PREDICTIVE MODELING OF THERMODYNAMIC PROPERTIES AND PHASE EQUILIBRIUM FOR PRAZIQUANTEL USING SAFT-γ MIE AND PC-SAFT FRAMEWORK

ANTONIO CAVALCANTE DE LIMA NETO

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Química, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Química.

Orientadores: Frederico Wanderley Tavares

Iuri Soter Viana Segtovich

Rio de Janeiro

Março, 2024

PREDICTIVE MODELING OF THERMODYNAMIC PROPERTIES AND PHASE EQUILIBRIUM FOR PRAZIQUANTEL USING SAFT-γ MIE AND PC-SAFT FRAMEWORK

Antonio Cavalcante de Lima Neto

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA QUÍMICA.

Orientadores: Frederico Wanderley Tavares

Iuri Soter Viana Segtovich

Aprovada por: Prof. Argimiro Resende Secchi

Prof. Marcelo Castier

RIO DE JANEIRO, RJ - BRASIL

MARÇO DE 2024

Cavalcante, Antonio

Predictive Modeling of Thermodynamic Properties for Praziquantel Using SAFT-γ Mie and Pc-SAFT Framework / Antonio Cavalcante de Lima Neto. – Rio de Janeiro: UFRJ/COPPE, 2024.

XXIV, 122 p.: il.; 29,7 cm.

Orientadores: Frederico Wanderley Tavares
Iuri Soter Viana Segtovich

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de Engenharia Química, 2024.

Referências Bibliográficas: p. 75-80.

1. Thermodynamics. 2. Phase Equilibrium. 3. Praziquantel. I. Tavares, Frederico Wanderley *et al*. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química. III. Título.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.).

MODELAGEM PREDITIVA DAS PROPRIEDADES TERMODINÂMICAS E EQUILIBRIO DE FASES DO PRAZIQUANTEL USANDO OS MODELOS SAFT-γ MIE E PC-SAFT

Antonio Cavalcante de Lima Neto

Março/2024

Orientadoes: Frederico Wanderley Tavares
Iuri Soter Viana Segtovich

Programa: Engenharia Química

Esta dissertação apresenta um estudo sobre as propriedades termodinâmicas de compostos, com especial enfoque em uma substância farmacêutica, praziquantel (PZQ) e um solvente orgânico amplamente utilizado, N-Metil-2-pirrolidona (NMP), utilizando os modelos SAFT-γ Mie e PC-SAFT. O objetivo principal era aumentar a acurácia preditiva desses modelos para comportamentos de misturas complexas, ajustando novos parâmetros, especificamente dois novos parâmetros de grupo para amidas cíclicas (cNcCO para NMP e PZQ, cNCO para PZQ) no contexto de SAFT-γ Mie, contabilizando 8 parâmetros de grupo *like* e 34 parâmetros de grupo *unlike*. Além disso, foram ajustados 18 parâmetros de interação binária (k_{ij}) entre PZQ e solventes no contexto do PC-SAFT. O estudo revelou que a SAFT-γ Mie, com um desvio relativo absoluto médio (AARD) de 5,72% e máximo de 27,38%, superou o modelo PC-SAFT, que teve AARD de 9,90% e máximo de 28,15%, na previsão da solubilidade do praziquantel em diversos solventes. Isto destaca as capacidades preditivas superiores dele e seu potencial para aplicações mais amplas na indústria farmacêutica. No entanto, a pesquisa encontrou limitações devido à variabilidade e incerteza nos dados experimentais, incluindo a ausência de réplica e variância, o que prejudicou a análise estatística dos resultados. Este trabalho abre um precedente para futuras investigações que visam superar as limitações atuais e expandir as aplicações destes modelos, particularmente na desafiadora área da purificação de enantiômeros, avançando assim no campo das ciências farmacêuticas.

Abstract of Qualifying Exam presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

PREDICTIVE MODELING OF THERMODYNAMIC PROPERTIES AND PHASE EQUILIBRIUM FOR PRAZIQUANTEL USING SAFT-γ MIE AND PC-SAFT FRAMEWORK

Antonio Cavalcante de Lima Neto

March/2024

Advisors: Frederico Wanderley Tavares

Iuri Soter Viana Segtovich

Department: Chemical Engineering

This dissertation presents a comprehensive study on the thermodynamic properties of compounds, with a special focus on a pharmaceutical substance, Praziquantel (PZQ) and a widely used organic solvent, N-Methyl-2-pyrrolidone (NMP), using the SAFT-γ Mie and PC-SAFT models. The primary aim was to enhance the predictive accuracy of these models for complex mixture behaviors by adjusting new parameters, specifically two new group parameters for cyclic amides (cNcCO for NMP and PZQ, cNCO for PZQ) in the context of SAFT-γ Mie, accounting for 8 like group parameters and 34 unlike group parameters. Moreover, it was adjusted 18 binary interaction parameters (k_{ij}) between PZQ and solvents in the context of PC-SAFT. The study revealed that the SAFT-γ Mie model, with an average absolute relative deviation (AARD) of 5.72% and maximum of 27.38%, outperformed the PC-SAFT model, which had an AARD of 9.90% and maximum of 28.15%, in predicting the solubility of praziquantel in various solvents. This highlights the SAFT-γ Mie model's superior predictive capabilities and its potential for broader applications in the pharmaceutical industry. However, the research encountered limitations due to the variability and uncertainty in experimental data, including the absence of replica and variance, which impaired the statistical analysis of the findings. This work sets a precedent for future investigations aimed at overcoming the current limitations and expanding the applications of these models, particularly in the challenging area of enantiomer purification, thereby advancing the field of pharmaceutical sciences.

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$k_{2_{ij}}$	PC-SAFT association binary interaction parameter
x_{s_k}	The fraction of segments of a specific type k
a_{0i}	PC-SAFT constants
a_i, b_i	PC-SAFT functions
C_p	Molar heat capacity at constant pressure
I_1, I_2	PC-SAFT integrals abbreviations
k_b	Boltzmann constant
k_{ij}	PC-SAFT binary interaction parameter
$\bar{m}$	Mean segment number in the system
r_{AB}^c	Cut-off range of the interaction between sites A and B
S_2	Reduced well width
S_{ji}	Multiplicity of site j within molecule i
X_{ji}	Fraction of molecules of type i not bonded to site j
Δ_{ijklab}	Strength of interaction due to the association of the site ‘a’ of the group ‘k’ in the molecule ‘i’ with the site ‘b’ of the group ‘l’ in the molecule ‘j’.
ε_{AB}	Association energy between sites A and B
F_{klab}	Mayer-f function of the square-well association potential due to the association of the site ‘a’ of the group ‘k’ with the site ‘b’ of the group ‘l’.
G_{ij}	NRTL’s coefficient
I_{klab}	Association integral for a Lennard-Jones monomer due to the association of the site ‘a’ of the group ‘k’ with the site ‘b’ of the group ‘l’.
K_{klab}	Bonding volume due to the association of the site ‘a’ of the group ‘k’ with the site ‘b’ of the group ‘l’.
N_{ST_k}	Distinct sites within group ‘k’

W_i	Number of moles, species i
X_{ika}	Proportion of unbounded molecules belonging to component ‘i’ with respect to group ‘k’ through the site ‘a’
a_1^s	SAFT-γ Mie function
a_{pkl}	Pairwise interactions between groups ‘k’ and ‘l’
c_i, c_n	SAFT-γ Mie coefficients
f_m	SAFT-γ Mie function
g_{ij}	Wilson’s interaction energy between components i and j
k_{ij}	Cubic equations of state interaction parameter
k_m	SAFT-γ Mie coefficients
l_{ij}	Cubic equations of state interaction parameter
n_i	Number of moles, species i
n_{ka}	Quantity of ‘a’-type sites within group ‘k’
q_1	Eos/GE constant
w_i	Mole fraction of component i
x_i	Mole fraction of component i
x_{ji}	Local mole fraction of molecules j around molecules i
y_i	Mole fraction of component i
z_{ki}	Fraction of group type k in the molecular structure of component i
Δ_{lkji}	Strength of interaction due to the association of site l of molecule k with site j of the molecule i
α_{ij}	Non-randomness parameter of molecules of type i around a molecule of type j
$\gamma_{c_{ii}}$	SAFT-γ Mie correction factor
ε^{lkji}	Association energy between the sites l and j per molecule k and i,

κ^{lkji}	Effective association volume between the sites l and j per molecule k and i,
λ^a	Attractive exponent of the Mie intermolecular potential
λ^r	Repulsive exponent of the Mie intermolecular potential
v_k^*	Number of identical segments comprising group k
v_{ki}	Number of groups k on component i
ρ_s	Total number of spherical segment density of the Mie fluid
ϕ_{mn}	SAFT-γ Mie coefficients
μ	Chemical potential
A	Margules' Adjustable parameters
ã	Reduced Helmholtz energy
C	Function of the Mie intermolecular potential exponents
g	Radial pair distribution function
h	Molar enthalpy
N	Total number of molecules
R	Universal gas constant
T	Temperature
Z	Coordination factor
ζ	Reduced density
η	Packing fraction
ϕ(r)	Potential energy–distance function
Ω	Cubic equations of state constant or Parameters to be adjusted
A	Helmholtz energy
B	SAFT-γ Mie function
C	Eos/GE constant

G	Gibbs energy
H	Enthalpy
I,J	SAFT-γ Mie functions
K	Isothermal compressibility
Nc	Number of components
Ng	Number of groups
Np	Number of phases
P	Pressure
S	Shape factor
U	Internal energy
VLE	Vapor-liquid equilibrium data
a, b	Cubic equations of state parameters
d	Temperature-dependent segment diameter
f	Fugacity
g	Molar Gibbs free energy
m	Number of segments per chain
r	Intermolecular distance
s	Molar entropy
v	Molar volume
Λ	Wilson's adjustable parameter
α	Cubic equations of state function
γ	Activity coefficient
ε	Dispersion energy parameter
ρ	Total number density of molecules

ς	Cubic equations of state constant
σ	Segment diameter
τ	NRTL's adjustable parameter
φ	Fugacity coefficient
χ	Correction factor
ψ	Cubic equations of state constant
ω	Acentric factor or Empirical weight
ϵ	Cubic equations of state constant

LIST OF SUPERSCRIPTS & SUBSCRIPTS

c	Critical property
α, β, p	Identify phases
V	Vapor phase
L	Liquid phase
sat	Saturation
s	Solid phase
fus	Fusion
t	Triple point
rac	Racemic mixture
m	Melting
k	Iteration step
res	Residual
hc	Hard-chain
$assoc$	Association
$disp$	Dispersive
$mono$	Monomer
eff	Effective
exp	Experimental
$calc$	Calculated
obj	Objective

LIST OF ACRONYMS

WHO	World health organization
PZQ	Praziquantel
vdW	Van der Waals
NRTL	Non-random two-liquid
API	Active pharmaceutical ingredients
PC	Perturbed chain
EoS	Equation of State
SAFT	Statistical associating fluid theory
SRK	Soave/redlich/kwong
RK	Redlich/kwong
PR	Peng/robinson
GE	Excess Gibbs free energy model
HV	Huron and Vidal
LJ	Lennard-Jones potential
SW	Square-well potential
TPD	Tangent plane distance
RDF	Radial distribution function
AARD	Average absolute relative deviation
AD	Automatic differentiation
JAX	Just after execution
JIT	Just-in-time compiler
NMP	N-methyl-2-pyrrolidone
VLE	Vapor-liquid equilibrium

VLLE	Vapor-liquid-liquid equilibrium
SLE	Solid-liquid equilibrium
NIST	National Institute of Standards and Technology

1. Introduction

Schistosomiasis, a parasitic disease caused by a type of flatworm known as schistosomes, poses a significant public health challenge, particularly in developing countries such as Brazil, where access to clean water is often limited. The World Health Organization (WHO) estimated that over 250 million people required treatment for schistosomiasis in 2021, yet only 75 million received the necessary care. Despite its high prevalence and substantial impact on public health, schistosomiasis is classified as one of the Neglected Tropical Diseases (NTDs), which predominantly affect impoverished and marginalized populations and suffer from inadequate attention and funding for research and development due to their low financial return.

Praziquantel, an anthelmintic drug, is widely used for its effectiveness against a broad range of parasitic worms, including schistosomes. It is recognized for its safety and tolerability, presenting minimal side effects (RASO *et al.*, 2004). However, a notable limitation of praziquantel is its chirality; it contains a chiral carbon, which can lead to issues regarding the drug's efficacy and safety. A chiral carbon atom in a molecule is bonded to four distinct groups, resulting in two mirror-image forms known as enantiomers. These enantiomers are chemically identical except for their mirror-image configuration, which can lead to significantly different biological activities despite their thermodynamic properties being remarkably similar.

Since praziquantel possesses a chiral carbon, the drug has two enantiomers: one is biologically active, contributing to the therapeutic effects, while the other is inactive, potentially contributing only to adverse effects such as nausea, vomiting, and dizziness. Pure enantiomers would be more effective and have fewer adverse effects than the racemic mixture. However, the production and purification of pure enantiomers are expensive and time-consuming, potentially limiting their availability.

The optimization of the crystallization process in pharmaceutical development critically relies on understanding and modeling the thermodynamic properties of Active Pharmaceutical Ingredients (APIs). Studies by SHEIKHOLESLAMZADEH and ROHANI (2012), BOUILLOT *et al.* (2011), DOHRN *et al.*(2021), BOCHMANN, NEUMANN, *et al.*, (2019) and VALAVI *et al.* (2017) have highlighted the importance of solubility prediction, solvent selection, and the understanding of crystallization from mixed solvents in enhancing this process. Utilizing thermodynamic models such as

UNIFAC, NRTL-SAC, the Hansen Flory Huggins model, and notably, the PC-SAFT model, these studies enable the selection of optimal solvents or solvent mixtures that significantly improve API solubility, yield, and purity. The PC-SAFT model, in particular, is showcased in the work of DOHRN *et al.* (2021), demonstrating its utility in predicting API solubility in solvent mixtures and avoiding unwanted phase separation. Accurate solubility predictions allow for the fine-tuning of crystallization conditions, ensuring efficient API recovery and enhancing stability and bioavailability through improved crystallization techniques. This synthesis of work underscores the pivotal role of thermodynamic modeling, including the PC-SAFT model, in advancing the efficiency and quality of pharmaceutical crystallization processes, ultimately influencing the performance of the final drug products.

Advancements in theoretical methods for predicting solubility and other thermodynamic properties have significantly expedited the development process and reduced costs by diminishing reliance on experimental assays. Among these methods, Equations of State (EoS) stand out as mathematical models that describe the thermodynamic properties of fluids based on the fluid's composition and state. These EoS models offer a comprehensive representation of fluid behavior across a broad range of temperatures, pressures, and compositions, making them invaluable tools in the chemical and pharmaceutical industries for optimizing separation processes, including the purification of enantiomers from chiral compounds.

EoS models excel over activity coefficient-based models in their ability to predict phase behavior and thermodynamic properties over extensive conditions. Unlike activity coefficient models, which are typically limited to describing deviations from ideal behavior in mixtures at low pressures, EoS models can accurately represent the physical state of substances under a much wider array of conditions. This broad applicability stems from their fundamental basis in the principles of thermodynamics and statistical mechanics, allowing them to capture the critical interactions and phase equilibria of complex systems. Furthermore, these models can be integrated into process simulation software to optimize the design and operation of separation units, thereby enhancing efficiency and reducing experimental trial and error.

The study by TSIVINTZELIS *et al.* (2009) exemplifies the power of EoS models for extrapolating beyond the range of available experimental data, a capability of particular importance in the pharmaceutical industry. In their work, they developed a

methodology for modeling the phase equilibria of complex chemical mixtures, specifically pharmaceutical-solvent mixtures that exhibit complex hydrogen bonding behavior. They applied the nonrandom hydrogen bonding (NRHB) EoS to accurately correlate the solubilities of three pharmaceuticals: acetanilide, phenacetin, and paracetamol. The accuracy of their correlations and the successful evaluation against predictions from the COSMO-RS model underscore the potential of EoS models to predict the solubility and stability of new drug compounds under a wide range of conditions. This approach not only aids in the successful formulation and process design by reducing reliance on extensive experimental assays but also demonstrates the exceptional ability of EoS models to perform extrapolations where direct experimental data might be scarce or unavailable, thus underlining the models' excellence in extrapolation.

One of the most common separation methods for enantiomers is chromatography, which involves separating a mixture of chiral compounds based on their differential interactions with a stationary phase. Thermodynamic models, not limited to Equations of State, can be used to optimize the chromatographic separation process by predicting the thermodynamic behavior of chiral compounds in the mobile phases. This approach is exemplified in the work by CASTELLS and CARR (2000), which studied the thermodynamics of chiral separations using cellulosetris(3,5-dimethylphenylcarbamate) coated on porous zirconia stationary phases over a range of temperatures. By analyzing differences in enthalpy and entropy of transfer from the mobile to the stationary phase, their study provides valuable insights into the driving forces for chiral recognition, showing that while most separations are enthalpy-driven, others are dominated by entropy. This comprehensive thermodynamic perspective enables the design of chromatographic columns that maximize separation efficiency and minimize the amount of mobile phase required for separation, illustrating the intricate relationship between thermodynamic interactions and chromatographic performance.

Furthermore, thermodynamic models, can be utilized to optimize the conditions for the recrystallization of enantiomers. Recrystallization, a prevalent method for purifying chiral compounds, involves dissolving the racemic mixture in a solvent and inducing the formation of crystals of one enantiomer, while the other remains in solution. The study by GLYNN *et al.* (1990) delves into equilibrium in binary solid-solution aqueous-solution systems after dissolution, precipitation, or recrystallization

processes, based on the composition and relative proportion of the initial phases. Their exploration of stoichiometric saturation offers insights into understanding the thermodynamic equilibrium points and reaction paths during recrystallization, which is crucial for optimizing solvent composition, temperature, and other process parameters. Although Glynn et al. focused on solid-solution systems of minerals, the thermodynamic principles they elucidated, particularly regarding equilibrium, could be similarly applied using EoS models, therefore, it could provide a more detailed and predictive framework for adjusting recrystallization parameters to enhance the purity and yield of the desired enantiomer.

A class of EoS extensively utilized for drug optimization is founded on the group contribution approach. This methodology posits that molecules can be conceptualized in terms of functional groups, allowing for the deduction that the properties of any system can be derived from the contributions of its constituent groups to the system's thermodynamic properties. The work by PANKOW and ASHER (2008) introduces the SIMPOL.1 group contribution method, developed for predicting the vapor pressure and enthalpy of vaporization of organic compounds as functions of temperature. This method assumes the vapor pressure of a compound can be predicted from the sum of contributions by its structural groups, incorporating a wide range of molecular functionalities. The group contribution approach is pivotal for extrapolating properties, as it enables predictions for molecules not yet synthesized or lacking experimental data, by leveraging known group properties rather than specific molecular configurations.

This capability of group contribution methods to extrapolate beyond the confines of experimentally characterized compounds underscores their utility in drug development. By using functional groups as the basis for prediction, these models can forecast the thermodynamic behavior of novel compounds, facilitating the exploration of a vast chemical space without the immediate need for laboratory synthesis and characterization. Such a predictive tool is invaluable for identifying promising drug candidates by assessing their solubility, stability, and other critical properties in silico, significantly accelerating the drug development process and reducing reliance on extensive empirical testing.

The SAFT-γ Mie model, devised by PAPAIOANNOU *et al.* (2014), is a powerful group contribution EoS used for modeling the thermodynamic properties of a wide range of fluids, including pharmaceuticals. This equation of state utilizes the mie

potential, a versatile adaptation of the Lennard-Jones potential, to capture the complex intermolecular interactions within fluids. The flexibility of the mie potential allows for the accurate prediction of properties across various interaction ranges and intensities. Specifically, FEBRA *et al.* (2021) applied this approach to predict the solubility of pharmaceutical compounds, demonstrating its utility in drug development for predicting the behavior of compounds even before their synthesis. This capability significantly aids in the optimization of drug formulations and manufacturing processes.

1.1. Objective

The main aim of this dissertation is to delve into the realm of thermodynamic modeling, with a particular focus on EoS, enriching both theoretical understanding and practical computational skills for future applications This study is particularly focused on employing the SAFT-γ Mie and PC-SAFT EoS models to investigate the solubility behaviors of praziquantel in diverse solvents and to assess the extrapolation capabilities of these models. This investigation marks a step toward elucidating the separation process of praziquantel's racemic mixture, aiming to refine its optimization. A critical component of this endeavor involves compiling a detailed database capturing the thermodynamic properties of each enantiomer. This database will facilitate the estimation and differentiation of parameters associated with the group that encompasses the chiral carbon of each enantiomer. While this research does not extend to the crystallization improvement of the drug or the differentiation of enantiomers, it establishes an essential groundwork for future exploration in the purification of enantiomers.

2. Thermodynamic Background

2.1. General Overview of Thermodynamic Models History

Process simulation is fundamental to numerous industrial applications. It hinges on a deep understanding of both physical and chemical phenomena, with thermodynamic modeling at its core. This modeling is crucial for analyzing and predicting system behaviors under different conditions. It is vital in driving product innovation, expanding production capacities, and improving quality and profitability.

In the field of chemical engineering, a profound understanding of the thermodynamic properties of complex molecules is essential for numerous processes. As a result, older and more limited models, such as those grounded in just observed data,

are becoming increasingly obsolete. However, due to their simplicity and ease of implementation, these classic models continue to find various applications in the industry.

The forthcoming sections are dedicated to providing a general overview of some excess Gibbs free energy models (GE) and equations of state (EoS). This exploration will span from Random-Mixing Activity Coefficient Models to more sophisticated ones, such as the SAFT (Statistical Associating Fluid Theory) family.

2.1.1. Random-Mixing Activity Coefficient Models

Activity coefficient models in chemical engineering describe the properties of condensed phases. These models are extensively used to detail how various components in a mixture interact at the molecular level, influencing properties such as volatility and solubility.

To understand a chemical system, it is interesting to comprehend the diverse effects on entropy and enthalpy resulting from the mixing of components. This understanding is complex, which is why activity coefficients are often modeled using excess Gibbs free energy, which is a function of both entropy and enthalpy. This approach is straightforward and simpler.

These models exhibit a high dependence on the composition of their components, a moderate dependence on temperature, and are generally independent of pressure. Notably, it is a justified assumption, considering that the behavior of solids and liquids are, in general, minimally affected by pressure changes, unless they are extremely high.

The Margules equations (MARGULES, 1895) represent a historical milestone as the first significant model for the activity coefficient with relevance in thermodynamics. Despite being developed nearly two centuries ago, they continue to be widely employed in the industry. ANICETO *et al.* (2012) explored the applicability of the Margules model among others for ion exchange equilibrium. ZARGARZADEH and ELLIOTT (2019) used the Margules activity model for solid-liquid equilibrium, calculating parameters from experimental phase equilibrium data for various systems. VATANI *et al.* (2012) applied Two-Suffix Margules models for ternary extraction ionic

liquid systems. This testifies to the robustness and ongoing relevance of these equations, even in the face of scientific advancements.

There are two main variations of the Margules equations: the single-parameter and the two-parameter version. Both forms are described for a binary mixture as follows:

$$g^E = g - g^{id} = \mathrm{A}x_1x_2 \quad 2\text{-}1$$

$$g^E = (\mathrm{A}_{21}x_1 + \mathrm{A}_{12}x_2)x_1x_2 \quad 2\text{-}2$$

Where g^E is the deviation of the Gibbs free energy of a solution (g) from that of an ideal solution (g^{id}), A is a matrix comprising the adjustable parameters and x are the mole compositions. An ideal solution is a theoretical construct, which refers to a solution where the intermolecular forces between unlike molecules are the same as those between like molecules. This implies no volume or temperature changes, when mixing at constant Pressure and adiabatically.

The Margules equations were developed to represent experimental data on the non-ideal behavior of mixtures using second-order polynomials. This approach is practical in many situations because the g^E often exhibits a parabolic behavior in many systems.

Another important model in thermodynamics is the Van Laar equation. It's built on the van der Waals equation, upon using the conventional vdW one fluid mixing rules. However, it uses two parameters that don't have a clear physical meaning, just as Margules. These parameters are used to help fit the model to real-world data, which means the Van Laar equation is also considered an empirical model. The Van Laar equation for a binary mixture is represented as follows:

$$g^E = x_1x_2\left[\frac{x_1}{\mathrm{A}^{21}} + \frac{x_2}{\mathrm{A}^{12}}\right]^{-1} \quad 2\text{-}3$$

It can be noticed that when the constants are equal, the equation simplifies to the single-parameter Margules equation. The Van Laar equation predicts the behavior of

mixtures similarly to the Margules equations. It works best for mixtures where the different substances are quite similar to each other in terms of their molecular properties.

2.1.2. Local Composition Activity Coefficient Models

One notable limitation of random-mixing models lies in their inherent assumption of randomness within mixtures. This assumption does not fully encapsulate the reality due to the presence of molecular forces, which introduce a degree of non-randomness. To address this, models that consider the oriented short-range nature of intermolecular forces within the mixture—termed Local Composition Models—have been developed to enhance the predictive accuracy of properties of mixtures. These models provide a more realistic depiction of phase behavior in complex mixtures compared to their random-mixing counterparts. Among the most prevalent Local Composition Models in use are the Wilson, NRTL, and UNIQUAC equations.

WILSON (1964) was the first one to assume a model that the local composition differs from the bulk composition in the liquid mixture. He derived a relation between the local mole fraction of molecules j around molecules i:

$$\frac{x_{ji}}{x_{ii}} = \frac{x_j e^{-\frac{g_{ji}}{\mathrm{RT}}}}{x_i e^{-\frac{g_{ii}}{\mathrm{RT}}}} \qquad 2\text{-}4$$

Where g_{ij} is the interaction energy between components i and j and T is the temperature. The fraction $e^{\frac{g_{ji}-g_{ii}}{RT}}$ is related to a weighting factor, relative to the overall compositions. If this fraction equals 1, the interaction energy is the same and it will mean that the solution is random. For multicomponent systems, the Wilson equations can be written as:

$$\frac{g^E}{RT} = -\sum_j x_j \ln\left(\sum_i x_i \Lambda_{ji}\right) \qquad 2\text{-}5$$

Where Λ_{ji} represents the matrix of adjustable binary parameters. It relates to the interaction energy between molecules as follows:

$$\Lambda_{ji} = \frac{v_i}{v_j} e^{-\frac{g_{ji}-g_{ii}}{RT}} \quad \text{2-6}$$

Where $\Lambda_{ii} = 1$, $\Lambda_{ji} = \Lambda_{ji}$ and v_j is the molar volume of molecule j. Notice that the specification of liquid molar volumes is not strictly essential because the Λ parameters for each pair of components within a multicomponent mixture are determined through experimental data. For instance, in a ternary system, the pairs would be Λ_{12}, $\Lambda_{13}$and Λ_{23}. This model exhibits a limitation in its inability to accurately characterize systems that exhibit partial miscibility. This constraint presents complexities beyond the model's descriptive capacity.

The NRTL (Non-Random Two-Liquid) model, introduced by RENON and PRAUSNITZ in 1968, extends the two-liquid theory of Scott and incorporates a non-randomness parameter that significantly improves the prediction of phase behavior in complex mixtures and, unlike Wilson's model, applies to partially miscible systems. The NRTL equations for the description of multicomponent systems are:

$$\frac{g^E}{\mathrm{RT}} = \sum_i \frac{x_i \sum_j x_j \tau_{ji} G_{ji}}{\sum_j x_j G_{ji}} \quad \text{2-7}$$

$$G_{ji} = e^{-\alpha_{ji}\tau_{ji}} \quad \text{2-8}$$

Where the adjustable parameters are τ and α. In this case, $\tau_{ii} = 0$ and α is the non-randomness parameter. α generally assumes values between 0.2 and 0.47, and it could be calculated (if not estimated) as 2/Z, where Z is the coordination factor. If α_{ji} is zero, while $\tau_{ji}$and τ_{ij} are not, the mixture becomes completely random, and the original equation is reduced to the Margules equation. Whenever α is above 0.426, the model will predict partially miscible liquids.

2.1.3. Cubic EoS

Cubic equations of state (EoS) stand out for their cubic correlation with molar volume, relating pressure, molar or specific volume, temperature, and fluid composition at equilibrium. Their ability to simultaneously represent both liquid and vapor phases

within a singular mathematical structure is a key advantage over activity coefficient models, which excel in characterizing only one phase. However, it's important to note that cubic EoS generally exhibit less precision in describing liquid phase behaviors compared to activity coefficient models.

The advent of the VAN DER WAALS (vdW) (1873) equation as the first practical cubic EoS was a breakthrough. It incorporated the concepts of finite molecular size and intermolecular forces into the equation of state, thus modifying the ideal gas law to better reflect real gas behaviors. This innovation laid the groundwork for the development of more types of EoS.

Following the vdW equation, other significant cubic EoS emerged. These include the REDLICH/KWONG (RK) (1948) equation, the SOAVE/REDLICH/KWONG (SRK) (1972) equation, and the PENG/ROBINSON (PR) (1976) equation. These various cubic equations can be generalized into a single form of a cubic equation of state:

$$P = \frac{\mathrm{RT}}{v - b} - \frac{a(\mathrm{T})}{(v + \epsilon b)(v + \varsigma b)} \qquad 2\text{-}9$$

In this equation, P represents the pressure, T the temperature, v the molar volume and R is the universal gas constant. The parameters a and b are specific to each gas. Here, a reflects the strength of intermolecular forces, which can be dependent of temperature in some models. The parameter b represents the volume occupied by the gas molecules. The terms ϵ and ς are additional numbers used to differentiate the equations.

The parameters b and $a(T)$ can be related to the critical pressure (P_c) and critical temperature (T_c) of a substance:

$$b = \Omega \frac{RT_c}{P_c} \qquad 2\text{-}10$$

$$a(T) = \psi \frac{\alpha(T, \omega) R^2 T_c^2}{P_c} \qquad 2\text{-}11$$

The terms ψ and Ω are also additional numbers used to differentiate the equations. The function $\alpha(T,\omega)$ is an empirical expression, specific to each EoS. ω is the acentric factor, a parameter specific to a given chemical species. It represents a measure of the non-sphericity of the molecule.

To extend these cubic EoS to mixtures, empirical mixing rules are used to link mixture parameters to pure-species parameters. A common approach to adapt cubic EoS for mixtures employs the van der Waals one-fluid (vdW1f) mixing rules. These rules involve quadratic mixing rules using an arithmetic combination for parameter b and a geometric combination for parameter a, as follows:

$$b = \sum_i \sum_j x_i x_j \frac{(b_i + b_j)}{2}(1 - l_{ij}) \qquad 2\text{-}12$$

$$a = \sum_i \sum_j x_i x_j \sqrt{a_i a_j}(1 - k_{ij}) \qquad 2\text{-}13$$

Where l_{ij} and k_{ij} are the interaction parameters, typically fitted to phase equilibrium data. While l_{ij} is often neglected (being set to zero), k_{ij} is particularly important for systems involving polar or complex molecules, where standard mixing rules may not suffice to capture the subtleties of intermolecular interactions. By adjusting these parameters, a cubic EoS can be fine-tuned to represent the behavior of a wide range of mixtures in various conditions.

2.1.4. EoS/GE mixing rule

Cubic equations of state have limitations in adequately describing the phase equilibria of mixtures with the presence of highly polar or associative molecules. While they are effective in predicting the behavior of nonpolar mixtures at both low and high pressures, a standard cubic EoS struggle to accurately capture the complex interactions between polar or associative molecules.

On the other hand, activity coefficient models perform well for polar liquid mixtures but are unable to describe complete phase envelopes. Moreover, activity coefficient models tend to lose accuracy at high pressure. This is because the most commonly used models, such as the ones discussed here, do not incorporate pressure as

a direct variable in their formulations. They are primarily developed for low-pressure conditions, where pressure has minimal impact on the activities.

To address these limitations, it can be used an approach that involves the integration of Gibbs excess energy models with cubic EoS through specific mixing rules. This hybrid method seeks to combine the strengths of both the excess Gibbs energy model and cubic EoS, offering a more thorough description of complex mixtures. The EoS/GE approach combines the proficiency of the Gibbs excess energy model in handling polar liquids with the efficacy of cubic EoS in modeling nonpolar mixtures under high-pressure conditions.

In 1979, HURON and VIDAL proposed a novel approach for determining the attractive term in the cubic EoS. This method involves aligning the excess Gibbs free energy of the EoS at the infinite pressure limit with the activity coefficient of the local composition models.

The mixing rule introduced in their framework is formulated as follows:

$$a = \left(\sum_i x_i b_i\right)\left(\sum_i x_i \frac{a_i}{b_i} - \frac{\bar{g}_\infty^E}{C}\right) \qquad 2\text{-}14$$

In this equation, $\bar{g}_\infty^E$ denotes the molar excess Gibbs free energy at the infinite pressure limit, and C is a dimensionless numerical constant characteristic of the specific EoS used.

In addition to the approach for calculating the attractive term at the infinite pressure limit, some methods use zero pressure as a reference. This approach was first proposed by MOLLERUP in 1986 and subsequently refined and applied to more practical systems by MICHELSEN in 1990. Due to its resemblance to the mixing rules developed by Huron and Vidal (HV), Michelsen's mixing rule is often referred to as the modified Huron-Vidal mixing rule. The formulation is as follows:

$$\alpha = \frac{1}{q_1}\left[\frac{\bar{g}_0^E}{RT} + \sum_i x_i \ln\left(\frac{b}{b_i}\right)\right] + \sum_i x_i \alpha_i \qquad 2\text{-}15$$

$$\alpha = \frac{a}{bRT} \qquad 2\text{-}16$$

Where g_0^E denotes the molar excess Gibbs free energy at the zero pressure limit and q_1 is also a dimensionless numerical constant characteristic of the specific EoS used.

2.1.5. The statistical associating fluid theory

In the original Statistical Associating Fluid Theory (SAFT), molecules are represented as chains of spheres that are tangentially connected. Each chain in this representation corresponds to a component within a mixture. It's important to note that within a given chain, the spheres share identical properties, such as matching diameters and energy of interaction. SAFT, like many other equations of state in engineering, is rooted in perturbation theory.

While several established engineering equations of state employ the hard sphere as a reference fluid, the CHAPMAN *et al.* (1990) version of SAFT takes a different approach by using a reference fluid that interacts via the Lennard-Jones potential (LJ) (JONES, 1924). Additionally, Chapman SAFT integrates considerations for chain length and molecular association. This involves initially computing the Helmholtz energy for the fluid mixture, considering it as comprised of spherical Lennard-Jones segments, before any bonding or association phenomena occur.

$$\phi^{\mathrm{LJ}}(\mathrm{r}) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \qquad 2\text{-}17$$

The parameter ε signifies the potential's minimum energy, often referred to as the potential well. This ε value denotes the strength of attractive interactions between segments within the potential. On the other hand, σ represents the distance at which the potential becomes zero, indicating the point where attractive forces balance out repulsive ones. Essentially, σ corresponds to the temperature-independent diameter of the spherical segments.

Originally devised to model the cohesive forces observed in liquid noble gases like argon, the LJ potential behaves distinctly based on the radius of the interacting particles. For smaller radii, the first term of the potential exerts a more substantial influence, causing an elevation in its overall value. In cases of $2^{\frac{1}{6}}\sigma$ of radii, the potential

tends towards its minimum value. However, as the radius increases, the potential gradually tends towards zero.

The Lennard-Jones potential stands as a cornerstone in classical many-body molecular simulations due to its widespread applicability. Its versatility allows it to describe interactions not just between atoms, but also among molecules and coarse-grained models representing various entities like organic molecules, proteins, and sometimes even larger particles such as nano-colloids.

Despite its prevalent use across diverse systems, there exists a notable absence of empirical evidence that unequivocally establishes the superiority of the LJ potential over other plausible alternatives for these specific systems. Nevertheless, in the realm of testing new simulation techniques, the LJ 12-6 potential consistently emerges as the initial model of choice. This trend underscores its reputation as a satisfactorily versatile and adaptable potential, making it a go-to option for a wide array of simulation scenarios (WANG *et al.*, 2020).

In addition to considering the attractive and repulsive forces described by the LJ potential, the SAFT EoS incorporates the influence of chain formation on the Helmholtz energy. This model employs multi-segmented chain molecules formed by enforced bonding, mimicking covalent-like bonds. Each LJ sphere can bond with up to two sites, compelled to form chains. Consequently, this approach allows for a more accurate representation of chain-like molecules compared to classical models, where molecules are typically depicted in a spherical manner. Importantly, in this theory, molecule spheres cannot bond together to form ring-like structures. For instance, representing cyclic molecules such as benzene involves utilizing the aliphatic chain of LJ spheres to better approximate the properties of the ring-like molecule.

One of the most distinctive aspects of the SAFT EoS lies in its ability to approximate hydrogen bonding among molecules, i.e., the short-range and highly orientation-dependent interactions between sites. This feature significantly enhances the model's ability to represent highly polar molecules like alcohols, organic acids, and water. To simulate the strength of association of these interactions, the SAFT EoS employs a square-well potential (SW) between associating sites. The SW potential is used to describe interactions between particles with a finite range of attraction beyond which there is no interaction. The square-well potential is characterized by a potential

energy that is zero when particles are beyond a certain distance from each other, infinite when they are very close (to model the hard-core repulsion), and a constant negative value (indicating attraction) when they are within a specific intermediate range. Spherical segments can possess multiple tangential positioned associating sites, which center the SW potential. Crucially, these spherical entities can only form an association bond when both the distance and orientation between them align favorably.

$$\phi_{AB}^{\mathrm{HB}}(\mathrm{r}) = \begin{cases} \infty, \mathrm{r} < 0 \\ -\varepsilon_{\mathrm{AB}}, 0 < \mathrm{r} < \mathrm{r}_{AB}^{\mathrm{c}} \\ 0, \mathrm{r} > \mathrm{r}_{AB}^{\mathrm{c}} \end{cases} \qquad 2\text{-}18$$

Equation 2-18 represents the SW potential utilized to model the interactions between associating molecules. In this equation, the parameter $\varepsilon_{\mathrm{AB}}$ characterizes the association energy (well depth), while the parameter $\mathrm{r}_{AB}^{\mathrm{c}}$ delineates the cut-off range of the interaction between sites A and B (well width), which is related to the association volume parameter.

Over recent years, the SAFT model has undergone numerous refinements and modifications. These adaptations encompass diverse methodologies aimed at calculating the Helmholtz energy of fluids. These methods often involve selecting a reference fluid, outlining association schemes, selecting different intermolecular potentials, and integrating additional terms to accommodate intermolecular forces. As a result, multiple types of the SAFT model have emerged, each with its unique approach.

For instance, variations like SAFT-SW (HUANG and RADOSZ, 1993) implement the square-well potential to calculate intermolecular forces. TPT-D (TAVARES *et al.*, 1995) utilizes a reference fluid composed of dimers. Meanwhile, the polar soft-SAFT (ALKHATIB *et al.*, 2020) introduces a Helmholtz energy term specifically tailored to capture polar forces.

In this dissertation, the focus will be on revisiting and examining two specific versions of the SAFT model: PC-SAFT, as outlined by GROSS and SADOWSKI (2001), and SAFT-γ Mie , a proposition presented by PAPAIOANNOU *et al.* (2014). These versions represent distinct adaptations of the SAFT framework, each offering its unique set of assumptions and methodologies for modeling fluid behavior.

2.2. Phase Equilibria

According to Gibbs, equilibrium refers to a state of stability where no driving forces exist to alter the system, leading to an observable constancy in its macroscopic properties. In the context of phase equilibrium, this stability is pertinent to mass and energy exchanges among the system's phases. Therefore, for a closed system with no reactions, no total mass variations, no external field (gravitational, electric, magnetic, etc.), no interface curvature, with Np phases and Nc components, the necessary condition for equilibrium is:

$$\begin{gathered} \mathrm{T}^1 = \mathrm{T}^2 = \cdots = \mathrm{T}^{Np} \\ P^1 = P^2 = \cdots = P^{Np} \\ \mu_1^1 = \mu_1^2 = \cdots = \mu_1^{Np} \\ \mu_2^1 = \mu_2^2 = \cdots = \mu_2^{Np} \\ \cdots \\ \mu_{Nc}^1 = \mu_{Nc}^2 = \cdots = \mu_{Nc}^{Np} \end{gathered} \qquad 2\text{-}19$$

In these expressions, T is ensuring thermal equilibrium; P is guaranteeing mechanical equilibrium; and μ is ensuring chemical equilibrium. These equalities are fundamental in defining the equilibrium conditions for intensive properties. When a system is in equilibrium, it's at a state where the thermodynamic potential is minimized for specified state variables. The chemical potential for each component i can be defined as:

$$\mu_i = \left.\frac{\partial U}{\partial n_i}\right|_{S,V,n_{j\neq i}} = \left.\frac{\partial H}{\partial n_i}\right|_{S,P,n_{j\neq i}} = \left.\frac{\partial G}{\partial n_i}\right|_{\mathrm{T},P,n_{j\neq i}} = \left.\frac{\partial A}{\partial n_i}\right|_{\mathrm{T},V,n_{j\neq i}} \qquad 2\text{-}20$$

The equality of the chemical potential for all phases in a multicomponent system at constant T and P, leads to the minimization of the Gibbs free energy. The Gibbs free energy has great importance in the calculation of thermodynamic properties, since many chemical and physical transformations occur at constant pressure and temperature. We can also define the chemical potential in terms of fugacity:

$$\mu_i - \mu_i^0 = R\mathrm{T}ln\left(\frac{f_i}{f_i^0}\right) \qquad 2\text{-}21$$

Where f_i, f_i^0 and μ_i^0 are the fugacity, fugacity in the reference state and the chemical potential in the reference state of component i, respectively. Fugacity accounts for the attractive and repulsive forces between molecules. It's an abstract concept representing a substance's tendency to "escape" from a particular phase. Although its unit is pressure, it can be conceptualized as another Gibbs energy scale and can be directly inferred from PVT data. Through fugacity, the chemical equilibrium can be articulated without resorting to chemical potential. We can also write this expression for a scenario with two phases, α and β:

$$\mu_i^\alpha(\mathrm{T}, P, \underline{x}^\alpha) - \mu_i^0 = \mu_i^\beta(\mathrm{T}, P, \underline{x}^\beta) - \mu_i^0$$

$$\mu_i^\alpha = \mu_i^\beta = R\mathrm{T}ln\left(\frac{f_i^\alpha}{f_i^0}\right) = R\mathrm{T}ln\left(\frac{f_i^\beta}{f_i^0}\right) \forall\ i = 1,2 \dots Nc \qquad 2\text{-}22$$

Therefore, the equation defining the chemical equilibrium of component i between two phases α and β, can be written in terms of only fugacity:

$$f_i^\alpha(\mathrm{T}, P, \underline{x}^\alpha) = f_i^\beta(\mathrm{T}, P, \underline{x}^\beta) \qquad \forall\ i = 1,2 \dots Nc \qquad 2\text{-}23$$

The determination of equilibrium properties in multiphase systems can be meticulously approached through several computational strategies, each tailored to specific conditions and known variables. Among these, bubble point, dew point calculations, stability analysis, and flash calculations, stand as the foundational techniques. In this work, the focal point will be primarily on bubble point calculations to elucidate the vapor-liquid equilibria. For the solid-liquid equilibria, we will incorporate relations involving the enthalpy of fusion to model the solid phase. Furthermore, to ensure the fidelity of the phase equilibria predictions, this work will employ additionally the tangent plane distance as a verification tool. This will ensure the phases predicted by the computational models are not only thermodynamically stable but also align with empirical observations.

