Markus Hays Nielsen

The Effect of Binary Interaction Parameters on Phase Behavior

Master's thesis

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of
Science and Technology

Markus Hays Nielsen

The Effect of Binary Interaction Parameters on Phase Behavior

NTNU

Norwegian University of
Science and Technology

To my friends and family, their support and ideas have made this possible.

Summary

Accurate fluid modeling has been a goal of reservoir engineers since the 1940's. Based on the advances in computing power since the 1970's, cubic equation of state (EOS) models have been used to accurately describe a wide range of complex reservoir fluid systems, including gas condensates, volatile oils, compositional grading fluids with saturated and critical gas-oil phase transitions, and miscibility that is often controlled by near-critical vaporization and condensation.

Without an accurate fluid model, the engineering of petroleum reservoirs would bear significant and unnecessary uncertainties in the estimation of recoveries and production forecasts. This thesis is intended to study one specific method for improving the modeling of petroleum fluids - the use of binary interaction parameters (BIPs) in a cubic EOS.

Very little literature addresses the quantitative impact of BIPs on calculated phase behavior. This thesis therefore studies and quantifies the cause-and-effect of BIPs on phase behavior of binary, ternary and multi-component hydrocarbon fluid systems. Specifically, the phase behavior quantified includes (1) dependence of *equilibrium ratios* (K-values) on pressure, temperature and composition, and (2) critical phase boundaries that include *critical p-T loci* and the negative flash equivalent of critical pressure, *convergence pressure*. A seemingly general relationship was found between the sign of BIPs and the shift in critical locus, as well as the shift in low pressure K-values.

The second part of the thesis deals with tuning an EOS model using the BIPs as primary regression variables. Synthetic data sets were generated with a known set of BIPs, with K-values being the only data used in the tuning process. An EOS with zero BIPs was used as starting values and the non-linear regression algorithm of PhazeComp was used to determine a BIPs matrix that gave a best fit of the exact synthetic K-values. The regression algorithm of PhazeComp was successful in finding the correct set of BIPs used to create the synthetic data, but only if the number of components was less than about 15. Furthermore, it was established that the run-time of regression for components > 30 would be prohibitive, even if the regression algorithm succeeded.

This led to a search for algorithms that would select and optimize a subset of the entire BIPs matrix, yielding an adequate best-fit of the phase behavior data (i.e. K-values). Three methods for ranking the importance of the BIPs were tried: a gradient based method, a Pearson's correlation coefficient (PCC) based method and a random forest based method. Heuristic logic was introduced during the studies presented in this thesis, and successful results were achieved with the gradient and PCC based methods.

Preface

I was introduced to the topic of Binary Interaction Parameters some 3 years ago in the form of a regression problem for a similar topic. This lead me down a path to learning about multivariate regression and machine learning. At the beginning of the final year of my Master's degree I was indulged in the full problem. I was not only introduced to the topic of BIP tuning, but also equation of state modeling. Through this year I have attended an advanced industry course on PVT (pressure-volume-temperature) and several internal courses on the same topic through Whitson AS. It has been a privilege for me and for which I am truly grateful.

Throughout my thesis I have had to learn how to use some of the many features of Phaze-Comp, a PVT software by Zick Technologies. This has been an amazing experience and I am thankful for the incredible tool that Aaron Zick has created PhazeComp. This work could truly not have been achieved without the hard work and expertise he has put into making the software, and for that I am thankful.

I have also had the pleasure of working with the software team at Whitson AS and have been able to use modules of their work for my thesis. This has been a great learning experience and I can truly say that without their assistance I could not have produced the results here. I would especially like to thank Sebastian Roll and Arnaud Hoffmann for their assistance and inputs.

However, without a sound framework for my thesis I would never have thought of the ideas presented in this work, and for that I would like to thank my advisor Milan Stanko. Your insight into a vast number of related topics is greatly appreciated. I would also like to thank you for your patience and allowing me to follow my own path with the thesis work. This has been a great learning experience for me, and I look forward to working together in the future.

I would also like to thank the person who introduced me to the problem that resulted in this work. That person is Curtis Hays Whitson (dad). Your guidance and thoughtful discussions have been a pleasure that I cannot explain in words. When the topic was introduced, you stated that if someone was able to solve this task you would be happy to retire. I am therefore glad that I was not able to finish the task, but rather get a foot in the door. Throughout the process of my thesis work, you have guided me and contributed more than you know.

I would also like to thank several fellow students and colleagues. First, Stian Mydland has been essential in the development of the PhazeComp files that are able to consistently and automatically bound the search used to estimate critical points of fluid mixtures. He has also helped with insightful discussion and reminded me clearly when I was heading too

Table of Contents

List of Tables

List of Figures

Abbreviations

Greek Symbols

$\alpha(T)$	-	EOS parameter correction term
γ	-	Specific gravity
δ	-	Small perturbation in BIP
$\boldsymbol{\Delta}$	-	Matrix of small perturbation in BIPs
ϵ	-	Small threshold value
ϕ	-	Fugacity coefficient
λ	-	Counting variable or constant value
μ	-	Chemical potential
ρ_{ij}	-	Importance of BIP between component i and j
$\rho(\circ, \circ)$	-	Pearson correlation coefficient
$\sigma(\circ)$	-	Standard deviation
$\sigma(\circ, \circ)$	-	Covariance operation
χ	-	Error in cost function
θ	-	Generic parameter
ω	-	PVT observation data or acentric factor
Ω	-	Set of observed data
Ω_a	-	EOS parameter
Ω_a	-	EOS parameter

