A molecular-based group contribution equation of state for the description of fluid phase behaviour and thermodynamic derivative properties of mixtures (SAFT-γ Mie)

Vasileios Papaioannou

A thesis submitted for the Degree of
Doctor of Philosophy and the Diploma of
Imperial College London

Department of Chemical Engineering
Imperial College London
London SW7 2AZ, United Kingdom

December 2012

Declaration

I hereby declare that all the material in this dissertation is own and has been otherwise properly acknowledged.

Vasileios Papaioannou,
December 2012

Στους γονείς μου,
Χρήστο & Μαρία

Acknowledgements

The work presented in this thesis would not have been possible hadn't it been for the contribution of a number of people. First and foremost I would like to express my sincere gratitude to my supervisors Professors Claire Adjiman, Amparo Galindo, and George Jackson for giving me the opportunity to pursuit a Ph.D. in the Molecular Systems Engineering Group in the first place, and for their restless guidance and support throughout the last 4 years. I would particularly like to thank Prof. Claire Adjiman for her invaluable contribution in both the theoretical developments and the numerical aspects of the work, Prof. Amparo Galindo for her help with understanding phase behaviour and the depth of the theory, and, last but not least, Prof. George Jackson for his passion about the work and his great help by conveying his statistical mechanics knowledge to me. It has been a real privilege for me to work in an interdisciplinary group under the guidance of such passionate and motivated scientists.

A special thanks goes to Tom, Carlos, and Felix for sharing their precious knowledge on SAFT with me, which was crucial at the beginning of my Ph.D. I would also like to thank the members (current and ex) of CPSE for the fun times inside and outside college, and the nice memories of life in London, and particular Frances, Jens, Heiko, Andrei, Simon, Ioscani, Julio, May, Francesco, David, Fausto, Eria, Xenia, Manolis, Eirini, Christianna and Alex. Of particular mention, and of very special contribution stands Apostolis, who has been a very good friend and helped me realise a lot about science and life in general. I will always cherish the memories of my time in Imperial during the last years because of you guys!

During my time in Imperial, I had the chance of meeting Esther, a truly wonderful person, who has been of great importance and help to me, in all aspects of life. I know you don't like hearing this, but I really mean it that I don't know to which extent all of this would have happened without your support and understanding.

Finally, I would like to thank all of dear friends from Greece, in particular Alex, Ntokos, Danos, Johnny, Nikos, John-John, Michalis and Sotiris who even from afar gave me a lot of strength and nice memories of summers in Greece! Last, but not least, my dear parents and brother, to which now I can finally be seen as (kind of) equal!

Abstract

An accurate knowledge of the thermophysical properties and phase behaviour of fluid mixtures is essential for the reliable design of products and processes across a wide range of chemical engineering applications, varying from the processing of petroleum fluids to the manufacturing of pharmaceuticals. Thermodynamic tools and, in the context of this work, group contribution (GC) methods are predictive approaches that are expected to play an important role in meeting these industrial needs. The principal focus of the work presented in this thesis is the development of a novel GC method based on the statistical associating fluid theory (SAFT): the SAFT-γ Mie approach. The method is developed based on a detailed molecular model and a realistic intermolecular potential, the Mie potential with variable attractive and repulsive ranges, for the description of interactions at a molecular level. Over the past decade, an increasing research effort has been devoted to developing formalisms that couple the accuracy of the SAFT equation of state (EoS) with the predictive capabilities of group contribution approaches. In the development of such methods one aims to overcome the limitations inherent to GC approaches based on activity coefficient models, such as in the well-established universal quasi-chemical functional group activity coefficient (UNIFAC) approach. A more recent landmark has been the development of heteronuclear methods within SAFT. The SAFT-γ EoS based on the square-well (SW) potential has been shown to describe accurately the phase behaviour of a wide variety of fluids. In the work presented in this thesis, SAFT-γ SW is applied to the study of the fluid phase behaviour of aqueous solutions of hydrocarbons. These mixtures are of high industrial relevance, and the accurate representation of their highly non-ideal nature is very challenging from a theoretical perspective. The SAFT-γ method is shown to perform comparatively well in predicting the behaviour of the systems examined. Nonetheless, some challenges are identified, such as the description of thermodynamic derivative properties and the description of near-critical fluid phase behaviour, where the performance of the method is shown to be less accurate. These challenges partially arise from the simplistic intermolecular square-well potential employed within SAFT-γ SW, which allows for a rigorous theoretical development, but fails to reproduce accurately finer aspects of the thermophysical behaviour of fluids, such as second-order derivative thermodynamic properties.

These challenges are tackled here with the development of the SAFT-γ Mie GC approach, based on the versatile Mie intermolecular potential and a third-order treatment of the thermodynamics of the monomer segments. The SAFT-γ Mie method is applied to the study of the properties of two chemical families, *n*-alkanes and 2-ketones, and it is shown that a significant improvement over existing SAFT-based group contribution approaches can be achieved in the description of the pure component phase behaviour of the compounds studied. Moreover, the application of

a realistic intermolecular potential is shown to allow for an excellent description of second-order derivative thermodynamic properties, and the accurate treatment of the intersegment interactions is shown to improve the performance of the method in the description of the near-critical fluid phase behaviour. The predictive capability of the method is demonstrated in the description of mixture fluid phase behaviour and excess thermodynamic properties in a predictive manner. Given the promising performance of the SAFT-γ Mie EoS, the method is applied to the case study of the solubility of two active pharmaceutical ingredients in organic solvents. The method is shown to satisfactorily predict the solubilities of the mixtures considered, based on limited experimental data for simple systems. Given the complexity of the mixtures studied, the performance of the SAFT-γ Mie is considered very encouraging and shows that there is great potential in the application of the method to this challenging field.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Thermodynamic tools are being continually developed and improved in order to meet the need for accurate property prediction required common to many sectors of the chemical industries. Predictive approaches have come to play an important role in the design of processes and the accuracy of their output can significantly affect process-design decisions [1, 2]. Despite the wide variety of methods that are available, industrial requirements in property prediction highlight the need for accurate approaches that can be applied in an ever-expanding range of applications, including, among others, polymer processing, biotechnology and solvent screening [3]. Aside from design purposes, advances in thermodynamic modelling have led to tools that have found novel applications, such as the integrated design of solvents and processes, where molecular characteristics of solvents are determined as part of the optimisation of the process [4, 5]. An important aspect in the development of thermodynamic methodologies is their predictive capability, which is commonly perceived as the capability to provide predictions of fluid phase behaviour and other bulk properties without the need for experimental data for the determination of the molecular model parameters.

One of the first examples that attests the strength of predictive molecular approaches in the study of fluid phase behaviour is the following quote in which Kammerlingh-Onnes credits van der Waals for his help in achieving the liquefaction of hydrogen [6]:

> *"In what I describe to you, your theory has been my guide. The calculations were performed entirely on the basis of the law of corresponding states. Guided by the law, I estimated to need 20 litres [of hydrogen]. Had I estimated a few litres fewer, the experiment would not have succeeded".*

Predictive thermodynamic models are particularly suited for applications in the general areas of computer-aided molecular design (CAMD) and fluid formulations. In CAMD,

the desired properties of a given substance are set as an input and computational tools are used in the task of determining the molecular structure that satisfies the given prerequisites. Although this methodology was originally executed by means of heuristics and costly experimental procedures, the availability of predictive (and more specifically group contribution) methods allows the formulation of the corresponding task in terms of a more precise optimisation problem [7–10]. These tools find particular application in the design of processes in the pharmaceutical industry [11]. A common problem to be addressed in this sector is the selection of the appropriate solvent or solvent blend as part of the design of the drug manufacturing process. Despite the development and improvement of measurement techniques, solvent selection (or solvent screening, as it is commonly referred to) is difficult to accomplish experimentally as a large number of experiments have to be conducted to cover a considerable range of solvents, even more so for solvent blends where different compositions of the blend have to be studied in a discrete way. An additional limitation to the experimental determination of optimal solvents is posed by the limited availability of the solute (drug) for laboratory studies, mainly due to the difficulty and cost of manufacture in the early stages of drug development [12]. These limitations can be overcome by the application of predictive thermodynamic tools for the selection of the optimal solvent/solvent blend, that can explore a large design space of compounds and, in the case of solvent blends, investigate the effects of composition in a continuous way.

A vast body of research has been devoted to the development of a specific class of predictive methodologies, the so-called group contribution (GC) approaches. Within GC methods, molecules are decomposed into chemically distinct functional groups, so that a mixture of components is now regarded as a solution (mixture) of chemical groups. The basic underlying assumptions of methodologies of this kind are that molecular properties can be calculated as an appropriate sum of contributions that are attributed to these functional groups, and the contributions of a given group are the same, regardless of the host molecule that it appears in. Typically, a set of group-specific parameters that fully characterises a functional group is obtained by regression to experimental data. Once a set of functional groups has been fully characterised, these groups can be combined to predict the properties of compounds and mixtures in a fully predictive manner. The predictive power and applicability of methodologies of this kind stems from the fact that once a small number of functional groups have been characterised, the methods can be applied to the study of the properties for a wide range of mixtures. The concept of solution of groups is empirical; however, a detailed statistical mechanical analysis of the underlying assumptions of such an approach within activity coefficient methods can be found in [13]. A critical review of the broad range of GC methods, spanning from applications for the prediction of pure

component properties to phase behaviour of mixtures is given in chapter 2.

Perhaps the group contribution methodology which is most widely employed in industrial applications is the universal quasi-chemical functional group activity coefficient (UNIFAC) method [14]. UNIFAC was developed based on the seminal work of Guggenheim on a quasi-chemical theory [15], where the combinatorial part is calculated based on a statistical derivation of the free energy of a lattice; the attractions between segments are described by a Wilson-type expression [16]. The popularity of the UNIFAC approach is due to the extensive parameter table available, which now covers more than 80 functional groups and 1200 group interaction parameters [17], and to the very good performance of the method in the description of fluid phase equilibria. More details of the theoretical basis of the method, the evolution and development of modifications to the approach, as well as some limitations are also presented in chapter 2. Despite its popularity, the UNIFAC method is subject to a number of restrictions, as for example the study of different types of phase behaviour (vapour-liquid, liquid-liquid), the application to polymer and electrolyte systems or the consideration of pressure effects (since the theory stems from a lattice-fluid model). These limitations have given rise to novel group contribution approaches, a very recent example being the formulation of the GC concept within the scope of the statistical associating fluid theory (SAFT) [18–20].

The statistical associating fluid theory (SAFT) [21–24] is a general framework for the development of equations of state based on the statistical thermodynamics of associating chain molecules. Within SAFT a detailed molecular model is employed and the development of the theory is facilitated without the restriction of a lattice. More detail on the theoretical background of SAFT approaches with a focus on the reformulation of SAFT within the spirit of a group contribution methodology is given in section 2.4. Despite the successful application of SAFT-type GC approaches to the modelling of highly non-ideal and thermodynamically challenging systems [25–28], challenges remain, such as the description of second-order thermodynamic derivative properties, that require for further development of the theory. Addressing some of these challenges is one theme of the work described in this thesis.

The thesis is structured as follows: in chapter 2, a critical literature review of group contribution methodologies is presented, where the broad range of methodologies are discussed in the context of different applications. The main focus is placed on the popular GC activity coefficient method, UNIFAC, and the recent formulation of GC approaches within the framework of SAFT.

The performance of a promising SAFT GC method, the SAFT-γ equation of state (EoS) [19, 29] based on the square-well intermolecular potential to describe the interactions between groups, is subsequently examined in chapter 3 in the context of the description of the fluid phase behaviour of aqueous solutions of hydrocarbons and alcohols. Systems of this kind are of great industrial interest and the accurate description of the highly non-ideal phase behaviour they exhibit is a challenge for the majority of common thermodynamic models. In the course of the application of the SAFT-γ EoS, the issue of group identification is also discussed; the effect that this can have on the quality of the predictions of the theory is examined in some detail. In the study of the performance of the SAFT-γ approach, several challenges are revealed, mainly associated with the description of the second-order derivative properties and the performance of the theory in the description of the near-critical region. The performance of the method in the description of derivative properties is a direct consequence of the simplified intermolecular potential employed for the description of intersegment interactions. The square-well potential features a hard wall of infinite repulsion at contact and a finite attraction range; this simple functional form allows for the development of a compact theory in a rigorous manner, however it has proven to be too simplistic to simultaneously describe a wide range of properties of real fluids. In particular, caloric properties, which are highly relevant to process design, cannot be predicted accurately based on models that represent the fluid phase behaviour (vapour-liquid equilibria) well.

The aforementioned challenges are addressed through the development of a theory based on a more realistic and versatile intermolecular potential, such as the Mie potential, a generalised Lennard-Jonesium potential of variable attractive and repulsive ranges. The functional form of the Mie potential makes the development of the theory somewhat more cumbersome. However, the application of such a potential has been shown to provide an accurate description of fluid phase behaviour and derivative properties [30, 31]. Furthermore, the use of a Mie potential comes with the added advantage of retaining a direct link between an analytical theory and molecular simulation; molecular simulation techniques can be employed for the study of elements of fluids and fluid mixtures difficult to access by means of analytical theories, such as interfacial properties and structured phases. In the case of the square-well potential this is difficult to achieve, mainly due to the discontinuity of the potential form.

The main contribution of this thesis has been the development of a novel group contribution approach, the SAFT-γ Mie method, based on the Mie intermolecular potential.

Details of the development of the theory are presented in chapter 4, together with an assessment of the performance of the theory in the description of the fluid phase behaviour and derivative properties of real compounds, and the study of binary mixtures in a predictive manner. A key application of the predictive theory presented is the study of solubility of complex molecules in solvents and solvent blends. An example of the application of the SAFT-γ Mie approach to the study of the solubility of complex organic molecules (used as active pharmaceutical ingredients) in organic solvents is presented in chapter 5.

The work presented in the thesis is summarised in chapter 6, highlighting the contributions that have been made. Recommendations and directions for future work are briefly discussed.

Chapter 2

Group Contribution Methods

Group contribution methodologies have a long history. The main applications of GC methods can be divided in the following three categories:

- the calculation of pure component properties;
- the calculation of the activity coefficients of the liquid phase(s) in mixtures;
- the coupling with equations of state to treat the liquid and vapour phase.

In the next sections, the three different approaches are thoroughly discussed in order to provide with an understanding of the basic characteristics of each of the aforementioned categories. Furthermore, a contrast between the different methods is drawn and needs for improvement in predictive capabilities are identified.

2.1 Pure Component GC Methods

In these methods empirical correlations of pure component experimental properties, such as the critical temperature T_c, pressure p_c, volume V_c, the acentric factor ω, boiling point T_b, etc., are proposed based on expressions that incorporate information on the chemical groups forming the molecule. To obtain descriptions of phase behaviour these properties can be used following the corresponding states principle or equations of state. So-called first- and higher-order methods have been developed.

2.1.1 First-order pure component methods

The sole input with first-order pure component GC methods are the types and number of functional groups that appear in a given substance; the position and, more importantly, the connectivity between the functional groups are neglected. These methods in principle

cannot be used to distinguish between isomers that consist of the same functional groups or to take into account proximity effects (chemical and physical changes in a group that may occur due to the proximity of another group leading to substantially different physicochemical properties).

One of the first very successful methods for estimating critical properties was presented by Lydersen in 1955 [32]. The contributions to the critical properties of 43 different functional groups including ring and non-ring carbon, oxygen, halogen, nitrogen and sulfur were calculated by regression to experimental data; the quoted average absolute deviations (AADs) for the critical temperature, pressure, and volume were 8.2 K, 3.3 bar, and 10 cm^3 mol^{-1} respectively for the more than 200 compounds tested for each property [33]. This method is mostly preferred for the prediction of the critical properties of hydrocarbons and more specifically of alkylcycloalkanes, branched alkenes, and alkynes [34]. Later, a group contribution method for the calculation of critical properties that demonstrated predictions of higher accuracy than the method of Lydersen was presented by Ambrose [35, 36]; the average absolute errors reported in this case were 4.3 K for T_c, 1.8 bar for p_c, and 8.5 cm^3 mol^{-1} for V_c with a database containing values for over 200 compounds for each property [33]. The main difference between the two methods is that in the method of Ambrose the Platt number (a measure of the branching in the molecule) is included.

Joback and Reid [37] reassessed Lydersen's method, added more functional groups, and extended the capabilities of group contribution approaches to the calculation of a wide range of properties, including the normal boiling T_b and freezing temperature T_f, the enthalpy of vapourisation at the normal boiling point $\Delta H_{vap,b}$, the room-temperature (298 K) enthalpy of formation $\Delta H_{f,298}$, the enthalpy of fusion at the triple point ΔH_{fus}, the room-temperature Gibbs free energy of formation $\Delta G_{f,298}$, the ideal gas heat capacity c_p^0, and the liquid viscosity η_L (the latter two properties given as functions of temperature). The method leads to significant average absolute errors in the prediction of T_b and T_f (on average 17.9 K and 24.7 K, respectively) and should be considered as approximate. Regarding the critical properties, and more specifically the critical temperature, this is within the method of Joback and Reid a secondary property (following the notation of Constantinou and Gani [38]), as its value depends on the molecular structure as well as on the normal boiling point, which in turn is a primary property. When the primary property (normal boiling point) used in the prediction of the critical temperature is the one predicted by the method, high deviations are expected (AAD of 25.01 K), whereas when an experimental value is used the accuracy of the approach is significantly improved

with an AAD of 6.65 K [39]. The performance of the method as presented by the authors for the critical properties is comparable to the method of Ambrose [35] with reported AADs of 4.8 K, 2.1 bar, and 7.5 cm^3 mol^{-1} for T_c, p_c, and V_c respectively. However the GC approach of Joback and Reid is easier to implement and has therefore gained more popularity than the more cumbersome method of Ambrose. The performance of the Joback-Reid method in the estimation of the normal boiling point was addressed in a later study by Devotta and Pendyala [40], where a reevaluation of some functional groups was presented, and new groups were introduced to improve the prediction of properties of perhalogenated substances, in particular perfluorinated compounds. The method provides a higher accuracy in the prediction of normal boiling point and can be used to differentiate between some isomers accurately. A new GC-based correlation for the calculation of T_b and T_c based on the assumption that the boiling temperature and critical temperature reach finite values at very high (infinite) molar mass has recently been presented [41]. Such an approach results in an percentage absolute average deviation (%AAD) for the boiling temperature T_b of 3.5% over a range of 942 compounds and for two pressures, and is considered more accurate than the method of Joback and Reid [37], which leads to an %AAD of 4.7% for the same compounds. The proposed critical temperature correlation results in a 2.6% AAD from experimental values.

Of particular interest is the work of Jalowka and Daubert [42], who examined whether the including of nearest-neighbour effects can improve on the predictions of the critical properties by first-order group contribution approaches. The inclusion of the nearest-neighbour effects is conducted by following the group definition of the earlier work of Benson and co-workers [43] for the prediction of the ideal gas heat capacity, where functional groups are defined starting from the central carbon atom together with a bond that indicates the ligands the carbon atom is bonded to. Corrections accounting for the structure of substances are also introduced, such as the correction for the ring formation, a cis correction for the alkenes and an ortho correction for aromatic compounds. The presented approach was shown to lead to an improvement in the accuracy of the predictions of the critical properties for the chemical families of cycloalkanes, alkenes and aromatics over the method of Ambrose [35]. The parameter table of the method was later extended to organic compounds containing oxygen, nitrogen, sulfur and halogens [44].

A common feature of the methods discussed thus far is the fact that the calculation of the critical properties, and specifically the critical temperature T_c, requires a knowledge of the normal boiling point. In most cases, the use of experimentally determined values of the boiling point leads to an acceptable prediction of the critical point, whereas when

T_b is also predicted with the GC approach, the deviations in T_c rise significantly. An effort to provide a more accurate prediction of the boiling point has been presented by Stein and Brown [45]. The functional groups originally presented by Joback and Reid are extended by subdividing or joining existing groups when this appears to result in a better representation of T_b. For example, Joback and Reid employ a general OH group for the description of alcohols, whereas within the Stein and Brown approach, different group parameters are regressed for the OH group of a primary, secondary of tertiary alcohol. New functional groups were also examined and the resulting group parameter table included 85 different groups with an overall AAD of 15.5 K for T_b. The new group contributions were derived from a much larger data set than that used by Joback and Reid [37] with an average percent error of 3.2%. The predictive capability of the method has been tested in the prediction of the boiling point for a set of more than 6500 compounds not included in the regression and results in an AAE of 20.4 K (this translates to a 4.3% AAD).

A fully predictive method for the estimation of the critical temperature of pure components based solely on the number of occurrences of its functional groups, i.e., without the use of T_b, has also been presented by Fedors [46], who developed contributions for over 30 different functional groups and reported a mean deviation of 15 K for T_c for around 200 different compounds.

2.1.2 Second-order pure component methods

An important shortcoming of GC approaches which are formulated at the first-order level is that the actual position of a group in a molecule and its connectivity are not accounted for. In an effort to overcome these two deficiencies and introduce a group contribution approach with better accuracy and wider range of application, Constantinou and Gani [38] presented an approach where the properties of pure compounds are calculated on a two-level basis according to the following general relation

$$f(X) = \sum_{i=1}^{N_{G,1}} N_i C_i + W \sum_{j=1}^{N_{G,2}} M_j D_j \ , \tag{2.1}$$

where $f(X)$ is a function of property X (e.g., for the critical temperature $f(X) = \exp(T_c/181.128)$), $N_{G,1}$ and $N_{G,2}$ the total number of first- and second-order groups, respectively, C_i is the contribution of a first-order group i with N_i the number of occurrences, and D_j and M_j the second-order contributions and occurrences of group type j, respectively. The second-order group contribution can be optionally included or discarded for the property prediction by assigning the corresponding value (i.e., 0 or 1) to the constant

Table 2.1: First- and second-order group identification of isomeric dimethylhexanes (C_8H_{18}) in the framework of the GC method of Constantinou and Gani [38]

	2,3-Dimethylhexane	2,4-Dimethylhexane	2,5-Dimethylhexane
	First-Order Groups		
CH_3	4	4	4
CH_2	2	2	2
CH	2	2	2
	Second-Order Groups		
$CH(CH_3)CH(CH_3)$	1	0	0
$(CH_3)_2CH$	0	1	2

W.

First-order groups are simple functional groups providing basic information on the molecular structure. The identification of first-order groups in the work of Constantinou and Gani [38] differs from the one applied in previous pure-component approaches (e.g., Joback and Reid [37], Ambrose [35]). The procedure followed is instead the one applied for the estimation of properties of mixtures (as, e.g., in the UNIFAC approach). Second-order groups are used to provide more information of the molecular structure including the connectivity. They are constructed using first-order groups as structural units and are primarily used to increase the accuracy of the method. One can use a second-order formulation to distinguish between isomers and to some extent capture proximity effects. The difference between first- and second-order groups and how this procedure helps distinguishing between isomers can be illustrated by means of the example given in table 2.1, in which it can be seen that three isomers of dimethylhexane have the same representation in terms of first-order groups but are characterised by different combinations of second-order groups.

The second-order groups are identified according to the principle of conjugation [47–49]. The method was initially formulated for the calculation of T_b, T_m, T_c, p_c, V_c, the enthalpy of vapourisation at 298 K, $\Delta H_{v,298}$, $\Delta H_{f,298}$, and $\Delta G_{f,298}$, and was subsequently extended to the calculation of the acentric factor ω and the room-temperature molar volume of the liquid $V_{l,298}$ [50]. The consideration of second-order groups improves the performance of the proposed pure-component predictive method typically resulting in significantly improved accuracy when compared to the earlier techniques [32, 35, 37]. More

specifically, Constantinou and Gani [38] report AADs of 4.85 K for T_c, 0.113 MPa for p_c, 6.00 cm^3 mol^{-1} for V_c, and of about 3 kJ mol^{-1} for both $\Delta H_{f,298}$ and $\Delta G_{f,298}$. The improvement in the accuracy of the prediction of the normal boiling and freezing points compared to the method of Joback and Reid [37] is also impressive, going from AADs of 13 K with the method of Joback and Reid to 5.35 K for T_b and from 22.6 K to 14 K for T_f. An added advantage of the approach of Constantinou and Gani is the ability to distinguish between isomers and the fact that it leads to a reasonable description in the limit of very long alkane chains for the critical temperature, pressure, and volume.

Compared to the method of Jalowka and Daubert [42], where Benson-type groups are employed for the accounting of nearest-neighbour effects, the Constantinou and Gani method has a less complicated scheme of group identification. Furthermore, all properties are calculated based solely on the contributions of functional groups as primary properties and do not require other determined properties, as in earlier approaches [35, 37, 42] where, e.g., the boiling point is required for the prediction of T_c and in some cases, T_c is required for p_c.

2.1.3 Higher-order pure component methods

Following a similar scheme as presented in section 2.1.2, a function of a certain property $f(X)$ at the third-order level can be described using [51]

$$f(X) = \sum_{i=1}^{N_{G,1}} N_i C_i + W \sum_{j=1}^{N_{G,2}} M_j D_j + Z \sum_{k=1}^{N_{G,3}} N_k E_k \ , \tag{2.2}$$

where the first and second terms on the right-hand side have the same meaning as in eq. (2.1); E_k is the contribution of a third-order group of type k that appears O_k times in a compound and Z is a constant that can take the value 1 or 0 depending on whether third-order contributions are included or not.

At the first-order level a large number of groups are characterised, which can provide an accurate description of a wide range of simple and monofunctional organic compounds. However, only a partial treatment of proximity effects is incorporated, and the theory cannot be used to distinguish between isomers that comprise the same functional groups. The consideration of second-order contributions allows one to achieve an improved treatment of more complex molecules, i.e., multifunctional, polar or non-polar molecules with an average size ranging from 3 to 6 carbon atoms and aromatic or cycloaliphatic compounds with only one ring. At this level a better description of proximity effects and increased

differentiation between isomers are possible. In the case of even more complex, heterocyclic (e.g., with fused rings) and large multifunctional acyclic compounds (comprising 7 to 60 carbon atoms) the application of third-order groups is required to allow satisfactory predictions. For a detailed explanation of the identification of groups of different order, the reader is directed to the original publication [51].

The properties for which third-order group contributions have been derived to date are: T_b, T_m, T_c, p_c, V_c, $\Delta G_{\mathrm{f},298}$, $\Delta H_{\mathrm{f},298}$, $\Delta H_{\mathrm{f},298}$ and standard enthalpy of fusion (at the melting temperature, T_m) $\Delta H_{\mathrm{fus,m}}$. The number of first-order groups identified was 182, whereas 122 second-order, and 46 third-order groups were used to improve the accuracy and reliability. The approach has an improved performance when compared to previous methods with AADs of 5.89 K , 4.87 K, 0.74 bar, and 7.25 $\mathrm{cm}^3\ \mathrm{mol}^{-1}$ for T_b, T_c, p_c, and V_c, respectively, calculated for a range of more than 500 compounds. Note, however, that the accuracy of the prediction does not show a significant improvement compared to the second-order method of Constantinou and Gani [38].

2.1.4 Further developments for pure component approaches

In addition to the predictive methods based on the contribution of groups already discussed, other techniques have been developed in an effort to increase the predictive accuracy of property estimation methodologies. Pure compound properties are calculated by taking into account aspects of the molecular structure in addition to the functional groups appearing and their number. One such method is the ABC approach [47], where the compounds are regarded as hybrids of different conjugates. Every property is then estimated by summing the atom and bond contributions of different conjugates. The difficulty in applying this technique lies in the generation of conjugates, which is far from trivial [52]. A similar approach has been presented by Marrero-Morejón and Pardillo-Fontdevila [53], where the prediction of pure compound properties is based on the group-interaction contribution (GIC) approach [54]. Here, the contributions of interactions between bonded groups instead of the contributions of simple groups are taken into consideration. Popular approaches for the prediction of pure component properties also include those based on *molecular descriptors*. These quantitative structure-property relationships (QSPR) apply quantum mechanical calculations and relate component properties to molecular descriptors (such as the molecular surface and electrostatics), see for example [55]. Connectivity indices offer an alternative to GC approaches by using graphical theoretical concepts to describe the topology of molecular structures [56]. In more recent work, Gani *et al.* [57] have shown how connectivity-index and GC methods can be used in a complementary fashion

to develop a more predictive hybrid technique, GC$^+$. The GC$^+$ offers a framework for the estimation of the contribution of functional groups in a predictive manner, based on the atoms that the group comprises and how these are bonded. The connectivity indices for atoms are obtained by regression to a large amount of experimental data; once obtained they can employed for the prediction of the contributions of functional groups appearing on molecules for which no experimental data is available. Even in cases where limited experimental data is available, the GC$^+$ method is expected to lead to group parameters of statistical significance compared to the parameterisation of the entire group. This is an important generic development that can significantly increase the predictive capabilities of group contribution methods.

2.2 Activity Coefficient GC Methods

Activity coefficient methods have proved particularly useful as GC methods for the prediction of the properties of mixtures, usually their fluid phase behaviour. In engineering applications the solution of phase equilibrium is often formulated by means of the isofugacity criterion of components, i.e., by equating the fugacities (or chemical potential) of each component i in all the phases at equilibrium at a given temperature T and pressure p [58]

$$\begin{aligned} f_i^a(T,p,\underline{x}_a) &= f_i^b(T,p,\underline{x}_b), \\ \text{for} \quad i &= 1,2,\ldots,N_\mathrm{C}\ ;\ a=1,\ldots,N_\mathrm{P}-1\ ;\ b=2,\ldots,N_\mathrm{P}\ ;\ b>a\ , \end{aligned} \tag{2.3}$$

where $\underline{x}_a$ and $\underline{x}_b$ are the compositions of the phases, N_C is the number of components, and N_C is the number of phases. In the case of vapour (V) - liquid (L) equilibrium, neglecting the Poynting correcting factor allows for the equation of the fugacity of pure species i as a liquid at the specified T and p with its vapour pressure ($f_i^\mathrm{L}(T,p) = p_i^\mathrm{vap}(T)$), so that the above expression is usually rewritten as

$$y_i p \phi_i^\mathrm{V}(T,p,\underline{y}) = x_i p_i^\mathrm{vap}(T)\gamma_i(T,p,\underline{x})\ ,\ i=1,\ldots,N_\mathrm{C}, \tag{2.4}$$

where γ_i is the activity coefficient of component i, ϕ_i^V its fugacity coefficient, $p_i^\mathrm{vap}(T)$ is the saturated vapour pressure of pure component i at the specified temperature, and $\underline{x}$ and $\underline{y}$ refer to the mole fractions of the liquid and the vapour phase, respectively. Under the assumption of an ideal gas phase (i.e., $\phi_i^\mathrm{V} = 1$) this reduces to the commonly used expression

$$y_i p = x_i p_i^\mathrm{vap}(T)\gamma_i(T,p,\underline{x})\ ,\ i=1,\ldots,N_\mathrm{C}, \tag{2.5}$$

which describes the vapour–liquid equilibrium of nonideal fluids, where γ_i is used to incorporate the nonidealities of the liquid phase.

Many different models have been proposed for the prediction of activity coefficients in multicomponent mixtures. Group contribution approaches find application in this area as they provide one with the ability to estimate the values of activity coefficients in cases where no experimental data are available. In accordance with the general proposition of group contribution approaches the molecules appearing in a mixture are decomposed in distinct functional groups. Inter- and intra-molecular interactions are then calculated as weighted averages of interactions between functional groups. The basic assumption here is that a group exhibits the same behaviour regardless of the molecule in which it appears. The fundamental advantage of this approach is that the number of different functional groups that appear in a set of mixtures is far smaller than the combinations of molecules. As an example, by obtaining the parameters describing the methyl CH_3 and methylene CH_2 groups and their group-based interactions one can describe all pure components mixtures of saturated linear hydrocarbons.

The two most successful methods for the estimation of activity coefficients within the scope of a group contribution approach are the analytical solution of groups (ASOG method) and the universal quasi-chemical functional group activity coefficient (UNIFAC) methods, mostly due to the size of the parameter table available. These approaches are discussed in some detail below. Other activity coefficient GC methods include DISQUAC [59], SUPERFAC [60], effective UNIFAC [61], and SIGMA [62], but we do not discuss these further as they are closely related to ASOG and UNIFAC. For a more complete discussion the reader is directed to the comprehensive review by Gmehling [63].

2.2.1 The ASOG method

The analytical solution of groups method was developed by a number of authors in several stages (see Papadopoulos and Derr [64], Redlich *et al.* [65], Wilson and Deal [66], and Derr and Deal [67]). In this approach, the activity coefficients are calculated as consisting of two parts, by means of the following equation

$$\ln \gamma_i = \ln \gamma_i^{\mathrm{S}} + \ln \gamma_i^{\mathrm{G}}, \text{ for } \quad i = 1, \ldots, \mathrm{NC} \ , \tag{2.6}$$

where NC denotes the total number of components in the mixture. The first term (superscript S for size) is the configurational contribution due to differences in molecular shape and the second term accounts for nonidealities arising from intermolecular forces. The

configurational term is calculated following the Flory-Huggins theory [68] as

$$\ln \gamma_i^{\mathrm{S}} = 1 - R_i + \ln R_i \ , \tag{2.7}$$

with

$$R_i = \frac{\sum_{k=1}^{N_{\mathrm{G}}} \nu_{k,i}}{\sum_{j=1}^{N_{\mathrm{C}}} \sum_{l=1}^{N_{\mathrm{G}}} x_j \nu_{l,j}}, \ \text{for} \quad i = 1, \ldots, N_{\mathrm{C}} \ , \tag{2.8}$$

where $\nu_{k,i}$ is the number of groups of type k in component i, x_i is the mole fraction of component i in the system, and N_{G} denotes the number of types of functional groups present in the system. The contribution to the activity coefficient arising from the intermolecular forces is calculated as

$$\ln \gamma_i^{\mathrm{G}} = \sum_{k=1}^{N_{\mathrm{G}}} \nu_{k,i} \ln \Gamma_k - \sum_{k=1}^{N_{\mathrm{G}}} \nu_{k,i} \ln \Gamma_{k,i}^* \ , \tag{2.9}$$

Eq. (2.9) requires a knowledge of the activity coefficients of group k in two different "states": the term Γ_k represents the activity coefficient of group k in the mixture and is calculated as

$$\ln \Gamma_k = -\ln \sum_{l=1}^{N_{\mathrm{G}}} X_l A_{kl} + \left(1 - \sum_{l=1}^{N_{\mathrm{G}}} \frac{X_l A_{lk}}{\sum_{m=1}^{N_{\mathrm{G}}} X_m A_{lm}} \right) \ , \tag{2.10}$$

with X_l the mole fraction of groups of type l in the mixture, given by

$$X_l = \frac{\sum_{i=1}^{N_{\mathrm{C}}} x_i \nu_{l,i}}{\sum_{j=1}^{N_{\mathrm{C}}} x_j \sum_{l=1}^{N_{\mathrm{G}}} \nu_{l,j}} \ . \tag{2.11}$$

The term $\Gamma_{k,i}^*$ refers to the value of the activity coefficient of group k in a standard state, defined as a mixture of groups in which only the groups appearing in substance i are present. For example, if component i is benzene there is only one kind of group (group aCH) and then $\ln \Gamma_{\mathrm{aCH,C_6H_6}}^*$ is zero. In the case of methanol, however, where two kinds of groups appear (CH_3 and OH), both $\ln \Gamma_{\mathrm{CH_3,CH_3OH}}^*$ and $\ln \Gamma_{\mathrm{OH,CH_3OH}}^*$ have an finite value. Generally the standard-state coefficients are calculated as

$$\ln \Gamma_{k,i}^* = -\ln \left(\sum_{l=1}^{N_{\mathrm{G}}} X_{l,i}^* A_{kl} \right) + \left(1 - \sum_{l=1}^{N_{\mathrm{G}}} \frac{X_{l,i}^* A_{lk}}{\sum_{m=1}^{N_{\mathrm{G}}} X_{m,i}^* A_{lm}} \right) \ . \tag{2.12}$$

In this case, the necessary quantity for the calculation is the fraction $X^*_{k,i}$ of group type k in component i (as opposed to the mole fraction of group type k in the mixture in eq. (2.11)). This fraction is calculated as

$$X^*_{k,i} = \frac{\nu_{k,i}}{\sum_{l=1}^{N_G} \nu_{l,i}} \ . \tag{2.13}$$

The group interactions are accounted for using the so-called Wilson parameters,

$$A_{kl} = \exp\left(m_{kl} + \frac{n_{kl}}{T}\right) \ , \tag{2.14}$$

which are based on temperature-independent group interaction parameters m_{kl} and n_{kl}. It is important to note that the group interaction parameters are non-zero only for unlike interactions between groups (i.e., $m_{CH_2-CH_2} = 0$ and $n_{CH_2-CH_2} = 0$, whereas $m_{CH_2-COO} \neq 0$ and $n_{CH_2-COO} \neq 0$). The temperature-dependent Wilson parameters are non-symmetric, i.e., $A_{kl} \neq A_{lk}$, due to the underlying assumption of local compositions [16]. These parameters can be obtained by regression to vapour-liquid equilibrium (VLE) experimental data and infinite dilution activity coefficients [67]. Once they are obtained for a specific group interaction, they can be used in a predictive manner for any mixture where that specific interaction appears.

The parameter table of the ASOG method initially comprised 31 groups [69], including groups describing (among others) the chemical families of alkanes, alkanols, ketones, amines, and aromatic hydrocarbons. A distinct group for water was also included. Later the original parameters were revised and extended to include a total of 43 structural groups and 341 group pairs [70, 71]. The accuracy of the method for the prediction (AAD) of VLE is $\Delta y = 0.0219$ for the gas composition, $\Delta T = 1.63$ K for the temperature, and $\Delta p = 14.56$ mmHg (≈ 0.02 atm) for the pressure over a range of about 7000 data sets [72].

2.2.2 The UNIFAC method

The UNIFAC method was initially presented by Fredenslund *et al.* [14] and has been the subject of numerous publications since then. It is the most widespread GC method, and is considered the tool of choice by many academic and industrial groups to predict the phase behaviour of non-ideal mixtures. An overview of the evolution of UNIFAC is presented in the reviews of Gmehling [63] and Fredenslund [73, 74]. In a similar way to ASOG, it incorporates the solution-of-groups assumption for the estimation of activity coefficients in

mixtures. In UNIFAC activity coefficients are calculated as the sum of two contributions, given by

$$\ln \gamma_i = \ln \gamma_i^{\mathrm{C}} + \ln \gamma_i^{\mathrm{R}}, \quad \text{for} \quad i = 1, \ \ldots, N_{\mathrm{C}} \ , \tag{2.15}$$

where N_{C} is again the total number of components in the mixture. The term labelled C represents the contribution of the combinatorial part, which accounts for the differences in the size and shape of the molecules, and the term labelled R represents the residual part which expresses the contribution of intermolecular forces to the nonideality of a mixture.

In contrast with the ASOG method, where the combinatorial contribution is described by means of the Flory-Huggins athermal equation (eq. (2.7)), in UNIFAC, the universal quasi-chemical (UNIQUAC) [75] expression based on Guggenheim's [15] theory is used, which is given by

$$\ln \gamma_i^{\mathrm{C}} = \ln \frac{\Phi_i}{x_i} + \frac{z}{2} q_i \ln \frac{\theta_i}{\Phi_i} + l_i - \frac{\Phi_i}{x_i} \sum_{j=1}^{N_{\mathrm{C}}} x_j l_j \ , \tag{2.16}$$

where

$$l_i = \frac{z}{2}(r_i - q_i) - (r_i - 1) \ . \tag{2.17}$$

Here x_i, Φ_i, θ_i are the mole, volume and area fractions of component i, with

$$\Phi_i = \frac{r_i x_i}{\sum_{j=1}^{N_{\mathrm{C}}} r_j x_j} \ , \tag{2.18}$$

and

$$\theta_i = \frac{q_i x_i}{\sum_{j=1}^{N_{\mathrm{C}}} q_j x_j} \ , \tag{2.19}$$

z is the coordination number of the underlying lattice (which is usually set to 10), and x_j is the mole fraction of component j. In these expressions r_i and q_i are measures of the molecular van der Waals volumes and molecular surface areas, respectively. These are calculated using a weighted sum of the group contributions for the overall molecular volume and surface as

$$r_i = \sum_{k=1}^{N_{\mathrm{G}}} \nu_{k,i} R_k \ , \tag{2.20}$$

and

$$q_i = \sum_{k=1}^{N_\mathrm{G}} \nu_{k,i} Q_k \ , \tag{2.21}$$

where $\nu_{k,i}$ is the number of occurrences of a group of type k in component i. The group specific parameters R_k and Q_k are obtained using the van der Waals group volumes and areas [76]. Regarding r_i and q_i as pure component parameters, the methodology presented thus far is common in both UNIQUAC and UNIFAC. The only difference lies in the expression for the calculation of the residual term. In UNIQUAC this term is calculated from

$$\ln \gamma_i^\mathrm{R} = q_i \left[1 - \ln \left(\sum_{j=1}^{N_\mathrm{C}} \theta_j \tau_{ji} \right) - \sum_{j=1}^{N_\mathrm{C}} \frac{\theta_j \tau_{ij}}{\sum_{w=1}^{N_\mathrm{C}} \theta_w \tau_{wj}} \right] \ , \tag{2.22}$$

where

$$\tau_{ij} = \exp \left(-\frac{u_{ij} - u_{jj}}{RT} \right) \ , \tag{2.23}$$

The parameters τ_{ij} are measures of the interactions between components i and j. As in the case of the ASOG method, the interactions are expressed in a way that $\tau_{ij} \neq \tau_{ji}$ (even though $u_{12} = u_{21}$), since the expression is derived based on local compositions. These parameters have to be evaluated by regression to experimental data.

In UNIFAC the expression for the residual contribution to the activity coefficient of component i is replaced by the solution of groups concept, giving

$$\ln \gamma_i^\mathrm{R} = \sum_{k=1}^{N_\mathrm{G}} \nu_{k,i} (\ln \Gamma_k - \ln \Gamma_k^{(i)}) \ , \tag{2.24}$$

where the summation is for each component over all group types. In eq. (2.24) Γ_k represents the residual activity coefficient of group k, and $\Gamma_k^{(i)}$ is the residual activity coefficient of the same group in a solution containing only molecules of component i. In a similar way as for the ASOG approach (cf. eq. (2.10)), the first term is calculated as

$$\ln \Gamma_k = Q_k \left[1 - \ln \left(\sum_{l=1}^{N_\mathrm{G}} \theta_l \Psi_{lk} \right) - \sum_{l=1}^{N_\mathrm{G}} \frac{\theta_l \Psi_{kl}}{\sum_{m=1}^{N_\mathrm{G}} \theta_m \Psi_{ml}} \right] \ , \tag{2.25}$$

where θ_l is in this case the area fraction of group l, as compared to eq. (2.19) where the area fraction is component based. Hence,

$$\theta_l = \frac{Q_l X_l}{\sum_{m=1}^{N_G} Q_m X_m} \ , \tag{2.26}$$

and

$$X_m = \frac{\sum_{i=1}^{N_G} \nu_m^{(i)} x_i}{\sum_{j=1}^{N_C} \sum_{n=1}^{N_G} \nu_n^{(j)} x_j} \ . \tag{2.27}$$

The residual part of the activity coefficient of group k in component i is given in eq. (2.28). Here, a different expression of the area fraction $\theta_m^{(i)}$ is used, calculated in a similar way as in eq. (2.26), but based only on the groups appearing in component i rather than on all the groups present in the mixture:

$$\ln \Gamma_k^{(i)} = Q_k \left[1 - \ln \left(\sum_{l=1}^{N_G} \theta_l^{(i)} \Psi_{lk} \right) - \sum_{l=1}^{N_G} \frac{\theta_l^{(i)} \Psi_{kl}}{\sum_{m=1}^{N_G} \theta_m^{(i)} \Psi_{ml}} \right] \ , \tag{2.28}$$

$$\theta_l^{(i)} = \frac{Q_l X_l^{(i)}}{\sum_{m=1}^{N_G} Q_m X_m^{(i)}} \ , \tag{2.29}$$

and

$$X_l^{(i)} = \frac{\nu_l^{(i)}}{\sum_{m=1}^{N_G} \nu_{m,i}} \ , \tag{2.30}$$

where $X_l^{(i)}$ is the fraction of groups of type l in component i. The group-group interactions are represented by the term Ψ_{mn} which is calculated based on a Wilson-type expression [16] as

$$\Psi_{mn} = \exp\left(-\frac{u_{mn} - u_{nn}}{RT} \right) = \exp\left(-\frac{a_{mn}}{T} \right) \ . \tag{2.31}$$

As before the parameters $a_{mn} \neq a_{nm}$ are temperature independent and groups of the same type are assumed not to interact, i.e., $a_{mm} = 0$. They are determined by estimation from

experimental VLE data and are listed in the parameter table of UNIFAC.

In addition to the similarities already highlighted between ASOG and UNIFAC for the calculation of activity coefficients, the two methods become almost identical as the coordination number of the underlying lattice, z, is increased; for large chain molecules and more specifically in the limit where $q_i/r_i \to 1$ the UNIFAC expression for the calculation of the combinatorial part of the activity coefficient reduces to the athermal Flory-Huggins equation (eq. (2.7)). Moreover, the residual contributions used in the two methods become identical when the contribution that a group makes to the molecular area is taken to be the same for all groups.

In UNIFAC group-group interactions are calculated on the basis of the main groups; subgroups are also identified as further distinctions of main groups that are assigned different values for their contribution to the molecular volume and area. Considering for example a linear alkane, the molecule is decomposed in the subgroups CH_3 and CH_2 with different values of the parameters R_{CH_3}, Q_{CH_3} and R_{CH_2}, Q_{CH_2}. When it comes to the binary mixture of an alkane with, for example, water (H_2O), the group-group interaction parameter a_{mn} has the same value for the interaction H_2O - CH_3 and H_2O - CH_2 since the functional units of the alkane fall into the same main group CH_n, which includes the subgroups CH_3, CH_2 and CH (note: H_2O is modelled as one main group). In the original form of UNIFAC the parameter table comprised 18 main groups (expanding to a total of 26 subgroups) [14].

As with other group contribution methods, UNIFAC is necessarily approximate since it is developed under the assumption that the behaviour of a group does not depend on the environment in which it is found. This approximation can be overcome by selecting increasingly specific groups, the limit being the representation of each molecule as a single group. However, as the number of group interactions to be determined increases dramatically, the decomposition into structural groups loses its relevance and advantage. Initially, the choice of functional groups was based on experience and sometimes trial-and-error, in the sense that the final decision relied on an assessment of the optimal description for the given system. As an example of this, Fredenslund *et al.* [77] point out that the addition of a CH_2 group to a smaller group, such as the OH group, often improves the correlation. However, such choices can sometimes have a negative effect on the predictive capability of the method. This is clear in the case of alcohols, where the optimal fit to experimental data was obtained for a two carbon-atom group (CH_2CH_2OH), a grouping that can only be applied to the modelling of primary alcohols with two consecutive CH_2 groups next to the OH group.

During the procedure of estimating the group parameters, the square of the deviation of the calculated phase behaviour (typically vapour-liquid equilibria) from the experimental values is usually used as the objective function to be minimised,

$$\min_{\mathbf{\Omega}} f_{\text{obj}} = \sum_{i=1}^{N_\text{C}} \sum_{j=1}^{N_\text{P}} \Delta\text{VLE}_{ij} \ , \tag{2.32}$$

where the summations are over all components (N_C) and experimental points (N_P) and $\mathbf{\Omega}$ is the vector of the estimated parameters [77]. The parameters estimated are only the ones describing the group-group interactions, i.e., $\mathbf{\Omega} = \{a_{mn}, a_{mn}\}$ for $m = 1, \ldots, N_\text{G}$ and $n = 1, \ldots, N_\text{G}$; the volume ($R_k$) and area ($Q_k$) parameters are obtained from the van der Waals volume and surface areas given by Bondi [76]. In the original UNIFAC approach the group group interactions are regressed one pair at a time. In, for example, the problem of obtaining the interaction parameters in a mixture of a carboxylic acid and an alcohol, the functional groups for the acid are CH_3, CH_2 and COOH and the separate OH group is considered for the alcohol. The interaction of the alkyl groups with the COOH and OH groups, respectively, would be determined first; then the OH-COOH interaction is considered. Mathematically, all interactions could, of course, be estimated simultaneously, and such a procedure may indeed yield a good description of the fluid phase equilibria of the alcohol+acid mixtures. Such a procedure would restrict the reliability of transferring the parameters to other mixtures, as the total of the obtained parameters would be biased to the optimal description of the specific type of mixtures. Furthermore, the partial revision or extension of the parameter table would be more difficult, as opposed to determining parameters in a sequential manner.

2.2.2.1 Limitations of the original UNIFAC

Generally UNIFAC is recognised as an accurate predictive method for the calculation of activity coefficients for a wide variety of mixtures. This has led to its implementation in many commercial process simulators and its use is widespread in industry. Nevertheless, the original UNIFAC method suffers from a number of limitations which result from the following issues [60]:

- Activity coefficient approach

  Within UNIFAC the nonideal nature of the liquid phase is accounted by the activity coefficient. The vapour phase is typically assumed to be ideal (i.e., the fugacity coefficient $\phi_i^V = 1$). This approach to the solution of vapour-liquid phase equilibria is commonly referred to as the γ - ϕ approach. The assumption of an ideal vapour

phase is accurate only for low or moderate pressures, i.e., below 10-15 atm; at higher pressures the vapour phase has to be treated explicitly by a complementary thermodynamic model, e.g., a cubic EoS, such as the SRK EoS. This requires the knowledge of additional molecular pure component parameters. In addition, since the γ-ϕ approach does not treat the vapour and liquid phases on equal footing, it cannot be applied to the description of mixtures at conditions where a vapour-liquid critical point appears. Furthermore, since UNIFAC is based on a local-composition model, it cannot be readily applied to describe phase equilibria in systems containing ions, salts, polymers (since density effects are not considered) or non-condensable gases.

- Solution of groups approach

  As has already been mentioned, UNIFAC is based on the solution of groups concept, which assumes that the behaviour of a structural group is independent of its immediate environment, i.e., the molecule in which it appears. In other words when describing a certain group one does not take specifically into account the neighbouring groups on the molecule. As a result, the method cannot be used to distinguish between isomeric forms of compounds that are described by the same groups. The same basic assumption is the reason why UNIFAC fails to capture proximity effects (the difference in the behaviour of, for example, two diols where the OH groups are located further or closer apart). The fact that neighbouring groups can be distinguished with analytical techniques such as infrared (IR) or nuclear magnetic resonance (NMR) spectroscopy testifies to the presence of proximity effects due to polarisation.

- The UNIFAC functional form

  The parameters used in UNIFAC are estimated from regression to experimental VLE data of appropriate mixtures. It is therefore not a surprise that the use of this parameter table often leads to an inaccurate description of liquid-liquid equilibria (LLE) [78]. This is a major deficiency of the UNIFAC approach, since different parameters (based on LLE data) need to be developed; the solution of three phase equilibria (VLLE) becomes an important issue as different parameter sets are employed for the solution of different types of equilibria. Moreover, the use of temperature independent parameters for the group-group interactions restricts the applicability of UNIFAC to a temperature range of 275 - 425 K and leads to an erroneous description of the temperature dependence of the activity coefficients. As a consequence only a qualitative agreement of the predicted excess enthalpies with experimental

data [79]. Finally, the assumptions inherent in the lattice fluid theory on which UNIFAC is based, do not allow for the explicit accounting of pressure effects; the method cannot describe, e.g., the variation of the excess enthalpy of a mixture with pressure.

- Experimental data

  The accuracy of the predictions with the UNIFAC platform depends critically on the availability and quality of the experimental data used for the determination of the group parameters. As examples of this, UNIFAC predictions are usually not very accurate in the prediction of infinite dilution activity coefficients, γ_i^{∞}, and phase behaviour of asymmetric systems, i.e., where the components are very different in size (e.g. ethane+n-eicosane), as for such systems scarce experimental data are available. The availability of experimental data in the regression of the group parameters may also lead to restrictions of the applicability of UNIFAC in thermodynamic conditions, as the restriction to the temperature region of 275 - 425 K, mentioned in the previous point.

2.2.2.2 Modifications of the method

In order to address some these limitations, a number of modifications to UNIFAC have been proposed. Early improvements of the method were based on the amount and quality of the experimental data used in the parameter estimation. The theoretical basis of the method remained unchanged; a series of publications concentrated on revising and extending the existing parameter table [80–83]. The use of the Dortmund Data Bank of experimental data, in the UNIFAC parameter estimation and the use of more elaborate objective functions than the one presented in eq. (2.32) were then introduced [84]. In a series of publications on the revision of UNIFAC [80–82] new parameters were reported based on the updated experimental database. The last publication reporting parameter values for the original UNIFAC was presented by Hansen *et al.* [83] and Fredenslund and Sørensen [74] and features 50 main groups (and 108 secondary groups). Among these are chlorofluorohydrocarbons (Freons) and mixtures of relevance to biotechnology. Today, the parameter table of UNIFAC contains more than 65 main groups [17]. The reported relative deviations for the predictions of the approach for 3300 VLE data sets are $\Delta p = 1.89$ kPa for the pressure, $\Delta T = 1.25$ K for the temperature and $\Delta \mathrm{y} = 1.54\%$ for the composition. The inadequacies of UNIFAC in predicting liquid-liquid equilibria have also been studied and a parameter table containing values for 32 main structural groups based on experimental LLE data has been presented [78]. The original UNIFAC methodology can be then applied successfully for the prediction of liquid-liquid equilibria. However, the

inconsistency of having two different parameter tables within a single approach can lead to complications, as for example in the case of VLLE calculations.

Modifications of UNIFAC have also been presented at the level of the theory. The two most popular approaches were the modified UNIFAC (Dortmund) [85] and the modified UNIFAC (Lyngby) [86]. Both methods were modified by introducing an empirical term in the calculation of the combinatorial contribution to the activity coefficient based on the improved performance of the method and a modification; the residual part was also modified within the two aforementioned methods by the addition of two temperature independent parameters, so that every group-group interaction is characterised by a total of six adjustable parameters. A comparative study between the Dortmund and Lyngby modifications of UNIFAC was presented by Gmehling *et al.* [87], where it was shown that the modified UNIFAC (Dortmund) performs better in the description of infinite dilution activity coefficients.

In the modified UNIFAC (Dortmund) the values of the parameters describing the volume and area contributions (R_k and Q_k) were not taken from Bondi [76] as in the original UNIFAC, but were calculated together with the group-group interaction parameters by regression to experimental data. The fitting procedure for the determination of the group parameters within the modified UNIFAC (Dortmund) was undertaken including vapour-liquid equilibrium data (VLE) at different conditions ($p-x$ data at constant T, $T-x$ data at constant p, and $x-y$ data at constant T), as well as data for other properties such as liquid-liquid equilibrium (LLE), excess enthalpies (h^E), and excess heat capacities (c_p^E). The estimated parameters, aside from the volume and surface contributions of the groups, include the ones used for the calculation of group-group interactions (a_{nm}, b_{nm}, and c_{nm}) by means of the following relation

$$\Psi_{nm} = \exp\left(-\frac{a_{nm} + b_{nm}T + c_{nm}T^2}{T}\right) . \tag{2.33}$$

This simultaneous fitting procedure provides the modified UNIFAC approach with the capability of predicting, in addition to the VLE of mixtures, excess enthalpies of mixing, activity coefficients at infinite dilution, liquid-liquid, and solid-liquid equilibria. Since it was first published [85] the modified UNIFAC (Dortmund) platform has been revised and its parameter table extended [72, 88–92]. In the most recent revision Jakob *et al.* [93] refer to the ability of combining the method with the conductor-like screening model for real solvents (COSMO-RS); the COSMO-RS approach is used to obtain pseudo experimental data from quantum-mechanical calculations. The modified UNIFAC approach uses

the resulting data to obtain parameters (by estimation) even in cases where experimental measurements are unavailable. At present the parameter table of this method contains parameters for 85 main groups and 1293 parameters describing group-group interactions [17]. The reported accuracy of the method is $\Delta p = 1.19$ kPa, $\Delta T = 0.93$ K, and $\Delta \mathrm{y} = 1.11\%$. The improvement of the accuracy of the method is more impressive in the predictions for activity coefficients at infinite dilution and excess enthalpies.

Extensions to polymeric systems by adding a free-volume term and residual activity coefficient terms [94–97], to electrolyte solutions by combination with the Debye-Hückel theory [98, 99], and to associating mixtures [100] have also been presented. Association interactions are accounted for by the addition of a term based on Wertheim's first-order perturbation theory [101–104] in the combinatorial and residual terms of UNIFAC. Also of relevance is the extension of UNIFAC to incorporate second-order groups [38, 105, 106], which are taken into consideration in the calculation of the molecular surface and volume (eq. (2.20)) and the residual term of the activity coefficient (eq. (2.24)). A further important development was made by Wu and Sandler [107, 108] who set out a theoretical basis for group identification, in contrast with the usual selection of groups according to experience or only on the basis of best fit to experimental data. They proposed group identification according to two criteria: (1) the geometry of a group should be independent of the molecule; (2) each atom in a group should have the same charge in all molecules and the group should be approximately electroneutral. The theory is supported by the increased accuracy of predictions when the groups used within UNIFAC result from this methodology. Finally, González *et al.* [109] presented a method for the extension of the application range of UNIFAC. The method, referred to as the Connectivity Index UNIFAC (CI-UNIFAC), is based on the work of Gani *et al.* [57] and is used to predict missing group-group interactions by "re-organising" of the existing UNIFAC groups. As a result, there is no need for new experimental data and nor for a regression procedure for the determination of any missing parameters.

2.3 GC methods in equations of state

UNIFAC and its modifications are widely used and considered to be the state-of-the-art GC tools when it comes to predictions of fluid phase behaviour. However, the limited range (both in temperature and pressure) of reliable applicability of the methods and the inconsistency in the description of the critical region by means of the γ - ϕ approach are important drawbacks. Furthermore, the use of activity coefficient methods is limited to the calculation of a subset of properties and cannot be used to provide values for

other properties (such as heat capacities, densities, etc.) which are often required in process design. In contrast, equations of state approaches are not limited, in principle, in their range of application, and more importantly can treat liquid and vapour phases on an equal footing, and, through thermodynamic relations, can be used to calculate other thermodynamic properties. In general however, equations of state (EoS) are less predictive as component-based parameters are needed for their successful application; the calculations of the properties of mixtures invariably require the introduction of combining rules and adjustable mixture parameters. In an effort to use equations of state in a predictive manner, they have been combined with group contribution methods. These developments are reviewed in this section.

2.3.1 EoS-g^E methods

Most equations of state require a knowledge of component-specific parameters for accurate calculations. In the case of pure components the parameter estimation problem is relatively straightforward and the values for each component can be calculated based on the critical constants, as in the case of simple cubic equations of state [39], or through parameter estimation based on, e.g., pure component phase behaviour data with the more sophisticated approaches. In the case of calculations for mixtures, the parameters are also mixture-specific and have to be calculated based on the pure component and binary interaction parameters with the use of mixing and combining rules. The simplest form of mixing rules for use in an equation of state are the seminal expressions presented by van der Waals [110] which employ quadratic functions of the pure component parameters and the composition of the mixture; combining rules (such as the Lorentz-Berthelot) are usually taken as arithmetic for the unlike size and geometric for the unlike energy parameters [111]. The quadratic form of the mixing rule has a theoretical basis in the composition dependence of the second virial coefficient, but the Lorentz-Berthelot combining rules are well-known to fail in the case of highly non-ideal mixtures [112].

An interesting approach to bring predictive capabilities to EoSs is the use of the excess Gibbs energy of mixing, g^E, of a system calculated by means of an equation of state equated to the value obtained by means of a direct (predictive) method such as UNIFAC; these are the so-called EoS-g^E methods. Initially the equation of state is used to obtain a relation between the excess Gibbs energy and the parameters appearing in the equation of state. Then, and under specific conditions (these differ according to the method used), g^E is equated to the value obtained by an activity coefficient model either in the limit of zero (g^E_0) or infinite (g^E_∞) pressure. An expression is then obtained that relates the

EoS-related parameter(s) to the composition of the mixture; this is typically established for the parameter that represents the intermolecular interactions within the equation of state, commonly referred to as, a_{mix}. The co-volume parameter, b_{mix}, is typically calculated by means of simple mixing rules (cf. eq. (2.35)). It is important to note, however, that only some of the binary interaction parameters required to model a given mixture can typically be obtained with this approach and that pure component parameters must be obtained by regression to experimental data by using reliable combining rules, where appropriate. Huron and Vidal [113] have used this EoS - g^E method with the SRK equation of state [114] combined with two different methods for the calculation of the excess Gibbs energy (at infinite pressure): the Redlich-Kister expansion [115] and the NRTL method [116]. The method has proved to be rather successful and various modifications of the mixing rules presented have followed [117–119]. For a review see [39, 120–122].

Of particular relevance to our discussion are the methods in which the calculation of the excess Gibbs energy of mixing is performed by means of a GC method. In this case, only the pure component parameter of the EoS have to be determined for the calculation of mixture properties; the binary interaction parameters are obtained in a predictive fashion from the GC method. One of the most widely applied predictive EoS-g^E methods is the predictive Soave-Redlich-Kwong (PSRK) [123]. In such an approach the SRK-EoS [114] with the modified Huron-Vidal mixing rule of a first-order approximation (MHV1) [118] is employed. A temperature dependence in the attractive interaction parameter a_i is incorporated and the following expressions are proposed

$$a = b\left[\frac{g_0^E}{A_1} + \sum_{i=1}^{N_C} x_i \frac{a_i}{b_i} + \frac{RT}{A_1}\sum_{i=1}^{N_C} x_i \ln\frac{b}{b_i}\right] , \tag{2.34}$$

and

$$b = \sum_{i=1}^{N_C} x_i b_i, \tag{2.35}$$

as mixing rules for the parameters a and b of the EoS. The zero pressure limit used in the MHV1 mixing rules allows for the calculation of g^E from UNIFAC without a recalculation of the existing parameters. The latter correction is necessary when the matching of g_0^E is done at the infinite pressure limit, as in, e.g., the original Huron-Vidal mixing rules [123]. In a later study, it was shown that the PSRK method fails in the description of highly asymmetric systems [124]. This shortcoming gave rise to a new g^E model, the volume-translated Peng-Robinson (VTPR) EoS [125–127], aiming to replace the PSRK methodology.

Another important EoS-g^E is the PPR78 EoS of Jaubert and Mutelet [128]. PPR78 (predictive Peng-Robinson, 1978) is based on the Peng-Robinson EoS [129], where the binary interaction parameters necessary for the modelling of mixtures are calculated based on contributions of functional groups, in a similar way to the work of Peneloux and coworkers [130–132]. It is important to mention that, within the PPR78 EoS a van Laar expression is employed as a g^E model; the existing UNIFAC parameter table cannot be used and the contributions of the functional groups have to be estimated based on experimental data. The PPR78 has been extensively applied to the study of petroleum fluids, such as hydrocarbons with carbon dioxide [133], nitrogen [134], and hydrogen sulfide [135]. Several other EoS - g^E models have been suggested, such as UNIWAALS [136], MHV2 [117, 137], LCVM [138], and the predictive equations of state of Lermite and Vidal [139] and Orbey *et al.* [140].

2.3.2 GC methods directly implemented in equations of state

A direct description of the parameters of an equation of state within a GC formalism is also possible. This idea has been used together with the simplified perturbed hard chain theory (SPHCT) [141], and in the lattice fluid (LF) approach [142] derived from the Panayiotou and Vera EoS [143]. Other approaches in this area include the chain-of-rotators group contribution (CORGC) EoS [144], the work of Coniglio *et al.* [145, 146] based on the popular Peng-Robinson (PR) EoS [129], and the GC-EoS based on the lattice fluid theory presented by Mattedi *et al.* [147]. More recent examples are the GC approaches of Elvassore *et al.* [148], and Elliott and Natarajan [149].

One of the most widely applied group contribution EoS models is the GC-EoS [150]. The method is based on a generalised van der Waals partition function and the group contribution formalism. The attractive contribution is calculated based on a group contribution NRTL expression [116], while the repulsive hard-sphere residual free energy is calculated by means of the Carnahan-Starling EoS [151]. The parameters for more than 35 different chemical groups were presented, and the method was revised and extended in a series of publications [152–155]. The GC-EoS has been combined with the association term based on the seminal work of Wertheim [101–104], as is employed within SAFT-type models, in the development of the GCA-EoS [156]. The GCA-EoS is one of the few examples of cubic equations of state that explicitly account for associating effects and has been tested in the description of highly nonideal associating and non-associating systems – for a recent review of the applications of the GCA-EoS see [157]. Another similar approach gave rise to the cubic-plus-association (CPA) method [158], where the SRK EoS was combined with

Wertheim's association term (in the case of non-associating compounds, the original SRK description is recovered). Despite the extensive application that the CPA EoS has found in the study of the phase behaviour (VLE, LLE, SLE) of complex systems [159–162], it has not be reformulated within the scope of a group contribution approach.

Of significant importance and relevance to this thesis is the application of a group contribution formalism within the framework of the well known SAFT EoS [22–24]. We devote the next section to a more detailed discussion of these approaches.

2.4 GC approaches within SAFT

The statistical associating fluid theory (SAFT) stems from the first-order thermodynamic perturbation theory (TPT1) of Wertheim [101–104] for associating and chain fluids, and the later work of Gubbins and co-workers [21–24] who reformulated it in the form of an engineering equation of state. Two decades after the publication of the first papers on SAFT, a vast amount of research has been devoted to the development of numerous versions of the general method and applications. For a comprehensive review see the works of Müller and Gubbins [25, 163], Economou [26], and the more recent reviews by Tan *et al.* [164] and McCabe and Galindo [28].

SAFT provides a framework that allows for the calculation of the bulk properties of mixtures of associating chain-like molecules based on the knowledge of a monomeric reference fluid. Different approaches in the determination of the reference fluid have been presented that gave rise to different versions of SAFT, varying from semi-empirical approaches, such as the one employed within the PC-SAFT EoS [165, 166] and the SAFT of Huang and Radosz [167], to correlation based on molecular simulation data, as employed within the soft-SAFT EoS [168, 169]. Of particular interest are the SAFT approaches where the monomer reference fluid is approached in an entirely theoretical way, as for the development of the framework for interaction potentials of variable range (SAFT-VR [170, 171]).

In SAFT molecules are modelled as chains of bonded spherical segments interacting via an attractive potential (within SAFT-VR hard spheres with a van der Waals-type mean-field, square-well (SW), Yukawa (Y), Sutherland (S), Lennard-Jones (LJ), and Mie potentials have been considered), with short-range sites placed on the segments to mediate aggregate formation as occurs in hydrogen bonding or in polar fluids. This is depicted for the specific case of a square-well attractive potential in figure 2.1. The description of a substance within the SAFT formalism is undertaken by means of two parameters describing the size

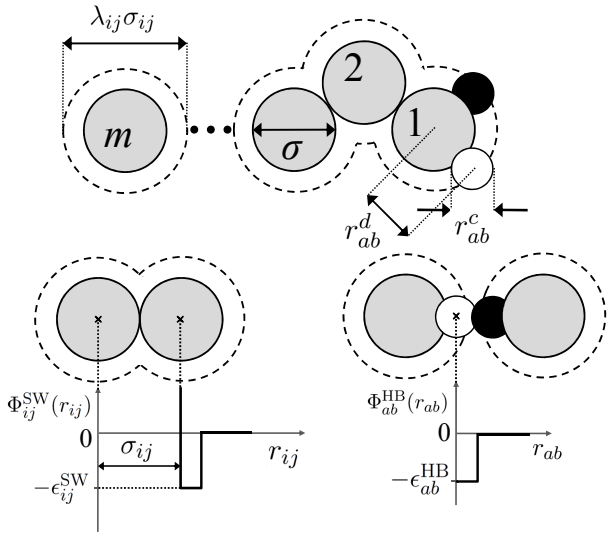

Figure 2.1: Schematic of the representation of molecules and molecular interactions within the framework of SAFT-VR with a square-well intermolecular potential. The parameters describing a molecule are the number of segments m, the segment diameter, σ_{ij}, and the energy (well-depth), $\epsilon^{\mathrm{SW}}_{ij}$, and range of dispersive interactions, λ_{ij}. Bonding sites are placed at a distance r^d_{ab} from the centre of the segment to mediate associating effects (with $\epsilon^{\mathrm{HB}}_{ab}$ and r^c_{ab} being the energy and range of association, respectively).

(usually the segment diameter σ_{ij} in hard potentials or the segment volume v^{00}_{ij} in soft potentials, and the chain length m_i), two further parameters accounting for the dispersion interactions (the potential depth ϵ_{ij} and range λ_{ij}, though in most implementations the potential range is fixed), and the number and type of association sites used to mediate hydrogen bond or aggregate formation. The sites are located at a distance r^d_{ab} from the centre of a segment, have an interaction cut-off range r^c_{ab}, and each $a-b$ site-site interaction is further characterised by an attractive energy parameter $\epsilon^{\mathrm{HB}}_{ab}$.

The SAFT equation of state is written in terms of the Helmholtz free energy of the system as a sum of different contributions

$$\frac{A}{Nk_{\mathrm{B}}T} = \frac{A^{\mathrm{ideal}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{mono.}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{chain}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{assoc.}}}{Nk_{\mathrm{B}}T} \ , \tag{2.36}$$

including an ideal term, A^{ideal}, and terms to account for the interactions between attractive segments (monomers) forming the molecules, $A^{\mathrm{mono.}}$, for the energy change due to the formation of chains of monomers, A^{chain}, and for the effect of association, A^{assoc}. Here, N is the total number of molecules in the mixture, T is the absolute temperature, and k_{B} is the Boltzmann constant. A pictorial representation of the different contributions to the free energy of a system accounted for within SAFT-type theories is shown in figure 2.2.(b).

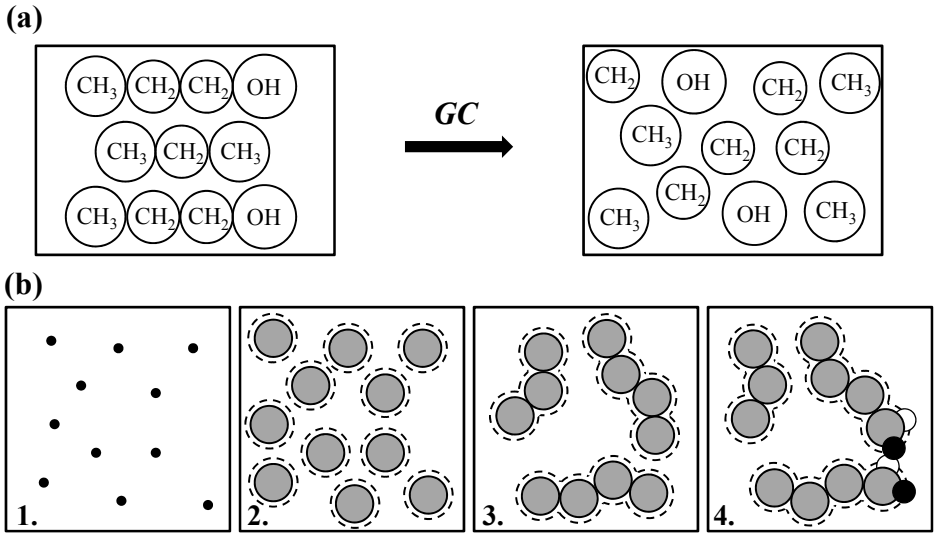

Figure 2.2: Comparison of the treatment of a binary system of n-propane+1-butanol at a composition of $x_{C_3H_8} = 0.333$: (a) within the scope of 1st order group contribution methods; and (b) within the framework of SAFT-type methods, where the contributions to the free energy of the system are represented (1: Ideal Gas, 2: Monomer Term, 3: Chain Term, 4: Association). Here, n-propane is modelled as a homonuclear chain of 3 segments, and 1-butanol as a homonuclear chain of 4 segments featuring 2 association sites.

Apart from the contributions shown in eq. (2.36), the general framework has been extended to include terms that explicitly account for other type of molecular interactions, such as dipole-dipole [172–174], quadrupole-quadrupole [175], intramolecular [176, 177] and ion-ion interactions [178–180]. Due to the explicit consideration of different molecular contributions SAFT approaches have been shown to provide an accurate representation of the fluid phase behaviour and thermodynamic properties of numerous pure compounds and mixtures, from small strongly associating molecules (such as hydrogen fluoride) to large polymers over wide ranges of conditions and types of systems. We refer to the reviews mentioned earlier for more details.