2.2.1. Vapor-Liquid Equilibrium

As previously mentioned, the isofugacity condition is crucial for accurately describing phase equilibria. To achieve this, two primary methodologies are generally employed: the phi-phi approach and the gamma-phi approach. Each method offers a distinct mechanism for representing the non-ideal behavior of components in various fluid phases, crucial for understanding complex systems where interactions between particles play a significant role.

The gamma-phi approach combines the concept of activity coefficients γ, calculated via activity coefficient models, with the fugacity coefficients ϕ, derived from equations of state (EoS). As discussed in the section 2.1.4, the activity coefficient models quantify deviations from ideal solution behavior in the liquid phase, capturing the essence of molecular interactions and mixtures' non-ideality. On the other hand, the fugacity coefficient accounts for the non-ideal behavior of gases, calculated using an EoS, which effectively encapsulates the real gas behaviors and phase interactions. This approach provides a comprehensive view by integrating these two aspects, making it suitable for systems where both liquid and vapor phases exhibit non-ideal behaviors.

In contrast, the phi-phi approach relies solely on the use of equations of state to compute the fugacity coefficients for all phases involved, be it vapor, liquid, or even supercritical. This method is deemed more symmetric and straightforward, as it applies a consistent thermodynamic framework across all phases, relying on the robustness of the chosen EoS to describe the system's behavior. It's particularly advantageous when the EoS accurately represents the phase behavior of the components across different conditions, making it a preferred choice for certain types of systems.

In the context of this study, the focus is exclusively on the phi-phi approach to compute the vapor-liquid equilibrium (VLE). Specifically, we are employing equations of state, namely PC-SAFT and SAFT-γ-Mie. These sophisticated EoS are known for their accuracy in modeling complex interactions within fluids, particularly for predicting thermodynamic properties and phase behavior in multi-component systems.

The phi-phi approach represents the fugacities in the following forms:

$$f_i^j = \varphi_i^j P w_i^j \qquad \forall i = 1,2 \dots Nc \quad and \quad \forall j = 1,2 \dots Np \tag{2-24}$$

where φ_i^j and w_i^j are the fugacity coefficients and mole fraction of component i in in the phase j, respectively.

This study determines either the limit pressure or temperature at which a nascent vapor phase begins to appear from a developed liquid mixture with a predetermined composition. This type of calculation, known as bubble point, also aims to ascertain the composition of the incipient vapor phase in the context of liquid-vapor equilibrium problems. Given the liquid phase composition, the working equations for bubble point calculation are the following:

$$P^V = P^L$$

$$\mathrm{T}^V = \mathrm{T}^L$$

$$y_i = K_i x_i \qquad \forall i = 1,2 \dots Nc$$

$$\ln\left(x_i \phi_i^L(\bar{x}, \mathrm{T}, P)\right) - \ln\left(y_i \phi_i^V(\bar{y}, \mathrm{T}, P)\right) = 0 \qquad \forall i = 1,2 \dots Nc \tag{2-25}$$

$$\sum_i^{Nc} y_i = 1 \qquad \forall i = 1,2 \dots Nc$$

In a scenario where the vapor phase cannot be represented by an ideal gas, and the liquid phase is far from an ideal mixture, it's impossible to arrive at an analytical expression to calculate the saturation temperature or pressure and the vapor phase composition. This difficulty arises because the distribution of the phase compositions depends on all the compositions involved. The computational strategy to calculate the bubble point, therefore, follows an iterative process. Numerical techniques for such calculations are described in numerous references, but in this study, we will use the bubble point algorithms described by SANDLER (2016).

2.2.2. Solid-Liquid Equilibrium for Pure Solids

Solubility is a chemical property that quantifies the ability of a substance, known as the solute, to dissolve in a solvent. It is typically expressed in terms of the maximum quantity (saturation) of solute that can dissolve in a specific amount of solvent under certain temperature and pressure conditions. Component 1, which forms a pure solid,

reaches saturation in the liquid phase at equilibrium. Therefore, this state can be characterized by the equality of the fugacities of this solute in both of the phases.

$$f_1^{sat,L} = f_1^s \tag{2-26}$$

In this context, the superscripts "L", "s" and "sat" denote the liquid phase, solid phase, and saturation, respectively. These two fugacities are independent of the nature of the solvent and we are treating the interactions in the solid phase as ideal. The fugacity of component 1 in the liquid phase under saturation conditions is given by the following expression:

$$f_1^{sat,l} = x_1^{sat,l}\gamma_1^{sat,l} f_1^l(x_1 = 1, \mathrm{T}) \tag{2-27}$$

Where $x_1^{sat,l}$ and $\gamma_1^{sat,l}$ are the mole fraction composition and the activity coefficient, respectively, of component 1 in the liquid phase under saturation conditions. f_1^l is the fugacity of pure component 1 in the sub-cooled liquid state at temperature T. The ratio $\frac{f_1^l}{f_1^s}$ can be related to the change in molar Gibbs energy (Δg) of the solid's fusion at this temperature. This change in Gibbs energy is a function of the enthalpic (Δh) and entropic (Δs) changes, as expressed by the relationship:

$$R\mathrm{T}\ln\left(x_1^{sat,l}\gamma_1^{sat,l}\right) = -R\mathrm{T}\ln\left(\frac{f_1^l}{f_1^s}\right) = \Delta g_1 = \Delta h_1 - \mathrm{T}\Delta s_1 \tag{2-28}$$

PRAUSNITZ *et al.*, (1999) provides a method to evaluate these enthalpy and entropy changes using the triple point, given by:

$$\Delta h = \Delta h_t^{fus} + \int_{T_t}^{T} \Delta C_p \, d\mathrm{T} \tag{2-29}$$

$$\Delta \mathrm{s} = \Delta s_t^{fus} + \int_{T_t}^{T} \frac{\Delta C_p}{T} d\mathrm{T} \tag{2-30}$$

Where,

$$\Delta C_p = C_p^L - C_p^s \quad \text{2-31}$$

Here, the subscript t indicates the condition at the triple point, and the superscript fus represents the property associated with fusion or melting. C_p^L is the molar heat capacity at constant pressure in the pure liquid state, and C_p^s is the equivalent in the crystalline state. The fusion entropy can be substituted by its relationship with the fusion enthalpy, which is a good approximation of the ΔC_p (NEAU *et al*, 1997). This relation is expressed as follows:

$$\Delta s = \frac{\Delta h^{fus}}{T} \cong \Delta C_p \quad \text{2-32}$$

When approximating the triple point temperature with the melting temperature at the system's pressure, while also assuming that the differences in heat capacities between the solid and liquid are temperature-independent, we can substitute equations 2-30, 2-31, 2-32 and 2-33 into 2-29. By performing the integrations with these assumptions, we obtain the following expression for solubility:

$$\ln\left(x_1^{sat,l}\gamma_1^{sat,l}\right) = \frac{\Delta h_1^{fus}}{RT}\left(\frac{T}{T_{m_1}} - 1\right) + \frac{\Delta C_{p_1}}{R}\left(\ln\left(\frac{T_{m_1}}{T}\right) - \frac{T_{m_1}}{T} + 1\right) \quad \text{2-33}$$

Note that Δh_1^{fus} represents the latent heat of fusion of component 1 at its melting temperature T_{m_1}. This equation describes the crystallization curve of a solution, provided that the solution forms no mixed crystals. It is reasonable to assume that the difference in heat capacities between the solid and liquid states is small enough to be negligible in the equation, besides, this term can be very difficult to measure.

2.2.3. Solid-Liquid Equilibrium for Racemic Mixtures

In the specific case of a racemic mixture, an equation similar to Equation 2-33 can be used. However, one should consider a solid composed by two molecules in the same proportion. Therefore, considering components 1 and 2 as the enantiomers, the

compositions in the liquid phase can be approximated as equal ($x_1^{sat,L} = x_2^{sat,L}$), while in the racemic solid ($x_1^{sat,s}$ and $x_2^{sat,s}$) they should be set to 0.5. Moreover, it is also required to replace the calorimetric properties with those of the racemic mixture ($T_{m_{rac}}$, $\Delta C_{p_{rac}}$ and Δh_{rac}^{fus}. Note that these parameters vary greatly with the composition of the solid. This means that a racemic mixture will have different properties from the enantiopure one. Therefore, substituting the new fugacities, the equation representing the solubility curve of a racemic solid is given by:

$$\ln\left(4\left(x_1^{sat,l}\right)^2\left(\gamma_1^{sat,l}\right)^2\right) = \frac{\Delta h_{rac}^{fus}}{RT}\left(\frac{T}{T_{m_{rac}}} - 1\right) + \frac{\Delta C_{p_{rac}}}{R}\left(\ln\left(\frac{T_{m_{rac}}}{T}\right) - \frac{T_{m_{rac}}}{T} + 1\right) \quad 2\text{-}34$$

2.2.4. Phase Stability and Tangent Plane Distance (TPD)

The analysis of phase stability and the calculation of thermodynamic equilibrium for a multicomponent mixture are crucial tools in the field of thermodynamics. As previously mentioned, equilibrium is achieved when there is an equality of the fugacities of each species in each phase. However, ensuring thermodynamic stability requires that the thermodynamic potential in question, typically Gibbs energy in a state defined by temperature and pressure, must converge to its global minimum while varying the compositions of each component in each phase.

A primary approach to finding the minimum of Gibbs energy in a system with np phases and at a given temperature T and pressure P can be described as follows:

$$\min G = \sum_{p=1}^{Np}\sum_{i=1}^{Nc} n_i^p \ln\left(x_i^p \gamma_i^p\right) \quad 2\text{-}35$$

Under the constraints:

$$\sum_{p=1}^{Np} n_i^p = Z_i \quad 2\text{-}36$$

$$n_i^p \geq 0 \qquad \forall\, i \; and \; p \quad 2\text{-}37$$

Where Z_i is the global mole number of component "i" and n_i^p is the total number of mols of component "i" in the phase "p".

The minimum of Gibbs energy is both a necessary and sufficient condition for equilibrium. However, locating this global minimum can be particularly challenging when the initial estimate is far from the actual solution, due to the complex and nonlinear nature of the function.

The method for phase stability analysis proposed by MICHELSEN in 1982 offers a practical algorithm to this challenge. It presents the tangent plane distance (TPD) of the Gibbs energy of the mixture, which is a powerful criterion in determining phase stability. Michelsen's method provides a systematic approach to evaluate the composition of a new phase at equilibrium.

Assuming a system that initially has one phase (α) with composition n and a Gibbs energy of $G^{(\alpha)}(n)$, let's consider a scenario where a new phase (β) emerges, removing an infinitesimal amount W from the first phase. The variation in the Gibbs energy of the system due to this development can be expressed as follows:

$$\Delta G = \left(G^{(\alpha)}(\bar{n} - \bar{W}) + G^{\beta}(\bar{W})\right) - G^{(\alpha)}(\bar{W}) \qquad 2\text{-}38$$

We can approximate the Gibbs energy of the first phase using a Taylor series expansion around the point just before the formation of the new phase:

$$G^{(\alpha)}(\bar{n} - \bar{W}) = G^{(\alpha)}(\bar{n}) - \sum_{i}^{Nc} \xi_i \left(\frac{\partial G^{(\beta)}}{\partial n_i}\right)_N \qquad 2\text{-}39$$

Given the definitions:

$$G^{(\alpha)}(\bar{n}) = \sum_{i}^{Nc} n_i \, \mu_i^{(\alpha)} \qquad 2\text{-}40$$

$$\frac{\partial G^{(\alpha)}}{\partial n_i} = \mu_i^{(\alpha)} \qquad 2\text{-}41$$

If we substitute Equations 2-39, 2-40 and 2-41 into 2-38, and then divide the entire expression by the total number of moles in the second phase, we obtain the expression for the Tangent Plane Distance (TPD) as described by BAKER *et al.* in 1982.

$$\frac{\Delta G}{\sum_i^{Nc} W_i} = F(w) = \sum_i^{Nc} w_i \left(\mu_i^{(\alpha)}(w) - \mu_i^{(\beta)}(z)\right) \qquad 2\text{-}42$$

In this equation, w and z represent the mole fractions of the newly formed phase and the initial phase, respectively. The function F is a representation of the Tangent Plane Distance (TPD). It measures the vertical distance from the composition point of the new phase to the tangent hyperplane at the composition point of the initial phase.

The significance of F lies in its ability to determine the stability of the new phase. If F is positive, it implies that the formation of the new phase has led to an increase in the Gibbs energy of the system, indicating that this new phase is not stable. For the initial phase to be the only stable phase, F must assume positive values for any composition w.

Michelsen, utilizing the above stability criterion, developed an equivalent version of the minimization problem with a single constraint. It is given by:

$$TPD(W_i) = \sum_i^{Nc} W_i \left[\ln(W_i) + \ln\left(\phi_i^{(\beta)}(W_i) - ln\left(z_i \phi_i^{(\alpha)}(z_i)\right)\right)\right] \qquad 2\text{-}43$$

In this equation, Michelsen introduces a new variable, W, which represents the total number of moles in the incipient phase (II). The single constraint is that $W_i > 0$. This formulation is useful when it is already known that the first phase (I) is stable. It allows for the use of a successive substitution method to compute the composition of the incipient phase (II). This method is iteratively applied and is represented as follows:

$$\ln\left(\mathrm{W_i^{[k+1]}}\right) = \ln(z_i) + \ln\left(\phi_i^{(\alpha)}(z_i)\right) - \ln\left(\phi_i^{(\beta)}\left(W_i^{[k]}\right)\right) \qquad 2\text{-}44$$

The superscript $[k]$ denotes the iteration step in the successive substitution method.

3. Methodology

3.1. PC-SAFT EoS

The Perturbed-Chain Statistical Association Fluid Theory (PC-SAFT) is an adaptation of the SAFT framework, devised by GROSS and SADOWSKI in 2001. This theory, like other perturbation theories, relies on two primary elements: the reference system and the perturbation expansion form. These elements are chosen based on numerical practicality and convergence efficiency. The reference system's properties should closely resemble the target system, facilitating simpler computation of the perturbation expansion compared to direct analysis of the target system. Despite these criteria, options remain broad. In contrast to SAFT's hard sphere fluid reference, PC-SAFT uses a hard-chain fluid for perturbation reference. The dispersion potential is specifically introduced between connected segments or chains, as opposed to the conventional SAFT approach that considers dispersion potential between individual spherical segments. This adjustment enables a more accurate representation of interactions between these linked molecular structures, thereby enhancing the model's predictive capabilities in systems where such unbranched chain-like organizations are prevalent (SMITH and NEZBEDA, 1983).

The formulation of the PC-SAFT Equation of State hinges on computing the residual Helmholtz energy, represented as A^{res} through the summation of various Helmholtz contributions arising from distinct intermolecular interactions. This computation is expressed by the following expression:

$$\frac{A^{res}}{Nk_bT} = \frac{A}{Nk_bT} - \frac{A^{ideal}}{Nk_bT} = \frac{A^{hc}}{Nk_bT} + \frac{A^{disp}}{Nk_bT} + \frac{A^{assoc}}{Nk_bT} + \ \ldots \qquad 3\text{-}1$$

Or in its dimensionless form:

$$\tilde{a}^{res} = \tilde{a} - \tilde{a}^{ideal} = \tilde{a}^{hc} + \tilde{a}^{disp} + \tilde{a}^{assoc} + \cdots \qquad 3\text{-}2$$

Where $\tilde{a} = \frac{A}{Nk_bT}$. The superscript notations *res*, *ideal*, *hc*, *disp*, and *assoc* correspond, respectively, to the residual property, ideal gas property, the contribution stemming from the presence of rigid molecular chains, the dispersive contribution, and the contribution to the Helmholtz energy arising from associations between molecules. In this notation the tilde indicates reduced properties which are obtained by normalizing

a given variable with respect to the Boltzmann's constant (k_b), the total number of molecules in the mixture (N) and the temperature.

3.1.1. Intermolecular Potential

The pair potential attributed to the segments forming a chain is characterized by a modified square well potential, a proposition introduced by KREGLEWSKI and CHEN (1980). This modified potential governs the interactions between the spherical segments composing the molecular chains, representing the manner in which these segments interact with each other within the chain structure.

$$\phi(r) = \begin{cases} \infty, r < (\sigma - 0.12\sigma) \\ 3\varepsilon, (\sigma - 0.12\sigma) \leq r < \sigma \\ -\varepsilon, \sigma \leq r < S_2\sigma \\ 0, r \geq S_2\sigma \end{cases} \quad 3\text{-}3$$

The function $\phi(r)$ characterizes the pair potential, with r denoting the radial distance between two segments. The parameter σ represents the segment diameter, providing insight into the characteristic size of the involved segments. S_2 is the reduced well width. ε symbolizes the depth of the potential well.

Moreover, the step function within this representation approximates a soft repulsion. This step function serves as a simplified yet rougher model akin to the Lennard-Jones potential, although possessing a more straightforward mathematical treatment. PRAUSNITZ *et al.*, (1999) highlighted this approach for its mathematical convenience despite its deviation from the intricacies of the Lennard-Jones potential.

The presented $\phi(r)$ potential is utilized to calculate the effective collision diameter of a specific substance i through the integration of the following equation:

$$d_i(T) = \int_0^{\sigma_i} \left(1 - \exp\left(-\frac{\phi(r)}{k_b T}\right)\right) dr \quad 3\text{-}4$$

Wich leads to:

$$d_i(T) = \sigma_i \left(1 - 0.12e^{-\frac{3\varepsilon_i}{k_b T}}\right) \quad 3\text{-}5$$

3.1.2. Hard Sphere Term

The term $\tilde{a}^{hs}$ signifies the alteration in Helmholtz free energy resulting from interactions among hard spheres at particular density and temperature conditions. This contribution can be characterized by the following expression:

$$\tilde{a}^{hs} = \frac{1}{\zeta_0}\left[\frac{3\zeta_1\zeta_2}{1-\zeta_3} + \frac{\zeta_2^3}{\zeta_3(1-\zeta_3)^2} + \left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0\right)\ln(1-\zeta_3)\right] \quad 3\text{-}6$$

The terms ζ_n represent the reduced densities, which is a vector proportional to molar density, as proposed by CHAPMAN *et al.* (1990), and its formulation is detailed as follows:

$$\zeta_n = \frac{\pi}{6}\rho\sum_i^{Nc} x_i m_i d_i^n \qquad n = \{0,1,2,3\} \quad 3\text{-}7$$

The symbol ρ denotes the density measured in molecules per volume unit. Nc represents the total number of components within the system. The parameter m signifies the number of segments pertaining to a specific component i, while x denotes the mole fraction corresponding to that particular component i. The exponent and subscript n takes the values 0, 1, 2, and 3. The term reduced density, or packing fraction (η), is used for n = 3.

3.1.3. Hard Chain Term

The term $\tilde{a}^{hc}$ represents the change in dimensionless Helmholtz energy due to the covalent bonds between the segments, providing the formation of chains. The expression described by BOUBLIK (1970) is represented as follows:

$$\tilde{a}^{hc} = \bar{m}\tilde{a}^{hs} - \sum_i^{Nc} x_i(m_i - 1)\ln\left(g_{ii}^{hs}\right) \quad 3\text{-}8$$

Where the term $\bar{m}$ is the mean segment number of the mixture, calculated in Equation 3-10. Additionally, the term g_{ii}^{hs} represents the radial distribution function of this hard sphere fluid, and it is expressed as follows:

$$g_{ij}^{hs} = \frac{1}{(1-\zeta_3)} + \left(\frac{d_i d_j}{d_i + d_j}\right)\left(\frac{3\zeta_2}{(1-\zeta_3)^2}\right)\left(\frac{d_i d_j}{d_i + d_j}\right)^2\left(\frac{2\zeta_2^2}{(1-\zeta_3)^3}\right) \quad 3\text{-}9$$

The radial pair distribution function is a measure of the probability of finding a particle of type i in a given distance from a fixed particle of type j in the fluid.

$$\bar{\mathrm{m}} = \sum_{i}^{Nc} x_i m_i \qquad 3\text{-}10$$

3.1.4. Dispersive Term

According to BARKER and HENDERSON (1967), it is possible to write the residual Helmholtz free energy as:

$$\tilde{\mathrm{a}}^{res} = \sum_{n=0}^{\infty} \frac{A'_n}{\mathrm{N}(k_b \mathrm{T})^n} \qquad 3\text{-}11$$

Where A_0 is the Helmholtz free energy of the reference fluid, in this case, composed of chains of hard spheres and A_n are the nth-order perturbation terms.

We can then compute the Helmholtz free energy related to the dispersive forces expressed as a second-order expansion. Gross and Sadowski's PC-SAFT expression for this term occurs as follows:

$$\frac{A^{disp}}{\mathrm{N k_b T}} = \frac{\mathrm{A_1}}{\mathrm{N k_b T}} + \frac{\mathrm{A_2}}{\mathrm{N k_b T}} = \tilde{\mathrm{a}}^{disp} = \tilde{\mathrm{a}}_1 + \tilde{\mathrm{a}}_2 \qquad 3\text{-}12$$

Where $\mathrm{A_1}$ and $\mathrm{A_2}$ are determined by the integrals:

$$\frac{\mathrm{A_1}}{\mathrm{N k_b T}} = -2\pi\rho \mathrm{m}^2 \left(\frac{\varepsilon}{\mathrm{k_b T}}\right) \sigma^3 \int_1^{-\infty} \left[\frac{\phi(\mathrm{r})}{\varepsilon} \mathrm{g}^{hc}\left(\mathrm{m}; \mathrm{r}\frac{\sigma}{\mathrm{d}}\right) \mathrm{r}^2\right] \mathrm{dr} \qquad 3\text{-}13$$

$$\frac{\mathrm{A_2}}{\mathrm{N k_b T}} = -\pi\rho \mathrm{m} \left(1 + \mathrm{Z}^{hc} + \rho \frac{\partial \mathrm{Z}^{hc}}{\partial \rho}\right)^{-1} \times \mathrm{m}^2 \left(\frac{\varepsilon}{\mathrm{k_b T}}\right)^2 \sigma^3 \frac{\partial}{\partial \rho} \left\{ \int_1^{-\infty} \left[\left(\frac{\phi(\mathrm{r})}{\varepsilon}\right)^2 \mathrm{g}^{hc}\left(\mathrm{m}; \mathrm{r}\frac{\sigma}{\mathrm{d}}\right) \mathrm{r}^2 \right] \mathrm{dr} \right\} \qquad 3\text{-}14$$

Where Z represents the compressibility factor, calculated as $\frac{P}{\rho k_b T}$, and g^{hc} represents the radial distribution function of the binding potential of a specific segment from any other chain.

The expression that combines the compressibility factor with its rate of change concerning density, as presented in Equation 3-14, can be derived using the following equation:

$$Z^{hc} + \rho\frac{\partial Z^{hc}}{\partial \rho} = \bar{m}\frac{8\eta - 2\eta^2}{(1-\eta)^4} + (1-\bar{m})\frac{20\eta - 27\eta^2 + 12\eta^3 - 2\eta^4}{\left((1-\eta)(2-\eta)\right)^2} \quad 3\text{-}15$$

The integrals in Equations 3-13 and 3-14 have analytical solutions, but for practicality, Gross and Sadowski solved these using series and denoted them as variables I_1 and I_2, respectively. They are calculated as follows:

$$\int_1^{-\infty}\left[\tilde{u}(r)^2 g^{hc}\left(m; r\frac{\sigma}{d}\right) r^2\right] = I_1(\eta, m) = \sum_{i=0}^{6} a_i \zeta_3{}^i \quad 3\text{-}16$$

$$\frac{\partial}{\partial \rho}\left\{\int_1^{-\infty}\left[\tilde{u}(r)^2 g^{hc}\left(m; r\frac{\sigma}{d}\right) r^2\right] dr\right\} = I_2(\eta, m) = \sum_{i=0}^{6} b_i \zeta_3{}^i \quad 3\text{-}17$$

Here, a_i and b_i are the series constants using experimental data from pure alkanes. The relationship describing the coefficients concerning the segment number has been deduced by LIU and HU (1997) as follows:

$$a_i = a_{0i} + \frac{m-1}{m}a_{1i} + \frac{m-1}{m}\frac{m-2}{m}a_{2i} \quad 3\text{-}18$$

$$b_i = b_{0i} + \frac{m-1}{m}b_{1i} + \frac{m-1}{m}\frac{m-2}{m}b_{2i} \quad 3\text{-}19$$

Where a_{ni} and b_{ni} for n = 0,1,2 are universal constants. Their values can be found in the original paper of GROSS and SADOWSKI (2001).

The first and second-order perturbation terms can be expressed for mixtures as follows:

$$\frac{A_1}{Nk_bT} = -2\pi\rho I_1 \overline{m^2\varepsilon\sigma^3} \tag{3-20}$$

$$\frac{A_2}{Nk_bT} = -2\pi\rho\bar{m}\left(1 + Z^{hc} + \rho\frac{\partial Z^{hc}}{\partial\rho}\right)^{-1} I_2\, \overline{m^2\varepsilon^2\sigma^3} \tag{3-21}$$

$$\overline{m^2\varepsilon\sigma^3} = \sum_i^{Nc}\sum_j^{Nc} x_i x_j m_i m_j \left(\frac{\varepsilon_{ij}}{k_bT}\right)\sigma_{ij}^3 \tag{3-22}$$

$$\overline{m^2\varepsilon^2\sigma^3} = \sum_i^{Nc}\sum_j^{Nc} x_i x_j m_i m_j \left(\frac{\varepsilon_{ij}}{k_bT}\right)^2\sigma_{ij}^3 \tag{3-23}$$

Where the parameters for a pair of different segments are obtained through the conventional BERTHELOT-LORENTZ (1877) combining rules, as described by the following equations:

$$\sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j) \tag{3-24}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_j\varepsilon_i}(1 - k_{ij}) \tag{3-25}$$

Where k_{ij} is one of the binary interaction parameters, related to the dispersive forces. It is determined through fitting procedures based on experimental data.

3.1.5. Association Term

As previously mentioned, SAFT-type equations of state can model associative systems involving hydrogen bonding. In this scenario, the term accounting for the contribution to the Helmholtz free energy due to hydrogen bond forces will be:

$$\bar{a}^{assoc} = \sum_i x_i \left(\sum_j \left(\ln(\mathrm{X}_{ji}) - \frac{\mathrm{X}_{ji}}{2} + \frac{1}{2} \right) \mathrm{S}_{ji} \right) \quad \text{3-26}$$

Where X_{ji} represents the fraction of molecules of type i not bonded to site j, and S_{ji} refers to the multiplicity of site j within molecule i. If we consider a mixture of water and ethanol, modeling both water and ethanol as having two types of sites—an electron donor and an electron acceptor—S will have dimensions of 2x2. The columns represent the components, while the rows denote the type of site. In a more realistic scenario, water could be modeled with 2 pairs of electrons and 2 hydrogen atoms available for bonding, whereas ethanol would have 2 pairs of electrons and only one hydrogen. In this case, it would give us the following site multiplicity matrix:

$$\begin{vmatrix} 2 & 2 \\ 2 & 1 \end{vmatrix} \quad \text{3-27}$$

We can iteratively calculate the fraction X_{ji} using the approach demonstrated in the studies by SOLMS, *et al.* (2003) and MICHELSEN and HENDRIKS (2001), following the expression:

$$X_{ji} = \frac{1}{1 + \rho \sum_i x_i \sum_l \sum_k S_{lk} X_{lk} \Delta_{lkji}} \quad \text{3-28}$$

Here we introduce the term Δ_{lkji}. It represents the strength of interaction due to the association of site l of molecule k with site j of the molecule i. It can be calculated as follows:

$$\Delta_{lkji} = \frac{\pi}{6} g_{ij}^{hs} \sigma_{ij}^3 \kappa^{\mathrm{lkji}} \left(exp\left(\frac{\varepsilon^{lkji}}{\mathrm{k_b T}} \right) - 1 \right) \quad \text{3-29}$$

Where ε^{lkji} and κ^{lkji} are the associating model parameters representing the association energy and effective association volume between the sites l and j per molecule k and i, respectively. These parameters can be obtained through fittings to phase equilibrium or other thermodynamic property data.

For the determination of the two cross-association parameters between two substances, one can utilize the combination rule proposed by WOLBACH and SANDLER (1998). They devised this rule based on an analysis of association in the gas phase under low-pressure conditions. The combination rules are provided as follows:

$$\kappa^{\mathrm{lkji}} = \sqrt{\kappa^{liji}\kappa^{\mathrm{lkjk}}}\left(\frac{\sqrt{\sigma_{\mathrm{ii}}\sigma_{\mathrm{kk}}}}{\frac{1}{2}(\sigma_{\mathrm{ii}}+\sigma_{\mathrm{kk}})}\right)^{3}(1-\mathrm{k}_{2_{ij}}^{asc}) \qquad 3\text{-}30$$

$$\varepsilon^{lkji} = \frac{\left(\varepsilon^{liji}+\varepsilon^{lkjk}\right)}{2} \qquad 3\text{-}31$$

These combination rules may or may not include an additional binary correction parameter, $k_{2_{ij}}^{asc}$.

3.2. SAFT-γ Mie EoS

The SAFT-γ Mie is a SAFT-type equation of state formulated by PAPAIOANNOU *et al.* (2014). It was built upon a group contribution approach of SAFT-VR-Mie. Within this model, molecules are characterized as connected chains of spheres. Diverging from the PC-SAFT model, the SAFT-γ Mie approximates a single molecule as spheres of varying sizes, interconnected with a certain degree of overlap, denoted by the parameter S_k. A reduced S_k value implies a smaller contribution of a chain to the Helmholtz energy (DUFAL *et al.*, 2014). Within this framework, which employs group contributions to compute properties within a mixture of compounds, each group 'k' is made by a chain of spheres sharing identical parameters. This uniformity spans diameters, dispersive energies, as well as attractive and repulsive exponents for segment-segment interactions. This functional group or chain 'k', comprises a count of ν_k spheres or segments, interacting through the Mie potential. Moreover, the multiplicity or frequency of occurrence of a group 'k', in a molecule 'i', is denoted by the parameter ν_{ki}. Analogous to the PC-SAFT approach, association sites may be incorporated to account for highly directional short-range forces. However, distinct from PC-SAFT, these sites are linked within a group.

The SAFT-γ Mie formulation deviates slightly from the PC-SAFT EoS in the computation of A^{res} :

$$\frac{\mathrm{A}^{res}}{\mathrm{N}\mathrm{k_b}\mathrm{T}} = \frac{\mathrm{A}^{mono}}{\mathrm{N}\mathrm{k_b}\mathrm{T}} + \frac{\mathrm{A}^{chain}}{\mathrm{N}\mathrm{k_b}\mathrm{T}} + \frac{\mathrm{A}^{assoc}}{\mathrm{N}\mathrm{k_b}\mathrm{T}} + \ \ldots \qquad 3\text{-}32$$

Or in its dimensionless form:

$$\tilde{\mathrm{a}}^{res} = \tilde{\mathrm{a}}^{mono} + \tilde{\mathrm{a}}^{chain} + \tilde{\mathrm{a}}^{assoc} + \cdots \qquad 3\text{-}33$$

Where $\tilde{\mathrm{a}} = \frac{\mathrm{A}}{\mathrm{N}\mathrm{k_b}\mathrm{T}}$. This formulation introduces two distinct subscripts: "*mono*" denotes the monomer term, while "*chain*" represents the chain formation term.

3.2.1. Intermolecular Potential

Unlike PC-SAFT, which employs a modified square well potential as an approximation of the Lennard-Jones potential, SAFT-γ Mie utilizes the Mie potential to characterize the forces between segments. The Mie potential is expressed as follows:

$$\phi_{\mathrm{kl}}^{mie}(r_{kl}) = \mathrm{C}_{kl}\ \varepsilon_{kl}\left(\left(\frac{\sigma_{kl}}{r_{kl}}\right)^{\lambda_{kl}^{r}} - \left(\frac{\sigma_{kl}}{r_{kl}}\right)^{\lambda_{kl}^{a}}\right) \qquad 3\text{-}34$$

Where r_{kl} signifies the distance between the center of segment ‘k’ and the center of segment ‘l’, ε_{kl} represents the combination of dispersive forces' energy between segments, λ_{kl}^{r} and λ_{kl}^{a} are the combinations of repulsive and attractive exponents of the segment-segment interaction, σ_{kl} denotes the average between the interacting diameters, and $\mathrm{C_{kl}}$ is a function of the interaction exponents used such that the minimum of the potential remains at ε_{kl} for all λ_{kl}^{r} and λ_{kl}^{a}. This function is formulated as follows:

$$\mathrm{C}_{kl} = \frac{\lambda_{kl}^{r}}{\lambda_{kl}^{r} - \lambda_{kl}^{a}}\left(\frac{\lambda_{kl}^{r}}{\lambda_{kl}^{a}}\right)^{\frac{\lambda_{kl}^{a}}{\lambda_{kl}^{r}-\lambda_{kl}^{a}}} \qquad 3\text{-}35$$

It's important to note that the Lennard-Jones potential constitutes a specific case of the Mie potential when the exponents λ_{kl}^{r} and λ_{kl}^{a} are set to 6 and 12, respectively. In the Lennard-Jones potential, the selection of the exponent "6" aligns with the lowest exponent in London dispersion forces among spherical, nonpolar molecules, while the choice of "12" remains arbitrary and is adopted for simplicity. Consequently, adjusting

these exponents within the Mie potential affords an enhanced portrayal of inter-segment forces overall, particularly when they deviate from being nonpolar and exhibit distinct shapes compared to spheres (HIRSCHFELDER *et al.*, 1964, PRAUSNITZ *et al.*, 1999)

3.2.2. Monomer Term

The SAFT-γ-Mie model employs perturbation theories to compute molecular interactions. These theories propose that molecule interactions can be calculated using attractive and repulsive contributions from the potential. Soft-core potentials, such as Mie or LJ, approximate the reference system as a fluid with no attraction forces, having hard repulsion, and a temperature-dependent diameter, $d_{kk}(\mathrm{T})$. This term resembles the one in Equation 3-4 in PC-SAFT, yet its calculation through the Equation 3-5 cannot be performed, due to the Mie potential's dependence on λ_{kl}^{r} and λ_{kl}^{a}, instead of the constants 6 and 12.

In this work, we numerically integrate Equation 3-4 using Simpson's method, for which an implementation was readily available. It worked satisfactorily and no other methods were tried in this work scope.

Given the assumption that each segment is a Mie Sphere, the first term from Equation 3-32, A^{mono}, can be formulated as a third-order high-temperature perturbation expansion over the hard spheres reference, A^{hs} :

$$\frac{\mathrm{A}^{mono}}{\mathrm{N k_b T}} = \frac{\mathrm{A}^{hs}}{\mathrm{N k_b T}} + \frac{\mathrm{A_1}}{\mathrm{N k_b T}} + \frac{\mathrm{A_2}}{\mathrm{N k_b T}} + \frac{\mathrm{A_3}}{\mathrm{N k_b T}} \qquad 3\text{-}36$$

Here, $\mathrm{A_1}$, $\mathrm{A_2}$ and $\mathrm{A_3}$ symbolize the dispersive forces and A^{hs} the repulsive ones. The A^{hs} expression of the mixture is given by:

$$\frac{A^{hs}}{\mathrm{N k_b T}} = \left(\sum_{i}^{Nc} \mathrm{x}_i \sum_{k}^{Ng} \nu_{ki} \nu_k^* S_k \right) \tilde{\mathrm{a}}^{hs} \qquad 3\text{-}37$$

Where Nc represents the total amount of components in the system, Ng is the number of different groups in the system and $\tilde{\mathrm{a}}^{hs}$ is the dimensionless contribution to the hard spheres per segment. Similar to the PC-SAFT EoS, the expression of Boublik is used for mixtures of the hard-sphere reference system:

$$\tilde{\mathrm{a}}^{\mathrm{hs}} = \frac{6}{\pi \rho_s} \left[\frac{3\zeta_1 \zeta_2}{1-\zeta_3} + \frac{\zeta_2^3}{\zeta_3 (1-\zeta_3)^2} + \left(\frac{\zeta_2^3}{\zeta_3^2} - \zeta_0 \right) \ln(1-\zeta_3) \right] \qquad 3\text{-}38$$

Where ρ_s, the total number of spherical segment density of the Mie fluid, is a function of the mixture density, ρ, by the following relation:

$$\rho_s = \rho\left(\sum_i^{Nc} \mathrm{x}_i \sum_k^{Ng} \nu_{ki}\nu_k^* S_k\right) \tag{3-39}$$

With ζ_n defined as:

$$\zeta_\mathrm{n} = \frac{\pi}{6}\rho_s \sum_k^{Ng} x_{s_k} d_{kk}^n \qquad n = \{0,1,2,3\} \tag{3-40}$$

Where x_{s_k} represents the fraction of segments of a specific type k group within the mixture, which is defined as:

$$x_{s_k} = \frac{(\sum_i^{Nc} \mathrm{x}_i \ \nu_{ki}\nu_k^* S_k)}{\left(\sum_j^{Nc} \mathrm{x}_j \ \sum_l^{Ng} \nu_{lj}\nu_l^* S_l\right)} \tag{3-41}$$

3.2.3. Dispersive Term

The third-order perturbation expansion used to calculate the attractive part of the spheres interactions is given as follows:

$$\frac{\mathrm{A}^{disp}}{\mathrm{Nk_bT}} = \frac{\mathrm{A_1}}{\mathrm{Nk_bT}} + \frac{\mathrm{A_2}}{\mathrm{Nk_bT}} + \frac{\mathrm{A_3}}{\mathrm{Nk_bT}} \tag{3-42}$$

$\mathrm{A_1}$ is the mean-attractive energy, while $\mathrm{A_2}$ and $\mathrm{A_3}$ are the fluctuations of the attractive energy within the system. Each perturbation term can be expressed as follows:

$$\frac{\mathrm{A}_p}{\mathrm{Nk_bT}} = \left(\frac{1}{\mathrm{k_bT}}\right)^p \left(\sum_i^{Nc} \mathrm{x}_i \sum_k^{Ng} \nu_{ki}\nu_k^* S_k\right)\tilde{\mathrm{a}}_p \qquad p = \{1,2,3\} \tag{3-43}$$

Where $\tilde{\mathrm{a}}_p$ is either the mean-attractive or fluctuation energy term per segment. Each of these terms can be obtained by the following summations:

$$a_p = \sum_{k}^{Ng} \sum_{l}^{Ng} x_{s_k} x_{s_l} a_{pkl} \qquad p = \{1,2,3\} \tag{3-44}$$

Where a_{pkl} are the pairwise interactions between groups 'k' and 'l' over all functional groups present in the system. The first term, a_{1kl}, can be calculated as follows:

$$a_{1kl} = C_{kl}\left(x_{0kl}^{\lambda_{kl}^a}\left(a_{1kl}^s(\lambda_{kl}^a) + B_{kl}(\lambda_{kl}^a)\right) - x_{0kl}^{\lambda_{kl}^r}\left(a_{1kl}^s(\lambda_{kl}^r) + B_{kl}(\lambda_{kl}^r)\right)\right) \tag{3-45}$$

Where x_{0kl} is defined as:

$$x_{0kl} = \frac{\sigma_{kl}}{d_{kl}} \tag{3-46}$$

The terms a_{1kl}^s and B_{kl} are functions of the exponents in the Mie potential. They can be calculated using either the repulsion or attraction term. B_{kl} is defined as:

$$B_{kl}(\lambda_{kl}) = 2\pi\rho_s d_{kl}^3 \varepsilon_{kl}\left(\frac{1-\zeta_x/2}{(1-\zeta_x)^3} I(\lambda_{kl}) - \frac{9\zeta_x(1+\zeta_x)}{2(1-\zeta_x)^3} J(\lambda_{kl})\right) \tag{3-47}$$

Where ζ_x represents the density of a hypothetical pure fluid calculated using the segment density and the unlike effective hard-sphere diameter d_{kl} :

$$\zeta_x = \frac{\pi\rho_s}{6} \sum_{k}^{Ng} \sum_{l}^{Ng} x_{s_k} x_{s_l} d_{kl}^3 \tag{3-48}$$

The unlike effective hard-sphere diameter (d_{kl}) is calculated from a simple arithmetic mean. The functions $I(\lambda_{kl})$ and $J(\lambda_{kl})$, similarly to a_{1kl}^s and B_{kl}, are functions of the exponents in the Mie potential. These terms are algebraic expressions used to simplify the first part of the integration of the first-order perturbation term for a Mie potential, as described by LAFITTE *et al.* (2013)

The $I(\lambda_{kl})$ and $J(\lambda_{kl})$ terms are calculated as follows:

$$I = \int_1^{x_{0kl}} \frac{x^2}{x^{\lambda_{kl}}} dx = \frac{1 - x_{0kl}^{3-\lambda_{kl}}}{\lambda_{kl} - 3} \quad 3\text{-}49$$

$$J = \int_1^{x_{0kl}} \frac{x^3 - x^2}{x^{\lambda_{kl}}} dx = \frac{1 - x_{0kl}^{4-\lambda_{kl}}(\lambda_{kl} - 3) + x_{0kl}^{3-\lambda_{kl}}(\lambda_{kl} - 4)}{(\lambda_{kl} - 3)(\lambda_{kl} - 4)} \quad 3\text{-}50$$

The term a_{1kl}^s presented in Equation 3-45 refers to the first-order perturbation term of a fluid composed by hard spheres of diameter d_{kl} interacting through the Sutherland potential, with well-depth potential energy of ε_{kl} and attractive interaction range of λ_{kl}^a. Originally, the calculation of this term requires the integration of the radial distribution function for the hard spheres reference fluid, however, Pappoiannou derived an analytical expression for this term using the mean-value theorem:

$$a_{1kl}^s = -2\pi\rho_s d_{kl}^3 \varepsilon_{kl} \left(\frac{1 - \frac{\zeta_x^{eff}}{2}}{\left(1 - \zeta_x^{eff}\right)^3 (\lambda_{kl} - 3)} \right) \quad 3\text{-}51$$

Where ζ_x^{eff} is the effective packing fraction, which was parameterized for exponents of the potential in the range $5 < \lambda_{kl} \leq 100$. It is calculated as follows:

$$\zeta_x^{eff.} = \sum_n^4 c_{nkl} \zeta_x^n \quad 3\text{-}52$$

Where the c_{ikl} coefficients are obtained as functions of the attractive or repulsive exponents of the potential:

$$\begin{pmatrix} c_{1kl} \\ c_{2kl} \\ c_{3kl} \\ c_{4kl} \end{pmatrix} = \begin{pmatrix} 0.81096 & 1.7888 & -37.578 & 92.284 \\ 1.0205 & -19.341 & 151.26 & -463.50 \\ -1.9057 & 22.845 & -228.14 & 973.92 \\ 1.0885 & -6.1962 & 106.98 & -677.64 \end{pmatrix} \times \begin{pmatrix} 1 \\ \lambda_{kl}^{-1} \\ \lambda_{kl}^{-2} \\ \lambda_{kl}^{-3} \end{pmatrix} \quad 3\text{-}53$$

An analytical expression for second-order perturbation term, a_{2kl}, is derived using the macroscopic compressibility approximation (MCA), along with a correction for soft potentials (ZHANG, 1999). a_{2kl} can be calculated as:

$$a_{2kl} = \frac{1}{2}K^{hs}(1+\chi_{kl})\varepsilon_{kl}C_{kl}^2\Bigg(x_{0kl}^{2\lambda_{kl}^a}\left(a_{1kl}^s(2\lambda_{kl}^a)+B_{kl}(2\lambda_{kl}^a)\right) - 2x_{0kl}^{\lambda_{kl}^a+\lambda_{kl}^r}\left(a_{1kl}^s(\lambda_{kl}^a+\lambda_{kl}^r)+B_{kl}(\lambda_{kl}^a+\lambda_{kl}^r)\right) + x_{0kl}^{2\lambda_{kl}^r}\left(a_{1kl}^s(2\lambda_{kl}^r)+B_{kl}(2\lambda_{kl}^r)\right)\Bigg) \quad 3\text{-}54$$

The value of K^{hs}, representing the isothermal compressibility of the hypothetical vdW one-fluid system, is obtained using the Carnahan and Starling expression, given as:

$$K^{hs} = \frac{(1-\zeta_x)^4}{1+4\zeta_x+4\zeta_x^2-4\zeta_x^3+\zeta_x^4} \quad 3\text{-}55$$

The correction factor χ_{kl} is obtained by the expression:

$$\chi_{kl} = f_1\zeta_x^* + f_2{\zeta_x^*}^5 + f_3{\zeta_x^*}^8 \quad 3\text{-}56$$

The terms f_m , for m = 1,2, … 6, are functions of α_{kl}, which is a dimensionless representation of the integrated van der Waals energy of the Mie potential, expressed as follows:

$$\alpha_{kl} = \frac{1}{\varepsilon_{kl}\sigma_{kl}^3}\int_\sigma^\infty \phi_{kl}^{\mathrm{Mie}}(r)r^2\mathrm{d}r = C_{kl}\left(\frac{1}{\lambda_{kl}^{\mathrm{a}}-3}-\frac{1}{\lambda_{kl}^{\mathrm{r}}-3}\right) \quad 3\text{-}57$$

Note that the term in equation 3-54 is not the same in equation 3-46, which uses de diameter d_{kl}. Instead, it is calculated using de unlike diameter σ_{kl}:

$$\sigma_x^3 = \sum_k^{Ng}\sum_l^{Ng} x_{\mathrm{s}_k}x_{\mathrm{s}_l}\sigma_{kl}^3 \quad 3\text{-}58$$

Therefore, can be calculated as follows:

$$\zeta_x^* = \frac{\pi\rho_s}{6}\sigma_x^3 \quad 3\text{-}59$$

The third-order perturbation term can be calculated using the following empirical expression:

$$\alpha_{3kl} = -\varepsilon_{kl}^{3} f_4 \zeta_x^* \; exp\left(f_5 \zeta_x^* + f_6 {\zeta_x^*}^2 \right) \tag{3-60}$$

Which is also dependent of ζ_x^*, computed as follows:

$$f_m = \frac{\sum_{n=1}^{4} \phi_{mn} \alpha_{kl}^{n}}{1 + \sum_{n=5}^{7} \phi_{mn} \alpha_{kl}^{n-3}} \qquad n = \{1, \dots, 7\} \tag{3-61}$$

Notice that this term has no dependence on temperature, as the diameter σ_{kl} is used for the packaging fraction ζ_x^*, instead of d_{kl}. The values for ϕ_{mn} can be found in Table 1.

