FUNDAMENTAL STUDIES AND ENGINEERING MODELING OF HYDROGEN BONDING

By

Aseel Mohamed Ahmed Bala Ahmed

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Chemical Engineering-Doctor of Philosophy

2018

ABSTRACT

FUNDAMENTAL STUDIES AND ENGINEERING MODELING OF HYDROGEN BONDING

By

Aseel Mohamed Ahmed Bala Ahmed

This project aims to enhance the engineering modeling of hydrogen bonding, or association, by blending ab initio quantum calculations, fundamental molecular level findings from experimental techniques, and thermodynamic models. Because of the ubiquity of hydrogen bonding, applications for an improved association model are extensive, ranging from drug design to plastics manufacturing. Therefore, a substantial amount of work has been aimed at improving traditional thermodynamic tools, which often fail to capture the behavior of associating systems accurately. To guide models, spectroscopic techniques have been leveraged to gain insight into the interactions between molecules in the liquid phase, but interpretation is difficult. Moreover, with the advancement of computational chemistry technology, molecular dynamics (MD) and quantum mechanical (QM) calculations have also been utilized to understand the characteristics of hydrogen bonded clusters. However, few studies have combined all 3 techniques (the thermodynamic model, spectroscopy and ab initio calculations) in a rigorous way. To this end, an activity coefficient model for association is developed using Wertheim's perturbation theory and its capabilities and limitations are explored with parameters from literature. Furthermore, a sequential MD and QM protocol is designed which facilitates the interpretation of the hydroxyl vibration in infrared spectroscopy and a method is developed to quantify the entire band. Finally, the methods are used to calculate the value of the association constant for an alcohol + alkane system.

This dissertation is dedicated to the memory of my mother, Amal Ziada, who gave everything and then some.

ACKNOWLEDGEMENTS

The work documented in this dissertation would not have been possible without the contributions and support of the following individuals. First and foremost, I owe my deepest gratitude to my advisor, Dr. Carl Lira, whose passions for thermodynamics and teaching have been a tremendous inspiration to me. His guidance, encouragement and patience were absolute, and I am thankful to have benefitted from his mentorship throughout my time as a graduate student. I would also like to thank my committee members, Drs. James "Ned" Jackson, Dennis Miller and Andre Lee for their valuable insight, assistance and contributions.

I acknowledge with gratitude the National Science Foundation for funding the project as well as our collaborators in industry, Drs. Tim Frank, Paul Mathias, Navin Patel, Eric Cheluget, Dung Vu, You Peng, and Suphat Watanasiri, for their input and advice. I am also grateful to have worked with my colleagues and friends, William G. Killian, Jackson Storer, Renming Liu and Dr. Lars Peereboom, who have assisted with data collection and contributed greatly to this work. In my time at Michigan State University, I have been fortunate enough to meet many wonderful

people who became my support system, sounding boards and friends. I would like to thank Amrita Oak, Souful Bhatia, Annalisa Grunwald, Iman Nezam, Neda Rafat, GracieLou Klinger and Aritra Sarkar for the many laughs, sound advice and great company.

Finally, I would like to express my sincerest gratitude for the life-long love and encouragement of my parents, Amal Ziada and Mohamed Ahmed Bala, my siblings, Asem, Leena, and Sarya, and my aunt, Iman Ziada. To this list, I add the friends that have played just as great a role in this journey: Sarah Gard and Fatima El-Hussain. I am thankful for their kindness, compassion and unwavering support throughout the many years of our friendship.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	X
KEY TO SYMBOLS AND ABBREVIATIONS	XV
Chapter 1. Introduction to Thermodynamic Modeling of Associating Systems	
1.1. Background	
1.1.1. Equations of state (EOS)	4
1.1.2. Activity coefficient models.	
1.1.3. Association modeling	ð
1.2. Research objectives	9
1.5. Dissertation outline	10
Chapter 2. Relation of Concentration-based Equilibrium Constants to Activity-based E	quilibrium
Constants in the Liquid Phase	
2.1. Introduction	
2.2. Gibbs energy and the activity-based equilibrium constant	
2.3. Gas phase activity-based equilibrium constant	14
2.4. Liquid phase activity-based equilibrium constant	
APPENDICES	
Appendix A: Development of the activity-based equilibrium constant for atherma	1 solutions
Annandix D: Effact of mass density on volume of mixing	
Appendix B. Effect of mass density on volume of mixing	
Chapter 3 Relation of Wertheim Association Constants to Concentration-based F	auilibrium
Constants for Mixtures with Chain-forming Components	22
3.1 Introduction	22
3.1.1. Motivation	
3.1.2. Historical perspective	
3.2. Theory and analysis	
3.2.1. Association term – Wertheim's theory	
3.2.2. Flory's concentration-based equilibrium constant	
3.2.3. Relationship between K_a , K_c and Δ	
3.2.4. Association term – Campbell's chemical theory model	
3.2.5. Algebraic equivalence of chemical and Wertheim's theory	
3.2.6. Calculating association parameters	
3.2.7. Numerical equivalence of chemical and Wertheim's theory	
3.3. Conclusions	
APPENDICES	
Appendix C: Derivation of the Wertheim association contribution to the activity	coefficient

Appendix D: Relationship between the fugacity coefficient and true compressi	bility factor $\Delta \Delta$
Appendix E: Relation of Campbell's concentration variable in terms of concentration of species	the molar
Chapter 4. Relation of Wertheim and Chemical Theories in Systems with Asso	ociation and
Solvation	
4.1. Introduction	
4.2. Theory and analysis	50
4.2.1. Implementation of Wertheim's perturbation theory	50
4.2.2. Implementation of chemical theory	
4.2.3. Numerical equivalence of chemical and Wertheim's theories	
4.2.4. Activity coefficients	64
4.2.5. Perspective on further generalization	66
4.3. Conclusions	67
APPENDICES	69
Appendix F: Simplification of double sums without a multiplication factor	
Appendix G: Simplification of double sums with a multiplication factor	71
Charten 5. Annlingtions of a Wardheim and initial and initial and finite set of initial set of the set	- 41 1
Chapter 5. Applications of a wertheim association activity coefficient model to m	ethanol and
ethanol-containing mixtures	
5.1. Introduction	
5.2. Model theory and parameters	
5.2.1. Association contribution	
5.2.2. Combinatorial contribution	
5.2.3. Residual contribution	
5.2.4. EOS model parameters	
5.2.5. Association parameters	
5.3. Model comparisons	
5.3.1. Binary system with no association	
5.3.2. Binary systems with self-association	
5.3.3. Binary systems with self- and cross-associating systems	
5.3.4. Ternary systems	
5.3.5. General findings and recommendations	
5.4. Association parameter value	
5.5. Conclusions	
APPENDIX	
Chapter 6. Quantitative Analysis of Infrared Spectra of n -Butanol + Cyclobevane wi	th Quantum
Chemical Calculations	126
6.1 Introduction	126
6.1.1 Analysis of the O-H stretching hand	120
6.1.2 Motivation	120
6.2 Computational methods	130
6.2.1 Molecular dynamics simulations	121
6.2.2. Quantum mechanics simulations	
0.2.2. Quantum mechanics sinulations	

6.3. Experimental methods	134
6.4. Results and discussions	135
6.4.1. Processing and preliminary analysis of experimental IR spectra	135
6.4.2. Results from MD+QM	138
6.4.3. Scaling and further analysis of experimental IR spectra	144
6.4.4. Universality of spectroscopic characteristics	148
6.5. Conclusions	150
Chapter 7. Integration of Quantum Calculations and Spectroscopy for Wertheim	Alcohol
Association	151
7.1. Introduction	151
7.1.1. Hydrogen bonding and infrared spectroscopy	151
7.1.2. Thermodynamic modeling of hydrogen bonding	155
7.1.3. Motivation	157
7.2. Experimental methods	157
7.3. Results and discussions	158
7.3.1. IR band assignments	164
7.3.2. IR peak fitting	167
7.3.3. IR calculation of association parameters	175
7.3.4. NMR calculations and results	178
7.4. Summary of future work	183
APPENDIX	185
Chapter 8. Conclusions and Future Directions	189
BIBLIOGRAPHY	193

LIST OF TABLES

Table 3.1: Concentration based equilibrium constants	. 30
Table 3.2: Equilibrium constant for primary alcohols	. 37
Table 3.3: Calculated parameters and intermediate values for Flory and Wertheim's theories three systems.	for . 39
Table 4.1: X^{B_i} Equations for sites in hypothetical system	. 52
Table 4.2: Key associations in the hypothetical system and their equilibrium constants	. 53
Table 4.3: Other species formed in hypothetical system	. 54
Table 4.4: Association parameters and component densities	. 57
Table 4.5: Calculated monomer fraction values	. 58
Table 4.6: Chemical and Wertheim's theory concentration equations and calculated values at x_{I} 0.2, x_{F} = 0.2 and $x_{G_{M}}$ =0.6	_E = . 59
Table 4.7: Calculated parameters and intermediate values for activity coefficient calculations	. 65
Table 5.1: CPA parameters used to calculate Δ^{AD}	. 83
Table 5.2: AspenPlus® databases used in calculations	. 85
Table 5.3: PC-SAFT pure component parameters	. 85
Table 5.4: Data used for parameter regression	. 86
Table 5.5: Total SSQ for systems studied in this work	. 87
Table 5.6: Model performance in order of best fit	. 88
Table 5.7: Legend for plots in Model Capabilities section	. 89
Table 5.8: Association contribution to the total activity coefficient at infinite dilution	116
Table H.1: Regressed parameters values.	123
Table H.2: Hayden-O'Connell η_{ij} parameters	125

Table 6.1: MD simulation details
Table 6.2: Number of each species and bond analyzed with QM calculations 138
Table 6.3: Average calculated vibrational frequencies and integrated absorption coefficients . 140
Table 6.4: Areas and standard deviations for the absorbance and scaled absorbance bands 147
Table 7.1: Frequency range defined as hydroxyl band for each spectrum
Table 7.2: Number of each species and bond analyzed with QM calculations for the distribution plot in Figure 7.7 and Figure 7.8
Table 7.3: Concentration of free ends $(\alpha + \beta)$ in mol/mL calculated from fitted Gaussian peaks
Table 7.4: Regressed parameters from NMR analysis and corresponding CPA parameters for comparison 182 182 182
Table 8.1: Accomplished and future goals for research project 190

LIST OF FIGURES

Figure 1.1: NRTL representation of VLE and LLE of methanol (1) – cyclohexane (2) 2
Figure 1.2: Vapor pressure as a function of temperature for ethanol calculated using two sets of CPA parameters
Figure 3.1: 2B association scheme
Figure 4.1: 2B association scheme
Figure 4.2: Hypothetical molecules for analysis
Figure 4.3: Bonding schemes for hydrogen bonding in a mixture of <i>E</i> , <i>F</i> and <i>G</i> 53
Figure 4.4: Counting species in chemical theory system
Figure 4.5: True monomer fraction of components in system of study and fractions of acceptors and donors
Figure 4.6: Distribution of species with component <i>E</i> when $x_F=0.2$
Figure 4.7: Distribution of species with component <i>F</i> when $x_F = 0.2$
Figure 4.8: Distribution of species with component <i>G</i> when $x_F = 0.2$
Figure 4.9: True mole fraction of methanol chains calculated with 4 values of $\Delta^{A_G D_G}$
Figure 4.10: Calculated values of the activity coefficient contributions
Figure 5.1: Phase equilibria diagram for <i>n</i> -heptane + cyclohexane at 101.3 kPa90
Figure 5.2: Heat of mixing for <i>n</i> -heptane + cyclohexane at T=298.14 K and 101.3 kPa90
Figure 5.3: Calculated and experimental activity coefficients for <i>n</i> -heptane + cyclohexane at 101.3 kPa
Figure 5.4: Percent errors of K-ratios for <i>n</i> -heptane + cyclohexane system
Figure 5.5: Phase equilibria diagram for methanol $+ n$ -heptane at 101.3 kPa
Figure 5.6: Heat of mixing for methanol + n -heptane at T=298.15 K and 101.3 kPa
Figure 5.7: Calculated and experimental [4] activity coefficients for methanol + <i>n</i> -heptane at 101.3 kPa

Figure 5.8: Percent errors of K-ratios for methanol $+ n$ -heptane system
Figure 5.9: Phase equilibria diagram for methanol + cyclohexane at 101.3 kPa
Figure 5.10: Heat of mixing for methanol + cyclohexane at T=416.29 K and 1900 kPa95
Figure 5.11: Calculated and experimental activity coefficients for methanol + cyclohexane at 101.3 kPa
Figure 5.12: Percent errors of K-ratios for methanol + cyclohexane system
Figure 5.13: Phase equilibria diagram for methanol + <i>n</i> -pentane at T=397.7 K
Figure 5.14: Calculated and experimental [145] activity coefficients for methanol + <i>n</i> -pentane at T=397.7 K
Figure 5.15: Percent errors of K-ratios for methanol + <i>n</i> -pentane system
Figure 5.16: Phase equilibria diagram for ethanol + cyclohexane at 101.3 kPa
Figure 5.17: Heat of mixing for ethanol + cyclohexane at T=298.14 K and 101.3 kPa
Figure 5.18: Calculated and experimental activity coefficients for ethanol + cyclohexane at 101.3 kPa
Figure 5.19: Percent errors of K-ratios for ethanol + cyclohexane system
Figure 5.20: Phase equilibria diagram for ethanol + n -decane at T=338.17 K 101
Figure 5.21: Phase equilibria diagram for ethanol + <i>n</i> -decane at 101.3 kPa
Figure 5.22: Heat of mixing for ethanol + n -decane at T=298.15 K and 101.3 kPa 102
Figure 5.23: Phase equilibria diagram for methanol + water at 101.3 kPa 104
Figure 5.24: Phase equilibria diagram for methanol + water at T=298.14 K (bottom) and T=308.14 K (top)
Figure 5.25: Calculated and experimental [146,147] activity coefficients for methanol + water at T=298.14 K, T=308.14 K and P=101.3 kPa from top to bottom
Figure 5.26: Percent errors of K-ratios for methanol + water system 106
Figure 5.27: Phase equilibria diagram for methanol + ethanol at T=298.14 K (bottom) and T=413.13 K (top)
Figure 5.28: Phase equilibria diagram for methanol + ethanol at T=298.14 K 108

Figure 5.29: Calculated and experimental activity coefficients for methanol + ethanol at T=298.14 K (bottom) and T=413.13 K (top)
Figure 5.30: Percent errors of K-ratios for methanol + ethanol system
Figure 5.31: Phase equilibria diagram for methanol + <i>n</i> -heptane + cyclohexane at 101.3 kPa and 298.15 K
Figure 5.32: K-ratios for methanol (a) + <i>n</i> -heptane (b) + cyclohexane (c) system at 101.3 kPa and 298.15 K
Figure 5.33: Phase equilibria diagram for methanol + ethanol + cyclohexane at 101.3 kPa and 298 K
Figure 5.34: K-ratios for methanol (a) + ethanol (b) + cyclohexane (c) system at 101.3 kPa and 298 K
Figure 5.35: Comparison of phase equilibria fits with two sets of CPA parameters for ethanol + cyclohexane at 101.3 kPa
Figure 5.36: Comparison of phase equilibria fits with two sets of CPA parameters for ethanol + <i>n</i> -decane at T=338.17 K
Figure 5.37: Δ^{AD} for the self-association of the primary alcohols at 298.15 K calculated by fitting to different properties
Figure 6.1: Types of covalent O-H bonds
Figure 6.2: Bond distribution in trimers calculated with MD simulations for an equimolar <i>n</i> -butanol + cyclohexane mixture
Figure 6.3: Experimental O-H IR band for a 0.1 mole fraction of <i>n</i> -butanol in cyclohexane 136
Figure 6.4: Total absorbance band area of the O-H band for <i>n</i> -butanol + cyclohexane data as a function of the mole fraction for 4 compositions
Figure 6.5: Hydroxyl stretching frequencies and integrated absorption coefficients for two concentrations of n -butanol + cyclohexane mixtures calculated from QM simulations
Figure 6.6: Hydroxyl stretching frequencies and integrated absorption coefficients for other systems calculated from QM simulations
Figure 6.7: Scaled absorbance spectra for <i>n</i> -butanol + cyclohexane at $x_{BuOH} = 0.1$
Figure 6.8: Scaled absorbance band area of the O-H band for <i>n</i> -butanol + cyclohexane data as a function of the mole fraction of 4 experimental compositions

Figure 6.9: Relationship between the covalent O-H bond length and vibrational frequency calculated for equimolar ethanol + cyclohexane mixture
Figure 6.10: Comparison of calculated NMR and IR characteristics for equimolar ethanol + cyclohexane mixture
Figure 7.1: Types of covalent OH bonds 153
Figure 7.2: Hydroxyl stretching frequencies and integrated absorption coefficients for <i>n</i> -butanol + and ethanol + cyclohexane mixtures calculated from QM simulations
Figure 7.3: Raw (top) and scaled (bottom) O-H IR band for a 0.0469 mole fraction of <i>n</i> -butanol in cyclohexane
Figure 7.4: Raw (top) and scaled (bottom) O-H IR band for a 0.0678 mole fraction of <i>n</i> -butanol in cyclohexane
Figure 7.5: Raw (top) and scaled (bottom) O-H IR band for a 0.0817 mole fraction of <i>n</i> -butanol in cyclohexane
Figure 7.6: Raw (top) and scaled (bottom) O-H IR band for a 0.100 mole fraction of <i>n</i> -butanol in cyclohexane
Figure 7.7: Smoothed normalized distributions of bond types from QM simulations of <i>n</i> -butanol + and ethanol + cyclohexane mixtures
Figure 7.8: Smoothed normalized distributions of bond types from QM simulations with γ from dimers separated from others
Figure 7.9: Peaks fitted to scaled IR spectra for <i>n</i> -butanol + cyclohexane mixture at x_{BuOH} =0.0469. 168
Figure 7.10: Peaks fitted to scaled IR spectra for <i>n</i> -butanol + cyclohexane mixture at x_{BuOH} =0.0678
Figure 7.11: Peaks fitted to scaled IR spectra for <i>n</i> -butanol + cyclohexane mixture at x_{BuOH} =0.0817
Figure 7.12: Peaks fitted to scaled IR spectra for <i>n</i> -butanol + cyclohexane mixture at x_{BuOH} =0.100. 171
Figure 7.13: Scaled band area as a function of apparent <i>n</i> -butanol concentration
Figure 7.14: Bond distributions calculated from IR peak fitting
Figure 7.15: Concentration of α bonds vs. mole fraction of <i>n</i> -butanol
Figure 7.16: Calculated and fitted concentrations of free end ($\alpha + \beta$) hydroxyls

Figure 7.17: Derived association parameter values from IR analysis (solid line) compared to CPA ΔAD values (dashed line) for the self-association of <i>n</i> -butanol
Figure 7.18: O-H proton chemical shift as a function of <i>n</i> -butanol mole fraction for <i>n</i> -butanol + cyclohexane mixtures at 26.3 °C
Figure 7.19: O-H proton chemical shift as a function of <i>n</i> -butanol mole fraction for <i>n</i> -butanol + cyclohexane mixtures at 34 °C
Figure 7.20: O-H proton chemical shift as a function of <i>n</i> -butanol mole fraction for <i>n</i> -butanol + cyclohexane mixtures at 41 $^{\circ}$ C

KEY TO SYMBOLS AND ABBREVIATIONS

а	interaction parameter for molecules in a mixture, calculated with mixing rules
a _i	interaction parameter for i number of molecules of a single component in Chapter 3; activity of species i otherwise
a _{ij}	interaction parameter for interactions between molecules of components i and j
<i>a</i> ₀	CPA parameter for attractive term
ã	fitted spectroscopic parameter
ã _{ij}	fitted residual term parameter
A	molar Helmholtz energy (extensive if accompanied by underbar); absorbance in Chapters 6, 7
A _i	absorbance of species i
Ã	scaled absorbance
A _{ij}	fitted residual term parameter
AVEC	association with variable equilibrium constant
b	size parameter of mixture, calculated with mixing rules
b_i	size parameter for <i>i</i> number of molecules of a single component
\widetilde{b}	fitted spectroscopic parameter
$ ilde{b}_{ij}$	fitted residual term parameter
B^{free}	free contribution to the second virial coefficient
<i>c</i> ₁	Intermediate CPA parameter for attractive term
ĩ	fitted spectroscopic parameter
C _i	concentration of component i (apparent unless accompanied by subscript T)
C'_i	apparent dimensionless concentration of component <i>i</i>

C'_{T,M_i}	dimensionless true monomer concentrations of component <i>i</i> in a mixture
$C_{T,M_i}^{\prime 0}$	dimensionless true monomer concentrations of component i in a pure solution of component i
CLAM	continuous linear association model
СРА	cubic-plus-association model
Ĩ	fitted spectroscopic parameter
DFT	density functional theory
EOS	equation of state
ECR	Elliott's combining rule
ESD	Elliott-Suresh-Donohue model
f	constant in Eq. (3.14)
f_i	fugacity of component <i>i</i>
\widetilde{f}	fitted spectroscopic parameter
g	function defined by Eq. (3.18) in Chapter 3; radial distribution function otherwise
\widetilde{g}	fitted spectroscopic parameter
G	molar Gibbs energy (extensive if accompanied by underbar)
h	function defined by Eq. (C.1)
\widetilde{h}	fitted spectroscopic parameter
Н	molar enthalpy (extensive if accompanied by underbar)
IR	infrared spectroscopy
k	Boltzmann's constant
k _f	reaction rate constant for forward reaction
k _r	reaction rate constant for reverse reaction

k _{ij}	binary interaction parameter
K	equilibrium constant
$K^{A_i D_j}$	fitted association parameter with units of volume
KLL	K-ratio in the liquid-liquid region
KVL	K-ratio in the vapor-liquid region
l	pathlength
L	generic notation for a regressed property
L _{OH}	length of the hydroxyl bond
LLE	liquid-liquid equilibrium
LACT	linear association with cyclic trimer
LCST	lower critical solution temperature
m	number of segments per chain in PC-SAFT
MD	molecular dynamics
MW _i	molecular weight of species <i>i</i>
n	number of moles
$n^{B_i,free}$	number of non-bonded sites of type B_i
$n^{B_i,total}$	total number of sites of type B_i in solution
<i>n</i> *	effective number of carbon atoms in an ether homomorph molecule
N _i	number of molecules of component <i>i</i>
N _{ij}	local composition, number of component i molecules around molecule j
$N_{B,i}$	number of sites identical to site i of type B on the site host
N _{matej}	synonymous to $N_{B,i}$, introduced for disambiguation in Eq. (4.1)
NMR	nuclear magnetic resonance spectroscopy

NPT	condition holding number of moles, pressure and temperature constant
NRTL	non-random two liquid model
NVT	condition holding number of moles, volume and temperature constant
Р	probability in Chapter 4, pressure otherwise
P _i	Gaussian peak fitted to IR
P_i^{vap}	vapor pressure of component <i>i</i>
PC-SAFT	perturbed chain – statistical associating fluid theory
PM1, PM2	CPA pure component parameters from Kontogeorgis <i>et al.</i> [1] and [2] respectively
PME	Particle Mesh Ewald
QM	quantum mechanics
<i>r</i> _i	ratio of reference molar density, ρ^{ref} (taken to be that of methanol at 303.15 K) to the molar density of component <i>i</i> , ρ_i
R	universal gas constant
stddev	standard deviation
S	molar entropy of solution
S _i	molar entropy of component <i>i</i>
SAFT	statistical associating fluid theory model
SG	Staverman-Guggenheim model
SH	Scatchard-Hildebrand model
SSQ	sum of squares error
t	time
Т	temperature
U	molar internal energy (extensive if accompanied by underbar)

UNIFAC	UNIQUAC functional-group activity coefficient model
UNIQUAC	universal quasichemical model
V_i	molar volume of pure component <i>i</i>
V	molar volume of solution, extensive volume if accompanied by underbar
VLE	vapor-liquid equilibrium
VLLE	vapor-liquid-liquid equilibrium
WAG	Wertheim association gamma model
x _i	apparent liquid mole fraction of component <i>i</i>
x _{ij}	mole fraction of component i molecules around a molecule of component j
X^{B_i}	fraction of sites identical to site i of type B that remain non-bonded at equilibrium
y_i	vapor phase mole fraction of species <i>i</i>
Ζ	true mole fraction (for monomer in Chapter 4)
Ζ	compressibility factor
Greek Symbols	
α	type of O-H bond in Chapters 6 and 7 given in Figure 6.1
α_{ij}	non-randomness parameter in NRTL model
$\alpha_{j,i}$	function defined by Eq. (E.2)
β	type of O-H bond in Chapters 6 and 7 given in Figure 6.1
$\beta^{A_i D_j}$	association volume parameter
γ	activity coefficient; type of O-H bond in Chapters 6 and 7 given in Figure 6.1
δ	type of O-H bond in Chapters 6 and 7 given in Figure 6.1
δ_i	solubility parameter in Scatchard-Hildebrand
Δ	Wertheim's association parameter
	xix

Δx_H	displacement of hydrogen atom in x-direction
Δy_H	displacement of hydrogen atom in y-direction
Δz_H	displacement of hydrogen atom in z-direction
$\Delta_{\widetilde{ u}}$	observed NMR chemical shift relative to monomer NMR chemical shift
$\Delta_{\widetilde{v}_N}$	NMR chemical shift of hydrogen bonded proton relative to monomer NMR chemical shift
ϵ	depth of pair potential in PC-SAFT
$\epsilon^{A_i D_j}$	fitted association interaction energy parameter
ε	molar absorption coefficient
ε_i	molar absorption coefficient of species <i>i</i>
ζ	type of O-H bond in Chapters 6 and 7 given in Figure 6.1
η	packing fraction; type of O-H bond in Chapters 6 and 7 given in Figure 6.1
θ	thermal correction factor
$\kappa^{A_i D_j}$	effective association volume in PC-SAFT
Λ_{ji}	rearrangement of Ω_{ij} used in Wilson's equation defined by Eq. (5.11)
ν_i	stoichiometric number of species <i>i</i>
$\tilde{\nu}$	wavenumber, $\tilde{\nu} = \text{frequency } (\nu)/\text{speed of light } (c)$
$\widetilde{ u}'$	NMR chemical shift
ρ	molar solution density (apparent unless accompanied by subscript T)
$ ho_i$	molar density of component <i>i</i> (apparent unless accompanied by subscript T)
$ au_{ij}$	interaction energy parameter between components i and j
σ	size parameter, <i>i.e.</i> contact distance
σ_{ii}	temperature-independent segment diameter in PC-SAFT

φ_i	fugacity coefficient of pure component <i>i</i>
Φ_i	volume fraction of component or species i (apparent unless accompanied by subscript T)
Φ'_i	modified apparent volume fraction of component or species <i>i</i>
ψ	integrated absorption coefficient
Ψ	total true dimensionless concentration of complexes (Campbell's work)
Ψ_i	true dimensionless concentration of all chains that end in i (Campbell's work)
Ω_{ij}	local composition parameter defined by Eq. (5.7)
	Subscripts
0	pertaining to apparent species
γ	activity-coefficient
a	activity-based

- *alc* pertaining to the alcohol
- avg averaged
- C concentration-based
- CPA pertaining to CPA
- Crit critical property
- Ddimerfpertaining to the formation of a reaction producthhypothetical fluid in Brandani's workhostpertaining to component hosting the sitemaxat maximum peak heightmate_jpertaining to site j that can bond to B_i
- mix mixture property

Μ	monomer
$n_{j \neq k}$	constant number of moles for all components except component k
Nmer	complex (hydrogen-bonded species)
р	constant power
Р	at constant pressure
Т	pertaining to true species or at constant temperature
Tri	trimer
V	at constant volume
x	mole fraction based
	Superscripts
,	dimensionless (used for Campbell's concentration)
0	standard state property
0	property for pure component
assoc	association (Wertheim's) contribution
att	attractive contribution
$B_i E_j$	interaction between site i of type B and site j of type E
С	cyclic species
chem	chemical theory
comb	combinatorial contribution
exp	experimental value
Ε	excess property
Flory	pertaining to or calculated with Flory's theory
hs	hard sphere

ig	ideal gas
is	ideal solution
l	linear species
L	pertaining to a liquid
Mod Flory	calculated with modified Flory's theory
Nagata	calculated with Nagata's model
NRTL	calculated with NRTL
p	constant for Eq. (5.6)
ref	pertaining to a reference state
rep	repulsive contribution
res	residual contribution
sat	property at saturation
SH	calculated with Scatchard-Hildebrand
Т	total
vap	pertaining to a vapor
Wilson	calculated with Wilson's equation
	Special Notations
as in <u>A</u>	extensive property
as in \overline{V}	partial molar property
$$ as in $\hat{\varphi}$	property of component in a mixture

[*i*] apparent concentration of component *i*

Chapter 1. Introduction to Thermodynamic Modeling of Associating Systems

Thermodynamic models play an essential role in the chemical industry. These tools provide valuable information on the qualitative and quantitative responses of a system to variations in temperature, pressure and composition. As such, they are necessary in the design of process equipment, allowing engineers to extrapolate the system's behavior to conditions that have not or cannot be investigated with bench-scale experiments. Of particular importance is the use of thermodynamic models for the design and operation of separation units, such as distillation columns, which can constitute as much as 70% of a plant's costs. Due to this high expense, the optimization of each component of a separation process, including the accuracy and robustness of the thermodynamic models, is an active area of research.

In today's chemical industry, process streams contain more polar components than ever before. Dow Chemical and Fluor manufacture large quantities of alcohols, ketones and esters. Honeywell UOP is interested in biorefining pyrolysis oils to fuels, a process that includes three-phase mixtures that are difficult to model. In cases like these where one or more components in a system can hydrogen bond, the relatively strong associations within and between molecules complicate thermodynamic modeling. Indeed, traditional models do not specifically account for hydrogen bonding between molecules, but instead lump effects into dispersion terms of the equations. For example, the non-random two liquid (NRTL) [3] model is commonly used in industry due to its relative simplicity and flexibility. This model represents molecular asymmetry with up to three parameters per binary, two for pair interactions at infinite dilution and one to adjust model to deviations from ideality at intermediate compositions. However, because the binary parameters are intended to represent relatively nondirectional dispersion interactions the model can fail to accurately represent components that hydrogen bond. This means that, when extrapolated to new sets of conditions, NRTL can perform poorly and can therefore not be used with confidence. This is illustrated in Figure 1.1 for the vapor-liquid (VLE) and liquid-liquid equilibria (LLE) of methanol in cyclohexane. When the VLE envelope is fitted (solid lines), the same parameters yield a poor prediction of the LLE, incorrectly indicating a three-phase vapor-liquid-liquid (VLLE) condition which would occur where the LLE binodal lines extrapolate to the VLE. Conversely, when the LLE is fitted (dashed lines), the model falsely predicts a higher temperature azeotrope and improperly shifts the VLE.



Figure 1.1: NRTL representation of VLE and LLE of methanol (1) – cyclohexane (2).Solid lines are fitted to the VLE envelope ($\tau_{12} = 661.8/T$, $\tau_{21} = 753.425/T$, and $\alpha_{12} = 0.422$) and dashed lines are fitted to the LLE envelope ($\tau_{12} = 0.7523 - 5.772/T$, $\tau_{21} = 147.8 + 2204/T$, and $\alpha_{12} = 0.2$). Markers are experimental VLE [4] and LLE [5] phase equilibria data.

To make models more flexible and robust, it has been demonstrated [6,7] that hydrogen bonding is best calculated explicitly as a separate term.

1.1. Background

Hydrogen bonding is a type of dipole-dipole interaction that occurs between a hydrogen atom bonded to a 'proton donor' and a neighboring strongly electronegative 'proton acceptor'. While weak compared to a covalent bond, this interaction is responsible for the higher melting and boiling points of water and alcohols relative to hydrocarbons with similar molecular weights.

In thermodynamics literature, the terms "association", "cross-association", and "solvation" are used to describe the phenomenon of hydrogen bonding. Association, or "self-association", refers to hydrogen bonding between two molecules of the same component, such as between two methanol molecules or between two water molecules. Cross-association describes hydrogen bonding between molecules of two different components which are also able to self-associate, such as an interaction between a water molecule and a methanol molecule. The term solvation is used in cases where one or both components involved in a hydrogen bond do not self-associate. Examples of solvating systems are water + ethers or chloroform + ethers. Though these definitions describe distinct behavior, the terms are often used loosely in literature. Frequently, the term "association" is used for convenience when referring to all three types of hydrogen bonding collectively and we adopt this convention in the present document.

To provide context for current state of the art in association modeling, we describe two categories under which most thermodynamic models fall: equations of state (EOS) and activity coefficient (γ) models. Thermodynamic modeling is based on quantifying departures or difference from idealized mixture states. The primary difference between the two classes is the choice of idealized state. Ideality, as it refers to ideal gases or solutions, is a simplified hypothetical state which is used as a conceptual reference point in thermodynamic model design. From this starting point,

contributions are incorporated to calculate the differences between ideal and real solution dynamics. In many cases, the major contribution to non-ideality is due to hydrogen bonding.

1.1.1. Equations of state (EOS)

Equations of state (EOSs) use departures from ideal gases to model mixture behavior. An ideal gas is a collection of point particles which do not interact with one another by collision, repulsion or attraction. The fugacity of each component of a gas mixture, \hat{f}_i , is given by [8]

$$\hat{f}_i = y_i \hat{\varphi}_i P \tag{1.1}$$

where y_i is the vapor phase mole fraction of component *i*, *P* is the total pressure. The component fugacity coefficient, $\hat{\varphi}_i$, quantifies the departure from an ideal gas mixture, and is unity for an ideal gas. Equations of state (EOS), such as the van der Waals and Peng-Robinson equations of state, calculate deviations from this idealized state by adding the effects of repulsive and attractive forces and association in solution. As such, the fugacity coefficient is separated as

$$\ln \hat{\varphi}_i = \ln \hat{\varphi}_i^{\ rep} + \ln \hat{\varphi}_i^{\ att} + \ln \hat{\varphi}_i^{\ assoc} - \ln Z$$
 1.2

where Z = PV/(RT) is the compressibility factor and can also be separated into the same contributions.

$$\ln \hat{\varphi}_i = \ln \hat{\varphi}_i^{rep} + \ln \hat{\varphi}_i^{att} + \ln \hat{\varphi}_i^{assoc} - \ln(1 + Z^{rep} + Z^{att} + Z^{assoc})$$
 1.3

Classical EOSs omit the association term entirely. However, newer models such as the Statistical Associating Fluid Theory (SAFT) [9], Elliott-Suresh-Donohue (ESD) [10] and Cubic-Plus-Association (CPA) [11] were developed to include hydrogen bonding explicitly. While the theoretical basis of these models is solid, and they are popular in academics, there are three barriers that impede their wide adoption in the chemical industry.

The first barrier that limits wide spread use is that the parameters for these models are determined by fitting to a large quantity of pure component vapor pressure and density experimental data, which are frequently unavailable to industry over a suitable span. When new compounds are developed, data is expensive to obtain over wide ranges, and limited data results in significant uncertainty in parameter fitting.

The second barrier to wider adoption is that dispersion forces and the association forces are both 'attractive', so proper attribution is difficult to discern when fitting parameters. This challenge is demonstrated for ethanol in Figure 1.2 using calculations of the pure component vapor pressure as a function of temperature.



Figure 1.2: Vapor pressure as a function of temperature for ethanol calculated using two sets of CPA parameters. Blue markers are experimental data, black dotted line are calculated with parameters from Kontogeorgis (2006) [1] (referred to in Chapter 5 as PM1) and dashed lines are calculated with parameters from Kontogeorgis (2010) [2] (referred to in Chapter 5 as PM2). The bottom plot is a zoom-in of the top plot.

The two black lines are both calculated using the CPA EOS and the only difference is the fitting method used to obtain the pure component parameters. For the dotted line, the parameters were fitted to pure vapor pressure and density (PM1) [1] and, for the dashed line, parameters were fitted to vapor pressure, density, and spectroscopic data (PM2) [2]. The numerical values of the parameters in PM2 are up to 37% different than those in PM1.

In both fits presented in Figure 1.2, the CPA parameters were fitted to data over a wide temperature range. Despite this, the PM1 parameters clearly fit the vapor pressure much better than PM2 parameters. Indeed, the PM2 fit underestimates the vapor pressure of ethanol by 22 kPa at T=338.17 K when the constraint is added to also represent spectroscopic data. The challenge in fitting indicates that several reasonable solutions exist for the fitting problem and that it is underspecified, and still may not represent all properties accurately. Even when fitted to a large amount of data, the parameters are sometimes not known with confidence. Making small adjustments to one parameter requires refitting of all parameters.

The third obstacle for more wide-spread adoption of EOSs is that the ability to represent mixtures is sensitive to accuracy of fitting the pure component property data. For example, if the parameters fit vapor pressure poorly at purity such as ethanol with the PM2 parameter set, the resulting mixture phase diagram will be skewed and fail to match experiment, even if the mixture deviations from mixture ideality are correct. This difficulty is further discussed in Chapter 5.

1.1.2. Activity coefficient models

In contrast to EOSs, activity coefficient (γ) models, such as NRTL and UNIversal QUAsiChemical (UNIQUAC) [12], calculate deviations from an ideal solution state and are commonly used for liquid states. For this approach, component fugacities of each component *i* in solution are calculated relative a standard state liquid fugacity, f_i° , often set to be that of a pure component.

6

$$\hat{f}_i = x_i \gamma_i f_i^{\circ} \qquad 1.4$$

where x_i is the mole fraction of component *i*. The ideal solution model requires that the activity coefficient, γ_i , is unity. More specifically, the ideal solution model requires several important constraints, including:

- 1. No excess volume, $V = \sum_i x_i V_i$ where V and V_i are the mixture molar volume and the component molar volume respectively, and no excess entropy, $S = \sum_i x_i S_i R \sum_i x_i \ln x_i$ where S and S_i are the mixture molar entropy and the component molar entropy respectively. These constraints are approached when molecules in a mixture have similar sizes and shapes.
- 2. No excess energy or $U^E = 0$. This constraint requires that the energy of all intermolecular interactions, including hydrogen bonding, are the same regardless of the molecule types involved across the composition range. This constraint is approached when mixing methanol and ethanol for example.

Deviations from these ideality constraints are corrected in activity coefficient models by combinatorial, residual and association γ contribution respectively. Models are typically developed for the logarithms of the contributions, which are added according to [8]:

$$\ln \gamma_i = \ln \gamma_i^{comb} + \ln \gamma_i^{res} + \ln \gamma_i^{assoc}$$
 1.5

Compared to EOSs, activity coefficient models are not limited by the first and third challenges outlined in Section 1.1.1. First, γ models can be used with limited pure component data because vapor pressure and density data are required only for the limited range where the models must be applied. Secondly, pure component fugacities and vapor pressures can be represented by correlations separate from the mixture model, decoupling the fitting of the pure component and mixture properties.

Despite these advantages, activity coefficient models that capture effects of hydrogen bonding have not gained wide acceptance. Until the mid-1990s, models primarily used a chemical theory approach where the association is represented by chemical species [13–16]. Chemical theory is difficult to generalize, which creates a barrier for wide-spread use. Fu and Sandler [6] demonstrated that the more general technique of Wertheim's theory could be used, but did not develop expressions that are readily applied in a generalized way. Development of activity models has languished while most development effort has focused on equations of state.

1.1.3. Association modeling

Association is most commonly modeled in thermodynamics with chemical theory or Wertheim's perturbation theory, which is the method used in CPA and the SAFT families of equations. While the chemical and Wertheim's theories are different, there are several parallels between them. One such similarity is that both methods have two key quantities: an association constant and another variable that is related directly to the fraction of "monomers" or molecules that remain non-bonded at equilibrium. A common challenge between EOS and activity coefficient models is discerning the value for these variables. In fact, published association constants can vary up to a factor of four to five depending on the source. A fundamentally satisfying method to obtain accurate association constants is to quantify the real monomer fraction in solution and fit the association parameters to the monomer fraction.

Some efforts have been made to determine monomer fractions from molecular dynamics simulations using atomistic force fields. While more rigorous than fitting to macroscopic density and vapor pressure data, the monomer fractions and hydrogen bonding depend significantly on the force field employed. Kwac and Geva [17] show that the relative population of monomers can vary by up to 13% between various force fields. A better approach is to probe the microscopic

nature of the interactions with spectroscopic tools such as infrared (IR) or nuclear magnetic resonance (NMR) spectroscopy. The advantage of these techniques lies in the measurement of the experimental response due to the types and relative quantities of microscopic species existing in solution. However, attainment of high-quality information from these techniques remains challenging. For example, this work will show that the IR extinction coefficient varies by approximately 20 times for different types of hydroxyl groups in solution. Due to overlapping peaks, individual species have not previously been determined experimentally with confidence. NMR provides a time-averaged measure of hydrogen bonding and, in cases with multiple equilibrating species, the measurements provide only broad insight into the specific distribution of hydrogen bonded species in solution.

1.2. Research objectives

The project developed in this dissertation integrates multiple tools to improve the thermodynamic modeling of associating systems: a new activity coefficient model, spectroscopy and MD + QM simulations. The complimentary relationship between these facets provides new quantitative insights that overcome the limitations and exceed the capabilities of the individual tools acting alone. First, a flexible thermodynamic activity coefficient model is developed that incorporates Wertheim's association term. Next, the key model quantity, the monomer fraction, is extracted from IR and NMR spectra. Finally, to overcome the challenges associated with spectral analysis, MD + QM simulations are used to estimate the extinction coefficients of various species and predict their distribution in solution. Ultimately, the objective of this work is to advance knowledge of hydrogen bonding by quantifying its effects at a molecular level to lend physical significance to engineering models for industry.

1.3. Dissertation outline

This remainder of this document is organized as follows: Chapter 2 describes the relation between the concentration-based equilibrium constant, K_c and the activity-based equilibrium constant, K_a . Equilibrium reactions and chemical theory are often described by K_c which is assumed to be independent of composition in the liquid phase. However, from the thermodynamic perspective, it is actually K_a that remains constant with composition as it is directly related to the Gibbs energy. The chapter explores the relationship between the two types of equilibrium constants, and their use in reaction engineering. In Chapters 3 and 4, the two most prevalent association theories in thermodynamics, chemical theory and Wertheim's perturbation theory, are shown to be numerically and analytically identical under reasonable approximations. We demonstrate a direct relationship between the concentration-based equilibrium constant K_c used in chemical theory, and the Wertheim association constant $\Delta^{A_i D_j}$ by relating them both to the activity-based equilibrium constant, K_a . Moreover, an activity coefficient association contribution (γ^{assoc}) model, named the Wertheim Association Gamma (WAG), is developed to provide generalized method to integrate association into thermodynamic γ models.

Chapter 3 considers associating species in an inert solvent. In Chapter 4, the analysis is extended to systems that solvate in addition to forming chains. Existing literature in this area has rarely detailed the species in chemical theory. Therefore, objectives of Chapters 3 and 4 are to elucidate how various species must be counted in chemical theory and to demonstrate the relationship with Wertheim theory.

