
by
Ryan Wilkins

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Biophysics

Guelph, Ontario, Canada
© Ryan Wilkins, September, 2019
ABSTRACT

ACCURATE AND PRECISE MOLECULAR SIMULATION ALGORITHMS FOR
STANDARD STATE CHEMICAL POTENTIALS AND ACTIVITY COEFFICIENTS IN
HENRY-LAW MODELS OF ELECTROLYTE SOLUTIONS

Ryan Wilkins
University of Guelph, 2019

Advisor:
Dr. William R. Smith

This thesis focuses on the development of a novel molecular simulation methodology to accurately calculate standard state chemical potentials and activity coefficients of electrolyte species for their ultimate implementation in thermodynamic models of liquid solutions. Molecular simulation is based on the use of simple mathematical forms (force fields) for the description of the forces among individual molecules.

For this study, the macroscopic Henry-Law (infinite dilution ideality) model for the chemical potential is connected to a particular ensemble used in statistical mechanics. We use a little-known explicit form for this connection, which yields an avenue for the calculation of $\mu^\dagger$ and activity coefficients by means of molecular simulation.

We developed a novel procedure to calculate $\mu^\dagger$ by extrapolation of residual chemical potentials obtained from molecular simulation to infinite dilution and allows for the subsequent calculation of $\ln \gamma$. Results are presented here for the simple system of NaCl, but this methodology can be applied to any electrolyte system with only minor modifications.
Acknowledgements

I would like to sincerely thank my supervisor Prof. William Smith for his guidance and unconditional support during the course of my degree. He provided bountiful advice and knowledge, without which this thesis would not be possible.

I would also like to thank my advisory committee, Drs. Hermann Eberl and Robert Wickham for their advice and guidance during the writing of this thesis, and Dr. Ashutosh Singh for serving on my examination committee.

I also extend my thanks to the Department of Mathematics and Statistics for providing a friendly and comfortable working environment, making life at the University of Guelph enjoyable and hassle-free.

I would finally like to thank my friends and family for providing much-needed distractions and life advice to keep me grounded and sane during this thesis.
Contents

Abstract ii

Acknowledgements iii

List of Tables ix

List of Figures xi

1 Introduction 1

1.1 Climate Change: A growing problem 1

1.2 Introduction to Carbon Capture 3

1.3 Carbon Capture by Reactive Absorption 4

1.4 Modelling of Candidate Solvents 6

1.5 Research Objectives 9

2 Thermodynamics of Solutions 11

2.1 Introduction to the Chemical Potential 11

2.2 Chemical Potential Models 13

2.2.1 The General Case 13

2.3 Translating Between Chemical Potential Models 15

2.3.1 \((T, P)\) and \((T, \rho)\) 15
3 Molecular Simulation

3.1 Statistical Mechanics of Molecular Simulation 23
3.2 Molecular Dynamics Simulations 29
   3.2.1 Boundary Conditions 31
   3.2.2 Long-Range Interactions 32
   3.2.3 Initialization and Energy Minimization 33
   3.2.4 Equations of Motion 34
   3.2.5 Temperature and Pressure Coupling 34
3.3 Algorithms for Calculation of Thermodynamic Properties from Simulation 35
3.4 Chemical Potential Calculations Using Molecular Simulations 37
   3.4.1 Free Energy Calculations 37
   3.4.2 Bennett Acceptance Ratio 39

4 Calculations of $\mu^\dagger$ and $\mu^{\text{res:H}}$ Using Molecular Simulation 41

4.1 Previous Attempts to Calculate $\mu^\dagger$ and $\ln \gamma$ 42
4.2 Calculations of $\mu^\dagger$ 46
   4.2.1 The Extrapolation Procedure for $\mu^\dagger_{\text{NaCl}}$ 46
   4.2.2 Low-Concentration Behaviour of Density 46
   4.2.3 Low Concentration Behaviour of $\ln \gamma_{\text{NaCl}}$ 48
   4.2.4 Accuracy of the Davies Equations 49
## List of Tables

1.1 Mixing ratios of atmospheric gases in dry air. [1] ........................................... 2
1.2 CO$_2$ emissions in megatons per year from different industrial point sources. [2] 3

2.1 Values of $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ and $\mu_{\text{NaCl}}^0(T, P^0)$ at different temperatures and pressures. $\mu_{\text{NaCl}}^0(T, P^0)$ values were obtained from JANAF. $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ were calculated using the procedures described in Appendix A using the Wagman [3] values as reference and Pitzer [4] thermodynamic data. ........................................... 22

2.2 Parameters from Rogers and Pitzer [5] used to calculate experimental activity coefficient values for low-concentration regions of aqueous sodium chloride. ........................................... 22

4.1 Values for $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ values obtained from previous attempts at calculating the solubility of NaCl in H$_2$O using Joung-Cheatham ion parameters and SPC/E water with molecular simulations. Corresponding $\mu_{\text{NaCl}}^\text{res;NVT}(T, P)$ calculated from Equation 2.50 are also shown. ........................................... 45

4.2 Linear and quadratic coefficients of each expanded ln $\gamma_{\text{NaCl}}$ (Equations 4.21 and 4.6) model calculated from experimental parameters at $T = 298.15\,\text{K}$, $373.15\,\text{K}$, and $473.15\,\text{K}$. For each model, the linear coefficient is $A_\gamma$. ........................................... 49

5.1 Force field parameters for the well depth $\epsilon$ and interaction radius $\sigma$ for the JC Na$^+$ and Cl$^-$ ions, as well as SPC/E water. ........................................... 54

5.2 $\lambda$-values used for scaling of the Lennard-Jones and Columb potentials in the NaCl+H$_2$O section of this study. ........................................... 56

6.1 Concentration-dependent terms of aqueous NaCl at dilute concentrations calculated from experimental results of Pitzer[6, 4, 5] and $\mu_{\text{NaCl}}^\dagger$ values obtained from Wagman [3]. Values of $\mu_{\text{NaCl}}^\text{res;NVT}(T, \rho[T, P]; m)$ at molalities indicated by an asterisk * were fit to linear and quadratic polynomials to extrapolate to $\mu_{\text{NaCl}}^\text{res;\infty}(T, P)$. ........................................... 58
6.2 Estimates of $\mu^{\text{res}}_{\text{NaCl}}$, $\mu^{\dagger}_{\text{NaCl}}$, and $\Delta\mu^{\dagger}_{\text{NaCl}}$ obtained from fitting experimental $\mu^{\text{res;NVT}}_{\text{NaCl}}$ values from fitting to the Davies-0.2, Davies-0.3, and Debye-Hückel equations at $m^* = 0.0111$, as used by Mester [7, 8] and $m^* = 0.0555$ kg mol$^{-1}$, as used by Young [9], as well as from fitting four low-molality points ($m = (0.0111, 0.093, 0.0056, 0.0037)$) to linear and quadratic polynomials. Uncertainties obtained from the fits are given in parentheses and are one standard deviation in the rightmost digit(s).

6.3 Estimates of $\mu^{\dagger}_{\text{NaCl}}$ and $A_\gamma$ obtained from fitting $\mu^{\text{res;NVT}}_{\text{NaCl}}$ values calculated from Wagman $\mu^{\dagger}_{\text{NaCl}}$ values to linear and quadratic polynomials in $\sqrt{m}$. Values of $\mu^{\dagger}_{\text{NaCl}}(T, P; m^0)$ were calculated from Equation 2.50. Uncertainties obtained from the fits are given in parentheses and are one standard deviation in the rightmost digits.

6.4 Simulation results of chemical potential calculations of the removal of a single NaCl particle from water boxes using Joung-Cheatham NaCl ion parameters and SPC/E water. Comparisons between chemical potential calculations using an ion pair and the sum of individual ion chemical potential calculations are presented.

6.5 Logarithmic density term and residual chemical potentials of aqueous NaCl at dilute concentrations calculated from molecular dynamics simulations. Uncertainties given in parentheses, are one standard deviation in the rightmost digit and are calculated from the results of 10 independent replicate runs.

6.6 Comparison of $\mu^{\text{res;\infty}}_{\text{NaCl}}(T, P)$, $\mu^{\dagger}_{\text{NaCl}}(T, P; m^0)$, and $\Delta\mu^{\dagger}_{\text{NaCl}}(T, P; m^0)$ obtained from fitting our simulation data using JC[10] NaCl force field parameters and SPC/E water parameters to linear and quadratic polynomials, and to the Davies and Debye-Hückel equations at $m^* = 0.0111$ kg mol$^{-1}$. We define the quadratic fit to our data to be the best available value for $\mu^{\dagger}_{\text{NaCl}}(T, P; m^0)$, thus set it as the “correct” value in calculating $\Delta\mu^{\dagger}_{\text{NaCl}}(T, P; m^0)$. Comparisons to literature values from Mester (Davies-0.2[8] and Debye-Hückel [7] at $m^* = 0.0111$ kg mol$^{-1}$) and Young (Davies-0.2[9] at $m^* = 0.0555$ kg mol$^{-1}$) are also presented. Values indicated by an asterisk in our fitting used the Young molality convention, where the molality corresponds to the midpoint of the molalities before and after removal of a particle. In addition, we also fit our data using Young’s extrapolation procedure by treating all of our simulation points as infinitely dilute.

6.7 Comparisons of $\mu^{\text{res;\infty}}_{\text{NaCl}}$ and $A_\gamma$ at 298.15K, 1 bar obtained from fitting residual chemical potentials calculated from molecular simulation with JC ion parameters and SPC/E water parameters using weighted linear and quadratic fits in $\sqrt{m}$, and corresponding $\mu^{\dagger}_{\text{NaCl}}$ values. The value indicated by double asterisks indicates our simulation data applied to Youngs’s extrapolation method treating all of our data points as a constant molality at infinite dilution. Both results using the standard molality convention, and the molality convention proposed by Young are presented in this table.
6.8 Box volumes, densities, and residual chemical potentials for boxes containing finite concentrations of NaCl in H$_2$O, calculated from molecular simulations at 1 bar, 298.15K using JC Na$^+$ and Cl$^-$, and SPC/E H$_2$O. The line in each molality containing $\infty$ NaCl and H$_2$O corresponds to the extrapolated finite concentration chemical potential. 72

6.9 ln $\gamma_{NaCl}$ calculated from $\mu_{NaCl}^{res,NVT}(T, \rho[T, P]; m)$ values obtained from molecular simulation using JC ion parameters and SPC/E water parameters at 298.15K, 1 bar. Uncertainties are to one standard deviation in the rightmost digits. Due to the small uncertainties in density, the uncertainties in the density term can be neglected. 72

6.10 Fitting parameters for activity coefficients using an extended Debye-Hückel equation for the JC NaCl force field parameters at 298.15K, 1 bar. 73
List of Figures

1.1 Carbon capture by reactive absorption. ........................................... 5
1.2 Primary amines being considered as reactive absorption solvents for carbon capture.[11] ................................................................. 7

3.1 A simplified schematic of a molecular dynamics algorithm. ............ 30
3.2 A 3D simulation box of sodium (blue) and chloride (teal) ions dissolved in water (hydrogen - white; oxygen - red). Such simulation boxes are finite in size and do not accurately represent macroscopic systems. ................. 31
3.3 Two layers of pure water (blue) and a layer of Na⁺ (yellow) and Cl⁻ (green) ions in water separated by semipermeable membrane. Water molecules in the simulation are allowed to pass through the membranes, but the solute is confined to the space in between the membranes. This set-up is used to calculate osmotic pressure [12]. ................................................................. 36

4.1 Molality-dependence of chemical potentials for Joung-Cheatham (green), Horinek [13](red), FF0 (blue), and FF1 (purple) NaCl force fields in comparison to the experimental value (black solid). Intersection of the chemical potential curve with the crystalline chemical potential (black dashed line) is used to calculate solubility of NaCl [14]. ................................................................. 42
4.2 Comparison of ln $\gamma_{\text{NaCl}}$(298.15 K, 1 bar; m) models in dilute solutions. Below 0.1 molal, the Davies equation with $c = 0.3$ is closest to the experimental Pitzer model. ................................................................. 50

6.1 Plots of experimental residual chemical potential values (diamond, left axis) and the logarithmic density term (black circles, right axis), to infinite dilution. Calculations used Equation 6.1 with Wagman $\mu^\dagger_{\text{NaCl}}(T,P;m^0)$ values, NIST-JANAF values for $\mu^0_{\text{NaCl}}$, and Pitzer [4] data for density and ln $\gamma_{\text{NaCl}}$ data. Points indicated by magneta diamonds were used in fitting to calculate $\mu^\text{res;=}_{\text{NaCl}}(T,P)$. ................................................................. 59
6.2 Linear (magenta curve) and quadratic (cyan curve) fits to obtain $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, P)$ from experimental $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m)$ values (blue curve), in comparison to $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, P)$ values calculated from the Davies-0.2 (black curves), Davies-0.3 (red curves) and Debye-Hückel equations (purple curves). Points indicated by circles were used to fit to linear and quadratic polynomials. Points indicated by diamonds were used to calculate $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, P)$ using the $\ln \gamma_{\text{NaCl}}$ approximations ($m^* = 0.0111$ kg mol$^{-1}$ indicated by solid lines; $m^* = 0.0555$ kg mol$^{-1}$ indicated by dashed lines). Subfigure (a) shows the full concentration range used, while Subfigure (b) is zoomed to better show the intercept.

6.3 Experimental values (blue curve) for $\Delta \mu_{\text{NaCl}}(T, P; m) = \mu_{\text{NaCl}}^{\text{approx}}(T, P; m) - \mu_{\text{NaCl}}^\text{corr}(T, P; m)$ obtained from using a quadratic fit (cyan curve), linear fit (magenta curve), and Davies-0.2 (black curves), Davies-0.3 (red curves), and Debye-Hückel (purple curves) equations to approximate $\ln \gamma_{\text{NaCl}}(T, P; m)$, and $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, P; m^0)$ calculated from each respective equation at $m^*_{\text{NaCl}} = 0.0555$ kg mol$^{-1}$ (Subfigure (a)) and $m^*_{\text{NaCl}} = 0.0555$ kg mol$^{-1}$ (Subfigure (b)). At $m = 0$ kg mol$^{-1}$, $\Delta \mu_{\text{NaCl}} = \Delta \mu_{\text{NaCl}}^\dagger$. 

6.4 Quadratic (cyan curve) and linear (magenta curve) fits of simulated $\mu_{\text{NaCl}}^{\text{res;NVT}}$ values calculated using JC ion parameters and SPC/E water parameters. Comparisons to the Davies-0.2 (black curve), Davies-0.3 (red curve), and Debye-Hückel (purple curve) are also presented. Uncertainty in points is to one standard deviation, and is calculated from 10 independent simulations. Uncertainties in the quadratic and linear fits are shown at the indicated abscissa values.

6.5 Activity coefficients calculated from Equation 4.29 in molecular dynamics simulations using the Joung-Cheatham NaCl force field and SPC/E water parameters. Results from this work (red) are compared to activity coefficient calculations from Mester and Panagiotopoulos [8], reanalyzed by Young and Panagiotopoulos (black) [9].
Chapter 1

Introduction

The work presented in this thesis is part of the Simulation, Modelling, and Experiment (SiModEx) project, a joint collaboration between the University of Guelph, Institut de Chimie de Clermont-Ferrand, and Gas Liquids Engineering Ltd. The project aims to utilise molecular simulation and macroscopic chemical thermodynamics experiments to develop thermodynamic models to screen candidate solvents for their suitability in post-combustion carbon capture in an effort to curb CO$_2$-induced climate change.

1.1 Climate Change: A growing problem

An enormous problem facing modern society is anthropogenic climate change, primarily driven by advancements in technology which have helped with the harvesting and utilization of fossil fuels. While the global temperature has been fluctuating for millions of years, the rapid introduction of human-made greenhouse gases into the atmosphere has resulted in an era of unprecedented warming of the Earth’s atmosphere and oceans. Thus, as far back as the mid-to-late 20$^{th}$ century, there has been a serious effort to model the impact of ever-increasing atmospheric greenhouse gas levels.

Compared to the major components of air (N$_2$, O$_2$, Ar), greenhouse gases are present only in trace amounts. Table 1.1 shows the concentrations of atmospheric gases in dry air. Despite being present in such low concentrations, gases like CO$_2$, CH$_4$, and N$_2$O are highly effective at absorbing infrared radiation. While CH$_4$ and N$_2$O are much more effective at trapping radiation, CO$_2$ is present in much higher concentrations, so it is considered a major contributor to the greenhouse effect.

Early studies have shown that a doubling of atmospheric greenhouse gas levels will effec-
Gas               Mole fraction
N₂               0.78
O₂               0.21
Ar               0.0093
CO₂              365×10⁻⁶
O₃               10×10⁻⁶
CH₄              1.7×10⁻⁶
N₂O              320×10⁻⁹

Table 1.1: Mixing ratios of atmospheric gases in dry air. [1]

tively increase global temperatures by upwards of 2°C relative to temperatures at the time [15]. More recent reports show that while maximum efforts worldwide to mitigate greenhouse gas emissions can minimize the temperature increase to than less than 1°C relative to 1986 temperatures, a failure to curb emissions is predicted to cause an increase of 4°C in the average global temperature by 2100 [2].

While a 2°C-4°C warming may not seem significant, the effects on the Earth's delicately balanced ecosystems can be quite drastic. Among currently-observed effects are an increase in forest fires [16], acidification of the world’s oceans [17], and accelerated rising of global sea levels [18]. Long term effects of an unregulated climate change are predicted to be dire. Flooding of low-lying population centres will result in the displacement of millions of humans from their homes and a large increase in world hunger due to loss of arable land [19]. Acidification of the world’s oceans will lead to a disruption of marine ecosystems, including the death of phytoplankton, which produce a large portion of atmospheric oxygen [20].

The mechanism by which these gases contribute to climate change is known as the greenhouse effect. Solar radiation enters the Earth’s atmosphere in the ultraviolet to near-infrared range. As the wavelength of emitted radiation is affected by the temperature of the emitting body, the absorbed solar radiation originating from the hot sun is re-emitted in the form of infrared radiation from the relatively cooler Earth. This infrared radiation matches the frequency of vibrational and rotational modes within select atmospheric gases, so it can be absorbed by the molecule, effectively trapping the energy in the Earth’s atmosphere [1].

Thus, in recent years there has been a push to mitigate climate change by reducing carbon emissions. Due to the diversity of CO₂-emitting sources (major contributors to CO₂ emissions can be seen in Table 1.2), no one solution is able to single-handedly curb carbon emissions. In recent years, there has been a shift to renewable zero-carbon or carbon-neutral energy sources, such as nuclear fission, hydroelectric, wind, and solar [21]. On the
consumer side, there has been an increase in electric vehicles [22], and a shift away from
carbon-intensive animal-sourced foods [23].

<table>
<thead>
<tr>
<th>CO₂ Source</th>
<th>CO₂ Emissions (MtCO₂ yr⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power</td>
<td>10539</td>
</tr>
<tr>
<td>Cement production</td>
<td>932</td>
</tr>
<tr>
<td>Refineries</td>
<td>798</td>
</tr>
<tr>
<td>Iron and steel industry</td>
<td>646</td>
</tr>
<tr>
<td>Petrochemical industry</td>
<td>379</td>
</tr>
<tr>
<td>Oil and gas processing</td>
<td>50</td>
</tr>
<tr>
<td>Other sources</td>
<td>33</td>
</tr>
</tbody>
</table>

Table 1.2: CO₂ emissions in megatons per year from different industrial point sources. [2]

However, current efforts are falling short. Implementation of alternative energy sources
has been slow due to a lack of technology, government roadblocks, and the dependence of
major economies on oil [24]. Thus, there is a need for shorter-term, cost-effective solutions.

1.2 Introduction to Carbon Capture

As CO₂ concentrations are highest at the source of emissions, one of the most effective
methods for reducing atmospheric carbon is removing CO₂ from industrial point sources
using a group of processes known as carbon capture. There are three main types of carbon
capture methods for these types of facilities: (1) pre-combustion carbon capture, (2) oxyfuel
combustion, and (3) post-combustion capture.

In pre-combustion carbon capture, carbon-based fuels are partially oxidized before being
combusted. This produces synthesis gas, a mixture of H₂, CO₂, CO, and other trace gases.
CO can be further reacted with water to form H₂ and CO₂. CO₂ can be separated out from
the mixture, while the H₂ is used as a clean fuel. While this is an effective method, capital
costs are high, so pre-combustion methods are not widely implemented [25].

Carbon capture by oxy-fuel combustion is an emerging technology for greenhouse gas
mitigation. In this process, O₂ is separated from the air before combustion. Burning fossil
fuels in an environment rich in oxygen and lean in nitrogen allows for more complete com-
bustion of the fuel into CO₂, and fewer undesirable compounds like NOₓ. The combusted
gases are recycled to continually further purify the stream. The resulting flue stream consists
of highly purified CO₂ and water vapour. The water can be condensed, while the CO₂ does
not need further separation before storage [25].
Post-combustion carbon capture is a method in which CO$_2$ is removed from a flue stream after combustion with air. Current methods for separating CO$_2$ from flue gas involve chemical, physical, and hybrid processes. Chemical processes involve reactive absorption with amine-based solvents or ionic liquids. Physical processes for sequestration utilise porous metallo-organic frameworks or cryogenic separation [25]. Post-combustion carbon capture is the most commonly implemented carbon capture methodology in industry today. This is mainly due to the ability to retrofit the technology to existing flue streams with relative ease and at a lower cost than other methods.