It is worth mentioning that the general family of SAFT-type approaches is particularly suited to the development of group contribution approaches. This is illustrated by means of a comparison between the treatment of a binary system within the general scope of group contribution methods and within the framework of SAFT-type approaches, as shown on figure 2.2. From the figure it can be seen that the decomposition of molecules into functional groups, as part of the solution of groups (or group contribution) treatment (figure 2.2.(a)), is very similar to the treatment of a system at the level of monomeric segments within SAFT (figure 2.2.(b)). This is an intrinsic feature of SAFT-type theories, which allows for such approaches to be reformulated within a group contribution formalism

in a very natural way.

2.4.1 Homonuclear approaches

In most implementations of the SAFT formalism molecules are regarded as composed of homonuclear segments, i.e., segments of the same size and characterised by identical energetic parameters (see figure 2.1). It is important to note, however, that the theory of Wertheim for the formation of chain-like fluids is not restricted to homonuclear models (indeed it is also not restricted to the use of spherical segments [181]). Further extensions of the fundamental molecular model proposed will be discussed later. Based on the original homonuclear model, one of the first attempts to combine a group contribution method with the SAFT methodology was presented by Lora *et al.* [182]. Using the parameters of low-molecular-weight propanoates presented by Huang and Radosz [167], which are homomorphs of the acrylate repeat group, Lora *et al.* presented calculations for the size and energy parameters of polyacrylates. Within this pseudo-group contribution methodology, the chain length (m) and segment volume (v^{00}) parameters for the polyalkylacrylates were calculated based on the parameters of the structural groups identified: CH_3, CH_2, CH, and the acrylate group ACgr. The method was not applied for the attractive energy of a segment v^0/k, which was instead calculated from monomer data or fitted directly to cloud curve data of the polymer solutions.

Following these ideas, Tobaly and co-workers [183, 184] and have presented a predictive GC implementation of the SAFT EoS. The authors presented two approaches for the predictive use of the equation of state for substances of the same chemical family, e.g., n-alkanes and branched alkanes. In the first approach, the parameters describing the members of a chemical family are estimated from experimental data (typically vapour pressure and saturated liquid density data for each compound are considered) and then relations are derived between the molecular parameters and molecular properties (e.g., molecular weight). In this case, each member of the n-alkane family is described by a different set of molecular parameters. This approach was based on previous findings [167, 185], where relations between the SAFT parameters and the molecular weight of the n-alkanes were reported. Paricaud *et al.* [186] presented how such a methodology can be applied to characterise linear polyethylene and showed that satisfactory results are obtained for the study of solubility of light gases in polymers. In the second approach, group parameters are determined under the assumption that molecules consist of identical segments. In the case of the n-alkanes, this means that the CH_3 and CH_2 chemical segments are described by the same group parameters, so that differences between members of the same chemical

family depend simply on the assignment of different values of the molecular weight and hence the chain length, m. The approach has been used to predict the phase behaviour of pure alkanes (linear and moderately branched), alkenes and ring compounds and mixtures of these, yielding reasonable results; the reported accuracy in the prediction of the vapour pressure of the series from n-C_9H_{20} to n-$C_{38}H_{78}$ is 13.61%. The same approach was later applied to the chemical family of alcohols [184]. A homonuclear model is used as before, and the group parameters are transferred from the study of the n-alkanes (i.e., those for the CH_3 and CH_2 groups are used to model the back-bone of the molecule), and association sites are included to model the hydroxyl hydrogen bond interaction. The parameters required to characterise the association sites are determined by fitting to experimental data; these are then attributed to the OH group alone. The predictive capability of this approach is rather limited, since it is not fully formulated within the scope of contributions of functional groups, being more of an effort to model the behaviour of substances within a certain homologous series by transferring parameters as applied in a similar manner in recent works [187, 188].

A more sophisticated group contribution approach was later introduced by the same group [18]. In this case, the molecules are decomposed into functional groups and distinct group contributions are estimated. The authors implemented a group contribution method within two different versions of SAFT: the original one presented by Chapman *et al.* [24], referred to as SAFT0, and the SAFT-VR approach of Gil-Villegas *et al.* [170, 171] (referred to as GC-SAFT-VR). While the concept of groups forming the molecules is introduced, so that group parameter tables are explicitly presented in the work of Tamouza *et al.* [18], the underlying molecular model was still homonuclear. Group parameters were combined to lead to a unique set of parameters characterising identical segments in the homonuclear model of the molecule of interest. Systems including n-alkanes, α-olefins, 1-alkanols, alkyl-benzenes and alkyl-cyclohexanes are studied and the predictive capability of the theory is tested by extrapolating to long-chain molecules of high molecular weight (e.g., n-$C_{32}H_{66}$) and later by studying mixtures of nonassociating and associating compounds [189]. In later work, the group contribution scheme was applied to another SAFT variant, the PC-SAFT EoS [165, 166]; the resulting methodology is commonly referred to as GC-PC-SAFT. The parameters for polar groups have since been revised using the empirical polar term of Gubbins and Twu [190] combined with the segment approach of Jog and co-workers [172, 173] and with experimental data of polar mixtures [191, 192]. The quality of the predictions with the resulting GC-PPC-SAFT compared to the previous approach was seen to be significantly improved. New parameters have also been obtained for alkyl-esters (R-COO-R') and formates (HCOO-R), with HCOO modelled as a separate group. The fluid phase

behaviour of mixtures of esters with n-alkanes, cyclohexanes, alkylbenzenes, and xylene has been studied in detail and the performance of the method was found to lie within an accuracy of a few percent in a fully predictive manner (the binary interaction parameters were calculated using standard combining rules). In a later study the predictive capability of the polar GC-SAFT was tested in mixtures of alkanes with methyl benzoate [193]. Both methods studied (i.e., the GC reformulation of SAFT-VR and the GC-PC-SAFT) were found to give satisfactory phase behaviour predictions (%AAD < 10 for the bubble-point pressure).

As a test of the predictive capability of GC-SAFT approaches, mixture properties are usually examined using the Lorenz-Berthelot combining rules without correcting parameters. However, the geometric mean of the intermolecular energy parameters is theoretically sound only in the case of components of identical size (diameter) and ionisation potential; as discussed in [112], high deviations from the Berthelot combining rule can be expected otherwise. In this spirit, Thi *et al.* [194] have presented a study of binary mixtures of n-alkanes with H_2 and CO_2 with GC-SAFT where the binary interaction parameters were also calculated within a group contribution formalism. Molecule-group parameters were calculated by regression to experimental mixture VLE data and were used to "build-up" the molecule-molecule interaction parameter. Following the ideas presented by Haslam *et al.* [112], Nguyen-Huynh *et al.* [195] also followed the London theory to develop an approach to predict the binary interaction parameters based on the pseudo-ionisation energy of the functional groups of the molecules present in the mixture. Initially the optimal unlike binary interaction parameter was obtained by regression to mixture data and the pseudo-ionization energy was then calculated for each functional group in the molecules. The group parameters obtained could then be transferred to the prediction of binary interaction parameters of other mixtures where the same functional groups appear. This method has been successfully tested for mixtures of CO_2+n-alkane, CH_4+n-alkane, C_2H_6+n-alkane [195], and in mixtures of light gases with hydrocarbons [196, 197], where VLE and LLE regions were satisfactorily predicted. Later work included the application of the method including the explicit treatment of polar groups (GC-PPC-SAFT) to study aqueous solutions of hydrocarbons [198], binary systems of methanol+n-alkanes [199], and, finally, systems of amines [200].

2.4.1.1 GC approaches in PC-SAFT

The formulation of the perturbed-chain SAFT approach (PC-SAFT) [165, 166] is slightly different to other versions of SAFT in that the free energy of a homonuclear hard chain

fluid (as opposed to that of the monomer fluid) is first considered, and the dispersion interactions are then obtained in a perturbation expansion with a hard-chain fluid as a reference. As such it is, in principle, not the best suited to be reformulated as a GC approach, but relations between the molecular model parameters and molecular properties (such as the molecular weight) can be derived so as to identify the contribution that different chemical groups make to the properties of molecules. For example, the intermolecular model parameters for a series of hydrofluoroethers, where experimental data (saturated liquid densities and vapour pressures) are available, have been determined and then used to calculate the contribution of each functional group (CH_3, CH_2, CF_3, CF_2, and O groups were considered) [201]. In a later study the chemical family of esters including dipole-dipole interactions using the perturbation expansion of Gubbins and co-workers [190, 202] was considered [203]; the fluid phase behaviour of compounds not included in the regression database were found to compare well with other GC EoS methods, although for high molecular weight esters, the deviations from the experimental data are rather significant (e.g., 50% standard deviation was found for $C_{23}COO$). The same method, with a term representing the quadrupole-quadrupole contribution, has also been used for the prediction of equilibrium properties of polycyclic aromatic hydrocarbons and their mixtures [204]. The vapour pressures of 19 chemical families including hydrocarbons, cyclic and aromatic hydrocarbons, alcohols, amines, nitriles, esters, ketones, and ethers, amongst others, have also been considered [205]. The PC-SAFT EoS has been also combined with the Constantinou and Gani [38] GC approach, that incorporates first- and second-order groups, in an attempt to distinguish between isomers and account for proximity effects [206–208].

At this point it is worth pointing out again that all approaches presented in section 2.4.1 thus far are based on a homonuclear model. Different groups are identified and characterised, each with different intermolecular parameters if necessary, but in treating multifunctional molecules, an average value for each parameter is obtained using the necessary groups; the average parameters are then used to characterise all of the identical segments forming the homonuclear molecule. The concept of group contribution is not applied on the level of the underlying theory, rather more as a predictive approach to determine the model parameters for different molecules. A different treatment is possible, where molecules are modelled comprising segments of different size and/or energy parameters, and each segment (or group of identical segments) represents a functional group. These so-called heteronuclear approaches are discussed in the following section.

2.4.2 Tangential heteronuclear segment models in SAFT

Following the original TPT1 approach of Wertheim heteronuclear models can be proposed, where the segments in a given model molecule are arbitrarily different (see figure 2.3), although it should be noted that the theory is strictly only appropriate for chains of tangentially bonded segments. The first application of the theory to deal with molecules formed from tangentially bonded heteronuclear spherical segments was a study of heteronuclear diatomic molecules [209] later extended to heteronuclear triatomics [210] and arbitrary polyatomic molecules [211, 212].

Adidharma and Radosz [213, 214] developed a version of the SAFT-1 theory where heterosegmented polymers are examined by means of a square-well potential describing the segment-segment interactions; reasonable predictions of the phase equilibria of small and large heterosegmented molecules that present no associating effects are obtained with this approach. McCabe *et al.* [215] generalised the SAFT-VR [170, 171] equation of state to treat heteronuclear molecules, where the chain term of the original SAFT-VR equation was replaced by a bond-term dealing with the formation of molecules from segments that are different in size and/or energy. The theory was validated by comparison with Monte Carlo simulation data for isotherms of heteronuclear diatomic models. The same theory [hetero-SAFT(VR)] was used to study symmetric and asymmetric diblock chains [216] and semifluorinated alkanes modelled as diblock chains [217].

The original SAFT expressions of Huang and Radosz [167] was also extended to model heteronuclear molecules with soft intermolecular potentials [218, 219]. This approach was applied to branched and linear copolymers, and was shown to account successfully for the effect of short-chain branching on the cloud-point pressures of certain mixtures. One should also point out that Blas and Vega [169] studied the performance of a heteronuclear

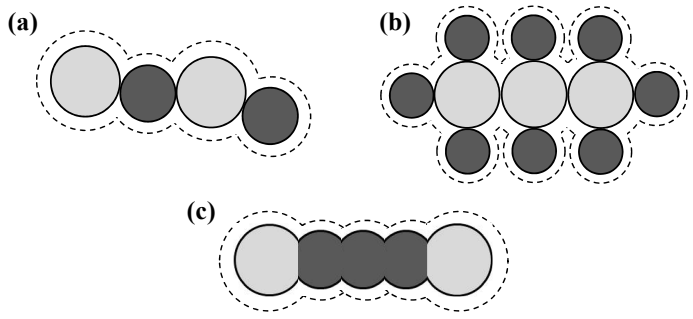

Figure 2.3: Heteronuclear models for SAFT-VR: (a) United-atom tangent model; (b) all-atom tangent model; and (c) united-atom *fused* model

soft (LJ)-SAFT in comparison with molecular simulation data for chains of varying length, and Gross *et al.* [220] have presented a heteronuclear version of PC-SAFT [165, 166] applied to co-polymers consisting of polar and non-polar units, where the bonding ratio between the different segments rather than the sequence of the segments in the chain is defined.

Based on the agreement of the heteronuclear theory with molecular simulation calculations, the performance of the heteronuclear SAFT-VR as a framework for a group contribution approach was examined. Lymperiadis *et al.* [19] showed that the tangent heteronuclear model fails to accurately describe the properties of real compounds, based on the results presented for the *n*-alkanes, where both an all-atom (figure 2.3.(a)) and a united-atom (figure 2.3.(b)) model were examined. It was suggested that the application of a *fused* heteronuclear molecular model (figure 2.3.(b)) is more appropriate. The application of such a molecular model gave rise to the development of three different heteronuclear GC approaches within the framework of SAFT, which are presented in detail in the following section.

2.4.3 Fused heteronuclear segment models

The SAFT-γ EoS [19, 29] was the first approach that addressed difficulties associated with tangential models in SAFT. The SAFT-γ approach [19, 29] is based on a generalisation of the SAFT-VR EoS [170, 171] to model heteronuclear chain molecules composed of fused segments. As in all other versions of SAFT, the contributions to the free energy due to the association interaction and irreversible formation of bonds are described following the first-order thermodynamic perturbation theory of Wertheim [101–104]. In the first paper [19] a detailed account of the inconsistencies arising from a representation of molecules formed from tangent segments, both in the united-atom (a group of atoms is represented by one segment - see figure 2.3.(a)) and in the all-atom (each segment represents a single atom - see figure 2.3.(b)) representations, is presented. It turns out that although the Wertheim TPT1 description of thermodynamic properties of molecules with tangentially bonded segments is very accurate (i.e., in comparison with molecular simulation data), a tangent model is not generally appropriate to model the properties of real molecules.

The introduction of the so-called shape factor, S_k, reflecting the contribution of each group k to the "overall" molecular properties, allows a for a description of molecules comprising *fused* segments (cf. figure 2.3.(c)); the different segments are used in the traditional group contribution manner to represent the different chemical groups. The shape factor

parameter S_k describes the contribution that a given segment k of diameter σ_k makes to the overall molecular geometry through the mean radius of curvature R_m, the surface area S_m and the volume V_m of the molecule. The physical significance of the shape factor has been discussed in detail by Lymperiadis *et al.* [29]. Using ethane as an example, it is shown that the molecular model proposed would correspond to a value of the bond length between 1.2 Å and 2.3 Å (compared to the experimental bond length of 1.54 Å). The exact value is very sensitive to the precise definition of the molecular volume, but the fact that the calculated and experimental values lie in the same physical range is encouraging.

Within the SAFT-γ treatment of molecules formed from fused square-well segments the thermodynamic properties attributable to a given group k are fully described by the number of identical segments that the group comprises, ν_k^*, a shape factor S_k, a segment diameter σ_{kk}, a dispersive energy ϵ_{kk}, and range λ_{kk}. In the case of associating groups, the number of associating site types of each group, N_{ST_k}, and the number of sites of each type, $n_{k,a}$, are usually determined based on the chemical structure of the group; two additional parameters are introduced for each $a-b$ site-site interaction, namely the energy $\epsilon^{\mathrm{HB}}_{kkab}$ and range r^{c}_{kkab}. For the description of the interactions between segments of different type, ϵ_{kl}, $\epsilon^{\mathrm{HB}}_{klab}$, r^{c}_{klab}, and in some cases λ_{kl}, need to be determined. In this approach, group parameters are determined by studying the smaller members of a given homologous series (n-alkanes, 1-alcohols, etc.) using experimental vapour pressure and saturated liquid density data usually up to the 10th member of the series. The SAFT-γ approach has been shown to provide a very accurate description of the fluid phase behaviour for chemical families containing the CH_3, CH_2, CH_3CH, aCH, $aCCH_2$, CH_2=, CH=, OH, COOH, and C=O groups (where aC denotes aromatic carbons). The overall %AAD for the chemical families considered so far is 3.34% for the vapour pressure and 0.92% for the saturated liquid density [19, 29].

The predictive capability of the method has been tested by assessing the description of the fluid phase behaviour of large molecular weight compounds that were not included in the determination of the group parameters. The result of this assessment is given in table 2.2. In this case the overall %AAD was found to be 8.18% for the pressure and 0.75% for the density, which is even lower than that for the fitted compounds. A key additional advantage of the SAFT-γ method is that mixtures can be treated in a fully predictive manner without the need to propose combining rules and adjustable parameters for the unlike interactions. The unlike interaction parameters necessary for the modelling of mixtures can, in some cases, be obtained from pure component data alone. This is a unique feature of group contribution approaches that employ a heteronuclear molecular model.

Table 2.2: Average absolute deviations (%AAD) for vapour pressures p_{vap} and saturated liquid densities ρ_{sat} of the SAFT-γ predictions compared to experimental data (where n is the number of data points) for the components not included in the parameter estimation database.

Compound	T range [K]	n	%AAD p_{vap}	T range [K]	n	%AAD ρ_{sat}
n-pentadecane	283-633	76	8.26	283-633	76	0.52
n-octadecane	301-671	75	12.48	301-671	75	2.07
n-tetracosane	293-723	87	19.51	293-723	87	0.42
2-methyltetradecane	403-537	9	3.53	283-598	11	0.36
2-methylhexadecane	428-568	9	4.84	293-643	11	0.38
2-methyloctadecane	451-595	9	8.12	293-653	10	0.39
n-dodecylbenzene	599-663	20	7.97	293-698	10	1.35
n-tridecylbenzene	343-463	13	3.74	-	-	-
n-tetradecylbenzene	298-627	3	15.31	283-375	10	1.05
1-dodecene	396-487	19	4.34	263-578	11	0.62
1-tetradecene	431-524	10	5.24	253-613	10	1.08
tridecan-2-one	424-546	15	2.94	-	-	-
dodecanoic acid	403-572	11	7.35	333-573	13	0.97
tetradecanoic acid	385-465	17	12.08	333-573	13	0.77
hexadecanoic acid	440-577	9	18.64	353-573	12	0.62
dodecan-1-amine	354-532	11	5.91	313-573	14	0.67
tetradecan-1-amine	382-564	11	4.14	-	-	-
hexadecan-1-amine	405-596	11	5.75	333-573	13	0.68
1-dodecanol	425-549	24	5.03	308-549	39	0.65
1-tetradecanol	424-569	12	6.84	313-573	14	0.94
1-octadecanol	435-518	27	15.87	353-573	12	0.74
Average %AAD	-	-	8.18	-	-	0.75

In this context the method has been shown to provide a good description of the fluid phase behaviour of mixtures including binary mixtures of n-alkanes, alkenes, alkanols, n-alkane+acids, n-alkanes+amines and even cases with highly non-ideal behaviour [19, 29], including liquid-liquid equilibrium (LLE) and polymer-solvent systems.

Another GC approach based on the same general theory (SAFT-VR) and the same *fused* heteronuclear molecular model as the SAFT-γ EoS, the hetero GC-SAFT-VR method has also been presented [20]. The hetero GC-SAFT-VR method was formulated based on previous work on the development of a heteronuclear EoS that employs the square-well intermolecular potential [215, 216]. Instead of the shape factor, a chain length m_k is used,

which, in common with the shape factor, is group specific. This chain length can take values less than 1 (as opposed to the chain length m in homonuclear approaches which for values of less than 1 is of ambiguous physical meaning). It is possible to relate SAFT-γ and hetero GC-SAFT-VR parameters to highlight the similarities between the two theories. First, note that

$$\nu_k^* S_k = m_k \ , \tag{2.37}$$

where ν_k^* is the number of identical segments a groups of type k comprises (as in [29]) and m_k is the chain length (as defined in [20]). Second, one can relate the definition of the fraction of segments of type k in the mixture, $x_{s,k}$, used in SAFT-γ and the fraction of segments of type k on a molecule of type i, $x_{s,ki}$, employed within the hetero GC-SAFT-VR method

$$x_{s,k} = \sum_{i=1}^{N_\mathrm{C}} x_{s,ki} \ , \tag{2.38}$$

where NC again denotes the number of components in the mixture. It then becomes clear [28] that the two theories are identical in what relates to the treatment of the ideal and the monomer terms. The key difference between the two approaches is in the treatment of the term related to the formation of molecules from the monomeric segments. In SAFT-γ the formal TPT1 expression is maintained using effective parameters, so that the contribution to the free energy due to molecule (chain) formation is a function of the number of segments that appear in the chain and the contact radial distribution function of an effective chain fluid in the following way:

$$\frac{A^\mathrm{chain}}{N k_\mathrm{B} T} = -\sum_{i=1}^{N_\mathrm{C}} x_i \left(\sum_{k=1}^{N_\mathrm{G}} \nu_{k,i} \nu_k^* S_k - 1 \right) \ln g_{ii}(\bar{\sigma}_{ii}) \ . \tag{2.39}$$

In eq. (2.39), x_i is the mole fraction of component i, $\nu_{k,i}$ the number of groups of type k in component i, and $g_{ii}(\bar{\sigma}_{ii})$ the value of the segment-segment radial distribution function at contact $\bar{\sigma}_{ii}$ of an effective chain fluid of component i [19]. In the case of the hetero GC-SAFT-VR approach, the heterogeneity of the segments is maintained at the level of the chain molecule formation, at the cost of proposing an empirical relation that involves contributions of the segments as well as those of the contacts:

$$\begin{aligned} \frac{A^\mathrm{chain}}{N k_\mathrm{B} T} = \ & - \sum_{i=1}^{N_\mathrm{C}} x_i \left(\sum_{k=1}^{N_\mathrm{G}} \left[\left(\nu_{k,i} \left(m_k - 1 \right) + \frac{1}{2} b_{ik,ik} \right) \ln g_{ik,ik}(\sigma_{ik,ik}) \right. \right. \\ & + \left. \left. \frac{1}{2} \sum_{l=1}^{N_\mathrm{G}} b_{ik,il} \ln g_{ik,il}(\sigma_{ik,il}) \right] \right) \ , \end{aligned} \tag{2.40}$$

where m_k is the equivalent to the shape factor in the SAFT-γ approach (cf. eq. (2.37)), $g_{ik,il}(\sigma_{ik,il})$ is the value of the radial distribution function at contact $\sigma_{ik,il}$ and $b_{ik,il}$ the number of bonds between segments of type k and segments of type l on component i [20]. The corresponding effective many-body correlation function of a molecule of type i in the hetero GC-SAFT-VR approach is approximated with the following empirical relation:

$$g_{ii}(1,\ldots,\mathrm{NG}) = \prod_{k=1}^{N_\mathrm{G}} [g_{ik,ik}(\sigma_{ik,ik})]^{[\nu_{k,i}(m_k-1)+\frac{1}{2}b_{ik,ik}]} \prod_{l=1}^{N_\mathrm{G}} [g_{ik,il}(\sigma_{ik,il})]^{\frac{1}{2}b_{ik,il}} \quad . \tag{2.41}$$

This empirical relation allows one to distinguish between isomers within hetero GC-SAFT-VR, but at the same time appears to lead to a deterioration in the description of the pure component phase behaviour of real compounds, at least in comparison with the SAFT-γ methodology. As an example, the reported average %AAD for the linear n-alkanes (from n-butane to n-decane) with hetero GC-SAFT-VR is 5.95% for the vapour pressure, compared to 3.99% for SAFT-γ, and 3.07% and 0.51%, respectively, for the saturated liquid density [19, 20]. The hetero GC-SAFT-VR approach was initially applied to the study of the phase behaviour of non-associating groups, where the chemical families of the n-alkanes, unsaturated and branched alkanes, alkylbenzenes and ethers. The average deviations for all chemical families and compounds considered were 5.46% for the vapour pressure and 2.81% for the saturated liquid density. In later work, the GC-SAFT-VR method was extended to the study of the phase behaviour in polymer+solvent systems [221] and associating compounds [222, 223].

Recently, the PC-SAFT EoS [165, 166] was reformulated within a GC formalism based on a heteronuclear molecular model [224]. This heterosegmented PC-SAFT was developed following the concept of the hetero GC-SAFT-VR EoS, so that the connectivity of the different segments of the molecule is explicitly accounted for in the expressions that determine the contribution to the free energy due to chain formation. The method was shown to accurately describe the phase behaviour of the chemical families of n-alkanes, branched and cyclic alkanes, with average deviations of 1.50% for the vapour pressure and 0.72% for the saturated liquid density. The predictive capability of the method was examined in the prediction of the fluid phase behaviour of compounds not included in the regression of the group parameters, where good agreement with the experimental data was shown (%AADs of 12% for the vapour pressure and 1.1% for the saturated liquid density).

These are promising approaches developed on a strong theoretical basis around a detailed molecular model without the dependency of an underlying lattice. In this way they possess the advantages inherent in the successful UNIFAC approach but overcome the difficulties

associated with the underlying lattice model of UNIFAC. These SAFT GC approaches are accurate over large pressure ranges and can be used reliably for liquid and vapour phases. Their formulation as continuum fluid theories based on a heteronuclear molecular model allows for the binary interaction parameters to be determined from pure component data alone. Finally, as the application of the GC formulation is applied to an equation of state, the presented resulting methodologies can be used for the calculation of all bulk thermodynamic properties without any limitation, as is the case for activity coefficient GC approaches.

2.5 Other Predictive Methods

Although these are not strictly GC methods, predictive approaches based on quantum mechanical techniques such as COSMO-RS [225–227] are also becoming increasingly popular. The COSMO-RS model combines methods from quantum chemistry with statistical mechanics so that one is able to predict *a priori* the thermophysical properties of a substance or a mixture based solely on its atomic structure, without the need for experimental data. The methodology is based on the COSMO approach [228] which involves a quantum mechanical calculation used to obtain the energy, geometry, and the screening charge density on the surface of a solute assuming it is found in an ideal conductor. It is important to note that this computationally expensive calculation has to be done only once for a given substance; after each calculation the screening charge profile can be stored in a database and directly accessed from there. This hypothetical state of a molecule in a virtual conductor is used as a reference point within the COSMO-RS approach, in which the interactions in the fluid are then described as pair interactions between "surface segments". The chemical potential of a solute in a real solvent is then calculated from the chemical potential of surface segments, based on statistical thermodynamics. In this context the COSMO-RS model has been used as a g^E model (calculating the activity coefficient of the liquid phase) for the prediction of fluid phase equilibria in various, including highly nonideal, mixtures [229]. A modification of the original theory in the spirit of a group contribution approach has been presented as the COSMO-segment activity coefficient (SAC) method [230]. Here, molecules are treated as consisting of equally sized surface segments and the molecular activity coefficients are obtained by summing the contributions of each segment-based activity coefficient. The accuracy of prediction of this method has been demonstrated in calculations of vapour-liquid equilibria and infinite dilution activity coefficients.

Another important predictive method that has been employed primarily for solubility

calculations is the non-random two-liquid segment activity coefficient model (NRTL-SAC) [231]. Within the framework of NRTL-SAC molecules are decomposed in segments mapped into three categories based on surface interactions: hydrophobic segments, hydrophilic segments, and polar (polarity and solvation) segments. The parameters that describe the interaction between the different increments were determined based on a vast amount of VLE and LLE experimental data for binary solvent+solute mixtures. Having determined the interaction parameters, a molecule is fully characterised by determining the measure of each increment, i.e., the segment numbers of hydrophobicity, hyrdophilicity, polarity, and solvation. This requires a substantial amount of experimental data, typically solubilities, in at least four different solvents, each with a chemical structure representative of one type of increment (e.g., hexane for hydrophobicity, acetonitrile for hydrophilicity, and water for polarity). The extrapolation to other solvents, previously characterised, is straightforward, as a single set of interaction values is used. This method is used routinely in the pharmaceutical industry in conjunction with experimental data for solvent screening. Both the COSMO-RS and NRTL-SAC methods have been used in an effort to predict the solubility of some commonly used substances in the pharmaceutical industry [232]. The approaches were found to give reasonable predictions, with NRTL-SAC providing predictions of higher accuracy. However, NRTL-SAC is more of a correlative approach, as the characterisation of new solvents/solutes requires a substantial amount of experimental data. On the other hand, the COSMO-RS method, based solely on *ab initio* calculations, is fully predictive, but is generally reported to be less accurate than other techniques, such as the UNIFAC approach [233].

The force fields developed within an all-atom or a united-atom group framework for use in molecular simulation constitute a final class of predictive methodologies. A force field typically contains information about the energy and size parameters of structural units of molecules that can be employed within molecular simulation techniques (Monte Carlo or molecular dynamics) for property prediction. Given the varying level of detail and computational cost of molecular simulation, several force fields have been developed for different applications. The most widely applied include the all-atom OPLS [234] and AMBER [235] force-fields, the united-atom TraPPE [236], the work of Klein and co-workers [237], and the recent work of Potoff and Bernard-Brunel [238], where a similar grouping as the one in GC-EoS models is typically applied, and the more coarse-grained MARTINI force field [239].

The development of the aforementioned force-fields is typically based on brute-force fitting, so that the extension of the parameters to the study of new chemical families is rather cumbersome and computationally expensive. In an attempt to overcome this shortcoming,

and of great relevance to the work presented in this thesis, the development of force-fields in conjunction with analytical equations of state has been presented [240–242]. The key idea is that assuming a molecular model (e.g., united-atom or coarse-grained representation), an analytical EoS is employed for the regression to experimental data and the determined parameters are subsequently used in molecular simulation techniques. In the work of van Westen *et al.* [240], the PC-SAFT EoS [165, 166] is employed for the regression of the parameters to experimental fluid phase behaviour data. The force-field parameters are developed assuming a Lennard-Jonesium interaction potential and a united-atom representation, as in the TraPPE force-field. The presented parameters are shown to provide better agreement with the experimental values compared to the well-established TraPPE force-field for the vapour density and heat of vapourisation. However, as the interaction potential employed within the theory (which was parameterised based on experimental data for the *n*-alkanes) does not exactly correspond to the LJ potential used in the simulations, an iterative procedure has to be followed until convergence of the parameters is achieved [240]. In the same spirit, Avendaño *et al.* [241] employed the SAFT-VR Mie EoS for the parameterisation of the force-field based on experimental data. The great advantage of this approach is that given the direct link between theory and simulation, the parameters obtained from the analytical EoS can be employed within molecular simulation techniques with no further adjustment. Furthermore, as within the SAFT-VR Mie the parameters define the functional form of the interaction potential of variable attractive and repulsive range (as opposed to determining the parameters for a potential of fixed form), the presented approach can be employed for the development of coarse-grained models. Models of this type are known to require flexible interaction potentials due to the different grouping of atoms and are particularly attractive from the aspect of computational efficiency [243]. The SAFT-γ Mie EoS (presented in detail in chapter 4) has also been employed for the development of heterosegmented simulation models, as presented by Lafitte *et al.* [242].

2.6 Conluding Remarks

Group contribution approaches offer great possibilities for the prediction of fluid properties of pure components and mixtures from a knowledge of the parameters characterising a comparatively small number of functional groups. Since the early studies, mostly involving correlations of pure component properties [32], activity coefficient models, UNIFAC [14] in particular, have been seen as the state-of-the-art GC tools in the prediction of fluid phase equilibria and solubility. More recently, however, there has been an increasing interest in reformulating molecular-based equations of state such as SAFT within a GC formalism.

The advantage of this is that EoSs are based on continuum fluid models treating the liquid and gas phases on an equal footing. Within such approaches one can also naturally incorporate detailed molecular interactions such as hydrogen bonding, and the approaches can be easily extended to the study of more complex systems, such as electrolyte solutions and systems of polymers. In contrast to most other theories, the recently developed heteronuclear group contribution approaches make it possible for the unlike mixture parameters to be obtained using pure compound data alone, a clear advantage in cases where there is limited availability of experimental data. In this context the combination of GC models and quantum mechanical calculations, such as in the COSMO-RS [225] and COSMO-SAC [230], as well as the development of force-fields in conjuction with equations of state [240–242], are also of much current interest.

SAFT-based GC methods are particularly well suited for the study of the complex phase behaviour that highly non-ideal systems exhibit. As an example can be seen the family of aqueous solutions of hydrocarbons; systems of these kind are of relevance to a wide variety of industrial applications, from reservoir processing to waste water treatment, and are at the same time considered highly challenging from a theoretical perspective. The highly non-ideal phase behaviour of systems of this kind includes heterogeneous azeotropes and liquid-liquid immiscibility, where the compositions in the two phases differ by several orders of magnitude. The successful modelling of systems of this kind requires models that explicitly account for the associating effects that water exhibits, which dominate the behaviour of aqueous solutions.

Chapter 3

Modelling aqueous solutions with the SAFT-γ group contribution approach

In the previous chapter a comprehensive review of predictive group contribution based approaches and thermodynamic models was presented. Particular focus was given to group contribution approaches that have been combined with SAFT-type methods, with the view of developing an accurate predictive methodology for the prediction of thermodynamic properties of highyl non-ideal systems. In this chapter the application of one of the these methodologies, the SAFT-γ group contribution approach [19, 29], to the modelling of aqueous solutions of hydrocarbons is presented.

Aqueous solutions of hydrocarbons are of great interest in many applications ranging from the petrochemical industry to biological processes and waste water treatment. The thermodynamic modelling of systems of this type is particularly challenging due to the highly non-ideal behaviour that they exhibit over a wider range of thermodynamic conditions. Characteristics of the fluid phase equilibrium of these systems include heterogeneous azeotropes bounding regions of vapour-liquid and liquid-liquid equilibria, where the respective solubilities in the two liquid phases can differ by many orders of magnitude. The extreme nature of the fluid phase behaviour exhibited by these systems is a direct consequence of the dominant hydrogen bonding interactions between the water molecules. It is therefore clear that the key to a successful description of aqueous solutions requires the application of a thermodynamic treatment that accounts explicitly for the effects of association. Examples of such approaches include the statistical association fluid theory (SAFT) [23, 24] and the cubic plus association (CPA) [158] equations of state (EoSs),

both of which have been used to describe the fluid phase equilibria in water+hydrocarbon mixtures [244–256]. For a more complete review the reader is directed to the book of Kontogeorgis and Folas [162].

The fluid phase equilibria of aqueous solutions of hydrocarbons has been studied with various group contribution approaches. The performance of the group contribution-associating-equation of state (GCA-EoS) [156] in the prediction of the phase behaviour of these systems has been presented in [257, 258]. The method was shown to provide an accurate description of the fluid phase behaviour of water+hydrocarbon mixtures, however, based on the GC formulation of the NRTL expressions for the dispersive interactions, a considerable amount of group interaction parameters is necessary. The determination of the group interaction parameters requires a substantial amount of experimental data for binary mixtures. Group contribution methods formulated within the framework of SAFT have not yet been assessed in the description of the fluid phase behaviour of water+hydrocarbon mixtures. A predictive study of this type presented in the literature is based on the sPC-SAFT [259] approach where generalised parameters were used to represent the family of 1-alkanols, and the performance of the method in predicting the phase behaviour of aqueous solutions of 1-alkanols was discussed [260]. The method was shown to provide a satisfactory description of the phase behaviour (VLE, LLE and SLE) of selected water+1-alkanol mixtures, however, the binary interaction parameters between the components of the mixture were obtained on a molecular basis, which limits the predictive capability of the method.

Here, the performance of the SAFT-γ group contribution approach in the prediction of the fluid phase behaviour of aqueous solutions including hydrocarbons is assessed. The ability of the method to describe accurately the extreme phase behaviour that these systems exhibit over a wide range of thermodynamic conditions is examined in detail. The predictive capability of the method lies mainly in the fact that the predictions of the thermodynamic properties of mixtures are based on pure component experimental data alone. Although such an approach is possible for a broad class of systems, it is not appropriate in the case of molecules that are described as single functional groups. Since the most appropriate model for water is as a single functional group, the extension of the method to aqueous solutions requires the determination of unlike group parameters between the water molecule and the other groups making up the solute molecule. The group contribution concept allows for the estimation of the unlike group parameters to be undertaken based on a minimal set of experimental data; a single set of transferable interaction parameters can then be used in the modelling of a wide range of systems.

3.1 The SAFT-γ GC Approach

The SAFT-γ method [19, 29] was developed as a generalisation of the SAFT-VR EoS [170, 171] to treat molecules formed of *fused* heteronuclear segments. The *fused* heteronuclear molecular model was found to provide the best description of the macroscopic properties of real compounds, while tangent molecular models (united-atom and all-atom) were found to be unsuitable for the task. Although in the current version of the theory the segment interactions are of the square-well form, the SAFT-VR treatment is general and can easily be extended to treat Lennard-Jones [261], Yukawa [262] or Mie [30, 263] segments. A square-well functional group k and its self-interaction is fully described by the number of identical segments in the group ν_k^*, and the segment parameters such as its shape factor S_k, the segment diameter σ_{kk}, the dispersive energy ϵ_{kk} and range λ_{kk}. In the case of associating groups, additional parameters are introduced, namely the number of types of associating sites of a group k, N_{ST_k}, the number of sites of each type, $n_{k,a}$, and the energy $\epsilon^{\mathrm{HB}}_{kkab}$ and range r^{c}_{kkab} of the interaction for each pair of site types; in this particular case we assume that sites are positioned half-way between the centre of the segment and the surface [21]. The unlike group parameters σ_{kl} and λ_{kl} can be calculated by means of size-related combining rules [$\sigma_{kl} = (\sigma_{kk}+\sigma_{ll})/2$ and $\lambda_{kl} = (\lambda_{kk}\sigma_{kk}+\lambda_{ll}\sigma_{ll})/(\sigma_{kk}+\sigma_{ll})$], whereas for the dispersive and associative unlike group interactions the additional parameters (ϵ_{kl}, $\epsilon^{\mathrm{HB}}_{klab}$ and r^{c}_{klab}) are typically determined from experimental data. Within the SAFT-γ formalism, the total Helmholtz free energy of the system is represented as a sum of different contributions, as follows:

$$\frac{A}{Nk_{\mathrm{B}}T} = \frac{A^{\mathrm{ideal}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{mono.}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{chain}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{assoc.}}}{Nk_{\mathrm{B}}T} \ , \tag{3.1}$$

where N is the total number of particles of the system, T the absolute temperature and k_{B} the Boltzmann constant. The contributions accounted for are the ideal Helmholtz free energy of the molecules, A^{ideal}, and the Helmholtz free energy contributions resulting from the interactions between monomeric functional groups, $A^{\mathrm{mono.}}$, the formation of molecules from the functional groups, A^{chain}, and the effect of association, $A^{\mathrm{assoc.}}$.

The contribution of the ideal Helmholtz free energy is calculated as [264]

$$\frac{A^{\mathrm{ideal}}}{Nk_{\mathrm{B}}T} = \left(\sum_{i=1}^{N_{\mathrm{C}}} x_i \ln\left(\rho_i \Lambda_i^3\right) \right) - 1 \ , \tag{3.2}$$

where x_i is the mole fraction, ρ_i the number density ($\rho_i = N_i/V$), N_i the number of molecules of component i, and V the volume of the system. Λ_i denotes the de Broglie wavelength of each molecule of component i, which incorporates the translational and

rotational kinetic contributions to the ideal Helmholtz free energy and does not need to be defined precisely in studies of phase equilibria.

The contribution of the group-group interactions to the Helmholtz free energy of the system is calculated by means of a sum of a reference term for a system of hard spheres (purely repulsive), and a dispersive (attractive) term expressed as a second-order high-temperature perturbation expansion (terms A_1 and A_2) over the reference term [265]

$$\frac{A^{\text{mono.}}}{Nk_{\text{B}}T} = \frac{A^{\text{HS}}}{Nk_{\text{B}}T} + \frac{A_1}{Nk_{\text{B}}T} + \frac{A_2}{Nk_{\text{B}}T} \ . \tag{3.3}$$

The hard-sphere Helmholtz free energy A^{HS} is obtained from the expression of Boublík [266] and [267] *et al.*. The mean-attractive energy per molecule [170, 171] is expressed as

$$\frac{A_1}{Nk_{\text{B}}T} = \frac{1}{k_{\text{B}}T}\left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i}\nu_k^* S_k\right)(a_1) \ , \tag{3.4}$$

where N_{C} is the number of components, N_{G} the number of groups, $\nu_{k,i}$ the number of groups of type k on component i, ν_k^* the number of identical segments that a group of type k comprises, and a_1 the sum of the pairwise attractive contribution between segments of groups. The pairwise attractive contribution is obtained by using the mean-value theorem and a mapping to the contact value of the pair distribution function of a hypothetical pure hard-sphere fluid at an effective packing fraction expressed in the Carnahan and Starling [151, 264] form, i.e., $g_{0,kl}^{\text{HS}} = g_{0,kl}^{\text{HS}}(\sigma_x, \zeta_{kl}^{\text{eff}})$, as in the SAFT-VR approach [170, 171]. The final expression for the mean attractive energy is given by

$$\frac{A_1}{Nk_{\text{B}}T} = -\frac{\rho}{k_{\text{B}}T}\sum_{i=1}^{N_{\text{C}}}\sum_{j=1}^{N_{\text{C}}} x_i x_j \sum_{k=1}^{N_{\text{G}}}\sum_{l=1}^{N_{\text{G}}} \nu_{k,i}\nu_{l,j}\nu_k^*\nu_l^* S_k S_l \alpha_{kl}^{\text{vdw}} g_{0,kl}^{\text{HS}} \ . \tag{3.5}$$

The second-order (fluctuation) term accounting for the dispersive interactions in the Helmholtz free energy per molecule is obtained from

$$\frac{A_2}{Nk_{\text{B}}T} = \left(\frac{1}{k_{\text{B}}T}\right)^2\left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i}\nu_k^* S_k\right)(a_2) \ , \tag{3.6}$$

where the fluctuation term per segment a_2 is given as a sum of the pair contributions between the segments of the groups. This fluctuation is described using the local compressibility approximation (LCA) [268, 269]. Using the Carnahan and Starling [151, 264] expression for the isothermal compressibility of the reference hard-sphere mixture, K^{HS}, leads to the final expression for the second-order perturbation term:

$$\frac{A_2}{Nk_{\rm B}T} = -\frac{K^{\rm HS}\rho}{2}\left(\frac{1}{k_BT}\right)^2 \sum_{i=1}^{N_{\rm C}}\sum_{j=1}^{N_{\rm C}} x_i x_j \sum_{k=1}^{N_{\rm G}}\sum_{l=1}^{N_{\rm G}} \nu_{k,i}\nu_{l,j}\nu_k^*\nu_l^* S_k S_l \epsilon_{kl}\alpha_{kl}^{\rm vdw} \times \left(g_{0,kl}^{\rm HS} + \zeta_3 \frac{\partial g_{0,kl}^{\rm HS}}{\partial \zeta_{kl}^{\rm eff}} \frac{\partial \zeta_{kl}^{\rm eff}}{\partial \zeta_3}\right) . \quad (3.7)$$

From expressions (3.5) and (3.7) it is clear that the Helmholtz free energy contribution is expressed in an explicit group contribution form as a sum over molecular components and over the molecule's constituent groups. In these expressions the term $\alpha_{kl}^{\rm vdw}$ is the van der Waals attractive parameter, which for square-well segments is simply

$$\alpha_{kl}^{vdw} = \frac{2\pi}{3}\epsilon_{kl}\sigma_{kl}^3(\lambda_{kl}^3 - 1) , \qquad k = 1, ..., N_{\rm G},\ l = 1, ..., N_{\rm G}. \quad (3.8)$$

where ϵ_{kl} is the well-depth of the dispersive interaction between segments k and l. It is the explicit inclusion of this unlike group interaction term in the expressions for the contribution of monomers to the Helmholtz free energy of the system that enables one to use the method to predict the thermodynamic properties of the appropriate mixtures based on pure component data alone, in cases where the parameters for all of the functional groups of the system are found in pure components. Examples include binary systems of n-alkanes with 1-alkanols, which are examined in detail in a later section. During the estimation of the parameters for the functional group of the 1-alkanols (e.g., the CH_2OH functional group) values for the interactions between a given group and the other functional groups of the molecules (i.e., the hydroxyl and the methyl and methylene, $\epsilon_{CH_2OH-CH_3}$ and $\epsilon_{CH_2OH-CH_2}$) are determined. The calculation of the thermodynamic properties of mixtures of n-alkanes with 1-alkanols then follows in a predictive manner, as all of the necessary parameters are determined from the regression to the data of the pure n-alkanes and 1-alkanols, without the need of considering mixture-specific experimental data.

For the calculation of the contribution to the Helmholtz free energy of the system due to the formation of molecules from the monomeric group segments, some mixing rules are introduced, namely for the diameter $\bar{\sigma}_{ii}$, square-well energy $\bar{\epsilon}_{ii}$ and range $\bar{\lambda}_{ii}$ [19]. The Helmholtz free energy due to chain molecule formation is expressed in terms of the number of different segments making up the molecular species based on the first-order thermodynamic perturbation theory (TPT1) of Wertheim [22, 270] by means of these parameters:

$$\frac{A^{\rm chain}}{Nk_{\rm B}T} = -\sum_{i=1}^{N_{\rm C}} x_i \left(\sum_{k=1}^{N_{\rm G}} \nu_{k,i}\nu_k^* S_k - 1\right) \ln g_{ii}^{\rm SW}(\bar{\sigma}_{ii};\zeta_3) . \quad (3.9)$$

In the case of the contribution to the Helmholtz free energy due to the association between the various sites on the groups comprising the molecules, one can write the expression in the standard form [21],

$$\frac{A^{\text{assoc.}}}{Nk_{\text{B}}T} = \sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i} \sum_{a=1}^{N_{\text{ST}_k}} n_{ka} \left(\ln X_{ika} + \frac{1 - X_{ika}}{2} \right) , \tag{3.10}$$

which corresponds to sums over the numbers of components N_{C}, group types N_{G}, and sites types N_{ST_k} on the segments of a group of type k. X_{ika} represents the fraction of component i not bonded at the site of type a which is located on the segments of a group of type k. The fractions of a group not bonded at a given site X_{ika} are obtained from an iterative solution of the corresponding mass action relations [21]. This represents the standard treatment of association employed in SAFT approaches where one can describe dimerisation between molecules, and the formation of chain and network-like structures. The methodology of Wertheim can also be extended to deal with the formation of ring-like molecular structures [271–273], double bonding between given sites [274], and bond-cooperativity [275]. More details of the theory can be found in the original publications [19, 29].

As has been outlined, within the SAFT-γ formalism a heteronuclear model is retained explicitly in the monomer term, while for the treatment of the molecular chain and association terms, a van der Waals like mixing rule is used to approximate the contact value of the heteronuclear molecule. Thus in these terms, each segment is characterised by a set of effective molecular parameters [29]. It is in the treatment of the chain and association terms that SAFT-γ can be distinguished from hetero GC-SAFT-VR [20]. In the treatment of the monomer contribution, as well as in the limit of molecules that are modelled as comprising a single group, e.g., ethane, water, carbon dioxide, the two theories are essentially identical [28]. In the hetero GC-SAFT-VR description, an approximation is used for the contribution of the chain formation to the free energy which has the advantage of retaining the connectivity in the description of the molecules; the same approach is followed in the recent development of the hs-PC-SAFT [224]. Thus within the hetero GC-SAFT-VR and the hs-PC-SAFT formulation, the discrimination between isomers of a given molecule is possible, while in SAFT-γ this discrimination is possible only for those isomers that are built from different functional groups.

3.1.1 Estimation of group parameters from pure component data

The parameter estimation procedure is at the very heart of the group contribution methodology, where in most cases the group parameters are obtained by regression to experimental data. Within the SAFT-γ framework, the like and unlike group interaction parameters

are obtained by estimating pure component data for a series of compounds of the same chemical family. The number of identical segments of each group k, $\nu_k^\star$, as well as the number of associating site types (and the sites of each type) of each group k, NST_k and n_{ka}, are determined by comparing the quality of fit obtained for discrete choices of these parameters. The values tested are chosen to be consistent with a physical representation of the groups, e.g., the number of associating sites is determined so as to correspond to the number of lone pairs of electrons and hydrogen atoms. As has already been emphasised an estimation based on pure component data allows for the determination of like and unlike group parameters for all the pairs of groups present in the molecules used in the estimation. For example, when estimating the parameters for the functional groups of the n-alkanes (CH_3 and CH_2) the value of the unlike dispersive energy ($\epsilon_{CH_3-CH_2}$) is also obtained.

Group parameters are determined from pure component vapour-liquid equilibrium experimental data, i.e., saturated liquid densities and vapour pressures, for a series of compounds belonging to the same chemical family. The temperature range of the data included in the estimation is from 0.3 T_c^{exp} to 0.9 T_c^{exp}, T_c^{exp} being the experimental value of the critical temperature of each component. The region close to the critical point is not included in the estimation procedure, due to the inability of equations of state of this type to represent the critical and subcritical regions satisfactorily with a single set of parameters. The objective function of the parameter estimation is a least-squares function of the generic properties R expressed as

$$\min_{\mathbf{\Omega}} f_{\mathrm{obj}} = \sum_{i=1}^{N_{\mathrm{C}}} \sum_{j=1}^{N_{\mathrm{V}_i}} \sum_{k=1}^{N_{\mathrm{P}_{ij}}} \left[\ln\left(\sigma_{ijk}^2\right) \frac{(R_{ijk}^{\mathrm{exp}} - R_{ijk}^{\mathrm{calc}})^2}{\sigma_{ijk}^2} \right] , \tag{3.11}$$

where $\mathbf{\Omega}$ is the vector of the estimated parameters and the three sums are over all components (N_{C}, index i), over all properties (N_{V_i}, index j), and over all experimental data points ($N_{\mathrm{P}_{ij}}$, index k). The uncertainty of the experimental data used in the estimations is accounted for through the variance of the k^{th} measurement of property j for component i, σ_{ijk}. For the experimental data used in the work presented here, a constant relative variance of 1% was used, which corresponds to a 1% relative experimental error (i.e., $\sigma_{ijk}^2 = (0.01 \cdot R_{ijk}^{\mathrm{exp}})^2$). The estimations were performed using numerical solvers provided by the commercial software package gPROMS© [276].

The deviation of the theoretical description from the experimental data is expressed in terms of the percentage absolute average deviation (%AAD) for each property R_{ij} of molecule i as follows:

$$\%\mathrm{AAD}R_{ij} = \frac{1}{N_{\mathrm{P}_{ij}}} \sum_{k=1}^{N_{\mathrm{P}_{ij}}} \left| \frac{R_{ijk}^{\mathrm{exp}} - R_{ijk}^{\mathrm{calc}}}{R_{ijk}^{\mathrm{exp}}} \right| \times 100 \ . \tag{3.12}$$

The procedure for the determination of the group parameters is initiated with the family of n-alkanes. The parameters for the functional groups CH_3 and CH_2, as well as the interaction between these two groups, are estimated by providing the optimal description of the pure component experimental data from ethane to n-decane. A total of 7 parameters are estimated simultaneously (the group diameter σ_{kk}, interaction energy ϵ_{kk} and range λ_{kk} for both groups, as well as the unlike interaction ϵ_{kl} between the two functional groups) from the available experimental data [277]. The values of the shape factor S_k for the CH_3 and the CH_2 groups are fixed to 2/3 and 1/3 respectively (where both groups are composed of one segment only, $\nu^{\star}_{\mathrm{CH_3}} = 1$ and $\nu^{\star}_{\mathrm{CH_2}} = 1$) , following an analogous strategy to that within the framework of the SAFT-VR EoS, where the chain length m is related to the number of the carbon atoms of an alkane ($m = 1/3(C-1)+1$) [215, 245]. The description of the pure component vapour-liquid equilibria of the correlated n-alkanes is shown in figures 3.1, for the coexistence densities, and 3.2 for the vapour pressures. It can be seen that the SAFT-γ method can accurately describe the phase behaviour of the correlated compounds, with the average error for all compounds included in the regression being 3.98% for the vapour pressure, p_{vap}, and 0.57% for the saturated liquid density, ρ_{sat} [19].

The optimised parameters are then transferred to the study of other chemical families for the determination of the group parameters of additional functional groups in a sequential manner. The current parameter table of the method contains the parameters for 11 functional groups, including parameters for the families of n-alkanes, n-alkylbenzenes, mono- and di-unsaturated hydrocarbons as well as associating compounds such as the 1-alkanols, primary amines, and carboxylic acids [19, 29]. The average deviations over all compounds of each chemical family studied in previous work for the two properties considered are summarised in table 3.1. The average %AADs over all chemical families are 3.34% for the vapour pressure and 0.92% for the saturated liquid density [19, 29], where it is worth noting that the smallest deviations are observed for the families comprising the OH and NH_2 associating groups (the primary alkanols and amines) possibly due to the dominant nature of the hydrogen bond on the thermophysical properties.

The predictive capability of the method was initially tested in predicting the pure component fluid phase equilibria of high molecular weight compounds that were not included

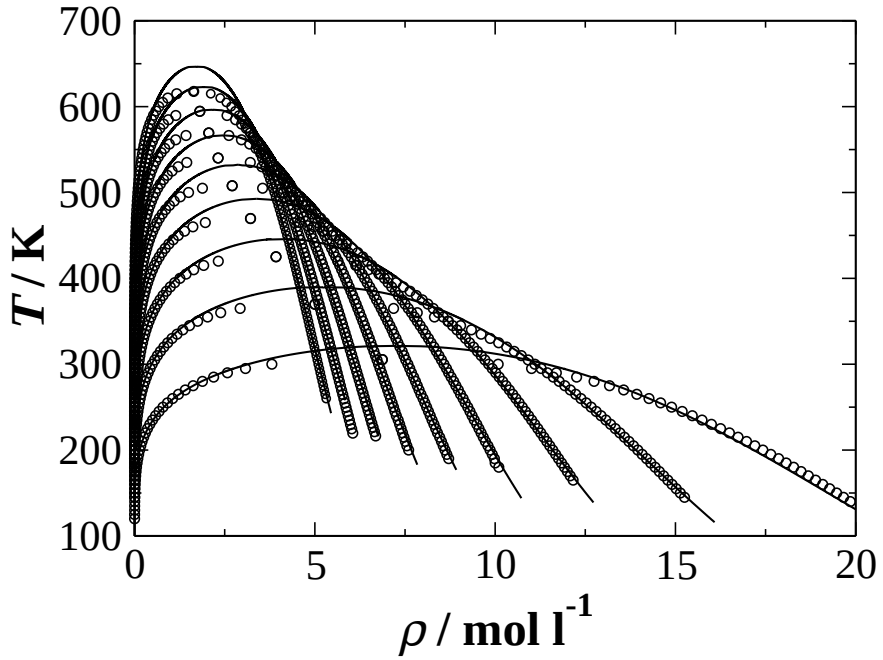

Figure 3.1: Description of the coexistence densities as a function of temperature for the linear alkanes (n-ethane to n-decane from bottom to top) included in the estimation of the CH_3 and CH_2 group parameters within the framework of the SAFT-γ group contribution approach. The symbols represent correlated experimental data from NIST [277] and the continuous curves the calculations with the theory.

Table 3.1: Overall average absolute deviations (%AAD) of the vapour pressures p_{vap} and saturated liquid densities ρ_{sat} within the SAFT-γ framework compared to experiment for all of the chemical families included in the database.

Chemical family	Number of Compounds	%AAD p_{vap}	%AAD ρ_{sat}
n-alkanes	9	3.98	0.57
branched alkanes	10	2.92	0.39
n-alkylbenzenes	9	4.12	1.49
alkenes	19	4.78	0.74
2-ketones	8	3.47	1.14
carboxylic acids	8	3.60	1.42
primary amines	9	1.64	0.53
1-alkanols	9	2.19	1.09
Average %AAD	-	3.34	0.92

in the parameter estimation procedure. The overall deviations for these predictions are 8.18% for p_{vap} and 0.75% for ρ_{sat} [19, 29]; the larger deviations observed in the vapour

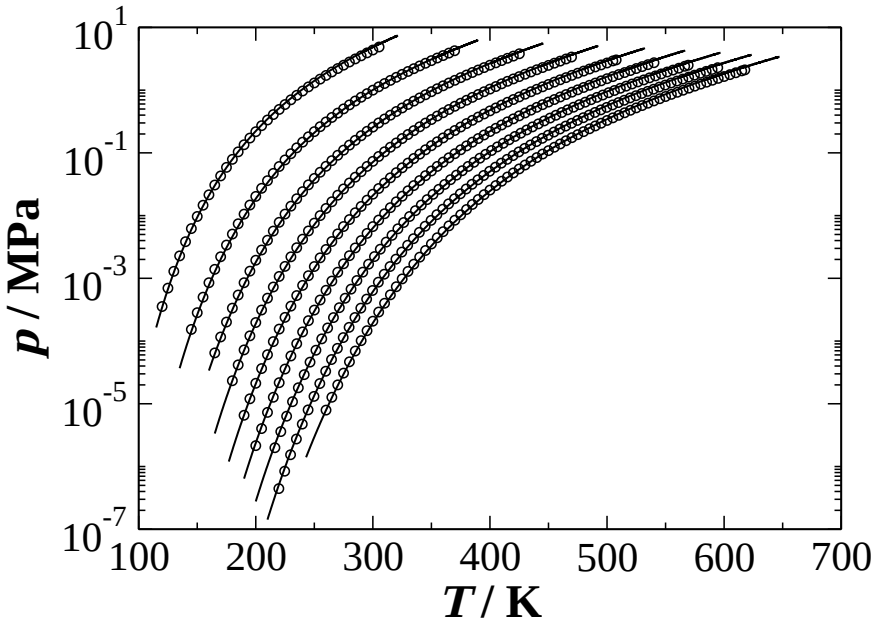

Figure 3.2: Description of the vapour pressure for the linear alkanes (n-ethane to n-decane from left to right) included in the estimation of the CH_3 and CH_2 group parameters within the framework of the SAFT-γ group contribution approach. The symbols represent correlated experimental data from NIST [277] and the continuous curves the calculations with the theory. The pressure is plotted in logarithmic scale to highlight both the high- and low-temperature regions.

pressure can be attributed to the extremely small values of the vapour pressures for the bigger compounds for the temperature range of the predictions.

3.2 Parameters Studied

The performance of the SAFT-γ method in the description of the fluid phase behaviour of aqueous solutions of hydrocarbons, and more specifically of n-alkanes and alkanols, is the subject of the current work. In order to perform such a study the necessary group parameters and unlike group-group interactions have to be determined from experimental data. This procedure is discussed in detail in the following sections.

3.2.1 Pure component parameters: water

The model for water used in the study of aqueous solutions with the SAFT-γ approach was the one developed by Clark *et al.* [278]. The model was originally formulated within the SAFT-VR approach, however the equivalence of the two approaches for molecules comprising a single spherical group allows for its use in SAFT-γ without further modifications. The shape factor, S_{H_2O}, was set to the value of the chain length of the original model,

which for the case of spherical groups in the SAFT-VR description is $S_{H_2O} = m_{H_2O} = 1$. In order to tackle the degeneracy of the molecular parameters when developing a model for water, a detailed examination of the parameter space was carried out as discussed in [278]. Regarding the association scheme, Clark *et al.* [278] proposed a model with 4 association sites, with two sites for the lone electron pair and two for the hydrogen atom, where association is allowed only between sites of different type. The specific values of the parameters for water are reported in table 3.2. The proposed model is found to correlate accurately the available data [277], as shown in figure 3.3, with corresponding %AADs of 0.99% for the vapour pressure and 1.28% for the saturated liquid densities for temperatures up to 0.9 T_c^{exp}.

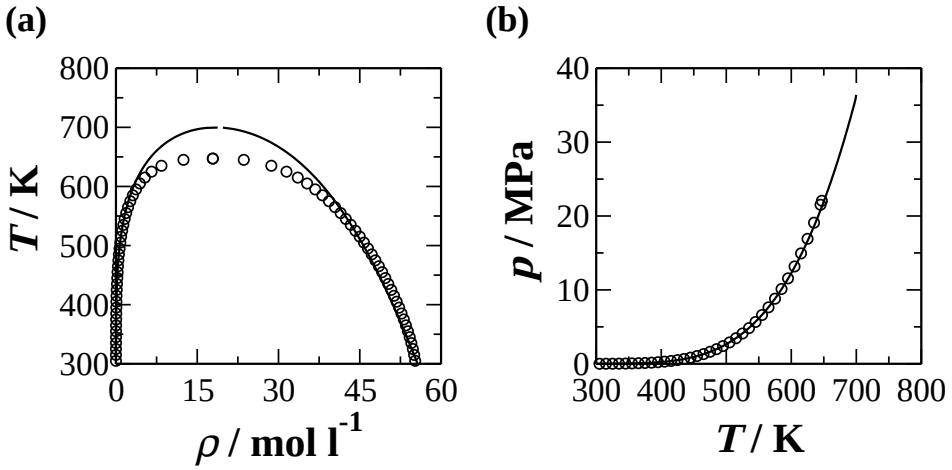

Figure 3.3: Comparison between the description of water with the model of Clark *et al.* [278] and the correlated experimental data from NIST [277] for (a) the coexisting liquid and vapour densities, and (b) the vapour pressure as a function of temperature.

3.2.2 Pure component parameters: 1-alkanols

A preliminary analysis of the existing treatment of the description used in the modelling of the chemical family of 1-alkanols is carried out first. The family of 1-alkanols was initially modelled by means of an OH functional group, as presented in [19]. The OH group employed by Lymperiadis *et al.* resulted to an accurate description of the pure component VLE of the 1-alkanols studied, with %AADs of 2.19% for the vapour pressure and 1.09% for the saturated liquid density. However, the accuracy of the representation of the fluid phase behaviour of binary mixtures of *n*-alkanes+1-alkanols was not found to be as good as one would have hoped. This is illustrated here by comparing the predictions of the theory and the experimental data for the vapour-liquid equilibria of the binary system of *n*-heptane+1-pentanol, as shown in figure 3.4. From the figure it can be seen that although

the theory reproduces the experimental composition of the vapour phase, the liquid side of the phase envelope and the location of the azeotrope exhibited by the mixture are not very well described.

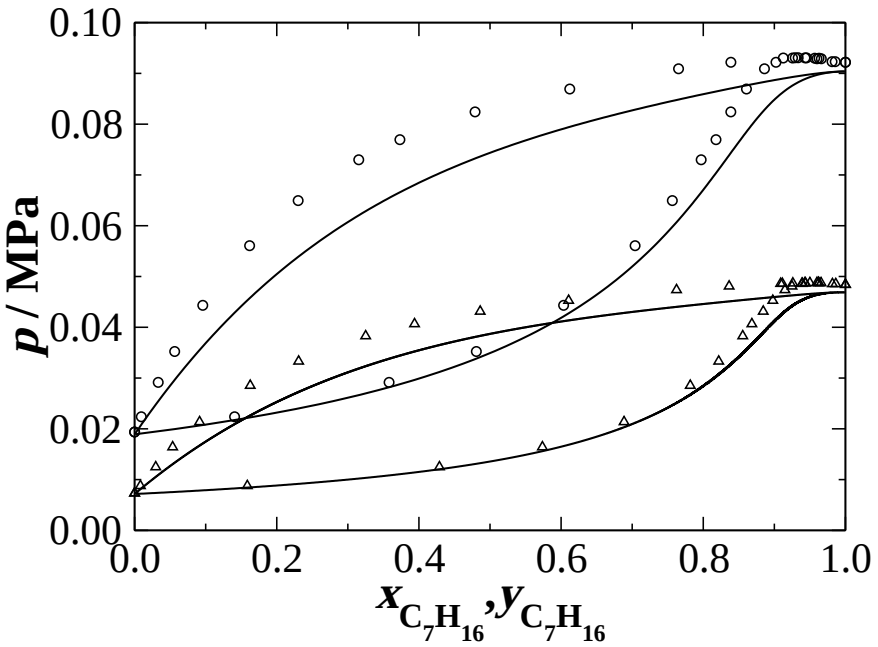

Figure 3.4: Prediction of the phase behaviour of the binary system *n*-heptane+1-pentanol as pressure-composition isotherms at two different temperatures. Circles represent the experimental data at 368.15 K and the triangles at 348.15 K [279], where the continuous curves are the predictions of the SAFT-γ EoS with the 1-alkanols modelled with an OH functional group.

In an attempt to improve the predictive capability of the method in the description of binary systems, the modelling of the chemical family of 1-alkanols was revisited. One has to bear in mind that the identification of groups is in most cases based on heuristics and basic chemical experience, with the final choice being usually the combination of groups that results in the best representation of the experimental data. The combination of atoms chosen to represent the functional groups identified on a molecule is chosen based on a balance between the predictive capability of the method and the number of parameters (mainly unlike group parameters) to be estimated; the use of large functional groups often increases the accuracy of the method but at the same time the number of cross interaction parameters to be determined would also increase. The OH functional group may be a common choice in modelling 1-alkanols, but there is no firm theoretical consideration dictating this option.

Wu and Sandler [107, 108] have presented a methodology for the identification of functional

Table 3.2: SAFT-γ square-well potential parameters for the functional groups of the n-alkanes, 1-alkanols and water. The CH_2OH group features 2 association sites of type a and 1 of type b, whereas the H_2O group has 2 sites of each type.

Group k	ν_k^*	S_k	λ_k	σ_k [Å]	ϵ_k/k_B [K]	$n_{k,a}$	$n_{k,b}$	ϵ^{HB}_{kkab}/k_B [K]	r^c_{kkab} [Å]
CH_3	1	0.667	1.413	3.810	252.601	-	-	-	-
CH_2	1	0.333	1.661	4.027	240.482	-	-	-	-
CH_2OH	1	0.566	1.652	4.317	399.959	2	1	2555.721	2.359
H_2O	1	1.000	1.789	3.034	250.000	2	2	1400.000	2.1082

groups based on the requirement that a group should have the same net charge regardless of the molecule in which it appears. They have suggested that a CH_2OH group is more appropriate than independent adjacent CH_2 and OH groups in representing the primary alcohols. From a chemical perspective the use of a CH_2OH group to model 1-alkanols is more physically sound than the use of an OH group, particularly when one considers the polarisation effect of the OH on the neighbouring carbon group. The presence of the OH functional group affects the neighbouring group of the alkyl chain in a way that the CH_n group exhibits different characteristics than the rest of the groups on the alkyl chain of the alkanol. This is the underlying concept for the choice of a CH_2OH functional group (in the case of primary alkanols), where the effects of association, related to the hydrogen bonding between the lone electron pairs of the oxygen and the hydrogen atom, and polarisation effects are incorporated in a single group. The choice of a CH_2OH functional group is also in line with the theoretical analysis of the solutions of groups concept of Currier and O'Connell [13], where based on a statistical mechanical analysis it is concluded that the inherent approximations of the theory dictate the grouping of molecules to be such that the dominant electronic and electrostatic effects are completely localised in the groups.

In revisiting the parameters for the family of 1-alkanols the modelling strategy proposed by Wu and Sandler [107] is followed and the modelling of 1-alkanols is done by means of a CH_2OH functional group. The CH_2OH group is represented as one segment (i.e., $\nu^*_{CH_2OH} = 1$) featuring 3 association sites (two to represent the two lone pairs on the oxygen atom and one for the hydrogen atom), following the 3B associating scheme proposed by Huang and Radosz [167]. The parameters obtained for the CH_2OH functional group, based on pure component vapour-liquid equilibrium data, are summarised in table 3.2, together with these for the alkyl groups and water.

The regression to experimental vapour pressure (p_{vap}) and saturated liquid density (ρ_{sat}) data for the series of 1-alkanols (from ethanol to 1-decanol) results in an average %AAD

Table 3.3: Percentage average absolute deviations (%AAD) of the vapour pressures p_{vap} and saturated liquid densities ρ_{sat} obtained with the SAFT-γ framework compared to experiment [280] (where n is the number of data points) for the 1-alkanols using a CH_2OH functional group.

Compound	T range [K]	n	%AAD p_{vap}	T range [K]	n	%AAD ρ_{sat}
C_2H_5OH	231-463	30	3.77	159-463	39	1.32
C_3H_7OH	280-483	25	3.81	169-483	38	0.68
C_4H_9OH	295-506	26	1.92	186-506	39	0.61
$C_5H_{11}OH$	278-508	30	1.82	278-508	29	0.54
$C_6H_{13}OH$	310-428	17	0.76	273-547	38	0.85
$C_7H_{15}OH$	343-445	14	0.72	273-563	38	1.01
$C_8H_{17}OH$	296-549	31	1.99	263-583	39	1.14
$C_9H_{19}OH$	366-481	15	1.39	293-596	38	1.43
$C_{10}H_{21}OH$	301-526	27	3.61	293-613	38	1.92
Average	-	-	2.20	-	-	1.06

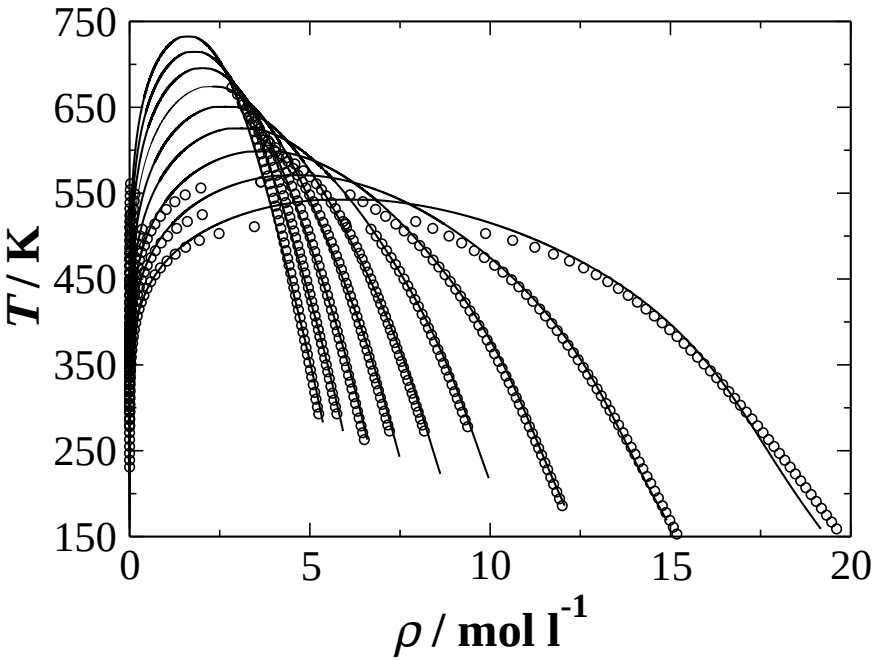

Figure 3.5: Description of the coexistence densities as a function of temperature for the family of 1-alkanols (ethanol to 1-decanol from bottom to top) included in the estimation of the CH_2OH group parameters within the framework of the SAFT-γ group contribution approach. The symbols represent correlated experimental data from [280] and the continuous curves the calculations with the theory.

of 2.20% for p_{vap} and 1.06% for ρ_{sat}. The deviations per compound as well as the temperature ranges of the experimental data used are shown in detail in table 3.3. The description

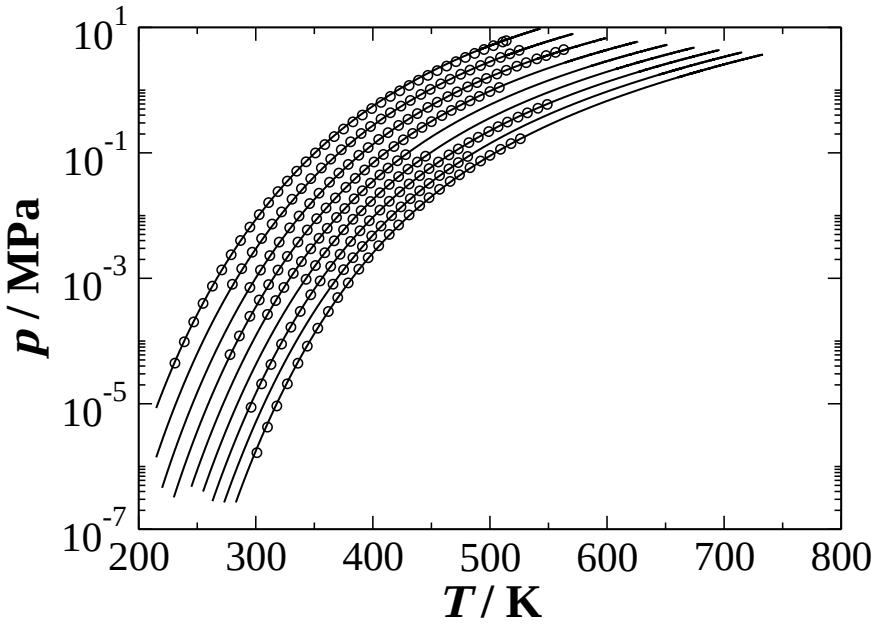

Figure 3.6: Description of the vapour pressure as a function of temperature (Clausius-Clapeyron representation) for the family of 1-alkanols (ethanol to 1-decanol from top to bottom) included in the estimation of the CH_2OH group parameters within the framework of the SAFT-γ group contribution approach. The symbols represent correlated experimental data from [280] and the continuous curves the calculations with the theory. The pressure is plotted in logarithmic scale to highlight both the high- and low-temperature regions.

of the pure component fluid phase behaviour of the correlated 1-alkanols is presented on figure 3.5 for the coexistence densities and figure 3.6 for the vapour pressure. The average deviations observed are essentially the same as the ones found when applying an OH functional group (%AADs of 2.19% and 1.09%, respectively [19]). From this comparison it is evident that the use of a CH_2OH functional group for the modelling of 1-alkanols does not result in a significantly improved description of the vapour-liquid equilibria of the pure components. Moreover, the application of such a functional group is limited only to primary alkanols; other functional groups such as CHOH and COH would have to be used to describe secondary and tertiary alkanols. Importantly, however, the use of this different group allows for a more accurate representation for the phase behaviour of binary mixtures containing 1-alkanols as will be demonstrated for binary mixtures of 1-alkanols+n-alkanes in a later section. This provides a strong motivation for introducing the CH_2OH functional group.