The WAG model developed in Chapter 3 is used in Chapter 5 to model the behavior of methanol and ethanol-containing binary and ternary systems of industrial significance. The purpose of this work is to demonstrate the capabilities and limitations of the functional form of the γ^{assoc} when combined with traditional physical models such as NRTL [3], Wilson [18], and Scatchard-Hildebrand [19,20] and compare with the SAFT and CPA models. To this end, the values of the association constants ($\Delta^{A_i D_j}$) are set to values determined from published CPA parameters.

In Chapters 6 and 7, fundamental studies are leveraged to obtain more physically significant values of $\Delta^{A_i D_j}$. First, a successive molecular dynamics and quantum mechanics (MD + QM) approach to investigating molecular association is outlined. Molecules are classified according to their participation in hydrogen bonding, and their response to excitation by infrared is studied. A universal map is created that relates the hydroxyl IR band characteristics and the bond length of the hydroxyl in alcohol monomers and clusters. Then, the QM maps are used to analyze experimental IR spectra to calculate $\Delta^{A_i D_j}$ for butanol + cyclohexane. In the final Chapter 8, the main conclusions of the project are summarized and recommendations for future work are outlined.

Chapter 2. Relation of Concentration-based Equilibrium Constants to Activity-based Equilibrium Constants in the Liquid Phase

Concentration-based equilibrium constants are used widely in modeling chemical kinetics for elementary reactions in both the gas phase and liquid phase. However, activity-based equilibrium constants are the rigorous values based on thermodynamic minimization of the Gibbs energy. The relation between the concentration-based constants and the activity-based constants is welldocumented in textbooks for the ideal gas phase. This research note shows the relation between the equilibrium constants in the liquid phase and discusses conditions under which the concentration-based equilibrium constant is independent of composition.

2.1. Introduction

Characterization of chemical equilibria is necessary for chemical reactor design, and of interest for understanding limitations of chemical conversion. A requirement for chemical equilibrium is that the Gibbs free energy of the reaction is minimized, and the standard state Gibbs energy change is related to the activity-based equilibrium constant, K_a [8]. As such, from a fundamental thermodynamic perspective, K_a is independent of composition for a given reaction. However, in practice, this compositional independence is commonly assumed for the concentration-based equilibrium constant, K_c , for reactions in both the gaseous and liquid phases. The use of concentrations in the liquid phase is often explained using the empirical law of mass action rather than on a basis of thermodynamic principles [21–23]. In this work, we show why this approach is valid for ideal gases and explore the conditions under which the approximation is reasonable for liquids.

2.2. Gibbs energy and the activity-based equilibrium constant

For the purposes of this communication, we consider the reaction components E and F combining into a product composed of eE and fF.

$$eE + fF \rightleftharpoons E_eF_f$$
 2.1

Commonly, reactions are generalized using stoichiometric coefficients, $|v_i|$ and stoichiometric numbers, v_i [21]. Thus, the stoichiometric coefficients for this reaction are e, f, and 1 respectively for the three species, and the stoichiometric numbers are $v_E = -e$, $v_F = -f$, and $v_{E_eF_f} = 1$. While there is only one product for this simple example, the results set forth here are not limited to single-product reactions.

Thermodynamic analysis focuses on the activity-based equilibrium constant defined by

$$K_a = \prod_{i=1}^{N} (a_i)^{\nu_i}$$

where a_i is the activity of species *i* and depends on the standard state. The activity-based equilibrium constant is dimensionless, independent of composition and depends only on temperature. The equilibrium constant value is calculated via

$$-RT\ln K_a = \Delta G^o = \sum_i \nu_i \Delta G^o_{f,i}$$
2.3

where ΔG^o is the standard Gibbs energy of reaction and $\Delta G_{f,i}^o$ is the standard state Gibbs energy of formation of component *i*. The value of the equilibrium constant depends on the standard state used for each component and the value will be different when ideal gas standards states are used compared to when liquid standards states are used. Commonly, all the components use the same standard state, though this is not required. In many cases, pure thermodynamic quantities are not
known, and the value of the equilibrium constant is determined empirically from applying Eq. (2.2) or approximations. The activity in Eq. (2.2) is given by

$$a_i = \hat{f}_i / f_i^o \tag{2.4}$$

where \hat{f}_i is the component fugacity and f_i^o is the standard state fugacity. A standard state is specified at a particular composition, state of aggregation, and typically chosen to be at the temperature of the reacting system and a specified standard state pressure, P^o .

2.3. Gas phase activity-based equilibrium constant

When the reactants and products are in the gas phase, the fugacity is calculated using an equation of state, $\hat{f}_i = y_i \hat{\varphi}_i P$, where y_i is vapor mole fraction and $\hat{\varphi}$ is the component fugacity coefficient. Using the common standard state of a pure ideal gas at a pressure of 1 bar, $f_i^o = P^o = 1$ bar, and for the reaction of Eq. (2.1), inserting activities from Eq. (2.4) into Eq. (2.2) gives

$$K_a = \frac{y_{E_eF_f}\hat{\varphi}_{E_eF_f}P/P^{\circ}}{(y_E\hat{\varphi}_EP/P^{\circ})^e(y_E\hat{\varphi}_FP/P^{\circ})^f}$$

$$2.5$$

where commonly P° is omitted from the expression and P is expressed in bar to assure that K_a remains dimensionless. When the gas phase is ideal, the fugacity simplifies by neglecting gas phase nonidealities, $\hat{\varphi}_i = 1$, and recognizing the molar concentration for species *i*, and denoting it by square brackets, $[i] = y_i P/(RT)$:

$$K_a^{ig} = \frac{y_{E_eF_f}}{y_E^{e} y_F^{f} P^{(e+f-1)}} = \frac{\left[E_eF_f\right]}{\left[E\right]^{e} \left[F\right]^{f} (RT)^{(e+f-1)}} = \frac{K_C}{(RT)^{(e+f-1)}}$$
2.6

The expression in Eq. (2.6) is only valid for ideal gases as noted with the superscript *ig*. The concentration-based equilibrium constant found within the equation is defined as

$$K_{C} \equiv \prod_{i=1}^{N} [i]^{\nu_{i}} = \frac{[E_{e}F_{f}]}{[E]^{e}[F]^{f}}$$
2.7

This relation is shown widely in many texts and developed only for ideal gases and/or used without proof for condensed phases [8,21–25]. Since K_a is independent of composition, it is obvious from Eq. (2.6) that K_c is independent of concentration for ideal gas species. However, for liquid phases, the ideal gas law does not apply and thus, $[i] \neq y_i P/(RT)$. Therefore, $(RT)^{(e+f-1)}$ is not the correct conversion factor between the liquid phase K_a and K_c values and the general relation for condensed phases is not obvious. The following section demonstrates the relation between K_a and K_c in condensed phases and explains the conditions under which liquid-phase K_c can be expected to be independent of composition.

2.4. Liquid phase activity-based equilibrium constant

One method to explore the composition dependence of K_c in the liquid phase is to use a standard state of 1 M in Eq. (2.3), and to use the concentration scale for activities in Eq. (2.4). However, activity relations for nonelectrolytes on the concentration scale are uncommon. This work will use the mole fraction scale to understand the relation between K_a and K_c and the conditions where K_c may be expected to be independent of composition in the liquid phase.

The use of K_c also arises from reaction kinetics. Chemical kinetics are routinely modeled using concentrations to represent the driving force [21–23] for the reaction rate. For the reaction in Eq. (2.1), the commonly-used concentration-based rate expression for an elementary reversible reaction in a batch reactor is

$$-\frac{1}{e}\frac{d[E]}{dt} = -\frac{1}{f}\frac{d[F]}{dt} = \frac{d[E_eF_f]}{dt} = k_f[E]^e[F]^f - k_r[E_eF_f]$$
2.8

where k_f and k_r are the forward and reverse rate constants, d[E]/dt and d[F]/dt are the rates of change of *E* and *F* concentrations, respectively. The term $d[E_eF_f]/dt$ is the rate of change for concentration of E_eF_f . Rate laws represent the non-equilibrium reaction phenomena, and a lengthy review regarding the use of the law of mass action is available [26]. For the purposes of this note, we accept the form of Eq. (2.8) as sufficient for engineering application for ideal gases. If valid, since the forward and reverse rates are equal at equilibrium, then

$$K_{C} = \frac{k_{f}}{k_{r}} = \frac{\left[E_{e}F_{f}\right]}{\left[E\right]^{e}\left[F\right]^{f}}$$
2.9

In the liquid phase, using a liquid standard state, the activity is $a_i = x_i \gamma_i f_i / f_i^o$ where f_i and f_i^o are evaluated according to $f_i = \varphi_i^{sat} P_i^{sat} \exp\left(\frac{V(P-P^{sat})}{RT}\right)$ and $f_i^o = \varphi_i^{sat} P_i^{sat} \exp\left(\frac{V(P^o-P^{sat})}{RT}\right)$ at P and P^o respectively. Ignoring the minor effect of pressure on liquid fugacity, the equilibrium constant for the example reaction of Eq. (2.1) results in [8]

$$K_a^L = \prod_{i=1}^N (x_i \gamma_i)^{\nu_i} = \frac{x_{E_e F_f} \gamma_{E_e F_f}}{(x_E \gamma_E)^e (x_F \gamma_F)^f}$$

$$2.10$$

To introduce concentrations for condensed phases, molar densities are typically used according to

$$x_i = [i]/\rho \tag{2.11}$$

where ρ is the mixture molar density. Thus, the relation between K_a and K_c is

$$K_a^L = K_C \rho^{(e+f-1)} K_\gamma \tag{2.12}$$

where

$$K_{\gamma} = \frac{\gamma_{E_e F_f}}{\gamma_E^e \gamma_F^f}$$
 2.13

Because K_a is independent of composition, and molar density and activity coefficients change with composition, then K_c may be expected to display composition dependence via Eq. (2.12), which is in conflict with the frequent empirical use of K_c for modeling of reactions that approach reaction equilibrium. Certainly, activity-based kinetics can model the composition dependence [27,28], but improved clarity of the relation between the K_c and K_a would benefit the kinetics and thermodynamics communities.

Within chemical engineering, activity coefficients are commonly modeled by [12,29]

$$\ln \gamma_i = \ln \gamma_i^{comb} + \ln \gamma_i^{res}$$
 2.14

where *comb* indicates the combinatorial entropy contribution, often given by Flory's equation [30] or some variant, and *res* indicates the energetic and other residual effects that are commonly modeled with adjustable parameters [12,29]. Note that Eq. (2.14) combined with Eq. (2.10) becomes

$$K_{a}^{L} = \frac{x_{E_{e}F_{f}}\gamma_{E_{e}F_{f}}^{comb}\gamma_{E_{e}F_{f}}^{res}}{\left(x_{E}\gamma_{E}^{comb}\gamma_{E}^{res}\right)^{e}\left(x_{F}\gamma_{F}^{comb}\gamma_{F}^{res}\right)^{f}} = \frac{x_{E_{e}F_{f}}\gamma_{E_{e}F_{f}}^{comb}}{\left(x_{E}\gamma_{E}^{comb}\right)^{e}\left(x_{F}\gamma_{F}^{comb}\right)^{f}}K_{\gamma}^{res} = K_{a}^{comb}K_{\gamma}^{res}$$

$$2.15$$

where the superscript *L* indicates liquid phase and K_a^{comb} is represented by the Flory equation, denoted explicitly in the superscript:

$$K_a^{Flory} = \frac{x_{E_eF_f} \gamma_{E_eF_f}^{comb}}{\left(x_E \gamma_E^{comb}\right)^e \left(x_F \gamma_F^{comb}\right)^f}; \quad K_{\gamma}^{res} = \frac{\gamma_{E_eF_f}^{res}}{\left(\gamma_E^{res}\right)^e \left(\gamma_F^{res}\right)^f}$$
2.16

Appendix A demonstrates that use of Flory's equation for configurational entropy under two conditions. First, Flory's equation assumes that the packing factor, which is the fraction of the total volume that is occupied by molecules and a measure of how "packed" molecules are in a given system, is universal. Second, the molar volumes are assumed to be additive on reaction, meaning that $V_{E_eF_f} = eV_E + fV_F$ which is satisfied if mass density of reaction products and reactants are all the same as shown in Appendix B. The calculation results in

$$K_a^{Flory} = K_C K_V \exp(e + f - 1)$$
 2.17

where $K_V = V_{E_eF_f}/(V_E^eV_F^f)$ with V_i representing the molar volume of species *i*. Note $K_V \exp(e + f - 1)$ is independent of concentration. Thus, we conclude that Flory's equation justifies density appearing in Eq. (2.12) and that

$$K_a^L = K_C K_V K_V^{res} \exp(e + f - 1)$$
2.18

Because K_a^L is independent of composition, dependence of K_c on composition is coupled to composition dependence of K_Y^{res} . We expect K_c to be independent of composition when K_Y^{res} is constant, defined by Flory as an athermal solution ($K_Y^{res} = 1$), and approximately independent of composition when residual contributions are small. Components often have large differences in molar volumes (thus molar densities) creating a composition dependence of molar volume (thus molar density). However, mass densities, ρ^* , are frequently found experimentally to be in the range of 0.8 kg/L $< \rho^* < 1$ kg/L. In cases where the reactants and products of a reaction have the same mass densities, the molar volumes are additive ($V_{E_eF_f} = eV_E + fV_F$) as shown in Appendix B. Then, under the assumption of Flory's equation with a constant K_Y^{res} , K_c will be constant. Note that the Flory configurational contribution to K_Y cancels the molar density appearing in Eq. (2.12). Also, this analysis indicates that when residual activity coefficients are used with K_c , the activity coefficient models should not include the combinatorial contribution to avoid double-counting the combinatorial effect. APPENDICES

Appendix A: Development of the activity-based equilibrium constant for athermal solutions Here we develop the expression for K_a for an athermal solution using Flory's equation [30] for the activity coefficients. Considering Flory's theory for the combinatorial activity coefficient, we write the following expression:

$$\ln x_i \gamma_i^{comb} = \ln x_i + \ln \left(\frac{\Phi_i}{x_i}\right) + 1 - \frac{\Phi_i}{x_i} = \ln \Phi_i + 1 - \frac{\Phi_i}{x_i}$$
A.1

where Φ_i is the volume fraction for component *i* defined as: $\Phi_i = x_i V_i / \sum_j x_j V_j$. Therefore, the equilibrium constant ratio for products and reactants is

$$\ln\left(\frac{x_{E_eF_f}\gamma_{E_eF_f}^{comb}}{\left(x_E\gamma_E^{comb}\right)^e\left(x_F\gamma_F^{comb}\right)^f}\right) = \ln\left(\frac{\Phi_{E_eF_f}}{\Phi_E^e \Phi_F^f}\right) - 1 + e + f - \frac{\Phi_{E_eF_f}}{x_{EF}} + \frac{e\Phi_E}{x_E} + \frac{f\Phi_F}{x_F} \quad A.2$$

If the molar volumes are assumed to be additive on product formation and $V_{E_eF_f} = eV_e + fV_f$, and recognizing the molar density $\rho = 1/\sum_j x_j V_j$, the equilibrium constant ratio can be written as:

$$\ln K_a^{Flory} = \ln \left(\frac{x_{E_eF_f} \gamma_{E_eF_f}^{comb}}{\left(x_E \gamma_E^{comb} \right)^e \left(x_F \gamma_F^{comb} \right)^f} \right)$$

$$= \ln \left(\frac{x_{E_eF_f}}{x_E^e x_F^f} \frac{V_{E_eF_f}}{V_E^e V_F^f} \frac{1}{\rho^{(e+f-1)}} \right) - 1 + e + f$$

$$\ln K_a^{Flory} = \ln \left(\frac{\left[E_eF_f \right]}{\left[E \right]^e \left[F \right]^f} \frac{V_{E_eF_f}}{V_E^e V_F^f} \exp(e + f - 1) \right)$$
A.4

$$K_a^{Flory} = K_C K_V \exp(e + f - 1)$$
 A.5

where we define for convenience a composition-independent term

$$K_V \equiv \frac{V_{E_eF_f}}{V_E^e \ V_F^f} \tag{A.6}$$

Appendix B: Effect of mass density on volume of mixing

Here we show that if mass density is constant upon mixing or reaction that excess volume is zero. If mass density of the mixture and components are all equal:

$$\frac{MW_{mix}}{V} = \frac{MW_E}{V_E} = \frac{MW_F}{V_F}$$
B.1

where MW_i and V_i is the molecular weight and molar volume of component *i* respectively. Using the definition of average molecular weight $MW_{mix} = x_E MW_E + x_F MW_F$, and selecting *E* as a basis for the right-hand side:

$$\frac{x_E M W_E + x_F M W_F}{V} = \frac{M W_E}{V_E}$$
B.2

Rearrange for *V*

$$\frac{(x_E M W_E + x_F M W_F) V_E}{M W_E} = V$$
B.3

Similarly, if the mass densities are equal, from Eq. (B.1):

$$V_j = \frac{V_i M W_j}{M W_i}$$
B.4

Substituting Eq. (B.4) in Eq. (B.3), shows that the excess volume is zero

$$x_E V_E + x_F \left(\frac{MW_F V_E}{MW_E}\right) = x_E V_E + x_F V_F = V_{mix}$$
B.5

Because most liquids have very similar mass densities, the excess volume can often be neglected in all but the most precise work.

Chapter 3. Relation of Wertheim Association Constants to Concentrationbased Equilibrium Constants for Mixtures with Chain-forming Components

Several association modeling approaches have been developed to accurately describe the properties of polar solutions. Chemical theory and Wertheim's perturbation theory are among the most popular of these and they have been shown to yield similar functional forms for the contributions of association to Helmholtz energy and activity coefficients. In this paper, we study Flory polymerization theory through the work of Campbell and elucidate its correlation to Wertheim's theory. A simple key relationship between the concentration-based equilibrium constant and Wertheim's association constant is developed for systems in which all associating components have one acceptor and one donor site. Algebraic and numerical proofs are given for the equivalence of Flory's polymerization theory and Wertheim's perturbation theory for pure fluids and mixtures. Additionally, a new generalized activity expression is developed for Wertheim's theory.

3.1. Introduction

3.1.1. Motivation

As global demand for renewable chemicals and fuels increases, new pathways for their manufacture are being proposed. These projects include upgrading of natural oils, and fermentation of pyrolysis products to chemicals and fuels. The ubiquitous presence of polar components in biobased process streams complicates the design of industrial separation units. Traditional thermodynamic models, such as non-random two liquid (NRTL) and UNIversal QUAsiChemical (UNIQUAC), are incapable of accurately representing LLE and VLLE in these systems without utilizing more parameters than can be determined with confidence of physical relevance [31].

The limitations of traditional models for polar systems are largely due to their crude representation of important molecular interactions; specifically, hydrogen bonding. Polar molecules that hydrogen bond (or associate) with each other form clusters of chains and/or rings which cause large deviations from ideal behavior. These interactions are responsible for higher boiling points and enthalpies of vaporization of hydrogen bonding compounds compared to molecules of comparable size and polarity that lack the ability to hydrogen bond. The extent and nature of association varies under different conditions including solution composition, temperature and pressure. Therefore, it is necessary for thermodynamic models to incorporate hydrogen bonding explicitly rather than lumping the contributions with nonpolar (van der Waals) and simple dipole-dipole interactions.

3.1.2. *Historical perspective*

Association models can be classified into three categories: chemical theories, lattice models and perturbation theories. Apelblat [32,33] presents a historical summary of the advancements made in association modeling from 1884 to 1984. Among the most noteworthy contributions is Dolezalek's [34] theory representing hydrogen bonding as an equilibrium reaction governed by an equilibrium constant. The original theory, often called 'ideal' chemical theory, considers the true species formed through hydrogen bonding (e.g. dimers, trimers etc.) to otherwise behave ideally. In other words, physical interactions between the complexes in solution are ignored. This simplification is difficult to justify and has proved inadequate for many systems especially in modeling liquid-liquid immiscibility. The Dolezalek approach was therefore enhanced by combining the chemical theory contribution with a conventional model, such as NRTL [35] or UNIQUAC [36], to capture the physical interactions.

With chemical theory, several decisions must be made a priori regarding the species that exist in solution and the equilibrium constants for various hydrogen bonds. The continuous association model, also known as the Mecke-Kempter model, makes the simplification that the value of the equilibrium constant does not depend on the chain or ring size formed. While there has been some discussion on the contradiction of this assumption with spectroscopic evidence [37], it is often adopted as a first-order representation of the complexation in solution.

Lattice models represent the system as a lattice of empty and occupied sites and investigate the extent of association by quantifying interactions between adjacent occupied sites. Perhaps the most prominent lattice model is the one presented by Guggenheim [38] which accounts for non-randomness in the solution. Hydrogen bonding has since been integrated into several lattice models [39–42].

Perturbation theories add the association to a non-associating reference fluid. Gubbins [43] recently presented a review of work done in this area. Several perturbation theories [44–46] have been developed but the simplest and most popular is Wertheim's theory [47–50], which is the basis of a number of equation of state (EOS) models including Elliott-Suresh-Donohue (ESD) [10,51], cubic plus association (CPA) [11,52] and the statistical associating fluid theory (SAFT) family of equations [9,53–57].

Numerous independent researchers [58–60] have compared association models. Economou and Donohue [58] found that for many systems, both chemical theory (using activity-based K_a) and Wertheim's theory yield essentially the same functional form for the association contribution, despite differences in their origins and derivations. Wolbach and Sandler derived relations between K_a and Wertheim's association constant for ideal gases and used quantum calculations to estimate values of the Wertheim association constant for molecules with multiple sites [59,61]. Chapman

et al. [9] noticed the similarity in polymer distributions using Wertheim's theory in the SAFT model compared to Flory's polymerization theory (using concentration-based K_c). Campbell rearranged the Flory polymerization theory and was able to find a similar algebraic form but the terms were not identical [16]. Thus, the analytical equivalence of the Wertheim theory and Flory's polymerization theory has not yet been proven. In this work, we compare the two approaches for systems of chain-forming species in which the associating components each have an acceptor site and a donor site to contribute to association. Examples of these systems include pure alcohols, alcohol + inerts and alcohol mixture solutions. We also provide algebraic and numerical proofs of the equivalence of Flory's theory and Wertheim's theory when the packing factor is universal, and the excess volume is zero.

3.2. Theory and analysis

A nonideal system's behavior is described using either deviations from ideal gas mixtures using fugacity coefficients, or deviations from ideal solutions using activity coefficients as discussed by Prausnitz and Tavares [62]. Fugacity coefficients are readily calculated with an equation of state through the compressibility factor, Z. It is common practice to separate contributions to Z relative to the ideal gas value, Z^{ig} , according to the nature of the nonideality causing the deviations.

$$Z - Z^{ig} = Z - 1 = Z^{rep} + Z^{att} + Z^{assoc}$$

$$3.1$$

Here Z^{rep} and Z^{att} represent effects of the repulsive and attractive forces and Z^{assoc} accounts for hydrogen bonding between molecules. Contributions to the Helmholtz energy departure at fixed Tand V are correspondingly

$$\left(A - A^{ig}\right)_{T,V} = A^{rep} + A^{att} + A^{assoc}$$
 3.2

In a model such as the SAFT EOS, a *bond* term is added to both $Z - Z^{ig}$ and $(A - A^{ig})_{T,V}$. An alternative approach is to consider the excess Gibbs energy (G^E), related directly to the activity

coefficient, γ , of each component and separate perturbations according to their sources. Deviations caused by differences in the size and shapes of molecules are accounted for by an entropic combinatorial contribution which ignores all attractive interactions and is the only term remaining at infinitely high temperatures. The association contribution represents non-idealities due to hydrogen bonding and the residual term accounts for other interactions that are not as well-understood and typically includes adjustable parameters. The contributions combine in a multiplicative form. However, it is conventional to calculate the logarithms of the activity coefficients and present the relationship as

$$\ln \gamma = \ln \gamma^{comb} + \ln \gamma^{assoc} + \ln \gamma^{res} \qquad 3.3$$

The focus of this work is the γ^{assoc} term. We begin by introducing the nomenclature used to describe the apparent and true concentrations in associating solutions. Here, the term 'apparent' is used to denote 'stoichiometric' or 'superficial' concentrations of species in a mixture which is the overall amount added in composing the mixture divided by the mixture volume. The speciation that occurs in solution results in 'true' species, which are model-dependent. For example, in a pure alcohol, the apparent mole fraction of alcohol is unity, but the solution can also be described in terms of concentrations of true monomers, dimers, trimers and so on. Notations vary in literature, and the distinction between apparent and true molarity is not always clear. In this work, a subscript *T* is used to denote the true quantities. Every solution will contain some 'free' or 'unassociated' molecular species which is described as monomer and denoted with the subscript M, which is understood to be a 'true' quantity.

3.2.1. Association term – Wertheim's theory

Wertheim's perturbation theory calculates the probabilities of various site interactions that result in the formation of complexes. Each molecule type in solution may have any number of acceptor (A) and/or donor (D) sites and bonding between each pair is governed by an association constant, $\Delta^{A_i D_j}$. Notation for the acceptor and donor sites subscripts differs in literature, thus comparisons must be done carefully. In early literature, the *i* and *j* denote the host component. However, as noted by Michelsen [63], the summations are greatly simplified by indexing the acceptor and donor sites rather than the site hosts because site hosts can have more than one type of each site. We write summations over sites in this work. The acceptor and donor terminology allows the model to lend itself to any type of complexation, including covalent bonding. Depending on the number and types of sites attributed to a molecule of a component, different association schemes can be used as reviewed by Kontogeorgis and Folas [64]. Alcohols, for example, are commonly modeled using the 2B association scheme which corresponds to having one acceptor and one donor site per molecule (Figure 3.1). We follow this convention in the present work. A detailed explanation of the theoretical development of Wertheim's theory is given by Chapman [9] and Zmpitas and Gross [65].



Figure 3.1: 2B association scheme. Hydrogen bonded oligomers with two bonding sites per molecule. The acceptor and donor sites are labeled on the leftmost molecule.

The Helmholtz energy departure at fixed temperature and volume due to association derived from Wertheim's theory is implemented in SAFT and CPA as

$$\frac{\left(\underline{A} - \underline{A}^{ig}\right)_{T,V}^{assoc}}{RT} = \sum_{\substack{all \ site \\ types \ B_i}} n_{0,B_i \ host} N_{B_i} \left(\ln(X^{B_i}) - \frac{X^{B_i}}{2} + \frac{1}{2} \right)$$
3.4

where B_i is generic notation for an acceptor, donor, or bivalent (acceptor/donor) site, $n_{0,B_i host}$ is the apparent number of moles of the component hosting site B_i , and N_{B_i} is the number of identical sites of type *B* on the host. The term $X^{B_i} = n^{B_i, free}/n^{B_i, total}$ is the fraction of sites of type *B* that remain non-bonded at equilibrium, calculated through the balance on the probability of binding:

$$X^{B_{i}} = \left(1 + \sum_{\substack{mating\\sites E_{j}}} x_{E_{j} host} N_{E_{j}} X^{E_{j}} \rho \Delta^{B_{i}E_{j}}\right)^{-1}$$

$$3.5$$

where E_j is generic notation for a site which mates with site B_i , $x_{E_j host}$ denotes the component apparent mole fraction of the site E_j host and $\Delta^{B_i E_j}$ is an association constant parameter between site *i* of type B and site *j* of type *E*. In this work ρ is molar density and thus $\Delta^{B_i E_j}$ has units of molar volume.

With chain-forming components, when each hydrogen bond formed, one acceptor (A) and one donor (D) site are consumed. Thus, in a pure fluid with the 2B association scheme, the fraction of molecules non-bonded at each acceptor site is the same as those free at the donor sites. Therefore, numbering acceptor and donor sites on the same host using the same subscripts,

For mixtures of associating species with the 2B model, Eq. (3.6) is contingent on $\Delta^{A_i D_j} = \Delta^{D_j A_i}$, which is true when certain cross-coefficients equations are used. The reader should recognize that site indexes *i* and *j* match the host indexes for the 2B model indexing discussed here. The combining rules we use to calculate cross-coefficients are consistent with Eq. (3.6) and are discussed in Section 3.2.6.

To derive a formula for γ_i^{assoc} using Wertheim's theory, we consider the work of Michelsen and Hendriks [66] in deriving a generalized function which reduces to Eq. (3.4) when maximized and results in the association contribution to chemical potential. This approach is extended to calculate the fugacity coefficient of each component in the mixture in Elliott and Lira [8] (Important corrections are made in the errata [67]). Calculating the activity coefficient from the fugacity coefficient for a given component for the case of conventional mixing rules, van der Waals Z^{rep} , a universal packing factor, and no excess volume yields (Appendix C)

$$\ln \gamma_k^{assoc} = \sum_{\substack{sites \ B_i \\ on \ k}} N_{B_i} \left[\ln \left(\frac{X^{B_i}}{X^{B_i,0}} \right) - \frac{1}{2} (1 - X^{B_i,0}) \right]$$

$$+ \frac{1}{2} \frac{\rho}{\rho_k} \sum_{\substack{all \\ sites \ B_i}} x_{B_i \ host} N_{B_i} (1 - X^{B_i})$$

$$(1 - X^{B_i})$$

$$(3.7)$$

where X^{B_i} is given by Eq. (1.4), and $X^{B_i,0}$ represents the fraction of sites free when the host for the site is pure. These constraints are identical to those used in Flory's polymerization theory. This activity coefficient equation is easier to apply than the equation of Fu and Sandler [6] because derivatives of X^{B_i} are not needed. A more general expression is given in Appendix C, Eq. (C.14). *3.2.2. Flory's concentration-based equilibrium constant*

In chemical theory, the energetic effects associated with hydrogen bonding are embedded within an equilibrium constant, K. This parameter could be based on the concentrations (K_c), mole fractions (K_x) or activities (K_a) of the products and reactants. Flory's well-known work on oligomer modeling used a concentration-based equilibrium constant and provided a discussion on the inconsistency of K_x with statistical mechanics [30]. Flory's original form of K_c and variations of it have since been applied in several publications. Chemical kinetics are commonly described in practice using concentrations, and for reversible reactions are consistent with K_c , which is frequently assumed to be independent of composition. This simplification is explored in Chapter 2 where we prove the assumption to be valid only when K_γ is invariant with composition. For pure alcohols or alcohol + inert systems, some of the variations of K_c for the formation of a chain of size *i*+1 from a complex of size *i* and a monomer are shown in Table 3.1. Here, ρ_{alc} represents the pure alcohol molar density. The quantity Φ_i is the volume fraction. The variable $\rho_T = n_T/\underline{V}$, is the true molar density which is smaller than the apparent molar density $\rho = n_0/\underline{V}$, because bonding reduces the number of moles, $n_T \leq n_0$. In Eq. (3.11), r_{alc} is a measure of the molecular size of the associating species relative to some reference ($r_{alc} = \rho^{ref}/\rho_{alc}$), chosen by Campbell to be the molar density of methanol at 303.15 K. The term z_i denotes the true mole fraction of component *i*.

Table 3.1: Concentration based equilibrium constants. Variations of the concentration-basedequilibrium constant used in literature. The subscript numeral on K_C differentiates between
models.

K _C	References	Eqn
$K_{C1} = \frac{C_{T,i+1}}{C_{T,i}C_M} = \frac{z_{i+1}}{z_i z_M} \frac{1}{\rho_T}$	Flory [30], Renon and Prausnitz [68,69] (eqn. 2)	3.8
$K_{C2} = \frac{\Phi_{T,i+1}}{\Phi_{T,i}\Phi_{M}} \frac{i}{i+1} = \frac{z_{i+1}}{z_{i}z_{M}} \frac{\rho_{alc}}{\rho_{T}} = K_{C1}\rho_{alc}$	Flory [30], Nagata [70], Nath and Bender [71], Renon and Prausnitz [68,69] (eqn. 1)	3.9
$K_{C3} = \frac{\Phi_{T,i+1}}{\Phi_{T,i}\Phi_{M}} = \frac{z_{i+1}}{z_{i}z_{M}}\frac{\rho_{alc}}{\rho_{T}}\frac{i+1}{i} = K_{C1}\rho_{alc}\frac{i+1}{i}$	Nagata [72]	3.10
$K_{C4} = \frac{\Phi_{T,i+1}}{\Phi_{T,i}\Phi_{M}}\frac{ir_{alc}}{i+1} = \frac{z_{i+1}}{z_{i}z_{M}}\frac{\rho^{ref}}{\rho_{T}} = K_{C1}\rho^{ref}$	Brandani [73,74], Campbell [16]	3.11

3.2.3. Relationship between K_a , K_c and Δ

To understand the relationship between K_a and K_c , we use equations for a pure component with one acceptor and one donor. These are later extended to mixtures using combining rules in Section 3.2.6. We begin with the thermodynamic definition of the activity-based equilibrium constant:

$$K_a = \frac{z_{i+1}}{z_i z_M} \frac{\hat{\varphi}_{i+1}}{\hat{\varphi}_i \hat{\varphi}_M} \frac{P^\circ}{P}$$
3.12

where P° is the standard state pressure, typically 1 bar. Inserting Eq. (3.8) into Eq. (3.12) yields

$$K_a = K_{C1} \rho_T \frac{\hat{\varphi}_{i+1}}{\hat{\varphi}_i \hat{\varphi}_M} \frac{P^\circ}{P}$$
3.13

Economou and Donohue [58] discuss different conventions used for equilibrium constants. By comparing the monomer fraction derived from various models, they conclude that Δ must be proportional to K_a at the same *T*. Wolbach and Sandler [59] explore the relationship further and show that in the case where the vapor phase is an ideal gas

$$\Delta^{ig} = f K_a \frac{RT}{P^{\circ}}$$
 3.14

where f is a constant that depends on the association scheme used to model the component in solution. For liquid components modeled with 2B, Eq. (3.14) becomes

$$\Delta = e^g K_a \frac{RT}{P^\circ} \tag{3.15}$$

where g is a function that depends on the equation of state used to calculate the fugacity coefficient ratio. The value of this function depends on the mixing rules used for the true solution of monomers and oligomers. The most accepted and widely used combining rules for complexes, introduced by Heidemann and Prausnitz [75], are

$$a_i = i^2 a_M \tag{3.16}$$

where a_M and b_M are the interaction and size parameters for monomers respectively. Heidemann and Prausnitz [75] show that when Eqs. (3.16) and (3.17) and the Lorentz-Berthelot mixing rules $(b = \sum_j x_j b_j, a = \sum_j \sum_k x_j x_k a_{jk})$ are followed, then g is given by Eq. (3.18).

$$g = \int_0^\eta \frac{Z^{rep}}{\eta} d\eta$$
 3.18

Here, η is the packing factor. From this equation, an important observation can be made. While K_a depends only on temperature, Wertheim's association constant Δ of Eq. (3.15) is also a function of the density of the solution due to the *g* function. Kontogeorgis and Folas [64] summarized the

assumptions taken in various models in literature and the resulting forms of g. When the van der Waals repulsive term is used, Eq. (3.18) gives the radial distribution function at contact distance, $g(\sigma)$. Moreover, applying the combining rules (Eqs. (3.16) and (3.17)) to the van der Waals equation for the component fugacity coefficients provides a useful relationship between the fugacity coefficient and true compressibility factor, Z_T , as shown in Appendix D. Calculating the ratio of the fugacity coefficients for the reaction of a complex of size i and a monomer to form a complex of size i+1 yields

$$\ln\left(\frac{\hat{\varphi}_i\hat{\varphi}_M}{\hat{\varphi}_{i+1}}\right) = g(\sigma) - \ln Z_T$$
3.19

or

$$e^{g(\sigma)} = \frac{P\underline{V}}{n_T RT} \frac{\hat{\varphi}_i \hat{\varphi}_M}{\hat{\varphi}_{i+1}}$$
3.20

Substituting Eq. (3.20) into Eq. (3.15), then converting to apparent density and inserting Eq. (3.13):

$$\Delta = \frac{P\underline{V}}{P^{\circ}n_{T}}\frac{\hat{\varphi}_{i}\hat{\varphi}_{M}}{\hat{\varphi}_{i+1}}K_{a} = \frac{PK_{a}}{P^{\circ}\rho}\frac{\hat{\varphi}_{i}\hat{\varphi}_{M}}{\hat{\varphi}_{i+1}}\frac{n_{0}}{n_{T}} = \frac{K_{C1}\rho_{T}}{\rho}\frac{n_{0}}{n_{T}}$$

$$3.21$$

Recognizing that $\rho_T / \rho = (n_T / \underline{V}) / (n_0 / \underline{V}) = n_T / n_0$ yields a remarkably simple $\Delta - K_C$ relationship for pure fluids based on the K_{C1} definition of Eq. (3.8).

$$\Delta = K_{C1} \qquad \qquad 3.22$$

This result is intuitive considering the conclusions arrived at by several investigators [58,60,66] on the mathematical similarity of chemical and Wertheim's theories. However, it is powerful in that it allows one to use and compare both association models with ease.

3.2.4. Association term – Campbell's chemical theory model

Pradhan et al. [76] developed a chemical theory model by integrating an equilibrium constant into the Flory-Huggins (FH) theory. Campbell [16] extended the expression to systems containing any number of alcohols. In this section, we clarify the relationship between this set of equations and other association models.

Following Flory [30], Campbell writes

$$\ln \gamma_i^{chem} = \ln \left(\frac{C'_{T,M_i}}{x_i C_{T,M_i}^{\prime 0}} \right) - r_i (\Psi - \Psi_i^0)$$
 3.23

where C'_{T,M_i} and C'^{0}_{T,M_i} are the dimensionless true monomer concentrations of component *i* in a mixture and at purity, respectively. All concentrations *C'*, are made dimensionless by dividing by ρ^{ref} , the liquid molar density of methanol at 303.15 K. Ψ is the total molar concentration of the true complexes in solution which Campbell showed is given by

$$\Psi = \sum_{i} \Psi_{i}; \qquad \Psi_{i} = \frac{C_{i}'}{1 + \sum_{j} K_{C4,ji} \Psi_{j}} \qquad 3.24$$

where the summations are over apparent species and C'_i is the apparent dimensionless concentration of component *i*. As shown in Appendix E, Campbell's Ψ_i can be recognized as the total molar concentration of species with a free end group *i*.

The Flory concentration-based $K_{C4,ji} = C'_{T(j,n+1)}/(C'_{T(i,n)}C'_{T(j,1)})$ is the equilibrium constant for the association between a chain of length n with a non-bonded acceptor i, forming a bond with monomer j donor, resulting in a chain of length n + 1 with a non-bonded j acceptor end. Campbell does not discuss acceptors or donors, but the nomenclature helps clarify how, when the alcohol monomer donor bonds to an end of a chain, it can still participate through its free acceptor. Note that the nomenclature for acceptors and donors could be flipped without changing the balances. Campbell provides mathematical derivations for Eqs. (3.23) and (3.24). The free end on a chain could be an acceptor or a donor. Considering the end to be an acceptor, Ψ_i represents the dimensionless molar concentration of non-bonded acceptors of type i. Because each alcohol molecule possesses one acceptor site, the dimensionless apparent molar concentration of all acceptors on component *i* in solution is equal to C'_i . Therefore, the fraction of unbonded acceptors is

$$\frac{\Psi_i}{C_i'} = \frac{n^{A_i}}{n^{A_i,T}} = X^{A_i}$$
 3.25

or $\Psi_i = C'_i X^{A_i}$. Substituting the dimensionless concentration in terms of molar densities $C'_i = \Phi_i \rho_i / \rho^{ref} = x_i \rho / \rho^{ref}$, we find that

$$\Psi_i = x_i \rho X^{A_i} / \rho^{ref} \qquad 3.26$$

Therefore, $\sum_{j} K_{Cji} \Psi_{j}$ can be written as $\sum_{j} x_{j} \rho K_{C4,ji} X^{D_{j}} / \rho^{ref}$. Simple manipulation will show that the right expression of Eq. (3.24) is the same as Eq. (1.4).

3.2.5. Algebraic equivalence of chemical and Wertheim's theory

For a binary alcohol (1) + inert (2) system, Wertheim's theory (Eq. (3.7)) gives

$$\ln \gamma_1^{assoc} = 2 \ln \left(\frac{X^A}{X^{A,0}} \right) - (1 - X^{A,0}) + \frac{\rho}{\rho_1} x_1 (1 - X^A)$$

$$= 2 \ln \left(\frac{X^A}{X^{A,0}} \right) + X^{A,0} - (1 - \Phi_1) - X^A \Phi_1$$

$$\ln \gamma_2^{assoc} = \frac{\rho}{\rho_2} x_1 (1 - X^A) = \frac{\Phi_2}{x_2} x_1 (1 - X^A)$$
3.28

For the same system, we begin the analysis of Campbell's model by calculating Ψ and Ψ_i^0 by combining Eqs. (3.24) and (3.26):

$$\Psi = \Psi_1 + \Psi_2 = \frac{x_1 \rho X^A}{\rho^{ref}} + \frac{x_2 \rho}{\rho^{ref}}$$

$$3.29$$

$$\Psi_1^0 = \frac{\rho_1 X^{A,0}}{\rho^{ref}}; \ \Psi_2^0 = \frac{\rho_2}{\rho^{ref}}$$
3.30

Next, we calculate the values of C'_{T,M_i} and C'^{0}_{T,M_i} . Campbell provides the relationship between the concentration of monomers ending with group *i* and C'_i given as $C'_{T,M_i} = \Psi_i^2 / C'_i$. Substituting into this equation with the definition of C'_i and Eq. (3.28), we find for both components in the mixture and at purity

$$C'_{T,M_1} = \frac{x_1 \rho}{\rho^{ref}} (X^A)^2; \quad C'_{T,M_2} = \frac{x_2 \rho}{\rho^{ref}}$$
3.31

$$C_{T,M_1}^{\prime 0} = \frac{\rho_1}{\rho^{ref}} \left(X^{A,0} \right)^2; \quad C_{T,M_2}^{\prime 0} = \frac{\rho_2}{\rho^{ref}}$$
 3.32

Therefore, by combining Eqs. (3.29-3.32), Eq. (3.23) for the alcohol becomes

$$\ln \gamma_1^{chem} = 2\ln\left(\frac{X^A}{X^{A,0}}\right) + \ln\left(\frac{\Phi_1}{x_1}\right) - X^A \Phi_1 - \frac{\Phi_1}{x_1} + \Phi_1 + X^{A,0}$$
 3.33

Adding and subtracting 1, and rearranging yields

$$\ln \gamma_{1}^{chem} = 2 \ln \left(\frac{X^{A}}{X^{A,0}} \right) + X^{A,0} - (1 - \Phi_{1}) - X^{A} \Phi_{1} + \ln \left(\frac{\Phi_{1}}{x_{1}} \right) + 1 - \frac{\Phi_{1}}{x_{1}}$$

$$= \ln \gamma_{1}^{assoc} + \ln \gamma_{1}^{Flory}$$
3.34

where Flory's equation [30] for the combinatorial contribution to activity coefficients is

$$\ln \gamma_i^{Flory} = \ln \left(\frac{\Phi_i}{x_i}\right) + 1 - \frac{\Phi_i}{x_i}$$
 3.35

Similarly, rearranging Eq. (3.23) and substituting Eqs. (3.29-3.32) for the inert solvent gives

$$\ln \gamma_2^{chem} = \frac{x_1 \rho}{\rho_2} (1 - X^A) + \ln \left(\frac{\Phi_2}{x_2}\right) + 1 - \frac{\Phi_2}{x_2} = \ln \gamma_2^{assoc} + \ln \gamma_2^{Flory}$$
 3.36

$$\ln \gamma_i^{chem} = \ln \gamma_i^{assoc} + \ln \gamma_i^{Flory}$$
 3.37

Thus, this provides analytical proof that Campbell's chemical theory incorporates combinatorial effects through Flory's equation. Consequently, when Wertheim's theory is compared to Campbell's model, the Flory term must be added separately. When calculating activity coefficients using Campbell's chemical theory, the combinatorial term is inherently included.

