the reaction Gibbs Ensemble Monte Carlo (RGEMC) method has been proposed as a means to more accurately model vapour-liquid phase equilibrium using a methodology closely related to the standard Gibbs Ensemble approach (Lísal *et al.* (1999)). The method is based on the Reaction Gibbs Ensemble method of Smith and Triska (1993) and has been used to study binary

mixtures containing isobutene, MTBE and *n*-butane (Lísal *et al.* (1999)), binary mixtures containing water, methanol, ethanol, carbon dioxide and ethane (Lísal *et al.* (2001)), as well as multicomponent and reacting multicomponent phase equilibrium. While this method appears to improve the simulated phase diagrams over conventional GEMC simulations, it only scales the simulated phase diagram by a simple pressure adjustment to the acceptance rules and, more importantly, RGEMC does so by shifting the true composition predicted by the simulation. Furthermore, the improved results given by the RGEMC may be achieved by simply applying the Gamma/Phi approach to VLE (Smith *et al.* (1996)) using a direct scaling methodology. In Appendix 4B of this Chapter, explicit details of this methodology and its derivation are given.

The RGEMC also does not address the issue presented by random fluctuations in the box volume in an NPT GEMC or NPT RGEMC simulation which can cause a simulation to fail due to the box length dropping below twice the cut-off radius (see Section (5.2)). This is because although the simulated pressures are now shifted closer to the experimental results, one still has to judiciously select the pressures for the Reaction Gibbs Ensemble simulation run. Thus, the RGEMC method does not obviate the difficulty of satisfying mass balance constraints imposed by the "lever-arm" rule referred to in Section (8.1) for binary simulations in which both N_1 and N_1 are specified *a priori*.

4.7 A Generalized Isothermal Gibbs Ensemble Formalism

Hill (1956) has developed the ensemble formalism behind a general case of a system in material, chemical and thermal contact with its surroundings. This 'generalized ensemble' effectively constitutes a μPT ensemble, i.e., an ensemble at constant chemical potential, temperature and pressure. Hill (1956) has, however, argued that this ensemble is of little practical use because T, P and all of the μ_i's cannot all be independent for reasons mentioned in Section (4.1). This 'generalized' ensemble does, however, have an important application. Indeed, it can be shown to provide a convenient starting point to 'derive' all other ensembles (which may therefore be regarded as special cases of the 'generalized' ensemble). It also can directly yield the characteristic thermodynamic potential for an ensemble which provides the link to macroscopic thermodynamics. In this Section, the generalized ensemble is given for the general case of multiple phases and multiple components. Several other ensembles are shown to be simplifications of this ensemble. Two new ensemble formalisms are also "derived" from this approach.

Escobedo (2001a) has argued that much still needs to be done to develop methodological frameworks that inter-relate different simulation approaches. Almost 15 years ago, Graben and Ray (1991) provided a unified treatment of adiabatic ensembles. In this analysis they provided the fundamental partition functions and fundamental entropy relations for four monophasic adiabatic ensembles. This analysis of adiabatic ensembles was significant because it suggested the formulation of a governing framework from which several classes of ensembles can be deduced. The analysis was, however, limited to pure components. Escobedo has invested considerable effort in providing generalized frameworks with which to consider pseudo-ensembles (Escobedo (1998)), multiphase pseudo-ensemble approaches collectively referred to as "virtual Gibbs Ensembles" (Shetty and Escobedo (2002)) and polydisperse systems (Escobedo (2001a)). The applications have been limited but have included the simulation of a simple Lennard-Jones mixtures in a bubble-point pseudo-ensemble (Shetty and Escobedo (2002), Escobedo (1998)) essentially equivalent to the Bubble Point Ensemble of Ungerer *et al.* (1999), the simulation of a simple Lennard-Jones mixture in a semi-open Gibbs Ensemble (Escobedo (2000)), as well as the study of some polydisperse fluid mixture behaviour (Escobedo (2003) and Escobedo (2001b)). More recently, Escobedo (2000b) has used pseudo-ensembles to study alkane mixtures that could be of relevance to distillation operations.

The aim of the following section of work is to use a generalized ensemble formalism to develop a methodological means by which to "generate" ensemble partition functions and the corresponding thermodynamic potentials.

The fundamental relation in thermodynamics may be expressed as (Rao (1994)):

$$U = U\left(S, V, N_1, N_2, ..., N_C\right) = U\left(\mathbf{X}'\right) \tag{4–20}$$

where U is total internal energy and is a function of extensive parameters only, S is the entropy, V is total volume, N_i is total number of molecules of species i and $\mathbf{X}'$ is the set of $C+2$ extensive variables in the entropy representation of the fundamental relation. Once the fundamental relation is known, the system under study is completely defined and all relevant thermodynamic properties are then obtainable (Rao (1994)). Since entropy is continuous, differentiable and single-valued, Equation (4–20) may alternatively be expressed as:

$$S = S\left(U, V, N_1, N_2, ..., N_C\right) = S\left(\mathbf{X}\right) \tag{4–21}$$

which is known as the entropy representation of the fundamental relation and where $\mathbf{X}$ is now the set of $C+2$ extensive parameters in the entropy representation of the fundamental relation. Note

that all extensive properties are total properties in the sense that they are sums over all subsystems such that $X_i = X_{I,i} + X_{II,i} + ...$ where X_i is a particular extensive variable, and that the number of components ranges from 1 to C.

The exclusive extensive variable dependence of Equation (4–21) or Equation (4–20) is inconvenient since they do not coincide with the understanding of Equation (3–24) which requires the specification of intensive variable parameters. Following the approach of Escobedo (2000a), if $\mathbf{X}$ is the set of all extensive variables such that $\mathbf{X} = \mathbf{X}(U, V, N_1, N_2, ..., N_C)$ and if $\mathbf{Y}$ is the set of all corresponding conjugate intensive variables such that $\mathbf{Y} = \mathbf{Y}(\beta, \beta P, -\beta\mu_1, -\beta\mu_2, ..., -\beta\mu_C)$, then it is possible to replace a subset $\tilde{\mathbf{X}} = \tilde{\mathbf{X}}(\tilde{X}_1, \tilde{X}_2, ...)$ of $\mathbf{X}$ consisting of $M_{\tilde{X}}$ extensive variables with a subset $\tilde{\mathbf{Y}} = \tilde{\mathbf{Y}}(\tilde{Y}_1, \tilde{Y}_2, ...)$ of $\mathbf{Y}$ consisting of $M_{\tilde{Y}}$ $(=M_{\tilde{X}})$ intensive variables where $\tilde{X}_i$ and $\tilde{Y}_i$ are particular conjugate extensive and intensive quantities respectively. This formally allows one to arrive at the familiar ensemble abbreviations (e.g. NPT, NVT, etc) which are a mixture of intensive and extensive parameters which satisfy the intensive property requirements of Equation (3–24).

Each intensive variable is related to its conjugate extensive variable by means of:

$$Y_i = \frac{1}{k_B}\left(\frac{\partial S}{\partial X_i}\right)_{X_{j\neq i}} \tag{4–22}$$

Details concerning the specifics of Legendre transforms of thermodynamic potentials are contained in Hill (1956) and Rao (1994). As pointed out by Escobedo (2000a), it is important to note that:

$$\mathbf{Y} = \bar{\mathbf{Y}} \cup \tilde{\mathbf{Y}} \tag{4–23}$$

and

$$\mathbf{X} = \bar{\mathbf{X}} \cup \tilde{\mathbf{X}} \tag{4–24}$$

where $\bar{\mathbf{X}}$ and $\bar{\mathbf{Y}}$ are the respective extensive and intensive subsets of $\mathbf{X}$ and $\mathbf{Y}$ not in $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$.

It is known that the entropy is a homogenous first order function of the extensive variables. Consequently, it is possible to derive the Euler relation for the entropy representation for a system of C components (Rao (1994)):

$$\frac{S}{k_B} = \beta U + \beta P V - \sum_{i=1}^{C} \beta\mu_i N_i \tag{4–25}$$

in which the internal entropy (or equivalently the energy after trivial re-arrangement of Equation (4–25)) is the sum of each intensive variable with its conjugate extensive variable. Clearly:

$$\frac{S}{k_B} = \mathbf{X}\cdot\mathbf{Y} \tag{4–26}$$

Using the Legendre transform rules (Rao (1994)), it becomes possible to derive a "Massieu" function in the entropy representation, $\mathcal{F}_{\tilde{Y}}$, to give an alternative representation of the fundamental relation. In particular:

$$\begin{aligned}\mathcal{F}_{\tilde{Y}} &= \frac{S}{k_B} - \tilde{\mathbf{X}}\cdot\tilde{\mathbf{Y}} = \frac{S}{k_B} - \sum_{i=1}^{M_\beta} \tilde{X}\tilde{Y} \\ &= \bar{\mathbf{X}}\cdot\bar{\mathbf{Y}} = \sum_{i=1}^{M_\beta} \bar{X}\bar{Y}\end{aligned} \tag{4–27}$$

The final line of Equation (4–27) arises because $\mathbf{X}\cdot\mathbf{Y} = \bar{\mathbf{X}}\cdot\bar{\mathbf{Y}} + \tilde{\mathbf{X}}\cdot\tilde{\mathbf{Y}}$. This may be seen directly from Equations (4–23) and (4–24).

At this point it is convenient to note that all the thermodynamic potentials for all of the ensembles described thus far contain only some of the terms present in Equation (4–25). At equilibrium, however, the system is uniquely specified by the $C+2$ extensive variables in $\mathbf{X}$, and the requirements of phase equilibrium embodied in Equation (4–1) must hold for all extensive parameters.

It should therefore be possible to arrive at a hypothetical general ensemble from which all ensemble partition functions may be readily determined by simplification, i.e. as special cases. In this case, one would desire to express the fundamental relation as:

$$\frac{S}{k_B} = \frac{S}{k_B}(\beta, \beta P, -\beta\mu_1, -\beta\mu_2, \ldots, -\beta\mu_C) = \frac{S}{k_B}(\mathbf{Y}) \tag{4–28}$$

where the entropy is now exclusively a function of intensive parameters, i.e. a total Legendre transform of all extensive quantities to intensive quantities. The partition function for this general ensemble must therefore involve integration or summation over all of the extensive variables present in Equation (4–25).

A convenient starting point for this ensemble is the 'generalized ensemble'. According to Hill (1956) the partition function for the generalized ensemble may be given by:

$$Q_{\mu PT} = \int dV Q_{\mu VT} \exp[-\beta PV] \tag{4–29}$$

where $Q_{\mu VT}$ is given by Equation (3–11) and is now integrated over all possible volumes. Since this is effectively the partition function for a μPT ensemble, Equation (4–29) is equivalent to integrating

any other ensemble over the remaining unspecified intensive variables, e.g. integrating the NPT ensemble over all possible numbers of molecules:

$$Q_{\mu PT} = \sum_N \left(Q_{NPT} \exp\left[\beta \mu N\right] \right) \tag{4–30}$$

Substituting Equation (3–11) for $Q_{\mu VT}$ in Equation (4–29) or substituting Equation (3–10) for Q_{NPT} in Equation (4–30), one arrives at:

$$Q_{\mu PT} = \sum_{N=1}^{\infty} \left\{ \begin{array}{l} \exp\left[\beta \mu N\right] \dfrac{1}{K_0 N! \Lambda^{3N}} \int dV \left(V^N \exp\left[-\beta P V\right] \right) \\ \times \int d\mathbf{s}^N \exp\left[-\beta U\left(\mathbf{s}^N\right) \right] \end{array} \right\} \tag{4–31}$$

where the integration over volume is performed over V and not $\ln[V]$. Expanding Equation (4–31) for the more general case of C components and F phases yields:

$$Q_{\mu PT} = \frac{1}{K_0} \sum_{N_1, N_2, \ldots N_C}^{\infty} \left\{ \begin{array}{l} \prod_{i=1}^{C} \dfrac{\exp\left[\beta \mu_i n_i\right]}{\Lambda_i^{3n_i} \prod_{j=1}^{F} n_{ij}!} \times \\ \int_{V_F} \cdots \int_{V_{II}} \int_{V_I} \prod_{j=1}^{F} \left\{ \begin{array}{l} dV_j \left(V_j^{n_j} \exp\left[-\beta P V_j\right] \right) \\ \times \int_j d\mathbf{s}_j^{n_j} \exp\left[-\beta U_j\left(\mathbf{s}_j^{n_j}\right) \right] \end{array} \right\} \end{array} \right\} \tag{4–32}$$

where the term K_0 is still arbitrary and its units are such as to make the integral of Equation (4–32) dimensionless. Since all terms exclusively involving n_i in Equation (4–32) are independent of all phase volumes V_j and dimensionless co-ordinates $\mathbf{s}^{n_j}$, all terms involving the phase volumes V_j are independent of the dimensionless co-ordinates $\mathbf{s}^{n_j}$, and all phase volumes and numbers of molecules of each species are independent of each other, Equation (4–32) may be rewritten as:

$$Q_{\mu PT} = \sum_{N_1, N_2, \ldots N_C}^{\infty} \left\{ \begin{array}{l} \int_{V_F} \cdots \int_{V_{II}} \int_{V_I} \prod_{j=1}^{F} dV_j \int_{s^{nF}} \cdots \int_{s^{nII}} \int_{s^{nI}} \\ \times \prod_{j=1}^{F} d\mathbf{s}_j^{n_j} \left\{ \dfrac{\prod_{j=1}^{F} V_j^{n_j}}{K_0 \prod_{i=1}^{C} \Lambda_i^{3n_i} \prod_{j=1}^{F} n_{ij}!} \exp\left[-\beta \Omega\right] \right\} \end{array} \right\} \tag{4–33}$$

where:

$$\Omega = \sum_{j=1}^{F} \left\{ U_j\left(\mathbf{s}^{N_j}\right) + P V_j - \sum_{i=1}^{C} \mu_i N_{ij} \right\} \tag{4–34}$$

The form of Equation (4–34) as a summation of energy terms arises due to the substitution of the product over exponentials in Equation (4–32) with equivalent summations over the exponentiated terms βPV_j, $\beta\mu_i N_i$ and $\beta U\left(\mathbf{s}^{N_j}\right)$. In particular, Equation (4–34) may be simplified to give:

$$\Omega = U_T\left(\mathbf{s}^{N_T}\right) + PV_T - \sum_{i=1}^{C}\mu_i N_{i,T} \tag{4–35}$$

where the summations over phases have been replaced by the total values for the extensive parameters. Comparison of Equation (4–35) with Equation (4–25) indicates that the dummy quantity Ω is in fact equivalent to TS where the partition function in Equation (4–33) is now a function of the intensive parameters only.

Expression of the generalized partition function over an arbitrary number of components and an arbitrary number of phases in Equation (4–33) provides a means to directly link back to the fundamental thermodynamic relation of Equation (4–21) or equivalently Equation (4–20).

Equation (4–33) provides an overall partition function from which it is possible to derive all other isothermal partition functions and from which the thermodynamic potential Ψ_{ens} may easily be discerned. This may be understood by writing Ω as the sum of two scalar products using Equation (4–26):

$$\begin{aligned}\Omega &= \mathbf{X}\bullet\mathbf{Y} \\ &= \bar{\mathbf{X}}\bullet\bar{\mathbf{Y}} + \tilde{\mathbf{X}}\bullet\tilde{\mathbf{Y}}\end{aligned} \tag{4–36}$$

where Ω is now the sum of the product of the conjugate pairs which have an extensive parameter which is fixed ($\bar{\mathbf{X}}\bullet\bar{\mathbf{Y}}$) and the conjugate pairs which have an extensive quantity which is allowed to vary ($\tilde{\mathbf{X}}\bullet\tilde{\mathbf{Y}}$). If one defines a particular ensemble with certain fixed extensive parameters in order to satisfy the equilibrium requirements of Equation (3–24), then the $\bar{\mathbf{X}}\bullet\bar{\mathbf{Y}}$ terms may be factored out of the integral in Equation (4–33) since the extensive quantities $\bar{\mathbf{X}}$ are constant:

$$\left[\mathcal{Q}_{\mu PT}\right]_{\bar{\mathbf{X}},\tilde{\mathbf{Y}}} = \exp\left[-\left(\bar{\mathbf{X}}\bullet\bar{\mathbf{Y}}\right)\right]\sum_{\mathbf{N}}^{\infty}\int_{\mathbf{X}\{1,\ldots,F\}} d\tilde{\mathbf{X}}\left\{\frac{f(\mathbf{N},\mathbf{V})}{\prod_{i=1}^{C}\Lambda_i^{3n_i}\times g(\mathbf{N})}\exp\left[-\left(\tilde{\mathbf{X}}\bullet\tilde{\mathbf{Y}}\right)\right]\right\} \tag{4–37}$$

where $f(\mathbf{N},\mathbf{V}) = \frac{1}{K_0}\prod_{j=1}^{F}V_j^{n_j}$, $g(\mathbf{N}) = \prod_{i=1}^{C}\prod_{j=1}^{F}n_{ij}!$, $\left[\mathcal{Q}_{\mu PT}\right]_{\bar{\mathbf{X}},\tilde{\mathbf{Y}}}$ denotes the special case of the μPT ensemble at constant $\bar{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$, $\mathbf{N} = \{n_I, n_{II},\ldots,n_F\}$ and $\mathbf{V} = \{V_I, V_{II},\ldots,V_F\}$. Note that the

integration is performed over all constrained and unconstrained extensive quantities which vary in all of the phases, i.e.:

$$\mathbf{X}\{1,...,\mathrm{F}\} = \left\{n_{1\mathrm{I}},...,n_{1\mathrm{F}},...,n_{\mathrm{C}1},...,n_{\mathrm{CF}},V_{\mathrm{I}},...,V_{\mathrm{F}},U_{\mathrm{I}}\left(\mathbf{s}^{n_1}\right),...,U_{\mathrm{F}}\left(\mathbf{s}^{n_{\mathrm{F}}}\right)\right\} \qquad (4\text{–}38)$$

It should be noted that integration over the set of $\{U_{\mathrm{I}},U_{\mathrm{II}},...,U_{\mathrm{F}}\}$ is in fact an integration over all of the molecular co-ordinates since $U = U\left(\mathbf{s}^N\right)$. For an ensemble at constant $\mathbf{N}$, the term $\prod_{i=1}^{C}\Lambda_i^{3n_i}\times g(\mathbf{N})$ factors out of the integral as a constant and so does not affect the formulation of the partition function or acceptance criteria, while for an ensemble at constant $\mathbf{N}$ and $\mathbf{V}$, the terms $\prod_{i=1}^{C}\Lambda_i^{3n_i}\times g(\mathbf{N})$ and $f(\mathbf{N},\mathbf{V})$ both factor out as constants. The integral term in Equation (4–37) is in fact the partition function for an ensemble at constant $\bar{\mathbf{X}}$ and $\bar{\mathbf{Y}}$. This shows that all partition functions and thermodynamic potentials may be derived as special cases of a generalized μPT ensemble:

$$\begin{aligned}\ln\left[\mathcal{Q}_{\mu PT}\right]\Big|_{\bar{\mathbf{X}},\bar{\mathbf{Y}}} &= -\left(\bar{\mathbf{X}}\bullet\bar{\mathbf{Y}}\right)+\ln\left[\mathcal{Q}\left(\bar{\mathbf{X}},\tilde{\mathbf{Y}}\right)\right]\\ &= -\left(\frac{S}{k_B}-\tilde{\mathbf{X}}\bullet\tilde{\mathbf{Y}}\right)+\ln\left[\mathcal{Q}\left(\bar{\mathbf{X}},\tilde{\mathbf{Y}}\right)\right]\end{aligned} \qquad (4\text{–}39)$$

where the final line of Equation (4–39) is derived from the relations given in Equation (4–26) and natural logarithms are taken of both sides of Equation (4–37) to begin to make a formal link with the thermodynamic potential defined in Equation (3–12). In order for Equation (4–39) to be useful, however, it is necessary to evaluate $\ln\left[\mathcal{Q}_{\mu PT}\right]\Big|_{\bar{\mathbf{X}},\bar{\mathbf{Y}}}$. For any given ensemble, the entropy may be related to probability by (Hill (1956)):

$$\langle S\rangle = -k_B\sum_{\hat{\mathbf{X}}}\mathcal{P}\left(\hat{\mathbf{X}}\right)\ln\left[\mathcal{P}\left(\hat{\mathbf{X}}\right)\right] \qquad (4\text{–}40)$$

where $\hat{\mathbf{X}}$ is the set of instantaneous values of the set of extensive variables $\tilde{\mathbf{X}}$ which are allowed to vary. It is more convenient to consider Equation (4–33) without the quasi-classical approximation such that:

$$\mathcal{Q}_{\mu PT} = \sum_{\hat{\mathbf{X}}}\exp\left[-T\hat{S}\right] \qquad (4\text{–}41)$$

where $\hat{S}$ is the instantaneous value of the entropy, the term $g(\mathbf{N})$ accounting for indistinguishability of the molecules is now implicitly accounted for in the summations over all

extensive variables and the need for the term $f(\mathbf{N},\mathbf{V})$ disappears since the partition function is now dimensionless. Using Equations (3–3) and (4–41), Equation (4–40) becomes:

$$\begin{aligned}\langle S\rangle &= -k_B \sum_{\bar{\mathbf{X}}} \wp\left(\hat{\mathbf{X}}\right)_{ens} \ln\left[\wp\left(\hat{\mathbf{X}}\right)_{ens}\right] \\ &= -k_B \sum_{S} \ln\left[\exp\left[-\beta TS\right] \Big/ Q_{\mu PT}\right] \exp\left(-\beta TS\right) \\ &= -k_B\left[-\left\langle S \Big/ k_B \right\rangle - \ln\left[Q_{\mu PT}\right]\right]\end{aligned} \qquad (4\text{–}42)$$

from which $-\ln\left[Q_{\mu PT}\right]=0$. This expression is analogous to the expression derived by Hill (1956). From Equation (3–12), this implies that the thermodynamic potential for the μPT ensemble, $\Psi_{\mu PT}=-\ln\left[Q_{\mu PT}\right]$, is zero This is consistent with the observation that the μPT does not describe a system of physical significance because no extensive parameter is fixed and therefore, since this ensemble should have no link to macroscopic thermodynamics, the μPT thermodynamic potential must be zero.

Equation (4–39) may thus be summarized as:

$$\ln\left[Q_{\mu PT}\right]\Big]_{\bar{\mathbf{X}},\bar{\mathbf{Y}}} = \Psi_{\bar{\mathbf{X}},\bar{\mathbf{Y}}} + \ln\left[Q\left(\bar{\mathbf{X}},\bar{\mathbf{Y}}\right)\right] = 0 \qquad (4\text{–}43)$$

Comparing Equations (4–39) and (4–43), it may be seen that:

$$\Psi_{\bar{\mathbf{X}},\bar{\mathbf{Y}}} = -\left(\bar{\mathbf{X}}\bullet\bar{\mathbf{Y}}\right) = -\left(\frac{S}{k_B} - \tilde{\mathbf{X}}\bullet\tilde{\mathbf{Y}}\right) \qquad (4\text{–}44)$$

It would be reassuring to note that the condensed formulation given in Equation (4–43) yields the correct partition function and thermodynamic potential for all ensembles. For this purpose, the application of Equation (4–43) to derive the partition function and the thermodynamic potential for the N_1N_2PT version of the Gibbs Ensemble is demonstrated in Appendix 4A at the end of this Chapter.

4.8 Appendix 4A: Application of μPT Generalized Ensemble to the Gibbs Ensemble

In this Appendix, the formalism of Equations (4–37) and (4–39) is used to alternatively 'derive' the partition function and the thermodynamic potential for the N_1N_2PT version of the Gibbs Ensemble to demonstrate the utility of the new formalism and to demonstrate that this new

formalism correctly generates known partition functions. The derivation of the thermodynamic potentials and partition functions for the NPT, NVT or μVT ensembles using this approach are similar.

4.8.1 The N_1N_2PT Gibbs Ensemble as a Special Case of the μPT Generalized Ensemble

The N_1N_2PT version of the Gibbs Ensemble is a two-phase and two-component ensemble at constant number of molecules of species 1, constant number of molecules of species 2, constant pressure and constant temperature. The sets $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$, $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ are then:

$$\begin{aligned}
\bar{\mathbf{X}} &= \left\{ n_{\mathrm{I},1} + n_{\mathrm{II},1}, n_{\mathrm{I},2} + n_{\mathrm{II},2} \right\} = \left\{ N_1, N_2 \right\} \\
\bar{\mathbf{Y}} &= \left\{ -\beta\mu_1, -\beta\mu_2 \right\} \\
\tilde{\mathbf{X}} &= \left\{ V_{\mathrm{I}}, V_{\mathrm{II}}, U_{\mathrm{I}}, U_{\mathrm{II}} \right\} \\
\tilde{\mathbf{Y}} &= \left\{ \beta_{\mathrm{I}} P_{\mathrm{I}}, \beta_{\mathrm{II}} P_{\mathrm{II}} = \beta_{\mathrm{I}} P_{\mathrm{I}}, \beta_{\mathrm{I}}, \beta_{\mathrm{II}} = \beta_{\mathrm{I}} \right\}
\end{aligned} \tag{4–45}$$

According to Equations (4–39) and (4–43), the thermodynamic potential for this ensemble is then:

$$\begin{aligned}
\Psi^{GE}_{N_1N_2PT} &= \beta\mu_1 N_1 + \beta\mu_2 N_2 \\
&= -S/k_B + \beta\left(U_{\mathrm{I}} + U_{\mathrm{II}}\right) + \beta P\left(V_{\mathrm{I}} + V_{\mathrm{II}}\right) \\
&= \beta G
\end{aligned} \tag{4–46}$$

where G is the Gibbs free energy. This thermodynamic potential is equivalent to the thermodynamic potential of the NPT ensemble (Equation (3–14)).

For the N_1N_2PT version of the Gibbs Ensemble, $C = 2$ and $F = 2$. Substituting for $\bar{\mathbf{X}} \bullet \tilde{\mathbf{Y}}$ from Equation (4–45), Equation (4–33) therefore becomes:

$$\left[\mathcal{Q}_{\mu PT}\right]^{GE}_{N_1N_2PT} = \sum_{N_1,N_2} \left\{ \begin{aligned}
&\int_{V_{\mathrm{II}}} \int_{V_{\mathrm{I}}} dV_{\mathrm{II}} dV_{\mathrm{I}} \int_{s^{n_{\mathrm{II}}}_{\mathrm{II}}} \int_{s^{n_1}_1} ds^{n_{\mathrm{II}}}_{\mathrm{II}} ds^{n_1}_1 \\
&\times \frac{V_{\mathrm{I}}^{n_1} V_{\mathrm{II}}^{n_{\mathrm{II}}}}{K_0 \Lambda_1^{3N_1} \Lambda_2^{3N_2} n_{\mathrm{I},1}! n_{\mathrm{II},1}! n_{\mathrm{I},2}! n_{\mathrm{II},2}!} \\
&\times \left\{ \exp\left[-\beta\left(U_{\mathrm{I}}\left(s^{n_1}_1\right) + U_{\mathrm{II}}\left(s^{n_{\mathrm{II}}}_{\mathrm{II}}\right) + PV_{\mathrm{I}} + PV_{\mathrm{II}} \right) \right] \right\}
\end{aligned} \right\} \tag{4–47}$$

Also, one has the relations $n_{\mathrm{I},1} + n_{\mathrm{II},1} = N_1$ and $n_{\mathrm{I},2} + n_{\mathrm{II},2} = N_2$ from which one can eliminate $n_{\mathrm{II},1}$ and $n_{\mathrm{II},2}$. Equation (4–47) becomes:

$$\left[Q_{\mu PT}\right]^{GE}_{N_1N_2PT} = \frac{1}{K_0\Lambda_1^{3N_1}\Lambda_2^{3N_2}} \sum_{N_{I,1},N_{I,2}} \frac{1}{n_{I,1}!n_{II,1}!n_{I,2}!n_{II,2}!} \int_{V_{II}}\int_{V_I} dV_{II}dV_I V_{II}^{n_{II}} V_I^{n_I}$$
$$\times \int_{s_{II}^{n_{II}}}\int_{s^{n_I}} ds_{II}^{n_{II}} ds_I^{n_I} \exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_I}\right)+U_{II}\left(\mathbf{s}_{II}^{n_{II}}\right)+PV_I+PV_{II}\right)\right]$$
$$= \frac{1}{K_0\Lambda_1^{3N_1}\Lambda_2^{3N_2}} \sum_{n_{I,1}=0}^{N_1}\sum_{n_{I,2}=0}^{N_2} \frac{1}{n_{I,1}!\left(N_1-n_{I,1}\right)!n_{I,2}!\left(N_2-n_{I,2}\right)!} \qquad (4\text{–}48)$$
$$\times\left\{\begin{array}{l} \int_{V_{II}=0}^{\infty} dV_{II}\exp\left[-\beta PV_{II}\right]V_{II}^{n_{II}} \int_{V_I=0}^{\infty} dV_I \exp\left[-\beta PV_I\right]V_I^{n_I} \times \\ \int_{s_{II}^{N-n_I}}\int_{s_I^{n_I}} ds_{II}^{N-n_I} ds_I^{n_I}\left\{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_I}\right)+U_{II}\left(\mathbf{s}_{II}^{N-n_I}\right)\right)\right]\right\} \end{array}\right\}$$

where both volumes are now changed independently and which yields the probability density of Equation (4–9) as required.

4.9 Appendix 4B: Generating RGEMC Results from Conventional GEMC Simulations

The RGEMC is able to achieve a shift in the simulated phase diagram by modifying the expression for the chemical potential in the vapour phase as follows (Lísal *et al.* (1999)):

$$\left(\mu_{g,i}\right)^{RG} = \mu_{g,i}\left(T,P^0\right)+\left(\mu_{g,i}^{ex}\right)^{GE}+k_BT\ln\left[\frac{P_{i,exp}^{sat}}{\left(P_i^{sat}\right)^{GE}}\right] \qquad (4\text{–}49)$$

where superscript *RG* denotes a property value for the reaction Gibbs Ensemble, superscript *GE* denotes a property value for the Gibbs Ensemble, superscript *ex* denotes an excess property, subscript exp denotes an experimentally measured quantity, $\mu_{g,i}$ denotes the chemical potential of species i in the vapour phase, $\mu_{g,i}\left(T,P^0\right)$ denotes the standard chemical potential in the ideal gas state, and P_i^{sat} is the pure component saturated vapour pressure of species i. Implicit in the derivation of Equation (4–49) is the assumption that:

$$\frac{P_{i,exp}^{sat}}{\left(P_i^{sat}\right)^{GE}} = \frac{f_{i,exp}^{sat}}{\left(f_i^{sat}\right)^{GE}} \qquad (4\text{–}50)$$

which in general is strictly only true for ideal mixtures. The term $P_{i,exp}^{sat}/\left(P_i^{sat}\right)^{GE}$ in Equation (4–49) may be viewed as an 'adjustment' to the reference chemical potential and is accounted for in the acceptance rule for a molecule transfer move in the RGEMC as follows:

$$acc(o \to n) = \min\left\{1, \left(\frac{P_{i,exp}^{sat}}{\left(P_i^{sat}\right)^{GE}} \frac{n_l V_g}{(N - n_l + 1)V_l} \exp\left[\frac{-(\Delta U_l + \Delta U_g)}{k_{BT}}\right]\right)\right\} \quad (4\text{–}51)$$

The term $P_{i,exp}^{sat} / \left(P_i^{sat}\right)^{GE}$ in Equation (4–51) is inverted in the acceptance criterion for the reverse transfer of a molecule from the liquid to the vapour phase. The only difference between Equation (4–51) and Equation (4–7) is the presence of the 'ideal' term $P_{i,exp}^{sat} / \left(P_i^{sat}\right)^{GE}$.

An analogous expression to Equation (4–49) may also be written for the liquid phase to ensure equality of chemical potentials between the two coexisting phases where the subscripts g of Equation (4–49) are replaced by subscripts l. Unlike in the vapour phase, the adjustment to the reference chemical potential is not accounted for in the acceptance rule of Equation (4–51). To understand how this term is accounted for and how the phase diagram is shifted using RGEMC, it is convenient to consider the gamma/phi formulation for VLE (Smith *et al.* (1996)):

$$y_i \hat{\phi}_i P = x_i \gamma_i f_i \quad (4\text{–}52)$$

or equivalently

$$y_i \Phi_i P = x_i \gamma_i P_i^{sat} \quad ; \quad \Phi_i = \hat{\phi}_i P_i^{sat} / f_i \quad (4\text{–}53)$$

where y_i is the mole fraction of component i in the vapour phase, x_i is the mole fraction of component i in the liquid phase, $\hat{\phi}_i$ is the fugacity coefficient of component i in solution in the vapour phase, γ_i is the activity coefficient of component i in the liquid, f_i is the fugacity of component i in solution, and Φ_i measures deviations from ideal behaviour in the vapour. The term on the right-hand side of Equation (4–52) is directly related to the chemical potential of the liquid phase. The link between the chemical potentials of a GEMC simulation and a RGEMC simulation may thus be written as:

$$\left(x_i \gamma_i f_i\right)^{GE} = \exp\left[\left(\mu_{l,i}\right)^{GE}\right] = \frac{\left(P_i^{sat}\right)^{GE}}{P_{i,exp}^{sat}} \exp\left[\left(\mu_{l,i}\right)^{RG}\right] \quad (4\text{–}54)$$

Therefore, Equation (4–53) may be rewritten as follows:

$$\begin{aligned} \left(y_i \Phi_i P\right)^{RG} &= \left(x_i \gamma_i P_i^{sat}\right)^{GE} \frac{P_{i,exp}^{sat}}{\left(P_i^{sat}\right)^{GE}} \\ &= \left(x_i \gamma_i\right)^{GE} P_{i,exp}^{sat} \end{aligned} \quad (4\text{–}55)$$

Since the scaling term $P_{i,exp}^{sat}/\left(P_i^{sat}\right)^{GE}$ represents an additional contribution to the reference chemical potential, both Φ_i and γ_i (which account for non-ideal vapour and liquid deviations respectively) are not functions of this new contribution. Thus, in general, for a given x_i a necessary condition for both the RGEMC and GEMC to yield similar compositions is that:

$$\frac{P^{RG}}{P_{i,exp}^{sat}} = \left(\frac{P}{P_i^{sat}}\right)^{GE} \tag{4–56}$$

where the above relation has been derived by a ratio of Equation (4–55) written for both reaction Gibbs Ensemble and the Gibbs Ensemble and where P may be evaluated for a binary mixture from:

$$P = \frac{x_1\gamma_1 P_1^{sat}}{\Phi_1} + \frac{(1-x_1)\gamma_2 P_2^{sat}}{\Phi_2} \tag{4–57}$$

by using the relation that the mole fractions in the vapour must sum to one. In general, Equation (4–56) is not satisfied by the RGMEC methodology unless both of the simulated vapour pressures of the pure components exhibit the same relative deviation from the experimental saturated vapour pressures of the pure components. For all other cases, a given x_i will not yield similar y_i values for both GEMC and RGEMC.

In addition, it is possible to generate the results from RGEMC using standard GEMC by use of the following formulae which may be derived by using Equation (4–55) and Equation (4–57) and by substituting the experimental saturated vapour pressures:

$$y_i = x_i\left(\frac{\gamma_i P_{i,exp}^{sat}}{\Phi_i}\right)^{GE}\frac{1}{P^{RG}} \tag{4–58}$$

$$P^{RG} = \sum_i x_i\left(\frac{x_i\gamma_i}{\Phi_i}\right)^{GE} P_{i,exp}^{sat} \tag{4–59}$$

where

$$\left(\frac{\gamma_i}{\Phi_i}\right)^{GE} = \left(\frac{y_i P}{x_i P_i^{sat}}\right)^{GE} \tag{4–60}$$

and may be evaluated from the GEMC results at the end of a simulation. Hence, the RGEMC approach does not address fundamental inaccuracies in any force field and thus serves to only qualitatively and not quantitatively shift the phase diagram. Any apparent improvement in the overall location of the phase diagram is achieved by compromising the true molar composition.

In order to demonstrate this procedure, the original P-x-y and x-y data for GEMC and RGEMC simulations (Lísal *et al.* (1999)) for the isobutene + MTBE system at 350K, as well as for the method described above are plotted in Figure 4-5 and Figure 4-6. Although there are only 3 mixture data points, it may be seen that the x-y data calculated from Equations (4–58) to (4–60) using the results from the Gibbs Ensemble simulations lie on the plot for the x-y data from the reaction Gibbs Ensemble simulations and that the vapour mole fractions are clearly shifted from the original Gibbs Ensemble results. This demonstrates the equivalence of the methodology proposed by Equations (4–58) to (4–60) and that in spite of the shift that occurs due to the RGEMC approach, this shift is still reproduced in Figure 4-6 by the methodology described above.

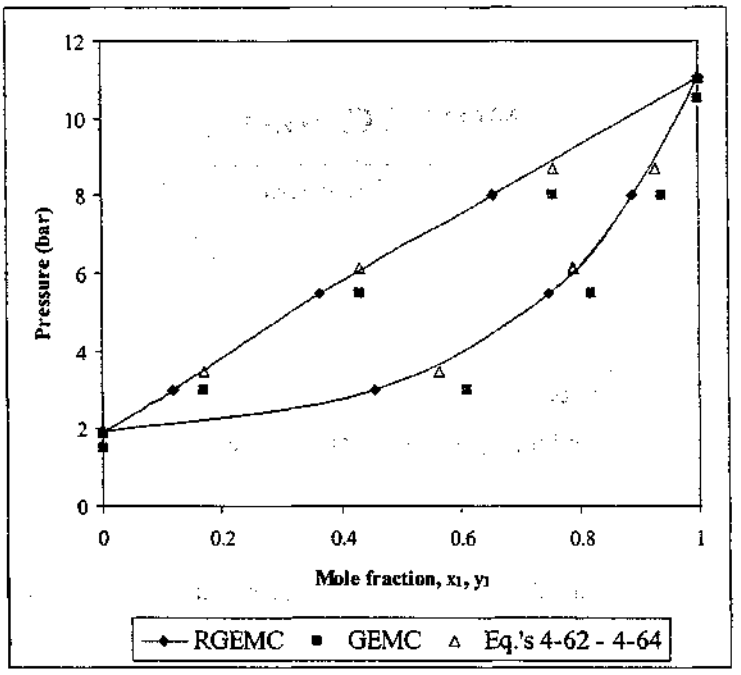


Figure 4-5: Simulated P-x-y diagrams for isobutene (1) + MTBE (2) at 350K from the original RGEMC and GEMC simulations, and using Equations (4–58) to (4–60).

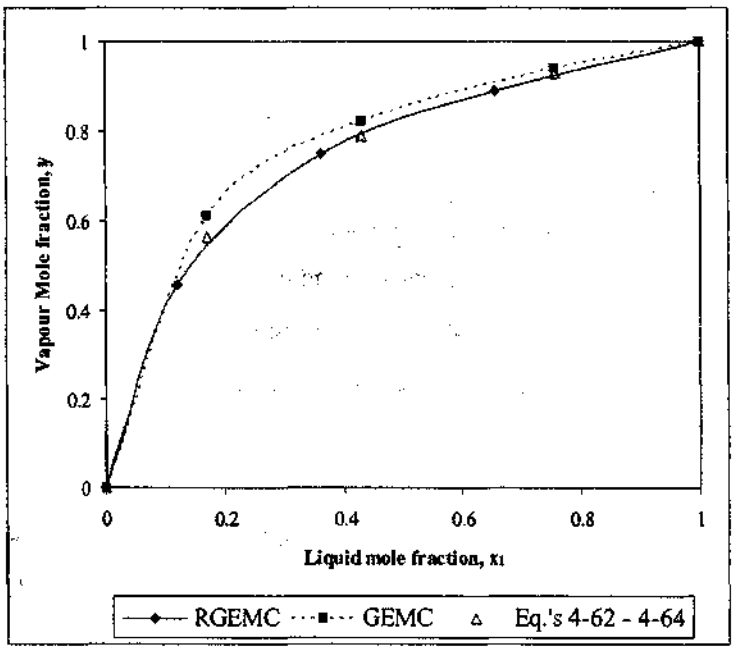


Figure 4-6: Simulated P-x-y diagrams for isobutene (1) + MTBE (2) at 350K from the original RGEMC and GEMC simulations, and using Equations (4–58) to (4–60).

CHAPTER 5:
MONTE CARLO SIMULATION METHODS

"Give me six hours to chop down a tree and I will spend the first four sharpening the axe" – Abraham Lincoln

5.1 Introduction

Recalling the Muybridge bet of Section (3.1), there were several practical difficulties to be confronted by Muybridge including the requirement for faster shutter speeds to overcome the limitations of 19th century wet photography that required long exposure times as well as the need to film dark horses against white backgrounds to diminish the effect of blurring. In an analogous manner, several practical considerations require special techniques for Monte Carlo simulations, for example when simulating charged molecules or for simulating molecules with an articulated structure. These methods include periodic boundaries and the minimum image convention, the calculation of long-range interactions and Configurational-Bias Monte Carlo (CBMC).

5.2 Periodic Boundaries and the Minimum Image Convention

Molecular simulations are usually performed using a small number of molecules, $10 \le N \le 10000$ (Allen and Tildesley (1987)). The size of the system under investigation is limited by the available computer memory and by the speed of execution of the program. Indeed, if all molecular interactions are assumed to be pairwise additive, then the time taken to complete the double loop required to evaluate all distinct interactions between pairs of interaction sites is proportional to N^2. Consequently, computational times scale rapidly with increasing size of the system under study and with model interaction complexity.

Unfortunately, in a three-dimensional and N–particle system with free boundaries, the number of molecules at the surface of a simulation cell is proportional to $N^{-1/3}$. Even for moderately large system sizes, this implies that a significant fraction of the molecules under study will be at the surface. Molecules on the surface will in general experience quite different forces to molecules in the bulk. It is thus necessary to choose boundaries which mimic the behaviour of an infinite bulk fluid. This is usually achieved by the use of periodic boundary conditions. The volume V

containing N molecules is treated as the primitive cell of an infinite lattice of identical cells. A two-dimensional version of a periodic system is shown in Figure 5-1:

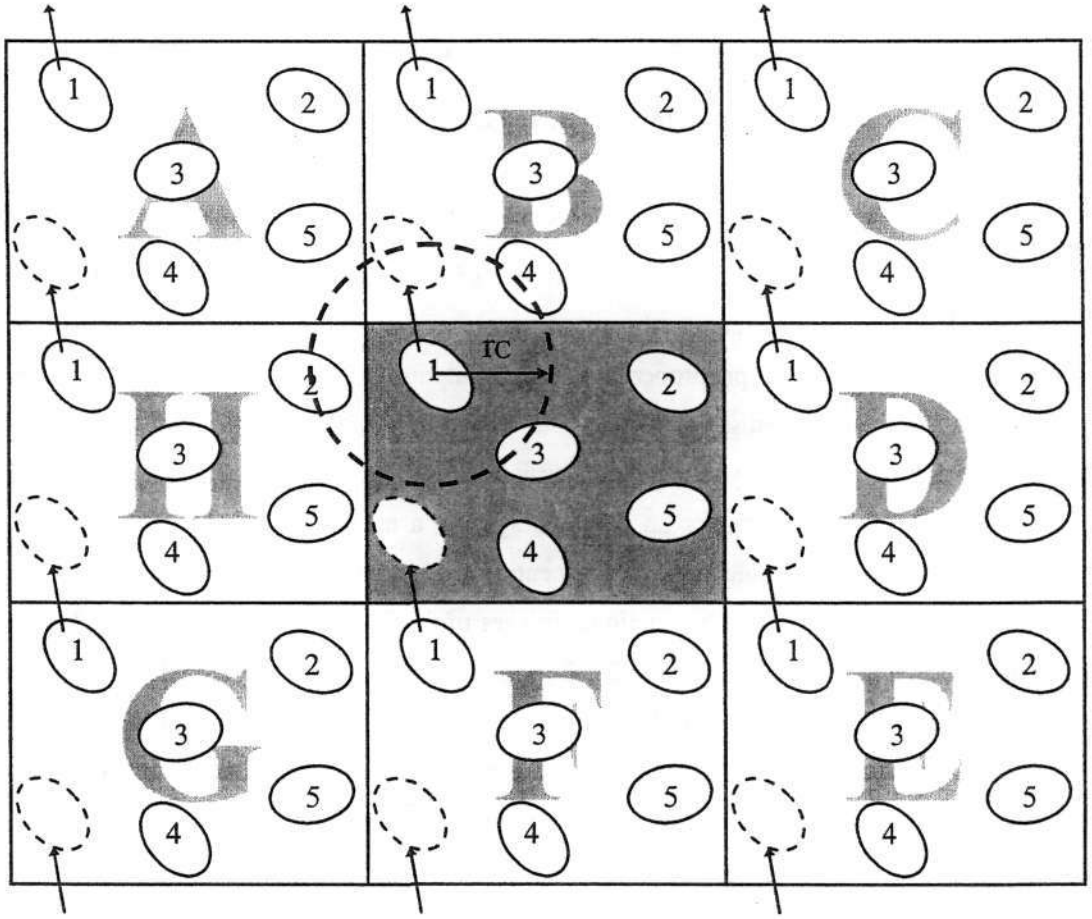


Figure 5-1: Two-dimensional periodic system (Allen and Tildesley (1987). The boxes labelled "A"-"H" are periodic images of the central grey box. Molecules may enter or leave each box across any of the four edges. The cut-off radius, r_C, for intermolecular interactions is shown using the dashed circle around molecule 1.

As a molecule 1 moves through a boundary, its images 1_A, 1_B, 1_C, etc. move across their corresponding boundaries. In this way, the number density of the central box is conserved and only the co-ordinates of molecules in the central box need to be recorded. If all the intermolecular interactions are taken to be pairwise additive, then the total potential energy of the N molecules for a three dimensional system may be given by (Frenkel and Smit (1996)):

$$U = \frac{1}{2}\sum_{\mathbf{i},\mathbf{j},\mathbf{b}}{}^{\dagger} U_{\mathbf{ij}}\left(\left|\mathbf{r}_{\mathbf{ij}} + \mathbf{b}L_b\right|\right) \qquad (5\text{–}1)$$

where the summation is carried over all pairs of interaction sites i and j, $\mathbf{r}_{ij}$ is a set of three integers describing the vector separation between sites interaction sites i and j , $\mathbf{b}$ is a set of three integers describing the vector co-ordinates of a periodic image, $U_{ij}\left(\left|\mathbf{r}_{ij}+\mathbf{b}L_b\right|\right)$ is the intermolecular potential energy function between two molecules, and the symbol † denotes that the term with i=j is excluded for $\mathbf{b}=0$.

The presence of the periodic image boxes has the effect of potentially introducing spurious correlations which are not present in the macroscopic system under study (Frenkel and Smit (1996)). The effect on macroscopic properties is thus a function of both the intermolecular potential and the properties under investigation (Allen and Tildesley (1987)). In general, therefore, the maximum wavelength of a fluctuation is then limited to be less than the simulation box length. Another factor to consider is that for an isotropic system, a molecule should not interact with a periodic image of itself (i.e. molecule 1 in the central grey box in Figure 5-1 should not interact with molecule 1_A, 1_B, etc.) or with two periodic images of the same molecule (i.e. molecule 1 may interact with 2_H in Figure 5-1 but not also 2_E, 2_F, 2_G, etc.) because this will introduce an artificial periodicity on the simulation. This is known as the minimum image convention and effectively implies that the interacting molecule should not interact with molecules outside of a box of the same size as the simulation box centred on the interacting molecule of interest.

5.3 The Lennard-Jones Potential

The well-known Lennard-Jones (LJ) potential for describing van der Waals interactions is the most important intermolecular potential energy model used in this work and is given by (Lennard-Jones (1924):

$$U_{LJ}\left(r_{ij}\right)=4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12}-\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \qquad (5\text{–}2)$$

where $U_{LJ}\left(r_{ij}\right)$ is the LJ interaction potential, ε_{ij} is the intermolecular potential well depth or energy parameter for the interaction between sites i and j, σ_{ij} is the intermolecular potential size parameter between sites i and j at which $U_{LJ}\left(r_{ij}\right)$ is zero, and r_{ij} is the separation distance between interaction sites i and j. Equation (5–2) may be thought of as consisting of a repulsive term proportional to r^{-12} which dominates at short separation distances and a dispersive or an

attractive term which is proportional to r^{-6} and which dominates at large separations. The physical relevance of ε_{ij} and σ_{ij} are given in Figure 5-2:

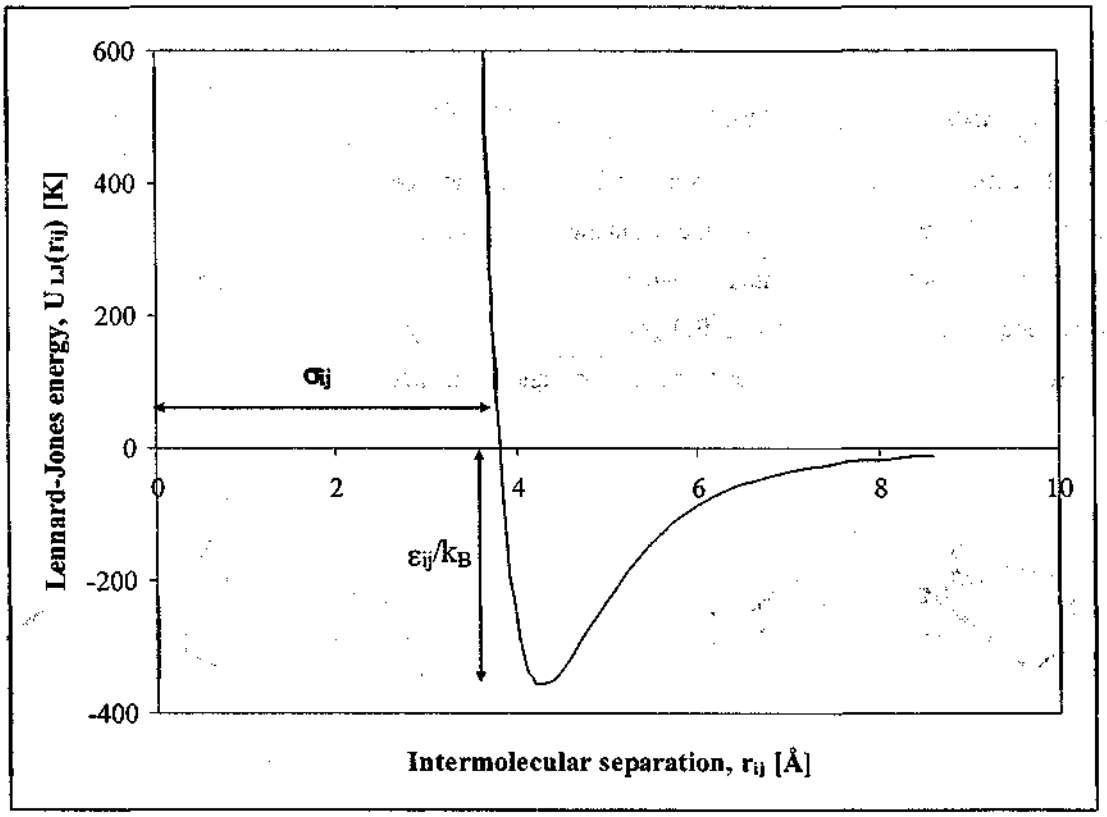


Figure 5-2: Graph showing the shapes of the intermolecular potential function for the LJ with $\varepsilon_{ij}/k_B = 359\text{K}$ and $\sigma_{ij} = 3.808\text{Å}$.

The section of the curve for which $dU/dr_{ij} > 0$ (i.e. the part of the curve to the left of the local minimum of U) represents the distance for which repulsive interactions dominate. The section of the intermolecular potential energy curve for which $dU/dr_{ij} < 0$ (i.e. the part of the curve to the right of the local minimum of U) contains the separation distances for which dispersion interactions dominate.

As pointed out by Wu and Sadus (2000), the LJ potential enjoys popularity because it is a continuous intermolecular potential that qualitatively captures the salient features of van der Waals intermolecular interactions. There are only two parameters (σ_{ij} and ε_{ij}) which need to be regressed

for a given interaction between sites i and j. The denominator powers of r_{ij} are also even which ensures that one can avoid the expensive square root function in any simulation code which would be required to calculate r_{ij} from molecular co-ordinates.

5.4 Intramolecular Potential Energy Models

In general, the application of spherically symmetric point potentials will only be valid for molecules of small size because the influence of molecular geometry on the intermolecular interactions are not well accounted for when considering larger molecules of varying shapes. Since the seminal work of Ryckaert and Bellemans (1975) and Ryckaert and Bellemans (1978), it has become common to view *n*-alkanes from the "ball and stick" perspective as seen in Figure 5-3:

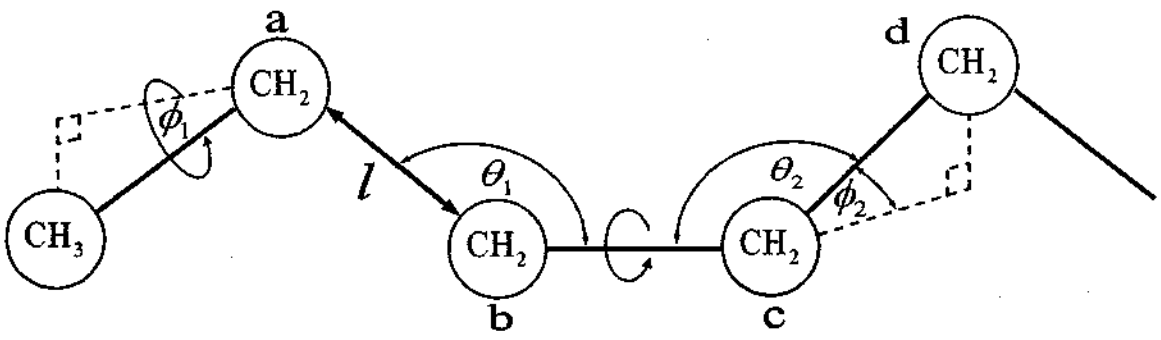


Figure 5-3: Diagram showing the representation of chain molecules. CH_2 groups labelled a , b and c are all in the same plane of the page. The angles θ_1 and θ_2 are the bond-bending angles between three adjacent functional groups. The angles ϕ_1 and ϕ_2 are the torsional angles between four adjacent functional groups and are defined as 0° for the *cis*-conformation. The torsion around the bond b-c is the angle ϕ_2 between the planes defined by the CH_2 groups a-b-c and b-c-d . l is the bond length between two adjacent functional groups.

There will be a conformational energy associated with a given molecular geometry. The intramolecular potential models which can be used to describe this interaction can become quite complicated and Sadus (1999) contains a discussion of some of the more complicated models. Unfortunately, the more complicated the intramolecular interactions, the longer the simulation will take and this effect increases with increasing chain length. For many years now since the work of

Ryckaert and Bellemans (1975), it has been common practice when simulating *n*-alkanes to only consider those intramolecular interactions arising from:

- bond stretching between two adjacent functional groups
- bond angle bending between three adjacent functional groups
- torsions between four functional groups
- dispersion interactions for functional groups separated by at least 4 bonds

The contribution from bond stretching for those models where the bond lengths are not constant may be described by a harmonic potential of the form:

$$U_{stretch} = {}^{k_s}\!/_{2}\,(l - l_0)^2 \tag{5–3}$$

where $U_{stretch}$ is the bond stretching potential energy, k_s is the bond stretching constant, l is the bond length of the bond and l_0 is the equilibrium bond length of the bond. Equation (5–3) indicates that positive or negative deviations from the equilibrium bond length lead to an increase in the intramolecular energy.

The contribution from bond angle bending for those models where the bond angles fluctuate may be described by a harmonic potential of the form (van der Ploeg and Berendsen (1982)):

$$U_{bend} = {}^{k_\theta}\!/_{2}\,(\theta - \theta_0)^2 \tag{5–4}$$

where U_{bend} is the bond angle bending energy, k_θ is the bond angle bending constant, $0 \le \theta \le \pi$ is the bond angle and θ_0 is the equilibrium bond angle. Equation (5–4) indicates that positive or negative deviations from the equilibrium bond angle also lead to an increase in the intramolecular energy.

Nearly all models for *n*-alkanes include a torsional potential to account for the out-of-plane rotations that occur about a central bond in a set of four adjacent functional groups or atoms. The work of Ryckaert and Bellemans (1975) and subsequently the work of Jorgensen *et al.* (1984) established the utility of using a cosine series to describe this contribution:

$$\begin{aligned} U_{tors} &= c_0 + c_1\left[1+\cos(\phi)\right] + c_2\left[1-\cos(2\phi)\right] + c_3\left[1+\cos(3\phi)\right] - \ldots \\ &= c_0 + \sum_{n=1,\ldots} c_n\left[1+(-1)^{n-1}\cos(n\phi)\right] \end{aligned} \tag{5–5}$$

where U_{tors} is the bond angle bending energy, $c_0, c_1, c_2, \ldots, c_n$ are constants, $0 \le \phi \le 2\pi$ is the torsional angle defined in Figure 5-3. Equation (5–5) captures the periodic character that this

potential should have because as the torsional angle rotates around the central bond (bond b-c in Figure 5-3) the intramolecular energy due this rotation must demonstrate local energy maxima or minima due to steric hindrance between the functional groups. The torsional angle is defined as 0.0° in the *cis* conformation for all torsions.

For interactions between functional groups separated by at least four bonds that do not involve charged species, the interactions are usually modelled using a dispersion interaction potential such as the LJ potential (Equation (5–2)).

Although these interactions are highly simplified, they provide an efficient compromise between computational expense and adequately describing intramolecular interactions.

5.5 Long-Range Interactions

In general, the molecular interactions are dominated by the interactions of each molecule with its nearest neighbours. It thus becomes practical to only consider those short-ranged interactions that are less than some cut-off distance, r_C. For the purposes of this work, this cut-off is always spherical. This cut-off also turns out to be computationally necessary since the time required increases proportional to $N_{\{int\}}{}^2$, where $N_{\{int\}}$ is the number of interaction sites, and the number of interaction sites contained in a sphere around a central interaction site increases proportional to $N_{\{int\}} = \rho V_C \propto V_C \propto r_C^3$ where V_C is cut-off volume arising due to the cut-off radius. The number of intermolecular interactions thus scales roughly as r_C^6.

With reference to Figure 5-2, the contribution from long-range interactions with $r_{ij} > 9.5\text{Å}$ may appear negligible but this is, in general, not the case particularly in the vicinity of the critical point. Powles (1984) has indicated that the inclusion of these long-range interactions must be undertaken with some care. Smit and Frenkel (1991) have shown that the phase diagram of the two-dimensional Lennard-Jones fluid depends largely on the details of the truncation while Smit (1992) has shown that the phase diagram of the three dimensional Lennard-Jones fluid is significantly underestimated when the long-ranged interactions are not correctly included.

If the total potential energy is not rigorously zero for $r_{ij} \geq r_C$ and if the interactions are pairwise additive, then it is possible to write the contribution of these long-range interactions to the total potential energy as (Frenkel and Smit (1996)):

$$U^T = \sum_{i<j} U^C\left(r_{ij}\right) + \frac{N\rho}{2}\int_{r_C}^{\infty} dr_{ij} U\left(r_{ij}\right) 4\pi r_{ij}^2 \qquad (5\text{–}6)$$

where U^T is the total potential energy, $U^C\left(r_{ij}\right)$ is the potential energy from two-body interactions evaluated up to the cut-off radius and the integral extends from the cut-off radius, r_C, to infinity. From Equation (5–6), it may be seen that unless $U\left(r_{ij}\right)$ decays faster than r_{ij}^{-3} (where the 3 arises from the 3-dimensional nature of the system), then the contribution from the long-range dispersion forces will be infinite since then $dr_{ij} r_{ij}^2 U\left(r_{ij}\right) > 1$. It thus becomes essential to judiciously select a suitable cut-off radius, r_C, and to have some means of accurately approximating the non-negligible contributions from long-range interactions.

5.5.1 Analytical Tail Corrections

The pair or "radial" distribution function, $g_2\left(r_{ij}\right)$, provides a convenient means to accurately estimate the contribution to the potential energy from the long-ranged interactions when $dr_{ij} r_{ij}^2 U\left(r_{ij}\right) < 1$ (Allen and Tildesley (1987)), i.e. when the dispersion interactions decay more rapidly than the dimensionality of the system. $g_2\left(r_{ij}\right)$ describes the probability of finding two atoms a distance r_{ij} apart relative to the probability expected for a completely random distribution at the same density. McQuarrie (1976) contains a more detailed description but in a generalized form, $g_2\left(r_{ij}\right)$ may be given by:

$$g_2\left(r_{ij}\right) = \frac{N(N-1)}{\rho^2 Q_{ens}} \int dr_3 dr_4 ... dr_N \exp\left[-\left(\tilde{\mathbf{X}}\left(r_1, r_2 ..., r_N\right) \cdot \tilde{\mathbf{Y}}\right)\right] \qquad (5\text{–}7)$$

where Q_{ens} is the partition function for the ensemble, the integration is performed over $N-2$ molecular co-ordinates and $\tilde{\mathbf{Y}}$ is the set of intensive parameters which have a conjugate extensive property $\tilde{\mathbf{X}}\left(r_1, r_2 ..., r_N\right)$ which is variable and, in general, is a function of $\left(r_1, r_2 ..., r_N\right)$. One may then write the total internal energy of a pure component as (Allen and Tildesley (1987)):

$$U^T = (3/2)Nk_BT + 2\pi N\rho\int_0^\infty dr_{ij}r_{ij}^2 g_2(r_{ij})U(r_{ij}) \qquad (5\text{–}8)$$

and the pressure of a pure component as (Allen and Tildesley (1987)):

$$PV = Nk_BT - (2/3)\pi N\rho\int_0^\infty dr_{ij}g_2(r_{ij})r_{ij}^2\frac{1}{3}\left(r_{ij}\frac{dU(r_{ij})}{dr_{ij}}\right) \qquad (5\text{–}9)$$

and the chemical potential of a pure component as (Allen and Tildesley (1987)):

$$\mu = k_BT\ln\left[\rho\Lambda^3\right] + 4\pi\rho\int_0^1 d\xi\int_0^\infty dr_{ij}r_{ij}^2 g_2(r_{ij},\xi)U(r_{ij}) \qquad (5\text{–}10)$$

where ξ is a coupling parameter that must be integrated over as well (McQuarrie (1976)).

So long as r_C is chosen sufficiently large, it is possible to assume that $g_2(r_{ij}) \approx 1$, i.e. the ratio of the molecule density around a central molecule in a sphere of radius r_{ij} to the average density is approximately unity for $r_{ij} \geq r_C$ such that Equations (5–8) to (5–10) for a pure component interacting via the LJ potential (Equation (5–2)) become (Allen and Tildesley (1987)):

$$\begin{aligned} U_{LJ}^T &\approx U_{LJ}^C + U_{LJ}^{LR} = U_{LJ}^C + 2\pi N\rho\int_{r_C}^\infty dr_{ij}r_{ij}^2 U_{LJ}(r_{ij}) \\ \therefore U_{LJ}^{LR} &= \frac{8}{3}\pi N\rho\varepsilon\left[\frac{1}{9}\left(\frac{\sigma^{12}}{r_C^9}\right) - \left(\frac{\sigma^6}{r_C^3}\right)\right] \end{aligned} \qquad (5\text{–}11)$$

$$\begin{aligned} (PV)_{LJ}^T &\approx (PV)_{LJ}^C + (PV)_{LJ}^{LR} \\ &= (PV)_{LJ}^C - (2/3)\pi N\rho\int_{r_C}^\infty dr_{ij}r_{ij}^2\frac{1}{3}\left(r_{ij}\frac{dU_{LJ}(r_{ij})}{dr_{ij}}\right) \\ \therefore P_{LJ}^{LR} &= \frac{16}{3}\pi\rho^2\varepsilon\left[\frac{2}{3}\left(\frac{\sigma^{12}}{r_C^9}\right) - \left(\frac{\sigma^6}{r_C^3}\right)\right] \end{aligned} \qquad (5\text{–}12)$$

$$\begin{aligned} \mu_{LJ}^T &\approx \mu_{LJ}^C + \mu_{LJ}^{LR} = \mu_{LJ}^C + 4\pi\rho\int_{r_C}^\infty dr_{ij}r_{ij}^2 U_{LJ}(r_{ij}) \\ \therefore \mu_{LJ}^{LR} &= \frac{16}{3}\pi\rho\varepsilon\left[\frac{2}{3}\left(\frac{\sigma^{12}}{r_C^9}\right) - \left(\frac{\sigma^6}{r_C^3}\right)\right] \end{aligned} \qquad (5\text{–}13)$$

where the subscript LR denotes a long-range contribution.