Table 1: ϕ_{mn} coefficients. They are utilized for the term α_{2kl} as presented in Equation 3-54, the term α_{3kl} as outlined in Equation 3-60, and $\gamma_{c,ii}$ of the g_2 term depicted in Equation 3-80. The symbol "-" indicates that a value is not necessary for those instances

n\m	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	7.5365557	-359.44	1550.9	-1.19932	-1911.28	9236.9	10
2	-37.60463	1825.6	-5070.1	9.063632	21390.175	-129430	10
3	71.745953	-3168	6534.6	-17.9482	-51320.7	357230	0.57
4	-46.83552	1884.2	-3288.7	11.34027	37064.54	-315530	-6.7
5	-2.467982	-0.82376	-2.7171	20.52142	1103.742	1390.2	-8
6	-0.50272	-3.1935	2.0883	-56.6377	-3264.61	-4518.2	-
7	8.0956883	3.709	0	40.53683	2556.181	4241.6	-

3.2.4. Chain Formation Term

The expression for the Helmholtz free energy contribution due to chain formation follows the original Wertheim TPT1 formulation, but adapted for fused heteronuclear segments. To do this, some average molecular parameters are defined, based on the group-specific parameters, namely the average molecular segment size $\bar{\sigma}_{ii}$, as well as the average exponent range $\bar{\lambda}_{ii}$ and well-depth $\bar{\varepsilon}_{ii}$ of dispersion interactions. These average parameters are obtained by means of appropriate combining rules based on the group-specific parameter. These parameters are function of molecular fraction z_{ki}, which is the fraction of group type k in the molecular structure of component i, defined as:

$$z_{ki} = \frac{\nu_{ki}\nu_k^* S_k}{\sum_l^{Ng} \nu_{li}\nu_l^* S_l} \quad 3\text{-}62$$

The effective molecular parameters are:

$$\bar{\sigma}_{ii}^3 = \sum_k^{Ng}\sum_l^{Ng} z_{ki} z_{li} \sigma_{kl}^3 \quad 3\text{-}63$$

$$\bar{d}_{ii}^3 = \sum_k^{Ng}\sum_l^{Ng} z_{ki} z_{li} d_{kl}^3 \quad 3\text{-}64$$

$$\bar{\varepsilon}_{ii} = \sum_k^{Ng}\sum_l^{Ng} z_{ki} z_{li} \varepsilon_{\mathrm{kl}} \quad 3\text{-}65$$

$$\bar{\lambda}_{ii} = \sum_k^{Ng}\sum_l^{Ng} z_{ki} z_{li} \lambda_{\mathrm{kl}} \quad 3\text{-}66$$

Where $\bar{\lambda}_{ii}$ is generalized for attractive or repulsive exponents. It's important to notice that none of these expressions are function of the compositions, which makes the computation simpler. Therefore, the resulting contribution to the free energy of mixture due to the formation of chains

$$\frac{\mathrm{A}^{chain}}{\mathrm{Nk_bT}} = -\sum_i^{Nc} x_i \sum_k^{Ng} (\nu_{ki}\nu^* S_k - 1)\ln\left(g_{ii}^{Mie}(\bar{\sigma}_{ii}, \zeta_{\mathrm{x}})\right) \quad 3\text{-}67$$

Where the function g_{ii}^{Mie} is the radial distribution function (RDF) at the effective contact distance $\bar{\sigma}_{ii}$ and the temperature-dependent packing fraction ζ_{x} of the mixture. As described by Laffite et al., the RDF can be approximated by a second-order expansion:

$$g_{ii}^{Mie} = g_d^{hs}(\bar{\sigma}_{ii}) exp\left(\frac{\frac{\bar{\varepsilon}_{ii} g_1(\bar{\sigma}_{ii})}{k_b T} + \left(\frac{\bar{\varepsilon}_{ii}}{k_b T}\right)^2 g_2(\bar{\sigma}_{ii})}{g_d^{hs}(\bar{\sigma}_{ii})}\right) \quad 3\text{-}68$$

Where the initial term in the expansion, denoted as g_d^{hs}, represents the RDF for a reference hard-sphere mixture, each with a diameter of $\bar{d}_{ii}$. This function is specifically assessed at a distance of $\bar{\sigma}_{ii}$ and a density of ζ_x. As detailed by Pappaioannou *et al.*, the function can be calculated by the following Boublík's expression:

$$g_d^{hs} = exp\left(\sum_{m=0}^{3} k_m \bar{x}_{0ii}^m\right) \quad 1 < \bar{x}_{0ii} < \sqrt{2} \quad 3\text{-}69$$

Where the coefficients k_m in this expression are obtained as:

$$k_0 = -\ln(1-\zeta_x) + \frac{42\zeta_x - 39\zeta_x^2 + 9\zeta_x^3 - 2\zeta_x^4}{6(1-\zeta_x)^3} \quad 3\text{-}70$$

$$k_1 = \frac{\zeta_x^4 + 6\zeta_x^2 - 12\zeta_x}{2(1-\zeta_x)^3} \quad 3\text{-}71$$

$$k_2 = \frac{-3\zeta_x^2}{8(1-\zeta_x)^2} \quad 3\text{-}72$$

$$k_3 = \frac{-\zeta_x^4 + 3\zeta_x^2 + 3\zeta_x}{6(1-\zeta_x)^3} \quad 3\text{-}73$$

The first-order term of the expansion 3-78, denoted as g_1, is estimated through approximation using its value at the point of contact ($g_1(\bar{\sigma}_{ii}) \approx g_1(\bar{d}_{ii})$):

$$g_1 = \frac{1}{2\pi\bar{\varepsilon}_{ii}\bar{d}_{ii}^3} \times \left[3\frac{\partial \bar{a}_{1,ii}}{\partial \rho_s} - \bar{C}_{ii}\bar{\lambda}_{ii}^a \bar{x}_{0ii}^{\bar{\lambda}_{ii}^a} \frac{\bar{a}_{1ii}^s(\bar{\lambda}_{ii}^a) + \bar{B}_{ii}(\bar{\lambda}_{ii}^a)}{\rho_s} + \bar{C}_{ii}\bar{\lambda}_{ii}^r \bar{x}_{0ii}^{\bar{\lambda}_{ii}^r} \frac{\bar{a}_{1ii}^s(\bar{\lambda}_{ii}^r) + \bar{B}_{ii}(\bar{\lambda}_{ii}^r)}{\rho_s}\right] \quad 3\text{-}74$$

In the given expression, $\bar{C}_{ii}$, $\bar{a}^{s}_{1ii}$ and $\bar{B}_{ii}$ are associated with the effective molecular interaction potential of component 'i'. They are calculated using the values of the average molecular repulsive or attractive exponents, denoted as $\bar{\lambda}^{\mathrm{r}}_{ii}$ and $\bar{\lambda}^{\mathrm{a}}_{ii}$, respectively. Additionally, these terms as well as $\bar{a}_{1,ii}$ and $\bar{x}_{0ii}$, are determined utilizing the effective molecular parameters, detailed in Equations 3-63 – 3-66.

$$\bar{B}_{ii}(\bar{\lambda}_{ii}) = 2\pi\rho_{\mathrm{s}}\bar{d}^{3}_{ii}\bar{\varepsilon}_{ii}\left[\frac{1-\zeta_{\mathrm{x}}/2}{(1-\zeta_{\mathrm{x}})^{3}}\bar{I}(\bar{\lambda}_{ii}) - \frac{9\zeta_{\mathrm{x}}(1+\zeta_{\mathrm{x}})}{2(1-\zeta_{\mathrm{x}})^{3}}\bar{J}(\bar{\lambda}_{ii})\right] \quad 3\text{-}75$$

$$\bar{a}_{1,ii} = \bar{C}_{ii}\left[\bar{x}^{\bar{\lambda}^{\mathrm{a}}_{ii}}_{0,ii}\left(\bar{a}^{\mathrm{s}}_{1,ii}(\bar{\lambda}^{\mathrm{a}}_{ii}) + \bar{B}_{ii}(\bar{\lambda}^{\mathrm{a}}_{ii})\right) - \bar{x}^{\bar{\lambda}^{r}_{ii}}_{0,ii}\left(\bar{a}^{\mathrm{s}}_{1,ii}(\bar{\lambda}^{\mathrm{r}}_{ii}) + \bar{B}_{ii}(\bar{\lambda}^{\mathrm{r}}_{ii})\right)\right] \quad 3\text{-}76$$

$$\bar{a}^{\mathrm{s}}_{1,ii}(\bar{\lambda}_{ii}) = -2\pi\rho_{\mathrm{s}}\left(\frac{\bar{\varepsilon}_{ii}\bar{d}^{3}_{ii}}{\bar{\lambda}_{ii}-3}\right)\frac{1-\bar{\zeta}^{\mathrm{eff.}}_{ii}/2}{\left(1-\bar{\zeta}^{\mathrm{eff.}}_{ii}\right)^{3}} \quad 3\text{-}77$$

The effective packing fraction, $\bar{\zeta}^{\mathrm{eff.}}_{ii}$, is calculated as follows:

$$\bar{\zeta}^{eff.}_{ii} = \sum_{n=1}^{4}\bar{c}_{nii}\zeta^{n}_{x} \quad 3\text{-}78$$

Where the $\bar{c}_{nii}$ coefficients are obtained as functions of the average molecular repulsive or attractive exponents:

$$\begin{pmatrix}\bar{c}_{1,ii}\\ \bar{c}_{2,ii}\\ \bar{c}_{3,ii}\\ \bar{c}_{4,ii}\end{pmatrix} = \begin{pmatrix}0.81096 & 1.7888 & -37.578 & 92.284\\ 1.0205 & -19.341 & 151.26 & -463.50\\ -1.9057 & 22.845 & -228.14 & 973.92\\ 1.0885 & -6.1962 & 106.98 & -677.64\end{pmatrix}\times\begin{pmatrix}1\\ \bar{\lambda}^{-1}_{ii}\\ \bar{\lambda}^{-2}_{ii}\\ \bar{\lambda}^{-3}_{ii}\end{pmatrix} \quad 3\text{-}79$$

Similar to the first-order term, g_1, the approximation for the second-order term, g_2, is also derived based on its value at the effective diameter, $\bar{d}_{ii}$ ($g_2(\bar{\sigma}_{ii}) \approx g_2(\bar{d}_{ii})$). The calculation of this term is derived by Lafitte, and it involves using an approximation expression for macroscopic compressibility and adding an empirical correction:

$$g_2 = (1+\gamma_{\mathrm{c_{ii}}})g^{\mathrm{MCA}}_{2} \quad 3\text{-}80$$

Where $\gamma_{c_{ii}}$ is the correction factor. It can also be calculated using the averaged parameters:

$$\gamma_{c,ii} = \phi_{7,1}\left(-\tanh\left(\phi_{7,2}(\phi_{7,3} - \bar{\alpha}_{ii})\right) + 1\right) \times \zeta_x^* \, e^{\frac{\bar{\varepsilon}_{ii}}{k_b T} - 1} \exp\left(\phi_{7,4}\zeta_x^* + \phi_{7,5}(\zeta_x^*)^2\right) \quad 3\text{-}81$$

Where the $\phi_{7,n}$ coefficients are given in the Table 1. The expression for $\bar{\alpha}_{ii}$ mirrors that of α_{kl}, however, it distinctively employs averaged parameters for its calculation:

$$\bar{\alpha}_{ii} = \bar{C}_{ii}\left(\frac{1}{\bar{\lambda}_{ii}^{a} - 3} - \frac{1}{\bar{\lambda}_{ii}^{r} - 3}\right) \quad 3\text{-}82$$

The term comprising the macroscopic compressibility approximation is also calculated using the averaged parameters and it is based on the fluctuation term of the Sutherland potential:

$$g_2^{MCA}(\bar{d}_{ii}) = \frac{1}{2\pi\bar{\varepsilon}_{ii}^2\bar{d}_{ii}^3}\left[3\frac{\partial}{\partial\rho_s}\left(\frac{\bar{a}_{2ii}}{1+\bar{\chi}_{ii}}\right) - \bar{\varepsilon}_{ii}K^{hs}\bar{C}_{ii}^2\bar{\lambda}_{ii}^{r}\bar{x}_{0ii}^{2\bar{\lambda}_{ii}^{r}}\frac{\bar{a}_{1ii}^{s}(2\bar{\lambda}_{ii}^{r}) + \bar{B}(2\bar{\lambda}_{ii}^{r})}{\rho_s} + \bar{\varepsilon}_{ii}K^{HS}\bar{C}_{ii}^2(\bar{\lambda}_{ii}^{r} + \bar{\lambda}_{ii}^{a})\bar{x}_{0ii}^{(\bar{\lambda}_{ii}^{r}+\bar{\lambda}_{ii}^{a})}\frac{\bar{a}_{1ii}^{s}(\bar{\lambda}_{ii}^{r} + \bar{\lambda}_{ii}^{a}) + \bar{B}(\bar{\lambda}_{ii}^{r} + \bar{\lambda}_{ii}^{a})}{\rho_s} - \bar{\varepsilon}_{ii}K^{HS}\bar{C}_{ii}^2\bar{\lambda}_{ii}^{a}\bar{x}_{0ii}^{2\bar{\lambda}_{ii}^{a}}\frac{\bar{a}_{1ii}^{s}(2\bar{\lambda}_{ii}^{a}) + \bar{B}(2\bar{\lambda}_{ii}^{a})}{\rho_s}\right] \quad 3\text{-}83$$

Where the expression for $\bar{\chi}_{ii}$ is:

$$\bar{\chi}_{ii} = f_1(\bar{\alpha}_{ii})\zeta_x^* + f_2(\bar{\alpha}_{ii})(\zeta_x^*)^5 + f_3(\bar{\alpha}_{ii})(\zeta_x^*)^8 \quad 3\text{-}84$$

Where the coefficients f_m are calculated similarly as in the Equation 3-61, but applying $\bar{\alpha}_{ii}$ instead of α_{kl}.

3.2.5. Association Term

The term of the contribution to Helmholtz energy related to the association between molecules through highly directional and short-range molecular interactions follows similar expressions to the work of WERTHEIM (1986), but is applied to group contribution methodology:

$$\frac{\mathrm{A^{assoc.}}}{\mathrm{Nk_BT}} = \sum_{i}^{N_c} x_i \sum_{k}^{N_g} v_{ki} \sum_{a}^{N_{ST_k}} n_{ka} \left(\ln X_{ika} + \frac{1 - X_{ika}}{2} \right) \quad 3\text{-}85$$

Where N_{ST_k} denotes count of distinct sites within group 'k', n_{ka} represents the quantity of 'a'-type sites within group 'k', and X_{ika} signifies the proportion of unbounded molecules belonging to component 'i' with respect to group 'k' through the site 'a'. The determination of X_{ika} can be achieved through an iterative solution to the following system of non-linear equations, as originally proposed by TAN *et al.* (2004), although adapted for application within the context of group contribution:

$$X_{ika} = \frac{1}{1 + \sum_{j}^{N_c} \sum_{l}^{N_g} \sum_{b}^{N_{ST_l}} \rho x_j v_{lj} n_{lb} X_{jlb} \Delta_{ijklab}} \quad 3\text{-}86$$

Here, similarly to the PC-SAFT, Δ_{ijklab} represents the overall strength of interaction due to the association of the site 'a' of the group 'k' in the molecule 'i' with the site 'b' of the group 'l' in the molecule 'j'. It can be calculated as follows:

$$\Delta_{ijklab} = \bar{\sigma}_{ij}^3 F_{klab} K_{klab} I_{klab} \quad 3\text{-}87$$

Where F_{klab} is the Mayer-f function of the square-well association potential, calculated as:

$$F_{klab} = \exp\left(\frac{\varepsilon_{klab}}{k_b T}\right) - 1 \quad 3\text{-}88$$

K_{klab}, defined as the bonding volume, can be calculated using the method outlined by LAFFITE *et al*. (2013). In this study, similar to the method employed by

DUFAL *et al.* (2014), K_{klab} is approached as an estimated parameter. $\mathrm{I_{ijklab}}$, also known as the association kernel, is the association integral for a Lennard-Jones monomer. It is expressed as a temperature-density polynomial correlation, derived from molecular simulations:

$$\mathrm{I_{ijklab}} = \sum_{p}^{10} \sum_{q}^{10-p} c_{pq} (\rho \sigma_x^3)^p \left(\frac{\mathrm{k_b T}}{\bar{\varepsilon}_{ij}} \right)^q \tag{3-89}$$

3.3. Automatic Differentiation

Differentiation is an essential practice in thermodynamics, particularly in problems involving minimization. These derivatives are crucial for use in Newton-based programs that perform phase equilibrium and other process engineering calculations. There are four main methodologies for computing derivatives in chemical process calculations. The first methodology involves manually deriving expressions and hard-coding the derivatives. This approach is the most straightforward and likely demands the least computational effort and time. Moreover, it is not subject to truncation errors. However, this method becomes increasingly complicated as the complexity of the function increases, making it more prone to human error. Furthermore, not all derivatives will have analytical expressions. This is because modeling results are often obtained as implicit solutions of complex systems of linear or non-linear equations, which do not allow for explicit and closed derivation of model results in relation to parameters or state variables.

Another alternative for obtaining derivatives is numerical derivation. This methodology is also straightforward and is the simplest among them. This is because it does not require any specific extra coding for each function. In fact, it utilizes the finite difference approximation using values of the original function evaluated at points near where we want to derive. For example, we can approximate the derivative of a function of one variable $f(x)$ using a center difference approximation:

$$\frac{\partial f(x)}{\partial x} \cong \frac{f(x+h) - f(x-h)}{2h} \tag{3-90}$$

where h is a small step value. Note that implementing this derivative is quite simple, but requires the function to be evaluated at least twice. For greater accuracy in the derivative value, the function must be evaluated at more points. Moreover, this is the case of a function with only one variable that returns a scalar. For a function f: $R^n \rightarrow R^m$, computing the Jacobian matrix of f would require at least 2mn evaluations.

Truncation errors occur when an infinite process (like a series) is 'truncated' or cut short to make calculations feasible. In the context of numerical differentiation, truncation errors arise when we approximate the derivative using discrete intervals (like h). The smaller the interval, the closer the approximation to the true derivative, but only to a certain extent. Rounding errors, on the other hand, are due to the limited precision of numerical representation in computers. As h becomes very small, the difference between $f(x+h)$ and $f(x-h)$ can become comparable to the precision limit of the computer, leading to significant rounding errors. This creates a trade-off in choosing h: while a smaller h reduces truncation error, it increases rounding error, and vice versa. Balancing these errors is crucial for achieving accurate numerical differentiation.

To elucidate this challenge in numerical differentiation, consider the modeling of methane gas at a temperature of 25°C and with density of 40mol/m^3 using PC-SAFT. The compressibility factor (Z), can be derived from the residual Helmholtz free energy as a function of density. The Z is defined by the following relationship:

$$Z(\rho, \mathrm{T}) = 1 + \rho \frac{\partial}{\partial \rho}\left(\frac{A^{res}(\rho, \mathrm{T})}{\mathrm{N k_B T}}\right) \qquad 3\text{-}91$$

This formulation entails computing the partial derivative of the residual Helmholtz free energy with respect to density.

Figure 1 provides a visual representation of the variation in computational error for the compressibility factor using the centered difference method for numerical differentiation at ρ = 40mol/m^3. The graph plots the error magnitude against the step size (h) and illustrates a clear trend of decreasing error with diminishing step size until a critical threshold is reached, beyond which the error sharply increases. In practice, a step size that is too large leads to high truncation errors, while a step size that is too small results in rounding errors due to finite numerical precision in computer

calculations. The red line shows the automatic differentiation (AD) error for comparison, which will be explained later.

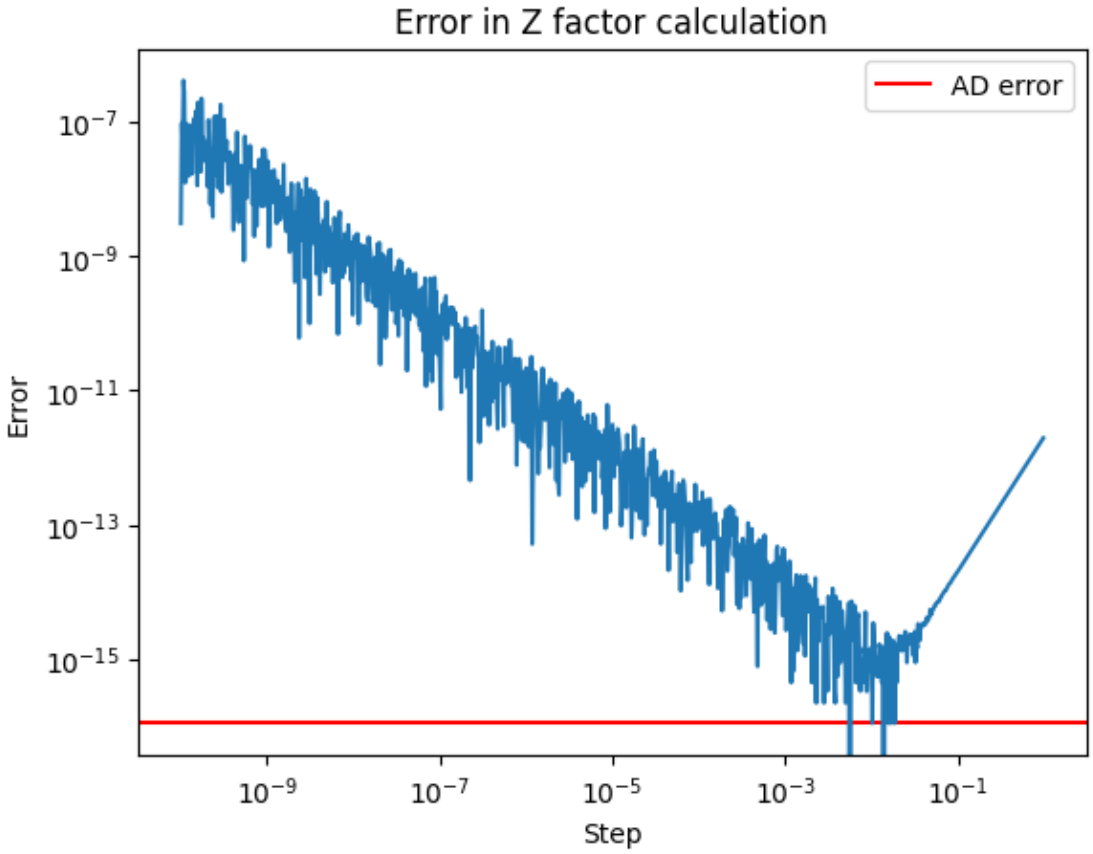

Figure 1: Error in center difference (Equation 3-90) approximations as a function of step size (h), for the calculation of Z (Equation 3-91) using derivative of A^{res}. Plotted errors are computed using the deviation from symbolically calculated Z at T = 25°C and ρ = 40mol/m^3.

Symbolic computation is the third methodology for calculating derivatives. This method essentially automates the manual differentiation process, utilizing computer programs to perform mathematical operations with symbols that represent values. These programs manipulate expressions, automatically applying transformation rules that correspond to the principles of differentiation. As a result, the process of differentiating an expression tree becomes a streamlined, mechanical task, well-suited for automation, and eliminates the need for repetitive manual calculations.

This approach mitigates the errors inherent in numerical derivation by enabling the computation of derivatives with the precision afforded by computer calculations. However, it is not without its challenges. A prevalent issue in symbolic computation is

"expression swell." This phenomenon occurs when the expressions for derivatives become significantly larger than the original expressions. This is often a result of certain differentiation rules, such as the product rule, which naturally lead to duplicated expressions. For instance, consider the recurrence relation known as the logistic map, depicted in Table 2, Taken from BAYDIN *et al.*, 2018.

Table 2: illustration on the progression of derivatives calculated using symbolic computation, specifically through the example of the logistic map. This table showcases how derivatives evolve from simple expressions to significantly more complex forms due to expression swell, a common challenge in symbolic differentiation.

n	l_n	$\frac{\partial}{\partial x} l_n$	$\frac{\partial}{\partial x} l_n$ (simplified)
1	x	1	1
2	$4x(1 - x)$	$4(1 - x) - 4x$	$4 - 8x$
3	$16x(1 - x)$ $(1 - 2x)^2$	$16(1 - x)(1 - 2x)^2$ $- 16x(1 - 2x)^2$ $- 64x(1 - x)(1 - 2x)$	$16(1 - 10x + 24x^2$ $- 16x^3)$
4	$64x(1 - x)$ $(1 - 2x)^2$ $(1 - 8x + 8x^2)^2$	$128x(1-x)(-8+ 16x)$ $(1- 2x)^2(1-8x+ 8x^2)$ $+64(1- x)$ $(1-2x)^2(1-8x+ 8x2)^2$ $-64x(1- 2x)^2$ $(1- 8x+8x^2)^2$ $-256x(1-x)(1-2x)$ $(1-8x+8x^2)^2$	$64(1 - 42x + 504x^2$ $- 2640x^3 + 7040x^4$ $- 9984x^5 + 7168x^6$ $- 2048x^7)$

When n is 1 or 2, the derivative's expression remains relatively simple. However, as further derivatives are taken, the resulting expression exponentially increases in size. Regularly applying simplification algorithms during differentiation can help manage expression swell. These algorithms identify and remove redundant terms, factor expressions, and perform other algebraic simplifications. Simplifying expressions for polynomials to match the size of the original function may be straightforward, as demonstrated in the third column of the Table 2. However. This kind of simplification is not always applicable for other types of expressions. This can be exemplified by the

softplus ReLu function, which is commonly used as an activation function in neural networks:

$$f(x; w, b) = \log(1 + e^{wx+b}) \quad 3\text{-}92$$

Where x is the input, b is the bias and w is the weight. When training a neural network, the derivative of the activation function is calculated concerning the weighted input, where the goal is to update the weights and the bias in the network to minimize the loss function. Let's consider a simple neural network with two layers (input layer and output layer) and one neuron in each layer. The final output would be:

$$g(x; w_1, w_2, b_1, b_2) = \log(1 + e^{w_2 \log(1+e^{w_1 x+b_1})+b_2}) \quad 3\text{-}93$$

The first derivative of this function in relation to the second weight in a neural network with merely two layers results in a large expression:

$$\frac{\partial g}{\partial w_2} = \frac{e^{b_2}(e^{b_1 + x\, w_1} + 1)^{w_2} \log(e^{b_1 + x\, w_1} + 1)}{e^{b_2}(e^{b_1 + x\, w_1} + 1)^{w_2} + 1} \quad 3\text{-}94$$

Where the subscript represents the layer of the neural network. As the complexity of the functions involved increases, the symbolic representation of derivatives demands significantly more memory. Given that neural network architectures commonly comprise dozens of layers, this method of differentiation becomes impractical.

Furthermore, symbolic differentiation only works in the functions that are expressed in closed forms, which restricts the ability to work with control flow mechanisms, such as conditionals or loops.

The final methodology explored here is automatic differentiation (AD). AD computes derivatives with precision comparable to symbolic differentiation but operates directly on the program of interest. Unlike producing an expression for the derivative, automatic differentiation yields the numerical value of the derivative. AD transforms a program that calculates numerical values of a function into one that computes numerical

values of the function's derivatives. The core mechanism of AD involves converting the text of a program into the text of a transformed program that computes the desired derivative values. It accomplishes these calculations by systematically employing the chain rule from elementary calculus to floating-point numerical values, bypassing the need for manipulation of symbolic expressions. This process inherently involves decomposing complex functions into elementary operations for which the derivatives are known. These operations include basic arithmetic (addition, subtraction, multiplication, division) and elementary functions (exponential, logarithmic, sine, cosine, etc.).

In the realm of AD, the computation of a function is approached by establishing an evaluation trace. The evaluation trace is essentially a step-by-step breakdown of the computational process, listing each elementary operation, the sequence of intermediate variables generated, and the operations that compute them. This process is not merely a replication of the function's computational graph but an enhancement of it, incorporating derivative computation simultaneously. To Illustrate this, let's consider the simple function taken from MARGOSSIAN (2019):

$$f(x,y,z) = -\frac{1}{2}\left(\frac{y-x}{z}\right)^2 - \log(z) - \frac{1}{2}\log(2\pi) \qquad 3\text{-}95$$

We aim to compute the derivatives of f with respect to x. First, let's see how the computer execute the sequence of operations for f when x = 5, y = 10 and z = 2 in the Table 3:

Table 3: Forward Mode Automatic Differentiation of $f(x,y,z) = -\frac{1}{2}\left(\frac{y-x}{z}\right)^2 - \log(z) - \frac{1}{2}\log(2\pi)$. This table demonstrates the process of forward mode AD for the function f, evaluated at the point (5,10,2) with the objective of computing $\frac{\partial f}{\partial x}$. The table is structured into two sections: the left side detailing the primal calculations and the right side outlining the corresponding derivative operations.

	Variables	Numerical value		Variables	Numerical value
v_1	y	10	v_1'	$\frac{\partial v_1}{\partial v_2}$	0
v_2	x	5	v_2'	$\frac{\partial v_2}{\partial v_2}$	1
v_3	z	2	v_3'	$\frac{\partial v_3}{\partial v_2}$	0
v_4	$v_1 - v_2$	5	v_4'	$\frac{\partial v_4}{\partial v_1}v_1' + \frac{\partial v_4}{\partial v_2}v_2'$	$0 + (-1) \times 1 = -1$
v_5	v_4/v_3	2.5	v_5'	$\frac{\partial v_5}{\partial v_4}v_4' + \frac{\partial v_5}{\partial v_3}v_3'$	$\frac{1}{2}(-1) + 0 = -0.5$
v_6	v_5^2	6.25	v_6'	$\frac{\partial v_6}{\partial v_5}v_5'$	$2 \times 2.5(-0.5) = -2.5$
v_7	$-0.5v6$	-3.125	v_7'	$\frac{\partial v_7}{\partial v_6}v_6'$	$-0.5\,(-2.5) = 1.25$
v_8	$log(v3)$	$log(2) = 0.30103$	v_8'	$\frac{\partial v_8}{\partial v_3}v_3'$	0
v_9	$v7 - v8$	$3.125 - 0.30103 = 2.82397$	v_9'	$\frac{\partial v_9}{\partial v_7}v_7' + \frac{\partial v_9}{\partial v_8}v_8'$	$1 \times 1.25 + 0 = 1.25$
v_{10}	$v_9 - 0.5\log(2\pi)$	$2.82397 - 0.5log(2\pi) = 2.4248$	v_{10}'	$\frac{\partial v_{10}}{\partial v_9}v_9'$	1.25

The evaluation trace, presented in Table 3, methodically lists a sequence of distinct mathematical variables, each representing specific functions and their respective numerical values. For each variable v_i in this sequence, we compute the derivative with respect to an independent variable, x, by accompanying a corresponding derivative variable, $v_i' = \frac{\partial v_i}{\partial x}$. This derivative computation is known as Forward mode, and efficiently executed by systematically applying the chain rule at each step of the trace. Additionally, it's important to note that another mode of AD, Reverse mode, exists which calculates derivatives in a different manner, backpropagating from the output to

the inputs. However, we are not using this mode in our work, primarily because it is not suitable for our needs. Specifically, we utilize JAX (just after execution) for computing differentials, and its reverse mode is incompatible with loops, which are used in the association calculation involved here. JAX is a Python library designed for high-performance numerical computation and machine learning research. Note that JAX has a JIT compiler (just-in-time compiler) that automatically compiles the python code into low level instructions to obtain performance similar to c/c++/fortran for the thermodynamic models while developing applications in python environments as parameter estimation and plotting.

A key aspect of this approach is the unique definition and consistent retention of values for each mathematical variable. This practice ensures computational efficiency by avoiding redundant evaluations of subexpressions, allowing their computed results to be reused throughout the process. This strategy distinguishes this methodology from symbolic differentiation.

To calculate the value of the function and its derivative while simultaneously enhancing the efficiency of AD, we employ operator overloading. This technique involves redefining and extending basic arithmetic operations, which typically only work with simple numerical values, to operate with a different object that encapsulates both the function's value and its derivative. These objects are often implemented as data structures or classes in programming languages, capable of storing and operating on multiple pieces of related data.

A particular methodology of automatic differentiation where operator overloading is applied is with dual numbers. Dual numbers are objects, like a tuple for instance, that carry the information of both values and derivatives of a function. We can understand how a dual number functions through an analogy with complex numbers. Just as a complex number has its imaginary part tied to an "i", dual numbers have their dual part tied to a new term: "ε". It then has a distinct property, where $\varepsilon^2 = 0$ but $\varepsilon \neq 0$. Thus, a dual number x can be represented as:

$$x = x_0 + x_1\varepsilon \tag{3-96}$$

where x_0 is the real part and x_1 is the dual part. Note that operations with dual numbers conveniently mirror the rules of symbolic differentiation. Consider, for example, the polynomial P(x):

$$P(x) = a_0 + \sum_{n}^{N} a_i x^n \tag{3-97}$$

If we apply a dual number as an argument to this polynomial, we would have:

$$P(x_0 + x_1\varepsilon) = a_0 + \sum_{n}^{N} a_i (x_0 + x_1\varepsilon)^n$$
$$P(x_0 + x_1\varepsilon) = a_0 + \sum_{n}^{N} a_i x_0^n + n\varepsilon x_1 \sum_{n}^{N} a_i x_0^{n-1} \tag{3-98}$$

Note that all terms multiplied by ε with a degree greater than 1 disappear, as they equal zero. Consequently:

$$P(x_0 + x_1\varepsilon) = P(x_0) + x_1\varepsilon P'(x_0) \tag{3-99}$$

Where P' is the derivative of P. In general, any real function can be evaluated with the exact Taylor expansion. Also, the chain rule applies to the composition of functions in this representation:

$$f(g(x_0 + x_1\varepsilon)) = f(g(x_0)) + f'(g(x_0))g'(x_0)x_1\varepsilon \tag{3-100}$$

A generalization of this methodology for higher-order derivatives and cross derivatives is the hyper-dual numbers(FIKE, ALONSO, 2011). In this mechanism, hyper-dual numbers possess N extra dimensions, with analogous properties:

$$\varepsilon_j\varepsilon_k = \frac{(j+k)!}{j!\,k!}\varepsilon_{j+k}\ if\ j + k < N\ and\ \varepsilon_i\varepsilon_j = 0\ if\ j + k > N \tag{3-101}$$

Where a term multiplying the ε_i represents the ith derivative.

3.4. Parameter Estimation

In a general manner, the development of a new group utilizing the SAFT-γ Mie approach entails the derivation of group parameters through the use of experimental data as part of a parameter estimation strategy. The determination of interaction parameters is conducted successively, where the newly characterized group will have its parameters informed by those of previously established groups. Within the SAFT-γ Mie equation of state framework, each functional group is considered independent of its position within the molecule, thereby ensuring that interactions between groups remain constant and, hence, are transferable across different systems.

In this study, the interaction parameters for both like and unlike interactions of two novel groups were estimated. Prior to the estimation, these new groups were defined. The first group, referred to as "cNcCO," represents an alkylated cyclic amide, in which both the nitrogen and the carbonyl components are part of the ring structure. The selection of this group was made following the approach of GMEHLING *et al.* (2002), who estimated the parameters for several groups, including this type of amide, within the UNIFAC model framework. Gmehling *et al.* noticed a significant difference in polarity among amides, whether cyclic, aliphatic, monoalkylated, or dialkylated, indicating that a single group encompassing multiple amide types would not yield satisfactory thermodynamic property predictions.

The Figure 1, provides an illustrative representation of the groups present in the NMP molecule, including the “cNcCO” group, colored as orange:

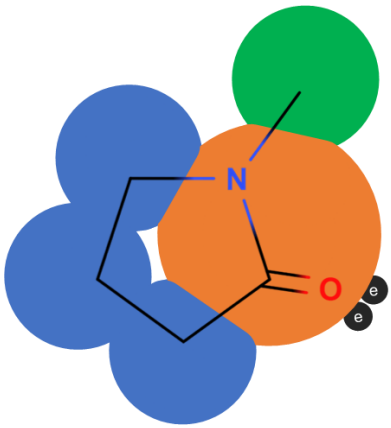

Figure 2: Stick figure and SAFT-γ Mie molecular modelling of N-Methyl-2-Pyrrolidone. This molecule is modelled by a combination of one CH_3, three cCH_2 and

one cNcCO group. Association sites are denoted by the smaller black circles, labelled e for electronegative (acceptor) sites.

For the parameter estimation of this group, liquid-vapor equilibrium data involving binary mixtures with N-Methyl-2-Pyrrolidone (NMP) and the saturation pressure data of NMP were utilized. The first objective function to be minimized was defined as:

$$\begin{aligned}\min_{\Omega} F_{obj1} = & \left(\left\|\left(\frac{\bar{T}_{VLE_0}^{exp} - \bar{T}_{VLE_0}^{calc}(\bar{\Omega})}{\bar{T}_{VLE_0}^{exp}}\right)\right\| + \left\|\left(\frac{\bar{y}_{1\,VLE_0}^{exp} - \bar{y}_{1\,VLE_0}^{calc}(\bar{\Omega})}{\bar{y}_{1\,VLE_0}^{exp}}\right)\right\|\right)\frac{\omega_0}{N_{VLE_0}} \\ & + \left(\left\|\left(\frac{\bar{P}_{VLE_1}^{exp} - \bar{P}_{VLE_1}^{calc}(\bar{\Omega})}{\bar{P}_{VLE_1}^{exp}}\right)\right\|\right. \\ & \left. + \left\|\left(\frac{\bar{y}_{1\,VLE_1}^{exp} - \bar{y}_{1\,VLE_1}^{calc}(\bar{\Omega})}{\bar{y}_{1\,VLE_1}^{exp}}\right)\right\|\right)\frac{\omega_1}{N_{VLE_1}} \\ & + \left\|\left(\frac{\bar{P}_{VLE_2}^{exp} - \bar{P}_{VLE_2}^{calc}(\bar{\Omega})}{\bar{P}_{VLE_2}^{exp}}\right)\right\|\frac{\omega_2}{N_{VLE_2}} + \left\|\left(\frac{\bar{P}_{sat}^{exp} - \bar{P}_{sat}^{calc}(\bar{\Omega})}{\bar{P}_{sat}^{exp}}\right)\right\|\frac{\omega_3}{N_{sat}} \\ & - \omega_4 \sum_{i}^{N_{par}} \min\left(\Omega_{\mathrm{i}} - \Omega_{\mathrm{i}}^{\min}; 0\right)\end{aligned} \qquad 3\text{-}102$$

Where $\|\bar{x}\|$ is the euclidean norm of vector $\bar{x}$, which is mathematically expressed as $\sqrt{\sum_i x_i^2}$. The term VLE specifically denotes vapor-liquid equilibrium data, distinguished by VLE_0 and VLE_1, which encompasses compositional information of both phases, in contrast to the VLE_2 that solely pertains to the liquid phase composition. The VLE_0 refers to isothermal data, while VLE_1 category refers to isobaric data. The computational determination of the vapor composition and the phase equilibrium pressure or temperature employs the bubble point algorithms as mentioned in section 2.2.1. The variable N represents the aggregate number of data points for each molecule. This methodology ensures equitable error distribution across each binary system by normalizing against the respective experimental data points, thereby mitigating biases towards any particular molecule arising from discrepancies in the volume of available data. The last part of the equation is used to prevent the parameters, symbolized by Ω,

to go under a threshold limit, $\Omega^{\min}$, by penalizing the objective function if so. N_{par} here represents total number of parameters adjusted.