3.2.6. Calculating association parameters

To clarify the comparison further, we provide some numerical calculations. Early work done in determining K_c for binary alcohol + inert systems used data reduction of mixture data to fit association constants [68,69]. Commonly, the value of K_c is assumed to be independent of solvent species or concentration. Therefore, models based only on the properties of the pure associating compound were developed. Nath and Bender [71] based K_c on the temperature and enthalpy of vaporization at the normal boiling point of the alcohol. Lobien [77] calculated smaller K_c values than others using limiting activity coefficient data. Because these approaches differ considerably, there is significant scatter in the resulting K_c values for the same alcohol [71]. For CPA and SAFT, the $\Delta^{A_i D_j}$ values are typically determined by fitting density and vapor pressure [9,11,78].

In this work, to provide numerical calculations, we follow the approach outlined by Nagata [70] in applying the Brandani [73]'s work. The model essentially assumes that the only differences between the vapor pressure of an alcohol, P_{alc}^{vap} , and that of a non-associating compound with an identical molecular weight and chemical formula, i.e. its homomorph, P_h^{vap} , are the consequences of hydrogen bonding. Therefore, K_c is determined by matching the two vapor pressures through Eq. (3.38).

$$P_{alc}^{vap} = \gamma_{alc,M}^0 x_{alc,M}^0 P_h^{vap} \exp\left(\frac{\left(P_{alc}^{vap} - P_h^{vap}\right)\left(V_{alc} - B_{alc}^{free}\right)}{RT}\right)$$
3.38

where B_{alc}^{free} is the free contribution to the second virial coefficient evaluated using Hayden and O'Connell's calculations [79] and $\gamma_{alc,M}^{0}$ is the activity coefficient of the alcohol monomer in pure alcohol as given by Flory's equation.

The volume and mole fractions of monomer in pure alcohol are calculated as [70]

$$\Phi_M^0 = \frac{1 + 2K_{C2} - \sqrt{1 + 4K_{C2}}}{2K_{C2}^2}$$
3.39

$$x_M^0 = 1 - K_{C2} \Phi_M^0 \tag{3.40}$$

Ethers were selected as homomorphs for alcohols and their vapor pressures were determined using Ambrose's [80] correlation with the same n^* (effective number of carbon atoms in the ether homomorph) values as Nagata. Our resulting values of K_{C2} are given in Table 3.2.

Alcohol	K _{C2}	$K_{C1} = \Delta$		
Methanol	111.0	4549		
Ethanol	81.48	4803		
<i>n</i> -Propanol	71.27	5381		
<i>n</i> -Butanol	69.88	6458		
<i>n</i> -Pentanol	57.09	6233		

Table 3.2: Equilibrium constant for primary alcohols. *K*_{C2} and *K*_{C1} calculated at 50 °C

We have refitted the K_{C2} values to be consistent with all of our physical properties. We assume Nagata's [70] values for the enthalpy for the formation of a hydrogen bond.

To extend the analysis to alcohol mixtures, the cross-association parameters for dissimilar components are related through combining rules. Common combining rules that have been developed and tested [11,51,54,61,81–83] are variations of an arithmetic or geometric mean of the pure components' association energy and volume parameters. Kontogeorgis [64] outlines a number of these and provides a brief summary of work done in identifying the most accurate rules for various classes of mixtures. We find that Flory's polymerization theory and Wertheim's perturbation theory yield identical results for alcohols described with a 2B model, regardless of the combining rules used if Eq. (3.22) is maintained for the cross coefficients. Therefore, further exploration of the effect of the combining rules lies beyond the scope and intentions of this work and for alcohol mixtures, we use Suresh and Elliott's [51] combining rule given by:

$$\Delta^{A_i B_j} = \Delta^{A_j B_i} = \sqrt{\Delta^{A_i B_i} \Delta^{A_j B_j}}$$

$$3.41$$

3.2.7. Numerical equivalence of chemical and Wertheim's theory

We implement the abovementioned approaches to provide numerical evidence for the equivalence of chemical theory and Wertheim's theory. Within the constraints discussed, we find that the two approaches produce exactly the same results for systems in which the associating components can be described by the 2B association scheme. This is illustrated in Table 3.3 for two alcohol + inert systems at 30 and 80 mol% of alcohol and one alcohol + alcohol system at 30 mol% of methanol. Both theories yield the same monomer fractions and activity coefficients.

These results prove that Campbell's observation of the similarities between his association model and Wertheim's perturbation theory are justified. Indeed, when combinatorial effects are added to the perturbation theory and the correct relationship between the association parameters of the two models is implemented, their mathematical forms become identical. Chapman's recognition that Wertheim's model gives the same distribution as Flory's model is also consistent with this work. Future work in this area will look at extending the analysis to components with multiple acceptors and donors.

Table 3.3: Calculated parameters and intermediate values for Flory and Wertheim's theories for three systems. The reference density of methanol used for Campbell's theory calculations is 0.02440 mol/cm³. For methanol + ethanol, the cross coefficient is $\Delta^{A_1D_2} = \Delta^{A_2D_1} = 2559 \text{ cm}^3/\text{mol or } K_{C4,ij} = 62.44$

System		Ethanol $(1) + $ <i>n</i> -heptane (2)		<i>n</i> -Pentanol (1) + <i>n</i> -heptane (2)		Methanol (1) + Ethanol (2)	
x ₁		0.3	0.8	0.3	0.8	0.3	
Т (К)		344.9	344.7	368.4	383.1	347.0	
Component <i>i</i>		1	1	1	1	1	2
Campbell's Chemical Theory	<i>KC</i> 4	67.37	67.65	53.60	40.30	60.76	64.16
	Ψ	0.2680	0.1790	0.2427	0.1419	0.1025	
	Ψ_i^0	0.0944	0.0942	0.0749	0.0849	0.1203	0.0966
	$C'_{i,M}$	0.0101	0.0123	0.0119	0.0184	0.0043	0.0095
	$C_{T,M_i}^{\prime 0}$	0.0128	0.0128	0.0149	0.0192	0.0145	0.0134
	$x_M \frac{n_0}{n_T}$	0.0301	0.0230	0.0396	0.0527	0.0056	0.0124
	ln γ ^{chem}	0.7195	0.06100	0.5280	0.03109	-0.002853	-0.0004456
Wertheim's Theory	$\Delta^{A_i D_i}$ (cm ³ /mol)	2761	2773	2197	1652	2490	2629
	$X^{A_i} = X^{D_i}$	0.3167	0.1696	0.3633	0.2566	0.1361	0.1330
	$x_M \frac{n_0}{n_T}$	0.0301	0.0230	0.0396	0.0527	0.0056	0.0124
	$\ln \gamma_i^{assoc}$	0.9281	0.09320	0.5513	0.03344	0.02999	0.004305
	$\ln \gamma_i^{Flory}$	-0.2086	-0.03220	-0.02337	-0.002351	-0.03285	-0.004751
	$\ln \gamma_i^{chem}$	0.7195	0.06100	0.5280	0.03109	-0.0028533	-0.0004455

3.3. Conclusions

In this work, we outline derivations for Wertheim's perturbation theory's activity coefficient model and chemical theory and compare the two models. The focus is for the 2B scheme where each associating component has one acceptor and one donor site per molecule such as alcohol + inert and alcohol + alcohol mixtures. In these cases, the association strength $\Delta^{A_i B_j}$ in Wertheim's theory is found to be equal to the concentration-based equilibrium constant K_{C1} used in Flory's polymerization theory. This elucidates the relationship between the two association models. Furthermore, we consider Campbell's implementation of chemical theory and show that the activity coefficient model is the sum of Flory's combinatorial term, calculated on an apparent basis, and the association contribution calculated with Wertheim's perturbation theory. The similarities drawn by Campbell between his theory and Wertheim's theory are therefore consistent with this work. The correlation between the association models also supports Chapman's understanding that the species distribution predicted by Flory's and Wertheim's models are identical. Finally, we provide in Eqs. (3.7) and (C.14) a new generalized expression for activity coefficients from Wertheim theory useful when the packing factor is universal and excess volume is zero.

APPENDICES

Appendix C: Derivation of the Wertheim association contribution to the activity coefficient

Following Michelsen and Hendricks [66] and to abbreviate notation, we define, at equilibrium,

$$h = \sum_{\substack{all \ site \\ types B_i}} x_{B_i \ host} N_{B_i} (1 - X^{B_i})$$
C.1

Elliott and Lira [8] show that

$$Z^{assoc} = -0.5 \left(1 + \left(\frac{\partial \ln g}{\partial \ln \eta_P} \right)_{T,n} \right) h$$
 C.2

$$\ln \hat{\varphi}_{k}^{assoc} = \sum_{\substack{sites B_{i} \\ only on k}} N_{B_{i}} \ln(X^{B_{i}}) - 0.5n_{0} \left(\frac{\partial \ln g}{\partial n_{k}}\right)_{T, \underline{V}, n_{j \neq k}} h$$
 C.3

We will use these equations to calculate and simplify the equation for γ^{assoc} . By definition, Helmholtz energy is a natural function of temperature and volume.

$$\underline{A}(\underline{V}, T, n_1, n_2, \dots, n_n)$$
C.4

For an activity coefficient, we need a derivative of Helmholtz energy at constant T and P [6]. At constant T,

$$d\underline{A} = \left(\frac{\partial \underline{A}}{\partial \underline{V}}\right)_{T,n} d\underline{V} + \sum_{k} \left(\frac{\partial \underline{A}}{\partial n_{k}}\right)_{T,\underline{V},n_{j\neq k}} dn_{k}$$
C.5

The expansion rule gives

$$\left(\frac{\partial \underline{A}}{\partial n_k}\right)_{T,P,n_{j\neq k}} = \left(\frac{\partial \underline{A}}{\partial \underline{V}}\right)_{T,n} \left(\frac{\partial \underline{V}}{\partial n_k}\right)_{T,P,n_{j\neq k}} + \left(\frac{\partial \underline{A}}{\partial n_k}\right)_{T,\underline{V},n_{j\neq k}}$$
C.6

The association contribution to the fugacity coefficient can be calculated from

$$\frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T, \underline{V}, n_{j \neq k}} = \ln \hat{\varphi}_k^{assoc}$$
C.7

and

$$\left(\frac{\partial \underline{A}^{assoc}}{\partial \underline{V}}\right)_{T,n} = -P^{assoc}$$
C.8

And making the appropriate substitutions into Eq. (C.6):

$$\frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} = -\frac{Z^{assoc}}{V} \overline{V}_k + \ln \hat{\varphi}_k^{assoc}$$
C.9

Substituting Eqs. (C.1 and C.2) yields the universal equation

$$\frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j\neq k}} = \sum_{\substack{sites B_i \\ only on k}} N_{B_i} \ln X^{B_i} + \frac{1}{2} h \left[\frac{\overline{V}_k}{V} \left(1 + \left(\frac{\partial \ln g}{\partial \ln \eta_P} \right)_{T,n} \right) - n_0 \left(\frac{\partial \ln g}{\partial n_k} \right)_{T,\underline{V},n_{j\neq k}} \right]$$
C.10

To simplify this further, we assume that Z^{rep} is given by the van der Waals equation. Applying Eq. (3.18) results in $g = 1/(1 - b\rho) = (1/(1 - \eta_P))$.

$$\left(\frac{\partial \ln g}{\partial \ln \eta_P}\right)_{T,n} = \eta_P / (1 - \eta_P)$$

$$n_0 \left(\frac{\partial \ln g}{\partial n_k}\right)_{T,\underline{V},n_{j\neq k}} = \frac{b_k \rho_k}{1 - \eta_P} \left(\frac{V_k}{V}\right)$$
C.11

Eq. (C.10) becomes

$$\frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} = \sum_{\substack{sites B_i \\ only on k}} N_{B_i} \ln(X^{B_i}) + \frac{1}{2} h \left[\frac{\overline{V}_k}{V} \left(\frac{1}{1 - \eta_P} \right) - \frac{V_k}{V} \left(\frac{b_k \rho_k}{1 - \eta_P} \right) \right]$$
C.12

If excess volume is zero and the packing factor is universal,

$$\frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} \sum_{\substack{sites B_i \\ only on k}} N_{B_i} \ln(X^{B_i}) + \frac{1}{2} \left(\frac{\rho}{\rho_k} \right) \sum_{\substack{all \ site \\ types B_i}} x_{B_i \ host} N_{B_i} (1 - X^{B_i})$$
C.13

Using this equation, γ^{assoc} can be calculated as

$$\ln \gamma_k^{assoc} = \frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} \bigg|_{mix} - \frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} \bigg|_{pure \ k}$$
C.14

This results in Eq. (3.7) in the text. Note that $x_i \rho / \rho_k = \Phi_i$ when excess volume is zero.

Appendix D: Relationship between the fugacity coefficient and true compressibility factor

Beginning with the fugacity coefficient relation, we substitute van der Waals equation of state

$$Z_T = \frac{P\underline{V}}{n_T RT} = 1 + \frac{b\rho_T}{1 - b\rho_T} - \frac{a\rho_T}{RT}$$
D.1

and its implementation into the Helmholtz departure equation

$$\frac{\left(\underline{A} - \underline{A}^{ig}\right)_{T,\underline{V}}}{n_T R T} = -\ln(1 - b\rho_T) - \frac{a\rho_T}{R T}$$
D.2

For a true associated species in a pure fluid

$$\ln \hat{\varphi}_{i} = -\ln(1 - b\rho_{T}) + \frac{b_{i}\rho_{T}}{1 - b\rho_{T}} - \frac{2\rho_{T}\Sigma_{j}x_{T,j}a_{ij}}{RT} - \ln Z_{T}$$
D.3

Applying Eq. (3.16) and (3.17), calculating Eq. (D.3) for a monomer and a dimer gives

$$\ln \frac{\hat{\varphi}_M^2}{\hat{\varphi}_D} = -\ln(1 - b\rho_T) - \ln Z_T$$
D.4

or

$$\frac{\hat{\varphi}_D}{\hat{\varphi}_M^2} = Z_T (1 - b\rho_T) = \frac{P\underline{V}(1 - b\rho_T)}{n_T RT} = \frac{\hat{\varphi}_{i+1}}{\hat{\varphi}_i \hat{\varphi}_M}$$
D.5

Appendix E: Relation of Campbell's concentration variable in terms of the molar concentration of species

Here, we provide proof that Ψ_i is the true concentration of all chains ending with *i*. Beginning with the definition of Ψ_i in Campbell's [16] paper:

$$\Psi_{i} = C_{i,1}' + \sum_{j} \alpha_{j,i} (C_{j,1}' + \sum_{k} \alpha_{k,j} (C_{k,1}' + \sum_{l} \alpha_{l,k} (C_{l,1}' + \cdots$$
 E.1

where

$$\alpha_{j,i} = K_{C4,ji}C'_{i,1}$$
 E.2

Multiplying through Eq. (E.1),

$$\Psi_{i} = C_{i,1}' + \sum_{j} \alpha_{j,i} C_{j,1}' + \sum_{j} \alpha_{j,i} \sum_{k} \alpha_{k,j} C_{k,1}' + \sum_{j} \alpha_{j,i} \sum_{j} \alpha_{k,j} \sum_{l} \alpha_{l,k} C_{l,1}' + \cdots$$
 E.3

In Campbell's appendix C, the concentration of all chains of length n ending in i is given by

$$C_{i,n} = \sum_{j} \alpha_{j,i} C_{j,n-1}$$
E.4

This equation provides a recursion for simplification. Recognizing Eq. (E.4) in the second, third, and fourth term on the right hand side of Eq. (E.3) can be expressed as

$$C'_{i,2} = \sum_{j} \alpha_{j,i} C'_{j,1}$$
 and $C'_{j,2} = \sum_{k} \alpha_{k,j} C'_{k,1}$ and $C'_{k,2} = \sum_{l} \alpha_{l,k} C'_{l,1}$ E.5

Then

$$\Psi_{i} = C'_{i,1} + C'_{i,2} + \sum_{j} \alpha_{j,i} C'_{j,2} + \sum_{j} \alpha_{j,i} \sum_{j} \alpha_{k,j} C'_{k,2} + \cdots$$
 E.6

Simplifying the third and fourth term on the right with Eq. (E.4) again,

$$\Psi_i = C'_{i,1} + C'_{i,2} + C'_{i,3} + \sum_j \alpha_{j,i} C'_{j,3} + \cdots$$
 E.7

This procedure applied continuously to all sums in Eq. (E.3) yields the result that Ψ_i is also the true concentration of all chains that end in *i*.

$$\Psi_i = C'_{i,1} + C'_{i,2} + C'_{i,3} + C'_{i,4} + \dots = \sum_j C'_{i,j}$$
E.8

Chapter 4. Relation of Wertheim and Chemical Theories in Systems with Association and Solvation

Hydrogen bonding is usually calculated in thermodynamic models using chemical or Wertheim's theory. Although the basis of the two methods is different, several researchers have noted parallels between them. In fact, our previous study derives analytical and numerical proofs of the equivalence of chemical and Wertheim's theories for the case of an alcohol + inert system. In this work, we extend this finding to a complex ternary system containing a chain-forming component and two non-chain-forming components. The non-chain-forming molecules have only acceptor sites or donor sites, but not both. One molecule has two acceptor sites, and the other a donor site. The main objectives of this work are to elucidate the complicated counting that must be done in chemical theory to correctly account for all the bonded species that can form in solution and to demonstrate the equivalence of the two approaches for this complex system. The work also demonstrates how Wertheim's free site fractions can be used to determine species concentrations.

4.1. Introduction

As mixtures containing water, alcohols, organic acids and esters occur frequently in the chemical industry, new methods are being developed to accurately model their thermodynamic behavior. The challenge with these types of components is their propensity for hydrogen bonding which has a significant effect on many of the mixture physical properties. For example, hydrogen bonding is responsible for the high melting and boiling points of water relative to other chemical compounds of similar molecular size. To robustly capture this behavior, it is now widely accepted that hydrogen bonding, or association, must be represented explicitly as a separate term in thermodynamic models.

Association can be modeled several ways and the most common methods fall into one of two categories: chemical theory or perturbation theory. Chemical theory is based on the representation of hydrogen bonding as a chemical reaction, described by an equilibrium constant K_c , which results in the formation of a new species. This concept was originally proposed by Dolezalek [34] and has since been implemented to understand on a wide variety of associating systems including liquid metals [16,37,68,75,84–91]. In particular, Nagata *et al.* have published an extensive body of work [84,92–94] applying chemical theory to a wide range of organic mixtures. With alcohols, a simplified model, called the Mecke-Kempter model, is often used which assumes that the value of the equilibrium constant is the same regardless of the size of the cluster formed. A challenge in chemical theory that has yet to be resolved is that the types of species that may be formed in solution must be known, or assumed, *a priori* and the equations for multicomponent mixtures with arbitrary sites are complex.

More recently, perturbation theories such as Wertheim's theory have become more common in thermodynamic modeling. Developed in the mid-1980s [47–50,65], Wertheim's perturbation theory is a statistical mechanics approach derived for molecules with a repulsive core and attractive "sites" and it calculates the probabilities of hydrogen bonds forming between sites. The method can be used for any type of association including charge transfer complexes. For hydrogen bonding, acceptor sites are allocated to the electronegative atom(s) which will accept a proton and donor sites are allocated to the protons. Huang and Radosz [55] tabulate schemes that have been developed for site allocations and the common 2B scheme used for alcohols is given in Chapter 3. and repeated here (Figure 4.1) for the reader's convenience.



Figure 4.1: 2B association scheme. The 2B scheme is commonly used to model alcohols by calculating the association between an acceptor site on the oxygen and a donor site on the hydrogen.

The probability of association is described by an association constant, $\Delta^{A_i D_j}$ between each binary pair. Wertheim's theory has been implemented into several equations of state such as CPA [11,52] and the SAFT family of equations [9,54–56,95]. In our previous work [96], we created an activity coefficient association contribution with Wertheim's theory by adapting methods developed by Michelsen and Hendricks [66]. We also show that with reasonable simplifications, the association constant $\Delta^{A_i D_j}$ is equivalent to the concentration-based equilibrium constant K_c and that the two association theories yield the same results for the activity coefficient for alcohol + inert and alcohol + alcohol systems. This aligns with prior work beginning with Elliott et al. [10] who were the first to document the relationship between the activity-based chemical association constant, K_{a} , and perturbation theories. Economou and Donohue [58] then used this equivalence analytically for pure and binary mixtures in which associating components possess 0, 1 or 2 sites each. Similarly, Campbell et al. [16,85] recognized the same parallels using the concentration-based equilibrium constant K_C algebraically for mixtures containing any number of alcohols and, later, for those containing any number of alkanes and any number of alcohols. Perhaps the most significant contribution of Campbell's earlier work is the novel method for counting clusters in solution by tracking the cluster size and end molecule identity of each cluster. While this scheme is innovative, it's implementation is time-consuming and cumbersome for more complex mixtures, such as water + alcohol, where the number of sites per molecule is greater.
In this work, we outline the calculations in chemical and Wertheim theories for a hypothetical ternary mixture in which one of the components possesses two identical acceptor sites and no donor sites. The presence of such a component complicates chemical theory because certain species resulting from hydrogen bonding can be formed in multiple ways. The main objective of the current work is to prove that the species calculation used in Wertheim's theory is in agreement with that of chemical theory. To this end, we implement a counting scheme and show that the monomer fractions obtained by the two association theories are identical.

4.2. Theory and analysis

The hypothetical system chosen for this work consists of a component with two identical acceptor sites, such as 1,4-dioxane, a component with one donor site, such as chloroform (where the electronegativity of the chlorine atoms makes the hydrogen atom susceptible to hydrogen bonding), and a component with one acceptor and one donor site, which is the "2B" scheme often used to model alcohols. We use *E*, *F*, *G* to designate components rather than *A*, *B*, *C* to avoid confusion in use of *A* for acceptor and *C* for concentration. The components are illustrated in Figure 4.2.



Figure 4.2: Hypothetical molecules for analysis. The system studied in this work is a ternary mixture of these three molecule types, where the acceptor sites are solid dots and the donors are open dots

4.2.1. Implementation of Wertheim's perturbation theory

At equilibrium, molecules which hydrogen bond form clusters of various geometries and sizes. We refer to the molecules that remain free without any sites bonded as "monomers", a term borrowed from polymer science. For both chemical theory and Wertheim's theory, the fraction of molecules that remain as monomers at equilibrium is the key term on which all subsequent calculations depend. Therefore, it is of particular importance that this quantity is known with certainty. When calculating mole balances, a distinction is important between a balance which counts individual molecules as if no association is occurring, the apparent mole balance, and that which counts each species formed after association has taken place in solution, the true mole balance. The number of species decreases when bonding occurs. Therefore, the true total mole count, n_T , is always less than the apparent mole count, n_0 .

To study the monomer fractions calculated by Wertheim's perturbation theory, the equation for the fraction of free sites, $X^{B_i} = n^{B_i,free}/n^{B_i,total}$, is used. This X represents the fraction of sites of type B_i that remain non-bonded at equilibrium, where $n^{B_i,free}$ is the number of moles of site B_i that are free and $n^{B_i,total}$ is the total number of moles of site B_i in solution. Variable *i* can be an index or can be a site host identity; in this work we use the site host identity because the two acceptor sites on *E* are equivalent. *B* is a generic variable that can represent an acceptor *A*, donor *D*, or carboxylic acid site *C* which describe a site that can bond with other *C* to form dimers. X^{B_i} is calculated via

$$X^{B_{i}} = \left(1 + \sum_{\substack{mating\\sites\ mate_{j}}} x_{mate_{j}\ host} N_{mate_{j}} X^{mate_{j}} \rho \Delta^{B_{i}mate_{j}}\right)^{-1}$$

$$4.1$$

where $mate_j$ is generic site notation that can represent an acceptor, donor, or *C*-site to which B_i can bond. Variable $x_{mate_j host}$ is the apparent mole fraction of the component which hosts $mate_j$, N_{mate_j} is the number of sites identical to $mate_j$ on the host for $mate_j$. ρ is the solution's apparent molar density and $\Delta^{B_i mate_j}$ is the association constant for an interaction between $mate_i$ and B_i . Applying Eq. (1.4) to the four unique sites in the hypothetical system yields the equations in Table 4.1.

Table 4.1: X^{B_i} Equations for sites in hypothetical system. The equations are used to calculate
the fraction of free sites of each type

Interactions involving	Wertheim X ^{B_i} Equation	Eq
$A_{E,}$ Acceptors on E	$1 - X^{A_E} = X^{A_E}(x_F \rho X^{D_F} \Delta^{A_E D_F} + x_G \rho X^{D_G} \Delta^{A_E D_G})$	4.2
A_{G} , Acceptor on G	$1 - X^{A_G} = X^{A_G}(x_F \rho X^{D_F} \Delta^{A_G D_F} + x_G \rho X^{D_G} \Delta^{A_G D_G})$	4.3
$D_{G_{i}}$ Donor on G	$1 - X^{D_G} = X^{D_G} (2x_E \rho X^{A_E} \Delta^{A_E D_G} + x_G \rho X^{A_G} \Delta^{A_G D_G})$	4.4
$D_{F,}$ Donor on F	$1 - X^{D_F} = X^{D_F} (2x_E \rho X^{A_E} \Delta^{A_E D_F} + x_G \rho X^{A_G} \Delta^{A_G D_F})$	4.5

Here, X^{A_E} , X^{A_G} , X^{D_G} and X^{D_F} are the free site fractions of the acceptors on component *E* and *G* and donors on *G* and *F* respectively. For a given molecule of component *E*, the monomer fraction is equal to the joint probabilities that the first acceptor is free, *and* the second acceptor is free. If interactions at the two sides are considered to be independent of one another and recognizing that X^{B_i} represents the probability that site B_i is free

$$\frac{n_{M,E}}{n_{0,E}} = P(Acceptor \ 1 \ free \ \cap \ Acceptor \ 2 \ free) = X^{A_E} \cdot X^{A_E} = (X^{A_E})^2$$

$$4.6$$

where $n_{M,E}$ is the moles of *E* monomer, n_E is the total apparent moles of *E* molecules and the *P* represents probability. Similarly, for components *F* and *G*:

$$\frac{n_{M,F}}{n_{0,F}} = P(Donor\ free\) = X^{D_F}$$

$$4.7$$

$$\frac{n_{M,G}}{n_{0,G}} = P(Acceptor\ free\ \cap\ Donor\ free) = X^{A_G} X^{D_G}$$

$$4.8$$

4.2.2. Implementation of chemical theory

Calculating monomer fractions with chemical theory is more complicated than Wertheim's theory for two reasons. First, chemical theory is a *species*-wise approach rather than a *site*-wise approach. Therefore, one must have *a priori* knowledge of all the possible species that may form in a solution of components. Figure 4.3 shows the bonding schemes that can form in the hypothetical ternary

mixture studied here. Because molecules of component G possess both an acceptor and a donor site, they can form chains in solution. Therefore, it should be noted that each molecule of component G in Figure 4.3 can be replaced by a chain of G molecules.



Figure 4.3: Bonding schemes for hydrogen bonding in a mixture of *E***,** *F* **and** *G***.** The bonding schemes circled and denoted by * can be formed in two ways.

As with Wertheim's theory, the species formed in solution can be described with chemical theory by four key types of association upon which the remaining associations are related. Each can be described by an equilibrium constant as shown Table 4.2.

$E + F \rightleftharpoons EF$	K _{C1}	$K_{C1} = \frac{C_{T,EF}}{2C_{T,E}C_{T,F}}$	4.9
$G_M + G_M \rightleftharpoons G_2$ $G_M + G_i \rightleftharpoons G_{i+1}$	К _{С2}	$K_{C2} = \frac{C_{T,G_2}}{C_{T,G_M}C_{T,G_M}} = \frac{C_{T,G_{i+1}}}{C_{T,G_M}C_{T,G_i}}$	4.10
$E + G_M \rightleftharpoons EG$	K _{C3}	$K_{C3} = \frac{C_{T,EG}}{2C_{T,E}C_{T,G_M}}$	4.11
$F + G_M \rightleftharpoons FG$	K _{C4}	$K_{C4} = \frac{C_{T,FG}}{C_{T,F}C_{T,G_M}}$	4.12

 Table 4.2: Key associations in the hypothetical system and their equilibrium constants.

A subscript of T on concentration indicates a true value. The common assumption is that the change in Gibbs energy for each bond type is independent of cluster size, thus logarithms of the equilibrium constants add, and the K_c for any compound is the product of the K_c 's for the constituent species. From these four interactions, the hydrogen bonded clusters can be formed as summarized in Table 4.3.

 Table 4.3: Other species formed in hypothetical system. This table outlines the equilibrium constants associated with each species formation and the relation to the key association constants.

$E + 2F \rightleftharpoons EF_2$	$K_{C5} = K_{C1}^2$	$K_{C5} = \frac{C_{T, EF_2}}{C_{T, E} C_{T, F}^2}$	4.13
$G_M + iG_M \rightleftharpoons G_{i+1}$	$K_{C6} = K_{C2}^i$	$K_{C6} = \frac{C_{T,G_{i+1}}}{C_{T,G_M} C_{T,G_M}^i}$	4.14
$E + iG_M \rightleftharpoons EG_i$	$K_{C7} = K_{C3} K_{C2}^{i-1}$	$K_{C7} = \frac{C_{T,EG_i}}{2C_{T,E}C_{T,G_M}^i}$	4.15
$F + iG_M \rightleftharpoons FG_i$	$K_{C8} = K_{C4} K_{C2}^{i-1}$	$K_{C8} = \frac{C_{T,FG_i}}{C_{T,F}C_{T,G_M}^i}$	4.16
$E + iG_M + jG_M \rightleftharpoons G_i EG_j$	$K_{C8} = K_{C3}^2 K_{C2}^{i+j-2}$	$K_{C8} = \frac{C_{T,G_i EG_j}}{C_{T,E} C_{T,G_M}^{i+j}}$	4.17
$E + iG_M + jG_M + F \iff FG_iEG_j$	$K_{C9} = K_{C3}^2 K_{C2}^{i+j-2} K_{C4}$	$K_{C9} = \frac{C_{T,FG_iEG_j}}{2C_{T,F}C_{T,G_M}^{i+j}C_{T,E}}$	4.18
$iG_M + E + F \rightleftharpoons G_i EF$	$K_{C10} = K_{C3} K_{C1} K_{C2}^{i-1}$	$K_{C10} = \frac{C_{T,G_i EF}}{2C_{T,G_M}^i C_{T,E} C_{T,F}}$	4.19
$E + iG_M + F \rightleftharpoons EG_iF$	$K_{C11} = K_{C3} K_{C4} K_{C2}^{i-1}$	$K_{C11} = \frac{C_{T,EG_iF}}{2C_{T,E}C_{T,G_M}^i C_{T,F}}$	4.20
$F + \overline{E + iG_M} + F \rightleftharpoons \overline{F}EG_iF$	$K_{C12} = K_{C1} K_{C3} K_{C4} K_{C2}^{i-1}$	$K_{C12} = \frac{C_{T,FEG_iF}}{2C_{T,E}C_{T,G_M}^i C_{T,F}^2}$	4.21
$E + iG_i + jG_j + 2F \rightleftharpoons FG_i EG_j F$	$K_{C13} = K_{C4}^2 K_{C3}^2 K_{C2}^{i+j-2}$	$K_{C13} = \frac{C_{T,FG_iEG_jF}}{C_{T,E}C_{T,G_M}^{i+j}C_{T,F}^2}$	4.22

The second complexity with chemical theory is correctly accounting for species that can be formed in multiple ways and which of the ways are indistinguishable by symmetry. In our hypothetical system, all species formed through an unsymmetrical bonding scheme, denoted by an asterisk in Figure 4.3, can be formed in two ways as illustrated for species EF in Figure 4.4. This occurs because E has two identical acceptor sites. To account for these multiple pathways, the equilibrium constant equation for these "reactions" contains a factor of 2 in Table 4.2 and Table 4.3.



Figure 4.4: Counting species in chemical theory system. Some species may be formed through multiple pathways.

Next, we outline a careful implementation of chemical theory to a mole balance of the three components in our system. To begin with, the mole balance for component E is given by

$$x_{E}\rho = C_{T,E} + C_{T,EF} + C_{T,EF2} + \sum_{i=1}^{\infty} (C_{T,EG_{i}} + C_{T,G_{i}EF} + C_{T,EG_{i}F} + C_{T,FEG_{i}F}) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (C_{T,G_{i}EG_{j}} + C_{T,FG_{i}EG_{j}} + C_{T,FG_{i}EG_{j}F})$$

$$4.23$$

The double sums in Eq. (4.23) are simplified to a converging series with a closed solution. This calculation is demonstrated in Appendix F for $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} C_{T,G_iEG_j}$. Substituting the equilibrium constant relations in Table 4.2 and Table 4.3 for the species concentrations and putting all K_C values in terms of K_{C1}, K_{C2}, K_{C3} and K_{C4} :

$$\frac{C_{T,E}}{x_E \rho} = \left(1 - K_{C2} C_{T,G_M}\right)^2 \left[\left(1 + K_{C1} C_{T,F}\right)^2 \left(1 - K_{C2} C_{T,G_M}\right)^2 + 2 K_{C3} C_{T,G_M} \left(1 - K_{C2} C_{T,G_M}\right) \left(1 + K_{C1} C_{T,F}\right) \left(1 + K_{C4} C_{T,F}\right) + \left(K_{C3} C_{T,G_M}\right)^2 \left(1 + K_{C4} C_{T,F}\right)^2 \right]^{-1}$$
4.24

Similarly, the mole balance for F and G can be calculated and rearranged as

$$\begin{aligned} x_{F}\rho &= C_{T,F} + C_{T,EF} + 2C_{T,EF2} + \sum_{i=1}^{\infty} (C_{T,FG_{i}} + C_{T,G_{i}EF} + C_{T,EG_{i}F} + 2C_{T,FEG_{i}F}) \\ &+ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (C_{T,FG_{i}EG_{j}} + 2C_{T,FG_{i}EG_{j}F}) \end{aligned}$$

$$\begin{aligned} 4.25 \\ &+ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (C_{T,G_{i}})^{2} \left[\left(1 + 2K_{C1}C_{T,E} \left(1 + K_{C1}C_{T,F} \right) \right) \left(1 - K_{C2}C_{T,G_{M}} \right)^{2} \\ &+ \left(1 - K_{C2}C_{T,G_{M}} \right) \left(K_{C4}C_{T,G_{M}} \left(1 + 2K_{C3}C_{T,E} \right) \\ &+ 2K_{C1}K_{C3}C_{T,E}C_{T,G_{M}} \left(1 + 2K_{C4}C_{T,F} \right) \right) \\ &+ 2K_{C3}^{2}K_{C4}C_{T,E}C_{T,G_{M}}^{2} \left(1 + K_{C4}C_{T,F} \right) \right]^{-1} \end{aligned}$$

$$\begin{aligned} x_{G}\rho &= \sum_{i=1}^{\infty} \left(iC_{T,G_{i}} + iC_{T,EG_{i}} + iC_{T,FG_{i}} + iC_{T,G_{i}EF} + iC_{T,EG_{i}F} + iC_{T,FEG_{i}F} \right) \\ &+ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left(i + j \right) \left(C_{T,G_{i}EG_{j}} + C_{T,FG_{i}EG_{j}} + C_{T,FG_{i}EG_{j}F} \right) \end{aligned}$$

$$\begin{aligned} 4.27 \\ &+ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left(i + j \right) \left(C_{T,G_{i}EG_{j}} + C_{T,FG_{i}EG_{j}} + C_{T,FG_{i}EG_{j}F} \right) \\ &- K_{C2}C_{T,G_{M}} \right) + 2K_{C3}^{2}C_{T,E}C_{T,G_{M}} \left(1 + K_{C4}C_{T,F} \right)^{2} \right]^{-1} \end{aligned}$$

The simplification of the double sums in Eqs. (4.25) and (4.27) is demonstrated in Appendix G.

4.2.3. Numerical equivalence of chemical and Wertheim's theories

To investigate parallels between the two association theories, it should be recognized that monomer fractions of each component are related by

$$\frac{C_{T,E}}{x_E\rho} = (X^{A_E})^2 \tag{4.29}$$

$$\frac{C_{T,F}}{x_F\rho} = X^{D_F} \tag{4.30}$$

$$\frac{C_{T,G_M}}{x_{G_M}\rho} = X^{A_G} X^{D_G}$$

$$4.31$$

where the left and right-hand sides pertain to chemical theory and Wertheim's theory respectively and the mixture density is calculated by assuming no excess volume of mixing, $\rho = 1/\sum_i (x_i/\rho_i)$. Based on the simplifications of our earlier work [96], the concentration-based equilibrium constant, as defined in Table 4.2 and Table 4.3, is equivalent to the association strength parameter, $\Delta^{B_iC_j}$. Therefore,

$$K_{C1} = \Delta^{A_E D_F}$$
 $K_{C2} = \Delta^{A_G D_G}$ $K_{C3} = \Delta^{A_E D_G}$ $K_{C4} = \Delta^{A_G D_F}$ 4.32

For this study, the chosen vales of the association constants do not represent values expected from experiment. Rather, we have chosen parameter values that yield concentrations which are visible on the plot axes (see Table 4.4). Standard liquid densities of dioxane, chloroform and methanol were adopted for components E, F, G respectively as summarized in Table 4.4.

Table 4.4: Association parameters and component densities. Concentration based equilibrium constants are equivalent to $\Delta^{A_i B_j}$ in Wertheim's theory with the simplifications used here

Association Parameters (cm ³ /mol)					
$K_{C1} = \Delta^{A_E D_F} $ 5000					
$K_{C2} = \Delta^{A_G D_G}$	10000				
$K_{C3} = \Delta^{A_E D_G}$	7000				
$K_{C4} = \Delta^{A_G D_F}$	12000				
Component Densities (mol/cm ³)					
$ ho_E$	1.167E-02				
$ ho_F$	1.242E-02				
ρ_G 2.479E-02					

Next, the values of the monomer fraction of each component is calculated. For both association theories, this involves solving simultaneous equations using successive substitution or other iterative techniques. With chemical theory, Eqs. (4.24), (4.26) and (4.28) are solved simultaneously to find values of $C_{T,E}$, $C_{T,F}$ and C_{T,G_M} . Similarly, for Wertheim's theory, the equations in Table 4.1 are solved simultaneously to yield values for X^{A_E} , X^{A_G} , X^{D_G} and X^{D_F} . Both methods yield identical results and this is demonstrated for a solution composition of $x_E = 0.2$, $x_F = 0.2$ and $x_{G_M} = 0.6$ in Table 4.5 and Table 4.6. The selected composition provides an environment where complex formation is extensive and only small fractions of the available sites

are free. Because the number of species is infinite, the chains are summed in Table 4.6 by bonding scheme.

Table 4.5: Calculated monomer fraction values. Component monomer fractions for a solution composition of $x_E = 0.2$, $x_F = 0.2$ and $x_{G_M} = 0.6$ where $\rho = 1.741E-02$ mol/cm³

	Chemical Theory	Wertheim's Theory	Monomer Fraction Value
Monomer fraction of <i>E</i>	$\frac{C_{T,E}}{x_E \rho}$	$(X^{A_E})^2$	7.50E-02
Monomer fraction of <i>F</i>	$\frac{C_{T,F}}{x_F \rho}$	X^{D_F}	2.92E-02
Monomer fraction of <i>G</i>	$\frac{C_{T,G_M}}{x_{G_M}\rho}$	$X^{A_G}X^{D_G}$	5.55E-03

While Wertheim's theory is easier to implement to find the monomer fractions, it does not directly provide species distributions. Table 4.6 demonstrates how species distributions can be obtained from both methods.