Latin Symbols

a	-	EOS parameter
A	-	EOS parameter or area of phase envelope
b	-	EOS parameter
B	-	EOS parameter
BWR	-	Benedict-Webb-Rubin
c	-	Volume shift
$C(\mathbf{K})$	-	Cost function for a given BIP matrix
d	-	Calculated PVT data
EOS	-	Equation of state
f	-	Fugacity of mixture
$f_M(PVT; EOS)$	-	BIP matrix
F_v	-	Molar vapor fraction
FI	-	Feature importance of random forest
g	-	Gibbs free energy
$h(\circ)$	-	Objective function / root function / random forest tree
$\hat{h}$	-	Trained random forest
i, j	-	Counting variables
k_{ij}	-	Binary interaction parameter (BIP)
K_i	-	Equilibrium ratio (K-value)
$\mathbf{K}$	-	BIP matrix
$\mathbf{K}^\circ$	-	Synthetic ("true" or reference) BIP matrix
$m(\omega)$	-	EOS parameter correction term
n	-	Amount of moles
p	-	Pressure
p_c	-	Critical pressure
p_k	-	Convergence pressure
p_v	-	Vapor pressure
$\hat{p}_v$	-	Apparent vapor pressure
PE	-	Phase envelope properties
PR	-	Peng-Robinson
PVT	-	Pressure-volume-temperature
r_i	-	Residual element in residual vector
SRK	-	Soave-Redlich-Kwong
v	-	Universal gas constant
T	-	Temperature
T_c	-	Critical temperature
u_i	-	Liquid, vapor or feed molar composition
v	-	Molar volume
V	-	Volume
x_i	-	Liquid molar composition
y_i	-	Vapor molar composition
z_i	-	Feed molar composition
Z	-	Compressiblity factor

Chapter 1

Introduction

1.1 Short History of Equations of State Modeling

Various models for describing petroleum fluid systems have been used historically. Early methods for describing simple fluid models, e.g. the real gas law, like the Standing-Katz [37] graphical description of the gas deviation factor, were developed in the 1940's. From the Standing-Katz chart, mathematical models were developed to describe similar behavior, using equations instead of charts. These models used Benedict-Webb-Rubin (BWR) equation of state (EOS), referring to the original work by Benedict, Webb and Rubin in 1940 [3]. The Hall-Yarborough [8] equation is one of the most applied BWR methods for real gases. Early attempts to develop a BWR model that was able to describe two-phase systems was presented by Starling in 1966 [38]. However, the early BWR EOS models used to describe phase behaviour was complex and therefore replaced by the cubic EOS models developed in the 1970's.

The introduction of the first cubic equation of state (EOS) models in the petroleum literature were developed by Redlich and Kwong (RK)[34] and by Peng and Robinson (PR)[29]. The general two-parameter cubic EOS model structure was not new, in fact the first two-parameter cubic EOS model was developed in 1873 by van der Waals [42] in his PhD dissertation. Even though the PR and RK were not the first cubic EOS models, the modifications that were made to the equations significantly improved the ability that they had to predict the phase behavior of petroleum systems. Modifications of both the PR and RK were made shortly after their initial development which further increased their ability to predict phase behavior of petroleum systems. Introductions of several other modifications, like the volume shifts by Peneloux [28] and BIPs by Chueh and Prausnitz [6], were essential to applying cubic EOS models in the petroleum industry. The introduction of the BIPs had a major impact on the ability of two parameter cubic EOS models (e.g. PR and SRK) to predict the behaviour of complex fluids, especially near critical fluids.

More complicated non-polynomial EOS models have also been developed like the Di-eterici EOS as well as higher order polynomial EOS models have been developed referred to as virial expansions or Kamerligh Onnes EOS models. There are also a range of exotic EOS models for relativistic fluids used in the modeling of stars and high speed gasses like the stiffened EOS or the Bose EOS. However, these more exotic EOS models are rarely, if ever, used in the petroleum industry. One branch of EOS modeling that is used in the petroleum industry that is outside the scope of the cubic EOS models are fluid-solids mod-els that predict the behavior of solids dropping out of the fluid.

1.2 Binary Interaction Parameters

BIPs were introduced as empirical correlation parameters to initial the mixing rule from van der Waals and were shown to have a significant impact on the prediction of cubic EOS models.

Traditional estimation of BIPs was based on tuning (regression) of the BIPs between spe-cific combinations of components determined by empirical data or previous knowledge. Katz and Firoozabadi [18] show that introducing a BIP between C_1 (lightest hydrocarbon component) and the C_{n+} (heaviest component) can have a significant impact on the EOS model's ability to predict complex fluid phase behavior (K-values). Using the C_1-C_{n+} BIP is a common approach. Other traditional methods for estimating the BIPs include em-pirical correlations like the one proposed by Chueh and Prausnitz [6] or the temperature dependent BIP correlations like [39]. However, these correlations are purely empirical methods.