1.3 Carbon Capture by Reactive Absorption

In carbon capture by reactive absorption, CO$_2$-rich flue gas enters an absorber column containing a solvent. In the absorber column, CO$_2$ reacts with the solvent at a low temperature and high pressure, forming carbamate and a protonated amine, removing CO$_2$ from the flue stream. The solution then enters a stripper column, where it is heated at a low pressure. This process drives the reaction in the reverse direction, separating out the CO$_2$, which is compressed and stored or used for industrial purposes. This step also regenerates the original solvent to be used again [26]. This process can be seen in Figure 1.1. In current implementations, aqueous monoethanolamine (MEA) solutions are the most common solvent used due to its effectiveness at removing CO$_2$ from emissions, and also due to historic reasons, as it was initially used to remove H$_2$S from natural gas. However, the solvent regeneration process is extremely energy intensive, making up approximately 70-80% of the total energy cost of the entire process [27].

This creates a need for better-performing solvents if post-combustion carbon capture is to be widely implemented, as solvents with increased solubility or less energy-intensive regeneration steps would increase efficiency and lower operating costs. However, with a large number of candidate solvents to be possibly considered, implementing production-scale testing is infeasible. Thus, there is currently a focus on screening compounds to predict their efficacy as reactive solvents before large-scale testing.

The most common method of screening candidates utilises experiments, whose data are fitted to macroscopic thermodynamic models, such as the universal quasichemical model (UNIQUAC) [28] and the non-random two liquid model (NRTL) [29]. However, these models are empirical in nature and rely primarily on correlation rather than prediction, so the complexity of candidate solvents results in questionable accuracy when used to predict prop-
Figure 1.1: Carbon capture by reactive absorption.

In recent years, increased computing power has brought computers to the forefront of science. As a result, there has been a focus on the development and implementation of simulation tools to study molecular systems [31] [32] [33]. Using such tools, it is possible to create more fundamental, theoretically-grounded models which can be used for solvent screening. Parameter sets consisting of atomic charges, masses, and interactions can be used to predict a variety of thermodynamic properties without the need for only limited amounts of experimental data.

Computational studies of MEA-CO$_2$-H$_2$O reactive absorption are thus far limited. Balaji et al. [34] used Reaction Ensemble Monte Carlo (RxMC) simulations combined with the continuous fractional component method (CFC) [35] to study the equilibrium compositions of the system. They note that for a complete description of equilibrium concentrations and speciation, all elementary reactions for the system must be considered. Their results show that the solvation of products and reactants is an essential component of the reactive absorption mechanism. However, as noted by Noroozi and Smith [30], this study neglected
significant contributions to the species chemical potentials from the respective ideal gas reference states, and the RxMC/CFC methodology does not allow for the calculation of minor species concentrations, nor the partial pressure of CO$_2$, without long simulation times or large system sizes.

Noroozi and Smith [30] developed a computationally efficient molecular-based chemical reaction equilibrium algorithm which can be used to calculate concentrations of even minor species present in the reactive absorption process of MEA. In addition, this algorithm is usable in conjunction with commonly-used molecular dynamics software packages, which can further benefit from GPU parallelization and require little or no modifications to use for this system.

1.4 Modelling of Candidate Solvents

Developing methodology to screen candidate solvents is not a straightforward process. The calculation of macroscopic thermodynamic properties from molecular simulation requires careful translation of properties between chemical potential models applicable to the respective reference systems.

Species of interest for carbon capture are structurally similar to monoethanolamine. Candidates possess the same N-C-C-O backbone present in MEA. Primary amines of interest were identified by adding methyl or ethyl groups to the carbon backbone atoms. A selection of candidate primary amines can be seen in Figure 1.2. In addition, secondary amines such as diethanolamine (DEA), and tertiary amines such as methyl diethanolamine (MDEA) and triethanolamine (TEA) were also of interest.

Alkanolamines are the candidate solvents for this project for a number of reasons. The presence of the hydroxyl (OH$^-$) group allows for hydrogen bonding with water, increasing surface tension, decreasing volatility, and therefore reducing solvent loss during the regeneration step. The amine group found in primary and secondary alkanolamines allows for carbamate formation in the presence of CO$_2$. These carbamates are generally quite stable, thus the high energy costs of regeneration. However, steric hindrance in some alkanolamines such as 2-amino-2-methyl-1-propanol favours the production of unstable carbamates from which CO$_2$ can be more easily recovered [36].

Modelling the thermodynamics of the species found in the carbon capture process requires the calculation of equilibrium compositions and related properties. The reaction equilibrium of a system containing $N$ distinct chemical species is determined by minimizing the Gibbs
free energy $\Delta G(T, P; x)$ at a temperature $T$, pressure $P$, and mole fraction $x$, subject to conservation of mass balance, resulting in a set of $R$ linearly independent stoichiometric equations:

$$\sum_{i=1}^{N} \nu_{ij} A_i = 0; j = 1, 2, ..., R$$

(1.1)

where $\nu_{ij}$ is the stoichiometric coefficient of a chemical species $A_i$ in reaction $j$, and:

$$R = N - \text{rank}(A),$$

(1.2)

where $A$ is a chemical formulae matrix of dimension $M \times N$ for $M$ elements.

Thus, it is necessary to create a set of chemical equations representing the chemical equilibrium reactions in the amine-H$_2$O-CO$_2$ mixture. Speciation studies of the reactive absorption mechanism of MEA have shown the presence of protonated MEA, carbamate, carbonate, bicarbonate, hydroxide, and hydronium ions [37]. A set of proposed reactions for
any arbitrary primary alkanolamine R is:

\[ 2RNH_2 + CO_2 = RNH_3^+ + RNHCOO^- \]
\[ RNH_3^+ + H_2O = H_3O^+ + RNH_2 \]
\[ RNHCOO^- + H_2O = RNH_2 + HCO_3^- \]
\[ RNH_2 + H_2O = RNH_3^+ + OH^- \]

It is important to note that these equations do not necessarily represent reaction mechanisms. However, it is also important to note that equilibrium composition of the system is independent of a reaction mechanism as it relates only to minimizing the free energy of the system.

To determine the equilibrium composition of the system, we minimize the Gibbs free energy, the first-order conditions for which are:

\[ \Delta G_j(T, P; x) \equiv \sum_{i=1}^{N} \nu_{ij} \mu_i(T, P; x) = 0 \quad ; \quad j = 1, 2, ..., R \]  

(1.3)

where \( \nu_{ij} \) is the stoichiometric coefficient of species \( i \) in reaction \( j \), and \( \mu_i \) is the chemical potential of species \( i \), a property used to quantify the free energy change per mole of species \( i \).

The chemical potential is an important thermodynamic property which is required to determine other important properties such as solubility limits and vapour pressures of species in solution. It can be expressed using a large variety of models, but all models quantify it as the sum of an ideal chemical potential model and a residual term:

\[ \mu_i = \mu_i^{id} + \mu_i^{res}. \]  

(1.4)

Three commonly-used models for the chemical potential are:

\[ \mu_i(T, P; x) = \mu_i^0(T, \rho^0) + RT \ln \left( \frac{\rho_i}{\rho^0} \right) + \mu_i^{res;\rho}(T, \rho; x) \]  

(1.5)

\[ \mu_i(T, P; x) = \mu_i^0(T, P^0) + RT \ln \left( \frac{x_i P}{P^0} \right) + \mu_i^{res;P}(T, P; x) \]  

(1.6)

\[ \mu_i(T, P; m) = \mu_i^\dagger(T, P) + RT \ln m_i + \mu_i^{res;H}(T, P; m) \]  

(1.7)

In each model, the first two terms on the right hand side of the equation are the underlying
ideality model, while the residual chemical potential term accounts for non-ideal interactions. Equations 1.5 and 1.6 both entail the use of an underlying ideal-gas chemical potential model, which have reference state chemical potentials of its pure ideal gas at \((T, \rho^0)\) and \((T, P^0)\) respectively. Equation 1.5 is a useful chemical potential model in statistical mechanics, and therefore in molecular simulation.

Equation 1.7 utilises an underlying Henry ideal solution model, where \(\mu^\dagger_i(T, P)\) is the chemical potential of a hypothetical ideal 1 molal solution of species \(i\). This chemical potential model is commonly used in experimental macroscopic thermodynamics of dilute electrolyte solutions, which are characterized by relatively low solute concentrations. As the species present in the alkanolamine reactive absorption reaction sets are all present in low concentrations, this model is useful for modelling their thermodynamics, and thus will be an important focus of this thesis.

### 1.5 Research Objectives

Of direct importance to the carbon capture project is the calculation of equilibrium compositions of the reaction of \(\text{CO}_2\) in aqueous alkanolamine solutions. As these species, their products, and their intermediates are present in dilute concentrations, it is convenient to implement the Henry chemical potential model found in Equation 1.7, thus knowledge of the reference standard state chemical potential \(\mu^\dagger_i(T, P)\) is required.

Until recently, there has been little progress in the statistical mechanics, and subsequently molecular simulation, of electrolyte solutions. The work presented in this thesis focuses on the development of a methodology to accurately and precisely calculate \(\mu^\dagger_i(T, P)\) and \(\ln \gamma_i(T, P; m)\) of each species \(i\) in the reaction solution through the use of molecular simulation in order to determine equilibrium compositions of the reaction of \(\text{CO}_2\) with aqueous alkanolamines. Our goal is to improve upon previous methods by calculating these properties with better accuracy and precision. As the ions found in the aforementioned reaction sets are quite large and complex, we focus on developing and testing this methodology using a simpler 1-1 electrolyte solution of \(\text{Na}^+\) and \(\text{Cl}^-\).

Chapter 2 of this thesis explores the macroscopic thermodynamics of solutions, leading up to describing the chemical potential models of electrolyte solutions. Chapter 3 covers the fundamentals of molecular simulation and how we can use it to calculate chemical potentials. Chapter 4 gives a literature review of previous attempts to calculate \(\mu^\dagger(T, P)\) and \(\mu_{\text{res:H}}(T, P; m)\) of aqueous NaCl from molecular simulation, and describes a new methodolog-
ical approach to calculating these quantities from the extrapolation of low-molality residual chemical potentials to infinite dilution, and the extension to calculate other properties, such as activity coefficients and solubilities. Chapter 5 describes the simulation parameters used in each step of the methodology described in the previous chapter. In Chapter 6, we present and discuss the results of our new approach to calculating $\mu^\dagger$ and the extension to calculating finite-concentration solution properties. Finally, in Chapter 7 we present concluding remarks and future recommendations.
Chapter 2

Thermodynamics of Solutions

In this chapter, we discuss aspects of chemical potential as it pertains to the thermodynamics of solutions. We start with a general definition of the chemical potential and we next consider how to model the chemical potential of species in ideal and non-ideal solutions, both in a general case and more specific forms. Finally, we explore how to translate between different chemical potential models and the application to electrolyte solutions. This last step is crucially important in translating simulation quantities to the corresponding experimental quantities.

2.1 Introduction to the Chemical Potential

The chemical potential is one of the most useful thermodynamic properties, as its knowledge is required to calculate a wide variety of other thermodynamic properties such as vapour pressure and solubility. This section uses a thermodynamic approach to defining chemical potential, first defining it from the internal energy $U$, before extending it to other thermodynamic potentials such as enthalpy $H$, Helmholtz free energy $A$, and Gibbs free energy $G$.

The internal energy is used to describe a closed, adiabatic, isochoric system. From the first law of thermodynamics, we can write a general form of the total internal energy $U = U(S, V; n)$ of any system containing $N$ distinguishable particle types as

$$dU(S, V; n) = TdS - PdV + \sum_{j=1}^{N} \mu_j dn_j.$$  \hspace{1cm} (2.1)
where $T, S, P,$ and $V$ are the system’s temperature, entropy, pressure, and volume, and $\mu_i$ and $n_i$ are the chemical potential and mole number of species $i$.

At constant entropy and volume, we can use Equation 2.1 to define chemical potential $\mu_i$ of a gas as:

$$\mu_i = \left( \frac{\partial U}{\partial n_i} \right)_{S,V,j \neq i}.$$  

(2.2)

Thus, chemical potential of a species $i$ can be defined as the change in internal energy of a system due to a change in the amount of species $i$.

However, this definition of chemical potential can be unwieldy, especially in experimental scenarios, due to difficulties in controlling entropy. It becomes convenient to apply Legendre transforms to switch the thermodynamic potential depending on a variable (e.g. entropy) to a new potential depending on a conjugate variable (e.g. temperature).

Applying the definition $H(S, P; n) = U + PV$, we get the following differential for enthalpy:

$$dH(S, V; n) = TdS + VdP + \sum_{j=1}^{N} \mu_j dn_j.$$  

(2.3)

Similarly, we can apply the definitions for Helmholtz free energy $A(T, V; n) = U - TS$ and Gibbs free energy $G(T, V; n) = A + PV = U - TS + PV$, resulting in:

$$dA(T, V; n) = -SdT - PdV + \sum_{j=1}^{N} \mu_j dn_j$$  

(2.4)

and

$$dG(T, P; n) = -SdT + VdP + \sum_{j=1}^{N} \mu_j dn_j,$$  

(2.5)

respectively.

Each of these differentials can then be used to express a thermodynamic definition relating
chemical potential to the original state functions, as seen in Equations 2.6 through 2.8:

\[ \mu_i = \left( \frac{\partial H}{\partial n_i} \right)_{S,P,n_j \neq i} \]  
\[ \mu_i = \left( \frac{\partial A}{\partial n_i} \right)_{T,V,n_j \neq i} \]  
\[ \mu_i = \left( \frac{\partial G}{\partial n_i} \right)_{T,P,n_j \neq i} \]  

For this thesis, the chemical potential obtained from the Helmholtz free energy in Equation 2.7 is the most important, as it is accessible through the use of molecular simulation.

### 2.2 Chemical Potential Models

#### 2.2.1 The General Case

Reality rarely follows ideality, so we must consider the chemical potential of species in non-ideal solutions. Due to interactions between species that are assumed to be non-existent in ideal systems, the chemical potential of a species will be different than the ideal chemical potential of the same species under the same conditions. We can express a general form of a chemical potential model of any non-ideal species \( i \) as the sum of a reference state and an activity term, as seen in Equation 2.9:

\[ \mu_i(\phi_1, \phi_2; z) = \mu_i^\ominus(\phi_1, \phi_2; z^0) + RT \ln a_i(\phi_1, \phi_2; z), \]  

where the activity \( a_i \) can be expressed as the product of the system composition and an activity coefficient:

\[ a_i(\phi_1, \phi_2; z) = z_i \gamma_i(\phi_1, \phi_2; z). \]  

This results in the total chemical potential expression:

\[ \mu_i(T, P; z_i) = \mu_i^\ominus(\phi_1, \phi_2; z^0) + RT \ln \frac{z_i}{z_i^0} + RT \ln \gamma_i(\phi_1, \phi_2; z) \]
The first two terms are simply the ideality model for the chemical potential. We further define the activity coefficient in general as:

$$\mu_{\text{res}}(\phi_1, \phi_2; z) \equiv RT \ln \gamma_i(\phi_1, \phi_2; z).$$  \hspace{1cm} (2.12)

Next, we must choose our thermodynamic variables, which depend on our choice of the underlying ideality model and the corresponding reference state chemical potential. We can choose any thermodynamic variables $\phi_1$ and $\phi_2$, any composition variable $z$, and any reference state value for the chemical potential $\mu_i^{\ominus}$ at the state $(\phi_1, \phi_2; z)$. The most common choices for thermodynamic variables are $T$ and $P$, as they are more easily controllable in a lab.

For the composition variable, convenient choices are generally mole fraction $x_i$:

$$x_i = \frac{n_i}{n_{\text{total}}} \hspace{1cm} (2.13)$$

$$= \frac{n_i}{n_i + n_{\text{solv}}} \hspace{1cm} (2.14)$$

where $n_i$ and $n_{\text{solv}}$ are the number of particles or moles of the solute and solvent, respectively, and $n_{\text{total}} = n_i + n_{\text{solv}}$; density $\rho_i$:

$$\rho_i = \frac{n_i}{V_{\text{soln}}} \hspace{1cm} (2.15)$$

$$= \frac{x_im_{\text{total}}}{V_{\text{soln}}} \hspace{1cm} (2.16)$$

$$= x_i \rho_{\text{soln}} \hspace{1cm} (2.17)$$

where $V_{\text{soln}}$ is the volume of the solution; or molality $m$:

$$m_i = \frac{n_i 1000}{n_{\text{solv}} M_{\text{solv}}} \hspace{1cm} (2.18)$$

where $n_{\text{solv}}$ and $M_{\text{solv}}$ are the number of moles or particles and the molecular weight of the solvent in kg mol$^{-1}$. In the limit $z \to z^0$, $a_i(\phi_1, \phi_2; z) \to 1$ and the chemical potential is equal to that of the reference state chemical potential.

The choice of the ideality model and reference state chemical potential depend on the system to be studied. For example, as molecular simulation is rooted in statistical mechanics, chemical potential models used in simulation are often of the form of Equation 1.5, where the underlying ideality model is an ideal gas mixture of density $\rho$. In the case of macroscopic
thermodynamics, the Henry (\(T, P\)) chemical potential model from Equation 1.7 most often used.

Comparing these equations to Equation 2.12, we define:

\[
\mu^\text{res,}\rho_i(T, \rho; x) = RT \ln \gamma^\rho_i(T, \rho; x) \quad (2.19)
\]

\[
\mu^\text{res,H}_i(T, \rho; m) = RT \ln \gamma^H_i(T, \rho; m) \quad (2.20)
\]

with limiting behaviour:

\[
\lim_{x_i \to 0^+} \ln \gamma^\rho_i(T, \rho; x) = 0 \lim_{m_i \to 0^+} \ln \gamma_i(T, \rho; m) = 0 \quad (2.21)
\]

### 2.3 Translating Between Chemical Potential Models

To be able to translate quantities obtained from simulation to those from experimental results, we must be able to translate between different chemical potential models. The basis for the ability to translate between models lies in the "state consistency principle." This principle states that for any state point at a given composition, when any two quantities in the set of (\(\rho, T, P\)) are specified, the third is also determined. This allows chemical potentials to be equated.

#### 2.3.1 \((T, P)\) and \((T, \rho)\)

It is sometimes convenient to translate between the \((T, P)\) chemical potential model and the \((T, \rho)\) model. We start by equating 1.5 and 1.6 for a pure ideal gas. As both models use an underlying ideal-gas model, we can neglect the residual chemical potentials in the conversion:

\[
\mu_i = \mu_i^\phi(T; \rho^0) + RT \ln \left( \frac{x_i \rho}{\rho^0} \right) = \mu_i^0(T; P^0) + RT \ln \left( \frac{x_i P}{P^0} \right). \quad (2.22)
\]

From here, we can use the state consistency principle, which allows us to substitute in the ideal-gas equation of state \(P = \rho RT\):

\[
\mu_i^\phi(T; \rho^0) + RT \ln \left( \frac{x_i \rho}{\rho^0} \right) = \mu_i^0(T; P^0) + RT \ln \left( \frac{x_i \rho RT}{P^0} \right). \quad (2.24)
\]
Solving for $\mu^\phi_i$:

$$\mu^\phi_i(T; \rho^0) = \mu^0_i(T; P^0) + RT \ln \left( \frac{RT \rho^0}{P^0} \right) \quad (2.25)$$

For a non-ideal system this yields:

$$\mu_i = \mu^0_i(T; P^0) + RT \ln \left( \frac{RT \rho^0}{P^0} \right) + RT \ln \left( \frac{\rho_i}{\rho^0} \right) + \mu_{i,\text{res};T,\rho,x}(T, \rho, x) \quad (2.26)$$

Equating this result with Equation 1.6, we derive an equation to translate between $\mu_{i,\text{res};\rho}$ and $\mu_{i,\text{res};P}$:

$$\mu_{i,\text{res};\rho}(T, \rho[T, P]; x) = RT \ln \left( \frac{P}{\rho RT} \right) + \mu_{i,\text{res};P}(T, P; x) \quad (2.27)$$

Care must be taken with the second density term: $\rho^0$ is required to have the same units as $\rho$, so the choice of units for the reference density will affect the value of the term. For example, at $T = 298.15$K with $\rho^0$ in mol L$^{-1}$, this term becomes:

$$RT \ln \left( \frac{RT \rho^0}{P^0} \right) = \frac{8.31446 \times 298.15}{1000} \ln \left( \frac{8.31446 \times 298.15 \times 1000}{10^5} \right) \quad (2.28)$$

$$= 7.958 \text{ kJmol}^{-1}$$

and for $\rho^0$ in units of mol m$^{-3}$, the term becomes:

$$RT \ln \left( \frac{RT \rho^0}{P^0} \right) = \frac{8.31446 \times 298.15}{1000} \ln \left( \frac{8.31446 \times 298.15}{10^5} \right) \quad (2.29)$$

$$= -9.165 \text{ kJmol}^{-1}.$$

### 2.3.2 Henry Model and $(T, \rho)$

The most important translation for the context of this thesis arises from linking the Henry law chemical potential model to the $(T, \rho)$ model, as it provides a relationship between the Henry standard state chemical potential $\mu^\dagger_i(T, P)$ and residual chemical potentials obtainable from molecular simulation. We can equate the two models which yields:

$$\mu^\dagger_i(T, P; m^0) + RT \ln \frac{m_i}{m^0} + \mu_{i,\text{res};H}(T, P; m) = \mu^0_i(T, P^0) + RT \ln \left( \frac{RT x_i \rho_{\text{soln}}}{P^0} \right) + \mu_{i,\text{res};\rho}(T, \rho[T, P]; m) \quad (2.30)$$
Rearranging to solve for $\mu_i^\dagger(T, P; m^0)$ and using the definitions of $x_i, m_i$ and $\rho_i$, we get:

$$\mu_i^\dagger(T, P; m^0) = \mu_i^0(T; P^0) + RT \ln \left( \frac{n_{\text{solv}} M_{\text{solv}} \rho_{\text{solv}} m^0}{1000(n_i + n_{\text{solv}}) P^0} \right) + \mu_{\text{res},T,\rho}(T, \rho; x) - \mu_{\text{res},H(T,P;m)}(T, P; m)$$

(2.31)

Next, we consider the behaviour of this for low concentrations of species $i$. In the infinite dilution limit of species $i$, $\mu_{\text{res},H(T,P;m)}(T, P; m) \rightarrow 0^+$ so:

$$\mu_i^\dagger(T, P; m^0) = \mu_i^0(T, P^0) + \Delta G_{i}^{\text{hyd}}(T, P)$$

(2.32)

$$\mu_i^\dagger(T, P; m^0) = \mu_i^0(T, P^0) + RT \ln \left( \frac{RT M_{\text{solv}} \rho_{\text{solv}}^0}{1000 P^0} \right) + \mu_{\text{res},\infty}^i(T, \rho_{\text{solv}}[T, P])$$

(2.33)

where $\Delta G_{i}^{\text{hyd}}(T, P)$ is the hydration free energy, or the energy requirement of inserting a single particle of species $i$ from the ideal gas phase at 1 bar per mol into the solution at unit molality, and $\mu_{\text{res},\infty}^i(T, P)$ is the infinite dilution or intrinsic chemical potential of species $i$. This equation proves useful: we now have an equation from which we can directly calculate $\mu_i^\dagger(T, P; m^0)$ using molecular simulation, provided we can calculate $\mu_{\text{res},\infty}^i(T, P)$.