5.6 Pressure Calculation

The pressure is an important quantity which must be calculated during simulation unless it is set to be constant like in an isobaric ensemble (e.g. the N_1N_2PT Gibbs Ensemble). For the general case of molecular fluids, Allen and Tildesley (1987) have shown using the virial theorem that the pressure may be written as the sum of the ideal and the excess contributions as:

$$\begin{aligned}\langle P_{vir}\rangle &= \langle P^{id}\rangle + \langle P^{C,ex}\rangle + \langle P^{LR}\rangle \\ &= \langle \rho k_B T\rangle + \langle \mathcal{W}/V\rangle\end{aligned} \qquad (5\text{–}14)$$

where P_{vir}, P^{id}, $P^{C,ex}$ and P^{LR} are the total virial pressure, the ideal pressure, the excess contribution to the pressure up to the cut-off radius and the long-range contribution to the pressure respectively; and $\mathcal{W}$ is the total molecular viral. For simplicity, the long-range contributions are included in the term $\mathcal{W}$ in Equation (5–14) but are easily separated out and $\langle \mathcal{W}/V\rangle$ may be evaluated from Equation (5–12). In addition for atomic fluids with pairwise additive intermolecular potentials, $\mathcal{W}$ may be written as (Allen and Tildesley (1987)):

$$\mathcal{W} = \tfrac{1}{3}\sum_i\sum_{j>i}\mathbf{r}_{ij}\cdot\mathbf{F}_{ij} = -\tfrac{1}{3}\sum_i\sum_{j>i}\mathbf{r}_{ij}\cdot\nabla_r U(\mathbf{r}_{ij}) = -\tfrac{1}{3}\sum_i\sum_{j>i} w(r_{ij}) \qquad (5\text{–}15)$$

where $\mathbf{r}_{ij}$ is the vector separating interaction sites i and j such that $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ and $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$; $\mathbf{F}_{ij}$ is the pair force exerted on molecule i by molecule j; ∇_r is the gradient vector defined by $\nabla_r(\chi) = \frac{d\chi}{dr_x}\hat{\mathbf{r}}_x + \frac{d\chi}{dr_y}\hat{\mathbf{r}}_y + \frac{d\chi}{dr_z}\hat{\mathbf{r}}_z$ where subscripts x, y and z denote 3 orthogonal directions; $\hat{\mathbf{r}}_x$, $\hat{\mathbf{r}}_y$ and $\hat{\mathbf{r}}_z$ are three orthogonal unit vectors; and $w(r_{ij})$ is the total intermolecular pair viral function such that:

$$w(r_{ij}) = r_{ij}\frac{dU(r_{ij})}{dr_{ij}} \qquad (5\text{–}16)$$

where $w(r_{ij})$ is a purely scalar quantity and only scalar quantities appear in Equation (5–16) since the dot product $\nabla_r U(\mathbf{r}_{ij})$ in Equation (5–15) must yield a scalar quantity. In Equation (5–15) it must also be understood that the function $\nabla_r U(\mathbf{r}_{ij})$ is, in general, orientation dependent for molecular fluids. Two options are then available when considering the definition of $w(r_{ij})$. Defining $w(r_{ij})$ as an atomic virial would define the distances r_{ij} as atomic distances and this would require the evaluation of intramolecular separations and forces between sites on the same

molecule as well as the expected intermolecular forces and distances between sites on different molecules. The other (considerably simpler) alternative is to define $w(r_{ij})$ as a molecular virial quantity where all the distances r_{ij} are explicitly intermolecular distances between interaction sites on different molecules. This is advantageous because it avoids the extra calculation of intramolecular forces. Akkermans and Ciccotti (2004) have demonstrated explicitly the equivalence of the pressure calculated by the atomic virial and the pressure calculated by the molecular virial. For the molecular virial, the vector distance $\mathbf{r}_{ij}$ must also now be interpreted as the vector separating the two co-ordinate centres of mass of the two molecules. The term $\mathbf{r}_{ij}$ must thus be replaced in Equation (5–15) with the projection of $\mathbf{r}_{ij}$ in the direction $\mathbf{r}_{iajb}$, $\text{proj}_{\mathbf{r}_{iajb}}\mathbf{r}_{ij}$ where $\mathbf{r}_{iajb} = \mathbf{r}_{jb} - \mathbf{r}_{ia}$ is the vector separating interaction site b on molecule j and interaction site a on molecule i. The term $\mathbf{F}_{ij}$ must also be replaced by $\sum_a \sum_b \mathbf{F}_{iajb}$ since the total force between two molecules must now be calculated as the sum between all sites on different molecules. The molecular virial becomes:

$$\begin{aligned} \mathcal{W} &= \tfrac{1}{3}\sum_i \sum_{j>i} \sum_a \sum_b \left(\text{proj}_{\mathbf{r}_{iajb}} \mathbf{r}_{ij}\right) \bullet \mathbf{F}_{iajb} \\ &= -\tfrac{1}{3}\sum_i \sum_{j>i} \sum_a \sum_b \frac{\left(\mathbf{r}_{iajb} \bullet \mathbf{r}_{ij}\right)}{\left|\mathbf{r}_{iajb} \bullet \mathbf{r}_{iajb}\right|} \mathbf{r}_{iajb} \bullet \nabla_r U\left(\mathbf{r}_{iajb}\right) \\ &= -\tfrac{1}{3}\sum_i \sum_{j>i} \sum_a \sum_b \frac{\left(\mathbf{r}_{iajb} \bullet \mathbf{r}_{ij}\right)}{r_{iajb}^2} w\left(r_{iajb}\right) \end{aligned} \tag{5–17}$$

where the molecular virial is now given by

$$w\left(r_{ij}\right) = \sum_a \sum_b \frac{\left(\mathbf{r}_{iajb} \bullet \mathbf{r}_{ij}\right)}{r_{iajb}^2} w^C\left(r_{iajb}\right) \tag{5–18}$$

and

$$w\left(r_{iajb}\right) = r_{iajb} \frac{dU\left(r_{iajb}\right)}{dr_{iajb}} \tag{5–19}$$

It must be noted that the density term, ρ, in Equation (5–14) is now the density of the molecules to coincide with the molecular virial and not the total density of all of the interaction sites on the molecules which coincides with the atomic virial.

From the preceding equations it becomes clear that the pressure is dependent on the slope of the potential energy curve, i.e. dependent on $dU(r)/dr$.

For the LJ potential, the expression for the molecular virial is easily evaluated as:

$$w_{LJ}\left(r_{iajb}\right) = r_{iajb}\frac{dU_{LJ}\left(r_{iajb}\right)}{dr_{iajb}} = -24\varepsilon_{ab}\left[2\left(\frac{\sigma_{ab}}{r_{iajb}}\right)^{12} - \left(\frac{\sigma_{ab}}{r_{iajb}}\right)^{6}\right] \qquad (5\text{–}20)$$

Another important calculation for pure component simulations is the latent heat of vaporization, ΔH_{vap}. The latent heat of vaporization may be calculated directly from the thermodynamic definition of the enthalpy for the liquid and vapour phases as:

$$\begin{aligned}\left\langle \Delta H_{vap}\right\rangle &= \left\langle H_g\right\rangle - \left\langle H_l\right\rangle = \left\langle U_g - \left(PV/N\right)_g\right\rangle - \left\langle U_l - \left(PV/N\right)_l\right\rangle \\ &= \left\langle U_g - U_l\right\rangle - \left\langle \left(PV/N\right)_g - \left(PV/N\right)_l\right\rangle\end{aligned} \qquad (5\text{–}21)$$

where subscript g denotes the vapour phase, subscript l denotes the liquid phase and where ΔH_{vap} and U are intensive properties. Strictly speaking, $\left\langle \Delta H_{vap}\right\rangle$ must be calculated as the ensemble average as indicated in Equation (5–21). This work has found that it is possible to apply the approximation that $\left\langle \left(PV/N\right)_g - \left(PV/N\right)_l\right\rangle \approx \left\langle P_g\left(\left(V/N\right)_g - \left(V/N\right)_l\right)\right\rangle$ without introducing significant error but the decrease in computational load is essentially negligible. Another possible approach is to suggest that one should only use ensemble averaged quantities at the end of the simulation as:

$$\left\langle \Delta H_{vap}\right\rangle = \left\langle U_g - U_l\right\rangle - \left\langle P_g\right\rangle\left\langle \left(V/N\right)_g\right\rangle + \left\langle P_l\right\rangle\left\langle \left(V/N\right)_l\right\rangle \qquad (5\text{–}22)$$

but this will in general not be valid since in general $\left\langle PV/N\right\rangle_g \neq \left\langle P_g\right\rangle\left\langle \left(V/N\right)_g\right\rangle$. Use of the Clausius-Clapeyron equation:

$$\Delta H_{vap} = R_G\frac{d\ln P}{d\left(1/T\right)} \qquad (5\text{–}23)$$

where R_G denotes the Universal Gas Constant, by making a plot of $\ln P$ against $1/T$ to give $\Delta H_{vap}/R_G$ will also only yield an approximate answer since it assumes ideality of vapour phase and it neglects the molar volume of the liquid as being negligible compare to that of the vapour which is not, in general, true. This is particularly the case for temperature near the critical point.

5.7 Configurational-Bias Monte Carlo Methods

A common bottleneck in attaining convergence by Monte Carlo methods which rely on molecule transfers is the low rates of acceptance for molecule transfer attempts. This is particularly true for dense fluids at low temperatures, for large molecules or for molecules with an articulated structure because any random move which tries to insert an entire molecule in a single step will struggle to find a sufficient gap in which to insert the large molecule. Smit *et* al. (1995) have found for a chain of n LJ atoms that the probability that an insertion will not result in an overlap is of the order of 0.005^n. Consequently, the probability of inserting a molecule becomes dramatically smaller for increasing chain lengths. The Configurational-Bias Monte Carlo (CBMC) method tackles this problem by biasing the selection of orientations such that configurations favourable to the insertion of a molecule are generated.

Based on the work of Siepmann and Frenkel (1992) which essentially described a CBMC scheme for chain molecules with a finite number of discrete confirmations and which was used to simulate polyethylene (de Pablo *et al.* (1992a)), Frenkel *et al.* (1992), Mooij *et al.* (1992) and Smit *et al.* (1995) and then separately de Pablo *et al.* (1992b) and Laso *et al.* (1992) developed the CBMC method for the more general off-lattice case. Frenkel *et al.* (1992) originally suggested that one should separate the calculation of the intermolecular interactions from the strong intramolecular interactions (like those in Equations (5–3) to (5–5) in the CBMC method for optimal efficiency. The method of de Pablo *et al.* (1992b) and Laso *et al.* (1992), however, does not separate these contributions. Smit *et al.* (1995) have shown explicitly as to why the method of Frenkel *et al.* (1992) is more efficient and for that reason the basic calculation of as suggested by Frenkel and Smit (1996) is outlined below. The CBMC algorithm is shown schematically in Figure 5-4.

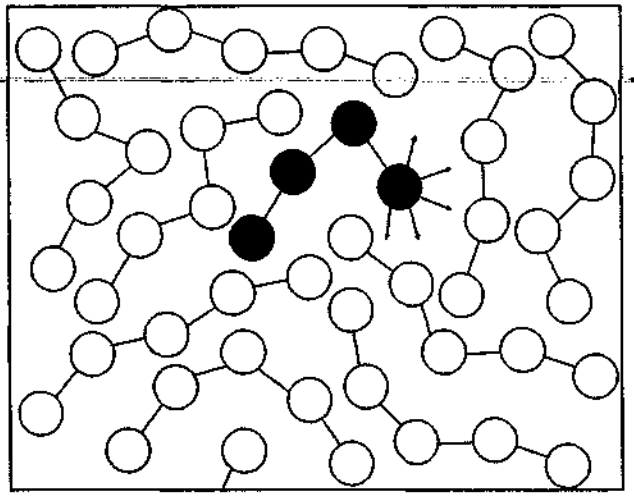


Figure 5-4: Schematic representation of the insertion of a chain molecule using CBMC. The arrows indicate the trial orientations in which to insert the fifth atom.

A chain is grown systematically segment by segment. For a chain with ℓ segments, a trial new conformation is generated as follows:

1) The first atom is placed at a random position and the intermolecular energy U_1^{ext} is calculated together with:

$$w_1^{new} = \exp\left[-\beta U_1^{ext}\right] \tag{5–24}$$

where w_1^{new} is the Rosenbluth factor of the first segment and U_1^{ext} is the intermolecular interaction energy of the first segment. For low density systems the placement of the first bead is usually not a problem but in many other cases the success rate can be quite low. For this reason, Esselink *et al.* (1995) have outlined a procedure in terms of which multiple sites are considered for the first segment. In that case, if $\{\mathbf{c}\}_{k_0} = \mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_{k_0}$ is a set of k_0 trial positions for the first segment then w_1^{new} becomes:

$$w_1^{new} = \sum_{h=1}^{k_0} \exp\left[-\beta U_1^{ext}(\mathbf{c}_h)\right] \tag{5–25}$$

2) From the k_0 trial first positions, one is selected where each first segment has probability:

$$p^{ext}(\mathbf{c}_h) = \frac{\exp\left[-\beta U_1^{ext}(\mathbf{c}_h)\right]}{w_1^{new}} \tag{5–26}$$

For the insertion of the next segment m, a set $\{\mathbf{c}\}_k = \mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_k$ of k trial orientations are again generated, but this time relative to the previous bead and according to a probability which is a function of the internal energy:

$$\mathcal{P}^{\text{int}}(\mathbf{c}_h) = \frac{\exp\left[-\beta U_m^{\text{int}}(\mathbf{c}_h)\right]}{\text{K}} \tag{5–27}$$

Mooij and Frenkel (1996) show how to optimize k for a particular application. For each of the k trial orientations generated according to Equation (5–27), the external energy $U(\mathbf{c}_h)_m^{\text{ext}}$ is calculated together with the Rosenbluth factor for the segment:

$$\text{w}_m^{\text{new}} = \sum_{h=1}^{k} \exp\left[-\beta U_m^{\text{ext}}(\mathbf{c}_h)\right] \tag{5–28}$$

From these k trial positions, one is selected with probability:

$$\mathcal{P}^{\text{ext}}(\mathbf{c}_h) = \frac{\exp\left[-\beta U_m^{\text{ext}}(\mathbf{c}_h)\right]}{\text{w}_m^{\text{new}}} \tag{5–29}$$

3) Step 2 is repeated $\ell - 1$ times until the entire molecule chain is grown and the total Rosenbluth weight, W^{new}, of the new chain may then be calculated. The Rosenbluth weight of the new chain is defined as:

$$\text{W}^{\text{new}} = \frac{\text{w}_1^{\text{new}} \prod_{m=2}^{\ell} \sum_{h=1}^{k} \exp\left[-\beta U_m^{\text{ext}}(\mathbf{c}_h)\right]}{k_0 \times k^{\ell-1}} \tag{5–30}$$

4) To complete a CBMC move, it is also necessary to calculate the Rosenbluth weight of the old configuration. The old chain is retraced in a similar way except that for each segment only $k-1$ trial orientations ($k_0 - 1$ trial orientations for the first segment) are considered. The k^{th} trial orientation ($k_0{}^{\text{th}}$ trial orientation for the first bead) is necessarily the old chain. The total Rosenbluth weight of the old chain, W^{old}, is defined as:

$$\text{W}^{\text{old}} = \frac{\text{w}_1^{\text{old}} \prod_{m=2}^{\ell} \sum_{h=1}^{k} \exp\left[-\beta U_m^{\text{ext}}(\mathbf{c}_h)\right]}{k_0 \times k^{\ell-1}} \tag{5–31}$$

5) To satisfy detailed balance, it has been shown (Frenkel *et al.* (1992)) that the new configuration is accepted with probability:

$$acc(o \to n) = \min\left\{1, \frac{\text{W}^{\text{new}}}{\text{W}^{\text{old}}}\right\} \tag{5–32}$$

The above acceptance criterion may be applied to the partial or full "regrowing" of a chain molecule. Note that the constant K which comes from the intramolecular interactions cancels off in the final derivation and does not appear in Equation (5–32). The method can be understood to be 'looking one segment ahead' before placing the next segment of the molecule. The Rosenbluth weights may be viewed as correcting for the bias introduced by preferentially considering the more favourable configurations. In this way the insertions of molecules are biased such that favourable configurations are generated. Smit *et al.* (1995) and Mooij *et al.* (1992) have shown that the acceptance criterion for the transfer of a molecule from phase II to phase I where the chain is grown by means of the CBMC algorithm is given by:

$$acc(o \to n) = \min\left\{1, \frac{V_I(N - n_I)}{(V - V_I)(n_I + 1)} \frac{W^{new}}{W^{old}}\right\} \qquad (5\text{–}33)$$

The difference between Equation (5–33) above and Equation (4–7) is the substitution of the internal energy difference term, $\exp\left[-\beta \Delta U_I^{new}\right]$, by the Rosenbluth weight, W^{new}. Also, the difference between the de Pablo *et al.* (1992b) and the Frenkel *et al.* (1992) methods is found in step 2 of the method for generating a new configuration. The de Pablo method does not separate the intramolecular interactions and generates new configurations according to Equation (5–27). For this reason, the intramolecular interactions will appear in the new Rosenbluth weight and in the acceptance criterion. This implies that the probability of accepting a new configuration will in general be less efficient for the de Pablo method than for the Frenkel method and increasingly less efficient for longer chains (Smit *et al.* (1995)).

Based on the CBMC method, Siepmann *et al.* (1993) successfully simulated the pure component phase behaviour up to the critical point of long-chain *n*-alkanes ranging from C_6 to C_{48} for the first time.

It is common practice to immediately reject those moves which would place two interactions sites at a distance less than some specified hard inner cut-off radius from one another. A hard inner cut-off value of 2Å is typical. The reasoning behind this may be deduced from Figure 5-2 where it may be seen that the intermolecular energy becomes rapidly positive for short separations which corresponds to intermolecular repulsions. The large positive intermolecular energy changes necessary to achieve these moves renders their probability of occurrence so negligibly low that these moves may be neglected. The idea that the intermolecular repulsions at short distances play a dominant role in determining the acceptance of attempted moves has been used by Vlugt *et al.*

(1998) to devise the Dual Cut-Off Configurational-Bias Monte Carlo (DC-CBMC) algorithm. The algorithm essentially divides the external energy of interaction of a molecule, $U^{ext}(r<r_C)$, into a shorter-range potential which is less expensive to calculate, $\bar{U}^{ext}(r<r_{DC})$, and the difference between the full and approximate potentials, $\delta U^{ext}(r_{DC}<r<r_C)$, as follows:

$$U^{ext}(r<r_C)=\bar{U}^{ext}(r<r_{DC})+\delta U^{ext}(r_{DC}<r<r_C) \qquad (5\text{–}34)$$

where r_{DC} is the shorter DC-CBMC cut-off radius. The acceptance criterion for accepting a molecule regrowth has been shown by Vlugt *et* al. (1998) to be given by:

$$acc(o\rightarrow n)=\min\left\{1,\frac{\bar{W}^{new}}{\bar{W}^{old}}\exp\left[-\beta\left(\delta U^{ext,new}-\delta U^{ext,old}\right)\right]\right\} \qquad (5\text{–}35)$$

where $\bar{W}$ is an approximate Rosenbluth weight calculated for $r<r_{DC}$. An optimal value for r_{DC} has been shown to be approximately 5Å. This algorithm has been found to speed up the simulation of *n*-octane by a factor of 4 and more for even longer molecules (Vlugt *et al.* (1998)).

The basic CBMC algorithm has been modified to include expanded ensembles (Escobedo and de Pablo (1995)) in which long-chain molecules are inserted gradually to improve insertion successes during a simulation. The recoil growth algorithm (Consta *et al.* (1999a), Consta *et al.* (1999b)) is another development in terms of which a chain molecule may recoil along its length while being grown in order to avoid configurations which hit dead-ends before the final configuration is accepted. Both of these algorithms greatly improve the efficiency of insertions for molecules with complicated articulated structures and for dense systems. Finally, Martin and Siepmann (1999a) have developed the coupled-decoupled CBMC algorithm for the simulation of branched molecules.

5.8 Chemical Potential Calculation

The evaluation of the chemical potential is an important calculation because the equality of the chemical potentials of a species in all phases for that species is a necessary requirement for total thermodynamic equilibrium. Smit and Frenkel (1989) have shown that the chemical potential of component i in phase I in an atomic system in the NVT Gibbs Ensemble for a simulation in which the probability that either of the boxes empties of molecules is small, and where the boxes do not change identities, is given by:

$$\mu_{I,i}=-k_BT\ln\left[\frac{1}{\Lambda^3}\left\langle\frac{V_I}{n_{I,i}+1}\exp\left(-\frac{\Delta U^{+}_{I,i}}{k_BT}\right)\right\rangle\right] \qquad (5\text{–}36)$$

where $\Delta U_{I,i}^{+}$ is the energy change associated with inserting a molecule of type i into phase I. Equation (5–36) has its origins in the test particle energy expression of Widom (1963). During a Gibbs Ensemble simulation, however, the "ghost-" or "test particle" insertions required by the Widom method are unnecessary. For the NVT or N_1N_2PT Gibbs Ensembles, the insertion energies, $\Delta U_{I,i}^{+}$, required by the ensemble average of Equation (5–36) are obtained from the actual energy changes associated with molecule swaps between the coexisting phases. Equation (5–36) is, however, only applicable for atomic systems. For articulated molecules simulated using the CBMC method, Mooij and Frenkel (1994) have shown that the chemical potential of a species i in phase I is given by:

$$\begin{aligned}\mu_{I,i} &= -k_B T \ln\left[\left(\frac{Q_i^{id}}{\Lambda^3}\right)\left\langle\left(\frac{V_I}{n_{I,i}+1}\right)W_{I,i}\right\rangle\right] \\ &= -k_B T \ln\left[\left(\frac{Q_i^{id}}{\Lambda^3}\right)\right] - k_B T \ln\left[\left\langle\left(\frac{V_I}{n_{I,i}+1}\right)W_{I,i}\right\rangle\right] \\ &= \mu_{I,i}^{id} + \mu_{I,i}^{ex}\end{aligned} \qquad (5\text{–}37)$$

where $W_{I,i}$ is the Rosenbluth factor of Equation (5–30) for the inserted molecule of type i in phase I and Q_i^{id} is the partition function of an ideal chain constructed using only intramolecular interactions and is given by:

$$Q_i^{id} = \frac{\prod_{m=2}^{\ell}\sum_{h=1}^{k}\exp\left[-\beta U_m^{int}(\mathbf{c}_h)\right]}{k^{\ell-1}} \qquad (5\text{–}38)$$

CHAPTER 6:

SIMULATING BINARY PHASE EQUIBRIUM USING TRANSFERABLE UNITED-ATOM FORCE FIELDS

"Somewhere, something incredible is waiting to be known." – Blaise Pascal

6.1 Transferable Intermolecular Force Fields

In recent years, a myriad of transferable force fields have emerged based on the approach of Section (5.3) for the intermolecular potential, Section (5.4) for the intramolecular potential and Section (5.5) for the long-range corrections. The majority of these have been directed towards the simulation of normal alkanes, branched alkanes and 1-alkenes. Prominent force fields for these hydrocarbons include:

- the TraPPE force field (Martin and Siepmann (1998), Martin and Siepmann (1999a), Chen and Siepmann (1999), Wick *et al.* (2000), Chen *et al.* (2001)),
- the NERD force field (Nath *et al.* (1998a), Nath and de Pablo(2000), Nath *et al.* (2001a)),
- the Spyriouni *et al.* olefin force field (Spyriouni *et al.* (1999)), and
- the *n*-alkane force field of Errington and Panagiotopoulos (1999a) (nAEP force field).

These force fields have made it possible to simulate a wide range of compounds, as well as mixtures that belong to certain homologous series of chemicals, without having to parameterise force fields for individual chemical compounds.

The above force fields all make the "(isotropic) united-atom" (UA) approximation in which functional groups of atoms are collapsed into a single interaction site. In this way, the $-CH_3$, $-CH_2-$ and $-CH-$ functional groups in the hydrocarbon chain are treated as single sites. Unlike the "explicit hydrogen" (EH) approach, these UA models assume that the explicit inclusion of hydrogen atom interactions is unnecessary and that hydrogen atom interactions can be lumped into LJ or Be6 parameters for functional groups of atoms. As pointed out by Chen and Siepmann (1999), the united-atom approximation reduces CPU calculation time by approximately one order of magnitude without significantly lowering the accuracy of the simulated phase equilibrium properties. The most significant deviations are found when simulating high pressure phase equilibrium or high-density liquids where the contributions of the hydrogen atom to molecular volume become very significant (Ryckaert *et al.* (1989), Moller *et al.* (1991)).

Mixtures of *n*-alkanes have been studied using the TraPPE (Martin and Siepmann (1997), Martin *et al.* (1999b), Martin *et al.* (1999c), Martin *et al.* (2000)) and NERD force fields (Nath *et al.* (1998b)) as well as the nAEP force field (Potoff *et al.* (1999)).

More recently, revised anisotropic united atom (AUA) force fields have also been parameterised for the simulation of *n*-alkanes (Ungerer *et al.* (2000)), branched alkanes (Bourasseau *et al.* (2002)) and 1-alkenes, 2-alkenes, isobutene and 1,3-butadiene (Bourasseau *et al.* (2003)). AUA force fields differ from UA force fields by allowing the interaction site on a functional group to be displaced off-centre relative to the actual location of the functional group in three-dimensional space. These force fields have also been applied to the simulation of *n*-alkanes and 1-alkene mixtures in zeolites (Pascual *et al.* (2003), Pascual *et al.* (2004)).

1-Alkenes (α-olefins) are an important class of compounds that find wide application in industrial processes (Grubbs and Chang (1998), Rouhi (2002)), and knowledge of their phase behaviour is vital for equipment design. In spite of this, there have been only a limited number of simulation studies of 1-alkene mixtures (Nath and de Pablo (1999), Nath *et al.* (2001a), Nath *et al.* (2001b), Lísal *et al.* (1999)).

6.2 Simulation of Binary Phase Envelopes

The NERD and TraPPE force fields as well as the 1-alkene force field of Spyriouni *et al.* (SA1) were used because these force fields have been shown to reproduce experimental pure component phase equilibrium data reasonably well in the original publications. Also, these force fields were developed using different approaches for describing the intramolecular geometries and different thermophysical properties were targeted when optimizing the intermolecular potential energy parameters. The intermolecular interactions have been shown previously to play a dominant role in determining the pure component phase equilibrium properties (Smit *et al.* (1995), Dubbeldam *et al.* (2004)). The SA1 force field has fixed bond lengths and bond angles but flexible torsional angles, the TraPPE force field has constant bond lengths but flexible bond angles and torsional angles, and the NERD force field has flexible bond lengths, bond angles and torsional angles. In addition, the parameterisation of the Lennard-Jones size σ and energy ε parameters for the SA1 force field differed to that of the NERD and TraPPE force fields. Specifically, when fitting the LJ parameters for the SA1 force field, emphasis was placed on the experimental vapour pressures, while fitting of the NERD and TraPPE force fields focused on reproducing the

experimental saturated liquid densities, critical temperatures and critical densities. It is therefore important to examine whether these differences will significantly affect the simulated phase diagrams of the mixtures studied here.

The recently developed AUA force fields mentioned previously have also been shown to reproduce the pure component phase equilibrium data well. These AUA force fields were, however, parameterised with less attention given to the equilibrium properties of the shorter *n*-alkanes, in particular ethane and propane. In addition, it remains to be seen whether a similar improvement in the equilibrium data obtained from current united-atom force fields could not be achieved by modifying the internal equilibrium geometries. This would amount to allowing the internal geometries to be different to the experimentally observed geometries to account for the effects of anisotropic interactions without having to resort to a full anisotropic potential. This is advantageous because AUA force fields require additional calculations to compute the interaction site, which is not located on the united-atom centre, as well as a correction which must be added to the Rosenbluth factor (Smit *et al.* (1995)). The results for 1-alkenes in Pascual *et al.* (2004) have also shown that the TraPPE-UA force field gives slightly better liquid density results while the AUA4 force field gives superior vapour pressure predictions. For these reasons and since the UA force fields are computationally cheaper, the NERD, TraPPE and Spyriouni *et al.* force fields only were used in this work since this would establish how these recently developed force fields compare in predicting binary VLE for *n*-alkane + 1-alkene mixtures and permits a comparison of how the different parameterisation targets (vapour pressures or liquid densities and critical properties) affect the P-x-y and x-y diagrams for the mixtures.

6.3 Simulation Details

CBMC Gibbs Ensemble simulations were undertaken for three *n*-alkane + 1-alkene mixtures:

- P-x-y data for ethane (1) + propene (2) at 277.6K
- P-x-y data for 1-hexene (1) + *n*-octane (2) at 328.15K
- T-x-y data for *n*-dodecane (1) + 1-octadecene (2) at 760mmHg

and two 1-alkene + 1-alkene mixtures:

- P-x-y data for propene (1) + 1-butene (2) at 294.3K
- P-x-y data for 1-butene (1) + 1-hexene (2) at 373K

The mixtures investigated were chosen since they are systems for which experimental binary VLE data are available (Dortmund Data Bank), and are representative of the types of mixtures which can

be encountered in industrial separation operations such as petroleum refining (McKay *et al.* (1951)), sub-atmospheric fractional distillation of petroleum (Jordan and van Winkle (1951)), and in the synthetic chemical industry (Laugier and Richon (1996), Goff *et al.* (1951)).

6.3.1 Force Field Parameters

All three force fields use the LJ 12-6 potential (Equation (5–2)) to calculate non-bonded intermolecular energies. The LJ parameters for interactions between like sites denoted as σ_{ii} and ε_{ii} are listed in Table 6-1 for all three force fields:

Table 6-1: Intermolecular Force field parameters for the NERD, TraPPE and SA1 force fields.

NERD			
Ethane	Propene		
$\varepsilon_{CH_3(sp^3)}/k_B = 100.6\text{K}$	$\varepsilon_{CH_3(sp^3)}/k_B = 100.0\text{K}$	$\varepsilon_{CH_2(sp^2)}/k_B = 92.5\text{K}$	$\varepsilon_{CH(sp^2)}/k_B = 46.0\text{K}$
$\sigma_{CH_3(sp^3)} = 3.825\text{Å}$	$\sigma_{CH_3(sp^3)} = 3.85\text{Å}$	$\sigma_{CH_2(sp^2)} = 3.72\text{Å}$	$\sigma_{CH(sp^2)} = 3.77\text{Å}$
Molecules with more than three carbon atoms			
$\varepsilon_{CH_3(sp^3)}/k_B = 104.0\text{K}$	$\varepsilon_{CH_2(sp^3)}/k_B = 45.8\text{K}$	$\varepsilon_{CH_2(sp^2)}/k_B = 92.5\text{K}$	$\varepsilon_{CH(sp^2)}/k_B = 46.0\text{K}$
$\sigma_{CH_3(sp^3)} = 3.91\text{Å}$	$\sigma_{CH_2(sp^3)} = 3.93\text{Å}$	$\sigma_{CH_2(sp^2)} = 3.72\text{Å}$	$\sigma_{CH(sp^2)} = 3.77\text{Å}$
TraPPE			
$\varepsilon_{CH_3(sp^3)}/k_B = 98.0\text{K}$	$\varepsilon_{CH_2(sp^3)}/k_B = 46.0\text{K}$	$\varepsilon_{CH_2(sp^2)}/k_B = 85.0\text{K}$	$\varepsilon_{CH(sp^2)}/k_B = 47.0\text{K}$
$\sigma_{CH_3(sp^3)} = 3.75\text{Å}$	$\sigma_{CH_2(sp^3)} = 3.95\text{Å}$	$\sigma_{CH_2(sp^2)} = 3.675\text{Å}$	$\sigma_{CH(sp^2)} = 3.73\text{Å}$
SA1			
$\varepsilon_{CH_3(sp^3)}/k_B = \varepsilon_{CH_2(sp^3)}/k_B = 47.66\text{K}$	$\varepsilon_{CH(sp^2)}/k_B = 81.69\text{K}$	$\varepsilon_{CH_2(sp^2)}/k_B = 89.93\text{K}$	
$\sigma_{CH_3(sp^3)} = \sigma_{CH_2(sp^3)} = \sigma_{CH(sp^2)} = 3.915\text{Å}$	$\sigma_{CH_2(sp^2)} = 3.905\text{Å}$		

Parameters for interactions between two sites of different type (e.g., between $-CH_3$ and $-CH_2-$) are calculated using the Lorentz-Berthelot mixing rules:

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad ; \quad \sigma_{ij} = \left(\sigma_{ii} + \sigma_{jj}\right)/2 \qquad (6\text{–}1)$$

The NERD and TraPPE force fields use the LJ 12-6 potential to calculate the potential energy of non-bonded interactions up to a specified cut-off radius of 13.8 Å and 14 Å respectively. For the SA1 force field, however, the LJ potential is used to calculate intermolecular interactions up to a first cut-off radius of $1.45\sigma_{ij}$, whereas a quintic spline is used between $1.45\sigma_{ij}$ and $2.30\sigma_{ij}$ (Spyriouni *et al.* (1999), Theodorou and Suter (1985)). For all force fields, analytical tail corrections are applied to estimate the long-range interactions arising from intermolecular separations greater than r_C (See Section (5.5.1)).

The intramolecular force field parameters for the bond lengths, the bond angles and torsions for each of the force fields used are contained in Table 6-2, Table 6-3, and Table 6-4 respectively.

Table 6-2: Bond length parameters for the NERD, TraPPE and SA1 force fields applicable to Equation 5–3.

NERD
$k_s/k_B = 96500\ \mathrm{K/Å^2}$ (C–C, C=C); $l_0 = 1.54$Å (C–C); $l_0 = 1.34$Å (C=C)
TraPPE
$k_s/k_B = 0\ \mathrm{K/Å^2}$ (C–C, C=C); $l_0 = 1.54$Å (C–C); $l_0 = 1.33$Å (C=C)
SA1
$k_s/k_B = 0\ \mathrm{K/Å^2}$ (C–C, C=C); $l_0 = 1.53$Å (C–C); $l_0 = 1.331$Å (C=C)

Table 6-3: Bond angle parameters for the NERD, TraPPE and SA1 force fields applicable to Equation 5–4.

NERD
$k_\theta/k_B = 62500\ \mathrm{K/rad^2}$ (C–C–C); $\theta_0 = 114.0°$ (C–C–C)
$k_\theta/k_B = 62500\ \mathrm{K/rad^2}$ (C–C=C); $\theta_0 = 124.0°$ (C–C=C)
TraPPE
$k_\theta/k_B = 62500\ \mathrm{K/rad^2}$ (C–C–C); $\theta_0 = 114.0°$ (C–C–C)
$k_\theta/k_B = 70420\ \mathrm{K/rad^2}$ (C–C=C); $\theta_0 = 119.7°$ (C–C=C)
SA1
$k_\theta/k_B = 0\ \mathrm{K/rad^2}$ (C–C–C); $\theta_0 = 112.0°$ (C–C–C)
$k_\theta/k_B = 0\ \mathrm{K/rad^2}$ (C–C=C); $\theta_0 = 124.0°$ (C–C=C)

Table 6-4: Torsional parameters for the NERD, TraPPE and SA1 force fields applicable to Equation 5–5.

NERD
$c_0/k_B = 0\text{K};\quad c_1/k_B = 355.04\text{K};\quad c_2/k_B = -68.19\text{K};\quad c_3/k_B = 791.32\text{K}\quad (\text{C}-\text{C}-\text{C}-\text{C})$
$c_0/k_B = 47.97\text{K};\quad c_1/k_B = 86.31\text{K};\quad c_2/k_B = -109.71\text{K};\quad c_3/k_B = 282.08\text{K}\quad (\text{C}-\text{C}-\text{C}=\text{C})$
TraPPE
$c_0/k_B = 0\text{K};\quad c_1/k_B = 355.03\text{K};\quad c_2/k_B = -68.19\text{K};\quad c_3/k_B = 791.32\text{K}\quad (\text{C}-\text{C}-\text{C}-\text{C})$
$c_0/k_B = 688.5\text{K};\quad c_1/k_B = 86.36\text{K};\quad c_2/k_B = -109.77\text{K};\quad c_3/k_B = 282.24\text{K}\quad (\text{C}-\text{C}-\text{C}=\text{C})$
SA1
$c_0/k_B = 0\text{K};\quad c_1/k_B = 355.03\text{K};\quad c_2/k_B = -68.19\text{K};\quad c_3/k_B = 791.32\text{K}\quad (\text{C}-\text{C}-\text{C}-\text{C})$
$c_0/k_B = 685.96\text{K};\quad c_1/k_B = 86.31\text{K};\quad c_2/k_B = -109.71\text{K};\quad c_3/k_B = 282.08\text{K}\quad (\text{C}-\text{C}-\text{C}=\text{C})$

6.3.2 Method of Simulation

The Configurational-Bias Monte Carlo method was used in conjunction with the Gibbs Ensemble Monte Carlo technique to simulate the four pressure-composition phase diagrams and the temperature-composition phase diagram presented below. The NVT version of the Gibbs Ensemble was used to simulate the two pure component compositions of each of the phase diagrams while the NPT version of the Gibbs Ensemble was used to simulate six intermediate compositions for each of the phase diagrams. The intermediate compositions used were 0.1, 0.2, 0.4, 0.6, 0.8 and 0.9 mole fraction of component 1. The total numbers of molecules used for each of the phase diagrams were as follows:

- 800 molecules for ethane + propene;
- 600 molecules for propene + 1-butene;
- 600 molecules for 1-butene + 1-hexene;
- 350 molecules for 1-hexene + *n*-octane; and
- 150 molecules for *n*-dodecane + 1-octadecene.

The above system sizes were found to give reproducible results without significantly increasing the CPU time required.

The types of moves performed during a simulation were (1) volume changes; (2) molecule identity changes using CBMC and transfers of molecules between simulation boxes; (3) full regrowing of chains and partial regrowing of chains using CBMC; (4) translation of the centre of mass (COM); and (5) rotation around the COM. Each of the moves was selected at random (Smit and Frenkel

(1991)) and the moves were performed with the following fixed probabilities: p1: p2: p3: p4: p5 = 0.006: 0.328: 0.222: 0.222: 0.222, where p1 refers the fixed probability of attempting a move of type 1, i.e. a volume change (Smit *et al.* (1995)). The maximum volume, translation and rotational displacements were chosen in such a way that 50% of all of these moves were accepted. This target of 50% is commonly used as it has been found to be approximately optimal for Monte Carlo simulations.

All the pure component simulations in the NVT Gibbs Ensemble were equilibrated for at least 100,000 cycles while the NPT GEMC mixture simulations were equilibrated for at least 150,000 cycles (a cycle consists of N attempted moves to the system where N is the total number of molecules in the simulation). Care was taken to ensure that the systems had reached equilibrium before production cycles for ensemble averages were started. The standard deviations of ensemble averages were estimated by the standard practice of dividing the production cycles of each run into five blocks and calculating the standard deviation from the averages obtained from each of these five blocks.

To increase simulation efficiency, multiple trial sites were chosen for the first segment of the chain (Esselink *et al.* (1995)). This ranged from 4 for the ethane + propene mixture, which had the shortest chains, to 15 for the *n*-dodecane + 1-octadecene mixture which had the longest chains. Also, a COM-based cut-off was used for the computation of intermolecular forces (Martin and Siepmann (1998)). In this time-saving approach for the interaction between sites on two different molecules A and B, the periodic image of an interaction site on the second molecule is evaluated as the same as the periodic image of the COM of B if the distance between the COM's of A and B is less than $d_{\mathrm{COM}}(\mathrm{A}) + d_{\mathrm{COM}}(\mathrm{B}) + r_C$, where $d_{\mathrm{COM}}(\mathrm{A})$ is the largest distance between the bead furthest from the COM of A. This also increased computational efficiency by reducing the calculation of periodic images of interactions sites for those interaction sites with the same periodic image as the COM of the molecule without compromising accuracy. Finally, the dual-cut-off CBMC algorithm of Vlugt *et al.* (1998) was used for split energy moves with an inner cut-off distance of 4.5Å. This was then corrected to the full potential with tail corrections in the acceptance rule.

It is important to choose initial volumes and molecule numbers that do not differ drastically from the eventual equilibrium values. It was observed that choosing values for the initial volumes or

initial compositions of the two boxes that differed considerably from the ultimate equilibrium values considerably increased the computational time required for reaching equilibrium.

A simple strategy to partially overcome this problem is to use the experimental pure component densities and the experimental compositions as input parameters to determine the initial volumes and initial compositions. The aim, however, was to use the NPT simulations to predict the vapour-liquid equilibrium curves without *a priori* knowledge of equilibrium compositions. A means to reasonably estimate the equilibrium compositions with a minimum of information was thus required.

Since the *n*-alkane and 1-alkene mixtures selected are almost ideal, it was decided to use Raoult's law to estimate the initial compositions. In terms of this very simple description of phase behaviour, the only two inputs required are the pure component saturated vapour pressures which yield an estimate for the vapour and liquid compositions. With the assumption that the molar volumes are linearly additive, it was then possible to estimate reasonable values for the initial molecule numbers and box volumes for the simulations. Using this above approach, suitable initial conditions could be set. The NPT GEMC simulation pressure was then adjusted so that at least 10% of the total number of molecules was in the vapour at equilibrium.

6.4 Transferability of the SA1 Force Field Parameters

Whereas the NERD and TraPPE force fields were used to simulate all mixtures studied, the SA1 force field was only used to simulate the 1-butene + 1-hexene and propene + 1-butene mixtures. Although the SA1 force field was parameterised to simulate the pure component properties of 1-alkenes only, it was attempted to simulate the *n*-alkane mixtures using the relevant functional group parameters from the SA1 force field parameters. It was found, however, that the SA1 force field is not suitable for simulating the *n*-alkane + 1-alkene mixtures. This is because the sp^3 hybridized $-CH_2-$ and $-CH_3$ functional group parameters in the SA1 force field are not transferable to *n*-alkanes (unlike the NERD and TraPPE force fields). For example, the simulation of pure 1-hexene and pure *n*-octane using the SA1 sp^3 hybridized $-CH_2-$ and $-CH_3$ parameters for the *n*-alkanes incorrectly predicted *n*-octane (simulated saturated vapour pressure of 90.2kPa at 328.15K) to be more volatile than 1-hexene (simulated saturated vapour pressure of 85.1kPa at 328.15K). No attempt was made to simulate the ethane + propene mixture using the SA1 force

field since it was expected that the simulated saturated vapour pressure for ethane would have been significantly worse than for the longer *n*-octane molecule discussed above.

6.5 Regression of Critical Temperatures and Critical Densities

An Arrhenius plot of the simulated saturated vapour pressure for 1-octadecene using the NERD, TraPPE and SA1 force fields is give in Figure 6-1 together with the available experimental data. Simulated coexistence liquid and vapour densities are shown in Figure 6-2 for the NERD, TraPPE and SA1 force fields. The critical temperature for 1-octadecene in Figure 6-2 from the simulations was estimated using a density scaling law (Rowlinson and Widom (1989)):

$$\rho_l - \rho_g = \mathbf{B}\left(T_{cr} - T\right)^{\beta} \tag{6–2}$$

while the critical density was estimated using the law of rectilinear diameters (Rowlinson and Swinton (1982)):

$$\frac{\rho_l + \rho_g}{2} = \rho_{cr} + \mathbf{A}\left(T_{cr} - T\right) \tag{6–3}$$

where ρ_l and ρ_g are the liquid and gas densities respectively, **A** and **B** are component-specific constants, **β** is the universal scaling exponent (usually assuming a value between 0.32-0.33), and ρ_{cr} and T_{cr} are the critical density and the critical temperature respectively.

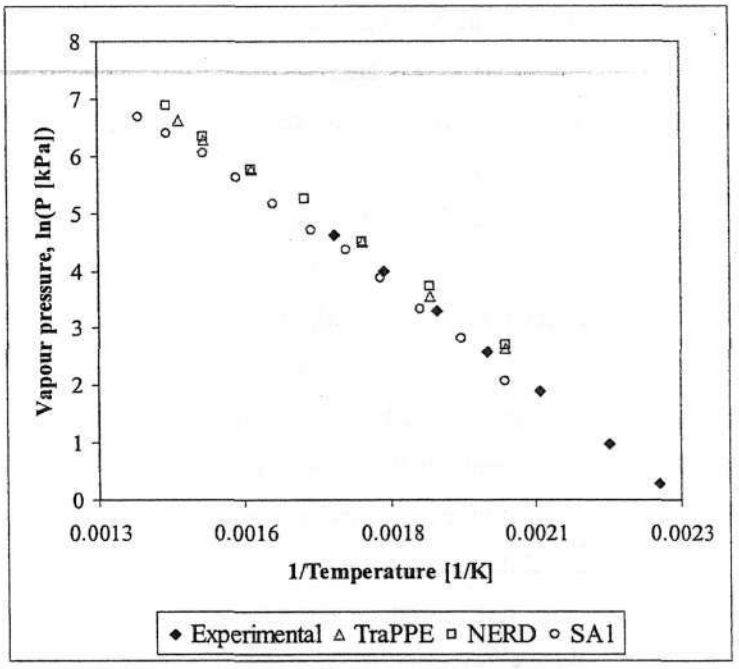


Figure 6-1: Vapour pressure for 1-octadecene as predicted by the TraPPE, NERD and SA1 force fields and the experimental 1-octadecene vapour pressure curve (Jordan and van Winkle (1951)).

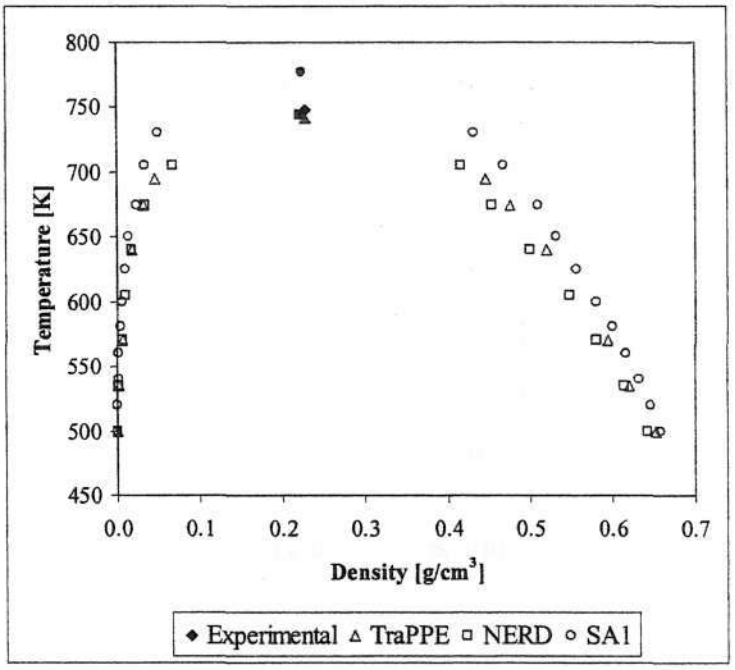


Figure 6-2: Simulated pure component coexistence densities for 1-octadecene from the NERD, TraPPE and SA1 force fields and the critical data of Wakeham *et al.* (2002) and Nikitin and Popov (1999).

In contrast to the poor transferability of the SA1 parameters to the *n*-alkane chains discussed above, preliminary results indicate that these parameters are transferable to the simulation of long 1-alkene molecules. In Figure 6-1 there is good agreement between the available experimental data and the vapour pressures simulated using the SA1 force field. The simulations for 1-octadecene using the NERD and TraPPE force fields are also included in Figure 6-1 and show similarly good agreement. The estimated critical density for 1-octadecene of 0.227g/cm^3, 0.221g/cm^3 and 0.224g/cm^3 for the NERD, TraPPE and SA1 force fields respectively all compare favourably with the extrapolated critical density of 0.227g/cm^3 of Wakeham *et al.* (2002). However, while the estimated critical temperatures of 744K and 742K for the NERD and TraPPE force fields respectively are in reasonable agreement with the experimental value of 748K determined by Nitikin and Popov (1999) using a pulse-heating technique, the estimated critical temperature for 1-octadecene from the SA1 force field off 777K is in substantial disagreement with the experimental value. It should be noted that the SA1 force field also does not reproduce the experimental saturated liquid and vapour densities well for the short 1-alkenes (Spyriouni *et al.* (1999)) while good agreement between the NERD and TraPPE force fields was found for critical temperatures and critical densities of the short 1-alkenes (Wick *et al.* (2000), Nath *et al.* (2001b)).

No additional parameters were required for the simulation of 1-octadecene for any of these united-atom force fields but additional intramolecular parameters are required when extending these force fields to other homologous series for chemicals with sp^2 hybridized double bonds, e.g. the TraPPE force field has been extended to the simulation of aromatic compounds including benzene (Wick *et al.* (2000)).

6.6 *n*-Alkane + 1-Alkene Mixtures

The P-x-y and x-y data for the ethane (1) + propene (2) mixture at 277.6 K are shown in Figure 6-3 and Figure 6-4; the P-x-y and x-y data for the 1-hexene (1) + *n*-octane (2) mixture at 328.15 K are shown in Figure 6-5 and Figure 6-6; and the T-x-y and x-y data for the *n*-dodecane (1) + 1-octadecene (2) mixture at 760 mmHg are shown in Figure 6-7 and Figure 6-8.

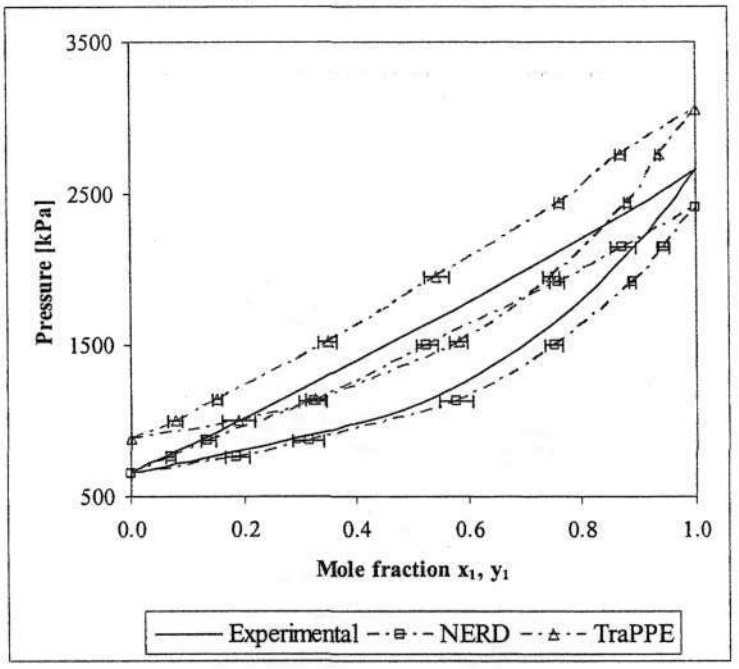


Figure 6-3: Simulated P-x-y data for the ethane (1) + propene (2) mixture at 277.6K and the corresponding experimental VLE (McKay *et al.* (1951)).

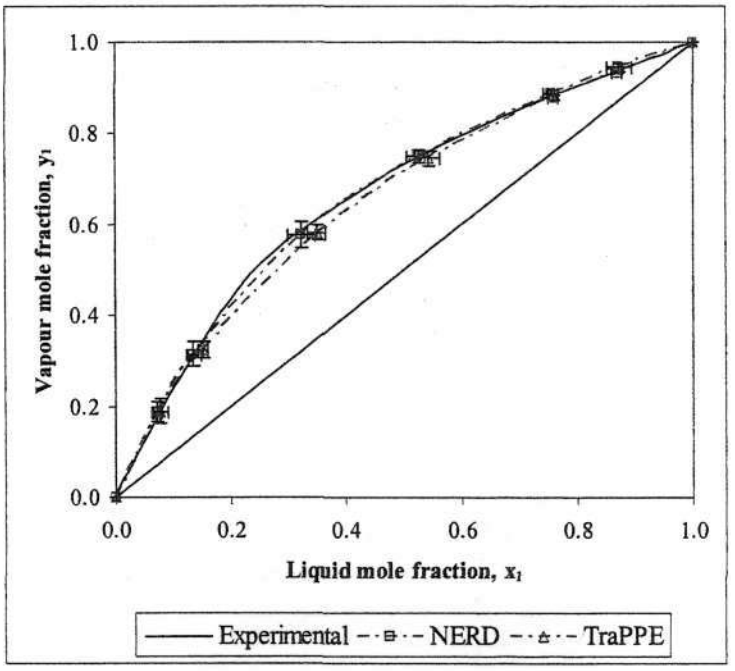


Figure 6-4: Simulated x-y data for the ethane (1) + propene (2) mixture at 277.6K and the corresponding experimental x-y data (McKay *et al.* (1951)).

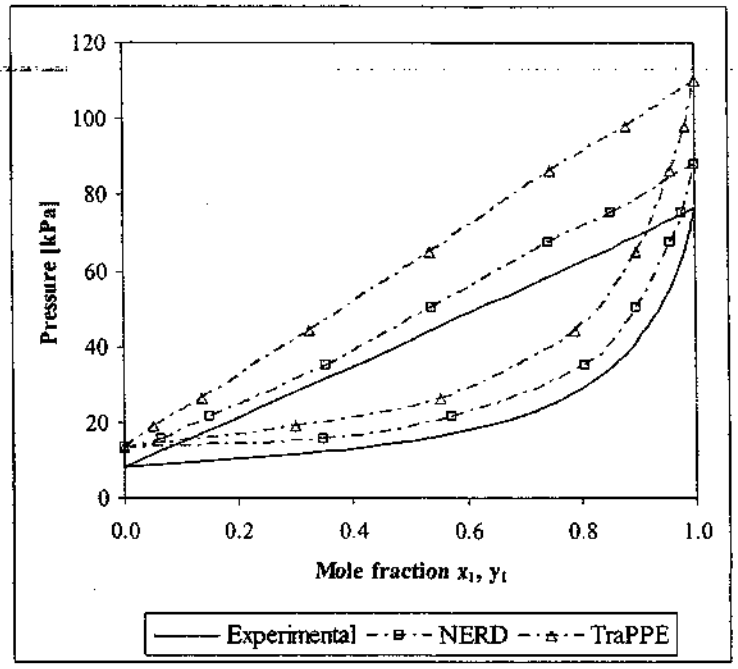


Figure 6-5: Simulated P-x-y data for the 1-hexene (1) + *n*-octane (2) mixture at 328.15K and the corresponding experimental VLE (Dortmund Data Bank).

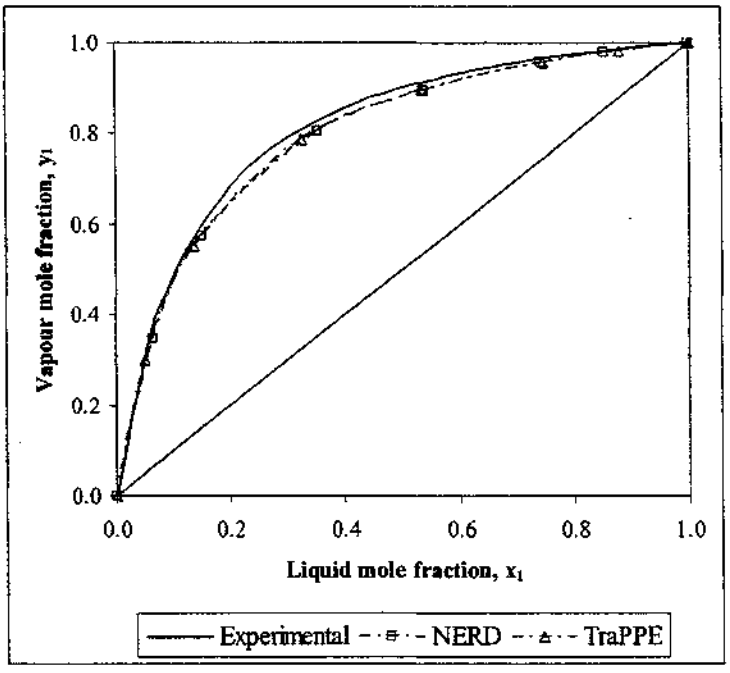


Figure 6-6: Simulated x-y data for the 1-hexene (1) + *n*-octane (2) at mixture 328.15K and the corresponding experimental x-y data (Dortmund Data Bank).

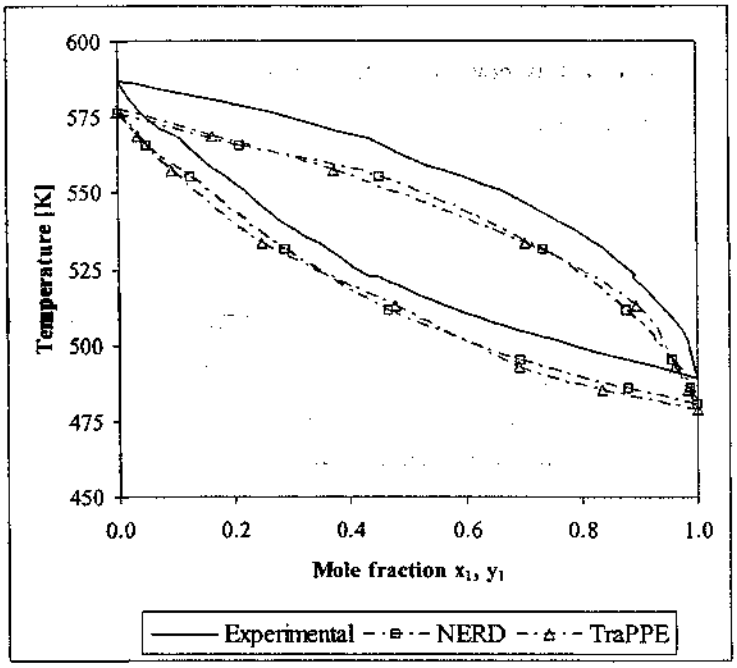


Figure 6-7: Simulated T-x-y data for the *n*-dodecane (1) + 1-octadecene (2) mixture at 760mmHg and the corresponding experimental VLE (Jordan and van Winkle (1951)).

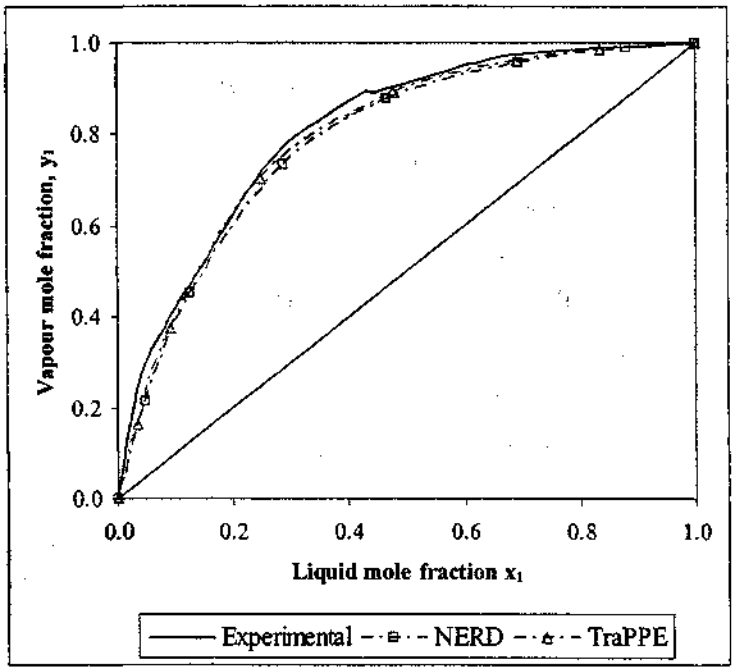


Figure 6-8: Simulated x-y data for the *n*-dodecane (1) + 1-octadecene (2) mixture at 760mmHg and the corresponding experimental x-y data (Jordan and van Winkle (1951)).

It is important to note that the connecting lines in Figure 6-3 to Figure 6-8 are only included to show the trends in the data. It is clear from the P-x-y diagram that the TraPPE force field overestimates the experimental phase envelope. This is a direct consequence of the feature that the TraPPE force field overestimates pure component saturated vapour pressures (Martin and Siepmann (1998), Wick *et al.* (2000)). The reason for this is that the TraPPE parameters were parameterised with the initial intent of reproducing all coexistence properties but eventually the authors settled on selecting experimental liquid coexistence densities, critical temperatures and critical densities only as the primary parameterisation criteria. In particular, the simulated ethane saturated vapour pressure overestimates the experimental value of 664.7kPa by 15% and the simulated propene saturated vapour pressure overestimates the experimental value of 2410kPa by 32%. This results in a simulated ethane + propene phase envelope that is shifted to higher pressures.

In contrast to the TraPPE force field results, the NERD force field yields an ethane + propene phase envelope that is shifted to lower pressures compared to experimentally measured data. Indeed, the simulated saturated vapour pressures of ethane and propene are lower than the corresponding experimental values by 9.2% and 2.1% respectively. As for the TraPPE results, this discrepancy between the simulated results and the experimental data may be because the NERD parameters were not fitted to experimental saturated vapour pressures, but rather the LJ parameters were primarily parameterised to reproduce experimental coexistence liquid densities, critical temperatures and critical densities as well as experimental second Virial coefficients for short hydrocarbon chains (e.g. ethane, ethene, propane and propene). Both force fields, however, yield an isotherm that has the correct shape. This is confirmed by the x-y plot in Figure 6-4 which shows that both the TraPPE and NERD force fields reproduce the experimental x-y plot to within the statistical deviation of the simulation results. Error bars are only included on the P-x-y and x-y diagrams for this mixture only, and are omitted from the remaining phase diagrams for the sake of clarity. The relative errors are, however, similar (in general approximately ±0.01 and never more than ±0.03 mole fraction) for all of the mixtures studied.

As for the ethane + propene mixture, the TraPPE force field overestimates the experimental phase envelope for the 1-hexene + *n*-octane mixture at 328.2 K as shown in Figure 6-5 and Figure 6-6. The experimental *n*-octane saturated vapour pressure of 8.4kPa is overestimated in the simulations by 58% and that of 1-hexene of 76.8kPa is overestimated by 43%. However, in contrast to the ethane + propene mixture, the NERD force field also overestimates the experimental phase envelope. The simulated *n*-octane saturated vapour pressure (obtained from the NERD force field)

is overestimated in the simulations by 16% and that of 1-hexene by 15%. Similar to the data in Figure 6-4, the simulated x-y diagrams for the NERD and TraPPE force fields agree with the experimental data in spite of the deviations in the P-x-y data.

As seen in Figure 6-7 and Figure 6-8, the phase envelope for the isobars of *n*-dodecane + 1-octadecene at 760 mmHg obtained from the NERD and TraPPE force fields are shifted to lower temperatures compared to experimental data. This observation is consistent with the trend that both force fields overestimate the pure component saturated vapour pressure for the longer *n*-alkane and 1-alkene chains. Indeed, the experimental normal boiling points of both 1-octadecene and *n*-dodecane of 587.3K and 488.8K respectively (Jordan and van Winkle (1951)) are underestimated in simulations by both force fields by approximately 10K. As in the case of the two mixtures discussed before, both force fields yield x-y data in good agreement with experiment (Figure 6-8).

Comparison of the simulated data in Figure 6-7 and Figure 6-8 with those in Figure 6-3 to Figure 6-6 reveals that the difference between the T-x-y results obtained from the NERD and TraPPE force fields is far smaller for the *n*-dodecane + 1-octadecene mixture than for the ethane + propene and 1-hexene + *n*-octane mixtures. It thus appears that the NERD and TraPPE force fields yield the most similar results for mixtures comprised of longer alkane and 1-alkene molecules. This is to be expected since, once the $-CH_3$ group had been parameterised for ethane, the $-CH_2-$ groups in the TraPPE force field were parameterised for all *n*-alkanes based on a fit to the vapour-liquid coexistence curve of *n*-octane (Martin and Siepmann (1998)). A similar procedure was adopted for the 1-alkene sp^2 hybridized groups (Wick *et al.* (2000)). In contrast to the TraPPE force field, the NERD force field uses a larger parameter set to describe the *n*-alkane and 1-alkene homologous series and thus only assumes functional group transferability for chain lengths of four carbon units and longer (Nath *et al.* (1998a), Nath and de Pablo (2000)). With increasing chain length, however, the larger number of sp^3 hybridized $-CH_2-$ groups dominate the intermolecular LJ interactions. Since the LJ size and energy parameters for the NERD and TraPPE models for this functional group are very similar (see Table 6-1), these force fields yield increasingly similar phase diagrams for longer alkane and 1-alkene chains.

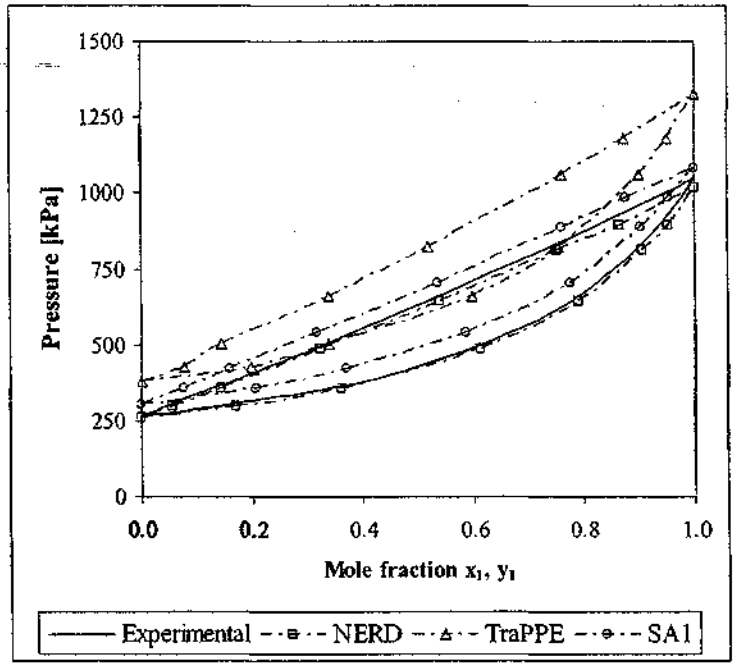


Figure 6-9: Simulated P-x-y data for the propene (1) + 1-butene (2) mixture at 294.3K and the corresponding experimental VLE (Goff *et al.* (1952)).

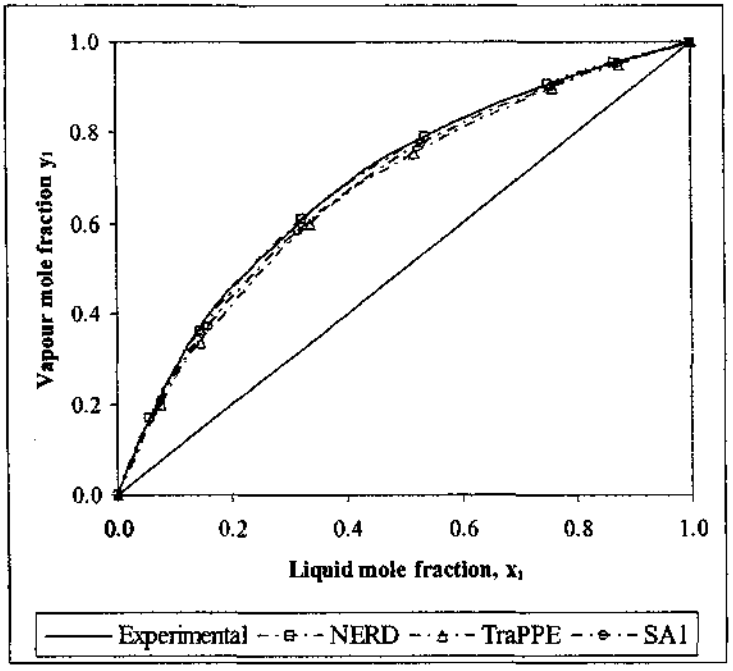


Figure 6-10: Simulated x-y data for the propene (1) + 1-butene (2) mixture at 294.3K and the corresponding experimental x-y data (Goff *et al.* (1952)).

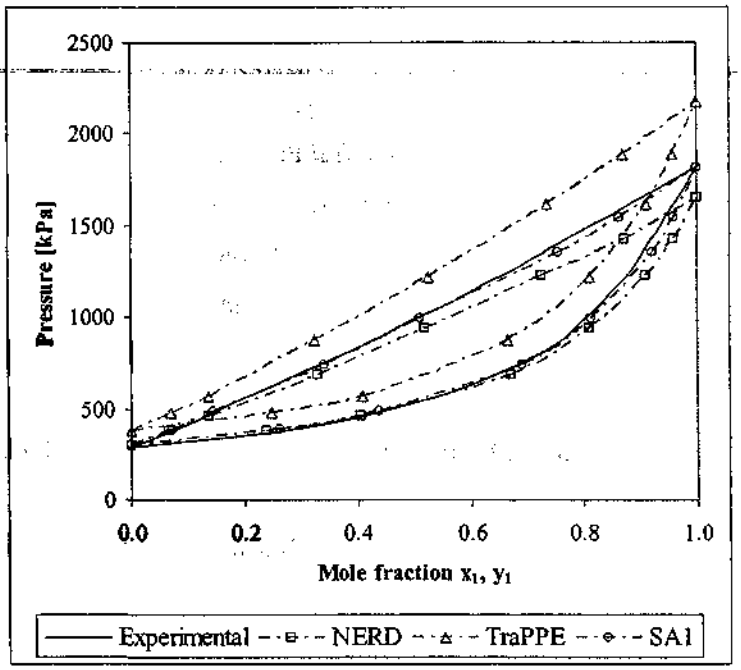


Figure 6-11: Simulated P-x-y data for the 1-butene (1) + 1-hexene (2) mixture at 373K and the corresponding experimental VLE (Laugier and Richon (1996)).