More recent attempts to define a first-principles correlation have shown some success at correlating temperature dependent BIPs. Jaubert et al. published a group contribution method (GCM) based correlation for BIPs for normal alkanes in 2004 [13]. The semi-analytic correlation is derived from several thermodynamic first principles and is tuned to phase behavior data (making it semi-analytical). Following the initial publication in 2004, a series of additional papers were published, expanding the GCM to paraffins, naph-tanes , aromats, non-hydrocarbon and psuedoized components (i.e. lumped components) [16, 40, 41, 31, 32, 15, 14, 45]. These methods do not appear to have gained general usage.

Machine learning methods have been applied to estimate the BIPs for a range of fluid sys-tems [1]. Much like the initial regression methods in the early implementation of BIPs, modern approaches are highly empirical and act somewhat like a black-box.

A pragmatic approach of engineers who work with EOS modeling in the petroleum in-dustry was described by Younus et al. in 2020 [46]. The paper explains that even experts working with EOS tuning have only limited tools for tuning the BIPs. A common approach for estimating BIPs is either by applying the Chueh-Prausnitz correlation, tuning individ-ual BIPs based on experience or a combination of the two. The approach is combined with a set of physical consistency checks of the tuned EOS model.

1.3   Flash Calculations

The flash calculation is an essential part of any phase behavior calculation. Based on the long history of the flash calculation in the petroleum and chemical engineering industry, most details of how to best calculate the results of the flash have been thoroughly worked out. Michelsen, who is a key figure in the development of computational thermodynamics, summarizes the details of different procedures needed to calculate phase equilibria in his book [22]. The flash calculation is, simply put, a combination of a component material balance defined by the Rachford-Rice equation [33] together with the constraint of thermodynamic equilibrium. A summary of the flash calculation is given in the following chapter.

In 1986 Michelsen developed a methodology for solving the flash calculation using a simplified flash approach based on the assumption that the EOS model has no BIPs [21]. The simplified flash solves the flash calculation without the need for iteration, which allows for a direct relationship between the EOS and the equilibrium ratios (equilibrium ratios are defined in the following chapter). The aim of Michelsen's simplified flash was to speed up programs which need to calculate the flash calculations multiple times.

Following Michelsen's publication of the simplified flash, several tried to expand the method to allow for EOS models containing a limited number of BIPs. The first paper to expand on Michelsen's simplified flash calculation was Jensen in 1987 [17] who introduced a single BIP to the EOS. Another formal description on the methodology was described as a theorem by Hendriks in 1988 [11] followed by a long list of further publications [10, 7, 19, 12, 25].

A critique of the applicability of the simplified flash method by Haugen in 2012 [9], followed by a comprehensive summary was given by Michelsen in 2013 [20]. The general conclusion can be summarized by the fact that BIPs are important and that most of the benefit of Michelsen's initial approach disappear as the number of BIPs increases.

1.4   Reduction Optimization

The structure of the BIPs introduce a unique challenge when trying to do EOS tuning because of the quadratic increase in the number of BIPs with the number of components (see section 2.1 for more details). As an example, a 50-component fluid system will yield 1225 different BIPs which is much more than traditional regression techniques can effectively handle. Another field of study has similar issues, namely machine learning (ML) and artificial intelligence (AI).

One specific method that tries to deal with the vast number of parameters is called principle component analysis (PCA). PCA is one of several methods that tries to reduce the number of parameters in the parameters space to achieve a reduced optimization scheme. Other, more traditional optimization schemes also try to reduce the parameter space, like the singular valued decomposition (SVD) where so-called null-effect variables are neglected. Other methods that avoid the issue of a large number of parameters are methods like Pow-

ell's method that heuristically choose subset of the parameter space to search as a means to reduce the size of the parameter space.

1.5 Layout of This Work

The main objective of this work is to investigate the topic of BIPs and its connection to hydrocarbon fluid phase behavior. The two goals that this work tries to achieve this is by (1) an advanced study into the effect of BIPs on binary and ternary phase behavior and (2) describe a methodology that allows for modifications of the BIPs to enhance the predictive power of an EOS fluid model. The secondary goal of this work is to emphasise the potential problems and complications that arise when modifying the BIPs of an EOS and show how these potential complications become apparent in the phase behavior. Pressure dependent equilibrium ratio behavior is the main source of describing the "phase behavior".

The two main advantages from this work are as follows. First, increased knowledge of a topic with little previous in-depth publications is obviously good, especially as it is shown to have a major impact on the fluid modeling. The second part is that the proposed methodology gives an approach to consistently approach BIP tuning. The work presented here acts as a starting point for further investigation into the topic of BIP tuning for EOS modeling that has, to the authors knowledge, not yet been utilized to its full potential.

In chapter 2 a review of the basic theory relevant to this work is given. The topics covered in this chapter is two-fold. First, a section on the methods used for EOS modeling, i.e. the specifics of different EOS models, the flash calculation and basic phase behavior are presented. The second half details the generic principles of the numerical methods used in this work. The methods in this section include Newton's method, gradient descent method, Powell's method, Pearson's correlation coefficient and the random forest algorithm.