To do this, we substitute our equation for $\mu_i^\dagger$ into Equation 2.31, yielding:

$$\mu_{\text{res},\rho}^i(T, \rho[T, P]; x) = \mu_{\text{res},\infty}^i(T, \rho[T, P]) - RT \ln \left( \frac{x_{\text{solv}} \rho_{\text{solv}}(T, P; m)}{\rho_{\text{solv}}^0(T, P)} \right) + RT \ln \gamma_i(T, P; m)$$

(2.34)

In the limit $m \rightarrow 0^+$, the last two terms on the right hand side of the equation go to 0, so this extrapolation yields an intercept equal to $\mu_{\text{res},\infty}^i(T, P)$. After calculating $\mu_{\text{res},\infty}^i(T, P)$, we can also use this equation to calculate the Henry-law activity coefficients from molecular simulation.

We can then finally insert back into the Henry law chemical potential model, resulting in:

$$\mu_i(T, P; m) = \mu_i^0(T, P^0) + RT \ln \left( \frac{RT \rho_{\text{solv}}(T, P) M_{\text{solv}} \rho^0 m^0}{1000 P^0} \right) + \mu_{\text{res},T,\rho,\infty}^i(T, \rho_{\text{solv}}) + RT \ln(m_i) + \mu_{\text{res},H(T,P;m)}(T, \rho[T, P]; m)$$

(2.35)
2.4 Chemical Potentials of Electrolyte Solutions

In determining the equilibrium compositions of the reactions between candidate alkanolamines, carbon dioxide, and water, there are many ionic species in the products of these reactions. The application of the chemical potential models and translations described in the previous sections to solutions containing electrolytes requires some adjustment to account for effects due to ionic charge and also due to the dissociation of compounds into their ions. One mole of an electrolyte $M_{\nu_i}X_{\nu_j}$ dissociates into ions in solution:

$$MX = M_{\nu_i}^{z_i^+} + X_{\nu_j}^{z_j^-}$$

and has a charge neutrality constraint of:

$$\nu_i^{z_i^+} + \nu_j^{z_j^-} = 0$$

producing $\nu_i$ moles of $M^{z_i^+}$ and $\nu_j$ moles of $X^{z_j^-}$.

2.4.1 Composition Variables

First and foremost, we need to modify the composition variables. For a 1-1 electrolyte MX which dissociates into $M^{z_i^+}$ and $X^{z_j^-}$, we must consider the simultaneous presence of both ions in an electrically neutral solution. Taking this into account, the mole fraction of an ion $x_i$ as:

$$x_i = \frac{\nu_i}{n_{\text{total}}} = \frac{n_i}{\nu_i + \nu_j + n_{\text{solv}}}.$$ 

(2.38)

The denominator is modified due to the presence of $\nu_i$ $M^{z_i^+}$ ions and $\nu_j$ $X^{z_j^-}$ ions in the solution. We also note that the mole fraction of the solvent is given by:

$$x_{\text{solv}} = \frac{n_{\text{solv}}}{n_{\text{total}}} = \frac{n_{\text{solv}}}{n_i + n_j + n_{\text{solv}}}.$$ 

(2.39)

Other composition variables such as molality and density are modified by multiplying the density of the electrolyte by
2.4.2 Electrolyte Solutions

We can write the Henry-law chemical potential of each individual ion in as:

\[
\begin{align*}
\nu^+ \mu_{M^+}(T, P; m) &= \nu^+ \left[ \mu^\dagger_{M^+}(T, P; m^0) + RT \ln m_+ + RT \ln \gamma_+(T, P; m) \right] \\
\nu^- \mu_{X^-}(T, P; m) &= \nu^- \left[ \mu^\dagger_{X^-}(T, P; m^0) + RT \ln m_- + RT \ln \gamma_-(T, P; m) \right]
\end{align*}
\]  

(2.40)

(2.41)

where \( m_+ = \nu_+ m \) and \( m_- = \nu_- m \) are the molalities of the cation and anion respectively. The total chemical potential can then be expressed as the sum of the individual ions:

\[
\mu_{MX} = \nu^+ \mu_{M^+} + \nu^- \mu_{X^-}
\]  

(2.42)

\[
\mu_{MX} = \nu^+ \mu^\dagger_{M^+} + \nu^- \mu^\dagger_{X^-} + RT \ln m_+ m_- + RT \ln \gamma_+ \gamma_-
\]  

(2.43)

From here, we can add the standard state chemical potentials of each ion:

\[
\mu^\dagger_{MX} = \nu^+ \mu^\dagger_{M^+} + \nu^- \mu^\dagger_{X^-}
\]  

(2.44)

as well as the individual ideal gas chemical potentials:

\[
\mu^0_{MX} = \nu^+ \mu^0_{M^+} + \nu^- \mu^0_{X^-}
\]  

(2.45)

and the intrinsic hydration free energies of each ion:

\[
\mu_{\text{res};\infty}^{\dagger}_{MX} = \nu^+ \mu_{\text{res};\infty}^{\dagger}_{M^+} + \nu^- \mu_{\text{res};\infty}^{\dagger}_{X^-}
\]  

(2.46)

We define the mean ionic molality for the electrolyte to be:

\[
m^\nu_\pm = (\nu^+ \nu^-) m^\nu,
\]  

(2.47)

where the total number of moles of ions from one mole of electrolyte is given by \( \nu = \nu_+ + \nu_- \). Finally, by defining the mean ionic activity coefficient of a 1-1 electrolyte solution as:

\[
\gamma^\nu_\pm = \gamma^\nu_+ \gamma^\nu_-,
\]  

(2.48)

we can express the the chemical potential of an electrolyte solution can be expressed as [38]:

\[
\mu_{MX} = \mu^\dagger_{MX} + \nu RT \ln m_\pm + \nu RT \ln \gamma_\pm.
\]  

(2.49)
For the simplest case, we can consider a 1-1 electrolyte such as aqueous NaCl, where the valency of each ion is \(|z^+| = |z^-| = 1|\), so \(\nu_+ = \nu_- = 1|\). We write the reference state chemical potential \(\mu^\dagger_{\text{NaCl}}\) of NaCl as:

\[
\mu^\dagger_{\text{NaCl}}(T, P; m^0) = \mu^0_{\text{NaCl}} + 2RT \ln \left( \frac{RTM_{\text{solv}}\rho^*_\text{solv}}{1000P^0} \right) + \mu^{\text{res};\infty}_{\text{NaCl}}(T, \rho^*_\text{solv}[T, P]),
\]

where:

\[
\begin{align*}
\mu^0_{\text{NaCl}}(T, P^0) &= \mu^0_{\text{Na}^+}(T, P^0) + \mu^0_{\text{Cl}^-}(T, P^0) \\
\mu^{\text{res};\infty}_{\text{NaCl}}(T, \rho^*_\text{solv}[T, P]) &= \mu^{\text{res};\infty}_{\text{Na}^+}(T, \rho^*_\text{solv}[T, P]) + \mu^{\text{res};\infty}_{\text{Cl}^-}(T, \rho^*_\text{solv}[T, P])
\end{align*}
\]

Therefore, the total chemical potential of aqueous NaCl is:

\[
\mu_{\text{NaCl}}(T, P; m) = \mu^\dagger_{\text{NaCl}}(T, P) + 2RT \ln m_{\text{NaCl}} + 2RT \ln \gamma_{\text{NaCl}}(T, P; m)
\]

### 2.4.3 Activity Coefficients of Electrolyte Solutions

From here, it is also necessary to implement functional models for the activity coefficients of the solution if we want to compare simulation quantities with experiment. For a 1-1 electrolyte solution, such as aqueous sodium chloride, the most common models are based on the Debye-Hückel model [39]:

\[
\ln \gamma_{\text{NaCl}}(T, P; m) = \frac{A_\gamma \sqrt{m}}{1 + B\sqrt{m}}.
\]

In Debye-Hückel models, \(B\) is an empirical parameter related to the ion size, and \(A_\gamma\) is a parameter accounting for the dielectric constant of the solvent, given by:

\[
A_\gamma = \left( \frac{2\pi N_A \rho^*_\text{solv}}{1000} \right)^{1/2} \frac{e^3}{(4\pi \epsilon_0)^{3/2} k^{3/2}} \left( \frac{1}{T^{3/2} \epsilon^{3/2}} \right)
\]

where \(N_A\) is Avagadro’s constant, \(e\) is the charge of an electron, and \(\epsilon\) and \(\epsilon_0\) are the permittivity of the solvent and free space, respectively. We note that in some literature, the density term has been neglected from the calculation of \(A_\gamma\) [7, 8]. At 298.15K this is fairly inconsequential, as the density of water is close to unity. However, at 473.15K, the value of
\( A, \) differs by approximately 10%.

Closely related are the Davies equations, given by:

\[
\ln \gamma_{\text{NaCl}}(T, P; m) = A_{\gamma} \left( \frac{\sqrt{m}}{1 + \sqrt{m}} - cm \right),
\]

(2.56)

where \( c \) is a constant, generally taken to have a value of 0.2 or 0.3. However, there is controversy in literature over which value is correct [40, 7].

Another option, which is suitable for larger concentration ranges, is the activity coefficient model given by Pitzer [4]:

\[
\ln \gamma_{\text{NaCl}}(T, P; m) = -|z_{M} + z_{X^{-}}| A_{\phi} \left( \frac{\sqrt{m}}{1 + B\sqrt{m}} + \frac{2}{B} \ln(1 + B\sqrt{m}) \right) \\
+ m \frac{2\nu_{M} + \nu_{X^{-}}}{\nu} \left( 2\beta_{\text{MX}}(0) + \frac{2\beta_{\text{MX}}(1)}{\alpha^2 m} \left[ 1 - \left( 1 + \alpha \sqrt{m} - \frac{\alpha^2 m}{2} \right) e^{-\alpha \sqrt{m}} \right] \right) \\
+ \frac{3m^2}{2} \left[ \frac{2(\nu_{M} + \nu_{X^{-}})^{3/2}}{\nu} C_{\text{MX}}^{\phi} \right] 
\]

(2.57)

The Debye-Hückel osmotic coefficient parameter \( A_{\phi} \) is used in this model. It differs from the \( A \) parameter used by Hamer and Wu [41] by a factor of 3. This is due to the use of the Debye-Hückel slope for activity coefficients, where:

\[
A_{\gamma} = 3A_{\phi}. 
\]

(2.58)

This model, while more complex than the Debye-Hückel and Davies equations, was obtained by fitting to experimental data.

**Experimental Parameters for Aqueous Sodium Chloride**

For future use, we give experimental values for \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) and \( \mu_{\text{NaCl}}^0(T, P^0) \) at three temperatures and pressures commonly found in literature, which can be seen in Table 2.1. The process for calculating \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) at different temperatures can be seen in Appendix A[42, 43]:
Table 2.1: Values of $\mu_{NaCl}^\dagger(T, P; m^0)$ and $\mu_{NaCl}^0(T, P^0)$ at different temperatures and pressures. $\mu_{NaCl}^0(T, P^0)$ values were obtained from JANAF. $\mu_{NaCl}^\dagger(T, P; m^0)$ were calculated using the procedures described in Appendix A using the Wagman [3] values as reference and Pitzer [4] thermodynamic data.

Parameters used in the model to calculate activity coefficients of aqueous sodium chloride from experimental results at the same three temperatures and pressures can be found in Table 2.2:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>298.15K, 1bar</th>
<th>373.15K, 1bar</th>
<th>473.15K, 15.5bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_\phi$</td>
<td>0.3915</td>
<td>0.4606</td>
<td>0.6228</td>
</tr>
<tr>
<td>$A_\gamma$</td>
<td>1.1745</td>
<td>1.3818</td>
<td>1.8684</td>
</tr>
<tr>
<td>$\epsilon_r$</td>
<td>78.38</td>
<td>55.51</td>
<td></td>
</tr>
<tr>
<td>$B$</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>$\nu_{M^+}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\nu_{X^-}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\nu$</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>$\beta_{MX}^0$</td>
<td>0.0754</td>
<td>0.1002</td>
<td>0.0938</td>
</tr>
<tr>
<td>$\beta_{MX}^1$</td>
<td>0.2770</td>
<td>0.3326</td>
<td>0.4175</td>
</tr>
<tr>
<td>$C_{MX}^{\phi}$</td>
<td>1.40E-3</td>
<td>-3.30E-3</td>
<td>-4.84E-3</td>
</tr>
</tbody>
</table>

Table 2.2: Parameters from Rogers and Pitzer [5] used to calculate experimental activity coefficient values for low-concentration regions of aqueous sodium chloride.
Chapter 3

Molecular Simulation

Molecular simulations are an extremely powerful tool in the physical sciences which allow calculation of properties from many-particle systems. Due to this inherent nature, the fundamentals of molecular simulation lies in statistical mechanics. Using statistical mechanics, it is possible to calculate a large range of properties of the system by computing averages over system states and particles.

In this chapter, we present the fundamentals behind molecular simulation. We start by explaining the importance of choosing appropriate simulation methodology to calculate properties of interest. Next we describe the fundamental statistical mechanics of molecular simulations. From here, we describe the main types of simulation approaches, Monte Carlo (MC) simulations and molecular dynamics (MD) simulations. Finally, we explore the theory behind free energy calculations in molecular simulations, and how we can use such calculations to calculate other thermodynamic properties of interest.

3.1 Statistical Mechanics of Molecular Simulation

Statistical mechanics is a branch of the physical sciences dealing with properties of bulk matter. As opposed to other fields which typically deal with with systems consisting of only a few particles, statistical mechanics generally deals with systems containing around $10^{23}$ particles. For these large systems, knowledge of each individual particle’s position and momentum is impossible, and essentially useless. Instead, what is important is averages of these properties over all particles for some time evolution of the system.

The fundamental assumption of statistical mechanics states that over an infinite time scale, all microstates of a physical system will appear with equal probability [44]. To con-
ceptualize this idea, we can create a thought experiment of any physical system with given properties. From this original system, we can create copies with the same properties, but differing microstates. Each of these copies of the system corresponds to a macrostate for the system. As macroscopic thermodynamic properties are related to averages of particle positions and velocities within the system, the precise microstate of the system is irrelevant. This is the concept of ensembles in statistical mechanics [45].

There are a variety of different ensembles important to statistical mechanics and subsequently molecular simulation. In molecular simulation, the choice of ensemble is dependent largely on the property of interest and simulation methodology of choice. Ensembles are described by the thermodynamic variables which are shared among all microstates in the system.

In the microcanonical (NVE) ensemble, the number of particles $N$, system volume $V$, and system energy $E$ is held constant. The NVE is considered the fundamental ensemble of statistical mechanics [46]. This ensemble corresponds to an equilibrated isolated system, where each microstate $q$ with particle positions $(r_1, r_2, ..., r_N)$ and energy $E_q$ equal to the macrostate energy $E$, occurs with equal probability, and other microstates occur with zero probability. Thus, we can write the probability for each microstate as:

$$P_q = \frac{\delta_{E_q,E}}{\Omega(N,V,E)},$$

(3.1)

where $\delta_{E_q,E}$ is the Kronecker delta:

$$\delta_{E_q,E} = \begin{cases} 
1; & E_q = E \\
0; & E_q \neq E
\end{cases},$$

(3.2)

and $\Omega(N,V,E)$ is a normalization quantity known as the partition function of the system, defined as:

$$\Omega(N,V,E) = \sum_q \delta_{E_q,E}.$$  

(3.3)

In the canonical (NVT) ensemble, the number of particles, system volume, and system temperature $T$ are fixed constants. We can write the probability for visiting any microstate $q$ in the NVT ensemble as:

$$P_q = \frac{\exp(-\beta H(q))}{Q(N,V,T)},$$

(3.4)
where $\beta = (k_B T)^{-1}$, $\mathcal{H}(q)$ is the Hamiltonian of microstate $q$, and $\mathcal{Q}(N, V, T)$ is the partition function of the system, defined as:

$$\mathcal{Q}(N, V, T) = \sum_q \exp(-\beta \mathcal{H}(q)).$$  \hspace{1cm} (3.5)

In Equation 3.4, the numerator on the right hand side is known as the Boltzmann factor.

Using these probability functions as a starting point, it becomes possible to compute an ensemble average of an observable quantity $\mathcal{O}$ by accounting for the distribution of microstates, allowing the translation of microstate properties to macroscopic thermodynamics. The ensemble average of the observable can be expressed as:

$$\langle \mathcal{O} \rangle = \sum_q \mathcal{O}_q \mathcal{P}_q$$  \hspace{1cm} (3.6)

Despite both being partition functions, $\Omega(N, V, E)$ and $\mathcal{Q}(N, V, T)$ are given different variables as the partition function for each ensemble can be used to calculate different thermodynamic potentials. If we compute the ensemble average energy in the NVT ensemble:

$$\langle E \rangle = \sum_q E_q \mathcal{P}_q$$

$$= \frac{\sum_q E_q \exp(-\beta E_q)}{\mathcal{Q}(N, V, T)}$$  \hspace{1cm} (3.7)

From this equation, it is clear that the numerator can be expressed as the derivative $-\partial \mathcal{Q}/\partial \beta$. Thus, it follows:

$$\langle E \rangle = \frac{1}{\mathcal{Q}} \left(-\frac{\partial \mathcal{Q}}{\partial \beta}\right)$$

$$= -\frac{\partial \ln \mathcal{Q}}{\partial \beta}$$  \hspace{1cm} (3.8)

Substituting in the definition $\beta = (k_B T)^{-1}$, we can write:

$$\partial \beta = -\frac{\partial T}{k_B T^2}$$  \hspace{1cm} (3.9)

and Equation 3.8 becomes:

$$\frac{\langle E \rangle}{T^2} = \frac{\partial(k_B \ln \mathcal{Q})}{\partial T}$$  \hspace{1cm} (3.10)
Comparing Equation 3.10 to the classical macroscopic thermodynamic relation

\[ \frac{E}{T^2} = -\frac{\partial(A/T)}{\partial T}, \] (3.11)

and assuming the macroscopic ensemble average energy \( \langle E \rangle \) is identical to the classical macroscopic energy value for a thermally equilibrated system, we finally get [46]:

\[ A(N, V, T) = -\beta^{-1} \ln Q(N, V, T), \] (3.12)

where \( A \) is the Helmholtz free energy of the system.

From Equation 3.12, it becomes desirable to compute the canonical partition function in order to calculate the Helmholtz free energy of the system. To do this, we factor the total partition function into its internal, translational, and configurational parts:

\[ Q_{\text{tot}} = Q_{\text{int}} Q_{\text{trans}} Q_{\text{config}}. \] (3.13)

The internal partition function \( Q_{\text{int}} \) is calculated from quantum mechanical simulations, and in the simplest case of a spherical particle, the translational partition function \( Q_{\text{trans}} \) is given by [47]:

\[ Q_{\text{trans}} = \frac{1}{N!} V^N \left[ \frac{2\pi mk_B T}{\hbar^2} \right]^{3N/2}. \] (3.14)

Together, these compose the ideal gas chemical potential, given by:

\[ A_{IG}(N, V, T) = -k_B T \ln Q_{\text{int}} Q_{\text{trans}}. \] (3.15)

For more complicated particles (e.g. molecules), it is also important to consider the contributions of internal vibrations, molecular rotations, electrons, and the nucleus to the internal partition function.

As the ideal gas components of the partition function arise from quantum mechanics, the important term to consider for molecular simulation is the configurational partition function, which arises from interactions between particles in the system. For a classical non-ideal gas, intermolecular interactions originate from van der Waals, Coulombic, and intermolecular interactions. The configurational Hamiltonian for the OPLS (Optimized Potentials for Liquid
Simulations) force field model commonly found in literature is given by:

\[ \mathcal{H}_{\text{config}} = \sum_{ij} [U_{\text{vdW}} + U_{\text{Coul}}] + \sum_{\text{bonds}} U_{\text{bonds}} + \sum_{\text{angles}} U_{\text{angles}} + \sum_{\text{dihedrals}} U_{\text{dihedrals}} \]  

(3.16)

where \( i \) and \( j \) are two distinct particles, \( U_{\text{vdW}} \) is the potential energy term given by van der Waals interactions between particles, \( U_{\text{Coul}} \) is the potential energy term arising from interactions between charged particles, \( U_{\text{bonds}} \) is the potential energy arising from bond stretching, \( U_{\text{angles}} \) is the potential energy arising from bond bending, and \( U_{\text{dihedrals}} \) is the potential energy arising from torsional twisting. Implementation of such interactions in molecular simulations requires the use of a functional form for the potential. Such functional forms are known as force field models.