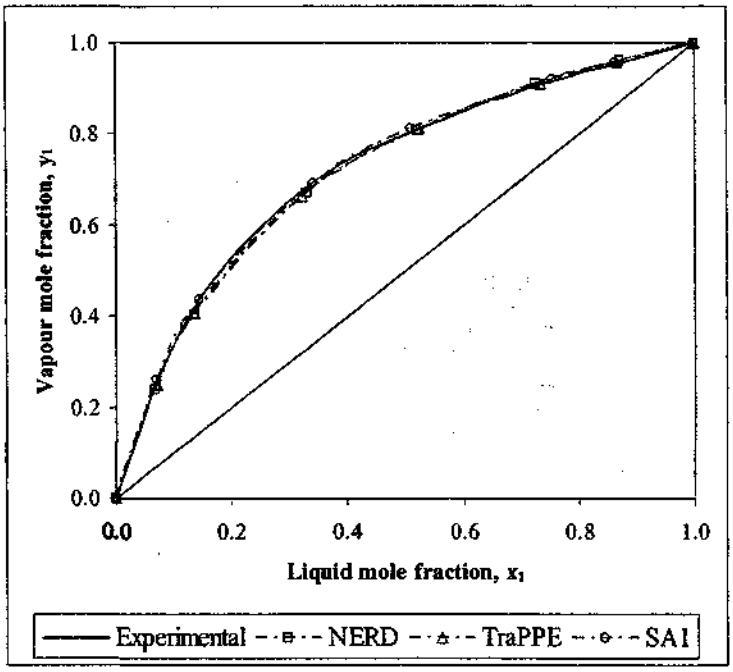


Figure 6-12: Simulated x-y data for the 1-butene (1) + 1-hexene (2) mixture at 373K and the corresponding experimental x-y data (Laugier and Richon (1996)).

6.7 1-alkene + 1-alkene mixtures

It is of interest to ascertain if the SA1 force field yields P-x-y data that is in better agreement with experiment than the data from the NERD and TraPPE simulations, and if this force field also yields simulated x-y results in good agreement with experimentally measured data. As pointed out by Nath *et al.* (2001), the available binary VLE data for 1-alkene mixtures are more limited than for *n*-alkanes. For this reason, only the almost ideal mixtures of propene + 1-butene at 294.3K and 1-butene + 1-hexene at 373K were simulated.

The simulated P-x-y and x-y diagrams for the propene + 1-butene mixture at 294.3K are shown in Figure 6-9 and Figure 6-10. As for the alkane + 1-alkene mixtures discussed above, the phase envelope obtained from the TraPPE force field is shifted to higher pressures compared to the experimental data (the experimental saturated vapour pressure of 1-butene of 263.4kPa is overestimated by 44%, and that of propene of 1049kPa is overestimated by 26%). The NERD force field yields a phase envelope in very good agreement with experimentally measured data, although it slightly underestimates the experimental data (the simulations underestimate the experimental saturated vapour pressure of 1-butene by 2% and that of propene by 3%). The SA1 force field also yields data in good agreement with the experimental P-x-y data, although the experimental saturated vapour pressure of 1-butene is overestimated by 15% and that of propene by 3%. This is what results in the slight overestimation of the phase envelope shown in Figure 6-9. Thus, even though the SA1 force field is parameterised to reproduce pure component saturated vapour pressures, it yields a 1-butene saturated vapour pressure that is higher than experiment. However, 1-butene was the shortest 1-alkene used in the parameterisation of the SA1 Lennard-Jones terms and no distinction is made between the sp^3 hybridized $-CH_2-$ and $-CH_3$ functional group parameters (Table 6-1). All three force fields reproduce the experimental x-y diagram, as was found for the alkane + 1-alkene mixtures discussed earlier.

The P-x-y and x-y diagrams for 1-butene + 1-hexene at 373K are shown in Figure 6-11 and Figure 6-12. Similar trends to those observed for the propene + 1-butene mixture are seen here. The TraPPE force field overestimates the phase envelope, with the experimental 1-butene and 1-hexene pure component saturated vapour pressures of 1822kPa and 292.6kPa being overestimated by 19% and 29% respectively. The NERD force field underestimates the 1-butene pure component saturated vapour pressure by 9% while the 1-hexene saturated vapour pressure is overestimated by 2%. The P-x-y envelope obtained from the NERD force field is thus in good agreement with

experiment. The SA1 force field yields P-x-y data that is in best agreement with experimental phase equilibria. For this force field the experimental pure component saturated vapour pressures of 1–hexene and 1-butene are overestimated by only 5% and 0.5%, respectively. Once again, in spite of the deviations in some of the P-x-y data, all force fields yield x-y diagrams that are in good agreement with the experimental data.

It is interesting to note that the NERD force field underestimates the pure component vapour pressure for some mixtures and overestimates the pure component vapour pressure for other mixtures. The TraPPE force field, however, consistently overestimates the pure component vapour pressure. This possibly indicates that the accuracy of the simulated vapour pressure from the NERD force field may depend on either the chain length or on the reduced temperature of the simulation.

In summary, for both sets of alkane + 1-alkene and 1-alkene + 1-alkene mixtures where both components are short chains the NERD force field yields P-x-y data that is in better agreement with experiment than the TraPPE data (Figure 6-3, Figure 6-5, Figure 6-9, and Figure 6-11). However, as discussed with reference to the *n*-dodecane + 1-octadecene phase diagram shown in Figure 6-7, the TraPPE and NERD force fields yield similar data for *n*-alkane + 1-alkene mixtures that contain longer chains, and both models underestimate the experimental isotherm. Similarly to the NERD force field, the SA1 potential also yields P-x-y data that are in good agreement with the experimental data for 1-alkene + 1-alkene mixtures that contain short chains. Also, in spite of deviations in some simulated P-x-y data from the experimental data, all force fields yield x-y data in close agreement with the experimental data for all mixtures studied.

6.8 Ideal Solution Behaviour of the Mixtures Studied

Figure 6-13 and Figure 6-14 show plots of the relative excess volumes of the liquid and vapour phases as a function of composition for the 1-hexene + *n*-octane mixture at 328K and the 1-butene + 1-hexene mixture at 373K respectively. Similar results were obtained for the other mixtures and are thus not shown here. The data sets are obtained for overall mole fractions of component 1 of 0.0, 0.1, 0.2, 0.4, 0.6, 0.8, 0.9 and 1.0 as one follows a data set from left to right in the figures.

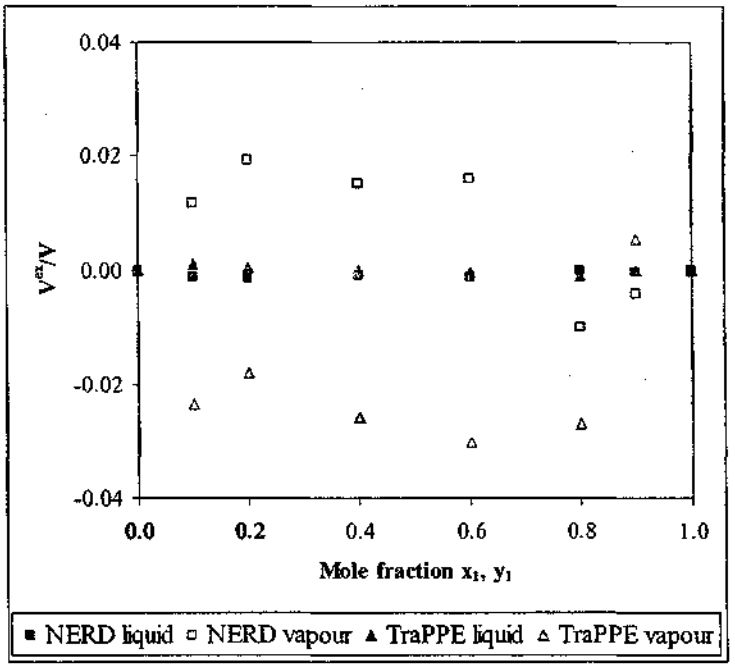


Figure 6-13: Plot of the ratio of the excess to the real volumes for both liquid and vapour phases for the 1-hexene + *n*-octane mixture at 328.15K

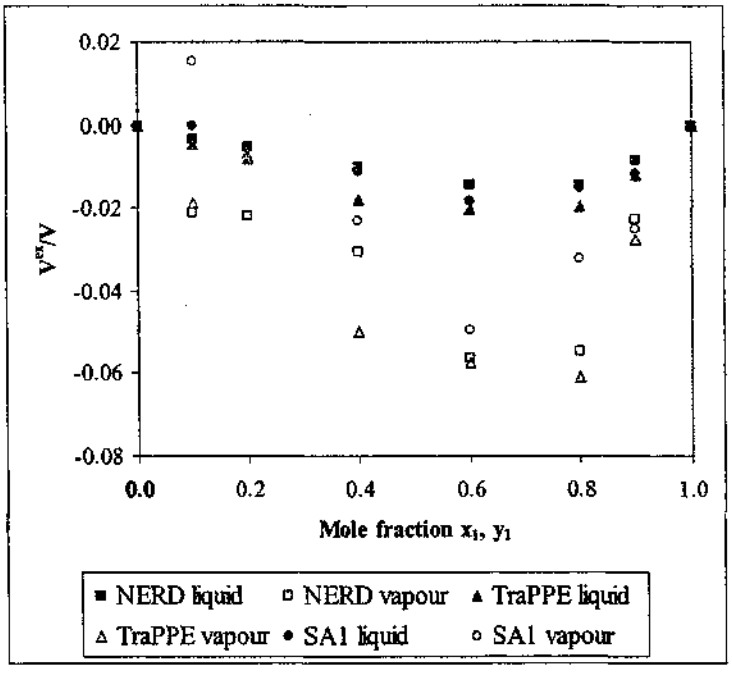


Figure 6-14: Plot of the ratio of the excess to the real volumes for both liquid and vapour phases for the 1-butene + 1-hexene mixture at 373K.

The proximity of the data about the $V^{ex}/V = 0$ line confirms that the force fields used at least qualitatively reproduce the ideal solution behaviour of these two mixtures. This indicates why the method of choosing the initial conditions for the simulations worked because this method inherently required a nearly ideal mixture. For both mixtures, however, it is noteworthy that the vapour phase deviates more from ideal solution behaviour than the liquid phase for all force fields. Both mixtures should, however, display ideal solution behaviour, particularly in the vapour phase. This curious deviation from what should be the expected trend is a result of the emphasis which was placed on the accurate reproduction of liquid phase densities during the pure component parameterisation of these force fields. All three force fields in general under predict the vapour densities in the original publications which leads to the overall negative deviation of the vapour excess volumes from ideal solution behaviour. The one exception is the NERD force field's deviation from ideal gas behaviour for the *n*-octane n+ 1-hexene mixture. This result may be understood, however, as being a consequence of the larger parameter set for the NERD force field as well as a result of the large emphasis that was placed on the parameterisation of the 1-alkene parameters using 1-hexene pure component data (Nath and de Pablo (1999), Nath *et al.* (2001)).

CHAPTER 7:

TESTING THE LIMITS OF TRANSFERABLE FORCE FIELDS

"Great things are not done by impulse but by a series of small things brought together." – Vincent van Gogh

7.1 Introduction

In view of the results stemming from the previous Chapter, it became interesting to determine whether the transferable force fields studied could be extended to different homologous series of chemicals for which the parameterised functional groups should be sufficient, but for which these transferable force fields were not originally and explicitly parameterised.

Conjugated alkenes represent a homologous series for which the intermolecular parameters listed in Table 6-1 and the intramolecular parameters listed in Table 6-2 to Table 6-4 should be largely sufficient. By far the most important conjugated alkene is 1,3-butadiene. Indeed, 1,3-butadiene ranks 36th in the top 50 most produced chemicals in the United States, and approximately 6 million metric tons are produced annually. Isolation of 1,3-butadiene is currently based on butadiene-containing C_4 fractions from the steam cracking of naphtha, gas oil, and other higher boiling hydrocarbon fractions to form ethylene and homologous compounds (Grub and Löser (2000)). This makes accurate knowledge of 1,3-butadiene mixture VLE data vital. The majority of 1,3-butadiene is used in the production of styrenebutadiene rubber copolymers (SBR) but other applications include its use as a polymer component for polybutadiene, styrene-butadiene latex and acrylonitrile-butadiene-styrene (ABS) resins. Several other conjugated alkenes also participate in industrially important chemical reactions. For example, the Diels-Alder reaction represents an important route towards the formation of cyclic hydrocarbon compounds (Grub and Löser (2000)). Consequently, it is also important to have an accurate description of the phase behaviour of the general homologous series of conjugated alkenes.

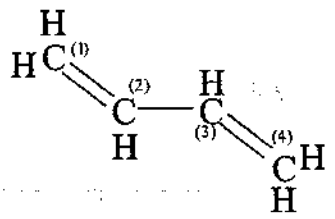
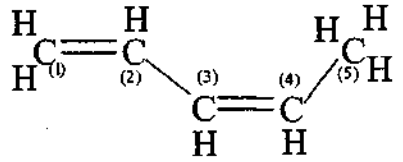
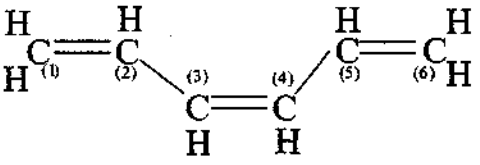
1,3-Butadiene has been simulated previously using the AUA4 anisotropic force field of Bourasseau *et al.* (2003) as part of a wider force field parameterisation for alkenes. Over the limited temperature range from 245K to 295K, the AUA4 olefin force field gave average relative errors of 2.6% on the liquid density, 1.1% on the latent heat of vaporisation, and 8.6% on the saturated vapour pressure. No emphasis was placed on correctly reproducing the vapour densities.

7.2 Intramolecular Parameters

The NERD and TraPPE force fields were considered for testing whether alkane and alkene parameters from a transferable force field could be used to extend a force field to the simulation of conjugated alkenes. The Spyriouni *et al.* olefin force field was not selected because this force field was shown to provide less accurate results in the previous Chapter.

Thermophysical property information for linear conjugated alkenes with more than 6 carbons is essentially unavailable. Consequently, 3 conjugated alkenes were selected for which there is thermophysical property data available (Vargaftik (1975), Smith and Srivastava (1986)). These are listed in Table 7-1:

Table 7-1: Names and molecular structures of the most stable conformers for the conjugated alkenes studied.

1,3-butadiene	
cis-1,3-pentadiene	
cis-1,3,5-hexatriene	

7.2.1 Torsional Parameters

The intramolecular torsional parameters in Table 6-4 are insufficient because there is no provision for a torsion for the internal *cis* double bond between two sp^2 hybridized functional groups (for *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene) and for the torsion about the carbon-carbon single bond in a $=CH-CH=$ sequence for all three components. The NERD olefin force field (Nath *et* al. (2001a)) has neither set of required parameters. The parameters for the *cis*-torsion around a double bond between two sp^2 hybridized functional groups were therefore borrowed from

the TraPPE force field (Wick *et al.* (2001)). Unlike for torsions around a carbon-carbon single bond, the rotation around a carbon-carbon double bond is strongly hindered and the TraPPE olefin force field uses a harmonic potential:

$$U_{tors} = k_\phi \left(\phi - \phi_0\right)^2 \qquad (7\text{–}1)$$

instead of the more familiar cosine series where k_ϕ is the harmonic torsional constant and ϕ_0 is the equilibrium torsional angle. The values of these parameters are contained in Table 7-2. The angle ϕ_0 is defined as zero in the *cis*-conformation and π radians in the *trans*-configuration. This borrowed torsional potential was found to work extremely well for the modified NERD force field because it was found to limit the structure of a *cis*-configuration to small perturbations about a torsional angle of zero as required.

Table 7-2: Intramolecular torsional parameters used for the simulation of the conjugated alkenes.

	TraPPE	**NERD**	**Bock *et al.***
***cis* torsion, Eq. 7–1**			
k_ϕ (K/rad^2)	12400	-	-
ϕ_0	0°	-	-
=CH–CH= **torsion, Eq. 7–2**			
c_0 (K)	-	-	0
c_1 (K)	-	-	249.84
c_2 (K)	-	-	1287.70
c_3 (K)	-	-	440.83
c_4 (K)	-	-	-129.33

Several investigators have studied the torsional energy arising from rotation about the =CH–CH= sequence (Nevins *et al.* (1996), Murcko *et al.* (1996), Karpfen *et al.* (1997)). In particular, the torsional potentials of Bock *et al.* (1979) and Szalay *et al.* (1989) are shown in Figure 7-1:

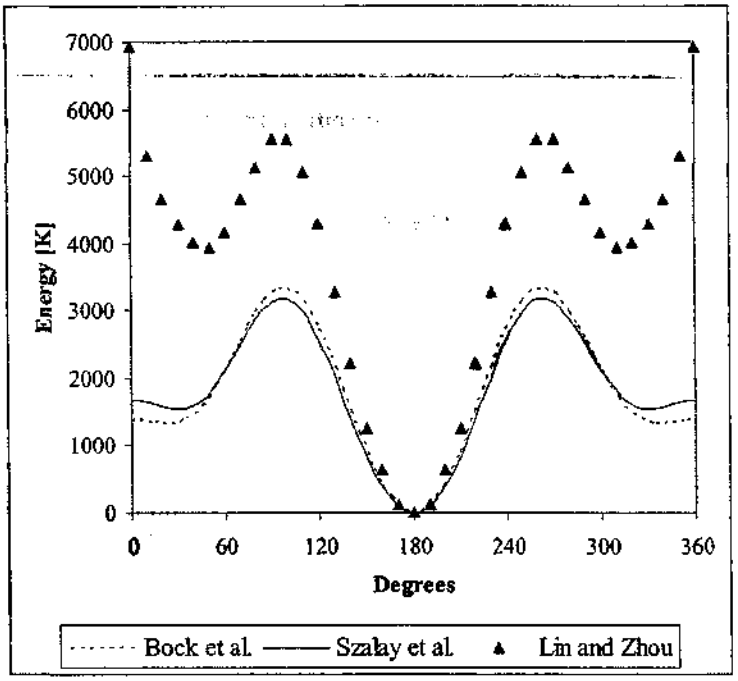


Figure 7-1: Comparison of the torsional potential of Bock *et* al. (1979) and Szalay *et* al. (1989) for 1,3-butadiene as well as the ***ab initio*** torsional data of Liu and Zhou (1993) for the =CH–CH= torsion of *cis*-1,3,5-hexatriene.

The *ab initio* calculations of Szalay *et al.* (1989) were used by Bourasseau *et al.* (2003) to describe the 1,3-butadiene torsional potential. The =CH–CH= torsional potential of Bock *et al.* (1979), however, was based on calorimetric and spectroscopic data as well as several sets of *ab initio* calculations and was therefore selected for the present work. It is given by a cosine series:

$$U_{tors} = c_0 + \sum_{n=1,4} c_n \left[1 + (-1)^{n-1} \cos(n\phi)\right] \qquad (7\text{–}2)$$

where the constants c_i are listed in Table 7-2. From Figure 7-1 it is clear that both torsional potentials display extrema of similar magnitude at similar torsional angles. Indeed, it was found that there was a negligible difference in the pure component results obtained from the two torsional potentials of Szalay *et al.* (1989) and Bock *et al.* (1979).

No data could be found concerning the =CH–CH= torsion of *cis*-1,3-pentadiene. Liu and Zhou (1993), however, did conduct an *ab initio* study of the two =CH–CH= torsions which are present in *cis*-1,3,5-hexatriene and data is also plotted in Figure 7-1. The energies associated with the conformation behaviour of =CH–CH= torsions in *cis*-1,3,5-hexatriene are considerably higher

than those arising from the same torsion in a 1,3-butadiene molecule. Both molecules, however, exhibit a zero torsional energy contribution at 180° consistent with a *trans* configuration of the =CH–CH= sequence and display qualitative agreement in the location of local torsional energy maxima at 90° and 270°. The reason for the local maxima observed at approximately 90° and 270° is due to electron correlation effects (Szalay *et al.* (1989). The 1,3-butadiene molecule, however, displays a torsional energy local minimum at 0° (360°) while the *cis*-1,3,5-hexatriene molecule displays a torsional energy local maximum at 0° (360°).

The apparent discrepancy between the two torsional descriptions can be explained in terms of the extra two carbons constituting the *cis*-1,3,5-hexatriene molecule. This is important because otherwise the approach described in Section (5.4) of using additive bond length, bond angle, torsional and intramolecular LJ contributions would be inconsistent with physical reality. Indeed, when a *cis*-1,3,5-hexatriene molecule rotates to a =CH–CH= *cis* configuration at 0°, and to a *gauche* configuration at 90° or 270°, the extra $C_{(5)}$ and $C_{(6)}$ carbons listed in Table 7-1 of *cis*-1,3,5-hexatriene are also brought closer to the $C_{(1)}$ carbon. These functional group centres must therefore result in a higher torsional energy associated in these configurations than that observed for the 1,3-butadiene molecule due to interatomic repulsions at short separations. This intramolecular energy contribution may therefore be correctly accounted for by the inclusion of the intramolecular LJ interactions between carbon atoms separated by more than 3 bonds as described in Section (5.4). As a result, the torsional potential of Bock *et al.* (1979) was also used to describe the =CH–CH= torsion for the *cis*-1,3,5-hexatriene and the *cis*-1,3-pentadiene molecules.

7.2.2 Bond Angle Parameters

The two sets of bond angle bending parameters for the =CH– bond angle for the NERD and TraPPE force fields in Table 6-3 differ significantly. There are also subtle differences between these values and the equilibrium values suggested by Bock *et al.* (1979) for 1,3-butadiene and by Liu and Zhou (1993) for *cis*-1,3,5-hexatriene as shown in Table 7-3. It is interesting to note, however, that the equilibrium value for θ_0 from the NERD force field is in closest agreement with the suggested equilibrium values from Bock *et al.* (1979) and Liu and Zhou (1993). It was therefore important to determine what effect, if any, that this would have on the pure component equilibrium properties.

Table 7-3: = CH – bond angle parameters for the NERD and TraPPE force fields as well as the suggested equilibrium values from Bock *et al.* (1979) and Liu and Zhou (1993).

	NERD	TraPPE	Bock *et al.*	Liu and Zhou
1,3-butadiene				
k_θ/k_b, K/rad^2	62500	70420	-	-
θ_0, ($C_{(1)}=C_{(2)}-C_{(3)}$)	124°	119.7°	123.3°-125.9°	-
***cis*-1,3,5-hexatriene**				
k_θ/k_b, K/rad^2	62500	70420	-	-
θ_0, ($C_{(1)}=C_{(2)}-C_{(3)}$)	124°	119.7°	-	122.1°
θ_0, ($C_{(2)}-C_{(3)}=C_{(4)}$)	124°	119.7°	-	125.9°

7.2.3 Bond Length Parameters

There are also subtle differences in the bond lengths between the functional groups suggested by the NERD and TraPPE force fields and the values reported by Bock *et al.* (1979) and Liu and Zhou (1993) as seen in Table 7-4.

Table 7-4: Bond length parameters for the NERD and TraPPE force fields as well as the suggested equilibrium values from Bock *et al.* (1979) and Liu and Zhou (1993).

	NERD	TraPPE	Bock *et al.*	Liu and Zhou
1,3-butadiene				
k_s/k_B, K/Å^2	96500	0	-	-
l_0 ($C_{(1)}=C_{(2)}$)	1.34Å	1.33Å	1.34Å	-
l_0 ($C_{(2)}-C_{(3)}$)	1.54Å	1.54Å	1.46Å -1.48Å	-
***cis*-1,3,5-hexatriene**				
k_s/k_B, K/Å^2	96500	0	-	-
l_0 ($C_{(1)}=C_{(2)}$)	1.34Å	1.33Å	-	1.34Å
l_0 ($C_{(2)}-C_{(3)}$)	1.54Å	1.54Å	-	1.46Å
l_0 ($C_{(3)}=C_{(4)}$)	1.34Å	1.33Å	-	1.36Å

Significantly, both the NERD and TraPPE force fields overestimate the C – C bond between two sp^2 hybridized carbons by about 0.08Å for both *cis*-1,3,5-hexatriene and 1,3-butadiene.

7.3 Simulation Methodology

The CBMC method was used in conjunction with the NVT version of the Gibbs Ensemble for the pure component simulations, while the CBMC method was used in conjunction with the NPT version of the Gibbs Ensemble for the binary simulation. The total numbers of molecules used for each of the phase diagrams were as follows:

- 500 1,3-butadiene molecules,
- 400 *cis*-1,3-pentadiene molecules, and
- 400 *cis*-1,3,5-hexatriene molecules
- 400 total molecules for the 1,3-butadiene + *n*-heptane binary mixture. Two pure component and six intermediate compositions corresponding to 0.1, 0.2, 0.4, 0.6, 0.8, and 0.9 overall mole fraction of 1,3-butadiene were simulated.

For the pure component simulations, the box volumes were adjusted so that the liquid and vapour box volumes at the end of a simulation were approximately equal in size. For the NPT simulations, the imposed pressure of the simulation was adjusted so that approximately 10% of the total molecules were in the vapour box at equilibrium. For the binary simulations, the estimation procedure described in the previous Chapter which assumed relatively ideal mixture behaviour was not used because the phase envelope for the 1,3-butadiene + *n*-heptane mixture was large compared to the mixtures studied in the Chapter Six. The simulation pressure was adjusted on a trial and error basis without introducing significant stability problems.

As for the simulations in the previous Chapter, five distinct types of moves were performed for both the pure component and binary simulations, namely

1) volume changes,
2) transfers of molecules between simulation boxes,
3) full regrowing of chains and partial regrowing of chains using CBMC,
4) translation of the centre of mass (COM), and
5) rotation around the COM.

Each of the moves was selected at random and, were performed with the fixed probabilities: p1: p2: p3: p4: p5 = 0.006: 0.328: 0.222: 0.222: 0.222. The maximum volume, translation and rotational displacements were chosen in such a way that 50% of all of these moves were accepted.

All the simulations were equilibrated for at least 50,000 cycles. The production runs consisted of at least 50,000 cycles for the pure component simulations and at least 100,000 cycles for the binary simulations. The standard deviations of ensemble averages were computed by dividing the production cycles of each run into five blocks and calculating the standard deviation from the averages of these five blocks.

For both the pure component and binary simulations the same and unaltered LJ parameters listed in Table 6-1 were used.

7.4 Pure Component Simulations of 1,3-Butadiene

The pure component properties of 1,3-butadiene over the temperature range of 223K – 403K were simulated using the appropriate NERD and TraPPE intermolecular parameters of Table 6-1 and the bond angle and bond length parameters of Table 7-3 and Table 7-4 but including the torsional potential of Bock *et al.* (1979) as described in Table 7-2. It was found, however, that while the NERD force field gave reasonable agreement with the experimental data (Smith and Srivastava (1986)), the TraPPE force field was considerably in error. The most significant difference between the intramolecular parameters of the NERD and TraPPE force fields resides in the description of the =CH– bond angle. The equilibrium =CH– bond angle of the NERD force field is, however, in closer agreement with the recommended values of Bock *et al.* (1979) and Liu and Zhou (1993). Since the intention was to simulate 1,3-butadiene pure component behaviour without adjusting the intermolecular parameters, but with flexibility regarding the intramolecular parameters, the NERD parameters for the =CH– bond angle were therefore also used for the 1,3-butadiene simulations using the TraPPE force field. This significantly improved the simulated results obtained from the TraPPE force field for 1,3-butadiene. The results for the NERD and modified TraPPE force fields denoted as "Version 1" for the liquid and vapour densities, the saturated vapour pressures and the latent heats of vaporization are shown in Figure 7-2 to Figure 7-5 and are listed in Table 7-7 to Table 7-9 in Appendix 7A at the end of this Chapter:

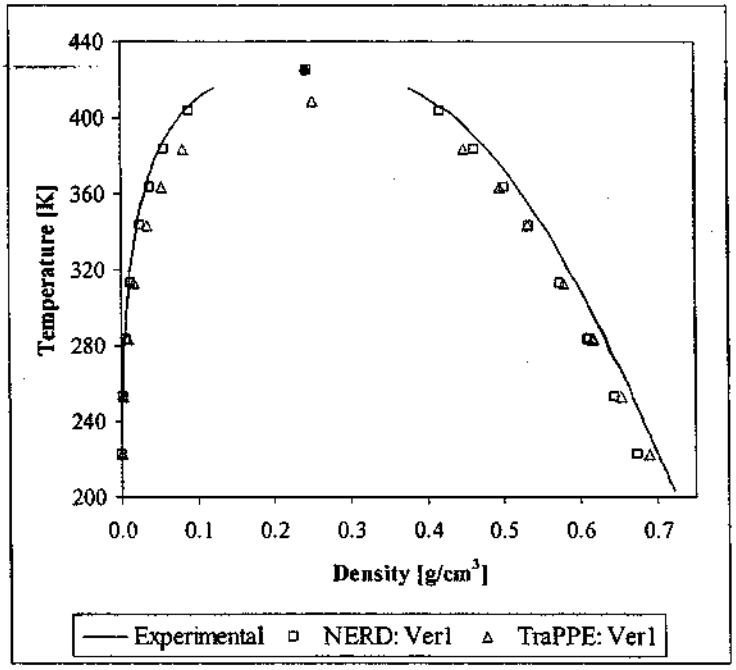


Figure 7-2: 1,3-Butadiene liquid and vapour coexistence densities from "Version 1" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986). Critical points are shown as solid symbols.

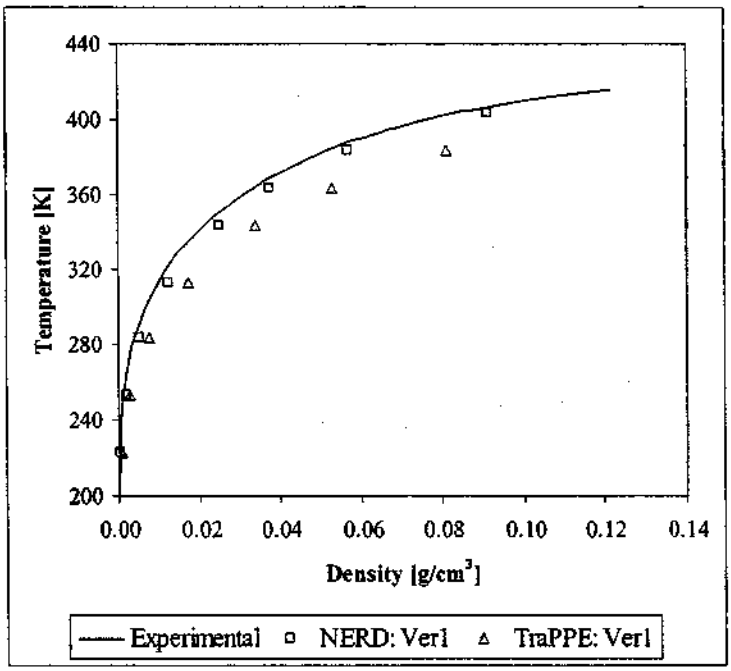


Figure 7-3: 1,3-Butadiene vapour coexistence densities from "Version 1" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

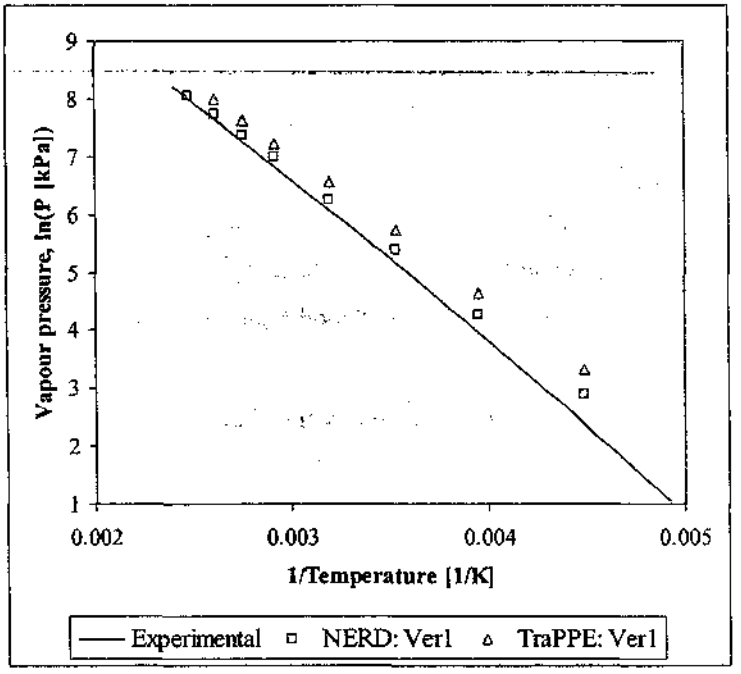


Figure 7-4: 1,3-Butadiene saturated vapour pressures from "Version 1" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

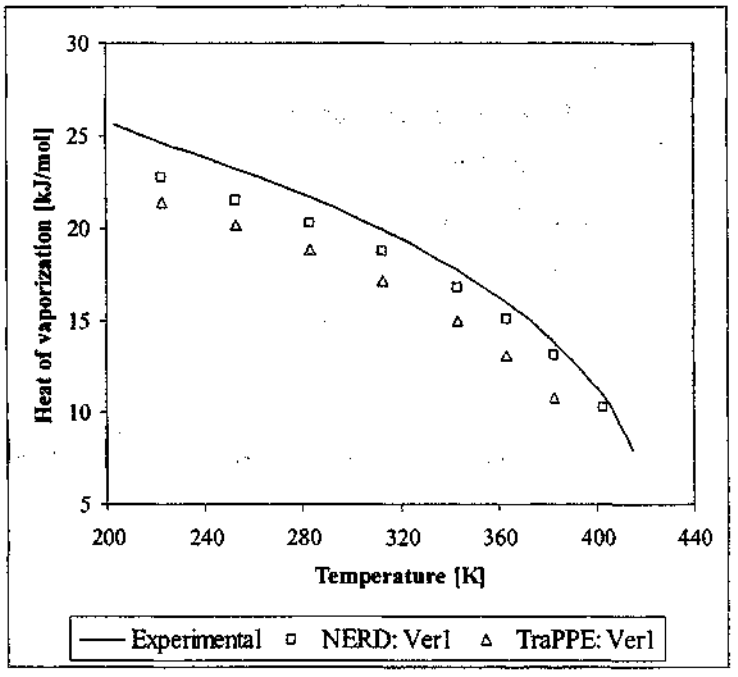


Figure 7-5: 1,3-Butadiene latent heats of vaporization from "Version 1" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

The NERD force field in general estimates the pure component saturated liquid and vapour densities (Figure 7-2 and Figure 7-3), vapour pressures (Figure 7-4) and heats of vaporization (Figure 7-5) of 1,3-butadiene better than the TraPPE force field except in the lower half of the temperature range considered where the TraPPE force field displays a better prediction of the pure component saturated liquid densities. The average differences between the experimental pure component data of 1,3-butadiene and the simulated data using the NERD and TraPPE force fields are listed in Table 7-5 together with the "Version 2" modifications.

The saturated liquid and vapour densities estimated by the TraPPE force field are in general shifted to lower temperatures. Consequently, while the critical temperature of 424K and the critical density of 0.2416g/cm^3 estimated by NERD force field using Equations (6–2) and (6–3) with $\beta = 0.32$ are close to the experimentally measured critical temperature and density of 425.15K and 0.245g/cm^3 respectively, the TraPPE force field provides a considerably poorer estimate of the critical temperature and the critical density of 409.1K and 0.250g/cm^3 respectively. It has been shown previously by Chen *et al.* (1998), Martin and Siepmann (1998), and Ungerer *et al.* (2000) that the critical temperature increases with increasing ε and that the critical density decreases with increasing σ. The NERD sp^2 hybridized LJ ε parameter for the –CH= group is 1K smaller than the corresponding group for the TraPPE force field but the NERD sp^2 hybridized LJ ε parameter for the CH_2= group is 7.5K larger than the same group of the TraPPE force field. In addition, both of the relevant sp^2 hybridized LJ σ parameters for the –CH= and the CH_2= groups have larger values in the NERD force field than they do in the TraPPE force field. The combination of the larger LJ σ and ε parameters for the NERD force field are what results in the better overall shape of the pure component properties for the NERD force field over the TraPPE force field. This suggests that the magnitudes of the NERD parameters are more suited to the simulation of conjugated alkene molecules.

Since a change of only a few degrees in the equilibrium bond bending angle achieved a considerable improvement in the agreement between the simulated and experimental pure component properties, it was decided to adjust the equilibrium bond lengths for both the NERD and TraPPE force fields to more correctly reproduce the molecular geometry of 1,3-butadiene. From Table 7-4, it may be seen that while both the TraPPE and NERD force fields use equilibrium values for the two $CH_2 = CH$ double bonds that are close to the 1.34Å suggested by Bock *et al.* (1979), the central $CH - CH$ single bond differs by as much as 0.08Å. The reason for this discrepancy resides in the way in which the transferable force fields are, in general, parameterized. Indeed, all

single bonds between sp^2 or sp^3 hybridized carbons are assumed to have a bond length of 1.54Å for both the TraPPE and the NERD (for molecules with more than three carbons) force fields. Clearly, the force exerted on a sp^2 hybridized –CH= group by a sp^3 hybridized CH_2= group along the single bond connecting them will be less than the force exerted by another sp^2 hybridized –CH = group because the second –CH= group has one fewer hydrogen atom attached. It is important to determine whether the assumption of using a single equilibrium bond length for all C – C bonds is valid. Consequently, the NERD and TraPPE force fields were adjusted to treat the central CH – CH single bond as having an equilibrium length of 1.46Å. The two double bonds retained their equilibrium lengths bonds as listed in Table 7-4 because these lengths are very similar to the value suggested by Bock *et al.* (1979). All other intermolecular LJ and intramolecular parameters were left unchanged. The results of these "Version 2" simulations are shown in Figure 7-6 to Figure 7-9 and the data is listed in Table 7-7 to Table 7-9 in Appendix 7A at the end of this Chapter:

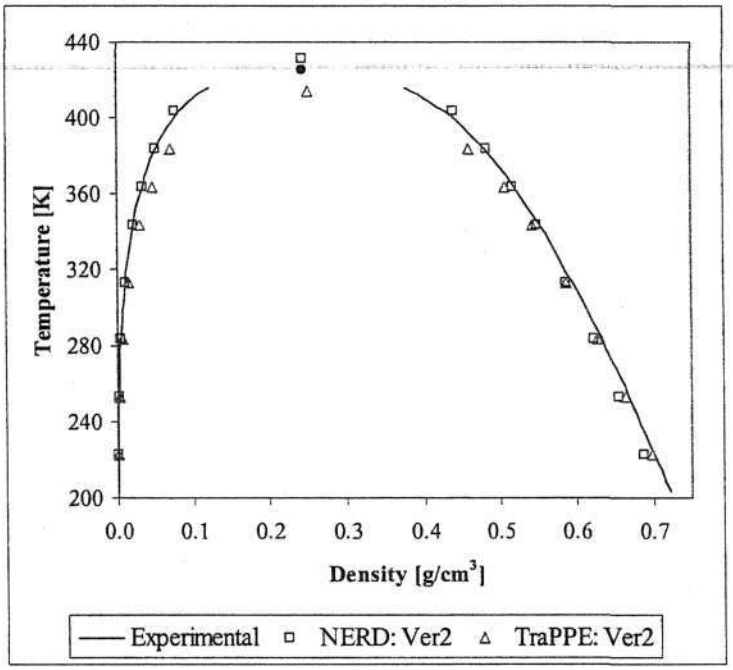


Figure 7-6: 1,3-Butadiene liquid and vapour coexistence densities from "Version 2" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986). Critical points are shown as solid symbols.

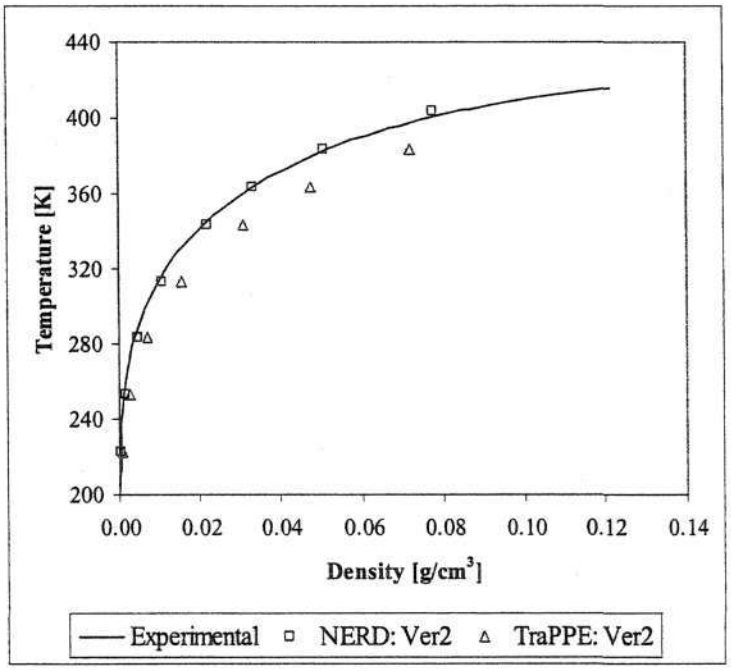


Figure 7-7: 1,3-Butadiene vapour coexistence densities from "Version 2" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

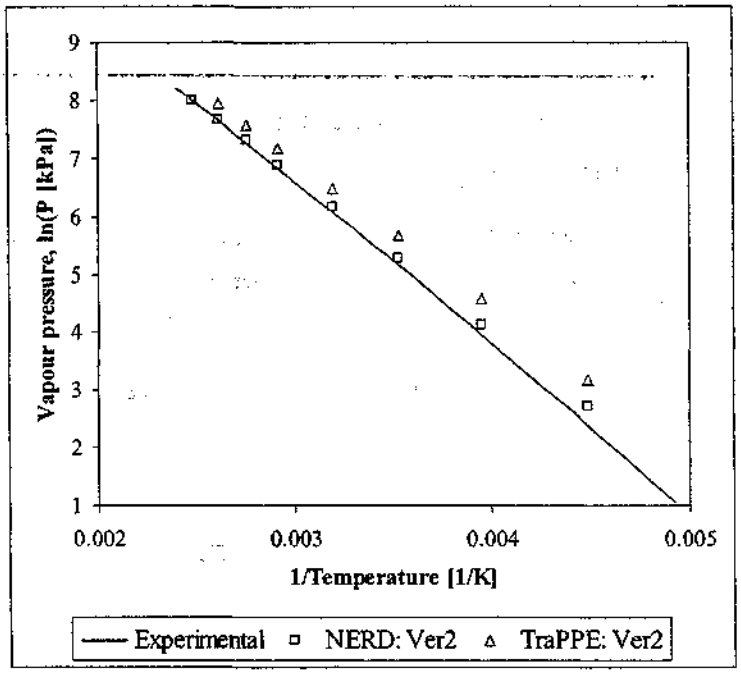


Figure 7-8: 1,3-Butadiene saturated vapour pressures from "Version 2" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

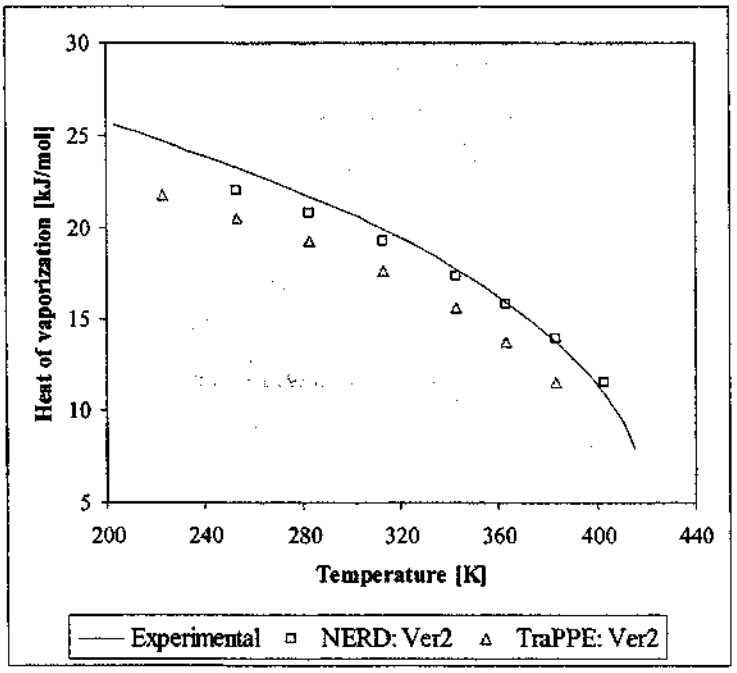


Figure 7-9: 1,3-Butadiene latent heats of vaporization from "Version 2" of the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

The NERD force field again demonstrates a better overall agreement with the experimental data than the TraPPE force field. This may be seen in Table 7-5 which shows the average differences between the experimental data of Smith and Srivastava (1986) and the simulated data from both sets of modifications to the NERD and TraPPE force fields:

Table 7-5: Average differences between the experimental saturated pure component properties of 1,3-butadiene and the simulated data from the "Version 1" and the "Version 2" modifications of the NERD and the TraPPE force fields.

	NERD			
Temperature range [K]	223-403			
Property	ρ_l	ρ_g	$\ln(P^{sat})$	ΔH_{vap}
Average difference [%], "Version 1"	3.4	25.3	5.3	6.2
Average difference [%], "Version 2"	1.3	11.2	2.8	3.3
Net % Improvement	2.1	14.1	2.5	2.9
	TraPPE			
Temperature range [K]	223-383			
Property	ρ_l	ρ_g	$\ln(P^{sat})$	ΔH_{vap}
Average difference [%], "Version 1"	3.5	82.5	12.9	15.5
Average difference [%], "Version 2"	1.7	65.6	11.1	12.6
Net % Improvement	1.8	16.9	1.8	2.9

The results in Table 7-5 also indicate that the shortening of the $CH-CH$ equilibrium bond length by only 0.08Å (a change of roughly 5%) significantly improved the pure component simulated data for both force fields. This is significant because it suggests that the internal molecular geometries must receive more serious consideration when parameterizing a pure component force field and that a wider set of intramolecular geometries could significantly improve the overall accuracy of any transferable force field. The other important feature of the improvement in physical property prediction from "Version 1" to "Version 2" in Table 7-5 is that both force fields showed an improvement of similar magnitude in each of the physical properties listed despite having different sets of intermolecular LJ parameters which produced different macroscopic properties. This suggests that the intramolecular potential parameters can affect the value of simulated physical properties largely independent of the value of the intermolecular parameters. This conclusion

would seem to contradict the observations of Smit *et al.* (1995), Chen *et al.* (1998) and Dubbeldam *et al.* (2004) who argued that the intramolecular interaction parameters have only a marginal influence on the prediction of physical properties. Recently, Ungerer *et al.* (2000) has suggested that intramolecular parameters can influence the suitability of the intermolecular parameters regressed for transferable force fields. It is the opinion of this work, however, that these two seemingly contradictory perspectives can be reconciled, particularly for harmonic potentials. Indeed, when additional simulations were conducted by varying the value of the bond stretching and bond angle bending constant, k_s and k_θ for the harmonic bond stretching and bond angle bending potentials, it was found that changes in the magnitude of up to 30% in these constants affected the final values of the simulated physical properties by less than 2% which supports the work of Smit *et al.* (1995), Chen *et al.* (1998) and Dubbeldam *et al.* (2004). When the values of the equilibrium bond angles, θ_0, and the equilibrium bond lengths, l_0, were modified, however, it was found that even small changes of 5% dramatically altered the simulated physical properties to a similar extent as those already seen listed for the CH – CH bond length which supports the work of Ungerer *et al.* (2000). Only the effect of the magnitude of the intramolecular interactions on the physical properties was actively explored in the works of Smit *et al.* (1995), Chen *et al.* (1998) and Dubbeldam *et al.* (2004). The above results therefore highlight the equilibrium geometries as the most crucial component of the intramolecular parameter sets that affect the simulated physical properties. Even though the magnitude of the intramolecular parameters are less significant than the magnitude of the interaction potential arising from intermolecular interactions, the equilibrium bond lengths and equilibrium bond angles will have a crucial role in positioning the intermolecular interaction sites in three-dimensional space. The role of the equilibrium geometries may therefore be understood to have an indirect, but still significant, role in determining macroscopic properties.

The NERD olefin force field of Nath *et al.* (2001a) utilized a –CH= equilibrium bond angle obtained from the work of Jorgensen *et al.* (1984) that was derived from microwave measurements of linear olefins. The TraPPE olefin force field of Wick *et al.* (2000), however, used equilibrium geometries obtained from the force field of Cornell *et al.* (1995) optimized for nucleic acids and proteins. The TraPPE olefin force field of Wick *et* al. (2000) does, however, provide a good description of the pure component liquid phase properties and the saturated vapour pressures. This has also been confirmed by Bourasseau *et al.* (2003). The regression of the intermolecular parameters for the TraPPE olefin force field proceeded by first regressing the sp^2 hybridized σ and ε parameters for the CH_2= group of ethylene to provide a good fit with the experimentally measured pure component data for ethylene. Thereafter, the σ and ε parameters for the sp^2

hybridized –CH= group were regressed to reproduce the pure component properties of longer 1-alkene molecules. Implicit within this "secondary regression", however, was the equilibrium bond angle of Cornell *et al.* (1995). It is, therefore, the perspective of this work that the final σ and ε parameters for the sp^2 hybridized –CH= group of the TraPPE force field compensated for the erroneous bond angle to the extent that their transferability to other functional groups was limited. This helps explain the superior results obtained using the NERD intermolecular parameters where an equilibrium bond angle closer to the experimentally measured value for linear olefins was utilized. This viewpoint is supported by the work of Chen *et al.* (1998) where it has already been reported that there can potentially exist several "optimal" intermolecular parameter sets which yield very similar macroscopic properties. An exploration of the effects of the equilibrium molecular geometry on the suitability of regressed intermolecular interaction parameters forms the basis of the work in Chapter Nine where it is considered in more detail.

What is particularly remarkable about the data from the NERD force field is that it shows good agreement with the experimental saturated vapour densities and the saturated vapour pressures. In the original publications, the simulated vapour data and heats of vaporization did not demonstrate this level of accuracy even for molecules with four carbons (Nath *et al.* (2001a)). The improvement in the prediction of the saturated vapour pressures and the vapour densities is also very important for the accurate simulation of the latent heats of vaporization as may be seen through the definition of the latent heat of vaporization in Equation 5–21. The good close agreement between the experimentally measured latent heats of vaporization as seen in Figure 7-9 and Table 7-5 therefore also indicates that the NERD force field parameters also provide a good estimate of the intermolecular potential energy, U. This further suggests that the values of the σ and ε parameters of the sp^2 hybridized –CH= and CH_2= groups of the NERD force field are close to their optimally transferable values.

7.5 Pure Component Simulations of *cis*-1,3-Pentadiene and *cis*-1,3,5-Hexatriene

The pure component simulation data for *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene are shown in Figure 7-10 to Figure 7-13 and the corresponding data is listed in Table 7-10 to Table 7-12 in Appendix 7A at the end of the Chapter.

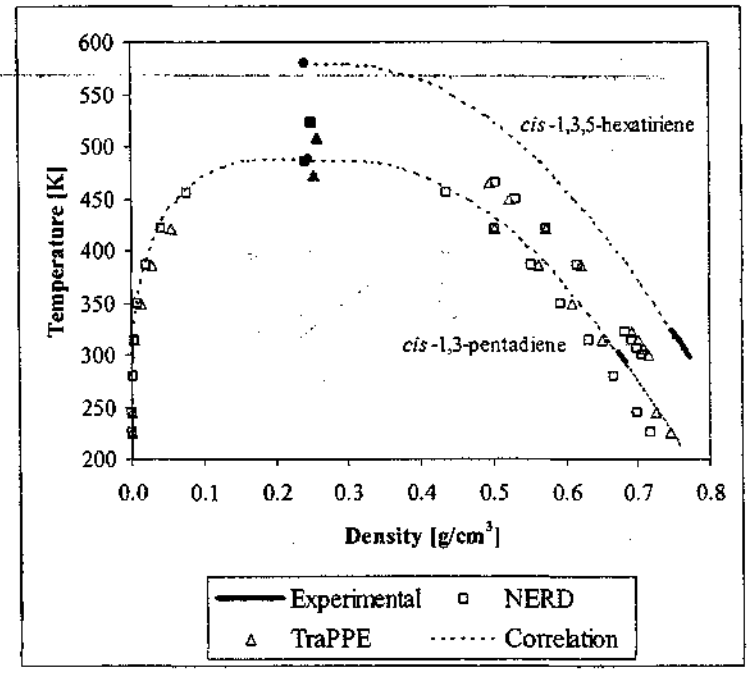


Figure 7-10: *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene liquid and vapour coexistence densities from the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986). Critical points are shown as solid symbols. Correlated data using the law of rectilinear diameters is shown as dotted lines.

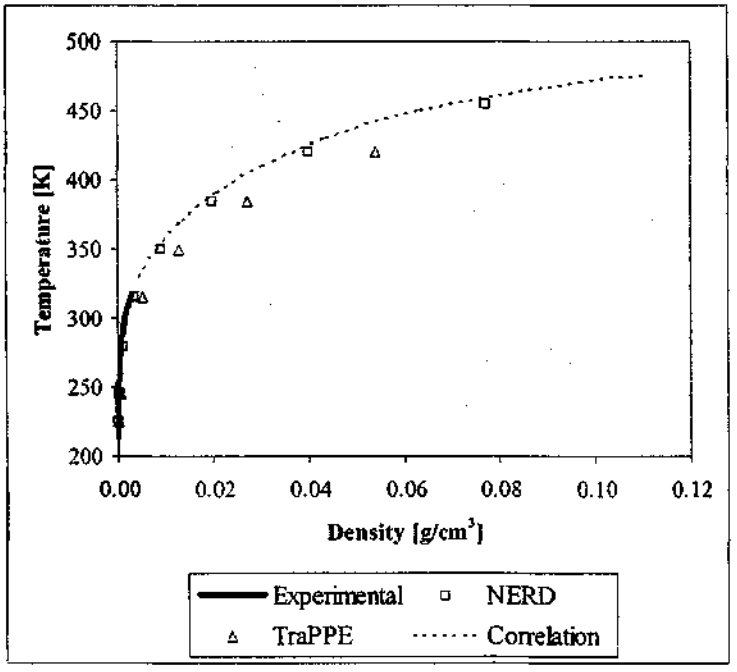


Figure 7-11: *cis*-1,3-pentadiene vapour coexistence densities for the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986). Correlated data using the law of rectilinear diameters is shown as dotted lines. Data for *cis*-1,3,5-hexatriene is omitted for clarity.

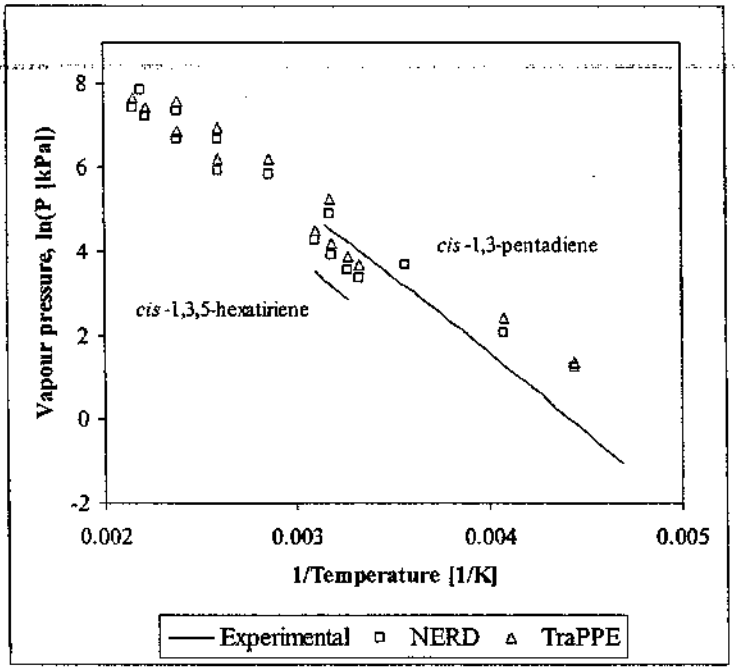


Figure 7-12: *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene saturated vapour pressures from the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986).

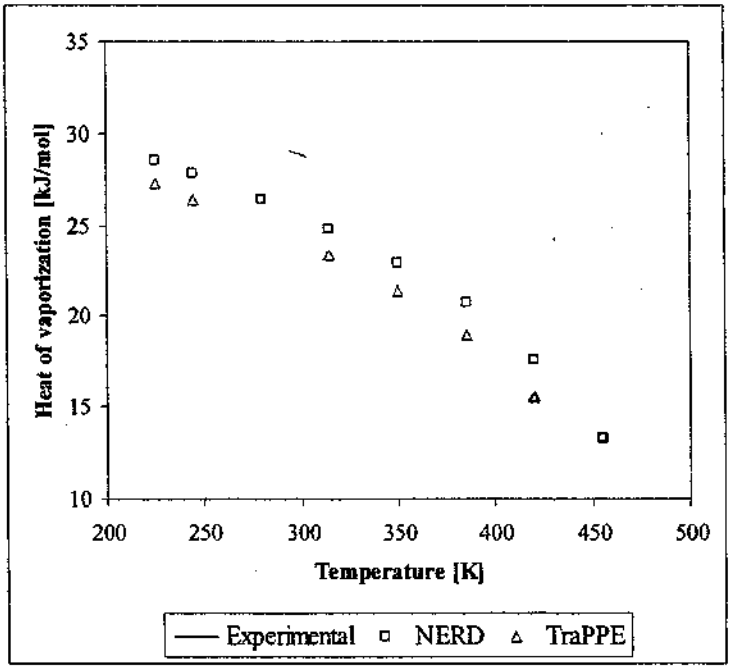


Figure 7-13: *cis*-1,3-pentadiene heats of vaporization from the NERD and TraPPE force fields and the experimental data of Smith and Srivastava (1986). Data for *cis*-1,3,5-hexatriene is omitted for clarity.

Given the superior results obtained by using the observed molecular geometries for 1,3-butadiene, it was decided to use the same approach when simulating *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene. Unfortunately, no experimental study of the equilibrium geometries for *cis*-1,3-pentadiene as shown in Table 7-1 could be found. The equilibrium lengths of the $C=C$ double bonds for both 1,3-butadiene and *cis*-1,3,5-hexatriene as listed in Table 7-4, however, all have values between 1.34 Å and 1.36Å. The internal $CH-CH$ bond length for both 1,3-butadiene and *cis*-1,3,5-hexatriene varies between 1.46Å and 1.48Å. Consequently, an equilibrium bond length of 1.34Å was used for the double bonds in *cis*-1,3,-pentadiene while an equilibrium bond length of 1.47Å was used for the internal single bond. For the $C^{(4)}-C^{(5)}$ single bond the relevant bond length parameters from Table 6-2 were used. For *cis*-1,3,5-hexatriene, the equilibrium bond lengths of Liu and Zhou (1993) as listed in Table 7-4 were used. The NERD simulations used a bond stretching constant of 96500K/Å^2. Both molecules exclusively have $=CH-$ bond angles. From Table 7-3, this equilibrium bond angle varies, in general, between 123° and 126° for both 1,3-butadiene and *cis*-1,3,5-hexatriene and therefore an equilibrium value of 124° was used for all of these bond angles together with a bond angle bending constant of 62500K/rad^2. From Table 7-2, the *cis* torsion of the TraPPE force field was used for both force fields to describe the *cis* torsion, while the torsional potential of Bock *et al.* (1979) was used for the $=CH-CH=$ torsion.

For *cis*-1,3-pentadiene, the available experimental data was limited to the critical temperature, the critical density, vapour densities and vapour pressures in the temperature range from 213-317K, and heats of vaporization and liquid densities in the temperature range from 294-303K. In order to estimate liquid densities in the temperature range between the critical temperature and the available experimental data, the modified Rackett equation of Spencer and Danner (1972) was used:

$$\frac{1}{\rho_l}=\left(\frac{R_G T_{cr}}{\rho_{cr}}\right) Z_{RA}^{\left[1+(1-T_r)^{2/7}\right]} \qquad (7\text{–}3)$$

where Z_{RA} is a modified compressibility factor. Using the known values of T_{cr} and ρ_{cr}, the available liquid densities as a function of temperature were then used to regress an optimal value of Z_{RA}. This allowed for the estimation of the liquid densities over the temperature ranges of 213-294K and 303-488K. Then using these generated liquid density values together with the available critical data and the vapour density data in the temperature range from 213-317K, it was possible to determine best-fit values for the constants **A** and **B** in Equations (6–2) and (6–3) with $\beta=0.32$. These two equations could then be used to estimate vapour densities in the temperature range 317-488K given the estimated liquid densities obtained from Equation (7–3).

The above approach could not, however, be confidently used for *cis*-1,3,5-hexatriene because of the lack of measured data for the critical temperature and the critical density. Saturated liquid density data was only available in the temperature range of 306-324K. The critical density and the critical temperature of *cis*-1,3,5-hexatriene were estimated by best-fit regression of these variables (together with Z_{RA}) using Equation (7–3) to be 0.243g/cm^3 and 579.6K. Since this critical temperature is 255K higher than the highest temperature of a measured liquid density, it was not reasonable to estimate the saturated vapour densities using Equations (6–2) and (6–3). For this reason, the simulated vapour densities of *cis*-1,3,5-hexatriene are not shown in Figure 7-10 or Figure 7-11 but are listed in Table 7-12 in Appendix 7A at the end of this Chapter.

From Figure 7-10 to Figure 7-13, it is clear that the simulated data of the NERD and TraPPE force fields show similar trends to those observed in the simulated data for 1,3-butadiene. The TraPPE force field shows a better description of the high liquid densities, the NERD force field is clearly superior in the predictions of the vapour densities over the whole range and the liquid densities in the near-critical temperature range, while the estimated critical temperature of 485K and critical density of 0.2423g/cm^3 (Table 7-11) for the NERD force field compare more favourably than the TraPPE force field with the measured critical temperature of 488K and the measured critical density of 0.2477g/cm^3. The vapour pressures are again overestimated by both force fields but the agreement with the measured data is worse than it was for 1,3-butadiene. This also leads to the poorer description of the latent heats of vaporization as seen in Figure 7-13. It is suspected that the origin of the apparently increasing deviation between the simulated results and the experimental data may lie in the description of equilibrium length of the $C^{(4)}-C^{(5)}$ single bond. Unlike for conventional *n*-alkanes, the sp^3 hybridized $-CH_3$ group now has a sp^2 hybridized =CH– group attached to it instead of the usual sp^3 hybridized $=CH_2$ group. With one less hydrogen, the sp^2 hybridized –CH= group will be able to exert a greater attractive force on the terminal $-CH_3$ group thereby reducing the equilibrium length of the $C^{(4)}-C^{(5)}$ bond. The observed difference between the measured physical property data and the simulated data therefore again suggests that the transferability of intramolecular bond lengths depends on the hybridization of the functional groups. Due to the lack of measured equilibrium geometries with which to confirm these suspicions and test modified bond lengths, however, no further simulations were conducted but the idea is explored further in Chapter Nine.

For *cis*-1,3,5-hexatriene, the measured vapour densities were very limited and measured heats of vaporization could not be found and are therefore not shown in Figure 7-11 and Figure 7-13. The

simulated liquid densities and the simulated vapour pressures are, however, in considerable disagreement with the limited experimental data. It was suspected that an incorrect model of the *cis*-torsional may have resulted in *trans*-1,3,5-hexatriene being simulated but simulations of *trans*-1,3,5-hexatriene did not significantly improve the data and the computer algorithm was confirmed to be correct. Consequently, with the equilibrium geometries of Liu and Zhou (1993) correctly used, the remaining possibility lay in the description of the $=CH-CH=$ torsional potential because the torsional potential of Bock *et al.* (1979) was used and it was assumed that the intramolecular LJ interactions would account for the local intramolecular energy maxima observed at 0° (360°). In order to confirm this, a histogram of the observed torsional angles for the $=CH-CH=$ torsion was recorded. For the above assumption to be true, then the following expression must hold true for all torsional angles:

$$\mathcal{P}_\phi = \frac{w_\phi}{\sum w_\phi} = \frac{\exp\left[-\beta U(\phi)_{\text{int}}\right]}{\sum_\phi \exp\left[-\beta U(\phi)_{\text{int}}\right]} \qquad (7\text{–}4)$$

where $\mathcal{P}_i$ is the probability of observing a configuration with a torsional angle of ϕ when the total intramolecular potential is $U(\phi)_{\text{int}}$, and w_ϕ is again a weighting function but this time it serves to describe the number of times that a particular torsional angle is observed for the histogram. For the Bock *et al.* (1979) torsional potential combined with the existing LJ parameters to have been adequate, then both the normalized histogram and the graph arising from the exponential function described in Equation (7–4) above must be equivalent over the range $0 \le \phi \le 2\pi$. The resulting curves are plotted in Figure 7-14 below. For clarity, a zoomed view is also given in Figure 7-15:

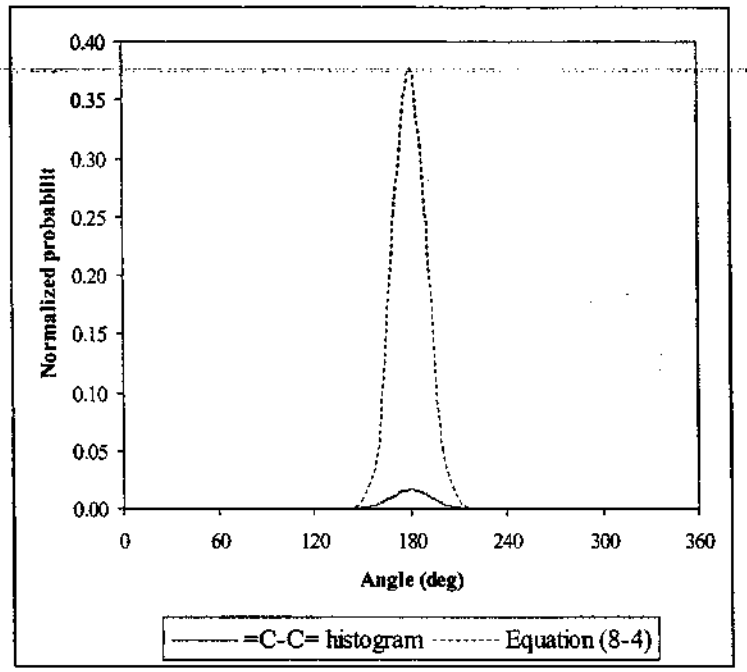


Figure 7-14: Normalized probability plot for the = CH – CH = torsion from the pure component simulation of *cis*-1,3,5-hexatriene and from Equation (7–4).

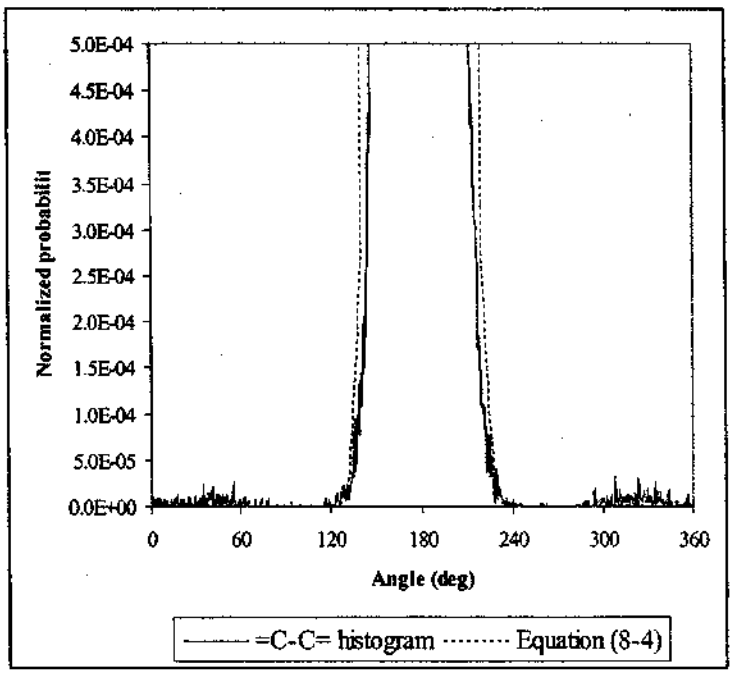


Figure 7-15: Zoomed view for the normalized probability plot for the = CH – CH = torsion from the pure component simulation of *cis*-1,3,5-hexatriene and from Equation (7–4).