3.2.3 Binary interaction parameters: water-alkyl groups

An advantage of the SAFT-γ group contribution approach lies in its predictive capability for the fluid phase behaviour and thermodynamic properties of mixtures based on pure component data alone, as long as the necessary information pertaining to the groups of the mixture can be gleaned from the pure components. Unfortunately, this is not the case for aqueous solutions, as water has to be treated as a single functional group. As a consequence, the unlike interaction parameters between water and other functional groups have to be determined based on available experimental data for the appropriate mixtures.

The study of mixtures containing water requires the determination of the unlike group interaction energy parameters for the unlike dispersion interactions and the unlike association, in cases when water is mixed with associating molecules. All other unlike group parameters (i.e., σ_{kl}, λ_{kl}, and r^{c}_{klab} where applicable) are calculated by means of standard combining rules [19]. The advantage of using a so-called solution-of-groups approach is that the unlike group interaction parameters are transferable, so that a minimal amount of experimental data is required. The parameters can be determined by regression to the experimental data for the fluid phase behaviour of a single mixture, and can then be transferred to the study of a series of mixtures comprising molecules formed from these groups. For the study presented here, the unlike group dispersion energies $\epsilon_{CH_3-H_2O}$ and $\epsilon_{CH_2-H_2O}$ are obtained by estimating the available experimental data for the vapour-liquid and liquid-liquid phase equilibria of the binary mixture of water+n-hexane [281–283]. This specific system was selected based on the chain length of hexane; a medium sized chain allows for the study of the transferability of the parameters for systems comprising both shorter and longer alkanes. This choice was also supported by the extensive experimental data available for the fluid phase behaviour of the system.

The VLE and LLE phase equilibria of the system are calculated by means of a (p, T) flash algorithm based on the work by [284], and every point is therefore calculated at a given temperature and pressure corresponding to the experimental data. The algorithm used for the estimation of the unlike group interactions is based on sequential quadratic programming with a numerical evaluation of the derivatives of the objective function [285]. The objective function used for the regression is a least-squares function characterising the absolute deviation of the predicted compositions of each phase from the experimental data for the system as:

$$\min_{\mathbf{\Omega}} f_{\text{obj}} = \left[\sum_{i=1}^{N_{\text{P}}} \left\{ (x_{i,\text{H}_2\text{O}}^{\text{exp}} - x_{i,\text{H}_2\text{O}}^{\text{calc}})^2 + (y_{i,\text{H}_2\text{O}}^{\text{exp}} - y_{i,\text{H}_2\text{O}}^{\text{calc}})^2 \right\} \right] . \tag{3.13}$$

In the estimation procedure, the available experimental compositions for water are used, both in the VLE and the LLE region of the phase envelope. It has to be noted that the very small values of the solubility of *n*-hexane in water are not included in the estimation. The aim of this study is to provide a generic model that describes the properties of the aqueous solutions of the entire homologous series of the *n*-alkanes. Previous studies of aqueous solutions of alkanes with the SAFT-VR EoS lead to the conclusion that the use of the same binary interaction parameters for all phases does not allow for a simultaneous description of the solubility of water in the alkane-rich phase and the solubility of alkanes in the water-rich phase [112]. However, SAFT as a general method has been shown to describe accurately the low solubilities encountered in water+*n*-alkane systems, when this feature is under specific investigation [112, 256]. The accurate description of the solubility of alkanes in the water-rich phase can also lead to satisfactory results in other challenging aspects such as the water-octanol partition coefficients of alkanes, which have been studied previously with group contribution methods [286], but this is not the aim of the current work.

The resulting SAFT-γ description of the fluid-phase behaviour for the water+*n*-hexane mixture is depicted in figure 3.7. It is evident that the estimated unlike group interaction parameters allow for an accurate description of both the vapour-liquid equilibria and the solubility of water in the alkane-rich phase (i.e., liquid-liquid equilibria) over a wide range of pressures (∼2 to 80 MPa). The calculated three-phase (vapour-liquid-liquid coexistence) line is in excellent agreement with the experimental data. As mentioned previously, the simultaneous accurate description of both the VLE and the LLE of a system remains a significant challenge for group contribution approaches. The optimal unlike group dispersion energy interaction parameters are found to be $\epsilon_{\text{CH}_3-\text{H}_2\text{O}}/k_{\text{B}} = 190.761$ K and $\epsilon_{\text{CH}_2-\text{H}_2\text{O}}/k_{\text{B}} = 174.481$ K (see table 3.4). The quality of the representation of the experimental data for the water+*n*-hexane mixture is included in table 3.5 where the %AADs in pressure and temperature are calculated as in eq. (3.12) and the absolute error in composition is defined as:

$$\Delta z = \frac{1}{n} \sum_{i=1}^{n} \left| z_i^{\text{exp.}} - z_i^{\text{calc.}} \right| \times 100 , \tag{3.14}$$

where z_i is the composition of component i in the phase of interest (i.e., x_i for the liquid phase and y_i for the vapour phase) and n is the number of data points for the mixture of

interest. The transferability of group parameters to the study of other aqueous solutions of alkanes of varying length is discussed in section 3.3.3.

3.2.4 Binary interaction parameters: water-hydroxymethyl group

The same approach as presented in section 3.2.3 for water+n-alkane mixtures is followed for the determination of the unlike group interaction parameters between H_2O and the CH_2OH group of the 1-alkanols homologous series. For the study of aqueous solutions of 1-alkanols, the parameters between water and the CH_3 and CH_2 functional groups of the alkanes are transferred from the previous study of aqueous alkanes, while the missing interaction parameters, i.e., the cross dispersion energy $\epsilon_{CH_2OH-H_2O}$ and association energy $\epsilon^{HB}_{CH_2OH-H_2O}$, are obtained by estimation to a small set of experimental data. The bonding volume between water and the alkanol is calculated by means of the following combining rule [171]:

$$K_{ijklab} = \left(\frac{K^{1/3}_{iikkab} + K^{1/3}_{jjllab}}{2} \right)^3 , \tag{3.15}$$

where the volume available for bonding between sites a and b on groups k and l of components i and j is calculated from

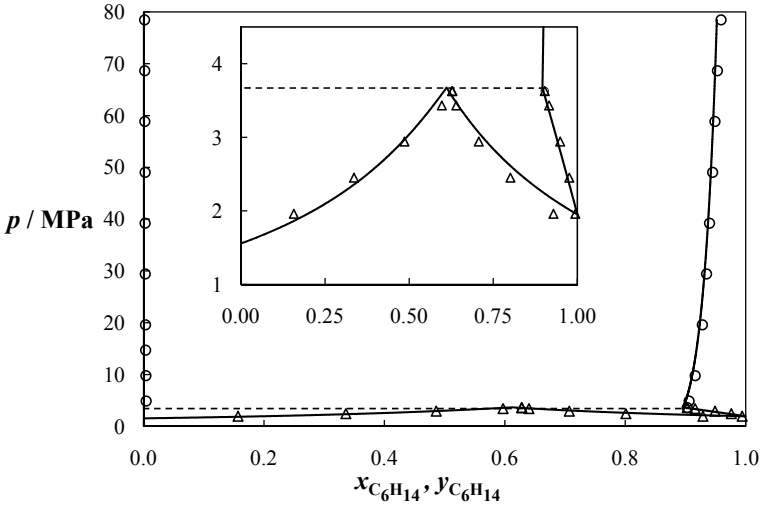

Figure 3.7: Isothermal pressure-composition phase diagram for the mixture water+n-hexane at 473.15 K used for the determination of the cross interaction parameters of water with the functional groups of the n-alkanes. The continuous curves represent the SAFT-γ calculations, the triangles the experimental VLE data [281], and the circles the experimental LLE data [282, 283]. The vapour-liquid phase envelope and the three phase region can be clearly seen in the inset image, where the dashed line denotes the three-phase vapour-liquid-liquid coexistence line.

Table 3.4: Unlike energetic group parameters for the dispersion and association interactions between functional groups. The parameters denoted with a are estimated from pure component data, whereas the parameters with b are obtained by regression to experimental data for the phase behaviour of mixtures. Any parameter not shown takes a value of zero (e.g., $\epsilon^{\mathrm{HB}}_{klab}/k_{\mathrm{B}} = 0$).

$\epsilon_{kl}/k_{\mathrm{B}}$ [K]	CH_3	CH_2	CH_2OH	H_2O
CH_3	252.60^a	261.520	279.939	190.761
CH_2	261.520^a	240.482^a	283.702	174.481
CH_2OH	279.939^a	283.702^a	399.959^a	431.345
H_2O	190.761^b	174.481^b	431.345^b	250.000^a
$\epsilon^{\mathrm{HB}}_{klab}/k_{\mathrm{B}}$ [K]	CH_3	CH_2	CH_2OH	H_2O
CH_3	-	-	-	-
CH_2	-	-	-	-
CH_2OH	-	-	2555.721^a	1988.154^b
H_2O	-	-	1988.154^b	1400.000^a

$$K_{ijklab} = \frac{4\pi\bar{\sigma}_{ij}^2}{72r_d^2}\left[\ln\left(\frac{r^c_{klab}+2r_d}{\bar{\sigma}_{ij}}\right)\left(6r^{c^3}_{klab}+18r_d r^{c^2}_{klab}-24r_d^3\right)\right.$$
$$+(r^c_{klab}+2r_d-\bar{\sigma}_{ij})\left(22r_d^2-5r_d r^c_{klab}-7r_d\bar{\sigma}_{ij}-8r^{c^2}_{klab}\right.$$
$$\left.\left.+r^c_{klab}\bar{\sigma}_{ij}+\bar{\sigma}_{ij}^2\right)\right]\ . \tag{3.16}$$

In eq. 3.16, r_d is the distance of the association site from the centre of the effective sphere of interaction and has a fixed value ($r_d/\bar{\sigma}_{ij} = 0.25$).

The regression is made based on the available VLE [287] and LLE [288] data for the binary mixture water+1-pentanol which are shown in figure 3.8. As in section 3.2.3, the solution of an alkanol of medium size for which sufficient experimental mixture data is available is selected for the estimation of the interactions between water and the CH_2OH functional group. The objective function used in the regression is given by eq. (3.13), where the experimental values of the composition of water in all studied phases are used. Given the increased solubility of alkanols in water (compared to the water+n-alkane mixtures), experimental data for both sides of the liquid-liquid equilibria of water+1-pentanol are included in the regression of the interaction parameters.

The values of the new estimated parameters are $\epsilon_{CH_2OH-H_2O}/k_{\mathrm{B}} = 431.345$ K for the unlike dispersion energy and $\epsilon^{\mathrm{HB}}_{CH_2OH-H_2O}/k_{\mathrm{B}} = 1988.154$ K for the unlike energy of the

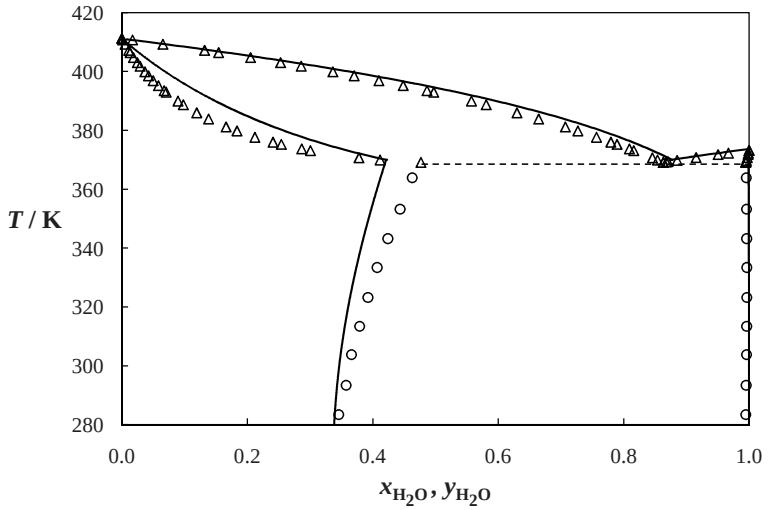

Figure 3.8: Isobaric temperature-composition phase diagram for the mixture water+1-pentanol at 101.3 kPa used for the determination of the cross interaction parameters of water with the CH_2OH group of the 1-alkanols. The continuous curves represent the SAFT-γ calculations, the triangles the experimental VLE data [287], the circles the experimental LLE data [288], and the dashed line denotes the three-phase vapour-liquid-liquid coexistence line.

association (cf. table 3.4). The resulting high value of the dispersive interaction between the CH_2OH functional group and water is justified as it also includes in an effective way the effect of the strong dipole-dipole interactions between these two groups and compensates for the weak attractive interaction between water and the alkyl part of the alkanol, as shown in section 3.2.3. The resulting group-averaged molecular dispersive interaction $\bar{\epsilon}_{ij}$ [19] between water and the alkanol lies close to the geometric mean of the like interactions, a finding consistent with previous studies of aqueous solutions of methanol [278] and ethanol within the SAFT-VR framework [187], as well as of aqueous solutions of longer alkanols with the CPA EoS [260]. The quality of the regression in terms of the deviation from the experimental data is summarised in table 3.5. From figure 3.8 one can see that the set of group parameters allows for a good description of the fluid phase behaviour of the system. The compositions of both phases are well described both in the VLE and LLE regions of the phase envelope, and the position and the compositions of the three-phase line are in good agreement with the experimental data. The performance of the method in the prediction of the fluid phase behaviour of aqueous systems of 1-alkanols of varying molecular weight is discussed in the following section.

Table 3.5: Results of the regression to the experimental data for the vapour-liquid and liquid-liquid equilibria of mixtures to obtain the SAFT-γ unlike group interactions between water and the functional groups of the n-alkanes and 1-alkanols. The values of the %AAD are obtained from eq. (3.12) and Δx, Δy from eq. (3.14). The values in brackets represent the %AAD for the coexisting compositions.

Mixture	Data	Ref.	T or p	%AAD(p or T)	$\Delta(y$ or $x_{\mathrm{I}})$ [a]	Δ (x or x_{II}) [b]
H_2O + C_6H_{14}	VLE	[281]	473.15 K	3.53	3.34 (20.39)	-
H_2O + C_6H_{14}	LLE	[282, 283]	473.15 K	-	0.45 (7.24)	0.30 (99.99)
H_2O + $C_5H_{13}OH$	VLE	[287]	101.3 kPa	1.25	3.56 (57.87)	-
H_2O + $C_5H_{13}OH$	LLE	[288]	101.3 kPa	-	2.63 (6.39)	0.41 (90.34)

[a] For LLE, refers to the solubility of water in the alkane/alkanol-rich phase

[b] For LLE, refers to the solubility of the alkane/alkanol in the water-rich phase

3.3 Predictions

3.3.1 Pure Components: 1-alkanols

A preliminary step in the validation of the parameters obtained for the functional groups is the examination of the performance of SAFT-γ in predictions of the VLE of pure compounds that were not included in the regression procedure. In the case of pure 1-alkanols the adequacy of the parameters obtained for the CH_2OH functional group is tested by examining long-chain 1-alkanols. The predictions for the pure component VLE of $C_{12}H_{25}OH$, $C_{14}H_{29}OH$, and $C_{18}H_{37}OH$ are depicted in figures 3.9(a) and 3.9(b). From the figures it can be seen that one is able to predict accurately the phase behaviour of the longer compounds of the chemical family of 1-alkanols. The %AADs characterising the predictive capability for the VLE of these compounds are 9.33% for the vapour pressure and 0.73% for the saturated liquid density; the larger deviations in pressure can be attributed to the fact that for compounds of higher molecular weight the values of the vapour pressures are extremely low over the temperature range studied, so that small absolute errors lead to large relative errors. The deviations in the prediction of the fluid phase behaviour of long compounds are similar to the description with an OH functional group (cf. %AADs or 9.25% for the vapour pressure and 0.78% for the saturated liquid density [19]). It is evident that the consideration of an CH_2OH functional group does not improve significantly the predictions of the fluid phase equilibria for pure 1-alkanols of high molecular weight not included in the parameter estimation data set. The advantage of using the CH_2OH functional group is mainly supported by the improved performance of the methodology for mixtures, as will become clear in the following section.

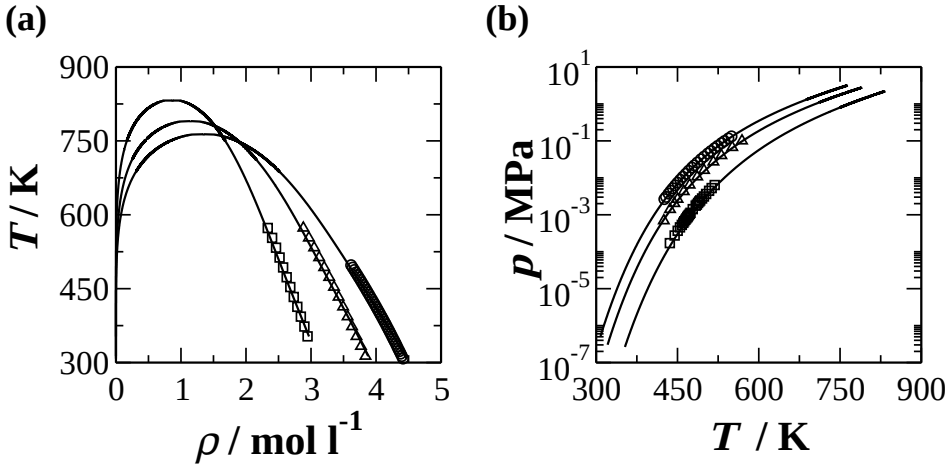

Figure 3.9: SAFT-γ predictions of the pure component vapour-liquid equilibria for 1-alkanols not included in the estimation procedure compared with experimental data: (a) temperature - coexistence density envelope; and (b) vapour-pressure curves shown in a logarithmic representation. The circles represent the experimental data for 1-dodecanol [289, 290], the triangles for 1-tetradecanol [291, 292], and the squares for 1-octadecanol [292, 293]. The continuous curves are the corresponding predictions.

3.3.2 Binary Mixtures: n-alkanes+1-alkanols

The predictive capability of the SAFT-γ GC method can be extended to the study of the phase behaviour and thermodynamic properties of multicomponent mixtures, as in some cases these can be predicted based on parameters obtained from pure component data alone. For the n-alkane+1-alkanol binary mixtures, all the necessary parameters can be obtained from pure component data of the two homologous series as all of the like and unlike group interactions under consideration in the mixtures are also present in the pure components. This is a very important advantageous feature of the method that distinguishes the SAFT-γ approach from other group contribution techniques. Activity coefficient models (such as UNIFAC and ASOG) are developed from a large database containing mixture-specific experimental data, and this tight dependency on the existing data is therefore inherent into EoS-g^{E} models that use such methods. One should also add that group contribution approaches developed within SAFT that are based on homonuclear molecular models offer a predictive capability for mixtures only when the values of the binary interaction parameters are estimated by means of approximate combining rules. It is important to note that within the SAFT-γ framework, the values of the unlike group interactions are not obtained with combining rules but are estimated from the pure component experimental VLE data. This is a common feature of the SAFT-γ [19, 29] and the hetero GC-SAFT-VR [20, 221] and hs-PC-SAFT [224] methods resulting from the explicit heteronuclear model that is employed within the aforementioned theories.

An example of the quality of the SAFT-γ predictions for the fluid phase behaviour for the systems in question can be seen in figure 3.10 for the *n*-heptane+1-pentanol binary mixture. A comparison of the prediction for the two modelling strategies of the hydroxyl group of 1-alkanols, namely the use of an OH or a CH_2OH functional group, is shown in the figure. It can be seen that the incorporation of a CH_2OH functional group provides a significant improvement in the prediction of the fluid phase behaviour of the mixture, especially for the liquid phase, and a more accurate representation of the minimum boiling azeotrope that the mixture exhibits at high concentrations of the hydrocarbon. The improved performance of the method in the prediction of binary mixtures, together with the physicochemical considerations of the molecule, lend strong support to the use of the CH_2OH functional group when modelling 1-alkanols. The SAFT-γ method is shown to compare well with the predictions of the original UNIFAC approach, shown on figure 3.10 as dashed-dotted curves. The UNIFAC calculations are based on the published parameters of Hansen *et al.* [83] for the calculation of the activity coefficients, whereas the gas phase is assumed to be ideal (i.e., $\phi^{\mathrm{V}} = 1$). It is important to highlight that the predictions of the SAFT-γ method are based on parameters from pure component data alone, while the UNIFAC parameters have been obtained by regression to binary mixture data.

One of the most important advantages of the formulation of a GC approach within an equation of state is that the methodology can be applied to a wide range of thermodynamic conditions. The range of applicability can be a serious limitation for methodologies like UNIFAC as discussed in section 2.2.2.1, where the performance of the method depends highly on the range of the experimental data used for the parameterisation of the model. This is illustrated in figure 3.11, where the predictions of the SAFT-γ approach are compared with calculations using the original UNIFAC approach for the fluid phase behaviour of the binary system of *n*-hexane+ethanol. The two isotherms shown (473.15 K and 483.15 K) are at temperatures close to the critical points of both components of the system ($T^{\mathrm{exp}}_{\mathrm{c;C_6H_{14}}} \simeq 507\ K$ and $T^{\mathrm{exp}}_{\mathrm{c;C_2H_5OH}} \simeq 514\ K$). From the figure it can be seen that the predictions with SAFT-γ are in better agreement with the experimental data for the composition of the dew curve of the phase envelope compared to the predictions with the original UNIFAC (parameters from [83] and ideal gas phase assumed). Significant deviations from the experimental data are seen on the right-hand side of the plot with both methodologies, corresponding to the pure component vapour pressure of ethanol at the presented temperatures. In the case of SAFT-γ this is due to the fact that the pure component data close to the critical point are not included in the estimation of the group parameters, whereas for UNIFAC the deviations arise from the coefficients of the Antoine

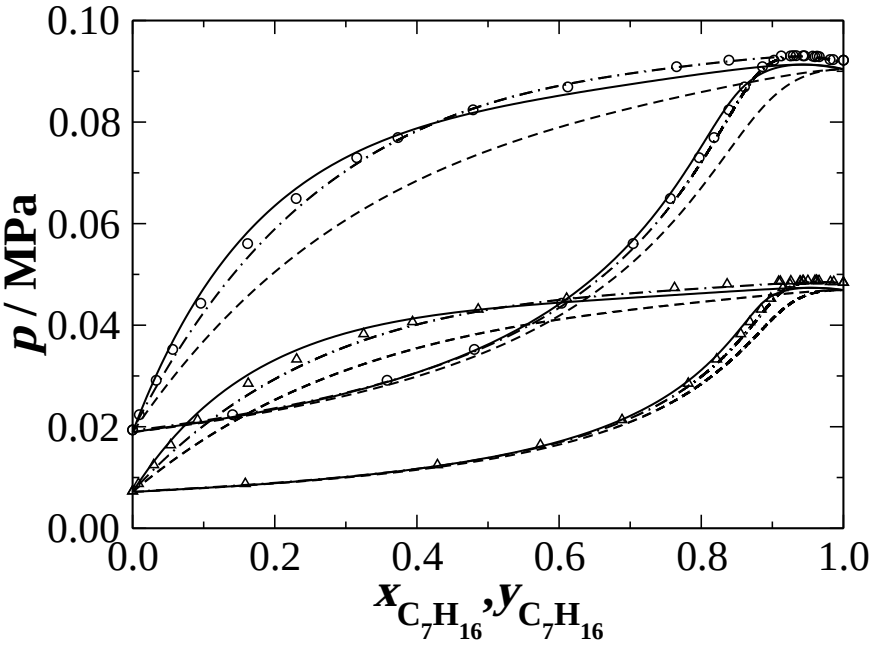

Figure 3.10: Prediction of the vapour-liquid phase behaviour of the binary mixture *n*-heptane+1-pentanol as pressure-composition isotherms at two different temperatures. The circles represent the experimental data at 348.15 K, and the triangles at 368.15 K [279]. The continuous curves are the SAFT-γ predictions modelling 1-alkanols with a CH_2OH functional group, the dashed lines are the predictions using an OH group, and the dashed-dotted lines are the UNIFAC predictions with parameters from Hansen *et al.* [83].

equations that are used to determine the pure component vapour pressure. This example attests to the physical robustness of the SAFT-γ approach to provide an accurate description of the mixtures of *n*-alkane+1-alkanols, over a wide range of thermodynamic conditions.

The performance of the SAFT-γ approach is also examined for other binary mixtures of *n*-alkanes and 1-alkanols. Such systems have also been studied with other group contribution approaches based on SAFT, namely the GC-SAFT-VR and GC-SAFT-0 formulations of Tamouza *et al.* [189] and the simplified PC-SAFT (sPC-SAFT) treatment with generalised parameters as presented by Grenner *et al.* [260]. It is important to note that in both of these methods, a homonuclear molecular model is used. Grenner *et al.* have used a generalised parameter approach for all of the parameters apart from the dispersion energy; the latter is obtained by estimation to experimental VLE data for each substance. Within both approaches, the unlike interaction parameters are approximated by means of combining rules (Lorentz-Berthelot) without subsequent correction. A comparison of the %AADs for the pressure and for the composition of the vapour phase (in cases where data

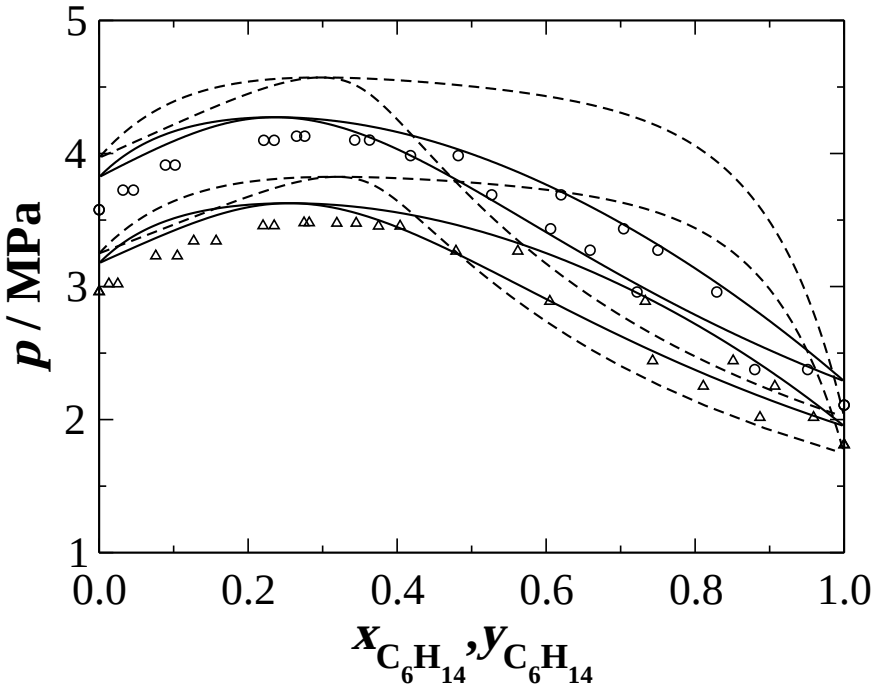

Figure 3.11: Prediction of the vapour-liquid phase behaviour of the binary mixture *n*-hexane+1-ethanol as pressure-composition isotherms at two different temperatures. The triangles represent the experimental data at 473.15 K, and the circles at 483.15 K [294]. The continuous curves are the SAFT-γ predictions modelling 1-alkanols with a CH_2OH functional group and the dashed curves are the predictions with the original UNIFAC approach using parameters from Hansen *et al.* [83].

for the latter are available) for a selection of the systems studied by Tamouza *et al.* and Grenner *et al.* is made in table 3.6. From the work of Tamouza *et al.* [189] only the performance of the GC-SAFT-VR is examined, since the comparisons are best suited to methods that are based on the same intermolecular potential. It is clear from the table that the overall predictive capability of the SAFT-γ method for these mixtures and range of conditions is very good and that in all cases the phase behaviour is predicted with deviations within a few percent. Moreover, the method is found to perform on average marginally better than the methods of Tamouza *et al.* [189] and Grenner *et al.* [260] for these systems.

Of further interest is the performance of the method for mixtures where at least one of the components was not included in the regression procedure for the determination of the parameters of the functional groups (recalling that in these cases the estimation is carried out with pure component VLE data alone). A few examples are presented in figures 3.12-3.14, where it can be seen that one is able to predict accurately the fluid phase behaviour of these binary mixtures. Apart from the vapour-liquid equilibria that has already been

presented for the binary mixtures of n-alkane+1-alkanol, the same parameter set can be used for the prediction of the liquid-liquid equilibria exhibited by highly asymmetric mixtures of this kind, i.e., mixtures where the components are very different in size. As an example, the predictions of the theory are compared with the liquid-liquid equilibrium data for the system of n-hexadecane+ethanol in figure 3.15. From the plot it can be seen that apart from the overprediction of the upper critical solution temperature, as is always the case for analytical classical approaches, the predictions of the theory are in good agreement with the experimentally determined phase behaviour of the system. It is important to stress that the same parameter set is used for the prediction of both the vapour-liquid and the liquid-liquid equilibria of these systems. In some cases, as for the original UNIFAC, different parameter tables have to be developed for the different types of fluid phase equilibrium [78].

The deviations of the predicted phase behaviour from the experimental data are summarised in table 3.7. The performance of the SAFT-γ method is compared with that of the GC-SAFT-VR approach of Tamouza *et al.* [189] and in the majority of the cases SAFT-γ is found to be slightly less accurate. A complete comparison cannot be established as the deviations of the GC-SAFT-VR method for the prediction of the composition of the vapour phase were not reported in [189].

Table 3.6: Comparison of the prediction of the vapour-liquid phase behaviour of mixtures of n-alkanes+1-alkanols with SAFT-γ and the predictive approaches of Tamouza *et al.* [189] and of Grenner *et al.* [260]. The values of the %AAD are obtained from eq. 3.12 and Δy from eq. 3.14.

					GC-SAFT-VR		pred. sPC-SAFT		SAFT-γ	
Mixture	Ref.	T [K]	n	p [Pa]	%AAD p	Δy	%AAD p	Δy	%AAD p	Δy
C_4H_9OH+				($8.0x10^3$)–						
C_7H_{16}	[295]	333.15	24	($3.1x10^4$)	7.33	0.97	7.9	1.7	4.29	2.26
$C_6H_{13}OH$+				($6.5x10^2$)–						
C_6H_{14}	[296]	323.15	23	($5.5x10^4$)	5.58	-	5.6	-	4.08	-
$C_8H_{17}OH$+				($2.5x10^3$)–						
C_8H_{18}	[297]	373.15	14	($4.7x10^4$)	5.78	-	6.7	-	5.26	-
$C_{10}H_{21}OH$+				($1.3x10^4$)–						
C_6H_{14}	[298]	323.15	10	($5.3x10^4$)	6.15	-	3.9	-	5.95	-
$C_{16}H_{33}OH$+				($7.9x10^5$)–						
C_6H_{14}	[299]	572.40	10	($4.3x10^6$)	7.13	2.37	7.3	1.16	6.66	1.40

3.3.3 Binary Mixtures: water+n-alkanes

Once the unlike group interaction parameters between H_2O and the CH_3 and CH_2 functional groups of the n-alkanes have been determined (as described in section 3.2.3), the transferability of the parameters can be assessed by predicting the fluid phase behaviour for a set of binary mixtures of n-alkanes and water. This includes systems of n-alkanes with a range of chain lengths, spanning from n-butane to n-hexadecane (n-C_4H_{10} to n-$C_{16}H_{34}$), over a wide range of conditions. The results of the predictions for the fluid phase equilibria of the selected mixtures are presented in figures 3.16-3.19, and the deviations from the experimental data are summarised in table 3.8. It is very gratifying that one is able to predict accurately Type III phase behaviour, according to the classification of van Konynenburg and Scott [303], as seen experimentally for the aqueous mixtures of alkanes. In general, the predictions are very good for both the VLE and the LLE regions of the systems examined as can be seen in figures 3.16-3.19; furthermore the position of the three phase line is accurately reproduced. Specific features of the phase behaviour shown in figures 3.16 and 3.19 are of particular interest. The first example is the $p-x$ slice of the phase diagram of the binary mixture water+n-butane at a temperature above the critical temperature of n-C_4H_{10} (cf. figure 3.16). The ability of the method to describe accurately

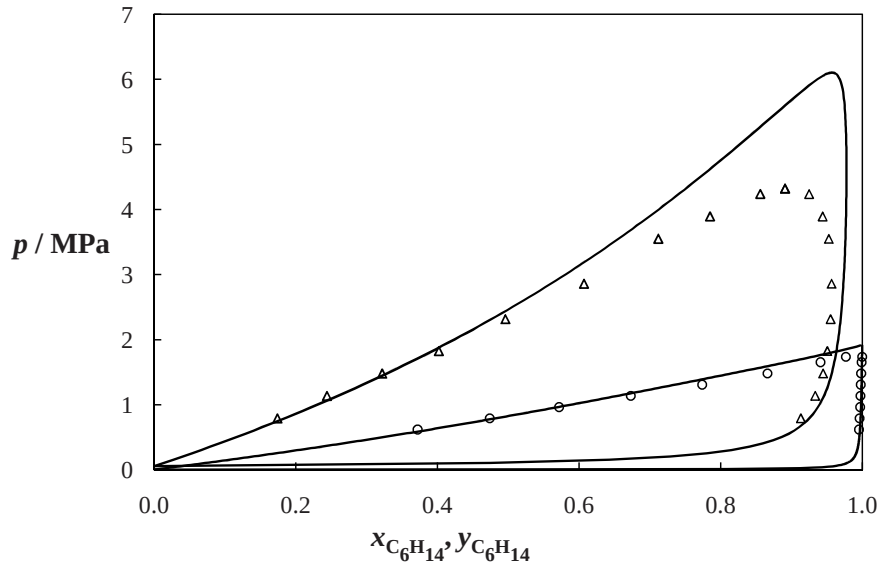

Figure 3.12: Prediction of the vapour-liquid phase behaviour of the binary mixture n-hexane+1-hexadecanol as pressure-composition isotherms at two different temperatures. The circles represent the experimental data at 472.1 K, and the triangles at 572.4 K [299]. The continuous curves are the SAFT-γ predictions modelling 1-alkanols with a CH_2OH functional group. The higher temperature is above the critical temperature of n-hexane ($T^{\text{exp}}_{\text{c},C_6H_{14}}$=507.82 K [277]), where an overprediction of the critical point of the mixture is noticeable.

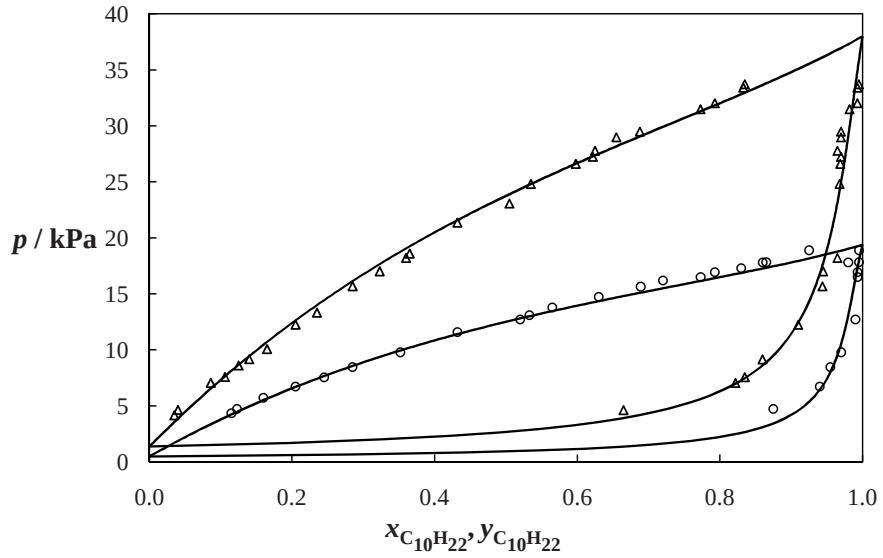

Figure 3.13: Prediction of the vapour-liquid phase behaviour of the binary mixture *n*-decane+1-dodecanol as pressure-compositions isotherms at two different temperatures. The circles represent the experimental data at 393.15 K, and the triangles at 413.15 K [300]. The continuous curves are the SAFT-γ predictions modelling 1-alkanols with a CH_2OH functional group.

the high-pressure fluid phase behaviour of the system demonstrates the wide range of reliable applicability in temperature and pressure of the SAFT-γ EoS compared to other

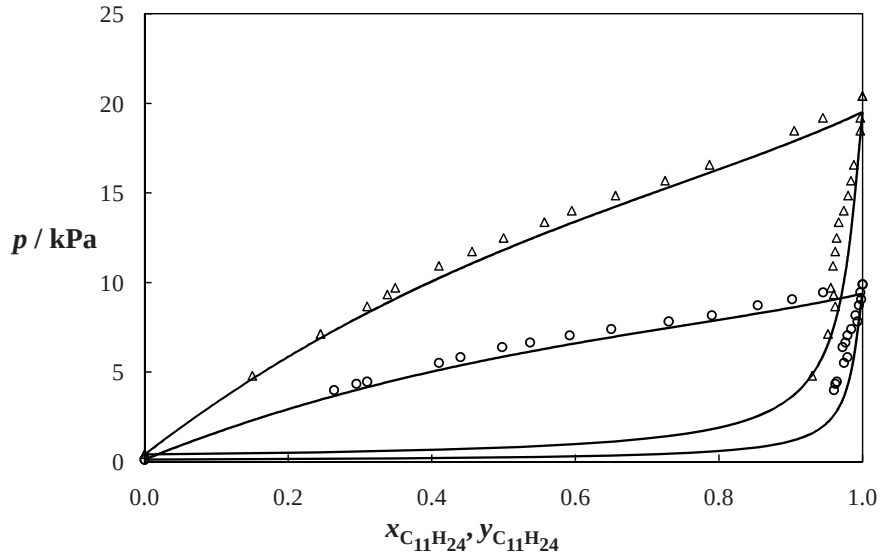

Figure 3.14: Prediction of the vapour-liquid phase behaviour of the binary mixture *n*-undecane+1-tetradecanol as pressure-composition isotherms at two different temperatures. The circles represent the experimental data at 393.15 K, and the triangles at 413.15 K [300]. The continuous curves are the SAFT-γ predictions.

predictive approaches, such as activity coefficient methods. Another feature of particular note is the water+n-hexadecane binary mixture depicted in figure 3.19; one should recall that n-$C_{16}H_{34}$ was not included in the regression procedure for the parameter estimation of the functional groups of the alkane series (the compounds considered were from C_2H_6 to n-$C_{10}H_{22}$ [19]). Nevertheless, the SAFT-γ GC method is still seen to describe the fluid phase behaviour of this mixture with great accuracy. This supports our previous assertion that the SAFT-γ EoS can be applied in a predictive manner to the calculation of the phase behaviour of mixtures when the unlike group interaction parameters are obtained from limited experimental data.

The models developed in this work are shown to provide with a very good description of the phase behaviour of a range of water+n-alkane binary mixtures over an extensive range of thermodynamic conditions. The performance of the theory so far has been assessed in the description of the vapour-liquid equilibria and the hydrocarbon-rich phase of the liquid-liquid equilibria that these systems exhibit. An example of the performance of the theory in the description of the compositions of the water-rich phase of the LLE region is shown in the inset image in figure 3.16. From the figure it can be seen that the predictions of the theory deviate from the experimental data by almost three orders of magnitude. This is a natural consequence of a modelling approach where both phases

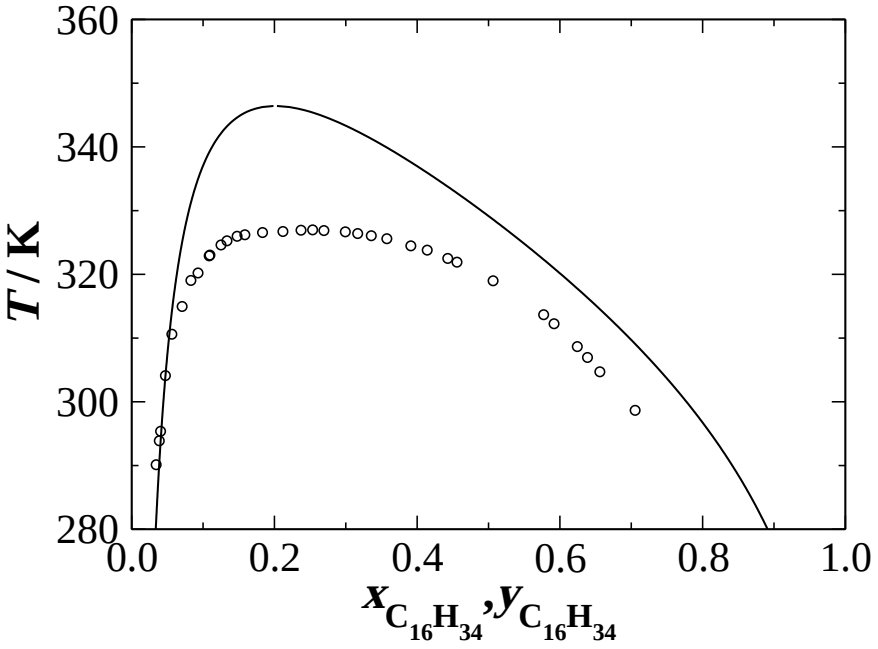

Figure 3.15: Prediction of the liquid-liquid phase behaviour of the binary mixture n-hexadecane+1-dodecanol as temperature-composition isobar at $p = 0.1013$ MPa. The circles represent the experimental data [301] and the continuous curves are the SAFT-γ predictions.

Table 3.7: Comparison of the prediction of the vapour-liquid phase behaviour of n-alkanes+1-alkanols binary mixtures with SAFT-γ and the GC approach of Tamouza *et al.*[189], where at least one of the compounds was not included in the parameter estimation procedure. The values of the %AAD are obtained from eq. 3.12 and Δy from eq. 3.14.

					GC-SAFT-VR		SAFT-γ	
Mixture	Ref.	T [K]	n	p [Pa]	%AAD p	Δy	%AAD p	Δy
$C_{12}H_{25}OH$+ $C_{10}H_{22}$	[302]	393.15	20	($4.3x10^3$)–($1.9x10^4$)	1.84	-	2.50	0.74
$C_{12}H_{25}OH$+ $C_{10}H_{22}$	[302]	393.15	25	($4.1x10^3$)–($3.3x10^4$)	3.25	-	3.51	0.68
$C_{12}H_{25}OH$+ $C_{14}H_{30}$	[302]	453.15	8	($8.6x10^3$)–($1.4x10^4$)	2.63	-	7.82	-
$C_{12}H_{25}OH$+ $C_{14}H_{30}$	[302]	473.15	8	($1.7x10^4$)–($2.6x10^4$)	2.45	-	7.06	-
$C_{14}H_{29}OH$+ $C_{11}H_{24}$	[300]	393.15	14	($3.9x10^3$)–($9.4x10^3$)	3.86	-	6.58	0.85
$C_{14}H_{29}OH$+ $C_{11}H_{24}$	[300]	413.15	15	($4.7x10^3$)–($1.9x10^4$)	2.45	-	4.38	0.86
$C_{16}H_{33}OH$+ C_6H_{14}	[299]	472.10	8	($6.1x10^5$)–($1.8x10^6$)	2.75	0.06	4.77	0.09

in coexistence (i.e., the water-rich and the hydrocarbon-rich phases) are treated with the same set of interaction parameters, despite their very different nature. This limitation can be overcome by employing different values of the interaction parameters for each phase, as shown by Haslam *et al.* [112].

As mentioned in the introduction, aqueous solutions of hydrocarbons have been studied previously by Soria et al. [258] with the GCA-EoS. The difference between the two approaches is that Soria *et al.* obtained the unlike group parameters for the mixture by correlating the experimental data for two mixtures of water with hydrocarbons (water+methane and water+ethane), while with SAFT-γ the unlike parameters are obtained from the data for a single mixure, namely water+n-hexane. The predictions of the two methods for the liquid-liquid equilibria of several binary mixtures of water and alkanes are compared in table 3.9. From the comparison it can be seen that both methods provide a comparable accuracy in the description of the fluid phase behaviour of these systems, with SAFT-γ performing slightly better.

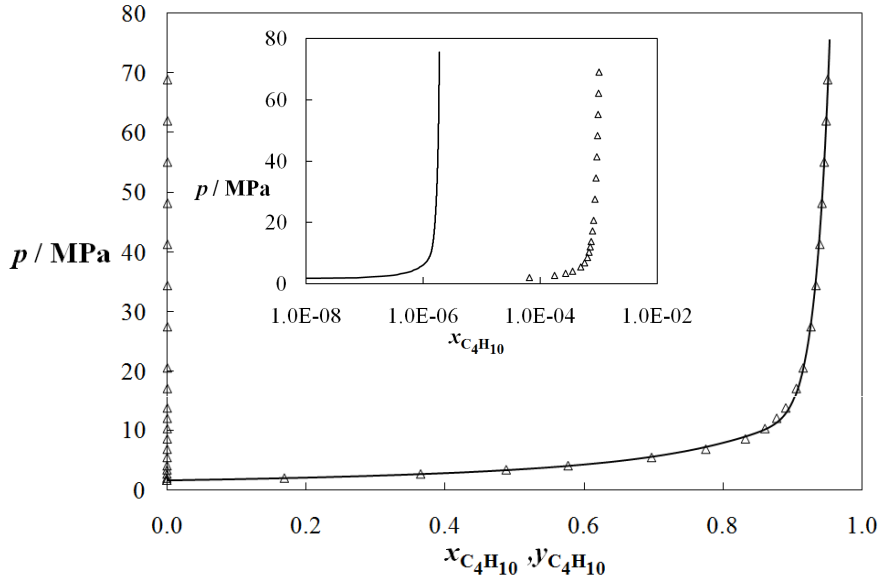

Figure 3.16: Prediction of the fluid phase behaviour of the binary mixture water+n-butane as a pressure-composition isotherm at 477 K, above the critical point of n-butane ($T_{c,C_4H_{10}}$=425.125 K [277]). The triangles represent the experimental data [304], and the continuous curves the corresponding SAFT-γ predictions. The predictions of the theory for the compositions of the water-rich phase can be seen in the inset image.

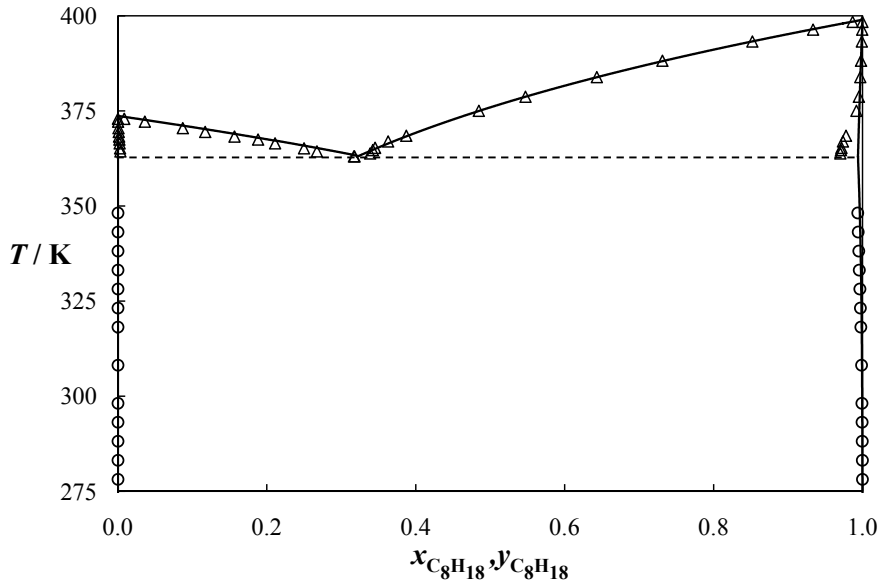

Figure 3.17: Prediction of the vapour-liquid and liquid-liquid phase behaviour of the binary mixture water+n-octane as a temperature-composition isobar at 101.3 kPa. The triangles represent the experimental VLE data [305], and the circles LLE data [304]. The continuous curves are the SAFT-γ predictions, and the dashed line denotes three-phase vapour-liquid-liquid coexistence line.

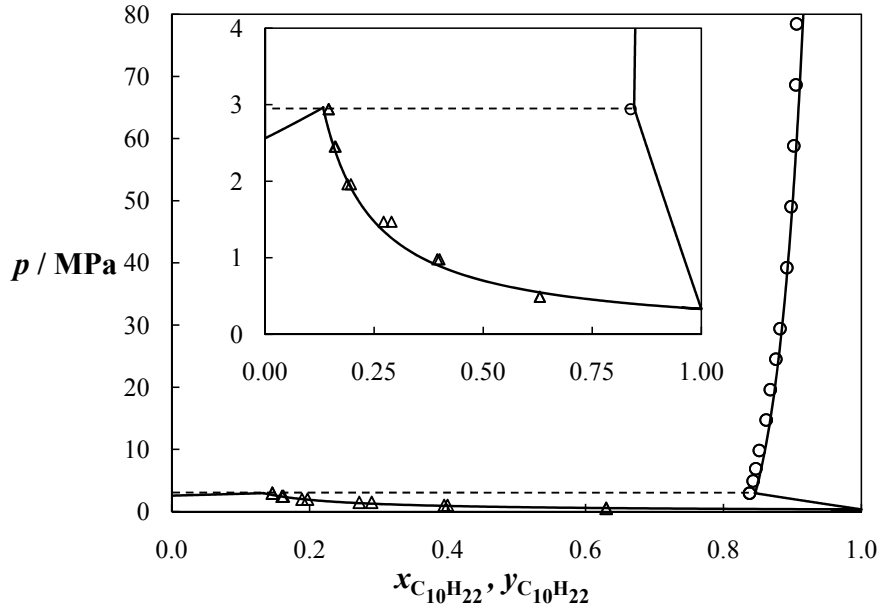

Figure 3.18: Prediction of the vapour-liquid and liquid-liquid phase behaviour of the binary mixture water+n-decane as a pressure-composition isotherm at 498.15 K. The triangles represent the experimental VLE data [281], and the circles LLE data [282]. The continuous curves are the SAFT-γ predictions, and the dashed line denotes the three-phase vapour-liquid-liquid coexistence line. A close-up of the vapour-liquid phase envelope and the three-phase region is shown in the inset image.

Table 3.8: SAFT-γ prediction of vapour-liquid and liquid liquid equilibria for binary water+n-alkane mixtures based on the transferable group interactions parameters estimated from the water+n-hexane binary mixture (cf. figure 3.7). Δx, Δy are calculated from eq. 3.14.

System	Data	Ref.	T or p	%AAD (p or T)	$\Delta($ y or $x_{\mathrm{I}})$	Δ (x or x_{II})
H_2O + C_6H_{14}	VLE	[281]	493.15 K	6.62	4.13	-
H_2O + C_6H_{14}	LLE	[282, 283]	493.15 K	-	0.42	0.46
H_2O + C_8H_{18}	VLE	[305]	101.3 kPa	0.048	0.91	-
H_2O + C_8H_{18}	LLE	[305]	101.3 kPa	-	0.12	3.18×10^{-5}
H_2O + C_8H_{18}	VLE	[282]	513.15 K	4.47	1.71	-
H_2O + C_8H_{18}	LLE	[282]	513.15 K	-	1.08	-
H_2O + $C_{10}H_{22}$	VLE	[281]	498.15 K	6.77	-	-
H_2O + $C_{10}H_{22}$	LLE	[282]	498.15 K	-	4.41	-
H_2O + $C_{16}H_{34}$	VLE	[306]	523.15 K	24.51	1.16	-
H_2O + $C_{16}H_{34}$	LLE	[282]	523.15 K	-	1.89	-
H_2O + $C_{16}H_{34}$	VLE	[306]	598.15 K	21.16	1.94	-
H_2O + $C_{16}H_{34}$	LLE	[282]	598.15 K	-	7.67	-

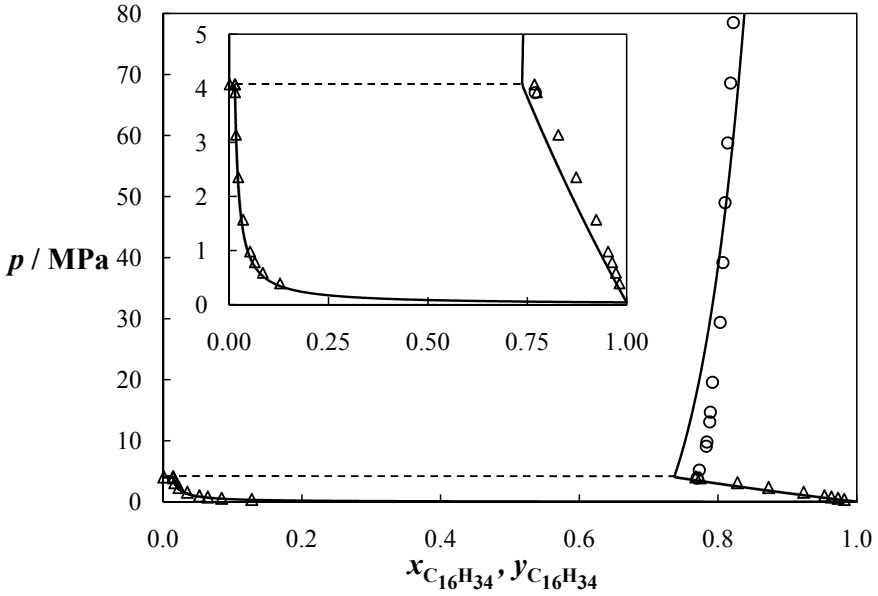

Figure 3.19: Prediction of the vapour-liquid and liquid-liquid phase behaviour of the binary mixture water+n-hexadecane as a pressure-composition isotherm at 523.15 K. The triangles represent the experimental VLE data [306], and the circles LLE data [282].The continuous curves are the SAFT-γ predictions, and the dashed line denotes the three-phase vapour-liquid-liquid coexistence line. A close-up of the vapour-liquid phase envelope and the three-phase region is shown in the inset image.

Table 3.9: SAFT-γ prediction of vapour-liquid and liquid liquid equilibria for binary water+n-alkane mixtures based on the transferable group interactions parameters compared to the corresponding results of Soria *et al.* [258] using the GCA-EoS. Δx and Δy are calculated from eq. 3.14.

				GCA-EoS [258]		SAFT-γ	
System	Ref.	T [K]	p [kPa]	Δx_{I}	Δx_{II}	Δx_{I}	Δx_{II}
H_2O + C_3H_8	[307, 308]	288-370	567-4398	6.00×10^{-2}	5.00×10^{-1}	3.06×10^{-2}	9.49×10^{-1}
H_2O + C_4H_{10}	[308, 309]	298-353	531-1059	1.60×10^{-2}	2.40	5.85×10^{-2}	1.00
H_2O + C_5H_{12}	[308, 310]	273-343	101-508	2.50×10^{-3}	5.00×10^{-2}	1.10×10^{-3}	3.38×10^{-3}
H_2O + C_6H_{14}	[308, 311, 312]	273-423	101-1255	1.60×10^{-3}	1.70	3.02×10^{-2}	1.14×10^{-1}
H_2O + C_8H_{18}	[313]	311-539	10-7410	5.00×10^{-3}	6.00	9.83×10^{-3}	5.07

3.3.4 Binary Mixtures: water+1-alkanols

The unlike interaction parameters between the CH_3 and CH_2 functional groups of the n-alkanes and water developed as described in section 3.2.3 can be transferred to the study of aqueous solutions of 1-alkanols. These systems pose an interesting challenge as they exhibit features in the fluid phase behaviour which are determined by the relative magnitudes of the unlike dispersion energy and hydrogen-bonding interactions. In a similar manner to the water+n-alkanes family, the aqueous alkanol systems feature both vapour-liquid and liquid-liquid equilibria, and heterogeneous azeotropy. The simultaneous description of these systems using a unique parameter set constitutes a stringent test.

The performance of SAFT-γ, using the cross interaction parameters obtained as described section 3.2.4, is assessed for the water+1-butanol and water+1-hexanol mixtures (figures 3.20 and 3.21). It can be seen that the predictions of SAFT-γ are in good agreement with the experimental fluid phase behaviour of these systems. Both the VLE and the LLE regions of the phase envelope are predicted well, together with the position of the three-phase vapour-liquid-liquid coexistence line. This is achieved while using a unique set of transferable interaction parameters for the water+n-alkane and water+1-alkanol systems. The overall accuracy of the SAFT-γ approach in describing the experimental fluid phase equilibria of the aqueous solutions of 1-alkanols is summarised in table 3.10 and is clearly seen to be good. The limited experimental data available for systems of this type does not allow for an extensive study of long-chain alkanols, as in the study of aqueous solutions of the alkanes.

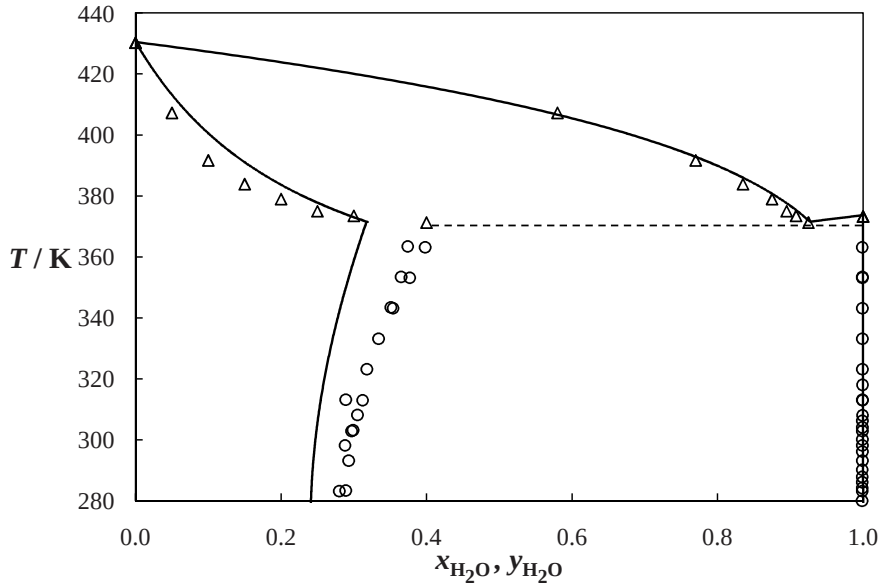

Figure 3.20: Prediction of the vapour-liquid and liquid-liquid phase behaviour of the binary mixture water+1-hexanol as a temperature-composition isobar at 101.3 kPa. The triangles represent the experimental VLE data [314], and the circles the LLE data [315]. The continuous curves are the SAFT-γ predictions, and the dashed line denotes the three-phase vapour-liquid-liquid coexistence line.

Table 3.10: SAFT-γ predictions of vapour-liquid and liquid-liquid equilibria for binary water+1-alkanol mixtures based on the transferable group interactions parameters. Δx, Δy are calculated from eq. 3.14.

System	Data type	Ref.	p	%AAD T	Δy or Δx_{I}	Δx or Δx_{II}
H_2O + C_4H_9OH	VLE	[316]	101.3 kPa	1.50	9.018	-
H_2O + C_4H_9OH	LLE	[317]	101.3 kPa	-	2.131	0.735
H_2O + $C_6H_{13}OH$	VLE	[314]	101.3 kPa	1.31	4.633	-
H_2O + $C_6H_{13}OH$	LLE	[315]	101.3 kPa	-	5.777	0.104
H_2O + $C_7H_{15}OH$	LLE	[315]	101.3 kPa	-	8.732	0.049
H_2O + $C_8H_{17}OH$	LLE	[315]	101.3 kPa	-	15.201	0.016
H_2O + $C_9H_{19}OH$	LLE	[315]	101.3 kPa	-	12.483	0.002
H_2O + $C_{10}H_{21}OH$	LLE	[315]	101.3 kPa	-	14.064	4×10^{-4}
H_2O + $C_{11}H_{23}OH$	LLE	[315]	101.3 kPa	-	14.165	-
H_2O + $C_{12}H_{25}OH$	LLE	[315]	101.3 kPa	-	13.718	1×10^{-5}

At this point, we should mention that the parameters developed here for the family of 1-alkanols can be generally applied to the prediction of pure component properties and mixtures containing 1-alkanols for the range studied (ethanol to 1-decanol) and extrapolated

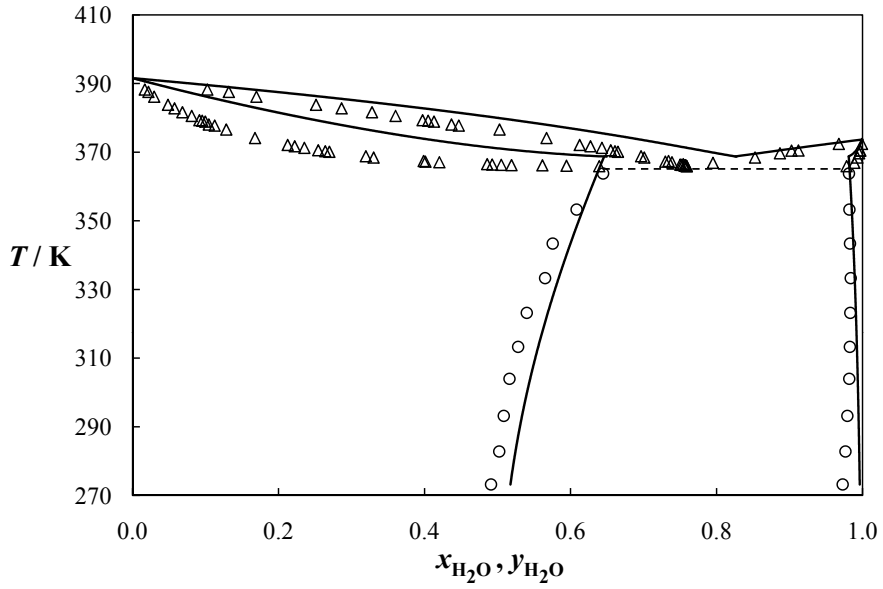

Figure 3.21: Prediction of the vapour-liquid and liquid-liquid phase behaviour of the binary mixture water+1-butanol as a temperature-composition isobar at 101.3 kPa. The triangles represent the experimental VLE data [316], and the circles LLE data [317]. The continuous curves are the SAFT-γ predictions, and the dashed line denotes the three-phase vapour-liquid-liquid coexistence line.

to longer compounds. It is important to note, however, that when it comes to modelling aqueous solutions of the shorter 1-alkanols (i.e., methanol, ethanol, and 1-propanol), the parameters presented here can lead to large deviations from the experimental data. The singular physicochemical features of these smaller molecules, such as pronounced polarisability effects lead to full miscibility in water, and must therefore be studied individually. The global fluid phase equilibria of generic model mixtures of water+1-alkanols has been studied with a mean-field version of the SAFT approach (SAFT-HS), where a detailed examination of the effect of the strength of the hydrogen-bonding interactions and chain length on the specific type of fluid phase behaviour that is observed was made [318, 319]. It is clear from the findings of this study that one cannot simultaneous reproduce the fluid phase behaviour of aqueous solutions of the shorter (VLE) and longer (VLE and LLE) with a single set of transferable intermolecular parameters [319].

3.3.5 **Ternary Mixtures:** water+n-alkane+1-alkanol

The SAFT-γ GC EoS can be used to describe ternary systems once all of the pair group interactions of the mixture have been determined. The SAFT-γ prediction of the fluid phase behaviour of two water+n-alkane+1-alkanol ternaries are shown in figures 3.22 and figure 3.23. In both cases reasonable agreement with the experimental data is observed. In

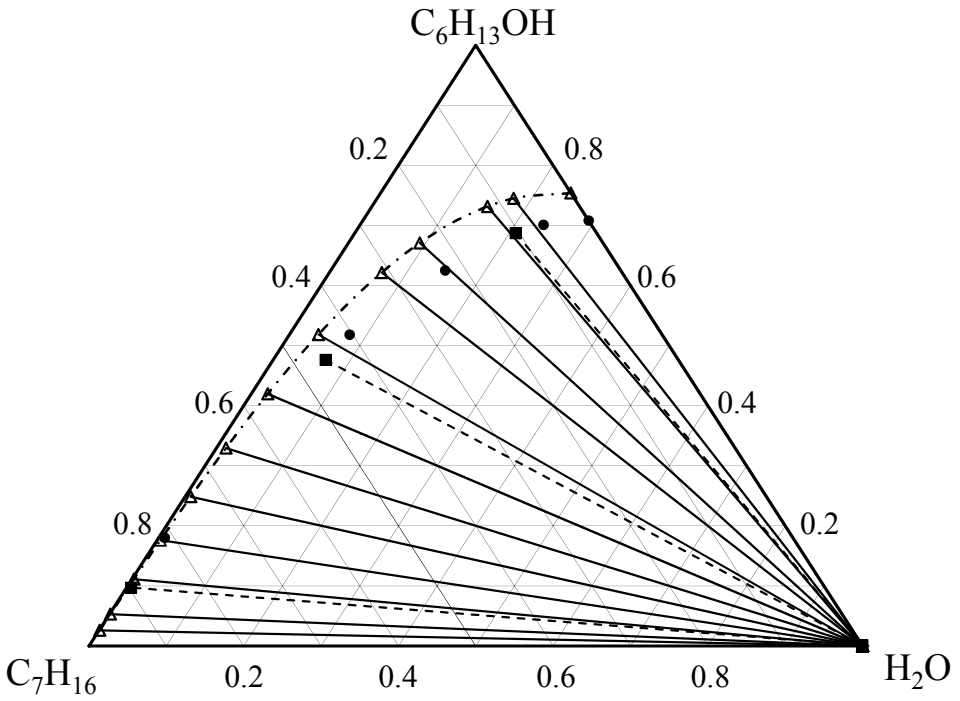

Figure 3.22: Prediction of the fluid phase behaviour of the ternary mixture water+1-heptane+1-hexanol at 101.3 kPa and 298.2 K. The circles represent the experimental data on the coexistence curve, the squares compositions of the conjugate solutions, and the dashed lines the experimental tie lines [320]. The continuous lines are the SAFT-γ predictions of the tie-lines, the triangles are the predicted compositions, and the dashed-dotted curve is the predicted coexistence curve.

the case of the water+*n*-hexane+1-octanol ternary mixture (figure 3.23) the deviation in the prediction of the boundary of the two-phase region is slightly more pronounced. This is due to the higher deviation observed in the prediction of the liquid-liquid equilibria for the water+1-octanol binary mixture at these conditions. Further improvement of the description of the LLE in this particular binary mixture is needed in order to provide a predictive capability for the determination of the water-octanol partition coefficient which is of particular practical importance.

3.4 Challenges

Based on the results presented thus far, the performance of the SAFT-γ methodology in the prediction of the phase behaviour of a wide range of systems, including the challenges presented by aqueous solutions of hydrocarbons, can be deemed very satisfactory. However, there are certain features that are not very accurately reproduced. A first issue is the description of fluid phase behaviour in the vicinity of the critical point. An example

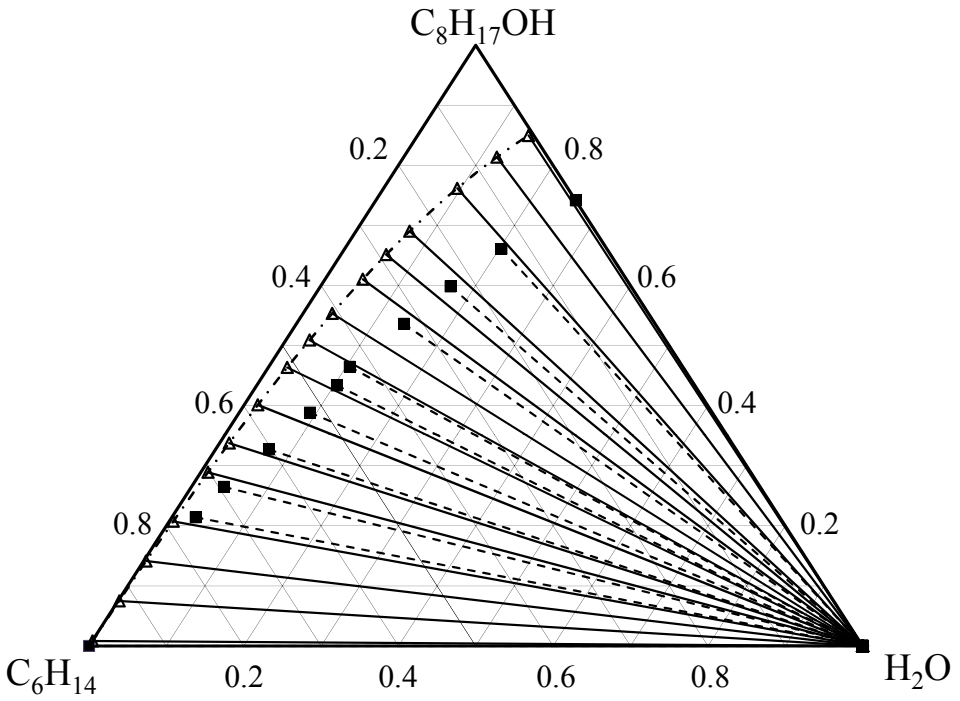

Figure 3.23: Prediction of the fluid phase behaviour of the ternary mixture water+1-hexane+1-octanol at 101.3 kPa and 293.15 K. The squares represent the compositions of the conjugate solutions, and the dashed lines the experimental tie lines [321]. The continuous lines are SAFT-γ predictions of the tie-lines, the triangles are the predicted compositions, and the dashed-dotted curve is the predicted coexistence curve.

of the performance of the SAFT-γ method obtained in the description of fluid phase behaviour for near-critical conditions and critical points of mixtures is shown on figure 3.24, where the predictions of the theory are compared to the experimentally determined phase behaviour of the binary n-butane+n-decane mixture for three isotherms, two of which are above the critical temperature of n-butane ($T^{\text{exp}}_{\text{c,C}_4\text{H}_{10}} \approx 425$ K). It is apparent that whereas for the lowest isotherm ($T = 377.59$ K) the theory describes the vapour-liquid phase behaviour of the system very accurately, for the isotherms at higher temperatures, the overprediction of the critical point of the mixture at high concentrations on n-butane is significant. The same applies for the case of liquid-liquid equilibria critical points, as shown for the case of the binary system of n-hexadecane+ethanol in figure 3.15.