The Coulombic potential term is an interaction term which accounts for electrostatic interactions between charged particles. It is given by the form in Equation 3.17, where \( \epsilon_0 \) is the vacuum permittivity, \( q_i \) is the charge of particle \( i \) (in C), and \( r_{ij} = |r_i - r_j| \) is the distance between particles \( i \) and \( j \). The charge of each particle \( q \) may take into account polarization effects and in the case the atom is part of a molecule, partial charges.

\[ U_{\text{Coul}} = \frac{1}{4\pi\epsilon_0} \sum_{i \neq j} \frac{q_i q_j}{r_{ij}} \]  

(3.17)

Van der Waals interactions, on the other hand, affect both charged and neutral particles. The most common form of the Van der Waals potential is the Lennard-Jones 6-12 interaction, as seen in Equation 3.18:

\[ U_{\text{LJ}} = \sum_{i=1}^{N} \sum_{i \neq j} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right], \]  

(3.18)

where \( \epsilon_{ij} \) is the depth of the potential well between particles \( i \) and \( j \), \( r_{ij} \) is the inter-particle distance, and \( \sigma_{ij} \) is the equilibrium distance of the interaction between particles (i.e. the distance at which the interaction between the two is zero). The \( r^{-6} \) term is predominantly responsible for attractive forces, while the \( r^{-12} \) contributes mainly to attraction.

Force field parameters \( \sigma_{ij} \) and \( \epsilon_{ij} \) are known as cross-term parameters, and are generally calculated from mixing rules of homogeneous parameters \( \sigma_{ii} \) and \( \epsilon_{ij} \). Care must be taken in choosing proper values for these cross-terms. A common choice for \( \epsilon_{ij} \) is the geometric
combination rule, where:

$$\epsilon_{ij} = \sqrt{\epsilon_{ii} \epsilon_{jj}}. \quad (3.19)$$

In some cases, $\sigma_{ij}$ can also use geometric combination rules. However, Lorentz-Berthelot combination rules are sometimes used. As seen in 3.20, these mixing rules are an arithmetic average of $\sigma$ for each species.

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad (3.20)$$

Homogeneous interaction parameters for each are determined by fitting simulation data to spectroscopic results and/or quantum mechanical calculations.

Bonded interactions account for energy terms arising from short-range interactions between bonded atoms within the same molecule. Due to the short-range nature of their interaction, bonded interaction terms also account for three-body and four-body interactions arising from angle-bending and torsional rotation in addition to two-body bond stretching. The short length scale of these interactions requires careful treatment of the quantum mechanics of the system to accurately calculate these force field parameters.

The simplest case of a bonded interaction is bond stretching. While bonds between atoms tend to be thought of as static, the presence of nearby particles can cause some distortion of the bond, squishing the bonded atoms together or stretching them apart. For small deviations from the bond equilibrium length $r_0$, we can use a harmonic approximation to model the bond potential:

$$U = \frac{k_{\text{bond}}}{2}(r - r_0)^2, \quad (3.21)$$

where $k_{\text{bond}}$ is the force constant for the potential and $r$ is the bond length.

The next bonded potential term to be considered is the bending of bond angles. Any three atoms $A, B$ and $C$ with fixed bond lengths $r_{AB}$ and $r_{BC}$ will form an angle. As this bond angle will similarly assume an equilibrium value $\theta_0$, deviations $\theta$ from this equilibrium value will require energy. Thus, it is also reasonable to approximate the angular bending term of a force field using a harmonic potential seen in Equation 3.22

$$U = \frac{k_{\text{angle}}}{2}(\theta - \theta_0)^2. \quad (3.22)$$

In comparison to the bond stretching potential, changing the energy of a bond requires about a tenth of the energy.

The final bonded potential term arises from the twisting of a dihedral between four atoms $A - B - C - D$. The dihedral angle for this system of atoms is the angle between the plane in
which atoms $A - B - C$ lie and the plane in which atoms $B - C - D$ lie. The functional form for a torsional potential is generally approximated using a cosine series, as seen in Equation 3.23

$$U = \sum_{i=1}^{4} \frac{V_i}{2} (1 + (-1)^i \cos(n\phi)) \quad (3.23)$$

Due to the distance between atoms $A$ and $D$, the torsional potential is not the only contribution to their interaction — there is also a contribution from the van der Waals and Coulombic potentials. Thus, it is common to scale the torsional potential by half.

Combining this and converting the sum to an integral, we can express the residual chemical potential arising from the configurational partition functional as:

$$A^{res}(N, V, T) = -k_B T \ln Q_{conf}(N, V, T) \quad (3.24)$$

$$= k_B T \ln \left( \int \cdots \int \exp(-\beta H_{\text{config}}) d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N \right) \quad (3.25)$$

$$= k_B T \ln \left( \int \cdots \int \exp \left[ -\beta \sum_{ij} [U_{\text{vdW}} + U_{\text{Coul}}] + \sum_{\text{bonds}} U_{\text{bonds}} + \sum_{\text{angles}} U_{\text{angles}} + \sum_{\text{dihedrals}} U_{\text{dihedrals}} \right] d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_N \right) \quad (3.26)$$

which finally gives an expression for the calculation of the total Helmholtz of the system:

$$A(N, V, T) = A^{IG}(N, V, T) + A^{res}(N, V, T) \quad (3.27)$$

$$= -k_B T \ln Q_{\text{int}} Q_{\text{trans}} - k_B T \ln Q_{\text{conf}} \quad (3.28)$$

### 3.2 Molecular Dynamics Simulations

A useful tool in the study of thermodynamic properties of aqueous solutions is the implementation of molecular dynamics (MD) simulations. Brought to the forefront of computational sciences by advances in computing power and simulation methodology, these simulations study the time evolution of $N$-body systems to calculate thermodynamic behaviour of molecular systems.

Molecular dynamics simulations describe inter- and intramolecular interactions using force fields, which can be used to calculate forces between particles. The simulation is then advanced by a small time step, and thermodynamic quantities of interest are calculated for
the updated particle positions. This is repeated for a specified number of steps. At the end of the simulation, a thermodynamic average of simple calculated thermodynamic quantities, such as density in an NPT simulation or pressure in an NVT simulation, can be taken to calculate the desired property for the system. A basic flowchart of a molecular dynamics simulation can be seen in Figure 3.1.

![Flowchart of molecular dynamics simulation](image)

Figure 3.1: A simplified schematic of a molecular dynamics algorithm.

The simulation methodology arising naturally from statistical mechanics is Monte Carlo simulations. However, molecular dynamics simulations make use of the ergodic hypothesis: if the Hamiltonian of the system is ergodic and a sufficient time scale is used, the time average (i.e. the average value of a property over time) of a property and the ensemble average (i.e. the average value of a property over phase-space) are assumed to be equal [48].

While these simulations are able to model macroscopic system properties using only a few thousands of particles, these systems are highly complex. As a result, extreme care must be taken in choosing appropriate input parameters so that simulation calculations accurately translate to macroscopic measurements. The following sections outline the theory and methodology behind proper implementation of molecular dynamics simulations to study the thermodynamics of solutions.
3.2.1 Boundary Conditions

Figure 3.2: A 3D simulation box of sodium (blue) and chloride (teal) ions dissolved in water (hydrogen - white; oxygen - red). Such simulation boxes are finite in size and do not accurately represent macroscopic systems.

A major pitfall in molecular dynamics simulations used to simulate bulk matter is the simulation box size. Even boxes containing $10^6$ molecules are considered small, and will be affected by the surrounding vacuum. In comparison, bulk matter often has system sizes on the order of $10^{23}$ particles. To accurately model these systems using only $10^6$ or fewer particles, we can impose periodic boundary conditions. This method surrounds the simulation box with spatially translated copies of itself.

The most common simulation box periodicity is $xyz$ periodic boundary conditions, where
each side of the box is surrounded by exact copies of the box, and each of these boxes is further surrounded by more copies ad infinitum, thus approximating an infinite system. When computing the interactions of a particle \( n_j \) on a particle \( n_i \), it is common to use the minimum-image convention, in which distances \( r_{ij} \) are computed between \( n_i \) and all images of \( n_j \) in the simulation box and neighbouring image boxes. The interaction term \( U_{ij} \) is then computed only for \( n_i \) and the nearest image of \( n_j \).

However, periodic boundary conditions are not perfect. As liquids and gases are not periodic in nature, introducing periodic boundary conditions is not an ideal approximation and introduces error. However, in comparison to the error from surrounding the simulation box with a vacuum, this error is small.

### 3.2.2 Long-Range Interactions

When considering van der Waals and electrostatic interactions, it is evident that their \( 1/r^n \) scaling would result in very small interaction potentials for highly separated particles. To speed up simulations, we can use cut-off algorithms to reduce the number of calculations required, and apply correction terms to account for these cut-offs.

The first and simplest case to consider is the long-range van der Waals interactions. For these cases, it is convenient to define a cut-off radius past which the potential is zero. In addition, some algorithms apply a twin-range cut-off scheme, which specifies two radii. Within the first radius, van der Waals interactions are computed explicitly. In between the first and the second radii, the particles are assumed to not move much, so their interactions are computed every specified number of \( M \) steps. Outside of the second radius, the contributions are zero. In a real system, the particle would still experience small potential contributions from particles past these radius. For large systems, the sum of these interactions can have a sizeable magnitude, so tail corrections should also be applied to account for this [32].

Next, we consider electrostatic interactions. Similar to the van der Waals interactions, long-range electrostatic interactions also have a defined cut-off radius. Within the radius, all electrostatic interaction terms are computed explicitly. However, the electrostatic interactions tend to be much greater in magnitude at longer distances due to their \( r^{-2} \) dependence, so more complicated methodology is desired to approximate these interactions.

One of the most commonly-used is the particle mesh Ewald (PME) summation. This method applied a spline interpolation to discretize the charges outside of the cut-off radius into a grid which can be transformed to \( k \)-space using a 3D-FFT algorithm, and the inverse
interaction term computed by a single sum over the entire grid. The potential can then be computed by the inverse FFT [49].

### 3.2.3 Initialization and Energy Minimization

A convenient and common method of initializing simulation boxes is by random insertions. Particles are initially assigned a position with constraints that it must be within the simulation box and that no particles can overlap within their atomic radii. Velocities are also assigned randomly and are generally chosen from the Maxwell-Boltzmann distribution at the system’s desired temperature.

This process can result in atoms being extremely close, leading to enormous forces between particles. If a standard production run is simulated immediately following particle initialization, the commonly used timesteps (e.g. 1 or 2 femtoseconds) for simulations on the nanosecond scale would result in the system blowing up. To avoid this, we can implement energy minimization runs.

A commonly used minimization algorithm is the iterative method of steepest descent:

\[
\mathbf{r}_{n+1} = \mathbf{r}_n + \frac{\mathbf{F}_n}{\max(|\mathbf{F}_n|)} h_n, \tag{3.29}
\]

where \( \mathbf{r}_n \) is the coordinate vector for particle \( n \), \( \mathbf{F}_n \) is the force on particle \( n \), with \( \max(\mathbf{F}_n) \), and \( h_n \) is the maximum displacement a particle can move in any given step. After each step in energy minimization, forces and energies are recalculated for the new particle positions. Step sizes are also adjusted to ensure minimization. This is repeated for a set number of iterations or until the force on any particle is less than a user-defined maximum force. Steepest descent guarantees convergence to a local minima, but convergence can be slow near the minima or in narrow valleys.

As energy minimization can be slow to converge near minima, the system may still not be near the absolute lowest point in a local minima after energy minimization runs. For this reason, it is also important to include equilibration runs in the simulation. Generally, equilibration runs are performed with identical or very similar parameters as production runs, but only run for a fraction of the time, usually around 10% of the total number of steps. In the end during analysis, no position or energy data from the equilibration runs are included.
3.2.4 Equations of Motion

Molecular dynamics simulations move particles in the system by numerically integrating Newton’s equations of motion for each particle over small time steps, and the system states at each time step allow ensemble averages of system properties to be calculated. For a system containing \( N \) particles, these equations are as follows:

\[
F_i = m_i \ddot{r}_i \quad ; \quad i = 1, 2, ..., N
\]

\[
m_i \ddot{r} = -\frac{\partial}{\partial r} \sum_{j=1}^{N} U_i(r_j) \quad ; \quad i = 1, 2, ..., N
\]

In these equations, \( F_i \) is the force exerted on molecule \( i \), \( m_i \) is the particle mass, \( \dot{r} \) and \( \ddot{r} \) are position and acceleration of the \( i^{th} \) particle respectively, and \( U_i \) the total potential energy of the system. This potential energy term sums inter- and intramolecular interactions between a particle \( i \) and every other particle \( j \) in the system. These interaction terms are mathematical functions known as force fields, and take into account atomic bonds, bond angles, dihedral torsion angles, and attractive/repulsive interactions.

For an \( N \)-body problem with \( N > 2 \), analytical solutions to this system of ODEs become impossible to obtain. For this reason, numerical quadrature is used to solve the motion over discretized time steps.

3.2.5 Temperature and Pressure Coupling

The natural ensemble arising from molecular dynamics simulations is one at constant number of particles, volume, and energy (NVE). However, as seen from the statistical mechanics from which molecular simulation is derived, many thermodynamic properties of interest arise from the NVT or NPT ensembles. Thus, as we require a constant pressure for NPT simulations, or constant temperature for NPT and NVT simulations, additional control loops must be implemented in the form of thermostats and barostats to maintain the desired system properties.

Thermostats are a set of molecular simulation algorithms to control system temperature. One of the most common thermostats, the Nosé-Hoover algorithm, controls temperature by introducing a thermal reservoir and friction term into the equations of motion for the system. This replicates the system being attached to a heat bath at temperature \( T \), and ensures
the system temperature oscillates near the temperature of the heat bath. The modified Hamiltonian for a system using a Nosé-Hoover thermostat is:

\[ \mathcal{H}(P, R, p, s) = \sum_i \frac{P_i^2}{2ms^2} + \frac{1}{2} \sum_{i,j,i\neq j} U(r_i - r_j) + \frac{p_s^2}{2Q} + gkT \ln(s) \] (3.32)

where the first two terms are the original Hamiltonian of the system, \( s \) is the extra degree of freedom accounting for the heat bath, \( g \) is the number of momentum degrees of freedom, and \( Q \) is an imaginary mass determining the degree to which the system is coupled to the heat bath.

Barostats, similar to how thermostats control system temperature, control the system pressure in an NPT ensemble. Again, this is achieved by coupling the box to a “pressure bath”. A common algorithm for a barostat in molecular simulation is the Parrinello-Rahman algorithm, in which the box vectors represented by matrix \( b \) are given by:

\[ \frac{d^2 b}{dt^2} = VW^{-1}b^{-1}(P - P_{\text{ref}}) \] (3.33)

Like the Nosé-Hoover algorithm, this barostat also modifies the equations of motion for the system.

### 3.3 Algorithms for Calculation of Thermodynamic Properties from Simulation

Even in the case of accurate simulation parameters for describing interactions in a system, molecular simulations are not just a black boxes that can spit out any thermodynamic properties of interest. Instead, proper care must be taken to ensure proper methodology is used to obtain certain properties.

Sometimes calculating solution properties has been impeded by a lack of a good methodological approach to the simulation. For example, osmotic pressure was once difficult to calculate accurately and efficiently. Luo and Roux developed a novel method, changing this. The model uses an enclosed volume of dissolved ions in solution sandwiched between two layers of pure solvent, as seen in Figure 3.3. Each layer is separated by a semipermeable membrane, which only allows solvent molecules to pass and impedes the movement of solute molecules. The solvent will flow out of the solution chamber until solvent chemical potential
is equal across the interface, meaning the system is in equilibrium. A force is exerted on the membranes from the excess water in the sections containing the ions, from which osmotic pressure can be extracted [50].

Figure 3.3: Two layers of pure water (blue) and a layer of Na\(^+\) (yellow) and Cl\(^-\) (green) ions in water separated by semipermeable membrane. Water molecules in the simulation are allowed to pass through the membranes, but the solute is confined to the space in between the membranes. This set-up is used to calculate osmotic pressure [12].

Another example, which continues to be problematic to simulate, is vapor pressure. One approach to calculating vapor pressure directly is via Gibbs Ensemble Monte Carlo (GEMC), in which chemical potentials are held constant by adding or removing particles. However, this method results in large statistical uncertainties in vapour pressure at ambient temperatures due to the low vapor pressures of the system [51]. At higher temperatures, this was not as big of an issue [52]. Another method claimed to be approximately the same accuracy as GEMC is the Grand Canonical Ensemble Monte Carlo (GCEMC), in which particle number is held constant in the simulation and chemical potential is fluctuated to calculate solution pressure [53, 54].

Thus, before any simulation type is chosen, it is extremely important to determine the methodology and algorithms which much be applied during simulation to calculate the de-
sired properties. Only then can we choose a simulation type.

3.4 Chemical Potential Calculations Using Molecular Simulations

As seen in Equation 2.7, we can express the chemical potential of a species $i$ in the canonical ensemble as the partial derivative of the Helmholtz free energy with respect to the number of particles in the system. As particle number is a discrete quantity, this derivative simplifies to:

$$
\mu_i = A(T, V, N + 1) - A(T, V, N) .
$$

(3.34)

As we are interested in calculating residual chemical potentials via molecular simulation, we can further extend this to:

$$
\mu_{i; \rho}^{res} (T, V, N) = A_{res; \rho} (T, V, N + 1) - A_{res; \rho} (T, V, N) .
$$

(3.35)

Using these control loops introduces an intrinsic error into the system for each control loop. For this reason, free energy calculations in molecular simulation are best done in an NVT ensemble, as the error originates only from the thermostat, and not the use of a barostat.

3.4.1 Free Energy Calculations

Free energy calculations are an extremely powerful tool in molecular simulation. Knowledge of the free energy change between two different states in a system can be used to calculate a wide variety of system properties. The free energy change between two states $A$ and $B$ for an $N$-particle system can be calculated exactly using the partition function of each state, as seen in Equation 3.36:

$$
\Delta A = A_B - A_A \\
= -k_B T \ln Q_B + k_B T \ln Q_A \\
= -k_B T \ln \left( \frac{\int \int dp^N dr^N \exp [-\beta H_B]}{\int \int dp^N dr^N \exp [-\beta H_A]} \right) .
$$

(3.36)

Multiplying the integrand of the numerator by a factor of $\exp(\beta H_A) \exp(-\beta H_A) = 1$
\[
\Delta A = -k_B T \ln \left( \frac{\int \int dp^N dr^N \exp \left[ -\beta (\mathcal{H}_B - \mathcal{H}_A) \right] \exp \left[ -\beta \mathcal{H}_A \right]}{\int \int dp^N dr^N \exp \left[ -\beta \mathcal{H}_A \right]} \right) \quad (3.37)
\]

\[
\Delta A = \langle \exp(-\beta \Delta \mathcal{H}) \rangle_A \quad (3.38)
\]

It can be seen that Equation 3.37 is simply an ensemble average of the difference in the Hamiltonians between states \( A \) and \( B \) over state \( A \). Evaluating this ensemble average assumes that the phase states of \( A \) and \( B \) overlap in some region. Efficient and accurate free energy calculations require significant overlap between the two states.

In the case of \( |\mathcal{H}_B - \mathcal{H}_A| \gg k_B T \), there is often poor overlap between the two states. This overlap can be improved upon by implementing a set of \( n \) coupling parameters \( \lambda \in (0, 1) \) which represent intermediate states between \( A(\lambda = 0) \) and \( B(\lambda = 1) \). Free energy can then be calculated using Equation 3.39,

\[
\Delta A = -k_B T \ln \left( \frac{Q_B}{Q_{\lambda_n}} \times \frac{Q_{\lambda_n}}{Q_{\lambda_{n-1}}} \times ... \times \frac{Q_{\lambda_1}}{Q_{\lambda_0}} \times \frac{Q_{\lambda_0}}{Q_A} \right)
\]

\[
= (A_B - A_{\lambda_n}) + (A_{\lambda_n} - A_{\lambda_{n-1}}) + ... + (A_{\lambda_1} - A_{\lambda_0}) + (A_{\lambda_0} - A_A)
\]

\[
= A_B - A_A,
\]

where \( Q_i \) is the partition function for state \( i \), and \( G_i \) is the relative free energy of state \( i \).

The coupling parameter \( \lambda \) influences simulations through the implementation of scaling potential functions for Lennard-Jones and Coulomb interactions. For this study, linear Coulomb scaling was chosen, which can be seen as follows:

\[
U_{\text{Coul}} = (1 - \lambda) U_A^{\text{Coul}} + \lambda U_B^{\text{Coul}},
\]

where \( U_A \) and \( U_B \) are the Coulomb potentials in states \( A \) and \( B \), respectively. On the other hand, a soft-core Lennard-Jones scaling was implemented, which can be seen in Equation 3.41:

\[
U_{\text{LJ}}^{SC}(r) = (1 - \lambda) U_A^{\text{LJ}}(r_A) + \lambda U_B^{\text{LJ}}(r_B),
\]

where \( U_A \) and \( U_B \) are again the hard-core potentials in states \( A \) and \( B \), and \( r_A \) and \( r_B \) are scaled interaction distances found using Equations 3.42 and 3.43:

\[
r_A = (\alpha \sigma_A^6 \lambda^6 + r^6)^{1/6}
\]

38
\[ r_B = (\alpha \sigma_B^6 (1 - \lambda)^p + r^6)^{\frac{1}{p}}, \quad (3.43) \]

where \( \alpha \) and \( p \) are user-defined parameters, and \( \sigma \) is the interaction radius.

### 3.4.2 Bennett Acceptance Ratio

To implement free energy calculations in a molecular dynamics simulation, there are a variety of methods available. One of the earlier and most widely used is the Bennett Acceptance Ratio (BAR) method [55]. Unlike exponential averaging and thermodynamic integration which require an ensemble average of a single state, implementing BAR requires configuration information from multiple states to estimate a free energy change.