For both Figure 7-14 and Figure 7-15 only the histogram resulting from the $=CH-CH=$ torsion over the first four carbons the liquid phase is plotted. It was found that the vapour phase yielded the same distribution and that both $=CH-CH=$ torsions yielded the same histogram. Both the normalized histogram from the simulation and the curve resulting from Equation (7–4) are sharply peaked around 180°. This corresponds to the *trans* configuration for the $=CH-CH=$ torsion which is the more stable configuration because then the first and fourth carbons of the torsional sequence are most widely separated. It would appear, however, that the origin for the discrepancy between the simulated results and the experimental data is because the measured data of Liu and Zhou reveals that the $=CH-CH=$ torsion should be much more sharply peaked around 180° (the *trans* configuration) than that simulated by the Bock *et al.* (1979) potential with the existing set of LJ parameters. Figure 7-15 shows a zoomed view and reveals the much higher concentration of *cis* configurations near 0° (360°) obtained in the simulations. The *cis* configuration will, in general, yield a lower density than the *trans* configuration because the more linear *trans* orientation will allow for a closer packing of molecules in the liquid phase than the more "boat-shaped" *cis* configuration.

The solution required to resolve the difference between the measured physical property data and the simulated data would therefore require a reparameterization of the sp^2 hybridized $CH_2=$ and $-CH=$ LJ parameters or the formulation of a new $=CH-CH=$ torsional potential for *cis*-1,3,5-hexatriene or a combination of both approaches. Indeed, if a combination of revised sp^2 hybridized ε and σ parameters were given different values, then the repulsion between $=CH-CH=$ from the intramolecular LJ interactions at short separations would force *cis*-1,3,5-hexatriene towards a higher fraction of *trans* configurations as required and potentially also go a long way towards improving the simulated results from the TraPPE force field. It was, however, the objective of this study not to modify the LJ parameters as this would require an extensive effort for only 3 molecules. Alternatively, the torsional potential of Bock *et al.* (1979) could be empirically modified to give a better distribution of torsional angles. It is not immediately apparent how this could be easily or reasonably achieved with the very limited temperature-dependent data available for *cis*-1,3,5-hexatriene against which to check the resulting torsional potential. Chapter Nine presents a more extensive study of a new intermolecular force field based on observations from this Chapter.

7.6 Binary Simulations for Conjugated Alkenes

To test the applicability of the revised NERD and TraPPE force fields for binary VLE simulations, a 1,3-butaidene + *n*-heptane mixture at 70°C was simulated. The 1,3-butadiene molecule was simulated using the "Version 2" modifications described in the previous Chapter while the *n*-heptane molecule was simulated using the conventional NERD and TraPPE parameters for *n*-alkanes as laid out in Table 6-1 to Table 6-4. The resulting P-x-y diagram is shown in Figure 7-16 and the corresponding x-y data is plotted in Figure 7-17. Figure 7-18 shows a plot of the ratio of the excess to the real volumes for both phases and for both the NERD and TraPPE force fields.

From Figure 7-16 the NERD force field displays a much better description of the P-x-y phase envelope than the TraPPE force field. This is, however, largely attributable to the superior pure component vapour pressure predicted by the modified NERD force field as evidenced by the pure component simulated properties and the corresponding experimental data of Smith and Srivastava (1986) in Table 7-6:

Table 7-6: Experimental and simulated pure component properties obtained using the NERD and TraPPE force fields for 1,3-butadiene and *n*-heptane. Subscripted values indicate the estimated error in the simulated results.

	1,3-butadiene			***n*-heptane**		
	Exp.	NERD, "Vers. 2"	TraPPE, "Vers. 2"	Exp.	NERD	TraPPE
P^{sat}, [kPa]	949	977_{42}	1308_{75}	40.4	$44.8_{7.6}$	$57.5_{2.1}$
ρ_l, [g/cm^3]	0.552	$0.548_{0.001}$	$0.544_{0.002}$	0.640	$0.629_{0.002}$	$0.644_{0.003}$
$\rho_g \times 10^4$, [g/cm^3]	212	$222_{8.9}$	315_{12}	14.0	$16.2_{3.2}$	$20.9_{0.74}$

Crucially, however, from Figure 7-17 it may be seen that the x-y data resulting from the simulations are in poor agreement with the experimental data because both force fields underestimate the value of the ratio of the 1,3-butadiene phase compositions, x_1/y_1. Physically, this translates to NERD and TraPPE force fields overestimating the amount of *n*-heptane in the vapour, underestimating the amount of *n*-heptane in the liquid, overestimating the amount of 1,3-butadiene in the liquid, or underestimating the amount of 1,3-butadiene in the vapour.

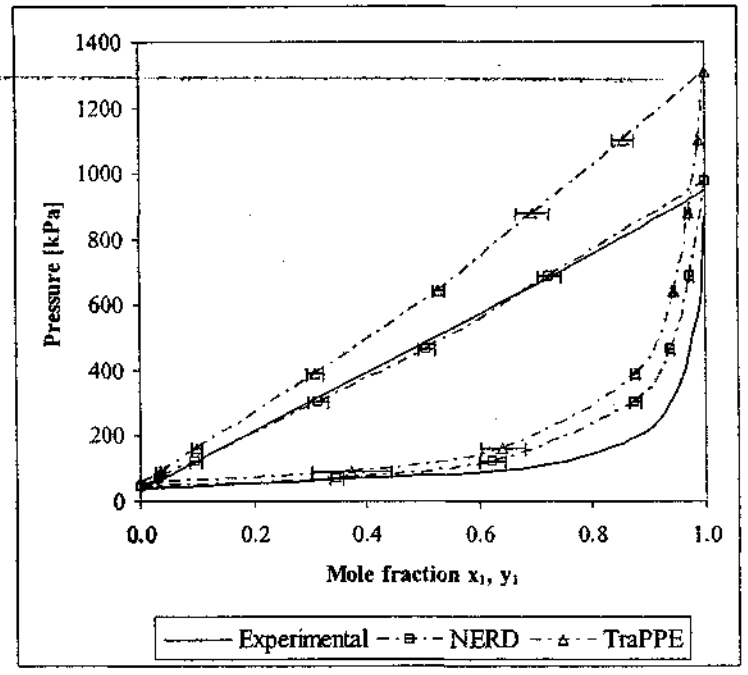


Figure 7-16: Simulated P-x-y data for the 1,3-butadiene (1) + *n*-heptane (2) mixture at 343.15K and the corresponding experimental VLE (Dortmund Data Bank). Estimated uncertainties are included.

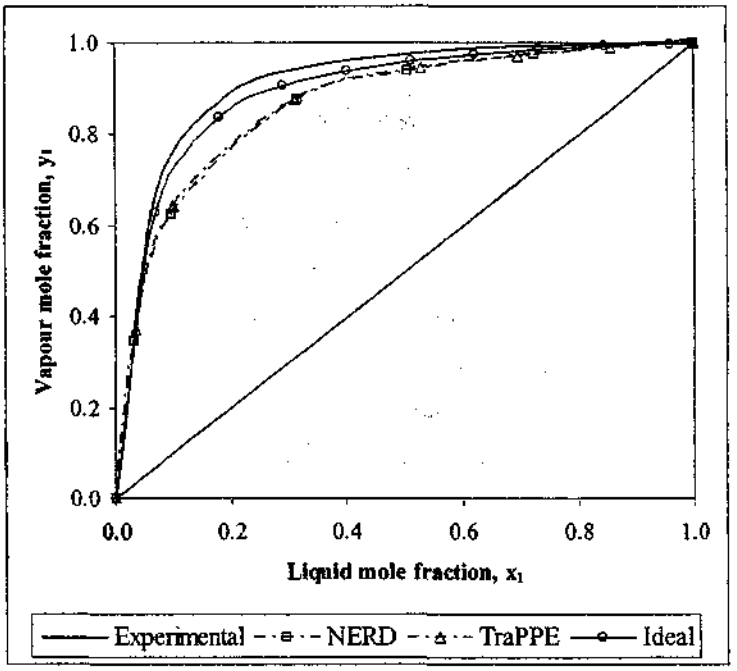


Figure 7-17: Simulated x-y data for the 1,3-butadiene (1) + *n*-heptane (2) at mixture 343.15K, the corresponding experimental x-y data (Dortmund Data Bank), and the ideal solution x-y data based on the experimental saturated vapour pressures of 1,3-butadiene and *n*-heptane. Uncertainties are omitted for clarity.

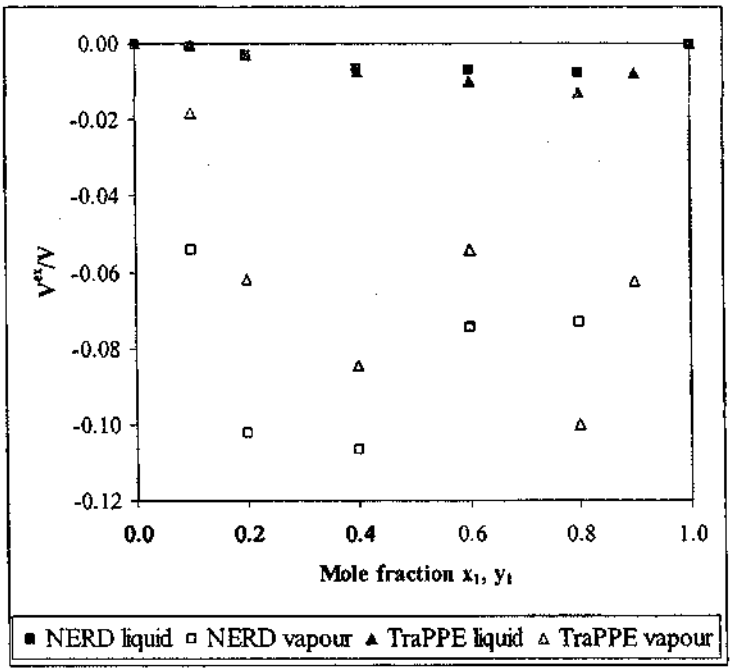


Figure 7-18: Plot of the ratio of the excess to the real volumes for both liquid and vapour phases for the 1,3-butadiene + *n*-heptane mixture at 343.15K.

From Table 7-6 it is possible to deduce that the pure component properties are not responsible for the discrepancy between the experimental and the simulated x-y curves. In terms of the saturated vapour pressure, the TraPPE force field overestimates both pure component saturated vapour pressures by approximately 50% while the NERD force field overestimates the 1,3-butadiene vapour pressure by roughly 3% and the *n*-heptane vapour pressure by about 10%. Thus, because the NERD force field describes the one vapour pressure well and the other badly, while the TraPPE force field describes both vapour pressures poorly, neither the quality of the pure component vapour pressures nor the relative inaccuracy between each of the pure component vapour pressures can be blamed for the discrepancy. Both force fields also describe the pure component liquid densities well and these can, therefore, not be the cause. The pure component vapour densities follow a similar trend to the vapour pressures and for the reason already mentioned excluding the vapour pressures as the cause, the inaccuracies in the simulated vapour densities can also not be the cause.

In Figure 7-17, the ideal mixture x-y curve based on Raoult's Law and using the experimental pure component vapour pressures is plotted. It may be seen that the force fields display a completely inaccurate trend away from the ideal curve because the ratio x_1/y_1 for the experimental data is

always higher than the ideal curve while the ratio x_1/y_1 from the force fields is always underestimated. In addition, the bubble point line connecting the pure component vapour pressures in Figure 7-16 is essentially linear for both the NERD and TraPPE force fields and for the experimental data. This confirms the ideality of the liquid phase for the experimental data and for the liquid phase simulated by both force fields (which is also supported by the V^{ex}/V ratio which displays only small deviationaround zero in Figure 7-18 and which indicates that the liquid excess volume is essentially zero for all compositions). The deviation would therefore appear to lie mostly in the vapour phase.

Recalling the x-y data of Chapter Seven, all the x-y diagrams (while still close to the experimental data) tended to slightly underestimate the x_1/y_1 ratio if they were not in perfect agreement with the experimental data. This phenomenon has essentially been enhanced for the 1,3-butadiene + *n*-heptane mixture studied. The 1,3-butadiene molecule consists exclusively of sp^2 hybridized functional groups whereas the *n*-heptane molecule consists exclusively of sp^3 hybridized functional groups. Since the discrepancy in the simulated and experimental data would seem to lie in the simulation of the vapour phase, this suggests that either the interactions between the 1,3-butadiene molecules in the liquid phase needs to be reduced by adjusting the values of LJ σ and ε parameters for the sp^2 hybridized –CH= and CH_2= groups to reduce the attraction between these groups at short separations, or by adjusting the values of the LJ σ and ε parameters for the sp^3 hybridized $-CH_2-$ and $-CH_3$ groups to increase the attraction between these groups at short separations in the liquid phase. The net effect would be an increase in the tendency of 1,3-butadiene to exist in the vapour and a decrease in the tendency of *n*-heptane to exist in the vapour.

Qualitatively, a decrease in the values of either or both of the LJ ε parameters for the sp^2 hybridized –CH= and CH_2= groups can be seen to be favourable based on the results obtained from the pure component NERD simulations of 1,3-butadiene. It is well known that the critical temperature and vapour pressures both decrease with decreasing ε. The "Version 2" NERD force field currently overestimates the critical temperature (Figure 7-6) and the saturated vapour pressures (Figure 7-8) and the reduction in the LJ ε parameters would therefore shift these properties towards the correct values. For *cis*-1,3,5-hexatriene, however, the critical temperature for *cis*-1,3,5-hexatriene is already too low and a decrease in the values of the LJ ε parameters might increase the fraction of *cis*-conformations by decreasing the repulsions at short separations for the intramolecular LJ interactions. This suggests that a simultaneous adjustment also needs be made to

increase the LJ sp^2 hybridized σ parameters because this can be used to increase repulsive interactions at short separations and hence the fraction of *trans* configurations as seen in Figure 5-2. It is also well known that the simulated vapour pressure decreases with increasing σ which will improve the situation for all conjugated alkenes studied. These and other considerations already mentioned are explored more fully in Chapter Nine.

7.7 Appendix 7A: Pure Component Simulation Data for the Three Conjugated Alkenes Studied

Table 7-7: Simulated coexistence densities for 1,3-butadiene for "Version 1" and "Version 2" of the NERD and TraPPE force fields and the corresponding experimental data. Subscripted figures signify the estimated error in the simulated value. All densities have units of g/cm^3.

Temperature	Experimental		NERD, Version 1		NERD, Version 2		TraPPE, Version 1		TraPPE, Version 2	
[K]	ρ_l	$\rho_g \times 10^4$	ρ_l	$\rho_g \times 10^4$	ρ_l	$\rho_g \times 10^4$	ρ_l	$\rho_g \times 10^4$	ρ_l	$\rho_g \times 10^4$
223	0.701	3.28	$0.676_{0.001}$	$5.38_{0.75}$	$0.688_{0.001}$	$4.47_{0.43}$	$0.690_{0.001}$	$8.18_{1.5}$	$0.698_{0.001}$	$7.14_{0.83}$
253	0.668	13.9	$0.644_{0.002}$	$19.1_{2.4}$	$0.656_{0.001}$	$16.4_{2.2}$	$0.655_{0.002}$	$27.9_{2.4}$	$0.664_{0.002}$	$26.4_{2.7}$
283	0.633	41.7	$0.610_{0.001}$	$54.0_{1.9}$	$0.624_{0.002}$	$46.9_{3.7}$	$0.618_{0.001}$	$75.5_{4.7}$	$0.628_{0.002}$	$70.8_{5.5}$
313	0.595	101	$0.574_{0.001}$	$123_{1.8}$	$0.588_{0.001}$	$110_{9.0}$	$0.578_{0.003}$	$173_{8.1}$	$0.587_{0.002}$	$156_{6.2}$
343	0.552	212	$0.533_{0.001}$	$249_{4.0}$	$0.548_{0.001}$	$222_{8.9}$	$0.531_{0.002}$	338_{14}	$0.542_{0.003}$	309_{15}
363	0.519	333	$0.500_{0.002}$	376_{10}	$0.517_{0.001}$	$335_{9.6}$	$0.494_{0.003}$	528_{11}	$0.506_{0.002}$	$476_{6.4}$
383	0.480	518	$0.464_{0.004}$	568_{24}	$0.483_{0.002}$	509_{16}	$0.448_{0.002}$	812_{29}	$0.459_{0.004}$	718_{36}
403	0.429	832	$0.418_{0.005}$	911_{51}	$0.440_{0.004}$	773_{39}				

Table 7-8: Critical temperatures and critical densities of 1,3-butadiene for "Version 1" and "Version 2" of the NERD and TraPPE force fields estimated using Equations (6–2) and (6–3) with $\beta = 0.32$.

	Experimental	NERD, Version 1	NERD, Version 2	TraPPE, Version 1	TraPPE, Version 2
T_{cr}, [K]	425.15	424.0	431.4	409.1	413.6
ρ_{cr}, [g/cm^3]	0.2450	0.2416	0.2435	0.2500	0.2497

Table 7-9: Simulated saturated vapour pressures and latent heats of vaporization of 1,3-butadiene for "Version 1" and "Version 2" of the NERD and TraPPE force fields and the corresponding experimental data. Subscripted figures signify the estimated error in the simulated value. All vapour pressures are measured in kPa and all heats of vaporization are measured in kJ/mol.

Temperature	Experimental		NERD, Version 1		NERD, Version 2		TraPPE, Version 1		TraPPE, Version 2	
[K]	$\ln(P^{sat})$	ΔH_{vap}	$\ln(P^{sat})$	ΔH_{vap}	$\ln(P^{sat})$	ΔH_{vap}	$\ln(P^{sat})$	ΔH_{vap}	$\ln(P^{sat})$	ΔH_{vap}
223	2.41	24.7	$2.90_{0.15}$	$22.7_{0.08}$	$2.72_{0.10}$		$3.33_{0.18}$	$21.4_{0.04}$	$3.18_{0.12}$	$21.7_{0.10}$
253	3.97	23.2	$4.28_{0.12}$	$21.5_{0.11}$	$4.14_{0.14}$	$22.0_{0.14}$	$4.66_{0.08}$	$20.2_{0.12}$	$4.60_{0.11}$	$20.5_{0.05}$
283	5.15	21.7	$5.40_{0.05}$	$20.3_{0.09}$	$5.26_{0.06}$	$20.8_{0.32}$	$5.72_{0.06}$	$18.8_{0.06}$	$5.67_{0.06}$	$19.3_{0.07}$
313	6.07	19.9	$6.26_{0.02}$	$18.7_{0.07}$	$6.16_{0.09}$	$19.2_{0.15}$	$6.57_{0.08}$	$17.1_{0.13}$	$6.48_{0.03}$	$17.6_{0.11}$
343	6.83	17.7	$6.98_{0.03}$	$16.8_{0.09}$	$6.88_{0.04}$	$17.4_{0.18}$	$7.23_{0.05}$	$15.0_{0.13}$	$7.17_{0.03}$	$15.6_{0.07}$
363	7.26	16.0	$7.38_{0.04}$	$15.1_{0.08}$	$7.29_{0.02}$	$15.9_{0.20}$	$7.63_{0.03}$	$13.1_{0.06}$	$7.58_{0.04}$	$13.8_{0.14}$
383	7.65	13.8	$7.75_{0.04}$	$13.1_{0.17}$	$7.66_{0.02}$	$14.0_{0.19}$	$8.00_{0.03}$	$10.8_{0.24}$	$7.93_{0.03}$	$11.5_{0.21}$
403	8.00	10.9	$8.07_{0.05}$	$10.2_{0.32}$	$8.01_{0.07}$	$11.5_{0.29}$				

Table 7-10: Simulated coexistence densities for *cis*-1,3-pentadiene and *cis*-1,3,6-hexatriene **for the modified NERD and TraPPE force fields. Subscripted figures** signify the estimated error in the simulated value. All densities have units of g/cm^3.

cis-1,3-pentadiene					*cis*-1,3,5-hexatriene				
Temperature	NERD		TraPPE		Temperature	NERD		TraPPE	
[K]	ρ_l	$\rho_g \times 10^4$	ρ_l	$\rho_g \times 10^4$	[K]	ρ_l	$\rho_g \times 10^4$	ρ_l	$\rho_g \times 10^4$
225	$0.717_{0.002}$	$1.26_{0.41}$	$0.745_{0/002}$	$1.39_{0.19}$	**300**	$0.706_{0.002}$	$9.42_{0.86}$	$0.715_{0.003}$	$13.0_{2.7}$
245	$0.699_{0.001}$	$2.66_{0.77}$	$0.724_{0/001}$	$3.69_{0.51}$	**306**	$0.699_{0.002}$	$11.5_{3.6}$	$0.710_{0.003}$	$15.3_{1.5}$
280	$0.667_{0.002}$	$11.8_{2.9}$			**314**	$0.692_{0.002}$	$15.6_{3.2}$	$0.701_{0.002}$	$21.1_{3.4}$
315	$0.632_{0.001}$	$36.1_{1.9}$	$0.652_{0/002}$	$52.7_{3.7}$	**323**	$0.684_{0.003}$	$22.1_{6.3}$	$0.692_{0.001}$	$27.6_{3.4}$
350	$0.594_{0.002}$	$90.6_{6.3}$	$0.610_{0/002}$	129_{11}	**385**	$0.618_{0.004}$	104_{12}	$0.621_{0.003}$	$137_{4.9}$
385	$0.552_{0.003}$	200_{13}	$0.562_{0/002}$	$272_{9.0}$	**420**	$0.575_{0.003}$	217_{11}	$0.574_{0.003}$	279_{13}
420	$0.504_{0.004}$	402_{27}	$0.504_{0/003}$	541_{21}	**450**	$0.532_{0.002}$	385_{10}	$0.524_{0.005}$	488_{32}
455	$0.437_{0.005}$	770_{22}			**465**	$0.506_{0.005}$	499_{48}	$0.495_{0.005}$	644_{33}

Table 7-11: Critical temperatures and critical densities of *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene for the NERD and TraPPE force fields estimated using Equations (6–2) and (6–3) with $\beta = 0.32$.

	cis-1,3-pentadiene			*cis*-1,3,5-hexatriene	
	Experimental	NERD	TraPPE	NERD	TraPPE
T_{cr}, [K]	488	485	473	523	508
ρ_{cr}, [g/cm^3]	0.2477	0.2423	0.2531	0.2517	0.2587

Table 7-12: Simulated saturated vapour pressures and latent heats of vaporization of cis-1,3-pentadiene and cis-1,3,5-hexatriene for the modified NERD and TraPPE force fields. Subscripted figures signify the estimated error in the simulated value. All vapour pressures are measured in kPa and all heats of vaporization are measured in kJ/mol.

cis-1,3-pentadiene					*cis*-1,3,5-hexatriene				
Temperature	NERD		TraPPE		Temperature	NERD		TraPPE	
[K]	$\ln\left(P^{sat}\right)$	ΔH_{vap}	$\ln\left(P^{sat}\right)$	ΔH_{vap}	[K]	$\ln\left(P^{sat}\right)$	ΔH_{vap}	$\ln\left(P^{sat}\right)$	ΔH_{vap}
225	$1.23_{0.33}$	$28.5_{0.07}$	$1.34_{0.13}$	$27.3_{0.11}$	**300**	$3.36_{0.10}$	$30.0_{0.12}$	$3.68_{0.20}$	$28.6_{0.12}$
245	$2.07_{0.29}$	$27.8_{0.18}$	$2.40_{0.14}$	$26.4_{0.11}$	**306**	$3.58_{0.30}$	$29.7_{0.15}$	$3.87_{0.09}$	$28.3_{0.24}$
280	$3.68_{0.24}$	$26.4_{0.13}$			**314**	$3.91_{0.20}$	$29.4_{0.20}$	$4.20_{0.16}$	$27.9_{0.09}$
315	$4.88_{0.06}$	$24.8_{0.22}$	$5.26_{0.06}$	$23.4_{0.14}$	**323**	$4.27_{0.27}$	$28.9_{0.10}$	$4.49_{0.12}$	$27.5_{0.08}$
350	$5.86_{0.07}$	$22.9_{0.15}$	$6.19_{0.08}$	$21.4_{0.14}$	**385**	$5.92_{0.11}$	$25.3_{0.23}$	$6.18_{0.04}$	$23.8_{0.20}$
385	$6.69_{0.08}$	$20.7_{0.14}$	$6.96_{0.06}$	$19.0_{0.20}$	**420**	$6.66_{0.06}$	$22.8_{0.28}$	$6.88_{0.04}$	$21.1_{0.22}$
420	$7.34_{0.06}$	$17.6_{0.14}$	$7.59_{0.03}$	$15.6_{0.17}$	**450**	$7.22_{0.06}$	$20.2_{0.23}$	$7.41_{0.06}$	$18.2_{0.35}$
455	$7.87_{0.05}$	$13.3_{0.33}$			**465**	$7.44_{0.06}$	$18.4_{0.30}$	$7.65_{0.06}$	$16.2_{0.24}$

CHAPTER 8:
ALTERNATIVE GIBBS ENSEMBLES FOR SIMULATING MULTICOMPONENT VLE

"Aut viam inveniam, aut faciam -- If I can't find a way then I'll make a way" – Hannibal

8.1 Practical Considerations

For conventional N_iN_2PT Gibbs Ensemble simulations to correctly simulate a two-phase mixture, it is necessary that the initial overall state of the mixture under study be specified within the two phase region of thermodynamic state space (i.e. within the phase envelope on a P-x-y or T-x-y diagram). This requires that the initial overall density and the initial overall composition of the mixture must lie in-between the equilibrium phase densities and compositions respectively. This is necessary to satisfy mass balance constraints because the total number of molecules of type 1, N_1, and the total number of molecules of type 2, N_2, cannot change for a N_iN_2PT Gibbs Ensemble simulation. This imposes a restriction in the sense that some prior knowledge of the phase envelope is required to satisfy these composition and density constraints. Consider Figure 8-1:

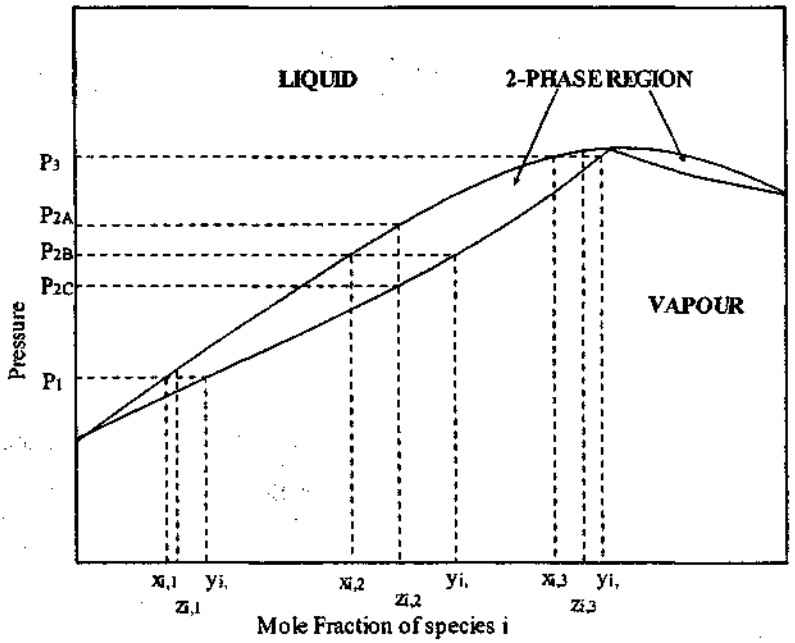


Figure 8-1: Sketch of an arbitrary azeotropic system to highlight the difficulty in simulating binary VLE for binary azeotropic mixtures or mixtures with narrow phase envelopes.

The following potential difficulties in simulating a phase envelope become evident:

- If the two-phase region is very narrow (e.g. for a mixture of close boiling components), it will be almost impossible to stipulate an initial state satisfactorily because the chance of stipulating an initial state within the two-phase region will then be very small. In Figure 8-1 this may be seen by considering the set of P_2, $x_{i,2}$, $y_{i,2}$ and $z_{i,2}$. In order to get a sensible equilibrium at point P_2, a value for the overall composition $z_{i,2}$ must be chosen in between $x_{i,2}$ and $y_{i,2}$. If a value for $z_{i,2}$ is chosen higher than $y_{i,2}$, then two vapour phases will result with approximately identical compositions of $z_{i,2}$. Contrarily, if a value for $z_{i,2}$ is chosen lower than $x_{i,2}$, then two liquid phases will result with approximately identical compositions of $z_{i,2}$. Alternatively, a pressure between P_{2A} and P_{2B} must be set for the N_iN_2PT Gibbs Ensemble simulation to achieve phase separation. Choosing the simulation pressure above P_{2A} will result in two liquid phases with composition $z_{i,2}$ while choosing P below P_{2B} will result in two vapour phases of identical composition of $z_{i,2}$. Suitable selections for z_i and P can often require several trial and error simulations for narrow phase envelopes.
- Azeotropic compositions or compositions in the infinitely dilute regions become virtually impossible to simulate because it is again very difficult to stipulate a satisfactory initial density, composition and simulation pressure. This may be seen in Figure 8-1 for the sets of P_1, $x_{i,1}$, $y_{i,1}$ and $z_{i,1}$ and P_3, $x_{i,3}$, $y_{i,3}$ and $z_{i,3}$ where the judicious choice of P and z_i becomes even more sensitive than that of P_2 and $z_{i,2}$ because of the narrowing of the two-phase region in these sections of the diagram.
- If the initial compositions are very close to one of the phase compositions, then most of the molecules in the simulation will tend to move into one of the simulation boxes. This might cause one of the simulation box sizes to drop in size such that one of the box side lengths is less than twice the cut-off radius. This will invalidate results from the simulation (see Section (5.5)). This is a consequence of the well known "lever-arm rule". From mass balance constraints, if N_i is the total number of molecules of species i, then the number of molecules of species i in the liquid for the set P_2, $x_{i,2}$, $y_{i,2}$ and $z_{i,2}$, in Figure 8-1 is given by the ratio $N_i \times (x_{i,2} - z_{i,2})/(x_{i,2} - y_{i,2})$. This implies that not only is it necessary to choose z_i between x_i and y_i and to choose P between its upper and lower limits in the two phase

region for a given z_i, but it is also necessary to choose z_i and P reasonably well inside these outer limits for the simulation not to risk falling below a minimum size restriction on the simulation box size. This can be considerably difficult for points like $z_{i,3}$ and P_3 or $z_{i,1}$ and P_1 in Figure 8-1.

- Escobedo (2000a) has pointed out that another limitation exists for the case when the incipient phase has negligible volume and the abundant phase had fixed composition such as bubble or dew point simulations. This is because it can be of considerable interest to calculate bubble or dew pressures corresponding to fixed compositions in the liquid or the vapour instead of calculating the liquid and vapour compositions arising from an imposed pressure and overall composition.

Essentially, the first three difficulties listed above concerning the application of the N_1N_2PT version of the Gibbs Ensemble to binary mixtures all originate from the requirement that one is forced to stipulate both P and z_i (where $z_i = {}^{N_i}/_{N_1+N_2}$) *a priori*. In Figure 8-1 this corresponds to specifying both the vertical (through P) and horizontal (through z_i) co-ordinates in the two-dimensional space of the phase diagram with the only freedom in this stipulation being constrained to the area between the dew and bubble point lines of the phase diagram. The smaller this area, the more unwieldy the N_1N_2PT Gibbs Ensemble becomes.

It therefore becomes prudent to consider alternative ensemble formulations or simulation approaches. Considerable attention has been paid in recent years to new ensemble formulations. In the next Section, several approaches are commented on which have been used with varying success for simulating binary mixtures before three new ensemble formulations are developed and applied.

Recalling Equation (3–24) and the discussion at the end of Section (3.3), the stipulation of any two intensive parameters should in principle be equivalently sufficient to guarantee phase separation. Thus it should not be necessary that P and T are the only intensive pairing that can be utilized for two-phase Gibbs Ensemble-type simulations with two intensive degrees of freedom. If an intensive pairing other than P and T is chosen, however, this will in general require that two different pieces of extensive variable information be specified other than N_1 and N_2 to place physical limits on the system. This alternative extensive parameter information could, for example, be a volume specification or an energy specification.

8.2 Formulation of a μ_1N_2VT Gibbs Ensemble for Binary Mixtures

It would be convenient to retain many of the advantages of the Gibbs Ensemble, including direct mechanical and chemical coupling of volume and mechanical fluctuations in the boxes, without having to resort to the approximations and interpolations inherent in pseudo-ensemble approaches. Escobedo (2000a) has alluded to the formulation of "osmotic" and "semi-open" Gibbs Ensembles. The semi-open variant in principle corresponds to the μ_1N_2VT Gibbs Ensemble described below. In this Section, the μ_1N_2VT Gibbs Ensemble partition function, the characteristic thermodynamic potential and the acceptance criteria for this ensemble using the methodology outlined in Section (4.7) are formally developed.

The proposed μ_1N_2VT Gibbs Ensemble is a two-phase and two-component ensemble at constant chemical potential of species 1, constant number of molecules of species 2, constant total volume and constant temperature. The reasoning behind the stipulation of μ_1 instead of P is that it will now permit some compositional freedom and allow one to obviate the mass balance constraints that necessarily must apply in the N_1N_2PT Gibbs Ensemble due to the fixed compositions. It should be noted that the total number of molecules N in this ensemble is not constant since the total number of molecules of species 1, n_1, fluctuates in accordance with the imposed chemical potential μ_1. The sets $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$, $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ are then:

$$\begin{aligned}
\bar{\mathbf{X}} &= \left\{n_{\mathrm{I},2}+n_{\mathrm{II},2},V\right\}=\left\{N_2,V\right\}\\
\bar{\mathbf{Y}} &= \left\{-\beta\mu_2,\beta P\right\}\\
\tilde{\mathbf{X}} &= \left\{n_{\mathrm{I},1}+n_{\mathrm{II},1},U_{\mathrm{I}},U_{\mathrm{II}}\right\}=\left\{n_1,U_{\mathrm{I}},U_{\mathrm{II}}\right\}\\
\tilde{\mathbf{Y}} &= \left\{-\beta\mu_1,\beta_1,\beta_{\mathrm{II}}=\beta_{\mathrm{I}}\right\}
\end{aligned} \tag{8–1}$$

According to Equations (4–39) and (4–43), the thermodynamic potential for this ensemble is then:

$$\begin{aligned}
\Psi^{G}_{\mu_1N_2VT} &= \beta\mu_2N_2-\beta PV\\
&= -S/k_B+\beta\left(U_{\mathrm{I}}+U_{\mathrm{II}}\right)-\beta\mu_1N_1
\end{aligned} \tag{8–2}$$

For the proposed μ_1N_2VT Gibbs Ensemble, $C=2$ and $F=2$. Substituting for $\tilde{\mathbf{X}}\cdot\tilde{\mathbf{Y}}$ from Equation (4–45), Equation (4–33) therefore becomes:

$$\begin{aligned}
\left[Q_{\mu PT}\right]^{GE}_{\mu_1N_2VT} = \sum_{n_1,n_2=0}^{\infty}\int_{V_{\mathrm{II}}}\int_{V_{\mathrm{I}}}dV_{\mathrm{II}}dV_{\mathrm{I}}\int_{s_{\mathrm{II}}^{n_{\mathrm{II}}}}\int_{s_{\mathrm{I}}^{n_{\mathrm{I}}}}ds_{\mathrm{II}}^{n_{\mathrm{II}}}ds_{\mathrm{I}}^{n_{\mathrm{I}}}\frac{V_{\mathrm{I}}^{n_{\mathrm{I}}}V_{\mathrm{II}}^{n_{\mathrm{II}}}}{K_0\Lambda_1^{3n_1}\Lambda_2^{3n_2}n_{\mathrm{I},1}!n_{\mathrm{II},1}!n_{\mathrm{I},2}!n_{\mathrm{II},2}!}\\
\times\left\{\exp\left[-\beta\left(U_{\mathrm{I}}\left(\mathbf{s}_{\mathrm{I}}^{n_{\mathrm{I}}}\right)+U_{\mathrm{II}}\left(\mathbf{s}_{\mathrm{II}}^{n_{\mathrm{II}}}\right)-\mu_1n_1\right)\right]\right\}
\end{aligned} \tag{8–3}$$

Also, one knows that $N_2 = n_{I,2} + n_{II,2} = \text{constant}$, $n_1 = n_{I,1} + n_{II,1}$ from which one can eliminate $n_{II,2}$. One also has the relation $V_I + V_{II} = V$ from which one can eliminate V_{II}. Equation (8–3) becomes:

$$\begin{aligned}\left[Q_{\mu PT}\right]^{GE}_{\mu_1 N_2 VT} &= \frac{1}{K_0 \Lambda_2^{3N_2}} \sum_{n_{I,1}=0}^{\infty} \sum_{n_{II,1}=0}^{\infty} \frac{1}{n_{I,1}!\, n_{II,1}!} \sum_{n_{I,2}=0}^{N_2} \frac{1}{\Lambda_1^{3n_1}\, n_{I,2}!\left(N_2 - n_{I,2}\right)!} \int_{V_0}^{V_B} dV_{II} \\ &\times \int_{V_I=0}^{V} dV_I V_I^{n_I} \left(V - V_I\right)^{n_{II}} \int_{s_{II}^{n_{II}}} \int_{s_I^{n_I}} ds_{II}^{n_{II}}\, ds_I^{n_I} \\ &\times \left\{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_I}\right) + U_{II}\left(\mathbf{s}_{II}^{n_{II}}\right) - \mu_1 n_1\right)\right]\right\} \\ &= \frac{1}{K_0 \Lambda_2^{3N_2}} \sum_{n_{I,1}=0}^{\infty} \sum_{n_{II,1}=0}^{\infty} \frac{1}{n_{I,1}!\, n_{II,1}!} \sum_{n_{I,2}=0}^{N_2} \frac{1}{\Lambda_1^{3n_1}\, n_{I,2}!\left(N_2 - n_{I,2}\right)!} \\ &\times \int_{V_I=0}^{V} dV_I V_I^{n_I} \left(V - V_I\right)^{n_{II}} \int_{s_{II}^{n_{II}}} \int_{s_I^{n_I}} ds_{II}^{n_{II}}\, ds_I^{n_I} \\ &\times \left\{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_I}\right) + U_{II}\left(\mathbf{s}_{II}^{n_{II}}\right) - \mu_1 n_1\right)\right]\right\}\end{aligned} \tag{8–4}$$

$V_{II,0} = \int_{V_{II}}^{V_B} dV_{II} = V_{II}$ may be included in the constant K_0 as for the NVT Gibbs Ensemble. The probability of observing a configuration around $s_I^{n_I}$ and $s_{II}^{N-n_I}$ with $n_{I,1}$ molecules in phase I with volume V_I, $n_{I,2}$ molecules of species 2 in phase I and $n_{II,2}$ molecules of species 2 in phase II is given by:

$$\begin{aligned}&\mathcal{P}\left(n_{I,1}, n_{I,2}, n_{II,2}, V_I, \mathbf{s}_I^{n_I}, \mathbf{s}_{II}^{n_{II}}\right)^{GE}_{\mu_1 N_2 VT} \\ &\propto \frac{V_I^{n_I}\left(V - V_I\right)^{n_{II}}}{\Lambda_1^{3n_1}\, n_{I,1}!\, n_{II,1}!\, n_{I,2}!\left(N_2 - n_{I,2}\right)!} \exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_I}\right) + U_{II}\left(\mathbf{s}_{II}^{n_{II}}\right) - \mu_1 n_1\right)\right]\end{aligned} \tag{8–5}$$

The probability distribution given in Equation (8–5) prescribes the same basic molecule translation, volume change and molecule swap moves given by Equations (4–5) to (4–7) for the NVT Gibbs Ensemble but with additional molecule creation and destruction moves for species 1:

1) Molecule creations of species 1 in phase I. The acceptance criteria for this move may be derived by combining Equations (3–21), (3–22) and (8–5):

$$acc(o \to n) = \frac{\left(\mathcal{P}^{GE}_{\mu_1 N_2 VT}\right)^{new}}{\left(\mathcal{P}^{GE}_{\mu_1 N_2 VT}\right)^{old}}$$

$$= \min\left\{1, \left(\frac{V_I^{n_1+1}}{\Lambda_1^{3(n_1+1)}\left(n_{1,I}+1\right)!}\frac{\Lambda_1^{3n_1} n_{1,I}!}{V_I^{n_1}} \times \frac{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_1+1}\right)-\mu_1\left(n_1+1\right)\right)\right]}{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_1}\right)-\mu_1 n_1\right)\right]}\right)\right\} \qquad (8\text{–}6)$$

$$= \min\left\{1, \left(\frac{V_I}{\Lambda_1^3\left(n_{1,I}+1\right)}\exp\left[-\beta\left(\Delta U_I-\mu_1\right)\right]\right)\right\}$$

where $\Delta U_1 = U_1\left(\mathbf{s}_1^{n_1+1}\right) - U_1\left(\mathbf{s}_1^{n_1}\right)$.

2) And molecule destructions of species 1 in phase I

$$acc(o \to n) = \frac{\left(\mathcal{P}^{GE}_{\mu_1 N_2 VT}\right)^{new}}{\left(\mathcal{P}^{GE}_{\mu_1 N_2 VT}\right)^{old}}$$

$$= \min\left\{1, \left(\frac{\Lambda_1^3 n_{1,1}}{V_I}\exp\left[-\beta\left(\Delta U_I+\mu_1\right)\right]\right)\right\} \qquad (8\text{–}7)$$

These acceptance criteria turn out to be identical to the molecule creation and destruction acceptance criteria for the μVT ensemble (Frenkel and Smit (1996)).

Unlike the N_1N_2PT Gibbs Ensemble in which P, N_1 and N_2 are specified *a priori* one now specifies μ_1, N_2 and V *a priori* in the μ_1N_2VT Gibbs Ensemble . The average total composition, z_1, is thus variable. The specification of μ_1 effectively determines the vertical co-ordinate on the phase diagram thus replacing the specification of P in the N_1N_2PT Gibbs Ensemble because each μ_1 corresponds to one unique pair of x_1 and y_1 on the phase diagram at equilibrium. The constant total volume V replaces the specification of the extensive variable N_1 and thus determines the horizontal co-ordinate on the phase diagram. As V increases, the average number of molecules in the liquid phase will increase. In terms of the "lever-arm" rule, this implies that as the value of V increases, the ensemble average overall composition approaches the ensemble average liquid composition.

8.3 Formulation of a ξ_1NVT Gibbs Ensemble for Binary Mixtures

It should be possible to construct a Semi-Grand variation of the basic Gibbs Ensemble using the ideas contained in Sections (8.2), (4.3) and (4.7). The advantage of this proposed ensemble would be that it would be able to maintain a constant total number of molecules while varying the number of each type of molecule. In the μ_1N_2VT Gibbs Ensemble the total number of molecules can vary which might result in a considerable increase in the length of a simulation if the number of molecules of species 1 increases substantially from the starting value.

The reasoning behind the formulation of the ξ_1NVT Gibbs Ensemble is that it should be possible to specify one intensive variable and one relationship among two or more intensive variables and still satisfy the requirement that two intensive degrees of freedom be specified for a two-component and two-phase system. As for the μ_1N_2VT Gibbs Ensemble of Section (8.2), it is most convenient to specify the temperature *a priori* which thus becomes the directly specified intensive variable. One may then choose a chemical potential difference to specify the remaining intensive degree of freedom. This effectively amounts to deriving a Gibbs Ensemble from the Semi-Grand Ensemble analogous to the way in which a constant chemical potential Gibbs Ensemble was effectively derived from the μVT ensemble in Section (8.2).

In the proposed ξ_1NVT Gibbs Ensemble, the total number of molecules of species 1 and species 2, n_1 and n_2 respectively, are both variable but subject to the constraint that $N = n_1 + n_2$. It is therefore convenient to write the contribution of $-\beta\mu_1 n_1$ and $-\beta\mu_2 n_2$ to Equation (4–25) for a binary system as:

$$\begin{aligned} -\beta\mu_1 n_1 - \beta\mu_2 n_2 &= -\beta\mu_1 n_1 - \beta\mu_2 (N - n_1) \\ &= -\beta\mu_2 N - \beta(\mu_1 - \mu_2) n_1 \end{aligned} \tag{8–8}$$

The total volume V is also held constant. The sets $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$, $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ can then become:

$$\begin{aligned} \bar{\mathbf{X}} &= \{N, V\} \\ \bar{\mathbf{Y}} &= \{-\beta\mu_2, \beta P\} \\ \tilde{\mathbf{X}} &= \{n_{\mathrm{I},1} + n_{\mathrm{II},1}, U_{\mathrm{I}}, U_{\mathrm{II}}\} = \{n_1, U_{\mathrm{I}}, U_{\mathrm{II}}\} \\ \tilde{\mathbf{Y}} &= \{-\beta(\mu_1 - \mu_2), \beta_{\mathrm{I}}, \beta_{\mathrm{II}} = \beta_{\mathrm{I}}\} \end{aligned} \tag{8–9}$$

According to Equations (4–39) and (4–43), the thermodynamic potential for this ensemble is then:

$$\begin{aligned} \Psi^{GE}_{\xi_1 NVT} &= \beta\mu_2 N - \beta PV \\ &= -S/k_B + \beta(U_{\mathrm{I}} + U_{\mathrm{II}}) - \beta(\mu_1 - \mu_2) n_1 \end{aligned} \tag{8–10}$$

It is interesting to note how similar the thermodynamic potential and sets of $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$, $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ are for both the ξ_1NVT and μ_1N_2VT Gibbs Ensembles. This is to be expected since both ensembles are similar in the sense that they involve molecule destructions and creations but differ in what type of molecule number specification remains constant, namely N for the ξ_1NVT Gibbs Ensemble and N_2 for the μ_1N_2VT Gibbs Ensemble.

For the proposed ξ_1NVT Gibbs Ensemble, $C=2$ and $F=2$. Substituting for $\tilde{\mathbf{X}}\cdot\tilde{\mathbf{Y}}$ from Equation (4–45), Equation (4–33) therefore becomes:

$$\begin{aligned}\left[Q_{\mu PT}\right]^{GE}_{\xi_1 NVT} = & \sum_{n_1,n_2,n_{1,1},n_{1,2}} \int_{V_{II}}\int_{V_I} dV_{II}\,dV_I \int_{s_{II}^{n_{II}}}\int_{s_I^{n_I}} ds_{II}^{n_{II}}\,ds_I^{n_I} \\ & \times \frac{V_I^{n_I} V_{II}^{n_{II}}}{K_0 \Lambda_1^{3n_1}\Lambda_2^{3n_2} n_{1,1}!\,n_{II,1}!\,n_{1,2}!\,n_{II,2}!} \\ & \times\left\{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_I}\right)+U_{II}\left(\mathbf{s}_{II}^{n_{II}}\right)-\left(\mu_1-\mu_2\right)n_1\right)\right]\right\}\end{aligned} \tag{8–11}$$

Also, one knows that $N=n_1+n_2=n_I+n_{II}$, $n_1=n_{I,1}+n_{II,1}$ and $n_2=n_{1,2}+n_{II,2}$ as mass balance constraints. One also has again $V_I+V_{II}=V$ from which one can eliminate V_{II}. The term $\dfrac{1}{n_{I,1}!\,n_{II,1}!\,n_{1,2}!\,n_{II,2}!}$ which accounts for indistinguishability (i.e. double counting) must also be retained and the molecules are free to swap identities from species 1 to species 2 or vice versa in both simulation boxes. For a binary mixture, Equation (8–11) becomes:

$$\begin{aligned}\left[Q_{\mu PT}\right]^{GE}_{\xi_1 NVT} = & \frac{1}{K_0\Lambda_2^{3N}} \sum_{n_1=0}^{N}\sum_{n_{1,1}=0}^{n_1}\sum_{n_{II,1}=0}^{N-n_1}\left(\frac{\Lambda_2}{\Lambda_1}\right)^{3n_1} \frac{\int_{V_0}^{V_{II}} dV_{II}\int_{V_I=0}^{V} dV_I V_I^{n_1}\left(V-V_I\right)^{N-n_1}}{n_{1,1}!\left(n_1-n_{1,1}\right)!\,n_{II,1}!\left(N-n_1-n_{II,1}\right)!} \\ & \times \int_{s_0^{N-n_1}}\int_{s_1^{n_1}} ds_{II}^{N-n_1}\,ds_I^{n_1}\left\{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_1}\right)+U_{II}\left(\mathbf{s}_{II}^{N-n_1}\right)-\left(\mu_1-\mu_2\right)n_1\right)\right]\right\} \\ = & \frac{1}{K_0\Lambda_2^{3N}} \sum_{n_1=0}^{N}\sum_{n_{1,1}=0}^{n_1}\sum_{n_{II,1}=0}^{N-n_1}\left(\frac{\Lambda_2}{\Lambda_1}\right)^{3n_1} \frac{\int_{V_I=0}^{V} dV_I V_I^{n_1}\left(V-V_I\right)^{N-n_1}}{n_{1,1}!\left(n_1-n_{1,1}\right)!\,n_{II,1}!\left(N-n_1-n_{II,1}\right)!} \\ & \times \int_{s_0^{N-n_1}}\int_{s_1^{n_1}} ds_{II}^{N-n_1}\,ds_I^{n_1}\left\{\exp\left[-\beta\left(U_I\left(\mathbf{s}_I^{n_1}\right)+U_{II}\left(\mathbf{s}_{II}^{N-n_1}\right)-\left(\mu_1-\mu_2\right)n_1\right)\right]\right\}\end{aligned} \tag{8–12}$$

Recalling Equations (4–13) and (4–14), and introducing the fugacity fraction as discussed in Equations (4–11) and (4–12), Equation (8–12) for a binary mixture re-arranges to become:

$$\left[Q_{\mu PT}\right]^{GE}_{\xi_1 NVT} = \frac{1}{K_0 \Lambda_2^{3N}} \sum_{n_1=0}^{N} \sum_{n_{1,1}=0}^{n_1} \sum_{n_{II,1}=0}^{N-n_1} \frac{\left(\frac{\xi_1}{1-\xi_1}\right)^{n_1} \int_{V_1=0}^{V} dV_1 V_1^{n_1} \left(V-V_1\right)^{N-n_1}}{n_{1,1}!\left(n_1-n_{1,1}\right)! n_{II,1}!\left(N-n_1-n_{II,1}\right)!}$$
$$\times \int_{s_{II}^{N-n_1}} \int_{s_1^{n_1}} ds_{II}^{N-n_1} ds_1^{n_1} \left\{\exp\left[-\beta\left(U_I\left(s_1^{n_1}\right)+U_{II}\left(s_{II}^{N-n_1}\right)\right)\right]\right\} \tag{8–13}$$
$$= \frac{1}{K_0 \Lambda_2^{3N}} \sum_{n_1=0}^{N} \sum_{n_{1,1}=0}^{n_1} \sum_{n_{II,1}=0}^{N-n_1} \frac{\left(\frac{\xi_1}{1-\xi_1}\right)^{n_1} \int_{V_1=0}^{V} dV_1 V_1^{n_1} \left(V-V_1\right)^{N-n_1}}{n_{1,1}!\left(n_1-n_{1,1}\right)! n_{II,1}!\left(N-n_1-n_{II,1}\right)!}$$
$$\times \int_{s_{II}^{N-n_1}} \int_{s_1^{n_1}} ds_{II}^{N-n_1} ds_1^{n_1} \left\{\exp\left[-\beta\left(U_I\left(s_1^{n_1}\right)+U_{II}\left(s_{II}^{N-n_1}\right)\right)\right]\right\}$$

where $\xi_2 = 1-\xi_1$ for a binary mixture. As pointed out by Frenkel and Smit (1996), the convenience behind introducing ξ_1 and ξ_2 for the Semi-Grand Ensemble arises because while $\ln\left[f_1/f_2\right]$ varies between $-\infty$ and ∞ as one goes from pure species 2 to pure species 1, ξ_2 varies between 0 and 1 in the same composition interval. From Equation (8–13), it is convenient to think of this new variant of the Gibbs Ensemble as a "two-phase Semi-Grand Ensemble" by comparing it with Equation (4–11) discussed previously. The probability of observing a configuration around $s_1^{n_1}$ and $s_{II}^{N-n_1}$ with n_1 total molecules in phase I with volume V_1, $n_{1,1}$ molecules in phase I, and $n_{II,2}$ molecules of species 2 in phase II with volume $\left(V-V_1\right)$ is given by:

$$\mathcal{P}\left(n_1, n_{1,1}, n_{II,1}, V_1, s_1^{n_1}, s_{II}^{N-n_1}\right)^{GE}_{\xi_1 NVT}$$
$$\propto \frac{V_1^{n_1}\left(V-V_1\right)^{N-n_1}\left(\frac{\xi_1}{1-\xi_1}\right)^{n_1}}{n_{1,1}!\left(n_1-n_{1,1}\right)! n_{II,1}!\left(N-n_1-n_{II,1}\right)!} \exp\left[-\beta\left(U_I\left(s_1^{n_1}\right)+U_{II}\left(s_{II}^{N-n_1}\right)\right)\right] \tag{8–14}$$

The probability distribution given in Equation (8–14) prescribes the same basic molecule translation, volume change and molecule swap moves given by Equations (4–5) to (4–7) for the NVT Gibbs Ensemble but with additional molecule identity changes given by:

1) Converting a molecule of type 2 into a molecule of type 1 in phase I. The acceptance criterion for this move is derived by combining Equations (3–21), (3–22) and (8–14):

$$acc(o \to n) = \frac{\left(\mathcal{P}^{GE}_{\xi_1 NVT}\right)^{new}}{\left(\mathcal{P}^{GE}_{\xi_1 NVT}\right)^{old}}$$

$$= \min\left\{1, \left(\frac{n_{1,1}!(n_1 - n_{1,1})!}{(n_{1,1}+1)!(n_1 - n_{1,1}-1)!}\left(\frac{1-\xi_1}{\xi_1}\right)^{n_1} \times \left(\frac{\xi_1}{1-\xi_1}\right)^{n_{1,1}+1} \frac{\exp\left[-\beta U_I\left(\mathbf{s}_1^{n_1, n_{1,1}+1}\right)\right]}{\exp\left[-\beta U_I\left(\mathbf{s}_1^{n_1, n_{1,1}}\right)\right]}\right)\right\} \qquad (8\text{–}15)$$

$$= \min\left\{1, \left(\frac{(n_1 - n_{1,1})}{(n_{1,1}+1)} \frac{\xi_1}{1-\xi_1} \exp\left[-\beta \Delta U_I\left(\mathbf{s}_1^{n_1, n_{1,1}+1}\right)\right]\right)\right\}$$

where $\Delta U_I\left(\mathbf{s}_1^{n_1, n_{1,1}+1}\right) = U_I\left(\mathbf{s}_1^{n_1, n_{1,1}+1}\right) - U_I\left(\mathbf{s}_1^{n_1, n_{1,1}}\right)$.

2) And converting a molecule of type 1 into a molecule of type 2 in phase I

$$acc(o \to n) = \frac{\left(\mathcal{P}^{G}_{\xi_1 NVT}\right)^{new}}{\left(\mathcal{P}^{G}_{\xi_1 NVT}\right)^{old}}$$

$$= \min\left\{1, \left(\frac{n_{1,1}}{(n_1 - n_{1,1}+1)} \frac{1-\xi_1}{\xi_1} \exp\left[-\beta \Delta U_I\left(\mathbf{s}_1^{n_1, n_{1,1}-1}\right)\right]\right)\right\} \qquad (8\text{–}16)$$

These acceptance criteria turn out to be similar to the molecule identity change steps for the monophasic Semi-Grand Ensemble (Frenkel and Smit (1996), Kofke and Glandt (1998)).

Unlike the N_1N_2PT Gibbs Ensemble in which P, N_1 and N_2 are specified *a priori*, one now specifies ξ_1, N and V *a priori*. The average total composition, z_1, is thus again variable. The specification of ξ_1 determines the vertical co-ordinate on the phase diagram thus replacing the specification of P in the N_1N_2PT Gibbs Ensemble because each ξ_1 corresponds to a single pair of x_1 and y_1 on the phase diagram. The ratio N/V replaces the specification of the extensive variables N_1 and N_2, and thus determines the horizontal co-ordinate on the phase diagram. The larger the value of N/V, the more molecules of the fixed total number molecules N will be found in the liquid phase. In terms of the "lever-arm" rule, this implies that as the value of N/V increases, the ensemble average overall composition approaches the ensemble average liquid composition.

The ξ_1NVT Gibbs Ensemble can be thought of as a biphasic Semi-Grand Ensemble. The linking of two phases by volume changes using the acceptance criterion given in Equation (4–6) in which the

total volume is conserved ensures that the specification of the intensive parameters ξ_1 and T fully parameterises an equilibrium ensemble without the need to specify the pressure.

Both the μ_1N_2VT and the ξ_1NVT Gibbs Ensembles obviate the difficulty associated with specifying N_1 and N_2 *a priori* by specifying μ_1 and N_2 in the μ_1N_2VT Gibbs Ensemble and by specifying ξ_1 and N in the ξ_1NVT Gibbs Ensemble. The imposed values of V and N_2 in the μ_1N_2VT Gibbs Ensemble and the imposed values of ξ_1 and N in the ξ_1NVT Gibbs Ensemble effectively determine the relative "size" of the simulated phases (i.e., how many molecules are in each of the phases at equilibrium) and hence the average value of z_1 in Figure 8-1.

8.4 Chemical Potential in the μ_1N_2VT and ξ_1NVT Gibbs Ensembles

In a manner analogous to the proofs outlined in Smit and Frenkel (1989) and Mooij and Frenkel (1994), it is possible to show that the expression give for the chemical potential in the NVT Gibbs Ensemble given in Equation (5–37) is applicable to the evaluation of the chemical potential in both the μ_1N_2VT and ξ_1NVT Gibbs Ensembles. This is the case because the molecule swap moves are the same for all types of Gibbs Ensembles and because the total volume is conserved in all three cases of the NVT, μ_1N_2VT and ξ_1NVT Gibbs Ensembles.

8.5 Simulations Using the μ_1N_2VT and ξ_1NVT Gibbs Ensembles

To assess the potential of the μ_1N_2VT and ξ_1NVT Gibbs Ensembles to simulate the VLE of mixtures with narrow phase envelopes that cannot reasonably be simulated using the N_1N_2VT Gibbs Ensemble, two close-boiling binary mixtures were simulated namely:

- 1-butene + *n*-butane mixture at 37.8°C (Laurance and Swift (1974))
- 1-hexene + *n*-hexane mixture at 55°C (Dortmund Data Bank)

The TraPPE force field parameters of Table 6-1, Table 6-2, Table 6-3 and Table 6-4 were used to simulate the two isotherms studied.

8.5.1 Simulation Methodology and Details

The CBMC method was used in conjunction with both the μ_1N_2VT and ξ_1NVT Gibbs Ensembles to simulate each of the isotherms studied. The N_1N_2PT Gibbs Ensemble could not be

used to study either of the isotherms for comparison because the small size of the phase envelopes made it impossible to simulate a phase coexistence point even with trial and error simulations. The NVT Gibbs Ensemble was used to simulate the pure component compositions for both of the isotherms studied. Values for μ_1 and ξ_1 were selected so as to yield approximate overall mole fractions of 0.05, 0.10, 0.20, 0.40, 0.60, 0.80, 0.90, and 0.95 of component 1 at equilibrium. For the ξ_1NVT Gibbs Ensemble simulations, the total numbers of molecules used for each of the phase diagrams were:

- 300 total molecules for the 1-butene (1) + *n*-butane (2) mixture
- 250 total molecules for the 1-hexene (1) + *n*-hexane (2) mixture

For the μ_1N_2VT Gibbs Ensemble simulations, it was obviously not possible to constrain the total number of molecules because N_1 is allowed to fluctuate. For comparison purposes, however, the values of μ_1 and V for a μ_1N_2VT Gibbs Ensemble simulation were adjusted so that the ensemble averaged total number of molecules at the end of a μ_1N_2VT Gibbs Ensemble simulation was similar to the total number of molecules used for the ξ_1NVT Gibbs Ensemble simulation for similar z_1 at equilibrium. The simulation of an isotherm using either the μ_1N_2VT or the ξ_1NVT Gibbs Ensembles for the mixture compositions therefore proceeded in the following general manner,

- Both the pure component compositions were simulated using the NVT Gibbs ensembles.
- The ensemble averaged chemical potentials for both pure components at the temperature of the isotherm were recorded.
- Several simulations of no more than 4,000 equilibration cycles and 20,000 production cycles were then used to generate intermediate values for μ_1 and ξ_1 such that approximately 20% of the molecules were located in the vapour phase at equilibrium. This was made possible by the low calculated statistical uncertainties associated with the calculated chemical potentials (typically less than 1%). The densities of each of the components in each of the phases were also recorded.
- Plots of μ_1 as a function of the overall mole fraction of component 1, z_1, and ξ_1 as a function of z_1 were then generated as shown in Figure 8-2 and Figure 8-3 for the 1-butene + *n*-butane mixture.
- From plots similar to those in Figure 8-2 used for the μ_1N_2VT Gibbs Ensemble and Figure 8-3 used for the ξ_1NVT Gibbs Ensemble, and using the approximate equilibrium component densities also recorded, it was possible to select values of μ_1 and ξ_1 and total volumes V

corresponding to the 8 required intermediate compositions such that approximately 20% of the molecules existed in the vapour phase at equilibrium.

It must be pointed out, however, that this preliminary simulation method is, in fact, unnecessary and was only undertaken in this instance to better facilitate a direct comparison of the relative performances of the μ_1N_2VT and ξ_1NVT Gibbs Ensembles. Even so, the method of simulation described above marked a considerable improvement over the essentially trial and error simulation approach that would had to have been used for conventional N_1N_2PT Gibbs Ensemble simulations similar to that described for the binary mixture simulations of Chapter Six.

The preliminary simulations revealed that the μ_1N_2VT and ξ_1NVT Gibbs Ensembles are considerably less sensitive to the values selected for the initial volumes and for the initial compositions than the N_1N_2PT Gibbs Ensemble. In general, it was found that an initial overall mole fraction of the dilute component of anything between 0.0 to 0.5 mole fraction could still yield a sensible simulation with phase separation. There was also considerable freedom with respect to values for the initial volumes. The total initial volumes were increased by up to as much as 100% in excess of the total volume required for approximately 20% of the molecules to exist in the vapour, and the simulations still converged to the same equilibrium phase densities (although with higher fractions of the total number of molecules in the vapour).

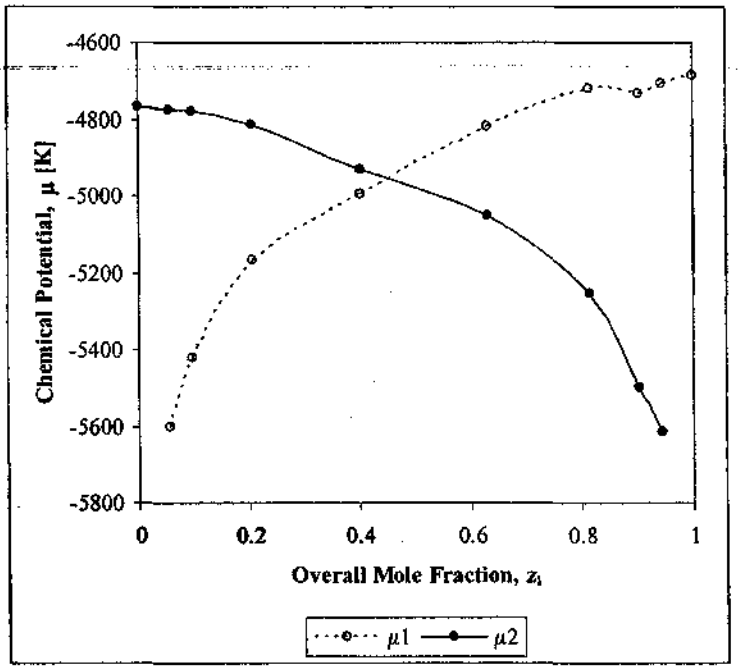


Figure 8-2: Plot of the simulated component chemical potentials as a function of the simulated overall mole fraction of 1-butene for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C using the μ_1N_2VT Gibbs Ensemble.

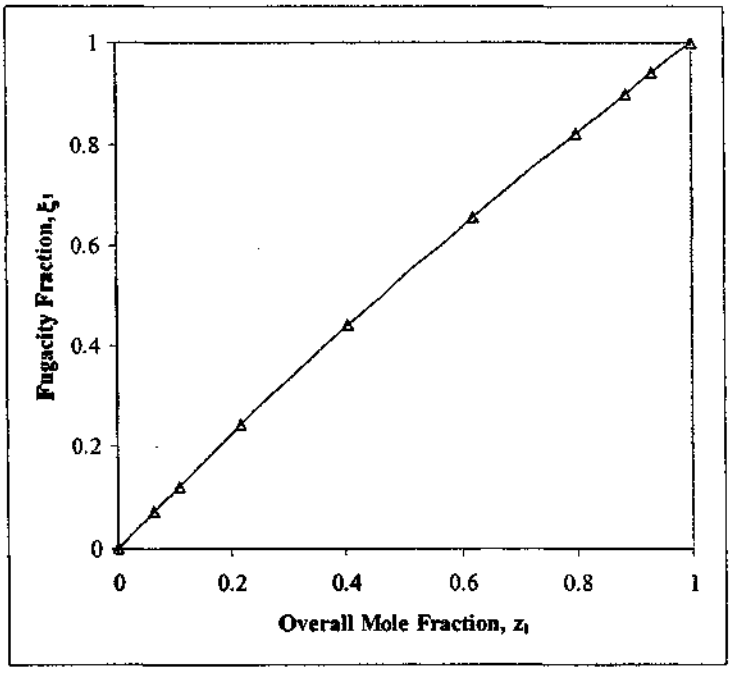


Figure 8-3: Plot of the simulated overall mole fraction as a function of the imposed fugacity fraction of 1-butene for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C using the ξ_1NVT Gibbs Ensemble.

As a general guideline, it will usually be better to stipulate the chemical potential of the component which is more easily transferred between simulation boxes for a μ_1N_2VT Gibbs Ensemble simulation because this will lead to a higher sampling efficiency. This will usually tend to be the smaller molecule if there is a size difference between the two components. For the mixtures studied, however, both components in both close-boiling mixtures have approximately equivalent molecular sizes and approximately equivalent intra- and intermolecular interactions. For these mixtures, it was found that it proved to be more efficient to stipulate the chemical potential for the more dilute component, i.e. to stipulate μ_1 for $z_1 \leq 0.5$ and then to stipulate μ_2 for $z_1 \geq 0.5$. This is because the addition of an extra molecule for the dilute component results in a larger deviation and hence better compositional sampling per insertion attempt, particularly for the infinitely dilute regions. For the ξ_1NVT Gibbs Ensemble, this is not applicable because two different molecules have to change identity in the same simulation box. The only requirement for binary phase envelope simulations using the ξ_1NVT Gibbs Ensemble is then that the imposed value of the fugacity fraction must span a range of values for $0 \leq \xi_1 \leq 1$ (or equivalently $0 \leq \xi_2 \leq 1$). Consequently, it is reasonable to expect that the μ_1N_2VT Gibbs Ensemble is likely to be more advantageous than the ξ_1NVT Gibbs Ensemble for the case where one of the molecules is not easily created or destroyed. This must be balanced against the feature that the ξ_1NVT Gibbs Ensemble can directly control the total number of molecules (and hence the simulation time) where this can only be indirectly controlled in the μ_1N_2VT Gibbs Ensemble.