Species	Chemical Theory $C_{T,k}$	Wertheim's Theory C _{T,k}	$C_{T,k} \times 10^4 (\mathrm{mol/cm^3})$
$C_{T,E}$	Eq. (4.24)	$x_E ho(X^{A_E})^2$	2.61
$C_{T,F}$	Eq. (4.26)	$x_F ho X^{D_F}$	1.02
C_{T,G_M}	Eq. (4.28)	$x_G \rho X^{A_G} X^{D_G}$	0.58
$\sum_{i=1}^{\infty} C_{T,G_i}$	$\frac{C_{T,G_M}}{1-C_{T,G_M}K_{C2}}$	$\frac{x_G \rho X^{A_G} X^{D_G}}{1 - r_G \rho X^{A_G} X^{D_G} \Lambda^{A_G D_G}}$	1.38
$C_{T,EF}$	$\frac{2C_{T,F}C_{T,E}K_{C1}}{2C_{T,F}C_{T,E}K_{C1}}$	$\frac{1}{2x_E\rho(X^{A_E})^2} x_F\rho X^{D_F} \Delta^{A_E D_F}$	2.65
$C_{T,FEF}$	$C_{T,F}^2 C_{T,E} K_{C1}^2$	$(x_F \rho X^{D_F})^2 x_E \rho (X^{A_E})^2 (\Delta^{A_E D_F})^2$	0.67
$\sum_{i=1}^{\infty} C_{T,FG_i}$	$\frac{C_{T,F}C_{T,G_M}}{1 - C_{T,G_M}K_{C2}}K_{C4}$	$\frac{(x_F \rho X^{D_F})(x_G \rho X^{A_G} X^{D_G})}{1 - x_G \rho X^{A_G} X^{D_G} \Delta^{A_G D_G}} \Delta^{A_G D_F}$	1.68
$\sum_{i=1}^{\infty} C_{T,EG_i}$	$\frac{2C_{T,E}C_{T,G_M}}{1 - C_{T,G_M}K_{C2}}K_{C3}$	$\frac{2(x_E\rho(X^{A_E})^2)(x_G\rho X^{A_G}X^{D_G})}{1-x_G\rho X^{A_G}X^{D_G}\Delta^{A_GD_G}}\Delta^{A_ED_G}$	5.04
$\sum_{i=1}^{\infty} C_{T,G_i EF}$	$\frac{2C_{T,G_M}C_{T,E}C_{T,F}}{1-C_{T,G_M}K_{C2}}K_{C1}K_{C3}$	$\frac{2(x_G\rho X^{A_G}X^{D_G})(x_E\rho (X^{A_E})^2)(x_F\rho X^{D_F})}{1-x_G\rho X^{A_G}X^{D_G}\Delta^{A_GD_G}}\Delta^{A_ED_F}\Delta^{A_ED_G}$	2.56
$\sum_{i=1}^{\infty} C_{T,EG_iF}$	$\frac{2C_{T,E}C_{T,G_M}C_{T,F}}{1 - C_{T,G_M}K_{C2}}K_{C3}K_{C4}$	$\frac{2(x_E\rho(X^{A_E})^2)(x_G\rho X^{A_G}X^{D_G})(x_F\rho X^{D_F})}{1-x_G\rho X^{A_G}X^{D_G}\Delta^{A_GD_G}}\Delta^{A_ED_G}\Delta^{A_GD_F}$	6.15
$\sum_{i=1}^{\infty} C_{T,FEG_iF}$	$\frac{2C_{T,G_M}C_{T,E}C_{T,F}^2}{1-C_{T,G_M}K_{C2}}K_{C1}K_{C3}K_{C4}$	$\frac{2(x_G\rho X^{A_G}X^{D_G})(x_E\rho (X^{A_E})^2)(x_F\rho X^{D_F})^2}{1-x_G\rho X^{A_G}X^{D_G}\Delta^{A_GD_G}}\Delta^{A_ED_F}\Delta^{A_ED_G}\Delta^{A_GD_F}$	3.12
$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}C_{T,G_iEG_j}$	$\frac{C_{T,G_M}^2 C_{T,E}}{\left(1 - C_{T,G_M} K_{C2}\right)^2} K_{C3}^2$	$\frac{(x_G \rho X^{A_G} X^{D_G})^2 (x_E \rho (X^{A_E})^2)}{(1 - x_G \rho X^{A_G} X^{D_G} \Delta^{A_G D_G})^2} (\Delta^{A_E D_G})^2$	2.43
$\sum_{i=1}^{\infty} \overline{\sum_{j=1}^{\infty} C_{T,FG_iEG_j}}$	$\frac{2C_{T,F}C_{T,G_M}^2C_{T,E}}{\left(1 - C_{T,G_M}K_{C2}\right)^2}K_{C3}^2K_{C4}$	$\frac{2(x_F \rho X^{D_F})(x_G \rho X^{A_G} X^{D_G})^2 (x_E \rho (X^{A_E})^2)}{(1 - x_G \rho X^{A_G} X^{D_G} \Delta^{A_G D_G})^2} (\Delta^{A_E D_G})^2 \Delta^{A_G D_F}$	5.94
$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} C_{T,FG_iEG_jF}$	$\frac{C_{T,F}^2 C_{T,G_M}^2 C_{T,E}}{\left(1 - C_{T,G_M} K_{C2}\right)^2} K_{C3}^2 K_{C4}^2$	$\frac{(x_F \rho X^{D_F})^2 (x_G \rho X^{A_G} X^{D_G})^2 (x_E \rho (X^{A_E})^2)}{(1 - x_G \rho X^{A_G} X^{D_G} \Delta^{A_G D_G})^2} (\Delta^{A_E D_G})^2 (\Delta^{A_G D_F})^2}$	3.62

Table 4.6: Chemical and Wertheim's theory concentration equations and calculated values at $x_E = 0.2$, $x_F = 0.2$ and $x_{G_M} = 0.6$.

Next, we explore the species distribution in solution by varying the composition of the ternary mixture. For the plots shown in Figures 4.5-4.8, x_F is fixed at 0.2 (with the exception of the pure *G* composition point), x_G is varied from 0 to 1 and the balance of the composition is x_E .



Figure 4.5: True monomer fraction of components in system of study and fractions of acceptors and donors. The mole fraction of component *F* is fixed at 0.2.



Figure 4.6: Distribution of species with component *E* when $x_F=0.2$. The bottom plot is a zoomed panel of the top plot



Figure 4.7: Distribution of species with component F when $x_F=0.2$. The bottom plot is a zoomed panel of the top plot



Figure 4.8: Distribution of species with component *G* when $x_F=0.2$. The bottom plot is a zoomed panel of the top plot

Figure 4.5 shows the mole fraction of monomers for each component and the percentage of acceptors and donors in solution. The successive three plots show how components E, F and G are distributed in solution.

The fraction of acceptors and donors in solution are calculated by the ratio of acceptors (or donors) to the total sites in solution. Acceptors are in excess to the left of $x_G = 0.7$ and because the association constants are large, almost all of the donors are bonded until above $x_G \approx 0.6$ where some *F* remains free. In Figure 4.6 the true mole fraction of monomer *E*, $(z_E = n_{T,E}/n_T = \rho_{T,E}/\rho_T)$, stays near the apparent mole fraction of *E* on the left side of the figure because, although the true moles of *E* decrease due to bonding, the true mole fractions are calculated using the total moles of all species after bonding, which is less than the apparent moles, $(n_T < n_0 \text{ or } \rho_T < \rho)$, used to calculate the apparent mole fraction $(x_E = n_E/n_0)$. For example, at the composition $x_G = 0$ and $x_E = 0.8$, $\rho_{T,E} = 7.26 \times 10^{-3} \text{ mol/cm}^3$, $\rho_T = 9.48 \times 10^{-3} \text{ mol/cm}^3$ and $\rho = 1.18 \times 10^{-2} \text{ mol/cm}^3$. This yields a z_E value of 0.766. Moreover, almost all the *F* is associated, and the dominant species are *E* and *EF*. At $x_G = 0.4$ the apparent *E* and apparent *G* are equimolar, so the *EG_i* compounds are favored. At $x_G = 0.6$, the ratios *F/G* and *E/G* are equal, and thus the dominant species are *EG_iF*, *FG_iEG_i*, and *EG_i*.

The figures show the sum of mole fractions for oligomers rather than oligomer distributions. For a pure solution of a chain-forming compound such as component G, it can be shown that the true mole fraction of monomers is always larger than the mole fraction of any oligomer species. The number of moles of a chain of G molecules of length i is [8,67]:

$$\frac{n_i}{n_T} = z_{G_M} (\rho \Delta^{A_G D_G} (X^{A_G})^2)^{i-1}$$
4.33

where n_T is the total true number of moles and z_{G_M} is the true monomer fraction of *G*. Note that for pure component *G*, $X^{A_G} = X^{D_G}$ and therefore, $X^{A_G}X^{D_G}$ can be written as $(X^{A_G})^2$. Further, it is shown that:

$$z_{G_M} = \frac{n_T}{n_0} = X^{A_G} \tag{4.34}$$

where n_0 is the total apparent number of moles. Combining Eqs. (4.33) and (4.34) yields:

$$\frac{n_i}{n_0} = (X^{A_G})^{2i} (\rho \Delta^{A_G D_G})^{i-1}$$
4.35

Using this equation, the relationship between n_i and n_{i+1} can be calculated to be:

$$\frac{n_{i+1}}{n_i} = \rho \Delta^{A_G D_G} (X^{A_G})^2 \tag{4.36}$$

Recognizing that $(X^{A_G})^2$ is the monomer fraction in solution, we note an interesting trend. As the value of $\Delta^{A_G D_G}$ is increased, molecules are more likely to form bonds and therefore $(X^{A_G})^2$ decreases. This occurs such that $\rho \Delta^{A_G D_G} (X^{A_G})^2$ is always a fraction and n_{i+1} is less than n_i . In other words, there are always more monomers in solution than dimers, more dimers than trimers and so on. This trend is calculated with Wertheim's theory in Figure 4.9 for pure methanol at $\rho_{MeOH} = 2.479 \times 10^{-2} \text{ mol/cm}^3$ for four values of $\Delta^{A_G D_G}$. As a reference, the cubic-plus-association (CPA) equation of state value for the $\Delta^{A_G D_G}$ for pure liquid methanol with itself is approximately 16000 cm³/mol at 298.15 K. More interestingly, as $\Delta^{A_G D_G}$ approaches infinity, the distribution of various cluster sizes becomes uniform.



Figure 4.9: True mole fraction of methanol chains calculated with 4 values of $\Delta^{A_G D_G}$

4.2.4. Activity coefficients

As further demonstration of the equivalence of Wertheim and chemical theory methods, we illustrate the calculation of activity coefficients. In our previous work [96], we demonstrated that the association contribution to the activity coefficients, assuming van der Waals correction to Flory's equation and no excess volume, results in:

$$\ln \gamma_k^{assoc} = \sum_{\substack{sites \ B_i \\ on \ k}} N_{B_i} \left[\ln \left(\frac{X^{B_i}}{X^{B_i,0}} \right) - \frac{1}{2} (1 - X^{B_i,0}) \right] + \frac{1}{2} \frac{\rho}{\rho_k} \sum_{\substack{all \\ sites \ B_i}} x_{B_i \ host} N_{B_i} (1 - X^{B_i})$$

$$4.37$$

where the superscript 0 represents the pure state for component k. In our previous work [96], the logarithm operator was missing on Equation A14 and it should have appeared as

$$\ln \gamma_k^{assoc} = \frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} \bigg|_{mix} - \frac{1}{RT} \left(\frac{\partial \underline{A}^{assoc}}{\partial n_k} \right)_{T,P,n_{j \neq k}} \bigg|_{pure k}$$

$$4.38$$

The activity coefficient from chemical theory is

$$\ln \gamma_k^{chem} = \ln \left(\frac{\rho_{k_M}}{x_k \rho_{k_M}^0} \right) + \frac{1}{\rho_k} \left(\rho_{T,k}^0 - \rho_T \right)$$

$$4.39$$

Note that the molar densities are equivalent numerically to the molar concentrations used above in the reaction equilibria. The previous work [96] showed that the two approaches are related by Flory's equation through $\ln \gamma_k^{chem} = \ln \gamma_k^{assoc} + \ln \gamma_k^{Flory}$. Table 4.7 summarizes calculation of the activity coefficients from the two approaches using the concentrations determined above. Figure 4.10 shows the activity coefficient contributions for mixtures with increasing compositions of component *G* when the mole fraction of *F* is fixed at 0.2. The association parameters and the pure component densities used in all calculations are given in Table 4.4.

Table 4.7: Calculated parameters and intermediate values for activity coefficient calculations. The selected composition is $x_E = 0.2$, $x_F = 0.2$ and $x_G = 0.6$ and $\rho = 1.741E-02$ mol/cm³, $\rho_T = 3.889E-03$ mol/cm³

Component k		Ε	F	G
	x_k	0.2	0.2	0.6
	$ ho_{k_M}$	2.62E-04	1.01E-04	5.80E-05
Chemical	$ ho_{k_M}^0$	1.17E-02	1.24E-02	9.38E-05
theory	$ ho_{T,k}^0$	1.17E-02	1.24E-02	1.53E-03
	$\ln \gamma_k^{chem}$	-1.52	-2.51	-6.65E-02
	X^A	0.274	-	0.190
	X^D	-	2.91E-02	2.93E-02
	$X^{A,0}$	1.00	-	6.15E-02
Wertheim's	$X^{D,0}$	-	1.00	6.15E-02
Theory	$\ln \gamma_k^{assoc}$	-1.43	-2.45	-1.08E-02
	$\ln \gamma_k^{Flory}$	-9.17E-02	-6.40E-02	-5.57E-02
	$\ln \gamma_k^{chem}$	-1.52	-2.51	-6.65E-02



Figure 4.10: Calculated values of the activity coefficient contributions. For each component, Wertheim's theory (---), chemical theory (---), and Flory's theory (----) contributions are shown as a function of the apparent mole fraction of *G*. The mole fraction of component *F* is fixed at 0.2.

4.2.5. Perspective on further generalization

It is worth noting that the system chosen for this study is intentionally rather simple compared to those encountered in the chemical industry. By definition, components which contain only one type of site, such as components E and F, limit the extent of association in solution. When a molecule of F binds to an acceptor site, it seals one end of the chain from further association. With mixtures of water, alcohols and/or glycols, hydrogen bonding is extensive and can create cyclic species that cannot be captured by Wertheim's theory [97]. Though chemical theory can capture cyclic species [13,98], specifying the species to include in solution is challenging. However, even

without considering cyclic species, the intricacy of chemical theory is in stark contrast with the simple elegance of Wertheim's theory where all site bonding is calculated using the site balance introduced in Eq. (1.4). Chemical theory requires correct identification of all possible bonding schemes and the number of combinations through which each can form. It is thus this complexity that has increased the popularity of Wertheim's theory in the area of thermodynamic modeling. The parallels between chemical and Wertheim's theories have been noted by several independent researchers [10,16,58]. Indeed, in our own previous work [96], the algebraic and numerical equivalence of the two theories is shown for simple mixtures containing chain-forming species, specifically pure alcohol and alcohol + inert type systems. However, in those cases, the alcohol, modeled in the same way as component *G* here, results in only chain equations that are more easily manipulated. The presence of species *E* creates a complexity that has not been previously demonstrated. To our knowledge, no preceding work has shown the numerical equivalence of the two methods for multicomponent non-chain-forming species with this rigor.

4.3. Conclusions

In this work, we explore the numerical equivalence of chemical and Wertheim's theory for a hypothetical system that contains a chain-forming component and molecules with multiple sites which can only solvate. For chemical theory, the calculations are outlined carefully to elucidate the oftentimes complicated counting procedure required to account for all associating species. Moreover, the effect of the value of the association parameters on derived monomer fraction values is studied. We demonstrated how Wertheim site free fractions can be used to calculate species concentrations. It is found that, when implemented carefully, Wertheim and chemical theory yield identical species distributions for the given system, providing that cyclic species are not present. Because the system of study is a mixture of generalized non-specific components, we expect that

the work presented here can be extended to all other chemical systems that do not form cyclic species. Although both theories are fundamentally the same, Wertheim's theory has a distinct advantage due to the simplicity of its universal equations and is therefore recommended as the superior method for determining monomer fractions.

APPENDICES

Appendix F: Simplification of double sums without a multiplication factor

Here, we demonstrate double sum simplifications such as that of

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} C_{T,G_i E G_j} = (C_{T,G_1 E G_1} + C_{T,G_1 E G_2} + C_{T,G_1 E G_3} + \cdots) + (C_{T,G_2 E G_1} + C_{T,G_2 E G_2} + C_{T,G_2 E G_3} + \cdots) + \cdots$$
F.1

Careful inspection of the equilibrium constants shows that whenever i + j sum to the same value, all concentrations where *G* is redistributed on both sides of *E* are identical, i.e.

$$C_{T,G_2EG_2} = C_{T,G_3EG_1} = C_{T,G_1EG_3} = \cdots$$
 F.2

Therefore, the double sum can be written as a single sum:

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} C_{T,G_i E G_j} = \sum_{i=1}^{\infty} i C_{T,G_1 E G_i} = C_{T,G_1 E G_1} + 2C_{T,G_1 E G_2} + 3C_{T,G_1 E G_3} + \cdots$$
F.3

Substituting the definition for the equilibrium constant for this reaction from Eq. (4.17):

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} C_{T,G_i E G_j} = K_{C3}^2 C_{T,E} C_{T,G_M}^2 + 2K_{C3}^2 C_{T,E} C_{T,G_M}^2 (K_{C2} C_{T,G_M})$$
F.4

$$+ 3K_{C3}^{2}C_{T,E}C_{T,G_{M}}^{2}(K_{C2}C_{T,G_{M}})^{2} + \cdots$$

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}C_{T,G_{i}EG_{j}} = K_{C3}^{2}C_{T,E}C_{T,G_{M}}^{2}\left(1 + 2K_{C2}C_{T,G_{M}} + 3(K_{C2}C_{T,G_{M}})^{2} + \cdots\right)$$
F.5

Recognizing the converging series

$$\sum_{i=0}^{\infty} (i+1)a^i = \frac{1}{(1-a)^2} = 1 + 2a + 3a^2 + \cdots$$
 F.6

results in the following simple term:

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} C_{T,G_i E G_j} = \frac{K_{C3}^2 C_{T,E} C_{T,G_M}^2}{\left(1 - K_{C2} C_{T,G_M}\right)^2}$$
F.7

Appendix G: Simplification of double sums with a multiplication factor

For the double sums involving an (i + j) coefficient, such as,

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left((i+j)C_{T,G_i E G_j} \right) = \left(2C_{T,G_1 E G_1} + 3C_{T,G_1 E G_2} + 4C_{T,G_1 E G_3} + \cdots \right) + \left(3C_{T,G_2 E G_1} + 4C_{T,G_2 E G_2} + 5C_{T,G_2 E G_3} + \cdots \right) + \left(4C_{T,G_3 E G_1} + \cdots \right) + \cdots$$
G.1

Applying Eqs. (F.2) and (4.17) again yields two single sums.

The first sum is addressed in Appendix A. The second sum is also a converging series.

$$\sum_{i=0}^{\infty} (i+1)^2 a^i = \frac{1+a}{(1-a)^3} = 1 + 4a + 9a^2 + \cdots$$
G.3

Applying Eqs. (F.6) and (G.3):

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (i+j) C_{T,G_i EG_j} = K_{C3}^2 C_{T,E} C_{T,G_M}^2 \left[\frac{1}{\left(1 - K_{C2} C_{T,G_M}\right)^2} + \frac{1 + K_{C2} C_{T,G_M}}{\left(1 - K_{C2} C_{T,G_M}\right)^3} \right] \qquad \text{G.4}$$
$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (i+j) C_{T,G_i EG_j} = \frac{2K_{C3}^2 C_{T,E} C_{T,G_M}^2}{\left(1 - K_{C2} C_{T,G_M}\right)^3} \qquad \text{G.5}$$

Chapter 5. Applications of a Wertheim association activity coefficient model to methanol and ethanol-containing mixtures

A new activity coefficient method for associating systems, Wertheim Association Gamma (WAG), is evaluated for methanol and ethanol-containing binary and ternary mixtures. The method combines Wertheim's perturbation theory, Flory's theory and a residual contribution. This residual contribution is represented using several existing thermodynamic models including NRTL, Scatchard-Hildebrand and a variation of Wilson's equation developed by Nagata. The capabilities and limitations of the models are compared to those of PC-SAFT and CPA. The WAG-NRTL and WAG-Nagata approaches capture the behaviors of these mixed systems well when modeled with four parameters. Meanwhile, WAG with Scatchard-Hildebrand (WAG-SH) out-performs CPA and PC-SAFT in several cases, with only two fitted parameters. Lastly, the limitations of existing methods for calculating association parameters are explored and shown to need improvement.

5.1. Introduction

Primary alcohols have many important uses in the chemical industry. Ethanol has been studied extensively as a promising alternative to petroleum-derived chemicals and fuels [99–101]. Methanol is often used as a raw material in the manufacture of important compounds, such as methyl-tert-butyl ether, and to prevent the formation of gas hydrates in the oil and gas industry [102]. In such cases, the thermodynamic behavior of the systems can be difficult to represent with traditional models and more advanced methods must be used. For alcohols, water, organic acids and other polar molecules, it is now well-established that models that include hydrogen bonding effects perform better than those that do not. For this reason, association equations of state, such as SAFT [9] (and its derivatives [54–56,95]), CPA [11] and ESD [10] provide better representation of these types of systems.

Though theoretically more rigorous, equations of state require fitting of several pure component parameters, requiring extensive vapor pressure and density data that are often not available for new compounds. Lack of data results in large uncertainty in parameter values. Further, the dispersion terms and association terms are both attractive and correct attribution of effects to the respective interactions is difficult when fitting limited macroscopic data. In contrast, activity coefficient (γ) models have been favored by industry for many applications (particularly those at low or moderate pressures) because the fitting of pure-component properties is only required in the range to be modeled, and attention can be focused on the mixture nonidealities. Additionally, the representation of the mixture is not biased by challenges in fitting the pure component behavior. Association has been modeled in activity coefficient models as early as 1969. Harris and Prausnitz [103] combined the van Laar equation with Dolezalek's [34] chemical theory to represent physical interactions and hydrogen bonding respectively. The Dolezalek approach describes the formation of a hydrogen bond as an equilibrium reaction guided by an equilibrium constant. More recently, Karachewski [13] combined this association theory with Scatchard-Hildebrand [19,20] and derived the value of the equilibrium constants of hydrogen bonding from NMR spectroscopy. Asprion et al. [7] demonstrated that use of chemical theory with local composition models can suppress false prediction of VLLE. An alternative method for calculating association is Wertheim's perturbation theory [47–50]. This model has become increasingly popular in recent years and is described in more detail in Section 5.2.1. Fu and Sandler [6] were the first to integrate Wertheim's theory for association into UNIQUAC [12], an activity coefficient model, achieving significant improvement for binary systems. The model was also extended to a predictive UNIFAC method [104] with good results, albeit with the association parameters regressed to experimental phase equilibria data without consideration of actual spectroscopically measured association.

Unfortunately, both of Fu's models include a derivative term that make them difficult to generalize for multicomponent systems. In our previous work [96], we developed a generalized γ expression for association which will henceforth be referred to as the Wertheim Association Gamma (WAG) method. This development eliminates the need for the complex derivative in Fu and Sandler's work yielding a general equation that can be applied readily in multicomponent systems. More importantly, the generalized γ association term can be attached to existing activity coefficient models to enhance their ability to model systems with hydrogen bonding. In this work, we explore the capabilities and limitations of three variations, WAG-NRTL, WAG-Scatchard-Hildebrand (WAG-SH) and WAG-Nagata, for various systems of industrial importance and then compare their performance to two association equations of state, CPA and PC-SAFT.

5.2. Model theory and parameters

In this section, we outline the models compared herein. In cases with the WAG models, the vapor phase non-idealities are calculated using the Hayden O'Connell [79] model which has long been recognized for its accurate modeling of the vapor phase. Liquid phase non-idealities are calculated using an activity coefficient model which is a product of three contributions:

$$\gamma = \gamma^{comb} \gamma^{res} \gamma^{assoc}$$
 5.1

or

$$\ln \gamma = \ln \gamma^{comb} + \ln \gamma^{res} + \ln \gamma^{assoc}$$
 5.2

The terms γ^{comb} , γ^{res} and γ^{assoc} are the combinatorial, residual and association contributions to the activity coefficient γ respectively. The combinatorial contribution calculates the entropic effects that arise due to differences in molecular sizes and shapes in solution. In theory, this would be the only term remaining at liquid densities and infinitely high temperature. The residual contribution accounts for molecular interactions between unlike molecules. In practice, depending on parameters used, this term can remain at liquid densities and infinitely high temperature. The association contribution calculates the effects of hydrogen bonding.

5.2.1. Association contribution

Association is often modeled using Wertheim's perturbation theory [47–50], a statistical mechanics approach that quantifies that probability of bonds forming between sites on molecules. Acceptor and donor sites are allocated on electronegative atoms that accept hydrogen bonds and the protons that can be donated respectively. In our previous work [96], we introduced a generalized activity coefficient contribution for association using Wertheim's perturbation theory. The model assumes conventional mixing rules, no excess volume, and a universal packing fraction. Further details regarding the development of the model are given in Appendix C of Chapter 3 and are omitted here. The resulting functional form is given by:

$$\ln \gamma_k^{assoc} = \sum_{\substack{sites \ A_i \\ on \ k}} N_{A_i} \left[\ln \left(\frac{X^{A_i}}{X^{A_i,0}} \right) - \frac{1}{2} (1 - X^{A_i,0}) \right] + \frac{1}{2} \frac{\rho}{\rho_k} \sum_{\substack{all \\ sites \ B_i}} x_{B_i \ host} N_{B_i} (1 - X^{B_i})$$
5.3

where X^{A_i} and $X^{A_i,0}$ represent the fraction of sites free in the mixture and when the host for the site is pure respectively. N_{A_i} is the number of identical A_i sites on a molecule of host component. X^{A_i} can be calculated using:

$$X^{A_{i}} = \left(1 + \sum_{D_{j}} x_{D_{j} host} N_{D_{j}} X^{D_{j}} \rho \Delta^{A_{i} D_{j}}\right)^{-1}$$
 5.4

where D_j represents any site that A_i can bond to and $\Delta^{A_i D_j}$ is the association constant for the two sites. The variable $x_{D_j host}$ is the apparent mole fraction of the component that hosts the site. Donor sites are solved by using a similar equation obtained by flipping the acceptor and donor identities in the equation.

5.2.2. Combinatorial contribution

The combinatorial contribution is commonly calculated using the well-known modified Flory's equation:

$$\ln \gamma_i^{Flory} = \ln \left(\frac{\Phi_i'}{x_i}\right) + 1 - \frac{\Phi_i'}{x_i}$$
 5.5

where Φ_i' is the volume fraction or modified volume fraction for component *i* calculated by assuming zero excess volume:

$$\Phi_i' = \frac{x_i V_i^p}{\sum_j x_j V_j^p}$$
 5.6

In Flory's original theory (and unless otherwise specified in this work), p = 1. Despite its extensive application, Flory's theory does have known shortcomings such as assuming a universal packing factor. This is reasonable for smaller molecules of similar size where the entropy effects are small. However, the assumption of constant packing factor is inaccurate for polymers in solvents where entropy of mixing is significant [105]. Various attempts to correct such weaknesses in the theory commonly use the Staverman-Guggenheim (SG) theory [38,106] which has been integrated into well-known models including UNIQUAC [12] and UNIFAC [29]. Recent work by Vahid *et al.* [105] compares the excess entropy term calculated using several methods with results from discontinuous molecular dynamics simulations. They found that Flory, both alone or corrected with SG, underestimated the excess entropy for certain values of packing factor and misrepresented its skewness with respect to equimolar composition.

Other semi-empirical variations of Flory's theory exist which replace the molar volumes in Eq. (5.6) with the UNIQUAC parameter r_i and/or using fractions for the power, p. Donohue and Prausnitz [107] were the first to suggest using r_i with a non-unity power in the Flory-Huggins model, which combines Flory's theory with Scatchard-Hildebrand. Kikic *et al.* [108] compared the excess Gibbs energy calculation from Donohue and Prausnitz model with others that corrected

Flory-Huggins with a modified SG term. They tested both $r_l^{2/3}$ and $V^{2/3}$ for the SG term and showed that both modifications far out-performed traditional SG for aliphatic hydrocarbons. Following these studies, two groups published modified UNIFAC models that incorporate a nonunity power in the Flory portion of the model. The first, which was developed at the University of Dortmund and is therefore often referred to as the Dortmund modified UNIFAC [109,110], sets p = 3/4 and retains the SG correction. The second, developed at the University of Denmark (Lyngby), is sometimes referred to as the Lyngby modified UNIFAC model [111,112] and follows Kikic *et al.*'s suggestion of using a power of 2/3. The Lyngby modified UNIFAC model also removed the SG correction completely, citing work by Sayegh and Vera [113] showing that the correction often has little effect or sometimes overestimates the combinatorial contribution leading to unrealistic excess entropies.

In our own work, we performed preliminary comparisons between:

- 1. Traditional Flory's theory (Eq. (5.6) with p=1)
- 2. Traditional Flory's theory (Eq. (5.6) with p=1) + SG
- 3. Modified Flory's theory (Eq. (5.6) with p = 2/3)

4. Traditional Flory's theory (Eq. (5.6) with p=1) + van der Waals entropy correction [8,67] A survey of several systems showed no appreciable differences in the fitting of binary systems studied in this work. In all cases, the residual term compensated for differences in the combinatorial term. Accepting that the original Flory and the Staverman-Guggenheim have some deficiencies, we follow Kikic *et al.* and Sayegh and Vera's recommendations and model the combinatorial contribution using Eq. (5.6) with p=2/3 and no additional correction. The decision to use molar volumes rather than the UNIQUAC r_i parameter was made based on the note by Kikic *et al.* that the former case showed a slight advantage over the latter in modeling linear alkanes. The use of volumes provides some temperature dependence in the Flory term.

5.2.3. Residual contribution

Three models for the residual contribution are tested in this work: non-random two liquid (NRTL) theory, a model developed by Nagata, and Scatchard-Hildebrand yielding the WAG-NRTL, WAG-Nagata and WAG-SH models respectively. In this section, we provide a brief overview of each of these $\ln \gamma_k^{res}$ models.

Activity coefficient models for liquid phase modeling may assume random mixing or consider variations in local compositions. In contrast to random mixing models, which assume that species are randomly dispersed in solution, local composition models represent the compositional differences between the solvent shell of a molecule and the bulk fluid. This variation is modeled by defining a constant, Ω_{ij} :

$$\Omega_{ij} = \frac{x_{ij}/x_{jj}}{x_i/x_j}$$
 5.7

where x_{ij} is the mole fraction of molecules of component *i* around a molecule of component *j*. Ω_{ij} represents a ratio of the local composition to the bulk composition of the fluid and is calculated differently in every model.

The NRTL equation, which was developed by Renon and Prausnitz [3] in 1968, calculates Ω_{ij} with adjustable parameters, α_{ij} and τ_{ij} :

$$\Omega_{ij} = \exp(-\alpha_{ij}\tau_{ij})$$
 5.8

Here, τ_{ij} is the interaction energy parameter between a molecule of component *i* and one of component j and $\alpha_{ij} = \alpha_{ji}$ is called the non-randomness parameter. The value of $\ln \gamma^{res}$ is then calculated through:

$$\ln \gamma_i^{NRTL,res} = \frac{\sum_j x_j \tau_{ji} \exp(-\alpha_{ji} \tau_{ji})}{\sum_k x_k \exp(-\alpha_{ki} \tau_{ki})} + \sum_j \frac{x_i \exp(-\alpha_{ij} \tau_{ij})}{\sum_k x_k \exp(-\alpha_{kj} \tau_{kj})} \left(\tau_{ij} - \frac{\sum_m x_m \tau_{mj} \exp(-\alpha_{mj} \tau_{mj})}{\sum_k x_k \exp(-\alpha_{kj} \tau_{kj})}\right)$$
5.9

Often, τ_{ij} is fitted by two parameters to increase the temperature dependence of the term:

$$\tau_{ij} = \tilde{a}_{ij} + \frac{\tilde{b}_{ij}}{T}$$
 5.10

NRTL is a commonly used model in the chemical industry because its functional form makes it remarkably flexible. However, its parameters lack the theoretical significance necessary to derive physical meaning or predict their values without regression. In this work, NRTL models (both with and without WAG) are called 2-parameter models if \tilde{a}_{ij} is set to 0 and only \tilde{b}_{ij} is fitted. The "2" indicates that there will be two \tilde{b}_{ij} values used to fit each binary system: \tilde{b}_{12} and \tilde{b}_{21} . The 4parameter model fits both \tilde{a}_{ij} and \tilde{b}_{ij} . The non-randomness parameter, α_{ij} , is not fitted to each individual system. Instead, $\alpha_{ij} = \alpha_{ji}$ was set to 0.2 when the system has a liquid-liquid immiscibility and 0.3 otherwise.

Another local composition model, Wilson [18], is the basis of what is referred to herein as the Nagata model. Wilson used the variable Λ_{ji} to represent local composition. Elliott and Lira [8,67] show that, with some rearrangement, Wilson's parameter Λ_{ji} is related to Ω_{ij} in Eq. (5.7) through:

$$\Omega_{ij} = \Lambda_{ji} = \frac{V_i}{V_j} \exp\left(-\frac{A_{ji}^{Wilson}}{RT}\right) = \frac{V_i}{V_j} \tau_{ij}$$
5.11

where A_{ji}^{Wilson} is a fitted parameter. With careful evaluation of the terms in Wilson's original equation, it is possible to separate Flory's theory [8,67]. For convenience, one can consider Flory to be the combinatorial contribution to the excess Gibbs energy, $G^{E,Wilson}$, and attribute the other terms to a residual contribution, $G^{E,Wilson res}$.

$$G^{E,Wilson} = G^{E,Flory} + G^{E,Wilson\,res}$$
5.12

$$G^{E,Flory} = RT \sum_{i} x_i \ln\left(\frac{\Phi_i}{x_i}\right)$$
5.13

$$G^{E,Wilson\,res} = -RT \sum_{i} x_{i} \ln\left(\sum_{j} \left(\Phi_{j} \exp\left(-\frac{A_{ij}^{Wilson}}{RT}\right)\right)\right)$$
 5.14

In their 1997 work, Nagata *et al.* [15] used this idea to combine Eq. (5.14), which they refer to as the physical contribution, with a chemical term based on Flory's theory. In the chemical term, the molar volume in Eq. (5.6) is replaced by a UNIQUAC volume parameter raised to 2/3, and the chemical contribution is calculated in terms of the true monomer activities. This approach is very similar to that developed by Campbell [16] and we show in our previous work that the resulting "true" Flory activity coefficient, γ^{chem} , is related to the traditional Flory activity coefficient, γ^{Flory} , through:

$$\ln \gamma^{chem} = \ln \gamma^{Flory} + \ln \gamma^{assoc} \qquad 5.15$$

Therefore:

$$\ln \gamma^{Nagata} = \ln \gamma^{Wilson, res} + \ln \gamma^{Flory} + \ln \gamma^{assoc}$$
 5.16

Nagata's model was shown to provide excellent fits of excess enthalpy, binary and ternary equilibria data. For this reason, we chose to test the capabilities of the "residual term" in Wilson and Nagata's work given by:

$$\ln \gamma_i^{Nagata, res} = -\ln\left(\sum_j \Phi_j \tau_{ji}\right) + \frac{\Phi_i}{x_i} - \sum_k \frac{x_k}{x_i} \frac{\Phi_i \tau_{ik}}{\sum_j \Phi_j \tau_{jk}}$$
 5.17

where $\Phi_i = x_i V_i / \sum_j x_j V_j$ is the volume fraction for component *i*. The quantity τ_{ij} is calculated through Eq. (5.10) and the 2- and 4-parameter implementations of this model are fitted in the same way as described above for NRTL. To avoid confusion, we refer to this as WAG-Nagata, rather than Wilson, to make it clear that only the residual portion of Wilson is used and the association and combinatorial contributions are calculated as described in Sections 5.2.1 and 5.2.2.

The third residual model evaluated here is Scatchard-Hildebrand (SH). Developed in the early 1930s [19,20], SH is a regular solution theory which extends the van der Waals equation of state to mixtures. The excess Gibbs energy and $\ln \gamma^{res}$ have the form:

$$G^{E,res} \approx U^{E,res} = V \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Phi_i \Phi_j A_{ij}^{SH}$$
 5.18

$$\ln \gamma_i^{res} = \frac{V_i}{RT} \left(\sum_{k=1}^n \varphi_k A_{ki}^{SH} - \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^n \Phi_k \Phi_j A_{kj}^{SH} \right)$$
 5.19

where *V* is the mixture molar volume calculated by assuming no excess volume, $V = \sum_{i} x_i V_i$, Φ_i is the volume fraction of component *i* and A_{ij}^{SH} is a combined term that can be related to the solubility parameter of each component δ_i through:

$$A_{ij}^{SH} = (\delta_i - \delta_j)^2 + 2\delta_i \delta_j k_{ij}$$
 5.20

The solubility parameter, δ_i , is a measure of the internal energy departure per molar volume of a component and is therefore often referred to as the cohesive energy density. k_{ij} of Eq. (5.20) is a binary interaction parameter that can be fitted to experimental mixture data. Because adjusting k_{ij} is equivalent to fitting A_{ij} , we regress A_{ij} as:

$$A_{ij}^{SH} = \tilde{a}_{ij}T + \tilde{b}_{ij}$$
 5.21

where \tilde{a}_{ij} and \tilde{b}_{ij} are the fitted parameters.

In summary, $\ln \gamma$ will be calculated using the following 3 WAG models:

$$\ln \gamma = \ln \gamma^{NRTL, res} + \ln \gamma^{Mod \ Flory} + \ln \gamma^{assoc}$$
 5.22

$$\ln \gamma = \ln \gamma^{SH,res} + \ln \gamma^{Mod \ Flory} + \ln \gamma^{assoc}$$
 5.23

$$\ln \gamma = \ln \gamma^{Nagata, res} + \ln \gamma^{Mod \ Flory} + \ln \gamma^{assoc}$$
 5.24

5.2.4. EOS model parameters

The WAG models developed in Eqs. (5.22-5.24) are compared to CPA and PC-SAFT. These EOS models are used as a reference and thus, their derivations are not provided here but can be found in references [11] and [95] for CPA and PC-SAFT respectively. We use the implementations in version 9 of AspenPlus® and we fit the binary interaction parameter, k_{ij} with the forms implemented in AspenPlus®:

$$k_{ij} = \tilde{a}_{ij} + \tilde{b}_{ij} \frac{T}{298.15 K} (CPA)$$

$$5.25$$

$$k_{ij} = \tilde{a}_{ij} + \tilde{b}_{ij} \frac{298.15 K}{T} (PC - SAFT)$$

5.2.5. Association parameters

Often, $\Delta^{A_i D_j}$ of Eqn. 5.4 is calculated by fitting to pure component properties. In this work, the WAG models use the CPA association parameter, given by:

$$\Delta^{A_i D_j} = g(\sigma) \beta_{CPA}^{A_i D_j} b_{CPA, ij} \left(\exp\left(\frac{\epsilon^{A_i D_j}}{kT}\right) - 1 \right)$$
 5.26

where $b_{CPA,ij}$ is the CPA covolume term, $\epsilon^{A_i D_j}$ is an interaction energy parameter, k is Boltzmann's constant, and $\beta_{CPA}^{A_i D_j}$ is a fitted parameter. The term $g(\sigma)$ is the radial distribution function at contact given by:

$$g(\sigma) = \frac{1}{1 - \frac{1.9}{4} b_{CPA} \rho}$$
 5.27

The values of CPA parameters used in this work are given in Table 5.1. The values of ϵ^{AD} and β^{AD} in the table refer to the self-association interactions, which occur between acceptors and donors on molecules of the same component. Using Huang and Radosz's [55] notation, alcohols are modeled with a 2B scheme (one acceptor site, one donor site) and water with 4C (two identical acceptor sites, two identical donor sites). For ethanol, one set of CPA parameters are used for all the WAG models and another is used for the CPA fits they are compared to. The reasoning for this choice is given in Section 5.4.

Table 5.1: CPA parameters used to calculate Δ^{AD} . The terms ϵ^{AD} and β^{AD} pertain to self-association interactions

Compound	T _{Crit} (K)	<i>b_{CPA}</i> (L/mol)	a _{0,CPA} (bar L ² mol ⁻²)	C _{1,CPA}	ϵ^{AD} (bar L mol ⁻¹)	β^{AD}	Reference
Methanol	512.6	0.03098	4.053	0.4310	245.91	0.0161	[1]
Ethanol (CPA)	513.9	0.04910	8.672	0.7369	215.32	0.0080	[1]
Ethanol (Other)	513.9	0.04690	6.701	0.7987	247.72	0.0127	[2]
Water	647.3	0.01452	1.228	0.6736	166.55	0.0692	[114]
<i>n</i> -Pentane	469.7	0.09101	18.20	0.7986			[114]
<i>n</i> -Heptane	540.2	0.1254	29.18	0.9137			[114]
<i>n</i> -Decane	617.7	0.1787	47.39	1.1324			[114]
Cyclohexane	553.6	0.09038	21.26	0.7427			[114]

Cross-association, which occurs between molecules of different components, is approximated by combining self-association parameters according to simplifying rules called combining rules. Elliott's combining rule (ECR) states that:

$$\Delta^{A_i D_j} = \sqrt{\Delta^{A_i D_i} \Delta^{A_j D_j}}$$
 5.28

For WAG models, we implement a modification. The ECR includes the radial distribution function of the pure components *i* and *j* in $\Delta^{A_i D_i}$ and calculates their geometric mean values of the association parameters. However, this eliminates any compositional dependence of $g(\sigma)$ for the mixture. Therefore, we modify the rule as follows:

$$\Delta^{A_i D_j} = g(\sigma) \sqrt{\frac{\Delta^{A_i D_i}}{g(\sigma_i)} \frac{\Delta^{A_j D_j}}{g(\sigma_j)}}$$
5.29

Where the pure component $g(\sigma_i)$ values are determined at the pure component densities at the same temperature. Assuming that $\exp\left(\frac{\epsilon^{AD}}{kT}\right) - 1 \cong \exp\left(\frac{\epsilon^{AD}}{kT}\right)$, Eq. (5.29) can be applied by setting:

$$\beta_{CPA}^{A_i D_j} b_{CPA,ij} = \sqrt{\beta_{CPA}^{A_i D_i} \beta_{CPA}^{A_j D_j} b_{CPA,i} b_{CPA,j}}$$
 5.30

$$\epsilon^{A_i D_j} = \frac{\epsilon^{A_i D_i} + \epsilon^{A_j D_j}}{2}$$
 5.31

where $\beta_{CPA}^{A_i D_j} b_{CPA,ij}$ is a constant in this work. For the CPA model, the CR-1 combining rules are applied for cross-association such that:

$$\beta_{CPA}^{A_i D_j} = \sqrt{\beta_{CPA}^{A_i D_i} \beta_{CPA}^{A_j D_j}}$$
 5.32

$$\epsilon^{A_i D_j} = \frac{\epsilon^{A_i D_i} + \epsilon^{A_j D_j}}{2}$$
 5.33

For the PC-SAFT model, parameters are obtained from literature for associating [115] and inert [95] components. The association parameter is given by:

$$\Delta^{A_i D_j} = g_{ij}^{hs} \kappa^{A_i D_j} \sigma_{ij}^3 \left(\exp\left(\frac{\epsilon^{A_i D_j}}{kT}\right) - 1 \right)$$
 5.34

where g_{ij}^{hs} is the radial distribution function for hard spheres calculated with the Carnahan-Starling [116] equation and: where $\kappa^{A_i D_j}$ has units of volume and σ_{ii} is the temperature-independent

$$\kappa^{A_i D_j} = \sqrt{\kappa^{A_i D_i} \kappa^{A_j D_j}} \left(\frac{\sqrt{\sigma_{ii} \sigma_{jj}}}{0.5(\sigma_{ii} + \sigma_{jj})} \right)^3$$
 5.35

segment diameter for component *i*. Eqs. (5.33) and (5.35) are the combining rules used to model cross-association in PC-SAFT.

5.3. Model comparisons

To isolate the capabilities of the residual model for all WAG models, the Wertheim association parameter was calculated using the CPA parameters from Table 5.1 consistently for all models except PC-SAFT. Only the residual term parameters were adjusted. Liquid molar volumes were calculated using DIPPR molar volumes as provided by the AspenPlus® databases given in Table 5.2. PC-SAFT parameters are given in Table 5.3.

Table 5.2: AspenPlus® databases used in calculations. WAG-* denotes all WAG models

Model	Databases in AspenPlus®
WAG-*	APV90 PURE35, NISTV90 NIST-TRC
CPA	APEOSV90 AP-EOS, APV90 PURE35, NISTV90 NIST-TRC
PC-SAFT	APV90 PURE35, APV90 PC-SAFT

Table 5.3: PC-SAFT pure component parameters. The terms ϵ^{AD} and κ^{AD} pertain to self-association interactions

Compound	m	σ _{ii} (Å)	<i>\epsilon (</i> K)	ϵ^{AD}/k (K)	κ^{AD}	Reference
Methanol	1.526	3.230	188.9	2900	0.03518	[115]
Ethanol	2.383	3.177	198.2	2653	0.03238	[115]
Water	1.066	3.001	366.5	2501	0.03487	[115]
<i>n</i> -Pentane	2.690	3.773	231.2			[95]
<i>n</i> -Heptane	3.483	3.805	238.4			[95]
<i>n</i> -Decane	4.663	3.838	243.9			[95]
Cyclohexane	2.530	3.850	278.1			[95]

Hayden-O'Connell parameters were used as provided by AspenPlus® ver. 9 (given in Appendix H) without adjustment. Excess volume was assumed to be zero, and all mixture molar volumes used for volume fractions were temperature-dependent. To facilitate regressions, the WAG and residual models were programmed as a user activity coefficient subroutine for AspenPlus®.