The weighting of each category of data, denoted by ω, was determined through empirical methods. Notably, the saturation pressure of NMP was accorded to have greater significance within the objective function in comparison to the binary data. This prioritization is justified by the observation that the vector of parameters, which exerts a more pronounced influence on the objective function as the mole fraction of NMP increases - NMP being the sole molecule characterized by the "cNcCO" group. In scenarios where the mole fraction of NMP approaches zero within a binary mixture, the error attributed to the objective function tends towards a constant value. This is attributable to the prediction of the second component properties relying solely on the parameters already established. In light of this consideration, the ω_3 was set to 100. The weights ω_1, ω_2 and ω_3 were assigned the value 1, while the ω_4 was set to 100000.

The initial step in the minimization involved selecting certain parameters to be fixed. It was decided that a single sphere would suffice to represent this group ($v^*_{cNcCO} = 1$). The repulsion potential exponent was fixed at 6, similar to the majority of other groups ($\lambda^a_{cNcCO} = 6$). Additionally, this group was assigned a single association site, formed by a pair of electrons ($Nst_{cNcCO} = 1$ and $e_1 = 2$).

The second group, termed "cNCO," represents an alkylated cyclic amide where only the nitrogen is part of the ring. The Figure 2, provides an illustrative representation of the groups present in the praziquantel molecule, including the "cNcCO" and the "cNCO" groups, colored as orange and green, respectfully:

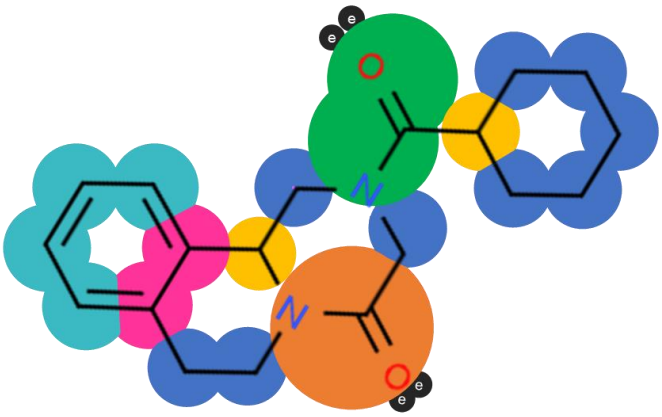

Figure 3 Stick figure and SAFT-γ Mie molecular modeling of Praziquantel. This molecule is modeled by a combination of nine cCH_2, two cCH, four aCH, two aC, one cNcCO group and one cNCO group. Association sites are denoted by the smaller black circles, labeled e for electronegative (acceptor) sites.

Solubility data of praziquantel in various solvents were employed for the estimation of this group's parameters. The second objective function to be minimized was defined as:

$$\min_{\Omega} F_{obj2} = \left(\left\| \left(\frac{\bar{x}_{sat}^{exp} - \bar{x}_{sat}^{calc}(\Omega)}{\bar{x}_{sat}^{exp}} \right) \right\| \right) \quad 3\text{-}103$$

Where $\|\bar{x}\|$ is the euclidean norm of vector $\bar{x}$, which is mathematically expressed as $\sqrt{\sum_i x_i^2}$. The term sat specifically denotes the solubility, or in other terms, the saturation of the Praziquantel in the liquid phase with different solvents. These predictions were made as presented in the Equations 2-35 in Section 2.2.3.

The initial step in the minimization was analogue to the first group. It was decided that two spheres could represent this group ($v_{cNcCO}^{*} = 2$), the repulsion potential exponent was fixed at 6 ($\lambda_{cNcCO}^{a} = 6$), and this group was assigned a single association site, formed by a pair of electrons ($Nst_{cNcCO} = 1$ and $e_1 = 2$).

A heuristic approach was adopted for parameter estimation to isolate contributions to the Helmholtz free energy from interactions between groups, thereby facilitating the objective function's minimization. Starting with the cNcCO group, the like parameters within the group, namely $\sigma_{cNcCO,cNcCO}$, S_{cNcCO}, $\varepsilon_{cNcCO,cNcCO}$, and

$\lambda^{r}_{cNcCO,cNcCO}$ along with the unlike parameters of this new group with other groups present in NMP, i.e., $\varepsilon_{CH_3,cNcCO}$ and $\varepsilon_{cCH_2,cNcCO}$ were estimated using only the saturation pressure data of pure NMP. This was achieved through the Particle Swarm Optimization (PSO) algorithm. This minimum served as the initial guess for the next step. All subsequent steps utilized the NELDER-MEAD algorithm. This procedure of using the previous result as the initial guess for the next step was repeated until the final estimation. The Unlike parameters between cNcCO and the remaining groups present in the molecules of the available VLE data, which do not associate, including $\varepsilon_{CH_3,cNcCO}$ and $\varepsilon_{cCH_2,cNcCO}$, were estimated/reestimated. This step was crucial as mixtures involving a nonpolar molecule, such as hexane, could only be described with the combination parameters between the new group and CH3 and CH2 groups, unlike a mixture with butanol, which would also be described by the association parameters with the CH2OH group, making the adjustment easier. Subsequently, the association parameters between the new group and the associative groups present in the experimental data were estimated individually. In the final stage, all parameters were reestimated together.

The second group, cNCO, the like parameters within the group, along with the unlike parameters of this new group with other groups present in the Praziquantel and hexane were estimated first using the Particle Swarm Optimization (PSO) algorithm. Similar to the other procedure, this minimum served as the initial guess for the next step and this was repeated. All subsequent steps utilized the NELDER-MEAD algorithm. The Unlike parameters between cNCO and the remaining groups present in the molecules of the available VLE data, which do not associate, were estimated/reestimated. Subsequently, the association parameters between the new group and the associative groups present in the experimental data were estimated individually. In the final stage, all parameters were reestimated together.

Finally, the binary interaction parameters (k_{ij}) for the PC-SAFT model were refined through an optimization process utilizing an objective function akin to the one employed for the second group parameter estimation. The keys distinction in this stage lies in the focus on optimizing the k_{ij} values rather than the group parameters themselves and the use of the Equation 2-34 detailed in section 2.2.2. The optimization of the k_{ij} parameters was conducted in a single iterative stage, starting with an initial

assumption of k_{ij}=0, thereby streamlining the parameter estimation process to efficiently enhance the model's predictive accuracy for solubility.

4. Results and Discussion

This section presents tables that delineate the estimated parameters for SAFT-γ Mie for both like and unlike group interactions, as well as the binary interaction parameters for PC-SAFT. It evaluates the accuracy of SAFT-γ Mie in describing experimental data on liquid-vapor and solid-liquid equilibria, alongside the efficacy of PC-SAFT in characterizing experimental data on solid-liquid equilibria for the constituents considered within this work. Furthermore, this section predicts the saturation pressure of N-Cyclohexyl-2-pyrrolidinone utilizing SAFT-γ Mie, despite being omitted from the parameter estimation process. The regression of group parameters took into account the phase equilibria of NMP and PZQ with various solvents. Tables 4 and 5 show the specific values for both like and unlike parameters for the groups under discussion.

Table 4: Group-Specific parameters within the SAFT-γ Mie framework for the groups considered in this work. Values marked with a CR are calculated using combining rules.

Group k	Group l	$\frac{\varepsilon_{kl}}{k_b T}$	λ_{kl}^r	Site a on group k	Site b on group l	$\frac{\varepsilon_{klab}}{k_b T}$	$\frac{K_{klab}}{Å^3}$
cNcCO	CH3	455.072	14.3919	-	-	-	-
cNcCO	CH2	488.016	16.4870	-	-	-	-
cNcCO	CH3COCH3	514.972	CR	-	-	-	-
cNcCO	[CH2][OCH2]	453.280	CR	-	-	-	-
cNcCO	[CH3][OCH2]	157.046	CR	-	-	-	-
cNcCO	cCH2	514.624	CR	-	-	-	-
cNcCO	CH2OH	CR	CR	e1	H	2082.45	208.854
cNcCO	CHOH	CR	CR	e1	H	76.5681	122803
cNcCO	CH3OH	CR	CR	e1	H	2689.52	40.3398
cNcCO	COOH	CR	CR	e1	H	4737.58	6.06928
cNcCO	H2O	450.885	CR	e1	H	2643.55	39.8381
cNCO	CH3	78.3501	CR	-	-	-	-
cNCO	CH2	823.014	CR	-	-	-	-
cNCO	CH3COCH3	716.484	CR	-	-	-	-
cNCO	[CH2][OCH2]	748.474	CR	-	-	-	-
cNCO	cNcCO	1217.06	CR	-	-	-	-
cNCO	COO	889.494	CR	-	-	-	-
cNCO	CH3CO	755.784	CR	-	-	-	-
cNCO	CH2OH	690.740	CR	e1	H	13.1408	900760
cNCO	CHOH	1065.54	CR	e1	H	2757.22	8.764696
cNCO	CH3OH	CR	CR	e1	H	30.2431	930735

Table 5: Group parameters for the functional groups cNcCO and cNCO within the SAFT-γ Mie framework.

group	ν^*	S_k	σ_{kk}	ε_{kk}	λ_{kk}^r	λ_{kk}^a	Nst_k	nH_k	$ne1_k$	$ne2_k$
cNcCO	1	0.998062	3.59719	753.120	13.7703	6	1	0	2	0
cNCO	2	0.75517	2.24814	1403.96	15.5354	6	1	0	2	0

As previously discussed, the Mie potential offers a significant advantage by allowing the modification of the interaction potential between segments through adjustments in the dispersive energy values and the repulsive and attractive exponents, as well as their intergroup combinations. This flexibility facilitates the detailed capture of interaction characteristics essential for accurately describing thermodynamic properties. However, employing a greater number of parameters in the regression process—specifically, estimating more unlike group combination parameters instead of calculating them—increases the risk of overfitting. Such overfitting can diminish the model's predictive accuracy. Therefore, a conscious strategy was adopted to minimize the number of unlike parameters in the regression, initially opting to keep the minimum number of unlike parameters with an acceptable error. The parameters utilized in this study are presented in Figure 3, with color coding to indicate their derivation: parameters estimated in this work are marked in green, those previously published are in blue (BURGER *et al.*, 2015, DUFAL *et al.*, 2014, FEBRA *et al.*, 2021, FERNANDES, 2023, HASLAM *et al.*, 2020, HUTACHAROEN *et al.*, 2017, PAPAIOANNOU *et al.*, 2016,

PAPAIOANNOU *et al.*, 2014, SADEQZADEH *et al.*, 2016), and orange denotes the use of combining rules.

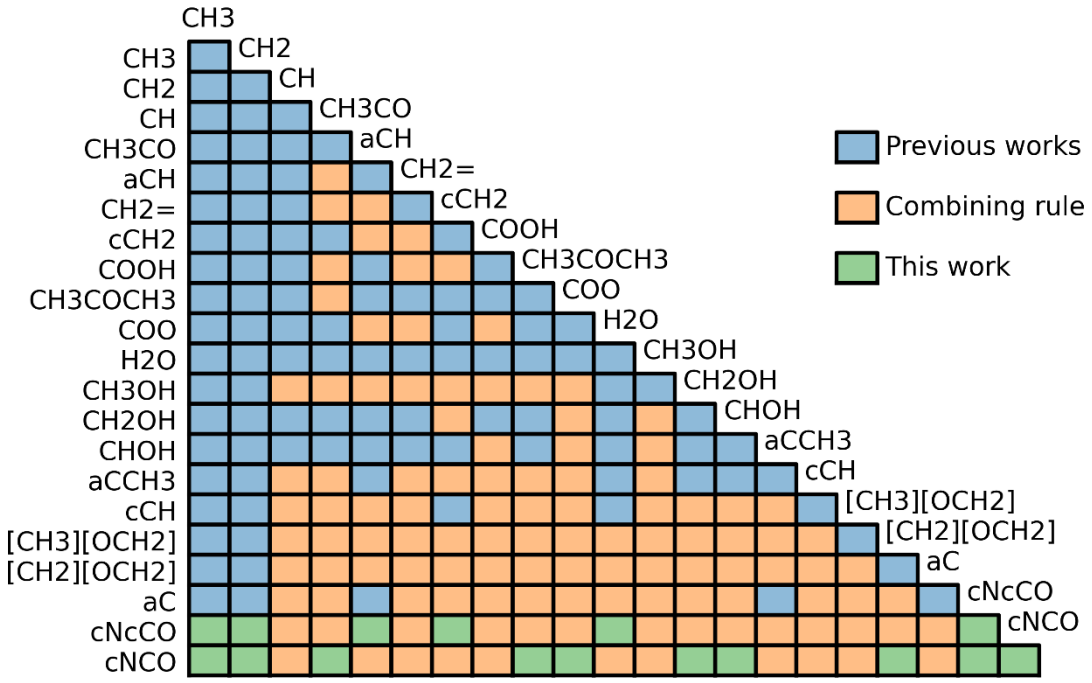

Figure 4: Groups Developed needed for use within the SAFT-γ Mie Approach applied to NMP and PZQ. The green shading indicates the intergroup interaction parameters estimated in this work, the blue shading indicates parameters estimated in previous works, while the orange shading indicates the unlike interaction parameters predicted using the combining rules.

4.1. Vapor-Liquid Equilibria

For the evaluation of liquid-vapor equilibria involving NMP or analogous molecules via the SAFT-γ Mie methodology, it is imperative to integrate the novel group parameters introduced in this study, specifically cNcCO, along with its detailed unlike interactions as cataloged in Tables 5 and 4. These parameters should be integrated with inter-group interactions documented in prior research. The outcomes of the saturation pressure calculations for pure NMP, alongside with experimental data sourced from the National Institute of Standards and Technology (NIST) database, are exhibited in Figure 4. The comparison between the predicted and experimental data underscores a high level of agreement, with the SAFT-γ Mie model achieving an average absolute relative

deviation (AARD) of 2.26% for this compound, highlighting the model's efficacy in replicating experimental findings.

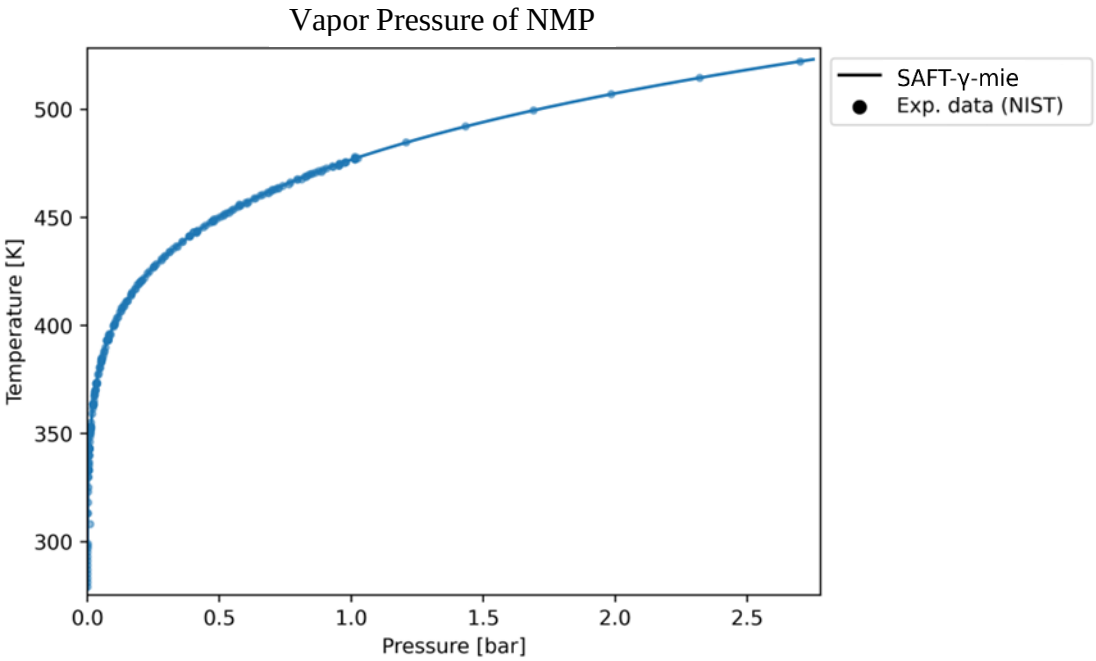

Figure 5. Saturation Pressure of NMP. The blue line illustrates the properties calculated using SAFT-γ Mie model, while the points repres ent experimental data taken from NIST database.

To evaluate the SAFT-γ Mie model's predictive power, the saturation pressure curve for N-Cyclohexyl-2-pyrrolidinone was computed and assessed. The outcomes of this analysis, alongside with experimental data from the NIST database, are depicted in Figure 5. This comparison shows an AARD of 9.59% for this compound. Despite this deviation, the relatively low AARD underscores the model's ability to make highly accurate predictions without necessitating additional fitting of unlike group parameters. It is important to note, however, that the experimental data presented in the figure exhibits some uncertainty in the reported physical property. The saturation curve appears to follow different trajectories for each set of experimental data, suggesting a high variance among the data points. This observation highlights potential inconsistencies within the experimental measurements, which may impact the interpretation of the model's predictive accuracy. Despite these uncertainties, the model's robustness in extrapolating to compounds not directly included in the parameter estimation phase is confirmed.

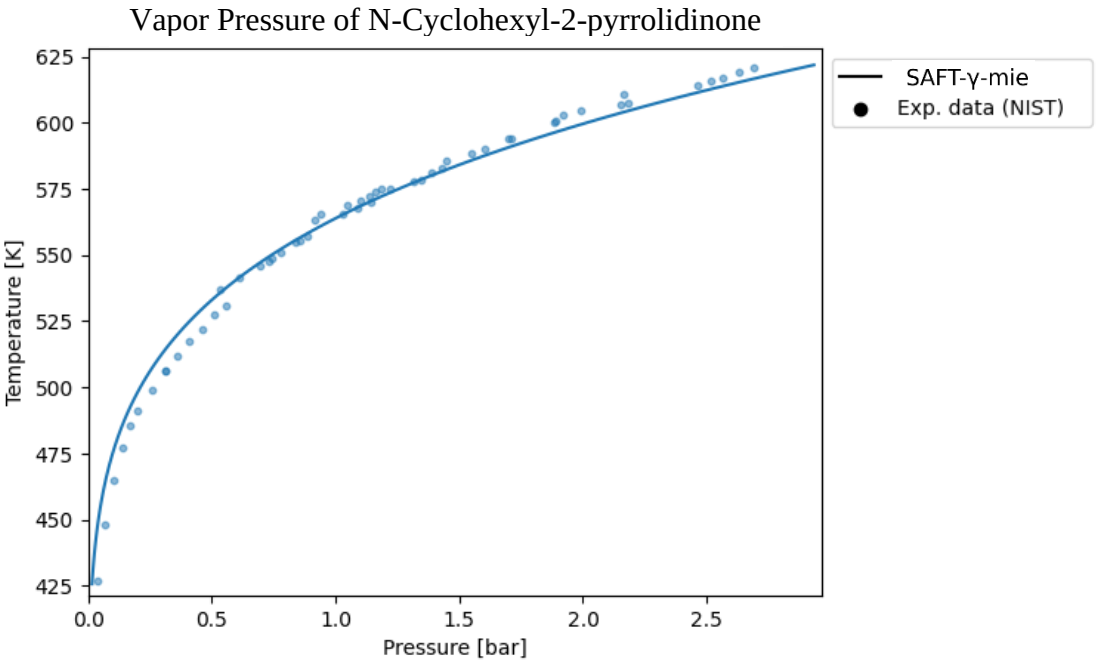

Figure 6: Saturation Pressure of N-Cyclohexyl-2-pyrrolidinone. The blue line illustrates the properties calculated using SAFT-γ Mie model, while the blue points represent experimental data taken from NIST database.

Tables 6, 7, and 8 present a comparison of the VLE correlation outcomes for 28 distinct binary mixtures involving NMP against experimental data, employing the SAFT-γ Mie methodology. The interaction parameters deduced in this section are shown to accurately reflect the fluid phase behavior of the mixtures incorporated into the regression analysis. However, one exception is noted in the VLE behavior of the NMP + dodecane mixture. For this specific combination, the SAFT-γ Mie model forecasts a vapor-liquid-liquid equilibrium (VLLE) region, a prediction that diverges from the observed experimental data as presented in Figure 7. This anomalous behavior can be elucidated by considering the inherent limitations of the EOS when applied to large molecules. As discussed by PAPAIOANNOU (2012), while the SAFT-γ Mie model provides a robust framework for predicting the behavior of complex mixtures, its accuracy might decrease for larger molecules, such as dodecane, particularly in predicting vapor pressures and handling compounds with significant polar interactions. The discrepancy in the prediction of the NMP + dodecane mixture's behavior underscores the importance of acknowledging and addressing the model's constraints, particularly in the context of mixtures involving large hydrocarbon molecules. This insight points to the necessity for further refinement of the model or the adoption of

complementary approaches when dealing with compounds exhibiting a wide range of molecular sizes to enhance the accuracy of predictions across all types of mixtures.

Table 6: Average absolute relative deviation (AARD) for bubble pressure (ΔP^{Bubble}) calculations, alongside temperature range ($\frac{T_{min}}{K}$ to $\frac{T_{max}}{K}$), pressure range ($\frac{P_{min}}{kPa}$ to $\frac{P_{max}}{kPa}$), and the number of experimental data points (N), for vapor-liquid equilibrium (VLE) studies of N-Methyl-2-Pyrrolidone (NMP) with various solvents. Calculations were performed using the SAFT-γ Mie Equation of State.

2nd Component	**AARD (%) ΔP^{Bubble}**	$\frac{T_{min}}{K}$	$\frac{T_{max}}{K}$	$\frac{P_{min}}{kPa}$	$\frac{P_{max}}{kPa}$	**N**
Ethanol	4.90	333.15	333.15	2.16	4.53	20
Methanol	1.04	333.15	393.19	2.69	640.52	40
Propanol	3.88	333.15	333.15	0.78	19.78	24
Ethane	9.59	263.20	393.20	101.30	7127.0	53
Propane	15.68	263.11	328.10	23.47	815.00	25
Hexane	1.58	342.83	363.58	9.22	192.72	56
Heptane	0.60	292.88	371.62	1.50	99.86	49
Cyclohexane	1.03	333.18	354.20	18.04	97.43	37
Acetone	1.89	333.15	333.15	28.60	95.70	9
Benzene	1.69	292.65	364.43	1.65	100.26	42
Toluene	2.10	313.15	436.81	0.79	101.32	46
Isopentane	6.87	363.55	363.55	181.70	546.40	13
Water	1.30	380.24	351.01	1.58	129.91	82
2-propanol	1.39	353.15	454.25	11.10	111.70	55
2-butanol	1.75	373.15	468.35	9.10	104.10	20
Isobutanol	4.91	382.55	465.65	95.30	95.30	8
Isoamyl	1.66	403.85	465.65	95.30	95.30	8

Table 7: Average absolute relative deviation (AARD) for bubble pressure (ΔP^{Bubble}) and vapor phase composition (Δy_1) calculation, alongside temperature range ($\frac{T_{min}}{K}$ to $\frac{T_{max}}{K}$), pressure range ($\frac{P_{min}}{kPa}$ to $\frac{P_{max}}{kPa}$), and the number of experimental data points (N), for vapor-liquid equilibrium (VLE) studies of N-Methyl-2-Pyrrolidone (NMP) with various solvents. Calculations were performed using the SAFT-γ Mie Equation of State.

2nd Component	AARD (%) ΔP^{Bubble}	AARD (%) Δy_1	$\frac{T_{min}}{K}$	$\frac{T_{max}}{K}$	$\frac{P_{min}}{kPa}$	$\frac{P_{max}}{kPa}$	N
Methylcyclohexane	5.65	21.82	373.25	354.15	22.73	94.9	21
Dodecane	1.54	28.87	393.25	393.25	6.37	10.69	12
Propanol	2.11	33.33	354.13	354.13	2.46	47.55	9
Hexanol	0.96	22.94	393.32	351.73	1.22	23.59	19
Cyclohexane	1.02	2.86	354.15	281.10	0.7522	97.33	51
Benzene	1.21	7.78	354.13	281.10	0.2628	96.65	75
Toluene	1.39	14.00	383.32	343.15	1.4	94.47	51
Heptane	1.19	19.25	365.00	340.00	5.43	82.65	27
Hexane	2.50	22.94	343.15	333.25	63.79	102.77	19
2-Methyl-2-propanol	2.75	14.91	373.15	353.15	1.665	50.989	45
Water	4.22	12.22	380.15	273.10	0.02	91.03	33
Dibutyl ether	4.57	28.87	373.15	373.15	11.3	27.9	12
Dipropyl ether	4.29	25.82	373.15	353.15	27	116.6	15

Table 8: Average absolute relative deviation (AARD) for bubble temperature (ΔT^{Bubble}) and vapor phase composition (Δy_1) calculations, alongside temperature range ($\frac{T_{min}}{K}$ to $\frac{T_{max}}{K}$), pressure range ($\frac{P_{min}}{kPa}$ to $\frac{P_{max}}{kPa}$), and the number of experimental data points (N), for vapor-liquid equilibrium (VLE) studies of N-Methyl-2-Pyrrolidone (NMP) with various solvents. Calculations were performed using the SAFT-γ Mie Equation of State.

2nd Component	AARD (%) ΔT^{Bubble}	AARD (%) Δy_1	$\frac{T_{min}}{K}$	$\frac{T_{max}}{K}$	$\frac{P_{min}}{kPa}$	$\frac{P_{max}}{kPa}$	N
Butanol	0.066	3.09	446.08	339.22	10	50	43
2-Methoxyethanol	0.114	2.72	470.71	398.19	95.3	95.3	13
2-Ethoxyethanol	0.374	8.15	470.95	409.46	95.3	95.3	13
Cyclohexane	0.222	10.26	418.20	355.25	101.33	101.33	11
Benzene	0.346	6.44	448.45	355.30	101	101.33	30
Toluene	0.225	21.37	443.65	387.15	101.3	101.32	15
Heptane	0.092	13.56	463.80	371.55	101.33	101.33	9
Hexane	0.378	10.42	384.40	342.00	101.3	101.33	27
Methylcyclopentane	1.047	28.65	373.87	346.51	101.3	101.3	18
Acetic acid	0.306	6.86	475.02	364.10	26.67	101.33	53
Water	0.217	7.64	380.15	343.15	9.01	91.03	24

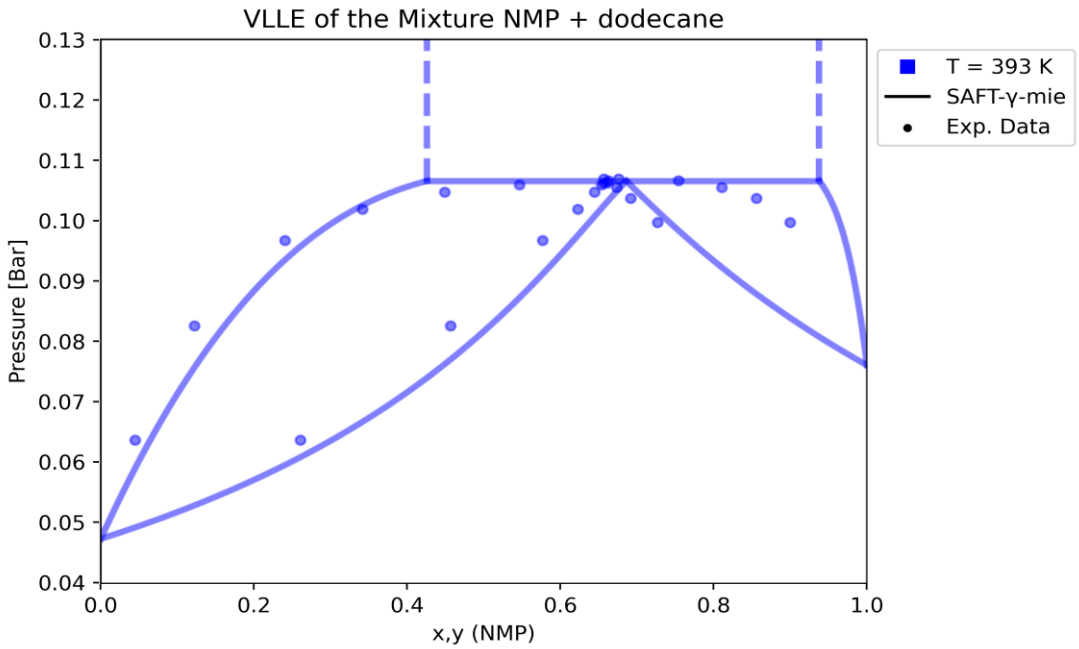

Figure 7: Pressure-composition relationship of the NMP + dodecane system at 393 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line for the VLE and dashed line for the liquid-liquid equilibrium.

To mitigate this kind of incoherent results, one could use the TPD algorithm to penalize the objective function whenever the EOS predicts different phases compared to experimental data. In the case discussed above, the TPD could be calculated with the mother phase being either liquid measured in experimental data or vapor predicted by the bubble point calculation. As for the incipient phase, it could be used an almost pure liquid composed by the most abundant component in the vapor phase or the least abundant component of the first liquid phase. If the TPD returns a negative value, it means that this second liquid reduces the Gibbs energy of the system, and therefore, is stable. The Equation 4-1 provides an example of a modification in the objective function using the TPD results as a penalty:

$$\underset{\Omega}{\text{Min}}\, F'_{obj1} = F_{obj1}(\Omega) - \omega_5 \sum_{i}^{N_{all}} \min\left(TPD(x_{exp_i}, z_{\text{guess}_i}, \Omega); 0\right) \qquad 4\text{-}1$$

Where z_{guess} is incipient liquid composition guess and N_{all} is the total number of the available VLE data. For qualitative comparison purposes, this modification in the objective function was applied to all the mixtures containing the NMP and hydrocarbons. The parameters after the modification in the objective function are presented in tables 9 and 10.

Table 9: Group parameters for the functional group cNcCO within the SAFT-γ Mie framework after the modification in the objective function.

group	S_k	σ_{kk}	ε_{kk}	λ_{kk}^r
cNcCO	0.993818	3.63074	794.263	13.4401

Table 10: Group-Specific Parameters within the SAFT-γ Mie Framework after the modification in the objective function.

Group k	**Group l**	$\frac{\varepsilon_{kl}}{k_b T}$	λ_{kl}^r
cNcCO	CH3	411.214	15.0082
cNcCO	CH2	554.844	18.5813

Figure 8 shows the VLE of the NMP + dodecane system after applying the modification in the objective function using the TPD. It can be seen that the model is not predicting a region with VLLE, which shows more agreement with the experimental data. However, the error in the calculated pressures were higher.

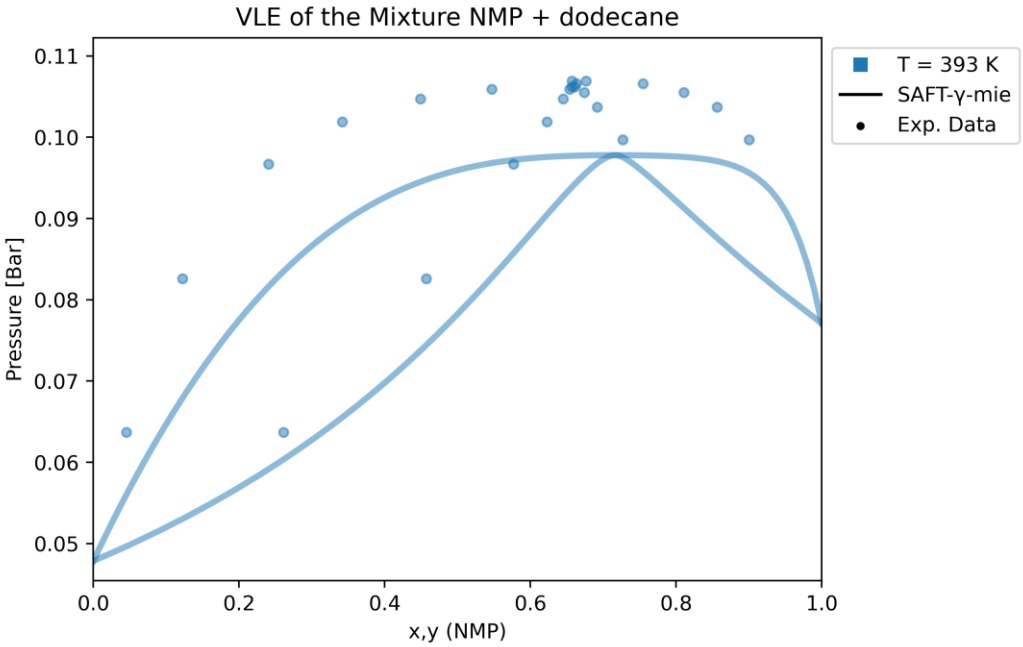

Figure 8: Pressure-composition relationship of the NMP + dodecane system at 393 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line for the VLE. For this system, it was used the estimated parameters after the modification in the objective function

In order to further evaluate the impact of the modification in the objective function as well as the extrapolability of the model, liquid-liquid equilibria (LLE) calculation was performed and compared to experimental data from NIST. Figures 9 and 10 show the LLE of the NMP + dodecane system before and after applying the modification in the objective function using the TPD, respectively. In both cases the SAFT-γ Mie model predicts LLE even though these experimental points were not included in the estimation. Moreover, it can be seen that the modified parameters reduced the error between the experimental and the calculated temperatures.

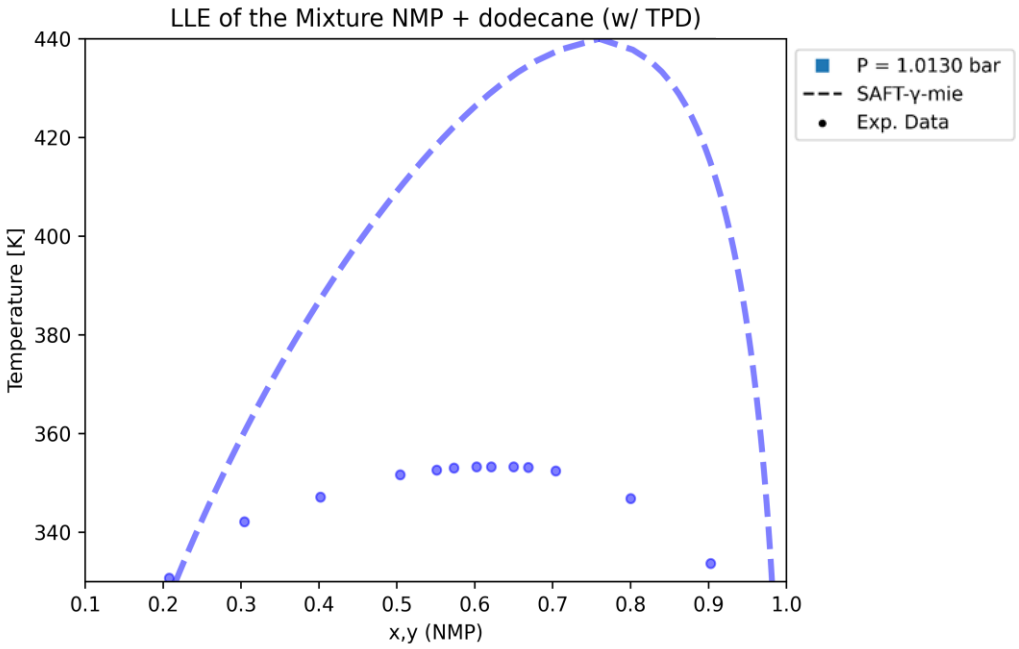

Figure 9: Temperature-composition relationship of the NMP + dodecane system at 1.013 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a dashed line for the LLE.

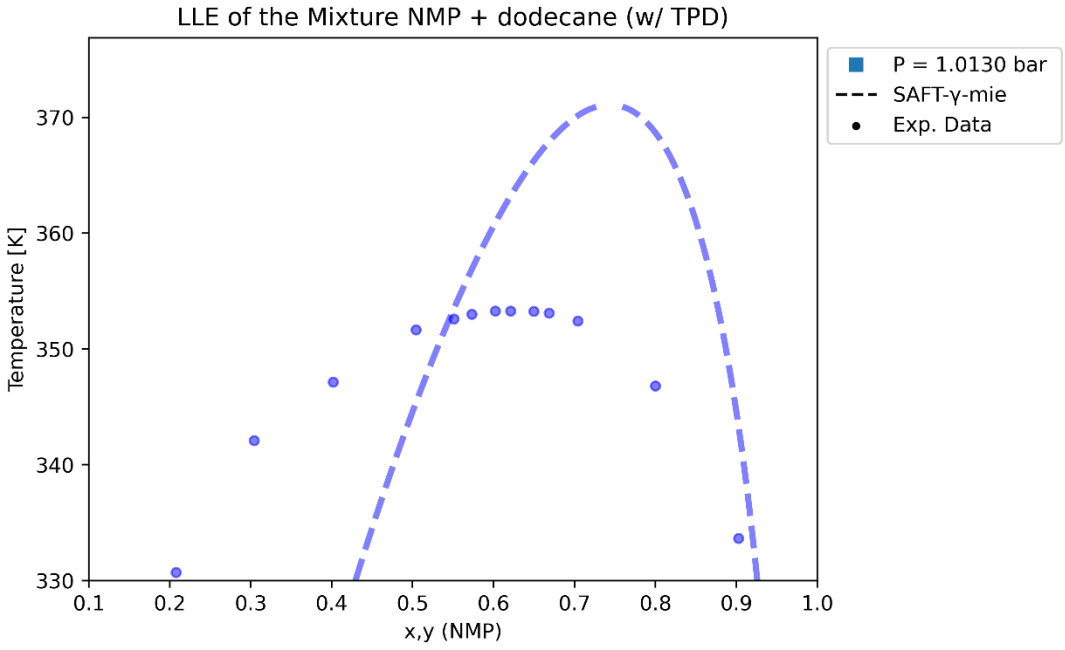

Figure 10: Temperature-composition relationship of the NMP + dodecane system at 1.013 bar. It is provided a direct comparison between experimental data, sourced from

NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a dashed line for the LLE. For this system, it was used the estimated parameters after the modification in the objective function

The evaluated metrics, including the average bubble-point pressure, bubble-point temperature, and vapor composition differences, yield AARDs of 3.175%, 0.31%, and 15.615%, respectively. These values indicate a high degree of accuracy in the model's replication of experimental data for binary mixtures. Nevertheless, it is observed that mixtures involving compounds like dibutyl ether or methylcyclohexane tend to exhibit less accurate results, with these pure components already demonstrating significantly poorer predictions (approximately 20%, in the case of, for example). Consequently, the group interaction parameters alone are insufficient to rectify the intrinsic errors present in previously estimated parameters. This insight suggests a limitation in the model's capacity to universally account for the complexities of all component interactions without specific adjustments or additional parameter estimations for certain mixtures.

4.2. Solid-Liquid Equilibria

The primary aim of this investigation was to forecast the solubility of praziquantel in a variety of solvents using the SAFT-γ Mie Equation of State and to undertake a comparative analysis with the predictions made by the PC-SAFT EoS. This comparison also focused on assessing their capabilities for extrapolation. The availability of comprehensive and reliable experimental data is critical for extending the modeling to pharmaceutical compounds. Given that complex molecules such as praziquantel tend to decompose rather than vaporize, direct data on praziquantel's saturation pressure are unattainable. Consequently, this study concentrated exclusively on binary SLE systems, with an inclusion of one case of ternary.

A literature review was conducted to gather experimental solubility data for praziquantel across 20 different solvents under atmospheric pressure, including variations across temperature ranges. This data collection is fundamental in assessing the efficacy of the SAFT-γ Mie and PC-SAFT EOS models in predicting the solubility of pharmaceutical substances.

In this study, parameters for pure components within the PC-SAFT framework were sourced from existing literature, ensuring a foundation built on previously validated data. Additionally, single temperature-independent binary interaction parameters (k_{ij}) for 18 mixtures were estimated. The fitting of k_{ij} values was conducted exclusively using SLE mixture data.

Just as with VLE predictions, forecasting solubilities with the SAFT-γ Mie approach necessitates the integration of two newly developed group parameters introduced in this study, specifically cNCO and cNcCO. These parameters, along with their associated unlike interactions, are detailed in Table 4 and 5. To ensure comprehensive modeling, these new parameters must be integrated with inter-group interactions that have been identified in previous research. This methodological approach allows for a nuanced representation of solubility behaviors by incorporating both novel and established interaction parameters, thereby enhancing the predictive capability of the SAFT-γ Mie model for praziquantel solubility calculations.

The SLE for praziquantel is determined utilizing the methodology outlined in Section 2.2.2, incorporating experimental data on the enthalpy of fusion and melting temperature of praziquantel. The calculation of the activity coefficient of the solute in the liquid phase is performed using either the SAFT-γ Mie or the PC-SAFT model. In employing the SAFT-γ Mie approach for solubility predictions, a specific equation (referred to as Equation 2-34) is applied, predicated on the assumption that the solid phase is composed of two solutes, with each enantiomer accounting for 50% of the mixture. This contrasts with the PC-SAFT approach, which conceptualizes the solid phase as a single, pure solute, as delineated in Equation 2-33. The assumption of an enantiomeric solid phase more accurately reflects this thermodynamic condition.

The distinction in methodologies allows for the use of pre-existing PC-SAFT parameters, facilitating a direct comparison with the work of SADOWSKI *et al.*(2021). For the SAFT-γ Mie model, the calculation includes terms accounting for the difference in heat capacity between the solid and liquid phases, as specified in Equation 2-32. The melting temperature and enthalpy of fusion for praziquantel were set at 410.6 K and 30.9 kJ mol^{-1}, respectively, in accordance with the findings of de MOARES et al. (yet to be published). These thermodynamic properties in the PC-SAFT model were aligned

with those reported by Sadowski, ensuring consistency in the comparative analysis of the solubility predictions provided by the two models.

The accuracy of the models was assessed through the reported errors for solubility computations across a range of organic solvents. Table 11 showcases a comprehensive comparison of SLE correlation results for 20 different solvents against experimental data, employing both the SAFT-γ Mie and PC-SAFT methodologies. The analysis reveals that both models, with the newly proposed parameter set in this study, deliver satisfactory outcomes. Notably, the SAFT-γ Mie model achieves better results, exhibiting an AARD of 5.72% comparing all the mixtures, while the PC-SAFT model reports an AARD of 9.90%. Moreover, SAFT-γ Mie model achieved smaller deviation in 15 of the 20 mixtures. The most significant deviations observed were a 27.37% AARD with the SAFT-γ Mie model for ethanediol and a 28.15% AARD with the PC-SAFT model for octanol. These results indicate a generally high level of accuracy in the solubility predictions for both models.