An accurate representation of the near critical region is a known limitation of classical EoS of this type requiring the application of a crossover treatment in order to account for the energetic, density and compositional fluctuations that arise. Such treatments have been applied previously to a wide range of thermodynamic approaches, from cubic equations of state to SAFT-type theories; the most relevant to the current study is the work of Mc-

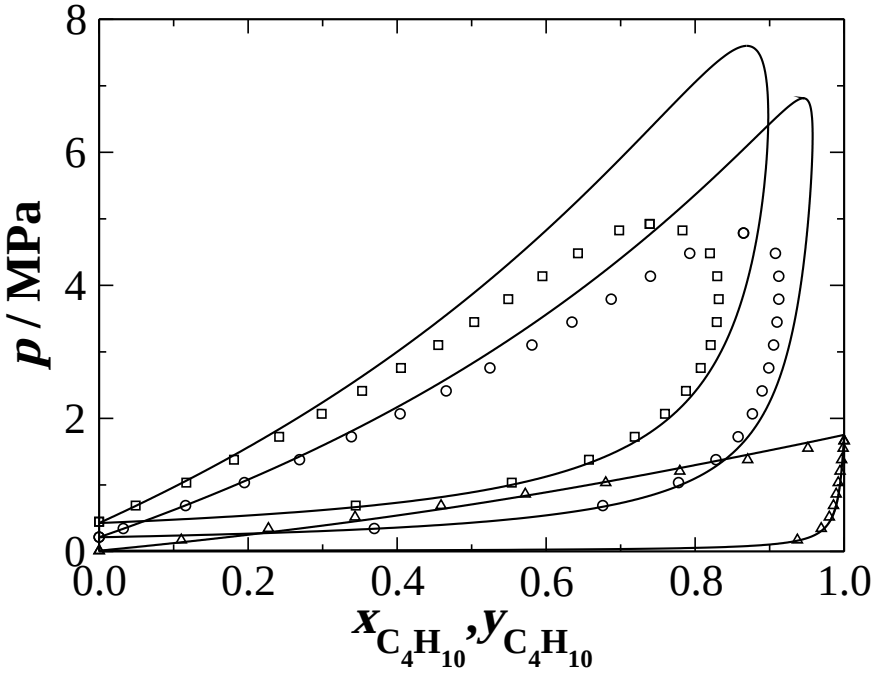

Figure 3.24: Pressure-composition ($p-x$) representation of the vapour-liquid phase behaviour of a binary mixture of n-butane + n-decane. The continuous curves represent the predictions of the SAFT-γ approach, and the symbols the experimental data at 377.59 K (triangles), 477.59 K (circles) and 510.93 K (squares) [322].

Cabe and Kiselev [323] and Forte *et al.* [324, 325] where a crossover treatment is applied to SAFT type equations with interaction potentials of variable range. Nevertheless, the improvement of the existing description of the near-critical phase behaviour of fluids and fluid mixtures whilst retaining an analytical theory still remains very challenging.

Another challenge is revealed when examining the performance of the SAFT-γ methodology in the prediction of single-phase thermodynamic properties other than the phase behaviour, e.g., the speed of sound and the heat capacities (isobaric and isochoric). The accurate representation of such properties, generally referred to as second-order thermodynamic derivative properties as they are obtained as second derivatives of the free energy, is a stringent test for most thermodynamic approaches, including sophisticated SAFT-type equations of state. An example of the performance of the SAFT-γ approach based on the square-well potential in the prediction of derivative properties is presented on figure 3.25, where the theoretical description is compared with the experimental data for the isothermal compressibility, k_T, and the speed of sound, u, for two linear alkanes, n-eicosane and n-pentadecane. It can be seen that, especially in the case of long-chain molecules, high errors are to be expected in the description of derivative properties, with parameters ob-

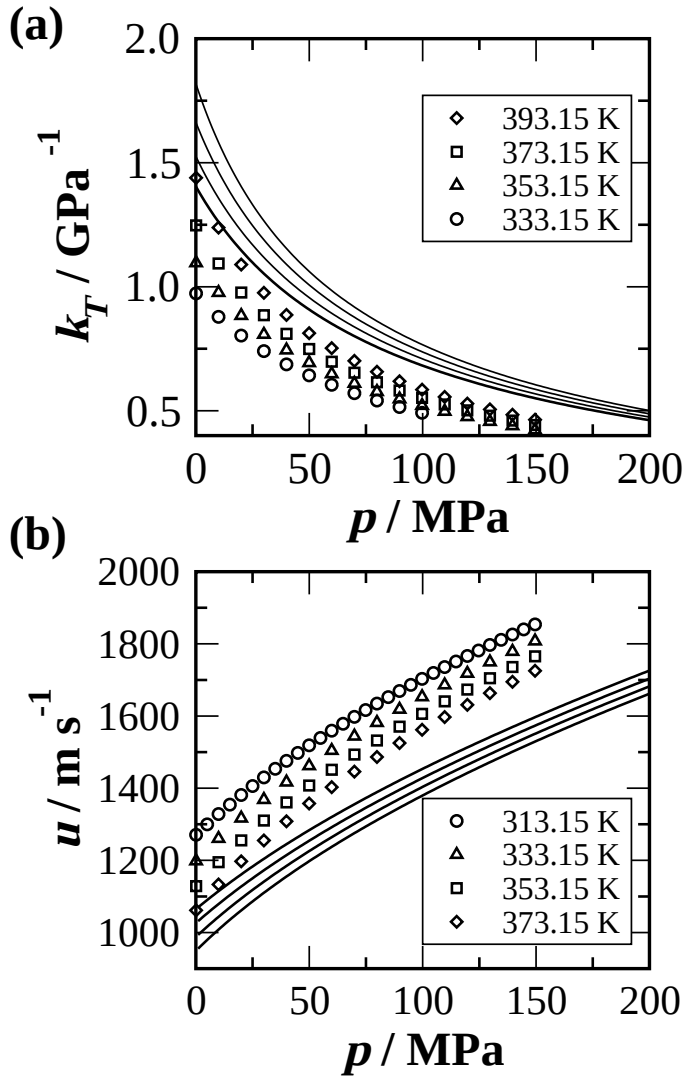

Figure 3.25: Predictions of the SAFT-γ approach compared with the experimental data for thermodynamic derivative properties of long-chain n-alkanes as a function of pressure at different temperatures: (a) Isothermal compressibility of n-eicosane [326]; and (b) speed of sound of n-pentadecane [327].

tained from data for the fluid phase behaviour.

The description of derivative properties with equations of state is a subject that has been attracting attention in recent years as attested by the number of topical publications on this subject, e.g., see references [30, 263, 328–332]. Focusing on the description of derivative properties by means of SAFT-type EoSs, Lafitte *et al.* [30, 263] have compared various approaches and highlighted the importance of using generic intermolecular potentials, such as the Mie potential (generalised Lennard-Jonesium potential of variable attractive and repulsive range), in accurately describing phase behaviour and derivative properties.

3.5 Concluding Remarks

In this chapter the extension of the SAFT-γ GC approach based on the square-well group-group interactions to the study of aqueous solutions of hyrdocabrons and 1-alkanols is presented. The functional group for the modelling of the chemical family of 1-alkanols is revisited following the ideas of Wu and Sandler [107], where a CH_2OH functional group is considered. The application of the CH_2OH group is shown to give a description of the pure-component fluid phase equilibria of the same quality as the previously employed OH group [19], but leads to a significantly improved description of the phase behaviour of binary mixtures of n-alkanes+1-alkanols. The robustness of the SAFT-γ approach is demonstrated in predictions of vapour-liquid and liquid-liquid equilibria for selected binary mixtures of n-alkanes+1-alkanols over wide ranges of thermodynamic conditions, with parameters obtained from pure component data alone. The level of accuracy of SAFT-γ is shown to be comparable to the well-established UNIFAC method.

Subsequently, the SAFT-γ EoS is applied in the description of the fluid phase equilibria of aqueous solutions, where the necessary interaction parameters are estimated from limited experimental data. The method is shown to accurately describe the highly non-ideal fluid phase behaviour that is exhibited by systems of this kind, including extensive regions of liquid-liquid immiscibility over wide pressure ranges and heterogeneous azeotropes (three-phase equilibria) at lower pressures. The accurate description of the phase behaviour for aqueous systems of this kind is a challenge for most thermodynamic methodologies, as the complex phase behaviour encountered is related to the strong association interactions (hydrogen bonding) in water. Regarding the predictive capability of the SAFT-γ approach, it is shown that a unique set of group interaction parameters can be successfully applied to the study of the fluid phase behaviour of binary mixtures of water+n-alkane for a broad range of chain lengths. In the case of water+1-alkanols, the deviations from the experimental fluid phase behaviour is the highest of the mixtures studied. The physical robustness of the estimated parameters is examined by studying the liquid-liquid equilibria of ternary mixtures of water+n-alkane+1-alkanol, where SAFT-γ is shown to provide a satisfactory performance for the systems considered.

Despite the impressive overall performance of the SAFT-γ approach in describing aqueous mixtures, two main challenges are identified, namely the improvement of the theory in the representation of the near-critical region within the scope of an analytical theory, and the improvement in the prediction of thermodynamic derivative properties. The performance of SAFT-γ (as implemented in this chapter) in the description of derivative

properties is a result of the simplified intermolecular potential (square-well) employed within the theory. The challenges highlighted here call for the development of a new theory employing a more realistic intermolecular potential for the description of the repulsive and attractive interactions between monomeric segments. The development of this new theory is presented in detail in the next chapter.

Chapter 4

SAFT-γ group contribution methodology for heteronuclear molecular models based on Mie (generalised Lennard-Jones) segments

Until recently the focus of a thermodynamic treatment has been mainly on the description of a limited range of properties, and specifically fluid phase behaviour. Over the past decade there has been an increasing number of publications on the performance of thermodynamic methods in the description of second-order derivative properties, e.g., the speed of sound, heat capacity, or isothermal compressibility. Derivative properties are properties of interest from a practical perspective; a known example is the importance of the Joule-Thomson inversion curve in the Linde technique as a standard process in the petrochemical industry. A precise description of such properties is however highly challenging from a theoretical perspective [333]. The performance of traditional cubic EoSs in the description of derivative properties has been shown to be relatively poor [333, 334], as these approaches fail to reproduce a number of the singularities that derivative properties exhibit, such as the density extrema in isothermal variations of the isochoric heat capacity.

An accurate representation of derivative properties is a stringent test even for a sophisticated thermodynamic treatment, such as the SAFT-type EoSs. One of the first detailed studies of the performance of SAFT-type methods in the description of such properties, is the work of Colina *et al.* [328]. The authors examined the Joule-Thomson inversion

curve as calculated with the soft-SAFT EoS [169, 335] and found that in the majority of cases only quantitative agreement is achieved. In a later study the description with several SAFT-variants (soft-SAFT [169, 335], SAFT-VR [170, 171], PC-SAFT [166]) were compared, and deviations of up to 20% from the experimental values were found in the description of the Joule-Thomson inversion curve [336]. Llovell *et al.* [337] have presented a detailed analysis on the performance of the soft-SAFT EoS for the derivative properties of pure components and mixtures [337] concluding that the accuracy of the description can range from 1 to 20% depending on the compound and the property studied, with the highest deviations typically obtained for the speed of sound of long chain *n*-alkanes. Another variant of SAFT that has been applied to the prediction of thermodynamic derivative properties is the SAFT-BACK EoS [338]. Manghari and co-workers [339, 340] have shown that the SAFT-BACK EoS can be used to represent the derivative properties of pure *n*-alkanes and mixtures of these with an accuracy of less than 5%; however, this is accomplished at the cost of introducing an empirical temperature dependence of the dispersion energy by means of a compound-specific constant. In more recent work, Diamantonis and Economou [330] have compared the PC-SAFT EoS [166] and the SAFT EoS of Huang and Radosz [167] in the description of the derivative properties for a number of carefully selected compounds, showing that PC-SAFT performs systematically better, with an accuracy within 10%. A comprehensive comparative study of the performance of SAFT-VR SW [170, 171], PC-SAFT [166] and SAFT-VR LJC [261] EoSs has also been presented and compared to the results obtained with the SAFT-VR Mie EoS developed in 2006 [30, 263]. In their study, Lafitte and co-workers [30, 263] showed that the versatility of the Mie (generalised Lennard-Jones) potential, on which the SAFT-VR Mie EoS is based, allows for a significant improvement in the description of derivative properties when compared to the other SAFT variants and popular cubic EoSs.

Lafitte *et al.* [31] have recently presented a new version of the SAFT-VR Mie EoS, where several approximations inherent in the underlying theory of the EoS presented in 2006 [30, 263] were revisited and improved. The new theory, which was validated by comparison with molecular simulations, was shown to constitute a very accurate methodology for the simultaneous description of the fluid phase behaviour and thermodynamic derivative properties of pure substances and mixtures. In view of the success of the latest SAFT-VR Mie EoS in describing the thermodynamic properties of real systems, the main focus of the work presented in this thesis is the reformulation of the SAFT-VR Mie EoS as a group contribution approach; SAFT-γ Mie.

4.1 Molecular model and intermolecular potential

In the SAFT-γ Mie group contribution approach, molecules are represented as *fused* heteronuclear models formed from Mie segments. Originally, the SAFT-γ method was developed for molecules formed from square-well (SW) segments [19, 29], as discussed in chapter 3 (see also figure 4.1.(a)). A description of the segment-segment interactions with a potential of the Mie form has been used for homonuclear models within the SAFT-VR framework by Davies *et al.* [261], as well as in the development of the SAFT-VR Mie EoS by Laffite *et al.* [30, 263]. The main difference between these approaches is in the treatment of the radial distribution function (RDF) of the Mie fluid, which in the approach of Davies *et al.* is approximated as the RDF of the Sutherland-6 potential. The accuracy of the representation of the RDF of the monomer reference system is of great importance in the development of EoSs based on the first-order thermodynamic perturbation theory of Wertheim (TPT1) [101–104], as the correlation functions are employed in the determination of the free energy of the formation of chain molecules and the treatment of the contribution due to association.

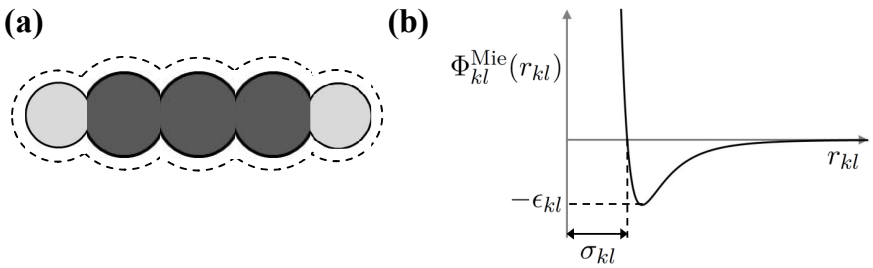

Figure 4.1: Pictorial representation of: (a) the *fused heteronuclear* molecular model; and (b) the intermolecular potential employed within the framework of the SAFT-γ Mie approach.

The Mie intermolecular potential [341], depicted on figure 4.1.(b), is a generalised Lennard-Jonnesium potential [342–344], where the attractive and the repulsive exponents which characterise the softness/hardness and the range of the interaction are allowed to vary freely. The pair interaction energy between segments k and l as a function of the interesegment distance r_{kl} is given by

$$\Phi_{kl}^{\mathrm{Mie}}(r_{kl}) = \mathcal{C}_{kl}\epsilon_{kl}\left[\left(\frac{\sigma_{kl}}{r_{kl}}\right)^{\lambda_{kl}^{r}} - \left(\frac{\sigma_{kl}}{r}\right)^{\lambda_{kl}^{a}}\right] , \tag{4.1}$$

where σ_{kl} is the segment diameter, ϵ_{kl} is the depth of the potential well, and λ_{kl}^{r} and λ_{kl}^{a} the repulsive and attractive exponents of the unlike interactions. The prefactor $\mathcal{C}_{kl}$ is a function of the exponents of the potential and ensures that the minimum of the interaction

is $-\epsilon_{kl}$:

$$\mathcal{C}_{kl} = \frac{\lambda_{kl}^r}{\lambda_{kl}^r - \lambda_{kl}^a} \left(\frac{\lambda_{kl}^r}{\lambda_{kl}^a} \right)^{\frac{\lambda_{kl}^a}{\lambda_{kl}^r - \lambda_{kl}^a}} \quad . \tag{4.2}$$

In common with other SAFT approaches, additional short range association sites can be placed on the molecular segments to mimic the association (hydrogen bonding) interactions present in some polar compounds. More specifically the association interactions are modelled by means of square-well sites, so that the interaction between a site of type a placed on a segment of type k and a site of type b placed on a segment of type l is given by

$$\Phi_{kl,ab}^{\mathrm{HB}}(r_{kl,ab}) = \begin{cases} -\epsilon_{kl,ab}^{\mathrm{HB}} & \text{if } r_{kl,ab} \leq r_{kl,ab}^{\mathrm{c}} \ , \\ 0 & \text{if } r_{kl,ab} > r_{kl,ab}^{\mathrm{c}} \ , \end{cases} \tag{4.3}$$

where $r_{kl,ab}$ is the centre-centre distance between sites a and b, $-\epsilon_{kl,ab}^{\mathrm{HB}}$ is the association energy, and $r_{kl,ab}^{\mathrm{c}}$ the cut-off range of the interaction between sites a and b.

Within the SAFT-γ Mie group contribution approach molecular properties are obtained by subdividing the molecules into distinct functional groups chosen to represent the chemical structure of a molecule, with appropriate summations over the contributions of all of the functional groups. A functional group can comprise one or multiple identical Mie segments described by the same set of group parameters, as in the case of the SAFT-γ approach based on SW segments [29]. The parameters that fully describe a functional group k are the number ν_k^* of identical segments that the group comprises, the segment diameter σ_{kk} of the segments of the group, the energy of interaction ϵ_{kk} between the segments of the group, and the values λ_{kk}^r and λ_{kk}^a of the repulsive and attractive exponents, respectively, that determine the form of the interaction potential. The extent to which the segments of a given group k contribute to the overall molecular properties is characterised with a key parameter of the methodology: the shape factor S_k. In the case of associating groups, the number NST_k of the different site types, the number of sites of each type, e.g., $n_{k,a}$, $n_{k,b}$, ..., $n_{k,\mathrm{NST}}$, together with the position $r_{kl,ab}^d$ of the site, and the energy $\epsilon_{kl,ab}^{\mathrm{HB}}$ and range $r_{kl,ab}^c$ of the association between different sites have to be determined.

4.2 SAFT-γ Mie

In the SAFT-γ Mie EoS the total Helmholtz free energy A of a mixture of heteronuclear associating molecules formed from Mie segments is written as the sum of four separate

contributions

$$\frac{A}{Nk_{\mathrm{B}}T} = \frac{A^{\mathrm{ideal}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{mono.}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{chain}}}{Nk_{\mathrm{B}}T} + \frac{A^{\mathrm{assoc.}}}{Nk_{\mathrm{B}}T} \ , \tag{4.4}$$

where A^{ideal} is the free energy of the ideal gas, $A^{\mathrm{mono.}}$ is the term accounting for interactions between monomeric Mie segments, A^{chain} is the contribution to the free energy for the formation of molecules from Mie segments, $A^{\mathrm{assoc.}}$ is the term accounting for the association interactions, N is the total number of molecules, k_{B} the Boltzmann constant, and T the absolute temperature. In the following sections, the separate contributions to the free energy will be discussed in detail.

4.2.1 Ideal Term

The ideal contribution to the free energy of the mixture is described by means of the following expression [264]:

$$\frac{A^{\mathrm{ideal}}}{Nk_{\mathrm{B}}T} = \left(\sum_{i=1}^{N_{\mathrm{C}}} x_i \ln \left(\rho_i \Lambda_i^3 \right) \right) - 1 \ , \tag{4.5}$$

where x_i is the mole fraction of component i in the mixture, $\rho_i = N_i/V$ is the number density of component i, N_i being the number of molecules of component i. The summation in eq. (4.5) is over all of the components NC present in the mixture. The ideal free energy incorporates the effects of the translational, rotational and vibrational contributions to the kinetic energy implicitly in the thermal de Broglie volume, Λ_i^3.

4.2.2 Monomer Term

The monomer term $A^{\mathrm{mono.}}$ describes the contribution of the reference monomeric Mie segments to the total Helmholtz free energy of the system. This contribution is obtained by applying the Barker and Henderson high-temperature perturbation theory [265] to third order. The intermolecular potential of eq. (4.1) is first decomposed into the sum of a reference repulsive contribution $\Phi_0(r_{kl})$, and a perturbation attractive contribution $\Phi_1(r_{kl})$:

$$\Phi_{kl}^{\mathrm{Mie}}(r_{kl}) = \Phi_0(r_{kl}) + \Phi_1(r_{kl}) \ , \tag{4.6}$$

where

$$\Phi_0(r_{kl}) = \begin{cases} \Phi_{kl}^{\mathrm{Mie}}(r_{kl}) & \text{if } r_{kl} < \sigma_{kl} \ , \\ 0 & \text{if } r_{kl} \geq \sigma_{kl} \ , \end{cases} \tag{4.7}$$

and

$$\Phi_1(r_{kl}) = \begin{cases} 0 & \text{if } r_{kl} \leq \sigma_{kl} \ , \\ \Phi_{kl}^{\text{Mie}}(r_{kl}) & \text{if } r_{kl} > \sigma_{kl} \ . \end{cases} \tag{4.8}$$

Barker and Henderson [269] have shown that the free energy of the system can then be obtained as a perturbation expansion in the inverse temperature relative to the reference system. In the case of soft core potentials, such as the Mie potential, the reference system can be approximated as that of an equivalent system of hard spheres with an effective diameter d_{kk}, since the properties of the reference system described by the potential in eq. (4.7) are not generally known. Following the application of the Barker and Henderson perturbation theory as employed within the recent SAFT-VR Mie EoS [31], the monomer term of the Helmholtz free energy is expressed as a third-order expansion. Having established the hard-sphere system as the reference system, the high-temperature perturbation expansion can be expressed as

$$\frac{A^{\text{mono.}}}{Nk_{\text{B}}T} = \frac{A^{\text{HS}}}{Nk_{\text{B}}T} + \frac{A_1}{Nk_{\text{B}}T} + \frac{A_2}{Nk_{\text{B}}T} + \frac{A_3}{Nk_{\text{B}}T} \ , \tag{4.9}$$

where A^{HS} is the free energy of the hard-sphere reference system of diameter d_{kk}. For a given group k, the effective hard-sphere diameter is obtained from [269]

$$d_{kk} = \int_0^{\sigma_{kk}} \left[1 - \exp\left\{ -\frac{\Phi_{kk}^{\text{Mie}}(r_{kk})}{k_{\text{B}}T} \right\} \right] dr \ . \tag{4.10}$$

The integral of eq. (4.10) is obtained by means of the Gauss-Legendre quadrature, a technique previously employed by Paricaud [345] who showed that a 5-point Gauss-Legendre procedure is adequate for an accurate representation of the effective hard-sphere diameter d_{kk}.

The hard-sphere Helmholtz free energy of the mixture is given by [19]

$$\frac{A^{\text{HS}}}{Nk_{\text{B}}T} = \left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i} \nu_k^* S_k \right) a^{\text{HS}} \ , \tag{4.11}$$

where a^{HS} is the dimensionless contribution to the hard-sphere free energy per segment, obtained using the expression of Boublík [266] and Mansoori *et al.* [267],

$$a^{\text{HS}} = \frac{6}{\pi \rho_s} \left[\left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0 \right) \ln(1 - \zeta_3) + 3 \frac{\zeta_1 \zeta_2}{1 - \zeta_3} + \frac{\zeta_2^3}{\zeta_3 (1 - \zeta_3)^2} \right] \ . \tag{4.12}$$

In eq. (4.12), ρ_s is the segment number density which is related to the molecular density ρ through

$$\rho_s = \rho \left(\sum_{i=1}^{N_\mathrm{C}} x_i \sum_{k=1}^{N_\mathrm{G}} \nu_{k,i} \nu_k^* S_k \right) \ , \tag{4.13}$$

and the reduced densities ζ_m are expressed as

$$\zeta_m = \frac{\pi \rho_s}{6} \sum_{k=1}^{N_\mathrm{G}} x_{s,k} d_{kk}^m \ , \ m = 0, 1, 2, 3 \ , \tag{4.14}$$

where the effective hard-sphere diameter d_{kk} of the reference fluid (cf. eq. (4.10)) is used. The summation of eq. (4.14) is expressed in terms of the fraction $x_{s,k}$ of segments of a group of type k in the mixture, which is defined as

$$x_{s,k} = \frac{\sum_{i=1}^{N_\mathrm{C}} x_i \nu_{k,i} \nu_k^* S_k}{\sum_{i=1}^{N_\mathrm{C}} x_i \sum_{l=1}^{N_\mathrm{G}} \nu_{l,i} \nu_l^* S_l} \ . \tag{4.15}$$

After substituting the expression of the group fraction $x_{s,k}$ (eq. (4.15)) in the definition of the reduced densities (eq. (4.14)) and expressing the reference hard-sphere energy per segment as a function of the molecular density, one obtains the following compact expression for the hard-sphere Helmholtz free energy per molecule:

$$\frac{A^\mathrm{HS}}{N k_\mathrm{B} T} = \frac{6}{\pi \rho} \left[\left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0 \right) \ln(1 - \zeta_3) + 3 \frac{\zeta_1 \zeta_2}{1 - \zeta_3} + \frac{\zeta_2^3}{\zeta_3 (1 - \zeta_3)^2} \right] \ , \tag{4.16}$$

which is identical to the form of the Helmholtz free energy of a hard-sphere mixture. The first-order term A_1 of the perturbation expansion corresponds to the mean-attractive energy, and as for the hard-sphere term it is obtained as a summation of the contributions to the mean-attractive energy per segment a_1:

$$\frac{A_1}{N k_\mathrm{B} T} = \frac{1}{k_\mathrm{B} T} \left(\sum_{i=1}^{N_\mathrm{C}} x_i \sum_{k=1}^{N_\mathrm{G}} \nu_{k,i} \nu_k^* S_k \right) a_1 \ . \tag{4.17}$$

The mean-attractive energy per segment is obtained by summing the pairwise interactions $a_{1,kl}$ between groups k and l over all functional groups present in the system,

$$a_1 = \sum_{k=1}^{N_\mathrm{G}} \sum_{l=1}^{N_\mathrm{G}} x_{s,k} x_{s,l} a_{1,kl} \ , \tag{4.18}$$

where it can be shown [31] that

$$
\begin{aligned}
a_{1,kl} &= \mathcal{C}_{kl}\left[x_{0,kl}^{\lambda_{kl}^a}\left(a_{1,kl}^s(\rho_s;\lambda_{kl}^a)+B_{kl}(\rho_s;\lambda_{kl}^a)\right)\right. \\
&- \left. x_{0,kl}^{\lambda_{kl}^r}\left(a_{1,kl}^s(\rho_s;\lambda_{kl}^r)+B_{kl}(\rho_s;\lambda_{kl}^r)\right)\right] \ .
\end{aligned} \tag{4.19}
$$

Here, $\mathcal{C}_{kl}$ is the pre-factor of the potential (cf. eq. (4.2)), $x_{0,kl}$ is defined as $x_{0,kl} = \sigma_{kl}/d_{kl}$, and B_{kl} is given by

$$
B_{kl}(\rho_s;\lambda_{kl}) = 2\pi\rho_s d_{kl}\epsilon_{kl}\left(\frac{1-\zeta_x/2}{(1-\zeta_x)^3}I(\lambda_{kl}) - \frac{9\zeta_x(1+\zeta_x)}{2(1-\zeta_x)^3}J(\lambda_{kl})\right) \ . \tag{4.20}
$$

The range λ_{kl} is a generalised notation which indicates that the expression can be evaluated for both the repulsive λ_{kl}^r and the attractive λ_{kl}^a exponents. In expression (4.20) ζ_x is the density of a hypothetical pure fluid, obtained based on the segment density ρ_s of the system by applying the van der Waals (vdW) one-fluid mixing rule:

$$
\zeta_x = \frac{\pi\rho_s}{6}\sum_{k=1}^{N_G}\sum_{l=1}^{N_G} x_{s,k}x_{s,l}d_{kl}^3 \ . \tag{4.21}
$$

The unlike effective hard-sphere diameter d_{kl} is obtained with an appropriate combining rule (cf. section 4.2.5). The quantities $I(\lambda_{kl})$ and $J(\lambda_{kl})$ are introduced in order to simplify the integration of the potential that leads to analytical expressions for the first-order perturbation term as discussed in detail in [31]. Both $I(\lambda_{kl})$ and $J(\lambda_{kl})$ are functions of the parameters of the intermolecular interaction potential alone and are calculated (for either λ_{kl}^r or λ_{kl}^a) as [31]

$$
I(\lambda_{kl}) = \int_1^{x_0}\frac{x^2}{x^{\lambda_{kl}}}dx = \frac{1-(x_0)^{(3-\lambda_{kl})}}{(\lambda_{kl}-3)} \ , \tag{4.22}
$$

and

$$
J(\lambda_{kl}) = \int_1^{x_0}\frac{x^3-x^2}{x^{\lambda_{kl}}}dx = \frac{1-(x_0)^{(4-\lambda_{kl})}(\lambda_{kl}-3)-(x_0)^{(3-\lambda_{kl})}(\lambda_{kl}-4)}{(\lambda_{kl}-3)(\lambda_{kl}-4)} \ . \tag{4.23}
$$

The free energy $a_{1,kl}^s(\rho_s;\lambda_{kl})$ appearing in eq. (4.19) corresponds to the first-order perturbation term of a Sutherland fluid characterised by a hard-core diameter d_{kl}, an interaction range exponent of λ_{kl}, and energy well-depth of ϵ_{kl}. The exact evaluation of this term requires a knowledge of the radial distribution function of the hard-sphere system over a range of separations. In order to derive an analytical expression in the same spirit as in the original SAFT-VR approach [170, 171], the mean-value theorem is applied in order to integrate over the radial distribution function by mapping it to its value at contact d_{kl} at an effective packing fraction $\zeta_x^{\text{eff.}}$ [31]. This procedure is shown to provide a description which is in excellent agreement with the evaluation of the $a_{1,kl}^s(\rho_s;\lambda_{kl})$ by full quadrature

using an integral equation theory for the radial distribution function [31]. The $a^s_{1;kl}(\rho_s;\lambda_{kl})$ term is then evaluated by means of the following compact expression:

$$a^s_{1,kl}(\rho_s;\lambda_{kl}) = -2\pi\rho_s \left(\frac{\epsilon_{kl}d^3_{kl}}{\lambda_{kl}-3}\right)\frac{1-\zeta_x^{\text{eff.}}/2}{(1-\zeta_x^{\text{eff.}})^3} \ . \tag{4.24}$$

The effective packing fraction was parametrised [31] for ranges of the exponents of $5 < \lambda_{kl} \leq 100$ and can be expressed as a function of the one-fluid packing fraction ζ_x as

$$\zeta_{kl}^{\text{eff.}} = c_{1,kl}\zeta_x + c_{2,kl}\zeta_x^2 + c_{3,kl}\zeta_x^3 + c_{4,kl}\zeta_x^4 \ , \tag{4.25}$$

where the coefficients, $c_{1,kl}, c_{2,kl}, c_{3,kl}, c_{4,kl}$, are obtained as functions of the generic exponent λ_{kl} as

$$\begin{pmatrix} c_{1,kl} \\ c_{2,kl} \\ c_{3,kl} \\ c_{4,kl} \end{pmatrix} = \begin{pmatrix} 0.81096 & 1.7888 & -37.578 & 92.284 \\ 1.0205 & -19.341 & 151.26 & -463.50 \\ -1.9057 & 22.845 & -228.14 & 973.92 \\ 1.0885 & -6.1962 & 106.98 & -677.64 \end{pmatrix} \begin{pmatrix} 1 \\ 1/\lambda_{kl} \\ 1/\lambda_{kl}^2 \\ 1/\lambda_{kl}^3 \end{pmatrix} \ . \tag{4.26}$$

The second-order perturbation term in the high-temperature expansion (cf. eq. (4.9)) represents the fluctuation of the attractive energy in the system and is obtained as the following sum:

$$\frac{A_2}{Nk_{\text{B}}T} = \left(\frac{1}{k_{\text{B}}T}\right)^2 \left(\sum_{i=1}^{N_{\text{C}}} x_i \sum_{k=1}^{N_{\text{G}}} \nu_{k,i}\nu_k^* S_k\right) a_2 \ , \tag{4.27}$$

where the fluctuation term per segment a_2 is obtained from the appropriate sum of the contributions of the pairwise interactions $a_{2,kl}$ between groups k and l as

$$a_2 = \sum_{k=1}^{N_{\text{G}}}\sum_{l=1}^{N_{\text{G}}} x_{s,k}x_{s,l}a_{2,kl} \ . \tag{4.28}$$

The expression for a_2 is obtained based on an improved macroscopic compressibility approximation (MCA) proposed by Zhang *et al.* [346] combined with a correction in the same spirit as that proposed by Paricaud [345] for soft potentials. The final expression for $a_{2;kl}$ is expressed as

$$\begin{aligned} a_{2,kl} \ = \ & \frac{1}{2}K^{\text{HS}}(1+\chi_{kl})\epsilon_{kl}\mathcal{C}_{kl}^2\Big[x_{0,kl}^{(2\lambda_{kl}^a)}\left(a^s_{1,kl}(\rho_s;2\lambda^a_{kl}) + B_{kl}(\rho_s;2\lambda^a_{kl})\right) \\ & -2x_{0,kl}^{(\lambda_{kl}^a+\lambda_{kl}^r)}\left(a^s_{1,kl}(\rho_s;\lambda^a_{kl}+\lambda^r_{kl}) + B_{kl}(\rho_s;\lambda^a_{kl}+\lambda^r_{kl})\right) \\ & +x_{0,kl}^{(2\lambda_{kl}^r)}\left(a^s_{1,kl}(\rho_s;2\lambda^r_{kl}) + B_{kl}(\rho_s;2\lambda^r_{kl})\right)\Big] \ , \end{aligned} \tag{4.29}$$

where K^{HS} is the isothermal compressibility of the hypothetical vdW one-fluid system (cf. eq. (4.21)) and is obtained based on the Carnahan and Starling expression [151] as

$$K^{\mathrm{HS}} = \frac{(1-\zeta_x)^4}{1+4\zeta_x+4\zeta_x^2-4\zeta_x^3+\zeta_x^4} \ . \tag{4.30}$$

The correction factor χ_{kl} is obtained from [31]

$$\chi_{kl} = f_1(\alpha_{kl})\zeta_x x_{0,kl}^3 + f_2(\alpha_{kl})(\zeta_x x_{0,kl}^3)^5 + f_3(\alpha_{kl})(\zeta_x x_{0,kl}^3)^8 \ , \tag{4.31}$$

where the quantities f_m, for $m = 1, 2, 3$, are functions of α_{kl}, a dimensionless form of the integrated van der Waals energy of the Mie potential:

$$\alpha_{kl} = \frac{1}{\epsilon_{kl}\sigma_{kl}^3}\int_\sigma^\infty \Phi_{kl}^{\mathrm{Mie}}(r) r^2 dr = \mathcal{C}_{kl}\left(\frac{1}{\lambda_{kl}^a-3}-\frac{1}{\lambda_{kl}^r-3}\right) \ . \tag{4.32}$$

The third-order perturbation term is obtained as a sum over the contribution per segment a_3 as:

$$\frac{A_3}{Nk_{\mathrm{B}}T} = \left(\frac{1}{k_{\mathrm{B}}T}\right)^3 \left(\sum_{i=1}^{N_{\mathrm{C}}} x_i \sum_{k=1}^{N_{\mathrm{G}}} \nu_{k,i}\nu_k^* S_k\right) a_3 \ , \tag{4.33}$$

where a_3 is obtained by summing the pairwise segment-segment contributions $a_{3,kl}$ on groups k and l as previously for the a_1 and a_2 terms:

$$a_3 = \sum_{k=1}^{N_{\mathrm{G}}}\sum_{l=1}^{N_{\mathrm{G}}} x_{s,k}x_{s,l}a_{3,kl} \ . \tag{4.34}$$

The contribution $a_{3,kl}$, is obtained using the following empirical expression [31]

$$a_{3;kl} = -\epsilon_{kl}^3 f_4(\alpha_{kl})\zeta_x x_{0,kl}^3 \exp\left(f_5(\alpha_{kl})\zeta_x x_{0,kl}^3 + f_6(\alpha_{kl})(\zeta_x x_{0,kl}^3)^2\right) \ . \tag{4.35}$$

This functional form was chosen to restrict the dependence to only the values of the repulsive and attractive exponents, the well-depth of the potential, and the density of the mixture. A temperature dependence is avoided by expressing the $a_{3,kl}$ term as a function of the product $\zeta_x x_{0,kl}^3$, as opposed to the packing fraction ζ_x which has an implicit temperature dependence through the effective diameter d_{kl}.

The functions f_m for $(m = 1, \ldots, 6)$ appearing in eqs. (4.31) and (4.35) are calculated by means of the following compact expression [31]:

$$f_m(\alpha_{kl}) = \frac{\sum_{n=0}^{3}\phi_{m,n}\alpha_{kl}^n}{1+\sum_{n=4}^{6}\phi_{m,n}\alpha_{kl}^{n-3}} \ , \text{ for } m = 1, \ldots, 6 \ . \tag{4.36}$$

The values of the coefficients $\phi_{m,n}$ are listed in table 4.1. The coefficients that appear in the modified MCA for the $a_{2,kl}$ term, i.e., f_m for $m = 1, 2, 3$, are obtained by analysis of the corresponding Monte Carlo simulation data for the fluctuation term and selected pure component vapour-liquid equilibrium data for selected Mie fluids (λ^r,λ^a). The coefficients employed for the calculation of the $a_{3,kl}$ term were obtained by comparison with simulation data for vapour-liquid equilibrium and critical points of several Mie (λ^r,λ^a) fluids as explained in [31]. It is important to note, that as the coefficients included in the calculation of the $a_{3,kl}$ are obtained from simulation data of the fluid phase behaviour of monomers, the final expression for the free energy accounts for several higher-order terms of the perturbation expansion of Barker and Henderson [269] (in fact, for the complete series), rather than just the third-order term. A retrospective analysis of the third-order pertubation term showed that the calculations using the empirical expression of eq. (4.35) are in good agreement with evaluations of the equivalent term using molecular simulation.

4.2.3 Chain Term

In the SAFT-γ formalism the treatment of the contribution to the free energy due to the formation of molecules from fused Mie segments (the so-called "chain" contribution in SAFT approaches) requires a knowledge of the contact value of the radial distribution function of the fluid at an effective diameter [19, 29]. In order to evaluate this, a number of average molecular parameters for each molecular species i in the mixture are introduced. The averaging of the molecular size and energy parameters, namely $\bar{\sigma}_{ii}$, $\bar{d}_{ii}$, $\bar{\epsilon}_{ii}$ and $\bar{\lambda}_{ii}$, is independent of the composition of the mixture and makes use of the fraction $z_{k,i}$ of a given group k on a molecule i:

$$z_{k,i} = \frac{\nu_{k,i}\nu_k^* S_k}{\sum_{l=1}^{N_G} \nu_{l,i}\nu_l^* S_l} \ . \tag{4.37}$$

The quantity $z_{k,i}$ is not to be confused with the fraction $x_{s,k}$ of a given group k in the mixture, which is composition dependent (cf. eq. (4.15)). The average molecular segment size $\bar{\sigma}_{ii}$ and the effective hard-sphere diameter $\bar{d}_{ii}$ are defined as

$$\bar{\sigma}_{ii}^3 = \sum_{k=1}^{N_G}\sum_{l=1}^{N_G} z_{k,i} z_{l,i} \sigma_{kl}^3 \ , \tag{4.38}$$

and

$$\bar{d}_{ii}^3 = \sum_{k=1}^{N_G}\sum_{l=1}^{N_G} z_{k,i} z_{l,i} d_{kl}^3 \ . \tag{4.39}$$

The averaging rule for the effective hard-sphere diameter $\bar{d}_{ii}$ is chosen such that the value of the one-fluid density of a mixture of *heteronuclear* monomeric segments as calculated in eq. (4.21), is the same when calculated for monomeric segments of average molecular size $\bar{d}_{ii}$, i.e., $\zeta_x = \frac{\pi\rho_s}{6}\sum_{k=1}^{N_\mathrm{G}}\sum_{l=1}^{N_\mathrm{G}} x_{s,k}x_{s,l}d_{kl}^3 = \frac{\pi\rho_s}{6}\sum_{i=1}^{N_\mathrm{C}}\sum_{j=1}^{N_\mathrm{C}} x_i(\sum_{k=1}^{N_\mathrm{G}}\nu_k^*\nu_{k,i}S_k)x_j(\sum_{k=1}^{N_\mathrm{G}}\nu_k^*\nu_{k,j}S_k)\bar{d}_{ij}^3$.

Other effective molecular parameters are obtained in the same way, so that the average interaction energy $\bar{\epsilon}_{ii}$ and exponents which characterise the range of the potential $\bar{\lambda}_{ii}$ are again obtained as

$$\bar{\epsilon}_{ii} = \sum_{k=1}^{N_\mathrm{G}}\sum_{l=1}^{N_\mathrm{G}} z_{k,i}z_{l,i}\epsilon_{kl} \ , \tag{4.40}$$

and

$$\bar{\lambda}_{ii} = \sum_{k=1}^{N_\mathrm{G}}\sum_{l=1}^{N_\mathrm{G}} z_{k,i}z_{l,i}\lambda_{kl} \ . \tag{4.41}$$

Relation (4.41), holds for both the repulsive, $\bar{\lambda}_{ii}^r$, and the attractive, $\bar{\lambda}_{ii}^a$, exponents.

The contribution to the free energy of the mixture due to the formation of the "chain" molecules is based on the thermodynamic perturbation theory of first order (TPT1) of Wertheim [22, 101–104] but using the effective molecular parameters:

$$\frac{A^\mathrm{chain}}{Nk_\mathrm{B}T} = -\sum_{i=1}^{N_\mathrm{C}} x_i \sum_{k=1}^{N_\mathrm{G}} (\nu_{k,i}\nu_k^* S_k - 1)\ln g_{ii}^\mathrm{Mie}(\bar{\sigma}_{ii};\zeta_x) \ , \tag{4.42}$$

where $g_{ii}^\mathrm{Mie}(\bar{\sigma}_{ii};\zeta_x)$ is the value of the radial distribution function (RDF) of the hypothetical one-fluid Mie system at a density ζ_x evaluated for the effective diameter $\bar{\sigma}_{ii}$. An accurate estimate of the contact value of the RDF for a Mie fluid can be obtained by means of a second-order expanion [31]

$$g_{ii}^\mathrm{Mie}(\bar{\sigma}_{ii};\zeta_x) = g_\mathrm{d}^\mathrm{HS}(\bar{\sigma}_{ii})\exp[\beta\bar{\epsilon}_{ii}g_1(\bar{\sigma}_{ii})/g_\mathrm{d}^\mathrm{HS}(\bar{\sigma}_{ii}) + (\beta\bar{\epsilon}_{ii})^2 g_2(\bar{\sigma}_{ii})/g_\mathrm{d}^\mathrm{HS}(\bar{\sigma}_{ii})] \ . \tag{4.43}$$

The zeroth-order term of the expansion, $g_\mathrm{d}^\mathrm{HS}(\bar{\sigma}_{ii})$, is the radial distribution function of a system of hard spheres of diameter $\bar{d}_{ii}$ evaluated at the distance $\bar{\sigma}_{ii}$ and density ζ_x. As shown in [31], a compact relation for this function can be obtained using the expression of Boublík [347],

$$g_\mathrm{d}^\mathrm{HS}(\bar{\sigma}_{ii}) = g_\mathrm{d}^\mathrm{HS}(\bar{x}_{0,ii}) = \exp(k_0 + k_1\bar{x}_{0,ii} + k_2\bar{x}_{0,ii}^2 + k_3\bar{x}_{0,ii}^3) \ , \tag{4.44}$$

which is valid for $1 < \bar{x}_{0,ii} < \sqrt{2}$ (with $\bar{x}_{0,ii} = \bar{\sigma}_{ii}/\bar{d}_{ii}$). In this expression, the coefficients k_m are obtained as functions of the one-fluid density ζ_x (cf. eq. (4.21)) of the hypothetical pure fluid as

$$k_0 = -\ln(1-\zeta_x) + \frac{42\zeta_x - 39\zeta_x^2 + 9\zeta_x^3 - 2\zeta_x^4}{6(1-\zeta_x)^3} \ , \tag{4.45}$$

$$k_1 = \frac{\zeta_x^4 + 6\zeta_x^2 - 12\zeta_x}{2(1-\zeta_x)^3} \ , \tag{4.46}$$

$$k_2 = \frac{-3\zeta_x^2}{8(1-\zeta_x)^2} \ , \tag{4.47}$$

and

$$k_3 = \frac{-\zeta_x^4 + 3\zeta_x^2 + 3\zeta_x}{6(1-\zeta_x)^3} \ . \tag{4.48}$$

The first-order term $g_1(\bar{\sigma}_{ii})$ of the expansion for the contact value of the RDF (eq. (4.43)) is approximated by its value at contact $\bar{d}_{ii}$, obtained by means of a self-consistent method for the calculation of pressure from the virial and the free energy routes [31, 170, 171]

$$\begin{aligned} g_1(\bar{\sigma}_{ii}) \approx g_1(\bar{d}_{ii}) \quad = \quad & \frac{1}{2\pi\bar{\varepsilon}_{ii}\bar{d}_{ii}^3}\Bigg[3\frac{\partial \bar{a}_{1,ii}}{\partial \rho_s} - \bar{\mathcal{C}}_{ii}\bar{\lambda}_{ii}^a \bar{x}_{0,ii}^{\bar{\lambda}_{ii}^a} \frac{\bar{a}_{1,ii}^s(\rho_s;\bar{\lambda}_{ii}^a) + \bar{B}_{ii}(\rho_s;\bar{\lambda}_{ii}^a)}{\rho_s} \\ & + \bar{\mathcal{C}}_{ii}\bar{\lambda}_{ii}^r \bar{x}_{0,ii}^{\bar{\lambda}_{ii}^r} \frac{\bar{a}_{1,ii}^s(\rho_s;\bar{\lambda}_{ii}^r) + \bar{B}_{ii}(\rho_s;\bar{\lambda}_{ii}^r)}{\rho_s}\Bigg] \ . \end{aligned} \tag{4.49}$$

In this expression $\bar{\mathcal{C}}_{ii}$ is the prefactor of the effective molecular interaction potential of component i (cf. eq. (4.2)) using the values of the average molecular repulsive and attractive exponents ($\bar{\lambda}_{ii}^r$ and $\bar{\lambda}_{ii}^a$), i.e. $\bar{\mathcal{C}}_{ii}\frac{\bar{\lambda}_{ii}^r}{\bar{\lambda}_{ii}^r - \bar{\lambda}_{ii}^a}\left(\frac{\bar{\lambda}_{ii}^r}{\bar{\lambda}_{ii}^a}\right)^{\frac{\bar{\lambda}_{ii}^a}{\bar{\lambda}_{ii}^r - \bar{\lambda}_{ii}^a}}$. The quantity $\bar{B}_{ii}(\rho_s;\bar{\lambda}_{ii})$ is obtained using eq. (4.20) as

$$\bar{B}_{ii}(\rho_s;\bar{\lambda}_{ii}) = 2\pi\rho_s\bar{d}_{ii}\bar{\epsilon}_{ii}\left(\frac{1-\zeta_x/2}{(1-\zeta_x)^3}\bar{I}(\bar{\lambda}_{ii}) - \frac{9\zeta_x(1+\zeta_x)}{2(1-\zeta_x)^3}\bar{J}(\bar{\lambda}_{ii})\right) \ , \tag{4.50}$$

but with $\bar{I}(\bar{\lambda}_{ii})$ and $\bar{J}(\bar{\lambda}_{ii})$ (cf. eq. (4.22) and (4.23)), now based on the effective exponents $\bar{\lambda}_{ii}^r$ and $\bar{\lambda}_{ii}^a$ and the ratio $\bar{x}_{0,ii} = \bar{\sigma}_{ii}/\bar{d}_{ii}$.

The first-order term of the radial distribution function further depends on the density derivative of the effective first-order perturbation term $\bar{a}_{1,ii}$ for the contribution of the monomeric interactions to the free energy per segment, which is calculated following expression (4.19) as

$$\begin{aligned}\bar{a}_{1,ii} &= \bar{\mathcal{C}}_{ii}\left[\bar{x}_{0,ii}^{\bar{\lambda}_{ii}^a}\left(\bar{a}_{1,ii}^s(\rho_s;\bar{\lambda}_{ii}^a)+\bar{B}_{ii}(\rho_s;\bar{\lambda}_{ii}^a)\right)\right. \\ &- \left.\bar{x}_{0,ii}^{\bar{\lambda}_{ii}^r}\left(\bar{a}_{1,ii}^s(\rho_s;\bar{\lambda}_{ii}^r)+\bar{B}_{ii}(\rho_s;\bar{\lambda}_{ii}^r)\right)\right] .\end{aligned} \tag{4.51}$$

The integrated energy of the Sutherland fluid $\bar{a}_{1;ii}^s(\rho_s;\bar{\lambda}_{ii})$ calculated for the effective molecular parameters, is obtained as

$$\bar{a}_{1,ii}^s(\rho_s;\bar{\lambda}_{ii}) = -2\pi\rho_s\left(\frac{\bar{\epsilon}_{ii}\bar{d}_{ii}^3}{\bar{\lambda}_{ii}-3}\right)\frac{1-\bar{\zeta}_{ii}^{\text{eff.}}/2}{(1-\bar{\zeta}_{ii}^{\text{eff.}})^3} , \tag{4.52}$$

where the effective packing fraction $\bar{\zeta}_{ii}^{\text{eff.}}$ used for the mapping of the radial distribution function at contact, is now calculated as

$$\bar{\zeta}_{ii}^{\text{eff.}} = \bar{c}_{1,ii}\zeta_x + \bar{c}_{2,ii}\zeta_x^2 + \bar{c}_{3,kl}\zeta_x^3 + \bar{c}_{4,ii}\zeta_x^4 . \tag{4.53}$$

The coefficients of eq. (4.53) are obtained based on the effective values of the exponents of the potential ($\bar{\lambda}_{ii}$):

$$\begin{pmatrix}\bar{c}_{1,ii}\\ \bar{c}_{2,ii}\\ \bar{c}_{3,ii}\\ \bar{c}_{4,ii}\end{pmatrix} = \begin{pmatrix}0.81096 & 1.7888 & -37.578 & 92.284\\ 1.0205 & -19.341 & 151.26 & -463.50\\ -1.9057 & 22.845 & -228.14 & 973.92\\ 1.0885 & -6.1962 & 106.98 & -677.64\end{pmatrix}\begin{pmatrix}1\\ 1/\bar{\lambda}_{ii}\\ 1/\bar{\lambda}_{ii}^2\\ 1/\bar{\lambda}_{ii}^3\end{pmatrix} . \tag{4.54}$$

The second-order term $g_2(\bar{\sigma}_{ii})$ of eq. (4.43) is also approximated by its value at the effective distance $\bar{d}_{ii}$. It is obtained based on the expression of the macroscopic compressibility approximation and an empirical correction [31],

$$g_2(\bar{\sigma}_{ii}) \approx g_2(\bar{d}_{ii}) = (1+\gamma_{c,ii})g_2^{\text{MCA}}(\bar{d}_{ii}) , \tag{4.55}$$

where $\gamma_{c,ii}$ is given as a function of temperature, density and the effective values of the exponents of the potential as

$$\gamma_{c,ii} = \phi_{7,0} - \tanh\left\{[\phi_{7,1}(\phi_{7,2}-\bar{\alpha}_{ii})]-1\right\}\zeta_x\theta\exp(\phi_{7,3}\zeta_x+\phi_{7,4}\zeta_x^2) , \tag{4.56}$$

where $\theta = \exp(\beta\bar{\epsilon}_{ii})-1$, the values of the coefficients $\phi_{7,0},\ldots,\phi_{7,4}$ are given in table 4.1, and $\bar{\alpha}_{ii}$ is obtained based on the effective exponents of the potential (cf. eq. 4.32), as

$$\bar{\alpha}_{ii} = \bar{\mathcal{C}}_{ii}\left(\frac{1}{\bar{\lambda}_{ii}^a-3}-\frac{1}{\bar{\lambda}_{ii}^r-3}\right) . \tag{4.57}$$

Table 4.1: Coefficients $\phi_{m,n}$ for the empirical corrections to the $a_{2,kl}$ term (eq. (4.31)), the $a_{3,kl}$ term (eq. (4.35)) and the correction $\gamma_{c,ii}$ of the g_2 term (eq. (4.55)). N/A denotes non-applicable values

n	$\phi_{1,n}$	$\phi_{2,n}$	$\phi_{3,n}$	$\phi_{4,n}$	$\phi_{5,n}$	$\phi_{6,n}$	$\phi_{7,n}$
0	7.5365557	-359.44	1550.9	-1.19932	-1911.28	9236.9	10
1	-37.60463	1825.6	-5070.1	9.063632	21390.175	-129430	10
2	71.745953	-3168.0	6534.6	-17.9482	-51320.7	357230	0.57
3	-46.83552	1884.2	-3288.7	11.34027	37064.54	-315530	-6.7
4	-2.467982	-0.82376	-2.7171	20.52142	1103.742	1390.2	-8
5	-0.50272	-3.1935	2.0883	-56.6377	-3264.61	-4518.2	N/A
6	8.0956883	3.7090	0	40.53683	2556.181	4241.6	N/A

The second-order term from the macroscopic compressibility approximation $g_2^{\mathrm{MCA}}(\bar{d}_{ii})$ of eq. (4.55) is obtained based on the fluctuation term of the Sutherland potential as

$$\begin{aligned} g_2^{\mathrm{MCA}}(\bar{d}_{ii}) &= \frac{1}{2\pi\bar{\epsilon}_i^2\bar{d}_{ii}^3}\left[3\frac{\partial}{\partial\rho_s}\left(\frac{\bar{a}_{2,ii}}{1+\bar{\chi}_{ii}}\right)\right. \\ &- \bar{\epsilon}_{ii}K^{\mathrm{HS}}\bar{\mathcal{C}}_{ii}^2\bar{\lambda}_{ii}^r\bar{x}_{0,ii}^{2\bar{\lambda}_{ii}^r}\frac{\bar{a}_{1,ii}^s(\rho_s;2\bar{\lambda}_{ii}^r)+\bar{B}(\rho_s;2\bar{\lambda}_{ii}^r)}{\rho_s} \\ &+ \bar{\epsilon}_{ii}K^{\mathrm{HS}}\bar{\mathcal{C}}_{ii}^2(\bar{\lambda}_{ii}^r+\bar{\lambda}_{ii}^a)\bar{x}_{0,ii}^{(\bar{\lambda}_{ii}^r+\bar{\lambda}_{ii}^a)}\frac{\bar{a}_{1,ii}^s(\rho_s;\bar{\lambda}_{ii}^r+\bar{\lambda}_{ii}^a)+\bar{B}(\rho_s;\bar{\lambda}_{ii}^r+\bar{\lambda}_{ii}^a)}{\rho_s} \\ &\left.- \bar{\epsilon}_{ii}K^{\mathrm{HS}}\bar{\mathcal{C}}_{ii}^2\bar{\lambda}_{ii}^a\bar{x}_{0,ii}^{2\bar{\lambda}_{ii}^a}\frac{\bar{a}_{1,ii}^s(\rho_s;2\bar{\lambda}_{ii}^a)+\bar{B}(\rho_s;2\bar{\lambda}_{ii}^a)}{\rho_s}\right]. \end{aligned} \tag{4.58}$$

where all parameters and free energy terms are again evaluated using the effective molecular parameters. Note that the empirical correction to the MCA expression is based on the effective parameters as

$$\bar{\chi}_{ii} = f_1(\bar{\alpha}_{ii})\zeta_x\bar{x}_{0,ii}^3 + f_2(\bar{\alpha}_{ii})(\zeta_x\bar{x}_{0,ii}^3)^5 + f_3(\bar{\alpha}_{ii})(\zeta_x\bar{x}_{0,ii}^3)^8 \ , \tag{4.59}$$

where the coefficients f_1, f_2, f_3 are obtained based on eq. (4.36) by using $\bar{\alpha}_{ii}$ instead of α_{kl}.

4.2.4 Association Term

The contribution to the Helmhotz free energy arising from the association of molecules through the sites is obtained as [21, 22, 101–104]

$$\frac{A^{\mathrm{assoc.}}}{Nk_{\mathrm{B}}T} = \sum_{i=1}^{N_{\mathrm{C}}} x_i \sum_{k=1}^{N_{\mathrm{G}}} \nu_{k,i} \sum_{a=1}^{N_{\mathrm{ST}_k}} n_{k,a}\left(\ln X_{i,k,a} + \frac{1-X_{i,k,a}}{2}\right) \ , \tag{4.60}$$

where N_{ST_k} is the total number of site types on a given group k, and $n_{k,a}$ the number of sites of type a on group k. $X_{i,k,a}$ is the fraction of sites of type a on group k of component i that are not bonded. It is obtained from the solution of the mass action equation as [19, 21, 104]

$$X_{i,k,a} = \frac{1}{1 + \sum_{j=1}^{N_{\mathrm{C}}} \sum_{l=1}^{N_{\mathrm{G}}} \sum_{b=1}^{N_{\mathrm{ST}_l}} \rho x_j n_{l,b} X_{j,l,b} \Delta_{ij,kl,ab}} \ . \tag{4.61}$$

Here, $\Delta_{ij,kl,ab}$ is the association strength between a site of type a on a group of type k of component i and a site of type b on a group of type l of component j, and is given by

$$\Delta_{ij,kl,ab} = \int g^{\mathrm{Mie}}(r) f_{kl,ab}(\mathbf{r}) \mathrm{d}\mathbf{r} \ , \tag{4.62}$$

where $f_{kl,ab} = \exp(-\beta \Phi^{\mathrm{SW}}_{kl,ab}) - 1$, is the Mayer-$f$ function of the association potential (cf. eq. (4.3)). By introducing the square-well bonding potential into eq. (4.62) and carrying out the angle average, the association strength can be expressed as

$$\Delta_{ij,kl,ab} = \bar{\sigma}^3_{ij} F_{kl,ab} I_{kl,ab} \ , \tag{4.63}$$

where $F_{kl,ab} = \exp(\beta \epsilon^{\mathrm{HB}}_{kl,ab}) - 1$, and $I_{kl,ab}$ is a dimensionless integral defined as [31]

$$I_{kl,ab} = \frac{\pi}{6 \bar{\sigma}^3_{ij} {r^d_{kl,ab}}^2} \int_{(2r^d_{kl,ab} - r^c_{kl,ab})}^{(2r^d_{kl,ab} - r^c_{kl,ab})} g^{\mathrm{Mie}}(r) (r^c_{kl,ab} + 2r^d_{kl,ab} - r)^2 (2r^c_{kl,ab} - 2r^d_{kl,ab}) r \mathrm{d}r \ . \tag{4.64}$$

The determination of the integral in eq. (4.64) requires a knowledge of the RDF of the reference Mie fluid over a range of distances. For a detailed discussion of the various options on how this can be calculated, see [31]. In the current work the value of the RDF is approximated based on a Barker-Henderson zeroth-order perturbation approach, so that $g^{\mathrm{Mie}}(r) \simeq g^{\mathrm{HS}}_d(r)$. After assuming that $r^2 g^{\mathrm{HS}}_d(r) \simeq d^2 g^{\mathrm{HS}}_d(d)$, an analytical form for the association contribution can be obtained, where the integral of eq. (4.64) is given by

$$I_{kl,ab} = g^{\mathrm{HS}}_d(\bar{d}_{ij}) K_{ij,kl,ab} \ , \tag{4.65}$$

and the bonding volume, $K_{ij,kl,ab}$ is obtained from [31]

$$\begin{aligned} K_{ij,kl,ab} \ = \ & \frac{\pi \bar{d}^2_{ij}}{6 \bar{\sigma}^3_{ij} {r^d_{kl,ab}}^2} \Big[-\frac{8}{3} {r^c_{kl,ab}}^3 - 7 {r^c_{kl,ab}}^2 r^d_{kl,ab} + 4 r^c_{kl,ab} {r^d_{kl,ab}}^2 + \frac{44}{3} {r^d_{kl,ab}}^3 \\ + \ & 2\ln(r^c_{kl,ab} + 2r^d_{kl,ab}) {r^c_{kl,ab}}^3 + 6\ln(r^c_{kl,ab} + 2r^d_{kl,ab}) {r^c_{kl,ab}}^2 r^d_{kl,ab} \\ - \ & 8\ln(r^c_{kl,ab} + 2r^d_{kl,ab}) {r^d_{kl,ab}}^3 - \frac{1}{3} \bar{d}^3_{ij} + 3 r^d_{kl,ab} \bar{d}^2_{ij} + 3 {r^c_{kl,ab}}^2 \bar{d}^2_{ij} + 3 {r^c_{kl,ab}}^2 \bar{d}_{ij} \\ - \ & 12 {r^d_{kl,ab}}^2 \bar{d}_{ij} - 2\ln(\bar{d}_{ij}) {r^c_{kl,ab}}^3 - 6\ln(\bar{d}_{ij}) {r^c_{kl,ab}}^2 r^d_{kl,ab} + 8\ln(\bar{d}_{ij}) {r^d_{kl,ab}}^3 \Big] \end{aligned} \tag{4.66}$$

4.2.5 Combining Rules

Combining rules are commonly employed within equations of state to facilitate the study of binary and multicomponent systems. In the specific case of the methodology presented in the current work, the interactions between groups of different kind also contribute to the description of pure components represented with a *heteronuclear* molecular model. The unlike segment diameter is obtained from a simple arithmetic mean (Lorentz rule [111]) as

$$\sigma_{kl} = \frac{\sigma_{kk} + \sigma_{ll}}{2} \ , \tag{4.67}$$

and the combining rule is applied for the calculation of the unlike effective hard-sphere diameter, so that d_{kl} is calculated as

$$d_{kl} = \frac{d_{kk} + d_{ll}}{2} \ . \tag{4.68}$$

Bearing in mind the definition of the Barker and Henderson hard-sphere diameter (cf. eq. (4.10)), a more rigorous way to obtain d_{kl} would be to integrate the potential of the unlike interaction between groups k and l as

$$d_{kl} = \int_0^{\sigma_{kl}} \left[1 - \exp\left\{-\beta u_{kl}^{Mie}(r)\right\}\right] dr \ . \tag{4.69}$$

However, such an approach would require extensive numerical calculations. Moreover, in agreement with earlier findings [31], the approximation of d_{kl} as an arithmetic mean is found to have minimal impact on the performance of the method, as judged by the quality of the description of properties of real compounds.

The unlike dispersion energy ϵ_{kl} between groups k and l, can be obtained by the application of an augmented geometric mean, to account for the asymmetry in size:

$$\epsilon_{kl} = \frac{\sqrt{\sigma_{kk}^3 \sigma_{ll}^3}}{\sigma_{kl}^3} \sqrt{\epsilon_{kk}\epsilon_{ll}} \ . \tag{4.70}$$

The combining rule for the repulsive λ_{kl}^r and the attractive λ_{kl}^a exponents of the unlike interaction is obtained by applying the geometric mean for the integrated van der Waals energy of a Sutherland fluid,

$$\alpha_{vdW;kl}^s = 2\pi\epsilon_{kl}\sigma_{kl}^3 \left(\frac{1}{\lambda_{kl} - 3}\right) \ , \tag{4.71}$$

and by imposing the Berthelot condition $\alpha_{vdW;kl}^s = \sqrt{\alpha_{vdW;kk}^s \alpha_{vdW;ll}^s}$ [111], which results in

$$\lambda_{kl} = 3 + \sqrt{(\lambda_{kk} - 3)(\lambda_{ll} - 3)} \ . \tag{4.72}$$

The combining rule of equation (4.70) provides a first estimate of the value of the unlike dispersion energy. It is known however that real systems often exhibit large deviations from simple or augmented combining rules, especially when the molecules comprise chemically different components and groups. As will be discussed in the following sections, the unlike dispersion energy ϵ_{kl} is in most cases treated as an adjustable parameter.

In the case of associating compounds, the unlike value of the association energy can be obtained by means of a simple geometric mean, as follows

$$\epsilon^{\text{HB}}_{kl,ab} = \sqrt{\epsilon^{\text{HB}}_{kk,aa}\epsilon^{\text{HB}}_{ll,bb}} \ , \tag{4.73}$$

while the unlike range of the association site-site interaction is obtained as

$$r^{c}_{kl,ab} = \frac{r^{c}_{kk,aa} + r^{c}_{ll,bb}}{2} \ . \tag{4.74}$$

A number of combining rules for the average molecular parameters required for the calculation of the chain and association contributions to the free energy also need to be considered. The unlike values for the effective segment size, $\bar{\sigma}_{ij}$ and $\bar{d}_{ij}$, dispersion energy, $\bar{\epsilon}_{ij}$, and repulsive and attractive exponents of the potential, $\bar{\lambda}_{ij}$, are obtained from

$$\bar{\sigma}_{ij} = \frac{\bar{\sigma}_{ii} + \bar{\sigma}_{jj}}{2} \ , \tag{4.75}$$

$$\bar{d}_{ij} = \frac{\bar{d}_{ii} + \bar{d}_{jj}}{2} \ , \tag{4.76}$$

$$\bar{\epsilon}_{ij} = \frac{\sqrt{\bar{\sigma}^3_{ii}\bar{\sigma}^3_{jj}}}{\bar{\sigma}^3_{ij}} \sqrt{\bar{\epsilon}_{ii}\bar{\epsilon}_{jj}} \ , \tag{4.77}$$

and

$$\bar{\lambda}_{ij} = 3 + \sqrt{(\bar{\lambda}_{ii} - 3)(\bar{\lambda}_{jj} - 3)} \ . \tag{4.78}$$

4.3 Estimation of group parameters

Within group contribution methods, the parameters that describe the contribution of each functional group to the molecular properties are typically estimated from appropriate experimental data. In the SAFT-γ Mie EoS, the contribution of a given group k is fully described by the number ν^*_k of identical segments it comprises and a set of five segment-specific parameters: the shape factor S_k; the segment diameter σ_{kk}; the energy of dispersion ϵ_{kk}; and the repulsive λ^r_{kk} and attractive λ^a_{kk} exponents of the interaction potential. In the case of associating compounds, a further two parameters (for each pair

of associating sites) have to be determined: the energy $\epsilon^{\mathrm{HB}}_{kk,ab}$ and the range $r^{\mathrm{c}}_{kk,ab}$ of the association interactions between sites of type a and b on group k. This can include interactions between sites of the same type, i.e. $\epsilon^{\mathrm{HB}}_{kk,aa}$ $r^{\mathrm{c}}_{kk,ab}$. The complete description of an associating functional group requires the knowledge of the number N_{ST_k} of site types as well as the number of sites of each type that a group comprises, $n_{k,a}$ for $a = 1, \ldots, N_{\mathrm{ST}_k}$, which are allowed to take only integer values. These parameters, together with ν^*_k are chosen *a priori* based on the chemical nature of each group (for the associating sites), by examining the different realistic possibilities with a trial-and-error approach.

As discussed in the previous section, the unlike segment diameter, σ_{kl}, and the values of the exponents of the unlike interaction potential, λ^r_{kl} and λ^a_{kl}, are calculated by means of the combining rules shown in eqs. (4.67) and (4.72), respectively. The value of the unlike dispersion energy, ϵ_{kl}, is typically treated as an adjustable parameter and is therefore obtained by regression to experimental data. In many cases, the value of the unlike interaction energy can be estimated from pure component data, a unique characteristic of *heteronuclear* models, as will be shown in the following section; this can allow for accurate predictions for properties of mixtures from pure component data alone. Mixture data is used where necessary to estimate the unlike energy, as demonstrated in previous work [348].

The group parameters are estimated simultaneously from experimental data for a series of pure substances belonging to a given chemical family. In most group contribution approaches, the parameter estimation procedure is started with the n-alkanes series, where the parameters for the CH_3 and the CH_2 groups are obtained. Once the parameters for these groups have been determined, they are transferred to the study of additional functional groups based on experimental data for a subsequent homologous series, e.g., the 2-ketones for the CH_3CO group. The parameters that describe each functional group within the SAFT-γ Mie EoS are obtained by regression to pure component vapour-liquid equilibrium data (i.e., vapour pressures p_{vap} and saturated liquid densities ρ_{sat}), including single-phase densities $\rho_{\mathrm{liq}}(T, p)$ at given temperature and pressure. The temperature range commonly used for vapour-liquid equilibrium data is between the triple point and 0.9 $T^{\mathrm{exp}}_{\mathrm{c}}$, with $T^{\mathrm{exp}}_{\mathrm{c}}$ the experimental critical temperature of the substance under study. Experimental data closer than 0.9 $T^{\mathrm{exp}}_{\mathrm{c}}$ are not included in the parameter estimation in the current work, despite the improved description of the near-critical region with the presented methodology. Since the equation of state is a classical theory it is characterised by mean-field critical exponents, and does not allow one to reproduce the density fluctuations in the critical region; including data closer to the critical point would bias the parameters towards a more accurate representation of the critical point and a sub-optimal

description of the sub-critical region. For the single phase density, experimental data in the high-temperature and pressure liquid and supercritical region are typically used. These data are included when available as they provide information of the compressibility for the fluid and can help towards the accurate prediction of derivative thermodynamic properties. The objective function used in the parameter estimations is given by

$$\begin{aligned} \min_{\mathbf{\Omega}} f_{\text{obj}} &= \frac{w_1}{N_{p_{\text{vap}}}} \sum_{i=1}^{N_{p_{\text{vap}}}} \left[\frac{p_{\text{vap}}^{\text{exp}}(T_i) - p_{\text{vap}}^{\text{calc}}(T_i;\mathbf{\Omega})}{p_{\text{vap}}^{\text{exp}}(T_i)} \right]^2 \\ &+ \frac{w_2}{N_{\rho_{\text{sat}}}} \sum_{j=1}^{N_{\rho_{\text{sat}}}} \left[\frac{\rho_{\text{sat}}^{\text{exp}}(T_j) - \rho_{\text{sat}}^{\text{calc}}(T_j;\mathbf{\Omega})}{\rho_{\text{sat}}^{\text{exp}}(T_j)} \right]^2 \\ &+ \frac{w_3}{N_{\rho_{\text{liq}}}} \sum_{k=1}^{N_{\rho_{\text{liq}}}} \left[\frac{\rho_{\text{liq}}^{\text{exp}}(T_k,p_k) - \rho_{\text{liq}}^{\text{calc}}(T_k,p_k;\mathbf{\Omega})}{\rho_{\text{liq}}^{\text{exp}}(T_k,p_k)} \right]^2 , \end{aligned} \tag{4.79}$$

where $\mathbf{\Omega}$ denotes the vector of the parameters to be estimated, the indices i, j and k allow for the summation over all experimental points for each property, denoted as $N_{p_{\text{vap}}}$, $N_{\rho_{\text{sat}}}$ and $N_{\rho_{\text{liq}}}$ for the vapour pressure, saturated liquid density, and single-phase density, respectively. The desired level of accuracy for each property can be adjusted by means of a weighting factor, w_1 for p_{vap}, w_2 for ρ_{sat} and w_3 for ρ_{liq}. The estimations were performed using numerical solvers provided by the commercial software package gPROMS© [276].

4.4 Results and discussion

4.4.1 SAFT-γ Mie group parameters

In the first instance models for the characterisation of the functional groups of two chemical families within the framework of SAFT-γ Mie are developed, and more specifically for the n-alkanes and the 2-ketones. In the following sections, the definitions of the functional groups for each chemical family, together with the detailed results of the regression to the experimental data are be presented. The metric used in this work to quantify the accuracy of the description of the experimental data for a property of a given compound R_i is the percentage average abolute deviation (% AAD) defined as:

$$\%\text{AAD} = \frac{1}{N_{R^{\text{exp.}}}} \sum_{i=1}^{N_{R^{\text{exp.}}}} \left| \frac{R_i^{\text{exp.}} - R_i^{\text{calc.}}}{R_i^{\text{exp.}}} \right| . \tag{4.80}$$

where $N_{R^{\text{exp.}}}$ is the number of data points of a property, $R_i^{\text{exp.}}$ the experimental value and $R_i^{\text{calc.}}$ the calculated value for the same property.

Table 4.2: Group parameters for the functional groups for the n-alkanes (CH_3 and CH_2) and for the 2-ketones (CH_3CO) within the SAFT-γ Mie framework.

Functional Group k	ν_k^*	S_k [-]	λ_{kk}^r [-]	λ_{kk}^a [-]	σ_{kk} [Å]	(ϵ_{kk}/k_B) [K]
CH_3	1	0.5725	15.05	6	4.077	256.766
CH_2	1	0.2293	19.871	6	4.88	473.389
CH_3CO	2	0.628	17.508	6	3.748	463.412

Table 4.3: Unlike dispersion interaction energies ϵ_{kl}/k_B for the functional groups for the n-alkanes (CH_3 and CH_2) and for the 2-ketones (CH_3CO) within the SAFT-γ Mie framework.