The different creation and destruction moves employed by the μ_1N_2VT and ξ_1NVT Gibbs Ensembles necessitated different relative probabilities of the different types of moves employed during a simulation. The types of moves performed during a μ_1N_2VT Gibbs Ensemble simulation were as follows:

1) volume changes(Equation (4–6)),
2) conservative molecule identity exchanges using CBMC (Equation (4–18)),
3) transfers of molecules between simulation boxes using CBMC (Equation (4–7)),
4) creations and destructions of the (dilute) species with the imposed chemical potential using CBMC (Equation (8–6) and Equation (8–7)),
5) full regrowing of chains at a new position in the same simulation box using CBMC (Equation (5–33)),
6) partial regrowing of a chain in the same simulation box using CBMC (Equation (5–33))
7) translation of the COM (Equation (4–5)), and
8) rotation around the COM (Equation (4–5)).

The type of move performed at each Monte Carlo step was selected at random. The moves were performed with the following fixed probabilities:

p1: p2: p3: p4: p5: p6: p7: p8 = 0.010: 0.050: 0.200: 0.300: 0.047: 0.100: 0.147: 0.146,

where p1 again refers to the probability of performing a move of type 1 listed above. To improve the sampling efficiency of each simulation, the molecule transfer move between boxes was only performed for the species whose chemical potential was not imposed. This is because the molecule with the imposed chemical potential could still sample different compositions through the creation and destruction moves of Equation (8–6) and Equation (8–7). Also, for the molecule with the imposed chemical potential, the choice between a creation or destruction attempt was made randomly as was the box identity in which a molecule was created or from which a molecule was deleted. A rough check on the consistency of a simulation was provided by comparing the number of successful and attempted creations and destructions of the species with the imposed chemical potential in either simulation box. At equilibrium, the number of creations and destructions in a given simulation box had to be more or less equal.

The types of moves performed during a ξ_1NVT Gibbs Ensemble simulation were as follows:

1) volume changes(Equation (4–6)),
2) conservative molecule identity exchanges using CBMC (Equation (4–18)),
3) Semi-Grand (non-conservative) molecule identity changes of a molecule using CBMC (Equation (8–15) and Equation (8–16))
4) transfers of molecules between simulation boxes using CBMC (Equation (4–7)),
5) full regrowing of chains at a new position in the same simulation box using CBMC (Equation (5–33)),
6) partial regrowing of a chain in the same simulation box using CBMC (Equation (5–33))
7) translation of the COM (Equation (4–5)), and
8) rotation around the COM (Equation (4–5)).

The type of move performed at each Monte Carlo step was selected at random. The moves were performed with the following fixed probabilities:

p1: p2: p3: p4: p5: p6: p7: p8 = 0.010: 0.050: 0.300: 0.200: 0.047: 0.100: 0.147: 0.146,

where p1 again refers to the probability of performing a move of type 1 listed above. For the molecule with the imposed fugacity, the choice between a Semi-Grand creation or destruction attempt from the other species was made randomly as was the box identity in which the molecule type was generated or removed. In this case too, a rough check on the consistency of a simulation was also possible. The number of creations of species i in simulation box I had to be

approximately equivalent to the number of destructions of the same species in the same simulation box at equilibrium.

The maximum volume, translation and rotation moves were again adjusted so that approximately 50% of all of these moves were accepted over the course of a simulation for both the μ_1N_2VT and ξ_1NVT Gibbs Ensemble simulations. Both the pure component and mixture simulations were equilibrated for at least 40,000 cycles. The production period for calculating ensemble averages consisted of at least 160,000 cycles. The standard deviations of ensemble averages were estimated by dividing the production cycles of each run into five blocks and calculating the standard deviation from the averages obtained from each of these five blocks.

8.5.2 Simulated Binary Phase Envelopes

The P-x-y and x-y diagrams for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C and the 1-hexene (1) + *n*-hexane (2) mixture at 55°C are shown in Figure 8-4 to Figure 8-7 and the tabulated data is given in Table 8-9 to Table 8-10 in Appendix 8A at the end of the Chapter.

To the knowledge of this work, this is the first time that such narrow phase envelope regions have been simulated with a variant of the Gibbs Ensemble. Both the P-x-y data of the 1-butene (1) + *n*-butane (2) mixture of Figure 8-4 and the P-x-y data of the 1-hexene (1) + *n*-hexane (2) mixture of Figure 8-6 were successfully simulated using both the μ_1N_2VT and ξ_1NVT Gibbs Ensembles. Although both of the simulated phase diagrams are shifted to higher pressures, this is not a failing of the simulation algorithms but rather a function of the overprediction of the vapour pressures for *n*-alkanes and 1-alkenes of the TraPPE force field as discussed in Chapter Six.

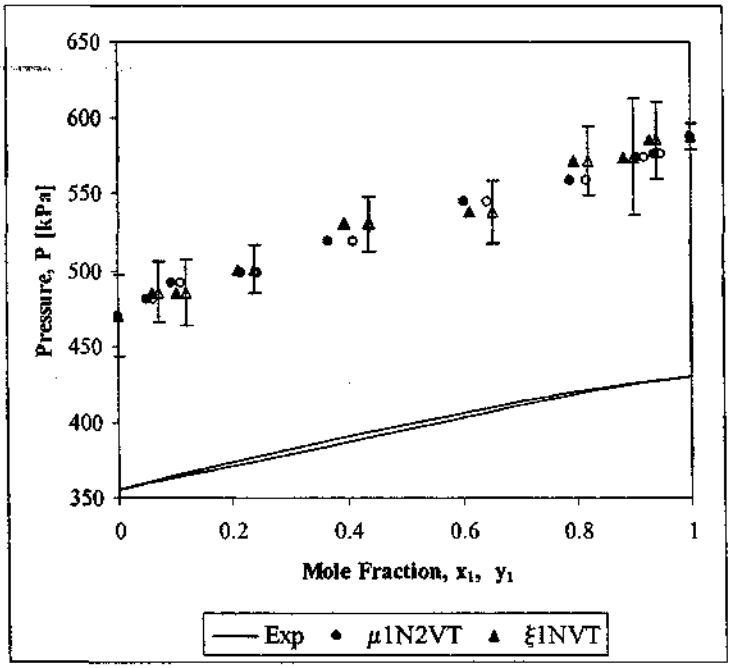


Figure 8-4: P-x-y diagram for the 1-butene (1) + n-butane (2) mixture at 37.8°C simulated using the μ_1N_2VT (circles) and the ξ_1NVT (triangles) Gibbs Ensembles and the corresponding experimental data of Laurance and Swift (1974). Open symbols denote dew points and closed symbols denote bubble points.

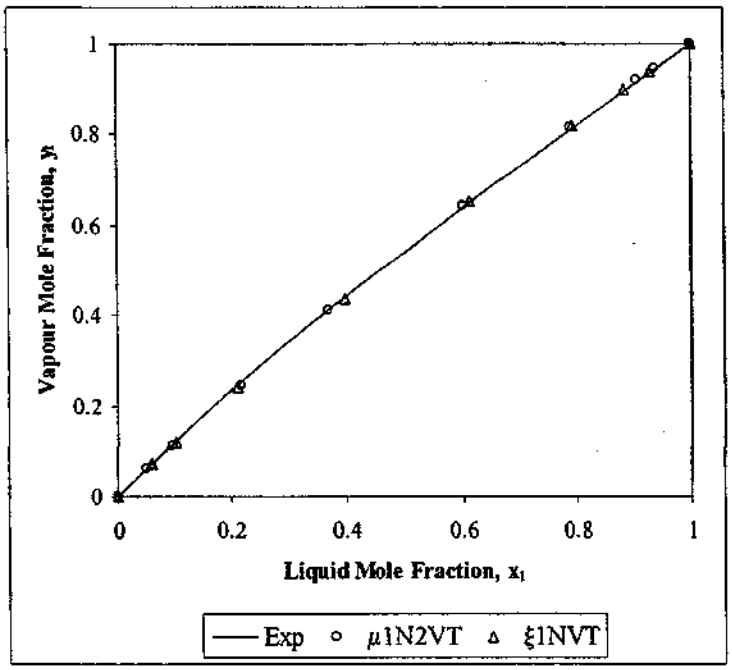


Figure 8-5: x-y diagram for the 1-butene (1) + n-butane (2) mixture at 37.8°C simulated using the μ_1N_2VT (circles) and the ξ_1NVT (triangles) Gibbs Ensembles and the corresponding experimental data of Laurance and Swift (1974).

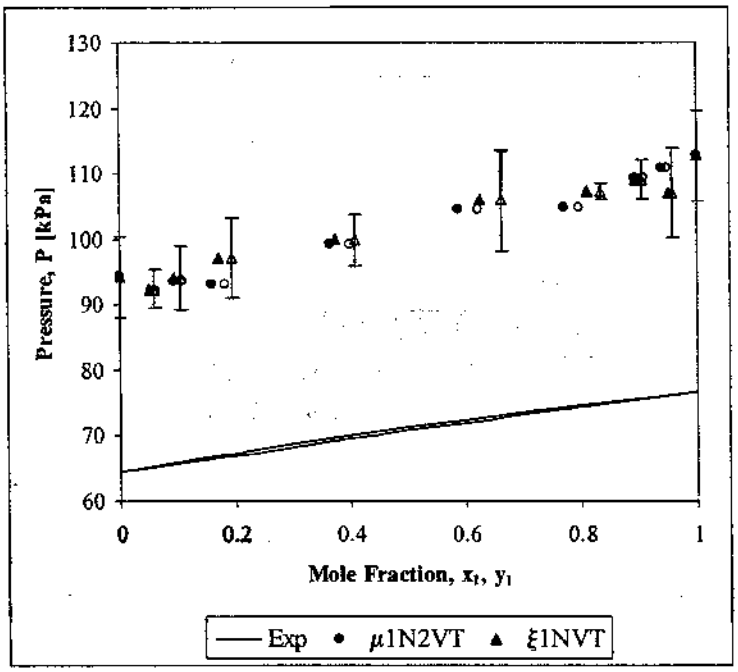


Figure 8-6: P-x-y diagram for the 1-hexene (1) + n-hexane (2) mixture at 55°C simulated using the μ_1N_2VT (circles) and the ξ_1NVT (triangles) Gibbs Ensembles and the corresponding experimental data of the Dortmund Data Bank. Open symbols denote dew points and closed symbols denote bubble points.

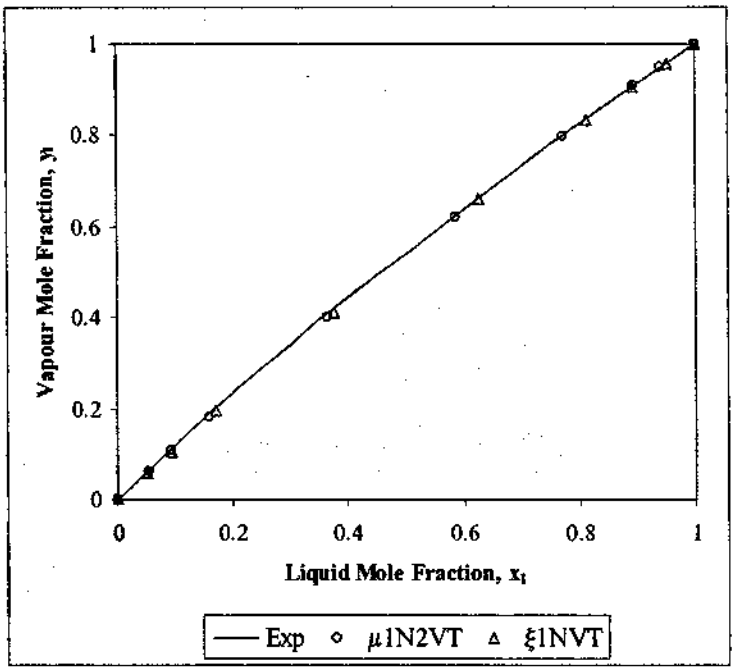


Figure 8-7: x-y diagram for the 1-hexene (1) + n-hexane (2) mixture at 55°C simulated using the μ_1N_2VT (circles) and the ξ_1NVT (triangles) Gibbs Ensembles and the corresponding experimental data of the Dortmund Data Bank.

Table 8-1: Simulated chemical potentials, overall compositions and numbers of molecules in each for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C obtained from the μ_1N_2VT Gibbs Ensemble. Subscripted values denote the calculated statistical uncertainty.

Imposed μ_1 [K]	Imposed μ_2 [K]	Simulated μ_1 [K]	Simulated μ_2 [K]	z_1	$N_{L,1}$	$N_{L,2}$	$N_{V,1}$	$N_{V,2}$
-	-	-	-4769.4_{15}	0		242.9		$57.1_{3.7}$
-5610		$-5601.5_{0.7}$	-4784.5_{10}	$0.054_{0.001}$	$13.1_{0.5}$	234.5	$3.7_{0.0}$	$55.5_{2.4}$
-5420		$-5420.0_{0.7}$	$-4795.9_{6.9}$	$0.099_{0.003}$	$25.1_{1.0}$	236.4	$6.7_{0.1}$	$53.6_{1.6}$
-5170		$-5170.0_{0.5}$	$-4843.8_{6.7}$	$0.222_{0.004}$	$59.3_{1.6}$	214.2	$15.0_{0.1}$	$45.8_{1.4}$
-5000		$-5000.1_{1.5}$	-4908.5_{11}	$0.377_{0.007}$	$95.2_{4.0}$	162.3	$26.3_{0.3}$	$37.7_{1.8}$
	-5050	$-4848.2_{5.6}$	$-5049.8_{1.2}$	$0.609_{0.004}$	$152.9_{0.8}$	101.8	$37.1_{0.8}$	$20.6_{0.1}$
	-5250	-4766.3_{17}	$-5249.9_{1.6}$	$0.794_{0.010}$	$201.5_{3.6}$	54.2	$48.5_{3.6}$	$10.9_{0.2}$
	-5500	-4721.9_{10}	$-5498.1_{1.4}$	$0.909_{0.003}$	$233.9_{2.4}$	24.3	$56.1_{2.4}$	$4.9_{0.0}$
	-5630	$-4711.7_{9.2}$	$-5617.7_{0.6}$	$0.940_{0.002}$	$231.8_{2.2}$	15.4	$58.2_{2.2}$	$3.2_{0.1}$
-	-	$-4686.6_{7.5}$	-	1	$240.6_{1.7}$		$59.4_{1.7}$	

Table 8-2: Simulated chemical potentials, overall compositions and numbers of molecules in each for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C obtained from the ξ_1NVT Gibbs Ensemble. Subscripted values denote the calculated statistical uncertainty.

Imposed ξ_1	Simulated μ_1 [K]	Simulated μ_2 [K]	z_1	$N_{L,1}$	$N_{L,2}$	$N_{V,1}$	$N_{V,2}$
0		-4769.4_{15}	0		$242.9_{3.7}$		$57.1_{3.7}$
0.072	$-5559.6_{9.6}$	-4785.5_{10}	$0.063_{0.0002}$	$14.6_{0.2}$	$225.5_{2.6}$	$4.3_{0.2}$	$55.7_{2.6}$
0.121	-5402.4_{13}	-4802.7_{13}	$0.106_{0.0003}$	$24.7_{0.3}$	$215.5_{2.9}$	$7.2_{0.4}$	$52.6_{2.8}$
0.243	$-5176.5_{8.1}$	$-4839.6_{8.3}$	$0.217_{0.0007}$	$50.2_{0.5}$	$187.8_{1.8}$	$15.0_{0.5}$	$47.1_{1.7}$
0.440	$-4975.4_{8.2}$	$-4917.0_{7.6}$	$0.404_{0.0009}$	$92.3_{1.0}$	$141.4_{1.8}$	$28.9_{1.2}$	$37.3_{1.5}$
0.655	-4846.9_{11}	$-5062.8_{9.3}$	$0.621_{0.0020}$	$142.5_{1.5}$	$90.4_{1.5}$	$43.7_{2.0}$	$23.3_{1.0}$
0.822	-4759.0_{11}	-5251.4_{11}	$0.800_{0.0008}$	$180.9_{2.6}$	$47.2_{0.8}$	$58.9_{2.7}$	$12.9_{0.6}$
0.900	-4729.7_{19}	-5428.8_{18}	$0.886_{0.0005}$	$200.9_{5.1}$	$26.8_{0.7}$	$65.0_{5.2}$	$7.3_{0.6}$
0.940	-4710.0_{11}	$-5578.9_{9.7}$	$0.931_{0.0004}$	$210.0_{3.3}$	$16.2_{0.3}$	$69.3_{3.4}$	$4.5_{0.2}$
1	$-4686.6_{7.5}$		1	$240.6_{1.7}$		$59.4_{1.7}$	

Table 8-3: Simulated chemical potentials, overall compositions and numbers of molecules in each for the 1-hexene (1) + *n*-hexane (2) mixture at 55°C obtained from the μ_1N_2VT Gibbs Ensemble. Subscripted values denote the calculated statistical uncertainty.

Imposed μ_1 [K]	Imposed μ_2 [K]	Simulated μ_1 [K]	Simulated μ_2 [K]	z_1	$N_{L,1}$	$N_{L,2}$	$N_{V,1}$	$N_{V,2}$
-	-		-5861.9_{16}	0		$198.4_{3.0}$		$51.6_{3.0}$
-6750		$-6736.2_{0.7}$	-5890.3_{19}	$0.055_{0.003}$	$11.8_{0.8}$	$202.9_{2.8}$	$3.2_{0.0}$	$47.1_{2.8}$
-6570		$-6568.6_{1.2}$	-5901.9_{17}	$0.094_{0.004}$	$20.3_{1.3}$	$194.4_{2.5}$	$5.5_{0.0}$	$45.6_{2.5}$
-6400		$-6399.9_{0.6}$	-5931.7_{27}	$0.160_{0.012}$	$32.3_{3.7}$	$168.4_{3.7}$	$9.3_{0.1}$	$41.6_{3.7}$
-6120		$-6120.0_{1.9}$	-6013.9_{35}	$0.365_{0.023}$	$79.6_{9.9}$	$137.5_{3.8}$	$21.8_{0.2}$	$32.5_{3.8}$
	-6150	$-5963.7_{4.5}$	$-6149.9_{1.3}$	$0.588_{0.032}$	$139.6_{4.3}$	99.0_{16}	$30.4_{4.3}$	$18.4_{0.2}$
	-6350	$-5878.0_{9.5}$	$-6350.3_{0.9}$	$0.771_{0.004}$	$181.0_{1.1}$	$54.0_{1.6}$	$39.0_{1.1}$	$10.0_{0.0}$
	-6600	-5823.6_{22}	$-6597.2_{0.9}$	$0.895_{0.007}$	$193.4_{3.8}$	$22.8_{2.0}$	$46.6_{3.8}$	$4.7_{0.1}$
	-6780	-5803.4_{16}	$-6757.5_{1.0}$	$0.940_{0.003}$	$200.6_{2.5}$	$12.9_{0.8}$	$49.4_{2.5}$	$2.7_{0.0}$
-	-	-5780.5_{15}		1	$196.8_{3.3}$		$53.2_{3.3}$	

Table 8-4: Simulated chemical potentials, overall compositions and numbers of molecules in each for the 1-hexene (1) + *n*-hexane (2) mixture at 55°C obtained from the ξ_1NVT Gibbs Ensemble. Subscripted values denote the calculated statistical uncertainty.

Imposed ξ_1	Simulated μ_1 [K]	Simulated μ_2 [K]	z_1	$N_{L,1}$	$N_{L,2}$	$N_{V,1}$	$N_{V,2}$
0		-5861.9_{16}	0		$198.4_{3.0}$		$51.6_{3.0}$
0.061	$-6756.6_{8.9}$	-5887.5_{11}	$0.053_{0.0003}$	$10.2_{0.1}$	$189.3_{1.5}$	$3.0_{0.1}$	$47.5_{1.6}$
0.110	-6572.4_{14}	-5899.8_{15}	$0.096_{0.0003}$	$18.5_{0.2}$	$180.1_{2.5}$	$5.5_{0.3}$	$46.0_{2.4}$
0.200	-6366.3_{19}	-5923.1_{19}	$0.176_{0.0008}$	$33.8_{0.4}$	$163.1_{3.0}$	$10.3_{0.6}$	$42.8_{2.8}$
0.420	-6112.3_{13}	-6017.6_{13}	$0.382_{0.0015}$	$73.2_{0.7}$	$122.1_{1.5}$	$22.4_{0.9}$	$32.3_{1.3}$
0.670	-5938.4_{25}	-6182.6_{25}	$0.635_{0.0012}$	$120.3_{2.6}$	$71.6_{1.7}$	$38.4_{2.9}$	$19.7_{1.4}$
0.840	$-5856.5_{6.2}$	$-6412.4_{6.4}$	$0.818_{0.0009}$	$155.3_{0.3}$	$35.8_{0.3}$	$49.1_{0.5}$	$9.8_{0.1}$
0.910	$-5824.2_{9.1}$	$-6593.7_{8.9}$	$0.896_{0.0004}$	$169.7_{1.5}$	$20.3_{0.2}$	$54.4_{1.6}$	$5.6_{0.2}$
0.960	-5811.6_{18}	-6836.1_{13}	$0.953_{0.0003}$	$182.0_{3.6}$	$9.2_{0.2}$	$56.3_{3.6}$	$2.4_{0.2}$
1	-5780.5_{15}		1	$196.8_{3.3}$		$53.2_{3.3}$	

8.5.2.1 Pressures from μ_1N_2VT or ξ_1NVT Gibbs Ensemble Simulations

The apparent scatter in the simulated pressures is, at first glance, also disappointing. In Figure 8-4 and Figure 8-6, only the statistical uncertainty in the pressure for the dew points for the ξ_1NVT Gibbs Ensemble simulations is shown for clarity because the uncertainty is of similar magnitude for all pure component and mixture simulation points. It must be noted, however, that the estimated statistical uncertainty in the simulated vapour pressures of the pure components using the NVT Gibbs Ensemble are of similar magnitude to the estimated uncertainty in the pressures calculated for the mixture points using either the μ_1N_2VT Gibbs Ensemble or the ξ_1NVT Gibbs Ensemble. This may be seen in Table 8-9 and in Table 8-10. Depending on the number of simulation production cycles, it is well known that the uncertainty in the simulation pressure from an isochoric Gibbs Ensemble is of the order of 5-10%. For the two mixtures studied, the small difference between the pure component vapour pressures magnifies this apparent uncertainty in Figure 8-4 and Figure 8-6. For an N_1N_2PT Gibbs Ensemble, the pressure for each mixture point is derived from the imposed pressure P used for each simulation point. For the μ_1N_2VT and ξ_1NVT Gibbs Ensembles, the pressure is not imposed which tends to yield the higher statistical uncertainty in the simulation pressure. This originates from the way in which new configurations are generated in conventional Monte Carlo simulations. Indeed, recalling Equation (3–16), it may be seen that new configurations are generated by considering the change in the energy between the old and new states only where configurations which lower the system energy are preferentially sampled.

No conventional algorithm for isochoric Monte Carlo simulations preferentially seeks to minimize the difference $|P_L - P_V|$ to rigorously satisfy the requirement for mechanical equilibrium as outlined by Equation (4–1). The NVT, μ_1N_2VT and ξ_1NVT Gibbs Ensembles achieve mechanical equilibrium by enforcing a conservation of total volume and requiring that $P_L = P_V$ in an average sense such that no pressure difference terms appear in the acceptance criterion of Equation (4–6). For isochoric Gibbs Ensembles, this is inescapable because of the way in which new configurations are generated by the Monte Carlo sampling algorithm. Thus while it is still correct, it can be argued that isochoric Gibbs Ensembles are, therefore, less rigorous in their establishment of mechanical equilibrium. Instantaneous configurations can (and do) display deviations from the mechanical requirement of Equation (4–1) which leads to the larger statistical uncertainty in the ensemble averaged pressure. Figure 8-8 and Figure 8-9 show the evolution of the ensemble average pressures for the simulated 1-butene + *n*-butane mixture for a point corresponding to $z_1 \approx 0.2$.

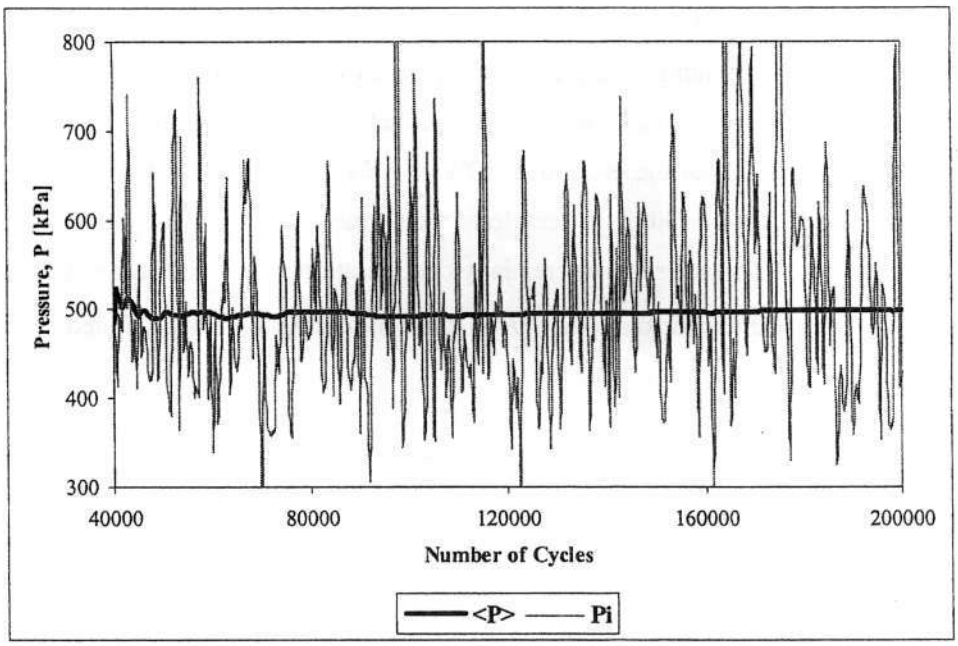


Figure 8-8: Evolution of the ensemble average pressure (dark line) as a function of the number of simulation cycles for the μ_1N_2VT Gibbs Ensembles for a point corresponding to $z_1 \approx 0.22$ for the 1-butene (1) + *n*-butane mixture at 37.8°C. The instantaneous pressures are shown as the lighter dotted line.

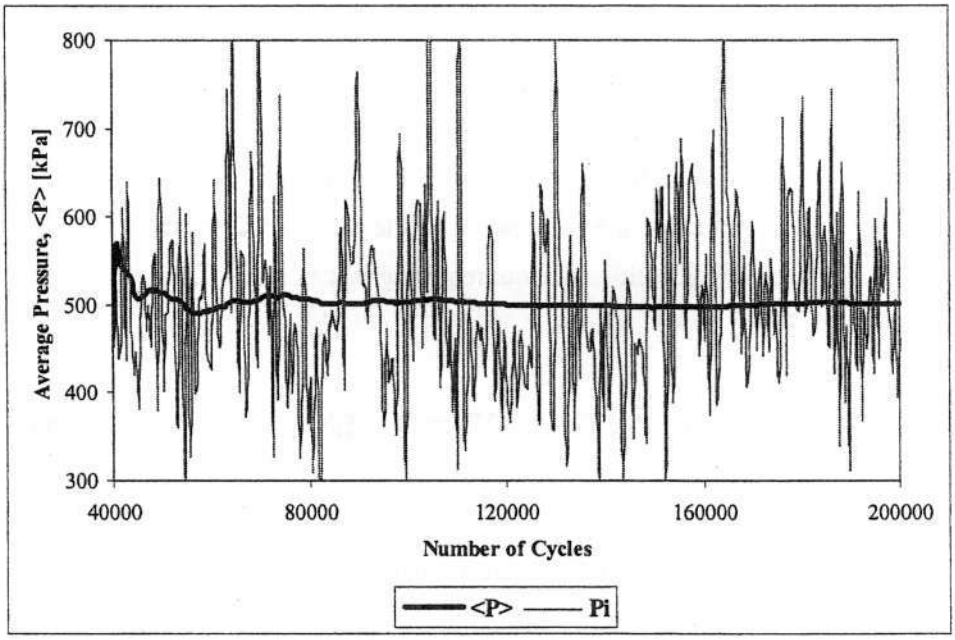


Figure 8-9: Evolution of the ensemble average pressure (dark line) as a function of the number of simulation cycles for the ξ_1NVT Gibbs Ensembles for a point corresponding to $z_1 \approx 0.22$ for the 1-butene (1) + *n*-butane mixture at 37.8°C. The instantaneous pressures are shown as the lighter dotted line.

Only configurations from 40,000 cycles onwards are shown in Figure 8-8 and Figure 8-9 as these then correspond to the production period for ensemble averages. As may be seen in these graphs, the averages pressures are composed of much larger instantaneous fluctuations which, in general, lead to statistical uncertainties of the order of 5-10% in the simulated average pressures. This is typical behaviour for isochoric Gibbs Ensembles. Consequently, further attention needs to be directed towards algorithms which still preferentially configurations which lower the overall system energy but which also minimize the difference $|P_L - P_V|$ in the configurations generated.

8.5.2.2 Compositions from μ_1N_2VT or ξ_1NVT Gibbs Ensemble Simulations

In spite of the apparent fluctuations in the pressures, Figure 8-5 and Figure 8-7 reveal that the x-y data is well described by the μ_1N_2VT and ξ_1NVT Gibbs Ensembles for both mixtures. The accurate description of the x-y data is largely a feature of the TraPPE force field which is able to reproduce this type of data well as described in Chapter Six. Most significantly, errors bars are excluded from Figure 8-5 and Figure 8-7 because they are less than the symbol size. This is confirmed by the data listed in Table 8-9 and Table 8-10 in Appendix 8A at the end of the Chapter where the estimated errors in the compositions of both phases for both mixtures is less than 10^{-2} for the μ_1N_2VT Gibbs Ensemble simulations and less than 10^{-3} for the ξ_1NVT Gibbs Ensemble simulations. This represents a considerable improvement over the estimated uncertainties in the compositions from N_1N_2PT Gibbs Ensemble simulations which tend to be of the order of 0.02-0.03. The reason for this improvement is because the μ_1N_2VT Gibbs Ensemble allows for the creation and destruction of molecules of one of the components while the ξ_1NVT Gibbs Ensemble is able to allow the molecules to change identities without requiring that the total number of each species be conserved.

The novel composition-based moves of the μ_1N_2VT and the ξ_1NVT Gibbs Ensembles result in their compositional sampling being better than that of the conventional N_1N_2PT Gibbs Ensemble. This is also because the overall composition can vary for both the μ_1N_2VT and the ξ_1NVT Gibbs Ensembles. The rapid convergence of the overall, vapour and liquid mole fractions for the 1-butene + *n*-butane mixture at 37.8°C for a point corresponding roughly to $z_1 \approx 0.22$ for the μ_1N_2VT and the ξ_1NVT Gibbs Ensembles are shown in Figure 8-10 and Figure 8-11 respectively. It is important to note that ensemble averaged compositions in Figure 8-11 from the ξ_1NVT Gibbs Ensemble simulations show considerably smaller fluctuations than the corresponding compositions in Figure 8-10 from the μ_1N_2VT Gibbs Ensemble simulations. This is supported by comparing the data listed

in Table 8-9 and Table 8-10. The reason why the ξ_iNVT Gibbs Ensemble has lower statistical uncertainties is because molecule identity changes are permissible for both species in the ξ_iNVT Gibbs Ensemble while molecule creations and destructions are only permissible for one of the species in the μ_1N_2VT Gibbs Ensemble. The second species in the μ_1N_2VT Gibbs Ensemble can only achieve chemical equilibrium by (less efficient) molecule transfers between the simulation boxes. The reason why molecule transfers will always be less efficient is because they require a simultaneous molecule destruction attempt in one simulation box and a corresponding molecule creation attempt in the other simulation box. The ξ_iNVT Gibbs Ensemble, therefore, has some advantages in the sense that it will, in general, ensure lower statistical uncertainties in the simulated compositions. This is evidenced in Table 8-5 to Table 8-8. For both mixtures simulated using the μ_1N_2VT Gibbs Ensemble, the highest acceptance rates ($\gtrsim$70%) are obtained for creations and destructions of molecules in the vapour phase where the probability of generating an overlap is lowest and where the energy associated with a creation or destruction attempt is lowest. It may also be seen that the probability of a molecule transfer is also less than or equal to the probability of a molecule creation or destruction in the μ_1N_2VT Gibbs Ensemble in the liquid phase. This is because a molecule transfer requires a molecule destruction in one simulation box and a simultaneous molecule creation in the remaining simulation box and the probability for both of these moves are separately less than unity as seen from the results of Table 8-5 and Table 8-7.

The importance of the "conservative" identity swap move of Equation (4–18) for exploring different composition configurations is shown by the high acceptances for this move for both the μ_1N_2VT and the ξ_iNVT Gibbs Ensembles in Table 8-5 to Table 8-8. Typically, the acceptance rates for this move are found to be above 30% for the mixtures studied. The high acceptance percentages for the "non-conservative" identity change move (which is exclusive to the ξ_iNVT Gibbs Ensemble) are, however, the important factor behind the low uncertainties of the compositions seen for the ξ_iNVT Gibbs Ensemble. This move is performed for both species in a binary mixture simulation and the acceptance percentages observed of over 55% for both mixtures are what guarantees the more efficient sampling of compositions for the ξ_iNVT Gibbs Ensemble. This is why the ξ_iNVT Gibbs Ensemble has the lowest estimated uncertainties in the simulated compositions.

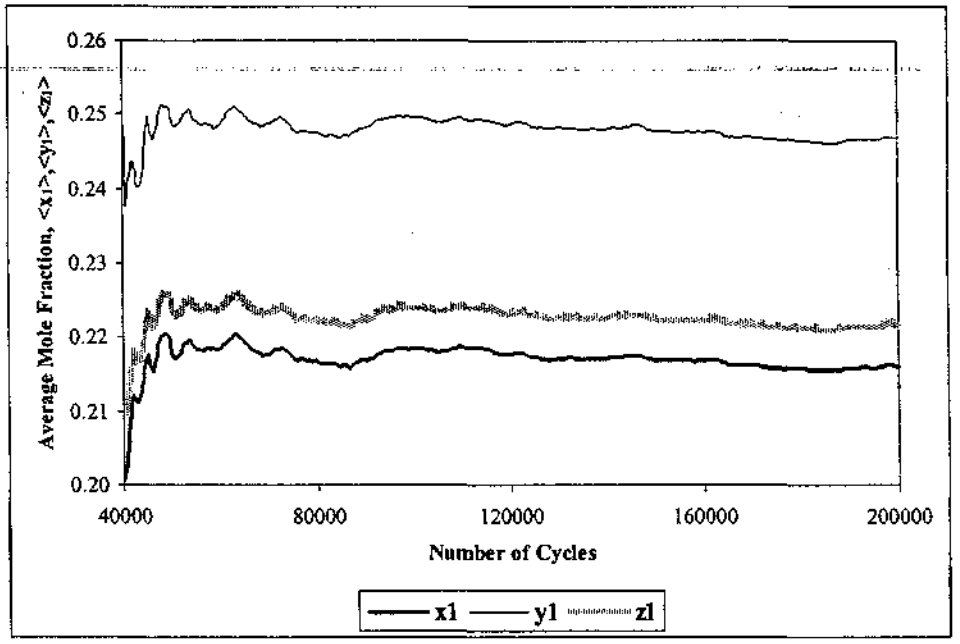


Figure 8-10: Evolution of the ensemble average liquid mole fraction (dark line), x_1, vapour mole fraction (plain line), y_1, and the overall mole fraction (dotted line), z_1, of 1-butene for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C determined using the μ_1N_2VT Gibbs Ensemble for $z_1 \approx 0.22$.

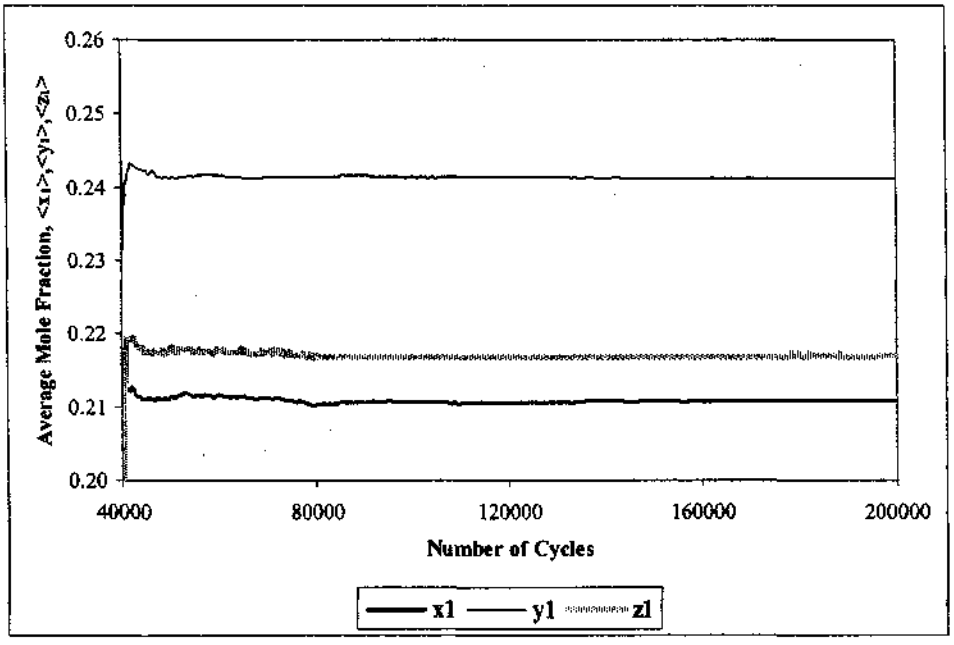


Figure 8-11: Evolution of the ensemble average liquid mole fraction (dark line), x_1, vapour mole fraction (plain line), y_1, and the overall mole fraction (dotted line), z_1, of 1-hexene for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C determined using the ξ_1NVT Gibbs Ensemble for $z_1 \approx 0.22$.

Table 8-5: Percentage acceptances for the composition-related moves in the μ_1N_2VT Gibbs Ensemble for the 1-butene (1) +*n*-butane (2) mixture at 37.8°C.

Imposed μ_1 [K]	Imposed μ_2 [K]	Acceptance [%]										
		Molecule Transfers		Creations Liquid		Destructions Liquid		Creations Vapour		Destructions Vapour		Conservative Identity Swaps Eq. (4–18)
		1-butene	*n*-butane	1-butene	*n*-butane	1-butene	*n*-butane	1-butene	*n*-butane	1-butene	*n*-butane	
-5610	-	-	4.2	5.3	-	5.3	-	75.4	-	75.4	-	31.5
-5420	-	-	4.1	5.2	-	5.2	-	80.0	-	79.9	-	32.3
-5170	-	-	4.2	5.3	-	5.3	-	84.0	-	84.0	-	33.2
-5000	-	-	4.2	5.2	-	5.2	-	85.5	-	85.5	-	33.3
-	-5050	5.2	-	-	4.2	-	4.2	-	84.1	-	84.0	33.2
-	-5250	5.2	-	-	4.3	-	4.3	-	81.6	-	81.6	32.8
-	-5500	5.2	-	-	4.2	-	4.2	-	76.8	-	76.8	31.7
-	-5630	5.2	-	-	4.3	-	4.3	-	73.3	-	73.2	31.0

Table 8-6: Percentage acceptances for the composition-related moves in the ξ_1NVT Gibbs Ensemble for the 1-butene (1) +*n*-butane (2) mixture at 37.8°C.

Imposed ξ_1	Acceptance [%]			
	Molecule Transfers		Conservative Identity Swaps, Eq. (4–18)	Non-Conservative Identity Swaps, Eq.'s (8–15) & (8–16)
	1-butene	*n*-butane		
0.072	5.0	4.2	31.6	56.0
0.121	5.1	4.2	32.4	57.9
0.243	5.1	4.1	33.1	59.7
0.440	5.1	4.1	33.3	60.5
0.655	5.2	4.2	33.3	60.4
0.822	5.1	4.1	32.8	59.3
0.900	5.1	4.1	32.2	57.8
0.940	5.2	4.0	31.5	56.1

Table 8-7: Percentage acceptances for the composition-related moves in the μ_1N_2VT Gibbs Ensemble for the 1-hexene (1) +*n*-hexane (2) mixture at 55°C.

Imposed μ_1 [K]	Imposed μ_2 [K]	Acceptance [%]										
		Molecule Transfers		Creations Liquid		Destructions Liquid		Creations Vapour		Destructions Vapour		Conservative Identity Swaps Equation (4–18)
		1-hexene	*n*-hexane	1-hexene	*n*-hexane	1-hexene	*n*-hexane	1-hexene	*n*-hexane	1-hexene	*n*-hexane	
-6750	-	-	2.1	2.6	-	2.6	-	74.7	-	74.6	-	41.6
-6570	-	-	2.1	2.6	-	2.6	-	79.5	-	79.4	-	42.7
-6400	-	-	2.1	2.7	-	2.7	-	82.8	-	82.8	-	43.6
-6120	-	-	2.0	2.6	-	2.6	-	86.2	-	86.3	-	44.5
-	-6150	2.6	-	-	2.1	-	2.1	-	86.3	-	86.3	44.2
-	-6350	2.6	-	-	2.1	-	2.0	-	83.6	-	83.6	43.7
-	-6600	2.5	-	-	2.0	-	2.0	-	78.4	-	78.4	42.1
-	-6780	2.5	-	-	2.0	-	2.0	-	73.1	-	73.0	41.0

Table 8-8: Percentage acceptances for the composition-related moves in the ξ_1NVT Gibbs Ensemble for the 1-hexene (1) +*n*-hexane (2) mixture at 55°C.

Imposed ξ_1	Acceptance [%]			
	Molecule Transfers		Conservative Identity Swaps, Eq. (4–18)	Non-Conservative Identity Swaps, Eq.'s (8–15) & (8–16)
	1-butene	*n*-butane		
0.061	2.5	2.1	41.3	60.4
0.110	2.5	2.1	42.6	62.5
0.200	2.6	2.1	43.7	64.3
0.420	2.6	2.0	44.4	65.5

Imposed ξ_1	Acceptance [%]			
	Molecule Transfers		Conservative Identity Swaps, Eq. (4–18)	Non-Conservative Identity Swaps, Eq.'s (8–15) & (8–16)
0.670	2.5	2.0	44.4	65.5
0.840	2.5	2.0	43.5	64.2
0.910	2.5	1.9	42.4	62.5
0.960	2.5	1.9	40.6	59.6

8.5.2.3 Equivalence of the μ_1N_2VT and ξ_1NVT Gibbs Ensembles

Figure 8-12 and Figure 8-13 demonstrate the statistical and thermodynamic consistency of the μ_1N_2VT and ξ_1NVT Gibbs Ensembles with respect to each other. For a given overall mole fraction of 1-butene, z_1, the chemical potentials of each species lie on the same curve, irrespective of the type of Gibbs Ensemble used. The data points do not lie at precisely the same point along the same curve because the different extensive parameter specifications required for the μ_1N_2VT and ξ_1NVT Gibbs Ensembles make it impossible to exactly duplicate the same thermodynamic conditions. Figure 8-12 and Figure 8-13 are, nonetheless, important because they confirm the equivalence of the two new types of Gibbs Ensembles because they yield the same values for intensive properties at chemical equilibrium. The equivalence of the two ensembles may also be compared in terms of the simulated numbers of each species in each phase at equilibrium. The total number of molecules present in a μ_1N_2VT Gibbs Ensemble simulation can vary because only N_2 is constrained to remain constant. From Table 8-1 to Table 8-4 listed previously, it may be seen that approximately equivalent chemical equilibrium specifications (simulated using approximately the same total volume, V) yield approximately equivalent numbers of each type of molecule in each phase at equilibrium. This result is also important because it confirms the extensive property equivalence of these two new ensembles.

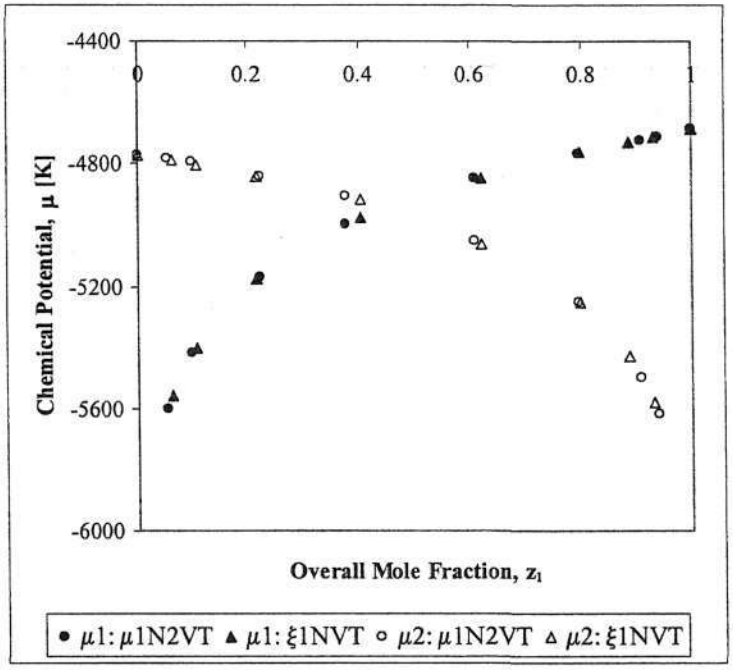


Figure 8-12: Chemical potentials as a function of the overall simulated mole fraction for the 1-butene (1) + n-butane (2) mixture at 37.8°C simulated using the μ_1N_2VT (circles) and the ξ_1NVT (triangles) Gibbs Ensembles. Closed symbols denote 1-butene and open symbols denote *n*-butane.

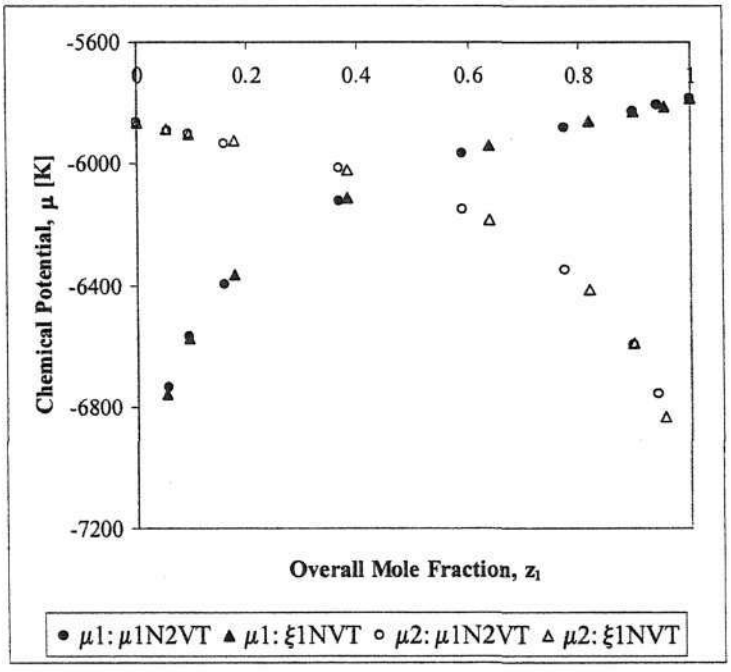


Figure 8-13: Chemical potentials as a function of the overall simulated mole fraction for the 1-hexene (1) + n-hexane (2) mixture at 55°C simulated using the μ_1N_2VT (circles) and the ξ_1NVT (triangles) Gibbs Ensembles. Closed symbols denote 1-hexene and open symbols denote *n*-hexane.

Since both the μ_1N_2VT and ξ_1NVT Gibbs Ensembles rely on composition-based Monte Carlo moves to achieve thermodynamic equilibrium, it is important to validate the application of Equation (5–37) for simulations in these ensembles. Figure 8-14 to Figure 8-17 in Appendix 8A at the end of this Chapter demonstrate the validity of Equation (5–37) and the ability of both the μ_1N_2VT and ξ_1NVT Gibbs Ensembles to converge rapidly to the final equilibrium chemical potentials. Indeed, both ensembles satisfy the requirement for the chemical equilibrium of each species of Equation (4–1) because the chemical potentials of each individual species in different phases closely approach similar values at the end of a simulation for both Gibbs Ensembles. Figure 8-14, however, reveals that the calculated chemical potential which is imposed for 1-butene for the μ_1N_2VT Gibbs Ensemble has a lower associated uncertainty than the chemical potential for the same species when calculated from a simulation in the ξ_1NVT Gibbs Ensemble as shown in Figure 8-15. This behaviour is to be expected though because the chemical potential is not strictly imposed for the ξ_1NVT Gibbs Ensemble. This is confirmed by the observation that the chemical potentials of *n*-butane calculated from both the μ_1N_2VT and ξ_1NVT Gibbs Ensembles in Figure 8-16 and Figure 8-17 have similar uncertainties.

8.6 Appendix 8A: P-x-y, x-y and chemical potential data from μ_1N_2VT and the ξ_1NVT Gibbs Ensemble Simulations

Table 8-9: Simulated composition and pressure data for the 1-butene (1) + *n*-butane (2) mixture at 37.8°C simulated using the μ_1N_2VT and the ξ1**NVT Gibbs** Ensembles and the corresponding experimental data of Laurance and Swift (1974). Subscripted values denote the calculated statistical uncertainty.

Experimental			μ_1N_2VT Gibbs Ensemble					ξ_1NVT Gibbs Ensemble			
P [kPa]	x_1	y_1	Imposed μ_1 [K]	Imposed μ_2 [K]	P [kPa]	x_1	y_1	Imposed ξ_1	P [kPa]	x_1	y_1
355.1	0	0	-	-	470.6_{27}	0	0	0	470.6_{27}	0	0
364.7	0.1	0.1206	-5610		480.9_{17}	$0.053_{0.001}$	$0.062_{0.002}$	0.072	485.7_{20}	$0.061_{0.0002}$	$0.071_{0.0003}$
373.9	0.2	0.2345	-5420		492.3_{11}	$0.095_{0.003}$	$0.112_{0.002}$	0.121	485.1_{22}	$0.103_{0.0002}$	$0.120_{0.0004}$
382.8	0.3	0.3425	-5170		$498.3_{9.3}$	$0.216_{0.004}$	$0.247_{0.004}$	0.243	501.0_{15}	$0.211_{0.0009}$	$0.241_{0.0006}$
391.3	0.4	0.4452	-5000		519.4_{15}	$0.369_{0.007}$	$0.410_{0.009}$	0.440	530.7_{18}	$0.395_{0.0009}$	$0.437_{0.0004}$
399.3	0.5	0.5426		-5050	$545.3_{5.6}$	$0.602_{0.004}$	$0.643_{0.004}$	0.655	538.2_{20}	$0.612_{0.0020}$	$0.652_{0.0008}$
406.7	0.6	0.6383		-5250	559.5_{29}	$0.789_{0.010}$	$0.816_{0.008}$	0.822	571.6_{23}	$0.793_{0.0008}$	$0.820_{0.0003}$
413.6	0.7	0.7302		-5500	574.4_{17}	$0.906_{0.002}$	$0.920_{0.003}$	0.900	574.5_{38}	$0.882_{0.0004}$	$0.899_{0.0003}$
420.0	0.8	0.8204		-5630	576.7_{17}	$0.938_{0.002}$	$0.947_{0.001}$	0.940	585.6_{25}	$0.928_{0.0005}$	$0.939_{0.0003}$
425.6	0.9	0.9099	-	-	$588.4_{8.7}$	1	1	1	$588.4_{8.7}$	1	1
430.6	1	1									

Table 8-10: Simulated composition and pressure data for the 1-hexene (1) + *n*-hexane (2) mixture at 55°C simulated using the μ_1N_2VT and the **ξ1NVT Gibbs** Ensembles and the corresponding experimental data of the Dortmund Data Bank. Subscripted values denote the calculated statistical uncertainty.

Experimental			μ_1N_2VT Gibbs Ensemble					ξ_1NVT Gibbs Ensemble			
P [kPa]	x_1	y_1	Imposed μ_1 [K]	Imposed μ_2 [K]	P [kPa]	x_1	y_1	Imposed ξ_1	P [kPa]	x_1	y_1
64.4	0	0	-	-	$94.3_{6.1}$	0	0	0	$94.3_{6.1}$	0	0
66.0	0.102	0.122	-6750		$92.2_{4.8}$	$0.055_{0.003}$	$0.063_{0.003}$	0.061	$92.5_{2.8}$	$0.051_{0.0003}$	$0.059_{0.0006}$
67.4	0.200	0.233	-6570		$93.6_{4.4}$	$0.094_{0.004}$	$0.108_{0.005}$	0.110	$94.1_{4.8}$	$0.093_{0.0004}$	$0.106_{0.0006}$
68.7	0.300	0.341	-6400		$93.2_{6.6}$	$0.160_{0.012}$	$0.182_{0.012}$	0.200	$97.1_{6.0}$	$0.171_{0.0009}$	$0.194_{0.0007}$
70.1	0.400	0.445	-6120		$99.2_{7.0}$	$0.365_{0.023}$	$0.401_{0.025}$	0.420	$99.9_{3.9}$	$0.375_{0.0016}$	$0.410_{0.0008}$
71.3	0.500	0.543		-6150	$104.5_{9.2}$	$0.588_{0.032}$	$0.622_{0.030}$	0.670	$105.9_{7.8}$	$0.627_{0.0006}$	$0.661_{0.0006}$
72.4	0.600	0.639		-6350	$104.8_{2.3}$	$0.771_{0.004}$	$0.797_{0.004}$	0.840	$107.2_{1.3}$	$0.813_{0.0013}$	$0.834_{0.0008}$
73.6	0.700	0.732		-6600	$109.2_{7.8}$	$0.895_{0.007}$	$0.908_{0.005}$	0.910	$109.1_{3.0}$	$0.893_{0.0003}$	$0.907_{0.0005}$
74.7	0.799	0.823		-6780	$110.9_{5.1}$	$0.940_{0.003}$	$0.948_{0.002}$	0.960	$107.0_{6.7}$	$0.952_{0.0003}$	$0.958_{0.0002}$
75.8	0.900	0.912	-	-	$112.7_{7.0}$	1	1	1	$112.7_{7.0}$	1	1
76.8	1	1									

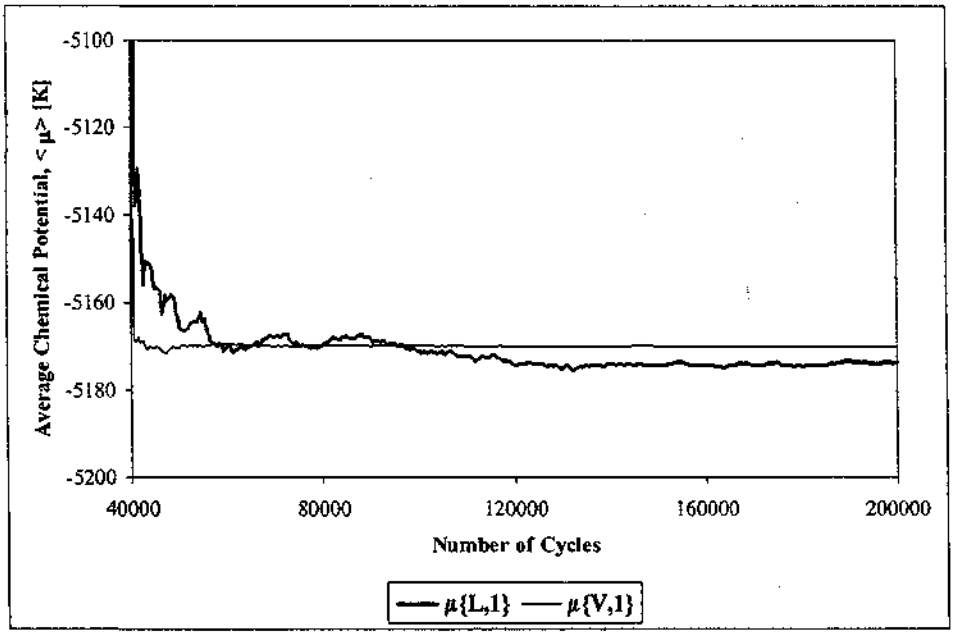


Figure 8-14: Evolution of the ensemble average liquid (dark line) and vapour (plain line) chemical potentials of 1-butene in the 1-butene (1) + *n*-butane (2) mixture at 37.8°C with $z_1 \approx 0.22$ using the μ_1N_2VT Gibbs Ensemble.

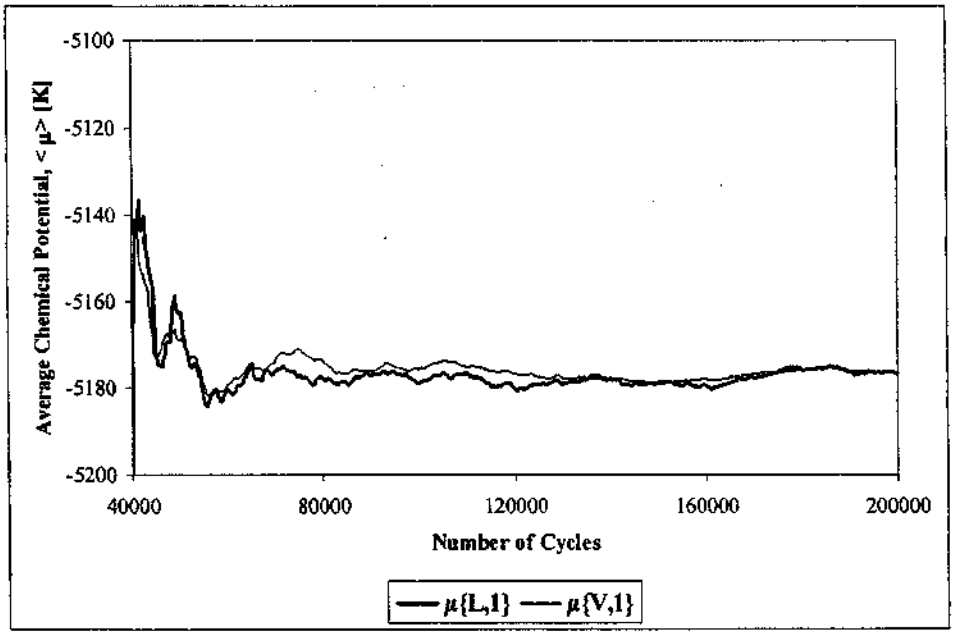


Figure 8-15: Evolution of the ensemble average liquid (dark line) and vapour (plain line) chemical potentials of 1-butene in the 1-butene (1) + *n*-butane (2) mixture at 37.8°C with $z_1 \approx 0.22$ using the ξ_1NVT Gibbs Ensemble.

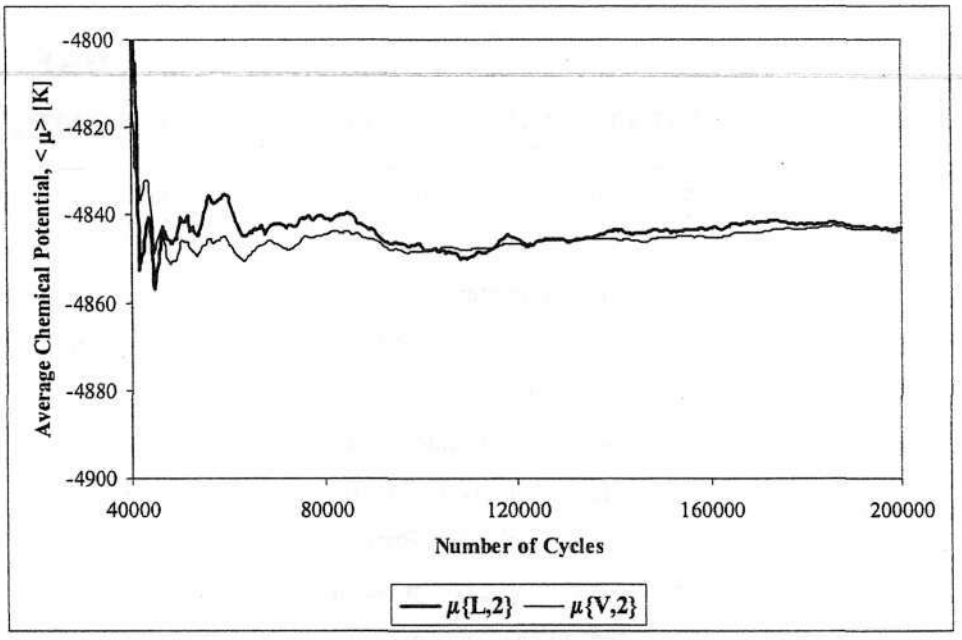


Figure 8-16: Evolution of the ensemble average liquid (dark line) and vapour (plain line) chemical potentials of *n*-butane in the 1-butene (1) + *n*-butane (2) mixture at 37.8°C with $z_1 \approx 0.22$ using the μ_1N_2VT Gibbs Ensemble.

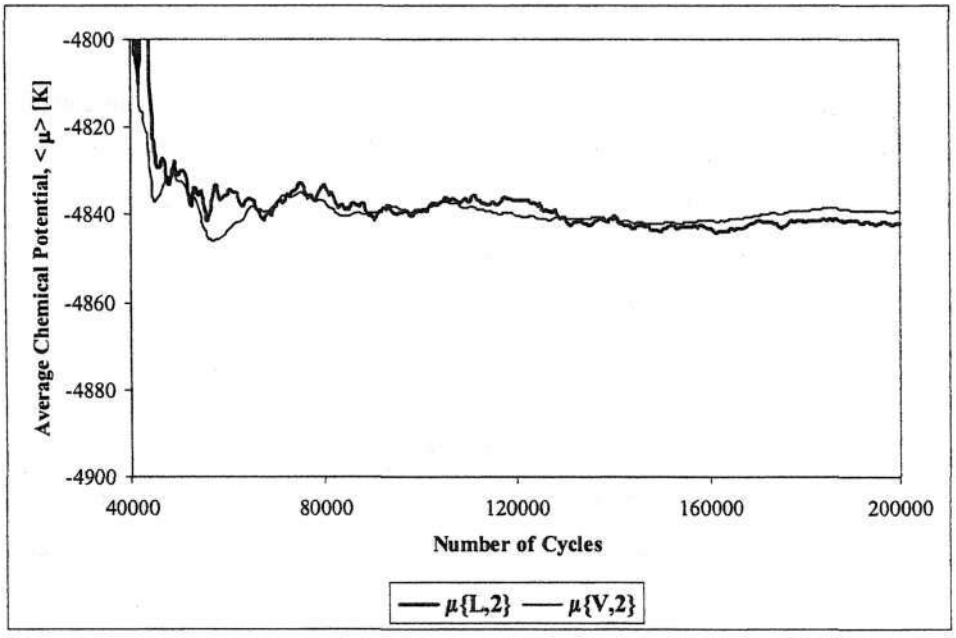


Figure 8-17: Evolution of the ensemble average liquid (dark line) and vapour (plain line) chemical potentials of *n*-butane in the 1-butene (1) + *n*-butane (2) mixture at 37.8°C with $z_1 \approx 0.22$ using the ξ_1NVT Gibbs Ensemble.

CHAPTER 9:

TRANSFERABILITY BASED ON RELAXING GEOMETRIES

"If you look for the good in something, you will surely find it" – Abraham Lincoln

A number of issues concerning the parameterization of intermolecular force fields and how the parameterization affects the simulated pure component properties and the simulated binary VLE have been alluded to in Chapter Six and Chapter Seven including:

1) If the saturated vapour pressure is to be reproduced accurately, more consideration needs to be given towards including this property in the fitting process for the intermolecular parameters. This view has been supported by Errington and Panagiotopoulos (1999a) and by Li *et al.* (2003). It was also seen in Chapter Seven that an improvement in the vapour phase densities for conjugated alkenes accompanied the improvement in the saturated vapour pressure. This was also observed by Errington and Panagiotopoulos (1999a) who produced what is arguably the most accurate united-atom force field for *n*-alkanes to date.
2) The complete transferability of intramolecular geometries should be reviewed. The work in Chapter Seven indicated that an improvement in the pure component properties can result if the equilibrium intramolecular geometries are selected to closely approximate their experimentally observed values. It may also be necessary to incorporate a wider range of equilibrium molecular geometries as part of the force field to achieve a good prediction of all pure component properties.
3) A wide range of molecules and homologous series should be included in the parameterization of intermolecular force field parameters to ensure the maximum transferability of the parameters fitted. This idea has been explored by Ungerer *et al.* (2000) and by Bourasseau *et al.* (2003).
4) Force fields should be tested to examine their prediction of binary VLE. Subtle inaccuracies in the regressed non-bonded interaction parameters can result in significant errors in the simulated P-x-y and x-y data when binary or multi-component VLE is simulated.

To examine these ideas, the following Chapter attempts to develop an improved force field for *n*-alkanes and 1-alkenes.

9.1 The Evolution of the Functional Form of Transferable Force Fields

The evolution of transferable united-atom force fields can be traced to the 1980's with the work of Jorgensen *et al.* (1984) on the OPLS force field. A number of molecular force fields already existed by this time. The work of Jorgensen and co-workers, however, was significant because:

- it introduced the basic functional form which has been widely employed by molecular simulation ever since, and
- it was the first study to attempt to simulate the vapour-liquid equilibrium behaviour of a range of hydrocarbons using a limited "base" set of intermolecular and intramolecular parameters

The main features of this force field included:

1) One common constant bond length and one common constant bond angle, but a cosine series for the torsional potential.
2) A Lennard-Jones description for the non-bonded interactions (including interaction sites on the same molecule separated by more than 3 bonds), and a Coulombic description for charges which were assigned to each of the interaction sites.
3) Distinct LJ ε_{ii} parameters for the $-CH_2-$ and $-CH_3$ functional groups. Unlike in previous polymer simulation studies (where differences in the interactions of terminal carbon groups can be treated as negligible compared to the bulk interactions of a molecular chain), the $-CH_2-$ and $-CH_3$ groups were now regarded as having interactions which were distinctly different in magnitude.
4) The LJ σ_{ii} and ε_{ii} parameters for the $-CH_2-$ group were regressed from simulations of liquid cyclopentane and were then regarded as common for all the branched and normal alkanes studied.
5) The LJ σ_{ii} and ε_{ii} parameters for the $-CH_3$ group had to be made distinct for different branched and normal alkanes to get a good fit to the data.
6) The LJ σ_{ii} parameters for the $-CH_2-$ and $-CH_3$ groups were set to the same value.
7) The LJ parameters were regressed to predict pure component *n*-alkane behaviour for a variety of physical properties well near 25°C

Requiring only 5 distinct LJ parameters for the *n*-alkanes, the OPLS force field was reasonably successful. The OPLS force field did not, however, give a good description of the vapour-liquid coexistence curve up the critical point. The development of CBMC techniques also revealed that

the OPLS predictions became progressively worse for longer *n*-alkanes. In response to this observation, Siepmann *et al.* (1993) derived the SKS force field which provided a considerably better description of the critical densities, the critical temperatures and the liquid densities over a wider temperature range for molecules from 6 to 48 carbon atoms. The main features of this force field included:

1) One common constant bond length, one common but now flexible bond angle described by a harmonic potential, and a cosine series for the torsional potential.
2) Distinct LJ σ_{ii} and ε_{ii} parameters for the $-CH_2-$ and $-CH_3$ functional groups. Only three LJ parameters were, however, required because while the ε_{ii} parameters for the $-CH_2-$ and $-CH_3$ functional groups were distinct, the σ_{ii} parameters for the $-CH_2-$ and $-CH_3$ functional groups had the same magnitude.
3) The LJ parameters were regressed to reproduce the pure component liquid densities and the pure component critical temperature and critical densities.