The principal behind BAR requires that for some configuration \( q \) at two separate states \( i \) and \( j \), there must exist a pathway from \( U_i(q) \) to \( U_j(q) \) with a potential difference \( \Delta U_{ij}(q) \). As both states are of the same configuration \( q \), there exists an exact relationship between the potential energy differences \( \Delta U_{ij}(q) \) sampled from state \( i \), and the reverse free energy differences \( \Delta U_{ji}(q) \) sampled from state \( j \).

To compute the free energy change \( \Delta A \) between states, Bennett showed that for two neighbouring states computed from states \( \lambda_l \) and \( \lambda_{l+1} \) using \( N_l \) and \( N_{l+1} \) coordinate frames, we can calculate \( \Delta A \) from:

\[
\Delta A (\lambda_l \rightarrow \lambda_{l+1}) = k_B T \left( \ln \frac{\langle f(U(\lambda_l) - U(\lambda_{l+1}) + C) \rangle_{\lambda_{l+1}}}{\langle f(U(\lambda_{l+1}) - U(\lambda_l) - C) \rangle_{\lambda_{l}}} \right) + C, \quad (3.44)
\]

where \( f(x) \) is the Fermi function:

\[
f(x) = (1 + \exp x/k_B T)^{-1}, \quad (3.45)
\]

and \( C \) is given by:

\[
C = k_B T \ln \frac{Q_l N_{l+1}}{Q_{l+1} N_l}. \quad (3.46)
\]

To find \( C \), Equation 3.44 is iterated until the numerator and denominator are equal. From here, the free energy change between states \( l \) and \( l + 1 \) can finally be calculated using:

\[
\Delta A (\lambda_l \rightarrow \lambda_{l+1}) = -k_B T \ln \frac{N_{l+1}}{N_l} + C. \quad (3.47)
\]

Finally, to calculate the free energy change between the initial state \( \lambda_i = 0 \) and end state
\( \lambda_j = 1 \) with \( N \) intermediate states, we can sum the free energy changes between intermediate states:

\[
\Delta A(i \rightarrow \lambda_j) = \sum_{t=0}^{N-1} \Delta A(\lambda_t \rightarrow \lambda_{t+1})
\]  

(3.48)

An extension of the BAR method is the Multistate Bennett Acceptance Ratio (MBAR) [56]. It differs from the normal BAR method by reweighing equilibrium states from all states so that any state \( i \) is overlapped with all other states \( i \neq j \), not just neighbouring states. This method is statistically optimal and results in the lowest uncertainty of free energy estimates of other available methods [56].

As seen from the BAR and MBAR methods, free energy calculations require the calculation of free energy changes between intermediate states. To make this possible, each intermediate state is run as an individual simulation using a single pair of \((\lambda_{\text{Coul}}, \lambda_{\text{LJ}})\) values. After the completion of all intermediate state simulations, the BAR and MBAR methods can be implemented using the freely-available Python script `alchemical-analysis.py` [57], available as part of the Pymbar package [56].
Chapter 4

Calculations of $\mu^\dagger$ and $\mu^{\text{res;H}}$ Using Molecular Simulation

Modelling the macroscopic thermodynamics of chemical reaction equilibrium requires intimate knowledge of the Henry-law chemical potential model from Equation 1.7. Two important parameters in this equation are the Henry-law standard state chemical potential $\mu^\dagger(T, P^m)$ and the Henry-law residual chemical potential, also known as the activity coefficient $\ln \gamma(T, P; m)$. As shown in Chapter 2, it is possible to calculate these macroscopic properties using properties obtained from chemical potential calculations in molecular simulation.

Accurate and precise calculations of $\mu^\dagger$ and $\ln \gamma$ using molecular simulation have thus far been a difficult task. Current efforts are impeded by two problems: (1) lack of appropriate force fields for the electrolytes of interest, and (2) improper simulation methodology. The following chapter gives a background of previous attempts to calculate $\mu^\dagger$ and $\ln \gamma$ via molecular simulation before proposing new simulation methodology to accurately and precisely calculate $\mu^\dagger$ and $\ln \gamma$ for any force field. This methodology is initially developed and tested using a relatively simple and frequently-studied system of aqueous Na$^+$ and Cl$^-$, with plans to be extended to species of interest in the RNH$_2$-CO$_2$-H$_2$O mixtures found in the carbon capture project in future work. The Joung-Cheatham force field is used for the ions, along with SPC/E water, due to their frequent use in other literature with comparable results.
4.1 Previous Attempts to Calculate $\mu^\dagger$ and $\ln \gamma$

Previous attempts to calculate $\mu^\dagger$ and $\ln \gamma$ were motivated by the calculation of the solubility of alkali halides, such as NaCl, KCl, and NaI. The solubility of a species $i$ is the solute composition $m$ where the solid and solution chemical potentials of the species are equal:

$$\mu_{\text{solid}}(T, P) = \mu_{\text{solution}}(T, P; m)$$ (4.1)

where $\mu_{\text{solid}}$ and $\mu_{\text{solution}}$ are the crystalline solid and Henry-law solution chemical potentials, respectively. As seen from Figure 4.1, the intersection of the solution-phase chemical potential with the solid chemical potential occur on a relatively flat region of the $\mu_{\text{soln}}$ curve, so solubility calculations are highly sensitive to the value of $\mu$, thus $\mu^\dagger$ and $\ln \gamma$ [58]. Therefore, an effort has been made by several groups to accurately and precisely calculate $\mu^\dagger$ and $\ln \gamma$ in order to better predict the solubility of NaCl.

Moučka et al. [14] used osmotic ensemble Monte Carlo (OEMC) to calculate the solubility of NaCl in water. In an osmotic ensemble, chemical potentials are treated as a fixed...
parameter, so OEMC can be used to calculate the molality corresponding to a specific chemical potential. By setting the simulation chemical potential to $\mu_{\text{solid}}$ for NaCl, the solubility limit can be obtained from OEMC. In addition, they also used OEMC to calculate molalities corresponding to a range of chemical potential values. Chemical potentials were plotted versus molality and fit using a weighted regression of Equation 4.2:

$$
\mu_i(T, P; m) = \mu^\dagger_i(T, P^0) + 2RT \ln m_i + 2RT \left[ \frac{A_{\gamma}(m)}{1 + B \sqrt{m}} + bm + cm^2 + dm^3 \right],
$$

(4.2)

where $A$ and $B$ were fixed to their experimental values, while $\mu^\dagger, b, c,$ and $d$ were used as fitting parameters. This equation is an empirical extension of the Debye-Hückel equation. In this approximation, the first term arises from ion-solvent interactions. As system size effects on $\mu_{i;\text{NVT}}$ are most noticeable for dilute systems, fitting across the entire concentration range reduces this error in $\mu^\dagger_i$ [9].

Mester and Panagiotopoulos [7, 8] calculated total chemical potentials of NaCl, $\mu_{\text{NaCl}}$ as a function of composition using Monte Carlo simulations in an NPT ensemble. The total chemical potentials were calculated at each composition using:

$$
\mu_{\text{NaCl}} = \mu^0_{\text{NaCl}} + 2RT \ln \frac{RN_{\text{NaCl}}}{P_0\langle V \rangle} + \mu_{\text{NaCl}}^{\text{res;NVT}}.
$$

(4.3)

This was motivated by attempts to calculate the solubility of NaCl using Equation 4.1. From here, $\mu^\dagger$ was calculated in two different ways using Equation 4.4 at the lowest simulated molality:

$$
\mu_{\text{NaCl}}(T, P; m) = \mu^\dagger_{\text{NaCl}}(T, P) + 2RT \ln(m_{\text{NaCl}}) + 2RT \ln \gamma_{\text{NaCl}}(T, P; m).
$$

(4.4)

In the first study, they employed a Debye-Hückel limiting law for $\ln \gamma_{\text{NaCl}}$ at $m = 0.038$ molal, where:

$$
\ln \gamma_{\text{NaCl}}(T, P; m) = -A_{\gamma} \sqrt{m}.
$$

(4.5)

In the second approach, they used the Davies equation [59] to calculate $\ln \gamma$ at $m = 0.0111$ molal:

$$
\ln \gamma_i = -A_{\gamma} \left( \frac{\sqrt{m}}{1 + \sqrt{m}} - 0.2m \right).
$$

(4.6)

The Davies equation is often chosen for its lack of fitting parameters and its accuracy in predicting $\ln \gamma$ in concentrations above 0.1 molal using experimental values for $A$ [8].
addition, as a range of concentration-dependent $\mu_{\text{NaCl}}^{\text{res;NPT}}$ values were calculated, they also extended their work to calculate concentration-dependent activity coefficients for the system by fitting total chemical potentials to Equation 4.2 in a similar approach to Moučka.

Young and Panagiotopoulos [9] also followed a similar approach to calculate $\mu^1$ by using it as a fitting parameter in the Davies equation. As noted in Chapter 3, a major caveat in molecular simulation is the small system size in comparison to those studied in macroscopic systems. To account for errors due to small system size, Young and Panagiotopoulos calculated chemical potentials at multiple system sizes per state point by scaling both $N_{\text{NaCl}}$ and $N_{\text{H}_2\text{O}}$ to maintain constant $m$, for $N_{\text{H}_2\text{O}} = 500, 1500, 2500, 5500$. Chemical potentials for each composition were extrapolated to infinite system size to give a chemical potential $\mu_i(\infty)$ at an infinitely scaled system at composition $m$. In the calculation of activity coefficients, they noted that past 0.3 molal, system size effects were negligible, so activity coefficients calculated in Mester and Panagiotopoulos using 500 $\text{H}_2\text{O}$ could be used without the need to extrapolate to infinite system size. Another important difference is the system composition ascribed to each chemical potential calculation.

Another approach by Benavides et al. [60] was also motivated by solubility calculations. In this work, the Gibbs free energy of a solution containing NaCl, $G_{\text{soln}}$ in $\text{H}_2\text{O}$ was calculated at a variety of compositions using Equation 4.7:

$$G_{\text{soln}} = A_{\text{soln}} + pV_{\text{soln}},$$

(4.7)

where $A_{\text{soln}}$ is the Helmholtz free energy of the solution, given by:

$$A_{\text{soln}} = A_{\text{soln}}^{\text{id}} + A_{\text{soln}}^{\text{res}}$$

$$= A_{\text{soln}}^{\text{id}} + A_{\text{soln}}^{\text{res;ref}} + \int_0^1 (U - U_{\text{ref}})_{N,V,T,\lambda} d\lambda$$

(4.8)

where $U_{\text{ref}}$ is the potential energy of an ideal Lennard-Jones fluid, $U$ is the potential energy. For these simulations, a pure Lennard-Jones fluid was used as a reference state, and LJ force field parameters were slowly perturbed to match the desired system composition using thermodynamic integration. From here, chemical potentials were calculated by numerically differentiating:

$$\mu_{\text{NaCl}} = \left( \frac{\partial G_{\text{soln}}}{\partial N_s} \right)_{T,p,N_{\text{H}_2\text{O}}},$$

(4.9)

44
After obtaining a range of composition-dependent $\mu_{\text{NaCl}}$, the values were fitted to Equation 1.7, using an extended Debye-Hückel equation for $\ln \gamma_{\text{NaCl}}$:

$$
\ln \gamma(T, P; m) = \left[ \frac{A \gamma \sqrt{m}}{1 + B \sqrt{m}} + \beta m \right].
$$

(4.10)

Using this method, $\mu^\dagger_{\text{NaCl}}$, $B$, and $\beta$ can be extracted as fitting parameters.

Comparisons between the aforementioned literature results for $\mu^\dagger$ and the corresponding $\mu_{\text{res};\infty}$ values can be seen in Table 4.1. All available literature values, while similar in magnitude, use approximations and fitting to calculate $\mu^\dagger_{\text{NaCl}}$, resulting in relatively large uncertainties. It is also important to note that the uncertainty reported by Mester and Panagiotopoulos [8] is inaccurate. As the work of Young and Panagiotopoulos re-analyzes this data, the true expected uncertainty will likely be similar to that reported by Young.

<table>
<thead>
<tr>
<th>Paper</th>
<th>$\mu^\dagger / \text{kJ mol}^{-1}$</th>
<th>$\mu_{\text{res};\infty}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young and Panagiotopoulos[9]</td>
<td>-390.8(9)</td>
<td>-740.86(90)</td>
</tr>
<tr>
<td>Mester and Panagiotopoulos[7]</td>
<td>-391.6(2)</td>
<td>-741.66(20)</td>
</tr>
<tr>
<td>Mester and Panagiotopoulos[8]</td>
<td>-391.7(2)</td>
<td>-741.76(20)</td>
</tr>
<tr>
<td>Moučka et al. [14]</td>
<td>-391.278</td>
<td>-741.34</td>
</tr>
<tr>
<td>Benavides et al. [60]</td>
<td>-391.70(60)</td>
<td>-741.76(60)</td>
</tr>
<tr>
<td>Experimental</td>
<td>-393.133(80)</td>
<td>-743.92</td>
</tr>
</tbody>
</table>

Table 4.1: Values for $\mu^\dagger_{\text{NaCl}}(T, P; m^0)$ values obtained from previous attempts at calculating the solubility of NaCl in H$_2$O using Joung-Cheatham ion parameters and SPC/E water with molecular simulations. Corresponding $\mu_{\text{NaCl};\infty}(T, P)$ calculated from Equation 2.50 are also shown.

The motivation for this work is to develop methodology to accurately and precisely calculate $\mu^\dagger_{\text{NaCl}}$ using residual chemical potential values obtainable by free energy calculations in molecular simulation for the Joung-Cheatham [10] NaCl force field. It is important to note that the resulting $\mu^\dagger_{\text{NaCl}}$ value corresponds only to the specific ion force field parameters, and will not necessarily match the experimental value cited throughout this work.
4.2 Calculations of $\mu_i^\dagger$

For this project we want to be able to translate from chemical potential calculations obtained from molecular simulation to experimental macroscopic thermodynamics. In this section, we propose a new methodology to precisely calculate $\mu_i^\dagger$ from molecular simulation.

4.2.1 The Extrapolation Procedure for $\mu_{\text{NaCl}}^\dagger$

Equation 2.50 provides an equation to calculate $\mu_i^\dagger$ from $\mu_i^{\text{res};\infty}$, so we need to develop a methodology to calculate $\mu_i^{\text{res};\infty}$ from simulation. This will also allow us to calculate $\ln \gamma_{\text{NaCl}}$ from molecular simulation.

Using the results of Equation 2.34 and considering NaCl is a 1-1 electrolyte, we can express the residual chemical potential of the canonical ensemble as:

$$
\mu_{\text{NaCl}}^{\text{res};\text{NVT}}(T, \rho[T, P]; m) = \mu_{\text{NaCl}}^{\text{res};\infty}(T, P) - 2RT \ln \left( \frac{x_{\text{solv}} \rho_{\text{solv}}(T, P; m)}{\rho_{\text{solv}}^*(T, P)} \right) + \mu_{\text{NaCl}}^{\text{res};\text{H}}(T, \rho[T, P]; m)
$$

In the limit $m \to 0^+$, $\mu_{\text{NaCl}}^{\text{res};\text{H}}(T, \rho[T, P]; m) \to 0^+$, as does the logarithmic density term. Thus, this extrapolation yields $\mu_{\text{NaCl}}^{\text{res};\infty}(T, P)$ as the intercept. As we are motivated to obtain $\mu_{\text{NaCl}}^{\text{res};\infty}(T, P)$ by extrapolating Equation 4.11 to the $m \to 0^+$ limit and therefore calculate $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$, we must look at the behaviour of the logarithmic density term and the excess Henry chemical potential in the low concentration region.

4.2.2 Low-Concentration Behaviour of Density

To observe the behaviour of $\mu_i^{\text{res};\infty}(T, P)$ for dilute solutions, we are also required to know the concentration dependence of the logarithmic density term. We can express a relationship between apparent molar volume $V_{\phi,2}$ and solution density as:

$$
V_{\phi,2} = 1000 \left( \frac{\rho_{\text{solv}}^*(T, P) - \rho_{\text{solv}}(T, P; m)}{m \rho_{\text{solv}}(T, P; m) \rho_{\text{solv}}^*(T, P)} \right) + \frac{M_s}{\rho_{\text{solv}}(T, P; m)},
$$

where $M_s$ is the molar mass of the solute. Solving Equation 4.12 for density, we obtain:

$$
\rho_{\text{solv}}(T, P; m) = \rho_{\text{solv}}^*(T, P) \left[ \frac{1000 + m M_s}{V_{\phi,2} m \rho_{\text{solv}}^*(T, P) + 1000} \right]
$$

We can use expressions from Trevani et al. [61] to obtain the concentration dependence
of \( V_{\phi,2} \), which can be expressed as a series expansion from extended Debye-Hückel theory as such:

\[
V_{\phi,2} = V_2^o + A_\nu \sqrt{m} + O(m), \tag{4.14}
\]

where \( V_2^o \) is the standard molar volume and \( A_\nu \) is the Debye-Hückel limiting slope for, where \( A_\nu = \partial A / \partial P \).

Substituting Equation 4.14 into Equation 4.13, we obtain:

\[
\rho_{\text{solv}}(T, P; m) = \rho_{\text{solv}}^*(T, P) \left[ 1 + \frac{mM_s}{1000} - \frac{\rho_{\text{solv}}^* V_2^o m}{1000} + O(m^{3/2}) \right]
\]

(4.15)

The concentration dependence of the solvent mole fraction, \( x_{\text{H}_2\text{O}} \) is determined as follows:

\[
x_{\text{H}_2\text{O}} = \frac{N_{\text{H}_2\text{O}}}{N_{\text{tot}}} = \frac{N_{\text{H}_2\text{O}}}{N_{\text{H}_2\text{O}} + 2N_{\text{ion}} + N_{\text{H}_2\text{O}}} = \frac{2mM_{\text{H}_2\text{O}} + 1000}{1000} = 1 - \frac{2mM_{\text{H}_2\text{O}}}{1000} + O(m^2)
\]

(4.16)

Therefore, we can express the product \( x_{\text{H}_2\text{O}} \rho_{\text{solv}} \) as:

\[
x_{\text{H}_2\text{O}} \rho_{\text{solv}} = \left( 1 - \frac{2mM_{\text{H}_2\text{O}}}{1000} \right) \left( \rho_{\text{solv}}^*(T, P) \left[ 1 + \frac{mM_s}{1000} - \frac{\rho_{\text{solv}}^* V_2^o m}{1000} \right] \right) + O(m^2)
\]

(4.17)

Using this result, we can finally express the logarithmic density term in Equation 4.11 as:

\[
\ln \left( \frac{x_{\text{soln}} \rho_{\text{solv}}(T, P; m)}{\rho_{\text{solv}}^*(T, P)} \right) = \ln \left( 1 - \frac{2mM_{\text{H}_2\text{O}}}{1000} + \frac{mM_s}{1000} - \frac{\rho_{\text{solv}}^*(T, P)V_2^o m}{1000} \right) + O(m^2)
\]

\[
= \left( \frac{2mM_{\text{H}_2\text{O}}}{1000} - \frac{mM_s}{1000} + \frac{\rho_{\text{solv}}^*(T, P)V_2^o m}{1000} \right) + O(m^2)
\]

(4.18)
4.2.3 Low Concentration Behaviour of $\ln \gamma_{\text{NaCl}}$

We also are required to look at the dilute behaviour for a 1-1 electrolyte solution, namely NaCl. For this system, we define $\mu_{\text{i}}^{\text{res;H}}(T, \rho[T, P]; m)$ to be:

$$\mu_{\text{NaCl}}^{\text{res;H}}(T, P; m) \equiv 2RT \ln \gamma_{\text{NaCl}}(T, P; m),$$  \hspace{1cm} (4.19)

where $\gamma_{\text{NaCl}}$ is the mean ionic activity coefficient of NaCl. As we are interested in the infinitely dilute limiting behaviour of $\mu_{\text{i}}^{\text{res;NVT}}(T, \rho[T, P]; m)$, we must choose an activity coefficient model suitable for this concentration region.

One possible choice is to use an extended Debye-Hückel model developed by Pitzer et al. [4], as seen in Equation 2.57, to express a concentration dependence for $\ln \gamma_{\text{NaCl}}(T, P; m)$:

$$\ln \gamma(T, P; m) = -|z_{M^+}z_X| \frac{A_\gamma}{3} \left( \frac{\sqrt{m}}{1 + B\sqrt{m}} + \frac{2}{B} \ln(1 + B\sqrt{m}) \right)$$

$$+ m \frac{2\nu_{M^+}\nu_X}{\nu} \left( 2\beta_{\text{MX}}^{(0)} + \frac{2\beta_{\text{MX}}^{(1)}}{\alpha^2 m} \left[ 1 - \left( 1 + \alpha \sqrt{m} - \frac{\alpha^2 m}{2} \right) e^{-\alpha \sqrt{m}} \right] \right)$$

$$+ \frac{3m^2}{2} \left[ \frac{2(\nu_{M^+}\nu_X^{-})^{3/2}}{\nu} C_{\text{MX}}^{\phi} \right].$$  \hspace{1cm} (4.20)

This model is a convenient choice because data from Pitzer [5, 6, 4] provides a self-consistent set of experimental density and activity coefficient data. Expanding this activity coefficient model, we find a quadratic dependence on $\sqrt{m}$ for $\ln \gamma_{\text{NaCl}}$ at low $m$:

$$\ln \gamma_{\text{NaCl}}(T, P; m) = -A_\gamma \sqrt{m} - m \left( \frac{2}{3} BA_\gamma - 2\beta_{\text{MX}}^{(0)} - 2\beta_{\text{MX}}^{(1)} \right) + \mathcal{O}(m^{3/2}).$$  \hspace{1cm} (4.21)

The linear coefficient of this expansion contains a factor of $3A_\phi$, which is in agreement with the discrepancy between Hamer and Wu’s value of logarithmically-scaled $A = 0.5108$ and the value of $A_\phi = 0.3915$ used by Pitzer et al. when accounting for the change between $\ln \gamma$ and $\log_{10} \gamma$.