ϵ_{kl}/k_B [T]	CH_3	CH_2	CH_3CO
CH_3	256.766	350.772	334.437
CH_2	350.772	473.389	405.882
CH_3CO	334.437	405.882	463.412

4.4.1.1 n-Alkanes

The chemical family of the n-alkanes is considered in order to obtain the intermolecular model parameters that describe the ubiquitous methyl (CH_3) and methylene (CH_2) functional groups. The parameters estimated for these functional groups, summarised in tables 4.2 and 4.3, are obtained by regression to experimental data for the vapour pressure, saturated liquid density, and single-phase density of linear alkanes from ethane to n-decane. The number of data points and the temperature range (and pressure range, for the single phase density) for each compound and property considered are given in tables 4.4 and 4.5. For the vapour pressure and the saturated liquid density correlated experimental data from NIST [277] are used, which are provided from multiparametric Helmholtz energy equations fitted to critically assessed and properly weighted experimental data. Given the availability of correlated data for the family of n-alkanes, the use of such data of high accuracy is preferred over actual experimental data of unknown uncertainty and/or quality.

The quality of the description of the pure component vapour-liquid equilibria of the n-alkanes included in the regression is depicted for the coexistence densities in figure 4.2 and for the vapour pressure in figure 4.3. From the figures, it can be seen that the SAFT-γ Mie group contribution approach allows for an excellent description of the phase behaviour of the correlated compounds, from ethane to n-decane. The deviations (%AAD) from the experimental data for the vapour pressure and saturated liquid density are summarised in table 4.4. The average deviation for all correlated compounds was found to be 1.55% for p_{vap} and 0.59% for ρ_{sat} and shows a significant improvement when compared to the results

Table 4.4: Percentage average absolute deviations (%AAD) of vapour pressures $p_{\text{vap}}(T)$ and saturated liquid densities $\rho_{\text{sat}}(T)$ for the n-alkanes obtained with the SAFT-γ Mie framework for the correlated experimental data from NIST [277], where n is the number of data points.

Compound	T range (K)	n	%AAD $p_{\text{vap}}(T)$	T range (K)	n	%AAD $\rho_{\text{sat}}(T)$
C_2H_6	125-275	31	2.24	125-275	31	1.48
C_3H_8	147-332	38	2.22	147-332	38	0.74
C_4H_{10}	170-385	44	1.27	170-385	44	0.37
C_5H_{12}	187- 422	48	1.90	187- 422	48	0.36
C_6H_{14}	201-456	52	1.68	201-456	52	0.27
C_7H_{16}	216-486	55	1.01	216-486	55	0.46
C_8H_{18}	227-512	58	1.22	227-512	58	0.54
C_9H_{20}	237-532	60	0.69	237-532	60	0.59
$C_{10}H_{22}$	245-555	63	1.75	245-555	63	0.52
Average	-	-	1.55	-	-	0.59

obtained by Lymperiadis *et al.* [19] for models based on the square-well potential (3.98% for p_{vap} and 0.57% for ρ_{sat}), especially for the vapour pressure. The level of accuracy of the SAFT-γ Mie approach constitutes a clear improvement in the description of the pure component phase behaviour of the n-alkanes with a Mie potential when compared to other group contribution approaches within SAFT. Tamouza *et al.* [18] have reported average deviations of 2.63% for the vapour pressure and 2.29% for the saturated liquid density with their homonuclear GC approach based on the SAFT-0 [167] and 1.66% and 2.23% with that based on SAFT-VR SW, while for the hetero GC-SAFT-VR of Peng *et al.* [20] deviations of 5.95% and 3.07% are found for the vapour pressure and saturated liquid density, respectively. It should be noted however, that though a comparison of the deviations of different methods provide a measure of the accuracy of each approach, they are based on different experimental data and temperature ranges, as well as, in some cases, different sets of compounds.

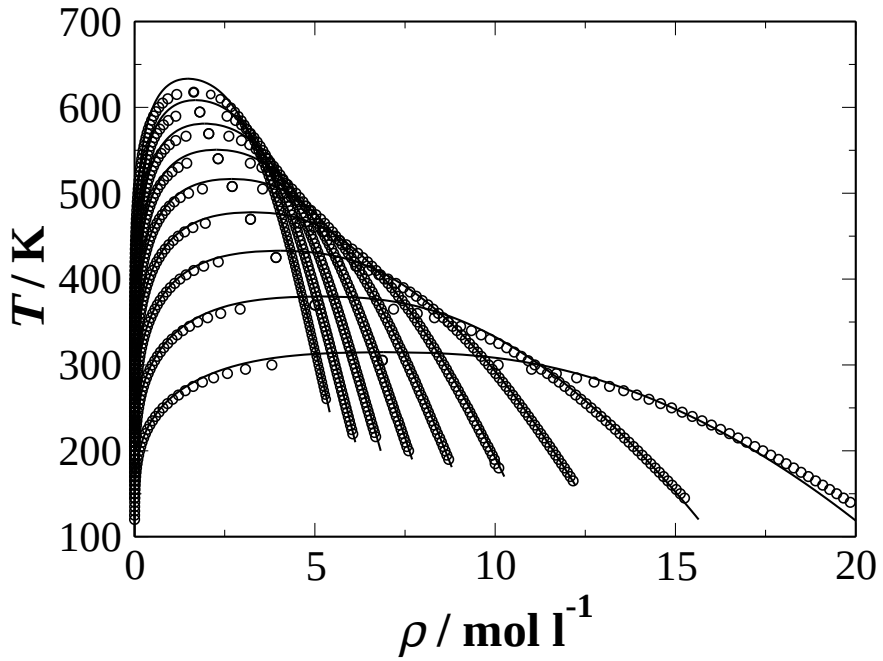

Figure 4.2: SAFT-γ Mie description of the coexistence densities as a function of temperature for the linear alkanes (n-ethane to n-decane from bottom to top) included in the estimation of the CH_3 and CH_2 group parameters presented in this work. The symbols represent the experimental data from NIST [277] and the continuous curves the calculations with the theory.

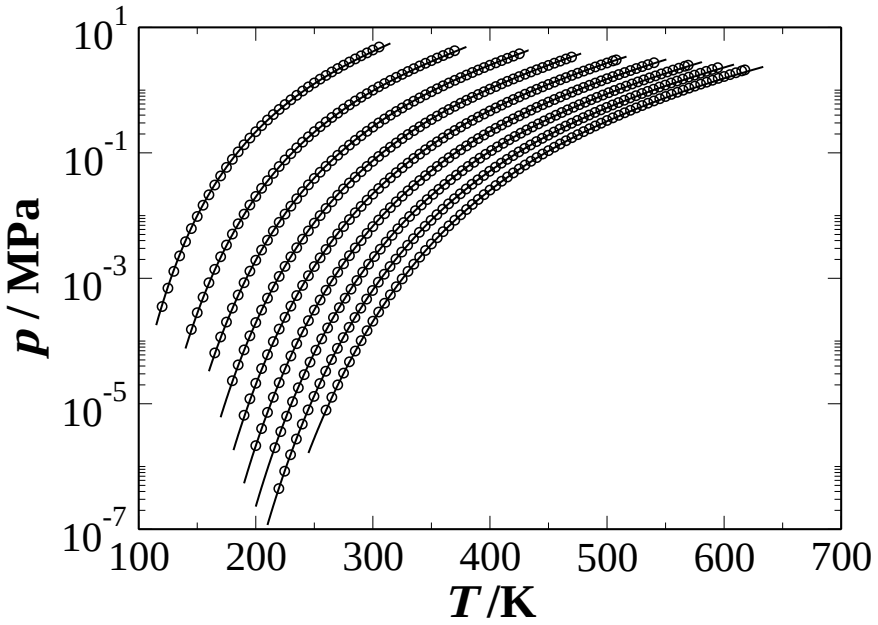

Figure 4.3: SAFT-γ Mie description of the vapour pressures for the linear alkanes (n-ethane to n-decane from left to right) included in the estimation of the CH_3 and CH_2 group parameters presented in this work. The symbols represent the experimental data from NIST [277], with the last point being the critical point for each compound, and the continuous curves the calculations with the theory.

It is also important to note that the deviations reported here are based on data within a temperature range up to 0.9 $T_{\text{crit}}^{\text{exp}}$ and as such fail to express the improved performance of the SAFT-γ Mie approach in the description of the near-critical region of pure substances. This can be seen when comparing graphically the vapour-liquid equilibria of selected n-alkanes obtained with our SAFT-γ Mie approach and the SAFT-γ SW of Lymperiadis *et al.* (cf. figure 4.4). From the comparison, it can be clearly seen that the SAFT-γ Mie EoS provides a significantly improved description of the near-critical region of systems of varying chain length, decreasing the overshoot of the critical point whilst retaining a highly accurate description of the fluid phase behaviour at temperatures far from the critical point. We reiterate that the approach, being an analytical mean-field type theory, fails to reproduce the critical scaling observed experimentally. An accurate representation of these cirtical properties would require the application of a renormalisation group treatment, e.g., as applied to the SAFT-VR EoS by McCabe *et al.* [323] and by Forte *et al.* [324, 325].

The description of the single phase densities ρ_{liq} summarised in table 4.5, is also excellent, with an average deviation for all compounds of 0.59%. When examining each compound it can be seen that the highest deviations are observed for ethane, as was the case for the

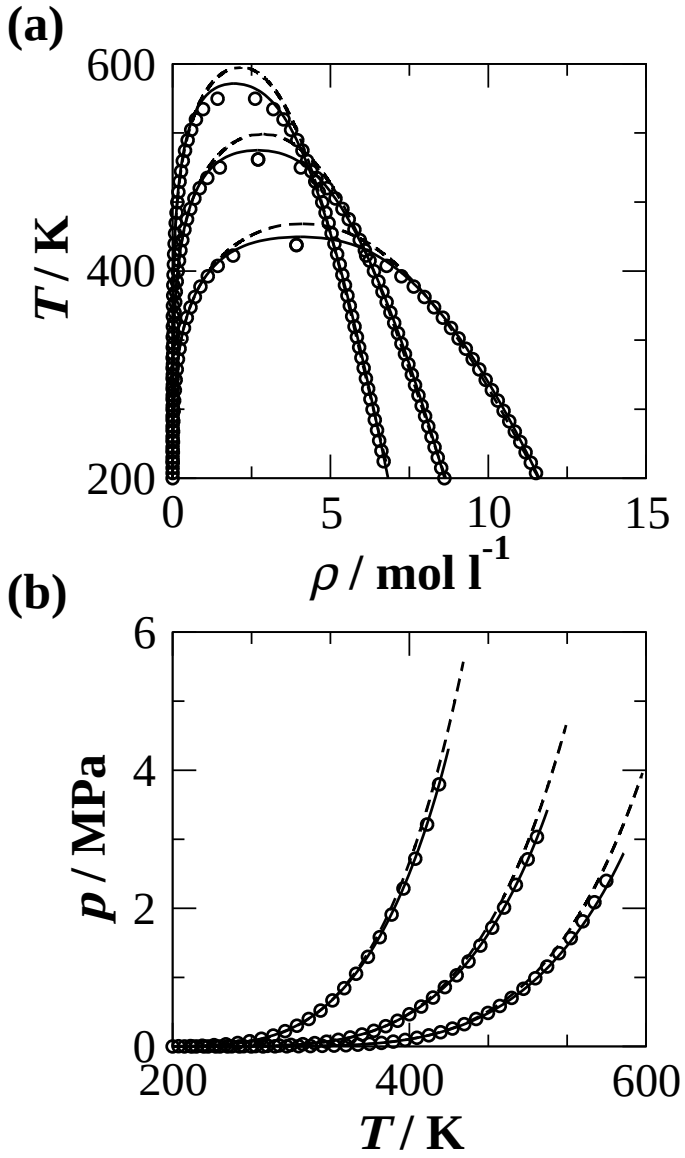

Figure 4.4: Comparison of the description of SAFT-γ SW [19] (dashed curves) and SAFT-γ Mie (solid curves) for the pure component vapour-liquid equilibria of n-butane, n-hexane and n-octane. (a) Coexistence densities (from bottom to top) and (b) vapour pressures (from left to right). The symbols represent the experimental data from NIST [277].

SAFT-γ square-well approach [19]. One may argue that ethane should not be included in the parameter estimation procedure, since ethane does not contain a CH_3-CH_2 interaction, while all longer n-alkanes do. Furthermore, group contribution techniques are not generally best suited to the study of small molecules as proximity effects are neglected. With this in mind the procedure was repeated omitting the data for ethane and the inclusion of experimental data for ethane was not found to bias the parameters of the groups in a way that would result in a sub-optimal description of the higher n-alkanes. When

Table 4.5: Percentage average absolute deviations (%AAD) of single-phase densities $\rho_{\text{liq}}(T,P)$, speed of sound $u(T,P)$, and isobaric heat capacity $c_p(T,P)$ for the n-alkanes obtained with the SAFT-γ Mie framework from the experimental data from NIST [277], where n is the number of data points.

Compound	T [K]	p [MPa]	n	%AAD $\rho_{\text{liq}}(T,p)$	%AAD $u(T,p)$	%AAD $c_p(T,p)$
C_2H_6	150-550	10-50	123	0.96	3.04	3.59
C_3H_8	150-500	10-50	108	0.49	1.65	2.19
C_4H_{10}	150-550	10-50	123	0.50	0.79	2.04
C_5H_{12}	150-550	10-50	120	0.60	1.08	1.76
C_6H_{14}	188-548	10-50	108	0.52	0.88	1.21
C_7H_{16}	193-553	10-50	108	0.62	1.42	0.99
C_8H_{18}	226-546	10-50	95	0.64	0.77	0.61
C_9H_{20}	230-550	10-50	95	0.50	2.12	0.48
$C_{10}H_{22}$	253-553	10-50	89	0.47	2.30	0.41
Average	-	-	-	0.59	1.56	1.48

a more accurate model is needed ethane can be described by a separate functional group (as would be the case for, e.g., water, methanol, etc.). This strategy is followed within the hetero GC-SAFT-VR of Peng *et al.* [20] and the GC-SAFT of Tamouza *et al.* [18] where n-propane was the first molecule considered for the regression of the parameters for the n-alkane series.

As has already been mentioned, a particular advantage of the Mie potential is that the form of the pair interactions between segments can be carried out by adjusting the values of the repulsive and attractive exponents. This allows one to capture the finer features of the interaction which are important in providing an accurate description of the thermodynamic derivative properties. Regarding the parameters that determine the form of the interaction potential between the CH_3 and CH_2 functional groups, the value of the attractive exponent is fixed to $\lambda^a = 6$ for both groups. This is based on the chemical nature of the groups, which given their non-polar nature are expected to interact via London dispersion forces which are characterised by an attractive exponent of six [112, 349]. The values of the repulsive exponents estimated from the experimental data for the n-alkanes are found to be $\lambda^r_{CH_3} = 15.050$ and $\lambda^r_{CH_2} = 19.871$. Potoff and Bernard-Brunel [238] have developed a force-field for the methyl and methylene groups based on the Mie potential for molecular simulation of the fluid phase behaviour of n-alkanes, finding that a good description can be obtained with values of the repulsive exponent of 16 for both groups. The values of the shape factors for the CH_3 and CH_2 groups are also obtained from the parameter

estimation procedure. The optimal values obtained are $S_{CH_3} = 0.5725$ and $S_{CH_2} = 0.2293$, which are found to be quite different from those used in previous work with the square-well potential [19]. In the previous work, the shape factors for the CH_3 and CH_2 groups were fixed to 1/3 and 2/3, respectively, as these yield a realistic molecular aspect ratio for the *n*-alkanes in a homonuclear model [27]. The segment size for the CH_3 is found to be smaller than the size of the CH_2 group (4.077 Å compared to 4.880 Å). Comparing the values of the segment diameters to the ones obtained by Lymperiadis *et al.* [19], one finds a significant increase; this will be discussed in more detail when examining the predictions of the theory for longer *n*-alkanes and the polyethylene polymer limit in section 4.4.2.1.

4.4.1.2 2-Ketones

Having determined the parameters for the CH_3 and CH_2 functional groups, these are transferred for the study of other homologous series of compounds to determine the parameters of additional functional groups. In the current work we develop parameters for the CH_3CO group based on experimental data for the 2-ketones. The chemical family of the terminal ketones (or 2-ketones) is modelled by defining the CH_3CO functional group, in accordance with the approach employed within UNIFAC [78]. Given the fact that this large functional group comprises the first neighbouring methyl group, we choose to model CH_3CO as comprising two identical segments ($\nu^*_{CH_3CO} = 2$), and treat it as non-associating as ketones are not expected to self-associate. It is important to note however, that associating sites will have to be included in order to capture the unlike association in some polar mixture of ketones, including, e.g., water or alkanols (see the work of Kleiner and Sadowski [350]). The value of the attractive exponent is again fixed to the value of $\lambda^a_{CH_3CO} = 6$, as for the groups of the *n*-alkanes, and the remaining parameters are estimated based on regression to a compilation of pure component and mixture experimental data. The pure component data used include experimental data for the saturated liquid density, vapour pressure and single phase liquid density for 2-propanone (acetone) to 2-nonanone. The specifics of the temperature and pressure range, the number of points for each property, as well as the deviations for each property and each compound are given in tables 4.6 and 4.7, and the estimated parameters are presented in table 4.2. In addition to the pure component data, we have included in the regression experimental data for the heats of mixing of three binary *n*-alkane+2-ketone mixtures: *n*-propane+acetone, *n*-hexane+acetone and *n*-octane+2-butanone. These experimental data are included in the estimation procedure so as to have a more complete characterisation of the unlike interactions between functional groups and to ensure an accurate description of the highly non-ideal phase behaviour that these mixtures exhibit. An advantage of the SAFT-γ *heteronuclear* approach is that in-

Table 4.6: Percentage average absolute deviations (%AAD) of vapour pressures $p_{\text{vap}}(T)$ and saturated liquid densities $\rho_{\text{sat}}(T)$ for the 2-ketones obtained with the SAFT-γ Mie framework for the experimental data, where n is the number of data points.

Compound	T [K]	n	%AAD p_{vap}	Ref.	T [K]	n	%AAD ρ_{sat}	Ref.
CH_3COCH_3	259-453	46	0.85	[293, 351]	183-453	15	2.21	[352, 353]
$CH_3CH_2COCH_3$	253-483	34	0.5	[354, 355]	213-483	11	1.01	[352]
$CH_3(CH_2)_2COCH_3$	263-505	43	5.37	[355, 356]	233-503	11	0.81	[352]
$CH_3(CH_2)_3COCH_3$	307-427	30	0.97	[354]	250-520	28	0.99	[357]
$CH_3(CH_2)_4COCH_3$	273-452	57	1.9	[354, 355]	250-550	31	0.89	[357]
$CH_3(CH_2)_5COCH_3$	263-461	30	2.96	[355]	260-570	32	0.96	[357]
$CH_3(CH_2)_6COCH_3$	283-466	20	2.51	[358]	270-590	33	1.66	[357]
Average	-	-	2.15	-	-	-	1.22	-

Table 4.7: Percentage average absolute deviations (%AAD) of single-phase densities $\rho_{\text{liq}}(T, p)$ for the 2-ketones obtained with the SAFT-γ Mie framework for the experimental data, where n is the number of data points. No data for 2-nonanone were available.

Compound	T [K]	p [MPa]	n	Ref.	%AAD $\rho_{\text{liq}}(T, p)$
CH_3COCH_3	323-423	1.6-65.5	56	[359]	0.97
$CH_3CH_2COCH_3$	293-473	0.1-160	16	[360]	1.80
$CH_3(CH_2)_2COCH_3$	273-473	0.1-78	38	[361]	2.22
$CH_3(CH_2)_3COCH_3$	292-304	0.1	11	[362]	0.79
$CH_3(CH_2)_4COCH_3$	293-360	0.1	6	[363]	0.94
$CH_3(CH_2)_5COCH_3$	273-566	0.1-79	46	[364]	1.26
Average	-	-	-	-	1.33

formation on the unlike interactions can be provided from pure component data alone, as has been successfully demonstrated in the past [19, 20, 29, 222, 348]. This is particularly useful in situations where data availability is an issue. However, in the current work we have chosen to include some experimental data for selected mixtures in order to increase the statistical significance of the interaction parameters.

A comparison between the SAFT-γ Mie description and the experimental data for the saturated liquid densities and vapour pressures of the 2-ketones is given in figures 4.5 and 4.6. The vapour pressures are shown in a logarithmic representation that highlights both the low- and the high-temperature region. The overall deviations (%AAD) for the correlated compounds (cf. table 4.6) are 2.15% for p_{vap} and 1.22% for ρ_{sat}. Regarding the reported deviations for the vapour pressures listed in table 4.6, a rather high value

is seen for the case of 2-pentanone; similar findings were reported with the SAFT-γ SW method [29], where for an average error in pressure over all correlated compounds of 3.47% the reported deviation for 2-pentanone was 7.44%. The experimental data used for 2-pentanone were assessed for consistency with an Antoine equation and parameters from reference [39], where a deviation of 1.64% is seen. This suggests that the experimental data used for 2-pentanone in the parameter estimation procedure are consistent.

A comparison of the SAFT-γ Mie description of the family of 2-ketones with the deviations reported by Lymperiadis *et al.* [29] suggests a clear improvement for the vapour pressure (%AADs of 3.37% for p_{vap} and 1.13% for ρ_{sat}). The difference in the description of the saturated liquid densities can be attributed to the different experimental data and temperature ranges used in the estimation; when comparing the predictions of the theory to the experimental data used by Lymperiadis *et al.* [29] a similar accuracy is found. Regarding the performance of other GC SAFT approaches for the family of ketones, Nguyen-Huyhn *et al.* [365] have applied three versions of SAFT (the SAFT-0 [167], SAFT-VR [170, 171] and PC-SAFT [166]) combined with a polar term and the deviations reported for the correlated compounds (acetone to 2-octanone) were, respectively, 3.50%, 3.13% and 3.00% for p_{vap} and 1.45%, 1.49% and 2.15% for ρ_{sat}. Although a precise comparison with the other methods is difficult due to the differences in the experimental data used, the average error for our SAFT-γ Mie approach based on the same compounds, are of 2.09 % and 1.14 % for p_{vap} and ρ_{sat}, respectively, which suggests a slight improvement. A comparison with the hetero GC-SAFT-VR [20] approach is somewhat more difficult, as the parameters for the C=O group of the ketones of Peng *et al.* are obtained based on intermediate ketones, as opposed to the terminal ketones employed in our current work. The accuracy in the description of the pure component fluid phase behaviour of the ketones is reported as of 5.75 % for p_{vap} and 1.57 % for ρ_{sat}.

The description of the single-phase densities of the 2-ketones used in the parameter estimation is found to be of high accuracy, with an average %AAD for all compounds considered of 1.33%. On analysing the deviations per compound, it can be seen that the highest deviations from the experimental data are found for the first member of the 2-ketones series, i.e., acetone, especially for the saturated liquid densities. An optimal description for acetone can be obtained by defining it as a single group (as would be the case for water, methanol, etc.). However, the fact that a satisfactory agreement can still be achieved within the group contribution approach and taking into account the importance of acetone as a solvent, we chose to include it in the regression procedure.

The limited mixture data included in the regression are also well correlated with SAFT-γ Mie, as shown in figure 4.7. The average deviation (%AAD) of the excess enthalpies for the three binary systems included in the estimation was found to be 14.37% (12.93% for *n*-propane+acetone, 17.41% for *n*-hexane+acetone, and 12.78% for *n*-octane+2-butanone), and the level of agreement was deemed satisfactory based on the sensitivity of the selected properties to the unlike interaction parameters between the functional groups and the small values of the excess enthalpies that the systems exhibit ($\sim$ 0.5-2 kJ/mol). In order to assess the level of accuracy, the description of the SAFT-γ Mie approach was compared with the description obtained with the modified UNIFAC (Dortmund) [85], within which group parameters are obtained by regression to data for different mixture properties including phase equilibria and excess enthalpies. From the comparison shown in figure 4.7, it is clear that the two methodologies provide an equivalent description of the experimental data. The modified UNIFAC gives a better description of the composition dependence of the excess enthalpies but it appears to underestimate the data in all cases, while the SAFT-γ Mie is seen to provide a better description of the magnitude of the excess enthalpies and the relative difference in the excess enthalpies of the mixtures examined.

The resulting form of the interaction potential of the CH_3CO group is described by an

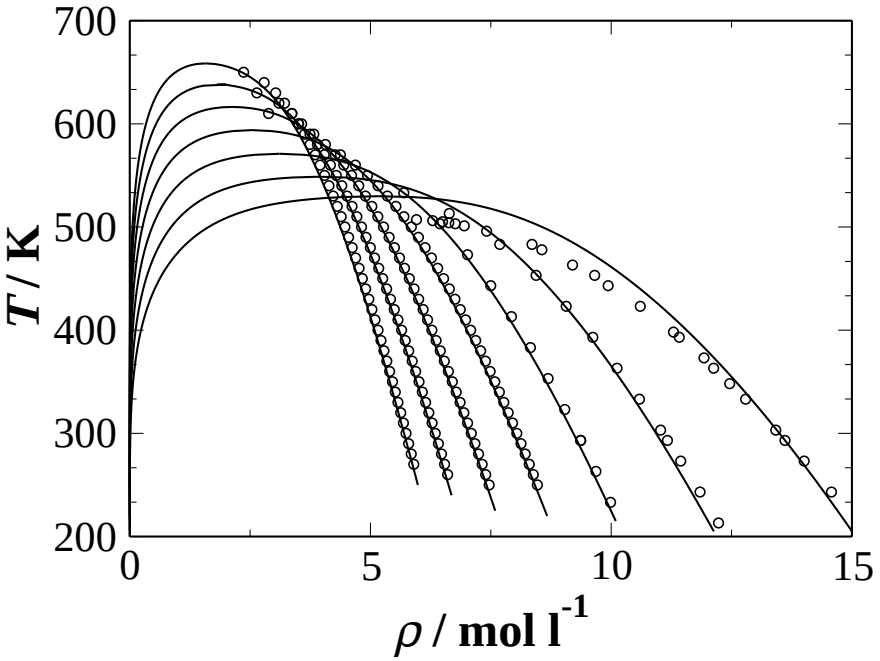

Figure 4.5: SAFT-γ Mie description of the coexistence densities as a function of temperature for the 2-ketones (2-propanone to 2-nonanone from bottom to top) included in the estimation of the CH_3CO group parameters. The symbols represent experimental data and the solid curves the calculations with the theory.

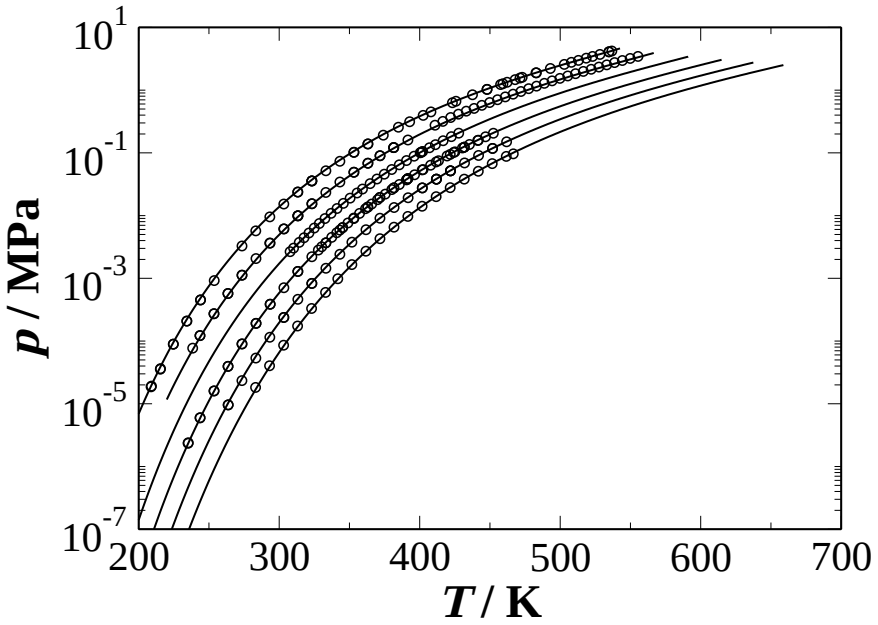

Figure 4.6: SAFT-γ Mie description of the vapour pressures for the 2-ketones (2-propanone to 2-nonanone from top to bottom) included in the estimation of the CH_3CO group parameters. The symbols represent experimental data, and the solid curves the calculations with the theory.

optimal value of the repulsive exponent of $\lambda^r_{CH_3CO} = 17.508$, while the attractive exponent is fixed to $\lambda^a_{CH_3CO} = 6$. No explicit treatment of the polarity of the ketone is included in the model developed within the current work; a term accounting for polar interactions has been considered in other modelling approaches for the ketones [365]. We should note however that in the case of the interaction of the CH_3CO group with other polar groups, such as the OH, NH_2, or H_2O groups, the strong electrostatic interactions are represented with additional site-site association contributions. The value of the segment size of the CH_3CO group obtained from the regression is $\sigma_{CH_3CO} = 3.748$ Å, which is slightly smaller than the size of the CH_3 group; bearing in mind that the CH_3CO group is modelled as two identical segments and the similar value of the shape factor of the CH_3CO group, it would appear that the CH_3CO group is roughly twice as big as the CH_3 group, which is in line with previous findings [20, 189, 365] where the segment size of the C=O group is found to be very similar to that of the CH_3 group. The energy of interaction characterising the CH_3CO group is found to be rather high, $\epsilon_{CH_3CO}/k_B = 463.412$, which is again consistent with previously reported values on this interaction within the SAFT-γ SW [29] and the hetero GC-SAFT-VR [20] frameworks.

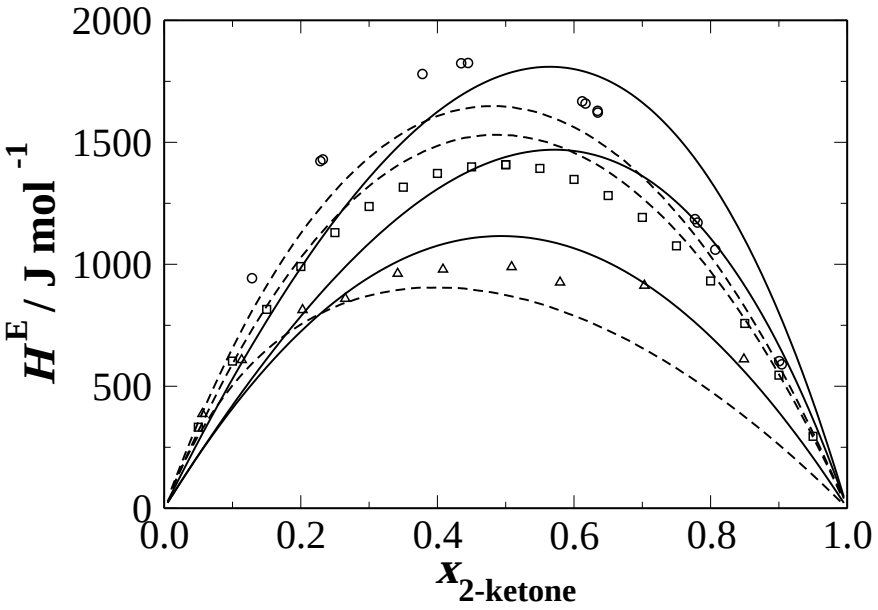

Figure 4.7: Excess enthalpies as a function of the composition of the 2-ketone of the binary systems included in the regression of the parameters of the CH_3CO group. The triangles are experimental data for the system of *n*-propane+acetone at 223.15 K [366], the squares for *n*-octane+2-butanone at 298.15 K [367], and the circles for *n*-hexane+acetone at 308.15 K [368]. The continuous curves are the calculations of the SAFT-γ Mie EoS and the dashed curves are the corresponding calculations of the modified UNIFAC (Dortmund) [85].

4.4.2 Predictions

4.4.2.1 Pure Components

An assessment of the adequacy of the parameters obtained for the CH_3 and CH_2 functional groups can be made by examining the performance of the theory in describing thermodynamic properties of compounds not included in the parameter estimation procedure. The application of the Mie intermolecular potential was previously shown to allow for an accurate simultaneous description of fluid phase equilibria and thermodynamic derivative properties [30, 31]. The prediction of derivative properties is a stringent test in any thermodynamic description, and high deviations are to be expected with most equations of state, be it with traditional cubic EoSs or the various SAFT variants [30, 31]. Here, we examine the performance of the SAFT-γ Mie approach in the description of thermodynamic derivative properties of representative n-alkanes. We compare the predictions obtained with our current approach to the correlated experimental data of NIST [277] for the speed of sound $u(T,p)$ and isobaric heat capacity $c_p(T,p)$, of the compounds included in the parameter estimation procedure (cf. table 4.5). The reported average deviations, 1.56% for $u(T,p)$ and 1.48% for $c_p(T,p)$ for the n-alkane homologous series from ethane to n-decane provide a good indication of the capability of the SAFT-γ Mie EoS to describe thermodynamic derivative properties whilst retaining an excellent description of the pure component fluid phase behaviour. The performance of SAFT-γ Mie in the prediction of thermodynamic derivative properties of selected n-alkanes is also depicted in figure 4.8, where in addition to the speed of sound and the isobaric heat capacity, the SAFT-γ Mie predictions for the isothermal compressibility $k_T(T,p)$ and the Joule-Thomson coefficient $\mu_{\mathrm{JT}}(T,p)$ are compared to correlated experimental data [369].

A key advantage of a group contribution formulation lies in the transferability of the group parameters. In order to demonstrate this, the parameters for the CH_3 and the CH_2 functional groups are applied to the prediction of pure component properties of systems not used to estimate the group parameters. Predictions with the SAFT-γ Mie EoS of the fluid phase behaviour for some heavier n-alkanes (n-$C_{15}H_{32}$, n-$C_{20}H_{42}$, n-$C_{25}H_{52}$ and n-$C_{30}H_{62}$) are compared with the experimental data in figure 4.9, and the corresponding deviations are summarised in table 4.8. From this analysis it is clear that the methyl and methylene group parameters can be successfully transferred to the prediction of the pure component phase behaviour of longer n-alkanes; the description of the saturated liquid density remains good, whereas higher deviations are seen in the vapour pressure. The latter could be related to the high uncertainty and low absolute values of the experimental data of the saturation pressure of long compounds at low temperatures. Furthermore apart from the

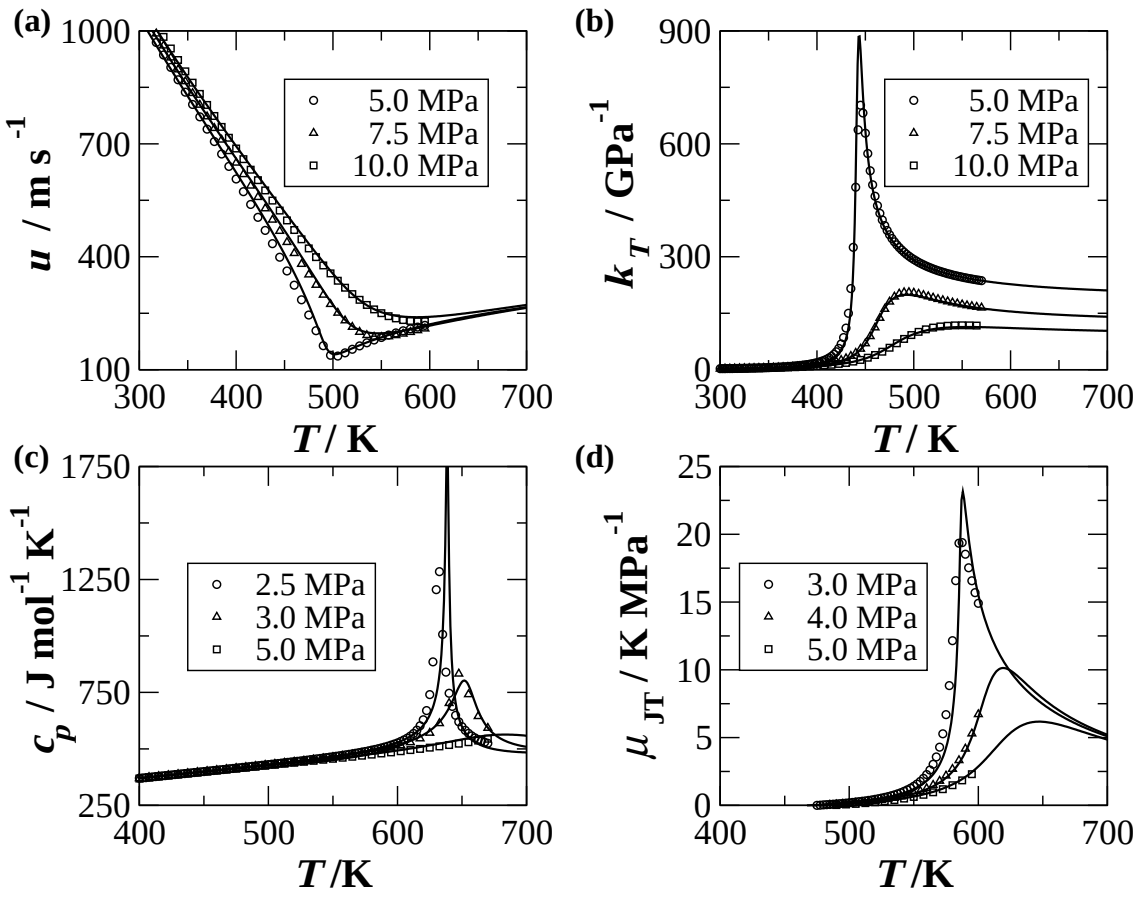

Figure 4.8: Prediction of thermodynamic derivative properties of selected n-alkanes with the SAFT-γ Mie approach: (a) speed of sound of n-pentane; (b) isothermal compressibility of n-butane; (c) isobaric heat capacity of n-decane; and (d) Joule-Thomson coefficient of n-octane, where the symbols are correlated experimental data from REFPROP [369] and the continuous curves the theoretical predictions.

phase behaviour, the CH_3 and CH_2 group parameters result in a satisfactory description of derivative properties of long n-alkanes, as shown in figure 4.10, for the the speed of sound of n-$C_{15}H_{32}$ [327] and the isothermal compressibility of n-$C_{20}H_{42}$ [326].

The formulation of a SAFT-type approach within a group contribution concept combined the wide range of applicability of EoSs and the predictive nature of GC methods. This can be illustrated in the description of the properties of polymers. The SAFT EoS has been successfully applied as a general methodology to the study of polymeric systems (e.g., see [186, 370, 371]), where the theory is particularly suited to polymers due to the explicit consideration of chain formation. Within the scope of the current work, and taking advantage of the group contribution formulation, the properties of polyethylene polymers

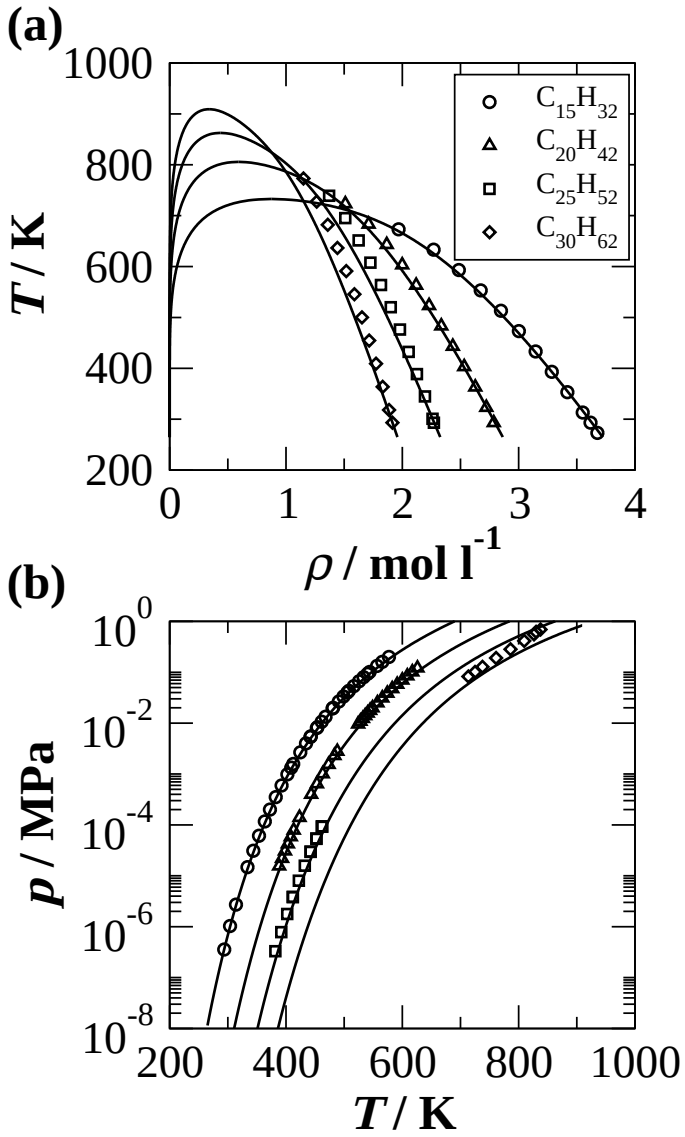

Figure 4.9: Comparison of the predictions of the SAFT-γ Mie approach and the experimental data for the pure component vapour-liquid equilibria of long-chain n-alkanes not included in the estimation of the group parameters. (a) Saturated liquid densities and (b) vapour pressures of n-pentadecane (n-$C_{15}H_{32}$), n-eicosane (n-$C_{20}H_{42}$), n-pentacosane (n-$C_{25}H_{52}$), and n-triacontane (n-$C_{30}H_{62}$).

care predicted, using the parameters estimated for the methyl and methylene groups from the lower n-alkanes (cf. section 4.4.1.1). The predictions of the SAFT-γ Mie approach are compared with the experimental single-phase densities of linear polyethylene [372] in figure 4.11. The parameters for the methyl and (predominantly) for the methylene group are seen to allow for an accurate description of linear polyethylene with a molecular weight of 126,000 g/mol over a wide range of temperatures and pressures. The level of agreement with the experimental data is particularly pleasing, especially when one recalls that

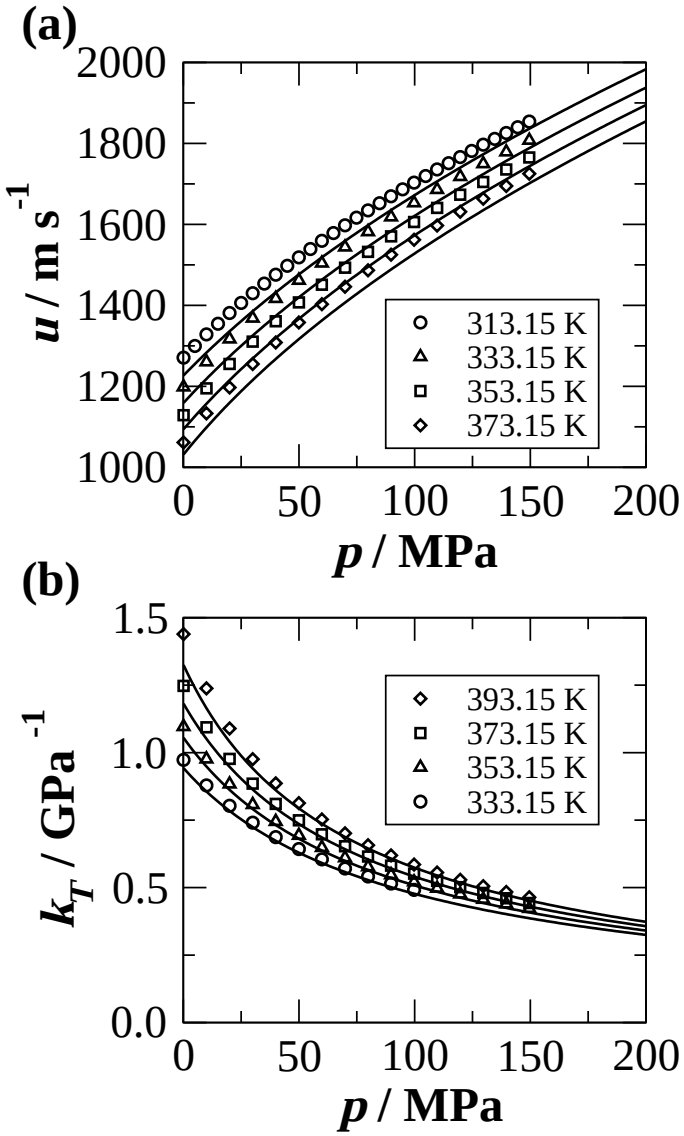

Figure 4.10: Comparison of the predictions of the SAFT-γ Mie approach and the experimental data for thermodynamic derivative properties of long-chain n-alkanes not included in the regression of the group parameters: (a) speed of sound of n-pentadecane at 313.15 K (circles), 333.15 K (triangles), 353.15 K (squares), 373.15 K (diamonds) [327]; and (b) isothermal compressibility of n-eicosane at 333.15 K (circles), 353.15 K (triangles), 373.15 K (squares) and 393.15 K (diamonds) [326].

the parameters for the CH_3 and CH_2 groups are obtained based on the properties of the much shorter n-alkanes, ranging from ethane to n-decane. The description obtained with the SAFT-γ SW method using published parameters [19] are also plotted in figure 4.11. From the comparison it is apparent that the description of the polymer with Mie segments provides a marked improvement to that obtained with square-well segments, and lends physical relevance to the values of the parameters for the CH_3 and CH_2 groups (cf.

table 4.2).

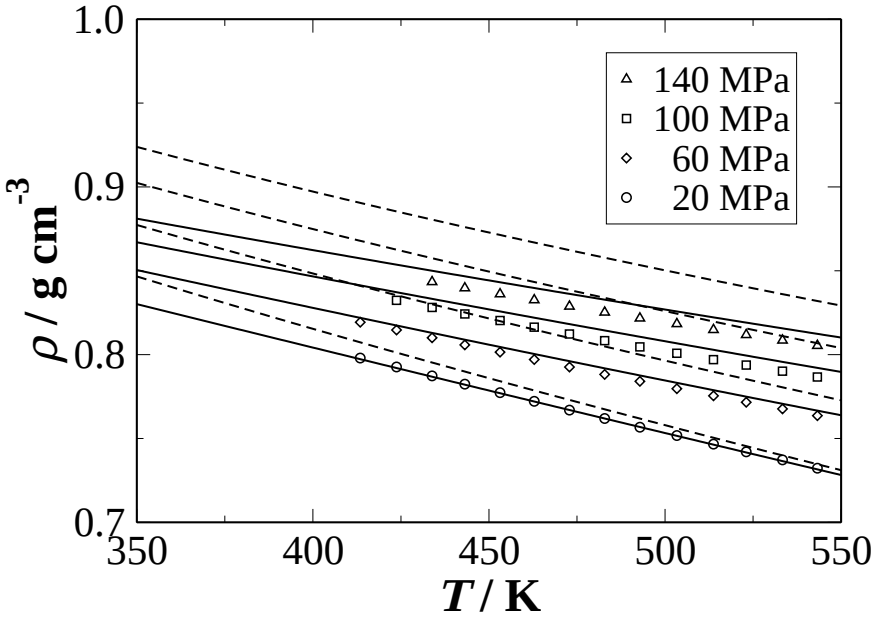

Figure 4.11: Comparison of the predictions of the SAFT-γ Mie approach (solid lines) and the SAFT-γ SW models (dashed lines) with the experimental data of the single phase densities of linear polyethylene [372] (MW = 126,000 g/mol).

A similar assessment is performed for the chemical family of the 2-ketones, where the performance of the methodology in the prediction of derivative properties is examined. In accordance with the findings reported in section 4.4.1.2 for the fluid phase equilibria, high deviations are expected for the case of acetone in the description of thermodynamic derivative properties. A high-fidelity model that accurately reproduces a wide range of properties of acetone has to be developed on a molecular-group basis. The predictions of the SAFT-γ Mie EoS are compared with the experimental data for the speed of sound of 2-butanone, and with the isothermal compressibility of 2-octanone in figure 4.12. From figure 4.12.(a) it can be seen that for 2-butanone the predictions of the theory are in qualitative agreement with the experimental data, however the level of accuracy is not as satisfactory as for the *n*-alkanes (cf. table 4.5 and figure 4.8). For increasingly long 2-ketones, the effect of the polar CH_3CO group decreases in importance, and as a consequence a very accurate representation of thermodynamic properties is recovered, as shown for the isothermal compressibility of 2-octanone, shown in figure 4.12.(b).

An assessment of the SAFT-γ Mie EoS in the prediction of the pure component phase behaviour for longer compounds not included in the parameter estimation procedure has been performed for 2-ketones. The parameters for the CH_3, CH_2 and CH_3CO group have been

combined to predict the phase behaviour of two high molecular weight 2-ketones, namely 2-undecanone and 2-tridecanone, and the deviations from the experimental data [354], available for the vapour pressure only, are summarised in table 4.8. It is apparent that the obtained group parameters can be successfully extrapolated to the accurate prediction of the pure component VLE of these longer compounds.

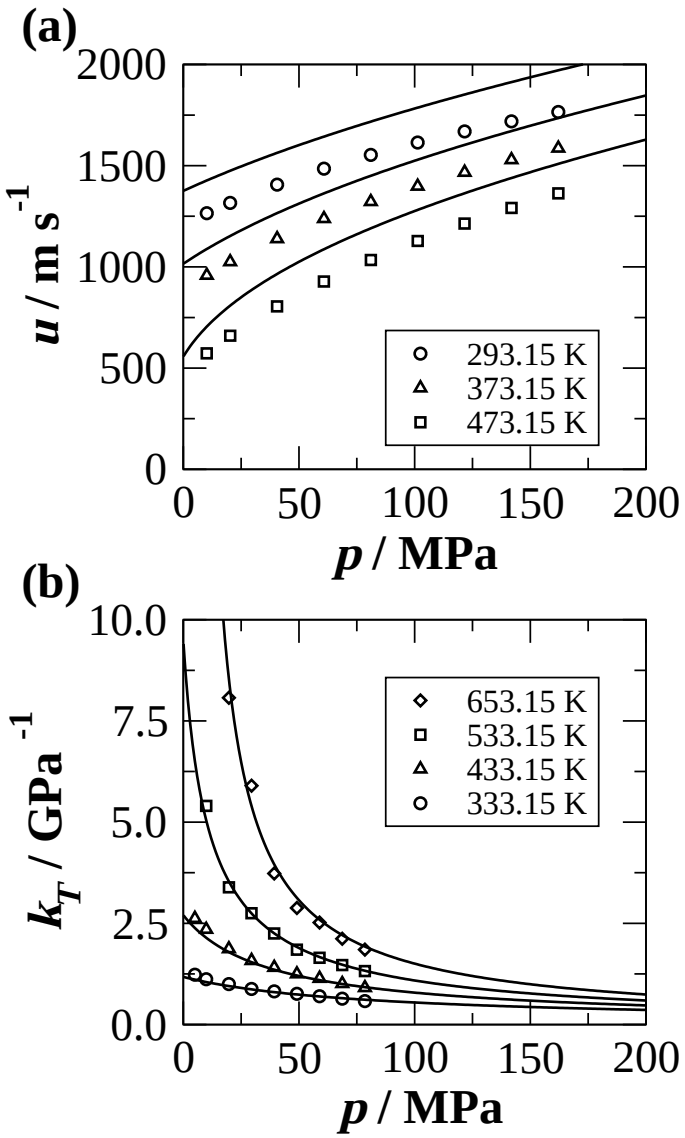

Figure 4.12: Prediction of thermodynamic derivative properties of selected 2-ketones with the SAFT-γ Mie approach: (a) speed of sound of 2-butanone at 293.15 K (circles), 373.15 K (triangles), and 473.15 K (squares) [373]; and (b) isothermal compressibility of 2-octanone at 333.15 K (circles), 433.15 K (triangles), 533.15 K (squares), and 633.15 K (diamonds) [374]. The continuous curves are the predictions with the SAFT-γ Mie approach.

Table 4.8: Percentage average absolute deviations (%AAD) of vapour pressures $p_{\text{vap}}(T)$ and saturated liquid densities $\rho_{\text{sat}}(T)$ for the predictions of SAFT-γ Mie EoS from the experimental data (where n is the number of data points) for long-chain n-alkanes and 2-ketones not included in the estimation of the group parameters.

Compound	T [K]	n	%AAD $p_{\text{vap}}(T)$	Ref.	T [K]	n	%AAD $\rho_{\text{sat}}(T)$	Ref.
$C_{15}H_{32}$	293-576	35	6.67	[375],[376]	273-633	11	0.70	[352]
$C_{20}H_{42}$	388-625	29	16.98	[377]	293-683	11	0.98	[352]
$C_{25}H_{52}$	381-461	13	29.85	[378]	293-695	11	3.10	[352]
$C_{30}H_{62}$	432-452	5	36.56	[378]	293-727	11	4.01	[352]
$CH_3(CH_2)_8COCH_3$	393-538	27	2.26	[354]	-	-	-	
$CH_3(CH_2)_{10}COCH_3$	424-546	15	2.18	[354]	-	-	-	
Average	-	-	15.75	-	-	-	2.20	-

4.4.2.2 Binary Systems

The ability to obtain information about the nature of the unlike dispersion interaction between segments belonging to the same molecule is an inherent advantageous feature of the SAFT-γ approach. For the specific case of the chemical families considered in this work, this means that the values of the unlike dispersion energies between the CH_3 and CH2 groups are obtained by pure component data alone (see table 4.3). This allows for calculations of phase behaviour and other properties of mixtures of n-alkanes in a purely predictive manner.

An example of the predictive capabilities of the SAFT-γ Mie approach is shown in figure 4.13, where the theoretical description is compared to experimental data for the phase behaviour of the binary system n-butane+n-decane at different temperatures [322]. It is clear that the SAFT-γ Mie predictions are in very good agreement with the experimental data for the vapour-liquid equilibria of the mixture. The improvement in the description of the near-critical fluid phase behaviour within the SAFT-γ Mie approach allows for an accurate representation of the system over a wide range of conditions, including the high-pressure critical region for isotherms at temperatures above the critical point of n-butane. The overshoot of the critical points of the phase envelopes of the binary system is significantly reduced in comparison to the results obtained with the SAFT-γ SW approach [19].

The same parameter set can be applied to the prediction of the fluid phase behaviour of binary mixtures of compounds not included in the regression of the group parameters. An example of this is shown in figure 4.14, where the theoretical predictions are

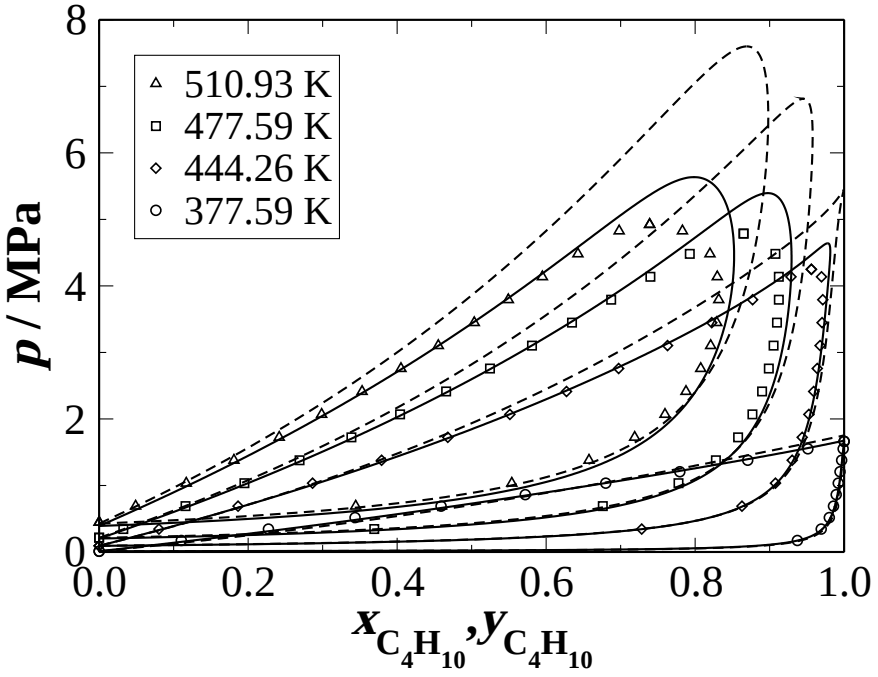

Figure 4.13: Pressure-composition ($p-x$) representation of the fluid phase behaviour (vapour-liquid equilibria) of the n-butane+n-decane binary mixture. The continuous curves represent the predictions with the SAFT-γ Mie approach, the dashed curves the corresponding predictions with SAFT-γ SW, and the symbols the experimental data at different temperatures [322].

compared with the available experimental data [379] for the liquid-liquid equilibria of the propane+n-hexacontane mixture for three different temperatures. This system was chosen as a representative example of the type of fluid phase equilibria that polymer+solvent mixtures exhibit [380]. n-$C_{60}H_{122}$ is the highest molecular weight monodisperse n-alkane commercially available, so the mixture can be modelled without the need to account for polydispersity [379]. Polydispersity has to be generally considered in representing polymer+solvent systems and it can have a significant effect, primarily in the description of liquid-liquid equilibria (e.g., see [186]). From figure 4.14 it can be seen that the SAFT-γ Mie predictions are in very good agreement with the experimentally determined phase behaviour of the mixture, accurately reproducing the width of the coexistence region at higher pressures, with only a minor overshoot of the liquid-liquid critical point. The proposed methodology provides a significantly improved description of the mixture in comparison to the results obtained with the SAFT-γ SW EoS based on the SW potential [19, 29]. It is important to note that the same parameter set can be used for the prediction of different types of phase behaviour: vapour-liquid equilibria for the n-butane+n-decane system and liquid-liquid equilibria for the propane+n-hexacontane system.

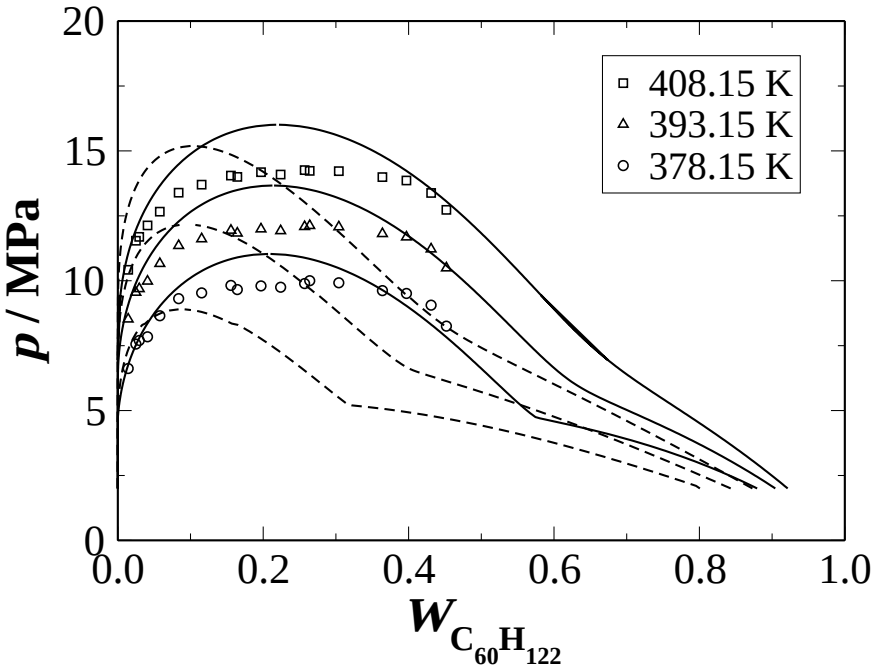

Figure 4.14: Pressure-weight fraction representation of the fluid phase behaviour (liquid-liquid equilibria) of the propane + n-hexacontane (n-$C_{60}H_{122}$) binary mixture. The continuous curves represent the predictions with the SAFT-γ Mie approach, the dashed curves the corresponding predictions with SAFT-γ SW [19], and the symbols the experimental data at different temperatures [379].

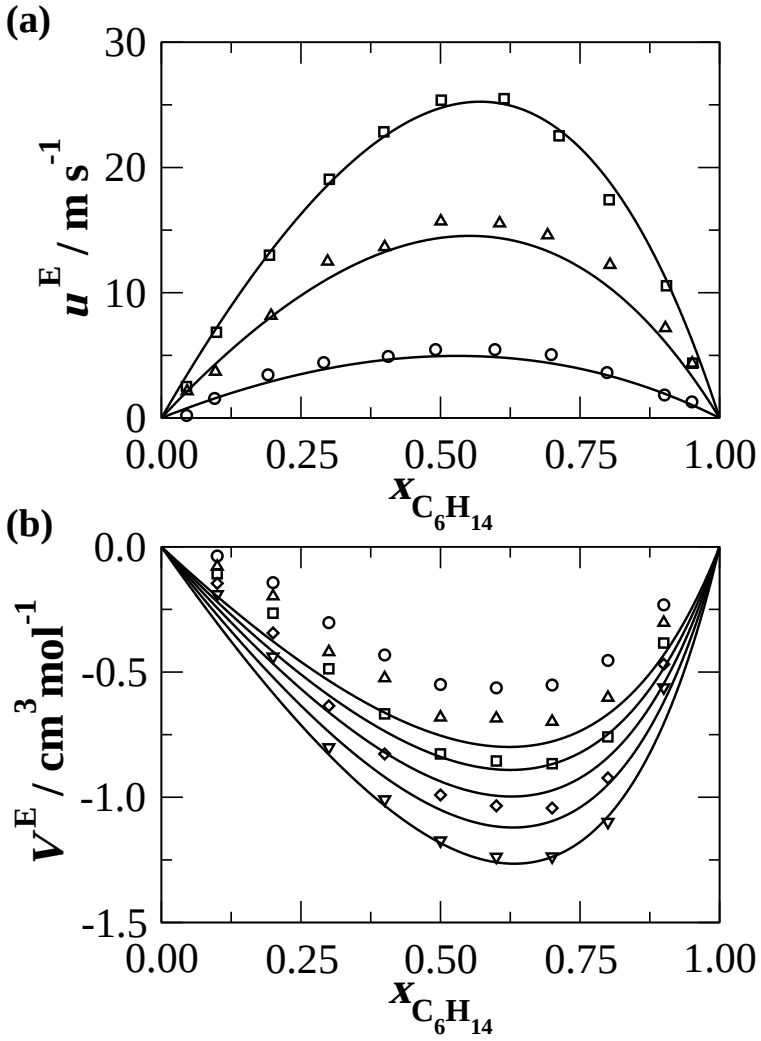

Figure 4.15: Predictions for selected excess thermodynamic properties of binary mixtures of n-alkanes: (a) excess speed of sound for n-hexane+n-dodecane (squares), n-hexane+n-decane (triangles) and n-hexane+n-octane (circles) at 298.15 K [381]; and (b) excess molar volumes for n-hexane+n-hexadecane at 293.15 K (circles), 303.15 K (triangles-up), 313.15 K (squares), 323.15 K (diamonds) and 333.15 K (triangles-down) [382]. The continuous curves represent the predictions with the SAFT-γ Mie approach.

The performance of SAFT-γ Mie has also to be assessed in the prediction of excess thermodynamic properties of binary mixtures of n-alkanes. Excess thermodynamic properties are a stringent test in the validation of thermodynamic methodologies, as they tend to be more sensitive to the molecular details and intermolecular parameters than fluid phase equilibria. Properties such as the excess volume and enthalpies have been the subject of numerous studies, including among others the seminal work of Flory and co-workers [383–385] and more recently the work of Blas and co-workers [177, 386, 387], within the SAFT framework. An example of the performance of the SAFT-γ Mie approach is illustrated in

figure 4.15, where the predictions of the theory are compared to experimental data for the excess speed of sound (u^E) of selected n-hexane+n-alkane binary mixtures [381] and the excess molar volumes (V^E) of the binary mixture of n-hexane+n-hexadecane at different temperatures [382]. As for the fluid phase behaviour, the predictions of the theory are in very good agreement with the experimental data. For the excess speed of sound, the predictions of the theory are seen to be in quantitative agreement with the experimental data; the trend of the data to lower values as the difference in the chain length of the components of the mixture decreases is correctly reproduced. Quantitative agreement for the excess molar volume is seen with the data at higher temperatures, and the trend of the data to lower values with decreasing temperatures is reproduced by the theory. Given the small magnitude of the data (especially at lower temperatures), and the fact that one of the components of the mixture was not included in the parameter estimation procedure (n-hexadecane), the overall agreement of the theory with the experimental data is deemed to be satisfactory. Despite an accurate overall performance, it has to be noted that the theory cannot reproduce the fine features of the excess enthalpies of binary systems of asymmetric n-alkanes, as these arise from conformational effects that the theory does not account for. Such effects can be incorporated in the general theoretical framework by including a term that accounts for intramolecular interactions as has been shown by dos Ramos and Blas [177].

The group parameters obtained for the CH_3CO group and the unlike interaction values shown in tables 4.2 and 4.3 allow for the prediction of properties of binary n-alkane + 2-ketone mixtures. A rather interesting mixture to examine in this regard is the binary mixture of n-hexane+2-propanone (acetone), due to the liquid-liquid immiscibility exhibited by the system at lower temperatures (see figure 4.16). The SAFT-γ Mie approach with a single set of parameters is seen to reproduce, in a satisfactory manner the experimental data for both the vapour-liquid [388] and the liquid-liquid equilibria [389]. Comparing the results with the predictions of the SAFT-γ SW approach [29], it can be seen that the proposed methodology leads to a reduction in the overprediction of the upper critical solution temperature of the liquid-liquid phase envelope, whilst retaining a similar accuracy in the description of the vapour-liquid phase equilibria. It is important to reiterate at this point that within a GC approach for ketones, acetone is expected to be an outlier, and that a specific model for acetone would provide a much improved description of the mixture fluid phase behaviour.

The same set of group parameters can be used for the prediction of the phase behaviour of a wide range of n-alkanes+2-ketones mixtures. In figure 4.17.(a) a comparison between

the predictions of the theory and the experimental data [390–392] for the phase behaviour of the n-heptane+acetone binary mixture is shown. It can be seen that the fluid phase behaviour of the mixture can be accurately reproduced with SAFT-γ Mie, including the homogeneous azeotrope that appears at high concentrations of acetone. In figure 4.17.(b) the predictions of the theory are compared to the experimental data [392, 393, 395] for three binary mixtures including 2-butanone with three alkanes of varying size (n-hexane, n-heptane and n-octane). The trend of the experimental data and the shift of the location of the homogeneous azeotrope to higher compositions of 2-butanone as the chain length of the alkane in the system increases is very well predicted by the theory.

One can also examine the adequacy of the theory in describing excess properties of mixing of binary systems of n-alkanes+2-ketones. The excess heat of mixing predicted with SAFT-γ Mie for three binary mixtures of 2-butanone with alkanes of varying size (n-pentane, n-hexane and n-heptane) is compared to the experimental data [367, 396] in figure 4.18.(a). It can be seen that the heat of mixing of the three selected mixtures is described well and the correct variation in the magnitude of the excess heat with varying chain length of the n-alkane is also predicted. This is very pleasing, given the magnitude of

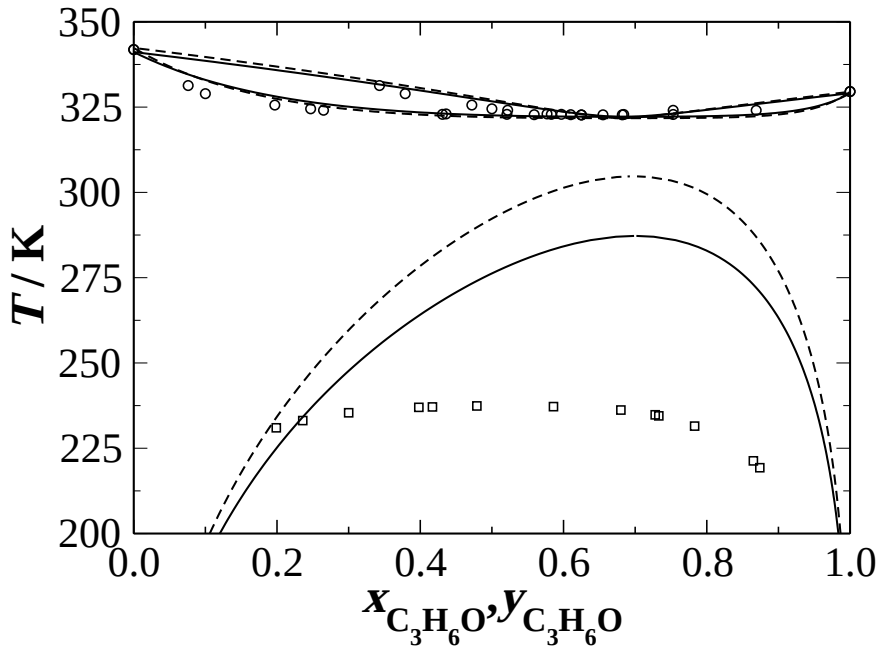

Figure 4.16: Predictions of the fluid phase behaviour of n-hexane+acetone as temperature-composition ($T - x$) isobar at 1 bar. The circles represent the experimental data for the vapour-liquid equilibria of the system [388], the squares the liquid-liquid equilibria [389], the continuous curves the predictions of the SAFT-γ Mie approach, and the dashed curves the corresponding predictions with the SAFT-γ SW approach [29].

the excess enthalpy for this mixtures ($\Delta H^{\mathrm{E}} \leq 1.5$ kJ/mol). The predictions for the excess volumes of two binary mixtures, *n*-heptane+2-butanone and *n*-heptane+2-pentanone, are compared to experimental data [398] in figure 4.18.(b). In this case, higher deviations are observed between the experimental values and the SAFT-γ Mie predictions. However, the performance of the proposed methodology is deemed as satisfactory given the correct representation of the trend of the excess volumes to smaller values as the components of the system become more similar in size is reproduced.

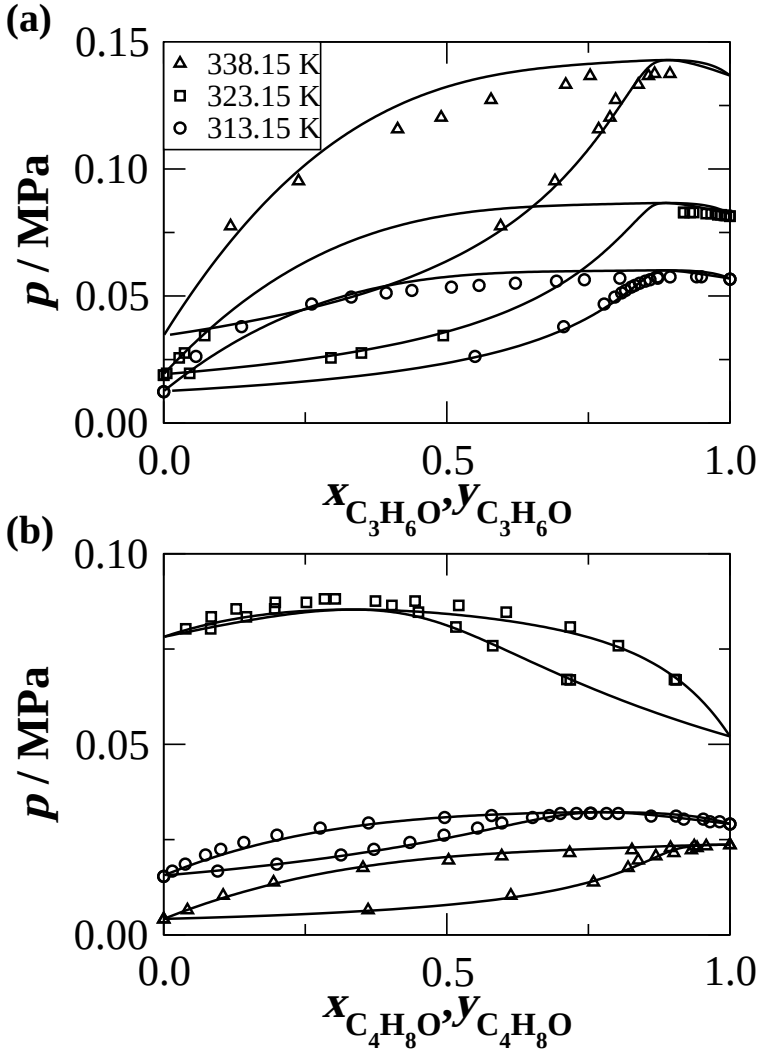

Figure 4.17: Predictions of the fluid phase behaviour of selected *n*-alkane+2-ketone binary mixtures. Pressure-composition ($p - x$) representation of the vapour-liquid equilibria of: (a) *n*-heptane+acetoneat 308.15 K [390] (circles), at 313.15 K [391] (squares), at 323.15 K [392] (triangles); and (b) *n*-octane+2-butanone at 313.15 K [393] (triangles), *n*-heptane+2-butanone at 318.15 K [394] (circles) and *n*-hexane+2-butanone at 333.15 K [395] (squares). The continuous curves are the predictions of the SAFT-γ Mie approach.