Data sources used for regression are summarized in Table 5.4 and the values of the regressed parameters are given in the Appendix. When multiple data sets are available from the source, all sets are used in the regressions.
System	Data fitted to
Methanol $+ n$ -heptane	VLE [117], LLE [118]
Methanol + cyclohexane	VLE [117], LLE [5], H ^E [119]
Methanol $+ n$ -pentane	VLE [120]
Methanol + water	VLE [121,122]
Methanol + ethanol	VLE [123–125], H ^E [126]
<i>n</i> -Heptane + cyclohexane	VLE [127] (point at $x_{Hep} = 0.984$ omitted), H ^E [128]
Ethanol + n -decane	VLE [129], H ^E [130]
Ethanol + cyclohexane	VLE [131], H ^E [132]

Table 5.4: Data used for parameter regression.

The goodness of fit for each regression is given in this work by the sum of squares value, SSQ, calculated by:

$$SSQ = \sum_{i=1}^{k} \sum_{j=1}^{m} \left(\frac{L_{ij} - L_{ij}^{exp}}{stddev_{ij}} \right)^{2}$$
 5.36

Here, *L* is generic notation for all properties used in the regression such as mole fractions or heats of mixing. The superscript '*exp*' refers to the experimental value and $stddev_{ij}$ is a standard deviation assumed for the experimental data; $stddev_{ij}$ is 0.1 K for temperature values in the VLE data sets and 0.01 K for those in LLE data sets, 0.1% for excess enthalpy data, 1% for vapor mole fractions, and 0.1% for pressure and liquid mole fraction data. These weights reflect the confidence with which each property can be obtained experimentally.

Table 5.5 gives the total SSQ of each model for all systems tested here with the exception of two. The methanol + pentane system was omitted because 4-parameter WAG models and NRTL were deemed unnecessary due to the satisfactory fits of the 2-parameter versions. The methanol + cyclohexane system was also omitted because the 2-parameter WAG-NRTL model did not converge. From these results, the four parameter versions of WAG-NRTL and WAG-Nagata provide the best overall fits. This is not surprising when compared to CPA and PC-SAFT since the equations of state were fitted with only two parameters each. However, it is encouraging that WAG-SH and the 2-parameter variation of WAG-Nagata, which are also fitted with just two parameters, perform better than the CPA and PC-SAFT for the systems tested.

Table 5.5: Total SSQ for systems studied in this work. Systems for which only some model
were fitted have been omitted from this table and (2) and (4) denote the number of fitted
parameters

Model	Number of fitted parameters	Total SSQ × 10 ⁻⁵		
WAG-NRTL(4)	4	1.12		
WAG-Nagata(4)	4	1.82		
NRTL(4)	4	2.13		
WAG-SH	2	8.14		
WAG-Nagata(2)	2	9.30		
CPA	2	11.9		
NRTL(2)	2	19.7		
WAG-NRTL(2)	2	27.2		
PC-SAFT	2	42.5		

While the total SSQ provides a reasonable overview for the fit, it is important to consider each regression individually to understand the limitations of each model. For example, the high SSQ value for PC-SAFT misrepresents the majority of fits. This SSQ is large due to the moderately poor representation of the methanol + n-heptane liquid-liquid equilibrium. To explore fits more carefully, Table 5.6 arranges the models according to SSQ value for each system. Each of the models will be represented by the line and/or color in Table 5.7. In some figures, certain lines may be obscured due to overlap with others.

Methanol + <i>n</i> -heptane	WAG- NRTL(4) 8.83	NRTL(4) 13.0	WAG- Nagata(4) 16.0	WAG-SH 75.2	WAG- Nagata(2) 87.0	CPA 103	NRTL(2) 183	WAG- NRTL(2) 265	PC- SAFT 419
Methanol + cyclohexane	NRTL(4) 17.6	CPA 20.6	WAG- Nagata(4) 23.8	WAG- NRTL(4) 24.5	WAG-SH 41.8	PC- SAFT 80.3	WAG- Nagata(2) 107	NRTL(2) 285	
Methanol + <i>n</i> -pentane	WAG-SH 0.170	WAG- Nagata(2) 0.215	WAG- NRTL(2) 0.241	PC- SAFT 0.279	CPA 0.303	NRTL(2) 0.597			
Methanol + water	WAG- Nagata(4) 0.133	NRTL(4) 0.169	WAG-SH 0.199	NRTL(2) 0.218	WAG- Nagata(2) 0.244	CPA 0.325	PC- SAFT 0.525	WAG- NRTL(4) 0.564	WAG- NRTL(2) 0.790
Methanol + ethanol	WAG- NRTL(4) 0.00872	NRTL(4) 0.00882	WAG- Nagata(4) 0.0108	WAG- NRTL(2) 0.0423	NRTL(2) 0.0431	WAG- Nagata(2) 0.0716	PC- SAFT 0.0993	CPA 0.618	WAG- SH 0.664
<i>n-</i> Heptane + cyclohexane	WAG- Nagata(2)* 0.0152	NRTL(2)* 0.0152	WAG- NRTL(2)* 0.0153	WAG- Nagata(4) 0.0172	NRTL(4) 0.0191	WAG- NRTL(4) 0.0193	WAG-SH 0.0881	PC- SAFT 0.0893	CPA 0.101
Ethanol + <i>n-</i> decane	WAG- NRTL(4) 1.51	WAG- Nagata(4) 1.68	PC-SAFT 4.35	WAG-SH 4.89	WAG- Nagata(2) 5.23	WAG- NRTL(2) 5.64	NRTL(4) 7.50	NRTL(2) 12.0	CPA 13.0
Ethanol + cyclohexane	WAG- NRTL(4) 0.305	WAG- Nagata(4) 0.318	WAG-SH 0.392	WAG- Nagata(2) 0.401	PC- SAFT 0.443	NRTL(4) 0.560	WAG- NRTL(2) 0.645	NRTL(2) 1.43	CPA 1.98

Table 5.6: Model performance in order of best fit. Numbers are SSQ $\times 10^{-4}$ values and (2) and (4) indicate the number of fittedparameters in the model. (*) denotes that only VLE (not excess enthalpy) data was fitted

Model	Notation in Figures				
WAG-NRTL (2 parameters)	• • • or •				
WAG-NRTL (4 parameters)	or •				
WAG-Nagata (2 parameters)	• • • or •				
WAG-Nagata (4 parameters)	or •				
WAG-SH	or •				
NRTL (2 parameters)	— — — or •				
NRTL (4 parameters)	or •				
PC-SAFT	— — — or o				
CPA	or •				

Table 5.7: Legend for plots in Model Capabilities section

In the upcoming sections, we show phase diagrams, heat of mixing plots (when applicable), activity coefficients and percent errors for the K-ratios calculated by the various models. As an additional measure of the model fits, the percent error of the K-ratios is calculated for each binary as:

$$\% \ error = \frac{KVL^{exp} - KVL}{KVL^{exp}} \times 100\%$$
 5.37

where KVL is the K-ratio in the vapor-liquid equilibrium (VLE) region and the superscript of *exp* indicates experimental values. For the liquid-liquid equilibrium (LLE) region, KVL is replaced by the K-ratio in the LLE region.

5.3.1. Binary system with no association

First, we test the capabilities of the residual and combinatorial activity coefficients with a simple binary mixture of two inert components, *n*-heptane and cyclohexane. The results are given in Figures 5.1-5.4. The 2-parameter WAG and NRTL models were unable to simultaneously fit the VLE and heat of mixing data well. Because the heat of mixing is small and inconsequential for nearly ideal systems such as this, the parameters were fitted only to the VLE data. Therefore, the lines for the 2-parameter WAG and NRTL models are shown in the VLE diagram but omitted from the heat of mixing diagram. All the models were able to describe VLE behavior equally well.



Figure 5.1: Phase equilibria diagram for *n*-heptane + cyclohexane at 101.3 kPa.



Figure 5.2: Heat of mixing for *n***-heptane** + **cyclohexane at T=298.14 K and 101.3 kPa**. The WAG-NRTL(4) line is obscured by the WAG-Nagata(4) and NRTL(4). The WAG-SH, CPA and PC-SAFT lines are overlapped.



Figure 5.3: Calculated and experimental activity coefficients for *n*-heptane + cyclohexane at 101.3 kPa. Data and fits correspond to the experimental data in Figure 5.1



Figure 5.4: Percent errors of K-ratios for *n***-heptane** + **cyclohexane system**. Figures (a) is the K-ratio of *n*-heptane and (b) is the K-ratio of cyclohexane in the VLE region at 101.3 kPa

5.3.2. Binary systems with self-association

Figures 5.5-5.8 show fits for the methanol + *n*-heptane system. The fits for heat of mixing exhibit a linear segment that spans the predicted LLE for each model. NRTL and WAG-NRTL(2) are unable to capture both phase equilibria simultaneously. NRTL's poor representation, both for the fitted phase equilibria and the predicted heat of mixing in Figure 5.6, can be explained by its lack of association. Including the Wertheim association term in WAG-NRTL improves the fit significantly. However, WAG-NRTL(2) does not have the temperature dependence necessary to fit the LLE region accurately and two more parameters must be included. PC-SAFT also shows a poor fit in the liquid-liquid region. The other models are able to represent the phase equilibria with comparable accuracy.



Figure 5.5: Phase equilibria diagram for methanol + *n*-heptane at 101.3 kPa. Experimental VLE and LLE data is shown by markers.



Figure 5.6: Heat of mixing for methanol + *n*-heptane at T=298.15 K and 101.3 kPa. Experimental data [133] is shown for the miscible region ($x_{MeOH} < 0.1464$ and $x_{MeOH} > 0.8899$).



Figure 5.7: Calculated and experimental [4] activity coefficients for methanol + *n*-heptane at 101.3 kPa. Data and fits correspond to the experimental VLE data in Figure 5.5



Figure 5.8: Percent errors of K-ratios for methanol + *n*-heptane system. Figures (a) and (c) are K-ratios in the VLE region and (b) and (d) are K-ratios in the LLE region

The fitting results of the methanol + cyclohexane system are very similar to those of methanol + *n*-heptane, which are given in Figures 5.9-5.12. NRTL and the two-parameter WAG models provide poorer fits, particularly in the LLE. The excess enthalpy data in Figure 5.10 is fitted well by the four parameter WAG models. While capable of capturing the magnitude well, CPA and WAG-SH are slightly inferior in capturing the minor asymmetry in the heat of mixing.



Figure 5.9: Phase equilibria diagram for methanol + cyclohexane at 101.3 kPa. Experimental VLE and LLE data is shown by markers.



Figure 5.10: Heat of mixing for methanol + cyclohexane at T=416.29 K and 1900 kPa.



Figure 5.11: Calculated and experimental activity coefficients for methanol + cyclohexane at 101.3 kPa. Data and fits correspond to the experimental VLE data in Figure 5.9



Figure 5.12: Percent errors of K-ratios for methanol + cyclohexane system. Figures (a) and (c) are K-ratios in the VLE region and (b) and (d) are K-ratios in the LLE region

The methanol + n-pentane system is described in Figures 5.13-5.15. The parameters were fitted to data at two temperatures, 372.7 and 397.7 K, and the results at both conditions were very similar. The higher temperature fits are shown here. It is clear that both CPA and PC-SAFT have difficulty describing the dew line for methanol + n-pentane. In contrast, WAG-Nagata and WAG-SH capture both phase lines very well with only two fitted parameters each.



Figure 5.13: Phase equilibria diagram for methanol + *n*-pentane at T=397.7 K



Figure 5.14: Calculated and experimental [120] **activity coefficients for methanol** + *n*-**pentane at T=397.7 K**. Data and fits correspond to the experimental VLE data in Figure 5.13.



Figure 5.15: Percent errors of K-ratios for methanol + *n*-pentane system. Figure (a) is the K-ratio of methanol and (b) is the K-ratio of *n*-pentane in the VLE region at T=397.7 K

Ethanol + cyclohexane fits are provided in Figures 5.16-5.19. This system has no liquid-liquid immiscibility. Three models perform particularly poorly for this system: NRTL(2), NRTL(4) and CPA. The first two models falsely predict an LLE and the last miscalculates the azeotropic composition. WAG-SH and the four parameter WAG models describe the VLE and heat of mixing data much better for this system.



Figure 5.16: Phase equilibria diagram for ethanol + cyclohexane at 101.3 kPa.



Figure 5.17: Heat of mixing for ethanol + cyclohexane at T=298.14 K and 101.3 kPa.



Figure 5.18: Calculated and experimental activity coefficients for ethanol + cyclohexane at 101.3 kPa. Data and fits correspond to the experimental VLE data in Figure 5.16.



Figure 5.19: Percent errors of K-ratios for ethanol + cyclohexane system. Figures (a) is the K-ratio of ethanol and (b) is the K-ratio of cyclohexane in the VLE region at 101.3 kPa

Figures 5.20-5.22 show the model fits for ethanol + n-decane. The parameters were fitted to the data in the figures and two more P-xy VLE data sets in Narasigadu *et al.* [129]. The P-xy fits at the other temperatures are similar and, therefore, omitted for brevity. Linear segments in the bubble line or the heat of mixing indicate prediction of VLLE.



Figure 5.20: Phase equilibria diagram for ethanol + *n*-decane at T=338.17 K



Figure 5.21: Phase equilibria diagram for ethanol + *n*-decane at 101.3 kPa



Figure 5.22: Heat of mixing for ethanol + *n*-decane at T=298.15 K and 101.3 kPa.

The binary ethanol + *n*-decane system is especially challenging for all models tested. NRTL (4) incorrectly predicts VLLE for the isothermal fits between 328 K and 348 K. WAG-NRTL (4) begins to incorrectly predict VLLE at a temperature slightly above 328 K and WAG-Nagata (4) does so at a temperature between 328 K and 298.15 K. WAG-SH and PC-SAFT do not incorrectly predict VLLE at 328 K but they cannot model the steep rise in the bubble line for the P-xy diagrams. The heat of mixing data shown in Figure 5.22 does not indicate phase splitting at 298.15 K, which would appear in the form of a linear section. Therefore, the VLLE predicted by almost all models tested here at temperatures above 298.15 K is incorrect and illustrates a shortcoming of the models. The CPA EOS generates a wider P-xy envelope than the other models due to the difference in the association parameters used for CPA. The Δ^{AD} value used in CPA is approximately 1/5th that of the value used in the WAG models and therefore, the hydrogen bonding is modeled to be weaker. The regressed residual portion must then compensate for the association

term in order to match the experimental data and the resulting residual contribution is overestimated. This demonstrates the importance of correctly quantifying the association effects on non-ideality and the difficulties that arise when the balance between γ^{assoc} and γ^{res} is incorrect.

Exploring the models' predictions at lower temperatures reveal another challenge with WAG-NRTL and WAG-Nagata. For ethanol + *n*-decane, the experimental data used in the parameter regression was between 445 and 328 K. However, when the phase behavior was extrapolated 60 K lower, both models incorrectly predict a lower critical solution temperature (LCST) near 260 K. Interestingly, although CPA, PC-SAFT and WAG-SH provided inferior fits of the experimental data, they did not predict an LCST when extrapolated to low temperatures. We attribute this advantage to the temperature dependence of the residual contribution in the EOS models. Thus, future work will focus on studying the temperature dependence of the residual contribution for the local composition models.

5.3.3. Binary systems with self- and cross-associating systems

Two systems with cross-association are modeled in this work. Figures 5.23-5.26 and 5.27-5.30 show the fits for methanol + water and methanol + ethanol respectively. For the first system, all the models describe the data well with no difficulty. However, the methanol + ethanol binary is more complicated because it is so close to an ideal solution. Almost all models can capture the phase behavior and heat of mixing for the system well except for CPA and WAG-SH. CPA is especially poor at modeling the system and Section 5.4 explores this limitation further.



Figure 5.23: Phase equilibria diagram for methanol + water at 101.3 kPa



Figure 5.24: Phase equilibria diagram for methanol + water at T=298.14 K (bottom) and T=308.14 K (top)



Figure 5.25: Calculated and experimental [121,122] activity coefficients for methanol + water at T=298.14 K, T=308.14 K and P=101.3 kPa from top to bottom. Data and fits correspond to the experimental VLE data in Figure 5.23 and Figure 5.24.



Figure 5.26: Percent errors of K-ratios for methanol + **water system**. Figures (a-c) are Kratios of methanol and (d-f) are K-ratios of water in the VLE regions at T=298.14 K, T=308.14 K and P=101.3 kPa from top to bottom.



Figure 5.27: Phase equilibria diagram for methanol + ethanol at T=298.14 K (bottom) and T=413.13 K (top)



Figure 5.28: Phase equilibria diagram for methanol + ethanol at T=298.14 K



Figure 5.29: Calculated and experimental activity coefficients for methanol + ethanol at T=298.14 K (bottom) and T=413.13 K (top). Data and fits correspond to the experimental VLE data in Figure 5.27.



Figure 5.30: Percent errors of K-ratios for methanol + ethanol system. Figures (a) and (c) are K-ratios of methanol and (b) and (d) are K-ratios of ethanol in the VLE regions at T=298.14 K (top) and T=413.13 K (bottom)

5.3.4. Ternary systems

The parameters that were obtained by regression to binary data are used to predict the behavior of the components in a ternary mixture. Mathias and Kister [134] have shown that including ternary LLE in the regression of binary parameters can lead to improved simultaneous representation. However, to explore the behavior for predictive capabilities, the simultaneous regressions were not performed in this work. The phase diagram for methanol + *n*-heptane + cyclohexane is given in Figure 5.31 and the K-ratios for each of the components are given in Figure 5.31. The predicted LLE compositions are most accurate for the four parameter WAG-NRTL and NRTL models. WAG-SH and CPA also provide a good prediction. However, PC-SAFT and the other two-parameter WAG variations struggle to capture the methanol and cyclohexane rich phase.



Figure 5.31: Phase equilibria diagram for methanol + *n*-heptane + cyclohexane at 101.3 kPa and 298.15 K. Parameters used were fitted only to binary data.



Figure 5.32: K-ratios for methanol (a) + *n*-heptane (b) + cyclohexane (c) system at 101.3 kPa and 298.15 K.

Similarly, the ternary diagram for methanol + ethanol + cyclohexane and K-ratios for the components is given in Figures 5.33 - 5.34. The performance of the models is much different for this system. NRTL's false prediction of an LLE between ethanol and cyclohexane creates the wrong envelope in the figure. While this is the worst description of data, the other models are also weak at predicting the behavior of this system. However, WAG-SH provides the best prediction, out-performing CPA and PC-SAFT.



Figure 5.33: Phase equilibria diagram for methanol + ethanol + cyclohexane at 101.3 kPa and 298 K. Parameters used were fitted only to binary data.



Figure 5.34: K-ratios for methanol (a) + ethanol (b) + cyclohexane (c) system at 101.3 kPa and 298 K.

5.3.5. General findings and recommendations

Some general conclusions can be drawn from the results shown in this work. The first is that 4parameter variations of WAG-NRTL and WAG-Nagata models provide excellent fits of experimental data for binary systems. The total SSQ of these two models is 47% and 15% smaller than that of 4-parameter NRTL, which is the model with the next smallest SSQ, and are therefore recommended as four parameter models. For systems that are difficult to model, the performance of WAG-NRTL model could be further improved by fitting the α_{ij} parameter. However, because the goal of this work is to compare WAG-Nagata and WAG-NRTL with the same number of adjustable parameters, this was not examined for the systems studied here, and is therefore left for future studies.

Amongst two-parameter models, the WAG-SH model developed here is a clear frontrunner. WAG-SH has been shown to correlate the behavior of the systems in this work better than the equations of state. It is particularly interesting to compare it to CPA. The functional forms of the energetic contributions in both WAG-SH and CPA are very similar but the SSQ for WAG-SH is only about 70% of that for CPA. The differences are that WAG-SH uses a 2/3rd power on the volume terms in Flory's theory whereas CPA does not, and the liquid volumes are calculated differently. Additionally, CPA includes a correction for Flory's theory derived from the van der Waals equation [8,67].

The parameter of \tilde{a}_{ij} is small for WAG-SH, CPA and PC-SAFT across all systems. However, the parameter b_{ij} of WAG-SH is consistently larger than that of CPA and PC-SAFT, in some cases by three orders of magnitude. This is because, the k_{ij} parameters in the EOS are corrections to the attractive parameter which have an imbedded temperature dependence in their functional form. In contrast, the fitted parameters in WAG-SH, and indeed all the WAG models, provide all of the

temperature dependence of the residual term. Therefore, the difference in the magnitude of the parameters between the WAG and EOS methods is expected. The small values of the fitted parameters in the equations of state indicate that the functional form of the equations adequately captures the temperature dependence of the experimental data. However, the fits generated by them are still weaker than those by the 4-parameter models for most systems. This is a consequence of the poor compositional dependence of CPA and PC-SAFT compared to WAG and NRTL. As indicated by Flemr [135], NRTL and Wilson (and by extension, Nagata's model) share a

common flaw: violation of local composition pair conservation (Flemr's condition) given by.

$$N_a N_{ba} = N_b N_{ab} 5.38$$

 N_i is the total number of *i* molecules and N_{ij} is the number of *i* molecules around *j*. Lin et al. [136] showed that the non-compliance of these models to Flemr's condition is due to the use of Boltzmann weighing factors for local compositions. Instead, they introduced a flexible weight, f_i , characteristic of each species, to enforce pair conservation. The resulting function, while more fundamentally correct, requires iterative solution. Moreover, similar to EOSs, the equation is less flexible with respect to composition than NRTL or Nagata. Despite this, an interesting future extension of this work can explore combining a residual contribution that obeys Flemr's condition with the combinatorial and association contributions used in this work.

The SSQ analysis reported here is only strictly applicable to the current set of systems and will therefore not necessarily represent all systems. However, the relative performances of the models are expected to yield similar trends. Further, some conclusions for this work may be dependent on the association parameter. An observation in this work is that experimental alcohol + inert infinite dilution activity coefficients increase rapidly at infinite dilution. When the phase behavior is fitted using a model without chemical association, false LLE can be predicted because the residual

models do not include this rapid rise at infinite dilution (except for Nagata). However, adding chemical association has a strong effect at infinite dilution, and suppresses the tendency for false LLE because the association term creates very large infinite dilution activity coefficients before LLE begins to occur. Larger values of association constants permit fitting of infinite dilution activity coefficients while diminishing the tendency for false LLE and warrants further study. Future work will also investigate the relative magnitudes of the contributions to the activity coefficient, γ . For the WAG models and the systems tested here involving alcohol + hydrocarbons, the activity coefficient is dominated by the association contribution. The ratios of γ^{assoc} to the total γ at infinite dilution compositions are given in Table 5.8.

	Component	$\gamma^{\infty,assoc}/\gamma^{\infty}$					
System	at infinite	WAG- WAG-		WAC SIL	WAG-	WAG-	
	dilution	NRTL(4)	Nagata(4)	WAG-SH	NRTL(2)	Nagata(2)	
Methanol +	Methanol	1.255	1.234	1.177	1.082	1.145	
<i>n</i> -Heptane T=335 K	<i>n</i> -Heptane	1.368	1.274	1.069	1.070	0.903	
Methanol +	Methanol	0.988	0.958	0.954	-	0.924	
Cyclohexane T=335 K	Cyclohexane	0.880	0.793	0.725	-	0.712	
Methanol +	Methanol	2.844	3.845	3.905	2.848	4.129	
Water T=350 K	Water	1.226	1.793	1.881	2.055	1.794	
Methanol +	Methanol	-	-	1.061	1.073	1.089	
<i>n</i> -Pentane T=370 K	<i>n</i> -Pentane	-	-	0.897	1.060	0.955	
Methanol +	Methanol	1.056	1.048	0.935	1.025	1.028	
Ethanol T=340 K	Ethanol	1.064	1.060	0.902	1.029	1.030	
Ethanol +	Ethanol	0.877	0.861	0.851	0.773	0.840	
Cyclohexane T=345 K	Cyclohexane	0.680	0.684	0.711	0.725	0.641	
Ethanol +	Ethanol	1.429	1.448	1.245	1.211	1.212	
<i>n</i> -Decane T=380 K	<i>n</i> -Decane	1.680	1.283	1.392	1.434	1.281	

 Table 5.8: Association contribution to the total activity coefficient at infinite dilution

In general, the ratios are very similar across all WAG models for every system. However, there is significant variation in the calculated $\gamma^{\infty,assoc}/\gamma^{\infty}$ values for the methanol + water system. Since the calculated values of γ^{assoc} and $\gamma^{Mod \ Flory}$ are constant for a given system and condition, this is an indication that the calculated energies of the dispersion interactions, described by γ^{res} , are very different depending on the calculation method used. Compared to Nagata and SH, NRTL calculates larger deviations from ideality at infinite dilution due to dispersion interactions. Thus, future work should include consideration of infinite dilution values.

5.4. Association parameter value

The parameters used to model ethanol with CPA were different than those used to calculate the ethanol association parameters in WAG. In the former case, the parameters, which will be referred to here as PM1, were calculated by fitting to vapor pressure and saturated liquid density data [1]. This procedure is typical of equations of state such as PC-SAFT and CPA. However, other studies have employed different methods for the same purpose. For example, Renon and Prausnitz [68] and Campbell [16] fit the association parameters directly to binary phase equilibrium data. Others [70,96] follow an approach developed by Brandani *et al.* [73] that calculates the association parameter by quantifying the difference between the vapor pressure of an alcohol and that of an ether homomorph with a similar molecular size. In a recent study, Kontogeorgis *et al.* [2] refitted CPA parameters to data that included the fraction of molecules that are non-bonded at equilibrium (monomers). This quantity was calculated from spectroscopy and the resulting parameter set, PM2, was found to be significantly different from those calculated previously. The values fitted to spectroscopy for ethanol give more association and thus a larger association contribution to infinite dilution. Unfortunately, the parameters also do not simultaneously represent the vapor pressure.



We illustrate this difference for ethanol + cyclohexane and ethanol + n-decane in Figures 5.35 -

5.36.

Figure 5.35: Comparison of phase equilibria fits with two sets of CPA parameters for ethanol + cyclohexane at 101.3 kPa. Solid lines are fits calculated using PM1 parameters [1] and dashed lines are calculated with PM2 parameters [2]



Figure 5.36: Comparison of phase equilibria fits with two sets of CPA parameters for ethanol + *n*-decane at T=338.17 K. Solid lines are fits calculated using PM1 parameters [1] and dashed lines are calculated with PM2 parameters [2]

The solid and dashed lines in Figures 5.35 - 5.36 are fits of the CPA model made using PM1 and PM2 parameters respectively. The end points for pure ethanol are completely missed by the PM2 parameters. Thus, PM1 parameters were implemented for CPA fits in this work. Exploratory work showed that most of the WAG models resulted in false LLE when regressed with PM1 for ethanol + cyclohexane. Also, because of the direct relationship between Δ^{AD} and monomer fraction, the PM2 parameters, which are fitted to the monomer fraction, are expected to be more fundamentally consistent with reality. Therefore, we elected to use PM2 parameters for ethanol to calculate the association parameter for WAG models.

The Δ^{AD} parameters for the self-association of primary alcohols as calculated by the various approaches described are shown in Figure 5.37.



Figure 5.37: Δ^{AD} for the self-association of the primary alcohols at 298.15 K calculated by fitting to different properties. Parameters were calculated by fitting to pure component properties (CPA [1,64], PC-SAFT [115]), binary VLE data (Renon [68], Campbell [16]) or vapor pressures of ether homomorphs (Bala [96], using methodology from Nagata [70]). The marker '×' shows the association parameter calculated with the PM2 parameters which are fitted to spectroscopic and pure component data

For the methods of Renon, Campbell and Bala, association was modeled using chemical theory and the calculated parameter was the equilibrium constant, K_C . In these cases, we use the finding of our earlier work [96] which showed that, with reasonable assumptions, $\Delta^{AD} = K_C$, to compare the values of the Δ^{AD} .

It is immediately clear from Figure 5.37 that the value of Δ^{AD} can vary by a factor of over 6 depending on the calculation method used. This is due, in part, to the fact that all these methods, with the exception of the one that yields the PM2 CPA parameter set, lack molecular-level experimental insight. Instead, they depend on macroscopic manifestations of hydrogen bonding. To address this situation, more fundamental techniques such as spectroscopy must be utilized in parameter calculations and we explore these methods in the following chapters. It is worth noting that Δ^{AD} calculated for the self-association of ethanol from the PM2 parameters, which include spectroscopic information, falls more in line with the values calculated from other methods and are close to that calculated by PC-SAFT. The work of Asprion *et al.* [7,137,138]uses spectroscopy, but is not directly useful in the Wertheim framework. Regardless of the uncertainty in the exact value of the association constant, this work demonstrates that inclusion of association is important to improve the performance of traditional activity coefficient models.

5.5. Conclusions

In this work, a Wertheim activity coefficient term is combined with existing thermodynamic models including NRTL, Scatchard-Hildebrand and the energetic part of Wilson's equation. The resulting models are used to fit phase equilibria and heat of mixing data for methanol and ethanol-containing systems and compared to CPA and PC-SAFT. The four parameter WAG models were found to perform the best, with sum of squares errors that were 3 to 22% of those of 2-parameter models. Among the two parameter models, WAG-SH provided comparable or, in some cases,

superior results to CPA and PC-SAFT. Including association improved the models' capability to fit simultaneous VLE and LLE data even though there was uncertainty in the association constant value. While CPA parameters from literature were used to calculate the association parameter values in this work, the resulting fits highlight the importance of incorporating spectroscopic findings in determining accurate values of the association parameter Δ^{AD} .
APPENDIX

		\tilde{a}_{ij}	\tilde{a}_{ji}	\tilde{b}_{ij}	$ ilde{b}_{ji}$	α_{ij}
	WAG-NRTL(2prms)	_	-	530	-318	0.2
	WAG-NRTL(4prms)	-2.71	0.233	1400	-426	0.2
+	WAG-SH	-0.05	505	24	.7	-
l (<i>i</i>) le (<i>j</i>	СРА	-0.02	.98	0.0336		-
ano	WAG-Nagata(2prms)	-	-	-140	-17.5	-
eth: -hej	WAG-Nagata(4prms)	4.85	-0.691	-1532	177	-
M n	NRTL(2prms)	0	0	541	387	0.2
	NRTL(4prms)	0.781	-3.96	335	1515	0.2
	PC-SAFT	0.08	30	-0.04	455	-
	WAG-NRTL(2prms)	_	-	374	-225	0.3
(j)	WAG-SH -0.128		28	57.7		-
iol (ane	СРА	-0.01	.86	0.05	557	-
han ent	WAG-Nagata(2prms)	-	-	-79.4	-22.2	-
Met <i>n</i> -p	NRTL(2prms) -		-	552	492	0.3
	PC-SAFT 0.0534		0.00770		-	
	WAG-NRTL(4prms)	-0.645	-0.643	732	-101	0.2
	WAG-SH	-0.09	065	46	.8	-
(<i>i</i>) + (<i>j</i>) = (<i>j</i>)	CPA** 0.0320		20	0.00782		-
iol (WAG-Nagata(2prms)	-	-	-14.1	-80.1	-
lhan lohe	WAG-Nagata(4prms)	1.62	-0.196	-502	-23.7	-
Met [yc]	NRTL(2prms)	-	-	354	459	0.2
	NRTL(4prms)	0.0803	-4.780	347	1906	0.2
	PC-SAFT	0.0726		-0.0220		-
	WAG-NRTL(2prms)	-	-	41.2	-20.4	0.2
	WAG-NRTL(4prms)	-6.61	5.56	2170	-1746	0.2
thanol (i) + t-decane (j)	WAG-SH -0.01		54		0	-
	CPA		22	0.122		-
	WAG-Nagata(2prms)	-	-	-5.63	-18.1	-
	WAG-Nagata(4prms)	5.635	-1.732	-1721	472	-
Ц	NRTL(2prms)	-	-	138	749	0.2
	NRTL(4prms)	1.35	-0.554	-63.0	690	0.2
	PC-SAFT	0.0634		-0.0363		-

Table H.1: Regressed parameters values. The parameter α_{ij} was not regressed but is tabulated for convenience

Table H.1 (cont'd)

		ã _{ij}	ã _{ji}	${ ilde b}_{ij}$	${ ilde b}_{ii}$	α_{ij}
	WAG-NRTL(2prms)	_	-	201	-55.6	0.3
	WAG-NRTL(4prms)	0.466	-0.529	292	-52.5	0.3
+ 🕄	WAG-SH	-0.0189		17.4		-
(j) ane	СРА	-0.0560		0.0947		-
nol hex	WAG-Nagata(2prms)	-	-	-152	-19.6	-
tha	WAG-Nagata(4prms)	0.3605	-0.103	-256	16.8	-
Cy E	NRTL(2prms)	-	-	187	794	0.3
	NRTL(4prms)	-0.00483	0.912	238	415	0.3
	PC-SAFT	0.06	546	-0.0225		-
	WAG-NRTL(2prms)*	0	0	-95.1	123	0.3
	WAG-NRTL(4prms)	0.0367	-0.494	-142	344	0.3
+ 5	WAG-SH	-0.0	202	8.4	41	-
e (i ane	СРА	0.01	123	-0.0	119	-
otan nex;	WAG-Nagata(2prms)*	0	0	-193	184	-
Hep cloł	WAG-Nagata(4prms)	-0.159	0.625	65.4	-259	-
<i>n-</i>] cy	NRTL(2prms)*	0	0	-123	149	0.3
	NRTL(4prms)	-0.0382	-0.435	-131	334	0.3
	PC-SAFT -0.00573		0.00787		-	
	WAG-NRTL(2prms)	-	-	-473	322	0.2
	WAG-NRTL(4prms)	12.7	-4.74	-2881	720	0.2
+	WAG-SH	-0.1	47	-26	5.6	-
(i)	СРА	-0.0881		0.0163		-
ano ater	WAG-Nagata(2prms)	-	-	130	59.6	-
eth W	WAG-Nagata(4prms)	1.72	-2.76	-489	1058	-
Μ	NRTL(2prms)	-	-	-111	345	0.2
	NRTL(4prms)	3.59	-3.53	-1224	1430	0.2
	PC-SAFT	0.00387		-0.0759		-
	WAG-NRTL(2prms)	-	-	67.8	-63.2	0.3
	WAG-NRTL(4prms)	-0.868	0.712	336	-291	0.3
anol (i) + anol (j)	WAG-SH	WAG-SH 0.01		1.28		-
	CPA**	0.01	155	-0.0	365	-
	WAG-Nagata(2prms)	-	-	-49.4	32.2	-
Eth	WAG-Nagata(4prms)	1.03	-0.660	-327	215	-
Z	NRTL(2prms)	-	-	62.8	-53.8	0.3
	NRTL(4prms)	-1.03	0.929	351	-320	0.3
	PC-SAFT	0.00926		-0.00830		-

		Comp j		
		Methanol	Ethanol	
np i	Methanol	1.63	1.55	
COL	Ethanol	1.55	1.4	

Table H.2: Hayden-O'Connell η_{ij} **parameters**. The value of η_{ij} , which is dimensionless, for alcohol + alkane binaries is zero

Chapter 6. Quantitative Analysis of Infrared Spectra of *n*-Butanol + Cyclohexane with Quantum Chemical Calculations

Hydrogen bonding has profound effects on the microscopic behavior of molecules. Infrared spectroscopy (IR) allows for the analysis of molecular vibrations by excitation with infrared radiation. However, quantitative analysis of the hydroxyl band in the IR spectrum, where hydrogen bonding is most prominently expressed, is non-trivial. Specifically, the broadness of the band and the range of variation of the O-H absorption coefficient complicate the analysis. In the present work, sequential MD and QM simulations are used to develop functions that relate the vibrational frequencies to the integrated absorption coefficients of hydroxyl bands. This relationship is then used to quantitatively calculate the mixture concentration of alcohol molecules from experimental IR spectra by integration across the entire hydroxyl band. Finally, universal maps to relate bond lengths, vibrations and NMR chemical shifts are proposed.

6.1. Introduction

In infrared (IR) spectroscopy, molecules are excited with light in the mid IR wavelength region and the amount of light absorbed by the sample is detected. Absorption of light at different wavelengths induces vibrations of different bonds within the molecule and analysis of the resulting spectra can yield information about the types of functional groups it possesses. IR spectroscopy has long been used in this way to identify chemical compounds through the interpretation of band patterns. Moreover, quantitative analyses of IR spectra can be used to gain insight into the concentration of functional groups in a solution. For example, Williams *et al.* [139] explored the relationship between absolute integrated intensities of the C-H stretching and bending bands of gas-phase alkanes. The authors compared density functional theory (DFT) calculation results to experimental IR spectra and found that the number of C-H bonds in the molecules studied is linearly correlated with the integrated intensities of C-H stretching and bending bands. IR is also frequently used in chemometrics [140]. Several researchers have been able to correlate intensities of various classes of compounds with physical characteristics of the molecules such as the number of methylene groups [141], molecular size [142] and degree of branching [143].

The O-H stretching bands have been the focus of many studies that aim to understand the complex effects of hydrogen bonding [37,144–151]. These bands occur between ~3200 cm⁻¹ and 3700 cm⁻¹ ¹ and their combinations appear in the form of two overall peaks: a sharp higher frequency peak and a broad lower frequency peak. The formation of hydrogen bonds is known to decrease the O-H stretching band frequency and increase its integrated intensity [152]. Wu et al. [146] investigated supercritical and liquid methanol and found that an isothermal increase in density causes the integrated spectroscopic area to increase and the hydroxyl vibrational band to shift to lower frequencies. Expectedly, isobaric heating has the opposite effect. Due to these widely varying measurement response factors compared to C-H bonds, vibrations associated with free and hydrogen-bonded hydroxyl groups are much more difficult to interpret in quantitative terms and have therefore largely been interpreted qualitatively. To overcome some of the challenges involved in IR analysis of hydroxyl peaks, computational tools such as molecular dynamics (MD) [153] and quantum mechanical (QM) [139,144,154–159] simulations are being leveraged to elucidate the effects of hydrogen bonding on IR peak characteristics. MD and QM calculations balance computational expense with modeling rigor. The former method is less demanding of computational time and resources but does not model electronic effects, which are the basis of hydrogen bonding and, in particular, result in the distributions of hydrogen bond types. Indeed, Kwac and Geva [17] show that, depending on the choice of empirical forcefield, the simulated

relative populations of species can vary by as much as 20%. For this reason, MD is often used in conjunction with QM [17,160–162] rather than as the primary tool for investigation.

A recent development in the computational sphere is empirical mapping. This approach, developed by Skinner *et al.* [163–167] for water, creates functions or "maps" that relate vibrational frequencies to properties that can be calculated with reasonable accuracy from MD calculations. In this way, one can obtain a meaningful fundamental understanding of a system's IR response without having to use excessive computational resources. Mesele and Thompson [168] extended these techniques to primary alcohols, developing several "universal" maps that relate the transition frequencies, dipole derivatives and position matrix elements to the electric field on the atoms.

In this work, we aim to present a combined computational and experimental approach which leverages the power of simulations to address the challenges of interpreting the infrared behavior of O-H bonds. Furthermore, we apply this technique to quantitatively analyze the entire hydroxyl band and calculate the relative and absolute concentrations of hydroxyl groups in the various contexts (monomers and clusters) existing in solution. In all our discussions, references to the hydroxyl vibrational bands pertain to the vibrations of the covalent O-H bond and not vibrations of the actual hydrogen bond (which appear at much lower frequencies [152]). Additionally, the term 'apparent concentration' will be used to refer to concentration of species in solution ignoring the clusters formed by hydrogen bonding – the molecules are counted individually. The term "apparent concentration" in chemical engineering is synonymous with "formal concentration" in chemistry literature.

6.1.1. Analysis of the O-H stretching band

Quantitative interpretation of infrared spectra begins with the Beer-Lambert law, which relates the observed IR absorbance of a solute to its concentration in solution according to:

$$A_i = \varepsilon_i C_i l \tag{6.1}$$

where A_i and C_i are the observed absorbance and concentration of the absorbing solute *i* respectively. The pathlength, *l*, is the length of the sample that the IR light beam passes through and ε_i is the molar absorption coefficient, also known as the extinction coefficient, of species *i*. The latter parameter is a measure of the IR absorbance of a single bond vibration at a selected wavenumber and is often assumed to be independent of concentration for the solute. In common practice, solutions of known concentrations are prepared and analyzed with IR spectroscopy. Then, a peak corresponding to a vibration in the solute molecules is chosen in an area where overlap with solvent bands is minimal. The absorbance peak height values are then plotted against the experimental concentrations of the solutions. Finally, the value of ε is calculated as the gradient of this plot, which is ideally linear, and used in subsequent studies to analyze solutions of unknown concentration.

However, the Beer-Lambert law is not universal; it fails under conditions where the relationship between A_i and C_i becomes non-linear. This occurs when the solute molecules interact significantly with one another or solvent molecules, such as in high concentration solutions. In these cases, the value of the molar absorption coefficient, ε , can vary dramatically for the same bond depending on the molecule conformation and its microenvironment. This contributes to the broadening of the observed IR band and complicates quantification. Moreover, there is disagreement in the literature concerning the assignment of vibrational bands. In earlier studies [137,145,169], vibrational bands were assigned to hydrogen bonded clusters and were distinguished based on the size of the cluster. Hall and Wood [170] were the first to propose that covalent O-H bond vibrations should be classified individually according to if and how they participate in hydrogen bonding. If the O-H is neither accepting nor donating a hydrogen atom, it is classified as an α bond. If it is accepting one hydrogen atom but not donating, it is classified as a β bond. These and the other 4 categories of bonds are illustrated in Figure 6.1. Using femtosecond IR pump probe studies of ethanol clusters, Woutersen *et al.* [171] were able to assign the α , β , γ and δ bonds to experimentally deconvoluted peaks as far back as 1997.



Figure 6.1: Types of covalent O-H bonds

This approach to IR band interpretation is more aligned with the fundamental basis of IR excitation as a bond perturbation rather than a molecular or cluster-wide phenomenon. Therefore, Hall and Wood's classification system has been adopted more frequently in recent work [147,148,150,172] and will also be used in this work.