Table 11: Percentage average absolute relative deviation (%AARD) for the solubilities of the Praziquantel in various solvents at atmospheric pressure obtained with the SAFT-γ Mie group contribution and PC-SAFT approach with respect to the experimental data points.

Solvent	AARD (%) SAFT-y-Mie	AARD (%) PC-SAFT
Hexane	1.51	12.81
Toluene	6.28	9.69
NMP	1.19	4.19
Methyl isobutyl ketone	4.26	7.90
Acetone	3.27	8.44
Butylacetate	1.22	11.85
Ethylacetate	5.13	11.00
Isopropylacetate	19.34	10.24
Methylacetate	1.68	10.35
Propylacetate	2.83	9.80
2-Ethoxyethanol	5.06	2.17
2-Propoxyethanol	3.51	8.10
2-Butoxyethanol	6.56	6.22
Butanol	2.00	1.46
Ethanol	5.68	24.89
Octanol	5.28	28.15
Propanol	2.63	3.09
Methanol	6.56	12.05
2-Butanol	1.35	3.10
2-Propanol	7.44	16.37
Ethanediol	27.38	6.06

Table 12: Binary PC-SAFT Interaction Parameters (k_{ij}) between Praziquantel and the Solvents.

Solvent	$\boldsymbol{k_{ij}}$	**Source**
Hexane	0.061416	this work
Toluene	0.004528	this work
NMP	-0.0368	this work
Methyl isobutyl ketone	0.013682	this work
Acetone	0.034236	this work
Butylacetate	0.019963	this work
Ethylacetate	0.030746	this work
Isopropylacetate	0.027724	this work
Methylacetate	0.041318	this work
Propylacetate	0.022393	this work
2-Ethoxyethanol	0.023591	this work
2-Propoxyethanol	0.023717	this work
2-Butoxyethanol	-0.00455	this work
Butanol	-0.0102	this work
Octanol	-0.00455	this work
Ethanol	0.02	SADOWSKI *et al.*(2021)
Propanol	0.007189	this work
Methanol	-0.00653	this work
2-Butanol	-0.00829	this work
2-Propanol	0.03	SADOWSKI *et al.*(2021)
Ethanediol	0.02098	this work

In this study, the predictive performance of the SAFT-γ Mie and PC-SAFT models for PZQ's solubility in a mixed solvent system of NMP and 2-propanol was evaluated, with no direct experimental data for this specific mixture included in the parameter estimation. This approach was aimed at testing the models' extrapolation capabilities. The Figures 6 and 7 depict the solubility predictions for PZQ in varying proportions of NMP and 2-propanol using both SAFT-γ Mie and PC-SAFT models.

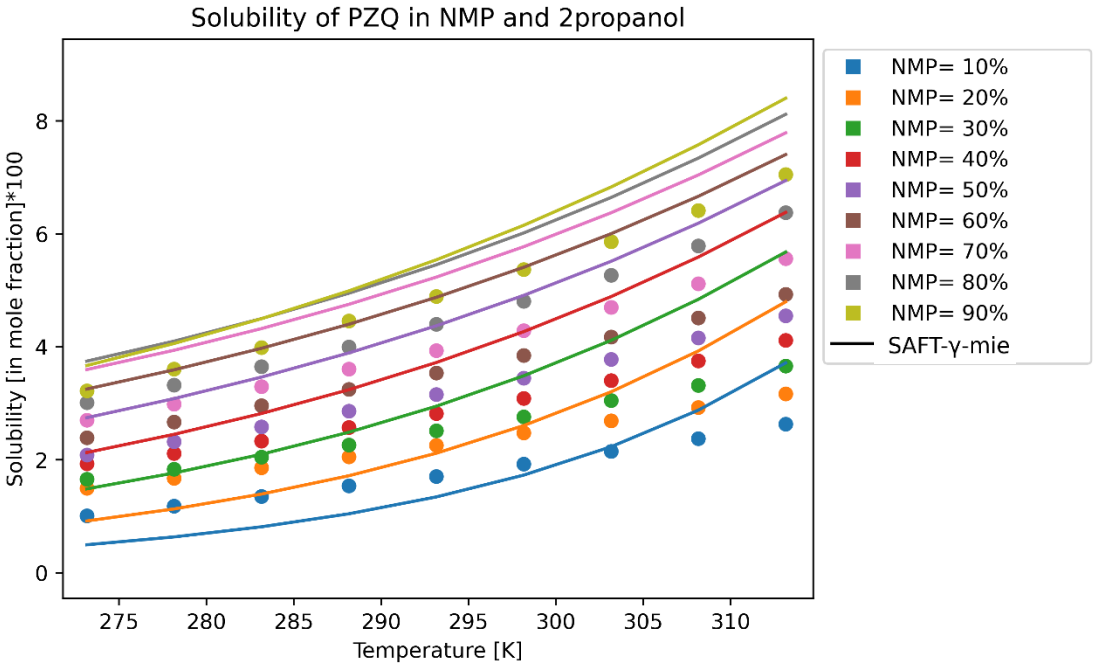

Figure 11: Comparison of Experimental and Predicted Solubilities of Praziquantel in Mixtures of NMP and 2-Propanol (2propanol). This figure illustrates the solubility predictions of praziquantel in mixed solvent systems of NMP and 2-propanol at varying proportions using the SAFT-γ Mie modeling. Different colors are used to denote various solvent mixture ratios. The continuous lines indicate the model's predicted solubilities, whereas the dots correspond to the actual experimental data points. Experimental data sourced from LI *et al.* (2020).

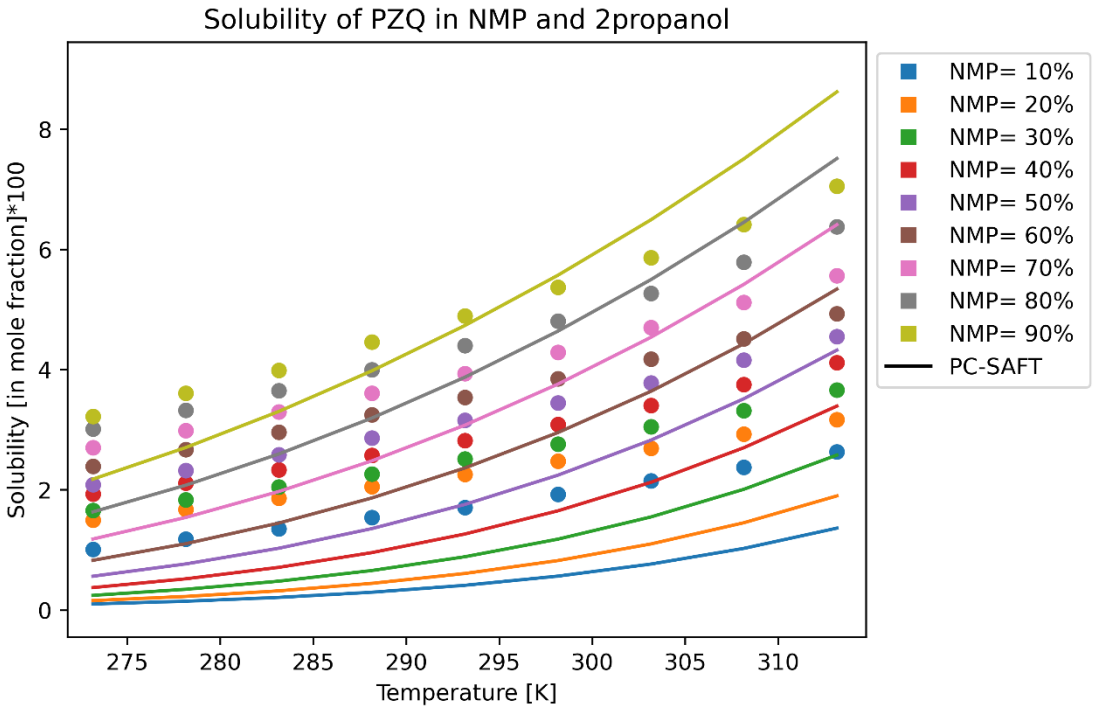

Figure 12: Comparison of Experimental and Predicted Solubilities of Praziquantel in Mixtures of NMP and 2-Propanol. This figure illustrates the solubility predictions of praziquantel in mixed solvent systems of NMP and 2-propanol at varying proportions using the PC-SAFT modeling. Different colors are used to denote various solvent mixture ratios. The continuous lines indicate the model's predicted solubilities, whereas the dots correspond to the actual experimental data points. Experimental data sourced from LI *et al.* (2020).

The SAFT-γ Mie model demonstrated superior accuracy, achieving an AARD of 3.59%, compared to the PC-SAFT model's AARD of 5.7%. This improved accuracy from the SAFT-γ Mie model was made possible after estimating the unlike parameter $\varepsilon_{CH_2OH,cNCO}$, which initially, calculated with combining rules, resulted in an AARD of 11.29%, significantly higher and indicative of potential inaccuracies in the model's predictions. The initial reluctance to estimate unlike parameters was driven by concerns over the risk of overfitting. The strategy was to refrain from adjusting unlike parameters when the error margin for the majority of compounds, which interact through these parameters, was below approximately 10%. Nonetheless, this instance demonstrated that, with ample experimental data, the fitting of unlike SAFT-γ Mie parameters could lead to more accurate outcomes and enhance the model's ability to extrapolate. The combining rules, in this context, failed to accurately capture the interaction energies

between groups, emphasizing the importance of parameter estimation in achieving higher predictive accuracy in modeling complex systems.

This observation aligns with results obtained from the VLE modeling of the NMP + water system using the SAFT-γ Mie model. Prior to the adjustment of the $\varepsilon_{H_2O,cNcCO}$ parameter, the model inaccurately predicted a three-phase region that was not supported by experimental evidence. This instance reinforces the notion that, with sufficient experimental data, the fitting of unlike parameters can lead to more accurate representations of thermodynamic properties. Furthermore, it emphasizes the potential of such adjusted models to extrapolate effectively to regions where experimental data may not be readily available, thereby enhancing the predictive power and reliability of the SAFT-γ Mie model in capturing complex system behaviors.

In the Supplementary Information accompanying this dissertation, comprehensive datasets detailing VLE and SLE of the components used in this study are presented. These datasets encompass extensive experimental and modeled equilibrium data crucial for understanding the thermodynamic behaviors observed in the study. The inclusion of this supplementary data ensures that interested readers have access to the full scope of research findings, facilitating further analysis and application of the results presented herein.

5. Conclusion

This dissertation has made significant strides in enhancing our understanding of VLE and SLE of compounds, particularly through the application of the SAFT-γ Mie and PC-SAFT models to the PZQ and NMP. Through experimental work and theoretical modeling, it has been demonstrated that the introduction of novel parameters in these models can substantially improve the predictive accuracy for the behavior of Praziquantel in complex mixtures. Notably, the SAFT-γ Mie model has been highlighted for its enhanced predictive and extrapolability capabilities across a diverse range of solvents, indicating its potential for wider application in the field of pharmaceutical research.

However, this research acknowledges several limitations. Among these, the low variability and not inclusion of uncertainties in the experimental data, present significant challenges. This uncertainty is compounded by the lack of replica and variance data in a large portion of the experimental results available, which would have provided a more

statistically robust basis for the analysis. The absence of this replica and variance data limited our ability to fully utilize statistical methods to enhance the reliability of our predictions. Furthermore, the necessity to use previously estimated parameters, which themselves were not derived using these variances, introduces additional layers of complexity and potential inaccuracy in the modeling efforts.

5.1. Future Work

To directly test the model's predictions, future work could involve conducting experiments to search for liquid-liquid equilibria (LLE) at the predicted location (Figure 42). Finding LLE in these areas would provide direct evidence of the model's high predictive and extrapolative power, further confirming its applicability in pharmaceutical research.

Additionally, this dissertation identifies the potential application of these thermodynamic models in the challenging area of enantiomer purification processes. The insights gained from the SLE studies lay the groundwork for future research aimed at modeling the purification of enantiomers, a critical step in the production of enantiopure pharmaceuticals. Addressing the limitations related to the experimental data and enhancing the statistical robustness of the modeling approaches could significantly impact the efficiency and reliability of enantiopure compound production.

Future research directions should focus on improving the accuracy and reliability of experimental data reporting, including the incorporation of variance data where available. Efforts should also be made to refine the estimation of unlike parameters, especially including the TPD calculations in the objective function to prevent the prediction of incoherent phases. Moreover, it could also be explored the application of the studied thermodynamic models to enantiomer purification, making possible the practical utility of these models in pharmaceutical sciences. Through continued research and methodological improvements, there is a promising pathway toward advancing the field of pharmaceutical sciences, particularly in the development and production of enantiopure compounds.

6. References

ALKHATIB, I. I. I., PEREIRA, L. M. C., TORNE, J., *et al.* "Polar soft-SAFT: Theory and comparison with molecular simulations and experimental data of pure polar fluids", **Physical Chemistry Chemical Physics**, v. 22, n. 23, p. 13171–13191, 2020. DOI: 10.1039/d0cp00846j. .

ANICETO, J. P. S., CARDOSO, S. P., FARIA, T. L., *et al.* "Modeling ion exchange equilibrium: Analysis of exchanger phase non-ideality", **Desalination**, v. 290, p. 43–53, 2012. DOI: 10.1016/j.desal.2012.01.001. .

BAKER, L. E., PIERCE, A. C., LUKS, K. D. "Gibbs Energy Analysis of Phase Equilibria.", **Society of Petroleum Engineers of AIME, (Paper) SPE**, p. 471–488, 1982. .

BARKER, J. A., HENDERSON, D. "Perturbation theory and equation of state for fluids. II. A successful theory of liquids", **The Journal of Chemical Physics**, v. 47, n. 11, p. 4714–4721, 1967. DOI: 10.1063/1.1701689. .

BAYDIN, A. G., PEARLMUTTER, B. A., AARDUL, A. A., *et al.* "Automatic Differentiation in Machine Learning: a Survey", **Journal of Machine Learning Research**, v. 18, p. 1–43, 2018. .

BOCHMANN, E. S., NEUMANN, D., GRYCZKE, A., *et al.* "Micro-scale solubility assessments and prediction models for active pharmaceutical ingredients in polymeric matrices", **European Journal of Pharmaceutics and Biopharmaceutics**, v. 141, p. 111–120, 2019. DOI: 10.1016/j.ejpb.2019.05.012. .

BOUBLÍK, T., BOUBLIK, T. "Hard-sphere equation of state [25]", **The Journal of Chemical Physics**, v. 53, n. 1, p. 471–472, 1 jul. 1970. DOI: 10.1063/1.1673824. Disponível em: https://doi.org/10.1063/1.1673824.

BOUILLOT, B., TEYCHENÉ, S., BISCANS, B. "An evaluation of thermodynamic models for the prediction of drug and drug-like molecule solubility in organic solvents", **Fluid Phase Equilibria**, v. 309, n. 1, p. 36–52, 2011. DOI: 10.1016/j.fluid.2011.06.032. Disponível em: http://dx.doi.org/10.1016/j.fluid.2011.06.032.

BAARDBURY, J., FROSTIG, R., HAWKINS, P., *et al.* **JAX: composable transformations of Python+NumPy programs**. . [S.l: s.n.]. Disponível em: http://github.com/google/jax. Acesso em: 9 mar. 2024. , 2018

BURGER, J., PAPAIOANNOU, V., GOPINATH, S., *et al.* "A Hierarchical Method to Integrated Solvent and Process Design of Physical CO2 Absorption Using the SAFT-γ Mie Approach", **AIChE Journal**, v. 61, n. 10, 2015. DOI: https://doi.org/10.1002/aic.14838. .

CASTELLS, C. B., CARR, P. W. "A study of the thermodynamics and influence of temperature on chiral high-performance liquid chromatographic separations using cellulose tris(3,5-dimethylphenylcarbamate) coated zirconia stationary phases", **Chromatographia**, v. 52, n. 9–10, p. 535–542, 2000. DOI: 10.1007/BF02789747. .

CHAPMAN, W. G., GUBBINS, K. E., JACKSON, G., *et al.* "New reference equation of state for associating liquids", **Industrial and Engineering Chemistry Research**, v. 29, n. 8, p. 1709–1721, 1990. DOI: 10.1021/ie00104a021. .

DOHRN, S., LUEBBERT, C., LEHMKEMPER, K., *et al.* "Solvent mixtures in

pharmaceutical development: Maximizing the API solubility and avoiding phase separation", **Fluid Phase Equilibria**, v. 548, p. 113200, 2021. DOI: 10.1016/j.fluid.2021.113200. Disponível em: https://doi.org/10.1016/j.fluid.2021.113200.

DUFAL, S., PAPAIOANNOU, V., SADEQZADEH, M., *et al.* "Prediction of thermodynamic properties and phase behavior of fluids and mixtures with the SAFT-γ mie group-contribution equation of state", **Journal of Chemical and Engineering Data**, v. 59, n. 10, p. 3272–3288, 2014. DOI: 10.1021/je500248h. .

FEBRA, S. A., BERNET, T., MACK, C., *et al.* "Extending the SAFT-γ Mie approach to model benzoic acid, diphenylamine, and mefenamic acid: Solubility prediction and experimental measurement", **Fluid Phase Equilibria**, v. 540, 2021. DOI: 10.1016/j.fluid.2021.113002. .

FERNANDES, D. T. **Equilíbrio De Fases De Misturas Complexas Reconstruídas Molecularmente Utilizando A Equação De Estado SAFT-γ Mie**. Msc Thesis. UFRJ, 2023.

FIKE, J., ALONSO, J. "The Development of Hyper-Dual Numbers for Exact Second-Derivative Calculations", n. January, 2011. DOI: 10.2514/6.2011-886. .

GLYNN, P. D., REARDON, E. J., PLUMMER, L. N., *et al.* "Reaction paths and equilibrium end-points in solid-solution aqueous-solution systems", **Geochimica et Cosmochimica Acta**, v. 54, n. 2, p. 267–282, 1990. DOI: https://doi.org/10.1016/0016-7037(90)90317-E. Disponível em: https://www.sciencedirect.com/science/article/pii/001670379090317E%0A.

GMEHLING, J., WITTIG, R., LOHMANN, J., *et al.* "A modified UNIFAC (Dortmund) model. 4. Revision and extension", **Industrial and Engineering Chemistry Research**, v. 41, n. 6, p. 1678–1688, 2002. DOI: 10.1021/ie0108043. .

GROSS, J., SADOWSKI, G. "Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules Joachim", **Industrial and Engineering Chemistry Research**, v. 40, p. 1244–1260, 2001. DOI: 10.1021/acs.iecr.9b01515. .

HASLAM, A. J., GONZÁLEZ-PÉREZ, A., DI LECCE, S., *et al.* "Expanding the Applications of the SAFT-γMie Group-Contribution Equation of State: Prediction of Thermodynamic Properties and Phase Behavior of Mixtures", **Journal of Chemical and Engineering Data**, v. 65, n. 12, p. 5862–5890, 2020. DOI: 10.1021/acs.jced.0c00746. .

HIRSCHFELDER, J. O., CURTISS, C. F., BIRD, R. B., *et al.* **Molecular Theory of Gases and Liquids**. New York, Wiley, 1964. v. 8. Disponível em: https://books.google.com.br/books?id=BJkuAAAAIAAJ. (Molecular Theory of Gases and Liquids).

HUANG, S. H., RADOSZ, M. "Equation of state for small, large, polydisperse, and associating molecules: extension to fluid mixtures. [Erratum to document cited in CA115(8):79950j]", **Industrial & Engineering Chemistry Research**, v. 32, n. 4, p. 762–762, 1993. DOI: 10.1021/ie00016a029. .

HURON, M.-J. J., VIDAL, J. "New Mixing Rules in Simple Equations of State For Representing Vapour-Liquid Equilibria of Strongly Non-Ideal Mixtures", **Fluid Phase Equilibria**, v. 3, n. 4, p. 255–271, 1979. DOI: 10.1016/0378-3812(79)80001-1. .

HUTACHAROEN, P., DUFAL, S., PAPAIOANNOU, V., *et al.* "Predicting the Solvation of Organic Compounds in Aqueous Environments: From Alkanes and Alcohols to Pharmaceuticals", **Industrial and Engineering Chemistry Research**, v. 56, n. 38, p. 10856–10876, 2017. DOI: 10.1021/acs.iecr.7b00899. .

JONES, J. E. "On the Determination o f Molecular Fields. II. From the Equation of State of a Gas", v. 106, n. 738, p. 463–477, 1924. DOI: 10.1098/rspa.1924.0082. Disponível em: https://ui.adsabs.harvard.edu/link_gateway/1924RSPSA.106..463J/doi:10.1098/rspa.1924.0082.

KREGLEWSKI, A., STEPHEN S., C. "Applications of the augmented Van der Waals theory of fluids", **Journal de Chimie Physique**, v. 77, n. 1, p. 441–444, 1980. DOI: 10.1051/jcp/1980770441. .

LAFITTE, T., APOSTOLAKOU, A., AVENDAÑO, C., *et al.* "Accurate statistical associating fluid theory for chain molecules formed from Mie segments", **Journal of Chemical Physics**, v. 139, n. 15, 2013. DOI: 10.1063/1.4819786. .

LI, R., CHEN, X., HE, G., *et al.* "The dissolution behaviour and thermodynamic properties calculation of praziquantel in pure and mixed organic solvents", **Journal of Chemical Thermodynamics**, v. 144, p. 106062, 2020. DOI: 10.1016/j.jct.2020.106062. Disponível em: https://doi.org/10.1016/j.jct.2020.106062.

LIU, H., HU, Y. "Molecular thermodynamic theory for polymer systems III. Equation of state for chain-fluid mixtures", **Fluid Phase Equilibria**, v. 138, n. 1–2, p. 69–85, 1997. DOI: 10.1016/s0378-3812(97)00103-9. .

LIU, Y., ZHANG, X., WANG, M., *et al.* "Uncovering the effect of solvents on solid-liquid phase equilibrium of praziquantel", **Journal of Molecular Liquids**, v. 297, p. 111917, 2020. DOI: 10.1016/j.molliq.2019.111917. Disponível em: https://doi.org/10.1016/j.molliq.2019.111917.

LORENTZ, H. A. "Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase", 1877. .

MARGOSSIAN, C. C. "A review of automatic differentiation and its efficient implementation", **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, v. 9, n. 4, p. 1–19, 2019. DOI: 10.1002/widm.1305. .

MARGULES, M. "Über die Zusammensetzung der gesättigten Dämpfe von Mischungen", **Sitzungsberichte der Kaiserliche Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche Classe**, v. 104, p. 1243–1278, 1895. .

MICHELSEN, M. L. "A modified Huron-Vidal mixing rule for cubic equations of state", **Fluid Phase Equilibria**, v. 60, n. 1–2, p. 213–219, 1990. DOI: 10.1016/0378-3812(90)85053-D. .

MICHELSEN, M. L. "The isothermal flash problem. Part I. Stability", **Fluid Phase Equilibria**, v. 9, n. 1, p. 1–19, 1982. DOI: 10.1016/0378-3812(82)85001-2. .

MICHELSEN, M. L., HENDRIKS, E. M. "Physical properties from association models", **Fluid Phase Equilibria**, v. 180, n. 1–2, p. 165–174, 2001. DOI: 10.1016/S0378-3812(01)00344-2. .

MOLLERUP, J. "A note on the derivation of mixing rules from excess Gibbs energy

models", **Fluid Phase Equilibria**, 1986. .

NEAU, S. H., BHANDARKAR, S. V., HELLMUTH, E. W. "Differential molar heat capacities to test ideal solubility estimations", **Pharmaceutical Research**, v. 14, n. 5, p. 601–605, 1997. DOI: 10.1023/A:1012148910975. .

PANKOW, J. F., ASHER, W. E. "SIMPOL.1: A simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds", **Atmospheric Chemistry and Physics**, v. 8, n. 10, p. 2773–2796, 2008. DOI: 10.5194/acp-8-2773-2008. .

PAPAIOANNOU, V. **A molecular-based group contribution equation of state for the description of fluid phase behaviour and thermodynamic derivative properties of mixtures (SAFT- γ Mie)**. PhD Thesis. 241 f. Imperial College London, 2012.

PAPAIOANNOU, V., CALADO, F., LAFITTE, T., *et al.* "Application of the SAFT-γ Mie group contribution equation of state to fluids of relevance to the oil and gas industry", **Fluid Phase Equilibria**, v. 416, p. 104–119, 2016. DOI: 10.1016/j.fluid.2015.12.041. .

PAPAIOANNOU, V., LAFITTE, T., AVENDAÑO, C., *et al.* "Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments", **Journal of Chemical Physics**, v. 140, n. 5, 2014. DOI: 10.1063/1.4851455. .

PENG, D. Y., ROBINSON, D. B. "A New Two-Constant Equation of State", **Industrial and Engineering Chemistry Fundamentals**, v. 15, n. 1, p. 59–64, 1976. DOI: 10.1021/i160057a011. .

PRAUSNITZ, J. M., LICHTENTHALEE, R. ., GOMES DE AZEVEDO, E. **Molecular Themrodynamics of Fluid-Phase Equilibria**. New Jersey, Prentice-Hall, Inc., 1999.

RASO, G., N'GORAN, E. K., TOTY, A., *et al.* "Efficacy and side effects of praziquantel against Schistosoma mansoni in a community of western Côte d'Ivoire", **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 98, n. 1, p. 18–27, 2004. DOI: 10.1016/S0035-9203(03)00003-8. .

REDLICH, O., KWONG, J. N. S. "On the thermodynamics of solutions. V. An equation of state. Fugacities of gaseous solutions", **Chemical Reviews**, v. 44, n. 1, p. 233–244, 1948. DOI: 10.1021/cr60137a013. .

RENON, H., PRAUSNITZ, J. M. M. "Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures", **AICHE Journal**, v. 14, n. 1, p. 135–144, 1968. DOI: 10.1002/aic.690140124. .

SADEQZADEH, M., PAPAIOANNOU, V., DUFAL, S., *et al.* "The development of unlike induced association-site models to study the phase behaviour of aqueous mixtures comprising acetone, alkanes and alkyl carboxylic acids with the SAFT-γ Mie group contribution methodology", **Fluid Phase Equilibria**, v. 407, p. 39–57, 2016. DOI: 10.1016/j.fluid.2015.07.047. Disponível em: http://dx.doi.org/10.1016/j.fluid.2015.07.047.

SADOWSKI, G., BRINKMANN, J., EXNER, L., *et al.* "Pc-saft modeling of phase equilibria relevant for lipid-based drug delivery systems", **Journal of Chemical and Engineering Data**, v. 66, n. 3, p. 1280–1289, 2021. DOI: 10.1021/acs.jced.0c00912. .

SANDLER, S. I. **Chemical, Biochemical and Egineering Thermodynamics**. Hoboken, John Wiley & Sons Inc, 2016.

Schistosomiasis. [S.d.]. Disponível em: https://www.who.int/news-room/fact-sheets/detail/schistosomiasis. Acesso em: 10 mar. 2024.

SHA, J., GONG, Y., CAO, Z., *et al.* "Solid-liquid phase equilibrium of praziquantel in eleven pure solvents: Determination, model correlation, solvent effect, molecular simulation and thermodynamic analysis", **Journal of Chemical Thermodynamics**, v. 154, 2021. DOI: 10.1016/j.jct.2020.106327. .

SHEIKHOLESLAMZADEH, E., ROHANI, S. "Solubility prediction of pharmaceutical and chemical compounds in pure and mixed solvents using predictive models", **Industrial and Engineering Chemistry Research**, v. 51, n. 1, p. 464–473, 2012. DOI: 10.1021/ie201344k. .

SMITH, W. R., NEZBEDA, I. V. O., "The Reference Average Mayer-Function (RAM) Perturbation Theory for Molecular Fluids". **Molecular-Based Study of Fluids**, Advances in Chemistry. [S.l.], AMERICAN CHEMICAL SOCIETY, 1983. v. 204. p. 11–235. DOI: doi:10.1021/ba-1983-0204.ch011. Disponível em: https://doi.org/10.1021/ba-1983-0204.ch011.

SOAVE, G. "Equilibrium constants from a modified Redlich-Kwong equation of state", **Chemical Engineering Science**, v. 27, n. 6, p. 1197–1203, 1972. DOI: 10.1016/0009-2509(72)80096-4. .

TAN, S. P., ADIDHARMA, H., RADOSZ, M. "Generalized Procedure for Estimating the Fractions of Nonbonded Associating Molecules and Their Derivatives in Thermodynamic Perturbation Theory", **Industrial and Engineering Chemistry Research**, v. 43, n. 1, p. 203–208, 2004. DOI: 10.1021/ie034041q. .

TAVARES, F. W., CHANG, J., SANDLER, S. I. "Equation of state for the square-well chain fluid based on the dimer version of wertheim's perturbation theory", **Molecular Physics**, v. 86, n. 6, p. 1451–1471, 1995. DOI: 10.1080/00268979500102851. .

TSIVINTZELIS, I., ECONOMOU, I. G., KONTOGEORGIS, G. M. "Modeling the solid-liquid equilibrium in pharmaceutical-solvent mixtures: Systems with complex hydrogen bonding behavior", **AIChE Journal**, v. 55, n. 3, p. 756–770, 2009. DOI: 10.1002/aic.11716. .

VALAVI, M., UKRAINCZYK, M., DEHGHANI, M. R. "Prediction of solubility of active pharmaceutical ingredients by semi- predictive Flory Huggins/Hansen model", **Journal of Molecular Liquids**, v. 246, p. 166–172, 2017. DOI: 10.1016/j.molliq.2017.09.073. Disponível em: http://dx.doi.org/10.1016/j.molliq.2017.09.073.

VATANI, M., ASGHARI, M., VAKILI-NEZHRAD, G. "Application of Genetic Algorithm to the calculation of parameters for NRTL and Two-Suffix Margules models in ternary extraction ionic liquid systems", **Journal of Industrial and Engineering Chemistry**, v. 18, n. 5, p. 1715–1720, 2012. DOI: 10.1016/j.jiec.2012.03.008. Disponível em: http://dx.doi.org/10.1016/j.jiec.2012.03.008.

VON SOLMS, N., MICHELSEN, M. L., KONTOGEORGIS, G. M. "Computational and physical performance of a modified PC-SAFT equation of state for highly asymmetric and associating mixtures", **Industrial and Engineering Chemistry**

Research, v. 42, n. 5, p. 1098–1105, 2003. DOI: 10.1021/ie020753p. .

WAALS, J. D. van der. **Over de Continuiteit van den Gas-en Vloeistoftoestand (On the Continuity of the Gas and Liquid State)**. PhD Thesis. 135 f. Sijthoff, 1873.

WANG, X., RAMÍREZ-HINESTROSA, S., DOBNIKAR, J., *et al.* "The Lennard-Jones potential: When (not) to use it", **Physical Chemistry Chemical Physics**, v. 22, n. 19, p. 10624–10633, 2020. DOI: 10.1039/c9cp05445f. .

WERTHEIM, M. S. "Fluids with highly directional attractive forces. IV. Equilibrium polymerization", **Journal of Statistical Physics**, v. 42, n. 3–4, p. 477–492, 1986. DOI: 10.1007/BF01127722. .

WILSON, G. M. "Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing", **Journal of the American Chemical Society**, v. 86, n. 2, p. 127–130, 1964. DOI: 10.1021/ja01056a002. .

WOLBACH, J. P., SANDLER, S. I. "Using Molecular Orbital Calculations To Describe the Phase Behavior of Cross-associating Mixtures", **Industrial & Engineering Chemistry Research**, v. 37, n. 8, p. 2917–2928, 1998. DOI: 10.1021/ie970781l. Disponível em: https://doi.org/10.1021/ie970781l.

ZARGARZADEH, L., ELLIOTT, J. A. W. "Comparison of the Osmotic Virial Equation with the Margules Activity Model for Solid-Liquid Equilibrium", **Journal of Physical Chemistry B**, v. 123, n. 5, p. 1099–1107, 2019. DOI: 10.1021/acs.jpcb.8b09285. .

ZHANG, B. J. "Calculating thermodynamic properties from perturbation theory. I. An analytic representation of square-well potential hard-sphere perturbation theory", **Fluid Phase Equilibria**, v. 154, n. 1, p. 1–10, 1999. DOI: 10.1016/S0378-3812(98)00431-2. .

7. Appendix

The following appendix presents supplementary figures with corresponding literature sources given by NIST, Liu *et al.*, (2020), Li *et al.* (2020) and de Moraes et al. (yet to be published).

7.1. Appendix I. VLE Systems

7.1.1. Isochores

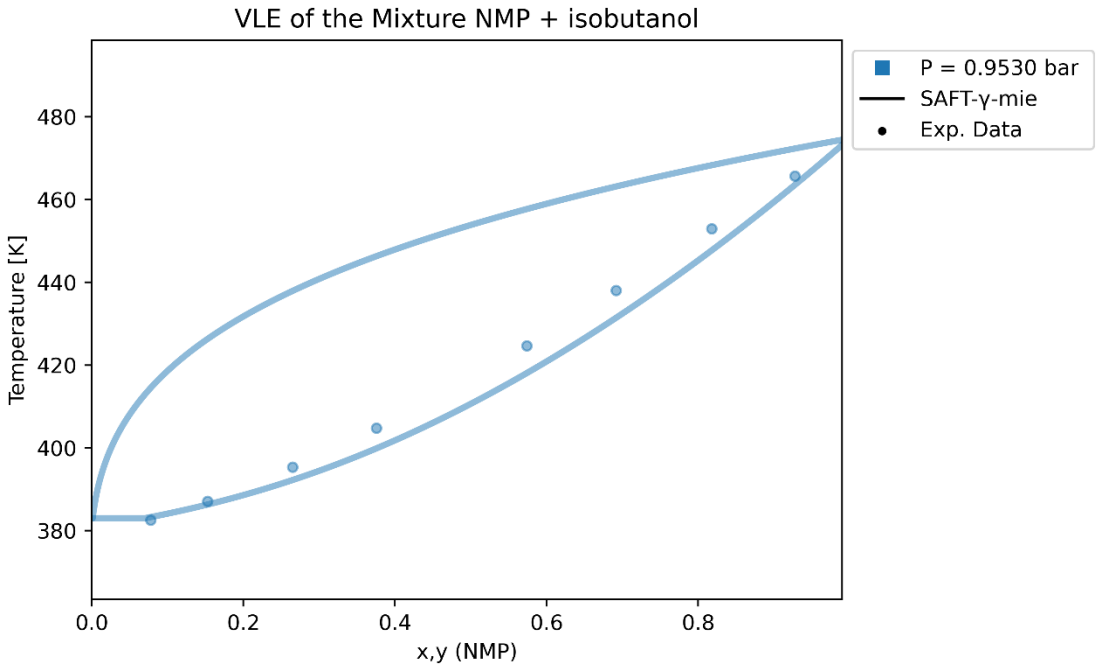

Figure 13: Temperature-composition relationship of the NMP + isobutanol system at 0.9530 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

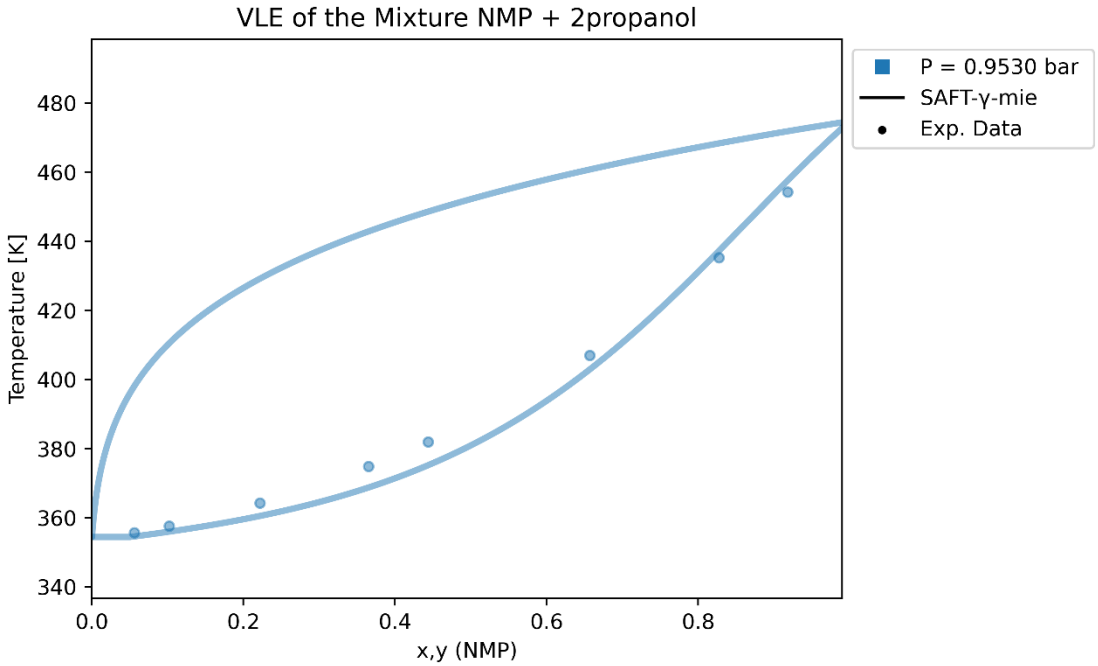

Figure 14: Temperature-composition relationship of the NMP + 2-propanol (2propanol) system at 0.9530 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

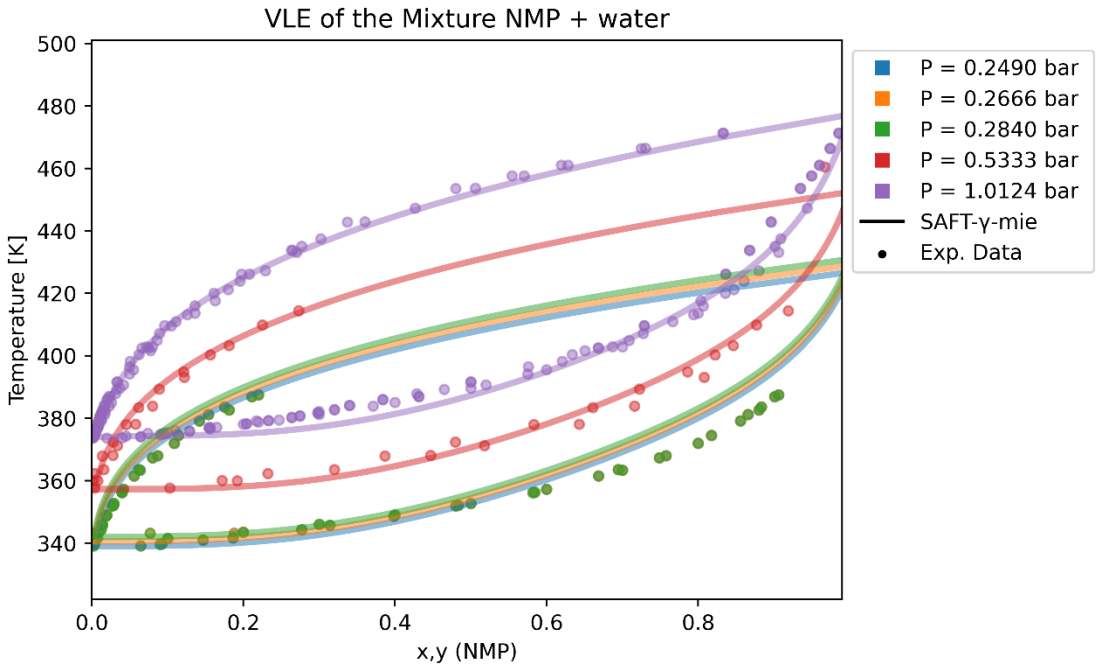

Figure 15: Temperature-Composition Relationship of the NMP + water System under pressures ranging from 0.249 to 1.0124 bar. Different colors are used to denote various isochores, each representing a pressure variation of ±5% around the indicated values. A direct comparison is made between experimental data from NIST, shown as discrete

points, and the predictions calculated using the SAFT-γ Mie EOS, illustrated through continuous lines.