The SKS force field provided a very good description for medium to long *n*-alkanes but gave no explicit attention to short *n*-alkanes with less than 6 carbons. The SKS force field did, however, demonstrate that a small group of transferable LJ parameters could provide a very good description of the pure component vapour-liquid coexistence curve over a wide range of temperatures for a range of carbon chain lengths.

The TraPPE force field improved upon the LJ parameters of the SKS force field and, importantly, was also applicable to short *n*-alkanes. The main features of this widely used force field include:

1) As for the SKS force field, one common constant bond length, one common flexible bond angle described by a harmonic potential, and a cosine series for the torsional potential.
2) Distinct σ_{ii} and ε_{ii} LJ parameters for the $-CH_2-$ and $-CH_3$ groups (a total of four non-bonded LJ parameters) which are fully transferable to all *n*-alkanes. The LJ parameters were parameterized sequentially. The σ_{ii} and ε_{ii} parameters for the $-CH_3$ group were regressed first to provide the best possible reproduction of the physical properties of ethane. Thereafter the σ_{ii} and ε_{ii} parameters for the $-CH_2-$ group were regressed to provide the best possible physical property predictions for *n*-octane.
3) As for the SKS force field, the LJ parameters were regressed to reproduce the pure component liquid densities and the pure component critical temperatures and critical densities.

Considering the limited intermolecular and intramolecular parameter set used, the TraPPE force field was remarkably successful even for *n*-alkanes longer than *n*-octane. This success was also evidenced in the TraPPE alkene force field of Wick *et al.* (2000) and served to confirm the usefulness of transferable united-atom force fields. It has, however, been found that while the critical densities, the critical temperatures and the liquid densities are well predicted, the saturated vapour pressures, vapour densities and heats of vaporization are not well reproduced.

More or less concurrent with the development of the TraPPE *n*-alkane and alkene force fields was the development of the NERD *n*-alkane (Nath *et al.* (1998)) and 1-alkene (Nath *et al.* (2001)) force field. The major features of this force field include:

1) One common, but fully flexible, bond length for *n*-alkanes which is described by a harmonic potential. As for the TraPPE force field, however, there is one common flexible bond angle described by a harmonic potential and a torsional potential described by a cosine series.
2) As for the TraPPE force field, there are distinct LJ σ_{ii} and ε_{ii} parameters for the $-CH_2-$ and $-CH_3$ functional groups. There is only one set of σ_{ii} and ε_{ii} parameters for the $-CH_2-$ group but there are distinct σ_{ii} and ε_{ii} parameters for the $-CH_3$ group for ethane, propane and molecules longer than propane to provide a larger set of LJ parameters for the $-CH_3$ group.
3) As for the SKS and the TraPPE force fields, the LJ parameters were regressed to predict pure component critical densities, critical temperatures and liquid densities.

It is interesting to note from the original publications that the NERD force field is, in general, able to provide a better prediction of the saturated vapour pressures, vapour densities and heats of vaporization of short hydrocarbons than the TraPPE force field and this can probably be attributed to the larger set of LJ non-bonded parameters.

Using Hamiltonian scaling grand canonical Monte Carlo and Histogram Reweighting techniques, Errington and Panagiotopoulos (1999a) parameterized an *n*-alkane force field (nAEP force field). The model was able to achieve good success in providing an accurate description of the pure component vapour pressures, vapour densities, liquid densities, critical temperatures and critical densities. The major features of this model include:

1) A Be6 potential is used to describe the non-bonded interactions.

2) Three different bond lengths of 1.839Å for the CH_3-CH_3 bond length in ethane, 1.687Å for the CH_3-CH_2 bond length in longer *n*-alkanes, and 1.536Å for the CH_2-CH_2 bond length in longer *n*-alkanes. The authors found this increase in the number of bond lengths necessary to obtain a good fit to the data.
3) One constant bond length and one fully flexible bond angle were used for all *n*-alkanes. A cosine series was again used for the torsional potential.
4) Distinct σ_{ii}, ε_{ii} and α_{ii} Be6 parameters for the $-CH_2-$ and $-CH_3$ functional groups are used (for a total of six non-bonded Be6 interaction parameters) which are fully transferable to all *n*-alkanes. The Be6 parameters were also parameterized sequentially. The σ_{ii}, ε_{ii} and α_{ii} parameters for the $-CH_3$ group were regressed to reproduce the physical properties of ethane. Thereafter, the σ_{ii}, ε_{ii} and α_{ii} parameters for the $-CH_2-$ group were regressed to provide the best possible physical property predictions for propane and *n*-hexane.
5) The Be6 parameters were regressed to provide a good fit to the experimental pure component liquid and vapour coexistence densities, the critical temperatures, the critical densities, and the saturated vapour pressures.

The force field was able to achieve a superior prediction of the pure component vapour densities and saturated vapour pressures when compared to the corresponding TraPPE and NERD predictions. The nAEP force field has, however, not been extensively used. This is perhaps because of the more complicated from of the non-bonded Be6 potential with the extra α_{ii} parameter as well as the exponential repulsive term which requires more computational overhead.

Most recently, anisotropic united-atom force fields for *n*-alkanes and alkenes have received renewed attention. The option of using anisotropic intermolecular interactions to improve the predictions of united-atom force fields had been alluded to the in first TraPPE publication of Martin and Siepmann (1998). The *n*-alkane force field of Ungerer *et al.* (2000) which optimized the parameter set of Toxvaerd (1997), and the alkene force field of Bourasseau *et al.* (2003) represent the most comprehensive AUA force field available. The main features of this force field include:

1) One constant bond length and one constant bond angle were used for all *n*-alkanes. A cosine series was used to describe the torsional potential.
2) Distinct LJ σ_{ii} and ε_{ii} parameters for both the $-CH_2-$ and $-CH_3$ groups, and two distinct anisotropic δ_{ii} parameters which describe the displacement of the interaction site from the

location of the carbon centre. These six parameters are fully transferable to all $-CH_3$ and $-CH_2-$ groups.

3) The intermolecular parameters were regressed to reproduce the pure component liquid densities, the heats of vaporization and the saturated vapour pressures.

The AUA4 *n*-alkane potential of Ungerer *et al.* (2000) was able to achieve good agreement with the experimental liquid densities, the heats of vaporization and the saturated vapour pressures for *n*-alkanes longer than *n*-pentane. The agreement with the critical densities and temperatures was, however, worse than for the NERD or TraPPE force fields and the AUA4 force field did not explicitly consider *n*-alkanes with 4 or fewer carbons in detail.

In light of the above discussion, the general approach to the formulation of the Transferability Based On Relaxing Geometries (TBORG) force field described below was as follows:

1) All bond lengths were taken to be fixed. While the NERD force field did achieve an improved description of the pure component VLE compared to the TraPPE force field, it is unlikely that this improvement was exclusively due to flexibility in the bond lengths. This improvement might have been the result of the larger set of regressed LJ parameters. Chapter Seven revealed that the pure component properties are more sensitive to the equilibrium geometry than the magnitude of the intramolecular potential. What is clear is that the nAEP and AUA4 *n*-alkane force fields were both able to achieve a good description of the pure component VLE while employing fixed bond lengths.
2) Different bond lengths were employed for bonds connecting distinct pairs of functional groups. Two functional groups were regarded as distinct from one another if they had different numbers of hydrogen atoms or different hybridizations. This marks a break from the NERD and TraPPE approach where different bond lengths were only employed to distinguish between single and double bonds. The AUA4 force field effectively amounts to a relaxing of the fully transferable bond length constraint because of the use of the anisotropic δ_{ii} parameter while the nAEP force field used three different bond lengths for the three possible connectivities of $-CH_3$ and $-CH_2-$ functional groups.
3) Flexible bond angles were employed. Bond angles were regarded as flexible because the alkene force field of Bourasseau *et al.* (2003) highlighted the lack of flexible bond angles as potentially the reason behind their poorer description of *trans*- and *cis*-2-butene.
4) As for the bond lengths, the constraint of full transferability of bond angles was relaxed. Bond angles were allowed to assume different equilibrium values for bond angles including three distinct functional groups. Again, this is effectively applied in the AUA4

force field because of the displacement of the interaction site from the carbon centre, but this approach is different to that in the TraPPE, NERD and nAEP force fields.

5) The isotropic LJ potential is used with distinct LJ ε_{ii} and σ_{ii} parameters for the required functional groups $-CH_3$, $-CH_2-$, $-CH=$, and $=CH_2$ for a total of 8 LJ parameters. These LJ parameters were, however, assumed to be fully transferable regardless of the connectivities of the functional groups. The flexibility in the internal equilibrium geometries is assumed to more accurately account for the effects of anisotropy introduced by the united-atom approximation.

9.2 Regression Methodology

The 8 required LJ parameters as well as the equilibrium geometries for the TBORG force field were regressed to obtain the closest possible fit to the following properties in the data compilations of Smith and Srivastava (1986):

- The liquid and vapour coexistence densities (ρ_l and ρ_g),
- The natural logarithm of the saturated vapour pressures ($\ln\left(P^{sat}\right)$), and
- The critical temperatures and the critical densities (ρ_{cr} and T_{cr}).

The following families of *n*-alkanes and alkenes were used to regress the 8 required LJ parameters to provide a good fit to the data listed above:

Table 9-1: Components used in the regression of the TBORG force field.

***n*-alkanes**	**alkenes**
ethane	ethene
propane	propene
n-butane	1-butene
n-octane	*cis*-2-butene
	trans-2-butene
	1,3-butadiene
	1-octene

Although several interesting observations were made in Chapter Seven concerning *cis*-1,3-pentadiene and *cis*-1,3,5-hexatriene, these were not included due the lack of measured data. It was

suspected that these might then bias the final parameters regressed. This was also the reason for truncating the maximum chain length used in the regression to 8 carbons because a complete set of the coexistence liquid and vapour densities, the critical temperature, the critical density and the saturated vapour pressures are, in general, unavailable for alkenes and *n*-alkanes with more than 9 carbons. This limited regression set was also used by the TraPPE and NERD force fields.

The intended flexibility in the intramolecular geometries necessitated that a more rigorous method of parameter optimization be employed than the trial search methods used by the NERD and TraPPE force fields. The computationally expensive method of Hamiltonian Scaling Monte Carlo employed by the Be6 force field was also not utilized. Consequently, to define an objective function, F_{obj}, required for the regression, the following formulation of Ungerer *et al.* (2000) was used:

$$F_{obj} = \frac{1}{\eta}\sum_{i=1}^{\eta}\frac{\left(J_i - J_{i,exp}\right)^2}{s_i^2} \qquad (9\text{–}1)$$

where s_i is the estimated statistical uncertainty in the measured simulated property J_i, $J_{i,exp}$ is the associated experimental value of the property under consideration, and η is the total number of measured properties (i.e. T_{cr} and ρ_{cr}, and all the discrete values for ρ_l, ρ_g and $\ln\left(P^{sat}\right)$ selected). The estimated statistical uncertainty was used in the denominator of Equation (9–1) in place of $J_{i,exp}$ because trial regressions revealed that the higher magnitudes of the liquid densities in the lower temperature regions tended to bias the final regressed parameter values. This proved unsatisfactory because the low temperature region also has the highest uncertainty in the simulated ρ_g and P^{sat} values. This is because the low acceptance rate of swap moves at low temperatures (high liquid densities) causes long period fluctuations in averaged values for ρ_g and P^{sat}.

The regression was performed sequentially in the sense that the intramolecular parameters and the LJ parameters for the sp^3 hybridized $-CH_2-$ and $-CH_3$ groups for the *n*-alkanes were regressed first. Thereafter, the intramolecular parameters and the LJ parameters for the $=CH-$ and $=CH_2$ groups for the alkenes were regressed. Bourasseau *et al.* (2003) conducted an extensive regression in which a sequential approach similar to that listed above was compared to the approach of using a single global optimization for all required parameters. The single global optimization approach is advantageous because it avoids the propagation of subtle errors in the regression of parameters to additional homologous series because only a single optimization is performed. The single global

optimization method was, unfortunately, found to be intractable due to the much larger set of intramolecular equilibrium geometries employed. There were too many parameters that would have to have been regressed simultaneously. For the *n*-alkanes, F_{obj} in Equation (9–1) was therefore a function of the parameters listed in Table 9-2 while for the alkenes, F_{obj} in Equation (9–1) was considered to be a function of the parameters in Table 9-3.

A total of 14 intra- and intermolecular parameters ($\lambda = 14$) must be regressed for the TBORG *n*-alkane force field while a further 25 intra- and intermolecular parameters ($\lambda = 25$) are required to simulate alkenes for the TBORG force field. If ν_j denotes parameter "j" to be regressed, then the set of all ν_j's given by $\mathbf{v}_\lambda = (\nu_1, \nu_2, ..., \nu_\lambda)$ corresponding to the optimal parameter set will be achieved when every partial derivative $\partial F_{obj}/\partial \nu_j$ of Equation (9–1) is set to zero:

$$\frac{\partial F_{obj}}{\partial \nu_j} = \frac{1}{\eta}\sum_{i=1}^{\eta} \frac{2\left(J_i - J_{i,exp}\right)}{s_i^2}\frac{\partial J_i}{\partial \nu_j} = 0 \qquad (9\text{–}2)$$

In the above equation, J_i was estimated by using a second order expansion of J_i about a central value $\mathbf{v}_\lambda^0$ as:

$$J_i\left(\mathbf{v}_\lambda^0 + \Delta\mathbf{v}_\lambda\right) = J_i\left(\mathbf{v}_\lambda^0\right) + \sum_{j=1}^{\lambda}\frac{\partial J_i\left(\mathbf{v}_\lambda^0\right)}{\partial \nu_j}\Delta\nu_j + \frac{1}{2}\sum_{j=1}^{\lambda}\frac{\partial^2 J_i\left(\mathbf{v}_\lambda^0\right)}{\partial \nu_j^2}\Delta\nu_j^2 \qquad (9\text{–}3)$$

Substituting Equation (9–3) into Equation (9–2), the condition to be satisfied for the optimal parameter set can therefore be expressed as:

$$\sum_{i=1}^{\eta}\frac{\left(J_i\left(\mathbf{v}_\lambda^0\right) - J_{i,exp} + \sum_{j=1}^{\lambda}\frac{\partial J_i\left(\mathbf{v}_\lambda^0\right)}{\partial \nu_j}\Delta\nu_j + \frac{1}{2}\sum_{j=1}^{\lambda}\frac{\partial^2 J_i\left(\mathbf{v}_\lambda^0\right)}{\partial \nu_j^2}\Delta\nu_j^2\right)\frac{\partial J_i}{\partial \nu_j}}{s_i^2} = 0 \qquad (9\text{–}4)$$

for each and every ν_j. For Equation (9–4) to be useful, $J_i\left(\mathbf{v}_\lambda^0\right)$, $\partial J_i\left(\mathbf{v}_\lambda^0\right)/\partial \nu_j$, $\partial^2 J_i\left(\mathbf{v}_\lambda^0\right)/\partial \nu_j^2$, $J_i\left(\mathbf{v}_j\right)$, and $\partial J_i\left(\mathbf{v}_j\right)/\partial \nu_j$ are required. Consequently, $J_i\left(\mathbf{v}_\lambda^0\right)$ was evaluated by conducting a "base simulation" with "central values" for $\mathbf{v}_\lambda = \mathbf{v}_\lambda^0$. In addition, "central values" for $\rho_l(T)$, $\rho_g(T)$, $P^{sat}(T)$ and $\Delta H_{vap}(T)$ for a range of discrete temperature values were stored as well as the estimated values for T_{cr} and ρ_{cr} from Equations (7–2) and (7–3) with $\beta = 0.32$.

Table 9-2: List of model parameters optimized for the TBORG *n*-alkane force field.

Lennard-Jones non-bonded interactions	
$\sigma_{CH_3(sp^3)}$, $\varepsilon_{CH_3(sp^3)}$, $\sigma_{CH_2(sp^3)}$, $\varepsilon_{CH_2(sp^3)}$	4
Bond Lengths	
CH_3-CH_3 (ethane), CH_3-CH_2 (propane), CH_3-CH_2 (butane), CH_3-CH_2 (octane) CH_2-CH_2 (butane), CH_2-CH_2 (octane)	6
Bond Angles	
$CH_3-CH_2-CH_3$ (propane), $CH_3-CH_2-CH_2$ (butane), $CH_3-CH_2-CH_2$ (octane) $CH_2-CH_2-CH_2$ (octane)	4

Table 9-3: List of model parameters optimized for the TBORG 1-alkene force field.

Lennard-Jones non-bonded interactions	
$\sigma_{CH_2(sp^2)}$, $\varepsilon_{CH_2(sp^2)}$, $\sigma_{CH(sp^2)}$, $\varepsilon_{CH(sp^2)}$	4
Bond Lengths	
$CH_2=CH_2$ (ethene), $CH_2=CH$ (1-propene), $CH_2=CH$ (1-butene), $CH_2=CH$ (1,3-butadiene), $CH_2=CH$ (1-octene), $CH=CH$ (*cis*-2-butene), $CH=CH$ (*trans*-2-butene), $CH-CH_3$ (1-propene), $CH-CH_2$ (1-butene), $CH-CH$ (1,3-butadiene), $CH-CH_2$ (1-octene), CH_3-CH (*cis*-2-butene), CH_3-CH (*trans*-2-butene)	13
Bond Angles	
$CH_2=CH-CH_3$ (propene), $CH_2=CH-CH_2$ (1-butene), $CH_2=CH-CH$ (1,3-butadiene), $CH_2=CH-CH_2$ (1-octene), $CH_3-CH=CH$ (*cis*-2-butene), $CH_3-CH=CH$ (*trans*-2-butene), $CH-CH_2-CH_3$ (1-butene), $CH-CH_2-CH_2$ (1-octene)	8

To determine $J_i(\mathbf{v}_\lambda)$, each of the v_j's were systematically varied to consider values for each of the v_j's above and below the central v_j^0 values. Once these simulations were completed, elementary second order polynomials of the following form were regressed to the simulated data:

$$J_i(\mathbf{v}_\lambda^0 + \Delta v_j) = K_1(i,T) + K_2(i,T)v_j + K_3(i,T)v_j^2 \qquad (9\text{–}5)$$

where the constants $K_1(i,T)$, $K_2(i,T)$ and $K_2(i,T)$ were functions of:

1) the particular property considered, i, and
2) the temperature at which that property was simulated, T.

For the saturated vapour pressure, however, $\ln(P^{sat}(1/T))$ was used instead of $P^{sat}(T)$ because this resulted in a nearly linear relationship between T and P^{sat}. $J_i(\mathbf{v}_\lambda^0)$, $\partial J_i(\mathbf{v}_\lambda^0)/\partial v_j$, $\partial^2 J_i(\mathbf{v}_\lambda^0)/\partial v_j^2$, $J_i(\mathbf{v}_j)$, and $\partial J_i(\mathbf{v}_j)/\partial v_j$ could all be calculated from evaluation of Equation (9–5) or suitable differentiation of Equation (9–5). This approach was expensive in terms of the number of simulations required but was essential in elucidating the dependence of physical properties on the varied parameters and the dependence of this variation with temperature as well.

Equation (9–4) may now be solved for the set of all Δv_j's. The optimized parameters may then each be calculated as:

$$v_j = v_j^0 + \Delta v_j \qquad (9\text{–}6)$$

In contrast, Ungerer *et al.* (2000) used a first order Taylor series expansion of the force field of Toxvaerd (1997) to approximate $J_i(\mathbf{v}_\lambda^0 + \Delta\mathbf{v}_\lambda)$ (i.e. Equation (9–3) without the third term on the right hand side of the equation), while Bourasseau *et al.* (2003) used statistical fluctuations (which implicitly only include the first order derivatives in Equation (9–3)) obtained during their regression simulations to evaluate the partial derivatives. The approach of Ungerer *et al.* (2000) could not be used because trial regressions employing Equation (9–4) revealed that the optimum geometries were very different to the values used in the TraPPE, NERD and Be6 united-atom potentials. In addition, trial regressions based on optimizing the properties of ethane with and without the second order derivatives in Equation (9–3) revealed that the inclusion of the second order derivatives provided regressions that were more accurate and more stable. This was the case because ρ_l, ρ_g, P^{sat} and ΔH_{vap} were all later found to have non-linear dependencies on the bond angles, and ρ_g in particular tended to have non-linear dependencies on most of the regressed parameters.

It should be noted that Equation (9–3) effectively amounts to a second order Taylor series which excludes the second order cross derivatives of the form $\partial^2 J_i\left(\mathbf{v}_\lambda^0\right)/\partial v_j \partial v_{k\neq j}$. The second order cross-derivatives were excluded because the large number of parameters considered would have created the requirement for an additional large number of simulations to determine the functional forms of these cross derivatives. The exclusion of the second order cross derivatives was, however, found to introduce negligible error. This conclusion was based on simulations conducted once the TBORG model parameters were optimized. It was found that the properties predicted from these "test" simulations matched the corresponding values derived from the regression simulations within the statistical uncertainty of the results. This, therefore, confirmed the validity of Equation (9–3).

9.3 Simulation Methodology

All of the regression simulations were conducted using the CBMC method in conjunction with the NVT version of the Gibbs Ensemble. The total numbers of molecules, the absolute temperature and reduced temperature ($T_r = T/T_{cr}$) ranges for each type of molecule as well as the number of CBMC trial directions when "growing" molecules used for the *n*-alkane and alkene parameter regressions are listed in Table 9-4.

Each molecule was simulated at six different temperatures spanning the range of temperatures listed in Table 9-4. Only for *cis*-2-butene was the selection of the temperature range severely constrained by the limited available data (Smith and Srivastava (1986), Vargaftik (1975)). For the rest of the molecules, the temperature range was selected to allow as wide a temperature range as possible to ensure maximum applicability of the final regressed set of parameters. The lower reduced temperature limit imposed on the temperature range of around 0.5 was chosen because it was found that below this reduced temperature, the number of accepted swap moves was so prohibitively low as to make equilibrium virtually impossible to achieve within a reasonable number of CBMC trial directions or within a reasonable number of cycles. The low temperature number of CBMC trial directions listed in Table 9-4 reflects the number of CBMC trial directions required to achieve a minimum acceptance rate of 0.1% for swap moves. It was found that 0.1% was the lowest allowable acceptance for swap moves which could be tolerated without significantly increasing the required number of equilibration cycles. The reason for the higher minimum reduced temperature for longer molecules was to compensate for the greater difficulty experienced in inserting these molecules, particularly in dense phases.

Table 9-4: Simulation details for each of the molecules considered in the TBORG parameter regressions.

Molecule		Temperature Range		CBMC trial directions	
Species	**Number**	**Absolute [K]**	**Reduced**	**Low Temp**	**High Temp**
ethane	500	125 – 275	0.41 – 0.90	100	10
propane	400	173 – 331	0.47 – 0.90	50	4
n-butane	350	210 – 386	0.49 – 0.91	50	4
n-octane	200	320 – 520	0.56 – 0.91	50	4
ethene	500	136 – 253	0.48 – 0.90	50	4
1-propene	400	176 – 327	0.48 – 0.89	50	4
1-butene	350	220 – 383	0.52 – 0.91	50	4
cis-2-butene	350	225 – 302	0.51 – 0.69	50	4
trans-2-butene	350	226 – 377	0.53 – 0.88	50	4
1,3-butadiene	350	216 – 383	0.51 – 0.90	50	4
1-octene	200	311 – 518	0.55 – 0.91	50	4

Only for ethane was a reduced temperature as low as 0.41 explored which required 100 CBMC trial directions to achieve a 0.1% swap move acceptance. The reason for this was to ensure that the LJ parameters for the $-CH_3$ functional group included data from high density regions because the effect of the $-CH_3$ group constitutes a smaller and smaller fraction of the intermolecular potential as chain length increases. The high density data for ethane would, therefore, serve to better guarantee the correctness of the $-CH_3$ LJ parameters by forcing them to better fit the ethane properties over a wide temperature range. The higher reduced temperature of about 0.9 was imposed to prevent the simulations from becoming unstable. If the simulations were conducted at a temperature much above this reduced temperature, the well-known instability of the Gibbs Ensemble near the critical point (Panagiotopoulos (2000)) caused the simulation box volumes to change phase identity during a simulation run thereby compromising the collection of ensemble averages for distinct phases.

As in the previous Chapters, the box volumes were adjusted so that the liquid and vapour box volumes were approximately equal in size at the end of a simulation. The same five distinct types of NVT moves listed previously were performed, namely (1) volume changes, (2) transfers of molecules between simulation boxes, (3) full regrowing of chains and partial regrowing of chains using CBMC, (4) translation of the COM, and (5) rotation around the COM. The maximum volume, translation and rotational displacements were chosen in such a way that roughly 50% of

each of these moves were accepted. For the four highest temperatures for each type of molecule, the moves were performed with the fixed probabilities: p1: p2: p3: p4: p5 = 0.006: 0.328: 0.222: 0.222: 0.222. For the two lowest temperatures considered for each molecule, the fixed probabilities for each of the moves were adjusted to favour an increased number of molecule insertion attempts per cycle (because of higher liquid densities) as p1: p2: p3: p4: p5 = 0.006: 0.500: 0.165: 0.165: 0.164. Trial simulations consisting of no more than 10,000 equilibration and 20,000 production cycles were used initially to identify approximate values for the model parameters around which to conduct further regressions. During the final regression simulations, the equilibration cycles lasted for at least 80,000 cycles while production runs included at least 100,000 cycles for averaging. The standard deviations of ensemble averages were computed by dividing the production cycles of each run into five blocks and calculating the standard deviation from the averages of these five blocks.

For each of the molecules, six base simulations corresponding to six different temperatures within the ranges listed in Table 9-4 were conducted using "base" parameter values. Thereafter, the phase envelope of each molecule was resimulated at all 6 temperatures while systematically varying each of the parameters which could affect the properties of a species under consideration. For each parameter considered, revised phase envelopes were simulated for three values below and for three values above the base value of each parameter while keeping the other relevant parameters at their base values. Table 9-5 lists the number of simulations conducted for each species considered:

Table 9-5: Details of the parameters varied for each component. In general, the number of simulations required for each molecule is $(6\times\nu_i)+1$.

Species	No. Distinct Bond Lengths	No. Distinct Bond Angles	No. LJ parameters ε_{ii} and σ_{ii}	Total No. Simulations
ethane	1	0	2	19
propane	1	1	4	37
n-butane	2	1	4	43
n-octane	2	2	4	49
ethene	1	0	2	19
1-propene	2	1	4	43
1-butene	2	2	4	49
cis-2-butene	2	2	4	49
trans-2-butene	2	2	4	49
1,3-butadiene	2	1	4	43
1-octene	2	2	4	49

Note that the sp^3 hybridized $-CH_3$ and $-CH_2-$ groups were not considered as variables for the alkenes nor were the CH_3-CH_2 bond, the CH_2-CH_2 bond, the $CH_3-CH_2-CH_2$ bend or the $CH_2-CH_2-CH_2$ bond angle because these were resolved from the *n*-alkane regression alone.

In Chapter Seven it was concluded that the equilibrium geometries played a more significant role in the determination of the physical properties than the magnitude of the intramolecular interactions. For this reason, a bond-bending constant of $k_\phi = 62500\text{K/deg}^2$ was used for all bond angles in the TBORG *n*-alkane and alkene force fields. This is similar to the values used by the NERD and TraPPE force fields (see Table 6-3). The only intramolecular parameters not subject to optimization were the torsional potentials and the bond bending constants. The TraPPE torsional parameters applicable for the $-CH_2-CH_2-$ and $=CH-CH_2-$ torsions in Table 6-4 were used to describe the corresponding torsions for the TBORG force field. The torsional potential of Bock *et al.* (1979) as described in Table 7-2 was used to describe the $=CH-CH=$ torsion for 1,3-butadiene while the harmonic torsional potential of Wick *et al.* (2000) for *cis* and *trans* torsions was used for the torsion in *cis*- and *trans*-2-butene. The parameters for the *cis* torsion $-CH=CH-$ are listed in Table 7-2 while the *trans* harmonic torsional parameters for Equation (7–1) are taken to be $\phi_0 = 180°$ and $k_\phi = 13400\text{K/deg}^2$.

For interactions between functional groups of different types, the Lorentz-Berthelot combining rules of Equation (6–1) for the TBORG force field to evaluate ε_{ij} and σ_{ij} were used. As for the TraPPE force field, non-bonded interactions were evaluated up to a cut-off radius of 14Å. The Lennard-Jones potential was used to describe intramolecular interactions between functional groups separated by more than 3 bonds on the same molecule.

9.4 Regression of sp^3 Hybridized $-CH_3$ and $-CH_2-$ Lennard-Jones Parameters for *n*-Alkanes

Figure 9-1 and Figure 9-2 demonstrate the need to use a second order polynomial and second order finite differences to properly describe the effect of varying the model parameters. In Figure 9-1, the non-linear dependence of the vapour density on the CH_3-CH_2 equilibrium bond length for all temperatures considered in the regression may be readily seen. Typically, the vapour density exhibited a non-linear dependence on all TBORG model parameters at all temperatures. Figure 9-2

shows the non-linear dependence of the liquid density on the equilibrium $CH_3-CH_2-CH_2$ bond angle. While the liquid density tended to be an approximately linear function for most other TBORG model parameters, changes in the equilibrium bond angle tended to cause the vapour density, the latent heat of vaporization and the saturated vapour pressure to vary non-linearly.

It is important to understand the dependence of the simulated physical properties on the variation of the TBORG model parameters. There are four basic groups of model parameters which can be optimized for the TBORG force field namely bond lengths, bond angles, LJ σ_{ii} parameters and LJ ε_{ii} parameters. The bond lengths and bond angles may be conveniently understood as affecting the macroscopic properties through steric or molecular geometry influences on each molecule. Alternatively, the LJ σ_{ii} and ε_{ii} parameters will affect the macroscopic properties through energetic attractions and repulsions between molecules in the bulk phase. In Figure 9-3 to Figure 9-18 the dependence of the liquid density, vapour density, saturated vapour pressure and the latent heat of vaporization of *n*-butane have been graphed to illustrate these dependencies. *n*-Butane was selected because it has at least one example of the four basic TBORG parameter groups.

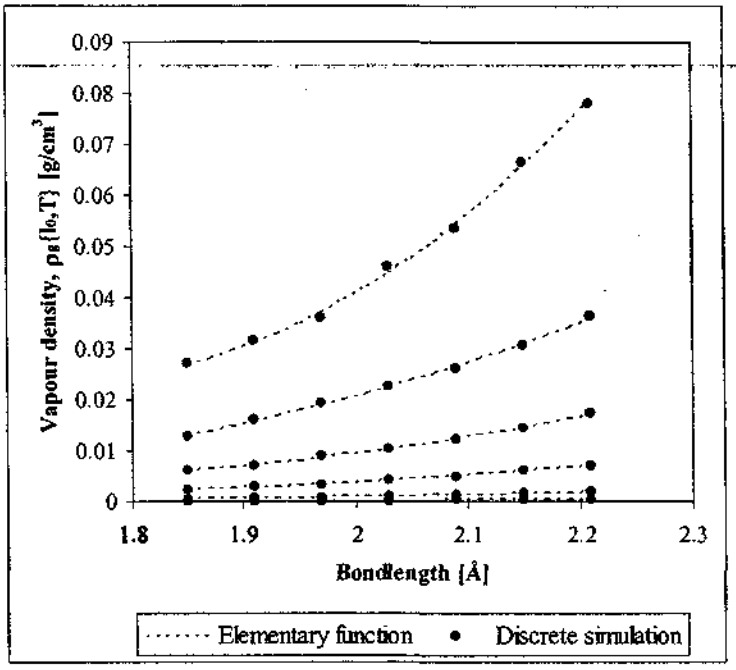


Figure 9-1: Variation in the vapour density of *n*-butane when varying the CH_3-CH_2 equilibrium bond length. Each parametric line represents the liquid density at a different temperature.

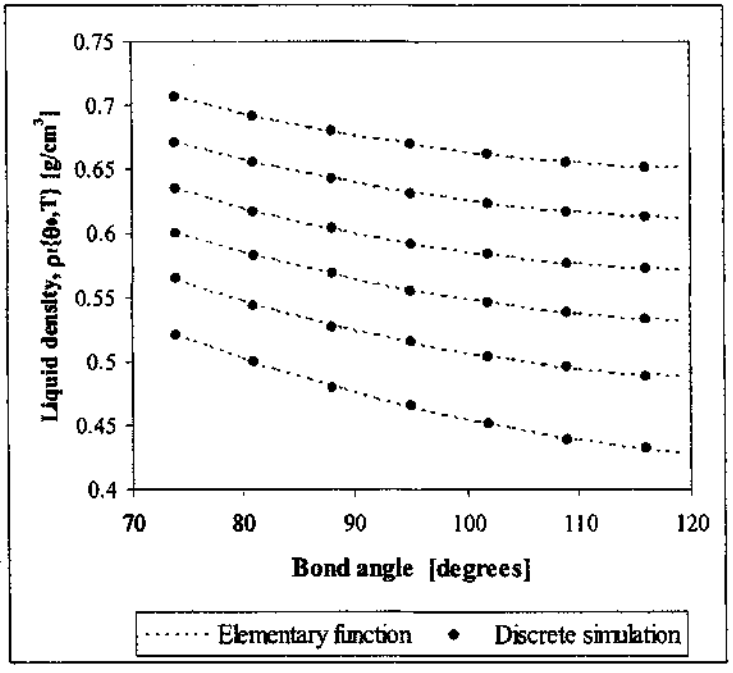


Figure 9-2: Variation in the liquid density of *n*-butane when varying the $CH_3-CH_2-CH_2$ equilibrium bond angle. Each parametric line represents the liquid density at a different temperature.

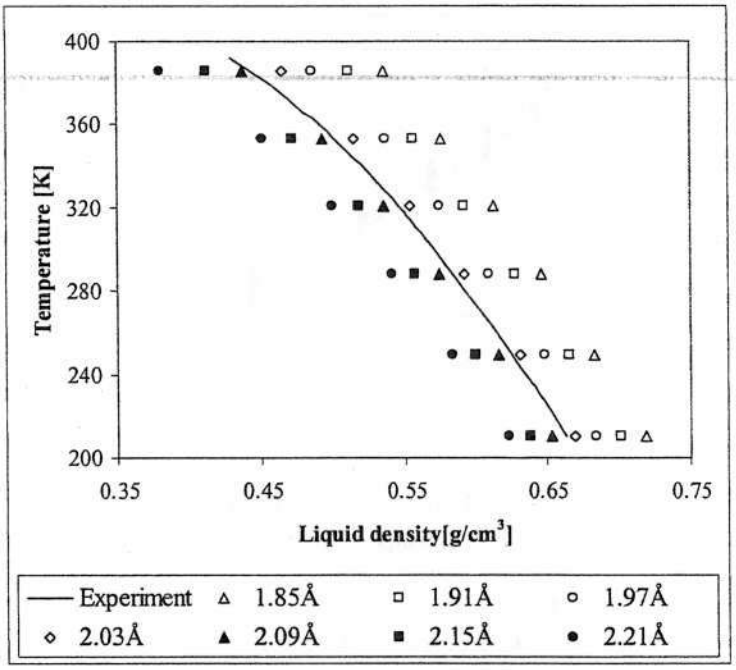


Figure 9-3: Effect of varying the CH_3-CH_2 bond length on the liquid density of *n*-butane.

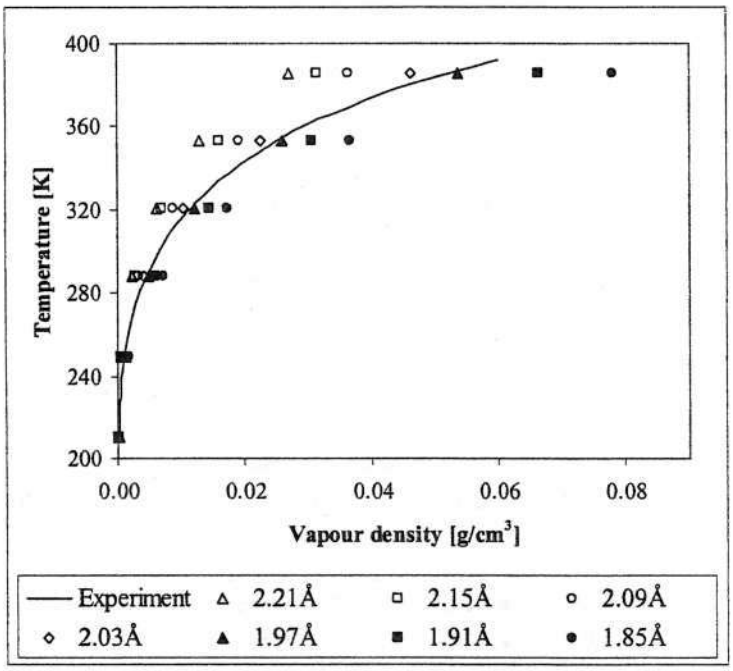


Figure 9-4: Effect of varying the CH_3-CH_2 bond length on the vapour density of *n*-butane.

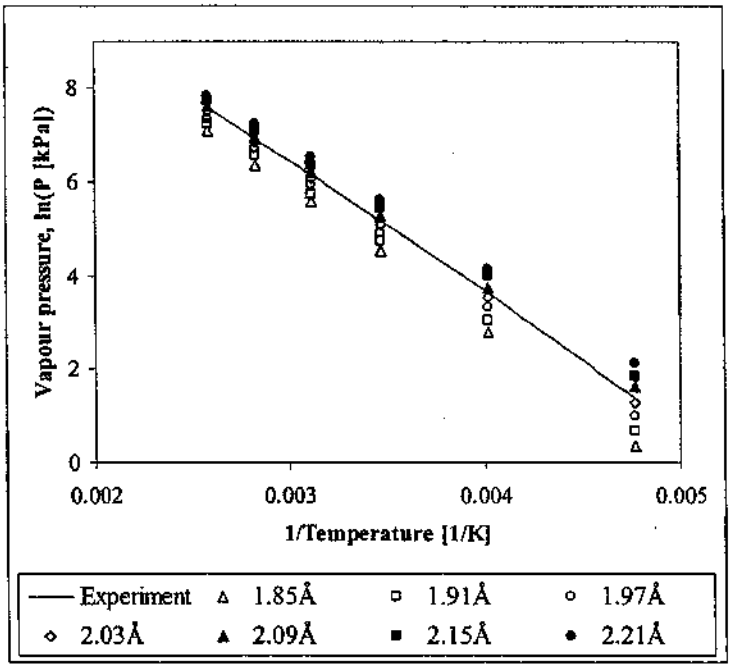


Figure 9-5: Effect of varying the CH_3-CH_2 bond length on the saturated vapour pressure of *n*-butane.

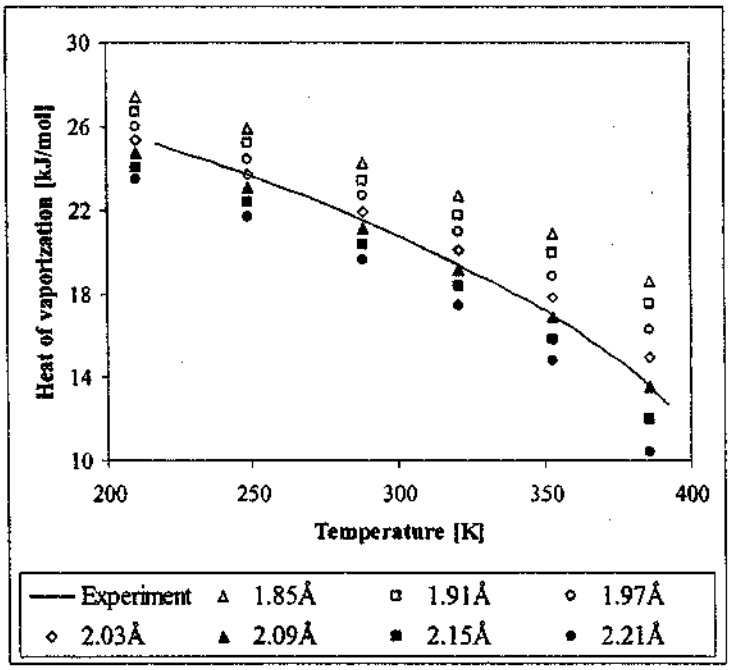


Figure 9-6: Effect of varying the CH_3-CH_2 bond length on the latent heat of vaporization of *n*-butane.

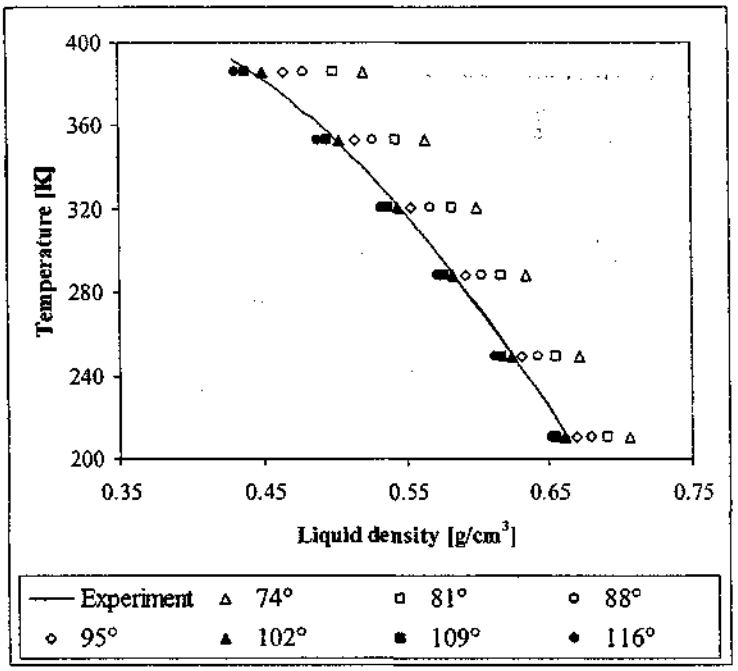


Figure 9-7: Effect of varying the $CH_3-CH_2-CH_2$ bond angle on the liquid density of *n*-butane.

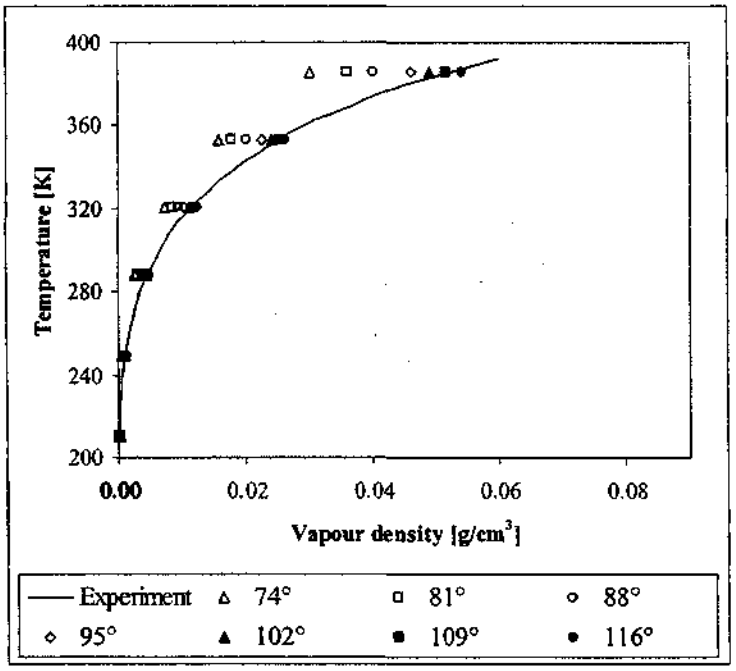


Figure 9-8: Effect of varying the $CH_3-CH_2-CH_2$ bond angle on the vapour density of *n*-butane.

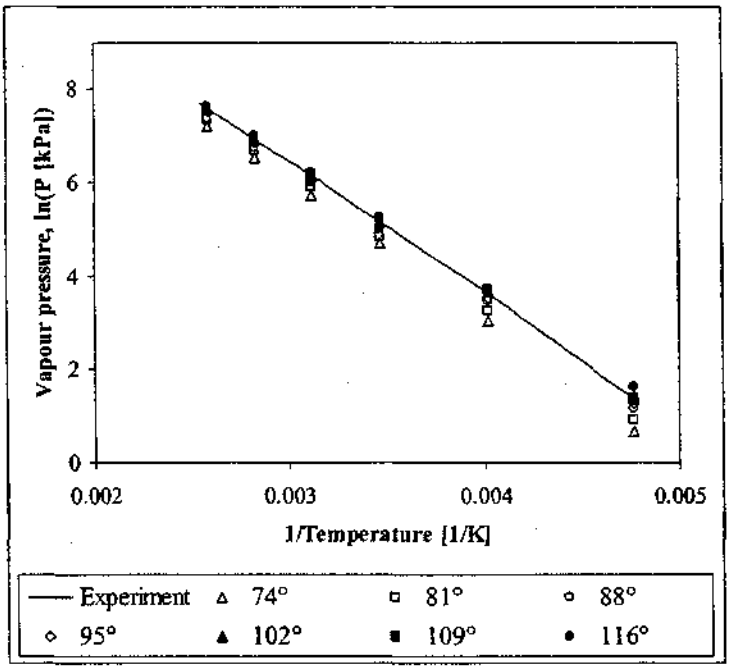


Figure 9-9: Effect of varying the $CH_3-CH_2-CH_3$ bond angle on the saturated vapour pressure of *n*-butane.

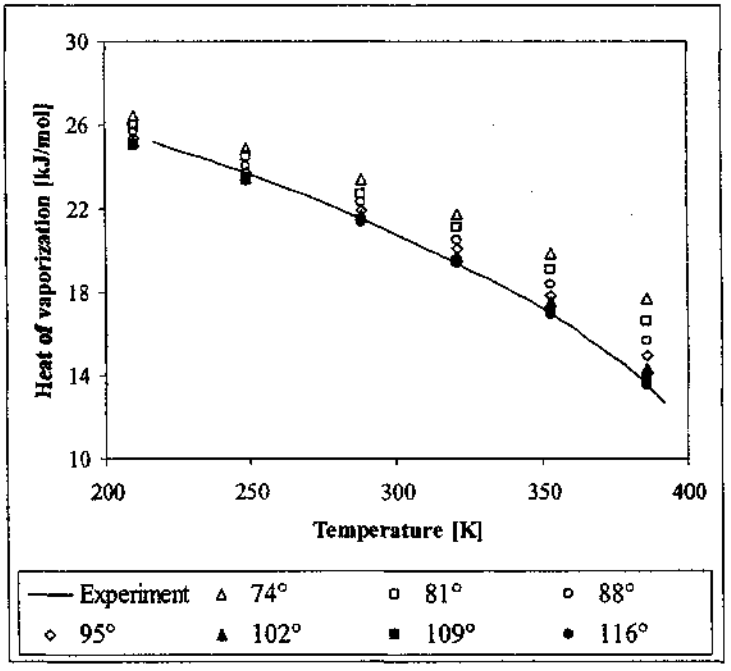


Figure 9-10: Effect of varying the $CH_3-CH_2-CH_3$ bond angle on the latent heat of vaporization of *n*-butane.

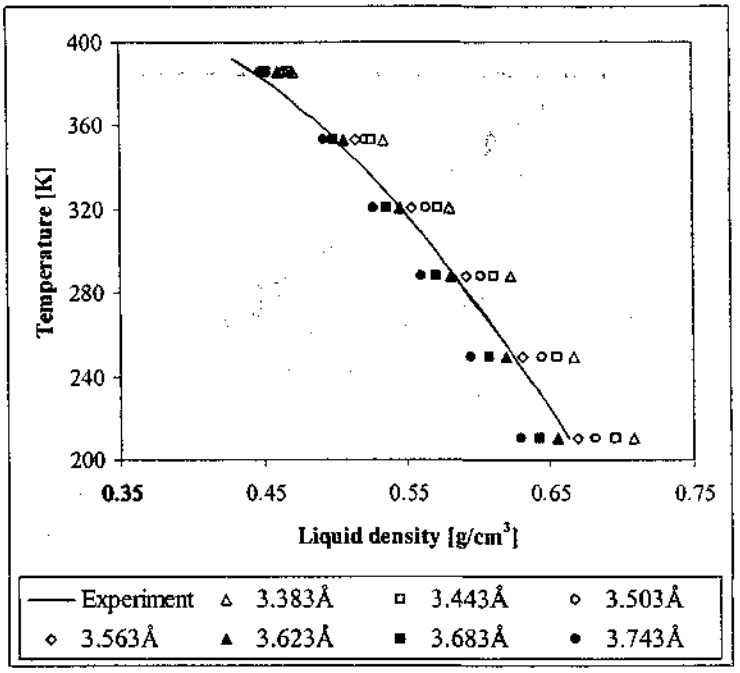


Figure 9-11: Effect of varying the Lennard-Jones σ_{ii} parameter for the $-CH_3$ group on the liquid density of *n*-butane.

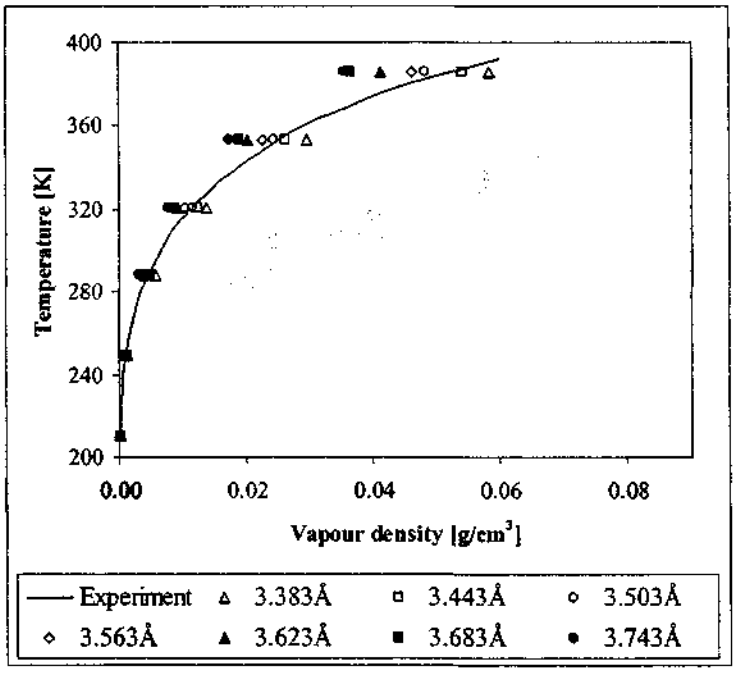


Figure 9-12: Effect of varying the Lennard-Jones σ_{ii} parameter for the $-CH_3$ group on the vapour density of *n*-butane.

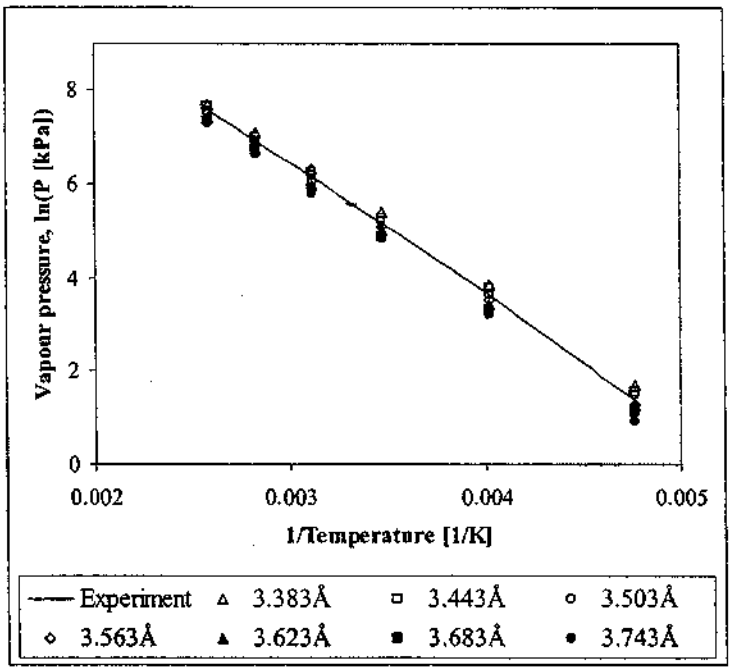


Figure 9-13: Effect of varying the Lennard-Jones σ_{ii} parameter for the $-CH_3$ group on the saturated vapour pressure of *n*-butane.

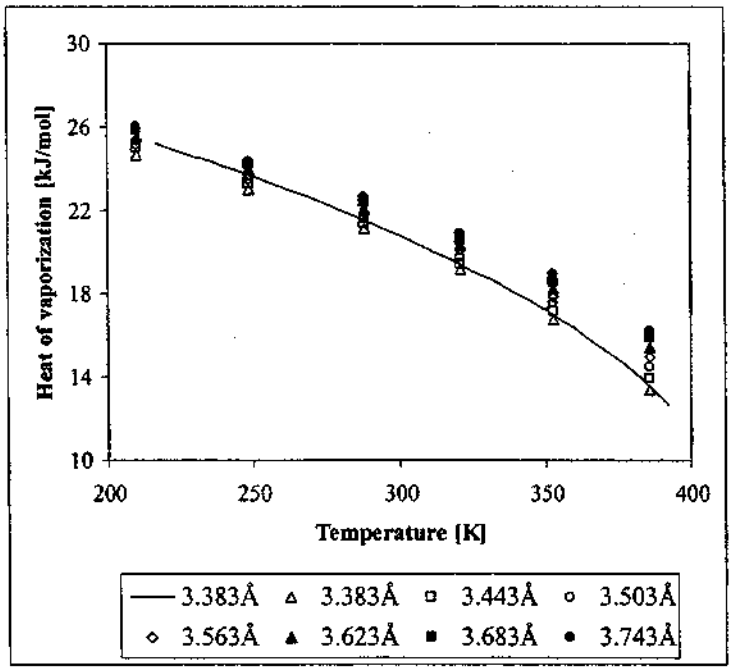


Figure 9-14: Effect of varying the Lennard-Jones σ_{ii} parameter for the $-CH_3$ group on the latent heat of vaporization of *n*-butane.

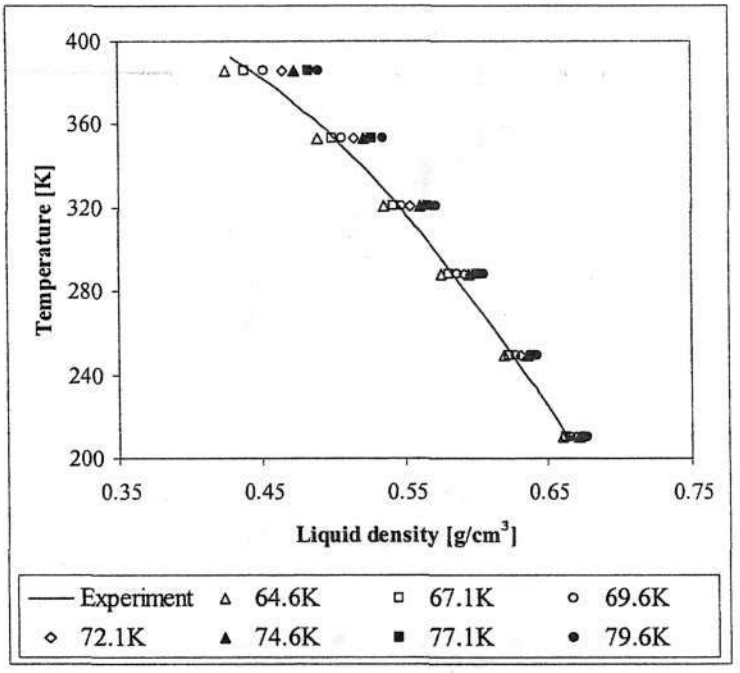


Figure 9-15: Effect of varying the Lennard-Jones ε_{ii} parameter for the $-CH_2-$ group on the liquid density of *n*-butane.

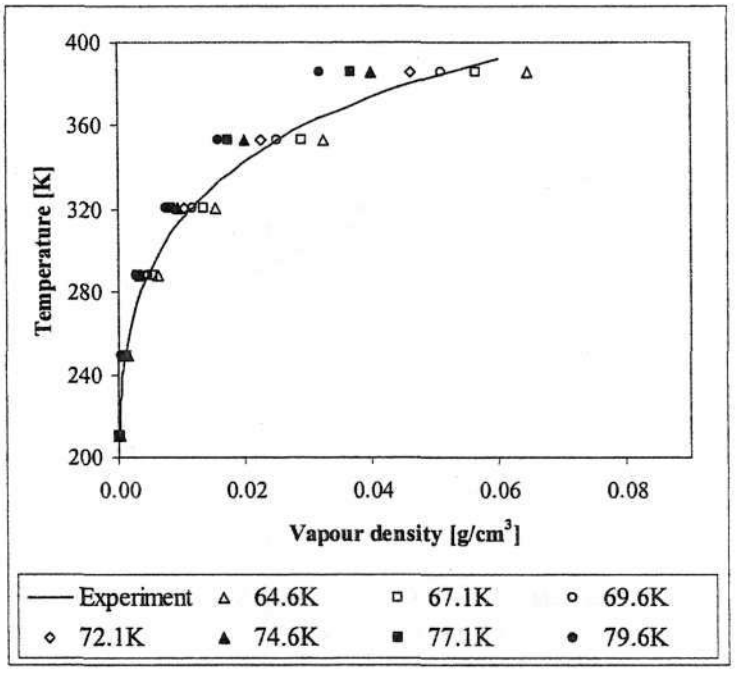


Figure 9-16: Effect of varying the Lennard-Jones ε_{ii} parameter for the $-CH_2-$ group on the vapour density of *n*-butane.

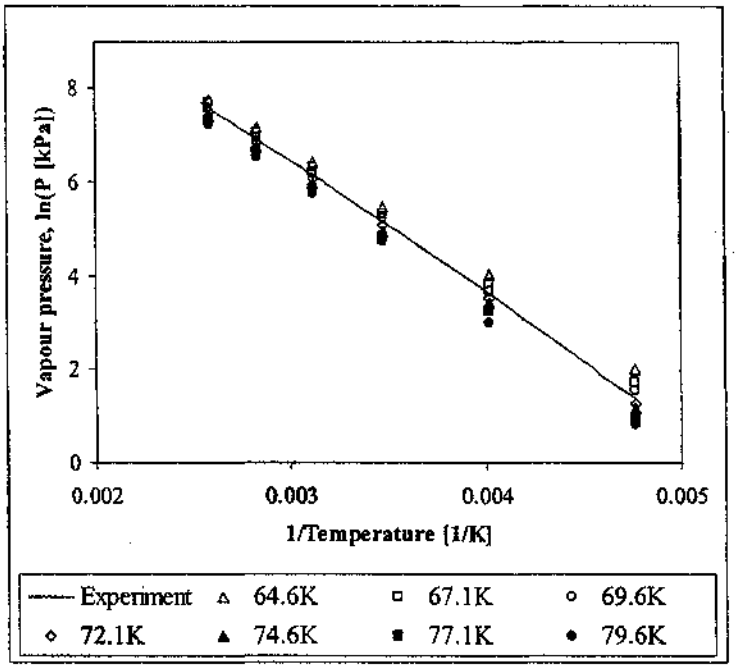


Figure 9-17: Effect of varying the Lennard-Jones ε_{ii} parameter for the $-CH_2-$ group on the saturated vapour pressure of *n*-butane.

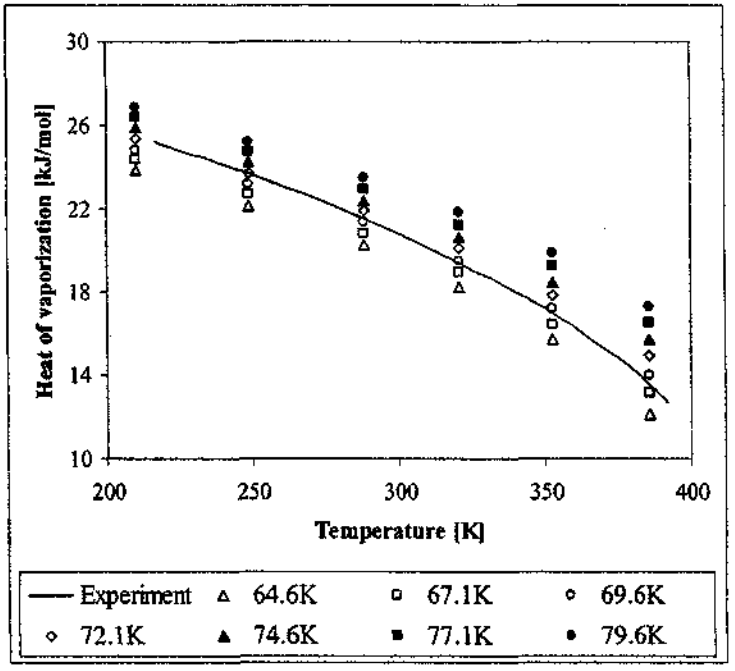


Figure 9-18: Effect of varying the Lennard-Jones ε_{ii} parameter for the $-CH_2-$ group on the latent heat of vaporization of *n*-butane.

9.4.1 Effect of the Equilibrium Bond Length, l_0

The graphs in Figure 9-3 to Figure 9-6 show the effect of varying the CH_3-CH_2 equilibrium bond length on the physical properties of *n*-butane. Figure 9-3 and Figure 9-4 reveal that an increase in the bond length leads to a decrease in the liquid densities and to an increase in the vapour densities. The decrease in the liquid densities can be explained in terms of the larger molecular volume that each molecule will occupy with a larger bond length. In the vapour phase, however, the effect of increasing the bond length is considerably smaller because of the larger distances between the molecules in the lower density phase. The increase in the density of the vapour phase with increasing the bond length is therefore a function of a larger molecular volume being more easily accommodated in a low density vapour phase than in the liquid phase. From Figure 9-4 it is also apparent that the while the bond length has a more or less constant relative shifting effect over the entire liquid density range, the effect on the vapour density is most pronounced in the near-critical region when the molecules in the vapour phase begin to approach higher densities. This combination implies that an increase in the bond length will also decrease the critical temperature. The logarithm of the saturated vapour pressure in Figure 9-5 increases with increasing bond length and the relative increase appears to be slightly larger at the lower temperatures. This effect can be explained in terms of the higher average repulsions between molecules with larger bond lengths. The saturated vapour pressure will, in general, increase with increasing intermolecular repulsions. These increased repulsive interactions will be most significantly felt in the low temperature region with the higher liquid densities. From Figure 9-6, it may be seen that the latent heat of vaporization increases with decreasing bond length. Recalling Equation (5–21), the latent heat of vaporization is a function of the difference between $(U+PV)_g$ for the vapour phase and $(U+PV)_l$ for the liquid phase and describes the enthalpy required to evaporate molecules from the liquid phase. The higher density of the liquid phase (due to lower molar volumes) combined with the lower vapour pressure for short bond lengths will in general require more energy to evaporate molecules in the liquid phase thus leading to the observed trend of an increasing latent heat of vaporization with a decreasing bond length.

9.4.2 Effect of the Equilibrium Bond Angle, θ_0

The graphs in Figure 9-7 to Figure 9-10 show the effect of varying the $CH_3-CH_2-CH_2$ equilibrium bond angle on the physical properties of *n*-butane. A decrease in the

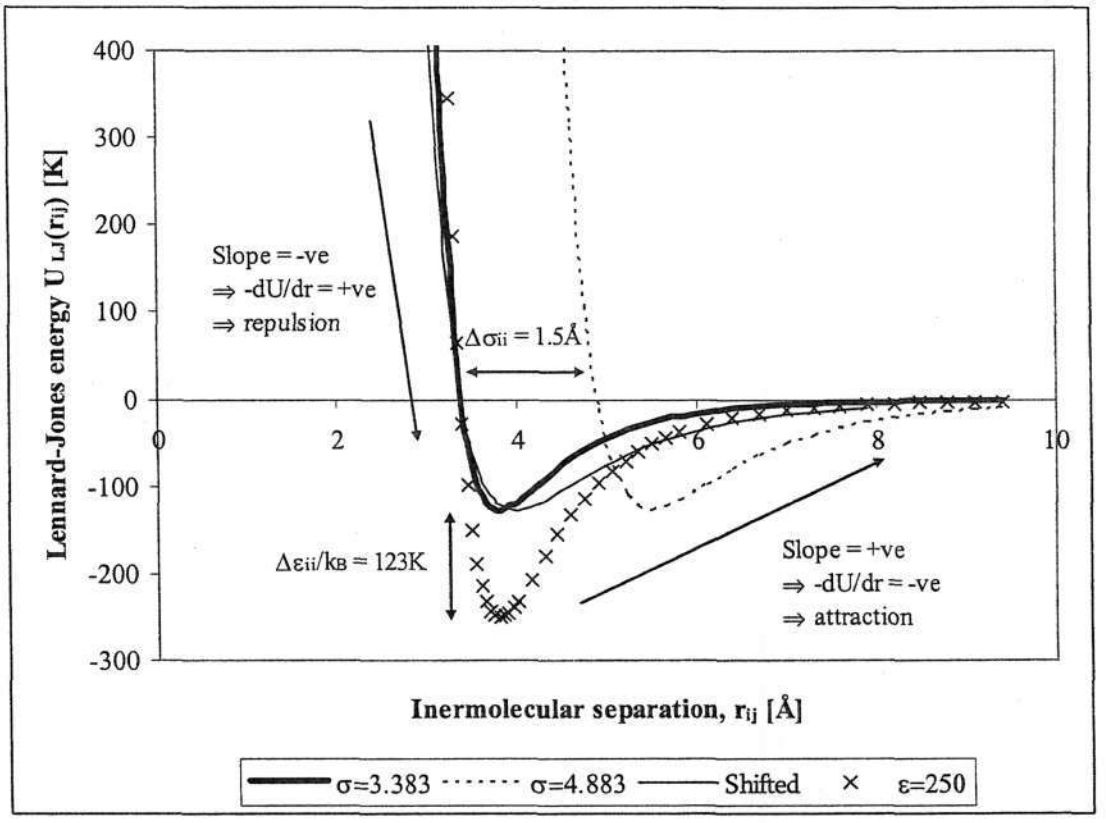


Figure 9-19: Graph showing the effect of the LJ σ_{ii} and ε_{ii} parameters on the shape of the intermolecular potential energy curve. The base curve ($\sigma = 3.383$) has $\sigma_{ii} = 3.383$Å and $\varepsilon_{ii} = 127.1$K. Parametric curves with $\sigma_{ii} = 4.883$Å and $\varepsilon = 250$K are also plotted. The Shifted parametric curve is the $\sigma_{ii} = 3.383$Å parametric curve but shifted horizontally to the left by an amount of 1.5Å.