Chapter 3 presents the results of an extensive study on the effects of BIPs on binary and ternary phase behavior. The proposed methodology aims to describe generalized trends of the effect by introducing positive and negative BIPs with a range of magnitudes. Methods for describing the effects of BIPs on the phase behavior are also detailed.

Chapter 4 presents the *selective tuning* approach implemented to preform BIP tuning on two fluid systems. Three methods were used to in the proposed solution for the BIP tuning; a gradient based method, a PCC based method and a random forest based method. The three proposed solutions were also compared to a base case generated by simultaneous regression on all the BIPs.

Chapters 5 and 6 present a discussion and conclusion respectively followed by a description of further work in chapter 7. Finally, an Appendix is given containing 6 parts (A through F). Appendix A, B and C catalog additional data for the results given in chapters 3 and 4. Appendix D example calculations are detailed for more complex procedures used in this thesis followed by Appendix E which contains a link to all the relevant Python code

and PhazeComp files used in this thesis. Finally, Appendix F contains a study of Phaze-Comp's regression tool for BIP tuning and contains an interesting example of unexpected K-value behavior when positive and negative BIPs are introduced to a multi-component fluid systems.

Chapter 2

Basic Theory

"Thus, I conceived the ideas that there is no essential difference between the gaseous and the liquid state of matter... And so the idea of continuity occurred to me"

- van der Waals

In the following sections the relevant theory for this thesis will be described. First a review of the equation of state will be described in section 2.1 followed by a description of the flash algorithm in section 2.2. Finally, a description of relevant regression techniques will be described in section 2.3.

2.1 Equation of State (EOS)

2.1.1 What is an EOS?

In general, an *equation of state* is any functional relationship that describes how a fluid behaves with relationship to the conditions of the system in terms of p, V, T and molalr amounts. One of the simplest examples of an equation of state is the *ideal gas law*

$$pV = nRT \tag{2.1}$$

The ideal gas law gives a simple relationship between pressure (p), volume (V), temperature (T) and the molar amount (n) of a gas. The topic associated with describing such systems is consequently called PVT (pressure-volume-temperature). As the name implies, the ideal gas law is only valid for gases at a specific and limited range of pressures and temperatures. Namely, the range is defined for low pressures and ambient or high temperatures where the main assumptions of the model hold. For the purposes of this work it is more convenient to rewrite equation (2.1) as follows.

$$\frac{pv}{RT} = 1 \tag{2.2}$$

where v is the molar volume defined by V/n.

A more general definition of equation of state models was introduced to increase the range of validity and allow for descriptions of both vapor and liquid phases by introducing the Z-factor as an independent variable. The definition of the Z-factor is given by

$$Z \equiv \frac{pv}{RT} \tag{2.3}$$

where the ideal gas law becomes a special case of the more general description when $Z = 1$.

Early attempts of defining graphical correlations for the Z-factor of hydrocarbon gases was made by Standing and Katz [37] based on van der Waals law of corresponding states. Using experimental data, look-up charts were developed to estimate the value of the Z-factor as a function of reduced temperature ($T_r = T/T_c$) and reduced pressure ($p_r = p/p_c$). These types of EOS models were further developed in the following years by introducing equations instead of the charts by Standing and Katz, called BWR EOS models after the initial work by Benedict, Webb and Rubin [3]. One of the most common BWR models used equations for describing the real gas deviation factor is the Hall-Yarborough correlation [8]. The first attempt to describe a two-phase fluid system by a BWR model was proposed by Starling in 1966 [38]. However, the BWR EOS models were superseded by the *cubic equation of state* models that were able to more easily predict the phase behavior with less complexity than its BWR counterpart.

2.1.2 Cubic Equation of State

The first *cubic equation of state* (CEOS) was developed by van der Waals in 1873 as part of his Ph.D dissertation

$$p = \frac{RT}{v - b} - \frac{a}{v^2} \tag{2.4}$$

where a can be thought of as a molecular "attraction" parameter and b can be thought of as a molecular "repulsion" parameter. The reason for the name cubic equation of state becomes apparent if equation (2.4) is re-written with the Z-factor as the main variable[1].

---

[1]The more common description of why these EOS models are called *cubic* is because a similar cubic expression can be derived by exchanging the Z-factor with the molar volume. However, it is the authors opinion that the form given in equation (2.5) is more practical and natural because this is the variable that is solved in the flash calculation.

$$Z^3 - (B+1)Z^2 + A \cdot Z - A \cdot B = 0 \qquad (2.5)$$

In equation (2.5) the coefficients A and B are defined in equation (2.6) and Z is defined in equation (2.3). Given the formulation of the Z-factor from equation (2.3), the equation is also cubic in volume.

$$A = a \frac{p}{(RT)^2} \qquad (2.6a)$$

$$B = b \frac{p}{RT} \qquad (2.6b)$$

In general, any two-parameter CEOS can be written as

$$Z^3 + A_2 \cdot Z^2 + A_1 \cdot Z + A_0 = 0 \qquad (2.7)$$

where A_0, A_1 and A_2 are constants defined by the two parameters a and b. For a CEOS to be accurate it is also assumed that the system satisfies the condition of chemical equilibrium (i.e. the chemical potential μ_i for each component for both phases are equal) [43]

$$\mu_{Vi} = \mu_{Li} \qquad (2.8)$$

which can be shown to be satisfied for the equal-fugacity constraint, $f_{Vi} = f_{Li}$, where the chemical potential is defined by

$$\mu_i = RT \cdot \ln(f_i) + \lambda_i(T) \qquad (2.9)$$

and the fugacity coefficient (ϕ_i) is defined by

$$\ln(\phi_i) = \ln\left(\frac{f_i}{u_i p}\right) = \frac{1}{RT} \int_V^\infty \left(\frac{\partial p}{\partial n_i} - \frac{RT}{V}\right) dV - \ln(Z) \qquad (2.10)$$