4.5 Concluding Remarks

In this chapter the development of the SAFT-γ Mie approach is presented in detail. The method is a reformulation of the SAFT-VR Mie EoS [31] within a group contribution formalism, where the interactions between monomeric segments are represented by means of the Mie pair potential of variable attractive and repulsive range. The molecular model employed for the description of pure substances and mixtures is a *fused* heteronuclear model, an idea following from previous studies [19]. Together with the implementation of

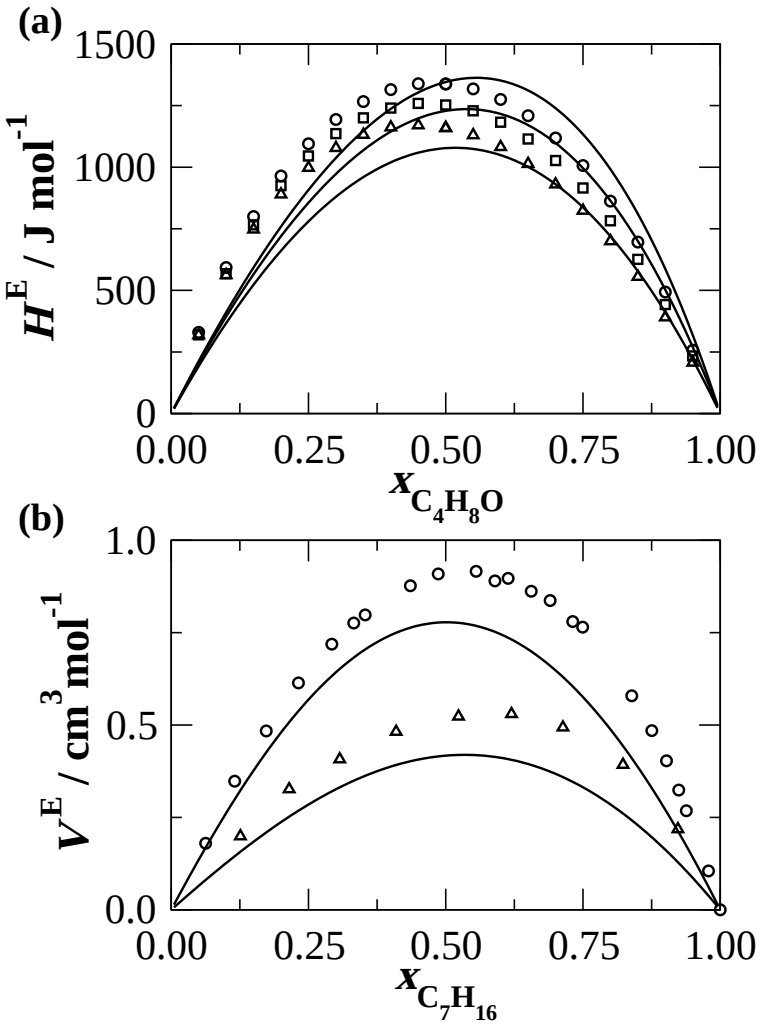

Figure 4.18: Predictions for selected excess properties of binary systems of *n*-alkanes+2-ketones: (a) excess heat of mixing for *n*-pentane+2-butanone (triangles), *n*-hexane+2-butanone (squares) and *n*-heptane+2-butanone (circles) at 298.15 K [367, 396]; and (b) excess volumes for *n*-heptane+2-butanone (circles) [397] and *n*-heptane+2-pentanone (triangles) [398] at 298.15 K. The continuous curves represent the predictions with the SAFT-γ Mie GC approach.

the new intermolecular potential, a key novelty of the theory lies in the treatment of the monomer contribution to the free energy, where a third-order perturbation expansion is employed. The performance of the theory in the description of real systems is examined in the study of two chemical families (the *n*-alkanes and the 2-ketones), where parameters for the corresponding functional groups (CH_3, CH_2 and CH_3CO) are obtained from fluid phase behaviour data and single-phase densities of the pure components. In the development of the CH_3CO group parameters (and the unlike parameters with the *n*-alkane groups) selected data for the excess enthalpies of binary mixtures are used, in order to better characterise the nature of the unlike interactions.

The SAFT-γ Mie approach is shown to allow for an excellent description of the pure component properties of the correlated compounds, with average relative errors of 1.85% for the vapour pressure, 0.91% for the saturated liquid density, and 0.96% for the single-phase density. It is also shown that the treatment of the monomer term proposed, leads to a significant improvement in the description of the near-critical region of fluids, compared to the SAFT-γ SW formulation [19]. Apart from fluid phase behaviour, second-order thermodynamic derivative properties are also considered. The Mie intermolecular potential model employed in the proposed methodology allows for an accurate representation of such properties (speed of sound, heat capacities, Joule-Thomson coefficient, etc.) as shown in detail for the *n*-alkanes, whereas slightly higher deviations are to be expected for the 2-ketones.

The predictive capability of the method is examined in the prediction of fluid phase equilibrium and, when available, derivative properties of high molecular weight compounds not included in the regression of the group parameters, where the performance of the SAFT-γ Mie approach is found to be very satisfactory. A predictive study in the limit of high molecular weight polymers confirms the robustness of the presented methodology and the corresponding group parameters obtained in this work. Finally, the performance of SAFT-γ Mie is examined in the prediction of the fluid phase behaviour and excess properties of binary mixtures of *n*-alkanes and *n*-alkane+2-ketone. The theory is seen to accurately describe the fluid phase behaviour of a variety of systems, including different types of equilibria (vapour-liquid and liquid-liquid), over a wide range of conditions, including the high-pressure critical points of the mixtures. The excess properties of binary mixtures are also very well described and, for the cases considered, the calculations of the SAFT-γ Mie are in better agreement compared to the calculations using the modified UNIFAC (Dortmund) GC approach. The description of such properties is a challenging task, which is known to be very sensitive to the specific molecular details, and the performance of

the SAFT-γ Mie demonstrates the future potential of the theory for the study of other chemical families.

Based on the very promising performance of the SAFT-γ Mie method presented here, the theory is applied to the modelling of solubilities of complex organic compounds in solvents. The thermodynamic framework for the modelling of solubilities together with an example of the application of the SAFT-γ Mie in the study of complex systems are presented in the following chapter.

Chapter 5

Modelling the solubility of complex organic molecules in organic solvents

In the work presented thus far, the performance of several thermodynamic methodologies in the description of vapour-liquid and liquid-liquid equilibria of fluid mixtures has been discussed. With a detailed analysis of SAFT-type GC methods it has been shown that an accurate representation of fluid phase equilibria for a variety of mixtures can be obtained. However, the performance of the method in studies of solid-liquid equilibria (solubilities) has not yet been examined.

The accurate modelling of the solubility of complex organic molecules in solvents and solvent blends is of great importance to the pharmaceutical and the agrochemical industry. The end product is in these cases typically obtained in pure solid form by means of a crystallisation process, where the quality of the solvent or solvent blend used can to a great extent determine the success or failure of the process [399]. The choice of the appropriate solvent, commonly referred to as solvent screening, can help manipulating the solubility of the solid product in order to achieve an optimal production process. For the specific case of active pharmaceutical ingredients (APIs), the importance of solubility modelling extends beyond the scope of production. In this case, the solubililty of the solid drug in aqueous media can dictate its bioavailability, while the permeation of the API through biomembranes can be related (in the context of sobulitily) to water/octanol partition coefficients. The study of solubilities can be very consuming and cumbersome to carry out experimentally, especially in the case of solvent blends, given the wide range of solvents and compositions of the blend that have to be examined. The number of experiments

that can be conducted is often limited by the small quantities of APIs typically produced in early stages of drug discovery and design. An interesting alternative is to carry out solvent screening and solubility study with thermodynamic tools that predict accurately the solubilities of interest. The importance of the application of such methodologies has been highlighted as a key element in the progress of the development of manufacturing processes within pharmaceutical industries [400].

The application of the SAFT-γ Mie GC method (presented in chapter 4) to the study of the solubility of complex organic molecules in organic solvents is developed in this chapter. The thermodynamic framework for calculating solubilities by means of fluid theories is presented next, followed by a review of the thermodynamic methodologies that have been applied to the study of solubilities of organic molecules, including APIs, in solvents and solvent blends. The SAFT-γ Mie parameter table is extended to include functional groups commonly encountered in APIs and solvents, and the performance of the novel approach in the challenging application of the prediction of solubilities of APIs in organic solvents is investigated.

5.1 Modelling solid-liquid equilibria

The solubility of a compound (the solute, *slt*) in a solvent or solvent blend is defined as the maximum quantity of the compound that can be dissolved in a given amount of a solvent (solid, gas, or liquid) to form a homogeneous solution. This can be determined by solving for the phase boundary of the solid-liquid equilibrium (SLE) of the system. Here, the thermodynamic framework for solid-liquid equilibria presented and discussed by Prausnitz *et al.* [401] is followed, starting from the conditions of fluid phase equilibrium, i.e., the equality of the fugacity f (chemical potential) of the solute (denoted by subscript *slt*) in the liquid (L), and in the solid (S) phase at the specified temperature T and pressure p:

$$f_{slt}^{\mathrm{S},*}(T,p,x_{slt}^{\mathrm{S}}=1)=f_{slt}^{\mathrm{L}}(T,p,x_{slt}^{\mathrm{L}})\ , \tag{5.1}$$

where the asterisk indicates a pure component; it is hereby assumed that the solid phase consists solely of pure solute (a good approximation in practice), while the liquid phase is a mixture of solute and solvent(s). Hence, the equality of fugacities applies only to the solute and not to all of the components of the mixture. The fugacity of the solute in the liquid phase can be calculated as:

$$f_{slt}^{\mathrm{L}}(T,p,x_{slt}^{\mathrm{L}})=x_{slt}^{\mathrm{L}}\gamma_{slt}^{\mathrm{L}}(T,p,\underline{x}^{\mathrm{L}})f_{slt}^{\mathrm{L},*}(T,p,x_{slt}^{\mathrm{L}}=1)\ , \tag{5.2}$$

where $\underline{x}^{\mathrm{L}}$ is the composition of the liquid mixture, x_{slt}^{L} is the composition of the solute in the liquid phase, and $\gamma_{slt}^{\mathrm{L}}(T, p, x_{slt}^{\mathrm{L}})$ its activity coefficient. For the study of solid-liquid equilibria in binary mixtures, the activity coefficient can be written as a function of the composition of the solute in the liquid phase, x_{slt}^{L}, i.e., $\gamma_{slt}^{\mathrm{L}}(T, p, x_{slt}^{\mathrm{L}})$. The fugacity $f_{slt}^{\mathrm{L},*}(T, p, x_{slt}^{\mathrm{L}} = 1)$ is that of the pure solute, as a subcooled liquid, at the specified temperature and pressure; for temperatures below the melting point of the solute, T_{m}, the fugacity of the subcooled liquid can be calculated by extrapolation of the liquid thermodynamic properties into the solid region [58]. Omitting the composition constraint $x_{slt}^{\mathrm{L}} = 1$ in the pure component fugacities for brevity, the equilibrium equation now becomes:

$$\frac{f_{slt}^{\mathrm{L},*}(T,p)}{f_{slt}^{\mathrm{S},*}(T,p)} = \frac{1}{x_{slt}^{\mathrm{L}}\gamma_{slt}^{\mathrm{L}}(T,p,\underline{x}^{\mathrm{L}})} \ . \tag{5.3}$$

The ratio on the left hand side of eq. (5.3) can be calculated from the thermodynamic cycle depicted in figure 5.1, where it is assumed that the pressure p is fixed. The thermodynamic cycle can be extended to account for differences in pressure, as shown in reference [402]. However, given that solubility data are usually reported at ambient pressure, the solid-liquid equilibria working equation is derived here neglecting pressure effects. The ratio of the solid and subcooled liquid fugacities can be related to the molar Gibbs free energy change going from the the solid (a) to the liquid (d) state as:

$$\Delta g_{a\to d} = RT\ln\frac{f_{slt}^{\mathrm{L},*}(T,p)}{f_{slt}^{\mathrm{S},*}(T,p)} \ . \tag{5.4}$$

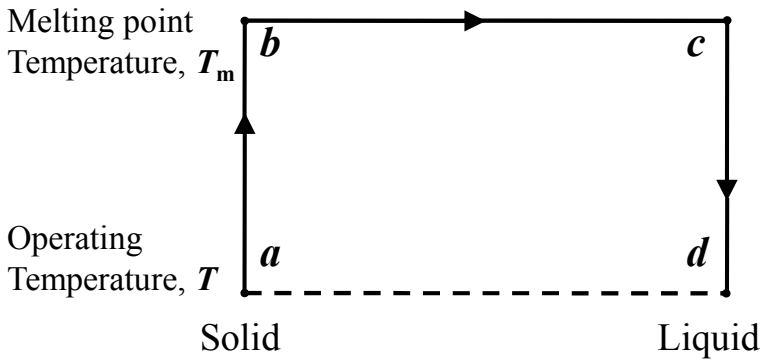

Figure 5.1: Thermodynamic cycle for the calculation of the fugacity of a pure subcooled liquid as in Prausnitz *et al.* [401].

Based on the standard thermodynamic relation $\Delta g_{a\to d} = \Delta h_{a\to d} - T\Delta s_{a\to d}$, the molar Gibbs free energy change is related to the molar enthalpy and entropy changes. The thermodynamic cycle presented in figure 5.1 is employed for the evaluation of these two quantities, according to which the difference along the path $a \to d$ can be calculated based

on the alternative path $a \rightarrow b \rightarrow c \rightarrow d$. The cycle describes the following 3-step constant pressure process [58]:

- the solid is heated at fixed p from the operating temperature T to its melting temperature T_{m} ($a \rightarrow b$);
- at T_{m} the solid is melted to form a liquid ($c \rightarrow c$);
- the liquid is cooled without solidification from T_{m} to the temperature of the mixture T ($c \rightarrow d$);

The change in enthalpy can be written as

$$\begin{aligned} \Delta h_{a\rightarrow d} &= \Delta h_{a\rightarrow b} + \Delta h_{b\rightarrow c} + \Delta h_{c\rightarrow d} \\ &= \int_{T}^{T_{\mathrm{m}}} c_p^{\mathrm{S}} dT + \Delta h_{\mathrm{fus.}}(T_{\mathrm{m}}) + \int_{T_{\mathrm{m}}}^{T} c_p^{\mathrm{L}} dT \\ &= \Delta h_{\mathrm{fus}}(T_{\mathrm{m}}) + \int_{T_{\mathrm{m}}}^{T} \Delta c_p(T) dT \ , \end{aligned} \tag{5.5}$$

where $\Delta h_{\mathrm{fus}}(T_{\mathrm{m}})$ is the enthalpy of fusion of the solute at the melting point, $\Delta c_p = c_p^{\mathrm{L}} - c_p^{\mathrm{S}}$ the difference between the isobaric heat capacity of the solute in liquid and solid state, and T_{m} the melting temperature of the solute. The same analysis for the entropy gives

$$\Delta s_{a\rightarrow d} = \Delta s_{\mathrm{fus}}(T_{\mathrm{m}}) + \int_{T_{\mathrm{m}}}^{T} \frac{\Delta c_p(T)}{T} dT \ . \tag{5.6}$$

Rearranging eqs. (5.4)-(5.6), making use of the relation between the entropy and enthalpy of fusion at the melting point ($\Delta s_{\mathrm{fus}} = \frac{\Delta h_{\mathrm{fus}}}{T_{\mathrm{m}}}$), and under the assumption that the heat capacity differences do not depend on temperature ($\Delta c_p = \text{const.}$) we obtain the following expression:

$$\ln \frac{f_{slt}^{\mathrm{L},*}}{f_{slt}^{\mathrm{S},*}} = \frac{\Delta h_{\mathrm{fus}}}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{m}}}\right) - \frac{\Delta c_p}{R}\left(\frac{T_{\mathrm{m}}}{T} - 1\right) + \frac{\Delta c_p}{R} \ln \frac{T_{\mathrm{m}}}{T} \ . \tag{5.7}$$

Typically eq. (5.7) is simplified in order to obtain the working equation for the calculation of solid-liquid equilibria, by neglecting the terms accounting for the difference between the heat capacity of the solid and the liquid phase, i.e., the last two terms on the right-hand side of eq. (5.7). In view of the typical absence of available experimental data for the heat capacity differences, these two terms are assumed to be of approximately equal contribution, and given the opposite sign, to cancel out, thereby introducing only slight error. The working equation is then written in its simplified form (where the fugacity ratio has been replaced using eq. (5.3)) as

$$\ln x_{slt}^{\mathrm{L}} = -\frac{\Delta h_{\mathrm{fus}}}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{m}}}\right) - \ln \gamma_{slt}^{\mathrm{L}}(T, p, \underline{x}^{\mathrm{L}}) \ . \tag{5.8}$$

Despite the fact that the contribution of the Δc_p terms is commonly neglected, it has been suggested that it can significantly affect the solubility calculations, especially for compounds with a significant difference between the heat capacity of the solid and the liquid [403, 404]. As an example, Gracin *et al.* [405] found that including Δc_p in the calculation of the solubility of paracetamol (for which $\Delta h_{\text{fus}} = 27.1$ kJ/mol [405] and $\Delta c_p = 98.9$ J/mol [406]) within the UNIFAC formalism can lead to a significant increase, in some cases exceeding 100% compared to the calculation neglecting the heat-capacity contribution.

From the working equation (cf. eq. (5.8)) it can be seen that the calculation of the solid-liquid phase boundary (solubility) of a pure solute and a liquid phase with composition $\underline{x}^{\text{L}}$ requires a knowledge of three quantities: the enthalpy of fusion Δh_{fus} and the melting temperature T_{m} of the pure solute, as well as the activity coefficient of the solute in the liquid phase at the specified conditions (T, p, and $\underline{x}^{\text{L}}$). The first two quantities are typically obtained from experimental data; when these are not available, their values can be approximated using, e.g., pure component GC approaches, such as that of Constantinou and Gani [38], the method of Chickos *et al.* [407], or the more recent approach of Oskoei and Keshavarz [408], and the GC+ approach of Gani and co-workers [409]. It should be noted that Δh_{fus} and T_{m} can be also used to distinguish between the solubility of different polymorphs of the same compound, as discussed in [403]. Polymorphs are different forms of a substance with similar free energies that might exist simultaneously at certain pressure and temperature conditions and, within the context of the approach presented here, can be distinguished based on the different values of Δh_{fus} and T_{m} that characterise each of these forms.

The activity coefficient of the solute in the liquid mixture quantifies the nonidealities arising from the interactions between the solute and the solvent. When the chemical nature of these compounds is very similar, the activity coefficient is tends to unit $\gamma^{\text{L}}_{slt}(T, p, \underline{x}^{\text{L}}) \rightarrow 1$, characterising an ideal mixture. The solubility calculated from eq. (5.8) is then the ideal solubility. For non-ideal systems, the activity coefficient is calculated by means of a thermodynamic approach for the liquid mixture. Several types of methods have been applied for the calculation of the solubility of APIs in solvents and solvent mixtures, a review of which is given in the following section.

5.2 Thermodynamic methodologies for the calculation of the solubility of APIs in solvents

A wide range of methodologies has been applied to describe the solubility of complex organic molecules and active pharmaceutical ingredients (APIs). Many empirical approaches have been developed where the solubility of a compound is predicted based on its structural properties; for a review see reference [410]. Most quantitative structure-property relationships (QSPR) have been developed for the study of the aqueous solubility and the water/octanol partition coefficient of drug-like molecules, mainly due to their importance both in the production of pharmaceuticals as well as their behaviour in the human body. As a consequence, QSPR approaches do not yield accurate solubility for other solvents. This can be overcome by the application of thermodynamic models, which are based on molecular interactions between chemically distinct species or functional groups and are, in principle, not limited to the study of a given class of solvents. A number of methodologies have been applied, varying from formalisms based on lattice-fluid theories, such as the non-random hydrogen bonding (NRHB) method [403, 404], to activity coefficient models, such as Wilson-type methods [411], approaches based on the Kirkwood-Buff fluctuation solution theory [412], and association theories, such as the CPA EoS [413, 414].

Of particular relevance to the current work is the application of SAFT-type methodologies to the prediction of solubility of pharmaceutical compounds in solvents. SAFT-type methods have been applied to the study of solid-liquid equilibria, e.g., the recent work with SAFT-VR on the study of the solubility of naphthalene and acetic acid in organic solvents [415]. As far as pharmaceutical compounds are concerned the PC-SAFT EoS [165, 166] has been recently applied to the prediction of solubility of, for example, paracetamol, ibuprofen and lovastatin [416, 417]. In these studies with PC-SAFT the parameters describing the API (typically 6 parameters per compound) have to be determined from solubility data, where commonly aqueous solubility data are used, as experimental vapour pressures and saturated liquid densities are rarely available; these compounds tend to be in the solid state at normal conditions. In addition to the pure component API parameters, solubility data are also used to determine the unlike interactions between the solutes and solvents. In order to make the approach more predictive, Ruether and Sadowski [416] proposed a scheme of classification of solvents according to their chemical nature, transferring the values of the unlike interaction parameters between the API and all solvents of a given class. Three representative solvents are used, namely ethanol for associating (hydrogen bonding) solvents, acetone for weakly polar solvents, and toluene for aromatic solvents, in order to determine the unlike dispersive energy between the API and

a class of solvents. It has to be noted that two solubility points at different temperatures are needed for each regression, as the unlike dispersive energy is assumed to have a linear temperature dependence. It is shown that the PC-SAFT EoS can accurately correlate the solubility data used in the regression, and that good predictions for the solubility in mixtures of mixed solvents used in the regression are obtained. However, when transferring the value of the unlike interaction parameters between solvents of the same class, i.e., the value obtained from regression to ethanol for associating solvents is used for the prediction of solubility in another associating solvent, e.g., 2-butanol, the agreement with experimental solubility data deteriorates significantly. In order to address the dependence of the pure solute parameters on the solubility data, Cassens *et al.* [418] subsequently presented a methodology that allowed for the estimation of the majority of the pure solute parameters based on quantum mechanical calculations. This approach reduced the dependency on experimental data, however it has to be noted that solubility data are still required for determining the solute-solvent interaction parameters. In more recent work, Spyriouni *et al.* [417] proposed a different scheme to determine the pure component parameters for APIs. Within their approach, solubility data in three different classes of solvents were used (a hydrophilic, a polar, and a hydrophobic liquid) in order to obtain average pure component parameters for the APIs, while the interaction energy between the API and solvents other than water was calculated by means of a standard geometric combining rule. It was shown that a satisfactory description of the solubility of APIs as a function of the composition of solvent mixtures can be achieved. However, for a number of pure solvents not included in the estimation of the API parameters, large deviations between the predictions of the theory and the experimental data were found.

One of the most widely applied thermodynamic models in the prediction of solubility of pharmaceutical compounds is the non-random two-liquid segment activity coefficient (NRTL-SAC) model. The NRTL-SAC model was developed by Chen and Song [231] based on a modification of the NRTL activity coefficient approach [116]. The main concept of NRTL-SAC rests on the assumption that the activity coefficient of a molecule can be calculated based on the molecular descriptors (segments) that characterise a molecule. Four distinct classes of molecular descriptors were identified, related to the chemical nature of the compound (hydrophobicity, polarity, solvation, and hydrophilicity); the interactions between the conceptual segments were assumed to be constant and were determined by regression to extensive experimental data for solid-liquid equilibria for a wide range of solvents [231]. Given the knowledge of the interaction parameters the characterisation of a compound within NRTL-SAC is undertaken based solely on the measure of each conceptual segment on the compound, i.e., to which extent it is hydrophobic, polar, solvating or

hydrophilic. These measures can be obtained from experimental data for the solubility of the compound(s) in question in pure solvents, where it has to be noted that, for a reliable characterisation, data for at least four solvents (each solvent representing one type of the four conceptual segments) is required. The reported average deviations in the prediction of the solubility of a selection of complex molecules in solvent and solvent mixtures with the NRTL-SAC method are of the order of 40-60% [419]. However, its heavy reliance on experimental solubility data is a significant drawback.

A common denominator in the methodologies mentioned thus far is that solubility data are typically used for the parameterisation of the models. This limits their application to systems for which experimental data is available. When this is not the case, predictive thermodynamic methodologies can be applied. The two predictive methods most widely applied to the study of solubility of APIs in solvents and solvent blends are the universal quasi-chemical functional group activity coefficient (UNIFAC) and the conductor-like screening model for real solvents - segment activity coefficient (COSMO-SAC) methods.

The performance of the UNIFAC method in the prediction of the solubility of a number of complex organic molecules (phenylacetic acid, p-hydroxyphenylacetic acid, ibuprofen, paracetamol) in a selection of eight pure solvents, including polar and hydrogen bonding compounds, was the focus of a study by Gracin *et al.* [405]. The solubility calculations presented were based on the original UNIFAC description of the activity coefficient of the liquid with the parameters of Hansen *et al.* [83]. The authors concluded that the original UNIFAC approach can lead to a varying degree of accuracy in the description of the solubility, ranging from 20% for ibuprofen to values exceeding 500% for paracetamol. In order to improve the performance of the method, a pharma-modified UNIFAC has recently been presented [420] where the fragmentation of groups was geared specifically to the study of APIs (e.g., the definition of the sulfonic acid amine group), and several solubility data were included in the regression of the parameters. The resulting parameterisation performs significantly better compared to the results obtained with the original UNIFAC database and provides more accurate predictions than the COSMO-SAC method described below. However, the method was developed by neglecting several group-group interactions and can only be applied to concentrations of APIs which are less than 10% mol/mol. Although this covers the range of concentrations encountered in most crystallisation purposes, the limited composition range can be considered a significant limitation of the pharma-modified UNIFAC. Of perhaps more critical importance, it should be noted that the underlying assumptions of the pharma-modified UNIFAC result in a reformulation of the theory, so that the existing UNIFAC parameter table cannot be directly used.

Another predictive thermodynamic methodology that has been applied to the study of the solubility of APIs in solvents is the COSMO-SAC model [230], developed based on the COSMO-RS methodology of Eckert and Klamt [421]. The residual part of the activity coefficient within this model is calculated based on the interactions between molecules through surface contacts. In this approach the nature of these interactions is expressed by using a screening charged density, typically referred to as the σ profile, which is obtained from *ab initio* calculations for a molecule in a perfect conductor [421], so that no experimental data are required for the parameterisation of the model. This, however, appears to affect the accuracy of the methodology, as it is generally considered less accurate than NRTL-SAC [232, 422]. In a recent study it has been shown that by combining COSMO-SAC with solubility data in pure solvents for the determination of the segment σ profile one can significantly improve the prediction of API solubility in mixed solvents, from average deviations of the order of 400% to 90% [423].

In summary, it may be appreciated at this point that the accurate description of solubilities of complex organic molecules in solvents remains a significant challenge. Given the promising results obtained with the SAFT-γ Mie approach presented in chapter 4 for the prediction of the fluid phase behaviour and derivative thermodynamic properties of a broad range of complex mixtures, the performance of the SAFT-γ Mie EoS in the prediction of the solubility of APIs in organic solvents is now considered as a case study for two APIs in the following sections.

5.3 Estimation of group parameters for the modelling of APIs

The main objective here is the preliminary assessment of the performance of the SAFT-γ Mie methodology in the description of the solubility of complex organic molecules and active pharmaceutical ingredients in organic solvents. The prototypical systems chosen for this assessment are the APIs phenylacetic acid and ibuprofen as solutes, and acetone and methyl-isobutyl ketone (MIBK) as solvents. The molecular structure of the components of the target molecules is shown in figure 5.2. Phenylacetic acid is used in the production of penicillin G and is also employed to treat type II hyperammonemia to help reduce the amount of ammonia in the bloodstream. It is modelled with 3 different functional groups (5 aCH groups, 1 $aCCH_2$ group, and 1 COOH group). Ibuprofen is a widely used nonsteroidal anti-inflammatory drug, and has a more complex molecular structure than phenylacetic acid. Within the context of the current work ibuprofen is modelled as

comprising 5 distinct functional groups: 4 aCH groups, 1 COOH group, 1 $CHCH_3$ group, 2 CH_3 groups, and 2 $aCCH_2$ groups; strictly speaking, ibuprofen contains an instance of the aCCH group, which in this instance is approximated as an additional $aCCH_2$ group. This is expected to introduce only slight error in the thermodynamics of the mixture, and reduces significantly the number of unlike interaction parameters that need to be determined to model the mixture of interest. The procedure followed to characterise the groups identified is presented in the next sections, however it is worth noting at this stage, that no experimental data regarding the APIs studied is included in the development of the model. A summary of the groups that the substances comprise and the instances of each group is given in table 5.1.

It is important to reiterate some of the assumptions of the study undertaken here in the representation of these complex systems. All substances are studied by applying the GC concept at the first-order level, i.e., accounting solely for the number of occurrences of each chemical functional group on a substance. In that way, proximity effects are not explicitly treated. In previous studies of systems of similar complexity using the UNIFAC approach based on quantum mechanical calculations, Gracin and co-workers [405] showed that for molecules such as ibuprofen, with multiple functional groups in close vicinity, the assumption that the behaviour of functional groups is independent of the molecular structure may be too crude an approximation; in such cases proximity effects should be taken into consideration. Furthermore, the only types of interactions between segments that are accounted for are repulsive and dispersion interactions and association, mediated by short-ranged interaction sites. Additional effects arising from the chemical nature of

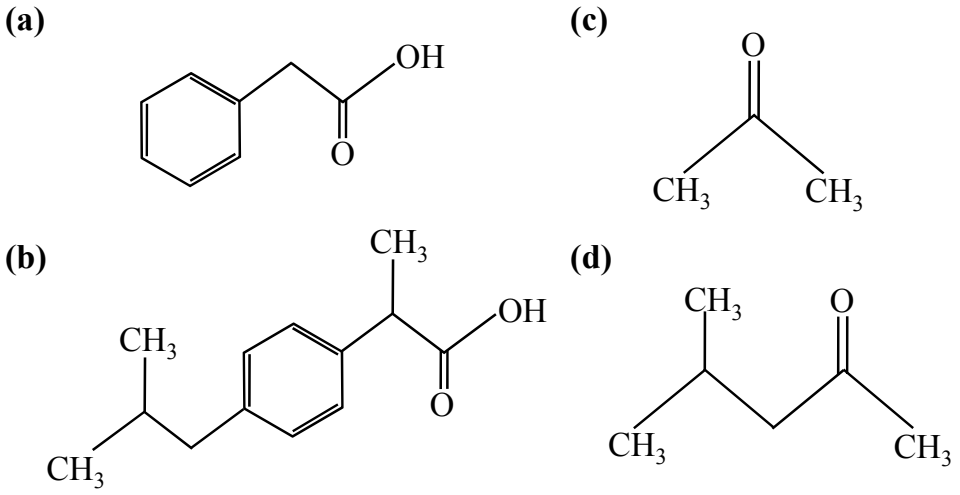

Figure 5.2: Molecular structures of the compounds in the solute+solvent mixtures studied: (a) phenylacetic acid, (b) ibuprofen, (c) acetone, and (d) methyl-isobutyl ketone (MIBK).

Table 5.1: Analysis of the two APIs (phenylacetic acid and ibuprofen) and the two solvents (acetone and methyl-isobutyl ketone (MIBK)) of the mixtures studied (cf. figure 5.2) in terms of the functional groups that these molecules comprise and the instances of each group. It should be noted that the aCCH group present in ibuprofen is approximated as an additional instance of the $aCCH_2$ group.

Group	Phenylacetic acid	Ibuprofen	Acetone	MIBK
CH_3	0	2	1	1
CH_2	0	0	0	1
CH_3CO	0	0	1	1
$CHCH_3$	0	1	0	1
aCH	5	4	0	0
$aCCH_2$	1	2	0	0
COOH	1	1	0	0

the compounds and functional groups studied, e.g., the polarity of the ketones or the aromaticity of the hydrocarbons, are treated only effectively by means of the aforementioned dispersion interactions, and are not explicitly accounted for. It is possible to account explicitly for such effects by including additional contributions for polar and quadrupolar interactions, see for example [172, 174, 175, 190].

From the molecular structure of the compounds of interest and the decomposition of these into functional groups, it can be seen that alongside the functional groups of the n-alkanes (CH_3 and CH_2) and the 2-ketones (CH_3CO), the parameters for the branched methyl group $CHCH_3$, the aromatic carbon (aCH) and methylene $aCCH_2$ groups, as well as for

Table 5.2: Group parameter matrix featuring the functional groups required for the modelling of the compounds of figure 5.2. The ticks denote the group parameters available from the work in chapter 4, while the dashes denote the parameters that have to be determined in order to describe the mixture of interest. The line demarcating the two parts of the matrix highlights the fact that the matrix is symmetric.

	CH_3	CH_2	CH_3CO	$CHCH_3$	aCH	$aCCH_2$	COOH
CH_3	✓	✓	✓	–	–	–	–
CH_2	✓	✓	✓	–	–	–	–
CH_3CO	✓	✓	✓	–	–	–	–
$CHCH_3$	–	–	–	–	–	–	–
aCH	–	–	–	–	–	–	–
$aCCH_2$	–	–	–	–	–	–	–
COOH	–	–	–	–	–	–	–

the carboxyl group (COOH) have to be determined, together with the unlike group interactions between the aforementioned functional groups. The group parameters and unlike group interactions required for the description of the solubility of ibuprofen in acetone are summarised in table 5.2.

5.3.1 Pure component parameters: branched alkanes

The branching of the alkyl chain for ibuprofen and MIBK is modelled by defining the $CHCH_3$ group (cf. figure 5.2 and table 5.1). The parameters for this group are obtained by regression to experimental vapour-liquid equilibrium data for selected methyl and dimethyl substituted alkanes. The group parameters estimated with the SAFT-γ Mie approach provide an excellent description of the fluid phase behaviour of the branched alkanes, with an average absolute deviation %AAD of 1.72% for the vapour pressure and 0.53% for the saturated liquid density for all of the compounds. A list of the compounds included in the parameter estimation procedure, together with the deviation per compound and per property and the details of the experimental data used in the regression are presented in table 5.3.

Table 5.3: Percentage average absolute deviations (%AAD) of vapour pressures $p_{\text{vap}}(T)$ and saturated liquid densities $\rho_{\text{sat}}(T)$ for the branched alkanes obtained within the SAFT-γ Mie framework from the experimental data [280] used in the regression, where n is the number of data points.

Compound	T range (K)	n	%AAD p_{vap}	T range (K)	n	%AAD ρ_{sat}
2-methylbutane	190-410	30	1.70	145-410	36	0.41
2-methylpentane	289-445	19	0.61	158-445	34	0.39
2-methylhexane	273-477	36	0.62	273-477	36	0.53
2-methylheptane	233-500	37	1.36	277-500	31	0.52
2-methyldecane	273-462	24	1.07	253-553	12	0.54
2-methylundecane	356-483	9	3.20	253-553	12	0.51
2-methyldodecane	373-502	9	2.98	273-588	11	0.53
2,4-dimethylpentane	289-463	32	1.76	273-463	35	0.62
2,5-dimethylhexane	246-495	37	0.63	273-495	33	0.57
2,7-dimethyloctane	293-432	28	3.10	253-523	11	0.69
Average	-	-	1.70	-	-	0.53

The description of the vapour-liquid phase behaviour of the branched alkanes is shown in graphical form in figures 5.3 and 5.4 for the coexistence densities and the vapour pres-

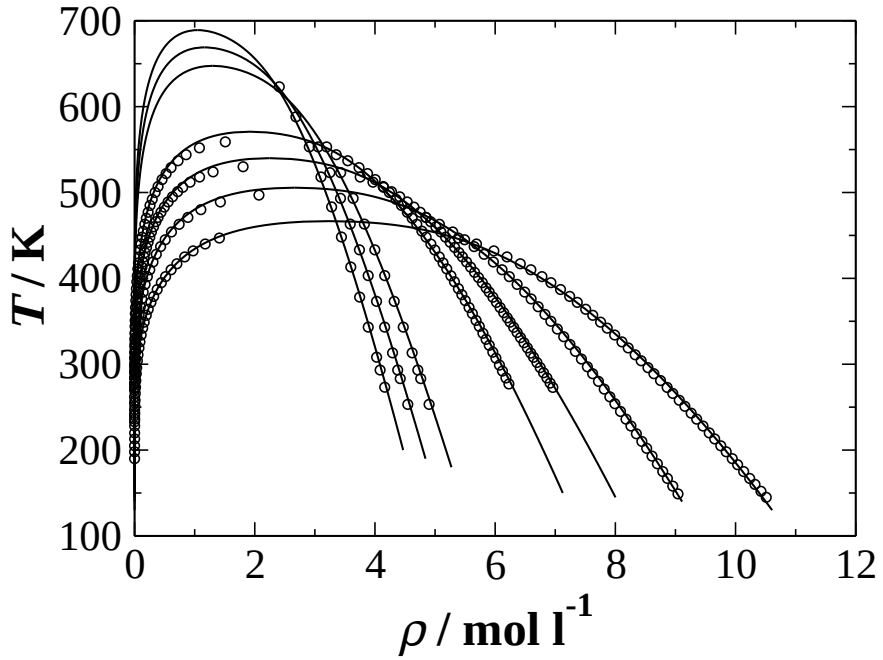

Figure 5.3: SAFT-γ Mie description of the coexistence densities for selected methyl branched alkanes (2-methylbutane to 2-methyldodecane from right to left) included in the estimation of the $CHCH_3$ group parameters. The symbols represent the experimental data [280] and the continuous curves the calculations of the theory.

sures in a logarithmic representation for the methyl branched alkanes and in figures 5.5.(a) and 5.5.(b) for the dimethyl branched alkanes.

The parameters obtained for the $CHCH_3$ functional group are presented in table 5.6. The intermolecular potential that describes the interactions between the segments of the $CHCH_3$ group was found to be determined by a value of the repulsive exponent of $\lambda^r_{CHCH_3} = 13.885$, whereas the attractive exponent was fixed to the London dispersion value of $\lambda^a_{CHCH_3} = 6$. The repulsive exponents of the CH_3 and CH_2 groups ($\lambda^r_{CH_3} = 15.050$ and $\lambda^r_{CH_2} = 19.871$), are seen to be larger than the optimal value of the repulsive exponent for the $CHCH_3$ group; this is in line with the finding that more coarse-grained models require lower values of the repulsive exponent ("softer" potentials) [424]. The segment size for the $CHCH_3$ is found to be larger than that of the CH_3 ($\sigma_{CHCH_3} = 4.575\,\text{Å}$ compared with $\sigma_{CH_3} = 4.077\,\text{Å}$), for a similar value of the shape factor, which is physically reasonable based on the number of carbon atoms that each group comprises. The optimal value of the interaction energy is $\epsilon_{CHCH_3}/k_B = 317.904\,\text{K}$, which as expected is larger than that for the CH_3 group, $\epsilon_{CH_3}/k_B = 256.766\,\text{K}$.

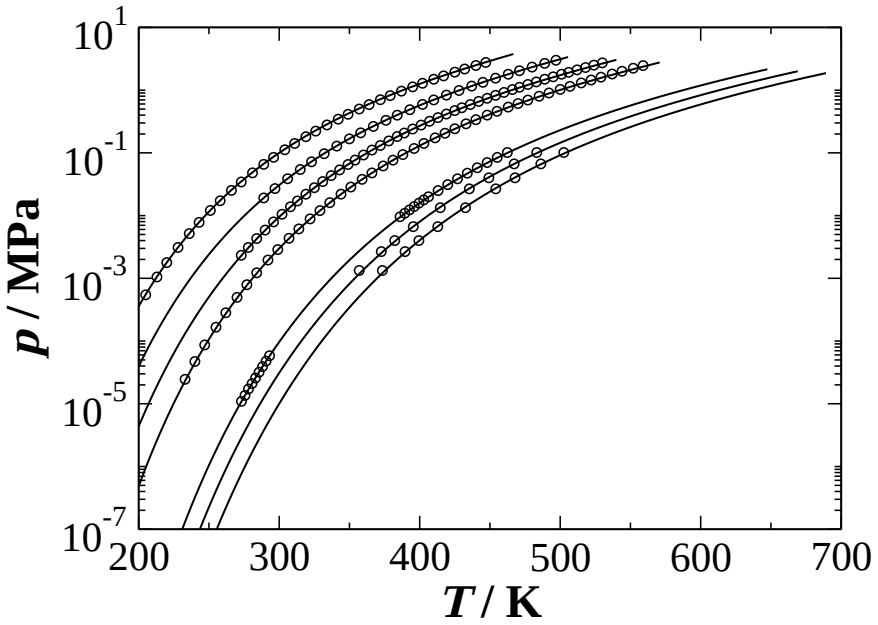

Figure 5.4: SAFT-γ Mie description of the vapour pressure for selected methyl branched alkanes (2-methylbutane to 2-methyldodecane from left to right) included in the estimation of the $CHCH_3$ group parameters. The symbols represent the experimental data [280] and the continuous curves the calculations of the theory.

5.3.2 Pure component parameters: alkylbenzenes

The aromatic part of the molecules considered here is modelled by means of two functional groups: an aCH group for the unsubstituted aromatic carbons of the ring, and an $aCCH_2$

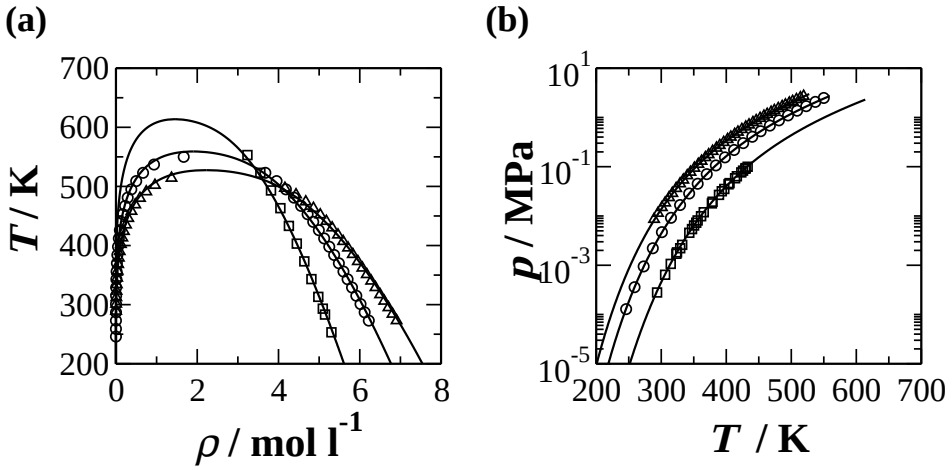

Figure 5.5: SAFT-γ Mie description of: (a) the coexistence densities; and (b) the vapour pressures for selected dimethyl branched alkanes included in the regression of the parameters of the $CHCH_3$ functional group. The symbols represent the experimental data [280] (triangles for 2,4-dimethylpentane, for 2,5-dimethylhexane, and squares for 2,7-dimethyloctane) and the continuous curves the calculations of the theory.

Table 5.4: Percentage average absolute deviations (%AAD) of vapour pressures $p_{\mathrm{vap}}(T)$ and saturated liquid densities $\rho_{\mathrm{sat}}(T)$ for the alkylbenzenes obtained within the SAFT-γ Mie framework from the experimental data [280], where n is the number of data points.

Compound	T range (K)	n	%AAD p_{vap}	T range (K)	n	%AAD ρ_{sat}
Benzene	280-508	58	0.51	280-508	58	0.51
Ethylbenzene	424-554	29	1.07	183-490	38	0.49
Propylbenzene	348-433	20	1.75	223-543	10	0.40
Butylbenzene	253-418	18	1.50	223-583	11	0.42
Pentylbenzene	233-477	12	1.41	233-593	35	3.96
Hexylbenzene	263-463	21	1.42	298-486	64	0.26
Heptylbenzene	309-513	20	5.29	281-367	11	0.22
Octylbenzene	303-462	17	5.48	243-648	11	1.00
Nonylbenzene	303-466	19	3.26	282-368	9	0.10
Decylbenzene	476-570	9	0.99	273-678	11	1.29
Average	-	-	2.27	-	-	0.87

group that incorporates the first methylene group in case of an alkyl substituted phenyl ring. The interaction parameters for both these groups are obtained by parameter esti-

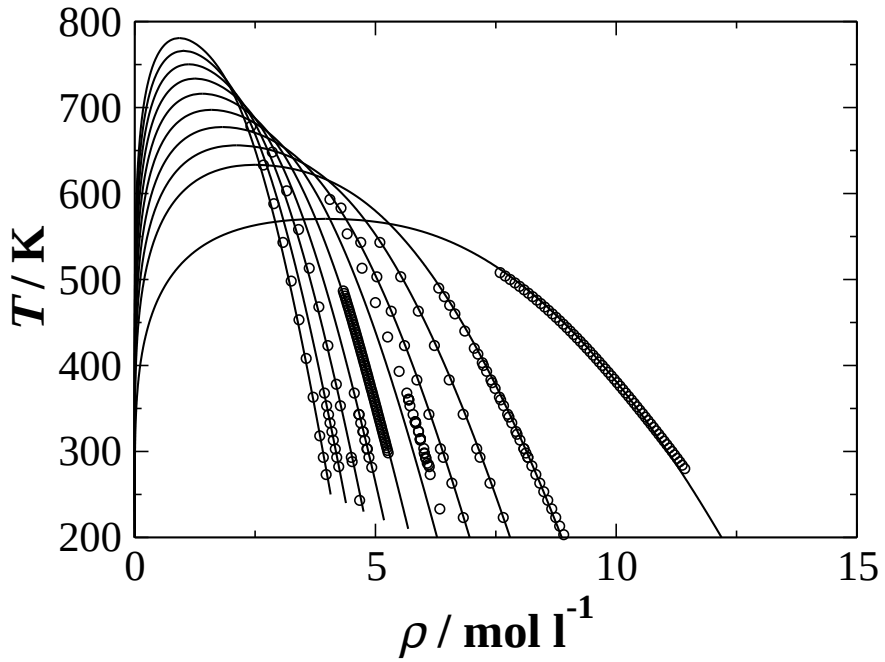

Figure 5.6: SAFT-γ Mie description of the coexistence densities as a function of temperature for benzene and selected alkylbenzenes (ethylbenzene to decylbenzene from right to left) included in the estimation of the aCH and $\mathrm{aCCH_2}$ group parameters. The symbols represent experimental data [280] and the continuous curves the calculations of the theory.

mation based on the pure component VLE data of selected alkylbenzenes, where benzene is also included in the estimation procedure. The identification of the functional groups for the modelling of alkylbenzenes is consistent with the one employed in previous work with the SAFT-γ method based on the square-well potential [19], and with the treatment within the framework of the original UNIFAC [83]. It is worth noting that the quadrupolar nature of benzene and the alkylbenzenes is not explicitly accounted for; it is described in an effective manner by means of dispersion forces. The theory provides an excellent description of the pure component phase behaviour of the correlated alkylbenzenes, with average deviations for all the compounds of 2.27% for the vapour pressure and 0.87% for the saturated liquid density. More details regarding the deviations per component (and per property) together with the specifics of the experimental data used are provided in table 5.4. From the results presented, it can be seen that the SAFT-γ Mie approach provides a significantly improved description of the pure component VLE of the correlated compounds compared to the corresponding results with the SAFT-γ SW approach [19]: for the same chemical family (excluding benzene) the reported %AADs with the SW version of the theory were 4.12% for the vapour pressure and 1.49% for the saturated liquid density [19]. The corresponding deviations with the SAFT-γ Mie approach are 2.47% for p_{vap} and 0.91% for ρ_{sat}.

The pure component fluid phase behaviour of the correlated alkylbenzenes is shown in figure 5.6 for the saturated liquid densities and figure 5.7 for the vapour pressure, in a logarithmic representation. The parameters for the aCH and aCCH_2 functional groups are summarised in table 5.6. For both groups, the value of the attractive exponent is fixed to $\lambda^a_{\text{aCH}} = \lambda^a_{\text{aCHCH}_2} = 6$, and the optimal values for the repulsive exponents were found to be $\lambda^r_{\text{aCH}} = 14.762$ and $\lambda^r_{\text{aCHCH}_2} = 9.117$, where it can be seen that as for the CHCH_3 group a more coarse-grained model of the aCCH_2 group leads again to lower values for the repulsive exponent. The segment size of the aCHCH_2 group is found to be significantly larger than the aCH group ($\sigma_{\text{aCCH}_2} = 5.041$ Å and $\sigma_{\text{aCH}} = 4.092$ Å), as expected based on the chemical structure of the groups.

5.3.3 Pure component parameters: carboxylic acids

A common feature of most active pharmaceutical ingredients is that they invariably feature one or more associating functional groups. One of the most frequently encountered groups is the carboxyl group (COOH), present in both APIs considered here (cf. figure 5.2). The parameters for the carboxyl group are obtained based on experimental data for the pure component fluid phase behaviour of selected members of the chemical family of carboxylic

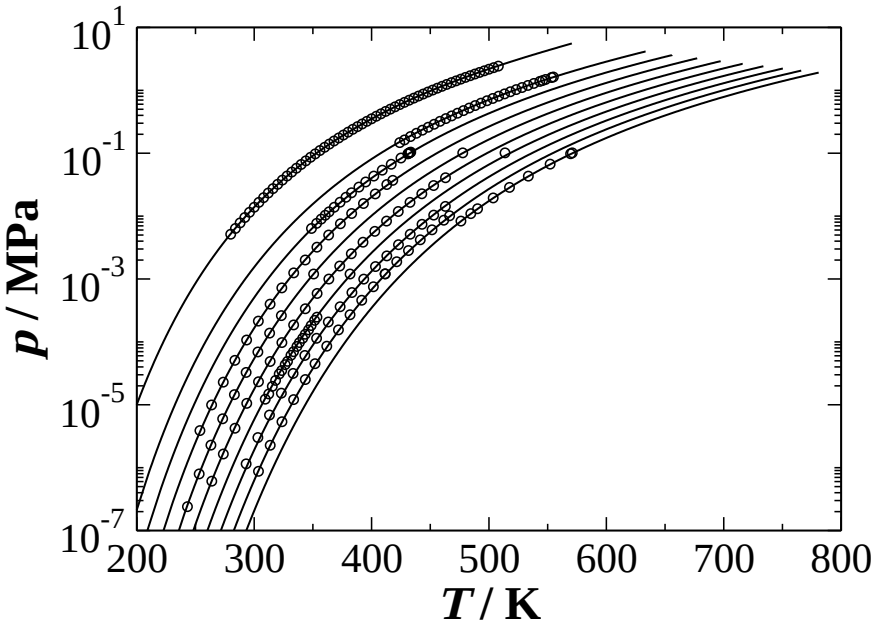

Figure 5.7: SAFT-γ Mie description of the vapour pressure as a function of temperature for benzene and selected alkylbenzenes (ethylbenzene to decylbenzene from left to right) included in the estimation of the aCH and $aCCH_2$ group parameters. The symbols represent experimental data [280] and the continuous curves the calculations of the theory.

acids, more specifically from propanoic to decanoic acid. The average deviations for the correlated compounds are 12.13% for the vapour pressure and 2.50% for the saturated liquid density. The variation in accuracy observed in the deviations for p_{vap} is related to the disparate temperature ranges of the experimental data used. For example, in the case of butanoic acid, data for the vapour pressure from 246 to 465 K are used, while for propanoic acid the temperature range is 328 to 438 K; the inclusion of data at lower temperatures combined with the lower absolute values of the vapour pressure of compounds of higher molecular weight at these conditions can lead to high relative errors. More details of the experimental data used and the deviations per property and per compound is given in table 5.5. The description of the pure component vapour-liquid equilibria of the correlated carboxylic acids is presented in figure 5.8 for the coexistence densities and figure 5.9 for the vapour pressure.

It is evident that the description of the carbocylic acids is not as accurate as for other chemical families presented so far. The large deviations could be related to the treatment of the association interactions where the approximation of the radial distribution function as that of an effective hard-core that allows for the simplified treatment of the association contribution presented in chapter 4, which is not as accurate for soft potentials. This

Table 5.5: Percentage average absolute deviations (%AAD) of vapour pressures $p_{\mathrm{vap}}(T)$ and saturated liquid densities $\rho_{\mathrm{sat}}(T)$ for the carbocylic acids obtained with the SAFT-γ Mie framework from the experimental data [280], where n is the number of data points.

Compound	T range (K)	n	%AAD p_{vap}	T range (K)	n	%AAD ρ_{sat}
propanoic acid	328-438	35	2.19	237-490	33	1.93
butanoic acid	293-452	39	6.06	273-563	21	2.13
pentanoic acid	246-465	33	24.08	233-553	26	2.53
hexanoic acid	355-478	38	16.02	273-588	27	2.64
heptanoic acid	348-494	40	11.21	263-583	16	2.64
octanoic acid	403-513	35	7.51	293-573	22	2.59
nonanoic acid	372-528	11	12.71	288-323	7	3.23
decanoic acid	350-543	13	17.23	313-573	14	2.30
Average	-	-	12.13	-	-	2.50

approximation is more appropriate for hard-core potentials, as testified by the level of accuracy achieved for associating compounds within the framework of the SAFT-γ method based on the square-well (hard-core) intermolecular potential, where specifically for the carboxylic acids the reported deviations were 3.60% for the vapour pressure and 1.42% for

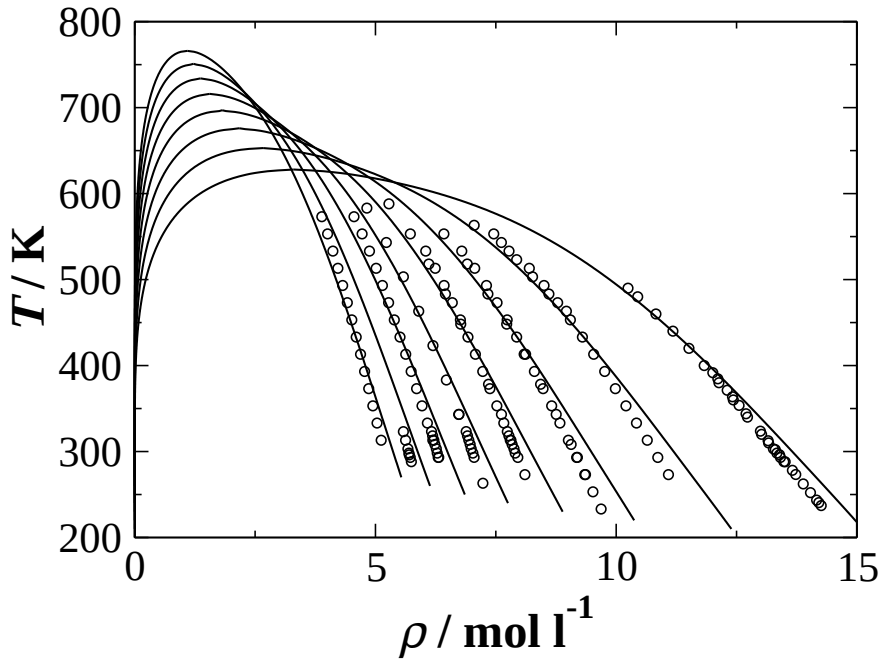

Figure 5.8: SAFT-γ Mie description of the coexistence densities for selected carboxylic acids (propanoic to decanoic acid from right to left) included in the estimation of the COOH group parameters. The symbols represent the experimental data [280] and the continuous curves the calculations of the theory.

the saturated liquid densities [29]. An alternative treatment of the association term is possible, e.g., by following the ideas employed in the development of the LJ-SAFT approach through the correlation of molecular simulation data to model associating systems [425]. This is however beyond the scope of the current work. The existing model is employed in the preliminary study of the solubilities presented here, since, as will be shown in the following section, it allows for a reasonable description of the nature of the interactions between functional groups and provides a good description of the fluid phase behaviour of mixtures.

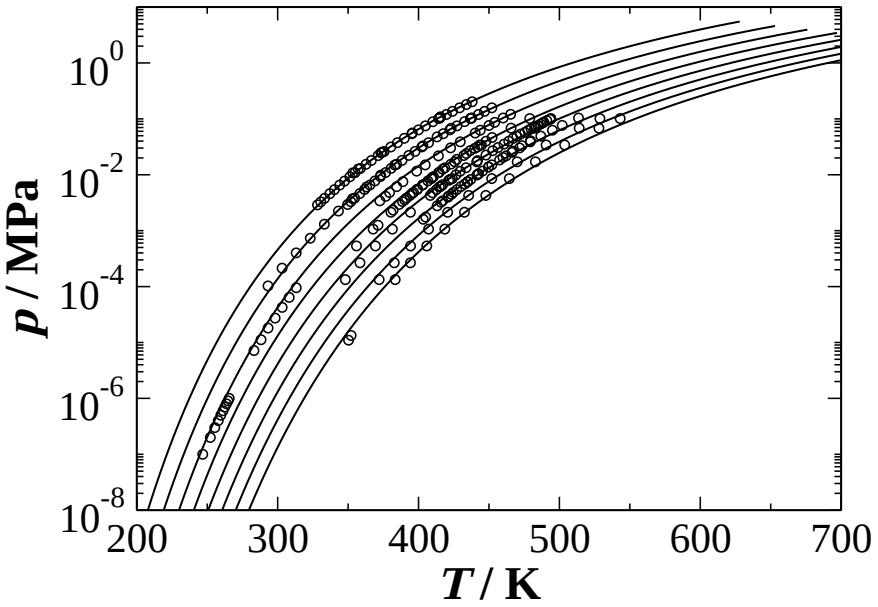

Figure 5.9: SAFT-γ Mie description of the vapour pressure for selected carboxylic acids (propanoic to decanoic acid from left to right) included in the estimation of the COOH group parameters. The symbols represent the experimental data [280] and the continuous curves the calculations of the theory.

The parameters that describe the COOH functional group are presented in table 5.5. The attractive exponent is fixed to the London dispersion value of $\lambda^a_{\mathrm{COOH}} = 6$ and the optimal value of the repulsive exponent is found to be $\lambda^r_{\mathrm{COOH}} = 14.679$. The COOH group is modelled as comprising 3 identical spherical segments ($\nu^*_{\mathrm{COOH}} = 3$), following the modelling approach of earlier work [19] in view of the relatively large size of the group. The functional group of the carboxylic acids is the only associating group considered in the current work, where the self-association between carboxylic groups is modelled with a single associating site, as carboxylic acids are known to dimerise. The additional association sites would allow for the formation of chains and network aggregates. It should be noted however, that in order to model the association with other components such as

water, additional association sites will have to be considered for the carboxylic acids in order to describe the complex unlike hydrogen-bonding.

5.3.3.1 Predictions of fluid phase behaviour from pure component data

The heteronuclear model employed within the SAFT-γ approaches allows for the unlike group interaction energy $\epsilon_{kl}/k_{\mathrm{B}}$ between groups k and l to be estimated directly from pure component data. This provides a capability for predicting the fluid phase behaviour of binary mixtures without the need for any experimental information of the mixtures, as previously shown in chapter 4 for binary mixtures of n-alkanes. The assessment of the predictions of the method for the fluid phase equilibria of binary mixtures is central to the validation of the physical robustness of the parameters estimated from pure component data alone and the group identification scheme employed. As shown in section 3.3.2, different modelling strategies for the functional group of a given chemical family can result in an equivalent accuracy in describing the pure component properties. It is the extrapolation to the representation of binary mixtures in a predictive manner that reveals the adequacy of the group identification procedure and the estimated parameter set, and that provides a test of the predictive capabilities of the approach. For the chemical families presented in this chapter, the respective group interaction parameters obtained are summarised in tables 5.6 and 5.7. The performance of the method in predicting the vapour-liquid phase behaviour of selected binary mixtures with parameters obtained from pure component

Table 5.6: Group parameters for the functional groups needed to model the solid-liquid equilibria of the target systems of figure 5.2. Each group is characterised by the number of identical segments the group comprises, ν_k^*, its shape factor, S_k, the repulsive and attractive exponents of the group-group interaction potential, λ_{kk}^r and λ_{kk}^a, the group segment size, σ_{kk}, and the group-group interaction energy, $\epsilon_{kk}/k_{\mathrm{B}}$. The association interactions are characterised by the number of sites of type a on group k, $n_{k,a}$, the energy, $\epsilon_{kkaa}^{\mathrm{HB}}/k_{\mathrm{B}}$, and the range, $r_{kkaa}^c/\bar{\sigma}_{ii}$, of the association interaction. The range is given in reduced units (reduced by the effective molecular segment diameter, $\bar{\sigma}_{ii}$).

Group k	ν_k^*	S_k	λ_{kk}^r [-]	λ_{kk}^a [-]	σ_{kk} [Å]	$(\epsilon_{kk}/k_{\mathrm{B}})$ [K]	$n_{k,a}$ [-]	$(\epsilon_{kk,aa}^{\mathrm{HB}}/k_{\mathrm{B}})$ [K]	$r_{kk,aa}^c/\bar{\sigma}_{ii}$ [-]
CH_3	1	0.572	15.050	6	4.077	256.766	-	-	-
CH_2	1	0.229	19.871	6	4.880	473.389	-	-	-
CH_3CO	2	0.628	17.508	6	3.748	463.412	-	-	-
$CHCH_3$	1	0.522	13.885	6	4.575	317.904	-	-	-
aCH	1	0.311	14.762	6	4.092	377.272	-	-	-
$aCCH_2$	1	0.245	9.117	6	5.041	501.642	-	-	-
COOH	3	0.941	14.679	6	2.524	375.360	1	2000	0.25

data alone is shown in figure 5.10. It is apparent from the figure that the theory provides accurate predictions of the vapour-liquid equilibria for a wide variety of mixtures, including polar and associating compounds. In particular, the predicted fluid phase behaviour of the binary mixture *n*-heptane+pentanoic acid (see figure 5.10.(d)) is in good agreement with the experimental data. This demonstrates that, despite the relatively high deviations observed in the description of the pure component carboxylic acids included in the estimation of the parameters for the COOH group, the values of the unlike COOH-CH_3 and COOH-CH_2 interactions allow for a satisfactory representation of the fluid phase behaviour of the mixture.

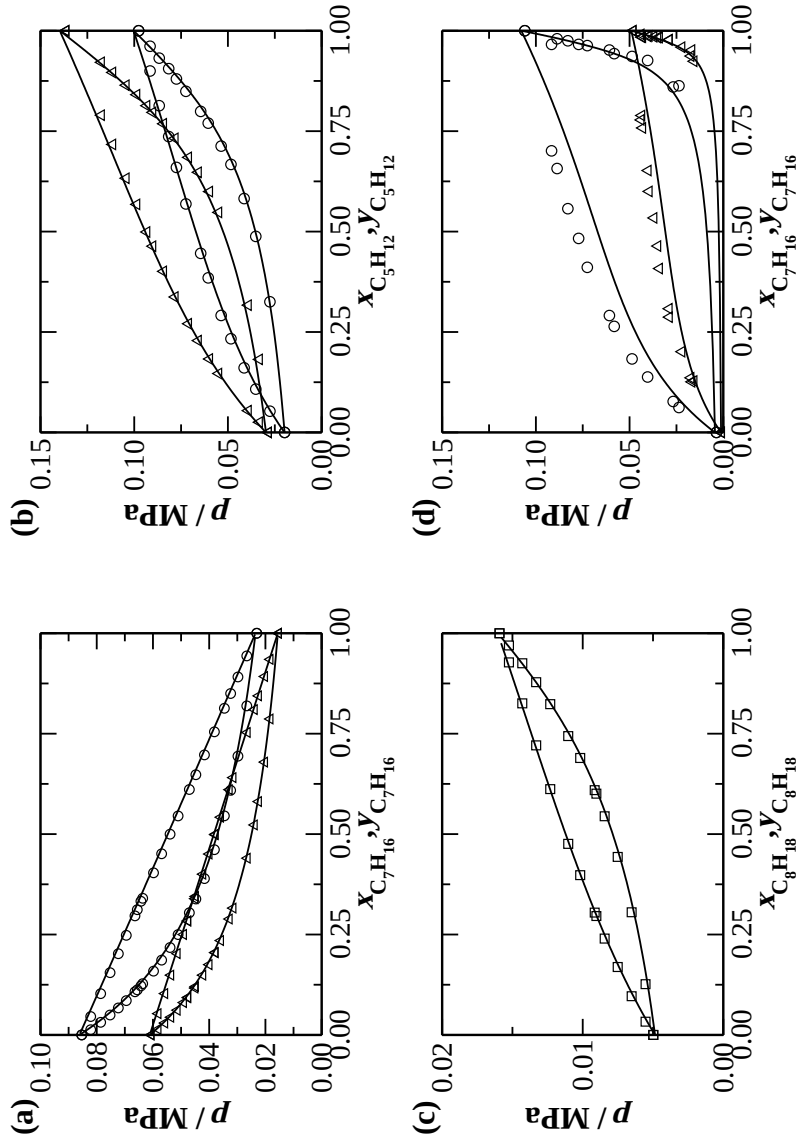

Figure 5.10: Comparison of the predictions of the SAFT-γ Mie approach and the experimental data for the vapour-liquid equilibria of selected binary mixtures: (a) 2-methylpentane+n-heptane at 318.15 K (triangles) and 328.15 K (circles) [426]; (b) benzene+n-pentane at 308.15 (circles) and 318.15 (triangles) [427]; (c) propylbenzene+n-octane at 343.15 K [428]; and (d) pentanoic acid+n-heptane at 348.15 (triangles) and 373.15 K (circles) [429]. In all figures, the symbols represent experimental data and the continuous curves the predictions of the theory.

5.3.4 Determination of unlike group interactions

The heteronuclear model employed within the SAFT-γ Mie formalism allows for the determination of the binary interaction energy $\epsilon_{kl}/k_{\mathrm{B}}$ between groups k and l from pure component data, when both groups k and l are present on the same molecule. The unlike interaction parameters obtained from pure component data during the characterisation of the functional groups presented in table 5.2, are given in table 5.7. Is it clear that in order to describe the systems of figure 5.2, a number of unlike group interaction parameters have to be determined. This is undertaken based on pure component data for molecules that feature multiple functional groups and an appropriate selection of mixture data.

5.3.4.1 Unlike group interactions from pure component data

For the determination of the unlike interactions between the branched alkyl, and the carbonyl and carboxyl groups, $CHCH_3$-CH_3CO and $CHCH_3$-COOH, experimental data for pure components are employed. This is mainly due to the availability of pure component data for compounds that feature the groups of interest, i.e., a carbonyl or carboxylic acid group and at the same time a branched alkyl chain. For each parameter pure component fluid phase behaviour data for a single compound can provide enough information, as only a single parameter has to be determined from each data set.

The compounds selected for this purpose are methyl-isobutyl ketone (MIBK) for the interaction between the $CHCH_3$ and CH_3CO functional groups and 2-methylbutanoic acid for the interaction between the $CHCH_3$ and COOH groups. The interaction parameters are

Table 5.7: Group interaction energies estimated for the functional groups needed to model the target system of figure 5.2. The values in bold denote parameters obtained from the experimental data of appropriate mixtures, while N/A denotes parameters that are not available. The symmetry of the group interaction matrix is highlighted by the diagonal boundary.

$\epsilon_{kl}/k_{\mathrm{B}}$ [K]	CH_3	CH_2	CH_3CO	$CHCH_3$	aCH	$aCCH_2$	COOH
CH_3	256.766	350.772	**334.437**	312.396	309.156	404.566	300.000
CH_2	350.772	473.389	**405.882**	396.016	419.959	439.841	412.991
CH_3CO	**334.437**	**405.882**	**463.412**	N/A	N/A	N/A	N/A
$CHCH_3$	312.396	396.016	N/A	317.904	N/A	N/A	N/A
aCH	309.156	419.959	N/A	N/A	377.272	410.567	N/A
$aCCH_2$	404.566	439.841	N/A	N/A	410.567	501.6416	N/A
COOH	300.000	412.991	N/A	N/A	N/A	N/A	375.360

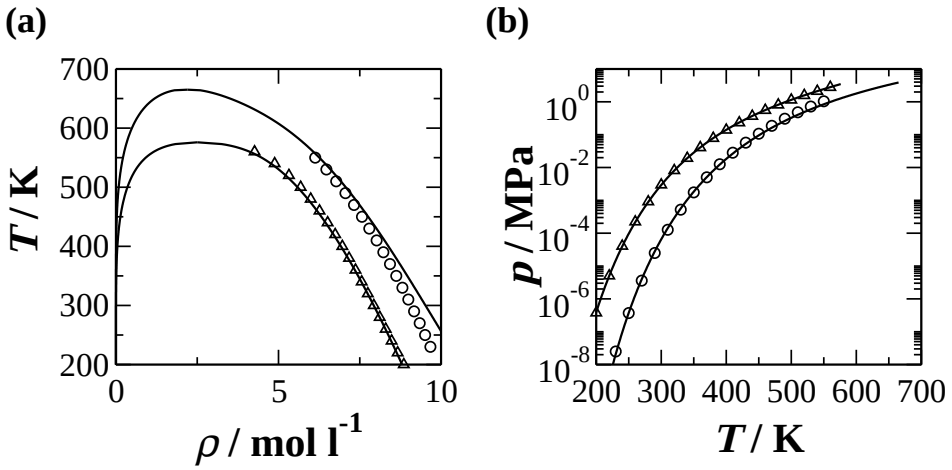

Figure 5.11: SAFT-γ Mie description of: (a) the coexistence densities, and (b) the vapour pressures for the pure components used in the estimation of unlike group interaction parameters. The triangles represent the experimental data for methyl-isobutyl-ketone (MIBK) [430], used to determine the $CHCH_3$–CH_3CO interaction, and the circles the data for 2-methylbutanoic acid [430], used for the $CHCH_3$–COOH interaction. The continuous curves are the calculations with the SAFT-γ Mie theory.

obtained based on vapour pressure and saturated liquid density data, using the objective function discussed in sections 3.1.1 and 4.3. The optimal values of the interaction energies are included in table 5.8, and the agreement between the calculations of the theory and the experimental data is shown in figure 5.11. The deviations of the SAFT-γ Mie calculations from the experimental data are 1.20% and 9.95% for the vapour pressure, and 0.95% and 3.59% for the saturated liquid density, for MIBK and 2-methylbutanoic acid, respectively. In line with the results obtained earlier for the carboxylic acids, the deviations for 2-methylbutanoic acid are somewhat higher than for MIBK.

5.3.4.2 Unlike group interactions from mixture data

For certain unlike group interactions, experimental data for pure compounds that feature the groups of interest can be difficult to find. In such situations, the estimation of unlike group interaction parameters can be obtained from experimental data for binary mixtures. The unlike interaction parameters for the following combinations of groups are obtained from data for the respective systems:

- CH_3CO – aCH from the acetone+benzene binary mixture;
- CH_3CO – $aCCH_2$ from the acetone+ethylbenzene binary mixture;
- $CHCH_3$ – aCH from the MIBK+benzene binary mixture;

Table 5.8: Group interaction energies estimated for the functional groups needed to model the target system of figure 5.2. The parameters in bold denote values obtained from the experimental data of appropriate mixtures. The symmetry of the group interaction matrix is highlighted by the diagonal boundary.

ϵ_{kl}/k_B [K]	CH_3	CH_2	CH_3CO	$CHCH_3$	aCH	$aCCH_2$	COOH
CH_3	256.766	350.772	**334.437**	312.396	309.156	404.566	300.000
CH_2	350.772	473.389	**405.882**	396.016	419.959	439.841	412.991
CH_3CO	**334.437**	**405.882**	**463.412**	329.042	**405.280**	**419.340**	**428.740**
$CHCH_3$	312.396	396.016	329.042	317.904	**329.340**	**153.700**	358.890
aCH	309.156	419.959	**405.280**	**329.340**	377.272	410.567	**375.800**
$aCCH_2$	404.566	439.841	**419.340**	**153.700**	410.567	501.641	**285.136**
COOH	300.000	412.991	**428.740**	358.890	**375.800**	**285.136**	375.360

- $CHCH_3$ – $aCCH_2$ from the MIBK +ethylbenzene binary mixture;
- COOH – aCH from the butanoic acid+benzene binary mixture;
- COOH – $aCCH_2$ from the propanoic acid+ethylbenzene binary mixture;
- COOH – CH_3CO from thebutanoic acid+butanone binary mixture.