Another suitable choice for low-molality modelling of activity coefficients is the Davies equation, an empirical extension of Debye-Hückel theory [59]:

$$\ln \gamma_{\text{NaCl}}(T, P; m) = -A_\gamma \left( \frac{\sqrt{m}}{1 + \sqrt{m}} + cm \right),$$  \hspace{1cm} (4.22)
where $c$ is a constant taken to be either 0.2 or 0.3. The Davies equation has the advantage of being useful for concentrations of up to 0.5 molal, whereas the Debye-Hückel equation is accurate up to 0.1 to 0.3 molal. Expanding the Davies equation, we get:

$$\ln \gamma_{\text{NaCl}}(T, P; m) = -A\gamma(\sqrt{m} - 1.3m) + O(m^{3/2})$$

or

$$\ln \gamma_{\text{NaCl}}(T, P; m) = -A\gamma(\sqrt{m} - 1.2m) + O(m^{3/2}),$$

depending on the value of $c$.

### 4.2.4 Accuracy of the Davies Equations

As an aside, we can also evaluate the linear and quadratic coefficients of Equations 4.26, 4.27, and 4.28 without the need for any data or fitting. This serves as a pre-check to determine which of the Davies equations is best suited for dilute systems. Quadratic coefficients for each $\ln \gamma_{\text{NaCl}}$ model can be seen in Table 4.2. A plot of $\ln \gamma_{\text{NaCl}}$ can be seen in Figure 4.2.

<table>
<thead>
<tr>
<th>Model</th>
<th>$A\gamma$</th>
<th>Quadratic Coefficient</th>
<th>Quadratic Coefficient/$A\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitzer</td>
<td>1.175</td>
<td>1.644</td>
<td>1.400</td>
</tr>
<tr>
<td>Davies-0.3</td>
<td>1.175</td>
<td>1.527</td>
<td>1.300</td>
</tr>
<tr>
<td>Davies-0.2</td>
<td>1.175</td>
<td>1.409</td>
<td>1.200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T = 298.15K &amp; P = 1bar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitzer</td>
<td>1.379</td>
<td>1.787</td>
<td>1.296</td>
</tr>
<tr>
<td>Davies-0.3</td>
<td>1.379</td>
<td>1.793</td>
<td>1.300</td>
</tr>
<tr>
<td>Davies-0.2</td>
<td>1.379</td>
<td>1.655</td>
<td>1.200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T = 373.15K &amp; P = 1bar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitzer</td>
<td>1.854</td>
<td>2.268</td>
<td>1.223</td>
</tr>
<tr>
<td>Davies-0.3</td>
<td>1.854</td>
<td>2.410</td>
<td>1.300</td>
</tr>
<tr>
<td>Davies-0.2</td>
<td>1.854</td>
<td>2.225</td>
<td>1.200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T = 473.15K &amp; P = 15.5bar</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Linear and quadratic coefficients of each expanded $\ln \gamma_{\text{NaCl}}$ (Equations 4.21 and 4.6) model calculated from experimental parameters at $T = 298.15K$, 373.15K, and 473.15K. For each model, the linear coefficient is $A\gamma$.

It can be seen that in the regions below $m = 0.1$ molal for $T = 298.15K$ that the Davies
model with \( c = 0.3 \) is a better fit for the experimental data given by Pitzer. Furthermore, as the quadratic coefficient of the Davies models are given by \( A_\gamma(-1 - c) \), an appropriate choice of \( c \) to fit the Pitzer data more precisely would be \( c = 0.4 \).

For higher temperatures, the Davies models more closely approximate experimental data. At \( T = 373.15 \text{K} \), the quadratic coefficient almost exactly matches that of the Davies-0.3 model. At \( T = 473.15 \text{K}, P = 15.5\text{bar} \), the quadratic coefficient is within 2% of the quadratic coefficient of the Davies equation with \( c = 0.2 \).

![Graph](image)

Figure 4.2: Comparison of \( \ln \gamma_{\text{NaCl}}(298.15 \text{K}, 1 \text{bar}; \mathbf{m}) \) models in dilute solutions. Below 0.1 molal, the Davies equation with \( c = 0.3 \) is closest to the experimental Pitzer model.

### 4.2.5 Final Extrapolation Expression for \( \mu^\dagger_{\text{NaCl}}(T, P; m^0) \)

Our next goal for this project is to calculate \( \mu^\dagger_{\text{NaCl}} \) using molecular simulation. This can be done using Equation 2.50:

\[
\mu^\dagger_{\text{NaCl}}(T, P; m^0) = \mu^0_{\text{NaCl}}(T, P^0) + 2RT \ln \left( \frac{RM_{\text{solv}} m^0 \rho^\ast_{\text{solv}}(T, P)}{1000 P^0} \right) + \mu_{\text{NaCl}}^{\text{res}}(T, P),
\]

(4.25)

as we can obtain \( \mu_{\text{NaCl}}^{\text{res}}(T, P) \) from an extrapolation of Equation 4.11 at low molalities.
To carry out this extrapolation procedure, we want to combine our results from 4.2.2 and 4.2.3 to express the concentration dependence of \( \mu_{\text{NaCl}}^{\text{res:NVT}}(T, \rho[T, P]; m) \) at low concentrations. Substituting Equations 4.18 and 4.21 into Equation 4.11, we get:

\[
\mu_{\text{NaCl}}^{\text{res:NVT}}(T, \rho[T, P]; m) = \mu_{\text{NaCl}}^{\text{res:}\infty}(T, P) + 2RT \left( \frac{2mM_{\text{H}_2\text{O}}}{1000} - \frac{mM_s}{1000} + \frac{\rho_{\text{solv}}^*(T, P)V_2^o m}{1000} \right)
- 2RT[A_\gamma \sqrt{m} - \frac{2}{3}BA_\gamma m + 2m\beta_0^0 + 2m\beta_1^1] + O(m^{3/2})
\]

(4.26)

for the Debye-Hückel expansion, and:

\[
\mu_{i}^{\text{res:NVT}}(T, \rho[T, P]; m) = \mu_{i}^{\text{res:}\infty}(T, P) + 2RT \left( \frac{2mM_{\text{H}_2\text{O}}}{1000} - \frac{mM_s}{1000} + \frac{\rho_{\text{solv}}^*(T, P)V_2^o m}{1000} \right)
- 2RTA_\gamma(\sqrt{m} - 1.2m) + O(m^{3/2})
\]

(4.27)

\[
\mu_{i}^{\text{res:NVT}}(T, \rho[T, P]; m) = \mu_{i}^{\text{res:}\infty}(T, P) + 2RT \left( \frac{2mM_{\text{H}_2\text{O}}}{1000} - \frac{mM_s}{1000} + \frac{\rho_{\text{solv}}^*(T, P)V_2^o m}{1000} \right)
- 2RTA_\gamma(\sqrt{m} - 1.3m) + O(m^{3/2})
\]

(4.28)

for each of the Davies equations.

Each of these equations is a quadratic polynomial in \( \sqrt{m} \). Furthermore, the only linear term is given by \( 2RTA_\gamma \). By fitting residual chemical potentials to a quadratic polynomial and extrapolating to \( \sqrt{m} \to 0^+ \), we can obtain \( \mu_{\text{NaCl}}^{\text{res:}\infty}(T, P) \) directly from the intercept, which
can then be used to calculate $\mu_{\text{NaCl}}^{\dagger}(T, P; m^0)$. $A_\gamma$ can then be calculated by dividing the coefficient of the linear polynomial term by $2RT$.

To calculate the standard state chemical potential $\mu_{\text{NaCl}}^{\dagger}$ computationally, we used molecular dynamics simulations to perform chemical potential calculations of the removal of one $\text{Na}^+$ or $\text{Cl}^-$ from a simulation box otherwise containing only $\text{H}_2\text{O}$. Calculating $\mu_{\text{NaCl}}^{\text{res};\text{NVT}}$ at multiple dilute system compositions allows an extrapolation in $\sqrt{m}$ to $m \to 0^+$. The intercept of this extrapolation corresponds to $\mu_{\text{NaCl}}^{\text{res};\infty}$ which can be used to calculate $\mu_{\text{NaCl}}^{\dagger}$ from 2.50. The linear coefficient of this extrapolation can be used as a consistency check as it yields the Debye-Hückel $A$ parameter, which can subsequently be used to calculate the dielectric constant of the solvent from Equation 2.55.

### 4.3 Activity Coefficient Calculations

Our next goal for this project is to calculate $\ln \gamma_{\text{NaCl}}$ using our results for $\mu_{\text{NaCl}}^{\text{res};\infty}(T, P)$ for a range of finite concentrations. Rearranging Equation 4.11, we obtain an equation to calculate $\ln \gamma_{\text{NaCl}}$ which requires only the calculation of residual chemical potentials and densities for the solution at finite concentrations:

$$2RT \ln \gamma(T, P; m) = \mu_{\text{NaCl}}^{\text{res};\text{NVT}}(T, \rho[T, P]; m) - \mu_{\text{NaCl}}^{\text{res};\infty}(T, P) + 2RT \ln \left( \frac{x_{\text{solv}} \rho_{\text{solv}}(T, P; m)}{\rho_{\text{solv}}^*(T, P)} \right).$$

From Equation 4.29, we now have a way to calculate $\ln \gamma$ entirely from simulation, as $\mu_{\text{NaCl}}^{\text{res};\text{NVT}}(T, \rho[T, P]; m), \rho_{\text{solv}}, x_{\text{solv}},$ and $\mu_{\text{NaCl}}^{\text{res};\infty}(T, P)$ are all variables obtainable directly from simulation or from our extrapolations in the previous section. In addition, this equation also allows for activity coefficient results of Mester [8] and Young [9] to be recalculated using updated values for $\mu_{i}^{\text{res};\infty}(T, P)$.

Unlike with the infinite dilution case, we cannot independently check calculations using experimental results, as no experimental values for $\mu_{\text{NaCl}}^{\text{res};\infty}$ exist in literature, and the ones calculated in the previous section required experimental $\ln \gamma$ values to be used.

Molalities for the chemical potential calculations at finite concentrations were chosen to match those found in Mester and Panagiotopoulos [8]. At low molalities, molecular dynamics simulations with periodic boundary conditions are subject to system size effects. To account for this, we followed the approach proposed by Young and Panagiotopoulos [9] using more dilute systems. In this approach, chemical potentials are calculated at a molality $m$ using multiple scaled boxes containing 5000, 6000, and 10000 $\text{H}_2\text{O}$. As each of these chemical
potentials are at the same molality but with differing box volumes $V$ (and therefore length $L$) due to the increasing number of $N_{\text{NaCl}}$ and $N_{\text{H}_2\text{O}}$, the chemical potentials at each box length $\mu(L)$ could be extrapolated to $1/L \to 0^+$ using a simple linear extrapolation:

$$
\mu_{\text{res; } \infty}(m; L) = \frac{a}{L} + \mu_{\text{res; } \infty}(m; \infty),
$$

(4.30)

where $a$ is the slope of the extrapolation, and $\mu(m; \infty)$ is the chemical potential at molality $m$ at infinite system size. From here, the extrapolated chemical potential at each molality was used to calculate the corresponding activity coefficients. We note that in Young and Panagiotopoulos, the total chemical potential $\mu_{\text{NaCl}}$ was extrapolated. However, this value was obtained by adding constant terms to the residual chemical potential calculated from simulations, which is the value which must be corrected to account for system size effects.
Chapter 5

Simulation Details

5.1 General Simulation Protocols

All molecular dynamics simulations for this work were performed using GROMACS 2016.3. The system temperature was set to 298.15K and controlled using a - thermostat. Periodic xyz boundary conditions were implemented to mirror the simulation box in each direction. Force field parameters used were obtained from Joung and Cheatham [10] for the ions, and SPC/E parameters were used for water [62]. Cross-interactions were determined using Lorentz-Berthelot mixing rules. Force field parameters can be found in Table 5.1:

<table>
<thead>
<tr>
<th>Atom</th>
<th>$\epsilon$ (kJ mol$^{-1}$)</th>
<th>$\sigma$ (nm)</th>
<th>$q$ (e Na$^+$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na$^+$</td>
<td>1.475463</td>
<td>0.21595</td>
<td>+1.0</td>
</tr>
<tr>
<td>Cl$^-$</td>
<td>0.053493</td>
<td>0.4700</td>
<td>-1.0</td>
</tr>
<tr>
<td>O</td>
<td>0.6502</td>
<td>0.3166</td>
<td>-0.8476</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
<td>0</td>
<td>-0.4238</td>
</tr>
</tbody>
</table>

Table 5.1: Force field parameters for the well depth $\epsilon$ and interaction radius $\sigma$ for the JC Na$^+$ and Cl$^-$ ions, as well as SPC/E water.

5.2 Volume Calculations

The first step in this work was to calculate the box volumes and densities for each composition of interest. Density calculations were performed using an NPT ensemble. System
temperature was fixed at 298.15 K using a Nose-Hoover thermostat. System pressure was fixed at 1 bar using a Parinello-Rahman barostat.

System configurations were initialized at the beginning of each run by randomizing particle positions and velocity. Velocities were sampled from a Maxwell-Boltzmann distribution at 298.15 K. This randomized initialization of system configuration ensured independent runs. However, it also had the potential to result in unphysical interactions between particles due to small distances between particles.

To account for this, the next step was to minimize the system energy. Energy minimization runs were performed using a steepest-descent algorithm for 1000 steps, using an initial step size of 0.0001 nm and a tolerance of 1.0 kJ mol$^{-1}$ nm$^{-1}$. For systems containing $N_{\text{H}_2\text{O}} = 10000$ or greater, the PME long-range electrostatics algorithm failed to converge during energy minimization, so a reaction-field algorithm was used instead.

After energy minimization, production runs were carried out. A leap-frog algorithm was used to integrate the equations of motion. A total of 11000000 time steps were used, with each step being 0.002 ps. The first 2 ns of each run was discarded to allow the system to properly equilibrate after energy minimization, and box volumes were calculated from a thermodynamic average over the remaining steps.

For the volume calculations for the infinite dilution extrapolation portion of this study, each state point was run for 10 replicates to obtain a mean and standard deviation for volume and density. However, only 5 replicates were used for the density calculations of finite concentration boxes, with only a small decrease in precision.

5.3 Extrapolation of Free Energy Calculations to Infinite Dilution

All free energy calculations were performed by gradually removing one solvated target ion from a simulation box containing the ion and $N_{\text{H}_2\text{O}}$ water. Each free energy calculation was performed in an NVT ensemble, where the volume of each box was set to the calculated volumes from the NPT volume calculations at the same composition.

5.3.1 Scaling and Staging

To improve the sampling of states during the removal of a particle from the simulation box, a wide range of intermediate states were used. For this work, a soft-core Lennard-Jones
potential was used to scale the van der Waals interactions, and Coloumb interactions were scaled linearly. The $\lambda$ values used for this study can be seen in Table 5.2:

<table>
<thead>
<tr>
<th>Potential</th>
<th>$\lambda$-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJ</td>
<td>0.0, 0.01, 0.02, ... 0.15, 0.2, 0.3, ... 0.7, 0.8, 0.85, ... 0.99, 1.0</td>
</tr>
<tr>
<td>Coulomb</td>
<td>0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0</td>
</tr>
</tbody>
</table>

Table 5.2: $\lambda$-values used for scaling of the Lennard-Jones and Coloumb potentials in the NaCl+H$_2$O section of this study.

In practice, each $\lambda$ value corresponds to one individual simulation at a state of ($\lambda_{LJ}$, $\lambda_{Coul}$). The $\lambda$ values selected correspond to slowly turning off the charge of the ion until there are no Coulomb interactions, and then slowly turning off the LJ interactions until the particle no longer interacts with any other particle in the simulation.

5.3.2 Equilibration and Production

At the start of each state point, a single run was performed where particle positions were initialized randomly in the simulation box. Molecule velocities were also randomly sampled from a Maxwell-Boltzmann distribution at 298.15 K. System energies were minimized as seen in the density calculations. For the initial run, 13ns of trajectories were simulated in 0.002ps time steps, with the first 3ns being discarded to allow the system to equilibrate. At the end of the run, the final positions of each particle was saved.

Each of the subsequent replicates at each state point used the saved positions from the initial run as a starting configuration. Particle positions were left untouched, while particle velocities were re-initialized from a Maxwell-Boltzmann distribution at 298.15 K. This allowed each run to be independent of the others, while reducing time needed for the system to reach equilibrium.

5.3.3 Analysis of Free Energy Results

Energy output files from GROMACS were analyzed using `alchemical-analysis.py` [56] [57], a free energy calculation tool with a built-in methodology to calculate free energy changes using MBAR. Estimates of statistical uncertainties were performed by performing 10 independent runs for each state point (T,P;m). In each independent run, particle positions and velocities were initialized randomly.
Chapter 6

Results and Discussion

In this chapter we present the results of leading up to the calculation of activity coefficients of an aqueous solution of NaCl from molecular simulation using our newly-developed methodology. We start by calculating $\mu_{\text{res};\text{NVT}}^{\text{NaCl}}(T, \rho[T, P]; m)$ using experimental data from literature to verify our methodology for the calculation of $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ from the extrapolation of experimental residual chemical potentials. Next, we follow the same extrapolation procedure used on the experimental data in Section 6.1 to calculate $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ using molecular simulation. Finally, we use our new values for $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ to calculate activity coefficients for an aqueous sodium chloride solution, both by reanalyzing past literature results, and using our own chemical potential calculations at finite concentration.

6.1 Calculation of $\mu_{\text{NaCl}}^\dagger$ Using Experimental Data

To demonstrate the validity of our methodology before applying it to molecular simulation, we consider its application to experimental data. Doing so requires a set of $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m)$ values calculated at various low-molality concentrations. Residual chemical potentials can be calculated at the molalities used in our simulations by rearranging Equation 2.31:

$$
\mu_{\text{NaCl}}^{\text{res;NVT}} = \mu_{\text{NaCl}}^\dagger(T, P; m^0) - \mu_{\text{NaCl}}^0(T, P^0) - 2RT \ln \left( \frac{RTM_{\text{solv}}x_{\text{solv}}\rho_{\text{solv}}}{1000P^0} \right) + 2RT \ln \gamma_{\text{NaCl}}(T, P; m).
$$

(6.1)

NIST-JANAF[63] values were used for the ideal gas chemical potentials of Na$^+$ and Cl$^-$, where $\mu_{\text{NaCl}}^0 = 334.15$ kJ mol$^{-1}$ [63], and values of $\mu_i^\dagger = -393.133$ were obtained from Wagman [3] values of formation free energy of Na$^+$ and Cl$^-$. The work of Pitzer [4, 5, 6], was cho-
sen to calculate \( \ln \gamma_{\text{NaCl}} \) and density since it contains a set of self-consistent equations for calculating solution density and activity coefficients. Residual chemical potentials, the density term, and activity coefficients were calculated at a range of system compositions, which correspond to molalities of 1 NaCl in \( N_{\text{H}_{2}\text{O}} = [500, 1000, 2000, 3000, 4000, 5000, 6000, 10000, 15000] \). For this section, we ignore the uncertainties in experimental data. A plot of the \( \mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m) \) and the density term can be seen in Figure 6.1.

The data for the terms in Equation 6.1 are given in Table 6.1:

<table>
<thead>
<tr>
<th>( N_{\text{H}_{2}\text{O}} )</th>
<th>( m ) (kg mol(^{-1}))</th>
<th>( \langle \rho \rangle / \text{kg m}^{-3} ) (kg m(^{-1}))</th>
<th>( 2RT \ln(x_{\text{solv}}\rho_{\text{solv}}/\rho_{\text{solv}}^{*}) ) (kJ mol(^{-1}))</th>
<th>( 2RT \ln \gamma_{\text{NaCl}} ) (kJ mol(^{-1}))</th>
<th>( \mu_{\text{NaCl}}^{\text{res;NVT}} ) (kJ mol(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>0.1110</td>
<td>1003.50</td>
<td>12.265×10(^{-3})</td>
<td>-1.289</td>
<td>-744.487</td>
</tr>
<tr>
<td>1000</td>
<td>0.0555</td>
<td>100.28</td>
<td>6.148×10(^{-3})</td>
<td>-1.017</td>
<td>-743.208</td>
</tr>
<tr>
<td>2000</td>
<td>0.0279</td>
<td>998.67</td>
<td>3.078×10(^{-3})</td>
<td>-0.782</td>
<td>-743.969</td>
</tr>
<tr>
<td>3000</td>
<td>0.0185</td>
<td>998.13</td>
<td>2.053×10(^{-3})</td>
<td>-0.662</td>
<td>-743.849</td>
</tr>
<tr>
<td>4000</td>
<td>0.0139</td>
<td>997.86</td>
<td>1.540×10(^{-3})</td>
<td>-0.586</td>
<td>-743.773</td>
</tr>
<tr>
<td>5000</td>
<td>0.0111*</td>
<td>997.70</td>
<td>1.232×10(^{-3})</td>
<td>-0.533</td>
<td>-743.719</td>
</tr>
<tr>
<td>6000</td>
<td>0.0093*</td>
<td>997.59</td>
<td>1.028×10(^{-3})</td>
<td>-0.492</td>
<td>-743.678</td>
</tr>
<tr>
<td>10000</td>
<td>0.0056*</td>
<td>997.37</td>
<td>0.616×10(^{-3})</td>
<td>-0.392</td>
<td>-743.578</td>
</tr>
<tr>
<td>15000</td>
<td>0.0037*</td>
<td>997.27</td>
<td>0.411×10(^{-3})</td>
<td>-0.326</td>
<td>-743.512</td>
</tr>
<tr>
<td>-</td>
<td>0</td>
<td>997.05</td>
<td>0</td>
<td>0</td>
<td>-743.185</td>
</tr>
</tbody>
</table>

Table 6.1: Concentration-dependent terms of aqueous NaCl at dilute concentrations calculated from experimental results of Pitzer\[6, 4, 5\] and \( \mu_{\text{NaCl}}^{I} \) values obtained from Wagman [3]. Values of \( \mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m) \) at molalities indicated by an asterisk * were fit to linear and quadratic polynomials to extrapolate to \( \mu_{\text{NaCl}}^{\text{res;\infty}}(T, P) \).
Figure 6.1: Plots of experimental residual chemical potential values (diamond, left axis) and the logarithmic density term (black circles, right axis), to infinite dilution. Calculations used Equation 6.1 with Wagman $\mu^\dagger_{\text{NaCl}}(T, P; m^0)$ values, NIST-JANAF values for $\mu^0_{\text{NaCl}}$, and Pitzer [4] data for density and $\ln \gamma_{\text{NaCl}}$ data. Points indicated by magneta diamonds were used in fitting to calculate $\mu^\text{res;}_\infty\text{NaCl}(T, P)$.