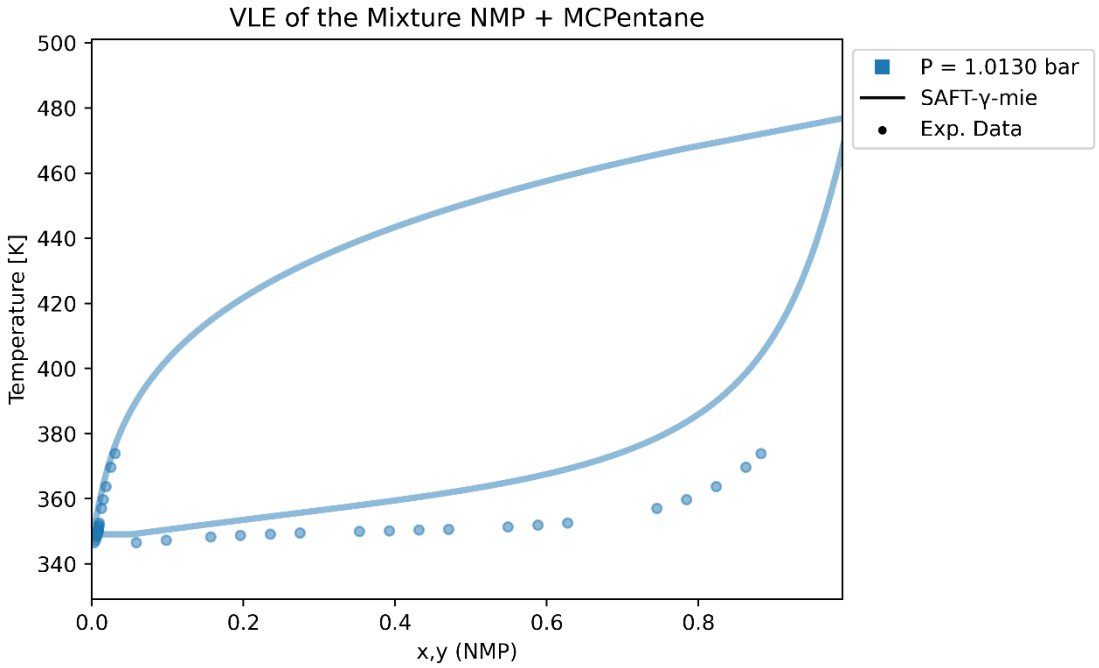

Figure 16: Temperature-composition relationship of the NMP + methylcyclopentane (MCPentane) system at 1.013 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

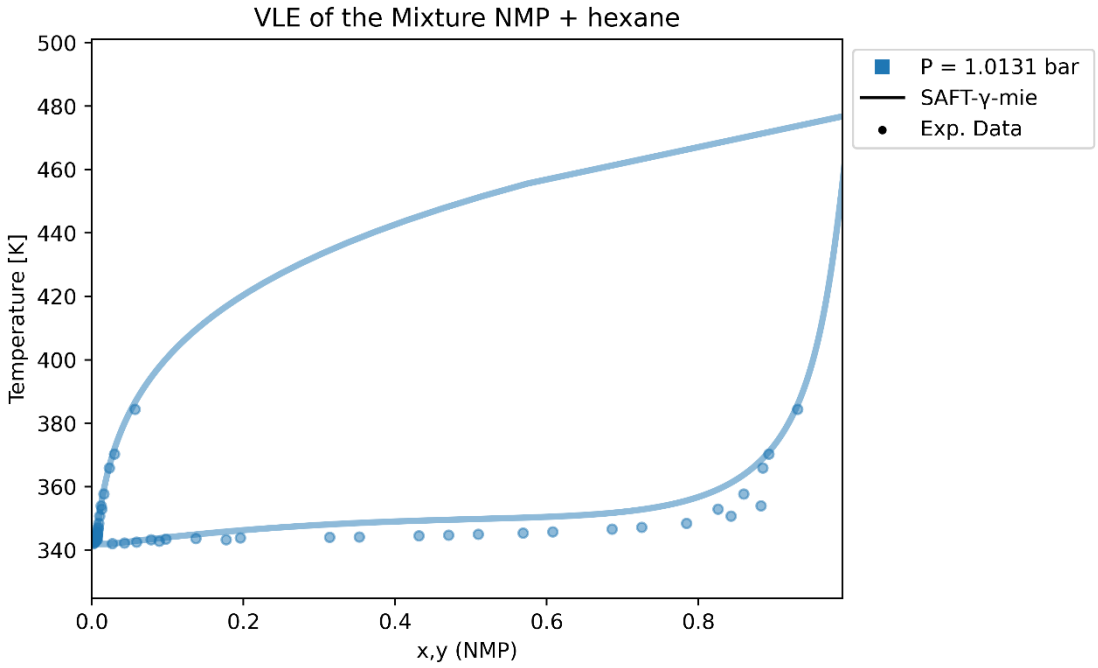

Figure 17: Temperature-composition relationship of the NMP + hexane system at 1.0131 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

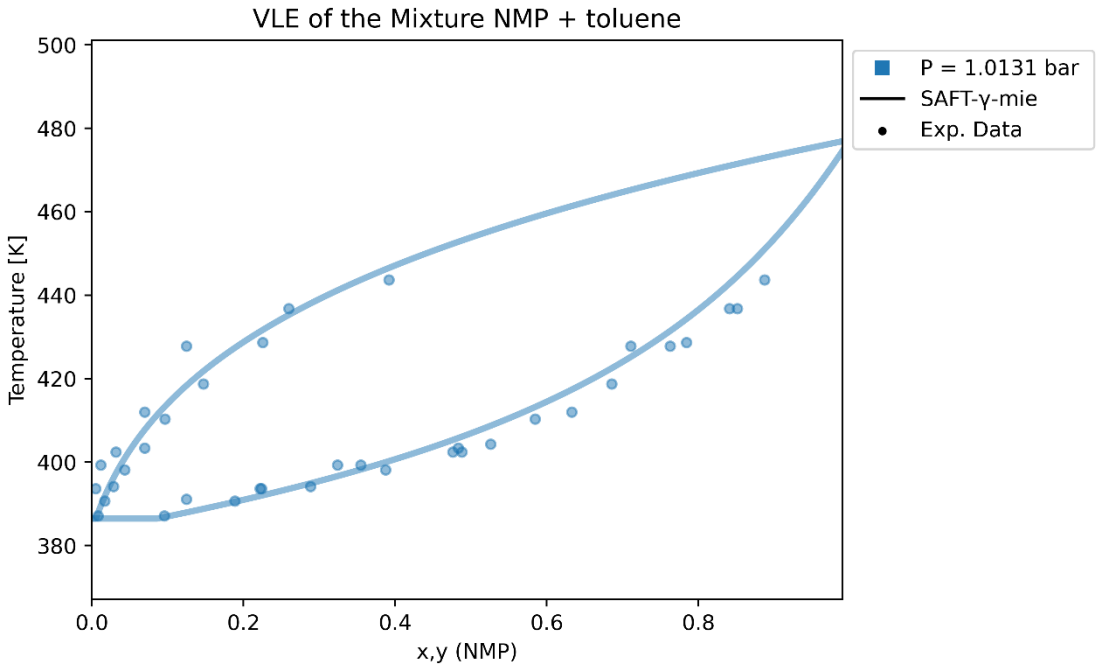

Figure 18: Temperature-composition relationship of the NMP + toluene system at approximately 1.0131 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

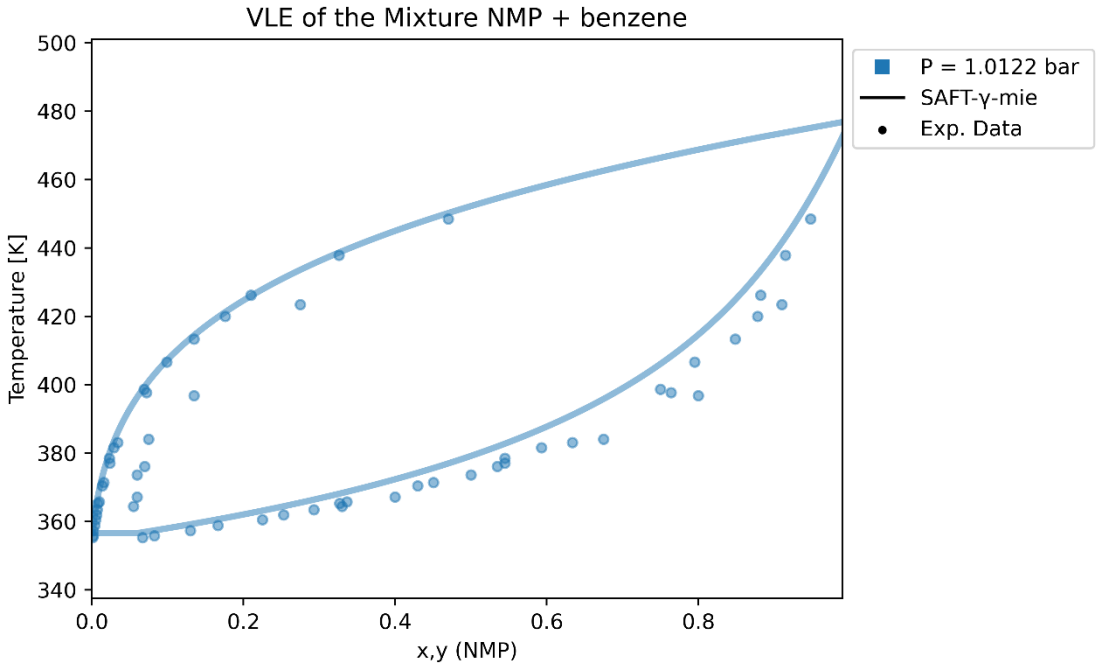

Figure 19: Temperature-composition relationship of the NMP + benzene system at approximately 1.0131 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

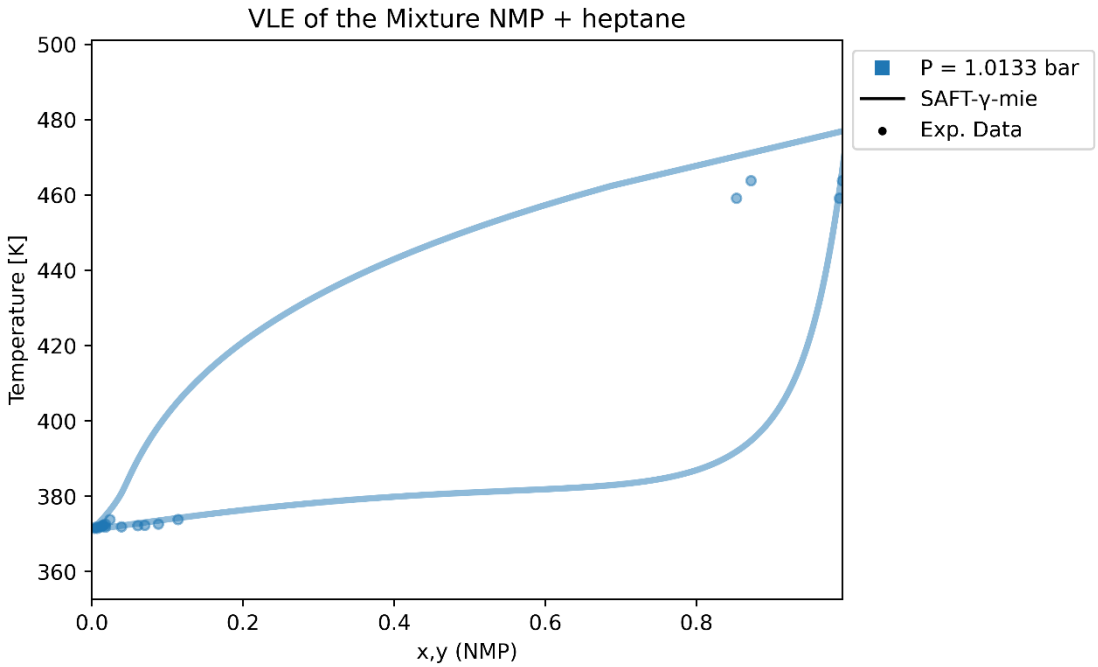

Figure 20: Temperature-composition relationship of the NMP + heptane system at 1.0133 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

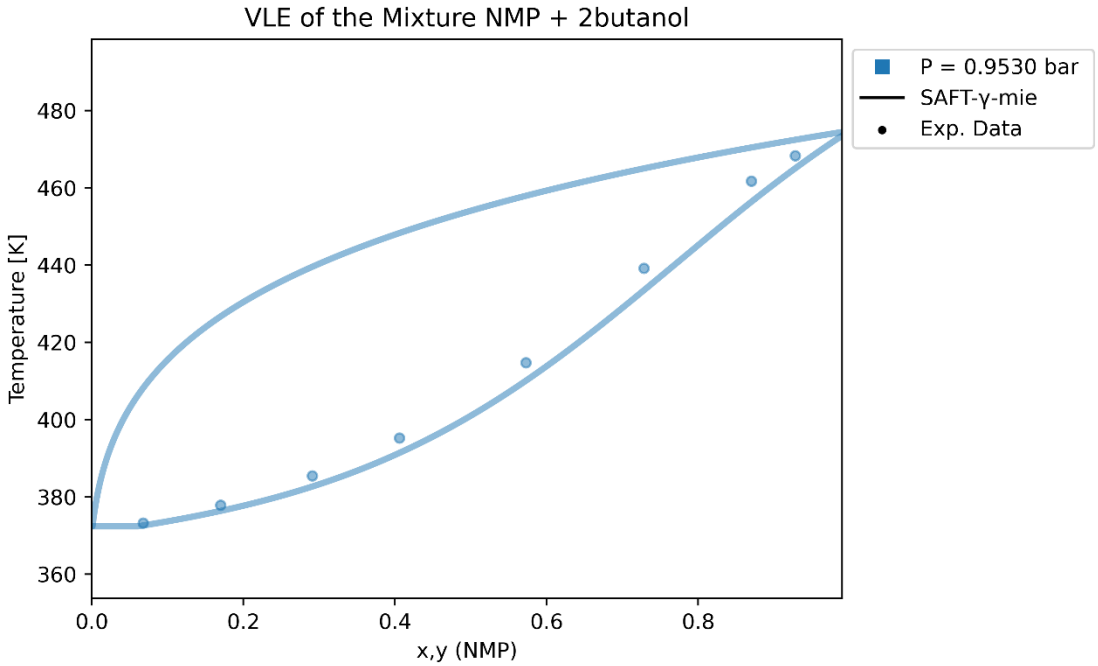

Figure 21: Temperature-composition relationship of the NMP + 2-butanol (2butanol) system at 0.953 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

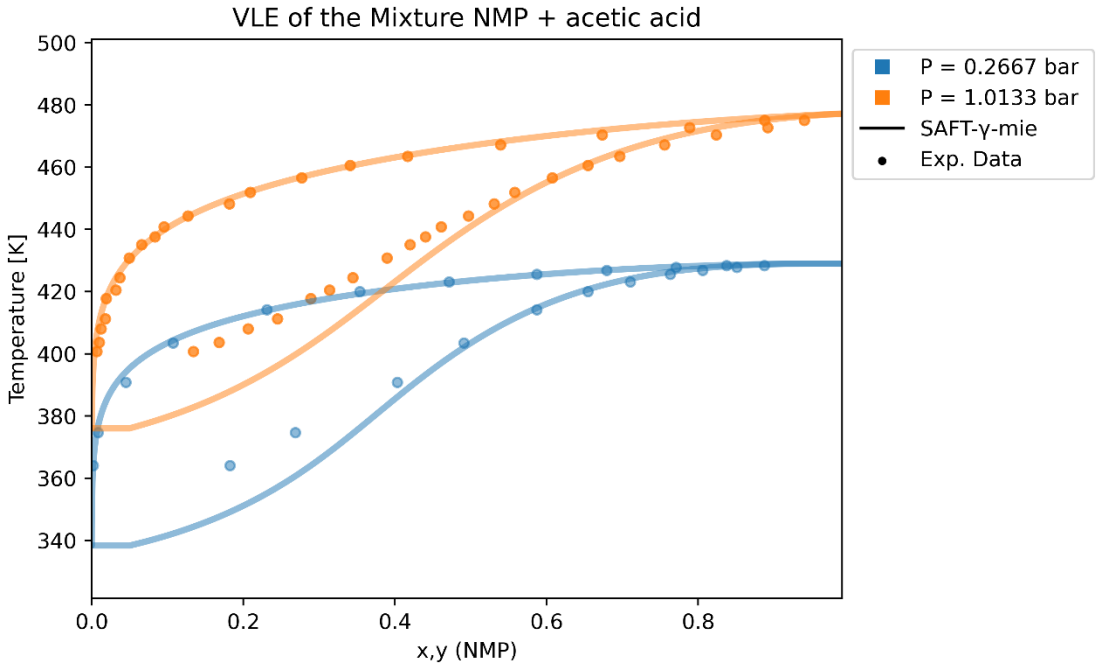

Figure 22: Temperature-composition relationship of the NMP + acetic acid system at 0.2667 and 1.0133 bar. It is provided a direct comparison between experimental data,

sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

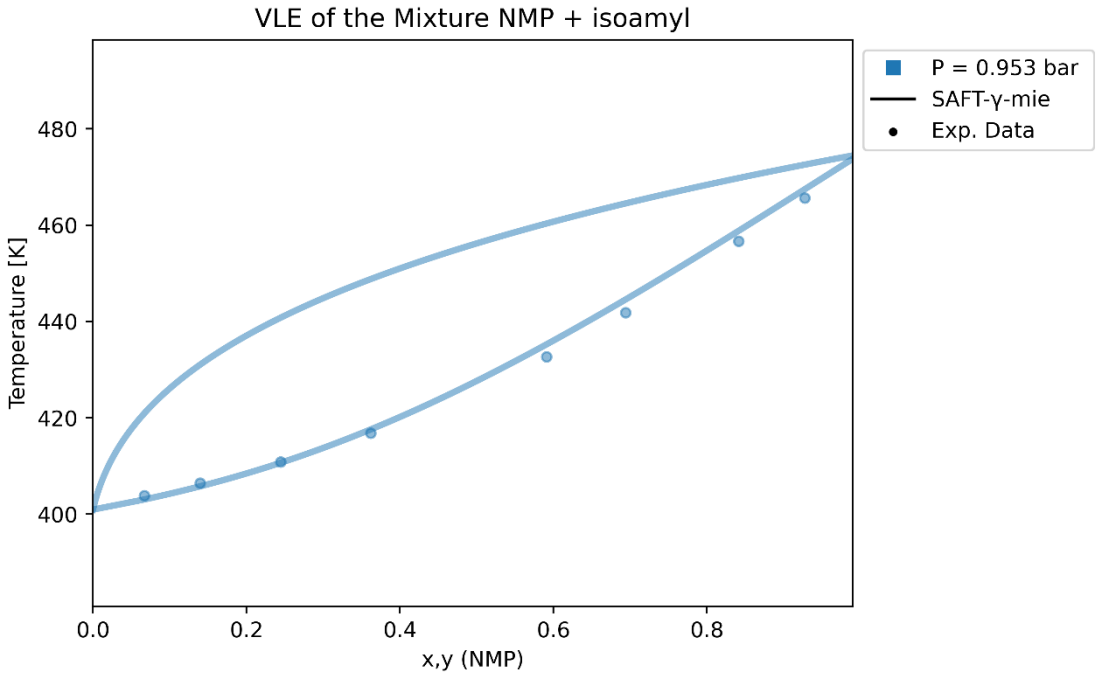

Figure 23: Temperature-composition relationship of the NMP + isoamyl alcohol system at 0.953 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

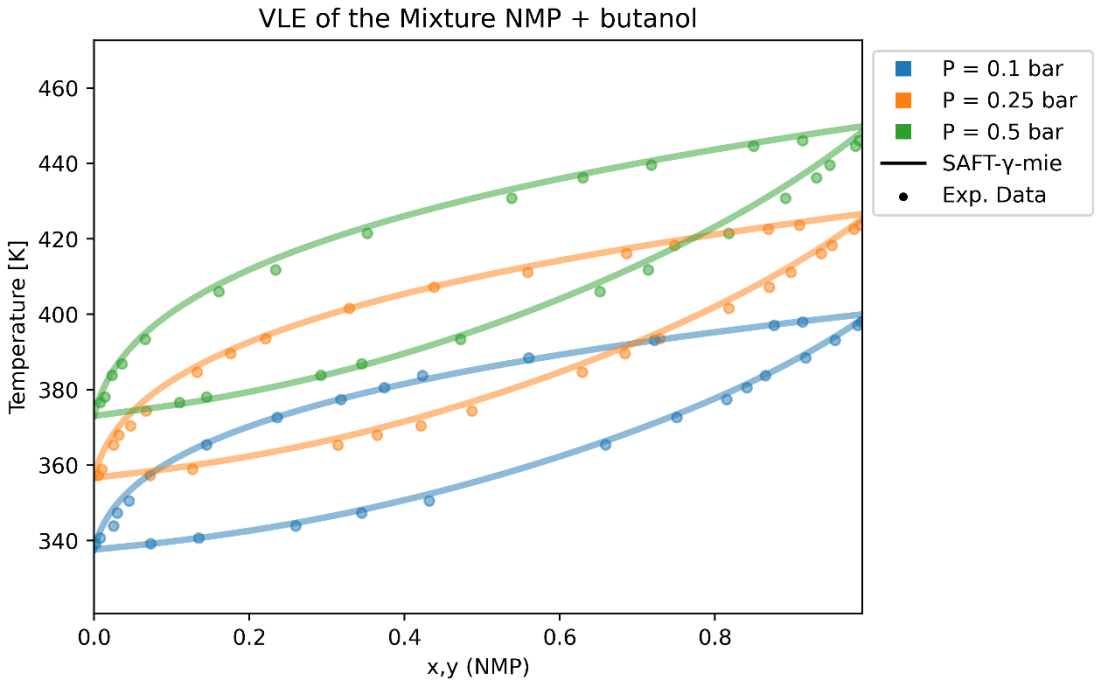

Figure 24: Temperature-composition relationship of the NMP + butanol system under pressures ranging from 0.1 to 0.5 bar. Different colors are used to denote various isochores. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

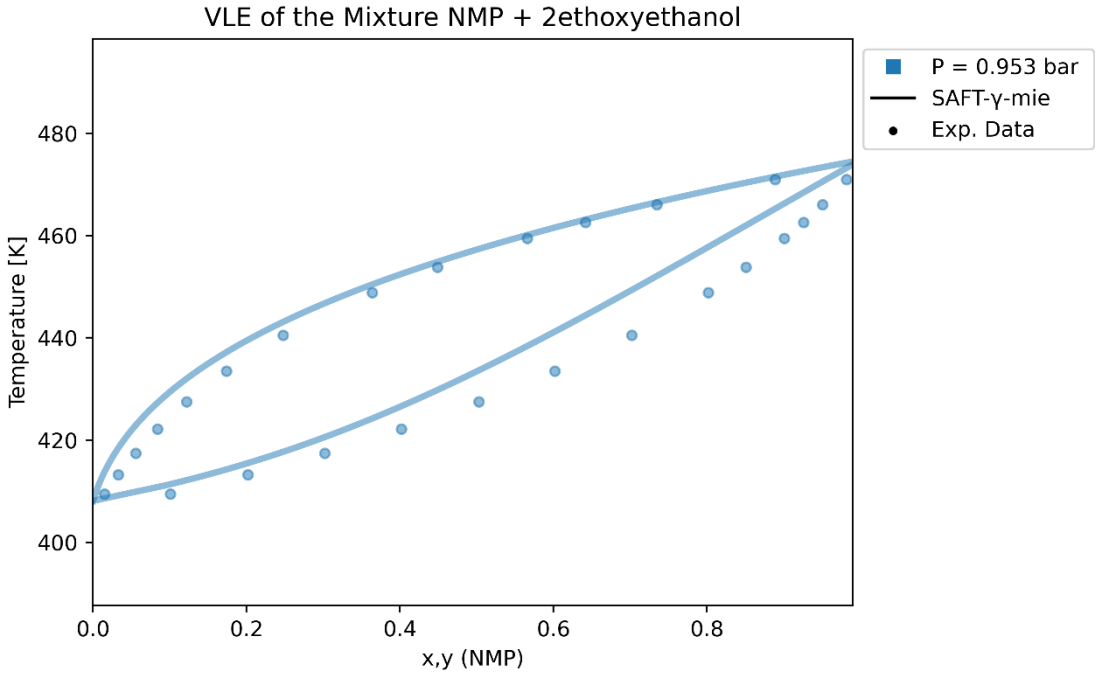

Figure 25: Temperature-composition relationship of the NMP + 2-ethoxyethanol (2 ethoxyethanol) system at 0.953 bar. It is provided a direct comparison between

experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

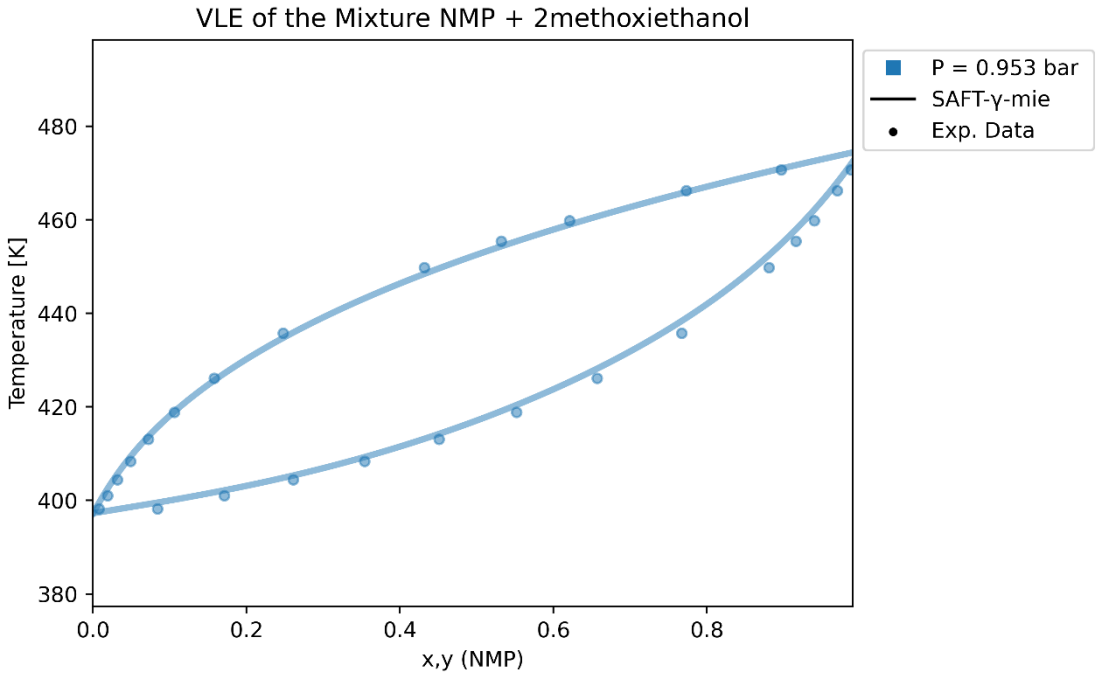

Figure 26: Temperature-composition relationship of the NMP + 2-methoxyethanol (2methoxyethanol) system at 0.953 bar. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

7.1.2. Isotherms

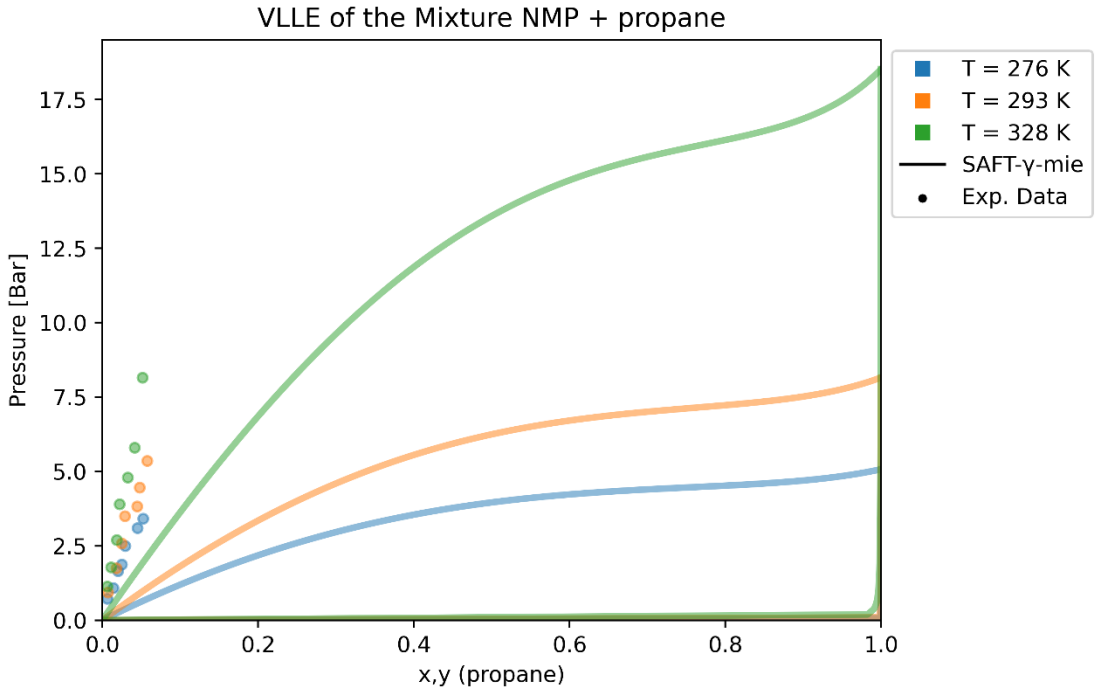

Figure 27: Pressure-composition relationship of the NMP + propane system under temperatures ranging from 276 to 328 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

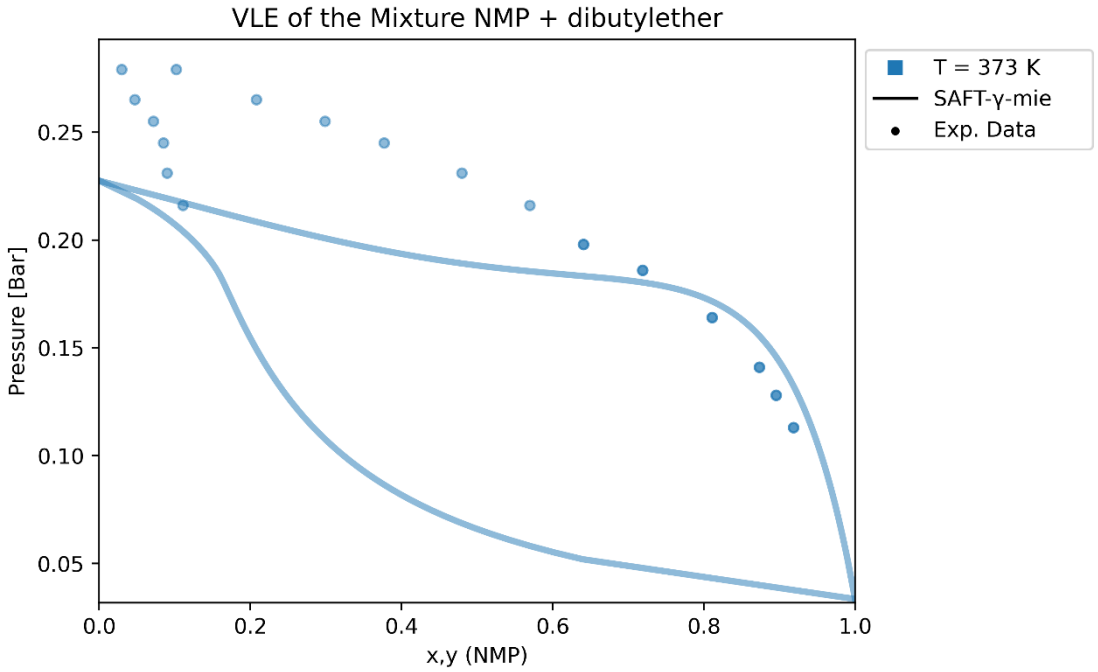

Figure 28: Pressure-composition relationship of the NMP + dibutyl ether (dibutylether) system at 373 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

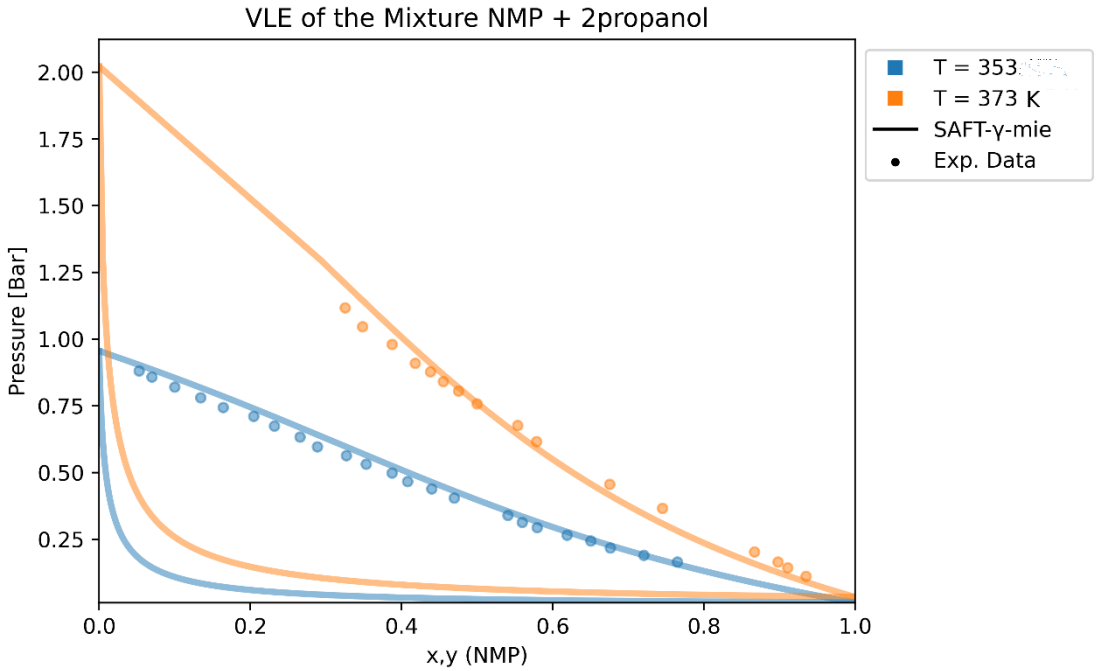

Figure 29: Pressure-composition relationship of the NMP + 2-propanol system at 353 and 373 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

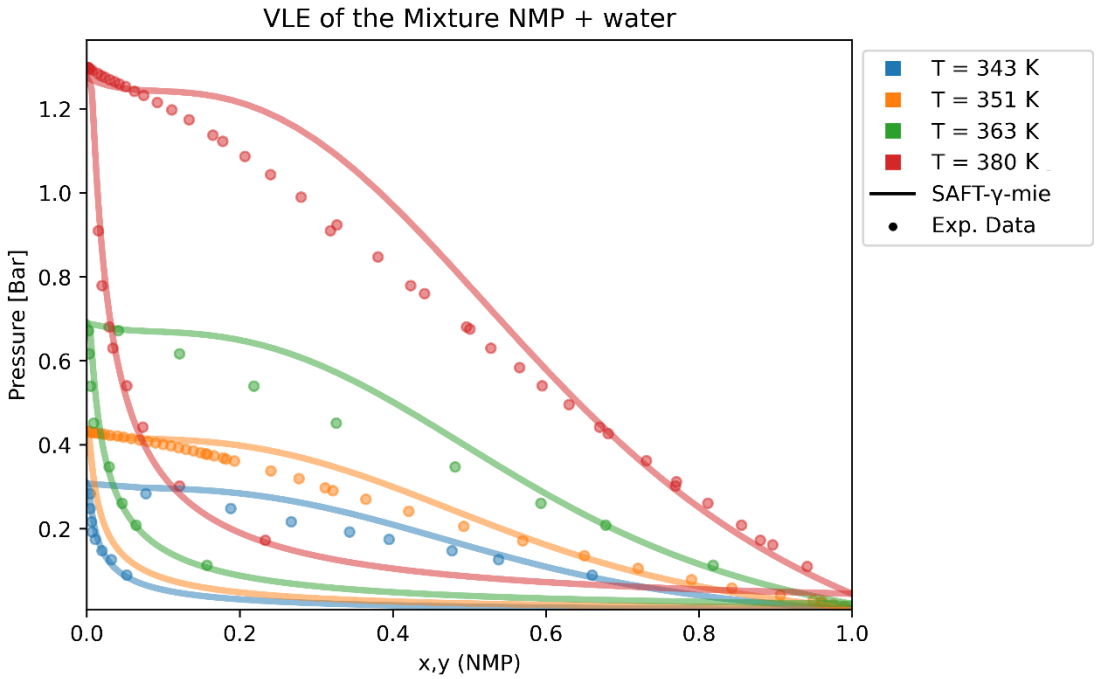

Figure 30: Pressure-composition relationship of the NMP + water system under temperatures ranging from 343 to 380 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

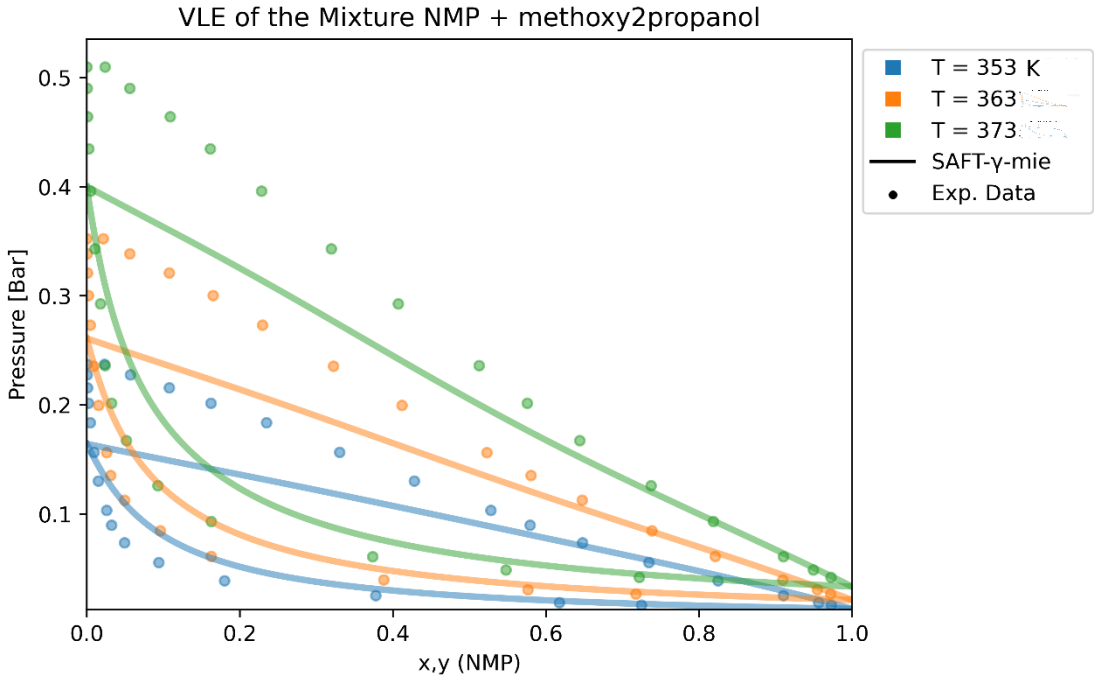

Figure 31: Pressure-composition relationship of the NMP + 1-methoxy-2-propanol (methoxy2propanol) system under temperatures ranging from 353 to 373 K. It is

provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

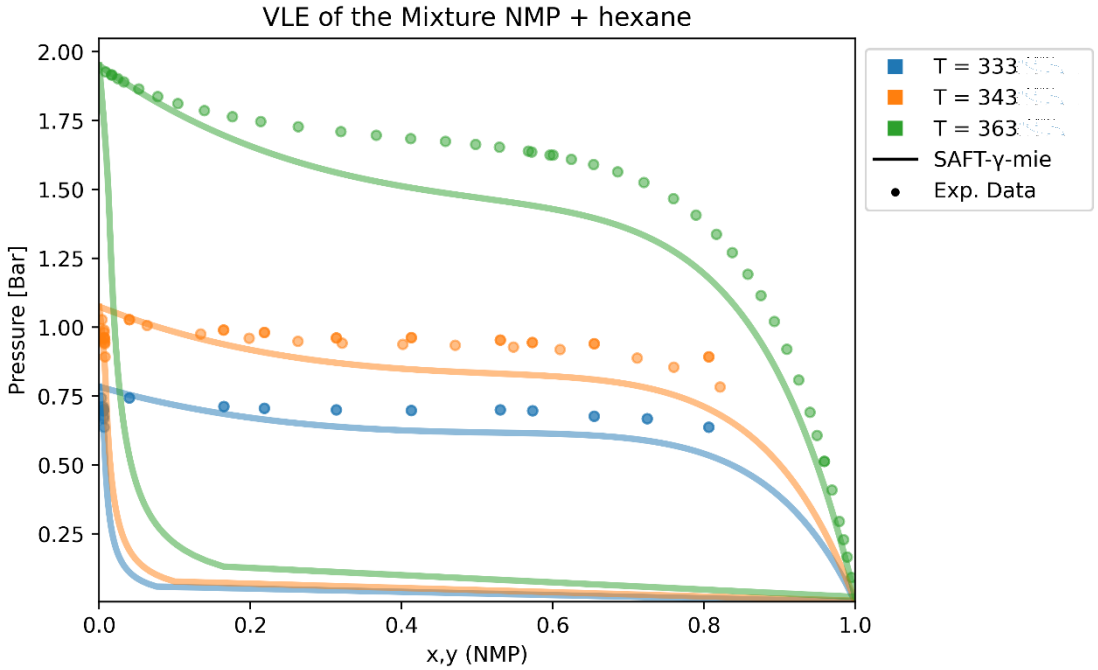

Figure 32: Pressure-composition relationship of the NMP + hexane system under temperatures ranging from 333 to 363 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS, illustrated through continuous lines.