In the majority of cases data of two binary mixtures are used to estimated each class of group parameters. This is appropriate when interaction parameters between a given group and the groups of the aromatic hydrocarbons (aCH, $aCCH_2$) are estimated, where typically a first interaction parameter between group k and aCH is obtained from a mixture with benzene, and the corresponding value is then transferred to represent a mixture with an aklylbenzene for the unlike interaction parameter with the $aCCH_2$ functional group. The agreement between the calculations of the SAFT-γ Mie approach and the experimental data for the mixtures in question can be assessed in figure 5.12. It is apparent that the interaction parameters estimated in this way provide an accurate representation of the fluid phase behaviour of the mixtures included in the regression. It should be noted that for the binary mixture of butanoic acid+2-butanone (see figure 5.12.(d)) the description of the theory is based on the assumption that there are no specific association interactions between the the carboxylic acid and the ketone groups, so that the only adjustable parameter in this case is the unlike dispersion energy, $\epsilon_{CH_3CO-COOH}$. The satisfactory representation of the fluid phase behaviour of the system at two temperatures supports this assumption. The values of all the unlike group interaction parameters estimated are summarised in table 5.8.

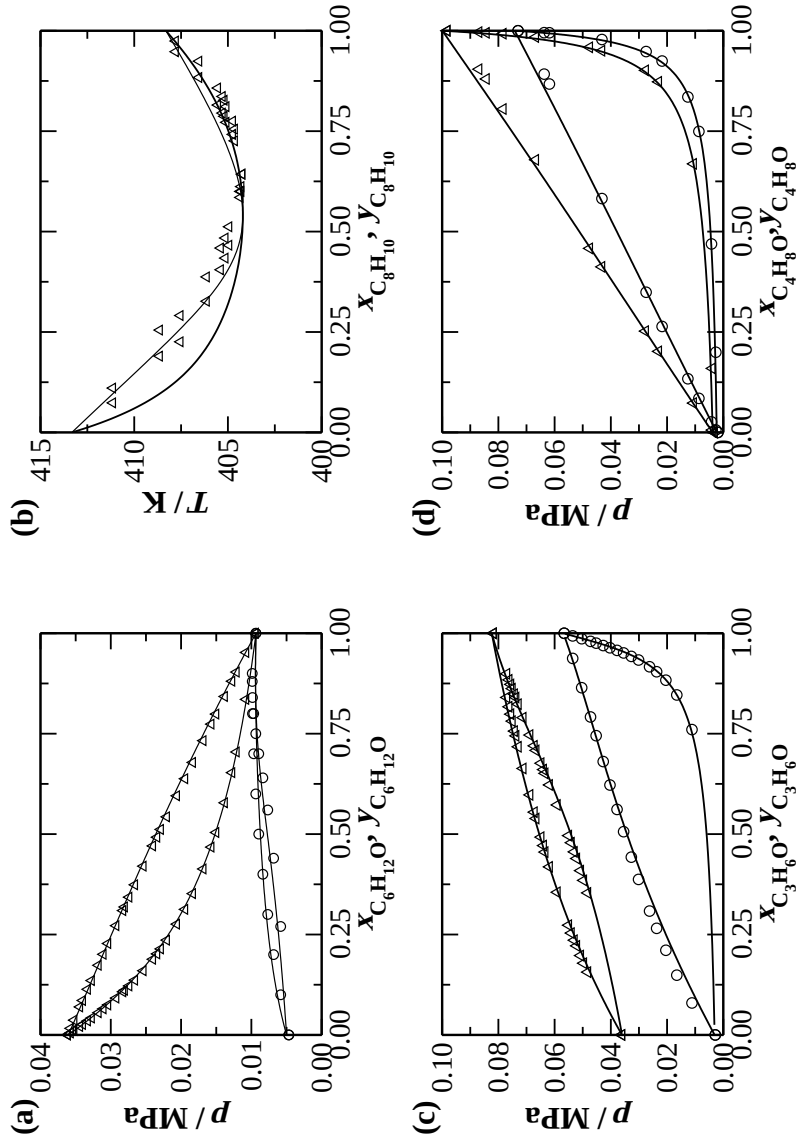

Figure 5.12: Comparison of the description of the SAFT-γ Mie approach and the experimental data for the fluid phase behaviour of selected binary mixtures used for the estimation of unlike group interaction parameters: (a) methyl-isobutyl ketone (MIBK)+benzene at 323.15 K (triangles) [431], and MIBK+ethylbenzene at 323.15 K (circles) [432]; (b) ethylbenzene+propanoic acid at 0.1 MPa [433]; (c) acetone+benzene at 323.15 K [431] (triangles), and acetone+ethylbenzene at 313.15 K [391] (circles); and (d) butanone+butanoic acid at 343.15 K (circles) and 352.15 K (triangles) [434]. In all the figures the symbols represent the experimental data and the continuous surves the calculations of the theory.

5.4 Predictions of solid-liquid equilibria

Solid-liquid equilibrium (SLE) is calculated with the approach outlined in section 5.1, where experimental values for the enthalpy of fusion and the melting temperature of each solute are used, and the activity coefficient of the solute in the liquid phase is calculated by means of the SAFT-γ Mie approach. In the derivation of eq. (5.8), the solid phase is assumed to consist only of pure solute and the contribution of the terms that arise due to the difference between the heat capacity in the solid and the liquid phase is neglected. The performance of the SAFT-γ Mie GC approach in the description of the SLE of real systems is now assessed based on the aforementioned assumptions. In figure 5.13 the predictions of the theory are compared with the experimental data for the solubility of the soluid form of long chain linear n-alkanes (n-dodecane, n-hexadecane and n-eicosane) in liquid n-hexane. The parameters for the CH_3 and CH_2 groups, and the unlike group interactions (cf. tables 5.6 and 5.7), that describe these mixtures are obtained entirely from pure component vapour-liquid equilibrium experimental data for the linear n-alkanes, from ethane to n-decane; this means that the corresponding calculations are fully predictive for the SLE of these mixtures. The ability to obtain information regarding the unlike interactions between groups from pure component data is a unique feature of the heteronuclear models inherent in SAFT-γ. From the figure it can be seen that the solubilities predicted with the SAFT-γ Mie EoS are in very good agreement with the experimental data for the solid-liquid equilibrium phase boundary, with parameters obtained from fluid phase properties. The experimental values for the heat of fusion and melting temperatures of the solutes examined in the current work are summarised in table 5.9.

As has been discussed, fluid theories can be employed to describe solid-liquid equilibria, by determining the activity coefficient of the solute in the solute+solvent liquid mixture (provided of course that T_m and Δh_{fus} are available). However, in some cases, further insight on the behaviour of the solid phase is necessary in order to achieve an accurate representation of solid-liquid equilibria. A case in point is the study of the solubility of compounds where the solute undergoes a solid-solid transition. This occurs, for example, in the chemical family of n-alkanes, mostly for compounds with more than 22 and an even number of carbon atoms, where solid-solid phase transitions occur at temperatures close to the melting point [437]. The enthalpy of this transition Δh_{SS} (occuring at a temperature T_{SS}) is of the same order as the heat of fusion (e.g., for n-$C_{22}H_{46}$, $\Delta h_{\text{fus}} \approx 48$ kJ mol^{-1} and $\Delta h_{\text{SS}} \approx 29$ kJ mol^{-1}) and therefore has a significant impact on the solid-liquid phase boundary. Such transitions can be accounted for with the current framework for the modelling of solid-liquid equilibria by extending the thermodynamic cycle described in

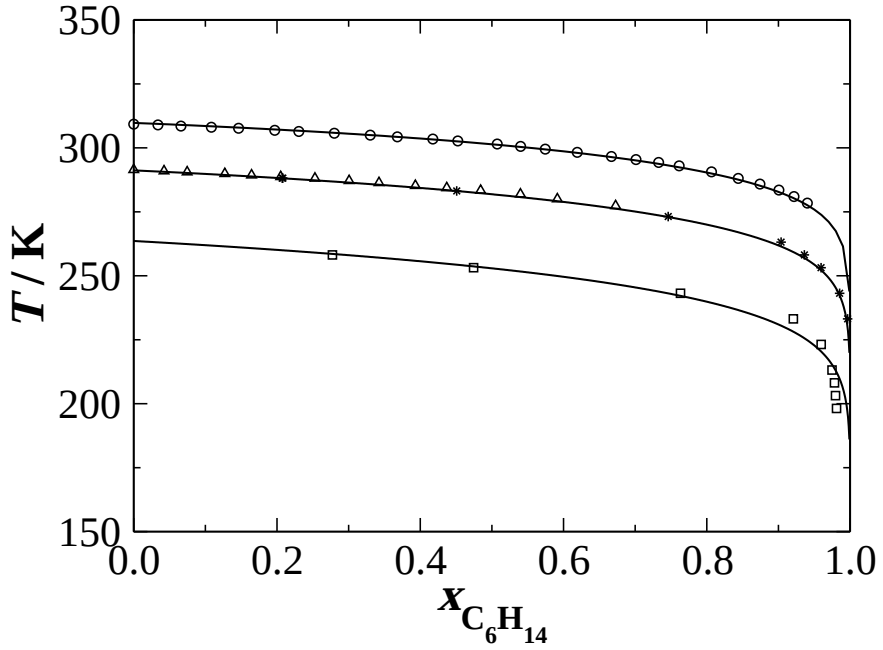

Figure 5.13: Prediction of temperature-composition representation of the solid-liquid equilibria of binary n-alkane mixtures at ambient pressure (1 atm). The composition $x_{C_6H_{14}}$ is that of the solvent shown in mole fraction and the symbols represent the experimental solid-liquid phase boundary for different solutes (squares for n-dodecane [435], triangles [436] and stars [435] for n-hexadecane, and circles [436] for n-eicosane). The continuous curves represent the predictions of the SAFT-γ Mie approach.

section 5.1, as shown in figure 5.14. On comparing the thermodynamic cycles of figures 5.1 and 5.14, it becomes apparent that the solid-solid phase transition is described by the path $a \rightarrow b \rightarrow c$ of figure 5.14.

Based on this thermodynamic cycle, the ratio of the fugacities between the liquid and the second solid phase can be calculated based on the change of the molar Gibbs free energy

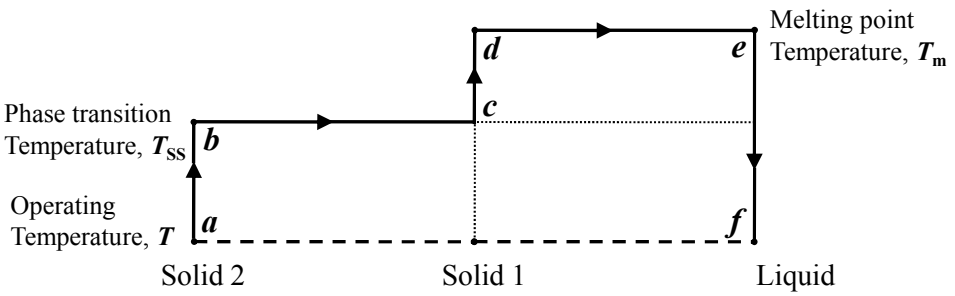

Figure 5.14: Thermodynamic cycle for the derivation of the working equation for the study of solid-liquid equilbria, where the compound undergoes a solid-solid phase transition.

along the path $a \rightarrow f$:

$$\Delta g_{a\rightarrow f} = RT\ln\frac{f_{slt}^{\mathrm{L},*}(T,p)}{f_{slt}^{\mathrm{S2},*}(T,p)} \ . \tag{5.9}$$

The molar Gibbs free energy is calculated based on the enthalpy and entropy changes along the path $a \rightarrow b \rightarrow c \rightarrow d \rightarrow e \rightarrow f$, in a similar way as in section 5.1. For the enthalpy,

$$\begin{aligned}
\Delta h_{a\rightarrow f} &= \Delta h_{a\rightarrow b} + \Delta h_{b\rightarrow c} + \Delta h_{c\rightarrow d} + \Delta h_{d\rightarrow e} + \Delta h_{e\rightarrow f} \\
&= \int_{T}^{T_{\mathrm{SS}}} c_p^{\mathrm{S2}} dT + \Delta h_{\mathrm{SS}}(T_{\mathrm{SS}}) + \int_{T_{\mathrm{SS}}}^{T_{\mathrm{m}}} c_p^{\mathrm{S1}} dT \\
&+ \Delta h_{\mathrm{fus}}(T_{\mathrm{m}}) + \int_{T_{\mathrm{m}}}^{T} c_p^{\mathrm{L}} dT \\
&= \Delta h_{\mathrm{fus}}(T_{\mathrm{m}}) + \Delta h_{\mathrm{SS}}(T_{\mathrm{SS}}) + \int_{T_{\mathrm{m}}}^{T_{\mathrm{SS}}} \Delta c_p^{\mathrm{I}} dT + \int_{T_{\mathrm{SS}}}^{T} \Delta c_p^{\mathrm{II}} dT \ ,
\end{aligned} \tag{5.10}$$

where two integrals relating to the heat capacity differences refer to the two distinct temperature ranges, defined as $\Delta c_p^{\mathrm{I}} = c_p^{\mathrm{L}} - c_p^{\mathrm{S1}}$ and $\Delta c_p^{\mathrm{II}} = c_p^{\mathrm{L}} - c_p^{\mathrm{S2}}$. Making use of the thermodynamic cycle for the calculation of the entropy changes, the following expression is derived:

$$\begin{aligned}
\ln x_{slt}^{\mathrm{L}} &= -\frac{\Delta h_{\mathrm{fus}}}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{m}}}\right) - \frac{\Delta h_{\mathrm{SS}}}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{SS}}}\right) \\
&- \frac{\Delta c_p^{\mathrm{I}}}{R}\left(\frac{T_{\mathrm{m}}}{T_{\mathrm{SS}}} - 1\right) + \frac{\Delta c_p^{\mathrm{I}}}{R}\ln\frac{T_{\mathrm{m}}}{T_{\mathrm{SS}}} \\
&- \frac{\Delta c_p^{\mathrm{II}}}{R}\left(\frac{T_{\mathrm{SS}}}{T} - 1\right) + \frac{\Delta c_p^{\mathrm{II}}}{R}\ln\frac{T_{\mathrm{SS}}}{T} \\
&- \ln\gamma_{slt}^{\mathrm{L}}(T,p,\underline{x}^{\mathrm{L}}) \ .
\end{aligned} \tag{5.11}$$

This expression can be simplified by neglecting the terms that account for the contribution due to the differences in the heat capacity between the liquid and the two solid phases (Δc_p^{I} and Δc_p^{II}). The simplified final equation is of the following form [438]:

$$\ln x_{slt}^{\mathrm{L}} = -\frac{\Delta h_{\mathrm{fus}}}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{m}}}\right) - \frac{\Delta h_{\mathrm{SS}}}{R}\left(\frac{1}{T} - \frac{1}{T_{\mathrm{SS}}}\right) - \ln\gamma_{slt}^{\mathrm{L}}(T,p,\underline{x}^{\mathrm{L}}) \ , \tag{5.12}$$

An example of the effect of such transitions on the solid-liquid phase boundary is shown in figure 5.15, where the theoretical predictions which neglect the solid-solid transition of *n*-dotriacontane (*n*-$C_{32}H_{66}$) are compared to the experimental solubility data for the *n*-dotriacontane+*n*-hexane mixture. The inclusion of the term accounting for the transition

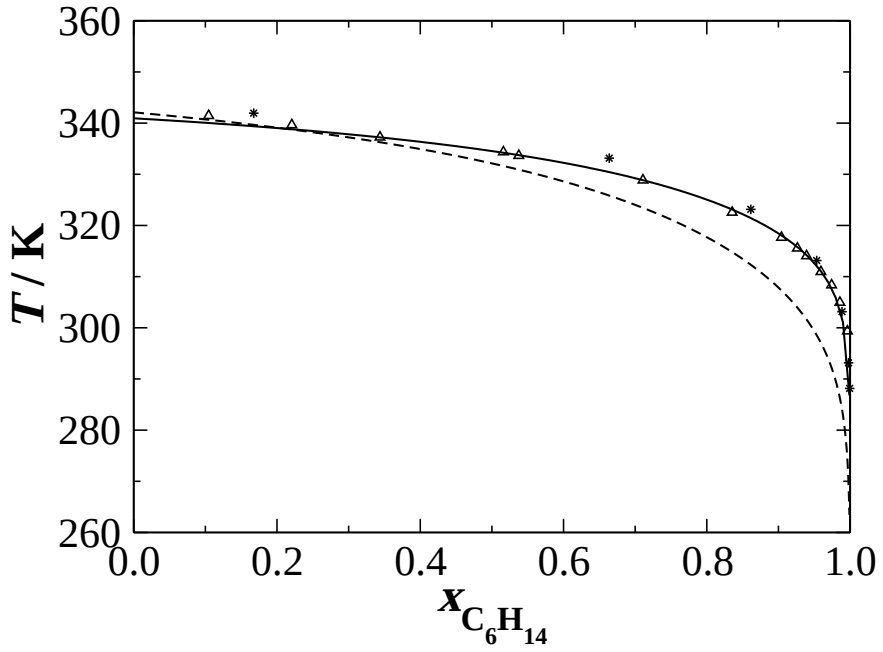

Figure 5.15: Comparison of the predictions of the SAFT-γ Mie approach for the solid-liquid equilibria of the binary n-hexane (solvent) + n-dotriacontane (solute) mixture with the experimental data at 1 atm. (triangles [439] and stars [435]). The dashed curves are predictions of the theory neglecting the solid-solid transition of the system (eq. 5.8), while the continuous curves represent predictions accounting for the transition (eq. 5.12).

in the solid phase is shown to improve significantly the accuracy of the calculations.

The solid-liquid equilibria of a wide range of systems can be determined predictively based on the parameters and the unlike group interaction values summarised in table 5.7, obtained for the chemical families described in sections 5.3.1-5.3.3. The description of the SAFT-γ Mie approach for the solid-liquid equilibria for selected binary systems is shown in figure 5.16. The experimental values of the heat of fusion and melting temperature for the solute in each of these systems are given in table 5.9; in cases where further information about transitions in the solid phase are available, eq. (5.12) is used. It is apparent that the SAFT-γ Mie approach provides a good predictive platform for the solid-liquid equilibria of a wide range of systems, including polar and associating compounds, in good agreement with the experimental data. Of particular mention is the example shown in figure 5.16.(b), where the theory is shown to accurately predict the SLE of long-chain n-alkanes in benzene, with the correct description of the location and the composition of the eutectic point of the mixtures. The slight deterioration of the theory for these mixtures on the branch of the SLE where benzene is the solute may be attributed to the chemical nature (quadrupole) of benzene which is expected to have a marked effect on the

solid-liquid equilibria of the system. As mentioned previously, multipolar interactions are not accounted for explicitly in our current theory, though, they can however be treated by including additional terms in the free energy (e.g. quadrupolar terms [190]). The decision to treat polar or quadrupolar compounds in an effective manner for all the interactions of the system through effective dispersion forces is justified based on the performance of the SAFT-γ Mie approach in the description of the vapour-liquid equilibria of pure components and mixtures (cf. chapter 4). Nevertheless, such effects are expected to be of greater importance in the solid phase than in the fluid (vapour and liquid) phases. The performance of the theory in describing the SLE of polar and quadrupolar compounds with the inclusion of a polar term in the free energy will be the subject of future work.

The predictions of the theory for the SLE in mixtures of carboxylic acids+n-hexane (see figure 5.16.(d)) are seen to be in good agreement with the experimental data. Despite the high deviations of the vapour-liquid equilibria reported when developing the parameters for the COOH group, the effect of the unlike interaction energies between the carboxyl and the methyl and methylene groups on the solid-liquid equilibrium appears to be adequately described by the theory. Finally, it is important to note that the predictions shown in figure 5.16 are obtained with parameters developed for the description of the vapour-liquid equilibria in the corresponding binary mixtures (cf. figure 5.10). The simultaneous description of different types of phase equilibria with a single set of parameters is a nontrivial task, and lends credence to the versatility and the physical robustness of the SAFT-γ Mie GC approach.

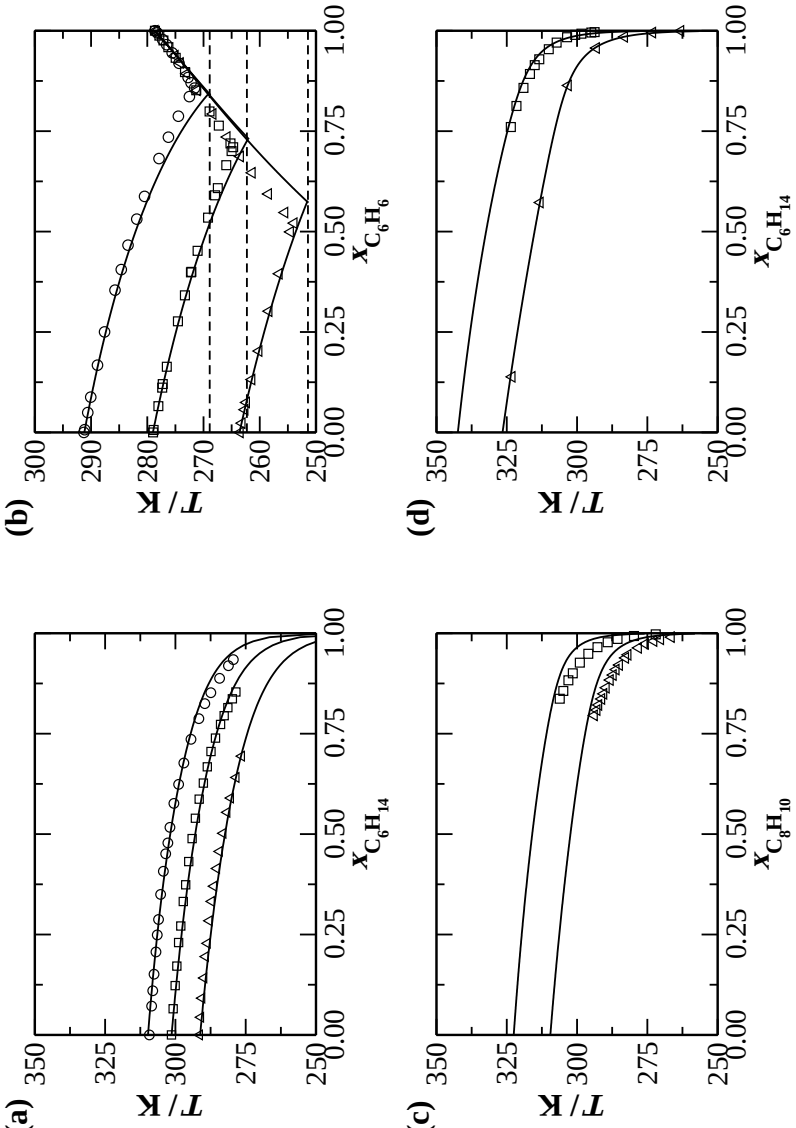

Figure 5.16: Comparison of the predictions of the SAFT-γ Mie approach and the experimental data for the solid-liquid equilibria of selected binary mixtures: (a) 3-methylpentane (solvent) + *n*-hexadecane (triangles), 3-methylpentane + *n*-octadecane (squares), 3-methylpentane + *n*-eicosane (circles) [436]; (b) benzene + *n*-dodecane (triangles), benzene + *n*-tetradecane (squares), benzene + *n*-hexadecane (circles) [440]; (c) ethylbenzene (solvent) + *n*-eicosane (triangles), ethylbenzene + *n*-tetracosane (squares) [441]; and (d) *n*-hexane (solvent) + tetradecanoic acid (triangles) [435], *n*-hexane + octadecanoic acid (squares) [442]. The dashed lines in subfigure (b) denote the predicted location of the eutectic point. In all figures the symbols represent the experimental data at 1 atm and the continuous curves the calculations of the theory.

Table 5.9: Melting points and heats of fusion for the solutes: benzene (C_6H_6), linear alkanes (from n-$C_{12}H_{26}$ to n-$C_{32}H_{66}$) and long chain carboxylic acids ($C_{13}H_{27}COOH$ and $C_{17}H_{35}COOH$). Where applicable the solid-solid transition temperature and enthalpy are also given.

Solute	T_m [K]	Δh_{fus} [J mol^{-1}]	T_{SS} [K]	Δh_{SS} [J mol^{-1}]	Ref.
C_6H_6	278.7	9934	-	-	[443]
$C_{12}H_{26}$	263.6	36755	-	-	[437]
$C_{14}H_{30}$	279.0	45030	-	-	[437]
$C_{16}H_{34}$	291.2	53332	-	-	[437]
$C_{18}H_{38}$	301.2	61306	-	-	[437]
$C_{20}H_{42}$	309.5	69730	-	-	[437]
$C_{24}H_{50}$	323.6	54396	320.7	31701	[437]
$C_{32}H_{66}$	342.5	75758	337.1	40835	[437]
$C_{13}H_{27}COOH$	326.5	45000	325.3	6400	[444]
$C_{17}H_{35}COOH$	342.4	63200	325.9	5700	[444]

5.5 Solubility predictions for APIs

The solubilities of two APIs, phenylacetic acid and ibuprofen, in two solvents, acetone and methyl-isobutyl ketone (MIBK), are considered as a preliminary assessment of the performance of the theory. The molecular structure of the four compounds is presented in figure 5.17 highlighting the identification of the constituent functional groups; the types of functional groups and instances of each group per molecule are summarised in table 5.1. For each system the solubilities are calculated by means of eq. (5.8), using the experimental values for the heat of fusion and the melting temperature [405]: for phenylacetic acid (PAA), $\Delta h_{fus,PAA} = 15.5$ kJ mol^{-1} and $T_{m,PAA} = 349.15$ K; and for ibuprofen, $\Delta h_{fus,Ibu} = 25.5$ kJ mol^{-1} and $T_{m,Ibu} = 347.15$ K.

Phenylacetic acid is modelled by identifying 3 different functional groups and a total of 8 occurrences of structural units (5 aCH groups, 1 $aCCH_2$ and 1 COOH group). A comparison of the solubility predictions with the SAFT-γ Mie approach and the experimental solubility data in acetone and MIBK is shown in figure 5.18. The temperature dependence of the solubility of the API is represented here in grams of solute per kilogram of solute-free solvent, a representation that allows for a clear distinction between the solubility in the two different solvents. From the figure it can be seen that the predictions of the SAFT-γ Mie approach for the solubility of phenylacetic acid are in good agreement with the experimental data for both solvents examined. Our approach correctly describes the significant increase in the solubility of phenylacetic acid in acetone, compared with that

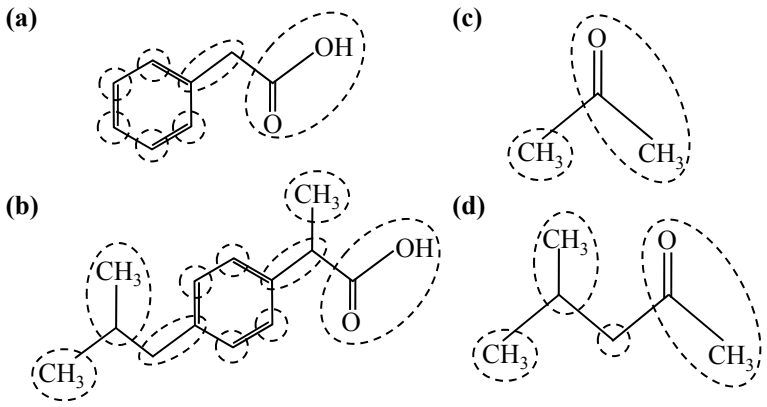

Figure 5.17: Molecular structures of the solute APIs, (a) phenylacetic acid and (b) ibuprofen, and the solvents, (c) acetone and (d) methyl-isobutyl ketone (MIBK). of the mixtures studied. The functional groups identified in the modelling of the compounds are highlighted with dashed curves.

in MIBK, at a given temperature. The quality of the description with the SAFT-γ Mie method compares favourably with the calculations using the original UNIFAC and the parameters of Hansen *et al.* [83] for the same mixtures, which are presented in figure 5.18. A fair comparison of the two methodologies is possible in this instance as only VLE data are employed in the development of the group parameters in both the SAFT-γ Mie and

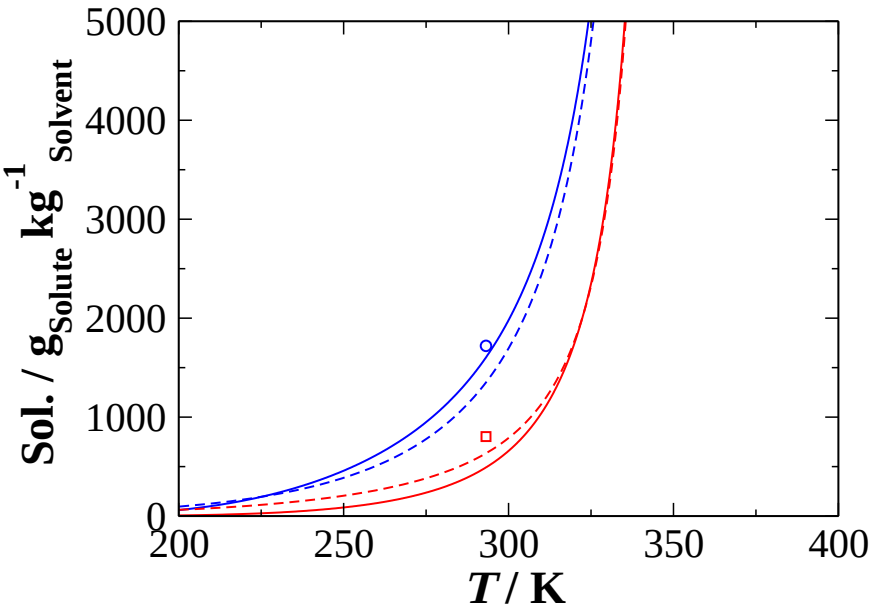

Figure 5.18: Comparison of the predictions of the SAFT-γ Mie approach (continuous curves) and the original UNIFAC method (dashed curves) with parameters from Hansen *et al.* [83] with the experimental data for the solubility of phenylacetic acid in acetone (circles) and methyl-isobutyl ketone [MIBK] (squares) [405] at 1 atm.

UNIFAC approaches. Admittedly, a large number of the group interaction parameters are obtained from pure component data alone in our SAFT-γ Mie formalism.

As a second preliminary case study, the solubility of the API ibuprofen was examined. It has a more complex molecular structure than phenylacetic acid, and within the context of the current work is modelled as comprising 5 distinct functional groups: 4 aCH group, 1 COOH group, 1 $CHCH_3$ group, 2 CH_3 groups, and 2 $aCCH_2$ groups (cf. table 5.1). The predictions with the SAFT-γ Mie approach are compared with the experimental data for the solubility of ibuprofen in acetone and MIBK in figure 5.19. From the figure it can be seen that the predictions of the theory are in good agreement with the experiments for the solubility of the API in MIBK, bearing in mind the complexity of the mixture which features a total of seven distinct functional groups. For the solubility in acetone, higher deviations between the predicted and the experimental values are apparent. Nevertheless, the ranking of the solvents in terms of their ibuprofen solubility is reproduced by the theory. The original UNIFAC method is shown to describe the solubility of ibuprofen in acetone more accurately; for MIBK the description with both methods is of the same quality. It has to be noted that within our current study the aCCH group present on ibuprofen is described as an additional instance of the $aCCH_2$, whereas within UNIFAC, the aCCH group is modelled explicitly.

The preliminary study of the solubility of two APIs presented here is an attempt to examine the performance of the SAFT-γ Mie approach in describing the solid-liquid equilibria of complex systems that feature multiple functional groups. It is important to reiterate that all the functional groups are developed based on experimental information of the fluid phase behaviour of pure components and selected binary mixtures. Including solid-liquid equilibrium data in the development of the group parameters is hoped to increase the accuracy of the predictions of the theory in the study of solubilities. However, the predictive capability of the approach presented here would still be retained, as no solubility data for the actual complex system under study is required. Based on the transferability of the parameters inherent in the GC formalism, solid-liquid equilibrium data for simpler systems that feature the functional groups of interest can be employed for the development of the group parameters.

5.6 Concluding Remarks

In this chapter the performance of the SAFT-γ Mie approach in predicting solid-liquid equilibria (SLE) is examined. The solubility limit is calculated following a thermodynamic

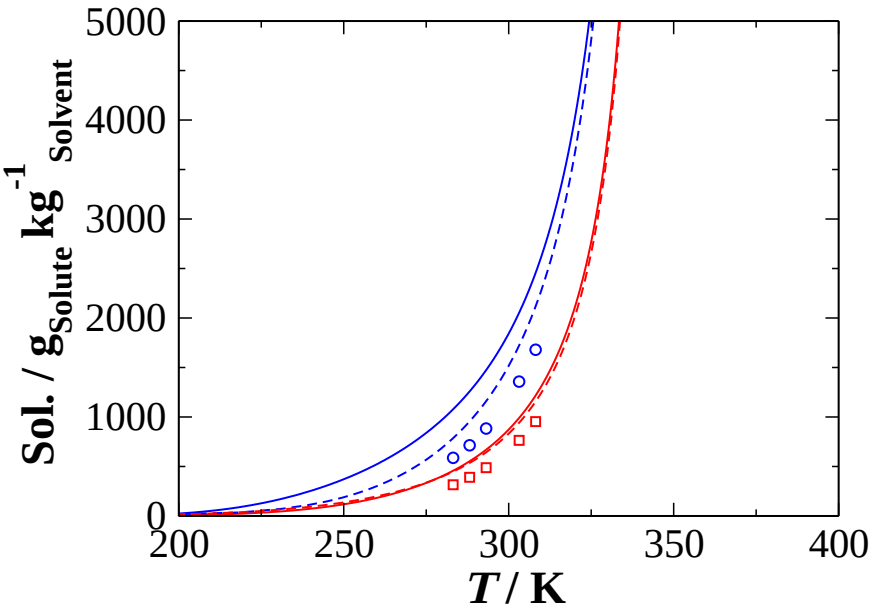

Figure 5.19: Comparison of the predictions of the SAFT-γ Mie approach (continuous curves) and the original UNIFAC method (dashed curves) with parameters from Hansen *et al.* [83] with the experimental data for the solubility of ibuprofen in acetone (circles) and methyl-isobutyl ketone [MIBK] (squares) [405] at 1 atm.

route, which requires a knowledge of the melting temperature T_m and the heat of fusion Δh_{fus} of the solute, as well as the activity coefficient of the solute in the solvent. The latter is calculated here by means of the SAFT-γ Mie approach. A simplified thermodynamic scheme is used, derived under the assumption that the difference in the heat capacities of the solute in the liquid and the solid phases have a negligible effect on its solubility. The SAFT-γ Mie approach is shown to provide an excellent description of a wide variety of binary systems of hydrocarbons, including polar (benzene and alkylbenzenes) and associating (carboxylic acids) compounds.

Given the promising performance of the SAFT-γ Mie approach in the description of the SLE of simple systems, the method is subsequently applied to the study of the solubility of active pharmaceutical ingredients (APIs) in organic solvents. For this preliminary study the solubilities of two APIs, phenylacetic acid and ibuprofen, in two solvents, acetone and methyl-isobutyl ketone, are assessed. The majority of the group parameters for the target systems are obtained from pure component vapour-liquid phase behaviour data for the chemical families of branched alkanes, alkylbenzenes, and carboxylic acids. For the families studied the average deviations from the correlated data are 5.36% for the vapour pressure and 1.30% for the saturated liquid density; the largest deviations are seen in the description of the carboxylic acids, due to approximations in the treatment of the associ-

ating effects. Whenever needed, unlike interaction parameters not available through pure component data (e.g., CH_3CO-aCH, CH_3CO-aCCH_2, $CHCH_3$-aCH, $CHCH_3$-aCCH_2) are obtained based on experimental data for the fluid phase behaviour of binary mixtures.

Based on the parameters estimated from the fluid phase properties, the performance of the SAFT-γ Mie theory is examined in the prediction of the solid-liquid equilibria of selected binary mixtures including alkanes, branched alkanes, alkylbenzenes, and carboxylic acids. It is shown that the theory can describe with equal accuracy the vapour-liquid and the solid-liquid equilibria exhibited by these binary mixtures with a single parameter set. This challenging task demonstrates the versatility and physical robustness of the SAFT-γ Mie approach for the description of non-ideal systems over a wide range of conditions and for different types of phase equilibria.

Finally the SAFT-γ Mie method is applied in a preliminary study of the solubility of complex organic pharmaceutical molecules in organic solvents, by transferring the previously determined group parameters. It is shown that the predictions of the theory for these challenging mixtures comprising multiple functional groups are in reasonable agreement with experiment, for phenylacetic acid and ibuprofen in acetone and methyl-isobutyl ketone. It is important to note that in both cases, the method predicts the correct ranking of the solubility in the two solvents, which is of particular relevance to solvent screening. The predictions of the theory are seen to compare well to the predictions of the well-established UNIFAC approach. Bearing in mind the complexity of these mixtures, the performance of the SAFT-γ Mie approach is deemed to be very promising. However, there are certain aspects that need to be considered in future work to improve the performance of the method in the prediction of more complex systems. From the perspective of obtaining group interaction parameters, it will be useful to assess how the inclusion of selected experimental data for the solid-liquid equilibria of selected binary mixtures can improve the description of the overall phase behaviour with the SAFT-γ Mie platform. A similar approach is followed for the parameter estimation within the Dortmund-modified UNIFAC which has been shown to improve the predictions of the theory in the study of SLE. The inclusion of SLE data can be undertaken in a manner that does not limit the predictive capability of the approach presented in this chapter, using experimental data for "simple" mixtures to determine the parameters which can then be transferred. Another avenue that can be followed to improve the predictions of the SAFT-γ Mie approach, which relates more to the theoretical background of the method, is the explicit inclusion of proximity effects between functional groups at the level of the theory. Proximity effects describe how the chemical nature of a given group is altered by the presence of a different functional group,

and can be treated, e.g., within the Constantinou and Gani approach [38]. Proximity effects are expected to play an important role in the prediction of solubility of complex organic molecules such as APIs since compounds of this kind typically comprise multiple functional groups of a different chemical nature in close vicinity to one another.

Chapter 6

Conclusions

Group contribution (GC) methods provide an excellent predictive thermodynamic tool that can yield an accurate description of the thermodynamic properties of fluids and fluid mixtures. Methodologies of this kind can also find extensive application in the general area of fluid formulation and computer-aided molecular design (CAMD), where the goal is to determine the molecular structure that satisfies a set of performance criteria. From the wide variety of GC methods available, a comprehensive review of which is provided in chapter 2, the UNIFAC approach and its modifications [14, 85] stand out as the current state-of-the-art methodology for the property prediction of mixtures. The popularity of UNIFAC is due to its accuracy in the prediction of fluid phase equilibria, its ease of use, and the extensive parameter table available. However, the UNIFAC approach suffers from a key limitation: being an activity coefficient method, UNIFAC treats the liquid phase alone, and an additional thermodynamic approach is required for the vapour phase. This limitation calls for the development of GC approaches within the framework of an equation of state for the fluid, a strategy which started to be followed in the 1980s, within the framework of cubic equations of state, and more recently with the application of the GC concept within the statistical associating fluid theory (SAFT) [21–24]. The development of such an approach is the core goal of the work presented in this thesis.

The SAFT-γ Mie approach presented in chapter 4 is based on a fused heteronuclear off-lattice molecular model, where the interactions between the monomeric segments making up the molecules are modelled by means of the Mie intermolecular potential of variable attractive and repulsive range. Early work on the formulation of SAFT as a group contribution approach has given rise to the homonuclear approaches [18, 201, 206], where the fundamental theory remains unchanged and the GC concept is applied at the level of determining the molecular parameters of the theory in a predictive manner. A more de-

tailed formulation was subsequently presented [19], where the molecular model employed explicitly accounts for the presence of different functional groups. Such a heteronuclear molecular model is at the heart of the SAFT-γ equation of state [19, 29] which was originally based on the square-well intermolecular potential, and was later employed in the development of other SAFT-based GC approaches [20, 224]. In chapter 3, the SAFT-γ method is applied to the study of the fluid phase behaviour of a wide variety of mixtures, including aqueous solutions of alkanes and alkanols [348]. Aqueous solutions are very challenging systems to model due to the highly non-ideal phase behaviour that they exhibit, featuring heterogeneous azeotropes and large regions of immiscibility. At the same time, they are of high industrial relevance for broad ranges of applications from waste water treatment to the study of solubilities of complex compounds. It is shown that the SAFT-γ approach provides a very good predictive description of aqueous solutions. During the course of the application of the method to the study of aqueous solutions the delicate issue of group identification is also discussed. A key finding presented in Chapter 3 is that different modelling strategies for a given family can change the performance of the theory in the prediction of fluid phase behaviour. Despite the overall satisfactory performance of the SAFT-γ approach based on the square-well potential form in the description of a wide range of mixtures, several challenges are identified, mainly related to the description of thermodynamic derivative properties and fluid phase behaviour in the vicinity of the critical point. Regarding the derivative properties, the performance of the SAFT-γ method is related to the simplistic form of the intermolecular potential model that is at the core of the SAFT-VR SW and SAFT-γ SW approaches.

The aforementioned challenges call for the development of a new theory, based on a more realistic representation of the interactions between molecules. In order to address this, the theory developed in chapter 4 is based on the Mie intermolecular potential of variable attractive and repulsive range. The versatility of this potential was previously shown to allow for an accurate description of phase behaviour and derivative properties within the framework of the homonuclear SAFT-VR Mie [31, 263] approach. A key feature of the SAFT-γ Mie representation is the treatment of the monomer term, where a third-order perturbation expansion is employed [31]; the inclusion of the higher-order terms is shown to significantly improve the description of fluid phase behaviour in the vicinity of the critical point. A central aspect of the extension of the theory to account for a heteronuclear molecular model is a detailed assessment of the combining and mixing rules employed, which are of critical importance to the accuracy of the theory. The resulting SAFT-γ Mie equation of state is then applied to the study of the fluid phase behaviour and derivative thermodynamic properties of the chemical families of the *n*-alkanes and the 2-ketones,

incorporating Mie models for the constituent functional groups: methyl, methylene and carbonyl. The formalism is shown to accurately describe the fluid phase behaviour of the correlated compounds, where a significant improvement compared to other SAFT-type GC approaches is seen. The effect of the new treatment of the monomer perturbation term in the description of the near-critical region is highlighted. Furthermore, it is shown that the application of a more detailed intermolecular potential allows for a significant improvement in the representation of the thermodynamic second derivative properties (such as the isobaric heat capacity, the speed of sound, the isothermal compressibility, and the Joule-Thomson coefficient), through an extensive study for the properties of the n-alkanes in comparison to previous work based on the square-well potential. The predictive capability of the approach presented in this thesis is examined through the description of the pure component properties for compounds not included in the regression of the group parameters, and to the study of the properties of binary mixtures. The predictions of the theory are, in all cases studied, in very good agreement with the experimental data, even in the case of the excess thermodynamic properties, the accurate representation of which is a challenge for most thermodynamic models.

Group contribution methodologies offer a particularly attractive platform for the application in the prediction of solubility of complex organic molecules (and more specifically, pharmaceutical compounds) in solvents and solvent blends. The study of solubilities of complex molecules, including active pharmaceutical ingredients (APIs), is of great relevance to the manufacturing process of pharmaceutical formulations. Since typically limited experimental data are available in this context, thermodynamic tools such as GC approaches with predictive capabilities are usually required. The solubilities of two APIs, phenylacetic acid and ibuprofen, in two organic solvents, acetone and methyl-isobutyl ketone, are predicted by application of the SAFT-γ Mie approach and a standard thermodynamic framework for the calculation of solid-liquid equilibria. It is shown, based on experimental data for the vapour-liquid equilibria of simple systems, that the method can provide a reasonable predictive capability for the preliminary systems studied and can be used to rank the solvent solubilities in agreement with the experimental data. Given the complexity of the mixtures considered, comprising molecules with multiple functional groups in close vicinity, the performance of the SAFT-γ Mie approach is deemed highly satisfactory. This preliminary study is conducted in order to demonstrate the potential of the theory in describing solid-liquid equilibria in a predictive manner. A more detailed analysis of these systems and other compounds, and the potential inclusion of proximity effects at the level of the theory can help towards improving our SAFT-γ Mie predictive platform.

6.1 Summary of the key contributions of the current work

The contributions of the work presented in this thesis can be summarised as follows:

- Group identification within group contribution methods. The impact of different modelling approaches based on the physicochemical nature of the functional groups, related to polarisability effects, on the performance of the theory is presented for the chemical family of 1-alkanols within the SAFT-γ EoS;

- Extension of the SAFT-γ method based on the square-well potential to the study of aqueous solutions of hydrocarbons. In this case, the ability of the theory to successfully describe the highly non-ideal phase behaviour of the mixtures studied is demonstrated. It is shown that the SAFT-γ method provides with an accurate simultaneous description of the vapour-liquid and the liquid-liquid equilibria that aqueous mixtures exhibit;

- Generalisation of the SAFT-VR Mie EoS to the SAFT-γ Mie group contribution approach, based on a *fused* heteronuclear molecular model and the Mie intermolecular potential of variable attractive and repulsive ranges for the description of the segment-segment interactions. The extension of the theory required a thorough examination of the performance and validity of the mixing and combining rules that are employed in the development of the theory for heteronuclear molecules;

- Modelling the thermodynamic properties and fluid phase behaviour of the chemical families of the *n*-alkanes and the 2-ketones with the SAFT-γ Mie GC approach, by developing parameters for the CH_3, CH_2, and CH_3CO functional groups. It is shown how the use of the Mie intermolecular potential allows for the simultaneous description of phase behaviour and thermodynamic derivative properties;

- Demonstration of the predictive capabilities of SAFT-γ Mie in the accurate description of the fluid phase behaviour (vapour-liquid and liquid-liquid) and excess thermodynamic properties of a range of binary mixtures, including *n*-alkanes, branched alkanes, 2-ketones, alkylbenzenes and carboxylic acids, in a predictive manner;

- Application of the SAFT-γ Mie approach to the preliminary study of the solubilities of the APIs phenylacetic acid and ibuprofen in organic solvents acetone and methyl-isobutyl ketone in a fully predictive manner. For the description of the APIs the group parameters for 7 distinct functional groups and all unlike group interactions were developed, as summarised on table 6.1.

Table 6.1: Parameter matrix of the pure group and unlike group interactions developed in the work presented. The line demarcating the two parts of the matrix highlights the fact that the matrix is symmetric.

Functional Group	CH_3	CH_2	CH_3CO	$CHCH_3$	aCH	$aCCH_2$	COOH
CH_3	✓	✓	✓	✓	✓	✓	✓
CH_2	✓	✓	✓	✓	✓	✓	✓
CH_3CO	✓	✓	✓	✓	✓	✓	✓
$CHCH_3$	✓	✓	✓	✓	✓	✓	✓
aCH	✓	✓	✓	✓	✓	✓	✓
$aCCH_2$	✓	✓	✓	✓	✓	✓	✓
COOH	✓	✓	✓	✓	✓	✓	✓

6.2 Directions for future work

As for every group contribution method, the predictive power of the SAFT-γ Mie approach depends primarily on the extent of the group parameters available. In order to enhance the predictive power of the methodology presented, the current parameter table, featuring the functional groups to represent the families of n-alkanes, branched alkanes, alkylbenzenes, 2-ketones and carboxylic acids, has to be extended to include other functional groups that will allow one to model other chemical families. Such examples are the family of amines (NH_2) and alkanols (OH), for example in the modelling of alkanolamines in processes of carbon capture, esters (COO) for biodiesel processes, as well as the modelling of charged groups in a group contribution fashion. The introduction of key molecules modelled as single groups, such as water (H_2O), carbon dioxide (CO_2) and methane (CH_4), as well as the binary interaction parameters of these groups with the functional groups in the larger multifunctional compounds, based on mixture data, will also expand the range of applicability of the method. It should to be noted however, that for the modelling of associating compounds, the current implementation of the association term has to be revisited.

A long-term aim of the work presented in this thesis is an extensive application of the method to the modelling of the solubilities of APIs in solvents and solvent blends. Based on the results presented in chapter 5, where the performance of the SAFT-γ Mie approach is examined in the description of the solubility of ibuprofen and phenylacetic acid in two solvents, the methodology is seen to offer great potential in the study of the phase behaviour of such complex molecules in a predictive manner. However, there are some aspects that need to be considered, especially in cases where high fidelity quantitative agreement between the calculations of the theory and the experimental data is needed.

From the perspective of the estimation of the group interaction parameters, it would be of great interest to assess the importance of including solid-liquid equilibrium data in the regression of the group parameters. The extension of the types of experimental data used in the parameter estimation is a strategy also employed within the modified UNIFAC Dortmund, and is expected to improve the predictions of the method in the study of solid-liquid equilibria and solubility. At the same time, from a theoretical perspective, the inclusion of higher-order group contributions (at the second- and third-order level as discussed in chapter 2), that account for proximity effects, is of great interest. Proximity effects, that describe how the chemical nature of a group is altered when a chemically different group is found in close vicinity, are expected to play an important role in the description of highly complex molecules that feature multiple distinct functional groups. Furthermore, higher-order effects can lead to a general improvement in the description of fluids and fluid mixtures and allow for the distinction between isomers, a common shortcoming of most GC approaches.

Having established a large database of group parameters and extended the application of the SAFT-γ Mie framework to the study of the solubility of several APIs in solvents, the method presented can serve as a platform for solvent screening. The aim in this case would be to construct a framework based on the SAFT-γ Mie method, that can optimise the components of a solvent blend and/or the composition of the blend based on a desired solubility of a given API.

Bibliography

[1] R. Dohrn and O. Pfohl. Thermophysical properties - Industrial directions. *Fluid Phase Equilib.* **15**, 194 (2002).

[2] W. B. Whiting. Effects of uncertainties in thermodynamic data and models on process calculations. *J. Chem. Eng. Data* **41**, 935 (1996).

[3] S. Gupta and J. D. Olson. Industrial needs in physical properties. *Ind. Eng. Chem. Res.* **42**, 6359 (2003).

[4] A. Bardow, K. Steur, and J. Gross. Continuous-molecular targeting for integrated solvent and process design. *Ind. Eng. Chem. Res.* **49**, 2834 (2010).

[5] F. E. Pereira, E. Keskes, A. Galindo, G. Jackson, and C. S. Adjiman. Integrated solvent and process design using a SAFT-VR thermodynamic description: High-pressure separation of carbon dioxide and methane. *Comput. Chem. Eng.* **35**, 474 (2011).

[6] J. Levelt-Sengers. *How fluids unmix. Discoveries by the school of van der Waals and Kamerlingh Onnes*. Royal Netherlands Academy of Arts and Science, The Netherlands, (2002).

[7] R. Gani and E. Brignole. Molecular design of solvents for liquid extraction based on UNIFAC. *Fluid Phase Equilib.* **13**, 331 (1983).

[8] O. Odele and S. Macchietto. Computer Aided Molecular Design: a novel method for optimal solvent selection. *Fluid Phase Equilib.* **82**, 47 (1993).

[9] S. Macchietto, O. Odele, and O. Omatsone. Design of optimal solvents for liquid-liquid extraction and gas absorption processes. *Tran. Inst. of Chem. Eng.* **68**, 429 (1990).

[10] P. M. Harper, R. Gani, P. Kolar, and T. Ishikawa. Computer-aided molecular design with combined molecular modeling and group contribution. *Fluid Phase Equilib.* **158-160**, 337 (1999).

[11] K. V. Gernaey and R. Gani. A model-based systems approach to pharmaceutical product-process design and analysis. *Chem. Eng. Sci.* **65**, 5757 (2010).

[12] C-C. Chen and P. A. Crafts. Correlation and prediction of drug molecule solubility in mixed solvent systems with the nonrandom two-liquid segment activity coefficient (NRTL-SAC) model. *Ind. Eng. Chem. Res.* **45**, 4816 (2006).

[13] R. P. Currier and J. P. O'Connell. An analysis of the solution of groups method for component activity coefficients. *Fluid Phase Equilib.* **33**, 245 (1987).

[14] A. Fredenslund, R. L. Jones, and J. M. Prausnitz. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AIChE J.* **21**, 1086 (1975).

[15] E. A. Guggenheim. *Mixtures.* Clarendon Press, Oxford, (1952).

[16] G. M. Wilson. Vapor-liquid equilibrium XI: A new expression for the excess free energy of mixing. *J. Am. Chem. Soc.* **86**, 127 (1964).

[17] The UNIFAC consortium. http://unifac.ddbst.de/, (2008).

[18] S. Tamouza, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Group contribution method with SAFT EoS applied to vapor liquid equilibria of various hydrocarbon series. *Fluid Phase Equilib.* **222-223**, 67 (2004).

[19] A. Lymperiadis, C. S. Adjiman, A. Galindo, and G. Jackson. A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT - γ). *J. Chem. Phys.* **127**, 234903 (2007).

[20] Y. Peng, K. D. Goff, M. C. dos Ramos, and C. McCabe. Developing a predictive group-contribution-based SAFT-VR equation of state. *Fluid Phase Equilib.* **277**, 131 (2009).

[21] G. Jackson, W. G. Chapman, and K. E. Gubbins. Phase equilibria of associating fluids. Spherical molecules with multiple bonding sites. *Mol. Phys.* **65**, 1 (1988).

[22] W. G. Chapman, G. Jackson, and K. E. Gubbins. Phase equilibria of associating fluids. Chain molecules with multiple bonding sites. *Mol. Phys.* **65**, 1057 (1988).

[23] W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz. SAFT: Equation of state solution model for associating fluids. *Fluid Phase Equilib.* **52**, 31 (1989).

[24] W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.* **29**, 1709 (1990).

[25] E. A. Müller and K. E. Gubbins. Molecular-based equations of state for associating fluids: A review of SAFT and related approaches. *Ind. Eng. Chem. Res.* **40**, 2193 (2001).

[26] I. G. Economou. Statistical Associating Fluid Theory: A successful model for the calculation of thermodynamic and phase equilibrium properties of complex fluid mixtures. *Ind. Eng. Chem. Res.* **41**, 953 (2002).

[27] P. Paricaud, A. Galindo, and G. Jackson. Recent advances in the use of the SAFT approach in describing electrolytes, interfaces, liquid crystals and polymers. *Fluid Phase Equilib.* **194**, 87 (2002).

[28] C. McCabe and A. Galindo. Chapter 8: SAFT Associating fluids and fluid mixtures. In A. R. H. Goodwin, J. V. Sengers, and C. J. Peters, editors, *Applied Thermodynamics of Fluids*, pages 215–279. Royal Society of Chemistry, UK, (2010).

[29] A. Lymperiadis, C. S. Adjiman, G. Jackson, and A. Galindo. A generalisation of the SAFT - γ group contribution method for groups comprising multiple spherical segments. *Fluid Phase Equilib.* **274**, 85 (2008).

[30] T. Lafitte, D. Bessieres, M. M. Pineiro, and J. L. Daridon. Simultaneous estimation of phase behavior and second-derivative properties using the Statistical Associating Fluid Theory with variable range approach. *J. Chem. Phys.* **124**, 024509 (2006).

[31] T. Lafitte, A. Apostolakou, C. Avendaño, A. Galindo, C. S. Adjiman, E. A. Müller, and G. Jackson. Accurate perturbation theory for chains of Mie soft-core segments (SAFT-VR Mie) for the description of vapour-liquid equilibria and derivative properties (2012). In preparation.

[32] A. L. Lydersen. Estimation of critical properties of organic compounds. *University of Wisconsin College Engineering, Eng. Exp. Stn. Rep. 3, Madison, WI* (1955).

[33] K. M. Klincewicz and R. C. Reid. Estimation of critical properties with group contribution methods. *AIChE J.* **30**, 137 (1984).

[34] C. F. Spencer and T. E. Daubert. A critical evaluation of methods for the prediction of critical properties of hydrocarbons. *AIChE J.* **19**, 482 (1973).

[35] D. Ambrose. Correlation and estimation of vapour-liquid critical properties. I. Critical temperatures of organic compounds. *National Physical Laboratory, Teddington, NPL Rep. Chem. 92* (1978).

[36] D. Ambrose. Correlation and estimation of vapour-liquid critical properties. II. Critical pressures and volumes of organic compounds. *National Physical Laboratory, Teddington, NPL Rep. Chem. 98* (1979).

[37] K. J. Joback and R. C. Reid. Estimation of pure-component properties from group-contributions. *Chem. Eng. Commun.* **57**, 233 (1987).

[38] L. Constantinou and R. Gani. New group contribution method for estimating properties of pure compounds. *AIChE J.* **40**, 1697 (1994).

[39] B. E. Poling, J. M. Prausnitz, and J. P. O'Connell. *The properties of gases and liquids*. Mc Graw Hill, 5th edition, (2007).

[40] S. Devotta and V. Rao Pendyala. Modified Joback group contribution method for normal boiling point of aliphatic halogenated compounds. *Ind. Eng. Chem. Res.* **31**, 2042 (1992).

[41] F. S. Emami, A. Vahid, and J. R. Jr. Elliott. Finitely limited group contribution correlations for boiling temperatures. *J. Chem. Thermodyn.* **41**, 530 (2009).

[42] J. W. Jalowka and T. E. Daubert. Group contribution method to predict critical temperature and pressure of hydrocarbons. *Ind. Eng. Chem. Process. Des. Dev.* **25**, 139 (1986).

[43] S. W. Benson and J. H. Buss. Additivity rules for the estimation of molecular properties. Thermodynamic properties. *J. Chem. Phys.* **29**, 546 (1958).

[44] T. E. Daubert and R. Bartakovits. Prediction of critical temperature and pressure of organic compounds by group contribution. *Ind. Eng. Chem. Res.* **28**, 638 (1989).

[45] S. E. Stein and R. L. Brown. Estimation of normal boiling points from group contributions. *J. Chem. Inform. Comput. Sci.* **34**, 581 (1994).

[46] R. F. Fedors. A relationship between chemical structure and the critical temperature. *Chem. Eng. Commun.* **16**, 149 (1982).

[47] M. L. Mavrovouniotis. Estimation of properties from conjugate forms of molecular structures: The ABC approach. *Ind. Eng. Chem. Res.* **29**, 1943 (1990).

[48] L. Constantinou, S. E. Prickett, and M. L. Mavrovouniotis. Estimation of thermodynamic and physical properties of acyclic hydrocarbons using the ABC approach and conjugation operators. *Ind. Eng. Chem. Res.* **32**, 1734 (1993).

[49] L. Constantinou, S. E. Prickett, and M. L. Mavrovouniotis. Estimation of properties of acyclic organic compounds using conjugation operators. *Ind. Eng. Chem. Res.* **33**, 395 (1994).

[50] L. Constantinou, R. Gani, and J. P. O'Connell. Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution method. *Fluid Phase Equilib.* **103**, 11 (1995).

[51] J. Marrero and R. Gani. Group-contribution based estimation of pure component properties. *Fluid Phase Equilib.* **183-184**, 183 (2001).

[52] S. E. Prickett, L. Constantinou, and M. L. Mavrovouniotis. Computational identification of conjugate paths for estimation of properties of organic compounds. *Mol. Simul.* **11**, 205 (1993).

[53] J. Marrero-Morejón and E. Pardillo-Fontdevila. Estimation of pure compound properties using group-interaction contributions. *AIChE J.* **45**, 615 (1999).

[54] E. Pardillo-Fontdevila and R. González-Rubio. A Group-Interaction contribution approach. A new strategy for the estimation of physico-chemical properties of branched isomers. *Chem. Eng. Commun.* **163**, 245 (1997).

[55] S. Grigoras. A structural approach to calculate physical properties of pure organic substances: The critical temperature, critical volume and related properties. *J. Comput. Chem.* **11**, 493 (1990).

[56] L. B. Kier and H. L. Hall. *Molecular connectivity in structure activity analysis*. John Wiley & Sons, New York, (1986).

[57] R. Gani, P. M. Harper, and M. Hostrup. Automatic creation of missing groups through connectivity index for pure-component property prediction. *Ind. Eng. Chem. Res.* **44**, 7262 (2005).

[58] S. I. Sandler. *Chemical and Engineering Thermodynamics*. Wiley, New York, 3rd. edition, (1997).

[59] H. V. Kehiaian. Group contribution methods for liquid mixtures: A critical review. *Fluid Phase Equilib.* **13**, 243 (1983).

[60] A. Fredenslund and P. Rasmussen. From UNIFAC to SUPERFAC - and back? *Fluid Phase Equilib.* **24**, 115 (1985).

[61] I. Nagata and J. Koyabu. Phase equilibria by effective UNIFAC group-contribution method. *Thermochim. Acta* **48**, 187 (1981).

[62] F. A. Ashraf and J. H. Vera. A simplified group method analysis. *Fluid Phase Equilib.* **4**, 211 (1980).

[63] J. Gmehling. Group contribution methods for the estimation of activity coefficients. *Fluid Phase Equilib.* **30**, 119 (1986).

[64] M. N. Papadopoulos and E. L. Derr. Group interaction. II. A test of the group model on binary solutions of hydrocarbons. *J. Am. Chem. Soc.* **81**, 2285 (1959).

[65] O. Redlich, E. L. Derr, and G. J. Pierotti. Group interaction. I. A model for interaction in solutions. *J. Am. Chem. Soc.* **81**, 2283 (1959).

[66] G. M. Wilson and C. H. Deal. Activity coefficients and molecular structure. Activity coefficients in changing environments - Solutions of Groups. *Ind. Eng. Chem. Fund.* **1**, 20 (1962).

[67] E. L. Derr and C. H. Deal. Analytical solutions of groups: correlation of activity coefficients through structural group parameters. *Int. Chem. Eng. Symp. Ser. No. 32* **3**, 44 (1969).

[68] P. J. Flory. Thermodynamics of polymer solutions. *Discuss. Faraday Soc.* **49**, 7 (1970).

[69] K. Kojima and K. Tochigi. *Prediction of vapor-liquid equilibria by the* ASOG *method, Physical sciences data 3.* Elsevier scientific publishing company, (1979).

[70] A. Correa, J. Tojo, J. M. Correa, and A. Blanco. New Analytical Solution of Groups method parameters for the prediction of vapor-liquid equilibrium. *Ind. Eng. Chem. Res.* **28**, 609 (1989).

[71] K. Tochigi, D. Tiegs, J. Gmehling, and K. Kojima. Determination of new ASOG parameters. *J. Chem. Eng. Jpn.* **23**, 453 (1990).

[72] J. Gmehling, J. Li, and M. Schiller. A modified UNIFAC model. 2. Present parameter matrix and results for different thermodynamic properties. *Ind. Eng. Chem. Res.* **32**, 178 (1993).

[73] A. Fredenslund. UNIFAC and related group-contribution models for phase equilibria. *Fluid Phase Equilib.* **52**, 135 (1989).

[74] A. Fredenslund and J. M. Sørensen. *Models for thermodynamic and phase equilibria caclulations*, chapter Group Contribution Estimation Methods. Marcel Dekker Inc., New York, (1994).

[75] D. S. Abrams and J. M. Prausnitz. Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE J.* **21**, 116 (1975).

[76] A. Bondi. *Physical properties of molecular crystals, liquids and glasses*. Wiley, New York, (1968).

[77] A. Fredenslund, J. Gmehling, and P. Rasmussen. *Vapor-Liquid Equilibria using* UNIFAC*: A G*roup-C*ontribution method*. Elsevier, Amsterdam, (1977).

[78] T. Magnussen, P. Rasmussen, and A. Fredenslund. UNIFAC parameter table for prediction of liquid-liquid equilibria. *Ind. Eng. Chem. Proc. Des. Dev.* **20**, 331 (1981).

[79] J. Gmehling. Present status and potential of group contribution methods for process development. *J. Chem. Thermodyn.* **41**, 731 (2009).

[80] J. Gmehling, P. Rasmussen, and A. Fredenslund. Vapor-liquid equilibria by UNIFAC group contribution. Revision and extension. 2. *Ind. Eng. Chem. Proc. Des. Dev.* **21**, 118 (1982).

[81] E. A. Macedo, U. Weidlich, J. Gmehling, and P. Rasmussen. Vapor-liquid equilibria by UNIFAC group contribution. Revision and extension. 3. *Ind. Eng. Chem. Proc. Des. Dev.* **22**, 676 (1983).

[82] D. Tiegs, J. Gmehling, P. Rasmussen, and A. Fredenslund. Vapor-liquid equilibria by UNIFAC group contribution. 4. Revision and extension. *Ind. Eng. Chem. Res.* **26**, 159 (1987).

[83] H. K. Hansen, P. Rasmussen, A. Fredenslund, M. Schiller, and J. Gmehling. Vapor-liquid equilibria by UNIFAC group contribution. 5. Revision and extension. *Ind. Eng. Chem. Res.* **30**, 2352 (1991).

[84] S. Skjold-Jørgensen, B. Kolbe, J. Gmehling, and P. Rasmussen. Vapor-liquid equilibria by UNIFAC group contribution. Revision and extension. *Ind. Eng. Chem. Proc. Des. Dev.* **18**, 714 (1979).

[85] U. Weidlich and J. Gmehling. A modified UNIFAC model. 1. Prediction of VLE, h^E and γ^∞. *Ind. Eng. Chem. Res.* **26**, 1372 (1987).

[86] B. L. Larsen, P. Rasmussen, and A. Fredenslund. A modified UNIFAC group-contribution model for prediction of phase equilibria and heats of mixing. *Ind. Eng. Chem. Res.* **26**, 2274 (1987).

[87] J. Gmehling, D. Tiegs, and U. Knipp. A comparison of the predictive capability of different group contribution methods. *Fluid Phase Equilib.* **54**, 147 (1990).

[88] J. Gmehling, J. Lohmann, A. Jakob, J. Li, and R. Joh. A modified UNIFAC (Dortmund) model. 3. Revision and extension. *Ind. Eng. Chem. Res.* **37**, 4876 (1998).

[89] J. Lohmann, R. Joh, B. Nienhaus, and J. Gmehling. Revision and extension of the group contribution method modified UNIFAC (Dortmund). *Chem. Eng. Technol.* **21**, 245 (1998).

[90] R. Wittig, J. Lohmann, R. Joh, S. Horstmann, and J. Gmehling. Vapor-liquid equilibria and enthalpies of mixing in a temperature range from 298.15 to 413.15 K for the further development of Modified UNIFAC (Dortmund). *Ind. Eng. Chem. Res.* **40**, 5831 (2001).

[91] J. Lohmann, R. Joh, and J. Gmehling. From UNIFAC to Modified UNIFAC (Dortmund). *Ind. Eng. Chem. Res.* **40**, 957 (2001).

[92] J. Gmehling, R. Wittig, J. Lohmann, and R. Joh. A modified UNIFAC (Dortmund) model. 4. Revision and extension. *Ind. Eng. Chem. Res.* **41**, 1678 (2002).

[93] A. Jakob, H. Grensemann, J. Lohmann, and J. Gmehling. Further Development of Modified UNIFAC (Dortmund): Revision and Extension 5. *Ind. Eng. Chem. Res.* **45**, 7924 (2006).

[94] T. Oishi and J. M. Prausnitz. Estimation of solvent activities in polymer solutions using a group-contribution method. *Ind. Eng. Chem. Proc. Des. Dev.* **17**, 333 (1978).

[95] J. Holten-Andersen, P. Rasmussen, and A. Fredenslund. Phase equilibria of polymer solutions by group contribution. 1. Vapor-liquid equilibria. *Ind. Eng. Chem. Res.* **26**, 1382 (1987).

[96] H. S. Elbro, A. Fredenslund, and P. Rasmussen. A new simple equation for the prediction of solvent activities in polymer solutions. *Macromolecules* **23**, 4707 (1990).

[97] G. M. Kontogeorgis, A. Fredenslund, and D. Tassios. Simple activity coefficient model for the prediction of solvent activities in polymer solutions. *Ind. Eng. Chem. Res.* **32**, 362 (1993).

[98] I. Kikic, M. Fermeglia, and P. Rasmussen. UNIFAC Prediction of vapor-liquid equilibria in mixed solvent-salt systems. *Chem. Eng. Sci.* **46**, 2775 (1991).

[99] C. Achard, C. G. Dussap, and J. B. Gros. Representation of vapour-liquid equilibria in water-alcohol-electrolyte mixtures with a modified UNIFAC group-contribution method. *Fluid Phase Equilib.* **98**, 71 (1994).

[100] A. C. Mengarelli, E. A. Brignole, and S. B. Bottini. Activity coefficients of associating mixtures by group contribution. *Fluid Phase Equilib.* **163**, 195 (1999).

[101] M. S. Wertheim. Fluids with highly directional attractive forces: I. Statistical thermodynamics. *J. Stat. Phys.* **35**, 19 (1984).

[102] M. S. Wertheim. Fluids with highly directional attractive forces: II. Thermodynamic perturbation theory and integral equations. *J. Stat. Phys.* **35**, 35 (1984).

[103] M. S. Wertheim. Fluids with highly directional attractive forces: III. Multiple attraction sites. *J. Stat. Phys.* **42**, 459 (1986).

[104] M. S. Wertheim. Fluids with highly directional attractive forces: IV. Equilibrium polymerization. *J. Stat. Phys.* **42**, 477 (1986).

[105] J. Abildskov, L. Constantinou, and R. Gani. Towards the development of a second-order approximation in activity coefficient models based on group contributions. *Fluid Phase Equilib.* **118**, 1 (1996).

[106] J. Abildskov, R. Gani, P. Rasmussen, and J. P. O'Connell. Beyond basic UNIFAC. *Fluid Phase Equilib.* **158-160**, 349 (1999).

[107] H. S. Wu and S. I. Sandler. Use of ab initio quantum mechanics calculations in group contribution methods. 1. Theory and the basis for group indentifications. *Ind. Eng. Chem. Res.* **30**, 881 (1991).

[108] H. S. Wu and S. I. Sandler. Use of ab initio quantum mechanics calculations in group contribution methods. 2. Test of new groups in UNIFAC. *Ind. Eng. Chem. Res.* **30**, 889 (1991).

[109] H. E. González, J. Abildskov, R. Gani, P. Rousseaux, and B. L. Bert. A method for prediction of UNIFAC group interaction parameters. *AIChE J.* **53**, 1620 (2007).

[110] J. D. van der Waals. *Over de Continuiteit van den Gas- en Vloeistoftoestand*. PhD thesis, Leyden, (1873). Translated and edited by J. S. Rowlinson in *On the Continuity of the Gaseous and Liquid States*, Dover Phoenix Editions, (2004).

[111] J. S. Rowlinson and F. L. Swinton. *Liquids and liquid mixtures.* Butterworths, 3^{rd} edition, (1982).

[112] A. J. Haslam, A. Galindo, and G. Jackson. Prediction of binary intermolecular potential parameters for use in modelling fluid mixtures. *Fluid Phase Equilib.* **266**, 105 (2008).

[113] M. J. Huron and J. Vidal. New mixing rules in simple equations of state for representing vapour-liquid equilibria of strongly non-ideal mixtures. *Fluid Phase Equilib.* **3**, 255 (1979).

[114] G. Soave. Equilibrium constants from a modified Redlich-Kwong equation of state. *Chem. Eng. Sci.* **27**, 1197 (1972).

[115] O. Redlich and A. T. Kister. On the thermodynamics of non-electrolyte solutions and its technical applications III. Systems with associated components. *J. Chem. Phys.* **15**, 849 (1947).

[116] H. Renon and J. M. Prausnitz. Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE J.* **14**, 135 (1968).

[117] S. Dahl and M. L. Michelsen. High-Pressure vapor-liquid equilibrium with a UNIFAC-based equation of state. *AIChE J.* **36**, 1829 (1990).

[118] M. L. Michelsen. A modified Huron-Vidal mixing rule for cubic equations of state. *Fluid Phase Equilib.* **60**, 213 (1990).

[119] D. S. H. Wong and S. I. Sandler. A theoretically correct mixing rule for cubic equations of state. *AIChE J.* **38**, 671 (1992).

[120] A. Anderko. Equation-of-State methods for the modelling of phase equilibria. *Fluid Phase Equilib.* **61**, 145 (1990).

[121] S. I. Sandler, H. Orbey, and B. Lee. *Models for thermodynamic and phase equilibria calculations*, chapter Equations of State. Marcel Dekker Inc., (1994).