6.1.1 Our Methodology

As we wish to calculate $\mu^\dagger_{\text{NaCl}}(T, P; m^0)$ from molecular simulation using Equation 2.50, we first want to obtain $\mu^\text{res;}_\infty\text{NaCl}(T, P)$. This can be done by following our procedure described in Section 4.2.1 applied to the data in Table 6.1.

The $\mu^\text{res;}_\text{NaCl}$ values at the four lowest molalities were chosen for the extrapolation as contributions from the logarithmic density term have a quadratic dependence in this region and the values of the term at these molalities are negligible in their contribution to $\mu^\text{res;NVT}_{\text{NaCl}}(T, \rho[T, P]; m)$. This region was fit using a linear and quadratic polynomial in $\sqrt{m}$ using Equations 4.26, 4.27, and 4.28. The results of fitting the data can be seen in Table 6.2. From the intercept $\mu^\text{res;}_\infty\text{NaCl}(T, P)$, $\mu^\dagger_{\text{NaCl}}(T, P; m^0)$ was calculated from Equation 2.50. $A_\gamma$ was simply calculated by dividing the linear coefficient by $2RT$. A plot of these same points with the linear and quadratic fits in comparison to other methodologies can be seen in Figure 6.2
Figure 6.2: Linear (magenta curve) and quadratic (cyan curve) fits to obtain $\mu_{\text{res}}^{\text{ref:∞}}(T, P)$ from experimental $\mu_{\text{NaCl}}^{\text{ref:NVT}}(T, \rho[T, P]; m)$ values (blue curve), in comparison to $\mu_{\text{NaCl}}^{\text{ref:∞}}(T, P)$ values calculated from the Davies-0.2 (black curves), Davies-0.3 (red curves) and Debye-Hückel equations (purple curves). Points indicated by circles were used to fit to linear and quadratic polynomials. Points indicated by diamonds were used to calculate $\mu_{\text{NaCl}}^{\text{ref:∞}}(T, P)$ using the ln $\gamma_{\text{NaCl}}$ approximations ($m^* = 0.0111$ kg mol$^{-1}$ indicated by solid lines; $m^* = 0.0555$ kg mol$^{-1}$ indicated by dashed lines). Subfigure (a) shows the full concentration range used, while Subfigure (b) is zoomed to better show the intercept.
6.1.2 Debye-Hückel and Davies Methodologies

Alternative methodologies used by Mester [7, 8] and Young [9], as described in Section 4.1, involving calculating \( \mu_{\text{NaCl}}(T, P; m^0) \) by fitting to the Debye-Hückel limiting and Davies-0.2 equations at a low concentration \( m^* = 0.0111 \text{ mol kg}^{-1} \). To calculate \( \mu_{\text{NaCl}}^\dagger(T, P; m^0) \) from this methodology, we solve Equation 2.53 for \( \mu_{\text{NaCl}}^\dagger(T, P; m^0) \):

\[
\mu_{\text{NaCl}}^\dagger(T, P; m^0) = \mu_{\text{NaCl}}(T, P; m^*) - 2RT \ln m^* - 2RT \ln \gamma_{\text{NaCl}}(T, P; m^*)
\]  

(6.2)

where:

\[
\ln \gamma_{\text{NaCl}} = -A_\gamma \sqrt{m^*}
\]  

(6.3)

for the Debye-Hückel equation and

\[
\ln \gamma_{\text{NaCl}} = -A_\gamma \left( \frac{\sqrt{m^*}}{1 + \sqrt{m^*}} - 0.2m^* \right)
\]  

(6.4)

for the Davies-0.2 equation. In addition, we can also use the Davies-0.3 equation:

\[
\ln \gamma_{\text{NaCl}} = -A_\gamma \left( \frac{\sqrt{m^*}}{1 + \sqrt{m^*}} - 0.3m^* \right)
\]  

(6.5)

for \( \ln \gamma_{\text{NaCl}}(T, P; m^*) \), as we have previously determined it to be more suitable in low-molality calculations than the other Davies equation. We also applied this methodology to the molality convention proposed by Young [9] by calculating \( \mu_{\text{NaCl}}^{\text{res:NVT}}(T, \rho[T, P]; m) \) at molalities corresponding to the midpoint molalities between \( N_{\text{NaCl}} = 0 \) and \( N_{\text{NaCl}} = 1 \) in \( \text{H}_2\text{O} \).

As seen in Section 4.2.4, the different Davies approximations are suitable for different concentrations. To this extent, we can evaluate the error introduced into \( \mu_{\text{NaCl}} \) by using each approximation. Using the correct value of \( \mu_{\text{NaCl}}^\dagger(T, P; m^0) \) from experiment, and approximate values of \( \mu_{\text{NaCl}}^\dagger(T, P; m^0) \) found by extrapolating the Davies and Debye-Hückel equations at \( m^* = 0.0111 \text{ kg mol}^{-1} \), we can express the total correct and approximate chemical potentials of aqueous NaCl as:

\[
\mu_{\text{NaCl}}^{\text{corr}}(T, P; m) = \mu_{\text{NaCl}}^{\dagger;\text{true}}(T, P; m_{\text{NaCl}}) + 2RT \ln m + 2RT \ln \gamma_{\text{NaCl}}^{\text{true}}(T, P; m)
\]  

(6.6)

\[
\mu_{\text{NaCl}}^{\text{approx}}(T, P; m) = \mu_{\text{NaCl}}^{\dagger;\text{approx}}(T, P; m_{\text{NaCl}}) + 2RT \ln m + 2RT \ln \gamma_{\text{NaCl}}^{\text{approx}}(T, P; m),
\]  

(6.7)

where the subscript ”approx” represents any approximate model of \( \ln \gamma_{\text{NaCl}} \). This allows us
to find the error in $\Delta \mu_{\text{NaCl}}$ from each approximation of $\ln \gamma_{\text{NaCl}}$:

$$
\Delta \mu_{\text{NaCl}}(T, P; m) = \mu_{\text{NaCl}}^{\dag: \text{true}}(T, P; m^*_\text{NaCl}) - \mu_{\text{NaCl}}^{\dag: \text{approx}}(T, P; m^*_\text{NaCl})
$$

$$
+ 2RT \ln \gamma_{\text{NaCl}}^{\text{true}}(T, P; m^*) - 2RT \ln \gamma_{\text{NaCl}}^{\text{approx}}(T, P; m^*)
$$

As Mester [7, 8] and Young [9] use these approximations to calculate $\mu_{\text{NaCl}}^{\dag}(T, P; m^0)$ at a specific concentration $m^*$ such that $\Delta \mu(T, P; m) = 0$, we express the intrinsic error in $\mu_{\text{NaCl}}^{\dag}(T, P; m^0)$ resulting from the approximation as:

$$
\Delta \mu_{\text{NaCl}}^{\dag}(T, P; m^0) = \mu_{\text{NaCl}}^{\dag: \text{true}}(T, P; m^*_\text{NaCl}) - \mu_{\text{NaCl}}^{\dag: \text{approx}}(T, P; m^*_\text{NaCl})
$$

$$
= 2RT \ln \gamma_{\text{NaCl}}^{\text{approx}}(T, P; m^*) - 2RT \ln \gamma_{\text{NaCl}}^{\text{true}}(T, P; m^*)
$$

Thus, we can express the error in $\mu_{\text{NaCl}}(T, P; m)$ resulting from $\ln \gamma_{\text{NaCl}}(T, P; m^*)$ approximations as:

$$
\Delta \mu_{\text{NaCl}}(T, P; m) = \Delta \mu_{\text{NaCl}}^{\dag}(T, P; m^0) + 2RT \ln \left( \frac{\gamma_{\text{NaCl}}^{\text{true}}(T, P; m)}{\gamma_{\text{NaCl}}^{\text{approx}}(T, P; m)} \right)
$$

A plot of the error introduced into chemical potential calculations from using the Davies and Debye-Hückel equations can be seen in Figure 6.3.
Figure 6.3: Experimental values (blue curve) for $\Delta \mu_{\text{NaCl}}(T, P; m) = \mu_{\text{NaCl}}^{\text{approx}}(T, P; m) - \mu_{\text{NaCl}}^{\text{corr}}(T, P; m)$ obtained from using a quadratic fit (cyan curve), linear fit (magenta curve), and Davies-0.2 (black curves), Davies-0.3 (red curves), and Debye-Hückel (purple curves) equations to approximate $\ln \gamma_{\text{NaCl}}(T, P; m)$, and $\mu_{\text{approx}}^{\text{res}; \infty}(T, P; m^0)$ calculated from each respective equation at $m_{\text{NaCl}}^* = 0.0555$ kg mol$^{-1}$ (Subfigure (a)) and $m_{\text{NaCl}}^* = 0.0555$ kg mol$^{-1}$ (Subfigure (b)). At $m = 0$ kg mol$^{-1}$, $\Delta \mu_{\text{NaCl}} = \Delta \mu_{\text{NaCl}}^{\dagger}$. 

63
6.1.3 Comparison of Methodologies

A comparison of results obtained for $\mu_{\text{NaCl}}^{\text{res} \infty}(T, P)$, $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$, and $\Delta \mu_{\text{NaCl}}^\dagger(T, P; m^0)$ from all methodologies are shown in Table 6.2:

<table>
<thead>
<tr>
<th>Source</th>
<th>$\mu_{\text{NaCl}}^{\text{res} \infty}$ kJ mol$^{-1}$</th>
<th>$\mu_{\text{NaCl}}^\dagger$ kJ mol$^{-1}$</th>
<th>$\Delta \mu_{\text{NaCl}}^\dagger$ kJ mol$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>-743.185</td>
<td>-393.133</td>
<td>-</td>
</tr>
<tr>
<td>$m^* = 0.0111$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debye-Hückel</td>
<td>-743.104</td>
<td>-393.052</td>
<td>0.081</td>
</tr>
<tr>
<td>Davies-0.2</td>
<td>-743.176</td>
<td>-393.124</td>
<td>0.009</td>
</tr>
<tr>
<td>Davies-0.3</td>
<td>-743.182</td>
<td>-393.130</td>
<td>0.003</td>
</tr>
<tr>
<td>$m^* = 0.0555$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Debye-Hückel</td>
<td>-742.830</td>
<td>-392.778</td>
<td>0.355</td>
</tr>
<tr>
<td>Davies-0.2</td>
<td>-743.156</td>
<td>-393.103</td>
<td>0.030</td>
</tr>
<tr>
<td>Davies-0.3</td>
<td>-743.189</td>
<td>-393.136</td>
<td>-0.003</td>
</tr>
<tr>
<td>Linear fit</td>
<td>-743.221(11)</td>
<td>-393.181(1)</td>
<td>-0.047</td>
</tr>
<tr>
<td>Quadratic Fit</td>
<td>-743.184(1)</td>
<td>-393.134(1)</td>
<td>-0.001</td>
</tr>
</tbody>
</table>

Table 6.2: Estimates of $\mu_{\text{NaCl}}^{\text{res} \rho}$, $\mu_{\text{NaCl}}^\dagger$, and $\Delta \mu_{\text{NaCl}}^\dagger$ obtained from fitting experimental $\mu_{\text{NaCl}}^{\text{res} \text{NVT}}$ values from fitting to the Davies-0.2, Davies-0.3, and Debye-Hückel equations at $m^* = 0.0111$, as used by Mester [7, 8] and $m^* = 0.0555$ kg mol$^{-1}$, as used by Young [9], as well as from fitting four low-molality points ($m = (0.0111, 0.093, 0.0056, 0.0037)$ to linear and quadratic polynomials. Uncertainties obtained from the fits are given in parentheses and are one standard deviation in the rightmost digit(s).

As seen from Figures 6.3 (a) and (b), using approximations for $\ln \gamma_{\text{NaCl}}$ introduce an intrinsic error in $\Delta \mu_{\text{NaCl}}$ and therefore $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$. At $m^* = 0.0111$ kg mol$^{-1}$, the Davies-0.3 approximates $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ the best. The Davies-0.2 approximation exhibits approximately twice the error, while the Debye-Hückel equation introduces a large error. At $m^* = 0.0555$ kg mol$^{-1}$, as used by Young [9], the Davies-0.2 equation introduces a substantial error in $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ of about 0.03 kJ mol$^{-1}$. Again, the Davies-0.3 results in a much better approximation. This is an intrinsic error present in calculations of $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ by Mester[7, 8] and Young [9], solely from their use of $\ln \gamma_{\text{NaCl}}(T, P; m)$ approximations. However this error, aside from the error from the Debye-Hückel approximation at $m^* = 0.0555$ kg mol$^{-1}$, is quite small, and all values agree with the experimental
\( \mu_{\text{NaCl}}^{\infty}(T, P) \) ans \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) values to one decimal place.

The linear and quadratic fits to experimental \( \mu_{\text{NaCl}}^{\text{res:NVT}}(T, \rho[T, P]; m) \) data resulted in intercepts of \( \mu_{\text{NaCl}}^{\infty}(T, P) = -743.22(1) \) and \( \mu_{\text{NaCl}}^{\infty}(T, P) = -743.184(1) \), respectively. These values correspond to \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) = -393.18(1) \) for the linear fit, and \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) = -393.134(1) \) for the quadratic fit. The error in \( \Delta \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) obtained from our quadratic fit is approximately the same as that found by the Davies-0.3 equation. The error introduced from the linear fit is higher than both the quadratic fit and Davies methodologies, but better than using the Debye-Hückel equation.

In our quadratic fit of low-molality data, \( \mu_{\text{NaCl}}^{\infty} \) agreed exactly with experimental values considering the uncertainty from the fit. Subsequently, \( \mu_{\text{NaCl}}^{\dagger} \) calculated from the \( \mu_{\text{NaCl}}^{\infty}(T, P) \) was also exactly equal to the experimental value. In comparison, using the Davies or Debye-Hückel equations do not reproduce the exact experimental \( \mu_{\text{NaCl}}^{\dagger} \) or \( \mu_{\text{NaCl}}^{\infty}(T, P) \), so this method provides a better estimate for calculating \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \).

As calculated values of \( \mu_{\text{NaCl}}^{\text{res:NVT}}(T, \rho[T, P]; m) \) at the same state point depend only on the experimental value of \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) used for calculating \( \mu_{\text{NaCl}}^{\text{res:NVT}}(T, \rho[T, P]; m) \), the resulting fits would differ only by shifting the value of the intercept by the difference in \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) between data sources. Thus, the choice of a specific experimental data set is not important, as long as consistency is maintained.

As the linear and quadratic fitting to low-molality data is a novel method presented by us, we can further focus on the fitting parameters obtained from each method, as shown in Table 6.3. Unlike the other methods for calculating \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \), in which \( A_{\gamma} \) is specified and used in the calculations, our method results in an estimate of \( A_{\gamma} \) as a fitting parameter. Comparisons with the correct value give an indication of the quality of the fit.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>( \mu_{\text{NaCl}}^{\infty} )</th>
<th>( \mu_{\text{NaCl}}^{\dagger} )</th>
<th>( A_{\gamma} )</th>
<th>Quadratic Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment</td>
<td>-743.183</td>
<td>-393.133</td>
<td>1.1745</td>
<td>8.153</td>
</tr>
<tr>
<td>Linear fit</td>
<td>-743.22(1)</td>
<td>-393.18(1)</td>
<td>0.94(1)</td>
<td>-</td>
</tr>
<tr>
<td>Quadratic Fit</td>
<td>-743.184(1)</td>
<td>-393.134(1)</td>
<td>1.14(3)</td>
<td>6.3(6)</td>
</tr>
</tbody>
</table>

Table 6.3: Estimates of \( \mu_{\text{NaCl}}^{\dagger} \) and \( A_{\gamma} \) obtained from fitting \( \mu_{\text{NaCl}}^{\text{res:NVT}} \) values calculated from Wagman \( \mu_{\text{NaCl}}^{\dagger} \) values to linear and quadratic polynomials in \( \sqrt{m} \). Values of \( \mu_{\text{NaCl}}^{\dagger}(T, P; m^0) \) were calculated from Equation 2.50. Uncertainties obtained from the fits are given in parentheses and are one standard deviation in the rightmost digits.

The experimental value of \( A_{\gamma} \) values calculated from our fit were within the error of the
fitted value of $A_\gamma$ for the quadratic fit, while the linear fit underestimated $A_\gamma$. With additional low-concentration molalities in the fit, this value could be obtained more accurately and precisely.

Important to note, however, is the failure to correctly recover the expected quadratic coefficient. In order to quantitatively calculate this coefficient using experimental data, calculations and fitting of residual chemical potentials at molalities corresponding to one ion pair in $10^8$ H$_2$O ($m = 5.55 \times 10^{-7}$ kg mol$^{-1}$) were necessary. Systems in this range are infeasible in molecular simulation, so we cannot expect to recover the quadratic coefficient when applying this methodology to simulation data.

### 6.2 Calculation of $\mu_{\text{NaCl}}^\dagger$ Using Molecular Simulation

In this section, we present the results of the calculation of $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ using $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m)$ values calculated from quadratic and linear fitting of molecular dynamics simulation data in comparison to other methodologies.

#### 6.2.1 Comparison of Raw Data with Literature Results

Initially, we focused on the removal of a single NaCl ion pair from simulation boxes containing $N_{\text{H}_2\text{O}} = (500, 1000, 2000, 3000, 4000, 5000, 10000)$ water molecules. However, preliminary results showed a failure for the free energy calculations of the ion pair calculations to converge. For a 80ns run time, the ion pair simulations failed to converge until approximately 50-60ns into the simulation. Conversely, the free energy calculations for single ion simulations converge in 3-5ns.

We first compare our simulation data with literature results. Joung and Cheatham [10] provide comparable $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m)$ results by performing single ion simulations, whereas Mester and Panagiotopoulos [7], and Young and Panagiotopoulos [9] used ion pairs. We performed initial simulations using both methodologies, the results of which are shown in Table 6.4 in comparison to other simulation results from Mester [8] and Young [9]:

66
Table 6.4: Simulation results of chemical potential calculations of the removal of a single NaCl particle from water boxes using Joung-Cheatham NaCl ion parameters and SPC/E water. Comparisons between chemical potential calculations using an ion pair and the sum of individual ion chemical potential calculations are presented.

<table>
<thead>
<tr>
<th>$N_{\text{H}_2\text{O}}$</th>
<th>m</th>
<th>$\mu^{\text{res};\rho}_{\text{NaCl}}$ (Ion Pairs)</th>
<th>Literature Value</th>
<th>$\mu^{\text{res};\rho}_{\text{NaCl}}$ (Single Ions)</th>
<th>Literature Value</th>
<th>$\mu^{\text{res};\rho}_{\text{NaCl}}$ (Ion Pairs)</th>
<th>Literature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kg mol$^{-1}$</td>
<td>Our Work</td>
<td>Literature Value</td>
<td>Our Work</td>
<td>Literature Value</td>
<td>Our Work</td>
<td>Literature Value</td>
</tr>
<tr>
<td>450</td>
<td>0.1230</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>743.50 [10]</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>500</td>
<td>0.1110</td>
<td>-743.85(29)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1000</td>
<td>0.0555</td>
<td>-743.70(54)[8]</td>
<td>-743.1(2)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2000</td>
<td>0.0279</td>
<td>-743.54(70)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3000</td>
<td>0.0185</td>
<td>-743.25(62)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4000</td>
<td>0.0139</td>
<td>-742.89(63)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5000</td>
<td>0.0111</td>
<td>-742.62(63)</td>
<td>-742.3(2)</td>
<td>-742.83(12)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6000</td>
<td>0.0093</td>
<td>-742.54(64)</td>
<td>-</td>
<td>-742.72(9)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10000</td>
<td>0.0056</td>
<td>-742.29(65)</td>
<td>-</td>
<td>-742.55(9)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15000</td>
<td>0.0037</td>
<td>-</td>
<td>-</td>
<td>-742.47(6)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

It is seen that our results agree with all literature results within their mutual simulation uncertainties. However, it is important to note where these uncertainties arise. Our data used 10 independent replicate runs at each state point to calculate uncertainty, whereas Mester [7, 8] used the uncertainty estimate from a single BAR run, and Young only used 5 independent runs.