From the fugacity and the composition, the Gibbs free energy can be defined by

$$g_L = \sum_{i=1}^N x_i \ln(f_{Li}) \qquad (2.11a)$$

$$g_V = \sum_{i=1}^{N} y_i \ln (f_{Vi}) \tag{2.11b}$$

and a mixture Gibbs free energy defined by

$$g_{mix} = F_V g_V + (1 - F_V) g_L \tag{2.12}$$

where F_V is the molar vapor fraction defined in section 2.2.

Several CEOS models were developed for modeling petroleum systems in the 1970's and 1980's. The most recognized among these are the Peng-Robinson (PR) [29] which was modified in 1978 [35] and Soaves modification (SRK) [36] to the Redlich-Kwong (RK) EOS [34]. The 1978 PR EOS and the SRK EOS will be discussed in detail in following sections, but before that the method used for determining the two parameters a and b for a given mixtures are described in section 2.1.3.

The PR and the SRK models are similar for the coefficients A, B, a and b

$$a = \Omega_a \frac{R^2 T_c^2}{p_c} \alpha(T) \tag{2.13a}$$

$$b = \Omega_b \frac{R T_c}{p_c} \alpha(T) \tag{2.13b}$$

$$\alpha(T) = [1 + m(\omega)(1 - \sqrt{\frac{T}{T_c}})]^2 \tag{2.13c}$$

where T_c and p_c are the critical properties of the pure components and ω is the acentirc factor and the equations for Ω_a, Ω_b, α and m are different for PR and SRK. From this point forward, any EOS will be assumed to be a CEOS unless it is specified differently.

2.1.3 Mixing Rules

van der Waals described a method for determining the two parameters of his EOS for a mixture of pure components by a set of mixing rules. These preliminary mixing rules are

$$a = \sum_{i=1}^{N} \sum_{i=1}^{N} u_i u_j \sqrt{a_i a_j} \tag{2.14}$$

$$b = \sum_{i=1}^{N} u_i b_i \tag{2.15}$$

where u_i can be the liquid phase composition (x_i), vapor phase composition (y_i) or total composition (z_i), and a_i & b_i are the pure component EOS parameters for component i.

A significant improvement to the prediction of the phase behavior was shown to occur by introducing an empirical correction factor between each binary pair in the mixing rule for the a parameter,

$$a = \sum_{i=1}^{N} \sum_{i=1}^{N} u_i u_j \sqrt{a_i a_j}(1 - k_{ij}) \tag{2.16}$$

where k_{ij} is the correction factor referred to as the *binary interaction parameter* (BIP) or sometimes referred to as the binary interaction coefficient (BIC). In this work, k_{ij} will be referred to as the binary interaction parameter, or BIP.

2.1.4 Peng Robinson EOS

One common type of two-parameter cubic EOS is the modified 1978 Peng-Robinson model (PR78) and is determined by how the pure component parameters a and b are calculated from Ω_a, Ω_b and $m(\omega)$. The method for calculating these properties are given by

$$\Omega_a = \frac{8(5X + 1)}{49 - 37X} = 0.45724... \tag{2.17a}$$

$$\Omega_b = \frac{X}{X + 3} = 0.07780... \tag{2.17b}$$

$$X = \frac{-1 + (6\sqrt{2} + 8)^{1/3} - (6\sqrt{2} - 8)^{1/3}}{3} \tag{2.17c}$$

$$m(\omega) = 0.37464 + 1.54226\omega - 0.26992\omega^2, \omega \leq 0.49 \tag{2.17d}$$

$$m(\omega) = 0.3796 + 1.485\omega - 0.1644\omega^2 + 0.01667\omega^3, \omega > 0.49 \tag{2.17e}$$

2.1.5 Soave Redlich Kwong EOS

Another common EOS used in the petroleum industry is Soave's modification to the Redlich-Kwong EOS (SRK) which has the same structure as the PR78, but with differing values for Ω_a, Ω_b and $m(\omega)$. These relationships are by

$$\Omega_a = \frac{1}{9(2^{1/3} - 1)} = 0.42748... \tag{2.18a}$$

$$\Omega_b = \frac{2^{1/3} - 1}{3} = 0.08664... \tag{2.18b}$$

$$m(\omega) = 0.480 + 1.574\omega - 0.176\omega^2 \tag{2.18c}$$

2.1.6 Volume Shift

Even though the more advanced EOS models were able to more accurately describe complex fluid systems, there were still some major drawbacks with the models. The primary problem with both the PR78 and SRK models was that they were not predicting the vapor pressure curve correctly while still being able to correctly predict the phase densities. Attempts were made in the late 1970's [26] to solve this problem by introducing volume translations to the molar volume, however the problem was not solved until 1989 by Peneloux et al. [28]. The concept of the volume shift was to take the molar volume calculated by the original EOS (v_L^{EOS}) and subtract it by some constant shift (c) to yield the new molar volume (v_L^{new}) as described by