From Figure 9-13 it may be seen that the saturated vapour pressure decreases with increasing σ_{ii}. In Figure 9-19, two LJ potential curves are plotted with $\sigma_{ii} = 3.383$Å and $\sigma_{ii} = 4.883$Å. A third curve with $\sigma_{ii} = 4.883$Å but shifted to the left by an amount of 1.5Å is also drawn. The force between two molecules is given by $-dU_{LJ}/dr_{ij}$ and is positive for repulsive forces and negative for attractive forces. For the attractive region both the $\sigma_{ii} = 3.383$Å and $\sigma_{ii} = 4.883$Å curves have similar slopes. For the attractive region, however, it may be seen that with the exception of the region immediately to the right of the potential minimum, the slope and hence the attraction between the molecules for the curve with $\sigma_{ii} = 4.883$Å is greater in magnitude than slope for the $\sigma_{ii} = 3.383$Å curve. This result can also be understood by considering the molecular virial contribution to the pressure of Equation (5–14). For the excess contribution to the pressure, it can

be seen that while $w(r_{ij}) \propto r_{ij}$ in Equation (5–16), one also has $P \propto 1/V = 1/r_{ij}^3$ in Equation (5–14). It has already been seen that the phase density decreases with increasing σ_{ii} and hence since P is then proportional to $1/r_{ij}^2$ overall, the saturated vapour pressure decreases with increasing σ_{ii}.

It is also found in Figure 9-14 that the latent heat of vaporization increases with increasing σ_{ii}. This is consistent with the increasing attractive forces between the molecules as σ_{ii} increases which requires more energy to cause a molecule in the liquid phase to be vaporized.

9.4.4 Effect of the Lennard-Jones ε_{ii} Parameter

The graphs in Figure 9-15 to Figure 9-18 demonstrate the effect of varying the LJ ε_{ii} parameter. In terms of Figure 9-19, the effect of increasing the value of ε_{ii} may be understood as increasing the well depth. This can be seen by comparing the base graph with $\varepsilon_{ii} = 127.1\text{K}$ and the parametric curve with $\varepsilon_{ii} = 250\text{K}$. The slope of the repulsive section of the LJ potential remains effectively unchanged but the slope of the attractive section of the potential curve now has a larger magnitude. For this reason, an increase in the LJ ε_{ii} parameter leads to an increase in the liquid density and a corresponding decrease in the vapour density. This implies that the critical temperature will also increase with increasing values of ε_{ii}. An increase in ε_{ii} with its associated increase in attractive forces between molecules in the liquid phase thus also results in a decrease in saturated vapour pressure as seen in Figure 9-17 and an increase in the latent heat of vaporization as seen in Figure 9-18.

Although the above discussion was limited to *n*-butane, the same general trends were observed for all other molecules used in both the *n*-alkane and the alkene parameter regressions. The magnitude of these trends, however, varied from molecule to molecule as the number of instances of a given parameter increased or decreased. In this way, the effect of the LJ σ_{ii} parameter for the $-CH_3$ group and the $CH_3-CH_2-CH_2$ bond angle on the simulated physical properties were more pronounced for *n*-butane than for *n*-octane because of the higher "per molecule concentration" of these parameters for the case of *n*-butane.

9.4.5 Selection of the Final *n*-Alkane Parameter Set

To determine an optimal set of *n*-alkane parameters for the TBORG force field, four sets of optimization constraints were considered in addition to the Final constraint set which was used to optimize the Final set of parameters. Each of these different regression constraint sets resulted in different parameter sets. The TBORG model parameters which resulted from the different sets of constraints are listed in Table 9-6 and their effect on the simulated physical properties are listed in Table 9-7.

Table 9-6: Table of the regressed *n*-alkane TBORG model parameters for the different regression constraints.

	Final	**Case I**	**Case II**	**Case III**	**Case IV**
CH_3-CH_3 (ethane) [Å]	2.207	2.145	2.089	2.206	2.281
CH_3-CH_2 (propane) [Å]	2.045	2.023	1.952	2.045	2.066
CH_3-CH_2 ($\geq C_4$) [Å]	2.045	2.062	1.952	2.045	2.066
CH_2-CH_2 (*n*-butane) [Å]	1.688	1.677	1.683	1.688	1.655
CH_2-CH_2 ($\geq C_5$) [Å]	1.688	1.693	1.683	1.688	1.655
$CH_3-CH_2-CH_3$ (propane) [deg]	84.1	84.6	88.0	84.1	84.4
$CH_3-CH_2-CH_2$ (*n*-butane) [deg]	105.2	98.7	106.9	105.2	110.1
$CH_3-CH_2-CH_2$ ($\geq C_5$) [deg]	105.2	103.3	106.9	105.2	110.1
$CH_2-CH_2-CH_2$ ($\geq C_5$) [deg]	115.4	113.7	108.1	115.4	119.3
$\sigma_{ii}(-CH_2-)$ [Å]	3.698	3.690	3.753	3.698	3.707
$\varepsilon_{ii}/k_B(-CH_2-)$ [K]	66.54	67.67	62.38	66.52	65.14
$\sigma_{ii}(-CH_3)$ [Å]	3.546	3.562	3.580	3.546	3.525
$\varepsilon_{ii}/k_B(-CH_3)$ [K]	130.35	127.36	125.18	130.30	133.62

Table 9-7: Deviation between the simulated and experimental physical properties for the Final parameter set of the TBORG *n*-alkane force field and for each of the variations of the Final set considered. Data in brackets correspond to the bracketed and more limited temperature range listed for each molecule.

	Final	Case I	Case II	Case III	Case IV		Final	Case I	Case II	Case III	Case IV
Ethane: 125-275K (149-275K)						***n*-Butane**: 210-386K (249-386K)					
% Avg. Error ρ_l	0.30	0.33	0.44	0.30	0.15	% Avg. Error ρ_l	0.34	0.52	0.23	0.34	0.19
% Avg. Error ρ_g	8.54 (4.49)	6.54	5.14	8.48	20.97	% Avg. Error ρ_g	7.16 (3.95)	6.91	20.60	7.17	14.46
% Avg. Error P^{sat}	3.46 (0.74)	5.51	3.58	3.49	3.63	% Avg. Error P^{sat}	0.30 (0.14)	1.08	2.37	0.32	1.09
% Avg. Error ΔH_{vap}	1.44	1.80	2.09	1.44	1.45	% Avg. Error ΔH_{vap}	1.81	1.64	1.58	1.81	1.59
$\lvert\rho_{cc} - \rho_{c,exp}\rvert$ [g/cm^3]	0.0011	0.0010	0.0015	0.0002	0.0012	$\lvert\rho_{cc} - \rho_{c,exp}\rvert$ [g/cm^3]	0.0027	0.0042	0.0026	0.0006	0.0008
$\lvert T_{cc} - T_{c,exp}\rvert$ [K]	0.61	0.61	0.61	2.19	0.61	$\lvert T_{cc} - T_{c,exp}\rvert$ [K]	0.85	0.85	0.85	5.23	0.85
Propane: 173-331K (205-331K)						***n*-Octane**: 320-520K (368-520K)					
% Avg. Error ρ_l	0.28	0.34	0.23	0.28	0.10	% Avg. Error ρ_l	0.49	0.45	0.86	0.49	0.41
% Avg. Error ρ_g	11.84 (6.89)	10.95	19.93	11.84	18.78	% Avg. Error ρ_g	2.87 (1.31)	3.03	17.51	2.88	3.50
% Avg. Error P^{sat}	3.97 (0.79)	4.36	5.33	3.99	4.01	% Avg. Error P^{sat}	4.35 (1.35)	4.18	5.02	4.35	3.84
% Avg. Error ΔH_{vap}	2.37	2.46	2.66	2.37	2.03	% Avg. Error ΔH_{vap}	3.81	3.54	1.90	3.80	4.34
$\lvert\rho_{cc} - \rho_{c,exp}\rvert$ [g/cm^3]	0.0020	0.0018	0.0020	0.0032	0.0026	$\lvert\rho_{cc} - \rho_{c,exp}\rvert$ [g/cm^3]	0.0005	0.0002	0.0035	0.0011	0.0004
$\lvert T_{cc} - T_{c,exp}\rvert$ [K]	0.74	0.74	0.74	2.98	0.74	$\lvert T_{cc} - T_{c,exp}\rvert$ [K]	1.14	1.14	2.55	2.61	1.14

The Final set of TBORG *n*-alkane model parameters had the following characteristics:

- A unique $CH_3 - CH_3$ bond length for ethane.
- A transferable $CH_3 - CH_2$ bond length valid for propane and longer molecules.
- A transferable $CH_2 - CH_2$ bond length valid for *n*-butane and longer molecules.
- A unique $CH_3 - CH_2 - CH_3$ bond angle for propane.
- A transferable $CH_3 - CH_2 - CH_2$ bond angle valid for *n*-butane and longer molecules.
- A transferable $CH_2 - CH_2 - CH_2$ bond angle valid for *n*-pentane and longer molecules.
- Transferable ε_{ii} and σ_{ii} LJ parameters for the $-CH_3$ group valid for all molecules.
- Transferable ε_{ii} and σ_{ii} LJ parameters for the $-CH_2-$ group valid for all molecules.
- Vapour density, saturated vapour pressure, liquid density, critical temperature and critical density data were used to regress the intermolecular and intramolecular parameters. Latent heat of vaporization data was excluded.

The four other constraint variations considered had the following characteristics:

- Case I: The same as for the TBORG Final set of parameters except that there were
  - distinct $CH_3 - CH_2$ bond lengths for propane, *n*-butane and molecules with at least than 5 carbons
  - distinct $CH_2 - CH_2$ bond lengths for *n*-butane and molecules with at least 5 carbons
  - distinct $CH_3 - CH_2 - CH_2$ bond angles for *n*-butane and molecules with at least 5 carbons
- Case II: The same as for the TBORG Final set of parameters except that
  - Latent heat of vaporization data was also used in the definition of the objective function
- Case III: The same as for the TBORG Final set of parameters except that
  - A good fit to estimated critical densities and critical temperatures was not enforced
- Case IV: The same as for the TBORG Final set of parameters except that
  - The vapour density data was excluded from the regression and not used in the definition of the objective function.

Comparing the accuracy of the simulated physical properties obtained from the Final set of parameters with the accuracy of the simulated physical properties resulting from the more expanded parameter set of Case I in Table 9-7, it may be seen that there is little difference for all molecules

studied. In addition, from Table 9-6 it can be seen that the optimal sizes regressed for the two CH_3-CH_2 bond lengths allowed differ by only 0.04Å while the optimal values for the two CH_2-CH_2 bond lengths allowed differ only by 0.02Å. The two optimal values considered for the $CH_3-CH_2-CH_2$ bond angle also differ by only 5°. Since there was no substantial benefit in using the more expanded intramolecular parameter set, the Final parameter set listed in Table 9-6 was preferred.

The regression for Case II which included data for the latent heat of vaporization in Equation ((9–4) was unsatisfactory because of the considerable deterioration in the agreement between the simulated and experimental vapour pressures particularly for propane, *n*-butane and *n*-octane. This can be traced back to the definition of the latent heat of vaporization in Equation (5–21) where the influence of the liquid molar volume will tend to dominate over the influence of the vapour molar volume in the regression because of the lower statistical uncertainty which tends to be found in the liquid densities. This is the case for propane and *n*-butane where an improvement in the liquid densities accompanied less accurate simulated vapour densities. The largest improvement in the latent heat of vaporization occurred for *n*-octane where the average discrepancy between the experimental and the simulated data was halved. It therefore appears that the effect of including the latent heat of vaporization data in the regression is most significantly felt for the larger molecules. This may also be seen for ethane where the agreement between the simulated and the experimental latent heat of vaporization data in fact deteriorated. In Table 9-12 in Appendix 9A it may be seen that the relative statistical uncertainty in the latent heat of vaporization is always lower than that of the corresponding vapour density. For this reason, the inclusion of latent heat of vaporization data in the objective function will in general lead to a lower accuracy in the simulated vapour density results.

For all regression cases except Case III, the model parameters regressed had to achieve values such that critical temperatures estimated through Equations (6–2) and (6–3) from the simulated data were constrained to be within 0.2% of the corresponding experimental critical temperature. This is why the predicted critical temperatures are the same for regressions cases other than Case III. For Case III, the constraint of accurately reproducing the critical data was relaxed. This was also the approach followed by Ungerer *et al.* (2000) and by Bourasseau *et al.* (2003). From Table 9-7 the error in the estimated critical temperature for all four molecules for Case III ranges from 2.2K to 5.2K. This was also the typical magnitude of the error experienced for the AUA4 *n*-alkane and alkene force fields of Ungerer *et al.* (2000) and Bourasseau *et al.* (2003) respectively. To a limited

extent, good estimations of the critical density and the critical temperature both require good predictions of the liquid and vapour densities. This result therefore argues in favour of producing a force field which provides good estimates of the vapour densities and which explicitly accounts for the critical properties in the regression objective function.

For Case IV, the experimental data for the vapour densities were excluded from the objective function. As a result, this mainly led to an improvement in the liquid densities and the saturated vapour pressures. For propane, ethane and *n*-butane the improvement in the latent heats of vaporization was marginal while for *n*-octane the simulated ΔH_{vap} data got slightly worse. This may be due to the slight deterioration in the accuracy of the simulated vapour densities for *n*-octane. It was seen previously for Case II that the relationship between ΔH_{vap} and ρ_g was most significant for *n*-octane.

The equilibrium bond lengths show the following trend:

$$CH_3-CH_3 \geq CH_3-CH_2 \geq CH_2-CH_2$$

This trend is qualitatively similar to the larger equilibrium bond lengths required by the nAEP force field to accurately reproduce the pure component VLE for *n*-alkanes. This observation would, therefore, seem to confirm the importance of allowing longer bond lengths to accurately reproduce pure component VLE. The relative sizes of the bond lengths above are also consistent with the effects of anisotropy. A larger off-centre intermolecular interaction for the $-CH_3$ groups is to be expected because of the terminal position of the $-CH_3$ group and the attachment of three hydrogen atoms. For the $-CH_2-$ group, however, there are an even number of hydrogen atoms attached (namely two) and the location of the $-CH_2-$ groups within the carbon backbone implies that the influence of anisotropic interactions are less pronounced. This is because the $-CH_2-$ group is somewhat symmetrical about the carbon centre. For this reason the CH_2-CH_2 equilibrium bond length should be the shortest and the optimized bond length parameters correctly reproduce the expected trend. The CH_2-CH_2 equilibrium bond length of 1.688Å is also reasonably close to the length used by the NERD and TraPPE force field for this bond of 1.54Å.

The three equilibrium bond angles are also consistent with the molecular geometry that is to be expected. For the NERD, TraPPE, nAEP and AUA4 force fields, an equilibrium bond angle of 114.0° was used for all of the bond angles as listed in Table 9-6. The TBORG $CH_3-CH_2-CH_3$ equilibrium bond angle for propane of 84.11° is, however, considerably smaller. This smaller

equilibrium bond angle when combined with the longer CH_3-CH_2 equilibrium bond length of 2.045Å does, however, provide a molecular volume for propane which is similar to that provided by the other force fields. This is also the case for the $CH_3-CH_2-CH_2$ equilibrium bond angle for *n*-butane and *n*-octane which, although also smaller than the 114.0° angle used by the NERD, TraPPE, nAEP and AUA4 force fields for this bond angle, does allow the TBORG force field to accommodate the longer bond lengths of 2.045Å and 1.688Å. The molecules in the TBORG force field, therefore, have a more "zig-zag" or "concertina" shape but still have approximately the same molecular volume as the volumes for these same molecules modelled by the NERD, TraPPE, nAEP and AUA4 force fields. This is encouraging because it implies that the parameters in the TBORG force, while empirical, still have a connection to physical reality. The flexible molecular geometries therefore serve to place the interaction sites in more favourable locations in an anisotropic sense similar to the AUA4 force field.

This effective intramolecular rearrangement has, however, been crucial in improving the simulated physical properties, particularly the saturated vapour pressures and vapour densities. Using the Final set of parameters in Table 9-6, the pure component VLE of ethane, propane, *n*-butane and *n*-octane were simulated. The pure component simulations results are shown in Figure 9-20 to Figure 9-23 and the corresponding data is given in Appendix 9A at the end of the Chapter. The average deviations in the predicted physical properties were similar to those obtained from the regression estimates of Table 9-7.

In terms of the liquid densities in Figure 9-20, it can be seen that the TBORG force field is able to provide as good a description of the liquid densities as the NERD and TraPPE force fields which are known to reproduce the liquid densities well. The TBORG force field is also able to provide good predictions of the critical temperature and the critical density because these properties are required in the regression. The NERD and TraPPE force fields are known to provide good estimations of the critical densities and these are, therefore, not shown for clarity. Critical temperatures and critical densities for ethane, propane, *n*-butane and *n*-octane were not quoted in the original AUA4 publication (Ungerer *et al.* (2000)) and are, therefore, not shown.

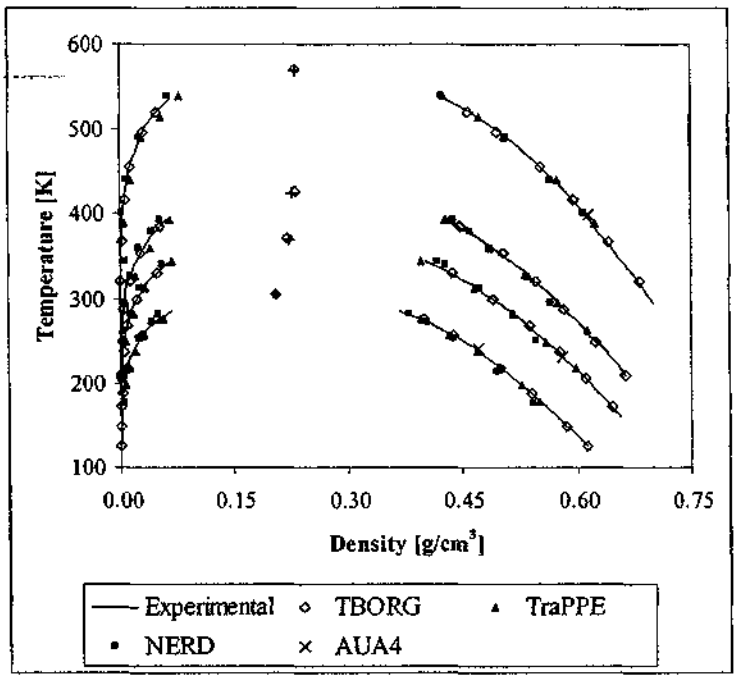


Figure 9-20: Liquid and vapour coexistence densities for ethane, propane, *n*-butane and *n*-octane. Simulated results from the TBORG, NERD, TraPPE and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986). Experimental critical points are shown as plus symbols.

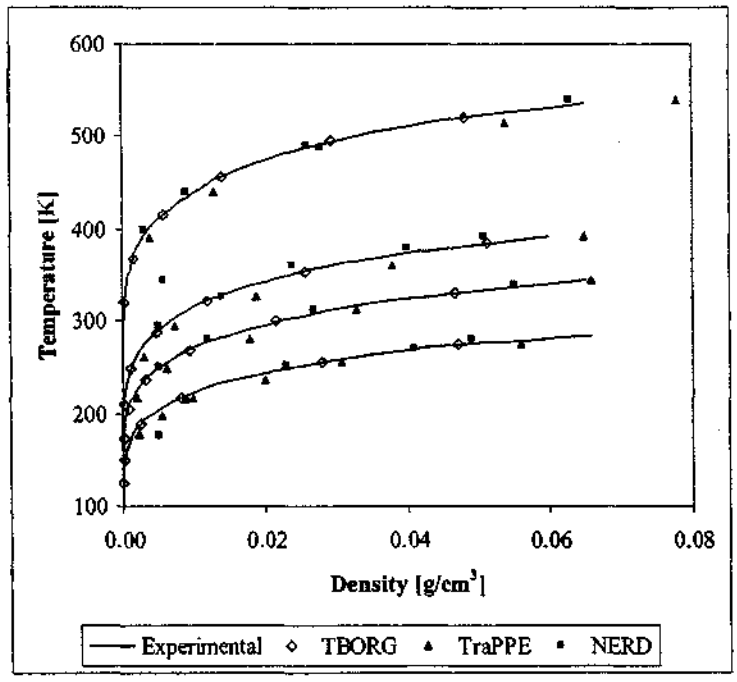


Figure 9-21: Vapour coexistence densities for ethane, propane, *n*-butane and *n*-octane. Simulated results from the TBORG, NERD and TraPPE force fields are shown with the experimental data of Smith and Srivastava (1986).

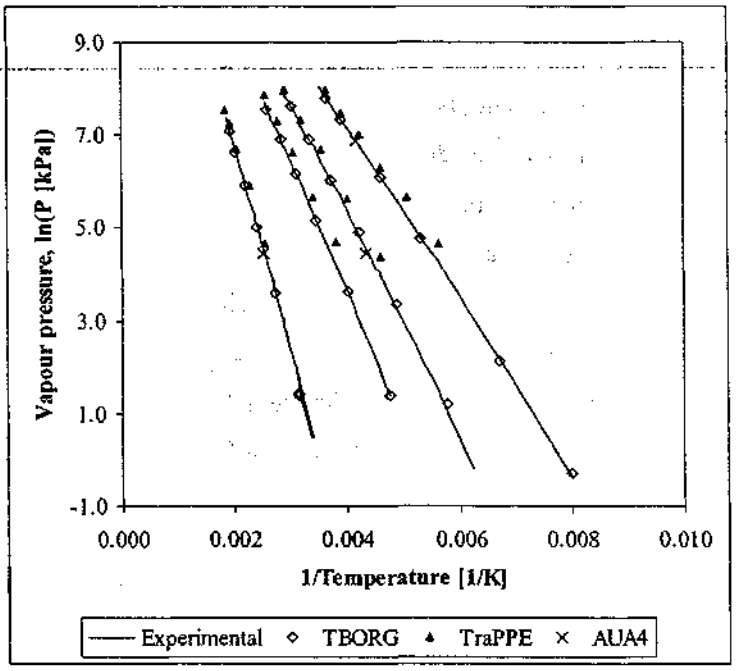


Figure 9-22: Saturated vapour pressures for ethane, propane, *n*-butane and *n*-octane. Simulated results from the TBORG, TraPPE and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986).

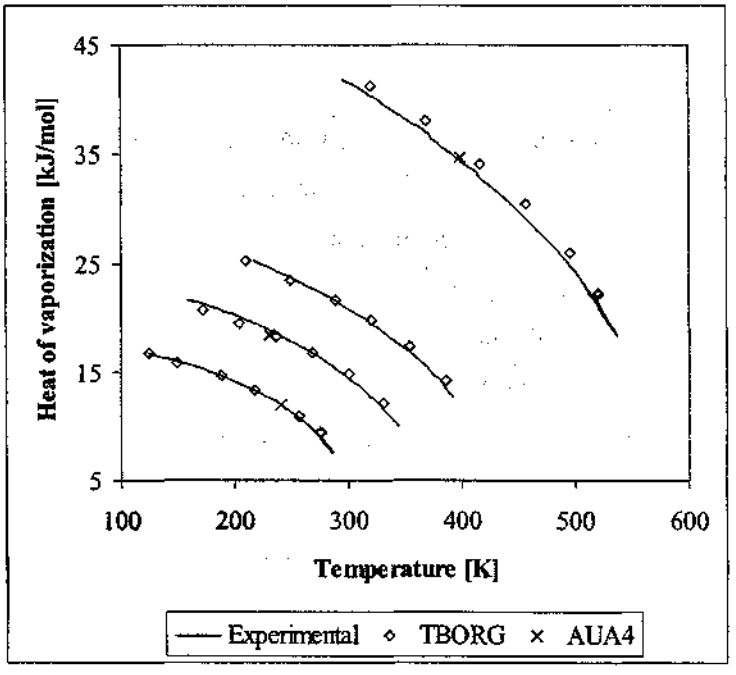


Figure 9-23: Latent heats of vaporization for ethane, propane, *n*-butane and *n*-octane. Simulated results from the TBORG and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986).

Figure 9-21 reveals the superior vapour densities predicted by the TBORG force field. The NERD force field tends to overestimate the vapour densities while the TraPPE force field tends to underestimate the vapour densities. This is largely a function of the greater emphasis placed on correctly reproducing the vapour densities during the parameterization of the TGORG force field. The saturated vapour pressures are also better estimated by the TBORG force field over the entire temperature range while the NERD and TraPPE force fields tend to overestimate the vapour pressure. The good fit to the saturated vapour pressures required by the TBORG parameterization has thus ensured superior pure estimates for the vapour pressures as well. Although the latent heats of vaporization were not included in the objective function for the regression of the TBORG *n*-alkane parameter set, it must be noted that the TBORG force field also provides a good estimate of the latent heats of vaporization for ethane, propane and *n*-butane. For *n*-octane, however, the agreement is less good but remains within an average of 5% of the experimental data.

Another important feature of the TBORG force is the temperature range over which the force has been parameterized and over which the force field gives accurate pure component properties. From Figure 9-20 to Figure 9-23 the larger effective range of the TBORG force field indicates its superior performance as when compared to the smaller effective temperature ranges of the NERD and TraPPE force fields. This is largely due to the deliberate simulation of pure component properties at lower reduced temperatures. It should be noted that the AUA4 force used only one coexistence point each for ethane, propane and *n*-octane during the parameterization of this force field. It has already been seen from the optimized geometries generated in Table 9-6 that the transferability of the intramolecular parameters is least for the shorter chain lengths. The TBORG force field explicitly accounted for this by the inclusion of additional intramolecular equilibrium geometries as required. This combined with the larger effective temperature range used for the parameterization of the TBORG force field would, therefore, suggest that greater confidence be placed in the TBORG force field when simulating these short and intermediate length *n*-alkanes over a wide range of temperatures.

9.5 Regression of sp^2 Hybridized $=CH_2$ and $=CH-$ Lennard-Jones Parameters for Alkenes

To maintain consistency with the optimization criteria used for the TBORG *n*-alkane force field, it was decided to enforce the transferability of intramolecular equilibrium geometries as far as

possible for the alkene parameters. Consequently, the Final set of TBORG alkene model parameters had the following characteristics:

- A unique bond length for the $CH_2=CH_2$ bond in ethene and a unique $CH-CH$ bond length for 1,3-butadiene.
- A transferable CH_3-CH bond length for propene, *cis*-2-butene and *trans*-2-butene. It was found that there was only a small impact on the physical properties when this bond length was required to be fully transferable.
- A transferable $CH_2=CH$ bond length valid for propene, 1-butene, 1,3-butadiene, 1-octene and longer molecules. It was again found that the values of an optimal independent set of $CH_2=CH$ bond lengths were very similar.
- A transferable $CH=CH$ bond length for *cis*-2-butene and *trans*-2-butene.
- A transferable $CH-CH_2$ bond length valid for 1-butene and longer molecules.
- A unique $CH_3-CH=CH_2$ bond angle valid for propene and a unique CH_3-CH_2-CH bond angle valid for 1-butene.
- A unique $CH_2=CH-CH$ bond angle valid for 1,3-butadiene.
- A transferable $CH_3-CH=CH$ bond angle valid for *cis*-2-butene and *trans*-2-butene
- A transferable $CH_2-CH=CH$ bond angle valid for molecules longer than 1-butene.
- A transferable $CH_2=CH-CH_2$ bond angle valid for 1-butene and longer molecules.
- Transferable ε_{ii} and σ_{ii} LJ parameters for the $=CH_2$ group valid for all molecules.
- Transferable ε_{ii} and σ_{ii} LJ parameters for the =CH– group valid for all molecules.
- Vapour density, saturated vapour pressure, liquid density, critical temperature and critical density data were used to regress the intermolecular and intramolecular parameters. Latent heat of vaporization data were again excluded.

The experimental data for *cis*- and *trans*-2-butene were more limited than the data available for the other alkenes studied. This was particularly true for *cis*-2-butene. Consequently, the Final TBORG model parameters were regressed to directly optimize the simulated physical property predictions for all of the alkenes except *cis*-2-butene. Bourasseau *et al* (2003) reported that the optimization of the AUA4 parameters for 2-alkenes could not provide the same quality of agreement with the experimental data as the TraPPE alkene force field of Wick *et al.* (2000). This was attributed to a deficiency in the AUA4 force field with respect to the 2-alkenes considered.

For the regression of the TBORG *n*-alkane parameters, it was found that it was possible to achieve a reasonable agreement between the critical temperatures and the critical densities estimated from simulation and the corresponding experimental critical temperatures and critical densities without having to strictly impose this requirement on the TBORG parameter optimization. For the alkenes, however, the agreement between the estimated critical temperatures and the critical densities was significantly worse than when this requirement was not imposed. This is evidenced in Table 9-10.

Consequently, the following three cases for the regression of the TBORG alkene parameters were also considered:

- Case I: The TBORG alkene regression was divided into two parts. In the primary regression, the model parameters necessary for all of the alkenes except *cis*- and *trans*-2-butene were optimized. In the next subsidiary regression, the CH_3-CH obtained for propene was imposed on *cis*- and *trans*-2-butene while the $CH=CH$ bond length and the $CH_3-CH=CH$ bond angle were regarded as common to both isomers but were optimized to achieve the best agreement with the experimental data.
- Case II: The TBORG alkene regression was again divided into two parts. In the primary regression, however, the model parameters necessary for all of the alkenes except *cis*-2-butene only were optimized. In the next subsidiary regression, the CH_3-CH obtained for propene and *trans*-2-butene were imposed on *cis*-2-butene while the $CH=CH$ bond length and the $CH_3-CH=CH$ bond angle for *cis*-2-butene were independently optimized to achieve the best agreement with the experimental data for *cis*-2-butene.
- Case III: The same as for the constraints used for the Final TBORG alkene parameter set except that a good agreement with the experimental critical temperatures and critical densities was not enforced.

For Equation (6–2) and (6–3) to provide a consistent estimate of the critical temperature and the critical density, coexistence density information in the near-critical region is required. The available data for *cis*-2-butene was limited to the reduced temperature range of $0.5T_{cr}-0.7T_{cr}$. Consequently, a good agreement with the critical density and the critical temperature of *cis*-2-butene was not enforced in any of the above cases or in the Final set of TBORG alkene parameters. The optimized parameters for the Final TBORG parameter set and for the three cases described above are listed in Table 9-8, the averages of the deviations of the simulated data from the

experimental data for the various species considered are listed in Table 9-9, while the effect of the various optimization cases on the estimated critical properties are listed in Table 9-10.

Table 9-8: Table of the regressed alkene TBORG model parameters for the different regression constraints.

	Final	**Case I**	**Case II**	**Case III**
$CH_2 = CH_2$ (ethene) [Å]	1.854	1.865	1.854	2.011
$CH_3 - CH$ ($\geq C_3$) [Å]	2.014	2.019	2.014	2.032
$CH_2 = CH$ ($\geq C_3$) [Å]	1.840	1.855	1.840	1.846
$CH - CH_2$ ($\geq C_4$) [Å]	1.753	1.751	1.753	1.771
CH – CH (1,3-butadiene) [Å]	1.809	1.806	1.809	1.865
CH = CH (*cis*-2-butene) [Å]	1.711	1.711	1.769	1.648
CH = CH (*trans*-2-butene) [Å]	1.711	1.711	1.711	1.648
$CH_3 - CH = CH_2$ (propene) [deg]	109.1	107.3	109.1	118.7
$CH_3 - CH_2 - CH$ (1-butene) [deg]	94.3	94.2	94.3	96.0
$CH_2 - CH_2 - CH$ ($\geq C_5$) [deg]	122.1	121.3	122.1	130.0
$CH_2 = CH - CH_2$ ($\geq C_4$) [deg]	116.1	115.9	116.1	117.8
$CH_2 = CH - CH$ (1,3-butadiene) [deg]	107.2	106.6	107.2	112.9
$CH_2 - CH = CH$ ($\geq C_5$) [deg]	116.1	115.9	116.1	117.8
$CH_3 - CH = CH$ (*cis*-2-butene) [deg]	117.1	117.1	115.0	117.4
$CH_3 - CH = CH$ (*trans*-2-butene) [deg]	117.1	117.1	117.1	117.4
$\sigma_{ii}(=CH_2)$ [Å]	3.503	3.499	3.503	3.4497
$\varepsilon_{ii}/k_B(=CH_2)$ [K]	108.40	108.80	108.40	114.53
$\sigma_{ii}(=CH-)$ [Å]	3.200	3.196	3.200	3.2161
$\varepsilon_{ii}/k_B(=CH-)$ [K]	100.22	101.09	100.22	97.83

Table 9-9: Deviation between the simulated and experimental physical properties for the Final parameter set of the TBORG alkene force field and for each of the variations of the Final set considered. Data in brackets corresponds to the bracketed and more limited temperature range listed for each molecule.

	Final	**Case I**	**Case II**	**Case III**		**Final**	**Case I**	**Case II**	**Case III**
Ethene: 136-253K (160-253K)					**1-Butene**: 216-383K (253-383K)				
% Avg. Error ρ_l	0.32	0.33	0.32	0.40	% Avg. Error ρ_l	0.55	0.55	0.55	0.53
% Avg. Error ρ_g	$16.19_{(11.19)}$	$14.68_{(10.17)}$	$16.19_{(11.19)}$	$5.31_{(4.28)}$	% Avg. Error ρ_g	$5.03_{(3.30)}$	$4.61_{(3.15)}$	$5.03_{(3.30)}$	$3.96_{(2.91)}$
% Avg. Error P^{sat}	$1.29_{(0.39)}$	$1.37_{(0.44)}$	$1.29_{(0.39)}$	$1.60_{(0.97)}$	% Avg. Error P^{sat}	$1.12_{(0.42)}$	$1.04_{(0.45)}$	$1.12_{(0.42)}$	$0.86_{(0.48)}$
% Avg. Error ΔH_{vap}	1.76	1.79	1.76	1.70	% Avg. Error ΔH_{vap}	1.33	1.38	1.33	1.71
$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0007	0.0005	0.0005	0.0002	$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0023	0.0023	0.0023	0.0027
$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	0.07	0.16	0.16	1.24	$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	0.84	0.84	0.84	1.10
Propene: 176-327K (207-327K)					**1,3-Butadiene**: 216-383K (253-383K)				
% Avg. Error ρ_l	0.37	0.39	0.37	0.15	% Avg. Error ρ_l	0.47	0.47	0.47	0.62
% Avg. Error ρ_g	$2.43_{(1.24)}$	$2.51_{(1.09)}$	$2.43_{(1.09)}$	$1.67_{(1.67)}$	% Avg. Error ρ_g	$13.20_{(11.51)}$	$12.93_{(11.35)}$	$13.20_{(11.51)}$	$5.00_{(4.61)}$
% Avg. Error P^{sat}	$0.80_{(0.17)}$	$0.74_{(0.17)}$	$0.80_{(0.17)}$	$0.68_{(0.23)}$	% Avg. Error P^{sat}	$1.70_{(0.98)}$	$1.58_{(0.95)}$	$1.70_{(0.95)}$	$1.34_{(0.84)}$
% Avg. Error ΔH_{vap}	1.86	1.90	1.86	2.24	% Avg. Error ΔH_{vap}	0.67	0.73	0.67	1.73
$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0022	0.0022	0.0022	0.0032	$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0018	0.0019	0.0019	0.0035
$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	0.73	0.73	0.73	0.18	$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	0.85	0.85	0.85	2.44

	Final	Case I	Case II	Case III		Final	Case I	Case II	Case III
***cis*-2-Butene**: 136-253K (160-253K)					**1-Octene**: 216-383K (253-383K)				
% Avg. Error ρ_l	0.10	0.12	0.15	0.09	% Avg. Error ρ_l	0.70	0.71	0.70	0.45
% Avg. Error ρ_g	$4.36_{(4.24)}$	$5.29_{(4.95)}$	$6.80_{(6.80)}$	$5.52_{(4.75)}$	% Avg. Error ρ_g	$15.56_{(10.01)}$	$15.04_{(9.76)}$	$15.56_{(10.01)}$	$16.79_{(13.22)}$
% Avg. Error P^{sat}	$6.15_{(5.29)}$	$6.50_{(5.93)}$	$6.36_{(6.30)}$	$4.77_{(3.73)}$	% Avg. Error P^{sat}	$4.48_{(2.46)}$	$4.44_{(2.33)}$	$4.48_{(2.46)}$	$4.20_{(3.05)}$
% Avg. Error ΔH_{vap}	2.20	2.38	1.97	1.89	% Avg. Error ΔH_{vap}	7.11	7.10	7.11	7.70
$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0038	0.0038	0.0029	0.0045	$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0023	0.0023	0.0023	0.0068
$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	11.54	11.54	12.31	11.15	$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	1.13	1.13	1.13	10.37
***trans*-2-Butene**: 176-327K (207-327K)									
% Avg. Error ρ_l	0.23	0.21	0.23	0.18					
% Avg. Error ρ_g	$2.27_{(2.30)}$	$3.14_{(3.03)}$	$2.27_{(2.30)}$	$3.63_{(4.00)}$					
% Avg. Error P^{sat}	$3.95_{(2.72)}$	$4.13_{(2.78)}$	$3.95_{(2.72)}$	$3.02_{(2.29)}$					
% Avg. Error ΔH_{vap}	4.72	4.89	4.72	4.44					
$\lvert\rho_{cr} - \rho_{cr,exp}\rvert$ [g/cm^3]	0.0020	0.0020	0.0020	0.0030					
$\lvert T_{cr} - T_{cr,exp}\rvert$ [K]	0.86	0.86	0.86	2.87					

Table 9-10: Comparison of the experimental critical temperature and densities with those estimated from the Final TBORG alkene parameter set and for the optimal parameter set obtained when good agreement with the experimental critical densities and temperatures was not imposed. The corresponding estimates from the AUA4 and TraPPE force fields are also listed.

	Exp.	Final	Case III	TraPPE	AUA4
Ethene					
ρ_{cr} [g/cm^3]	0.2141	0.2149	0.2140	0.215	0.212
T_{cr} [K]	282.3	282.3	283.6	283	280
Propene					
ρ_{cr} [g/cm^3]	0.2234	0.2256	0.2266	0.230	0.2257
T_{cr} [K]	365.6	366.3	365.7	363	373.5
1-Butene					
ρ_{cr} [g/cm^3]	0.2338	0.2361	0.2365	0.241	0.2362
T_{cr} [K]	419.6	420.4	420.7	414	419.6
1,3-Butadiene					
ρ_{cr} [g/cm^3]	0.2450	0.2432	0.2353	-	237.5
T_{cr} [K]	425.2	424.3	446.7	-	422.4
***cis*-2-Butene**					
ρ_{cr} [g/cm^3]	0.2398	0.2360	0.2353	0.246	0.2382
T_{cr} [K]	435.6	447.1	446.7	435	448
***trans*-2-Butene**					
ρ_{cr} [g/cm^3]	0.2357	0.2337	0.2327	0.238	0.234
T_{cr} [K]	428.6	429.5	431.5	426	439.4
1-Octene					
ρ_{cr} [g/cm^3]	0.2338	0.2361	0.2406	0.238	0.2409
T_{cr} [K]	567.0	568.1	577.4	567	559.2

There is only a small difference between the optimal parameters listed for the Final TBORG alkene parameter set and for Case I in Table 9-8. This is encouraging because it suggests that there is a substantial degree of transferability of the Final TBORG alkene parameter set to both 1-alkenes and 2-alkenes. In Case II in Table 9-8, the Final and independently regressed equilibrium bond lengths

and equilibrium bond angles for *cis*-2-butene and *trans*-2-butene are also very similar to the values obtained when *trans*-2-butene was included within the main parameter regression. These equilibrium bond lengths and equilibrium bond angle parameters are also very similar for both *cis*-2-butene and *trans*-2-butene and therefore confirm the simplifications of using transferable bond length values and bond angle values for *cis*-2-butene and *trans*-2-butene. For Case III in Table 9-8, however, the optimal parameter set is significantly different to the values obtained for the Final parameter set. Most of the equilibrium bond lengths for Case III are higher than the corresponding equilibrium bond lengths for the Final alkene parameter set. This is also true for the equilibrium bond angles for Case III which tend to be higher than the equilibrium bond angles for the Final alkene parameter set. This confirms the perspective of Chen *et al.* (1998) in which it was suggested that multiple parameter sets exist for which various "optimal" force fields can be obtained. From the above results and from the regressions obtained for the TBORG *n*-alkane parameters, the key feature in resolving the final and optimal parameter set resides in the regression constraints and the optimization criteria used.

The result of relaxing the constraint of a good prediction of the critical temperatures and critical densities is evidenced in Table 9-10 where the critical densities and critical temperatures for Case III show a larger deviation from the corresponding experimental values than for the Final parameter set, Case I and Case II. Table 9-9, however, shows that the simulated physical properties for Case III show a lower average deviation from the corresponding values. This was found to be largely attributable to a general improvement in the simulated physical properties at low temperatures which occurs when the constraint of a good estimation of the critical properties is relaxed. This explains why the AUA4 force field is generally able to provide a better prediction of the pure component densities, vapour pressures and heats of vaporization than the NERD or TraPPE force fields while the NERD and TraPPE force fields yield superior critical property estimations than the AUA4 force field.

It may be seen in Table 9-8 that all of the carbon-carbon double bonds are significantly larger than the 1.34Å used by the NERD and TraPPE force fields. This is again consistent with the TBORG force field parameters displacing the interaction centres to more optimal locations to effectively account for anisotropic interactions. The effect on the regressed equilibrium bond length is again most pronounced for bonds incorporating a terminal group, this time the sp^2 hybridized $-CH_2$ group. This increase in the equilibrium bond length may be understood in terms of a required displacement of the functional group centre to a location intermediate between the carbon centre

and its two attached hydrogen atoms. For ethane, the TBORG force field required an increase of the CH_3-CH_3 equilibrium bond length of 0.66Å over the same equilibrium bond length used for the NERD and TraPPE force fields due to the three hydrogen atoms present in the $-CH_3$ group. For ethene, the increase in the $CH_2=CH_2$ bond length for the TBORG force field is only 0.51Å due to the two hydrogen atoms present in the $=CH_2$ group. This is reassuring because it suggests that the regressed equilibrium bond lengths are not simply empirical but rather demonstrate the correct trend of increasing anisotropy with increasing numbers of attached hydrogen atoms. In addition, the $CH_2=CH_2$ equilibrium bond length for ethene is less than the CH_3-CH_3 equilibrium bond length for ethane as expected. The regressed carbon-carbon double bond lengths show the following trend:

$$CH_2=CH_2 \geq CH_2=CH \geq CH=CH$$

The difference between largest and smallest regressed values of the equilibrium lengths for the carbon-carbon double bonds in Table 9-8 is smaller than for the carbon-carbon single bonds of the TBORG *n*-alkane bond lengths Table 9-6. This is because of the greater anisotropy of the sp^2 hybridized –CH= group with only one attached hydrogen atom as compared to the $-CH_2-$ group with two attached hydrogen atoms. The net result is a $CH_2=CH$ bond length which is very similar in length to the $CH_2=CH_2$ bond length and a regressed equilibrium bond length for the $CH=CH$ bond length which is larger than the CH_2-CH_2 equilibrium bond length even though one would expect carbon-carbon double bonds to be shorter than carbon-carbon single bonds.

The regressed equilibrium lengths of the carbon-carbon single bonds in Table 9-8 follow the following trend:

$$CH_3-CH \geq CH-CH \geq CH-CH_2$$

The CH_3-CH bond length is the longest due to the anisotropy of the $-CH_3$ group while the anisotropic contribution of two –CH= groups in the $CH-CH$ bond results in it requiring a longer bond length than the $CH-CH_2$ bond with only one –CH= group. It is also encouraging to note that the anisotropy of the –CH= group is also results in longer carbon-carbon single bond lengths for the alkene regression as compared to the carbon-carbon single bonds regressed from the TBORG *n*-alkane force field. The only exception is the CH_3-CH_2 equilibrium bond length in propane as compared to the CH_3-CH equilibrium bond length in propene, but in this case the effects of the terminal $-CH_3$ group as well as the regressed propane and propene bond angles dominate.

It is interesting to note that there is less than an 8° difference between the regressed equilibrium values for all of the bond angles except those for the $CH_2=CH-CH$, CH_3-CH_2-CH and CH_2-CH_2-CH bond angles. This similarity of the Final equilibrium bond angle values indicates how the NERD and TraPPE force fields were able to achieve a reasonable agreement with the experimental data with only a single equilibrium bond angle value. As for the bond angles regressed for the TBORG *n*-alkane force field, the regressed equilibrium bond angles area all lower than the 124° used by both the NERD and TraPPE force field for bond angles with a carbon-carbon double bond. This is again largely a function of the longer bond lengths regressed for the TBORG alkene force field and which lead to a "concertina" structure for the alkene molecules as well. This is most strongly seen in the CH_3-CH_2-CH and $CH_3-CH=CH_2$ bond angles where the longer bond lengths for bond including terminal groups leads to the smallest values for the regressed bond angles.

Several features of the TBORG alkene force field are, however, less satisfactory than was found for the TBORG *n*-alkane force field:

- The CH_2-CH_2-CH equilibrium bond angle is larger than the 114° that is used by the NERD and TraPPE force fields for this bond angle.
- Comparing the agreement between the simulated and experimental results for 1-octene in Table 9-9 and the agreement between the simulated and experimental results for *n*-octane in Table 9-7, it may be seen that the simulated data is better for *n*-octane than for 1-octene, particularly for the vapour densities and the latent heats of vaporization.
- In Table 9-9, the deviation between the experimental and simulated vapour densities for ethene, 1,3-butadiene and 1-octene all tend to be greater than 10% and are larger than for any of the other molecules studied. This situation could not be improved by increased the number of independent intramolecular geometry parameters.
- The simulated latent heats of vaporization for the alkenes tend to be worse than the simulated latent heats of vaporization for the *n*-alkanes.

The disagreement between the vapour densities for ethene is most noticeable because ethene has only a $=CH_2$ group and an independent bond length. This therefore suggests that the value of the values of the Lennard-Jones parameters obtained for the $=CH_2$ group using the TBORG optimization constraints are perhaps not fully optimal. Based on the simulated results, it is believed that this might be attributable to assumptions inherent in the TBORG force field, namely

- The application of the united-atom assumption to the –CH= group, and
- The use of transferable torsional angles

The –CH= group is the least symmetrical of all of the functional groups for which Lennard-Jones parameters had to be regressed. This is because of the presence of only one hydrogen atom attached to the carbon centre. For all of the molecules studied, the –CH= group is attached by a carbon-carbon single bond to an adjacent functional group. Rotation around this carbon-carbon single bond is not hindered as for rotation about the carbon-carbon double bond. This implies that the hydrogen atom can effectively rotate about the central carbon in the functional group which effectively shifts the interaction site in a torsional sense. This is not the case for the $-CH_3$, $-CH_2-$ or the $=CH_2$ because the displaced centre tends to be averaged out by the presence of more than one attached hydrogen atom. This is mostly strongly understood for 1,3-butadiene where the $CH-CH$ equilibrium bond length should vary between a maximum value when the two hydrogen atoms on the two carbon centres are *trans* relative to one another and a minimum value when the two hydrogen atoms on the two carbon centres are *cis* relative to one another. The TBORG force field, however, uses a fixed bond length for the $CH-CH$ equilibrium bond length. This suggests why the results for the vapour densities for 1,3-butadiene are less accurate and provides another explanation as to why the flexible bond lengths of the modified NERD force field gave better results for the pure component properties of 1,3-butadiene in Chapter Seven. The carbon-carbon double bond in *trans*- and *cis*-2-butene also possesses two –CH= groups for which the $CH=CH$ bond length could vary between maximum and minimum values depending on the location of the attached hydrogen atoms relative to one another. This would explain the poorer estimate of the latent heats of vaporization for *trans*- and *cis*-2-butene as seen in Table 9-9. These results would therefore suggest one of two possible remedies for the –CH= group, namely:

1) Explicit and separate LJ parameters for the carbon and the hydrogen atoms in the –CH= functional group
2) Inclusion of a flexible bond length for bonds with a –CH= group to account for rotation of the hydrogen atom about the carbon centre

The other assumption which then also needs to be questioned is the transferability and origin of the torsional angle potentials. In the above regressions it has been seen that distinct bond angle and bond length values enabled the TBORG force field to more accurately simulate a range of pure component properties over a larger range of temperatures than previous united-atom force fields. For the torsions, however, it is common practice when parameterizing united-atom force fields to

treat a torsional potential as transferable so long as the bond between the two central functional groups is the same. In Chapter Seven it was argued that the $CH_2 = CH - CH = CH_2$ torsional potential of 1,3-butadiene appeared to not be applicable to the simulation of 1,3-pentadiene and 1,3,5-hexatriene. In light of the slightly poorer results obtained for *n*-octane relative to the other *n*-alkanes and in light of the poorer pure component properties also obtained for 1-octene, it would appear that this might be the case for torsions used in the parameterization of the TBORG force field as well. To an extent, the regressed values for the LJ ε_{ii} and σ_{ii} *n*-alkane parameters would compensate for slight errors in the torsional potential for *n*-octane. These errors would, however, propagate through the alkene parameterization because the parameterization necessarily had to be performed sequentially. For 1-octene, however, the regressed values for the LJ ε_{ii} and σ_{ii} for the $=CH_2$ and –CH= groups would be least able to compensate for errors in the torsional potential because of the lower occurrence of $=CH_2$ and –CH= groups. This also explains the observed inaccuracy of the ethene vapour densities because the Final values of the LJ parameters for the $=CH_2$ would have been shifted to partially compensate for small errors in the torsional potentials. The simultaneous regression of all of the required LJ parameters together with the additional parameters for distinct torsions would massively increase the number of regression simulations required, but the above results for *n*-octane and 1-octene would suggest that this is likely to be required to produce a truly transferable united-atom force field.

The simulated data obtained using the Final TBORG alkene parameter set are graphed in Figure 9-24 to Figure 9-31 where they are also compared to the results from other force fields, while Table 9-13 in Appendix 9A lists the simulation data used to generate Figure 9-24 to Figure 9-31. As listed in Table 9-9, the liquid densities are in general well reproduced for all the alkenes studied in Figure 9-24 and Figure 9-26 and are comparable to the estimates of the TraPPE force field. The vapour densities are also in reasonable agreement with the experimental data except for the cases of ethene, 1,3-butadiene and 1-octene as shown in Figure 9-25 and Figure 9-27 and are generally superior to both the NERD and TraPPE force fields because this was included in the optimization criteria. In Figure 9-28 and Figure 9-29 it may be seen that the simulated vapour pressures of the TBORG force field are also general superior the estimates of the TraPPE and AUA4 force fields. In Figure 9-30 and Figure 9-31, however, the estimates of the latent heats of vaporization of the TBORG force field are worse than those of the AUA4 force field, particularly for 1-octene but this is because a good prediction of the latent heats of vaporization was required by the AUA 4 force and not by the TBORG optimization criteria.

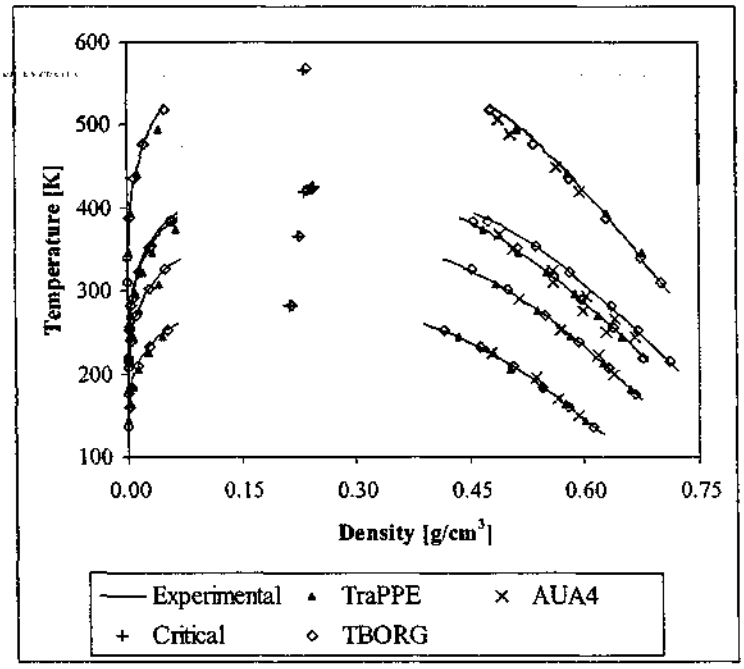


Figure 9-24: Liquid and vapour coexistence densities for ethene, propene, 1-butene, 1,3-butadiene and 1-octene. Simulated results from the TBORG, TraPPE and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986). Experimental critical points are shown as plus symbols.

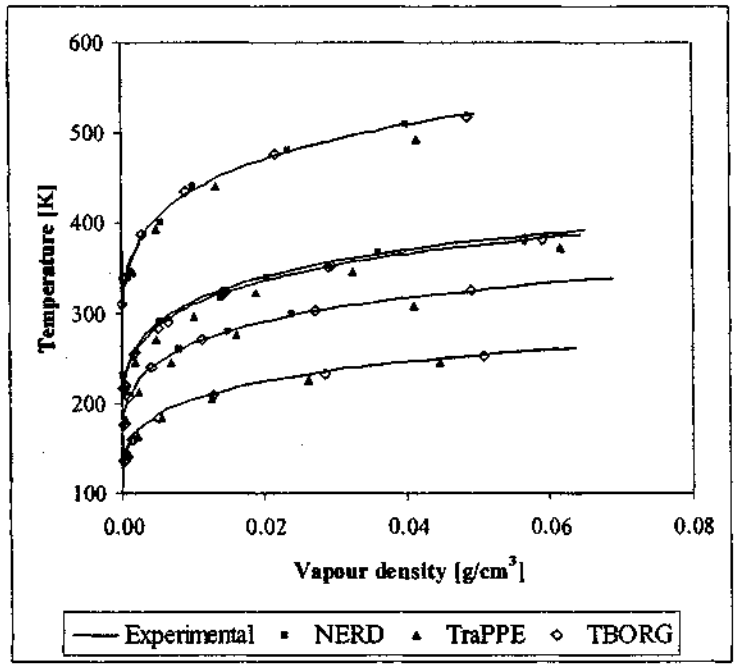


Figure 9-25: Vapour coexistence densities for ethene, propene, 1-butene, 1,3-butadiene and 1-octene. Simulated results from the TBORG and TraPPE force fields are shown with the experimental data of Smith and Srivastava (1986).

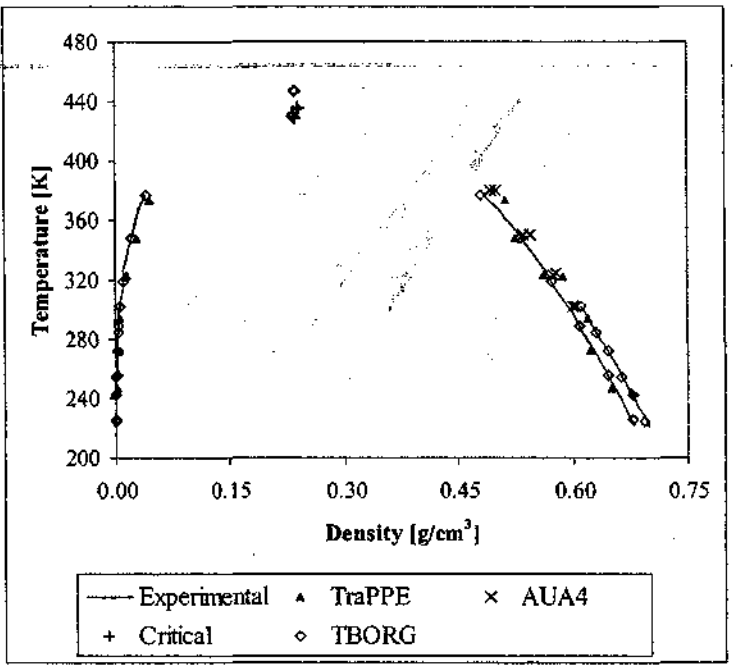


Figure 9-26: Liquid and vapour coexistence densities for *cis*-2-butene and *trans*-2-butene. Simulated results from the TBORG, TraPPE and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986). Experimental critical points are shown as plus symbols.

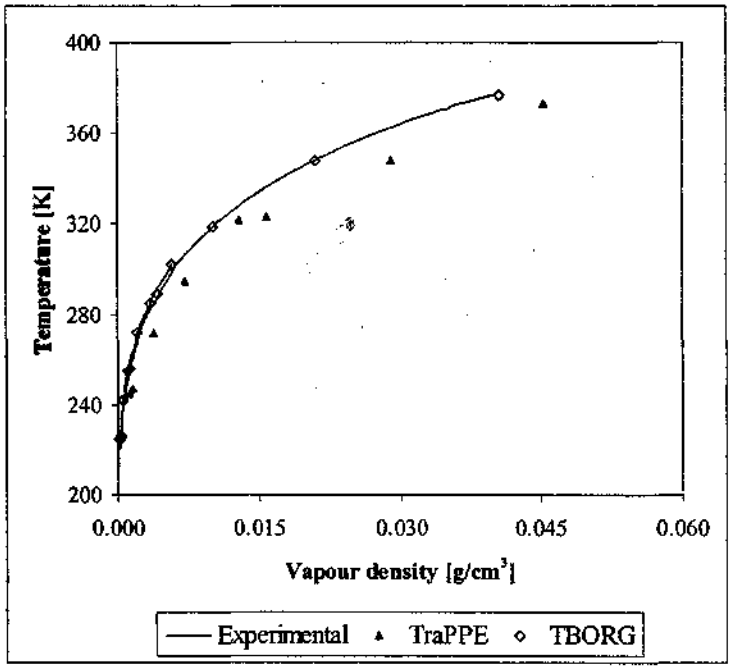


Figure 9-27: Vapour coexistence densities for *cis*-2-butene and *trans*-2-butene. Simulated results from the TBORG and TraPPE force fields are shown with the experimental data of Smith and Srivastava (1986).

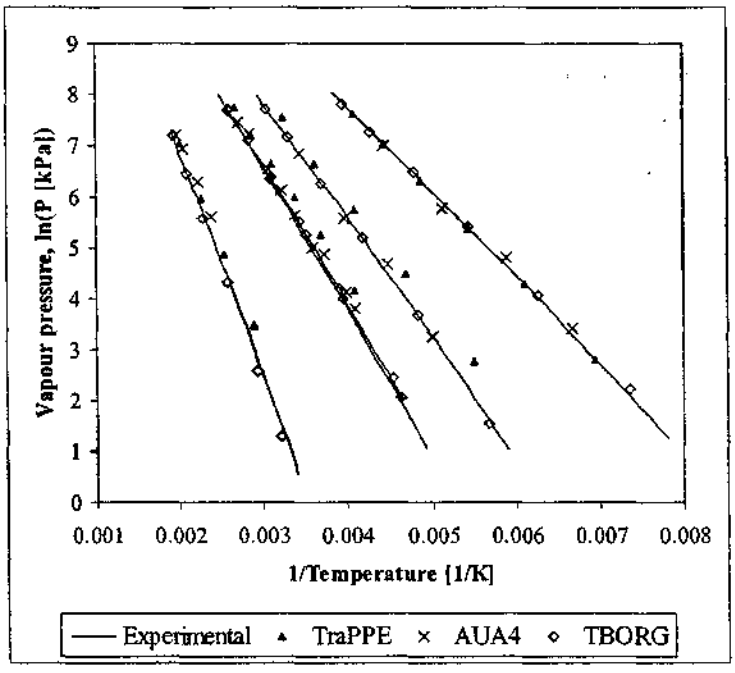


Figure 9-28: Saturated vapour pressures for ethene, propene, 1-butene, 1,3-butadiene and 1-octene. Simulated results from the TBORG, TraPPE and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986).

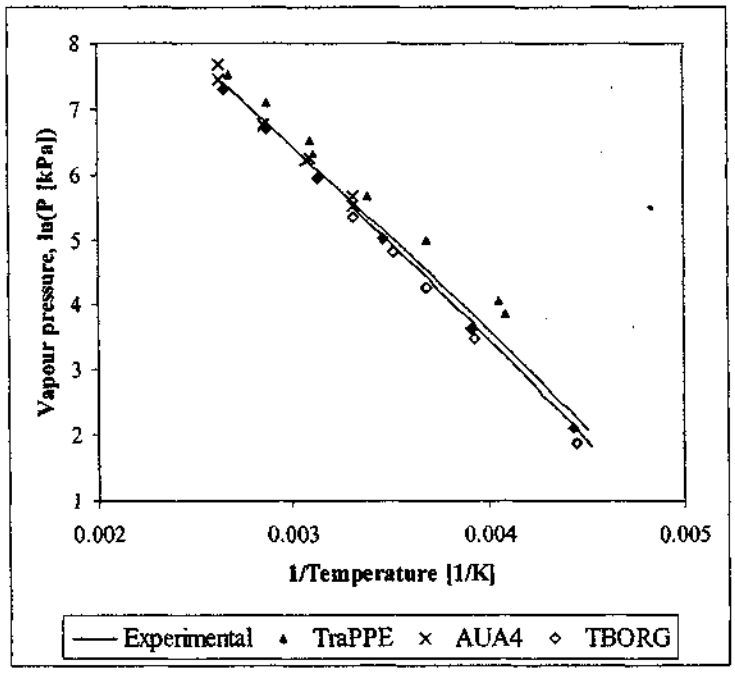


Figure 9-29: Saturated vapour pressures for *cis*-2-butene and *trans*-2-butene. Simulated results from the TBORG, TraPPE and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986).

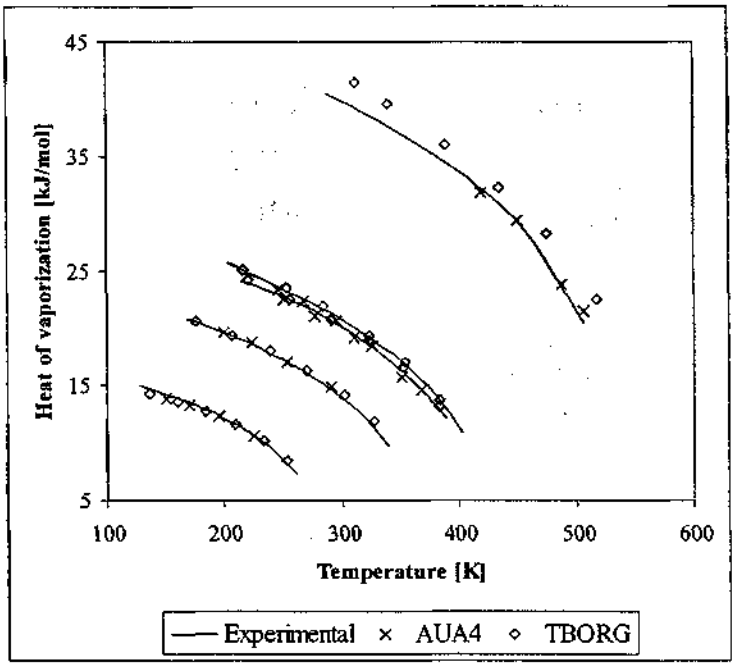


Figure 9-30: Latent heats of vaporization for ethene, propene, 1-butene, 1,3-butadiene and 1-octene. Simulated results from the TBORG and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986).

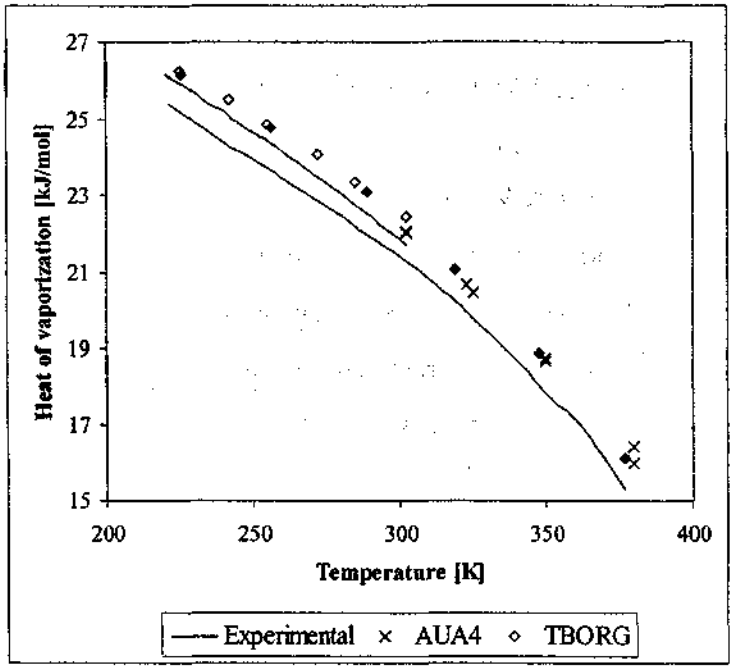


Figure 9-31: Latent heats of vaporization for *cis*-2-butene and *trans*-2-butene. Simulated results from the TBORG and AUA4 force fields are shown with the experimental data of Smith and Srivastava (1986).

9.6 Comparison of the Regressed Lennard-Jones Parameters

The values for the LJ ε_{ii} and σ_{ii} parameters for the sp^3 hybridized $-CH_3$ and $-CH_2-$ groups and for the sp^2 hybridized $=CH_2$ and $-CH=$ groups of the TraPPE, AUA4 and TBORG fields are compared in Figure 9-32 and Figure 9-33. From Figure 9-32 it may be seen that both the TBORG force field (which implicitly accounts for anisotropy) and AUA4 force field (which explicitly accounts for anisotropy) have LJ ε_{ii} parameters which are larger than the corresponding values for the TraPPE force field (which does not account for anisotropy). From Figure 9-19 the effect of a larger LJ ε_{ii} parameter is to increase the magnitude of the intermolecular attractions and repulsions. Larger LJ ε_{ii} parameters are required by the AUA4 and TBORG force fields relative to the TraPPE force field because the interaction sites for the AUA4 and TBORG force fields are displaced relative to their locations when displaced by the TraPPE force field. This requires a deeper potential energy well and stronger repulsions at short distances in order to provide the similar simulated values for the liquid densities in particular. For all three force fields, the value of the LJ ε_{ii} parameter for a given hybridization (sp^3 or sp^2) increases with increasing numbers of attached hydrogen atoms. This is because a united-atom functional group with a larger number of atoms should, in general, lead to an intermolecular interaction which is larger in magnitude.

For the LJ σ_{ii} parameter in Figure 9-33, the TraPPE force field tends to have larger σ_{ii} values than either the AUA4 or TBORG force fields. This again relates back to the effects of anisotropy because a larger σ_{ii} value in Figure 9-19 causes intermolecular repulsions to be shifted to larger separation distances. These larger σ_{ii} values are required by the TraPPE force field because there is no accounting of the displacement of the interaction site from the carbon centre and the larger σ_{ii} values ensure that neighbouring functional groups are still situated at average distances which yield the correct phase densities. It is interesting to note that for the TraPPE force field, the LJ σ_{ii} parameter increases with decreasing numbers of attached hydrogen atoms for a given hybridization while the opposite trend is observed for the AUA force field. Bourasseau *et al.* (2003) argued that the LJ σ_{ii} parameter should increase with increasing number of attached hydrogen atoms for a given hybridization because the increased molecular volume of more hydrogen atoms should increase the distance at which intermolecular repulsions are experienced. Martin and Siepmann (1998), however, argued that it was reasonable for the $-CH_2-$ methylene group to possess a larger LJ σ_{ii} parameter than the $-CH_3$ methyl group value because of the two bond lengths shared by the

$-CH_2-$ group with neighbouring functional groups. For the TBORG force field, it may be seen from Figure 9-33 that the LJ σ_{ii} parameter increases with increasing numbers of hydrogen atoms for sp^3 hybridized functional groups but that it decreases with increasing numbers of hydrogen atoms for sp^2 hybridized functional groups. This is because the TBORG force field implicitly accounts for the effects of anisotropy through optimization of the equilibrium bond lengths and the equilibrium bond angles. As discussed previously, the increases in the lengths of the carbon-carbon single bond lengths for the TBORG force field are larger than for the carbon-carbon double bond lengths. This bond length effect leads to the different trends in the values of the LJ σ_{ii} parameters for the sp^3 and sp^2 hybridized functional groups of the TBORG force field.