We conclude that our simulation protocols produce the most accurate results for single ion calculations; protocols used by others [7, 9] do not allow enough time for free energy calculations to converge. Due to this as well as the lower uncertainty from the individual ion calculations, we take the results from these simulations as the best values for our extrapolation methodology. Thus, we instead focused on the removal of a single Na$^+$ or Cl$^-$ ion from the simulation boxes, resulting in obtaining $\mu^{\text{res};\rho}_{\text{Na}^+}$ and $\mu^{\text{res};\rho}_{\text{Cl}^-}$ individually, which could then be added. The resulting data could then be used in our extrapolation procedure.
6.2.2 Extrapolation Procedure for Calculating $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$

Results of density and chemical potential calculations at these molalities are given in Table 6.5. Means and uncertainties, given in parentheses, were calculated from 10 independent replicates of each state point.

<table>
<thead>
<tr>
<th>$m$ / kg mol$^{-1}$</th>
<th>$\rho_{\text{soln}}$ / kg m$^{-3}$</th>
<th>$2RT \ln(x_{\text{solv}}\rho_{\text{soln}}/\rho_{\text{solv}}^*)$</th>
<th>$\mu_{\text{NaCl}}^{\text{res,NVT}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01110</td>
<td>999.78(2)</td>
<td>$1.5317 \times 10^{-4}$</td>
<td>-742.83(12)</td>
</tr>
<tr>
<td>0.00925</td>
<td>999.11(3)</td>
<td>$1.1943 \times 10^{-4}$</td>
<td>-742.72(9)</td>
</tr>
<tr>
<td>0.00555</td>
<td>998.93(2)</td>
<td>$6.566 \times 10^{-4}$</td>
<td>-742.55(9)</td>
</tr>
<tr>
<td>0.00370</td>
<td>998.83(2)</td>
<td>$3.148 \times 10^{-4}$</td>
<td>-742.47(6)</td>
</tr>
</tbody>
</table>

Table 6.5: Logarithmic density term and residual chemical potentials of aqueous NaCl at dilute concentrations calculated from molecular dynamics simulations. Uncertainties given in parentheses, are one standard deviation in the rightmost digit and are calculated from the results of 10 independent replicate runs.

Following the same methodology in fitting residual chemical potentials as found in Section 6.1, we fit residual chemical potentials corresponding to $N_{\text{H}_{2}\text{O}} = 5000, 6000, 10000,$ and $15000$ to linear and quadratic polynomials in $\sqrt{m}$. A plot of simulated data points and the accompanying fitted equations can be seen in Figure 6.4 in comparison to the other approximate literature results.

Our value of $\mu_{\text{NaCl}}^{\text{res,\infty}}(T, P) = -742.26(21)$ at 298K and 1 bar, obtained from a weighted fitting of $\mu_{\text{NaCl}}^{\text{res,NVT}}(T, \rho[T, P]; m)$ to a quadratic polynomial is between 0.5 and 1.5 kJ higher than previous values obtained from Mester [7, 8] and Young [9]. This corresponds to a value of $\mu_{\text{NaCl}}^\dagger(T, P; m^0) = -393.20(21)$. The uncertainty of the quadratic fit is almost 5 times smaller than that reported by Young, and approximately the same as that reported by Mester [7, 8]. However it is important to note that this uncertainty reported by Mester is likely inaccurate as it does not account for system size effects at this low concentration. Our method accounts for system size effects by extrapolating to an infinite box size from multiple low-concentration simulations. From a linear fit to our data, we obtained a value of $\mu_{\text{NaCl}}^{\text{res,\infty}}(T, P) = -741.95(3)$. From this, we calculated $\mu_{\text{NaCl}}^\dagger(T, P; m^0) = -391.89(3)$. The uncertainty of the linear fit is much smaller than both our own quadratic fit, and previous calculations by Mester [7, 8] and Young [9]. All $\ln \gamma_{\text{NaCl}}$ approximations used by Mester and Young introduce an error in $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ smaller than the simulation uncertainty.

We also repeated our extrapolation procedure using the molality convention proposed by
Figure 6.4: Quadratic (cyan curve) and linear (magenta curve) fits of simulated $\mu_{\text{NaCl}}^{\text{res};\text{NVT}}$ values calculated using JC ion parameters and SPC/E water parameters. Comparisons to the Davies-0.2 (black curve), Davies-0.3 (red curve), and Debye-Hückel (purple curve) are also presented. Uncertainty in points is to one standard deviation, and is calculated from 10 independent simulations. Uncertainties in the quadratic and linear fits are shown at the indicated abcissa values.

Young and Panagiotopoulos. Regardless of the convention used to ascribe chemical potential calculations to a molality, the intercept $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ of the fit, and therefore $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$, were exactly the same. Next, we applied Young’s extrapolation method to our simulation data by treating each of our simulation data points as constant molality at infinite dilution. This resulted in $\mu_{\text{NaCl}}^{\text{res};\infty} = -741.99(32)$, which has higher uncertainty than our method.

Detailed results from fitting simulation data using the standard molality convention and the Young molality convention, as well as $\mu_{\text{NaCl}}^\dagger(T, P; m^0)$ values calculated using Equation 2.50, can be seen in Table 6.7.
<table>
<thead>
<tr>
<th>Method</th>
<th>Our Work $\mu_{\text{NaCl}}^{\text{res; } \infty}$ kJ mol$^{-1}$</th>
<th>Mester Data $\mu_{\text{NaCl}}^{\text{res; } \infty}$ kJ mol$^{-1}$</th>
<th>Young Data $\mu_{\text{NaCl}}^{\text{res; } \infty}$ kJ mol$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debye-Hückel</td>
<td>-742.0(1) -392.0(1)</td>
<td>-741.7(2) -391.6(2)</td>
<td>-740.5(9) -390.4(9)</td>
</tr>
<tr>
<td>Davies-0.2</td>
<td>-742.2(1) -392.1(1)</td>
<td>-741.8(2) -391.7(2)</td>
<td>-741.0(9) -390.9(9)</td>
</tr>
<tr>
<td>Davies-0.3</td>
<td>-742.2(1) -392.1(1)</td>
<td>-741.8(2) -391.7(2)</td>
<td>-741.0(9) -390.9(9)</td>
</tr>
<tr>
<td>Linear fit</td>
<td>-791.95(3) -391.89(3)</td>
<td>-                                   -</td>
<td>-</td>
</tr>
<tr>
<td>Quadratic Fit</td>
<td>-742.3(2) -392.2(2)</td>
<td>-                                   -</td>
<td>-</td>
</tr>
<tr>
<td>Linear fit*</td>
<td>-741.95(3) -391.89(3)</td>
<td>-                                   -</td>
<td>-</td>
</tr>
<tr>
<td>Quadratic Fit*</td>
<td>-743.2(2) -392.2(2)</td>
<td>-                                   -</td>
<td>-</td>
</tr>
<tr>
<td>Young methodology</td>
<td>-742.0(3) -391.9(3)</td>
<td>-                                   -</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6.6: Comparison of $\mu_{\text{NaCl}}^{\text{res; } \infty}(T, P)$, $\mu_{\text{NaCl}}^{\dagger}(T, P; m^0)$, and $\Delta \mu_{\text{NaCl}}^{\dagger}(T, P; m^0)$ obtained from fitting our simulation data using JC[10] NaCl force field parameters and SPC/E water parameters to linear and quadratic polynomials, and to the Davies and Debye-Hückel equations at $m^* = 0.0111$ kg mol$^{-1}$. We define the quadratic fit to our data to be the best available value for $\mu_{\text{NaCl}}^{\text{res; } \infty}(T, P; m^0)$, thus set it as the “correct” value in calculating $\Delta \mu_{\text{NaCl}}^{\dagger}(T, P; m^0)$. Comparisons to literature values from Mester (Davies-0.2[8] and Debye-Hückel [7] at $m^* = 0.0111$ kg mol$^{-1}$) and Young (Davies-0.2[9] at $m^* = 0.0555$ kg mol$^{-1}$) are also presented. Values indicated by an asterisk in our fitting used the Young molality convention, where the molality corresponds to the midpoint of the molalities before and after removal of a particle. In addition, we also fit our data using Young’s extrapolation procedure by treating all of our simulation points as infinitely dilute.

<table>
<thead>
<tr>
<th>Source</th>
<th>Standard Molality</th>
<th>Young Molality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_{\text{NaCl}}^{\text{res; } \infty}$</td>
<td>$\mu_{\text{NaCl}}^{\dagger}$</td>
</tr>
<tr>
<td>True Value</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Linear fit</td>
<td>-741.95(3) -391.89(3)</td>
<td>1.62(9)</td>
</tr>
<tr>
<td>Quadratic fit</td>
<td>-742.26(21) -392.20(21)</td>
<td>0.2(11)</td>
</tr>
<tr>
<td>Young Extrapolation</td>
<td>-741.99(32) -391.93(32)</td>
<td>0.82(4)</td>
</tr>
</tbody>
</table>

Table 6.7: Comparisons of $\mu_{\text{NaCl}}^{\text{res; } \infty}$ and $A_\gamma$ at 298.15K, 1 bar obtained from fitting residual chemical potentials calculated from molecular simulation with JC ion parameters and SPC/E water parameters using weighted linear and quadratic fits in $\sqrt{m}$, and corresponding $\mu_{\text{NaCl}}^{\dagger}$ values. The value indicated by double asterisks indicates our simulation data applied to Young's extrapolation method treating all of our data points as a constant molality at infinite dilution. Both results using the standard molality convention, and the molality convention proposed by Young are presented in this table.
Our value of $A_\gamma$ obtained from fitting to a quadratic polynomial agrees within uncertainty with the theoretical value obtained from Equation 2.55 using a value of $\epsilon = 73(2)$ for SPC/E water [9], while the linear fit overestimated $A_\gamma$. The linear coefficient obtained from the fits was found to vary slightly between the two conventions, however the uncertainty in the linear coefficient of the fitted quadratic polynomial was approximately 30% higher using the Young convention.

6.3 Calculation of Electrolyte Activity Coefficients

We can use our calculated value of $\mu^{\text{res;NVT}}_{\text{NaCl}}(T, P)$ to further calculate activity coefficients for aqueous NaCl at a range of compositions using Equation 4.29. As system size effects arising from a finite-sized simulation box are prevalent for molalities less than 0.5 molal, residual chemical potentials at compositions below this concentration must be calculated by performing chemical potential calculations at multiple sets of $(N_{\text{NaCl}}, N_{\text{H}_2\text{O}})$ at constant molality, and extrapolated to infinite box size using Equation 4.30.

Results of chemical potential calculations from molecular simulation at various scaled system compositions and the corresponding extrapolations to infinite system size are given in Table 6.8:

As we used large simulation boxes for each set of $m(N_{\text{NaCl}}, N_{\text{H}_2\text{O}})$, all of our simulations at low molalities overlapped within the uncertainties of the extrapolated values at infinite box size. This shows that for large enough boxes, this extrapolation results in minimal improvement for the increased simulation time.

From here, $\ln \gamma_{\text{NaCl}}$ can be calculated using Equation 4.29:

$$2RT \ln \gamma(T, P; m) = \mu^{\text{res;NVT}}_{\text{NaCl}}(T, \rho[T, P]; m) - \mu^{\text{res;NVT}}_{\text{NaCl}}(T, P) + 2RT \ln \left( \frac{x_{\text{solv}} \rho_{\text{solv}}(T, P; m)}{\rho^*_{\text{solv}}(T, P)} \right)$$

(4.29)

Results of these calculations are shown in Table 6.9. Uncertainties to one standard deviation were calculated from the square root of the sum of errors of the error in $\mu^0_{\text{NaCl}}(T, P; m^0)$ and $\mu^{\text{res;NVT}}_{\text{NaCl}}(T, \rho[T, P]; m)$. Due to the small uncertainties in density, the uncertainties in the density term can be neglected. Values were fitted to an extended Debye-Hückel equation:

$$\ln \gamma_{\text{NaCl}}(T, P; m) = -\frac{A_k \sqrt{m}}{1 + B \sqrt{m}} + cm,$$

where $B$ and $c$ are fitting parameters. Fitted parameters are given in Table 6.10.
Table 6.8: Box volumes, densities, and residual chemical potentials for boxes containing finite concentrations of NaCl in H$_2$O, calculated from molecular simulations at 1 bar, 298.15K using JC Na$^+$ and Cl$^-$, and SPC/E H$_2$O. The line in each molality containing $\infty$ NaCl and H$_2$O corresponds to the extrapolated finite concentration chemical potential.

<table>
<thead>
<tr>
<th>$N_{\text{NaCl}}$</th>
<th>$N_{\text{H}_2\text{O}}$</th>
<th>$m$ kg mol$^{-1}$</th>
<th>$\langle V \rangle$ nm$^{-3}$</th>
<th>$\rho_{\text{solv}}$ kg mol$^{-1}$</th>
<th>$\mu_{\text{NaCl}}^\text{res;NVT}$ kJ mol$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5000</td>
<td>0.056</td>
<td>149.875(2)</td>
<td>1001.26(2)</td>
<td>-743.2(2)</td>
</tr>
<tr>
<td>6</td>
<td>6000</td>
<td>-</td>
<td>179.859(2)</td>
<td>1001.21(2)</td>
<td>-742.8(2)</td>
</tr>
<tr>
<td>10</td>
<td>10000</td>
<td>-</td>
<td>299.719(3)</td>
<td>1001.31(3)</td>
<td>-743.2(2)</td>
</tr>
<tr>
<td>$\infty$</td>
<td>$\infty$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-742.8(4)</td>
</tr>
<tr>
<td>15</td>
<td>5000</td>
<td>0.167</td>
<td>150.083(5)</td>
<td>1006.33(5)</td>
<td>-743.3(1)</td>
</tr>
<tr>
<td>18</td>
<td>6000</td>
<td>-</td>
<td>180.108(4)</td>
<td>1006.28(4)</td>
<td>-743.1(1)</td>
</tr>
<tr>
<td>30</td>
<td>10000</td>
<td>-</td>
<td>300.164(4)</td>
<td>1006.34(4)</td>
<td>-743.3(1)</td>
</tr>
<tr>
<td>$\infty$</td>
<td>$\infty$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-743.5(5)</td>
</tr>
<tr>
<td>50</td>
<td>5000</td>
<td>0.556</td>
<td>150.879(2)</td>
<td>1023.54(2)</td>
<td>-743.3(2)</td>
</tr>
<tr>
<td>60</td>
<td>6000</td>
<td>-</td>
<td>181.092(1)</td>
<td>1023.48(1)</td>
<td>-743.5(1)</td>
</tr>
<tr>
<td>100</td>
<td>10000</td>
<td>-</td>
<td>301.808(1)</td>
<td>1023.49(1)</td>
<td>-743.5(2)</td>
</tr>
<tr>
<td>$\infty$</td>
<td>$\infty$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-744.0(7)</td>
</tr>
<tr>
<td>90</td>
<td>5000</td>
<td>1.00</td>
<td>151.854(2)</td>
<td>1042.53(2)</td>
<td>-743.4(2)</td>
</tr>
<tr>
<td>180</td>
<td>5000</td>
<td>2.00</td>
<td>154.508(4)</td>
<td>1081.15(1)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6.9: $\ln \gamma_{\text{NaCl}}$ calculated from $\mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m)$ values obtained from molecular simulation using JC ion parameters and SPC/E water parameters at 298.15K, 1 bar. Uncertainties are to one standard deviation in the rightmost digits. Due to the small uncertainties in density, the uncertainties in the density term can be neglected.

<table>
<thead>
<tr>
<th>$m$ kg mol$^{-1}$</th>
<th>$\mu_{\text{NaCl}}^{\text{res;NVT}}(T, \rho[T, P]; m)$ kJ mol$^{-1}$</th>
<th>$2RT \ln(\rho_{\text{solv}}/\rho_{\text{solv}}^*)$ kJ mol$^{-1}$</th>
<th>$\ln \gamma_{\text{NaCl}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0555</td>
<td>-742.8(4)</td>
<td>0.0012</td>
<td>-0.101(101)</td>
</tr>
<tr>
<td>0.0167</td>
<td>-743.6(5)</td>
<td>0.0035</td>
<td>-0.243(120)</td>
</tr>
<tr>
<td>0.5551</td>
<td>-744.0(7)</td>
<td>0.0102</td>
<td>-0.349(154)</td>
</tr>
<tr>
<td>1.00</td>
<td>-74</td>
<td>0.0163</td>
<td>-0.186(70)</td>
</tr>
</tbody>
</table>

A plot of the fitted $\ln \gamma_{\text{NaCl}}$ v.s. $\sqrt{m}$ in comparison to literature values from Young and Panagiotopoulos can be seen in Figure 6.5.

All of our calculated activity coefficients fall above those calculated by Young. This can in part be accounted for by the difference in $\mu_{\text{NaCl}}^{\dagger}(T, P; m^0)$ used to calculate each set of activity coefficients. All activity coefficients calculated in concentrations below one molal overlap within the mutual uncertainties to those calculated by Young [9]. However, the uncertainties in our calculations are lower than those reported by Young.
Table 6.10: Fitting parameters for activity coefficients using an extended Debye-Hückel equation for the JC NaCl force field parameters at 298.15K, 1 bar.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$</td>
<td>2.677(856)</td>
</tr>
<tr>
<td>$c$</td>
<td>0.182(171)</td>
</tr>
</tbody>
</table>

Figure 6.5: Activity coefficients calculated from Equation 4.29 in molecular dynamics simulations using the Joung-Cheatham NaCl force field and SPC/E water parameters. Results from this work (red) are compared to activity coefficient calculations from Mester and Panagiotopoulos [8], reanalyzed by Young and Panagiotopoulos (black) [9].
Chapter 7

Conclusions

In this thesis, we develop a new methodology for calculating $\mu_i^\dagger$ and $\ln \gamma_{NaCl}$ from molecular simulation by the extrapolation of $\mu_i^{res:NVT}$ values at very low concentrations to the infinitely dilute regime. We first show that for experimental data from macroscopic experiments of aqueous NaCl, this new method recovers the exact $\mu_{NaCl}^\dagger$, while previous methods employed in the literature do not. In addition, we assess the intrinsic error introduced into $\mu^\dagger$ calculations from the use of $\ln \gamma$ approximations.

Next, we apply the new methodology to molecular simulations using the JC NaCl force field parameters and SPC/E water to obtain a new value for $\mu_{NaCl}^\dagger$. This new value allows for the direct calculation of more accurate and precise values of $\ln \gamma_{NaCl}$ for NaCl in water.

For future work, we hope to follow two pathways. First, we want to continue to use this methodology to calculate $\mu_{NaCl}^\dagger$ and activity coefficients at other temperatures commonly found in literature, as well as using other NaCl and H$_2$O force field parameters. Finally, we wish to extend this methodology to calculate $\mu_i^\dagger$, activity coefficients, and other thermodynamic quantities accessible from these properties for the ionic species found in the reaction of candidate alkanolamines with water and CO$_2$.

We note that in doing this, accurate force fields must be obtained for some of the ions found in these reactions, or a switch to hybrid quantum mechanics/molecular modelling must be made to negate the need for force fields. In addition, the Debye-Hückel and Davies equations used to calculate $\ln \gamma_{NaCl}$ are generally only valid for spherical particles of approximately the same size, so other activity coefficient models may also be required.
Bibliography


[38] L. L. Lee, Molecular Thermodynamics of Electrolyte Solutions. WORLD SCIENTIFIC, 2008.


Appendix A

Calculation of $\mu_{^{\text{NaCl}}}^{\dagger}(T, P; m^0)$ at Different Temperatures

While thermodynamic tables often conveniently report Gibbs formation energies at $T_{\text{ref}} = 298.15K$ and $P_{\text{ref}} = 1\text{bar}$ [3], from which we can calculate $\mu_i^{\dagger}(T_{\text{ref}}, P_{\text{ref}})$, it is sometimes desirable to obtain values at other temperatures and pressures. In this section, we describe the calculation of $\mu_i^{\dagger}(T, P)$ using just a reference value and data obtainable from accessible thermodynamic tables [4].

We begin using the definition:

$$\mu_i^{\dagger}(T, P) = \Delta_f H^0(T_{\text{ref}}, P_{\text{ref}}) + [H(T, P) - H(T_{\text{ref}}, P_{\text{ref}})] + T[S(T_{\text{ref}}, P_{\text{ref}}) - (S(T, P) - S(T_{\text{ref}}, P_{\text{ref}})],$$

where $\Delta_f H^0(T_{\text{ref}}, P_{\text{ref}})$ is the enthalpy of formation of the species from its elements at $T_{\text{ref}}$ and $P_{\text{ref}}$. From here, we can add $T_{\text{ref}}S(T_{\text{ref}}, P_{\text{ref}}) - T_{\text{ref}}S(T_{\text{ref}}, P_{\text{ref}}) = 0$:

$$\mu_i^{\dagger}(T, P) = \Delta_f H^0(T_{\text{ref}}, P_{\text{ref}}) + [H(T, P) - H(T_{\text{ref}}, P_{\text{ref}})] + T[S(T_{\text{ref}}, P_{\text{ref}}) - (S(T, P) - S(T_{\text{ref}}, P_{\text{ref}})]$$

Next, grouping like terms, we can simplify:

$$\mu_i^{\dagger}(T, P) = \mu_i^{\dagger}(T_{\text{ref}}, P_{\text{ref}}) + [H(T, P) - H(T_{\text{ref}}, P_{\text{ref}})] + (T_{\text{ref}} + T)S(T_{\text{ref}}, P_{\text{ref}}) - T[S(T, P) + S(T_{\text{ref}}, P_{\text{ref}})]$$

81
where $\mu_i^\dagger(T_{\text{ref}}, P_{\text{ref}})$ is a known value at a reference temperature and pressure, such as those values found in Wagman [3].

This format is convenient, as thermodynamic tables such as those presented by Pitzer [4] contain values for $[H(T, P) - H(T_{\text{ref}}, P_{\text{ref}})]$, $S(T, P)$, and $S(T_{\text{ref}}, P_{\text{ref}})$, allowing for a new $\mu_i^\dagger(T, P)$ to be easily calculated at a large number of temperatures and pressures.