[122] S. I. Sandler and H. Orbey. Chapter 9: Mixing and Combining Rules. In J. V. Sengers, R. F. Kayser, C. J. Peters, and H. J. Jr. White, editors, *Equations of state for fluids and fluid mixtures, part 1*, pages 321–359. Elsevier, Amsterdam, (2000).

[123] T. Holderbaum and J. Gmehling. PSRK: A group contribution equation of state based on UNIFAC. *Fluid Phase Equilib.* **70**, 251 (1991).

[124] G. M. Kontogeorgis and P. M. Vlamos. An interpretation of the behaviour of EoS/g^E models for asymmetric systems. *Chem. Eng. Sci.* **55**, 2351 (2000).

[125] J. Ahlers and J. Gmehling. Development of a universal group contribution equation of state. I. Prediction of liquid densities for pure compounds with a volume translated Peng-Robinson equation of state. *Fluid Phase Equilib.* **191**, 177 (2001).

[126] J. Ahlers and J. Gmehling. Development of a universal group contribution equation of state. II. Prediction of vapor-liquid equilibria for asymmetric systems. *Ind. Eng. Chem. Res.* **41**, 3489 (2002).

[127] J. Ahlers and J. Gmehling. Development of a universal group contribution equation of state. III. Prediction of vapor-liquid equilibria, excess enthalpies, and activity coefficients at infinite dilution with the VTPR model. *Ind. Eng. Chem. Res.* **41**, 5890 (2002).

[128] J. N. Jaubert and F. Mutelet. VLE predictions with the Peng-Robinson equation of state and temperature dependent k_{ij} calculated through a group contribution method. *Fluid Phase Equilib.* **224**, 285 (2004).

[129] D. Y. Peng and D. B. Robinson. A new two-constant equation of state. *Ind. Eng. Chem. Fund.* **15**, 59 (1976).

[130] E. Rauzy and A. Peneloux. Vapor-liquid equilibrium and volumetric properties calculations for solutions in the supercritical carbon dioxide. *Int. J. Thermophys.* **7**, 635 (1986).

[131] A. Peneloux, W. Abdoul, and E. Rauzy. Excess functions and equations of state. *Fluid Phase Equilib.* **47**, 115 (1989).

[132] W. Abdoul, E. Rauzy, and A. Peneloux. Group-contribution equation of state for correlating and predicting thermodynamic properties of weakly polar and non-associating mixtures. Binary and multicomponent mixtures. *Fluid Phase Equilib.* **68**, 47 (1991).

[133] S. Vitu, R. Privat, S. N. Jaubert, and F. Mutelet. Predicting the phase equilibria of CO_2 + hydrocarbon systems with the PPR78 model (PR EoS and k_{ij} calculated through a group contribution method). *J. Supercrit. Fluids* **45**, 1 (2008).

[134] R. Privat, J. N. Jaubert, and F. Mutelet. Use of the PPR78 Model to predict new equilibrium data of binary systems involving hydrocarbons and nitrogen. Comparison with other GC-EoS. *Ind. Eng. Chem. Res.* **47**, 7483 (2008).

[135] R. Privat, F. Mutelet, and J. N. Jaubert. Addition of the Hydrogen Sulfide group to the PPR78 model (Predictive 1978, Peng-Robinson Equation of State with Temperature Dependent k_{ij} calculated through a group contribution method). *Ind. Eng. Chem. Res.* **47**, 10041 (2008).

[136] P. A. Gupte, P. Rasmussen, and A. Fredenslund. A new group-contribution equation of state for vapor-liquid equilibria. *Ind. Eng. Chem. Fund.* **25**, 636 (1986).

[137] S. Dahl, A. Fredenslund, and P. Rasmussen. The MHV2 Model: A UNIFAC-based equation of state model for prediction of gas solubility and vapor-liquid equilibria at low and high pressures. *Ind. Eng. Chem. Res.* **30**, 1936 (1991).

[138] C. Boukouvalas, N. Spiliotis, P. Coutsikos, N. Tzouvaras, and D. Tassios. Prediction of vapor-liquid equilibrium with the LCVM model: a linear combination of the Vidal and Michelsen mixing rules coupled with the original UNIFAC and the t-mPR equation of state. *Fluid Phase Equilib.* **92**, 75 (1994).

[139] C. Lermite and J. Vidal. A group contribution equation of state for polar and non-polar compounds. *Fluid Phase Equilib.* **72**, 111 (1992).

[140] H. Orbey, S. I. Sandler, and D. S. H. Wong. Accurate equation of state predictions at high temperatures and pressures using the existing UNIFAC model. *Fluid Phase Equilib.* **85**, 41 (1993).

[141] G. K. Georgeton and A. S. Teja. A group contribution equation of state based on the **S**implified Perturbed Hard Chain Theory. *Ind. Eng. Chem. Res.* **27**, 657 (1988).

[142] M. S. High and R. P. Danner. A group contribution equation of state for polymer solutions. *Fluid Phase Equilib.* **53**, 323 (1989).

[143] C. Panayiotou and J. H. Vera. An improved lattice-fluid equation of state for pure component polymeric fluids. *Polym. Eng. Sci.* **22**, 345 (1982).

[144] J. D. Pults, R. A. Greenkorn, and K. C. Chao. Chain-of-rotators group contribution equation of state. *Chem. Eng. Sci.* **44**, 2553 (1989).

[145] L. Coniglio, E. Rauzy, and C. Berro. Representation and prediction of thermophysical properties of heavy hydrocarbons. *Fluid Phase Equilib.* **87**, 53 (1993).

[146] L. Coniglio, L. Trassy, and E. Rauzy. Estimation of thermophysical properties of heavy hydrocarbons through a group contribution based equation of state. *Ind. Eng. Chem. Res.* **39**, 5037 (2000).

[147] S. Mattedi, F. W. Tavares, and M. Castier. Group contribution equation of state based on the lattice fluid theory: Alkane - alkanol system. *Fluid Phase Equilib.* **142**, 33 (1998).

[148] N. Elvassore, A. Bertucco, and M. Fermeglia. Phase-equilibria calculation by group-contribution Perturbed-Hard-Sphere-Chain equation of state. *AIChE J.* **48**, 359 (2002).

[149] J. R. Jr. Elliott and R. N. Natarajan. Extension of the Elliott-Suresh-Donohue equation of state to polymer solutions. *Ind. Eng. Chem. Res.* **41**, 1043 (2002).

[150] S. Skjold-Jørgensen. Gas solubility calculations. II. Application of a new group-contribution equation of state. *Fluid Phase Equilib.* **16**, 317 (1984).

[151] N. F. Carnahan and K. E. Starling. Equation of state for nonattracting rigid spheres. *J. Chem. Phys.* **51**, 635 (1969).

[152] S. Skjold-Jørgensen. Group contribution equation of state (GC-EoS): A predictive method for phase equilibrium computations over wide ranges of temperature and pressures up to 30 MPa. *Ind. Eng. Chem. Res.* **27**, 110 (1988).

[153] S. Espinosa, G. M. Foco, A. Bermudez, and T. Fornari. Revision and extension of the group contribution equation of state to new solvent groups and higher molecular weight alkanes. *Fluid Phase Equilib.* **172**, 129 (2000).

[154] S. Espinosa, T. Fornari, S. B. Bottini, and E. A. Brignole. Phase equilibria in mixtures of fatty oils and derivatives with near critical fluids using the GC-EoS model. *J. Supercrit. Fluids* **23**, 91 (2002).

[155] B. Breuce, S. B. Bottini, G.-J. Witkamp, and C. J. Peters. Thermodynamic modeling of the phase behavior of binary systems of ionic liquids and carbon dioxide with the group contribution equation of state. *J. Phys. Chem. B.* **111**, 14265 (2007).

[156] H. P. Gros, S. Bottini, and E. A. Brignole. A group contribution equation of state for associating mixtures. *Fluid Phase Equilib.* **116**, 537 (1996).

[157] F. A. Sánchez, S. Pereda, and E. A. Brignole. GCA-EoS: A SAFT group contribution model - Extension to mixtures containing aromatic hydrocarbons and associating compounds. *Fluid Phase Equilib.* **306**, 112 (2011).

[158] G. M. Kontogeorgis, E. Voutsas, I. V. Yakoumis, and D. P. Tassios. An equation of state for associating fluids. *Ind. Eng. Chem. Res.* **35**, 4310 (1996).

[159] G. K. Folas, , S. Derawi, M. L. Michelsen, E. H. Stenby, and G. M. Kontogeorgis. Recent applications of the cubic-plus-association (CPA) equation of state to industrially important systems. *Fluid Phase Equilib.* **228-229**, 121 (2005).

[160] G. M. Kontogeorgis, M. L. Michelsen, G. K. Folas, S. Derawi, N. von Solms, and E. H. Stenby. Ten years with the CPA (Cubic-Plus-Association) equation of state. Part 1. Pure compounds and self-associating systems. *Ind. Eng. Chem. Res.* **45**, 4855 (2006).

[161] G. M. Kontogeorgis, M. L. Michelsen, G. K. Folas, S. Derawi, N. von Solms, and E. H. Stenby. Ten years with the CPA (Cubic-Plus-Association) equation of state. Part 2. Cross-associating and multicomponent systems. *Ind. Eng. Chem. Res.* **45**, 4869 (2006).

[162] G. M. Kontogeorgis and G. K. Folas. *Thermodynamic models for industrial applications. From classical and advanced mixing rules to association theories.* Wiley, UK, (2010).

[163] E. A. Müller and K. E. Gubbins. Chapter 12: Associating fluids and fluid mixtures. In J. V. Sengers, R. F. Kayser, C. J. Peters, and H. J. Jr. White, editors, *Equations of state for fluids and fluid mixtures, part 2*, pages 435–479. Elsevier, Amsterdam, (2000).

[164] S. P. Tan, H. Adidharma, and M. Radosz. Recent advances and applications of Statistical Associating Fluid Theory. *Ind. Eng. Chem. Res.* **47**, 8063 (2008).

[165] J. Gross and G. Sadowski. Application of perturbation theory to a hard-chain reference fluid: an equation of state for square-well chains. *Fluid Phase Equilib.* **168**, 183 (2000).

[166] J. Gross and G. Sadowski. Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules. *Ind. Eng. Chem. Res.* **40**, 1244 (2001).

[167] S. H. Huang and M. Radosz. Equation of state for small, large, polydisperse, and associating molecules. *Ind. Eng. Chem. Res.* **29**, 2284 (1990).

[168] J. K. Johnson, E. A. Müller, and K. E. Gubbins. Equation of state for Lennard-Jones chains. *J. Phys. Chem.* **98**, 6413 (1994).

[169] F. J. Blas and L. F. Vega. Thermodynamic behaviour of homonuclear and heteronuclear Lennard-Jones chains with association sites from simulation and theory. *Mol. Phys.* **92**, 135 (1997).

[170] A. Gil-Villegas, A. Galindo, P. J. Whitehead, S. J. Mills, G. Jackson, and A. N. Burgess. Statistical Associating Fluid Theory for chain molecules with attractive potentials of variable range. *J. Chem. Phys.* **106**, 4168 (1997).

[171] A. Galindo, L. A. Davies, A. Gil-Villegas, and G. Jackson. The thermodynamics of mixtures and the corresponding mixing rules in the SAFT-VR approach for potentials of variable range. *Mol. Phys.* **93**, 241 (1998).

[172] P. K. Jog and W. G. Chapman. Application of Wertheim's thermodynamic perturbation theory to dipolar hard sphere chains. *Mol. Phys.* **97**, 307 (1999).

[173] P. K. Jog, S. G. Sauer, J. Blaesing, and W. G. Chapman. Application of dipolar chain theory to the phase behavior of polar fluids and mixtures. *Ind. Eng. Chem. Res.* **40**, 4641 (2001).

[174] J. Gross and J. Vrabec. An equation of state contribution for polar components: Dipolar molecules. *AIChE J.* **52**, 1194 (2005).

[175] J. Gross. An equation of state contribution for polar components: Quadrupolar molecules. *AIChE J.* **51**, 2556 (2005).

[176] C. Vega, L. MacDowell, and A. López-Rodríguez. Excess properties of mixtures of n-alkanes from perturbation theory. *J. Chem. Phys.* **111**, 3192 (1999).

[177] M. C. dos Ramos and F. J. Blas. Examination of the excess thermodynamic properties of n-alkane binary mixtures: a molecular approach. *J. Phys. Chem. B* **109**, 12145 (2005).

[178] A. Galindo, A. Gil-Villegas, G. Jackson, and A. N. Burgess. SAFT-VRE: Phase behavior of electrolyte solutions with the statistical associating fluid theory for potentials of variable range. *J. Phys. Chem. B.* **103**, 10272 (1999).

[179] A. Gil-Villegas, A. Galindo, and G. Jackson. A statistical associating fluid theory for electrolyte solutions. *Mol. Phys.* **99**, 531 (2001).

[180] L. F. Cameretti, G. Sadowski, and J. M. Mollerup. Modeling of aqueous electrolyte solutions with perturbed-chain statistical associated fluid theory. *Ind. Eng. Chem. Res.* **44**, 3355 (2005).

[181] R. P. Sear and G. Jackson. Theory and computer simulation of hard-sphere site models of ring molecules. *Mol. Phys.* **81**, 801 (1994).

[182] M. Lora, F. Rindfleisch, and M. A. McHugh. Influence of the alkyl tail on the solubility of poly(alkyl acrylates) in ethylene and CO2 at high pressures: Experiments and modeling. *J. Appl. Polym. Sci.* **73**, 1979 (1999).

[183] S. Benzaghou, J. P. Passarello, and P. Tobaly. Predictive use of a SAFT EoS for phase equilibria of some hydrocarbons and their binary mixtures. *Fluid Phase Equilib.* **180**, 1 (2001).

[184] J. P. Passarello and P. Tobaly. Modeling some alcohol/alkane binary systems using the SAFT equation of state with a semipredictive approach. *Ind. Eng. Chem. Res.* **42**, 5383 (2003).

[185] C. McCabe and G. Jackson. SAFT-VR modelling of the phase equilibrium of long-chain n-alkanes. *Phys. Chem. Chem. Phys.* **1**, 2057 (1998).

[186] P. Paricaud, A. Galindo, and G. Jackson. Modeling the cloud curves and the solubility of gases in amorphous and semicrystalline polyethylene with the SAFT-VR approach and Flory theory of crystallization. *Ind. Eng. Chem. Res.* **43**, 6871 (2004).

[187] N. Mac Dowell, F. Llovell, C. S. Adjiman, G. Jackson, and A. Galindo. Modeling the fluid phase behavior of carbon dioxide in aqueous solutions of monoethanolamine using transferable parameters with the SAFT-VR approach. *Ind. Eng. Chem. Res.* **49**, 1883 (2010).

[188] J. Rodriguez, N. Mac Dowell, F. Llovell, C. S. Adjiman, G. Jackson, and A. Galindo. Modeling the fluid phase behavior of aqueous mixtures of multifunctional alkanolamines and carbon dioxide using transferable parameters with the SAFT-VR approach. *Mol. Phys.* **110**, 1325 (2012).

[189] S. Tamouza, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Application to binary mixtures of a group contribution SAFT EoS (GC-SAFT). *Fluid Phase Equilib.* **228-229**, 409 (2005).

[190] K. E. Gubbins and C. H. Twu. Thermodynamics of polyatomic fluid mixtures. I: Theory. *Chem. Eng. Sci.* **33**, 863 (1978).

[191] D. Nguyen-Huynh, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Application of GC-SAFT EoS to polar systems using a segment approach. *Fluid Phase Equilib.* **264**, 62 (2008).

[192] D. Nguyen-Huynh, A. Falaix, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Predicting VLE of heavy esters and their mixtures using GC-SAFT EoS. *Fluid Phase Equilib.* **264**, 184 (2008).

[193] D. Nguyen-Huynh, J. P. Passarello, and P. Tobaly. In situ determination of phase equilibria of methyl benzoate + alkane mixtures using an infrared absorption method. Comparison with polar GC-SAFT predictions. *J. Chem. Eng. Data* **54**, 1685 (2009).

[194] C. Le Thi, S. Tamouza, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Modeling phase equilibrium of H_2 + *n*-alkane and CO_2 + *n*-alkane binary mixtures using a group contribution statistical association fluid theory equation of state (GC-SAFT-EoS) with a k_{ij} group contribution method. *Ind. Eng. Chem. Res.* **45**, 6803 (2006).

[195] D. Nguyen-Huynh, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Modeling phase equilibria of asymmetric mixtures using a group-contribution SAFT (GC-SAFT) with a k_{ij} correlation method based on London's theory. 1. Application to CO_2 + *n*-alkane, Methane + *n*-alkane and ethane + *n*-alkane systems. *Ind. Eng. Chem. Res.* **47**, 8847 (2008).

[196] D. Nguyen-Huynh, T. K. S. Tran, S. Tamouza, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Modeling phase equilibria of asymmetric mixtures using a group-contribution SAFT (GC-SAFT) with a k_{ij} correlation method based on London's theory. 2. Application to binary mixtures containing aromatic hydrocarbons, *n*-alkanes, CO_2, N_2 and H_2S. *Ind. Eng. Chem. Res.* **47**, 8859 (2008).

[197] T. K. S. Tran, D. Nguyen-Huyhn, N. Ferrando, J. P. Passarello, J. C. de Hemptinne, and P. Tobaly. Modeling VLE of H_2 + Hydrocarbon mixtures using a group contribution SAFT with a k_{ij} correlation method based on London's theory. *Energy Fuels* **23**, 2658 (2009).

[198] D. Nguyen-Huynh, J. C. de Hemptinne, R. Lugo, J. P. Passarello, and P. Tobaly. Modeling liquid-liquid and liquid-vapor equilibria of binary systems containing water with an alkane, an aromatic hydrocarbon, an alcohol or a gas (Methane, Ethane, CO_2, or H_2S), using Group Contribution Polar Perturbed-Chain Statistical Associating Fluid Theory. *Ind. Eng. Chem. Res.* **50**, 7467 (2011).

[199] M. Mourah, D. Nguyen-Huyhn, J P. Passarello, J. C. de Hemptinne, and P. Tobaly. Modelling LLE and VLE of methanol + n-alkane series using GC-PC-SAFT with a group contribution k_{ij}. *Fluid Phase Equilib.* **298**, 154 (2010).

[200] J. Rozmus, J. C. de Hemptinne, and P. Mougin. Application of GC-PPC-SAFT EoS to amine mixtures with a predictive approach. *Fluid Phase Equilib.* **303**, 15 (2011).

[201] J. Vijande, M. M. Piñeiro, D. Bessières, H. Saint-Guirons, and J. L. Legido. Description of PVT behaviour of hydrofluoroethers using the PC-SAFT EoS. *Phys. Chem. Chem. Phys.* **6**, 766 (2004).

[202] C. G. Gray and K. E. Gubbins. *Theory of Molecular Fluids.* Clarendon Press, Oxford, (1984).

[203] T. X. N. Thi, S. Tamouza, P. Tobaly, J. P. Passarelo, and J. C. de Hemptinne. Application of group contribution SAFT equation of state (GC-SAFT) to model phase behaviour of light and heavy esters. *Fluid Phase Equilib.* **238**, 254 (2005).

[204] D. Nguyen-Huynh, M. Benamira, J. P. Passarello, P. Tobaly, and J. C. de Hemptinne. Application of GC-SAFT EoS to polycyclic aromatic hydrocarbons. *Fluid Phase Equilib.* **254**, 60 (2007).

[205] F. S. Emami, A. Vahid, J. R. Jr. Elliott, and F. Feyzi. Group contribution prediction of vapour pressure with Statistical Associating Fluid Theory, Perturbed-Chain Statistical Associating Fluid Theory and Elliott-Suresh-Donohue equations of state. *Ind. Eng. Chem. Res.* **47**, 8401 (2008).

[206] A. Tihic, G. M. Kontogeorgis, N. von Solms, M. L. Michelsen, and L. Constantinou. A predictive group-contribution simplified PC-SAFT equation of state: Application to polymer systems. *Ind. Eng. Chem. Res.* **47**, 5092 (2008).

[207] A. Tihic, N. von Solms, M. L. Michelsen, G. M. Kontogeorgis, and L. Constantinou. Application of sPC-SAFT and group contribution sPC-SAFT to polymer systems - Capabilities and limitations. *Fluid Phase Equilib.* **281**, 70 (2009).

[208] J. Vijande, M. M. Piñeiro, J. L. Legido, and D. Bessières. Group-contribution method for the molecular parameters of the PC-SAFT equation of state taking into account the proximity effect. Application to nonassociated compounds. *Ind. Eng. Chem. Res.* **49**, 9394 (2010).

[209] A. L. Archer and G. Jackson. Theory and computer simulations of heteronuclear diatomic hard-sphere molecules (hard dumbbells). *Mol. Phys.* **73**, 881 (1991).

[210] M. D. Amos and G. Jackson. BHS theory and computer simulations of linear heteronuclear triatomic hard-sphere molecules. *Mol. Phys.* **74**, 191 (1991).

[211] M. D. Amos and G. Jackson. Bonded hard-sphere (BHS) theory for the equation of state of fused hard-sphere polyatomic molecules and their mixtures. *J. Chem. Phys.* **96**, 4604 (1992).

[212] A. L. Archer, M. D. Amos, G. Jackson, and I. A. McLure. The theoretical prediction of the critical points of alkanes, perfluoroalkanes, and their mixtures using bonded hard sphere (BHS) theory. *Int. J. Thermophys.* **17**, 201 (1996).

[213] H. Adidharma and M. Radosz. Prototype of an engineering equation of state for heterosegmented polymers. *Ind. Eng. Chem. Res.* **37**, 4453 (1998).

[214] H. Adidharma and M. Radosz. Square-well SAFT equation of state for homopolymeric and heteropolymeric fluids. *Fluid Phase Equilib.* **158-160**, 165 (1999).

[215] C. McCabe, A. Gil-Villegas, G. Jackson, and F. del Río. The thermodynamics of heteronuclear molecules formed from bonded square-well (BSW) segments using the SAFT-VR approach. *Mol. Phys.* **97**, 551 (1999).

[216] Y. Peng, H. Zhao, and C. McCabe. On the thermodynamics of diblock chain fluids from simulation and heteronuclear Statistical Associating Fluid Theory for potentials of variable range. *Mol. Phys.* **104**, 571 (2006).

[217] P. Morgado, H. Zhao, F. J. Blas, C. McCabe, L. P. N. Rebelo, and E. J. M. Filipe. Liquid phase behavior of perfluoroalkylalkane surfactants. *J. Phys. Chem. B* **111**, 2856 (2007).

[218] M. Banaszak, C. K. Chen, and M. Radosz. Copolymer SAFT equation of state. Thermodynamic perturbation theory extended to heterobonded chains. *Macromolecules* **29**, 6481 (1996).

[219] M. Banaszak and M. Radosz. Molecular dynamics study on homonuclear and heteronuclear chains of Lennard-Jones segments. *Fluid Phase Equilib.* **193**, 179 (2002).

[220] J. Gross, O. Spuhl, F. Tumakaka, and G. Sadowski. Modeling copolymer systems using the Perturbed-Chain SAFT equation of state. *Ind. Eng. Chem. Res.* **42**, 1266 (2003).

[221] Y. Peng, K. D. Goff, M. C. dos Ramos, and C. McCabe. Predicting the phase behaviour of polymer systems with the GC-SAFT-VR approach. *Ind. Eng. Chem. Res.* **49**, 1378 (2010).

[222] M. C. dos Ramos, J. D. Haley, J. R. Westwood, and C. McCabe. Extending the GC-SAFT-VR approach to associating functional groups: Alcohols, aldehydes, amines and carboxylic acids. *Fluid Phase Equilib.* **306**, 97 (2011).

[223] M. C. dos Ramos and C. McCabe. On the prediction of ternary mixture phase behaviour from the GC-SAFT-VR approach: 1-Pentanol + dibutyl ether + *n*-nonane. *Fluid Phase Equilib.* **302**, 161 (2011).

[224] K. Paduszyński and U. Domańska. Heterosegmented Perturbed-Chain Statistical Associating Fluid Theory as a robust and accurate tool for modeling of various alkanes. 1. Pure fluids. *Ind. Eng. Chem. Res.* **51**, 12967 (2012).

[225] A. Klamt. Conductor-like Screening Model for Real Solvents: A new approach to the quantitative calculation of solvation phenomena. *J. Phys. Chem.* **99**, 2224 (1995).

[226] A. Klamt. COSMO and COSMO-RS *in* Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer.* Wiley, New York, (1998).

[227] A. Klamt and F. Eckert. COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilib.* **172**, 43 (2000).

[228] A. Klamt and G. Schüürmann. COSMO - A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **5**, 799 (1993).

[229] O. Spuhl and W. Arlt. COSMO-RS predictions in chemical engineering - A study of the applicability to binary VLE. *Ind. Eng. Chem. Res.* **43**, 852 (2004).

[230] S. T. Lin and S. I. Sandler. A priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model. *Ind. Eng. Chem. Res.* **41**, 899 (2002).

[231] C-C. Chen and Y. Song. Solubility modeling with a Nonrandom Two-Liquid Segment Activity Coefficient Model. *Ind. Eng. Chem. Res.* **43**, 8354 (2004).

[232] H-H. Tung, J. Tabora, N. Variankaval, D. Bakken, and C-C. Chen. Prediction of pharmaceutical solubility *via* NRTL-SAC and COSMO-SAC. *J. Pharm. Sci.* **97**, 1813 (2008).

[233] Z. Xue, T. Mu, and J. Gmehling. Comparison of the a priori COSMO-RS models and group contribution methods: original UNIFAC, modified UNIFAC (Do), and modified UNIFAC (Do) consortium. *Ind. Eng. Chem. Res.* **51**, 11807 (2012).

[234] W. L. Jorgensen and J. Tirado-Rives. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **110** (1988).

[235] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Foz, J. W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* **117**, 5179 (1995).

[236] M. G. Martin and J. I. Siepmann. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *J. Phys. Chem. B* **102**, 2569 (1998).

[237] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein. A coarse grain model for n-alkanes parameterized from surface tension data. *J. Chem. Phys.* **119**, 7043 (2003).

[238] J. J. Potoff and D. A. Bernard-Brunel. Mie potentials for phase equilibria calculations: Applications to alkanes and perfluoroalkanes. *J. Phys. Chem. B* **113**, 14725 (2009).

[239] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812 (2007).

[240] T. van Westen, T. J. H. Vlugt, and J. Gross. Determining force field parameters using s physically based equation of state. *J. Phys. Chem. B* **115**, 7872 (2011).

[241] C. Avendaño, T. Lafitte, A. Galindo, C. S. Adjiman, G. Jackson, and E. A. Müller. SAFT-γ force field for the simulation of molecular fluids. I. A single-site coarse grained model of carbon dioxide. *J. Phys. Chem. B* **115**, 11154 (2011).

[242] T. Lafitte, C. Avendaño, V. Papaioannou, A. Galindo, C. S. Adjiman, G. Jackson, and E. A. Müller. SAFT-γ force field for the simulation of molecular fluids. 3. Coarse-grained models of benzene and hetero-group models of n-decylbenzene. *Mol. Phys.* **110**, 1189 (2012).

[243] S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein. Coarse grain models and the computer simulation of soft materials. *J. Phys.: Condens. Matter* **16**, R481 (2004).

[244] D. G. Green and G. Jackson. Theory of phase equilibria for model aqueous solutions of chain molecules: water + alkane mixtures. *J. Chem. Soc., Faraday Trans.* **88**, 1395 (1992).

[245] A. Galindo, P. J. Whitehead, G. Jackson, and A. N. Burgess. Predicting the high-pressure equilibria of water + n-alkanes using a simplified SAFT theory with transferable intermolecular interaction parameters. *J. Phys. Chem.* **100**, 6781 (1996).

[246] B. H. Patel, P. Paricaud, A. Galindo, and G. C. Maitland. Prediction of the salting-out effect of strong electrolytes on water + alkane solutions. *Ind. Eng. Chem. Res.* **42**, 3809 (2003).

[247] I. V. Yakoumis, G. M. Kontogeorgis, E. C. Voutsas, E. M. Hendriks, and D. P. Tassios. Prediction of phase equilibria in binary aqueous systems containing alkanes, cycloalkanes, and alkenes with the cubic-plus-association equation of state. *Ind. Eng. Chem. Res.* **37**, 4175 (1998).

[248] G. M. Kontogeorgis, I. V. Yakoumis, H. Meijer, E. Hendriks, and T. Moorwood. Multicomponent phase equilibrium calculations for water-methanol-alkane mixtures. *Fluid Phase Equilib.* **158-160**, 201 (1999).

[249] E. C. Voutsas, I. V. Yakoumis, and D. P. Tassios. Prediction of phase equilibria in water/alcohol/alkane systems. *Fluid Phase Equilib.* **158-160**, 151 (1999).

[250] E. C. Voutsas, G. C. Boulougouris, I. G. Economou, and D. P. Tassios. Water/hydrocarbon phase equilibria using the thermodynamic perturbation theory. *Ind. Eng. Chem. Res.* **39**, 797 (2000).

[251] A. Grenner, J. Schmelzer, N. von Solms, and G. M. Kontogeorgis. Comparison of two association models (Elliott-Suresh-Donohue and simplified PC-SAFT) for complex phase equilibria of hydrocarbon-water and amine-containing mixtures. *Ind. Eng. Chem. Res.* **45**, 8170 (2006).

[252] W. Yan, G. M. Kontogeorgis, and E. H. Stenby. Application of the CPA equation of state to reservoir fluids in presence of water and polar chemicals. *Fluid Phase Equilib.* **276**, 75 (2009).

[253] X. S. Li and P. Englezos. Vapor-liquid equilibrium of systems containing alcohols, water, carbon dioxide and hydrocarbons using SAFT. *Fluid Phase Equilib.* **224**, 111 (2004).

[254] E. K. Karakatsani, G. M. Kontogeorgis, and I. G. Economou. Evaluation of the truncated perturbed chain-polar Statistical Associating Fluid Theory for complex mixture fluid phase equilibria. *Ind. Eng. Chem. Res.* **45**, 6063 (2006).

[255] N. M. Al-Saifi, E. Z. Hamad, and P. Englezos. Prediction of vapor-liquid equilibrium in water-alcohol-hydrocarbon systems with the dipolar perturbed-chain SAFT equation of state. *Fluid Phase Equilib.* **271**, 82 (2008).

[256] L. F. Vega, F. Llovell, and F. J. Blas. Capturing the solubility minima of n-alkanes in water by Soft-SAFT. *J. Phys. Chem. B* **113**, 7621 (2009).

[257] H. P. Gros, S. B. Bottini, and E. A. Brignole. High pressure phase equilibrium modeling of mixtures containing associating compounds and gases. *Fluid Phase Equilib.* **139**, 75 (1997).

[258] T. M. Soria, F. A. Sánchez, S. Pereda, and S. B. Bottini. Modeling alcohol+water+hydrocarbon mixtures with the group contribution with association equation of state GCA-EoS. *Fluid Phase Equilib.* **296**, 116 (2010).

[259] N. von Solms, M. L. Michelsen, and G. M. Kontogeorgis. Computational and physical performance of a modified PC-SAFT equation of state for highly asymmetric and associating mixtures. *Ind. Eng. Chem. Res.* **42**, 1098 (2003).

[260] A. Grenner, G. M. Kontogeorgis, N. von Solms, and M. L. Michelsen. Modeling phase equilibria of alkanols with the simplified PC-SAFT equation of state and generalized pure component parameters. *Fluid Phase Equilib.* **258**, 83 (2007).

[261] L. A. Davies, A. Gil-Villegas, and G. Jackson. Describing the properties of chains of segments interacting via soft-core potentials of variable range with the SAFT-VR approach. *Int. J. Thermophys.* **19**, 675 (1998).

[262] L. A. Davies, A. Gil-Villegas, and G. Jackson. An analytical equation of state for chain molecules formed from yukawa segments. *J. Chem. Phys.* **111**, 8659 (1999).

[263] T. Lafitte, M. M. Piñeiro, J. L. Daridon, and D. Bessieres. A comprehensive description of chemical association effects on second derivative properties of alcohols through a SAFT-VR approach. *J. Phys. Chem. B* **111**, 3447 (2007).

[264] L. L. Lee. *Molecular thermodynamics of nonideal fluids.* Butterworths, Boston, (1988).

[265] J. A. Barker and D. Henderson. What is "liquid" ? Understanding the states of matter. *Rev. Mod. Phys.* **48**, 587 (1976).

[266] T. Boublík. Hard-sphere equation of state. *J. Chem. Phys.* **53**, 471 (1970).

[267] G. A. Mansoori, N. F. Carnahan, K. E. Starling, and T. W. Jr. Leland. Equilibrium thermodynamic properties of the mixture of hard spheres. *J. Chem. Phys.* **54**, 1523 (1971).

[268] J. A. Barker and D. Henderson. Perturbation theory and equation of state for fluids: the square-well potential. *J. Chem. Phys.* **47**, 2856 (1967).

[269] J. A. Barker and D. Henderson. Perturbation theory and equation of state for fluids.II. A successful theory of liquids. *J. Chem. Phys.* **47**, 4714 (1967).

[270] M. S. Wertheim. Thermodynamic perturbation theory of polymerization. *J. Chem. Phys.* **87**, 7323 (1987).

[271] R. P. Sear and G. Jackson. Thermodynamic perturbation theory for association into chains and rings. *Phys. Rev. E* **50**, 386 (1994).

[272] D. Ghonasgi, V. Perez, and W. G. Chapman. Intermolecular association in flexible hard chain molecules. *J. Chem. Phys.* **101**, 6880 (1994).

[273] R. P. Sear and G. Jackson. The ring integral in a thermodynamic perturbation theory for association. *Mol. Phys.* **87**, 517 (1996).

[274] R. P. Sear and G. Jackson. Thermodynamic perturbation-theory for association into doubly bonded dimers. *Mol. Phys.* **82**, 1033 (1994).

[275] R. P. Sear and G. Jackson. Thermodynamic perturbation theory for association with bond cooperativity. *J. Chem. Phys.* **105**, 1113 (1996).

[276] Process Systems Enterprise Ltd. gPROMS v. 3.2.0. http://www.psenterprise.com/, (2009).

[277] http://webbook.nist.gov/, (2010), National Institutes of Standards and Technology WebBook.

[278] G. N. I. Clark, A. J. Haslam, A. Galindo, and G. Jackson. Developing optimal Wertheim-like models of water for use in Statistical Associating Fluid Theory (SAFT) and related approaches. *Mol. Phys.* **104**, 3561 (2006).

[279] I. Máchová, J. Linek, and I. Wichterle. Vapour-liquid equilibria in the heptane - 1-pentanol and heptane - 3-methyl-1-butanol systems at 75, 85 and 95 ° C. *Fluid Phase Equilib.* **41**, 257 (1988).

[280] B. D. Smith and R. Srivastava. *Thermodynamic data for pure compounds part B: Halogenated compounds and alcohols.* Elsevier, Amsterdam, (1986).

[281] A. Yu. Namiot, V. G. Skripka, and G. Yu. Lotter. Phase equilibria in hydrocarbon - water systems at high temperatures. *Deposited Doc. VINITI* **1213-76**, 1 (1976). (cited in DETHERM).

[282] R. G. Sultanov and V. G. Skripka. Solubility of water in n-alkanes at elevated temperatures and pressures. *Deposited Doc. VINITI* **4386-72**, 1 (1972). (cited in DETHERM).

[283] R. G. Sultanov and V. G. Skripka. The solubility of water in n-hexane, cyclohexane and benzene at elevated temperatures and pressures. *Deposited Doc. VINITI* **2347-73**, 1 (1973). (cited in DETHERM).

[284] F. E. Pereira, G. Jackson, A. Galindo, and C. S. Adjiman. A duality-based optimisation approach for the reliable solution of (p, T) phase equilibrium in volume-composition space. *Fluid Phase Equilib.* **299**, 1 (2010).

[285] http://www.nag.co.uk/numeric/fl/manual20/pdf/E04/e04ucf.pdf, (2010).

[286] S. O. Derawi, G. M. Kontogeorgis, and E. H. Stenby. Application of group contribution models to the calculation of the octanol-water partition coefficient. *Ind. Eng. Chem. Res.* **40**, 434 (2001).

[287] T-H. Cho, K. Ochi, and K. Kojima. Isobaric vapor-liquid equilibria for binary systems with limited miscibility, water-*n*-amyl alcohol and water-isoamyl alcohol. *Kagaku Kogaku Ronbunshu* **10**, 181 (1984). (cited in DETHERM).

[288] R. Stephenson, J. Stuart, and M. Tabak. Mutual solubility of water and aliphatic alcohols. *J. Chem. Eng. Data* **29**, 287 (1984).

[289] D. Ambrose and C. H. S. Sprake. Thermodynamic properties of organic oxygen compounds XXV. Vapour pressures and normal boiling temperatures of aliphatic alcohols. *J. Chem. Thermodyn.* **2**, 631 (1970).

[290] E. C. Ihmels. Viscosity measurements with a Ubbelohde-viscosimeter and density measurements with a resonance frequency transducer prototype. Diplomarbeit, (1998). (cited in DETHERM).

[291] H. R. Kemme and S. I. Kreps. Vapor pressure of primary n-alkyl chlorides and alcohols. *J. Chem. Eng. Data* **14**, 98 (1969).

[292] J. M. Costello and S. T. Bowden. The temperature variation of orthobaric density difference in liquid-vapor systems. III. Alcohols. *Recl. Trav. Chim. Pays-Bas* **77**, 36 (1958).

[293] D. Ambrose, J. H. Ellender, and C. H. S. Sprake. Thermodynamic properties of organic oxygen compounds XXXV. Vapour pressures of aliphatic alcohols. *J. Chem. Thermodyn.* **6**, 909 (1974).

[294] J. Seo, J. Lee, and H. Kim. Isothermal vapor-liquid equilibria for the system ethanol and *n*-hexane in the near critical region. *Fluid Phase Equilib.* **182**, 199 (2001).

[295] C. Berro and A. Peneloux. Excess Gibbs energy and excess volumes of 1-butanol+*n*-heptane and 2-methyl-1-propanol+*n*-heptane binary systems. *J. Chem. Eng. Data* **29**, 206 (1984).

[296] H. Wolff and A. Shadiakhy. The vapor pressure behavior and association of mixtures of 1-hexanol and *n*-hexane between 293 and 373 K. *Fluid Phase Equilib.* **7**, 309.

[297] Z. Plesnar, P. Gierycz, and A. Bylicki. Vapour-liquid equilibrium and solid-liquid equilibrium in the system formed by 1-octanol and 1-octane. *Thermochim. Acta* **128**, 93 (1988).

[298] S. A. Wieczorek. Vapour pressures and thermodynamic properties of decan-1-ol+*n*-hexane between 283.160 and 333.151 K. *J. Chem. Thermodyn.* **11**, 239.

[299] P. C. Joyce and M. C. Thies. Vapor-liquid equilibria for the hexane+hexadecane and hexane+1-hexadecanol systems at elevated temperatures and pressures. *J. Chem. Eng. Data* **43**, 819 (1998).

[300] J. Schmelzer, I. Lieberwirth, M. Krug, and R. Pfestorf. Vapor-liquid equilibria and heats of mixing in alkane-alcohol systems. I. Vapor-Liquid equilibria in 1-alcohol - undecane systems. *Fluid Phase Equilib.* **11**, 187 (1983).

[301] H. Matsuda and K. Ochi. Liquid-liquid equilibrium data for binary alcohol + *n*-alkane (C_{10}-C_{16}) systems: methanol + decane, ethanol + tetradecane and ethanol + hexadecane. *Fluid Phase Equilib.* **224**, 31 (2004).

[302] J. Schmelzer, V. Creutziger, I. Lieberwirth, and R. Pfestorf. Vapour-liquid equilibria and heats of mixing in n-alkane-1-alcohol systems. III. Vapour-liquid equilibria in n-alkane-1-dodecanol systems. *Fluid Phase Equilib.* **15**, 107 (1983).

[303] P. H. van Konynenburg and R. L. Scott. Critical lines and phase equilibria in binary van der Waals mixtures. *Philos. Trans. R. Soc. London, Ser. A* **298**, 495 (1980).

[304] H. H. Reamer, B. H. Sage, and W. N. Lacey. Phase equilibria in hydrocarbon systems: n-butane - water system in the two-phase region. *Ind. Eng. Chem.* **44**, 609 (1952).

[305] M. Tu, D. Fei, Y. Liu, and J. Wang. Phase equilibrium of the partially miscible system of octane-water. *Gaoxiao Huaxue Gongcheng Xuebao* **12**, 325 (1998). (cited in DETHERM).

[306] R. G. Sultanov, V. G. Skripka, and A. Yu. Namiot. Phase equilibria in the system water-n-hexadecane at elevated temperatures. *Deposited Doc. VINITI* **3139-71**, 1 (1971). (cited in DETHERM).

[307] R. Kobayashi and D. Katz. Vapor-Liquid equilibria for binary hydrocarbon-water systems. *Ind. Eng. Chem.* **45**, 440 (1944).

[308] S. Mokraoui, C. Coquelet, A. Valtz, P. E. Hegel, and D. Richon. New solubility data of hydrocarbons in water and modeling concerning vapor-liquid-liquid binary systems. *Ind. Eng. Chem. Res.* **46**, 9257.

[309] H. H. Reamer, R. H. Olds, B. H. Sage, and W. N. Lacey. Phase equilibria in hydrocarbon systems. *n*-Butane Ű water system in three-phase region. *Ind. Eng. Chem.* **36**, 381 (1944).

[310] H. D. Nelson and C. L. de Ligny. Interpretation of solubility data on the *n*-alkanes and some other substances in the solvent water. *Recl. Trav. Chim. Pays-Bas* **87**, 623 (1968).

[311] B. A. Englin, A. Plate, V. M. Tugolkov, and M. A. Pryanishnikova. Solubiliy of water in individual hydrocarbons. *Khim. Tekhnol. Topliv Masel* **10**, 42 (1965). (cited in DETHERM).

[312] C. Tsonopoulos and G. M. Wilson. High-temperature mutual solubilities of hydrocarbons and water. Part I: Benzene, cyclohexane and *n*-hexane. *AIChE J.* **29**, 990 (1983).

[313] J. L. Heidman, C. Tsonopoulos, C. J. Brady, and G. M. Wilson. High-temperature mutual solubilities of hydrocarbons and water. Part II: Ethylbenzene, ethylcyclohexane, and *n*-octane. *AIChE J.* **31**, 376 (1985).

[314] S. P. Tunik, T. M. Lesteva, and V. I. Chernaya. Phase equilibria in the water-alcohol-formaldehyde systems. I. The liquid-vapour equilibria in the systems water-C6 and C7 alcohols. *Deposited Doc. VINITI* **435-77**, 1 (1977). (cited in DETHERM).

[315] M. Góral, B. Wiśniewska-Cocłowska, and A. Maczyński. Recommended liquid-liquid equilibrium data. Part 4. 1-alkanol-water systems. *J. Phys. Chem. Ref. Data* **35**, 1391 (2006).

[316] K. Iwakabe and H. Kosuge. Isobaric vapor-liquid-liquid equilibria with a newly developed still. *Fluid Phase Equilib.* **192**, 171 (2001).

[317] R. Stephenson and J. Stuart. Mutual binary solubilities: Water-alcohols and water-esters. *J. Chem. Eng. Data* **31**, 56 (1986).

[318] D. G. Green and G. Jackson. Theory of phase equilibria and closed-loop liquid–liquid immiscibility for model aqueous solutions of associating chain molecules: Water+alkanol mixtures. *J. Chem. Phys.* **97**, 8672 (1992).

[319] I. Nezbeda, J. Pavlíček, J. Kolafa, A. Galindo, and G. Jackson. Global phase behavior of model mixtures of water and n-alkanols. *Fluid Phase Equilib.* **158**, 193 (1999).

[320] T. M. Letcher, S. Wootton, B. Shuttleworth, and C. Heyward. Phase equilibria for (*n*-heptane+water+an alcohol) at 298.2K. *J. Chem. Thermodyn.* **18**, 1037 (1986).

[321] A. M. Kiryukhin, T. M. Lesteva, and N. P. Markuzin. Phase equilibrium in hydrocarbon-alcohol-water systems. *Promst. Sint. Kauch.* **12**, 12 (1981). (cited in DETHERM).

[322] H. H. Reamer and B. H. Sage. Phase equilibria in hydrocarbon systems. Phase behavior in the n-butane - n-decane System. *J. Chem. Eng. Data* **9**, 24 (1964).

[323] C. McCabe and S. B. Kiselev. Application of crossover theory to the SAFT-VR equation of state: SAFT-VRX for pure fluids. *Ind. Eng. Chem. Res.* **43**, 2839 (2004).

[324] E. Forte, F. Llovell, L. F. Vega, J. P. M. Trusler, and A. Galindo. Application of a renormilization-group treatment to the statistical associating fluid theory for potentials of variable range (SAFT-VR). *J. Chem. Phys.* **134**, 154102 (2011).

[325] E. Forte, F. Llovell, J. P. M. Trusler, and A. Galindo. Application of the statistical associating fluid theory for potentials of variable range (SAFT-VR) coupled with renormalisation-group (RG) theory to model the phase equilibria and second-derivative properties of pure fluids. *Fluid Phase Equilib.* **337**, 274 (2013).

[326] S. Dutour, J.-L. Daridon, and B. Lagourette. Speed of sound, density and compressibilities of liquid eicosane and docosane at various temperatures and pressures. *High Temp. High Pressures* **33**, 371 (2001).

[327] J.-L. Daridon, H. Carrier, and B. Lagourette. Pressure Dependence of the thermophysical properties of *n*-pentadecane and *n*-heptadecane. *Int. J. Thermophys.* **33**, 697 (2002).

[328] C. M. Colina, L. F. Turrens, K. E. Gubbins, C. Oliveira-Fuentes, and L. F. Vega. Predictions of the Joule-Thomson inversion curve for the n-alkane series and carbon dioxide from the Soft-SAFT Equation of State. *Ind. Eng. Chem. Res.* **41**, 1069 (2002).

[329] F. Llovell and L. F. Vega. Prediction of thermodynamic derivative properties of pure fluids through the soft-SAFT equation of state. *J. Phys. Chem. B* **110**, 11427 (2006).

[330] N. I. Diamantonis and I. G. Economou. Evaluation of Statistical Associating Fluid Theory (SAFT) and Perturbed Chain-SAFT equations of state for the calculation of thermodynamic derivative properties of fluids related to carbon capture and sequestration. *Energy Fuels* **25**, 334 (2011).

[331] A. Diedrichs, J. Rarey, and J. Gmehling. Prediction of liquid heat capacities by the group contribution equation of state VTPR. *Fluid Phase Equilib.* **248**, 56 (2006).

[332] A. M. A. Dias, F. Llovell, J. A. P. Coutinho, I. M. Carrucho, and L. F. Vega. Thermodynamic characterization of pure perfluoroalkanes, including interfacial and second order derivative properties, using the crossover soft-SAFT EoS. *Fluid Phase Equilib.* **286**, 134 (2009).

[333] J. Gregorowitz, J. P. O'Connell, and C. J. Peters. Some characteristics of pure fluid properties that challenge equation-of-state models. *Fluid Phase Equilib.* **116**, 94 (1996).

[334] M. R. Faradonbeh, J. Abedi, and T. G. Harding. Comparative study of eight cubic equations of state for predicting thermodynamic properties of alkanes. *Can. J. Chem. Eng.* (2011). in press.

[335] F. J. Blas and L. F. Vega. Prediction of binary and ternary diagrams using the Statistical Associating Fluid Theory (SAFT) equation of state. *Ind. Eng. Chem. Res.* **37**, 660 (1998).

[336] F. Castro-Marcano, C. G. Olivera-Fuentes, and C. M. Colina. Joule-Thomson inversion curves and third virial coefficients for pure fluids from molecular-based models. *Ind. Eng. Chem. Res.* **47**, 8894 (2008).

[337] F. Llovell, C. J. Peters, and L. F. Vega. Second-order thermodynamic derivative properties of selected mixtures by the soft-SAFT equation of state. *Fluid Phase Equilib.* **248**, 115 (2006).

[338] J. Chen and J.-G. Mi. Equation of state extended from SAFT with improved results for non-polar fluids across the critical point. *Fluid Phase Equilib.* **186**, 165 (2001).

[339] A. Maghari and M. S. Sadeghi. Prediction of sound velocity and heat capacities of n-alkanes from the modified SAFT-BACK equation of state. *Fluid Phase Equilib.* **252**, 152 (2007).

[340] A. Maghari and M. Hamzehloo. Second-order thermodynamic derivative properties of binary mixtures of n-alkanes through the SAFT-CP equation of state. *Fluid Phase Equilib.* **302**, 195 (2011).

[341] G. Mie. Zur kinetischen Theorie der einatomigen Körper. *Ann. Phys.* **316**, 657 (1903).

[342] J. E. Jones. On the determination of molecular fields. I. From the variation of the viscosity of a gas with temperature. *Proc. Roy. Soc. London A* **106**, 441 (1924).

[343] J. E. Jones. On the determination of molecular fields. II. From the equation of state of a gas. *Proc. Roy. Soc. London A* **106**, 463 (1924).

[344] J. E. Lennard-Jones. Cohesion. *Proc. Phys. Soc.* **43**, 461 (1931).

[345] P. Paricaud. A general perturbation approach for equation of state development: Applications to simple fluids, *ab initio* potentials, and fullerenes. *J. Chem. Phys.* **124**, 154505 (2006).

[346] B.-J. Zhang. Calculating thermodynamic properties from perturbation theory I. An analytic representation of square-well potential hard-sphere perturbation theory. *Fluid Phase Equilib.* **154**, 1 (1999).

[347] T. Boublík. Background correlation functions in the hard sphere systems. *Mol. Phys.* **59**, 775 (1986).

[348] V. Papaioannou, C. S. Adjiman, G. Jackson, and A. Galindo. Simultaneous prediction of vapour-liquid and liquid-liquid equilibria (VLE and LLE) of aqueous mixtures with the SAFT-γ group contribution approach. *Fluid Phase Equilib.* **306**, 82 (2011).

[349] F. London. The general theory of molecular forces. *Trans. Faraday Soc.* **33**, 8 (1937).

[350] M. Kleiner and G. Sadowski. Modeling of polar systems using PCP-SAFT: An approach to account for induced-association interactions. *J. Phys. Chem. C* **111**, 15544 (2007).

[351] A. N. Campbell and R. M. Chatterjee. Orthobaric data of certain pure liquids in the neighborhood of the critical point. *Can. J. Chem.* **46**, 575 (1968).

[352] G. Liessmann, W. Schmidt, and S. Reiffarth. Recommended Thermophysical Data. *Data compilation of the Saechsische Olefinwerke Boehlen Germany* (1995). (cited in DETHERM).

[353] I. A. Tugarev, Z. I. Avdus, and V. F. Nozdrev. An experimental study of the p-V-T parameters of the liquid phase in the acetone-benzene mixtures along the boundary curve. *Zh. Fiz. Khim.* **49**, 1256 (1975). (cited in DETHERM).

[354] D. Ambrose, J. H. Ellender, E. B. Less, C. H. S. Sprake, and R. Townsend. Thermodynamic properties of organic oxygen compounds XXXVIII. Vapour pressures of some aliphatic ketones. *J. Chem. Thermodyn.* **7**, 453 (1975).

[355] T. Guetachew, I. Mokbel, J. P. Meille, and J. Jose. Vapor pressures and sublimation pressures of naphthalene and of five alkan-2-ones (C4, C5, C7, C8, C9) at pressures in the range from 0.3 Pa to 160 kPa. *ELDATA Int. Electron. J. Phys. Chem. Data* **1**, 249 (1995). (cited in DETHERM).

[356] K. A. Kobe, H. R. Crawford, and R. W. Stephenson. Industrial design data – Critical properties and vapor pressures of some ketones. *Ind. Eng. Chem.* **47**, 1767 (1955).

[357] DIPPR 801 Tables, Thermophysical Properties Database, http://dippr.byu.edu/, (2010).

[358] M. A. Espinosa Dìaz, T. Guetachew, P. Landy, J. Jose, and A. Voilley. Experimental and estimated saturated vapour pressures of aroma compounds. *Fluid Phase Equilib.* **157**, 257 (1999).

[359] H. Pöhler and E. Kiran. Volumetric properties of Carbon Dioxide + Acetone at high pressures. *J. Chem. Eng. Data* **42**, 379 (1997).

[360] V. A. Atoyan and I. A. Mamedov. Investigation of the acoustic and viscosity characteristics of some liquid ketones with V=const and as a function of the molecular weight. *Zh. Fiz. Khim.* **54**, 856 (1980). (cited in DETHERM).

[361] T. A. Apaev, A. M. Kerimov, and N. Dzhanakhmedov. Studies in p-V-T dependence of methyl propyl ketone within a temperature range and pressure range. *Ukr. Fiz. Zh. Russ. Ed.* **22**, 408 (1977). (cited in DETHERM).

[362] F. Comelli. Calorimetric study of binary mixtures containing 1-chloronapthalene + aliphatic ketones. *Chim. Ind. Milan* **73**, 269 (1991). (cited in DETHERM).

[363] D. M. Cowan, G. H. Jeffery, and A. I. Vogel. Physical Properties and chemical constitution. Part V. Alkyl Ketones. *J. Chem. Soc.* , 171 (1940).

[364] T. A. Apaev, A. M. Kerimov, and N. Dzhanakhmedov. Observation of p-V-T dependence of methyl hexyl ketone in a wide interval of temperatures and pressures. *Ukr. Fiz. Zh. Russ. Ed.* **25**, 166 (1978). (cited in DETHERM).

[365] D. Nguyen-Huynh, J. P. Passarello, J. C. de Hemptinne, and P. Tobaly. Extension of polar GC-SAFT to systems containing some oxygenated compounds: Applications to ethers, aldehydes and ketones. *Fluid Phase Equilib.* **307**, 142 (2011).

[366] K. Schaefer. Excess enthalpies of acetone + *n*-alkane (C3-C6) binary mixtures. *Int. Data Ser. Sel. Data Mixtures Ser. A* **1**, 74 (1978). (cited in DETHERM).

[367] O. Kiyohara, Y. P. Handa, and G. C. Benson. Thermodynamic properties of binary mixtures containing ketones III. Excess enthalpies of *n*-alkanes + some aliphatic ketones. *J. Chem. Thermodyn.* **11**, 453 (1979). (cited in DETHERM).

[368] P. P. Singh, R. Malik, S. Maken, W. E. Acree, and S. A. Tucker. Thermochemical investigations of associated solutions. 10. Excess enthalpies and excess volumes of ternary acetone + bromoform + n-hexane Mixtures. *Thermochim. Acta* **162**, 291 (1990). (cited in DETHERM).

[369] E. W. Lemmon, M. L. Hubert, and M. O. McLinden. NIST Standard reference database 23: Reference fluid thermodynamic and transport properties - REFPROP, Version 9.0. National Institute of Standards and Technology, Standard Reference Data Program, Gaithersburg, (2010).

[370] C.-S. Wu and Y.-P. Chen. Calculation of vapor-liquid equilibria of polymer solutions using the SAFT equation of state. *Fluid Phase Equilib.* **100**, 103 (1994).

[371] J. Gross and G. Sadowski. Modeling polymer systems using the Perturbed-Chain Statistical Associating Fluid Theory equation of state. *Ind. Eng. Chem. Res.* **41**, 1084 (2002).

[372] P. Zoller and D. J. Walsh. *Standard pressure-volume-temperature data for polymers.* Technomic, Lancaster-Basel, (1995).

[373] V. A. Atoyan and I. A. Mamedov. Setting for measuring the velocity of sound in the methylethylketone by the high pressures and the temperatures. *Izv. Akad. Nauk Az. SSR Ser. Fiz. Tekh. Mat. Nauk* **6**, 123 (1975). (cited in DETHERM).

[374] T. A. Apaev, N. Dzhanakhmedov, and K. D. Gyulmamedov. Some thermodynamical properties of methyl hexyl ketone in a wide range of temperature and pressure. *Izv. Vyssh. Uchebn. Zaved. Neft Gaz* **8**, 37 (1991). (cited in DETHERM).

[375] S. Salerno, M. Cascella, D. May, P. Watson, and D. Tassios. Prediction of vapor pressures and saturated volumes with a simple cubic equation of state: Part I. A reliable data base. *Fluid Phase Equilib.* **27**, 15 (1986).

[376] C. Viton, M. Chavret, E. Behar, and J. Jose. Vapor pressure of normal alkanes from decane to eicosane at temperatures from 244 K to 469 K and pressures from 0.4 Pa to 164 kPa. *Int. Electron. J. Phys. Chem. Data* **2**, 215 (1996). (cited in DETHERM).

[377] R. D. Chirico, A. Nguyen, W. V. Steele, M. M. Strube, and C. Tsonopoulos. Vapor pressure of n-alkanes revisited. New high-precision vapor pressure data on *n*-decane, *n*-eicosane and *n*-octacosane. *J. Chem. Eng. Data* **34**, 149 (1989).

[378] T. Sawaya, I. Mokbel, N. Ainous, E. Rauzy, C. Berro, and J. Jose. Experimental vapor pressures of six n-alkanes (C21, C23, C25, C27, C29, C30) in the temperature range between 350 K and 460 K. *J. Chem. Eng. Data* **51**, 854 (2006).

[379] C. E. Schwarz and I. Nieuwoudt. Phase equilibrium of propane and alkanes part II: Hexatriacontane through hexacontane. *J. Supercrit. Fluids* **27**, 145 (2003).

[380] A. R. Shultz and P. J. Flory. Phase equilibria in polymer-solvent systems. *J. Am. Chem. Soc.* **74**, 4760 (1952).

[381] A. Touriño, M. Hervello, V. Moreno, M. Iglesias, and G. Marino. Thermodynamics of binary mixtures of aliphatic linear alkanes (C_6-C_{12}) at 298.15 K. *Phys. Chem. Liq.* **42**, 37 (2004).

[382] M. F. Bolotnikov, Y. A. Neruchev, Y. F. Melikhov, V. N. Verveyko, and M. V. Verveyko. Temperature dependence of the speed of sound, densities and isentropic compressibilities of hexane + hexadecane in the range of (293.15 to 373.15) K. *J. Chem. Eng. Data* **50**, 1095 (2005).

[383] P. J. Flory, R. A. Orwoll, and A. Vrij. Statistical thermodynamics of chain molecule liquids. I. An equation of state for normal paraffin hydrocarbons. *J. Am. Chem. Soc.* **86**, 3507 (1964). 17.

[384] P. J. Flory, R. A. Orwoll, and A. Vrij. Statistical thermodynamics of chain molecule liquids. II. Liquid mixtures of normal paraffin hydrocarbons. *J. Am. Chem. Soc.* **86**, 3515 (1964). 17.

[385] R. A. Orwoll and P. J. Flory. Thermodynamic properties of binary mixtures of *n*-alkanes. *J. Am. Chem. Soc.* **89**, 6822 (1967). 26.

[386] F. J. Blas and I. Fujihara. Excess properties of Lennard-Jones binary mixtures from computer simulation and theory. *Mol. Phys.* **100**, 2823 (2002).

[387] F. J. Blas. Excess thermodynamic properties of chainlike mixtures. II. Self-associating systems: predictions from soft-SAFT and molecular simulation. *Mol. Phys.* **100**, 2221 (2002).

[388] S. K. Ogorodnikov, V. B. Kogan, and M. S. Nemtsov. Liquid-vapor equilibrium in binary systems of hydorcarbons with acetone. *J. Appl. Chem. USSR* **34**, 313 (1961). (cited in DETHERM).

[389] G. Spinolo and R. Riccardi. Liquid-liquid equilibria in propanone and *n*-alkane mixtures. *Int. Data Ser. Sel. Data Mixtures Ser. A* **2**, 91 (1977). (cited in DETHERM).

[390] V. O. Maripuri and G. A. Ratcliff. Measurement of isothermal vapor-liquid equilibriums for acetone - *n*-heptane mixtures using modified Gillespie still. *J. Chem. Eng. Data* **17**, 366 (1972).

[391] G. Kolasinska, M. Goral, and J. Giza. Vapor-liquid equilibria and excess gibbs free energy in binary systems of acetone with aliphatic and aromatic hydrocarbons at 313.15 K. *Z. Phys. Chem. Leipzig* **263**, 151 (1982). (cited in DETHERM).

[392] H. Pluddemann and K. Schaefer. Investigations on the liquid mixture acetone - n-heptane at low and high concentrations. *Z. Elektrochem.* **63**, 1024 (1959). (cited in DETHERM).

[393] F. Ratkovics and B. Palagyi-Fenyes. Physico chemical properties of binary mixtures: 2-Butanone + n-alkane systems. *Private Communication* (1987). (cited in DETHERM).

[394] M. Takeo, K. Nishi, T. Nitta, and T. Katayama. Isothermal vapor-liquid equilibria for two binary mixtures of heptane with 2-butanone and 4-4-methyl-2-pentanone measured by a dynamic still with a pressure regulation. *Fluid Phase Equilib.* **3**, 123 (1979).

[395] D. O. Hanson and M. Van Winkle. Alteration of the relative volatility of hexane-1-hexene by oxygenated and chlorinated solvents. *J. Chem. Eng. Data* **12**, 319 (1967).

[396] O. Kiyohara, G. C. Benson, and J. P. E. Grolier. Thermodynamic properties of binary mixtures containing ketones I. Excess enthalpies of some aliphatic ketones

+ n-hexane, + benzene, and + tetrachloromethane. *J. Chem. Thermodyn.* **9**, 315 (1977). (cited in DETHERM).

[397] A. V. Ramallo, J. L. Legido, and J. Tojo. Densities , refractive indexes and excess molar volumes of $x_1CH_3COCH_2CH_3$ + $x_2$cyclo-C_6H_{12} + (1-x_1-x_2)$CH_3(CH_2)_5CH_3$ at the temperature 298.15 K. *J. Chem. Thermodyn.* **25**, 993 (1993). (cited in DETHERM).

[398] J. P. E. Grolier and G. C. Benson. Thermodynamic properties of binary mixtures containing ketones. VIII. Heat capacities and volumes of some n-alkanone + n-alkane mixtures at 298.15 K. *Can. J. Chem.* **62**, 949 (1984). (cited in DETHERM).

[399] H. Modaressi, E. Conte, J. Abildskov, R. Gani, and P. Crafts. Model-based calculation of solid solubility for solvent selection - a review. *Ind. Eng. Chem. Res.* **47**, 5234 (2008).

[400] P. McKenzie, S. Kiang, J. Tom, A. E. Rubin, and M. Futran. Can pharmaceutical process development become high tech? *AIChE J.* **52**, 3990 (2006).

[401] J. M. Prausnitz, R. N. Lichtentaler, and E. G. de Azevedo. *Molecular Thermodynamics of Fluid-Phase Equilibria.* Prentice-Hall PTR, New Jersey, 3rd edition, (1999).

[402] C. Pan and M. Radosz. Modeling of solid-liquid equilibria in naphthalene, normal-alkane and polyethylene solutions. *Fluid Phase Equilib.* **155**, 57 (1999).

[403] I. Tsivintzelis, I. G. Economou, and G. M. Kontogeorgis. Modeling the solid-liquid equilibrium in pharmaceutical-solvent mixtures: Systems with complex hydrogen bonding behavior. *AIChE J.* **55**, 756 (2009).

[404] I. Tsivintzelis, I. G. Economou, and G. M. Kontogeorgis. Modeling the phase behaviour in mixtures of pharmaceuticals with liquid or supercritical solvents. *J. Phys. Chem. B* **113**, 6446 (2009).

[405] S. Gracin, T. Brinck, and Å. C. Rasmuson. Prediction of solubility of solid organic compounds in solvents by UNIFAC. *Ind. Eng. Chem. Res.* **41**, 5114 (2002).

[406] S. H. Neau, S. V. Bhandarkar, and E. W. Hellmuth. Differential molar heat capacities to test ideal solubility estimations. *Pharm. Res.* **14**, 601 (1997).

[407] J. S. Chickos, C. M. Braton, and D. G. Hesse. Estimating entropies and enthalpies of fusion of organic compounds. *J. Org. Chem.* **56**, 927 (1991).

[408] Y. M. Oskoei and M. H. Keshavarz. Improved method for reliable predicting enthalpy of fusion of energetic compounds. *Fluid Phase Equilib.* **326**, 1 (2012).

[409] A. S. Hukkerikar, B. Sarup, A. T. Kate, J. Abildskov, G. Sin, and R. Gani. Group-contribution$^+$ (GC^+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilib.* **321**, 25 (2012).

[410] A. Jouyban, M. A. A. Fakhree, and A. Shyanfar. Solubility prediction methods for drug/drug like molecules. *Recent Pat. Chem. Eng.* **1**, 220 (2008).

[411] H. Matsuda, K. Kaburagi, K. Kurihara, K. Tochigi, and K. Tomono. Prediction of solubilities of pharmaceutical compounds in water + co-solvent systems using an activity coefficient model. *Fluid Phase Equilib.* **290**, 153 (2010).

[412] M. Ellegaard, J. Abildskov, and J. P. O'Connell. Method for predicting solubilities of solids in mixed solvents. *AIChE J.* **55**, 1256 (2009).

[413] F. L. Mota, A. J. Queimada, S. P. Pinho, and E. A. Macedo. Water solubility of drug-like molecules with the cubic-plus-association equation of state. *Fluid Phase Equilib.* **298**, 75 (2010).

[414] F. L. Mota, A. J. Queimada, S. P. Pinho, and E. A. Macedo. Solubility of drug-like molecules in pure organic solvents with the CPA EoS. *Fluid Phase Equilib.* **303**, 62 (2011).

[415] J. Cuevas, F. Llovell, A. Galindo, V. Vesovic, H. Segura, and J. R. Pérez-Correa. Solid-liquid equilibrium using the SAFT-VR equation of state: Solubility of naphthalene and acetic acid in binary mixtures and calculation of phase diagrams. *Fluid Phase Equilib.* **306**, 137 (2011).

[416] F. Ruether and G. Sadowski. Modeling the solubility of pharmaceuticals in pure solvents and solvent mixtures for drug process design. *J. Pharm. Sci.* **98**, 4205 (2009).

[417] T. Spyriouni, X. Krokidis, and I. G. Economou. Thermodynamics of pharmaceuticals: Prediction of solubility in pure and mixed solvents with PC-SAFT. *Fluid Phase Equilib.* **302**, 331 (2011).

[418] J. Cassens, F. Ruether, K. Leonhard, and G. Sadowski. Solubility calculation of pharmaceutical compounds - A priori parameter estimation using quantum-chemistry. *Fluid Phase Equilib.* **299**, 161 (2010).

[419] F. L. Mota, A. P. Carneiro, A. J. Queimada, S. P. Pinho, and E. A. Macedo. Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling. *Eur. J. Pharm. Sci.* **37**, 499 (2009).

[420] A. Diedrichs and J. Gmehling. Solubility calculation of active pharmaceutical ingredients in alkanes, alcohols, water and their mixtures using various activity coefficient models. *Ind. Eng. Chem. Res.* **50**, 1757 (2011).

[421] F. Eckert and A. Klamt. Fast solvent screening via Quantum Chemistry: COSMO-RS Approach. *AIChE J.* **48**, 369 (2002).

[422] E. Mullins, Y. A. Liu, A. Ghaderi, and S. D. Fast. Mixtures for organic pharmacological compounds with COSMO-based thermodynamic methods. *Ind. Eng. Chem. Res.* **47**, 1707 (2008).

[423] C.-C. Shiu and S.-T. Lin. Prediction of drug solubility in mixed solvent systems using the COSMO-SAC activity coefficient model. *Ind. Eng. Chem. Res.* **50**, 142 (2011).

[424] A. A. Louis, P. G. Bolhuis, J. P. Hansen, and E. J. Meijer. Can polymer coils be modeled as "soft collods" ? *Phys. Rev. Lett.* **85**, 2522 (2000).

[425] E. A. Müller and K. E. Gubbins. An Equation of State for Water from a Simplified Intermolecular Potential. *Ind. Eng. Chem. Res.* **34**, 3662 (1995).

[426] C. Berro, F. Laichoubi, and E. Rauzy. Isothermal vapor-liquid equilibria and excess volumes for the systems n-hexane + ethylbenzene, 2-methylpentane + n-heptane, and 2-methylpentane + n-octane. *J. Chem. Eng. Data* **36**, 474 (1991).

[427] J. L. H. Wang and B. C. Y. Li. Vapor-liquid equilibrium data for the n-pentane - benzene system. *J. Appl. Chem. Biotechnol.* **21**, 297 (1971).

[428] H. I. Paul, J. Krug, and H. Knapp. Measurements of VLE, h^E and v^E for binary mixtures of n-alkanes with n-alkylbenzenes. *Thermochim. Acta* **108**, 9 (1986).

[429] S. J. Lodl and W. A. Scheller. Isothermal vapor-liquid equilibrium data for the system n-heptane - n-valeric acid at 50, 75 and 100 C. *J. Chem. Eng. Data* **12**, 458 (1967).

[430] http://wtt-pro.nist.gov/, (2012), NIST Standard Reference Subscription Database 3 - Professional Edition, Version 2-2012-1-Pro.

[431] M. Diaz Pena, A. Crespo Colin, and A. Compostizo. Isothermal liquid-vapor equilibria 2. The binary systems formed by benzene + acetone + methyl ethyl ketone + methyl propyl ketone and + methyl isobutyl ketone. *J. Chem. Thermodyn.* **10**, 1101 (1978).

[432] S. M. Ashraf and D. H. L. Prasad. Isothermal phase equilibria in the binary systems formed by methyl isobutyl ketone with alkylbenzenes. *Phys. Chem. Liq.* **40**, 345 (2002).

[433] S. C. Jain, O. P. Bagga, and K. S. N. Raju. Isobaric vapor-liquid equilibrium data for the systems ethylbenzene - propionic acid and p-xylene - propionic Acid. *Indian J. Technol.* **15**, 174 (1977).

[434] P. Rasmussen adn L. Kierkegaard and A. Fredenslund. Vapor-liquid equilibria in ketone-organic acid mixtures. *Chemical Engineering with Per Søltoft, Teknisk Forlag* , 129 (1977). (cited in DETHERM).

[435] C. W. Hoerr and H. J. Hardwood. Solubilities of high molecular weight aliphatic compounds in *n*-hexane. *J. Org. Chem.* **16**, 779 (1951).

[436] U. Domanska and K. Knia. Solid-liquid equilibria of normal alkanes (C16, C18, C20) + hexane, + 3-methylpentane, + 2,2-dimethylbutane, or + cyclohexane. *Int. Data Series, Ser. A., Sel. Data Mixtures* **2**, 83 (1990). (cited in DETHERM).

[437] M. Dirand, M. Bouroukba, A.-J. Briard, V. Chevallier, and D. Petitjean. Temperatures and enthalpies of (solid+solid) and (solid+liquid) transitions of *n*-alkanes. *J. Chem. Thermodynamics* **34**, 1255 (2002).

[438] R. B. Choi and E. McLaughlin. Effect of a phase transition on the solubility of a solid. *AIChE J.* **29**, 150 (1983).

[439] F. W. Seyer. Mutual solubilities of hydrocarbons II. The freezing point curves of dotriacontane in dodecane, decane, octane, hexane, cyclohexane and benzene. *J. Am. Chem. Soc.* **60**, 827 (1938).

[440] J. B. Ott and J. R. Goates. Solid + liquid phase equilibria in binary mixtures containing benzene, a cycloalkane, an n-alkane, or tetrachloromethane. An equation for representing solid + liquid phase equilibria. *J. Chem. Thermodyn.* **15**, 267 (1983).

[441] P. M. Chogomu, J. Dellacherie, and D. Balesdent. Solubility of normal paraffin hydrocarbons (C20 to C24) and some of their binary mixtures (C22 + C24) and (C23 + C24) in ethylbenzene. *J. Chem. Thermodyn.* **21**, 925 (1989).

[442] B. Calvo and E. A. Cepeda. Solubilities of stearic acid in organic solvents and in azeotropic solvent mixtures. *J. Chem. Eng. Data* **53**, 628 (2008).

[443] C. Fiege, R. Joh, M. Petri, and J. Gmehling. Solid-liquid equilibria for different heptanones with benzene, cyclohexane, and ethanol. *J. Chem. Eng. Data* **41**, 1431 (1996).

[444] E. Moreno, R. Cordobilla, T. Calvet, M. A. Cuevas-Diarte, G. Gbabode, P. Negrier, D. Mondieig, and H. A. J. Oonk. Polymorphism of even saturated carboxylic acids from n-decanoic to n-eicosanoic acid. *New J. Chem.* **31**, 947 (2007).