6.1.2. *Motivation*

While research in this area is extensive, we are unaware of any work demonstrating quantification of the entire hydroxyl IR band area for alcohols and relating the area to the apparent concentration. Such a relationship has profound implications for the study and modeling of hydrogen bonding. Indeed, in the development of engineering models for the association of an alcohol in an inert solvent, the key parameter that defines the extent of hydrogen bonding is the fraction of hydroxyl protons that remains nonbonded at equilibrium. Once this value is determined, some assumptions can be made about the types of species present in solution and the energy of their formations. A simple model that has often been adopted [98,173,174] for alcohol + inert systems is the CLAM, or Continuous Linear Association Model. As the name suggests, CLAM accounts only for linear species in solution and assumes that each hydrogen bond has the same energy of formation regardless of the size of the cluster it forms. The motivation behind the current study is to develop a procedure capable of accurately determining the fraction of free hydroxyl protons and conduct a thorough quantitative analysis of the O-H IR bands.

6.2. Computational methods

To analyze the vibrational spectra derived from infrared spectroscopy, a combined molecular dynamics (MD) and quantum mechanics (QM) approach was used. In this section, these computational techniques are outlined.

6.2.1. Molecular dynamics simulations

Molecular dynamics simulations were carried out using the AMBER 14 package [175]. The AMBER94 force field was implemented with the AM1-BCC charge method and no modification to the force field parameters. For each concentration, a cubic box of *n*-butanol and cyclohexane molecules was created at the ideal solution density using PACKMOL [176] and the energy of the system was minimized within 1500 steps. Next, the box was heated up with a 40 ps NVT run, using a time step of 2 fs, in two stages. The temperature was ramped up from 0 to 283.15 K during the first 9000 steps then maintained at that temperature for the remaining 11000 steps. Finally, a 10.4 ns NPT production run was conducted at 1 bar and a time step size of 2 fs. The temperature and pressure were controlled using the Langevin thermostat (with a collision frequency of 2 fs) and Berendsen barostat respectively; both implemented with default parameters. During the NVT and NPT simulations, periodic boundary conditions were enforced in the x,y and z coordinates.

The cutoff for non-bonded interactions was set at 8 Å and a continuum model was used for Lennard-Jones interactions beyond this range. Electrostatic interactions were calculated using the Particle Mesh Ewald (PME) summation method with default settings and parameters. Only bond lengths involving hydrogen were constrained using the SHAKE algorithm. This method was repeated for *n*-butanol + cyclohexane and ethanol + cyclohexane at the equimolar composition. The simulation details that varied between systems are given in Table 6.1.

 Table 6.1: MD simulation details. Box and run for *n*-butanol + cyclohexane systems studied in this work

System	<i>n</i> -Butanol + Cyclohexane		Ethanol + Cyclohexane	
Alcohol mole fraction	0.1	0.5	0.5	
Number of alcohol molecules	26	128	168	
Number of cyclohexane molecules	230	128	168	
Initial cubic box length (Å)	40	40	36	
Temperature (K)	283.15	283.15	298.15	
Production run length (ns)	10.4	10.4	13.0	

The production runs were longer than 10 ns in all cases, giving all the systems ample time to equilibrate. Furthermore, only the last 2.4 ps of the simulations were analyzed for hydrogen bonding information.

The hydrogen bond criteria were defined as an O-O distance < 3.2 Å and an O-H…O bond angle > 130°. Using this definition, hydrogen bonded clusters of various sizes were identified, and each bond was assigned a class based on the categories in Figure 6.1. For the current study, only linear species are analyzed. Furthermore, the low incidence of structures including η and ζ bonds suggested that they play a minor role; they are thus neglected.

To verify thermal equilibration, Figure 6.2 shows the number of times each *n*-butanol molecule appears as a β , δ or γ bond in trimers in approximately 590 evenly spaced frames collected every 40 fs in the last 2.4 ps of the production run. Each molecule appears in a trimer in about 15% of

the frames. The *x*-axis is the unique identifying number of each *n*-butanol molecule in the box. The distribution is random for the three bond types across all molecules in solution.



Figure 6.2: Bond distribution in trimers calculated with MD simulations for an equimolar *n*-butanol + cyclohexane mixture. The distribution illustrates that molecules are equally likely to sample all bond types.

6.2.2. Quantum mechanics simulations

Once the hydroxyl bonds were classified, each cluster was prepared to undergo QM calculations. Careful implementation of QM simulations is a balance between accuracy and computational expense. For this reason, performing extensive calculations over the entire box was not feasible and an alternative was necessary. In this work, all molecules with atoms that fell within 5 Å of the hydroxyl hydrogen atom in the hydrogen bonded cluster were retained in the simulation environment. All other molecules were excluded. This cutoff effectively ensures that the sampled environments were representative of the whole box while minimizing the computational effort required for quantum calculations. To this end, we also constrained all atoms and bonds except for the -CH₂-O-H groups in each alcohol molecule of the hydrogen bonded cluster of interest.

Gaussian 09 [177] was used to optimize the geometries of each cluster and perform the frequency calculations. The B3LYP [178] level of theory and 6-31G* basis set were chosen shown [154,159] to capture the effects of hydrogen bonding accurately for a reasonable computational cost.

Finally, the IR vibrational information calculated underwent two tests. In the first, the final atom positions were checked to ensure that the covalent bond classifications made earlier had not changed during QM optimization. Secondly, vibrational coupling between multiple -O-H bonds was estimated by calculating the displacement of each -O-H hydrogen atom in the cluster using:

$$displacement = \Delta x_{H}^{2} + \Delta y_{H}^{2} + \Delta z_{H}^{2}$$

$$6.2$$

where Δx_H , Δy_H and Δz_H is the displacement of the hydrogen atom in the *x*, *y*, and *z* directions respectively. Several studies have shown that vibrational coupling between different groups have an effect on the spectra [149,172,179,180]. However, in this work, we are interested in comparing the behaviors of individual bond types. Therefore, to reduce the effects of coupling, vibrations were excluded from analyses if they could not be confidently assigned to only one –O-H bond. We define confidence as having one hydrogen atom's displacement in a vibration being at least 0.3 Å² greater than any other's in the optimized geometry.

6.3. Experimental methods

Experimental IR data collection was performed by William G. Killian, a member of the Lira Lab at Michigan State University and will be included in his Ph.D. dissertation [181]. Therefore, this section provides a basic outline of the experimental methodology for context. Cyclohexane, ethanol and *n*-butanol were purchased in anhydrous forms at purities of 99.5%, 99.5% and 99.8% respectively and further dried using 3 Å molecular sieves for at least 72 hours. Samples were

prepared volumetrically in a glove box and then transferred to scintillation vials to be removed from the inert environment for analysis.

Infrared absorbance spectra were collected on a JASCO FT/IR-6600 spectrometer using a Specac demountable temperature-controlled liquid flow cell (GS20582) with CaF₂ windows and PTFE spacers. The system was allowed to stabilize for 10 minutes at each temperature. The pathlength for all measurements was 0.01092 cm. A background of the empty cell was taken at each temperature and subtracted from the sample data. For each reported spectrum, 128 scans were accumulated with a resolution of 2 cm⁻¹.

6.4. Results and discussions

6.4.1. Processing and preliminary analysis of experimental IR spectra

The raw IR spectra were processed as follows. The hydroxyl band region is determined to be 3049.9 to 3755.2 cm⁻¹. First, the solvent bands were removed from the spectra by subtracting concentration-weighted spectroscopic data of neat cyclohexane at the same temperature and pathlength of the sample. The concentration was calculated by assuming ideal volume of mixing using temperature-dependent volumes through $V = \sum_i x_i V_i$. The processed experimental IR spectra for *n*-butanol in cyclohexane in the region of the hydroxyl stretching band is given in Figure 6.3 for $x_{BuOH} = 0.1$. As the temperature increases, the peak at ~3650 cm⁻¹ and shoulder at ~3540 cm⁻¹, P_1 and P_2 respectively, increase in absorbance while the broad peak centered at ~3320 cm⁻¹, P_3 , diminishes.



Figure 6.3: Experimental O-H IR band for a 0.1 mole fraction of *n***-butanol in cyclohexane.** Data was collected at five temperatures between 30 °C and 70 °C

While band assignment is a topic of great interest to the scientific community, we defer such discussion to the following chapter. In this work, we instead conduct a quantitative analysis of the entire hydroxyl band. To this end, we begin with an investigation of the physical significance of the absorbance band area. Figure 6.4 shows the relationship between the area under the entire O-H absorbance band and the apparent concentration for four concentrations at five temperatures.



Figure 6.4: Total absorbance band area of the O-H band for *n*-butanol + cyclohexane data as a function of the mole fraction for 4 compositions. Spectra were collected at x_{BuOH} = 0.0469,0.0678, 0.0817 and 0.100 and the lines indicate the line of best fit given by the equations on the figure. A thermal correction for density is applied.

When temperature rises, it is expected that an increasing number of hydrogen bonds will break, forming smaller chains and monomers. Thus, the distribution of the different bond types will change. However, at a given composition, the total quantity of hydroxyl groups should remain constant across the temperature range studied here. To correct concentration for the effect of temperature on the density, the band areas plotted in Figure 6.4 are scaled with a thermal correction factor which is defined as:

$$\theta = \frac{\rho^{ref}}{\rho} \tag{6.3}$$

where ρ and ρ^{ref} are the mixture densities at the temperature of the experiment and a reference temperature (298.15 K) respectively. The mixture density is calculated by assuming no excess volume. Therefore, $\rho = (\sum_i x_i V_i)^{-1}$ where V_i is the molar volume of component *i*. It is immediately clear from Figure 6.4 that the assumed linearity of A_i and C_i that is traditionally assumed for a given absorbance in the Beer-Lambert law does not apply for the overall hydroxyl band. If it did, all of the points (at all temperatures and compositions) would be reasonably well represented with a single straight line that passes through zero on both axes. However, while data at each temperature is quite linear, the intercept varies significantly between temperatures, indicating that the changing distribution of bond types is not being captured. This discrepancy is expected as hydrogen bonding is known to conflict with the assumptions that underlie the Beer-Lambert law. Figure 6.4 is evidence that the molar absorption coefficient, ε , must vary at different vibrational frequencies. Next, the results of the QM/MM analysis are shown to provide key insights into this variability, resulting in relationships between absorption intensity, wavenumber, and covalent bond length.

6.4.2. Results from MD+QM

In this section, we investigate the calculated IR characteristics of the hydroxyl stretch. Because of the computational expenses associated with each QM calculation, the scope is limited to monomers, dimers, trimers, and tetramers and Table 6.2 lists the numbers of each bond and species type analyzed.

Table 6.2: Number of each species and bond analyzed with QM calculations. All species larger than a monomer have one β and one γ bond. Trimers and tetramers also possess one and two δ bonds respectively. Cyclic clusters are not considered.

		$x_{BuOH} = 0.1$	$x_{BuOH} = 0.5$	Total
cies	Monomer	653	366	1019
	Dimer	153	230	383
Spe	Trimer	53	52	105
9 1	Tetramer	46	23	69
	α	653	366	1019
Bonds	β	252	305	557
	δ	145	98	243
	γ	252	305	557

The calculated "intensity" from Gaussian09 is the integrated absorption coefficient, ψ , in units of km/mol, which is defined by the equation:

$$\psi = \int_{band} \varepsilon(\tilde{\nu}) d\tilde{\nu}$$

$$6.4$$

In this function, the integral spans the entire absorption band. Like the molar absorption coefficient, the integrated absorption coefficient is directly related to the derivative of the dipole moment and is a measure of the intensity of a transition. Literature on the relationship between the integrated absorption coefficient, ψ , and the molar absorption coefficient, ε , is somewhat scarce. However, Spanget-Larson [182] provides a simple interpretation of the process which involves assuming a function for $\varepsilon(\tilde{v})$. When a Lorentzian function is assumed for $\varepsilon(\tilde{v})$, then ψ is related to the peak height selected for ε_{max} via a constant. For the purposes of this work, the results will be scaled by a constant and thus avoiding the need for numerical conversion or specification of the Lorentzian parameter values. We use ψ reported by Gaussian09 as a measure of each hydroxyl's IR absorbance, with the understanding that it is proportional to the peak height ε_{max} . Following standard practice, the vibrational frequencies resulting from Gaussian are scaled by a factor of 0.96 to correct for the limitations of the B3LYP/6-31G* model and better match experimental IR data [183].

To visualize general trends, the average vibrational wavenumbers and integrated absorption coefficients for different bonds and species are given in Table 6.3. The first observation is that with an increase in the apparent concentration of alcohol in the simulated box, there is a slight red-shift in almost all the vibrational frequencies, most prominently those of the α and β bonds. In general, there is also a red-shifting tendency for bond vibrations as the size of the chain they are on grows. Indeed, the difference is most prominent between dimers and larger oligomers with dimer β and γ bonds having vibrational frequencies that are, on average, ~9 and ~50 cm⁻¹ greater

than those of trimers and tetramers respectively. Table 6.3 follows a trend that increases in the order α , β , γ , δ though the differences between species are more complex.

Gradian	Bond	$\tilde{\nu}_{avg} \ (\mathrm{cm}^{-1})$		ψ_{avg} (km/mol)		
Species		$x_{BuOH}=0.1$	$x_{BuOH}=0.5$	$x_{BuOH}=0.1$	$x_{BuOH} = 0.5$	
Monomer	α	3602	3587	31.90	53.97	
D.:	β	3601	3583	52.33	72.32	
Dimer	γ	3479	3471	376.5	411.0	
	β	3589	3572	68.31	96.21	
Trimer	γ	3418	3414	464.1	525.2	
	δ	3410	3412	583.9	537.0	
	β	3586	3574	68.39	87.65	
Tetramer	γ	3436	3426	426.8	428.1	
	δ	3387	3363	628.8	675.9	

 Table 6.3: Average calculated vibrational frequencies and integrated absorption coefficients

Figure 6.5 shows the vibrational frequency and integrated absorption coefficient associated with each hydroxyl vibration at both apparent concentrations, $x_{BuOH} = 0.1$ and $x_{BuOH} = 0.5$. When the two concentrations are considered separately, we observe little difference in the trends, indicating that apparent concentration has little effect on the $\tilde{\nu}$ - ψ relationship. Therefore, they are plotted together in Figure 6.5 and the different markers denote the four bond types studied here: α (diamonds), β (triangles), γ (squares), and δ (circles).



Figure 6.5: Hydroxyl stretching frequencies and integrated absorption coefficients for two concentrations of *n*-butanol + cyclohexane mixtures calculated from QM simulations. Diamond, triangle, square and circle markers denote α , β , γ , and δ bond vibrations respectively. The solid line is a line of best fit determined visually.

It is evident that the two bonds in which the hydrogen atom is free, α and β , overlap completely and are responsible for the sharp higher frequency peak in the O-H stretching band. Together, these two bonds constitute all the free hydroxyl protons in solution. This is convenient since the key parameter required for the thermodynamic modeling of alcohol + inert systems is the fraction of the free O-H protons. Moreover, there is significant disagreement in the literature concerning the α - β overlap in IR spectra. Several authors have assumed that β bond vibrations do not contribute significantly to the sharp free O-H proton peak, instead allocating it entirely/predominantly to the α vibration [184–186]. In these cases, it is assumed that most clusters in solution are in cyclic form (resulting in few β bonds), that the β peak occurs at different frequency altogether, or that the intensity of the β absorbance is significantly less than that of the α . However, the results of our QM calculations are consistent with more recent work in this area [2,150]. While there is a redshift in the β vibration due to charge redistribution when an oxygen atom accepts another proton, this shift is very slight. This indicates that the electronic structure is only weakly affected when an oxygen in a hydroxyl hydrogen bonds. The next observation is that the γ bonds appear at a lower frequency and have greater integrated absorption coefficients, (approximately 3-10 times that of the α and β bonds). The δ bonds follow the same pattern with respect to γ bonds with integrated absorption coefficients that are 3-20 times that of the α and β bonds). The frequency trend is easily explained by recognizing that, as the hydroxyl protons become more "shared" due to hydrogen bonding, the covalent bond is weakened. As a result, the potential energy well in which the hydrogen in moving is effectively broadened, causing the vibrational frequency of the hydroxyl to red shift. The increase in intensity is caused by an increase of the dipole moment due to charge distribution that occurs through the hydrogen bond as the proton vibrates between the two oxygen atoms.

The most valuable finding of this work and, in fact, the most obvious is that the relationship between the integrated absorption coefficient of an O-H bond and its vibrational wavenumber follows a curve that is independent of the bond category. The relationship can be described well by an inverted Gaussian function of the form:

$$\psi = \tilde{a} - \tilde{b} \exp\left(-\left(\frac{(\tilde{v} - \tilde{c})^2}{2\tilde{d}^2}\right)\right) = 1500 - 1455 \exp\left(-\left(\frac{(\tilde{v} - 3597)^2}{2(200)^2}\right)\right)$$
 6.5

where the parameters \tilde{a} , \tilde{b} , \tilde{c} and \tilde{d} are visually determined. This fit is plotted as a black line in Figure 6.5. It is worth noting that a polynomial function was insufficient in capturing the curve characteristics well. Previous studies in this area have recognized patterns in vibrational characteristics. For example, as early as 1956, Huggins and Pimentel [187], using notation X-H ... Y to indicate that a variety of electronegative groups, found similarly interesting patterns for a wide range of hydrogen-bonding systems. They show that a monotonic relationship between the vibrational X-H frequency and ΔB , which is defined as the difference between the intensity in the solvent minus the intensity in chloroform, a weak hydrogen bond donor. More recently, Mesele and Thompson [168] conducted DFT calculations on neat alcohols and showed that the empirical maps that relate transition frequencies, position matrix elements, dipole derivatives and the electric field are surprisingly linear. Moreover, these relationships were identical for all 4 primary alcohols tested and were predicted to hold to all other alcohols. Having independently uncovered this pattern, we further explore the applicability of Figure 6.5 and Eq. (6.5) by repeating the described procedure for an equimolar mixture of ethanol and cyclohexane at 298.15 K. The resulting plot of vibrational frequencies and integrated extinction coefficient overlaps completely with data in Figure 6.5 as shown in Figure 6.6 for all bond types.

Further, we also include the results of Murdoch *et al.*'s work [144] in which the vibrational behavior of ethanol clusters from monomers to hexamers was calculated with DFT methods. The clusters were optimized in vacuum and the effects of conformation (anti and gauche) were also investigated. In Figure 6.6, the linear species are included as filled squares. For the cyclic species, we noted that, while there was a wide scatter in the intensities of the bond vibrations, each high intensity vibration was coupled with a low intensity vibration. From an experimental perspective, an IR experiment will collect the average intensity. Thus, we averaged the values of frequency and integrated extinction coefficient for all the molecules in each cluster and included them in Figure 6.6 as filled triangles. The fit of these calculations with our own is striking. Though untested thus far, Figure 6.6 is a promising indication that the inverted Gaussian fit may be representative for other alcohols.



Figure 6.6: Hydroxyl stretching frequencies and integrated absorption coefficients for other systems calculated from QM simulations. Unfilled diamond, triangle, square and circle markers denote α , β , γ , and δ bond vibrations of ethanol + cyclohexane respectively (this work). Filled small circles are butanol + cyclohexane data (this work). Filled squares and triangles denote linear and averaged cyclic ethanol oligomers respectively (Murdoch *et al.* [144]). The solid line is a line of best fit determined visually.

6.4.3. Scaling and further analysis of experimental IR spectra

To quantify the number of bonds in the hydroxyl band in Figure 6.3, we revisit the Beer-Lambert law and use the integrated extinction coefficient, ψ in place of ε . At the wavenumber of every data point, a value for the ψ is calculated and used to scale the absorbance. In addition to this, the wavenumber quantities found using QM calculations must be applied carefully when interpreting experimental data. It should be recognized that relative relationships between *ab initio* calculated IR features are more reliable than absolute values. Therefore, we adjust the inverted Gaussian function developed in Eq. (6.5) such that the wavenumber of the minimum, which is decided by parameter \tilde{c} , matches the free-end peak center in the experimental spectra. Therefore, we set \tilde{c} to 3645 cm⁻¹ and ψ is calculated according to:

$$\tilde{A}(\tilde{v}) = \frac{A(\tilde{v})}{\psi(\tilde{v})l} = \frac{A(\tilde{v})/l}{1500 - 1455 \exp\left(-\left(\frac{(\tilde{v} - 3645)^2}{2(200)^2}\right)\right)}$$
6.6

where \tilde{A} represents the scaled absorbance at wavenumber $\tilde{\nu}$ and represents a concentration response by the system to IR irradiation. Plotting the resulting scaled spectra yields Figure 6.7 which is visually very different from the original spectra in Figure 6.3. Peaks P_1 and P_2 now dominate in intensity while P_3 is diminished considerably.



Figure 6.7: Scaled absorbance spectra for *n*-butanol + cyclohexane at $x_{BuOH} = 0.1$. Spectra are scaled with the integrated absorption coefficient function

Calculating the area under the new scaled spectra and comparing the results with the experimental mole fraction provides Figure 6.8, which is analogous to Figure 6.4.



Figure 6.8: Scaled absorbance band area of the O-H band for *n*-butanol + cyclohexane data as a function of the mole fraction of 4 experimental compositions. Spectra were collected at $x_{BuOH} = 0.0469, 0.0678, 0.0817$ and 0.100 and the line indicate the line of best fit given by the equations on the figure. A thermal correction for density is applied.

The relationship between the band area and the apparent mole fraction is corrected significantly. First, at a given composition (and correcting for temperature effects on density due to thermal expansion), the integrated concentration response of the system is remarkably close regardless of temperature for all the compositions studied. This is especially clear when the sample standard deviation between temperatures is compared as in Table 6.4. The standard deviation is calculated using the equation:

$$stddev = \sqrt{\frac{\sum(Area - Mean Area)^2}{s - 1}}$$
 6.7

where s is 5, the number of temperature data points at each composition.

	T(°C)	Absorbance Band, A		Scaled Absorbance Band, \widetilde{A}	
x _{BuOH}		Area	Area Std Dev	Area	Area Std Dev
	30	94.94		17.73	1.0013
	40	82.05		17.74	
0.0469	50	68.35	22.35	17.54	
	60	53.85		16.98	
	70	38.40		15.37	
	30	154.1		26.78	0.4453
	40	138.1		27.08	
0.0678	50	120.2	27.91	26.82	
	60	102.0		26.52	
	70	83.99		25.91	
	30	187.1	32.48	30.69	0.5723
	40	168.9		31.11	
0.0817	50	149.2		30.87	
	60	126.7		30.16	
	70	105.6		29.69	
	30	229.9		35.04	
	40	209.2	34.68	35.37	
0.1	50	186.3		35.27	0.3828
	60	165.1		35.97	
	70	142.3		35.80	

Table 6.4: Areas and standard deviations for the absorbance and scaled absorbance bands

Figure 6.8 demonstrates that, with the scaled spectra, the Beer-Lambert law can be applied successfully to the entire O-H band permitting quantification of the hydroxyl band to determine the alcohol concentration in solution. The functional form of the inverted Gaussian function developed in Eq. (6.5) results adequate scaling. To our knowledge, no other work has performed a comparable quantitative analysis of the O-H band due to its known complexity. The work presented here provides a powerful tool that maps the band characteristics to the concentration in solution, effectively placing the absorptions of all O-H sites, regardless of their context, on a quantitatively equal footing.

6.4.4. Universality of spectroscopic characteristics

An interesting and important universality emerges from the QM calculations of various properties demonstrated here for an equimolar mixture of ethanol + cyclohexane at T = 298.15 K. The same computational technique described in Section 6.2 were carried out on the system. Figure 6.9 shows how the length of the optimized O-H bond length correlates with its vibrational frequency.



Figure 6.9: Relationship between the covalent O-H bond length and vibrational frequency calculated for equimolar ethanol + cyclohexane mixture. Diamond, triangle, square and circle markers denote α , β , γ , and δ bond vibrations of ethanol + cyclohexane respectively

The relationship is described extremely well by a linear function of the form:

$$L_{OH} = -5.764 \times 10^{-5} \tilde{\nu} + 1.178$$

where L_{OH} is the length of the hydroxyl bond in Angstroms and $\tilde{\nu}$ is the wavenumber in cm⁻¹.

Furthermore, the IR vibration also shows a distinct relationship with the NMR chemical shift of

the hydroxyl calculated with Gaussian09 as shown in Figure 6.10.



Figure 6.10: Comparison of calculated NMR and IR characteristics for equimolar ethanol + cyclohexane mixture. Diamond, triangle, square and circle markers denote α , β , γ , and δ bond vibrations of ethanol + cyclohexane respectively

While some relationship between these different molecular characteristics is expected, the direct correlation is striking. Hydrogen bonding is often considered to be a complex phenomenon with random effects on mixture properties. These figures show that, with careful implementation and analysis, computational simulations have the power to create models that relate the bond length, vibrational behavior and NMR response. These universal "maps" allow for the prediction of one of these elements from the others and can have far-reaching impacts on the analysis of the molecular behavior of alcohols.

The visual fit of Eq. 6.5 may require refinement as more data become available at lower concentrations, or with other alcohols. Typically, *ab initio* calculations at the level of theory implemented here would be expected to require additional scaling beyond the frequency shift that is implemented. However, with the currently available data, such scaling is difficult to justify

because the current constants provide scaled areas with excellent linear correlation to experimental data.

In future work, we aim to leverage these maps to gain quantitative insight into the speciation that occurs due to hydrogen bonding. It will be important to study the behavior of cyclic species and include them in model development. Our preliminary work with Murdoch [144]'s data for cyclic species in Figure 6.6 indicates that bonds in cyclic clusters behave very similarly to δ bonds. For this reason, we expect that they will also be well-represented with the QM correlation developed here. However, the behavior of η and ζ bonds is less easy to anticipate without further calculations. Therefore, future directions will explore the infrared responses of these bonds. Finally, we aim to calculate the values of association parameters for a thermodynamic model developed in an earlier work [96] from spectroscopic findings.

6.5. Conclusions

In this work, the hydroxyl vibrational band is analyzed quantitatively for alcohol + inert mixtures using insights from QM simulations. First, MD simulations were carried out to generate sample environments around O-H bonds. Then, selected clusters were further analyzed with QM calculations to gain an understanding of their spectroscopic characteristics. It was found that the integrated absorption coefficient and vibrational frequency of a group can be related by an inverted Gaussian function. Through a variation of the Beer-Lambert law, this function can be used to calculate the concentration of hydroxyl groups in solution from IR spectra. Finally, universal "maps" between the O-H bond length, its vibrational frequency, integrated absorption coefficient and NMR chemical shift were developed.

Chapter 7. Integration of Quantum Calculations and Spectroscopy for Wertheim Alcohol Association

In this work, we provide preliminary calculations and results for the derivation of Wertheim association parameters from spectroscopic findings. IR and NMR spectra for mixtures of *n*-butanol + cyclohexane mixtures were collected and analyzed using a combination of thermodynamic modeling and insight from MD+QM calculations. While some results are discussed, the challenges encountered herein lie beyond the scope of this dissertation. As such, recommendations are given for future directions that will help advance the current work and address its limitations.

7.1. Introduction

7.1.1. Hydrogen bonding and infrared spectroscopy

Several analytical tools have been applied to further the understanding and modeling of hydrogen bonding including Raman scattering [153,188–190], neutron diffraction [191–193] and NMR spectroscopy [13,14,194]. Infrared spectroscopy [146,150,169,195–198] has been especially instrumental in uncovering the behavior of molecules that hydrogen bond. There is an established understanding that the effects of hydrogen bonding are most evident in the changes occurring in the O-H vibration region. For this reason, research is often focused on investigating the effects of various conditions, such as temperature and pressure, on the characteristics of the hydroxyl vibrational band [146]. As hydrogen bonded cluster sizes increase, the high frequency O-H vibrational peak decreases and a second broad O-H peak increases at lower frequencies [144,150,154]. The spectrum is fitted with curves to represent the experimental data. The contributions of the hydroxyl bonds to the band is known in general terms [199]. However, the curve fitting procedure has historically been conducted without recognizing the importance of variations in extinction coefficient with wavenumber. Deconvolution typically uses several symmetrical peaks that are assigned to specific bonds or species based on the absorbance dependence on temperature or pressure. As indicated by Barlow [200], a common challenge with this approach is that the problem is under-constrained; comparable fits can be obtained from several different interpretations depending on the initial guess.

Computational simulations are increasingly being leveraged to lend more insight to hydrogen bonding in alcohols [17,139,154,155,160] and its effect on their IR vibrations. Studies have included molecular dynamics (MD) [153], quantum mechanical (QM) [139,144,154–159] and hybrid multi-level (QM/MD) simulations [17,160–162]. While MD simulations are less computationally expensive than QM calculations, they neglect quantum effects which are important in hydrogen bonding. Another shortcoming of MD was highlighted by Kwac and Geva [17]. Using simulations of CH₃OD in carbon tetrachloride, they found that the relative populations of hydrogen bonded clusters calculated are dependent on the harmonicity, polarizability and damping of the empirical MD forcefield. The relative populations of the different bonds varied by up to 20% percent depending on the choice of forcefield. Furthermore, in their MD simulations with ethanol + water mixtures, Gereben *et al.* [201,202] were not able to identify a forcefield model that represented experimental scattering X-ray structure factors well across the composition range.

In Chapter 6, we developed a sequential MD + QM method to calculate the IR vibrational frequencies and integrated absorption coefficients of different types of O-H bonds. These were classified according to how their atoms participate in hydrogen bonding as shown in Figure 7.1.



Figure 7.1: Types of covalent OH bonds

Though all the bonds in Figure 7.1 occur in solution, it is often assumed, especially in chemical engineering thermodynamics, that η and ζ bonds are formed in minimal quantities and are thus ignored. Moreover, cyclic species are usually neglected. Gereben *et al.* recently extended their work with ethanol + water to study cluster formations [203] and found that 60% (for an 80 mol% ethanol in water mixture) to 95% (pure water) of hydrogen bonded molecules were bonded in cyclic clusters. Considering this, the assumption of no cyclic species is difficult to justify. However, the thermodynamic models which use this assumption are often capable of representing experimental data reasonably well [14,68,173]. Therefore, as a first approximation, we ignore cyclic species in this work and defer their analysis to future work.

The work covered in Chapter 6 aimed to quantify concentration of alcohol molecules in an alcohol + alkane mixture using the infrared O-H vibration. To this end, QM calculations were conducted to find relationships between several spectroscopic properties including the vibrational frequency of an O-H bond and NMR shift of its proton. A key finding was the correlation shown in Figure

6.6 (and repeated in Figure 7.2 for the reader's convenience) between the integrated absorption coefficient, ψ , and the wavenumber of O-H vibrations.





It was shown that the relationship is very well modeled by an inverted Gaussian function given by:

$$\psi = \tilde{a} - \tilde{b} \exp\left(-\left(\frac{(\tilde{v} - \tilde{c})^2}{2\tilde{d}^2}\right)\right) = 1500 - 1455 \exp\left(-\left(\frac{(\tilde{v} - 3597)^2}{2(200)^2}\right)\right)$$
 7.1

When used to analyze the experimental spectra, the minimum of the Gaussian function, where α and β bonds fall, was shifted to match the free-end peak wavenumber in the spectrum by adjusting parameter \tilde{c} to 3645 cm⁻¹. The quantity ψ was used with the Beer-Lambert law to calculate a "scaled" absorbance that is directly related to the total O-H concentration. In this work, we extend this method to calculate the concentrations of α , β , γ and δ bonds. Further, this cluster distribution

is used to calculate association (or hydrogen bonding) parameters for thermodynamic modeling of the alcohol + alkane systems.

7.1.2. Thermodynamic modeling of hydrogen bonding

Capturing the effects of hydrogen bonding is important for the thermodynamic modeling of many systems encountered in the chemical industry. In our previous work [96], we developed a model that calculates deviations of the real solution from an ideal solution reference straight via an activity coefficient term, γ , which is unity for an ideal solution and non-unity otherwise. Deviations due to hydrogen bonding were calculated using Wertheim's perturbation theory [47–50], a statistical mechanics approach which assigns sites on molecules which act as hydrogen bond acceptors and/or donors. A common scheme used for alcohols is the "2B" scheme in which the oxygen is designated as an acceptor site and the hydrogen as the donor site. Mathematically, this behavior is captured through:

$$\ln \gamma_k^{assoc} = \sum_{\substack{sites \ B_i \\ on \ k}} N_{B_i} \left[\ln \left(\frac{X^{B_i}}{X^{B_i,0}} \right) - \frac{1}{2} (1 - X^{B_i,0}) \right] + \frac{1}{2} \frac{\rho}{\rho_k} \sum_{\substack{all \\ sites \ B_i}} x_{B_i \ host} N_{B_i} (1 - X^{B_i})$$
 7.2

where X^{B_i} and $X^{B_i,0}$ are the fraction of sites B_i which remain free (not bonded) in the mixture and when the site host is pure, respectively. Variable $x_{B_i host}$ is the apparent mole fraction of the molecule that hosts site B_i . The number of identical B_i sites on a host is given by N_{B_i} . For example, the two hydrogen atoms in water are modeled as identical sites due to the symmetry of their host molecule and the assumption that they will interact in the same way with their microenvironments, independent of bonding order.

Careful consideration on the definition of the term X^{B_i} in Wertheim's theory, given in Eq. (7.3), reveals a physical significance relevant to IR spectroscopy.

$$X^{B_i} = \frac{n^{B_i, free}}{n^{B_i, total}}$$

$$7.3$$

Here, $n^{B_i,free}$ and $n^{B_i,total}$ are the moles of sites with free B_i sites and the moles of total B_i sites respectively. For alcohols and alcohol + inert (non-hydrogen bonding component) mixtures, X^{B_i} represents two sites simultaneously, the fraction of free acceptor sites (oxygen atoms), X^A , and the fraction of free donor sites (hydrogen atoms), X^D , at equilibrium. Because each hydrogen bond involves one acceptor and one donor site, the material balance when the alcohol is the only associating species yields $X^A = X^D$. Coupling this with Eq. (7.3) yields:

$$X^{A} = X^{D} = \frac{n^{D, free}}{n^{D, total}} = \frac{C_{\alpha+\beta}}{C_{alc}}$$

$$7.4$$

where $C_{\alpha+\beta}$ is the concentration of α and β bonds and C_{alc} is the apparent concentration of the alcohol. Moreover, Wertheim's theory includes a function that relates the fraction of free sites to an association parameter, Δ^{AD} . For the case considered here where the alcohol is dissolved in a non-associating solvent:

$$X^{A} = 2/\left(1 + \sqrt{1 + 4x_{alc}\rho\Delta^{AD}}\right)$$

$$7.5$$

where x_{alc} is the mole fraction of alcohol in the mixture and ρ is the solution molar density. Therefore, by using Eq. (7.4) and Eq. (1.4), it is possible to regress spectroscopic measurements of X^A to find the value of the association parameter Δ^{AD} . The generalized model and form of Eq. (3.7) are provided in our previous work [96].

Wertheim's theory has been incorporated into several thermodynamic equation of state models including the cubic-plus-association (CPA) [11,52] and the statistical associating fluid theory (SAFT) family of equations [9,54–56,95]. In these models, the association parameter is normally found by regressing macroscopic experimental data such as phase equilibria data, vapor pressures

and/or saturated liquid densities. However, as shown in Figure 5.37 in Chapter 5, these methods can yield values of Δ^{AD} that differ by factors of up to four to five depending on the data used for the fitting. This discrepancy is greater than the expected dependence on solvent. An alternative approach is to derive association parameter values from data collected using fundamental tools such as IR and NMR spectroscopy. Studies that leverage these tools often calculate a concentration-based equilibrium constant, K_C (or a directly related form of the equilibrium constant) instead of Δ^{AD} . In these cases, association is modeled using chemical theory which is a distinct form but, if applied with reasonable assumptions, equivalent to Wertheim's theory [58,96]. 7.1.3. Motivation

In this work, we extend the quantitative MD + QM analysis methods developed in Chapter 6 to calculate the concentrations of different types of O-H bonds in solution. Further, the aim is to describe the methodology and provide preliminary results for the calculation of association parameters for the thermodynamic activity coefficient model in Eq. (3.7). This primary method will involve interpretation of the infrared spectra of *n*-butanol + cyclohexane mixtures. Nuclear magnetic resonance (NMR) spectroscopy is used as a complementary tool to confirm findings from IR.

7.2. Experimental methods

The complete methodology for the computational and IR studies is identical to that in Chapter 6 and details can be found in Sections 6.2 and 6.3. For the NMR experiments, two protocols were used. In the first, we followed Karachewski's [98] NMR procedure closely and collected data at T=34 °C. Samples were prepared volumetrically in a glove bag under a continuous flow of nitrogen gas. Anhydrous *n*-butanol, cyclohexane and deuterated cyclohexane at purities of 99.8, 99.5% and 99% respectively, were purchased from Sigma-Aldrich and used with no further purification. For
dilute samples with an alcohol mole fraction less than 0.03, the solvent was completely replaced with deuterated cyclohexane (Sigma Aldrich– isotopic purity \geq 99.6 atom % D) for the purpose of diminishing the solvent proton shifts such that no significant overlap occurs with the alcohol O-H proton shift. For samples at compositions greater than 3 mol% *n*-butanol, a coaxial tube system was used in which the inner tube contained the prepared sample and the outer tube contained deuterated cyclohexane (for locking) and TMS (as an internal reference). For dilute samples, the outer tube was left empty and the inner tube contained the sample with 0.5 mole% TMS. Data was obtained using a Varian 600 MHz superconducting NMR Spectrometer and the temperature was calibrated with ethylene glycol.

For data collected at other temperatures, a very similar procedure was used but the coaxial tube system was replaced with a simpler single tube configuration and all samples were prepared using deuterated cyclohexane as the solvent. In these cases, data was obtained using a Varian 500 MHz NMR Spectrometer.

7.3. Results and discussions

In IR spectroscopy, the consequences of hydrogen bonding are most evident in the region of the O-H stretching vibration, in which the bands combine and appear as two peaks between ~3200 and 3700 cm⁻¹. As mentioned in Chapter 6, the raw spectra are processed to remove solvent bands and the O-H bands are determined to begin and end at local minima on either side of the two peaks. Furthermore, the method developed earlier to create more physically significant spectra by scaling the absorbance with the integrated absorption coefficient function in Figure 7.2 is repeated here. However, in this chapter, the integration ranges determined for each individual spectrum by finding the minima on either end of the hydroxyl band. These ranges are provided in Table 7.1.

Table 7.1: Frequency range defined as hydroxyl band for each spectrum

158

24	Temperature (°C)					
х BuOH	30	40	50	60	70	
0.0469	3055 - 3750	3064 - 3752	3091 - 3750	3092 - 3750	3153 - 3752	
0.0678	3048 - 3751	3053 - 3751	3067 - 3751	3076 - 3752	3089 - 3752	
0.0817	3050 - 3751	3058 - 3751	3066 - 3751	3076 - 3751	3088 - 3751	
0.100	3049 - 3752	3050 - 3755	3056 - 3752	3066 - 3752	3078 - 3752	

Although the integrated ranges are not consistent across all temperatures and compositions, the difference between the integrated areas calculated using this method compared to that in Chapter 6 (in which a fixed range was integrated for all spectra) is only 1.2%. In the future, we plan to conduct the analysis described in this chapter over a fixed wavenumber range for a wider variety of systems and conditions. However, the minor change in the integration limits would not affect the preliminary results provided here.

Figures 7.3-7.6 show the raw and scaled spectra for *n*-butanol + cyclohexane at four concentrations. There are some features of the O-H vibrational band that are well-documented which are evident in the raw spectra. First, the band consists of two peaks: a sharp, high-frequency peak and a broad, lower-frequency peak. The former peak is often attributed to free end group protons, α and β bonds [146,150,153,154,204], though there is some disagreement on whether β contributes significantly to this peak [185,186]. As the temperature is increased, the sharp free end peak increases in absorbance while the broad O-H peak diminishes. This is evidence of shifting bond distributions as hydrogen bonds break at higher temperature [144,146] and more monomers form in the solution. Moreover, as the proton in a bond becomes more tethered to others by hydrogen bonding, its vibrations occur at lower frequencies and higher intensities. This is evident in our own work in Figure 7.2 as well as that of others [153,205,206].



Figure 7.3: Raw (top) and scaled (bottom) O-H IR band for a 0.0469 mole fraction of *n***-butanol in cyclohexane.** Data was collected at five temperatures between 30 °C and 70 °C at a pathlength of 0.01092 cm.



Figure 7.4: Raw (top) and scaled (bottom) O-H IR band for a 0.0678 mole fraction of *n***-butanol in cyclohexane.** Data was collected at five temperatures between 30 °C and 70 °C at a pathlength of 0.01092 cm



Figure 7.5: Raw (top) and scaled (bottom) O-H IR band for a 0.0817 mole fraction of *n***-butanol in cyclohexane.** Data was collected at five temperatures between 30 °C and 70 °C at a pathlength of 0.01092 cm.



Figure 7.6: Raw (top) and scaled (bottom) O-H IR band for a 0.100 mole fraction of *n***-butanol in cyclohexane.** Data was collected at five temperatures between 30 °C and 70 °C at a pathlength of 0.01092 cm.

7.3.1. IR band assignments

It has been recognized that IR vibrational bands are more accurately attributed to bonds [147,150,172,207], as first proposed by Hall and Wood [170], rather than to entire hydrogen bonded clusters. For this reason, we begin peak assignments by studying how the α , β , γ and δ bonds vibrate in the QM simulations. The number of the different species and bond types analyzed to generate the QM calculated distributions in this work is given in Table 7.2.

Table 7.2: Number of each species and bond analyzed with QM calculations for the distribution plot in Figure 7.7 and Figure 7.8. All species larger than a monomer have one β and one γ bond. Trimers and tetramers also possess one and two δ bonds respectively.

		Ethanol + Cyclohexane	<i>n</i> -Butanol + Cyclohexane
		$x_{EtOH} = 0.5$	$x_{BuOH} = 0.1$ and 0.5
Species	Monomer	859	1019
	Dimer	413	383
	Trimer	236	105
	Tetramer	223	70
Bonds	α	859	1019
	β	872	558
	δ	682	245
	γ	872	558

Figure 7.7 shows the distribution of vibrational frequencies of α , β , γ and δ bonds calculated through QM simulations of ethanol + and *n*-butanol + cyclohexane mixtures. Findings from the two systems are plotted together because separate plots showed no discernable difference between them. To create the figure, the vibrations of each bond type were binned in 5 cm⁻¹ wide bins and normalized, and the resulting lines were smoothed with a 7-bin moving average and normalized to display the relative populations for each bond type as a function of wavenumber. These processing steps were carried out to refine the shape of the distributions and display relative information about

them more clearly. It should be emphasized that only semi-quantitative information should be obtained from this figure from the peak shapes. The peak heights should not be compared between bond types because the data are insufficient to reflect the distribution of bond types at the simulated concentrations. Instead, the value of Figure 7.7 lies in its ability to reveal approximate peak shapes and a qualitative description of the vibrational frequencies of different bond types relative to one another.

In agreement with other studies [146,150,153,154,204], Figure 7.7 shows that α and β bonds vibrate at the same frequency and show significant overlap. This confirms that the β bond contribution to the sharp high-frequency O-H peak should not be ignored. Moreover, the QM calculations show that some of α and β hydroxyls are found at frequencies red-shifted from the high frequency peak. This finding is noteworthy as most previous work has considered only the high frequency peak to represent all the α and β hydroxyls [150,153].