$$v_L^{new} = v_L^{EOS} - c \tag{2.19}$$

The effect of the volume translation was simply to shift the entire pressure-volume curve along the molar volume axis, subsequently shifting the vapor pressure curve. The simplicity and effectiveness of the method had a significant impact on EOS modeling. A dimensionless variation of the volume shift (s) was also introduced as

$$s = c/b \tag{2.20}$$

2.1.7 What Does an EOS Look Like?

In the sections above, the formal definition of the most common equations of states for petroleum systems are given, but what does an equation of state actually look like as a product? Typically, an EOS is used in some PVT software, process simulator, pipe-flow simulator or a compositional reservoir simulator. In this case the thing that is referred to as the EOS is a set of two lists of properties. First, the component properties that are needed for any internal calculations of PVT calculations (EOS and other correlations like the LBC viscosity correlation) have to be listed. A simple example of some component properties is given in Table 2.1 containing molecular weights (MW), critical pressures (p_c), critical temperatures (T_c), acentric factors (ω) and volume shifts (s).

Table 2.1: An example of component properties molecular weight, critical temperature, critical pressure, acentric factor and dimensionless volume shift.

Component Properties					
Component Name	MW	T_c (°F)	p_c (psia)	ω	s
C1	16.04	-116.66	667.0	0.011	-0.150
C2	30.07	89.91	706.6	0.099	-0.063
C3	44.10	206.02	616.1	0.152	-0.064
i-C4	58.12	274.5	527.9	0.186	-0.062
n-C4	58.12	305.5	550.6	0.200	-0.054
i-C5	72.15	369.0	490.4	0.229	-0.056
n-C5	72.15	385.8	488.8	0.252	-0.029
C6	82.42	464.4	490.0	0.240	-0.003
C7	95.63	529.0	455.3	0.274	0.013
C8	109.2	585.2	420.4	0.312	0.028

The component properties shown above are only a few of the different types that can be given in a typical PVT software, but are the required properties.

The second list of properties are the BIPs. Because each binary pair has a specific value for the BIP, the number of BIPs for a large number of components will be significantly higher than the number of component properties. The BIP matrix, as it is sometimes referred to, has certain properties that must be satisfied. The diagonal elements of the matrix are all equal to zero ($k_{ii} = 0$) because it is assumed there is no interaction between a component and itself. The second criteria is that the matrix is symmetric (i.e. $k_{ij} = k_{ji}$). An example of a BIP matrix is given in Table 2.2 for the same system as Table 2.1.

Table 2.2: An example of a BIP matrix using the Chueh-Prausnitz correlation.

Binary Interaction Parameters (BIPs)									
	C1	C2	C3	i-C4	n-C4	i-C5	n-C5	C6	C7	C8
C1	0	0.0021	0.007	0.013	0.012	0.018	0.018	0.021	0.025	0.030
C2	0.002	0	0.001	0.005	0.004	0.008	0.008	0.010	0.013	0.016
C3	0.007	0.001	0	0.001	0.001	0.003	0.003	0.004	0.006	0.008
i-C4	0.013	0.005	0.001	0	0.000	0.000	0.000	0.001	0.002	0.004
n-C4	0.012	0.004	0.001	0.000	0	0.001	0.001	0.001	0.002	0.004
i-C5	0.018	0.008	0.003	0.000	0.001	0	0.000	0.000	0.001	0.002
n-C5	0.018	0.008	0.003	0.000	0.001	0.000	0	0.000	0.001	0.002
C6	0.021	0.010	0.004	0.001	0.001	0.000	0.000	0	0.000	0.001
C7	0.025	0.013	0.006	0.002	0.002	0.001	0.001	0.000	0	0.000
C8	0.030	0.016	0.008	0.004	0.004	0.002	0.002	0.001	0.000	0

Tables 2.1 and 2.2 constitute an EOS model along with the equations given in the previous sections.

2.1.8 EOS Tuning

A simple explanation of EOS tuning, is that some of the properties in the two EOS tables are changed to match measured PVT data. Mostly the C_{n+} properties and BIPs including C_{n+} are modified in the EOS modeling. However, this explanation is somewhat oversimplified as there are a whole list of internal consistency checks that must be held because of physical restrictions. One paper that goes into the details of the entire process of tuning an EOS for a basin or field was published by Younus et al. [46]. To summarize, the objective of EOS tuning is to minimize some objective or cost function (as described in equation (2.21)) by changing some sub-set of the component properties in the EOS.

$$C(\mathbf{d};\theta) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\frac{d_i(\theta) - d_{exp,i}}{d_{ref,i}})^2} \tag{2.21}$$

In equation (2.21) $d_i(\theta)$ is the value calculated by the EOS, θ are the properties of the EOS, d_{exp} are the experimental values from PVT experiments, d_{ref} is some reference value dependent on the type of objective function and N is the total number of datapoints in all experiments.