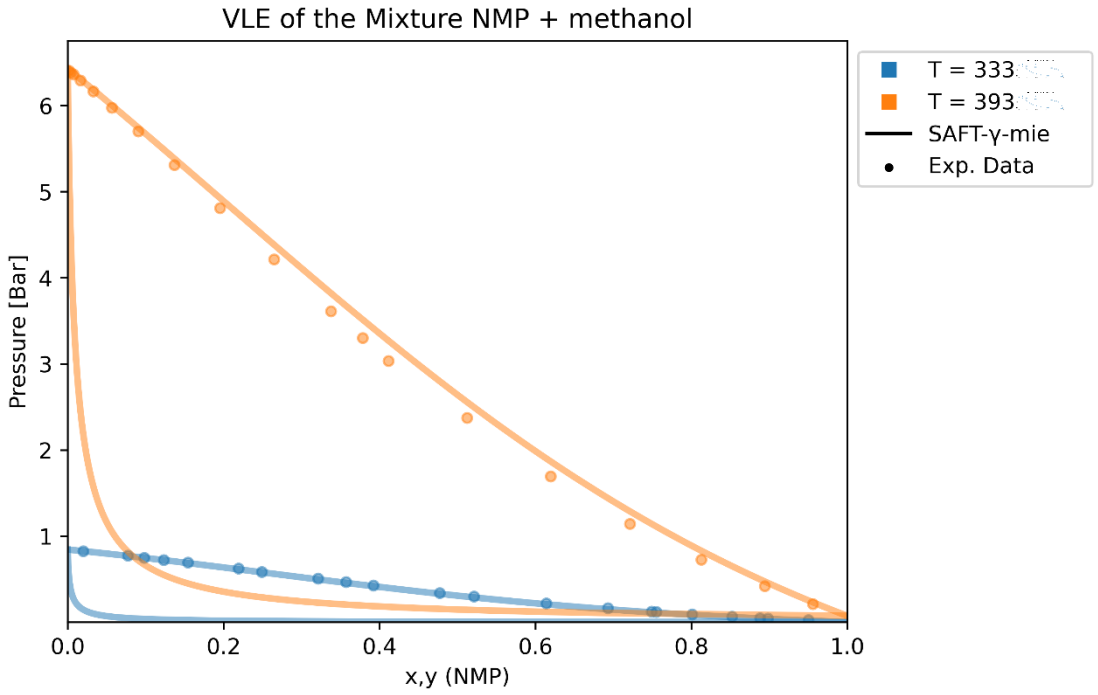

Figure 33: Pressure-composition relationship of the NMP + methanol system at 333 and 393 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

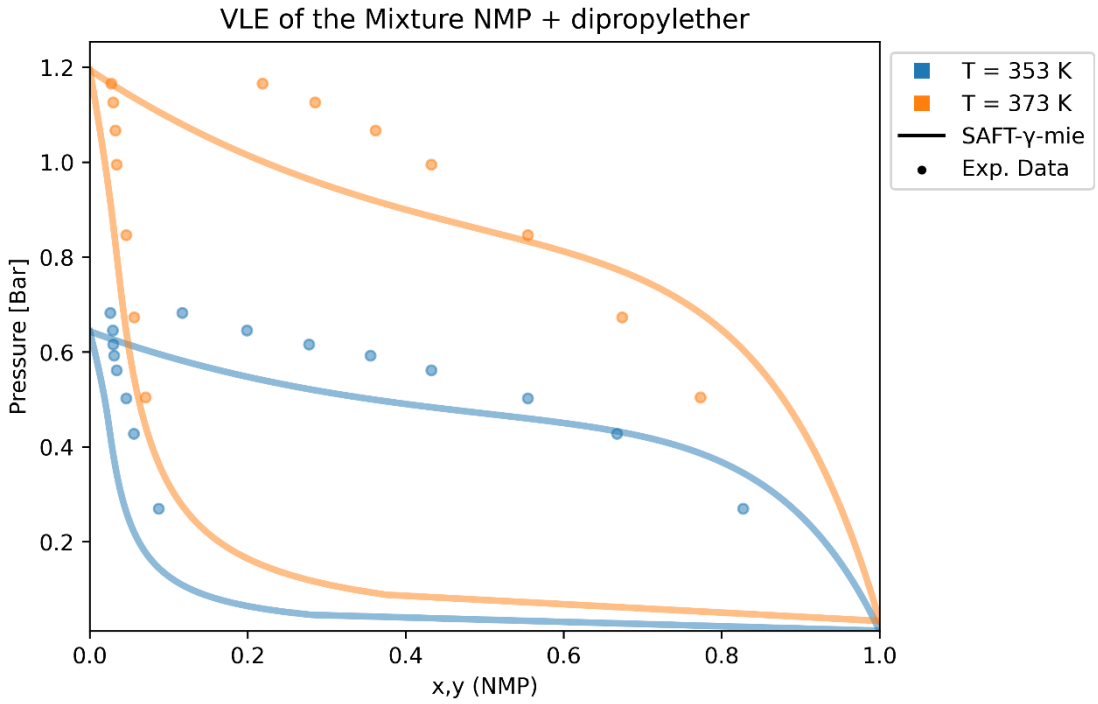

Figure 34: Pressure-composition relationship of the NMP + dipropyl ether (dipropylether) system at 353 and 373 K. It is provided a direct comparison between

experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

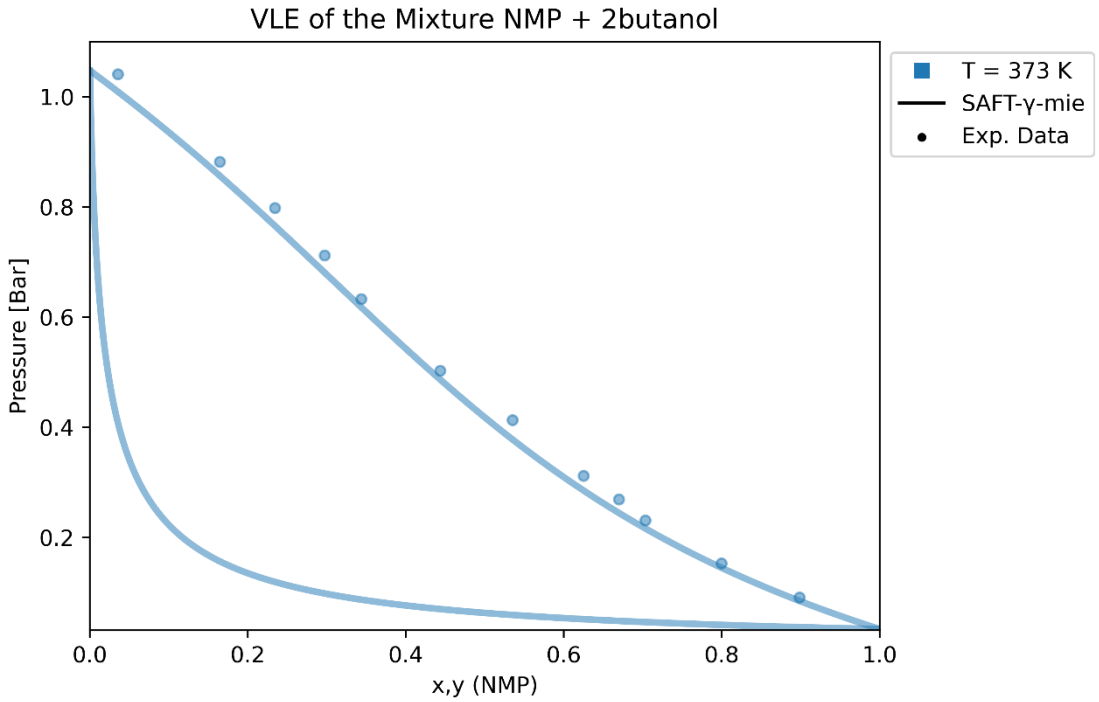

Figure 35: Pressure-composition relationship of the NMP + 2-butanol system at 373 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

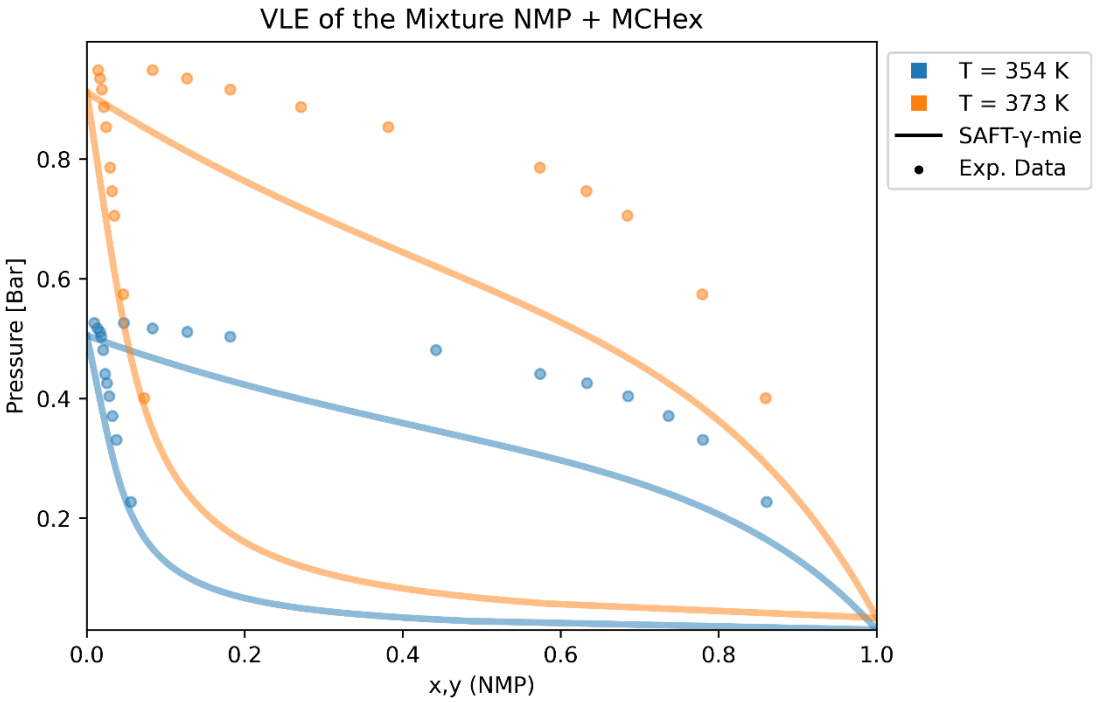

Figure 36: Pressure-composition relationship of the NMP + methylcyclohexane (MCHex) system at 354 and 373 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

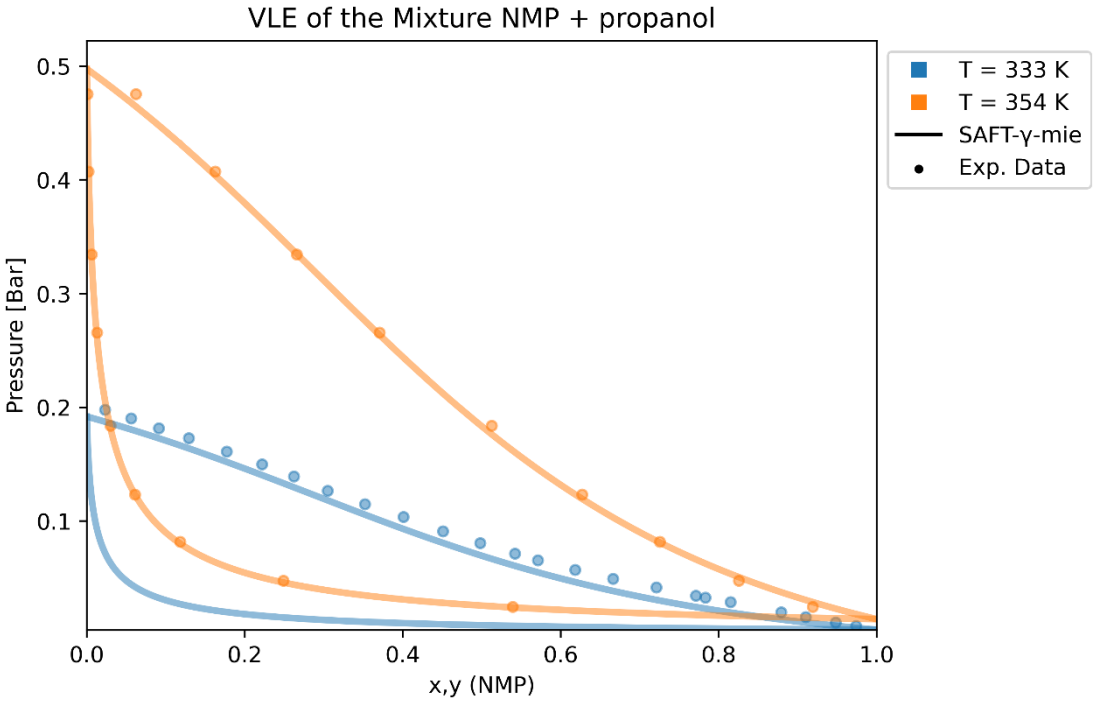

Figure 37:Pressure-composition relationship of the NMP + propanol system at 333 and 354 K. It is provided a direct comparison between experimental data, sourced from

NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

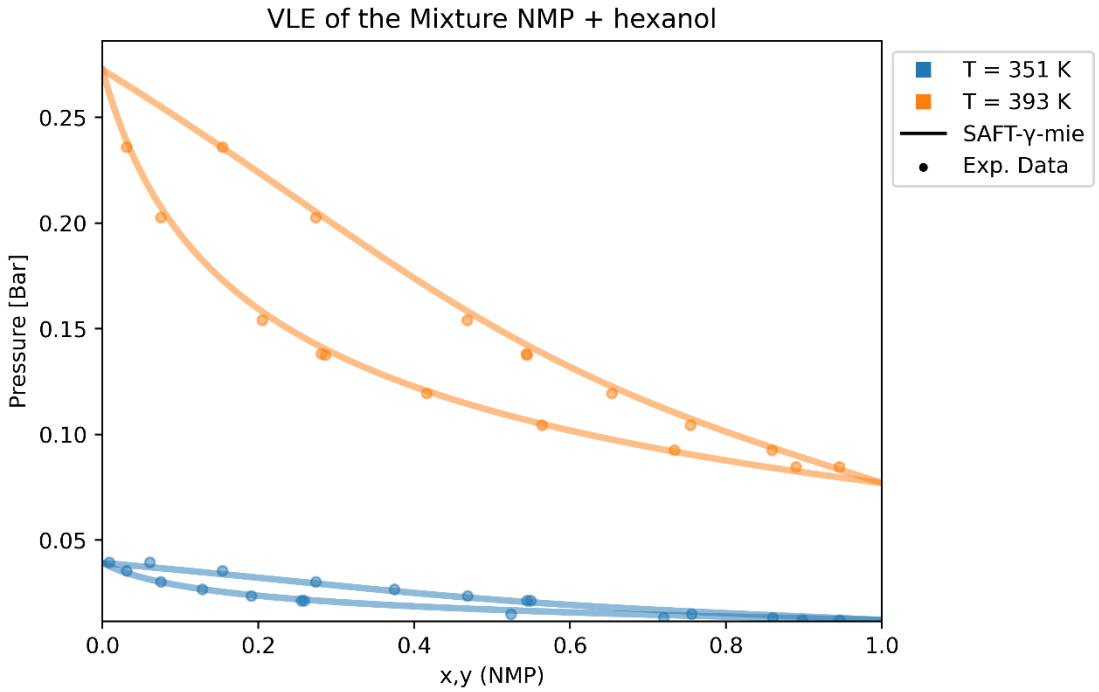

Figure 38: Pressure-composition relationship of the NMP + hexanol system at 351 and 393 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

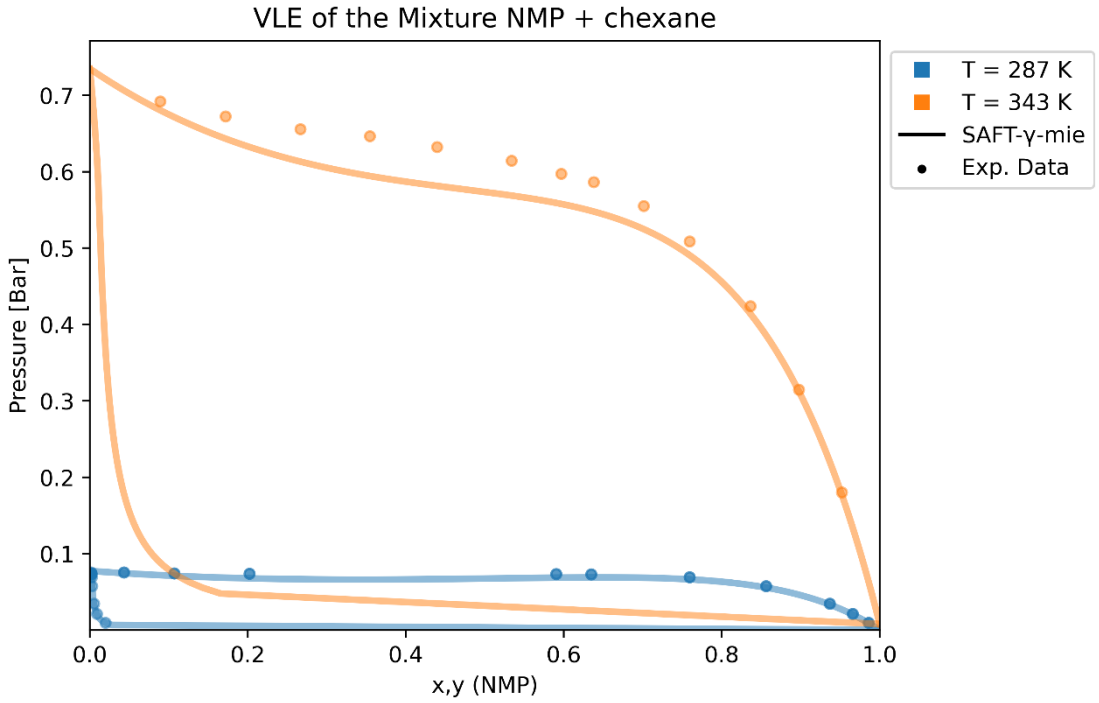

Figure 39: Pressure-composition relationship of the NMP + cyclohexane (chexane) system at 287 and 343 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

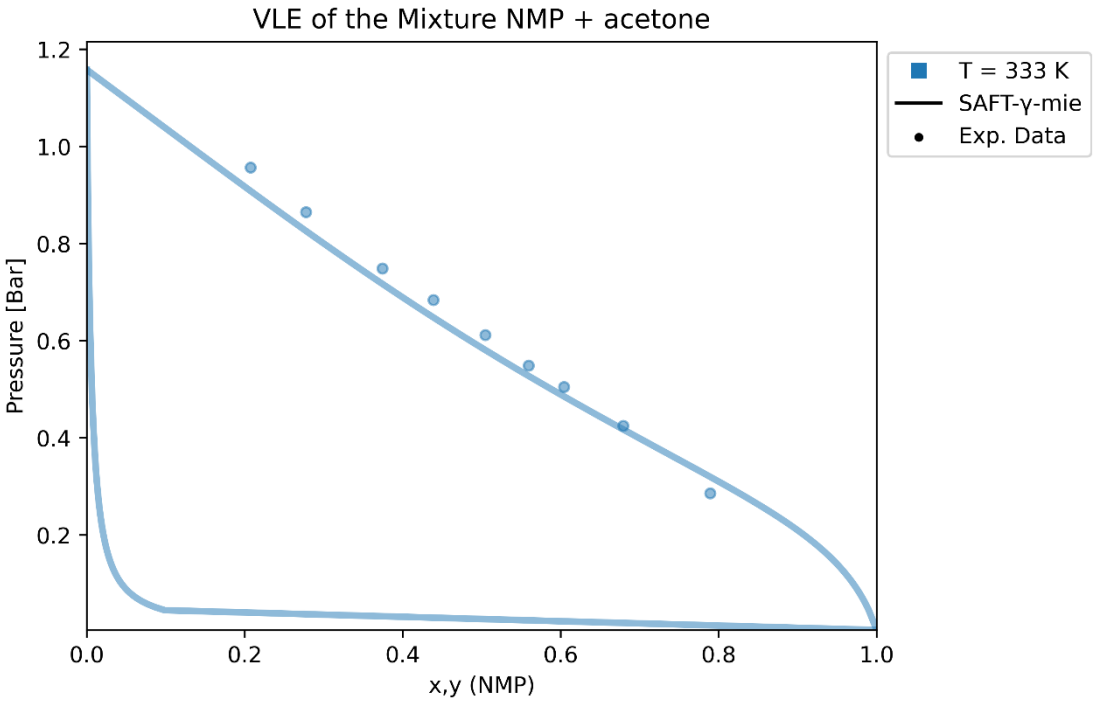

Figure 40: Pressure-composition relationship of the NMP + acetone system at 333 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

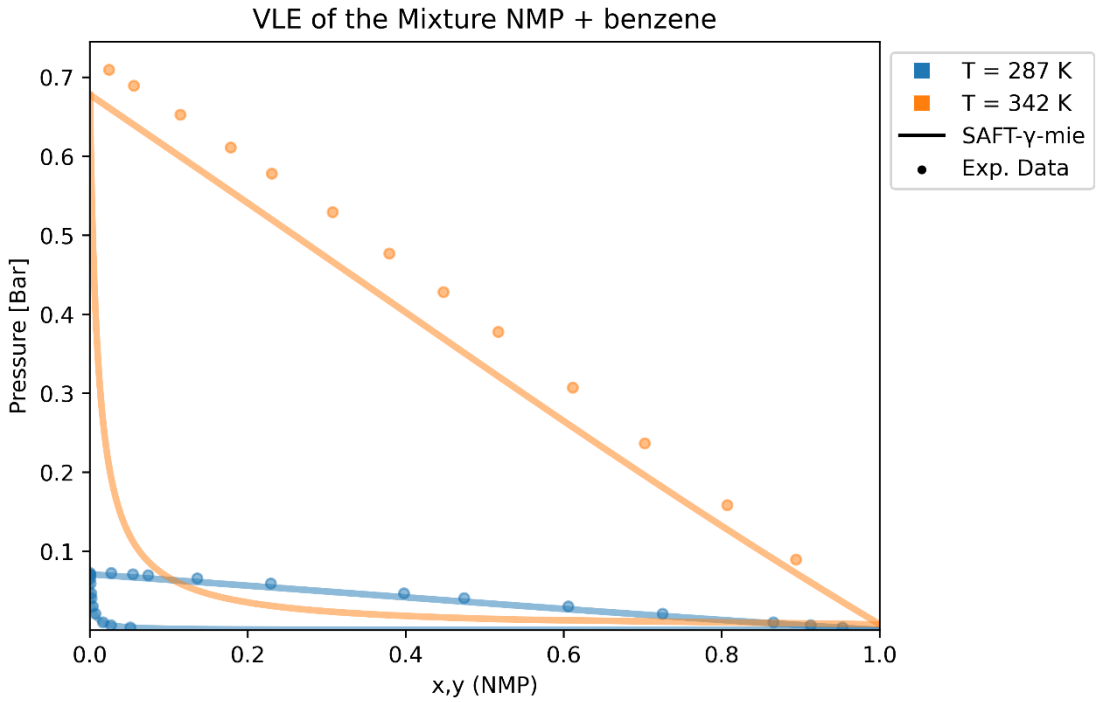

Figure 41: Pressure-composition relationship of the NMP + benzene system at 287 and 342 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

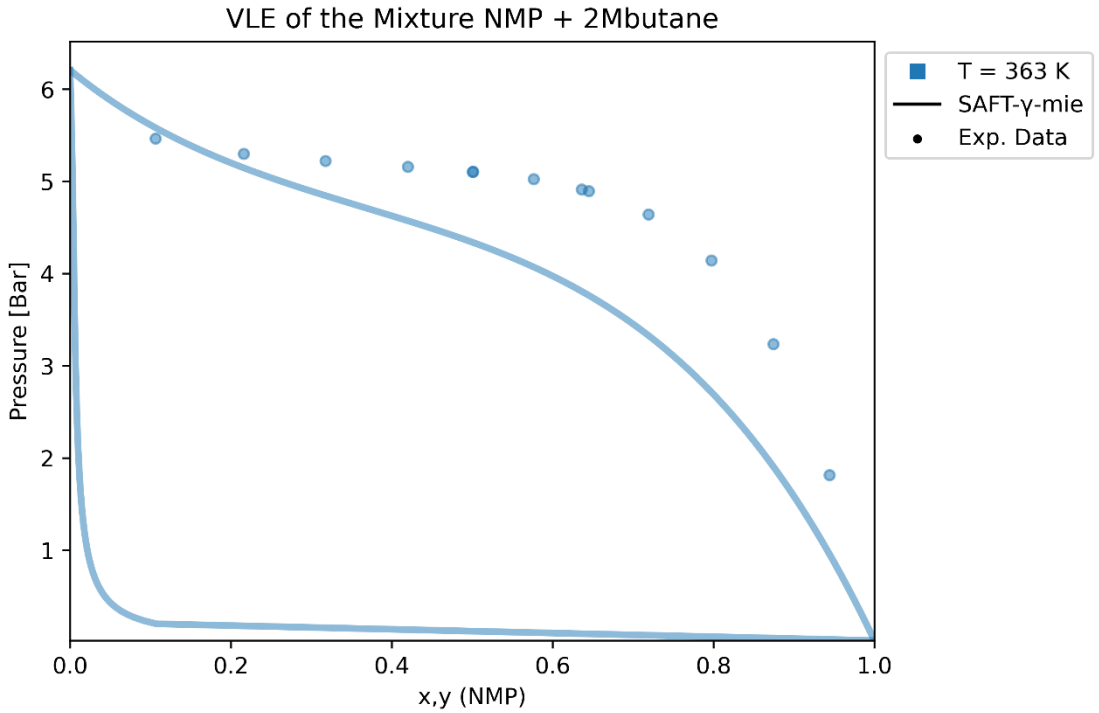

Figure 42: Pressure-composition relationship of the NMP + 2-methylbutane (2Mbutane) system at 363 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

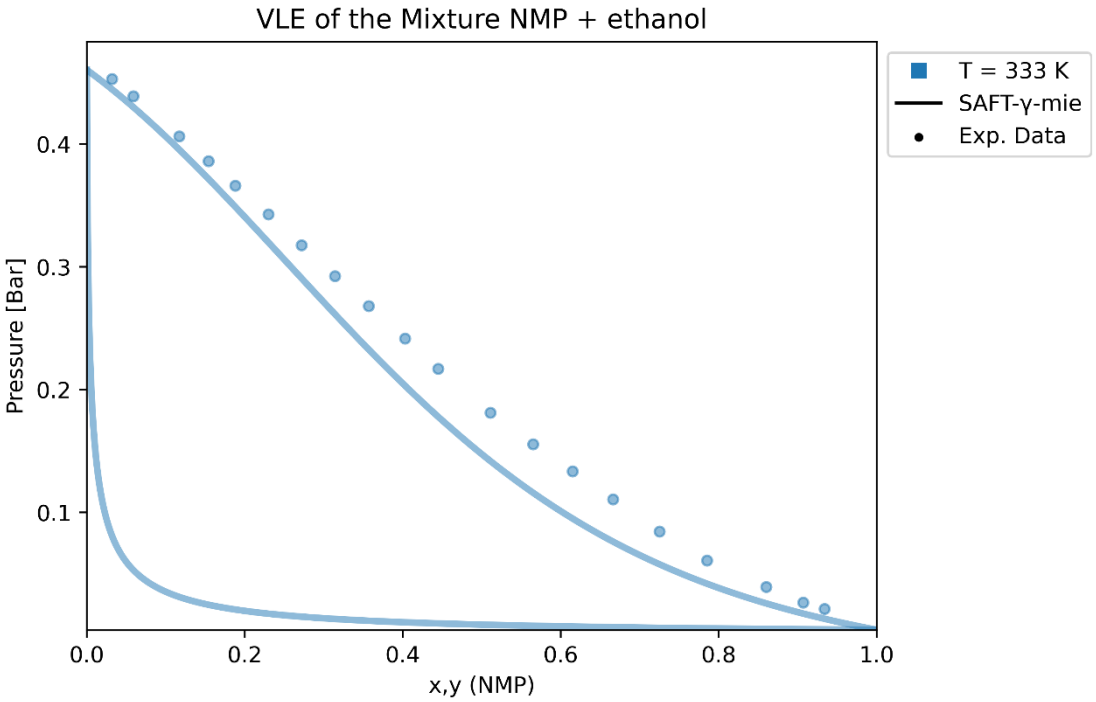

Figure 43: Pressure-composition relationship of the NMP + ethanol system at 363 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

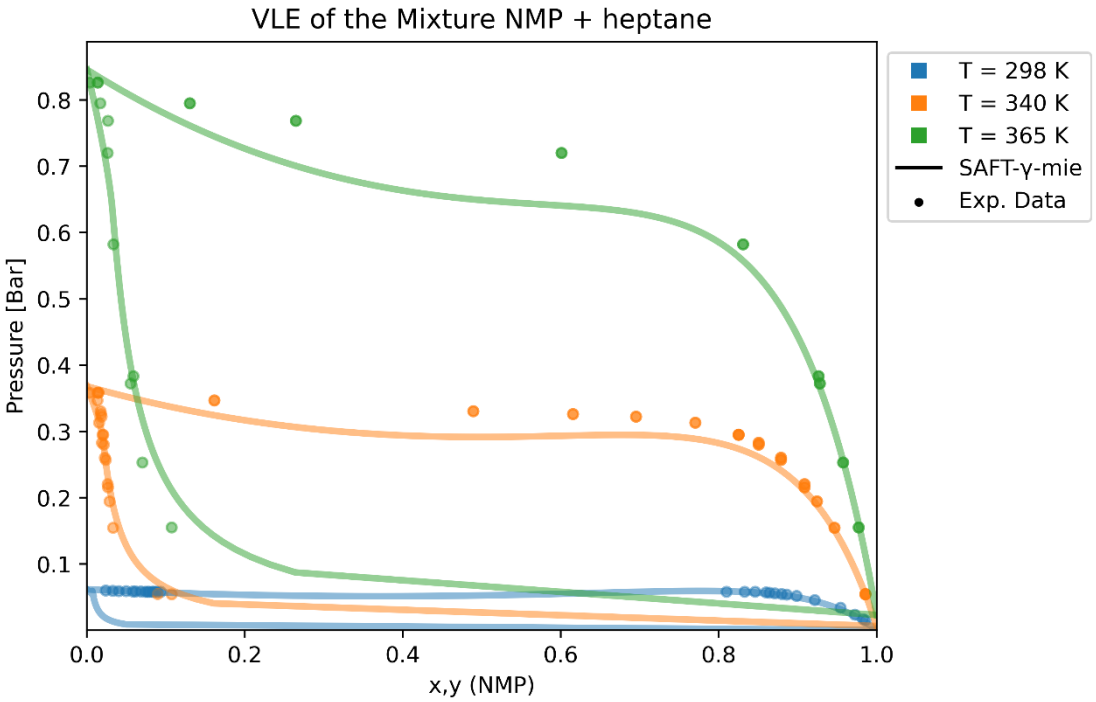

Figure 44: Pressure-composition relationship of the NMP + heptane system under temperatures ranging from 298 to 365 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

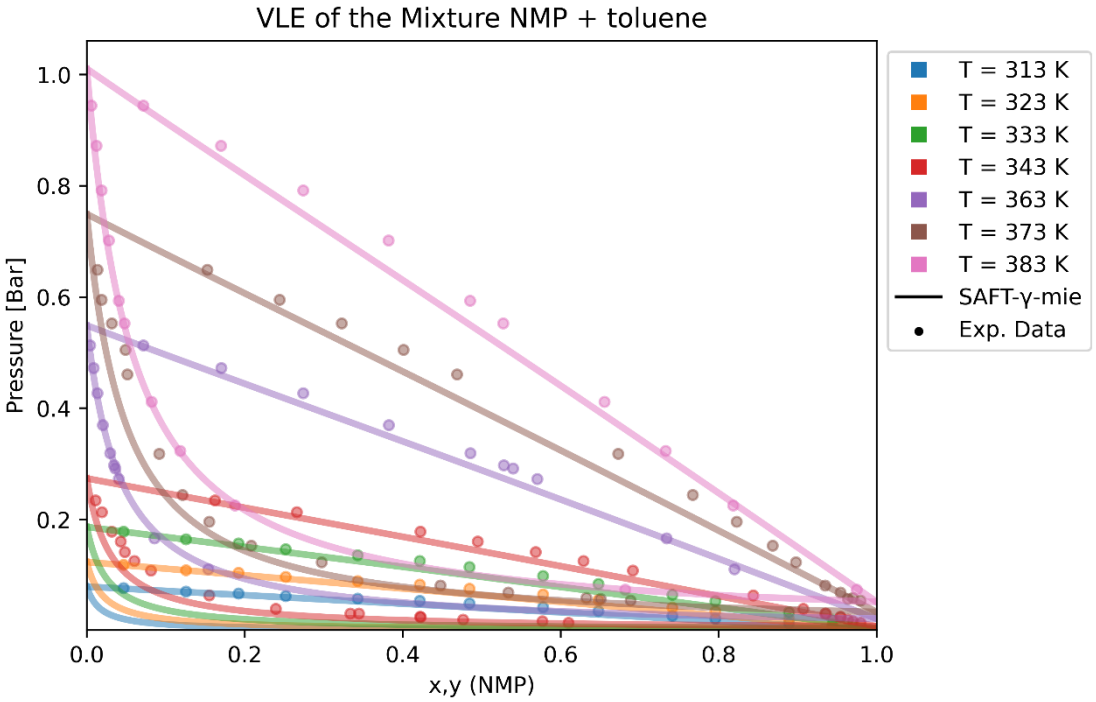

Figure 45: Pressure-composition relationship of the NMP + toluene system under temperatures ranging from 313 to 383 K. It is provided a direct comparison between

experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line.

7.2. Appendix II. VLLE System

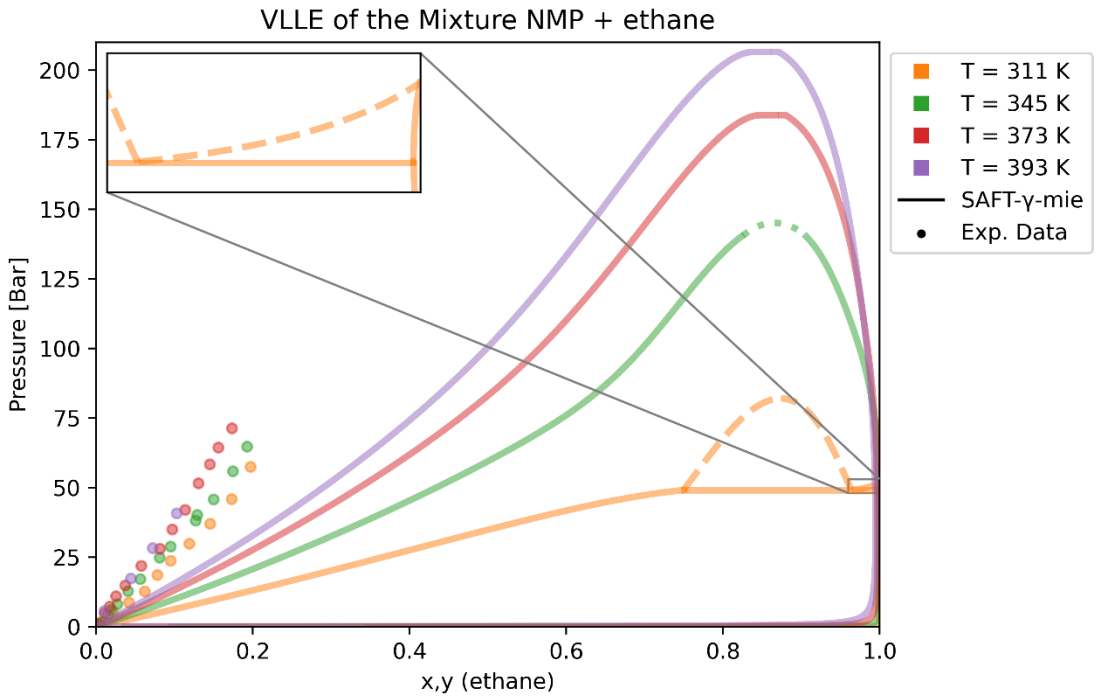

Figure 46: Pressure-composition relationship of the NMP + ethane system under temperatures ranging from 311 to 393 K. It is provided a direct comparison between experimental data, sourced from NIST, represented by discrete points, and predictions obtained using the SAFT-γ Mie EOS depicted as a continuous line for the VLE and dashed line for the liquid-liquid equilibrium. Due to difficulties in the convergence of some points, the dotted line describes an extrapolation of the VLE without using the SAFT-γ Mie model, but based on the thermodynamics of the mixture.

7.3. Appendix III. SLE Systems

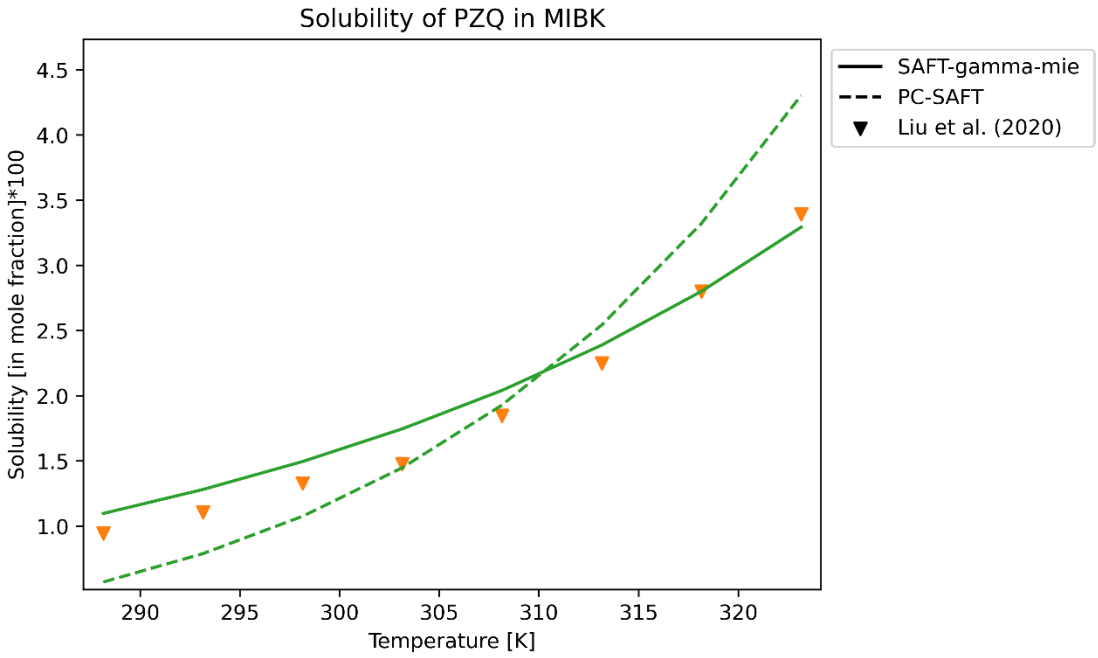

Figure 47: Solubility temperature curve of PZQ in methyl isobutyl ketone (MIBK) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by triangles, were determined and reported by Liu *et al.*, (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

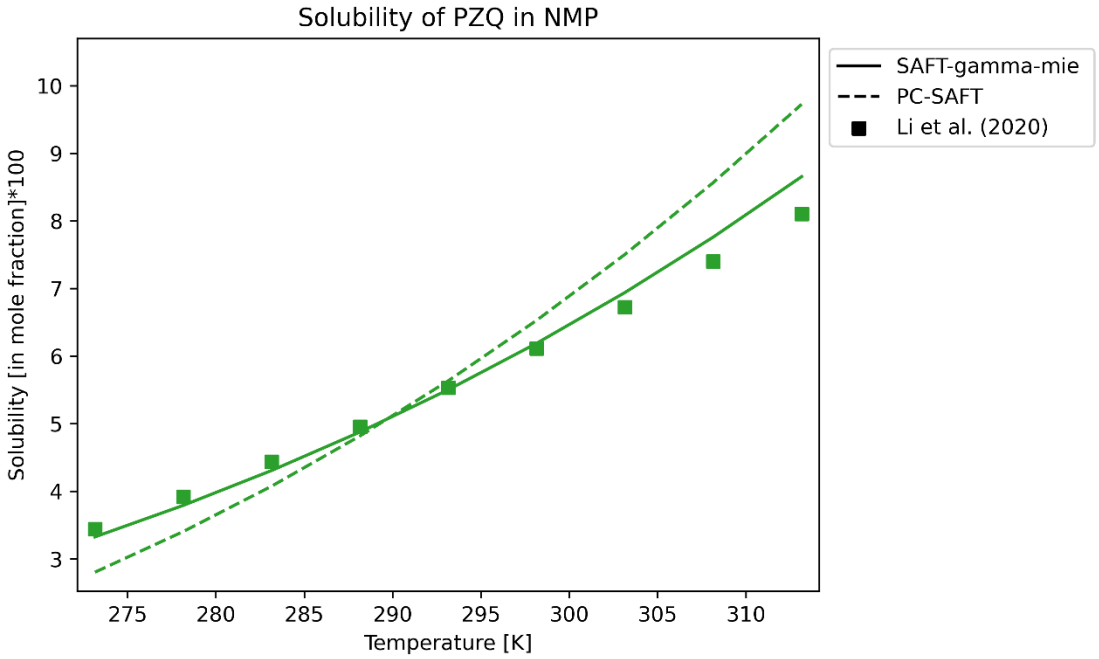

Figure 48: Solubility temperature curve of PZQ in NMP at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by squares, were determined and reported by Li *et al.* (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

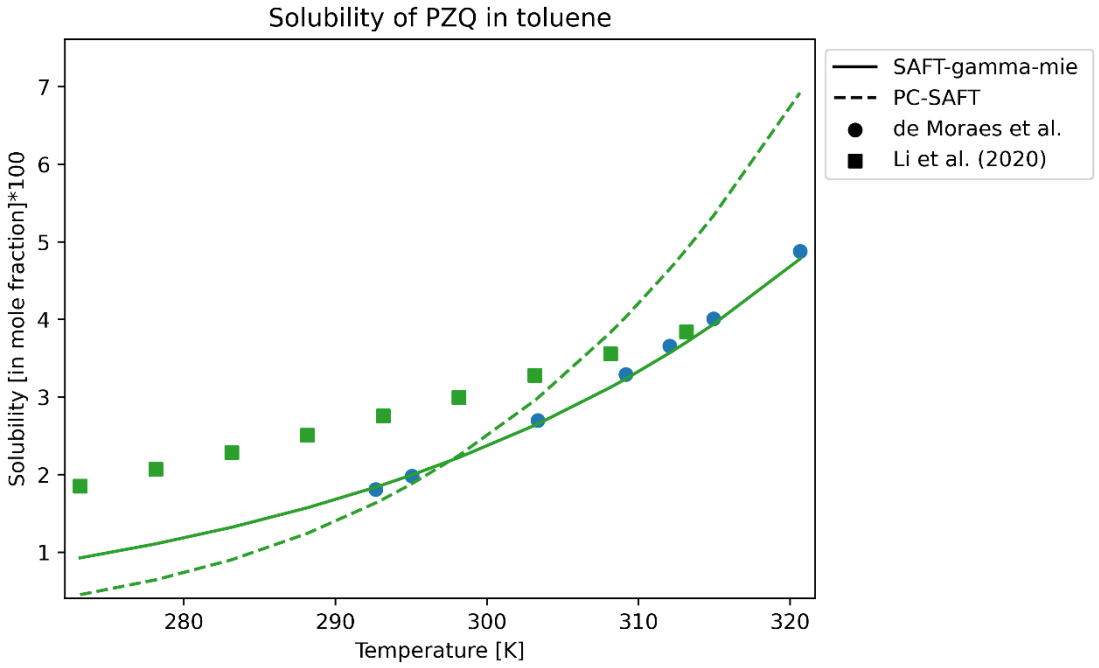

Figure 49: Solubility temperature curve of PZQ in toluene at ambient pressure (101.1 KPa). Experimental solubility temperatures identified by squares were determined and

reported by Li *et al.* (2020), while the circles were determined by de Moraes et al. (yet to be published). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

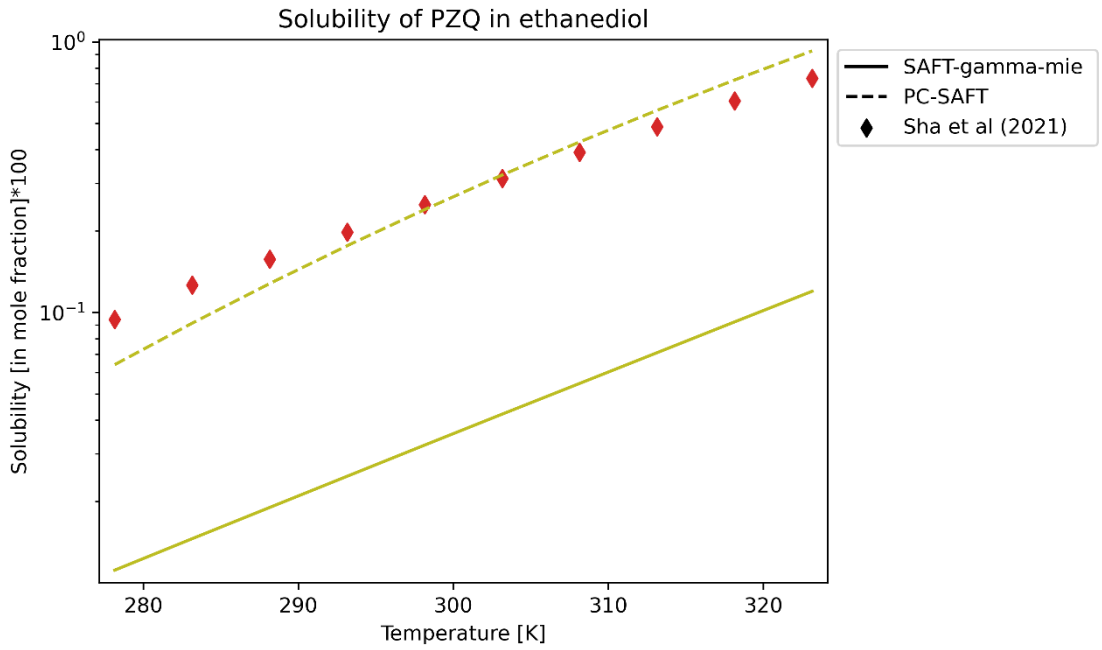

Figure 50: Solubility temperature curve of PZQ in ethane-1,2-diol (ethanediol) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by diamonds, were determined and reported by Sha *et al.*, (2021). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