Figure 7.7: Smoothed normalized distributions of bond types from QM simulations of *n*-butanol + and ethanol + cyclohexane mixtures. Vibrational frequencies of α (n=1878), β (n=1430), γ (n=1430), and δ (n=927) bonds are binned in 5 wavenumber bins and smoothed with a 7-bin moving average



Figure 7.8: Smoothed normalized distributions of bond types from QM simulations with γ from dimers separated from others. Vibrational frequencies of α (n=1878), β (n=1430), γ_{Dim} (n=796), γ_{Nmer} (n=634) and δ (n=927) bonds are binned in 5 wavenumber bins and smoothed with a 7-bin moving average

Another interesting observation is made when the vibrations of the γ bond on dimer clusters, γ_{Dim} is separated from that of γ bonds on larger oligomers, γ_{Nmer} . As shown in Figure 7.8, γ_{Dim} vibrations are blue-shifted by about 50 cm⁻¹ compared to γ_{Nmer} . Interestingly, β bonds on dimers display no similar shift compared to those on larger clusters. We hypothesize that the reason for the shift in γ is the small size of the dimer molecule. This causes the frequencies of protons in the γ_{Dim} bonds to be higher as they have characteristics that are a hybrid of those of the β hydrogen it is attached to and those of γ bonds on larger clusters, γ_{Nmer} . For species larger than dimers, the γ_{Nmer} proton is separated from β proton by one or more δ bonds, so the vibrational mixing lowers its frequency. To our knowledge, this is the first time this distinction has been made and we use this finding to guide the peak fitting and assignment procedure described in the next section.

7.3.2. IR peak fitting

Quantitative fitting of infrared spectroscopy requires the overlap between constituent peaks to be accurately determined [208]. As alcohol + inert mixtures become more concentrated in alcohol, extensive hydrogen bonding causes the broad O-H peak to have more and more overlap with the free end peak. Therefore, as indicated by Wandschneider *et al.* [155], it becomes difficult to calculate the concentration of free ends when the mole fraction of alcohols exceeds 0.3. For this reason, we limit our study to the mixtures shown in the spectra in Figures 7.3 - 7.6 where the alcohol mole fractions are below 0.100.

Based on the line shapes found from the QM calculated vibrational frequency distributions, we fit each scaled experimental IR spectrum with 6 Gaussian peaks: one for α , β (visually determined to require 2 peaks), γ_{Dim} , γ_{Nmer} and δ . Each Gaussian peak, P_i is given by the function:

$$P_{i} = \tilde{f}_{i} \exp\left(-\left(\frac{(\tilde{\nu} - \tilde{g}_{i})}{\tilde{h}_{i}}\right)^{2}\right)$$
7.6

where \tilde{f}_i , \tilde{g}_i and \tilde{h}_i are fitted parameters. To focus the search for correct parameters, the regression is constrained such that the concentration of β bonds is equal to that of γ bonds. This physical constraint holds as long as no branched chains exist in solution as is assumed in our work. The area under a Gaussian curve is given by $\tilde{f}_i \tilde{h}_i \sqrt{\pi}$,

$$\int P_{\beta 1} d\nu + \int P_{\beta 2} d\nu = \int P_{\gamma_{Dim}} d\nu + \int P_{\gamma_{Nmer}} d\tilde{\nu}$$

$$7.7$$

$$\tilde{f}_{\beta_1}\tilde{h}_{\beta_1} + \tilde{f}_{\beta_2}\tilde{h}_{\beta_2} = \tilde{f}_{\gamma_{Dim}}\tilde{h}_{\gamma_{Dim}} + \tilde{f}_{\gamma_{Nmer}}\tilde{h}_{\gamma_{Nmer}}$$

$$7.8$$

Initial guesses for the wavenumbers of the peak apexes were determined from the shifts of the β , γ_{Dim} , γ_{Nmer} and δ peaks in Figure 7.8 relative to the α peak. The upper and lower bounds set for



the wavenumbers was ± 80 cm⁻¹ from the initial guesses. Figures 7.9-7.12 show the results of the fitting to the spectra at the four lower temperatures for each of three concentrations studied here.

Figure 7.9: Peaks fitted to scaled IR spectra for *n*-butanol + cyclohexane mixture at x_{Bu0H} =0.0469. Figures a, b, c, d and e indicate data at T= 30, 40, 50, 60, 70 °C respectively



Figure 7.10: Peaks fitted to scaled IR spectra for *n*-butanol + cyclohexane mixture at x_{BuOH} =0.0678. Figures a, b, c, d and e indicate data at T= 30, 40, 50, 60, 70 °C respectively



Figure 7.11: Peaks fitted to scaled IR spectra for *n*-butanol + cyclohexane mixture at x_{BuOH} =0.0817. Figures a, b, c, d and e indicate data at T= 30, 40, 50, 60, 70 °C respectively



Figure 7.12: Peaks fitted to scaled IR spectra for *n*-butanol + cyclohexane mixture at x_{BuOH} =0.100. Figures a, b, c, d and e indicate data at T= 30, 40, 50, 60, 70 °C respectively

The concentration of each of the bond types is calculated by multiplying the areas under the corresponding peaks by the constant relating scaled peak areas to apparent concentrations given in Figure 7.13. The scaling factor is 1/37.47 = 0.0290.



Figure 7.13: Scaled band area as a function of apparent *n***-butanol concentration**. The equation for the line of best fit provides a convenient method of relating the calculated apparent concentrations from the scaled band area.

The hydroxyl species distributions calculated from the scaling factor are given in shown in Figure 7.14. In plots were a gray bar appears, the fitted spectrum was not able to fit the experimental concentration exactly and therefore there was an "unassigned" portion. Conversely, plots in which the bars exceed unity indicate that the fitted peaks over predict the prepared concentration value. These slight discrepancies are especially prominent in the plots for the two most dilute sample and are likely due to experimental error in preparing the samples. The exception to this, though, is the fit of the data at T=70 °C and $x_{BuOH} = 0.0469$ which has an unassigned percentage of 13.1%. The peak shape for the raw spectra in Figure 7.3 has unexpected features that are indicative of a problem with the experiment, such as evaporation.





In general, however, the plots show several encouraging trends. First, the percent of α bonds increases with temperature by an average of 16.1% from 30 °C to 70 °C. This is consistent with the fact that more hydrogen bonds break as temperature is increased, converting large clusters to smaller oligomers and monomers. For the same reason, the fractions of δ bonds should and, for the most part, do decrease with increasing temperatures. The trends for β and γ are less predictable because their concentrations depend on the rate at which larger clusters break down to form smaller clusters (which increases the number of β and γ bonds) compared to the rate of dimers breaking to form monomers (which decreases the number of β and γ bonds). However, from Figure 7.14, it is observed that the $\beta + \gamma$ concentrations decrease slightly but consistently for the three most dilute samples, suggesting that the rate of large cluster disintegration to smaller (non-monomer) species is greater. For the highest concentration, the $\beta + \gamma$ concentrations decrease until the T = 50 °C point then increases again. From this, it may be deduced that the two rates balance at approximately 50 °C at this concentration.

An interesting trend is noted when the concentration of α bonds is plotted as a function of *n*-butanol mole fraction (see Figure 7.15):



Figure 7.15: Concentration of α bonds vs. mole fraction of *n*-butanol.

There is a linear decrease at all temperatures except for the data point at T=70 °C and x_{BuOH} =

0.0469. This is further evidence that the data point is flawed and should be measured again and

refitted. However, the relationship shown in the figure for the other data points is striking.

To calculate the association parameters in the following section, the concentration of free ends must be known. Therefore, the calculated values for $C_{\alpha+\beta}$ are given in Table 7.3.

Table 7.3: Concentration of free ends $(\alpha + \beta)$ in mol/mL calculated from fitted Gaussian peaks

**			T (° C)		
X	30	40	50	60	70
0.0469	2.16E-04	2.36E-04	2.51E-04	2.63E-04	2.43E-04
0.0678	3.25E-04	3.51E-04	3.70E-04	3.92E-04	4.03E-04
0.0817	3.65E-04	3.95E-04	4.14E-04	4.31E-04	4.50E-04
0.1001	4.04E-04	4.37E-04	4.61E-04	4.96E-04	5.20E-04

7.3.3. IR calculation of association parameters

To extract quantitative information from IR spectroscopy, it is necessary to assume some form of model for the species in solution and assign bands accordingly. Recently, Chen *et al.* [209] calculated the dimerization equilibrium constant for 3-ethyl-2-methyl-3-pentanol in tetrachloroethylene for quantitative analysis of infrared spectra. This system was intentionally chosen for the study because steric hindrance was thought to limit the extent of association to dimers. The analysis was consistent with this notion, but simple computer simulations point to little trouble going from dimers to linear and cyclic trimers. Such limited hydrogen bonding systems are, of course, unusual among alcohols which hydrogen bond extensively. For association of alcohols, models and band assignments vary widely in literature. A common model for alcohol + inert systems assumes that chains of infinite sizes can form continuously in solution. Such models have been used with some success to interpret associating systems [14,37,68,89,210–212].

In contrast, Asprion *et al.* [7,137,138] published several papers which assume only monomer, dimer and one other *n*-mer formation (1-2-*n*) where *n* was a fitted parameter. The work supplemented an existing thermodynamic model, UNIQUAC, with chemical theory and derived a value of K_c from infrared spectra. Many other studies obtain best fits of their experimental data by assuming at least one cyclic species (most often a cyclic trimer [150,213] or tetramer [144,185,186,214]) is stable enough to exist in appreciable quantities at equilibrium.

In this work, we aim to derive a value for the Wertheim association constant, Δ^{AD} and therefore consider only linear species. First, we regress the values of Δ^{AD} to best fit the calculated free end concentrations in Table 7.3 using Eq. (1.4). The resulting fit is given in Figure 7.16.



Figure 7.16: Calculated and fitted concentrations of free end $(\alpha + \beta)$ hydroxyls. Markers indicate the concentrations derived from IR at T= 30 (crosses), 40 (triangles), 50 (circles), 60 (filled circles), 70 (diamonds) °C respectively.

Then, the Δ^{AD} value is further separated using the form used in CPA:

$$\Delta^{AD} = g(\sigma) K^{AD} \left(\exp\left(\frac{\epsilon^{AD}/k}{T}\right) - 1 \right)$$
7.9

The K^{AD} and ϵ^{AD}/k are regressed to best fit the Δ^{AD} found in the first step. The quantity $g(\sigma)$ in Eq. (7.9) is the radial distribution function defined as:

$$g(\sigma) = \frac{1}{1 - \frac{1.9}{4} b_{CPA} \rho}$$
7.10

where $b_{CPA} = \sum x_i b_{CPA,i}$ is the CPA covolume term and ρ is the mixture apparent molar density. The final fit of Δ^{AD} is given in Figure 7.17.



Figure 7.17: Derived association parameter values from IR analysis (solid line) compared to CPA Δ^{AD} values (dashed line) for the self-association of *n*-butanol. Markers indicate the Δ^{AD} values from fitting the IR measurements.

While we plot the CPA value of Δ^{AD} in the figure as well, it should be recognized that the CPA calculations are not the standard to which other values of Δ^{AD} must be compared. However, it does provide reference values. It can be seen that the association parameter calculated from the IR fits has a weaker temperature dependence with its ϵ^{AD}/k parameter being 18% smaller than that of CPA. This parameter is a measure of the enthalpy of the hydrogen bond and the value calculated

from IR (2065 K) corresponds to an energy value of 4.1 kcal/mol which is in the correct range for hydrogen bonds. The IR-derived K^{AD} parameter is about 12% larger than that of CPA though the physical significance of K^{AD} is less defined. In the next section, we leverage another spectroscopic tool, NMR, for the calculation of association parameters.

7.3.4. NMR calculations and results

A useful complement to IR is proton nuclear magnetic resonance (¹H NMR) spectroscopy. Chemical shifts of proton spins observed in NMR spectra are a result of differences in the electron shielding of hydrogen atoms on a molecule due to their immediate environment. Unlike IR and NIR, the timescale of an NMR measurement is long relative to the lifetime of a hydrogen bond. Therefore, NMR spectra illustrate a time-averaged observation rather than an instantaneous one. However, there are no issues with varying extinction coefficients as there are with IR. Furthermore, with modern equipment, this form of spectroscopy is highly sensitive [98] and accurate to 0.1 ppb [215].

For alcohol + hydrocarbon systems, the work of Gutowsky and Saike [216] is heavily cited for the derivation of the relationship between the observed shift, \tilde{v} , for an O-H site and the weighted average of the shift of the free OH groups and the associated OH groups given by:

$$\tilde{v}' = \frac{C_M}{C_{alc}} \tilde{v}'_M + \sum_{i=2}^{\infty} \left[\left(\frac{C_i^l}{C_{alc}} \right) \tilde{v}'_M + \left((i-1) \frac{C_i^l}{C_{alc}} \right) \tilde{v}'_{Nmer} \right]$$

$$7.11$$

where \tilde{v} , \tilde{v}_M and \tilde{v}'_{Nmer} are the observed chemical shift, the monomer chemical shift and the chemical shift of hydrogen bonded protons respectively. C_i^l is the concentration of linear clusters of size *i*. It is assumed that end-groups and monomers have the same shift and all complexed species (linear and cyclic produce the same shift.

In this section, we derive the values of the self-association Δ^{AD} parameter for *n*-butanol from NMR. To this end, we test out two models: the 1-*K* Continuous Linear Association Model (CLAM) that was implemented by several studies [68,98,173] and a new model that we have developed, called the 2-*K* CLAM. As indicated by its name, the 1-*K* CLAM model neglects non-linear species and assumes that the formation of every hydrogen bond is governed by the same equilibrium constant, *K_C*. While the simplicity of this type of model is appealing, its often insufficient in capturing real behavior [98]. For this reason, Karachewski *et al.* developed several more sophisticated models including the LACT [98] (linear association with cyclic trimer) model and AVEC [13] (association with variable equilibrium constant) models. These models require more fitted parameters and, in the case of the AVEC model, have no closed form. This means that "infinite association possibilities" must instead be approximated by selecting some very large oligomer size as a cutoff. While this is unlikely to affect the physical significance of the model parameters significantly, any *a priori* assumptions that must be made are a disadvantage.

For the 2-*K* CLAM model, we recognize an important phenomenon regarding the energetic and entropic effects that accompany the formation of a dimer. The formation of the first hydrogen bond releases less energy than subsequent bonds which benefit from a "cooperative effect" [155]. When a hydroxyl group is already accepting or donating in a hydrogen bond, the charge redistribution is such that the bond becomes more receptive to another association. As such, it seems justified to include a different K_c for dimer formation than other oligomers. The 2-*K* CLAM does this by:

- 1. Combining a complete mole balance of alcohol molecules with the law of mass action using the equilibrium constant for larger oligomer formation, $K_{C,Nmer}$
- 2. Subtracting the dimer term from the mole balance
- 3. Re-introducing the dimer term with a different equilibrium constant, $K_{C,D}$

Similar to our calculations with IR spectra, we derive the association parameters from NMR spectra by fitting the CPA Δ^{AD} equation in Eq. (7.6). Therefore, the 1-*K* CLAM requires fitting 3 parameters: K^{AD} , ϵ^{AD}/k and the chemical shift of hydrogen-bonded protons, \tilde{v}'_{Nmer} . The monomer shift is obtained experimentally as the shift visible for an extremely dilute *n*-butanol + cyclohexane mixture (x_{BuOH} <0.002). The 2-*K* CLAM model requires fitting 5 parameters: K_D^{AD} , ϵ_D^{AD}/k , K_{Nmer}^{AD} , ϵ_{Nmer}^{AD}/k and \tilde{v}'_{Nmer} .

Figures 7.18-7.20 show the observed chemical shifts relative to the monomer shift for various mole fractions of *n*-butanol in cyclohexane at three temperatures. The black lines show the fits obtained by the 1-*K* and 2-*K* CLAM models when all the parameters, including the ϵ^{AD}/k parameters, are fitted.



Figure 7.18: O-H proton chemical shift as a function of *n*-butanol mole fraction for *n*butanol + cyclohexane mixtures at 26.3 °C. Inset shows a close-up of the dilute region. The markers, dashed line and solid line indicate experimental data, the 1-*K* CLAM model and the 2-*K* CLAM model respectively. The black lines include the ϵ^{AD} parameter in the regression while red lines fix ϵ^{AD} at the CPA value of 2526 K.



Figure 7.19: O-H proton chemical shift as a function of *n*-butanol mole fraction for *n*butanol + cyclohexane mixtures at 34 °C. Inset shows a close-up of the dilute region. The markers, dashed line and solid line indicate experimental data, the 1-*K* CLAM model and the 2-*K* CLAM model respectively. The black lines include the ϵ^{AD} parameter in the regression while red lines fix ϵ^{AD} at the CPA value of 2526 K.



Figure 7.20: O-H proton chemical shift as a function of *n*-butanol mole fraction for *n*-butanol + cyclohexane mixtures at 41 °C. Inset shows a close-up of the dilute region. The markers, dashed line and solid line indicate experimental data, the 1-*K* CLAM model and the 2-*K* CLAM model respectively. The black lines include the ϵ^{AD} parameter in the regression while red lines fix ϵ^{AD} at the CPA value of 2526 K.

The shortcomings of the 1-*K* CLAM are very evident in both the concentrated and dilute region. In contrast, the 2-*K* CLAM is much more capable of describing the dilute region where dimers are a dominant species. The regressed parameters calculated from NMR are given in Table 7.4.

 Table 7.4: Regressed parameters from NMR analysis and corresponding CPA parameters for comparison

		IR derived		2-K CLAM (NMR)	
	СРА	K _C	$\begin{array}{c c} \mathbf{I} - \mathbf{K} & \mathbf{CLAM} \\ (\mathbf{NMR}) & \mathbf{Dimer}, \mathbf{K}_{D}^{AD} & \mathbf{N} \end{array}$		$N-mer,\ K^{AD}_{Nmer}$
KAD	6.535E-01	7.31E-01	3.135E-03	1.722E+00	5.118E-02
$\epsilon^{AD}/k(K)$	2.526E+03	2.07E+03	4.182E+03	1.712E+03	3.525E+03
\widetilde{v}'_{Nmer}			6.136	5.281	

The parameters are derived from NMR are very different from both CPA values and those derived from the IR fitting. The ϵ^{AD}/k parameter of the 1-*K* CLAM model is almost double that of derived from the IR fits (a difference in the enthalpy of the hydrogen bond of approximately 4.2 kcal/mol) and the K^{AD} parameter is less than three orders of magnitude smaller. With the 2-*K* CLAM, the dimer enthalpy parameter is very low (3.4 kcal/mol) and that of larger oligomers is very high (7 kcal/mol). The NMR algorithm can yield multiple solutions with similar fits to the data. To demonstrate this, we fixed the value of ϵ^{AD}/k to the CPA value of 2526 K for both the 1-*K* CLAM and 2-*K* CLAM and refitted the other parameters. This yields the red lines in Figures 7.18-7.20. The values of the shift ν_N remained almost the same but K^{AD} , K_D^{AD} and K_{Nmer}^{AD} were found to be 0.696, 0.1166 and 1.341 respectively. These quantities are closer to CPA and IR-derived values. This indicates that a more reasonable solution exists but that the regression program must be constrained in order to find it. More work must be done to correct for this and we outline this and other future work in Section 7.4.

7.4. Summary of future work

This chapter contains preliminary analyses and results for the calculation of Wertheim association parameters from IR and NMR spectroscopy. As an ongoing project, there are several areas that will be further pursued in the future. Firstly, the fitting of the scaled infrared spectra should be revisited. Specifically, the spectrum at T=70 °C and $x_{BuOH} = 0.0469$ produced strange trends as discussed earlier. This may be due to errors in the experimental IR data collection because evaporation can occur at high temperatures. However, this may also be due to the low wavenumber cutoff which marks the end of the integration area. This is most clearly visible in Figure 7.3 where the line is discontinued at a lower wavenumber than the other spectra. Currently, the range of the plot is determined by finding the minima on either side of the O-H band. However, in the future, data should be integrated between band limits that are consistent for all temperatures and concentrations.

Though the QM species distributions in Figure 7.8 suggest that the α hydroxyl, like the β hydroxyl, is best described by two Gaussian peaks instead of one, adding another peak to the fitting procedure may result in difficulties in discerning differences between the two types. For this work, the β hydroxyls were represented by a broad low peak in the high frequency region and the α hydroxyls were represented by a tall narrow peak. However, because the α and β distributions are so similar in Figure 7.8 separation of the species may not be correct. Rather, it may be more realistic combine the two peaks together as the sum $\alpha + \beta$, eliminate the area constraint between β and γ hydroxyls, and, instead. focus on improved initial guesses for the γ hydroxyls.

Regarding the shape of the fitted peaks, it may be worthwhile to explore other line shapes beyond Gaussian such as Lorentzian or gamma distribution functions. Some researchers have found success by combining two functions such as with the Gaussian-Lorentzian curves [200]. However, this requires including more parameters for unclear benefit. Instead, it is recommended that Lorentzian functions be used for the peaks that convolute to form the sharp high frequency peak and Gaussians be used for the others. In this way, the same total number of parameters is maintained but the two peaks are analyzed differently.

Secondly, we encountered some challenges in calculating association parameters from NMR spectra. The key issue is that many combinations of parameters can yield reasonably accurate representations of the experimental data. In such cases, it is important that the stability of the final solution be tested by perturbing the initial guesses and ensuring that the same result is reached [200]. To correct for this issue, it is recommended that more constraints be defined to remove degrees of freedom.

Moreover, although the 2-K CLAM model was shown to represent NMR data more accurately, there is no equivalent form of the model within Wertheim's theory. A model with the accuracy of 2-K CLAM but flexibility and practicality of Wertheim's theory would be very powerful. Therefore, this endeavor warrants further study. Lastly, none of the models tested in this work considered cyclic species. With linear species, an *n*-mer contains *n*-1 hydrogen bonds whereas a cyclic *n*-mer contains only *n* hydrogen bonds. This counting discrepancy is not accounted for with Wertheim's theory, which serves as the basis of our thermodynamic model. While it may be difficult to adapt the functional form of Wertheim's theory for cyclic species, the idea is worth pursuing because, if successful, the resulting model will be a valuable contribution to the field. Additionally, hydrogen bonding in these types of clusters can be studied using the MD+QM method developed in Chapter 6. This insight can help improve chemical theory models by providing a physical basis for the assumptions used.

APPENDIX

Appendix I: Derivation of the 2-K CLAM model

To derive the 2-*K* CLAM model, we start with an apparent mole balance written in terms of concentration by dividing both sides by total volume, \underline{V} ,

$$\frac{n_{0,alc}}{\underline{V}} = \frac{n_M + 2n_D + 3n_{Tri} + \cdots}{\underline{V}}$$
 I.1

$$C_{alc} = C_M + 2C_D + 3C_{Tri} + \dots = C_M + \sum_{i=2}^{\infty} iC_i$$
 I.2

Defining two equilibrium constants, one for the formation of dimers and another for all subsequent oligomers:

$$K_{C,D} = \frac{C_D}{C_M^2} = \frac{\Phi_D / 2V_M}{(\Phi_M / V_M)^2} = \frac{1}{2} \frac{\Phi_D}{\Phi_M^2} V_M$$
 I.3

$$K_{C,Nmer} = \frac{c_{Tri}}{c_D c_M} = \dots = \frac{c_N}{c_{(N-1)} c_M} = \frac{(N-1)}{N} \frac{\Phi_N}{\Phi_{N-1} \Phi_M} V_M \text{ for } N > 2 \qquad \text{I.4}$$

Combining Eqs. (I.3 and I.4) and rearranging yields

$$C_N = K_{C,D} K_{C,Nmer}^{N-2} C_M^N \qquad \text{for } N > 1 \qquad \qquad \text{I.5}$$

Eq. (I.5) can be substituted into the mole balance in Eq. (I.2),

$$C_{alc} = C_M + \sum_{N=2}^{\infty} NK_{C,D} K_{C,Nmer}^{N-2} C_M^N = C_M + 2K_{C,D} C_M^2 + 3K_{C,D} K_{C,Nmer} C_M^3 + \cdots$$
 I.6

Rearranging:

$$C_{alc} = C_M \left(1 + 2K_{C,Dim}C_M + \frac{K_{C,D}}{K_{C,Nmer}} \left(3\left(K_{C,Nmer}C_M\right)^2 + 4\left(K_{C,Nmer}C_M\right)^3 + \cdots \right) \right)$$
 I.7

Recognizing the converging series

$$\sum_{i=0}^{\infty} (i+1)a^i = \frac{1}{(1-a)^2} = 1 + 2a + 3a^2 + \dots$$
 I.8

Eq. (I.7) can be written as

$$C_{alc} = C_M \left(1 + 2K_{C,D}C_M + \frac{K_{C,D}}{K_{C,Nmer}} \left(\frac{1}{\left(1 - K_{C,Nmer}C_M\right)^2} - 1 - 2K_{C,Nmer}C_M \right) \right)$$
 I.9

$$C_{alc} = C_{M} \left(1 + \frac{K_{C,D}}{K_{C,Nmer}} \left(\frac{1}{\left(1 - K_{C,Nmer} C_{M} \right)^{2}} - 1 \right) \right)$$
I.10

Substituting

$$\Phi_{Nmer} = C_{Nmer} V_{Nmer}$$
 I.11

and recognizing that the apparent molar volume of an alcohol, V_{alc} is the volume of the monomer V_M yields

$$\Phi_{alc} = \Phi_M \left(1 + \frac{K_{C,D}}{K_{C,Nmer}} \left(\frac{1}{\left(1 - K_{C,Nmer} \frac{\Phi_M}{V_M} \right)^2} - 1 \right) \right)$$
 I.12

Rearranging in terms of Φ_M , Eq. (I.12) can be written in the form of a cubic equation

$$a_0 \Phi_M^3 + a_1 \Phi_M^2 + a_2 \Phi_M + a_3 = 0 I.13$$

where a_0 , a_1 , a_2 and a_3 are:

$$a_0 = \frac{K_{C,D} K_{C,Nmer} - K_{C,Nmer}^2}{V_M^2}$$
 I.14

$$a_1 = \frac{2(K_{C,Nmer} - K_{C,D})}{V_M} + \left(\frac{K_{C,Nmer}}{V_M}\right)^2 \Phi_{alc}$$
 I.15

$$a_2 = -1 - 2\frac{K_{C,Nmer}}{V_M}\Phi_{alc}$$
 I.16

$$a_3 = \Phi_{alc}$$
 I.17

The chemical shift is then calculated from Eq. (7.11):

$$\tilde{v}' = \frac{C_M}{C_{alc}} \tilde{v}'_M + \sum_{i=2}^{\infty} \left[\left(\frac{C_i^l}{C_{alc}} \right) \tilde{v}'_M + \left((i-1) \frac{C_i^l}{C_{alc}} \right) \tilde{v}'_{Nmer} \right]$$
 I.18

Substituting Eq. (I.2) for C_M and rearranging yields:

$$\tilde{v}' = \left(1 - \sum_{i=2}^{\infty} (i-1)\frac{C_i^l}{C_{alc}}\right)\tilde{v}'_M + \sum_{i=2}^{\infty} \left((i-1)\frac{C_i^l}{C_{alc}}\right)\tilde{v}'_{Nmer}$$
 I.19

or

$$\tilde{v}' - \tilde{v}'_M = \Delta_{\tilde{v}} = \sum_{i=2}^{\infty} (i-1) \frac{C_i^l}{C_{alc}} (\tilde{v}'_{Nmer} - \tilde{v}'_M) = \sum_{i=2}^{\infty} (i-1) \frac{C_i^l}{C_{alc}} \Delta_{\tilde{v}_N}$$
 I.20

where $\Delta_{\tilde{v}}$ and $\Delta_{\tilde{v}_N}$ represent the observed chemical shift and the chemical shift of hydrogen-bonded protons relative to the monomer shift respectively. Written in terms of volume fractions and assuming that the volume of an *N*-mer is given by $V_i = iV_M$:

$$\Delta_{\tilde{\nu}} = \sum_{i=2}^{\infty} (i-1) \frac{\Phi_{Nmer}/V_{Nmer}}{\Phi_{alc}/V_{M}} \Delta_{\tilde{\nu}_{N}} = \sum_{i=2}^{\infty} \frac{(i-1)}{i} \frac{\Phi_{Nmer}}{\Phi_{alc}} \Delta_{\tilde{\nu}_{N}}$$
 I.21

Finally, substituting the definitions of the equilibrium constants in terms of volume fractions (Eqs. (I.3 and I.4)) and rearranging gives:

$$\Delta_{\tilde{\nu}} = \left[\frac{K_{C,D}}{V_M}\Phi_M^2 + \frac{K_{C,Nmer}}{V_M}\Phi_M(\Phi_{alc} - \Phi_M)\right]\frac{\Delta_{\tilde{\nu}_N}}{\Phi_{alc}}$$
 I.22

Chapter 8. Conclusions and Future Directions

The objective of this project is to combine spectroscopic tools and computational techniques to improve the thermodynamic modeling of systems with hydrogen bonding. To this end, this document begins by outlining two common methods for modeling association in thermodynamics: chemical theory and Wertheim's theory. Noting a gap in literature on the relationship between concentration- and activity-based equilibrium constants for the liquid, we first explore this question in Chapter 2. It is shown that the two equilibrium constants can be related by using Flory's theory [30] and that the concentration-based equilibrium constant, K_c , can only be assumed to be independent of composition when residual contributions to non-ideality are also constant. In Chapter 3, K_c , which is the association parameter in chemical theory, is shown to be identical to $\Delta^{A_i D_j}$, the association parameter in Wertheim's theory for alcohol + inert type systems under reasonable assumptions of zero excess volume and universal packing factor. A Wertheim activity coefficient model (WAG) is also introduced based on methods developed by Michelsen et al. [66]. Furthermore, algebraic and numerical proofs of the equivalence of the two theories are derived for these systems when only linear species are present. This finding is extended to other systems that contain both associating and solvating components in Chapter 4. The activity coefficient model developed in Chapter 3 is tested in Chapter 5 using literature (CPA) $\Delta^{A_i D_j}$ values and its performance is compared to that of existing association equations of state. It is found that, when combined with NRTL or Nagata's residual term and fitted with four parameters, the WAG models are able to capture the behavior of methanol and ethanol-containing systems well, generating fits with the lowest overall SSQ value. Among the two parameter models, WAG-SH, which combines WAG with Scatchard-Hildebrand provided the best fits and predictions of the experimental data. Indeed, WAG-SH performed similarly or better than the association equations of state tested, PC-

SAFT and CPA. Chapters 6 and 7 describe work on the analysis of IR spectra for the calculation of $\Delta^{A_i D_j}$ values. In the former, the hydroxyl vibrational region in IR spectra of alcohol + alkane systems is quantified for the first time and related to the apparent concentration of alcohol in the mixture. This is accomplished by using relationships between IR properties found by MD + QM calculations of hydrogen bonded clusters. In Chapter 7, preliminary calculations for the value of $\Delta^{A_i D_j}$ from the analysis of IR spectra of *n*-butanol + cyclohexane mixtures are shown. This deconvolution procedure lies beyond the scope of this thesis. However, detailed recommendations for future directions in the IR analysis portion of this project are outlined. Other future work is categorized and summarized in Table 8.1.

	Project Goal	Completed	Future Work
	Explore similarities and differences between association modeling theories		
	Develop an activity coefficient model for Wertheim's theory (WAG model)		
hermodynamic Model	Test modeling capabilities of WAG for VLE and LLE of alcohol + inert systems		
	Test predictive power for LLE of ternary systems of alcohol + inert systems		
	Test modeling capabilities of WAG for systems with association and solvation		
	Test modeling capabilities of WAG for components that only solvate		
	Develop a 2-parameter WAG model analogous to 2-K CLAM		
L	Create an AspenPlus® user model for industrial partners to explore		
	Release AspenPlus® user model for public use		
oscopy	Establish reliable IR protocols and benchmark methods against literature		
	Establish reliable NMR protocols and benchmark methods against literature		
ectro	Assemble database of spectroscopic data for alcohol + inert systems		
Spe	Assemble database of spectroscopic data for other challenging systems		

Table 8.1: Accomplished and future goals for research project

Table 8.1 (cont'd)

	Project Goal	Completed	Future Work
	Develop scripts and run MD simulations		
MD + QM	Develop scripts to analyze MD simulations for hydrogen bonding information		
	Develop scripts to conduct QM calculations on linear clusters		
	Develop scripts to conduct QM calculations on non-linear clusters		
	Run MD+QM calculations for alcohol + inert systems		
	Run MD+QM calculations for other challenging systems		
Overall	Use MD+QM findings to quantify IR peak		
	Use MD+QM findings to quantify concentration of free ends		
	Calculate association parameters from spectroscopy (IR, NMR) and MD + QM		
	Test modeling and predictive power of WAG with new association parameters		

For the thermodynamic model development, the derivation of the WAG function should be revisited to relax some of its underlying assumptions such as the use of a constant packing fraction. Moreover, the WAG model should be tested for more systems that contain solvating components. An especially complex case occurs when components that do not associate but are capable of solvating, such as dioxane and chloroform, are mixed. In these cases, the combining rules introduced in Chapter 5 cannot be used and association parameters must be estimated through other means. Furthermore, for consistency across all phases, it will be beneficial to extend the WAG model to be able to capture association in the vapor phase by adapting the Hayden O'Connell model [79]. Once this is completed, the WAG user model, which has already been developed to function with AspenPlus®, a process design software package, can be released for commercial use. In this work, we developed a chemical theory variation with two equilibrium constants: one for the formation of a dimer and another for subsequent association. The model is shown to be

more capable of fitting NMR chemical shifts than single parameter models. However, Wertheim's theory is more easily generalizable than chemical theory. Therefore, future efforts should be invested in developing a Wertheim association theory that allows for two association parameters. In terms of computational experimentation, future work should include MD + QM calculations like those developed in Chapter 6 for different classes of systems. Additionally, the characteristics of non-linear clusters, such as ring and branched structures, should be investigated.

BIBLIOGRAPHY
BIBLIOGRAPHY

- G.M. Kontogeorgis, M.L. Michelsen, G.K. Folas, S. Derawi, N. von Solms, E.H. Stenby, Ten Years with the CPA (Cubic-Plus-Association) Equation of State. Part 1. Pure Compounds and Self-Associating Systems, Ind. Eng. Chem. Res. 45 (2006) 4855–4868. doi:10.1021/ie051305v.
- [2] G.M. Kontogeorgis, I. Tsivintzelis, N. von Solms, A. Grenner, D. Bogh, M. Frost, A. Knage-Rasmussen, I.G. Economou, Use of monomer fraction data in the parametrization of association theories, Fluid Phase Equilibria. 296 (2010) 219–229. doi:10.1016/j.fluid.2010.05.028.
- [3] H. Renon, J.M. Prausnitz, Local compositions in thermodynamic excess functions for liquid mixtures, Aiche J. 14 (1968) 135-. doi:10.1002/aic.690140124.
- [4] L.S. Budantseva, T.M. Lesteva, M.S. Nemtsov, Liquid-vapor equilibrium in systems methanol-C7-8 hydrocarbons of different classes, Zh. Fiz. Khim. 49 (1975) 1844.
- [5] H. Katayama, Liquid–liquid equilibria of two ternary systems: methanol–cyclohexane including 1,3-dioxolane or 1,4-dioxane in the range of 277.79–308.64 K, Fluid Phase Equilibria. 164 (1999) 83–95. doi:10.1016/S0378-3812(99)00241-1.
- [6] Y.-H. Fu, S.I. Sandler, H. Orbey, A Modified UNIQUAC Model That Includes Hydrogen Bonding, Industrial & Engineering Chemistry Research. 34 (1995) 4351–4363. doi:10.1021/ie00039a027.
- [7] N. Asprion, H. Hasse, G. Maurer, Application of IR-spectroscopy in thermodynamic investigations of associating solutions, Fluid Phase Equilibria. 205 (2003) 195–214. doi:10.1016/S0378-3812(02)00181-4.
- [8] J.R. Elliott, C.T. Lira, Introductory Chemical Engineering Thermodynamics, International Edition, 2nd, International ed., Pearson Education, Inc., Upper Saddle River, NJ, 2012.
- [9] W.G. Chapman, K.E. Gubbins, G. Jackson, M. Radosz, New reference equation of state for associating liquids, Industrial & Engineering Chemistry Research. 29 (1990) 1709–1721. doi:10.1021/ie00104a021.
- [10] J.R. Elliott, S.J. Suresh, M.D. Donohue, A simple equation of state for nonspherical and associating molecules, Ind. Eng. Chem. Res. 29 (1990) 1476–1485. doi:10.1021/ie00103a057.
- [11] G.M. Kontogeorgis, E.C. Voutsas, I.V. Yakoumis, D.P. Tassios, An Equation of State for Associating Fluids, Industrial & Engineering Chemistry Research. 35 (1996) 4310–4318. doi:10.1021/ie9600203.

- [12] D.S. Abrams, J.M. Prausnitz, Statisical Thermodynamics of Liquid Mixtures: A New Expression for Excess Gibbs Energy of Partly or Completely Miscible Systems, AIChE Journal. 21 (1975) 116–128. doi:10.1002/aic.690210115.
- [13] A.M. Karachewski, W.J. Howell, C. a. Eckert, Development of the AVEC model for associating mixtures using NMR spectroscopy, AIChE Journal. 37 (1991) 65–73. doi:10.1002/aic.690370106.
- [14] A.M. Karachewski, M.M. McNiel, C.A. Eckert, A study of the hydrogen-bonding in alcoholsolutions using NMR spectroscopy, Industrial & Engineering Chemistry Research. 28 (1989) 315–324. doi:10.1021/ie00087a010.
- [15] I. Nagata, K. Tamura, K. Gotoh, Association model of fluids. Phase equilibria and excess enthalpies in mixtures containing alcohol and acetonitrile, Z. Phys. Chemie-Int. J. Res. Phys. Chem. Chem. Phys., Frankfort. 199 (1997) 1–23.
- [16] S.W. Campbell, Chemical theory for mixtures containing any number of alcohols, Fluid Phase Equilibria. 102 (1994) 61–84.
- [17] K. Kwac, E. Geva, A mixed quantum-classical molecular dynamics study of the hydroxyl stretch in methanol/carbon tetrachloride mixtures: equilibrium hydrogen-bond structure and dynamics at the ground state and the infrared absorption spectrum., The Journal of Physical Chemistry. B. 115 (2011) 9184–94. doi:10.1021/jp204245z.
- [18] G.M. Wilson, Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing, Journal of the American Chemical Society. 86 (1964) 127–130.
- [19] G. Scatchard, Equilibria in Non-electrolyte Solutions in Relation to the Vapor Pressures and Densities of the Components., Chem. Rev. 8 (1931) 321–333. doi:10.1021/cr60030a010.
- [20] J.H. Hildebrand, S.E. Wood, The Derivation of Equations for Regular Solutions, The Journal of Chemical Physics. 1 (1933) 817–822. doi:10.1063/1.1749250.
- [21] H.S. Fogler, Elements of Chemical Reaction Engineering, Pearson Education, 2016. https://books.google.com/books?id=uQ3rsgEACAAJ.
- [22] O. Levenspiel, Chemical Reaction Engineering, Wiley, 1999. https://books.google.com/books?id=X4svAQAAIAAJ.
- [23] D.A. McQuarrie, J.D. Simon, Physical Chemistry: A Molecular Approach, University Science Books, 1997.
- [24] K.G. Denbigh, The Principles of Chemical Equilibrium: With Applications in Chemistry and Chemical Engineering, Cambridge University Press, 1981. https://books.google.com/books?id=cwWT48wL238C.
- [25] S.I. Sandler, Chemical, Biochemical, and Engineering Thermodynamics, John Wiley & Sons, 2006. https://books.google.com/books?id=4MXDAgAAQBAJ.

- [26] M. Pekař, Thermodynamics and foundations of mass-action kinetics, Progress in Reaction Kinetics and Mechanism. 30 (2005) 3–113. doi:10.3184/007967405777874868.
- [27] R. Rönnback, T. Salmi, A. Vuori, H. Haario, J. Lehtonen, A. Sundqvist, E. Tirronen, Development of a kinetic model for the esterification of acetic acid with methanol in the presence of a homogeneous acid catalyst, Chemical Engineering Science. 52 (1997) 3369– 3381. doi:10.1016/S0009-2509(97)00139-5.
- [28] A.K. Kolah, N.S. Asthana, D.T. Vu, C.T. Lira, D.J. Miller, Reaction kinetics of the catalytic esterification of citric acid with ethanol, Industrial and Engineering Chemistry Research. 46 (2007) 3180–3187. doi:10.1021/ie060828f.
- [29] A. Fredenslund, R.L. Jones, J.M. Prausnitz, Group-contribution estimation of activity coefficients in nonideal liquid mixtures, AIChE Journal. 21 (1975) 1086–1099. doi:10.1002/aic.690210607.
- [30] P.J. Flory, Thermodynamics of Heterogeneous Polymers and Their Solutions, The Journal of Chemical Physics. 12 (1944) 425. doi:10.1063/1.1723887.
- [31] G.K. Folas, G.M. Kontogeorgis, M.L. Michelsen, E.H. Stenby, Application of the cubic-plusassociation equation of state to mixtures with polar chemicals and high pressures, Industrial & Engineering Chemistry Research. 45 (2006) 1516–1526.
- [32] A. Apelblat, The concept of associated solutions in historical development. Part 1. The 1884– 1984 period, Journal of Molecular Liquids. 128 (2006) 1–31. doi:10.1016/j.molliq.2006.02.005.
- [33] A. Apelblat, Erratum to "The concept of associated solutions in historical development. Part 1. The 1884-1984 period.," Journal of Molecular Liquids. 130 (2007) 133–162.
- [34] F. Dolezalek, Zur Theorie der Binären Gemische und Konzentrierten Lösungen, Z. Phys. Chem. 64 (1908) 727–747.
- [35] I. Nagata, K. Tamura, S. Tokuriki, Excess enthalpies and complex formation of acetonitrile with acetone, chloroform and benzene, Thermochimica Acta. 47 (1981) 315–331.
- [36] I. Nagata, Y. Kawamura, Thermodynamics of Alcohol-Unassociated Active Component Liquid Mixtures, Chemical Engineering Science. 34 (1979) 601–611.
- [37] N.D. Coggeshall, E.L. Saier, Infrared Absorption Study of Hydrogen Bonding Equilibria, Journal of the American Chemical Society. 73 (1951) 5414–5418. doi:10.1021/ja01155a118.
- [38] E.A. Guggenheim, Statistical thermodynamics of mixtures with zero energies of mixing, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences. 183 (1944) 203–212.
- [39] P. Muthukumaran, R.L. Brinkley, R.B. Gupta, Lattice-fluid equation of state with hydrogenbond cooperativity, Aiche Journal. 48 (2002) 386–392. doi:10.1002/aic.690480221.