9.7 Simulation of Binary VLE for the TBORG Force Field

To test the TBORG force field for the simulation of binary VLE, an *n*-pentane (1) + *n*-hexane (2) mixture at 25°C was selected (Dortmund Data Bank). The P-x-y data was simulated using the ξ_1NVT Gibbs Ensemble because it was shown in the previous Chapter that the ξ_1NVT Gibbs Ensemble yields lower uncertainties in the simulated compositions for both the TraPPE and the TBORG *n*-alkane force fields. The simulation methodology employed was essentially the same as for the ξ_1NVT Gibbs Ensemble simulations as described in Section (8.5.1). The CBMC method was used for the simulations and a total system size of 280 molecules was used as this was found to be sufficient to give reproducible results. The NVT Gibbs Ensemble was used for the pure component compositions. Values for ξ_1 were again selected give approximate overall mole fractions of 0.05, 0.10, 0.20, 0.40, 0.60, 0.80, 0.90 and 0.95 of *n*-pentane at equilibrium. The same types of Monte Carlo moves as listed in (8.5.1) for the ξ_1NVT Gibbs Ensemble simulations of the 1-butene + *n*-butane and 1-hexene + *n*-hexane mixtures were performed with the same relative probabilities. All of the simulations were equilibrated for 40,000 cycles and the production period for ensemble averages consisted of at least 160,000 cycles. As for previous simulations, the standard deviations of ensemble averages were computed by dividing the production period into five blocks. The standard deviation of an ensemble averaged quantity was then set as the standard deviation of the averages obtained from each of these five blocks. The resulting P-x-y data and x-y data are shown in Figure 9-34 and Figure 9-35 respectively and the data is tabulated in Table 9-11.

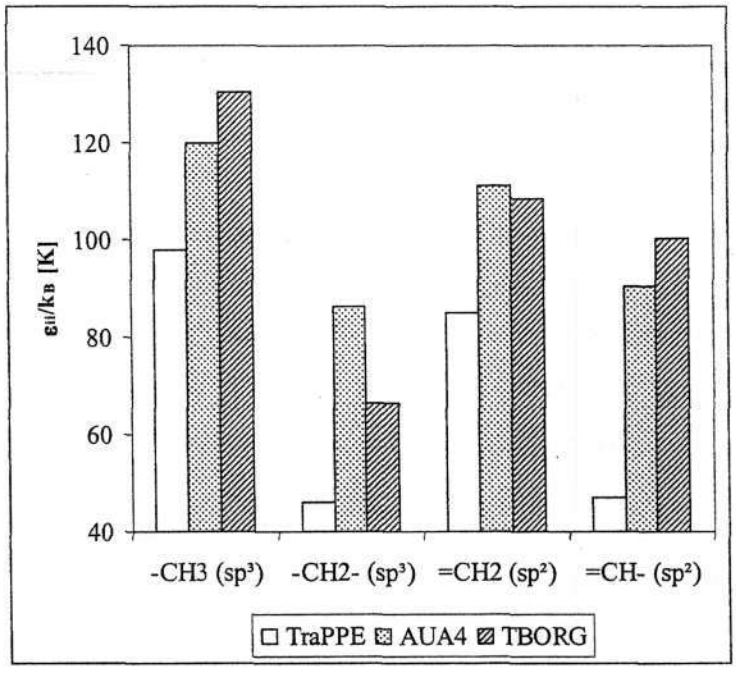


Figure 9-32: Comparison of the LJ ε_{ii} parameters for the sp^3 hybridized $-CH_3$ and $-CH_2-$ groups and for the sp^2 hybridized $=CH_2$ and $-CH=$ groups of the TraPPE, AUA4 and TBORG fields.

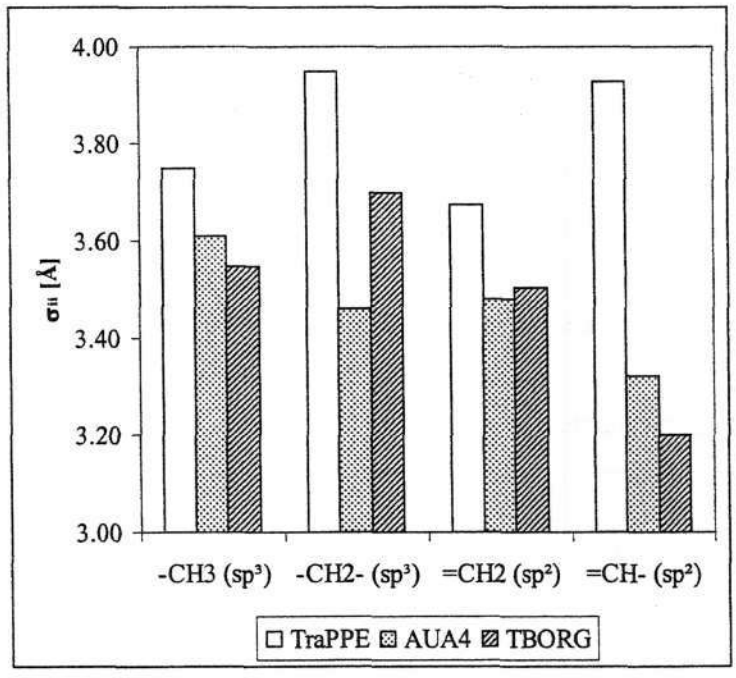


Figure 9-33: Comparison of the LJ σ_{ii} parameters for the sp^3 hybridized $-CH_3$ and $-CH_2-$ groups and for the sp^2 hybridized $=CH_2$ and $-CH=$ groups of the TraPPE, AUA4 and TBORG fields.

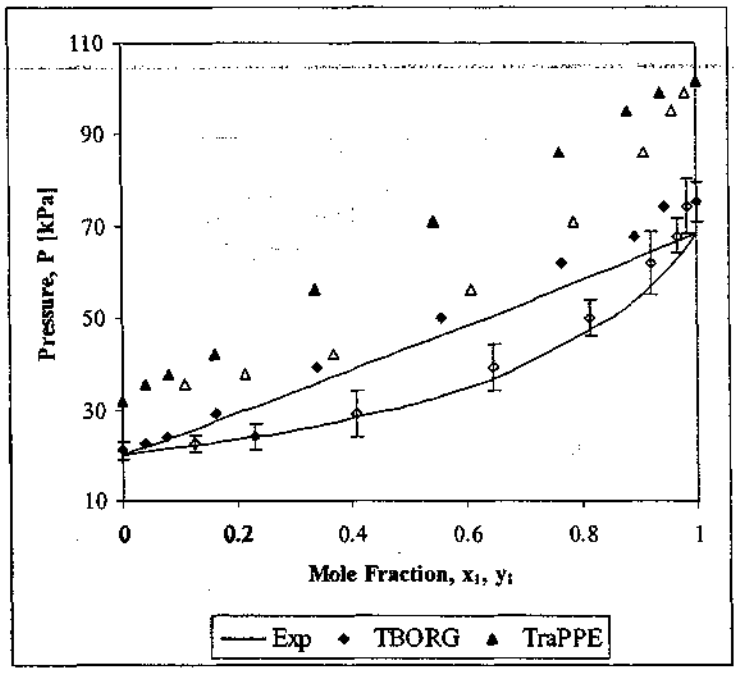


Figure 9-34: P-x-y data for the *n*-pentane (1) + n-hexane (2) mixture at 25°C from the TBORG (diamonds) and TraPPE (triangles) force fields obtained using the ξ_1NVT Gibbs Ensemble and the experimental data of the Dortmund Data Bank. Open symbols denote dew points and closed symbols denote bubble points.

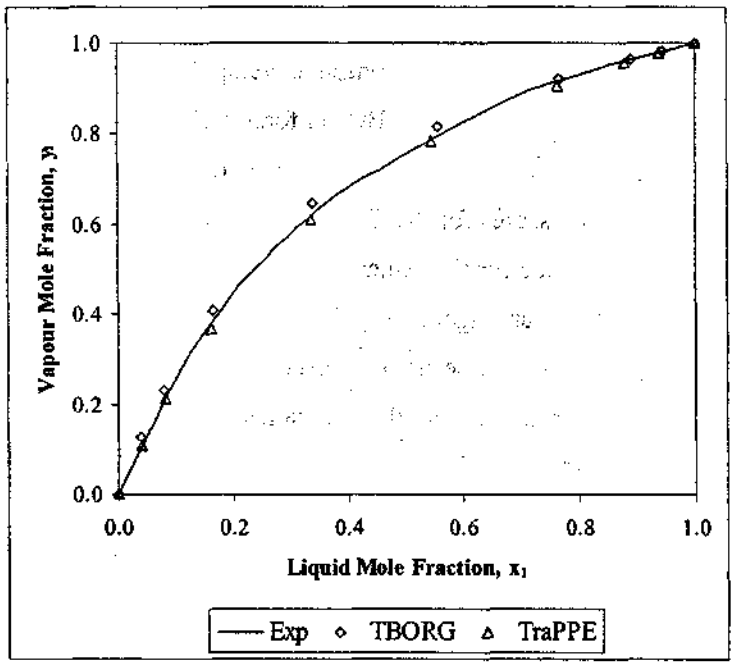


Figure 9-35: x-y data for the *n*-pentane (1) + n-hexane (2) mixture at 25°C from the TBORG (diamonds) and TraPPE (triangles) force fields obtained using the ξ_1NVT Gibbs Ensemble and the experimental data of the Dortmund Data Bank.

Table 9-11: P-x-y data for the ξ_1NVT Gibbs Ensemble simulations of the *n*-pentane (1) + n-hexane (2) mixture at 25°C from the TBORG and TraPPE force fields. Subscripts denote the estimated statistical uncertainty in the simulated result.

Pressure [kPa]		x_1		y_1	
TBORG	TraPPE	TBORG	TraPPE	TBORG	TraPPE
$21.1_{1.9}$	$31.9_{2.6}$	0_0	0_0	0_0	0
$22.7_{1.8}$	$35.7_{3.2}$	$0.039_{0.001}$	$0.039_{0.000}$	$0.125_{0.000}$	$0.109_{0.001}$
$24.2_{2.8}$	$37.8_{0.6}$	$0.078_{0.001}$	$0.081_{0.001}$	$0.231_{0.000}$	$0.213_{0.000}$
$29.2_{5.0}$	$42.0_{1.8}$	$0.162_{0.001}$	$0.160_{0.002}$	$0.408_{0.001}$	$0.369_{0.001}$
$39.1_{5.1}$	$56.1_{3.6}$	$0.339_{0.002}$	$0.337_{0.001}$	$0.645_{0.002}$	$0.609_{0.001}$
$50.0_{4.1}$	$71.1_{6.7}$	$0.556_{0.001}$	$0.544_{0.004}$	$0.815_{0.000}$	$0.785_{0.001}$
$62.1_{6.8}$	$86.2_{5.6}$	$0.766_{0.001}$	$0.763_{0.002}$	$0.920_{0.000}$	$0.908_{0.000}$
$67.9_{3.8}$	$95.1_{8.0}$	$0.890_{0.001}$	$0.879_{0.001}$	$0.966_{0.000}$	$0.957_{0.000}$
$74.3_{6.0}$	$99.3_{2.7}$	$0.943_{0.001}$	$0.938_{0.001}$	$0.983_{0.000}$	$0.979_{0.000}$
$75.3_{4.4}$	$101.7_{4.8}$	1_0	1_0	1_0	1_0

The benefit in the improvement of the pure component vapour pressures of the TBORG force field may be seen in Figure 9-34 where the TBORG force field results are shifted significantly closer to the correct phase envelope than the TraPPE force field results. As may be seen in Figure 9-35, this improvement in the overall location of the phase envelope is achieved without significantly compromising the x-y data predictions of the TBORG force field which agree well with the results from the TraPPE simulations and with the experimental data. Only the estimated statistical uncertainty for the dew point pressures for the TBORG force field are indicated in Figure 9-34 for clarity but are of similar magnitude for the remaining symbols. The estimated uncertainties in the liquid and vapour compositions are again very low for the ξ_1NVT Gibbs Ensemble and are consistently less than 0.004 for both force fields as seen in Table 9-11. This binary mixture result would therefore seem to confirm that the *n*-alkane parameters for the TBORG force field are also suitable for the simulation of binary VLE.

9.8 Methods to Improve United-Atom Force Fields

Based on the results of this Chapter, the following considerations would therefore appear to be important for the parameterization of improved united-atom force fields:

1) The correct simulation of critical data should be a criterion enforced during the optimization of the force field.
2) Vapour density and vapour pressure data should be used during the parameterization to ensure a good fit this data.
3) The torsional potentials should be re-examined and potentially optimized to specific force fields.
4) Sequential optimization of different LJ functional groups can lead to the propagation of errors through a force field. More attention should then be given towards simultaneous parameterization of all required functional groups.
5) A combination of displaced interaction sites using explicit anisotropic parameters and flexible equilibrium geometries may have to implemented to achieve a truly optimal united-atom force field
6) For certain functional groups (e.g. the –CH= functional group), the validity of the united-atom approximation may have to be re-examined.

9.9 Appendix 9A

Table 9-12: Simulated physical properties from the TBORG *n*-alkane force field for ethane, propane, *n*-butane and *n*-octane. Experimental data from Smith and Srivastava (1986). Estimated simulation errors are given as subscripts.

Molecule	Temperature	ρ_l [g/cm^3]		$\rho_g \times 10^4$ [g/cm^3]		$\ln(P^{sat})$ [kPa]		ΔH_{vap} [kJ/mol]	
	[K]	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**
Ethane	**125**	0.614	$0.614_{0.001}$	0.202	$0.259_{0.03}$	-0.362	$-0.300_{0.18}$	16.8	$16.8_{0.02}$
	149	0.587	$0.586_{0.001}$	2.16	$2.39_{0.08}$	2.18	$2.13_{0.04}$	16.0	$16.0_{0.06}$
	188	0.540	$0.540_{0.001}$	24.2	$25.3_{1.4}$	4.80	$4.76_{0.06}$	14.6	$14.6_{0.04}$
	217	0.501	$0.500_{0.001}$	80.7	$83.9_{1.8}$	6.09	$6.08_{0.02}$	13.3	$13.3_{0.10}$
	256	0.437	$0.439_{0.001}$	281	$282_{6.1}$	7.34	$7.32_{0.03}$	10.7	$10.7_{0.11}$
	275	0.396	$0.400_{0.004}$	488	472_{21}	7.82	$7.80_{0.04}$	8.86	$8.86_{0.04}$
Propane	**173**	0.645	$0.645_{0.001}$	0.853	$1.16_{0.28}$	1.02	$1.22_{0.24}$	21.2	$21.2_{0.08}$
	205	0.612	$0.611_{0.001}$	7.04	$8.17_{1.0}$	3.29	$3.35_{0.12}$	20.1	$20.1_{0.05}$
	236	0.577	$0.577_{0.001}$	29.7	$33.2_{2.5}$	4.83	$4.89_{0.07}$	18.6	$18.6_{0.08}$
	268	0.537	$0.538_{0.002}$	90.4	$94.4_{6.5}$	6.01	$6.03_{0.06}$	16.7	$16.7_{0.10}$
	300	0.490	$0.492_{0.002}$	220	$218_{5.1}$	6.91	$6.89_{0.03}$	14.4	$14.4_{0.03}$
	331	0.434	$0.438_{0.006}$	472	469_{47}	7.62	$7.60_{0.07}$	11.5	$11.5_{0.15}$

Molecule	Temperature	ρ_l [g/cm^3]		$\rho_g \times 10^4$ [g/cm^3]		$\ln\left(P^{sat}\right)$ [kPa]		ΔH_{vap} [kJ/mol]	
	[K]	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**
***n*-Butane**	**249**	0.664	$0.664_{0.001}$	1.33	$1.64_{0.29}$	1.38	$1.39_{0.24}$		$25.2_{0.14}$
	288	0.626	$0.625_{0.002}$	10.7	$11.6_{1.3}$	3.62	$3.62_{0.13}$	23.7	$23.5_{0.14}$
	321	0.585	$0.584_{0.001}$	45.6	$47.8_{2.2}$	5.17	$5.16_{0.05}$	21.6	$21.7_{0.05}$
	353	0.545	$0.546_{0.001}$	117	$120_{3.2}$	6.15	$6.15_{0.03}$	19.4	$19.7_{0.10}$
	386	0.501	$0.504_{0.004}$	251	257_{13}	6.92	$6.90_{0.04}$	16.9	$17.4_{0.14}$
	249	0.443	$0.448_{0.003}$	523	515_{35}	7.57	$7.56_{0.07}$	13.5	$14.2_{0.27}$
***n*-Octane**	**320**	0.681	$0.683_{0.003}$	2.48	$2.21_{0.26}$	1.75	$1.41_{0.18}$	40.3	$41.2_{0.52}$
	368	0.639	$0.643_{0.002}$	15.2	$15.3_{1.9}$	3.68	$3.59_{0.18}$	36.9	$38.1_{0.11}$
	416	0.595	$0.597_{0.003}$	57.4	$58.2_{5.5}$	5.08	$5.01_{0.13}$	33.0	$34.2_{0.51}$
	456	0.552	$0.554_{0.002}$	139	$140_{5.9}$	5.98	$5.92_{0.04}$	29.2	$30.5_{0.30}$
	496	0.501	$0.497_{0.003}$	302	295_{12}	6.72	$6.63_{0.06}$	24.7	$26.1_{0.63}$
	520	0.462	$0.459_{0.002}$	473	481_{33}	7.11	$7.06_{0.08}$	21.3	$22.3_{0.57}$

Table 9-13: Simulated physical properties from the TBORG alkene force field for ethene, propene, 1-butene, *cis*-2-**butene,** *trans*-2-**butene, 1,3-butadiene and** 1-octene. Experimental data from Smith and Srivastava (1986). Estimated simulation errors are given as subscripts.

Molecule	Temperature	ρ_l [g/cm^3]		$\rho_g \times 10^4$ [g/cm^3]		$\ln(P^{sat})$ [kPa]		ΔH_{vap} [kJ/mol]	
	[K]	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**
Ethene	**136**	0.614	$0.612_{0.001}$	2.03	$2.86_{0.14}$	2.10	$2.22_{0.06}$	14.8	$14.4_{0.05}$
	160	0.581	$0.580_{0.001}$	12.1	$14.7_{0.88}$	4.03	$4.06_{0.06}$	13.9	$13.6_{0.05}$
	184	0.546	$0.546_{0.001}$	43.2	$49.8_{2.5}$	5.40	$5.43_{0.06}$	12.9	$12.7_{0.06}$
	209	0.506	$0.507_{0.001}$	119	$129_{4.9}$	6.47	$6.47_{0.04}$	11.7	$11.6_{0.04}$
	233	0.460	$0.463_{0.001}$	266	$286_{5.6}$	7.28	$7.27_{0.02}$	10.1	$10.1_{0.07}$
	253	0.413	$0.415_{0.005}$	493	508_{38}	7.83	$7.79_{0.05}$	8.21	$8.46_{0.13}$
Propene	**176**	0.668	$0.667_{0.001}$	1.47	$1.35_{0.17}$	1.63	$1.56_{0.11}$	20.6	$20.6_{0.05}$
	207	0.631	$0.631_{0.001}$	10.0	$10.2_{0.23}$	3.69	$3.68_{0.02}$	19.3	$19.4_{0.05}$
	239	0.592	$0.592_{0.001}$	40.4	$40.6_{2.6}$	5.20	$5.18_{0.07}$	17.8	$18.0_{0.04}$
	270	0.549	$0.550_{0.002}$	113	$112_{4.0}$	6.26	$6.26_{0.04}$	16.1	$16.4_{0.05}$
	302	0.498	$0.501_{0.003}$	269	271_{15}	7.15	$7.15_{0.03}$	13.8	$14.2_{0.04}$
	327	0.447	$0.452_{0.005}$	505	490_{25}	7.71	$7.70_{0.03}$	11.3	$11.8_{0.09}$

Molecule	Temperature	ρ_l [g/cm^3]		$\rho_g \times 10^4$ [g/cm^3]		$\ln(P^{sat})$ [kPa]		ΔH_{vap} [kJ/mol]	
	[K]	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**
1-Butene	**220**	0.678	$0.677_{0.001}$	3.26	$3.71_{0.27}$	2.35	$2.46_{0.09}$	23.9	$24.2_{0.09}$
	256	0.638	$0.637_{0.001}$	17.7	$18.3_{0.87}$	4.18	$4.19_{0.05}$	22.5	$22.5_{0.08}$
	291	0.595	$0.597_{0.001}$	59.7	$63.4_{2.0}$	5.47	$5.52_{0.03}$	20.6	$20.7_{0.04}$
	322	0.554	$0.556_{0.001}$	141	$144_{4.5}$	6.36	$6.39_{0.04}$	18.6	$18.9_{0.10}$
	352	0.508	$0.512_{0.003}$	287	290_{18}	7.06	$7.08_{0.05}$	16.2	$16.6_{0.12}$
	383	0.448	$0.455_{0.007}$	573	591_{44}	7.67	$7.70_{0.05}$	12.9	$13.2_{0.13}$
***cis*-2-Butene**	**225**	0.697	$0.695_{0.001}$	2.45	$2.33_{0.37}$	2.09	$1.87_{0.22}$	25.9	$26.2_{0.06}$
	242	0.679	$0.678_{0.002}$	6.04	$5.13_{0.63}$	3.06	$2.70_{0.14}$	25.1	$25.5_{0.09}$
	255	0.665	$0.664_{0.001}$	10.9	$10.8_{0.94}$	3.70	$3.49_{0.11}$	24.4	$24.9_{0.09}$
	272	0.646	$0.646_{0.001}$	21.6	$21.7_{1.6}$	4.43	$4.25_{0.09}$	23.5	$24.1_{0.07}$
	285	0.631	$0.632_{0.002}$	34.2	$35.0_{0.94}$	4.92	$4.80_{0.03}$	22.7	$23.4_{0.12}$
	302	0.612	$0.612_{0.004}$	58.5	$57.5_{5.5}$	5.49	$5.34_{0.12}$	21.7	$22.4_{0.16}$
***trans*-2-Butene**	**226**	0.677	$0.680_{0.003}$	3.09	$3.03_{0.76}$	2.33	$2.09_{0.27}$	25.1	$26.2_{0.24}$
	256	0.646	$0.646_{0.002}$	13.0	$12.2_{1.3}$	3.87	$3.64_{0.11}$	23.6	$24.8_{0.15}$
	289	0.609	$0.610_{0.002}$	43.0	$41.5_{2.7}$	5.16	$5.01_{0.07}$	22.0	$23.1_{0.09}$
	319	0.573	$0.572_{0.002}$	102	$102_{7.4}$	6.07	$5.95_{0.07}$	20.2	$21.1_{0.07}$
	348	0.533	$0.533_{0.002}$	210	$209_{8.9}$	6.79	$6.70_{0.04}$	18.0	$18.9_{0.09}$
	377	0.484	$0.481_{0.001}$	399	405_{32}	7.40	$7.29_{0.06}$	15.3	$16.1_{0.22}$

Molecule	Temperature	ρ_l [g/cm^3]		$\rho_g \times 10^4$ [g/cm^3]		$\ln(P^{sat})$ [kPa]		ΔH_{vap} [kJ/mol]	
	[K]	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**	**Exp.**	**TBORG**
1,3-Butadiene	**216**	0.709	$0.712_{0.004}$	2.17	$2.64_{0.65}$	1.97	$2.07_{0.25}$	25.0	$25.1_{0.24}$
	253	0.668	$0.671_{0.002}$	13.9	$15.5_{2.4}$	3.97	$4.00_{0.15}$	23.2	$23.4_{0.07}$
	283	0.633	$0.636_{0.004}$	41.7	$49.7_{6.0}$	5.15	$5.26_{0.11}$	21.7	$21.9_{0.11}$
	323	0.582	$0.581_{0.002}$	131	$141_{8.4}$	6.34	$6.37_{0.05}$	19.2	$19.4_{0.08}$
	353	0.536	$0.536_{0.004}$	267	294_{22}	7.05	$7.11_{0.07}$	16.9	$17.0_{0.12}$
	383	0.480	$0.473_{0.011}$	518	564_{94}	7.65	$7.69_{0.09}$	13.8	$13.8_{0.20}$
1-Octene	**311**	0.700	$0.699_{0.002}$	2.00	$1.13_{0.32}$	1.52	$1.30_{0.27}$	39.0	$41.5_{0.32}$
	340	0.675	$0.674_{0.003}$	6.61	$5.33_{1.2}$	2.80	$2.59_{0.25}$	37.4	$39.6_{0.18}$
	388	0.631	$0.629_{0.002}$	30.8	$26.6_{2.1}$	4.44	$4.31_{0.09}$	34.4	$36.0_{0.27}$
	435	0.587	$0.583_{0.002}$	97.2	$98.8_{3.3}$	5.64	$5.57_{0.04}$	30.5	$32.3_{0.23}$
	476	0.543	$0.535_{0.003}$	220	$216_{9.3}$	6.47	$6.54_{0.04}$	25.1	$28.2_{0.19}$
	518	0.487	$0.479_{0.011}$	456	486_{57}	7.17	$7.19_{0.09}$		$22.5_{0.41}$

CHAPTER 10: CONCLUSIONS

"I never saw a wild thing sorry for itself. A small bird will fall frozen dead from a bough without ever having felt sorry" – D.H. Lawrence

The Beowulf cluster has indeed come of age and the maintenance and use of the Beowulf cluster yoda at the University of KwaZulu-Natal is an example of how supercomputing resources are now readily available to all academic researchers. The Beowulf cluster must now be considered as a serious proposition for everyday academic use rather than being considered as the domain of isolated groups of computer engineers or high-end users.

This work has demonstrated the utility of an accessible Beowulf cluster through the simulation of a variety of pure component and binary systems. Through a series of simulations using the TraPPE, NERD and SA1 force fields, this work was able to demonstrate the ability of transferable united-atom force fields to simulate *n*-alkane and 1-alkene binary mixtures of industrial relevance. There are, however, limitations to the range of applicability and the accuracy which can be obtained using these force fields. The SA1 force field was shown to be applicable to the simulation of the vapour pressures of long 1-alkene molecules through pure component simulations of the phase behaviour of 1-octadecene. The SA1 force field was, however, found to be inadequate when it comes to simulating *n*-alkane properties and the liquid densities of 1-alkenes. In contrast the NERD and TraPPE force fields were observed to provide a reasonable description for the five binary mixtures considered. Both force fields, however, displayed a tendency to overestimate the dew point and bubble point pressures of the phase envelope because of the overestimation of the pure component vapour pressures by both of these force fields. This observation emphasised the need for transferable united-atom force fields which correctly reproduce the pure component vapour pressures.

The x-y data revealed that in spite of discrepancies between the simulated and experimental P-x-y data, the x-y data for all binary mixtures was well described by all three force fields used. This is a significant result as it indicates that the force fields describe the correct shape of the phase envelope and can be regarded with confidence when simulating x-y data. By considering the excess volumes predicted by both force fields, however, it was found that the NERD, TraPPE and SA1 force fields cannot describe the expected ideal behaviour of the vapour phase. This observation motivated the

need for a transferable united-atom force field which correctly describes the pure component vapour densities.

With the mixed success obtained using the NERD, TraPPE and SA1 force fields, the following section of work focused on testing how transferable the TraPPE and NERD united-atom force fields were by examining their prediction of the pure component and binary phase behaviour of conjugated alkenes. The pure component properties of 1,3-butadiene; 1,3-pentadiene and 1,3,5-hexatriene were simulated using two sets of modified NERD and TraPPE parameters. It was found, however, that the current sets of parameters were insufficient to model the pure component phase behaviour accurately. Once the internal CH – CH equilibrium bond length had been modified to more accurately represent the actual equilibrium length of the internal single bond in 1,3-butadiene, however, a considerable improvement in the pure component simulated properties of 1,3-butadiene was observed, particularly for the modifications to the NERD force field. A dedicated torsional potential was also required for 1,3-butadiene. The results for the pure component simulations of 1,3-pentadiene and 1,3,5-hexatriene were, however, less satisfactory. It is believed that this is a combination of inadequate torsional potential and inadequate equilibrium intramolecular geometries. The precise root cause of the deviation of the simulated properties of 1,3-pentadiene and 1,3,5-hexatriene from the corresponding experimental data could not be established because of a lack of experimental data. The major result stemming from this section of this work was that more attention needs to be directed towards understanding the role and effect of the intramolecular equilibrium geometries on the simulated pure component physical properties.

The simulation of a binary mixture of *n*-heptane + 1,3-butadiene revealed further deficiencies in the parameter set which was used for the simulation of 1,3-butadiene because both the P-x-y and x-y data were not well represented. This is in contrast to the earlier observation that the NERD and TraPPE force fields provided a good description of the experimental x-y data in spite of poorer predictions for the P-x-y data. This observation suggested that the force field parameters might require a more fundamental investigation and a revision of the parameter values.

The development of a new transferable united-atom force was thus undertaken. The TBORG force field sought to improve the simulated pure component vapour pressures and vapour densities while not sacrificing an accurate description of the pure component liquid densities, critical temperatures and critical densities. This objective was met by relaxing the constraint of transferable intramolecular geometries to attempt to more faithfully account for subtle bond length and bond

angle variations for the different hybridizations and connectivities of functional groups that can be found in conjugated alkenes, *n*-alkanes, 1-alkenes and 2-alkenes. The assumption of fully transferable Lennard-Jones functional groups was, however, retained. A formal optimization program was undertaken through which it was also possible to elucidate and quantify the role of the bond lengths, the bond angles as well as the LJ ε_{ij} and σ_{ij} parameters on the vapour densities, the liquid densities, the heats of vaporization and saturated vapour pressures. Importantly, it was shown that the equilibrium geometries have a significant effect on the physical properties which confirmed the suspicions of earlier Chapters.

The TBORG force field was largely successful in achieving its goal of more accurately reproducing a larger range of physical properties over a wide temperature range, particularly for the *n*-alkanes. The agreement between the simulated and the experimental pure component properties for the alkenes was, unfortunately, less satisfactory. This is believed to have been due to a propagation of subtle errors in the parameterization of the 1-alkene parameters which argues in favour of a regression procedure which simultaneously regresses as large a number of functional groups as possible. In the process, the requirement for a more thorough consideration of the effects of anisotropy also became apparent because the optimization of the intramolecular geometries effectively corresponded to a shifting of the interaction centres to positions that would be consistent with displaced anisotropic interaction centres. It is therefore the opinion of this work that a combination of a larger anisotropic parameter set and a larger intramolecular geometry set might provide the largest improvement for the prediction of physical properties by united-atom force fields. It was also shown how the temperature range of the physical properties against which the model parameters are regressed should be as a large as possible and that as wide a set as possible of physical properties needs to be considered.

For the binary mixtures studied, several difficulties were experienced with using the N_1N_2PT Gibbs Ensemble because of mass balance limitations that are inherent to this type of simulation. To overcome these difficulties, a unifying formalism was developed based on an extension of the μPT ensemble to multiple phases. The result was a governing equation for isothermal Gibbs Ensembles that can be used to derive both the partition functions and thermodynamic potentials of novel Gibbs Ensembles. Using this framework, two new isothermal Gibbs Ensembles were developed which overcome all of the mass balance limitations of the N_1N_2PT Gibbs Ensemble. The μ_1N_2VT Gibbs Ensemble is characterised by the stipulation of one of the chemical potentials of one of the species while the ξ_1NVT Gibbs Ensemble is characterised by the stipulation of a fugacity fraction. Both of

these new ensembles are isochoric. To the knowledge of this work, these two new ensembles represent the first time that a Gibbs Ensemble has been able to simulate a close boiling mixture directly. To this end, two close-boiling mixtures of an *n*-alkane + 1-alkene were successfully simulated. The only slight drawback with the implementations of these two ensembles resides in the need to calculate the pressure during a simulation. The uncertainties in the computed pressures tend to be of the order of 5-10% which is typical for isochoric Gibbs Ensemble simulations. Both ensembles, however, are able to achieve a considerably lower uncertainty in the computed compositions than the N_1N_2PT Gibbs Ensemble, particularly the ξ_1NVT Gibbs Ensemble whose compositions can exhibit uncertainties two orders of magnitude smaller than those found in simulations using the N_1N_2PT Gibbs Ensemble.

The aim of this work was to address the simulation of hydrocarbon system of industrial relevance. To that end, a Beowulf cluster called yoda was assembled and installed. Several binary and pure component simulations were conducted which raised several important questions regarding the force fields and simulation algorithms used. In order to address these issues, a new force field called the TBORG force field was developed and a unifying framework and two new Gibbs Ensembles were characterised to permit the more efficient simulation of a range of binary mixtures.

CHAPTER 11: RECOMMENDATIONS

"Opportunity is missed by most people because it is dressed in overalls and looks like hard work." – Thomas Alva Edison

Based on the work described in the preceding Chapters, the following suggestions are extended for future work:

1) Investigation should take place into the development of transferable united-atom force fields which explicitly account for anisotropy as well as incorporating a larger intramolecular equilibrium geometry parameter set. While the TBORG *n*-alkane and alkene force fields demonstrate a general improvement in the accuracy of the pure component property predictions over a wider temperature range, some deficiencies, most notably the latent heat of vaporization, were identified. The agreement between the experimental and simulated data for the alkene force field was also less than that of the *n*-alkane force field and should be addressed.
2) More investigation should take place into methods which could efficiently facilitate simultaneous regression of intra- and intermolecular force field parameters for a number of homologous series. The propagation of subtle errors in the regressed force field parameters is not easily overcome with current regression techniques.
3) The principles underpinning the TBORG force field should be tested for a wider range of homologous series to ascertain whether or not a similar improvement in the physical properties can also be obtained for a wider range of compounds. This further testing of the TBORG force field should also include pure component simulations for carbon chains longer than C_8 to evaluate the transferability of the force field to longer *n*-alkane and alkene carbon chains.
4) More extensive testing of the TBORG force field needs to be undertaken for a larger number of mixtures to more thoroughly explore its application to the simulation of binary and ternary VLE.
5) Investigation should take place into the role of the torsional potential on the simulated pure component properties. An extensive study into the effect of the torsional parameters on the macroscopic phase behaviour has, to the knowledge of this work, never been undertaken. The apparent influence of the torsional potential on the conjugated alkenes

would, therefore, seem to justify a more formal investigation into the parameters of this intramolecular potential energy term.

6) The μ_1N_2VT and ξ_1NVT Gibbs Ensembles should be used to simulate azeotropic systems. This would mark a considerable step forward in terms of proving that a much wider range of binary systems, such as alcohols, can be simulated by the Gibbs Ensemble.

7) This work dealt primarily with the unification of isothermal Gibbs Ensemble under a governing framework. Using existing adiabatic ensemble theory, it will also be possible to unify a novel subset of adiabatic Gibbs Ensembles. This would then directly facilitate the formulation of Gibbs Ensembles which can be used for isobaric VLE without the mass balance constraints inherent in the N_1N_2PT Gibbs Ensemble. This work should then be further pursued to permit the formulation of a truly unifying formalism which provides for the unification of adiabatic and isothermal Gibbs Ensembles under a single governing framework.

8) The current Monte Carlo sampling methods result in an undesirably large uncertainty in the calculated pressure. This is unsatisfactory for the simulation of close boiling mixtures and is likely to also cause undesirably large uncertainties in the temperatures that would have to be calculated from any adiabatic Gibbs Ensemble simulations. New Monte Carlo sampling algorithms, therefore, need to be developed to more rigorously sample configurations which minimize the pressure and temperature differences between phases. This is explicitly in accordance with the mechanical and thermal equilibrium components of total thermodynamic equilibrium and will ensure lower statistical deviations in the calculated intensive properties.

CHAPTER 12:

PUBLICATIONS RESULTING FROM THIS WORK

"It is common sense to take a method and try it. If it fails, admit it frankly and try another. But above all, try something" – Franklin D Roosevelt

T McKnight, D Ramjugernath, K Bolton, and P Ahlström, "Phase equilibrium by computer experiments", *Chemical Technology*, April 2002, pg 28.

K Bolton, P Ahlström, T McKnight, and D Ramjugernath: "Studying the dynamics of chemical systems by computer simulations", Part 2", *Chemical Technology*, November 2002, pg 29.

T McKnight, D Ramjugernath, M Starzak, and K Bolton, "A Comparison of the NERD, TraPPE and Buckingham exp-6 united atom potentials for the simulation of vapour-liquid phase equilibrium", *Poster at the 17th IUPAC Conference on Chemical Thermodynamics*, ICCT, Rostock, Germany, 28 July - 2 August, 2002.

T McKnight, R Harris, D Ramjugernath, M Starzak, JD Raal, and K Bolton, "Molecular simulation and experiment measurement: A comparison", *Presentation at the 6th Italian Conference on Chemical and Process Engineering*, Pisa, Italy, 8-11 June, 2003.

T McKnight, D Ramjugernath, K Bolton, and M Starzak, "Comparing the experimental and simulation approaches for predicting binary VLE", *Presentation at the National Meeting of the South African Institute of Chemical Engineers (SAIChE) Congress 2003*, Sun City, 3-5 September 2003.

T McKnight, D Ramjugernath, K Bolton, P Ahlström, and M Starzak "Molecular simulation of binary vapour liquid equilibrium", *Presentation at the 1st Workshop on Applied Molecular Modelling and Simulations*, 22 – 23 April, 2004, Parys, South Africa.

T McKnight, D Ramjugernath, P Ahlström, and K Bolton, "Simulation of Conjugated Alkenes Using Transferable United Atom Force Fields", *Poster at the 3rd International Conference of "Computational Modeling and Simulation of Materials"*, Acireale, Italy, May 30 - June 4, 2004.

N du Preez, M Martin, T McKnight, and D Ramjugernath, "Pure component coexistence properties for long-chain 1-alkenes and 1-alcohols by molecular simulation using transferable force fields", *Poster at the 18th IUPAC Conference on Chemical Thermodynamics*, Beijing, China, 17-21 August 2004.

T McKnight, TJH Vlugt, D Ramjugernath, M Starzak, P Ahlström, and K Bolton "Simulation of 1-alkene and n-alkane binary vapour-liquid equilibrium using different united-atom transferable force fields", *Fluid Phase Equilibria*, 232(1-2): 136-148, **2005**

CHAPTER 13: REFERENCES

"Books serve to show a man that those original thoughts of his aren't very new at all" – Abraham Lincoln

Akkermans, R L C, and Ciccotti, G, On the equivalence of atomic and molecular pressure, *Journal of Physical Chemistry B*, 108 (21): 6866-6869, **2004**

Allen, M P, and Tildesley, D J, **1987**, "Computer Simulation of Liquids", Oxford University Press, Oxford

Attard, P, A Grand Canonical simulation technique for dense and confined fluids with application to a Lennard-Jones fluid, *Journal of Chemical Physics*, 107 (8): 3230-3238, **1997**

Bock, C W, George, P, Trachtman, M, and Zanger, M, Reappraisal of the structure of the 2nd stable conformer of buta-1,3-diene, *Journal of the Chemical Society-Perkin Transactions*, 2 (1): 26-34, **1979**

Boda, D, Liszi, J, and Szalai I, An extension of the NPT plus test particle method for the determination of the vapour-liquid equilibria of pure fluids, *Chemical Physics Letters*, 235 (1-2): 140-145, **1995**

Boda, D, Liszi, J, and Szalai I, A new simulation method for the determination of vapour-liquid equilibria in the Grand Canonical Ensemble, *Chemical Physics Letters*, 256 (4-5): 474-482, **1996**

Boulougouris, G C, Errington, J R, Economou, I G, Panagiotopoulos, A Z, and Theodorou, D N, Molecular simulation of phase equilibria for water-*n*-butane and water-*n*-hexane mixtures, *Journal of Physical Chemistry B*, 104 (20): 4958-4963, **2000**

Bourasseau, E, Ungerer, P, Boutin, A, and Fuchs, A H, Monte Carlo simulation of branched alkanes and long chain *n*-alkanes with anisotropic united-atoms intermolecular potential, *Molecular Simulation* 28 (4): 317-336, **2002**

Bourasseau, E, Haboudou, M, Boutin, A, Fuchs, A H, and Ungerer, P, New optimization method for intermolecular potentials: optimization of a new anisotropic united-atoms potential for olefins: Prediction of equilibrium properties, *Journal of Chemical Physics* 118 (7): 3020-3034, **2003**

Camp, P J, and Allen M P, Phase coexistence in a pseudo Gibbs Ensemble, *Molecular Physics*, 88 (6): 1459-1469, **1996**

Chen, B, Martin, M G, and Siepmann, J I, Thermodynamic properties of the Williams, OPLS-AA, and MMFF94 all-atom force fields for normal alkanes, *Journal of Physical Chemistry B*, 102 (14): 2578-2586, **1998**

Chen, B, and Siepmann, J I, Transferable potentials for phase equilibria. 3. Explicit-hydrogen description of normal alkanes, *Journal of Physical Chemistry B*, 103 (25): 5370-5379, **1999**

Chen, B, Potoff, J J, and Siepmann, J I, Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols, *Journal of Physical Chemistry B*, 105: 3093-3104, **2001**

Chialvo, A A, and Cummings, P T, Simple transferable intermolecular potential for the molecular simulation of water over wide ranges of state conditions, *Fluid Phase Equilibria*, 150: 73–81, **1998**

Conrad, P B, and de Pablo, J J, Comparison of Histogram Reweighting techniques for a flexible water model, *Fluid Phase Equilibria*, 150: 51-61, **1998**

Consta, S, Wilding, N B, Frenkel, D, and Alexandrowicz, Z, Recoil growth: an efficient simulation method for multi-polymer systems, *Journal of Chemical Physics*, 110 (6): 3220-3228, **1999a**

Consta, S, Vlugt, T J H, Hoeth, J W, Smit, B, and Frenkel, D, Recoil growth algorithm for chain molecules with continuous interactions, *Molecular Physics* 97 (12): 1243-1254, **1999b**

Cornell, W D, Cieplak, P, Bayly, C I, Gould, I R, Merz, K M, Ferguson, D M, Spellmeyer, D C, Fox, T, Caldwell, J W, and Kollman, P A, A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules, *Journal of the American Chemical Society*, 117 (19): 5179-5197, **1995**

de Pablo, J J, Laso, M, and Suter, U W, Simulation of polyethylene above and below the melting-point, *Journal of Chemical Physics*, 96 (3): 2395-2403, **1992a**

de Pablo, J J, Laso, M, and Suter U W, Estimation of the chemical potential of chain molecules by simulation, *Journal of Chemical Physics*, 96 (8): 6157-6162, **1992b**

Dodd, L R, and Theodorou D N, Atomistic Monte-Carlo simulation and continuum mean-field theory of the structure and equation of state properties of alkane and polymer melts, *Advances in Polymer Science*, 116: 249-281, **1994**

Dortmund Data Bank Vapour-Liquid Equilibrium Data Compilation, DDBST GmbH, Oldenburg, 1973

Dubbeldam, D, Calero, S, Vlugt, T J H, Krishna, R, Maesen, T L M, and Smit, B ,United-atom force field for alkanes in nanoporous materials, *Journal of Physical Chemistry B*, 108 (33): 12301-12313, **2004**

Eppinga R, and Frenkel, D, Monte Carlo study of the isotropic and nematic phases of infinitely thin hard platelets, *Molecular Physics*, 52: 1303-1334, **1984**

Errington J R, and Panagiotopoulos A Z, Phase equilibria of the modified Buckingham exponential-6 potential from Hamiltonian Scaling Grand Canonical Monte Carlo, *Journal of Chemical Physics*, 109(3): 1093-1100, **1998a**

Errington J R, and Panagiotopoulos A Z, A fixed point-charge model for water optimized to the vapour-liquid coexistence properties, *Journal of Physical Chemistry B*, 102: 7470-7475, **1998b**

Errington J R, Boulougouris, G C, Economou, I G, Panagiotopoulos A Z, and Theodorou, D N, Molecular simulation of phase equilibria for water-methane and water-ethane mixtures, *Journal of Physical Chemistry B*, 102: 8865-8873, **1998c**

Errington J R, and Panagiotopoulos A Z, A new intermolecular potential model for the *n*-alkane homologous series, *Journal of Physical Chemistry B*, 103 (30): 6314-6322, **1999a**

Errington J R, and Panagiotopoulos A Z, New intermolecular potential models for benzene and cyclohexane, *Journal of Chemical Physics*, 111 (21): 9731-9738, **1999b**

Escobedo, F A, and de Pablo J J, Monte Carlo simulation of the chemical potential of polymers in an expanded ensemble, *Journal of Chemical Physics*, 103 (7): 2703-2710, **1995**

Escobedo, F A, and de Pablo, J J, Pseudo-ensemble simulations and Gibbs-Duhem integrations for polymers, *Journal of Chemical Physics*, 106 (2): 793-810, **1997**

Escobedo, F A, Novel pseudo-ensembles for simulation of multicomponent phase equilibria, *Journal of Chemical Physics*, 108 (21): 8761-8772, **1998**

Escobedo, F A, Tracing coexistence lines in multicomponent fluid mixtures by molecular simulation, *Journal of Chemical Physics*, 110 (24): 11999-12010, **1999**

Escobedo, F A, Simulation and extrapolation of coexistence properties with single-phase and two-phase ensembles, *Journal of Chemical Physics*, 113 (19): 8444-8456, **2000**

Escobedo, F A, Molecular and macroscopic modelling of phase separation, *AIChE Journal*, 46 (10): 2086-2096, **2000b**

Escobedo, F A, Simulation of bulk, confined, and polydisperse systems. I. A unified methodological framework, *Journal of Chemical Physics*, 115 (12): 5642-5652, **2001a**

Escobedo, F A, Simulation of bulk, confined, and polydisperse systems. II. Application to chain systems, *Journal of Chemical Physics*, 115 (12): 5653-5661, **2001b**

Escobedo, F A, A simulation study of the lyotropic isotropic-nematic phase transitions in polydisperse chain systems, *Journal of Chemical Physics*, 118 (22): 10262-10275, **2003**

Esselink, K, Loyens L D J C, and Smit B, Parallel Monte-Carlo simulations, *Physical Review E*, 51 (2): 1560-1568, **1995**

Ferrenberg, A M, and Swendsen, R H, New Monte Carlo technique for studying phase transitions, *Physical Review Letters*, 61 (23): 2635-2638, **1988**

Ferrenberg, A M, and Swendsen, R H, Optimized Monte Carlo data analysis, *Physical Review Letters*, 63 (12): 1195-1198, **1989**

Frenkel, D, Mooij, G C A M, and Smit, B, Novel scheme to study structural and thermal properties of continuously deformable molecules, *Journal Of Physics-Condensed Matter*, 4 (12): 3053-3076, **1992**

Frenkel, D, and Smit, B, **1996**, "Understanding Molecular Simulation: From Algorithms to Applications", Academic Press, San Diego

Goff, G H, Farrington, P S, and Sage, B H, Volumetric and phase behaviour of propene-1-butene system, *Industrial & Engineering Chemistry*, 42 (4): 735-743, **1950**

Graben, H W, and Ray, J R, Unified treatment of adiabatic ensembles, *Physical Review A*, 43 (8): 4100-4103, **1991**

Grub, J, and Löser, E, Butadiene, *Ullmann's Encyclopedia of Industrial Chemistry, Electronic Release,* 7th ed., Wiley-VCH, Weinheim, **2000**

Grubbs, R H, and Chang, S, Recent advances in olefin metathesis and its application in organic synthesis, *Tetrahedron*, 54 (18): 4413-4450, **1998**

Hill, T L, **1956**, "Statistical Mechanics, Principles and Selected Applications", Dover Publications, New York

Jordan, B T, and van Winkle, M, Vapor-liquid equilibria at subatmospheric pressures - dodecane-octadecene system, *Industrial & Engineering Chemistry* 43 (12): 2908-2912, **1951**

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

Kamath, G, Cao F, and Potoff J J, An improved force field for the prediction of the vapour-liquid equilibria for carboxylic acids, *Journal of Physical Chemistry B*, 108 (37): 14130-14136, **2004**

Karpfen, A, Choi, C H, and Kertesz, M, Single-bond torsional potentials in conjugated systems: a comparison of ab initio and density functional results, *Journal of Physical Chemistry A* , 101 (40): 7426-7433, **1997**

Kofke, D A, and Glandt, E D, *Molecular Physics*, 64: 1105-1131, **1988**

Kofke, D A, Gibbs-Duhem integration: a new method for the direct evaluation of phase coexistence by molecular simulations, *Molecular Physics*, 78: 1331-1336, **1993a**

Kofke, D A, Direct evaluation of phase coexistence by molecular simulation via integration along the coexistence line, *Journal of Chemical Physics*, 98: 4149-4162, **1993b**

Laso, M, de Pablo, J J, and Suter, U W, Simulation of phase equilibria for chain molecules, *Journal of Chemical Physics*, 97 (4): 2817-2819, **1992**

Laugier, S, and Richon, D, High-pressure vapor-liquid equilibria for ethylene plus 4-methyl-1-pentene and 1-butene plus 1-hexene, *Journal of Chemical and Engineering Data*, 41 (2): 282-284, **1996**

Laurance, D R, and Swift G W, Vapour-liquid-equilibria in 3 binary and ternary systems composed of *n*-butane, 1-butene, and 1,3-butadiene, *Journal of Chemical and Engineering Data*, 19 (1): 61-67, **1974**

Lennard-Jones, J E, On the determination of molecular fields. II. From the equation of state of a gas, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 106: 463-477, **1924**

Li, H C, McCabe, C, Cui, S T, Cummings, P T, and Cochran, H D, On the development of a general force field for the molecular simulation of perfluoroethers, *Molecular Physics*, 101 (14): 2157-2169, **2003**

Lísal, M, Smith, W R, and Nezbeda, I, Accurate computer simulation of phase equilibrium for complex fluid mixtures. Application to binaries involving isobutene, methanol, methyl tert-butyl ether and *n*-butane, *Journal of Physical Chemistry B*, 103: 10496-10505, **1999**

Lísal, M, Smith, W R, and Nezbeda, I, Molecular simulation of multicomponent reaction and phase equilibria in MTBE ternary system, *AIChE Journal*, 46 (4): 866-875, **2000**

Lísal, M, Smith, W R, and Nezbeda, I, Accurate vapour-liquid equilibrium calculations for complex systems using the reaction Gibbs Ensemble Monte Carlo simulation method, *Fluid Phase Equilibria*, 181: 127-146, **2001**

Liu, R F, and Zhou, X F, Theoretical-study on the structure and vibrations of cis-1,3,5-hexatriene, *Journal of Physical Chemistry*, 97 (9): 1850-1855, **1993**

Lofti, A, Vrabec, J and Fischer, J, Vapour-liquid equilibria of the Lennard-Jones fluid from the NPT + test particle method, *Molecular Physics*, 76: 1319-1333, **1992**

Manousiouthakis V I, and Deem, M W, Strict detailed balance is unnecessary in Monte Carlo simulation, *Journal of Chemical Physics*, 110 (6): 2753-2756, **1999**

Martin, M G, and Siepmann, J I, Predicting multicomponent phase equilibria and free energies of transfer for alkanes by molecular simulation, *Journal of the American Chemical Society*, 119: 8921-8924, **1997**

Martin, M G, and Siepmann, J I, Transferable potentials for phase equilibria. 1. United-atom description of *n*-alkanes, *Journal of Physical Chemistry B*, 102 (14): 2569-2577, **1998**

Martin, M G, and Siepmann, J I, Novel Configurational-Bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes, *Journal of Physical Chemistry B*, 103 (21): 4508-4517, **1999a**

Martin, M G, Siepmann, J I, and Schure, M R, Simulating retention in gas-liquid chromatography, *Journal of Physical Chemistry B*, 103 (50): 11191-11195, **1999b**

Martin, M G, Zhuravlev, N D, Chen, B, Carr, P W, and Siepmann, J I, Origins of the solvent chain-length dependence of Gibbs free energies of transfer, *Journal of Physical Chemistry B*, 103 (15): 2977-2980, **1999c**

Martin, M G, Chen, B, and Siepmann, J I, Molecular Structure and Phase Diagram of the Binary Mixture of *n*-Heptane and Supercritical Ethane: A Gibbs Ensemble Monte Carlo Study, *Journal of Physical Chemistry B*, 104: 2415-2423, **2000**

Mckay, R A, Reamer, H H, Sage, B H, Lacey, W N, Volumetric and Phase Behavior in the Ethane-Propene System, *Industrial & Engineering Chemistry*, 43 (9): 2112-2117, **1951**

McQuarrie, D A, **1976**, "Statistical Mechanics", Harper and Row, New York

Mehta, M, and Kofke, D A, Molecular Simulation in a Pseudo Grand-Canonical Ensemble, *Molecular Physics*, 86 (1): 139-147, **1995**

Metropolis, N, Rosenbluth, A W, Rosenbluth, M N, Teller, A N, and Teller, E, Equation of state calculations by fast computing machines, *Journal of Chemical Physics*, 21: 1087-1092, **1953**

Miller, M A, Amon, L M, and Reinhardt, W P, Should one adjust the maximum step size in a Metropolis Monte Carlo simulation?, *Chemical Physics Letters*, 331 (2-4): 278-284, **2000**

Miyano, Y, Direct Determination of fluid phase equilibria by simulation in the Gibbs Ensemble, *Fluid Phase Equilibria*, 144: 137-144, **1997**

Mooij, G C A M, Frenkel, D, and Smit B, Direct simulation of phase equilibria of chain molecules, *Journal of Physics-Condensed Matter*, 4 (16): L255-L259, **1992**

Mooij, G C A M, and Frenkel, D; The overlapping distribution method to compute chemical-potentials of chain molecules, *Journal of Physics-Condensed Matter*, 6 (21): 3879-3888, **1994**

Mooij, G C A M, and Frenkel D, A systematic optimization scheme for Configurational-Bias Monte Carlo, *Molecular Simulation*, 17 (1): 41-55, **1996**

Moller, M A, Tildesley, D J, Kim, K S, and Quirke, N, Molecular-Dynamics simulation of a Langmuir-Blodgett-film, *Journal of Chemical Physics*, 94 (12): 8390-8401, **1991**

Möller, D, and Fischer, J, Vapour-liquid equilibrium of a pure fluid from test particle method in combination with NPT Molecular Dynamics simulations, *Molecular Physics*, 69 (3): 463-473, **1990**

Moore, G E, Cramming more components onto integrated circuits, *Electronics*, 38 (8), 114-117, **1965**

Murcko, M A, Castejon, H, and Wiberg, K B, Carbon-carbon rotational barriers in butane, 1-butene, and 1,3-butadiene, *Journal of Physical Chemistry* 100 (40): 16162-16168, **1996**

Nath, S K, Escobedo, F A, and de Pablo, J J, On the simulation of vapor-liquid equilibria for alkanes, *Journal of Chemical Physics*, 108 (23): 9905-9911, **1998a**

Nath, S K, Escobedo, F A, de Pablo, J J, and Patramai, I, Simulation of vapor-liquid equilibria for alkane mixtures, *Industrial & Engineering Chemistry Research*, 37 (8): 3195-3202, **1998b**

Nath, S K, and de Pablo, J J, Solubility of small molecules and their mixtures in polyethylene, *Journal of Physical Chemistry B*, 103 (18): 3539-3544, **1999**

Nath, S K, and de Pablo, J J, Simulation of vapour-liquid equilibria for branched alkanes, *Molecular Physics*, 98 (4): 231-238, **2000**

Nath, S K, Banaszak, B J, and de Pablo, J J, A new united atom force field for alpha-olefins, *Journal of Chemical Physics*, 114 (8): 3612-3616, **2001a**

Nath, S K, Banaszak, B J, and de Pablo, J J, Simulation of ternary mixtures of ethylene, 1-hexene, and polyethylene, *Macromolecules*, 34 (22): 7841-7848, **2001b**

Neumann, M, Dipole moment fluctuation formulas in computer simulations of polar systems, *Molecular Physics*, 50 (4): 841-858, **1983**

Nevins, N, Lii, J H, and Allinger, N L, Molecular mechanics (MM4) calculations on conjugated hydrocarbons, *Journal of Computational Chemistry* 17 (5-6): 695-729, **1996**

Nikitin, E D, and Popov, A P, Critical temperatures and pressures of linear alk-1-enes with 13 to 20 carbon atoms using the pulse-heating technique, *Fluid Phase Equilibria*, 166 (2): 237-243, **1999**

Noman, G E, and Filinov, V S, Investigation of phase transitions by a Monte Carlo method, *High Temperature* (USSR), 7: 216-222, **1969**

Panagiotopoulos, A Z, Direct Determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Molecular Physics*, 61: 813-826, **1987**

Panagiotopoulos, A Z, Quirke N, Stapleton, M, and Tildesley, D J, Phase equilibrium by computer simulations in the Gibbs Ensemble: Alternative derivation , generalization and applications to mixtures and membrane equilibria, *Molecular Physics*, 63: 527-545, **1988**

Panagiotopoulos, A Z, Exact calculations of fluid phase equilibria by Monte Carlo simulation in a new statistical ensemble, *International Journal of Thermophysics*, 10 (2): 447-457, **1989**

Panagiotopoulos, A Z, Direct Determination of fluid phase equilibria by simulation in the Gibbs Ensemble, *Molecular Simulation*, 9: 1-23, **1992**

Panagiotopoulos, A Z, Monte Carlo methods for phase equilibria of fluids, *Journal of Physics: Condensed Matter*, 12: R25-R52, **2000**

Pascual, P, Ungerer, P, Tavitian, B, Pernot, P, and Boutin, A, Development of a transferable guest-host force field for adsorption of hydrocarbons in zeolites - I. Reinvestigation of alkane adsorption in silicalite by Grand Canonical Monte Carlo simulation, *Physical Chemistry Chemical Physics*, 5 (17): 3684-3693, **2003**

Pascual, P, Ungerer, P, Tavitian, B, and Boutin, A, Development of a transferable guest-host force field for adsorption of hydrocarbons in zeolites. II. Prediction of alkenes adsorption and alkane/alkene selectivity in silicalite, *Journal of Physical Chemistry B*, 108 (1): 393-398, **2004**

Potoff, J J, and Panagiotopoulos, A Z, Critical point and phase behaviour of the pure fluid and a Lennard-Jones mixture, *Journal of Chemical Physics*, 109 (24): 10914-10920, **1998**

Potoff, J J, Errington, J R, and Panagiotopoulos, A Z, Molecular simulation of phase equilibria for mixtures of polar and non-polar components, *Molecular Physics*, 97 (10): 1073-1083, **1999**

Potoff, J J, and Siepmann, J I, Vapour-liquid equilibria of mixtures containing alkanes, carbon dioxide and water, *AIChE Journal*, 47 (7): 1676-1682, **2001**

Powles, J G, The liquid vapour coexistence line for Lennard-Jones-type fluids, *Physica*, 126: 289-299, **1984**

Rao, Y V C, **1994**, "Postulational and Statistical Thermodynamics", Allied Publishers, New Delhi

Ray, J R, and Graben, H W, Fundamental treatment of the isoenthalpic-isobaric ensemble, *Physical Review A*, 34 (3): 2517-2519, **1986**

Rouhi, A M, Olefin metathesis: Big-deal reaction, *Chemical & Engineering News*, 80 (51): 29-33, **2002**

Rowlinson, J S, and Swinton, F L, **1982**, "Liquid and Liquid mixtures, 3rd Edition", Butterworth Scientific, London

Rowlinson, J S, and Widom, B, **1989**, "Molecular Theory of Capillarity", Oxford University Press, New York,

Rull, L F, Jackson, G, and Smit, B, The condition of microscopic reversibility in Gibbs Ensemble Monte Carlo simulations of phase equilibria, *Molecular Physics*, 85 (3): 435-447, **1995**

Ryckaert, J P, and Bellemans, A, Molecular Dynamics of liquid normal-butane near its boiling point, *Faraday Discussions*, 30 (1): 123-125, **1975**

Ryckaert, J P, and Bellemans, A, Molecular Dynamics of liquid alkanes, *Faraday Discussions*, 66: 95-106, **1978**

Ryckaert, J P, Mcdonald, I R, and Klein, M L, Disorder in the pseudohexagonal rotator phase of normal-alkanes - Molecular-Dynamics calculations for tricosane, *Molecular Physics*, 67(5): 957-979, **1989**

Sadus, R J, **1999**, "Molecular Simulation of Fluids: Theory, Algorithms and Object-Orientation", Elsevier, Amsterdam

Shetty, R, and Escobedo, F A, On the application of virtual Gibbs Ensembles to the direct simulation of fluid-fluid and solid-fluid phase coexistence, *Journal of Chemical Physics*, 116 (18): 7957-7966, **2002**

Siepmann, J I, and Frenkel, D, Configurational-Bias Monte-Carlo - a new sampling scheme for flexible chains, *Molecular Physics*, 75 (1): 59-70, **1992**

Siepmann, J I, Karaborni, S, and Smit B, Simulating the critical-behavior of complex fluids, *Nature*, 365 (6444): 330-332, **1993**

Siever, E, Spainhour, S, Figgins, S, and Hekman, J P, **2000**, Linux in a Nutshell: A Desktop Quick Reference, 3rd Edition, O'Reilly, Sebastopol, USA

Smit, B, and Frenkel, D, Calculation of the chemical potential in the Gibbs Ensemble, *Molecular Physics*, 68 (4): 951-958, **1989**

Smit, B, de Smedt, P, and Frenkel ,D, Computer simulations in the Gibbs Ensemble, *Molecular Physics*, 68: 931-950, **1989**

Smit, B, and Frenkel, D, Vapour-liquid equilibria of the two-dimensional Lennard-Jones fluid(s), *Molecular Physics*, 68: 951-958, **1991**

Smit, B, Phase diagrams of Lennard-Jones fluids, *Journal of Chemical Physics*, 96: 8639-8640, **1992**

Smit, B, Karaborni, S, and Siepmann, J I, Computer simulations of vapour-liquid phase equilibria of *n*-alkanes, *Journal of Chemical Physics*, 102 (5): 2126-2140, **1995**

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

Smith, J M, van Ness, H C, and Abbott, M M, **1996**, "Introduction to Chemical Engineering Thermodynamics", 3rd edition, Mcgraw-Hill, Singapore

Smith, W R, and Triska, B, The Reaction Ensemble method for the computer simulation of chemical and phase equilibria. I. Theory and basic examples, *Journal of Chemical Physics*, 100 (4): 3019-3027, **1993**

Spencer, C F, and Danner, R P, improved equation for prediction of saturated liquid density, *Journal of Chemical and Engineering Data*, 17 (2): 236-241, **1972**

Spyriouni, T, Economou, I G, and Theodorou, D N, Molecular simulation of alpha-olefins using a new united-atom potential model: Vapor-liquid equilibria of pure compounds and mixtures, *Journal of the American Chemical Society*, 121 (14): 3407-3413, **1999**

Sterling, T, **2001**, "Beowulf Cluster Computing with Linux", The MIT Press, Cambridge, Massachusetts, USA

Stubbs, J M, Potoff J J, and Siepmann J I, Transferable potentials for phase equilibria. 6. United-atom description for ethers, glycols, ketones, and aldehydes, *Journal of Physical Chemistry B*, 108 (45): 17596-17605, **2004**

Swendsen, R H, Modern methods of analyzing Monte Carlo compute simulations, *Physica A*, 194: 53-62, **1993**

Szalay, P G, Lischka H, and Karpfen, A, Relative stabilities of the s-*cis* and gauche structures of 1,3-butadiene, *Journal of Physical Chemistry*, 93 (18): 6629-6631, **1989**

Tavares, F W, and Sandler, S I, Vapour-liquid equilibria of exponential-six fluids, *Molecular Physics*, 87(6): 1471-1476, **1996**

Theodorou, D N, and Suter, U W, Detailed molecular-structure of a vinyl polymer glass, *Macromolecules*, 18 (7): 1467-1478, **1985**

Toxvaerd, S, Equation of state of alkanes .2, *Journal of Chemical Physics*, 107 (13): 5197-5204, **1997**

Ungerer, P, Boutin, A, and Fuchs, A H, Direct calculation of bubble points by Monte Carlo simulation, *Molecular Physics*, 97 (4): 523-539, **1999**

Ungerer, P, Beauvais, C, Delhommelle, J, Boutin, A, Rousseau, B, and Fuchs, A H, Optimization of the anisotropic united atoms intermolecular potential for *n*-alkanes, *Journal of Chemical Physics*, 112 (12): 5499-5510, **2000**

Ungerer, P, Boutin, A, and Fuchs, A H, Direct calculation of bubble points for alkane mixtures by molecular simulation, *Molecular Physics*, 99 (17): 1423-1434, **2001**

Ungerer, P, From organic geochemistry to statistical thermodynamics: the development of simulation methods for the petroleum industry, *Oil & Gas Science and Technology*, 58 (2): 271-297, **2003**

van der Ploeg, P, and Berendsen, H J C, Molecular Dynamics simulation of a bilayer-membrane, *Journal of Chemical Physics*, 76 (6): 3271-3276, **1982**

Vargaftik, N B, **1975**, "Tables on the Thermophysical Properties of Liquid and Gases, 2nd Edition", Halstead Press, New York

Vlugt, T J H, Martin, M G, Smit, B, Siepmann, J I, and Krishna, R, Improving the efficiency of the Configurational-Bias Monte Carlo algorithm, *Molecular Physics*, 94 (4): 727-733, **1998**

Vlugt, T J H, personal communication, **2002**

Vrabec, J, and Hasse, H, Grand Equilibrium: vapour-liquid equilibria by a new molecular simulation method, *Molecular Physics*, 100 (21): 3375-3383, **2002**

Wakeham W A, St Cholakov, G, and Stateva, R P, Liquid density and critical properties of hydrocarbons estimated from molecular structure, *Journal of Chemical and Engineering Data*, 47 (3): 559-570, **2002**

Welsh, M, Dalheimer, M K, and Kaufman, L, **1999**, Running Linux, 3rd Edition, O'Reilly, Sebastopol, USA

Widom, B, Some topics in the theory of fluids, *Journal of Chemical Physics*, 39: 2802-2812, **1963**

Wick, C D, Martin, M G, and Siepmann, J I, Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and alkylbenzenes, *Journal of Physical Chemistry B*, 104 (33): 8008-8016, **2000**

Wood, W W, Monte Carlo simulation for hard disks in the isothermal isobaric ensemble, *Journal of Chemical Physics*, 48: 415-434, **1968**

Wu, G, and Sadus, R J, Molecular simulation of the high pressure phase equilibria of binary atomic fluid mixtures using the exponential-6 intermolecular potential, *Fluid Phase Equilibria*, 170(2): 269-284, **2000**

"Ah. Don't say that you agree with me. When people agree with me I always feel that I must be wrong" – Oscar Wilde