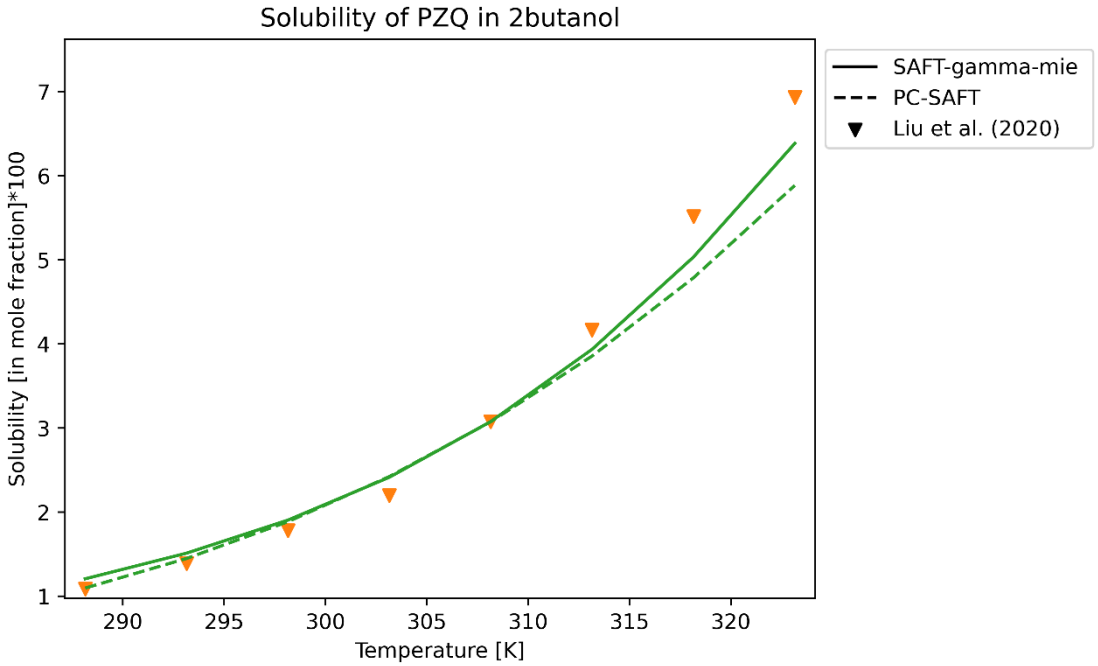

Figure 51: Solubility temperature curve of PZQ in 2-butanol at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by triangles, were determined and reported by Liu *et al.*, (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

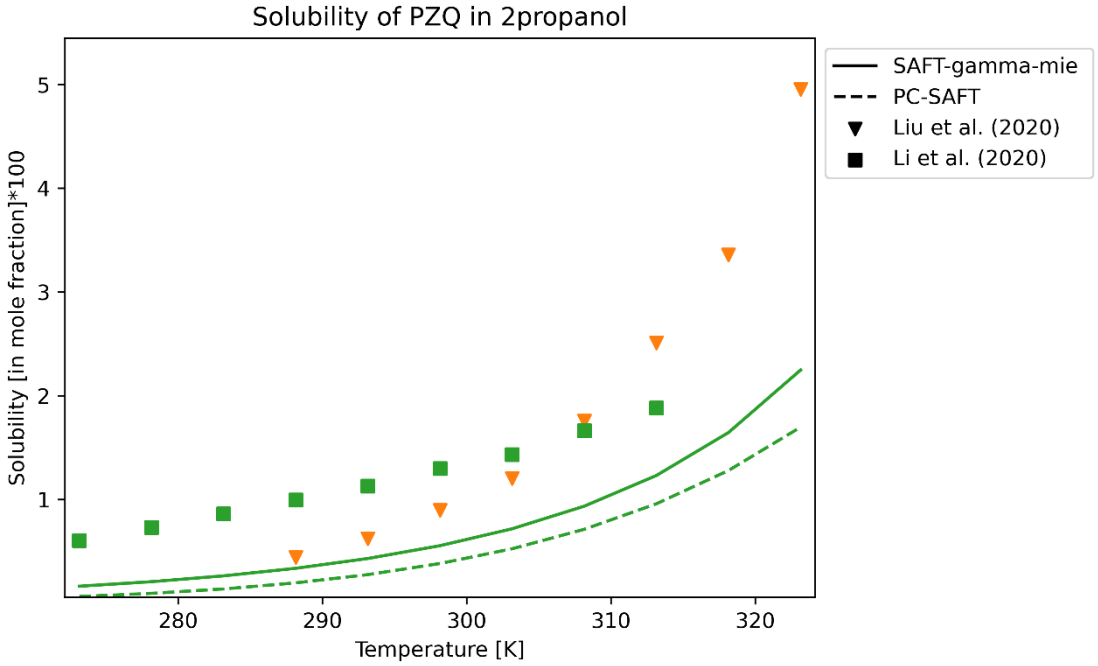

Figure 52: Solubility temperature curve of PZQ in 2-propanol at ambient pressure (101.1 KPa). Experimental solubility temperatures identified by squares were determined and

reported by Li *et al.* (2020), while the triangles were determined and reported by Liu *et al.*, (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

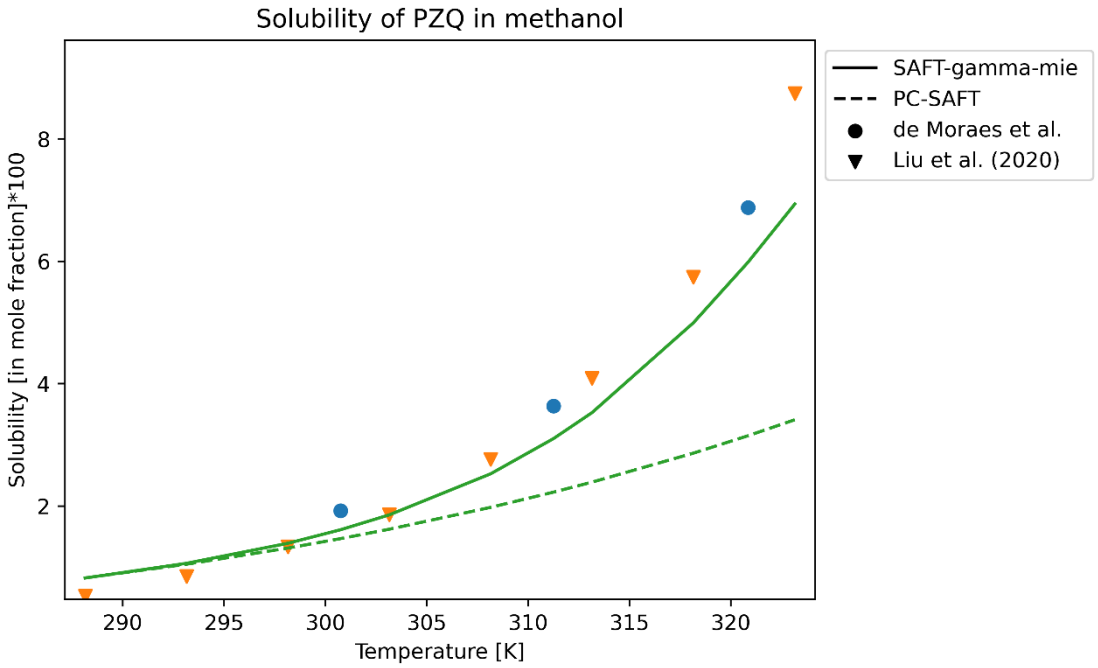

Figure 53: Solubility temperature curve of PZQ in methanol at ambient pressure (101.1 KPa). Experimental solubility temperatures identified triangles were determined and reported by Liu *et al.*, (2020), while the circles were determined by de Moraes et al. (yet to be published). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

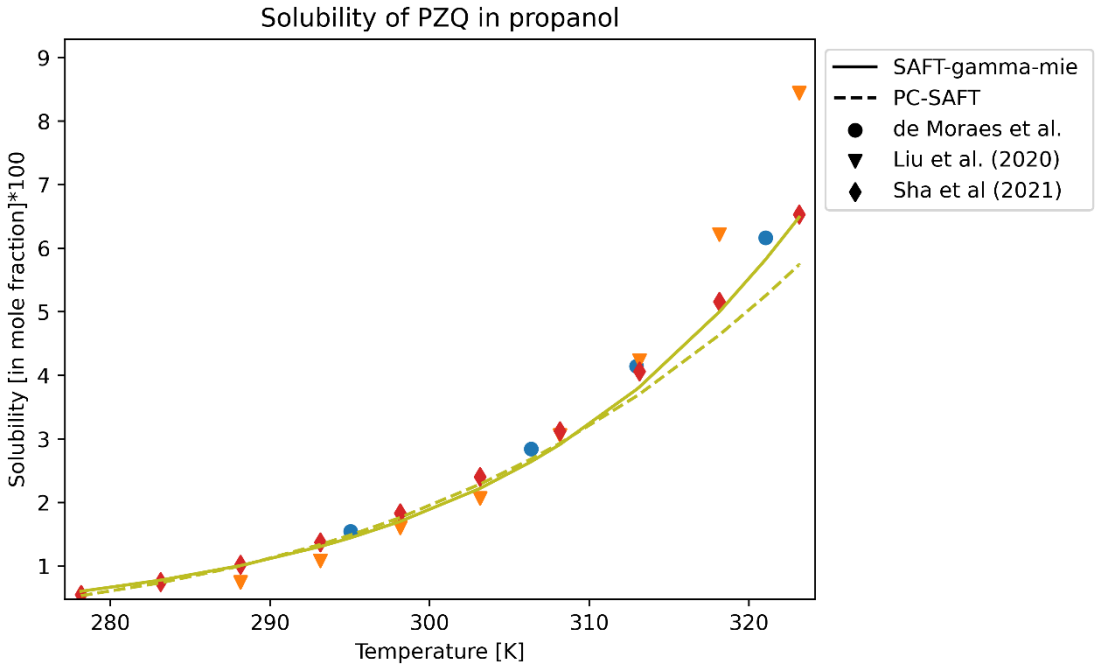

Figure 54: Solubility temperature curve of PZQ in propanol at ambient pressure (101.1 KPa). Experimental solubility temperatures are denoted by triangles from the study by Liu *et al.*, (2020), diamonds from Sha *et al.*, (2021), and circles from de Moraes et al. (yet to be published). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

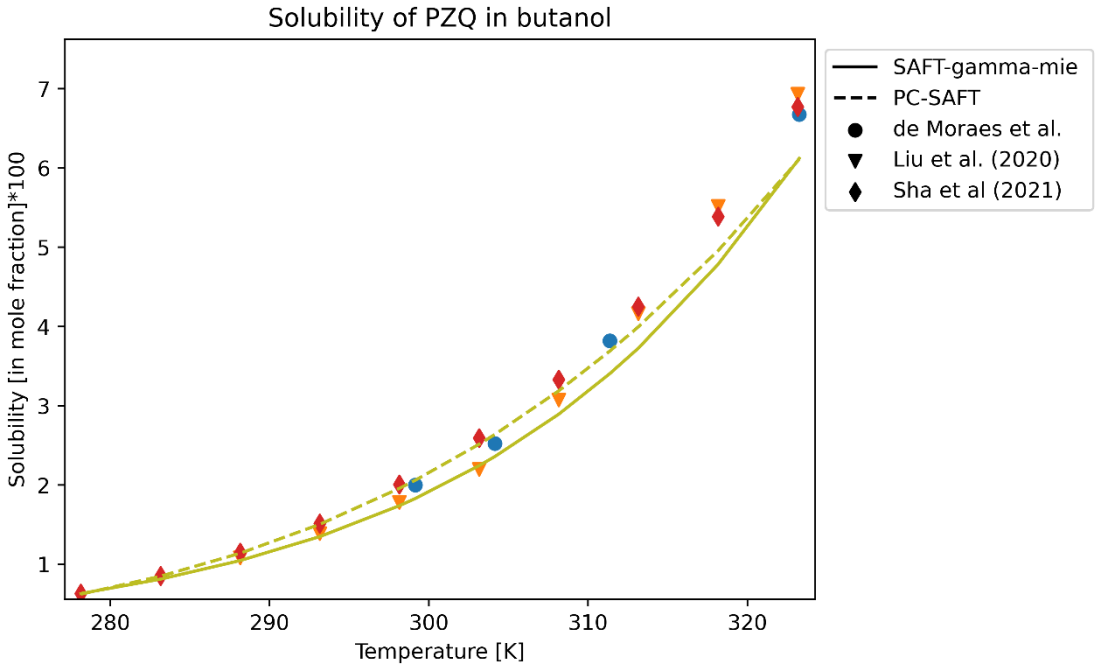

Figure 55: Solubility temperature curve of PZQ in butanol at ambient pressure (101.1 KPa). Experimental solubility temperatures are denoted by triangles from the study by Liu *et al.*, (2020), diamonds from Sha *et al.*, (2021), and circles from de Moraes et al. (yet to be published). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

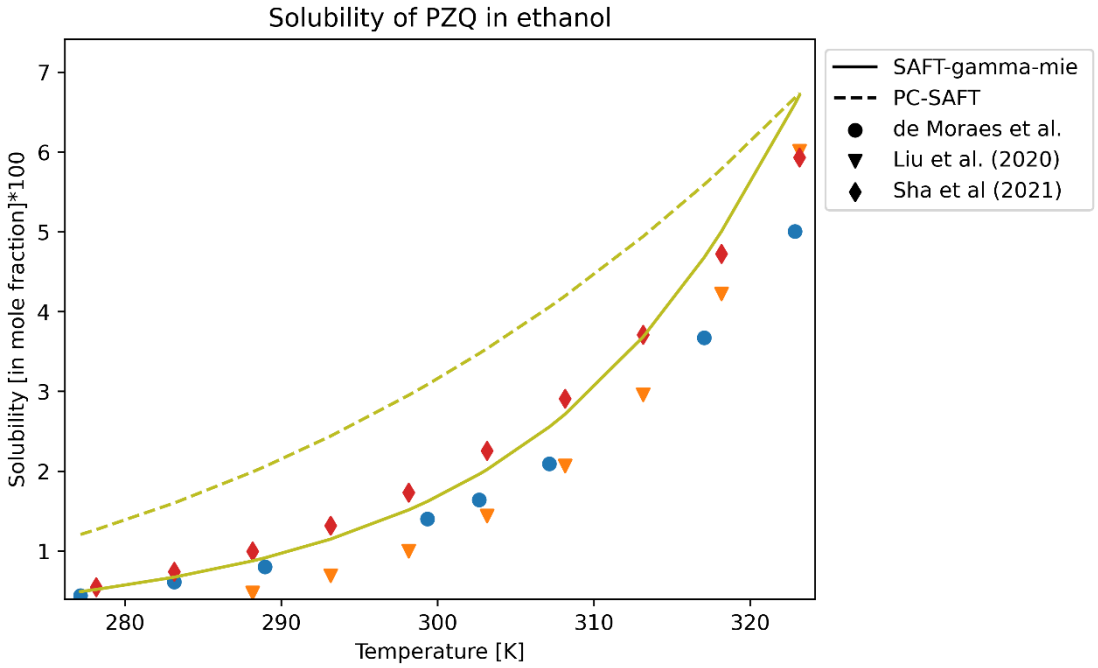

Figure 56: Solubility temperature curve of PZQ in ethanol at ambient pressure (101.1 KPa). Experimental solubility temperatures are denoted by triangles from the study by Liu *et al.*, (2020), diamonds from Sha *et al.*, (2021), and circles from de Moraes et al. (yet to be published). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

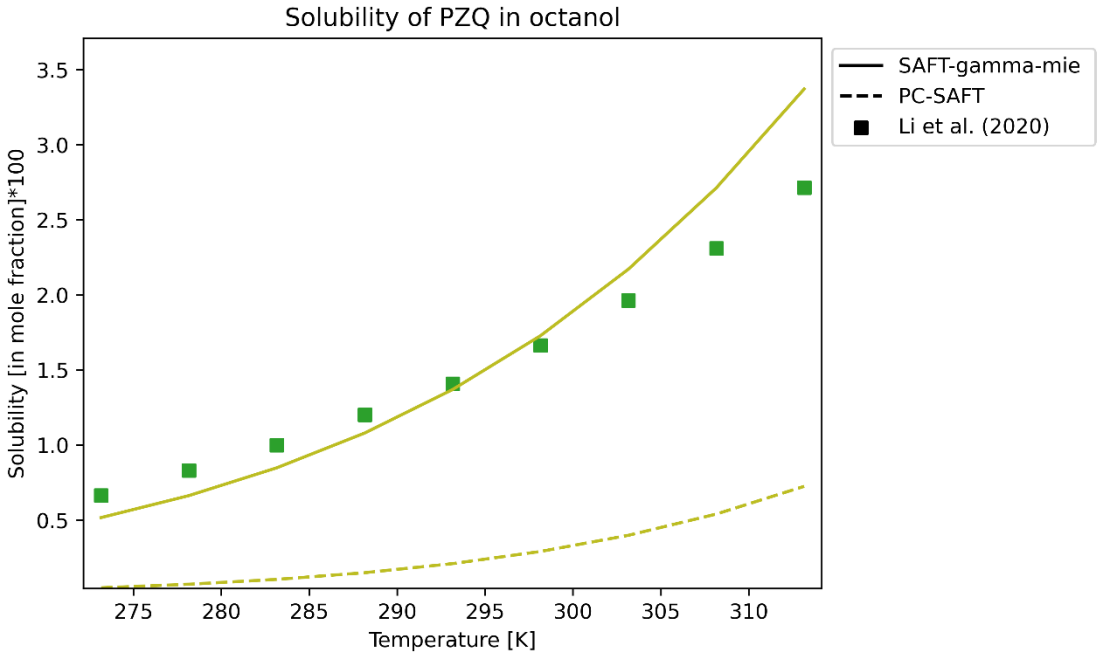

Figure 57: Solubility temperature curve of PZQ in octanol at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by squares, were determined and reported by Li *et al.* (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

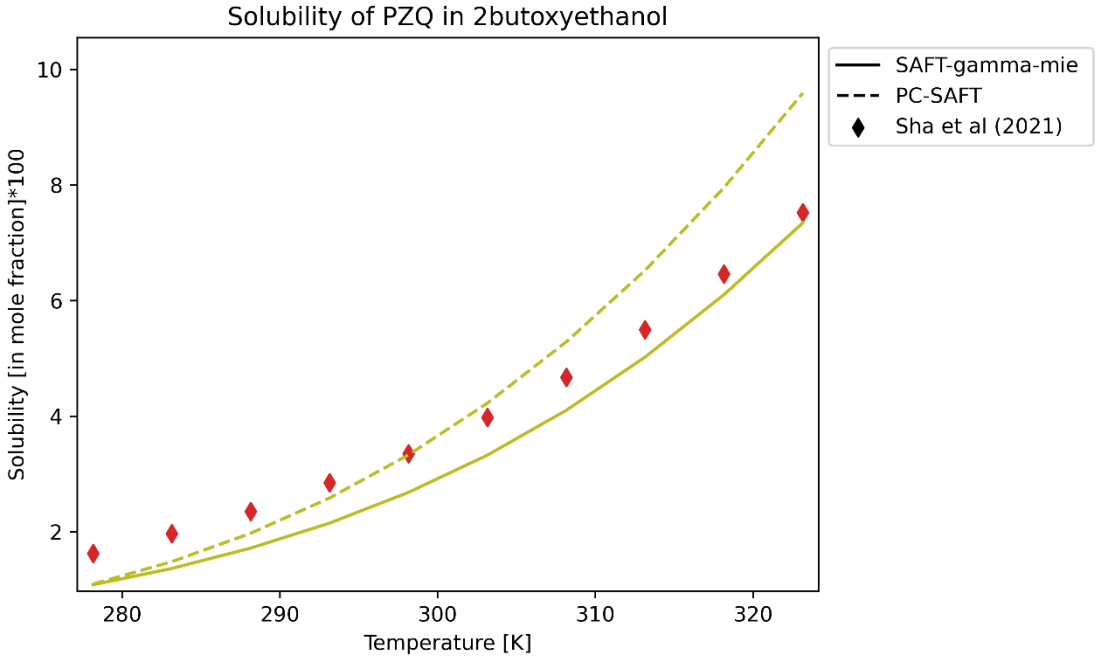

Figure 58: Solubility temperature curve of PZQ in 2-butoxyethanol (2butoxyethanol) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by

diamonds, were determined and reported by Sha *et al.*, (2021). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

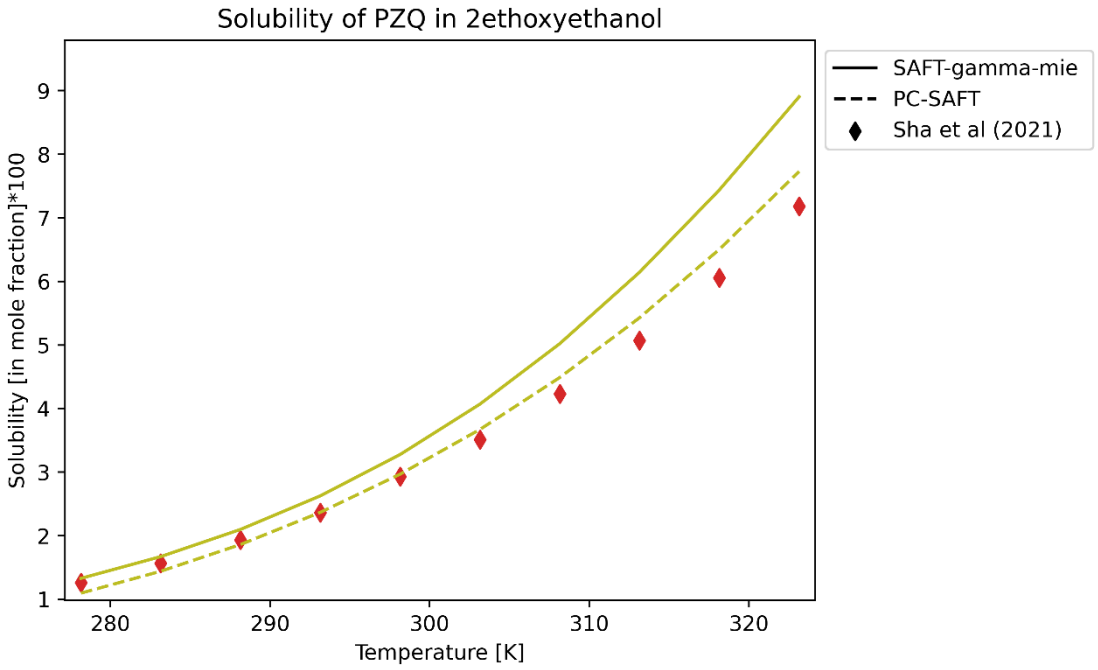

Figure 59: Solubility temperature curve of PZQ in 2-ethoxyethanol (2ethoxyethanol) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by diamonds, were determined and reported by Sha *et al.*, (2021). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

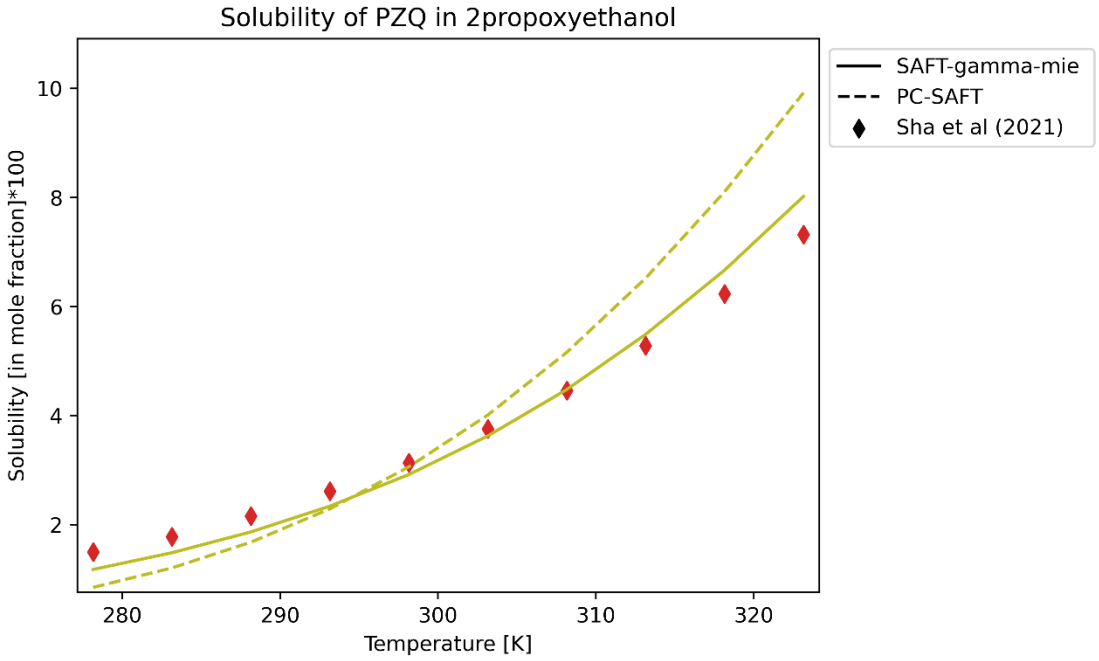

Figure 60: Solubility temperature curve of PZQ in 2-propoxyethanol (2propoxyethanol) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by diamonds, were determined and reported by Sha *et al.*, (2021). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

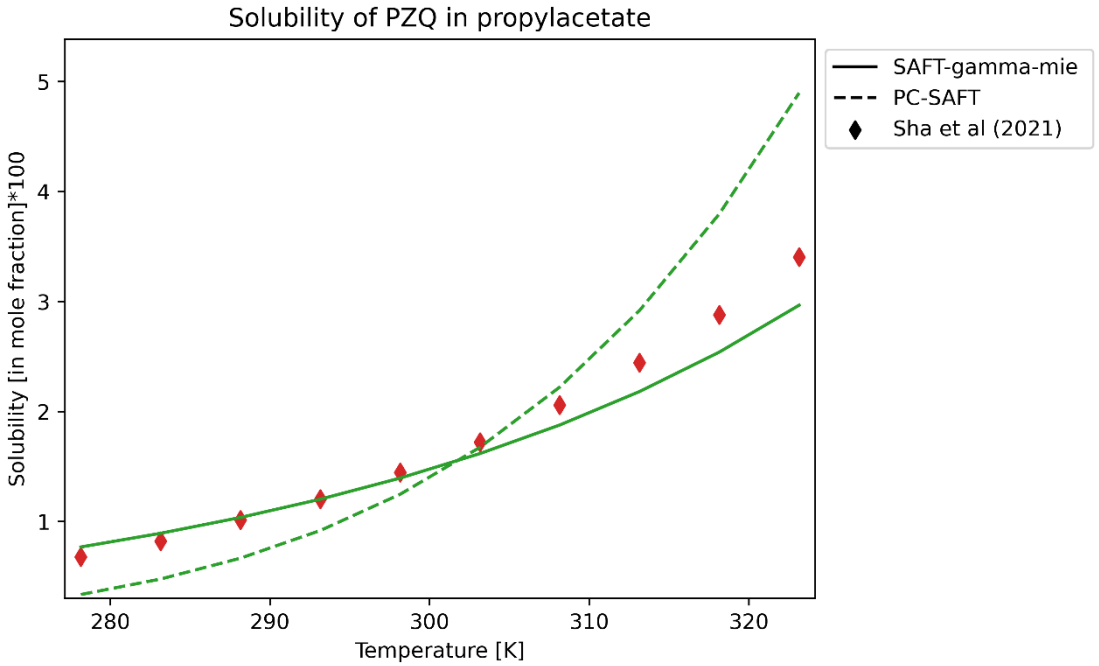

Figure 61: Solubility temperature curve of PZQ in propyl acetate (propylacetate) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by diamonds, were determined and reported by Sha *et al.*, (2021). The solid lines illustrate

solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

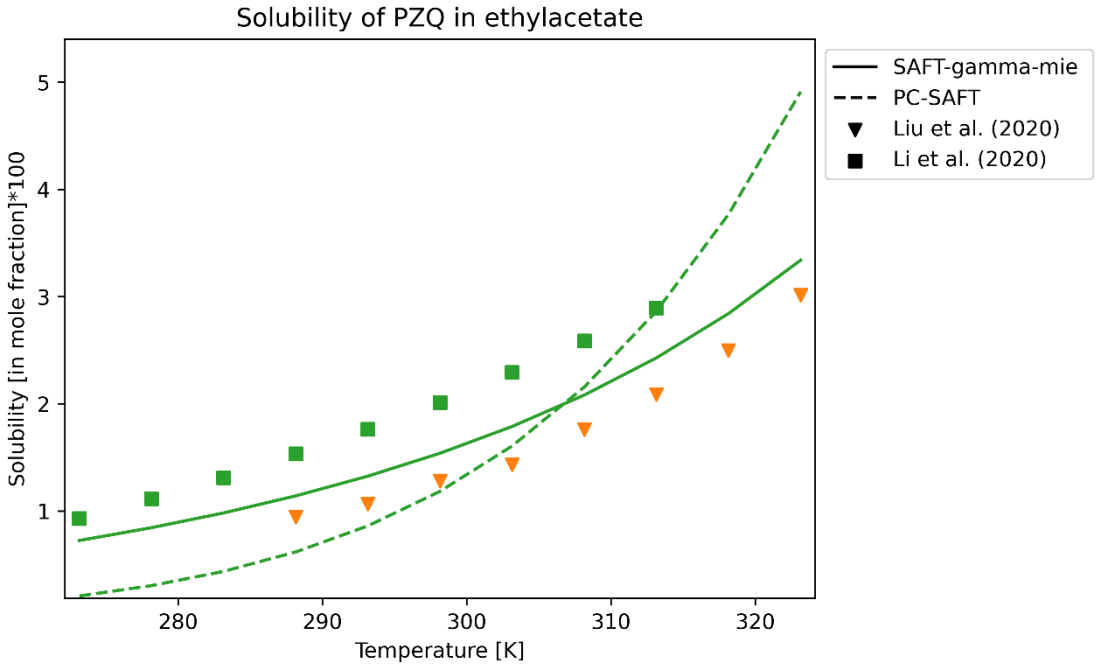

Figure 62: Solubility temperature curve of PZQ in ethyl acetate (ethylacetate) at ambient pressure (101.1 KPa). Experimental solubility temperatures identified by squares were determined and reported by Li *et al.* (2020), while the triangles were determined and reported by Liu *et al.*, (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

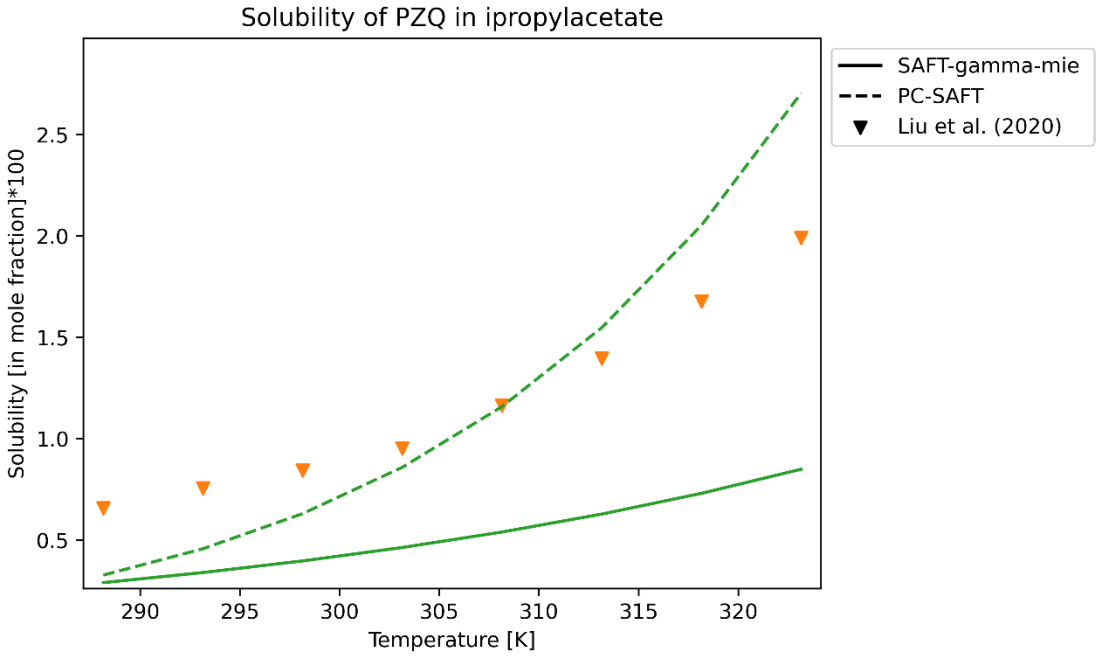

Figure 63: Solubility temperature curve of PZQ in isopropyl acetate (ipropylacetate) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by triangles, were determined and reported by Liu *et al.*, (2020). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

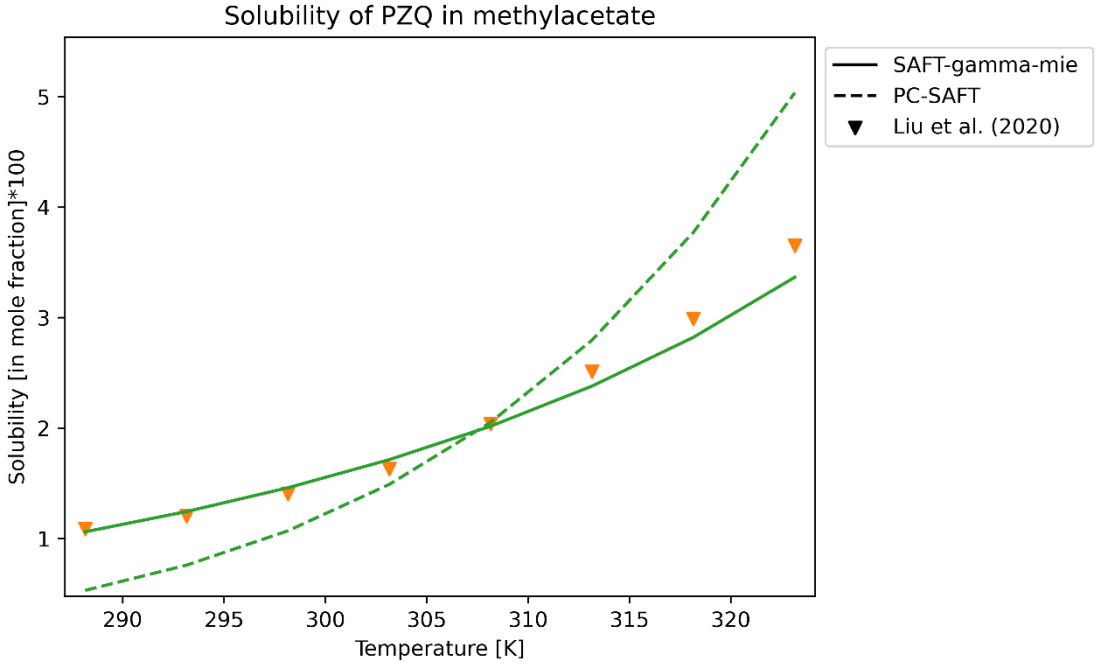

Figure 64: Solubility temperature curve of PZQ in methyl acetate (methylacetate) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by triangles, were determined and reported by Liu *et al.*, (2020). The solid lines illustrate

solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

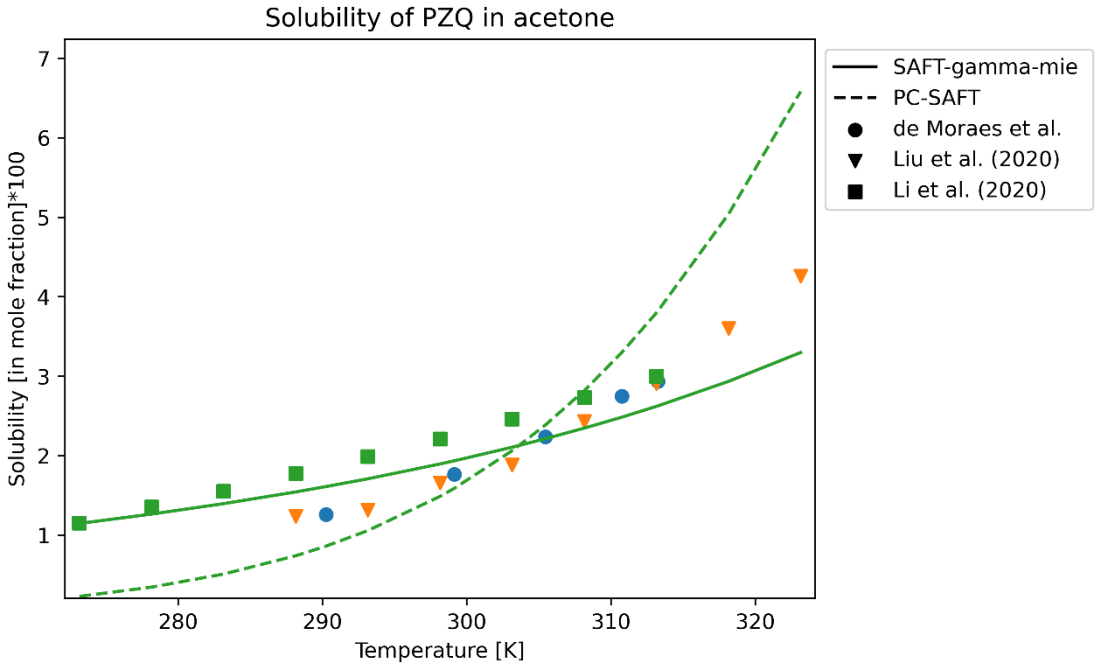

Figure 65: Solubility temperature curve of PZQ in acetone at ambient pressure (101.1 KPa). Experimental solubility temperatures are denoted by triangles from the study by Liu *et al.*, (2020), squares from Li *et al.* (2020), and circles from de Moraes et al. (yet to be published). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

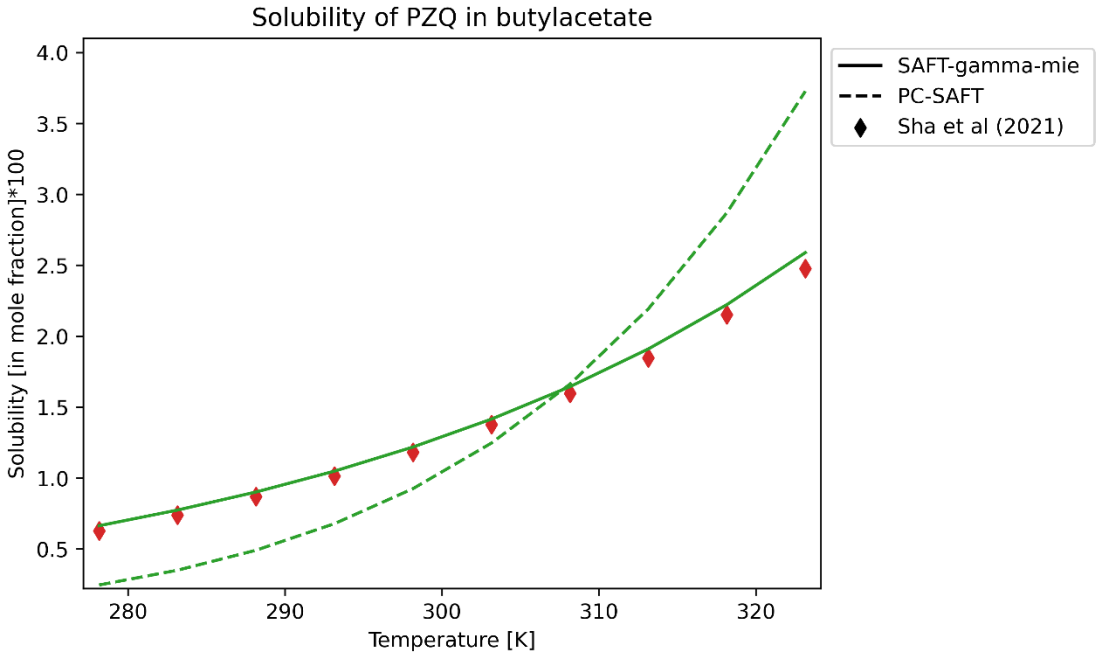

Figure 66: Solubility temperature curve of PZQ in butyl acetate (butylacetate) at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by diamonds, were determined and reported by Sha *et al.*, (2021). The solid lines illustrate solubility predictions made using the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.

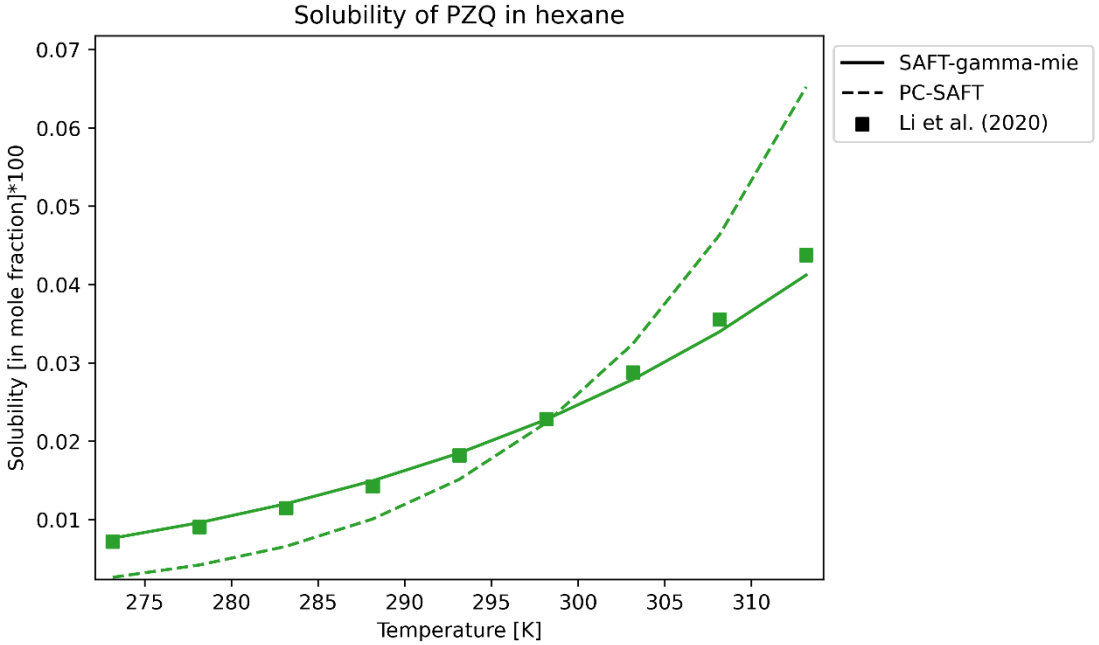

Figure 67: Solubility temperature curve of PZQ in hexane at ambient pressure (101.1 KPa). Experimental solubility temperatures, identified by squares, were determined and reported by Li *et al.* (2020). The solid lines illustrate solubility predictions made using

the SAFT-γ Mie EOS. In contrast, dashed lines show the solubility predictions as calculated by the PC-SAFT EOS.