- [40] R.B. Gupta, R.L. Brinkley, Hydrogen-bond cooperativity in 1-alkanol plus n-alkane binary mixtures, Aiche Journal. 44 (1998) 207–213. doi:10.1002/aic.690440122.
- [41] C. Panayiotou, M. Pantoula, E. Stefanis, I. Tsivintzelis, I.G. Economou, Nonrandom Hydrogen-Bonding Model of Fluids and Their Mixtures. 1. Pure Fluids, Ind. Eng. Chem. Res. 43 (2004) 6592–6606. doi:10.1021/ie040114+.
- [42] C. Panayiotou, I. Tsivintzelis, I.G. Economou, Nonrandom Hydrogen-Bonding Model of Fluids and Their Mixtures. 2. Multicomponent Mixtures, Ind. Eng. Chem. Res. 46 (2007) 2628–2636. doi:10.1021/ie0612919.
- [43] K.E. Gubbins, Perturbation Theories of the Thermodynamics of Polar and Associating Liquids: A Historical Perspective, Fluid Phase Equilibria. (2015) 1–15. doi:10.1016/j.fluid.2015.12.043.
- [44] P.T. Cummings, G. Stell, Statistical mechanical models of chemical reactions, Molecular Physics. 51 (1984) 253–287. doi:10.1080/00268978400100191.
- [45] G. Stell, Y. Zhou, Chemical association in simple models of molecular and ionic fluids, The Journal of Chemical Physics. 91 (1989) 3618–3623. doi:10.1063/1.456894.
- [46] Y. Zhou, G. Stell, Chemical association in simple models of molecular and ionic fluids. II. Thermodynamic properties, The Journal of Chemical Physics. 96 (1992) 1504–1506. doi:10.1063/1.462872.
- [47] M.S. Wertheim, Fluids with highly directional attractive forces.1. Statistical thermodynamics, Journal of Statistical Physics. 35 (1984) 19–34. doi:10.1007/bf01017362.
- [48] M.S. Wertheim, Fluids with highly directional attractive forces. 2. Thermodynamic perturbation-theory and integral-equations, Journal of Statistical Physics. 35 (1984) 35–47. doi:10.1007/bf01017363.
- [49] M.S. Wertheim, Fluids with highly directional attractive forces. 3. Multiple attraction sites, Journal of Statistical Physics. 42 (1986) 459–476. doi:10.1007/bf01127721.
- [50] M.S. Wertheim, Fluids with highly directional attractive forces. 4. Equilibrium polymerization, Journal of Statistical Physics. 42 (1986) 477–492. doi:10.1007/bf01127722.
- [51] S.J. Suresh, J.R. Elliott, Multiphase Equilibrium Analysis via a Generalized Equation of State for Associating Mixtures, Industrial & Engineering Chemistry Research. 31 (1992) 2783– 2794.
- [52] I. V Yakoumis, G.M. Kontogeorgis, E.C. Voutsas, E.M. Hendriks, D.P. Tassios, Prediction of Phase Equilibria in Binary Aqueous Systems Containing Alkanes, Cycloalkanes, and Alkenes with the Cubic-plus-Association Equation of State, Industrial & Engineering Chemistry Research. 37 (1998) 4175–4182.

- [53] J. Gross, G. Sadowski, Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules, Industrial & Engineering Chemistry Research. 40 (2001) 1244– 1260. doi:10.1021/ie0003887.
- [54] Y.-H. Fu, S.I. Sandler, A Simplified SAFT Equation of State for Associating Compounds and Mixtures, Industrial & Engineering Chemistry Research. 34 (1995) 1897–1909. doi:10.1021/ie00044a042.
- [55] S.H. Huang, M. Radosz, Equation of state for small, large, polydisperse, and associating molecules, Industrial & Engineering Chemistry Research. 29 (1990) 2284–2294. doi:10.1021/ie00107a014.
- [56] A. Gil-Villegas, A. Galindo, P.J. Whitehead, S.J. Mills, G. Jackson, A.N. Burgess, Statistical associating fluid theory for chain molecules with attractive potentials of variable range, The Journal of Chemical Physics. 106 (1997) 4168–4186. doi:10.1063/1.473101.
- [57] T. Kraska, K.E. Gubbins, Phase Equilibria Calculations with a Modified SAFT Equation of State. 1. Pure Alkanes, Alkanols, and Water, Industrial & Engineering Chemistry Research. 35 (1996) 4727–4737. doi:10.1021/ie9602320.
- [58] I.G. Economou, M.D. Donohue, Chemical, quasi-chemical and perturbation theories for associating fluids, AIChE Journal. 37 (1991) 1875–1894. doi:10.1002/aic.690371212.
- [59] J.P. Wolbach, S.I. Sandler, Using molecular orbital calculations to describe the phase behavior of hydrogen-bonding fluids, Industrial & Engineering Chemistry Research. 36 (1997) 4041–4051. doi:10.1021/ie9607255.
- [60] E.M. Hendriks, J. Walsh, A.R.D. van Bergen, A general approach to association using cluster partition functions, Journal of Statistical Physics. 87 (1997) 1287–1306. doi:10.1007/BF02181285.
- [61] J.P. Wolbach, S.I. Sandler, Using molecular orbital calculations to describe the phase behavior of cross-associating mixtures, Industrial & Engineering Chemistry Research. 37 (1998) 2917–2928. doi:10.1021/ie9707811.
- [62] J.M. Prausnitz, F.W. Tavares, Thermodynamics of Fluid-Phase Equilibria for Standard Chemical Engineering Operations, AIChE Journal. 50 (2004) 739–761. doi:10.1002/aic.10069.
- [63] M.L. Michelsen, Robust and efficient solution procedures for association models, Industrial & Engineering Chemistry Research. 45 (2006) 8449–8453. doi:10.1021/ie060029x.
- [64] G.M. Kontogeorgis, G.K. Folas, Thermodynamic Models for Industrial Applications, John Wiley & Sons, 2010.
- [65] W. Zmpitas, J. Gross, Detailed pedagogical review and analysis of Wertheim's thermodynamic perturbation theory, Fluid Phase Equilibria. 428 (2016) 121–152. doi:10.1016/j.fluid.2016.07.033.

- [66] M.L. Michelsen, E.M. Hendriks, Physical properties from association models, Fluid Phase Equilibria. 180 (2001) 165–174. doi:10.1016/S0378-3812(01)00344-2.
- [67] J.R. Elliott, C.T. Lira, Errata: Introductory chemical engineering thermodynamics, 2nd ed., 2012, Errata, (2016). http://chethermo.net/errata (accessed July 13, 2016).
- [68] H. Renon, J.M. Prausnitz, On the thermodynamics of alcohol-hydrocarbon solutions, Chemical Engineering Science. 22 (1967) 299–307. doi:10.1016/0009-2509(67)80116-7.
- [69] H. Renon, J.M. Prausnitz, On the thermodynamics of alcohol-hydrocarbon solutions. Errata., Chemical Engineering Science. 22 (1967) 1891. doi:10.1016/0009-2509(67)80223-9.
- [70] I. Nagata, Thermodynamics of alcohol solutions. Phase equilibria and excess molar enthalpies of mixtures containing two alcohols, Thermochimica Acta. 107 (1986) 199–217.
- [71] A. Nath, E. Bender, On the thermodynamics of associated solutions. I. An analytical method for determining the enthalpy and entropy of association and equilibrium constant for pure liquid substances, Fluid Phase Equilibria. 7 (1981) 275–287. doi:10.1016/0378-3812(81)80012-X.
- [72] I. Nagata, K. Gotoh, Thermodynamics of alcohol solutions, Thermochimica Acta. 258 (1995) 77–107.
- [73] V. Brandani, A continuous linear association model for determining the enthalpy of hydrogen-bond formation and the equilibrium association constant for pure hydrogen-bonded liquids, Fluid Phase Equilibria. 12 (1983) 87–104.
- [74] V. Brandani, F. Evangelista, The UNIQUAC associated-solution theory: vapor-liquid equilibria of binary systems containing one associating and one inert or active component, Fluid Phase Equilibria. 17 (1984) 281–302.
- [75] R. a Heidemann, J.M. Prausnitz, A van der Waals-type equation of state for fluids with associating molecules., Proceedings of the National Academy of Sciences of the United States of America. 73 (1976) 1773–6.
- [76] A.G. Pradhan, V.R. Bhethanabotla, S.W. Campbell, Vapor-liquid equilibrium data for ethanol-n-heptane-1-propanol and ethanol-n-heptane-2-propanol and their interpretation by a simple association model, Fluid Phase Equilibria. 84 (1993) 183–206. doi:10.1016/0378-3812(93)85123-4.
- [77] G.M. Lobien, J.M. Prausnitz, Correlation for the ratio of limiting activity coefficients for binary liquid mixtures, Fluid Phase Equilibria. 8 (1982) 149–160. doi:10.1016/0378-3812(82)80032-0.
- [78] W.A. Fouad, L. Wang, A. Haghmoradi, S.K. Gupta, W.G. Chapman, Understanding the Thermodynamics of Hydrogen Bonding in Alcohol-Containing Mixtures: Self Association, The Journal of Physical Chemistry B. 119 (2015) 14086–14101. doi:10.1021/acs.jpcb.5b08285.

- [79] J.G. Hayden, J.P. O'Connell, A generalized method for predicting second virial coefficients, Industrial & Engineering Chemistry Process Design and Development. 14 (1975) 209–216. doi:10.1021/i260055a003.
- [80] D. Ambrose, J.H. Ellender, C.H.S. Sprake, R. Townsend, Thermodynamic properties of organic oxygen compounds: Vapour pressures of some ethers, Journal of Chemical Thermodynamics. 8 (1976) 165–178.
- [81] A. Grenner, G.M. Kontogeorgis, N. von Solms, M.L. Michelsen, Application of PC-SAFT to glycol containing systems – PC-SAFT towards a predictive approach, Fluid Phase Equilibria. 261 (2007) 248–257. doi:10.1016/j.fluid.2007.04.025.
- [82] J. Suresh, E.J. Beckman, Prediction of liquid-liquid equilibria in ternary mixtures from binary data, Fluid Phase Equilibria. 99 (1994) 219–240. doi:10.1016/0378-3812(94)80033-2.
- [83] E.C. Voutsas, I.V. Yakoumis, D.P. Tassios, Prediction of phase equilibria in water/alcohol/alkane systems, Fluid Phase Equilibria. 158–160 (1999) 151–163.
- [84] I. Nagata, K. Tamura, Thermodynamics of Solutions of Acetonitrile with Propanols, Thermochimica Acta. 98 (1986) 147–158.
- [85] C.A. Schmidt, S.W. Campbell, An association model for mixtures containing any number of alkanes and any number of alcohols: Application to alcohol-alkane binary systems, International Journal of Thermophysics. 16 (1995) 1299–1307. doi:10.1007/BF02081297.
- [86] W. Acree, Thermodynamic properties of nonelectrolyte solutions, Academic Press Inc., 1984.
- [87] J.M. Prausnitz, R.N. Lichtenthaler, E.G. de Azevedo, Molecular thermodynamics of fluidphase equilibria, 2nd Edition, Pearson Education, 1986.
- [88] J. Gmehling, D.D. Liu, J.M. Prausnitz, High-pressure vapor—liquid equilibria for mixtures containing one or more polar components: Application of an equation of state which includes dimerization equilibria, Chemical Engineering Science. 34 (1979) 951–958. doi:10.1016/0009-2509(79)85006-X.
- [89] G.D. Ikonomou, M.D. Donohue, Extension of the associated perturbed anisotropic chain theory to mixtures with more than one associating component, Fluid Phase Equilibria. 39 (1988) 129–159. doi:10.1016/0378-3812(88)85002-7.
- [90] I. Nagata, K. Tamura, S. Tokuriki, Excess enthalpies and complex formation of acetonitrile with acetone, chloroform, and benzene, Thermochimica Acta. 47 (1981) 315–331. doi:10.1016/0040-6031(81)80110-4.
- [91] W.J. Howell, C.T. Lira, C.A. Eckert, Linear chemical-physical theory model for liquid metal solution thermodynamics, AIChE Journal. 34 (1988) 1477–1485. doi:10.1002/aic.690340909.

- [92] I. Nagata, Thermodynamic properties of alcohol-inert component mixtures, Zeitschrift Für Physikalische Chemie, Leipzig. 259 (1978) 1151–21159.
- [93] I. Nagata, K. Gotoh, Association model of fluids. Phase equilibria in sulphur dioxide mixtures, Fluid Phase Equilibria. 120 (1996) 69–80.
- [94] I. Nagata, K. Gotoh, Thermodynamics of amine and acetonitrile solutions, Thermochimica Acta. 274 (1996) 81–99.
- [95] J. Gross, G. Sadowski, Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules, Ind. Eng. Chem. Res. 40 (2001) 1244–1260. doi:10.1021/ie0003887.
- [96] A.M. Bala, C.T. Lira, Relation of Wertheim association constants to concentration-based equilibrium constants for mixtures with chain-forming components, Fluid Phase Equilibria. 430 (2016) 47–56. doi:10.1016/j.fluid.2016.09.015.
- [97] D. Ghonasgi, W.G. Chapman, Theory and simulation for associating fluids with four bonding sites, Molecular Physics. 79 (1993) 291–311. doi:10.1080/00268979300101221.
- [98] A.M. Karachewski, M.M. McNiel, C. a. Eckert, A study of hydrogen bonding in alcohol solutions using NMR spectroscopy, Industrial & Engineering Chemistry Research. 28 (1989) 315–324. doi:10.1021/ie00087a010.
- [99] A. Santhanakrishnan, A. Shannon, L. Peereboom, C.T. Lira, D.J. Miller, Kinetics of mixed ethanol/n-butanol esterification of butyric acid with amberlyst 70 and p-toluene sulfonic acid, Industrial and Engineering Chemistry Research. 52 (2013) 1845–1853. doi:10.1021/ie302267s.
- [100] T.L. Jordison, C.T. Lira, D.J. Miller, Condensed-phase ethanol conversion to higher alcohols, Ind. Eng. Chem. Res. 54 (2015) 10991–11000. doi:10.1021/acs.iecr.5b02409.
- [101] T.L. Jordison, L. Peereboom, D.J. Miller, Impact of Water on Condensed Phase Ethanol Guerbet Reactions, Ind. Eng. Chem. Res. 55 (2016) 6579–6585. doi:10.1021/acs.iecr.6b00700.
- [102] F.E. Anderson, J.M. Prausnitz, Inhibition of gas hydrates by methanol, AIChE Journal. 32 (1986) 1321–1333. doi:10.1002/aic.690320810.
- [103] H.G. Harris, J.M. Prausnitz, Thermodynamics of Solutions with Physical and Chemical Interactions. Solubility of Acetylene in Organic Solvents, Ind. Eng. Chem. Fund. 8 (1969) 180–188. doi:10.1021/i160030a001.
- [104] Y.-H. Fu, H. Orbey, S.I. Sandler, Prediction of Vapor Liquid Equilibria of Associating Mixtures with UNIFAC Models That Include Association, Industrial & Engineering Chemistry Research. 35 (1996) 4656–4666. doi:10.1021/ie950545f.

- [105] A. Vahid, N.H. Gray, J.R. Elliott, Trends in the Athermal Entropy of Mixing of Polymer Solutions, Macromolecules. 47 (2014) 1514–1531. doi:10.1021/ma402593m.
- [106] A.J. Staverman, The entropy of high polymer solutions.Generalization of formulae, Recueil Des Travaux Chimiques Des Pays-Bas. 69 (1950) 163–174.
- [107] M.D. Donohue, J.M. Prausnitz, Combinatorial Entropy of Mixing Molecules that Differ in Size and Shape. A Simple Approximation for Binary and Multicomponent Mixtures, Can. J. Chem. 53 (1975) 1586–1592. doi:10.1139/v75-224.
- [108] I. Kikic, P. Alessi, P. Rasmussen, A. Fredenslund, On the combinatorial part of the UNIFAC and UNIQUAC models, The Canadian Journal of Chemical Engineering. 58 (1980) 253–258. doi:10.1002/cjce.5450580218.
- [109] U. Weidlich, J. Gmehling, A modified UNIFAC model. 1. Prediction of VLE, hE, and gamma infinity, Industrial & Engineering Chemistry Research. 26 (2012) 1372–1381. doi:10.1021/ie00067a018.
- [110] U. Weidlich, Experimentelle und theoretische Untersuchungen zur Erweiterung der Gruppenbeitragsmethode UNIFAC, (1985).
- [111] B.L. Larsen, P. Rasmussen, A. Fredenslund, A Modified UNIFAC Group-Contribution Model for Prediction of Phase Equilibria and Heats of Mixing, Industrial and Engineering Chemistry Research. 26 (1987) 2274–2286.
- [112] B.L. Larsen, Predictions of phase equilibria and heat effects of mixing with a modified UNIFAC model, PhD Thesis, Danmarks tekniske Højskole, 1986.
- [113] S.G. Sayegh, J.H. Vera, Lattice-Model Expressions for the Combinatorial Entropy of Liquid Mixtures: A Critical Discussion, The Chemical Engmeering Journal, 19 (1980) 1– 10.
- [114] G.M. Kontogeorgis, G.K. Folas, Chapter 9, Appendix A, in: Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories, John Wiley & Sons, 2009. https://www.wiley.com/go/kontogeorgis.
- [115] J. Gross, G. Sadowski, Application of the Perturbed-Chain SAFT Equation of State to Associating Systems, Ind. Eng. Chem. Res. 41 (2002) 5510–5515. doi:10.1021/ie010954d.
- [116] N.F. Carnahan, K.E. Starling, Equation of State for Nonattracting Rigid Spheres, The Journal of Chemical Physics. 51 (1969) 635–636. doi:10.1063/1.1672048.
- [117] L.S. Budantseva, T.M. Lesteva, M.S. Nemtsov, Liquid-vapor equilibrium in methanol-C6 hydrocarbons of different classes, Zh. Fiz. Khim. 49 (1975) 260.
- [118] H. Katayama, M. Ichikawa, Liquid-liquid equilibria of three ternary systems: methanolheptane including 1, 3-dioxolane, 1, 4-dioxane and tetrahydropyran in the range of 253.15 to 303.15 K, Journal of Chemical Engineering of Japan. 28 (1995) 412–418.

- [119] J. Gmehling, B. Krentscher, Excess enthalpies of 12 binary liquid mixtures containing cyclohexane at elevated temperatures and pressures (up to 416 K and 1.9 MPa), ELDATA Int. Electron. J. Phys.-Chem. Data. 1 (1995) 181–190.
- [120] R.A. Wilsak, S.W. Campbell, G. Thodos, Vapor—liquid equilibrium measurements for the n-pentane—methanol system at 372.7, 397.7 and 422.6 K, Fluid Phase Equilibria. 33 (1987) 157–171.
- [121] Fu, J., Wang, K., Hu, Y., Studies on vapor-liquid and liquid-liquid vapor equilibria for the ternary system methanol-methyl methacrylate-water (I) three binary systems, Chinese Journal of Chemical Engineering. 4 (1989) 1–13.
- [122] C. Yang, X. Yin, S. Ma, Organic salt effect of tetramethylammonium bicarbonate on the vapor–liquid equilibrium of the dimethyl carbonate+ methanol system, Journal of Chemical & Engineering Data. 57 (2011) 66–71.
- [123] A. Arce, J. Martínez-Ageitos, E. Rodil, A. Soto, Measurement and prediction of isobaric vapour–liquid equilibrium data of the system ethanol+ methanol+ 2-methoxy-2methylpropane, Fluid Phase Equilibria. 146 (1998) 139–153.
- [124] Z.S. Kooner, D.V. Fenby, Vapour pressure study of the deuterium exchange reaction in methanol-ethanol systems: equilibrium constant determination, Australian Journal of Chemistry. 33 (1980) 1943–1946.
- [125] K.L. Butcher, W.I. Robinson, An apparatus for determining high-pressure liquid-vapour equilibrium data. I. The methanol—ethanol system, Journal of Chemical Technology and Biotechnology. 16 (1966) 289–292.
- [126] H.D. Pflug, A.E. Pope, G.C. Benson, Heats of mixing of normal alcohols at 25. deg., Journal of Chemical & Engineering Data. 13 (1968) 408–410.
- [127] H.S. Myers, VL equilibrium for naphthenes and paraffins, Petroleum Refiner. 36 (1957) 175–178.
- [128] J.M. Sturtevant, P.A. Lyons, Application of flow calorimetry to the determination of the enthalpies of mixing of organic liquids, The Journal of Chemical Thermodynamics. 1 (1969) 201–209.
- [129] C. Narasigadu, K. Moodley, J.D. Raal, Vapor–Liquid Equilibrium Measurements of Ethanol and n-Nonane or n-Decane Binary Mixtures with Large Relative Volatility, Journal of Chemical & Engineering Data. 63 (2018) 1240–1248.
- [130] L. Wang, G.C. Benson, B.C.-Y. Lu, Excess enthalpies of (ethanol+ hexane+ decane or dodecane) at the temperature 298.15 K, The Journal of Chemical Thermodynamics. 24 (1992) 1135–1143.
- [131] K.S. Yuan, J.C.K. Ho, A.K. Koshpande, B.C.-Y. Lu, Vapor-Liquid Equilibria., J. Chem. Eng. Data. 8 (1963) 549–559. doi:10.1021/je60019a024.

- [132] I. Nagata, K. Kazuma, Heats of mixing for the ternary system ethanol-1-propanolcyclohexane at 25. degree. C, Journal of Chemical and Engineering Data. 22 (1977) 79–84.
- [133] E. Tusel-Langer, J.M.G. Alonso, M.A.V. Olfos, R.N. Lichtenthaler, Excess enthalpies of mixtures containing n-heptane, methanol and methyl tert-butyl ether (MTBE), J Solution Chem. 20 (1991) 153–163. doi:10.1007/BF00651647.
- [134] P.M. Mathias, H.Z. Kister, Effect of Phase-Equilibrium Uncertainties on Ethyl Acetate Purification, J. Chem. Eng. Data. 62 (2017) 2872–2883. doi:10.1021/acs.jced.7b00172.
- [135] V. Flemr, Note on excess Gibbs energy equations based on local composition concepts., Collection of Czechoslovak Chemical Communications. 41 (1976) 3347–3349.
- [136] S.-T. Lin, M.-K. Hsieh, C.-M. Hsieh, C.-C. Hsu, Towards the development of theoretically correct liquid activity coefficient models, The Journal of Chemical Thermodynamics. 41 (2009) 1145–1153. doi:10.1016/j.jct.2009.05.002.
- [137] N. Asprion, H. Hasse, G. Maurer, FT-IR spectroscopic investigations of hydrogen bonding in alcohol-hydrocarbon solutions, Fluid Phase Equilibria. 186 (2001) 1–25. doi:10.1016/s0378-3812(01)00363-6.
- [138] N. Asprion, H. Hasse, G. Maurer, Thermodynamic and IR spectroscopic studies of solutions with simultaneous association and solvation, Fluid Phase Equilibria. 208 (2003) 23–51. doi:10.1016/S0378-3812(02)00317-5.
- [139] S.D. Williams, T.J. Johnson, S.W. Sharpe, V. Yavelak, R.P. Oates, C.S. Brauer, Quantitative vapor-phase IR intensities and DFT computations to predict absolute IR spectra based on molecular structure: I. Alkanes, Journal of Quantitative Spectroscopy and Radiative Transfer. 129 (2013) 298–307. doi:10.1016/j.jqsrt.2013.07.005.
- [140] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, Chemometrics tools used in analytical chemistry: an overview, Talanta. 123 (2014) 186–199. doi:10.1016/j.talanta.2014.02.003.
- [141] R.N. Jones, The intensities of the infra-red absorption bands of n-paraffin hydrocarbons, Spectrochimica Acta. 9 (1957) 235–251. doi:10.1016/0371-1951(57)80137-4.
- [142] A.S. Wexler, Infrared determination of structural units in organic compounds by integrated intensity measurements: Alkanes, alkenes and monosubstituted alkyl benzenes, Spectrochimica Acta. 21 (1965) 1725–1742. doi:10.1016/0371-1951(65)80085-6.
- [143] R. Řeřicha, P. Jarolímek, M. Horák, Determination of the degree of branching in alkanes by infrared spectroscopy. II. Variation of the intensity of the antisymmetrical stretching vibration bands of C-H bonds in CH₃ and CH₂ groups, Collect. Czech. Chem. Commun., CCCCC. 32 (1967) 1903–1912. doi:10.1135/cccc19671903.
- [144] K.M. Murdoch, T.D. Ferris, J.C. Wright, T.C. Farrar, Infrared spectroscopy of ethanol clusters in ethanol-hexane binary solutions, The Journal of Chemical Physics. 116 (2002) 5717–5724. doi:10.1063/1.1458931.

- [145] E.E. Tucker, E.D. Becker, Alcohol association studies. II. Vapor pressure, 220-MHz proton magnetic resonance, and infrared investigations of tert-butyl alcohol association in hexadecane, The Journal of Physical Chemistry. 77 (2012) 1783–1795. doi:10.1021/j100633a012.
- [146] X. Wu, Y. Chen, T. Yamaguchi, Hydrogen bonding in methanol studied by infrared spectroscopy, Journal of Molecular Spectroscopy. 246 (2007) 187–191. doi:10.1016/j.jms.2007.09.012.
- [147] J.T. Reilly, A. Thomas, A.R. Gibson, C.Y. Luebehusen, M.D. Donohue, Analysis of the Self-Association of Aliphatic Alcohols Using Fourier Transform Infrared (FT-IR) Spectroscopy, Ind. Eng. Chem. Res. 52 (2013) 14456–14462. doi:10.1021/ie302174r.
- [148] K. Ohta, K. Tominaga, Vibrational population relaxation of hydrogen-bonded phenol complexes in solution: Investigation by ultrafast infrared pump–probe spectroscopy, Chemical Physics. 341 (2007) 310–319. doi:10.1016/j.chemphys.2007.07.025.
- [149] K. Mazur, M. Bonn, J. Hunger, Hydrogen Bond Dynamics in Primary Alcohols: A Femtosecond Infrared Study, J. Phys. Chem. B. 119 (2015) 1558–1566. doi:10.1021/jp509816q.
- [150] A. Choperena, P. Painter, An infrared spectroscopic study of hydrogen bonding in ethyl phenol: A model system for polymer phenolics, Vibrational Spectroscopy. 51 (2009) 110– 118. doi:10.1016/j.vibspec.2008.11.008.
- [151] A. Choperena, P. Painter, Hydrogen Bonding in Polymers: Effect of Temperature on the OH Stretching Bands of Poly(vinylphenol), Macromolecules. 42 (2009) 6159–6165. doi:10.1021/ma900928z.
- [152] N. Sheppard, Infrared Spectroscopy and Hydrogen Bonding Band-Widths and Frequency Shifts, in: D. Hadži (Ed.), Hydrogen Bonding: Papers Presented at the Symposium on Hydrogen Bonding Held at Ljubljana, 29 July–3 August 1957, Pergamon, 1959: pp. 85–105. doi:10.1016/B978-0-08-009140-2.50013-0.
- [153] J.-M. Andanson, J.-C. Soetens, T. Tassaing, M. Besnard, Hydrogen bonding in supercritical tert-butanol assessed by vibrational spectroscopies and molecular-dynamics simulations., The Journal of Chemical Physics. 122 (2005) 174512. doi:10.1063/1.1886730.
- [154] I. Doroshenko, V. Pogorelov, V. Sablinskas, V. Balevicius, Matrix-isolation study of cluster formation in methanol: O–H stretching region, Journal of Molecular Liquids. 157 (2010) 142–145. doi:10.1016/j.molliq.2010.09.003.
- [155] D. Wandschneider, M. Michalik, A. Heintz, Spectroscopic and thermodynamic studies of liquid n-butanol+n-hexane and +cyclohexane mixtures based on quantum mechanical ab initio calculations of n-butanol clusters, Journal of Molecular Liquids. 125 (2006) 2–13. doi:10.1016/j.molliq.2005.11.011.

- [156] P.E. Hansen, J. Spanget-Larsen, On prediction of OH stretching frequencies in intramolecularly hydrogen bonded systems, Journal of Molecular Structure. 1018 (2012) 8– 13. doi:10.1016/j.molstruc.2012.01.011.
- [157] M. Koné, B. Illien, C. Laurence, J. Graton, Can Quantum-Mechanical Calculations Yield Reasonable Estimates of Hydrogen-Bonding Acceptor Strength? The Case of Hydrogen-Bonded Complexes of Methanol, The Journal of Physical Chemistry A. 115 (2011) 13975– 13985. doi:10.1021/jp209200w.
- [158] K. Ohno, T. Shimoaka, N. Akai, Y. Katsumoto, Relationship between the broad OH stretching band of methanol and hydrogen-bonding patterns in the liquid phase., The Journal of Physical Chemistry. A. 112 (2008) 7342–8. doi:10.1021/jp800995m.
- [159] J. Spanget-Larsen, B.K.V. Hansen, P.E. Hansen, OH stretching frequencies in systems with intramolecular hydrogen bonds: Harmonic and anharmonic analyses, Chemical Physics. 389 (2011) 107–115. doi:10.1016/j.chemphys.2011.09.011.
- [160] A. Staib, D. Borgis, A quantum multi-mode molecular dynamics approach to the vibrational spectroscopy of solvated hydrogen-bonded complexes, Chemical Physics Letters. 271 (1997) 232–240. doi:10.1016/S0009-2614(97)00470-3.
- [161] M.K. Ghosh, J. Lee, C.H. Choi, M. Cho, Direct Simulations of Anharmonic Infrared Spectra Using Quantum Mechanical/Effective Fragment Potential Molecular Dynamics (QM/EFP-MD): Methanol in Water, J. Phys. Chem. A. 116 (2012) 8965–8971. doi:10.1021/jp306807v.
- [162] J. Wang, R.J. Boyd, A. Laaksonen, A hybrid quantum mechanical force field molecular dynamics simulation of liquid methanol: Vibrational frequency shifts as a probe of the quantum mechanical/molecular mechanical coupling, The Journal of Chemical Physics. 104 (1996) 7261–7269. doi:10.1063/1.471439.
- [163] S.A. Corcelli, J.L. Skinner, Infrared and Raman Line Shapes of Dilute HOD in Liquid H2O and D2O from 10 to 90 °C, J. Phys. Chem. A. 109 (2005) 6154–6165. doi:10.1021/jp0506540.
- [164] S.A. Corcelli, C.P. Lawrence, J.L. Skinner, Combined electronic structure/molecular dynamics approach for ultrafast infrared spectroscopy of dilute HOD in liquid H2O and D2O, The Journal of Chemical Physics. 120 (2004) 8107–8117. doi:10.1063/1.1683072.
- [165] J.R. Schmidt, S.A. Corcelli, J.L. Skinner, Pronounced non-Condon effects in the ultrafast infrared spectroscopy of water, The Journal of Chemical Physics. 123 (2005) 044513. doi:10.1063/1.1961472.
- [166] B. Auer, R. Kumar, J.R. Schmidt, J.L. Skinner, Hydrogen Bonding and Raman, IR, and 2D-IR Spectroscopy of Dilute HOD in Liquid D₂O, Proceedings of the National Academy of Sciences of the United States of America. 104 (2007) 14215–14220.

- [167] S.M. Gruenbaum, C.J. Tainter, L. Shi, Y. Ni, J.L. Skinner, Robustness of Frequency, Transition Dipole, and Coupling Maps for Water Vibrational Spectroscopy, J. Chem. Theory Comput. 9 (2013) 3109–3117. doi:10.1021/ct400292q.
- [168] O.O. Mesele, W.H. Thompson, A "Universal" Spectroscopic Map for the OH Stretching Mode in Alcohols, J. Phys. Chem. A. 121 (2017) 5823–5833. doi:10.1021/acs.jpca.7b05836.
- [169] K.B. Whetsel, J.. Lady, Self-association of phenol in nonpolar solvents, in: R.A. Friedel (Ed.), Spectrometry of Fuels, Plenum Press, New York, 1970.
- [170] A. Hall, J.L. Wood, Origin of the structure of the OH stretching band of some phenols in solution, Spectrochimica Acta Part A: Molecular Spectroscopy. 23 (1967) 2657–2667. doi:10.1016/0584-8539(67)80157-0.
- [171] S. Woutersen, U. Emmerichs, H.J. Bakker, A femtosecond midinfrared pump-probe study of hydrogen-bonding in ethanol, The Journal of Chemical Physics. 107 (1997) 1483–1490. doi:10.1063/1.474501.
- [172] K. Shinokita, A.V. Cunha, T.L.C. Jansen, M.S. Pshenichnikov, Hydrogen bond dynamics in bulk alcohols, The Journal of Chemical Physics. 142 (2015) 212450. doi:10.1063/1.4921574.
- [173] C.B. Kretschmer, R. Wiebe, Thermodynamics of Alcohol-Hydrocarbon Mixtures, The Journal of Chemical Physics. 22 (1954) 1697–1701. doi:10.1063/1.1739878.
- [174] D. Deng, H. Li, J. Yao, S. Han, Simple local composition model for 1H NMR chemical shift of mixtures, Chemical Physics Letters. 376 (2003) 125–129. doi:10.1016/S0009-2614(03)00996-5.
- [175] D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E., Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X., Wu and P.A. Kollman (2014), AMBER 14, University of California, San Francisco, n.d.
- [176] L. Martínez, R. Andrade, E.G. Birgin, J.M. Martínez, PACKMOL: A package for building initial configurations for molecular dynamics simulations, Journal of Computational Chemistry. 30 (n.d.) 2157–2164. doi:10.1002/jcc.21224.
- [177] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L.

Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT, 2013.

- [178] A.D. Becke, Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals, The Journal of Chemical Physics. 107 (1997) 8554–8560. doi:10.1063/1.475007.
- [179] R. Zheng, Y. Sun, Q. Shi, Theoretical study of the infrared and Raman line shapes of liquid methanol, Phys. Chem. Chem. Phys. 13 (2011) 2027–2035. doi:10.1039/C0CP01145B.
- [180] H. Torii, Time-Domain Calculations of the Infrared and Polarized Raman Spectra of Tetraalanine in Aqueous Solution, J. Phys. Chem. B. 111 (2007) 5434–5444. doi:10.1021/jp070301w.
- [181] Killian W.G.; Unpublished experiments and calculations from Lira Lab, Michigan State University; Personal Communication, 2018, (n.d.).
- [182] J. Spanget-Larsen, Infrared Intensity and Lorentz Epsilon Curve from "Gaussian" FREQ Output, (2015). doi:10.13140/RG.2.1.4181.6160.
- [183] NIST Computational Chemistry Comparison and Benchmark Database; NIST Standard Reference Database Number 101; Release 19, April 2018, Editor: Russell D. Johnson III; http://cccbdb.nist.gov/; DOI:10.18434/T47C7Z, n.d.
- [184] C. Van Ness, B. Hollingerlb, Infrared Spectra and the Thermodynamics of Alcohol-Hydrocarbon Systems, The Journal of Physical Chemistry. 71 (1967) 1483–1494.
- [185] A.N. Fletcher, C.A. Heller, Self-association of alcohols in nonpolar solvents, J. Phys. Chem. 71 (1967) 3742–3756. doi:10.1021/j100871a005.
- [186] K. Shinomiya, T. Shinomiya, An Equilibrium Model of the Self-Association of 1- and 3-Pentanols in Heptane, The Chemical Society of Japan. 63 (1990) 1093–1097.
- [187] C.M. Huggins, G.C. Pimentel, Systematics of the Infrared Spectral Properties of Hydrogen Bonding Systems: Frequency Shift, Half Width and Intensity, J. Phys. Chem. 60 (1956) 1615–1619. doi:10.1021/j150546a004.
- [188] T. Ebukuro, A. Takami, Y. Oshima, S. Koda, Raman spectroscopic studies on hydrogen bonding in methanol and methanol/water mixtures under high temperature and pressure, The Journal of Supercritical Fluids. 15 (1999) 73–78. doi:10.1016/S0896-8446(98)00126-0.
- [189] D. Vojta, K. Dominković, S. Miljanić, J. Spanget-Larsen, Intramolecular hydrogen bonding in myricetin and myricitrin. Quantum chemical calculations and vibrational spectroscopy, Journal of Molecular Structure. 1131 (2017) 242–249. doi:10.1016/j.molstruc.2016.11.069.

- [190] F. Palombo, M. Paolantoni, P. Sassi, A. Morresi, R.S. Cataliotti, Spectroscopic studies of the "free" OH stretching bands in liquid alcohols, Journal of Molecular Liquids. 125 (2006) 139–146.
- [191] T. Yamaguchi, C.J. Benmore, A.K. Soper, The structure of subcritical and supercritical methanol by neutron diffraction, empirical potential structure refinement, and spherical harmonic analysis, The Journal of Chemical Physics. 112 (2000) 8976–8987. doi:10.1063/1.481530.
- [192] T. Yamaguchi, K. Hidaka, A.K. Soper, The structure of liquid methanol revisited: a neutron diffraction experiment at -80 °C and +25 °C, Molecular Physics. 96 (1999) 1159–1168. doi:10.1080/00268979909483060.
- [193] C.J. Benmore, Y.L. Loh, The structure of liquid ethanol: A neutron diffraction and molecular dynamics study, The Journal of Chemical Physics. 112 (2000) 5877–5883. doi:10.1063/1.481160.
- [194] C. Czeslik, J. Jonas, Pressure and temperature dependence of hydrogen-bond strength in methanol clusters, Chemical Physics Letters. 302 (1999) 633–638. doi:10.1016/S0009-2614(99)00170-0.
- [195] G.C. Pimentel, A.L. McClellan, The hydrogen bond, 2nd Edition, W.H. Freeman, San Francisco, 1960.
- [196] P. Schuster, G. Zundel, C. Sandorfy, eds., The hydrogen bond : recent developments in theory and experiments, North-Holland Publishing Company, Amsterdam ; New York, 1976.
- [197] S.N. Vinogradov, R.H. Linnell, Hydrogen bonding, Van Nostrand Reinhold, New York, 1971.
- [198] D.S. Bulgarevich, K. Otake, T. Sako, T. Sugeta, Y. Takebayashi, C. Kamizawa, D. Shintani, A. Negishi, C. Tsurumi, Hydrogen bonding in supercritical methanol studied by infrared spectroscopy, The Journal of Chemical Physics. 116 (2002) 1995–2003. doi:10.1063/1.1431585.
- [199] R. Böhmer, C. Gainaru, R. Richert, Structure and dynamics of monohydroxy alcohols— Milestones towards their microscopic understanding, 100 years after Debye, Physics Reports. 545 (2014) 125–195. doi:10.1016/j.physrep.2014.07.005.
- [200] S.J. Barlow, G.V. Bondarenko, Y.E. Gorbaty, T. Yamaguchi, M. Poliakoff, An IR Study of Hydrogen Bonding in Liquid and Supercritical Alcohols, J. Phys. Chem. A. 106 (2002) 10452–10460. doi:10.1021/jp0135095.
- [201] O. Gereben, L. Pusztai, Investigation of the Structure of Ethanol–Water Mixtures by Molecular Dynamics Simulation I: Analyses Concerning the Hydrogen-Bonded Pairs, J. Phys. Chem. B. 119 (2015) 3070–3084. doi:10.1021/jp510490y.

- [202] O. Gereben, L. Pusztai, Correction to "Investigation of the Structure of Ethanol–Water Mixtures by Molecular Dynamics Simulation I: Analyses Concerning the Hydrogen-Bonded Pairs," J. Phys. Chem. B. 119 (2015) 4564–4564. doi:10.1021/acs.jpcb.5b02167.
- [203] O. Gereben, Ring structure analysis of ethanol-water mixtures, Journal of Molecular Liquids. 211 (2015) 812–820. doi:10.1016/j.molliq.2015.08.010.
- [204] Y. Vaskivskyi, I. Doroshenko, Y. Chernolevska, V. Pogorelov, G. Pitsevich, Spectroscopic studies of clusterization of methanol molecules isolated in a nitrogen matrix, Low Temperature Physics. 43 (2017) 1415–1419. doi:10.1063/1.5012794.
- [205] R. Laenen, C. Rauscher, A structural model for associated liquid ethanol developed from transient spectroscopy, The Journal of Chemical Physics. 107 (1997) 9759–9763. doi:10.1063/1.475273.
- [206] T. Häber, U. Schmitt, M.A. Suhm, FTIR-spectroscopy of molecular clusters in pulsed supersonic slit-jet expansions, Physical Chemistry Chemical Physics. 1 (1999) 5573–5582. doi:10.1039/A907264K.
- [207] M. Paolantoni, P. Sassi, A. Morresi, R.S. Cataliotti, Infrared study of 1-octanol liquid structure, Chemical Physics. 310 (2005) 169–178. doi:10.1016/j.chemphys.2004.10.027.
- [208] P. Gans, Data fitting in the chemical sciences: by the method of least squares, Wiley, Chichester, 1992.
- [209] J.-S. Chen, K.-T. Yeh, D.-Y. Kao, J.K. Baird, Toward a Better Determination of Infrared Molar Absorptivities and Dimer Formation Constants in Self-association Through Hydrogen Bonding: 3-Ethyl-2-methyl-3-pentanol in Tetrachloroethylene as an Example, Journal of the Chinese Chemical Society. 64 (n.d.) 1156–1163. doi:10.1002/jccs.201600865.
- [210] H. Kempter, R. Mecke, Spektroskopische Bestimmung von Assoziations-Gleichgewichten, Die Naturwissenschaften. 27 (1939) 583–584. doi:10.1007/BF01496162.
- [211] G.D. Ikonomou, M.D. Donohue, Thermodynamics of hydrogen-bonded molecules: The associated perturbed anisotropic chain theory, AIChE Journal. 32 (n.d.) 1716–1725. doi:10.1002/aic.690321015.
- [212] G.D. Ikonomou, M.D. Donohue, Compact: a simple equation of state for associated molecules, Fluid Phase Equilibria. 33 (1987) 61–90. doi:10.1016/0378-3812(87)87004-8.
- [213] J.M. Bakke, I. Hampson, S. Kumar, J. Gallagher, R. Datema, J. Chattopadhyaya, Hydrogen Bonding of t-Butyl Alcohol at Low Temperatures and Concentrations., Acta Chemica Scandinavica. 39b (1985) 15–23. doi:10.3891/acta.chem.scand.39b-0015.
- [214] G. Brink, L. Glasser, Dielectric studies of molecular association. A model for the association of ethanol in dilute solution, J. Phys. Chem. 82 (1978) 1000–1005. doi:10.1021/j100498a007.

- [215] C.A. Eckert, M.M. McNiel, B.A. Scott, L.A. Halas, NMR Measurements of chemical theory equilibrium constants for hydrogen-bonded solutions, AIChE Journal. 32 (1986) 820– 828. doi:10.1002/aic.690320512.
- [216] H.S. Gutowsky, A. Saika, Dissociation, Chemical Exchange, and the Proton Magnetic Resonance in Some Aqueous Electrolytes, The Journal of Chemical Physics. 21 (1953) 1688– 1694. doi:10.1063/1.1698644.