This explanation is also somewhat oversimplified because it does not cover in detail the complexities of the EOS tuning process (e.g. does not describe the process of fitting a gamma model to the fluid nor describe the tuning of the viscosity model). However, for this work the full description of EOS tuning is not necessary and is therefore left out. For a more detailed review, the reader is advised to read the paper by Younus et al.

2.2 Flash Calculation

Let n be the total amount of moles in an overall composition (z_i) then it is possible to derive equations (2.22) and (2.23) from the component material balance.

$$z_i = F_V y_i + (1 - F_V)x_i \tag{2.22}$$

and

$$0 = \sum_{i=1}^{N}(y_i - x_i) \tag{2.23}$$

where F_V is the vapor fraction defined by the ratio of the molar amount in the vapor phase with the total molar amount in the system ($F_V = n_V/n$).

The material balance equations given in equations (2.22) and (2.23) yield an objective function ($h(F_V)$) that, when solved for the vapor fraction (F_V), yields a material balance. This method was first developed by Muskat and McDowell in 1949 [24], but the more well known equation was rediscovered by Rachford and Rice in 1952 [33] given by

$$h(F_V) = \sum_{i=1}^{N} \frac{z_i(K_i - 1)}{1 + F_V(K_i - 1)} \qquad (2.24)$$

where K_i is the equilibrium ratio (K-value) defined by

$$K_i = \frac{y_i}{x_i} \qquad (2.25)$$

The minimum and maximum values of the vapor fraction (F_V) that yield positive compositions are defined by

$$F_{v,min} = \frac{1}{1 - K_{max}} \qquad (2.26a)$$

$$F_{v,max} = \frac{1}{1 - K_{min}} \qquad (2.26b)$$

With the equations above it is possible to calculate the composition of each phase (x_i and y_i) from

$$x_i = \frac{z_i}{F_V(K_i - 1) + 1} \qquad (2.27a)$$

$$y_i = \frac{z_i K_i}{F_V(K_i - 1) + 1} = x_i K_i \qquad (2.27b)$$

As mentioned in the previous section 2.1, for the EOS to be applied correctly, the system has to be in chemical equilibrium. The criterion for equilibrium is given in equation (2.8) and can be shown to be equivalent to equal fugacities for each component of both phases. In equation (2.10) the expression for the fugacity is given, and is a function of the Z-factor (Z) and the phase composition. This means that an outer iteration loop must be solved for the K-values of the system. A procedure for solving the entire flash calculation can be found in the SPE monograph *Phase Behavior* by Whitson and Brulé [43] and is summarized below.

1. Estimate K-value

2. Calculate K_{min} and K_{max}

3. Solve for the vapor fraction (F_V) in the Rachford-Rice equation (2.24) within the range of $F_{V,min}$ and $F_{V,max}$.

4. Calculate the compositions of both phases (x_i and y_i) using equation (2.27).

5. Calculate the Z-factor from the EOS for each phase.

6. Calculate the component fugacities for each phase (f_{Li} and f_{Vi}).

7. Calculate the Gibbs free energy for each phase (g_{Li} and g_{Vi}) from equation (2.11) and calculate the *correct* root for the Z-factor if multiple roots exists, then calculate the mixture Gibbs free energy from equation (2.12).

8. Check the equal-fugacity constraint ($\sum_{i=1}^{N}(\frac{f_{Li}}{f_{Vi}} - 1)^2 < \epsilon$) where ϵ is some threshold value.

9. (a) If convergence is reached, then stop. (b) If convergence is not reached, update the K-value by some method.

10. Check if the converged solution is a trivial solution ($\sum_{i=1}^{N}(\ln K_i)^2 < \epsilon$).

11. If the solution is trivial, return to step 2. Otherwise, optionally confirm the trivial solution with a stability test[2].

2.2.1 Equilibrium Ratios

The equilibrium ratio (K-value), defined in equation (2.25), is one of the most important thermodynamic quantities for petroleum systems. The reason for this is the direct connection it has on the flash calculation, which is a key phase behavior calculation. Furthermore, the K-values have an intuitive and simple physical meaning. Given a K-value for a specific component at certain conditions, the magnitude will determine the affinity the component has to be in the vapor phase (K-value greater than 1) or the liquid phase (K-value smaller than 1). Theoretically, the K-values can be measured in the lab and an argument can be made that accurate K-value measurements is the best source of PVT data used in EOS modeling.

The shape of the K-value is typically divided into the low-pressure region and the high-pressure region. For the low pressure region the K-values tend to be inversely proportional to the pressure ($K_i \propto 1/p$) yielding a -1 slope on a log-log plot and is said to follow Raoult and Dalton's Law

---

[2]The stability test is an algorithm similar to a flash calculation with the purpose of establishing whether an equilibrium at a pressure and temperature is single phase or multi-phase, and if multi-phase it gives a set of K-value estimates
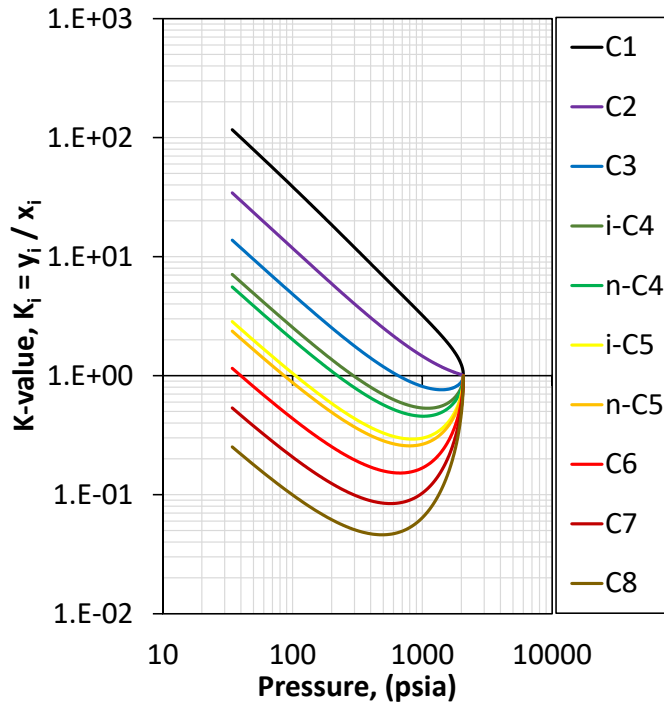
Figure 2.1: An example of pressure dependent K-values at a given temperature and composition.

$$K_i \approx \frac{p_{vi}(T)}{p} \tag{2.28}$$

where $p_{vi}(T)$ is the vapor pressure of component i. In the high-pressure region, all the K-values appear to converge to unity at the *convergence pressure* (described in detail in the following section). An example of a K-value plot for is shown in Figure 2.1.

2.2.2 Phase Envelope

The definition of a phase envelope can be described by; a closed curve in a pressure-volume-temperature-composition (PVT-z) space defined by the saturation pressures. For this work the pressure-temperature (PT) phase envelope will be used, assuming constant composition. The saturation pressures can either be *bubble-points* (typically marked by a green points) or *dew-points* (typically marked by a red points). An example of the phase envelope with some additional properties that will be described in the following sections is given in Figure 2.2. This figure was calculated using PhazeComp for uniform mixture with the EOS shown in Table 2.1 and 2.2.
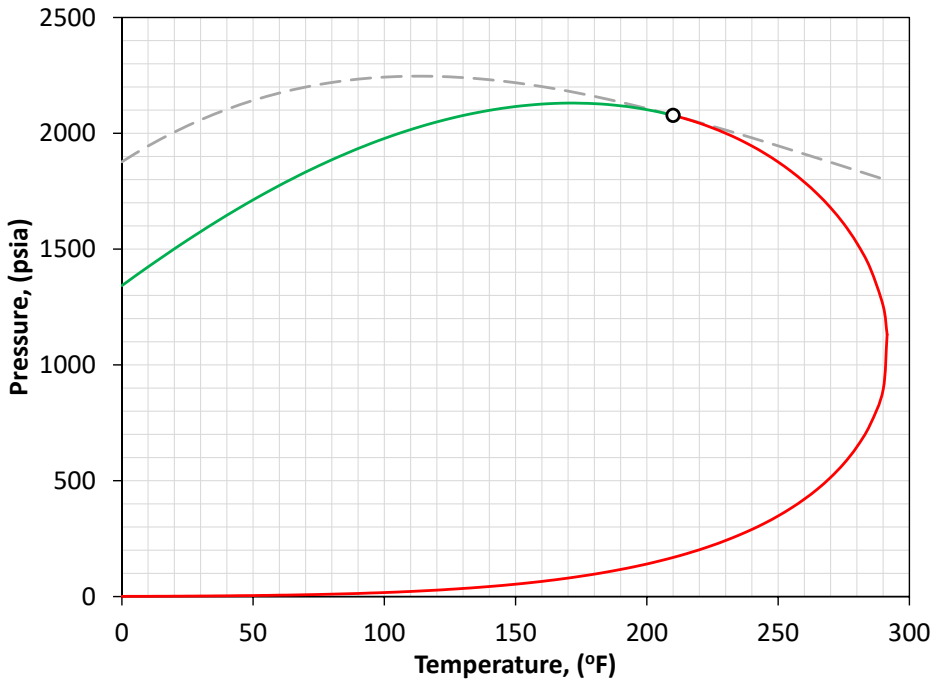
Figure 2.2: An example of a PT phase envelope (bubble point in green and dew point in red) for a uniform mixture with the EOS given in Table 2.1 and 2.2 as well as the convergence pressure as a function of temperature (grey dashed line).

2.2.3 Positive Flash

Conventionally, the accepted region of feasible vapor fraction values (F_V) was assumed to be $0 \leq F_V \leq 1$. At first glance, this seems like an intuitive assumption as any vapor fraction outside this range would return a negative molar amount of either the vapor phase ($F_V < 0$) or the liquid phase ($F_V > 1$). The region where the vapor fraction is within this range will be referred to as the *positive flash* region and can be associated with the two-phase region of the specific mixture composition. An example of the positive flash region is shown in Figure 2.3 for a temperature of 100 °F for a uniform mixture with the EOS given in Table 2.1 and 2.2. A special case of the positive flash where the vapor fraction (F_V) can be any value in the positive flash range is the *critical point*. The critical point is the point where the bubble-point and dew-point lines meet and all the component properties tend to the same value.
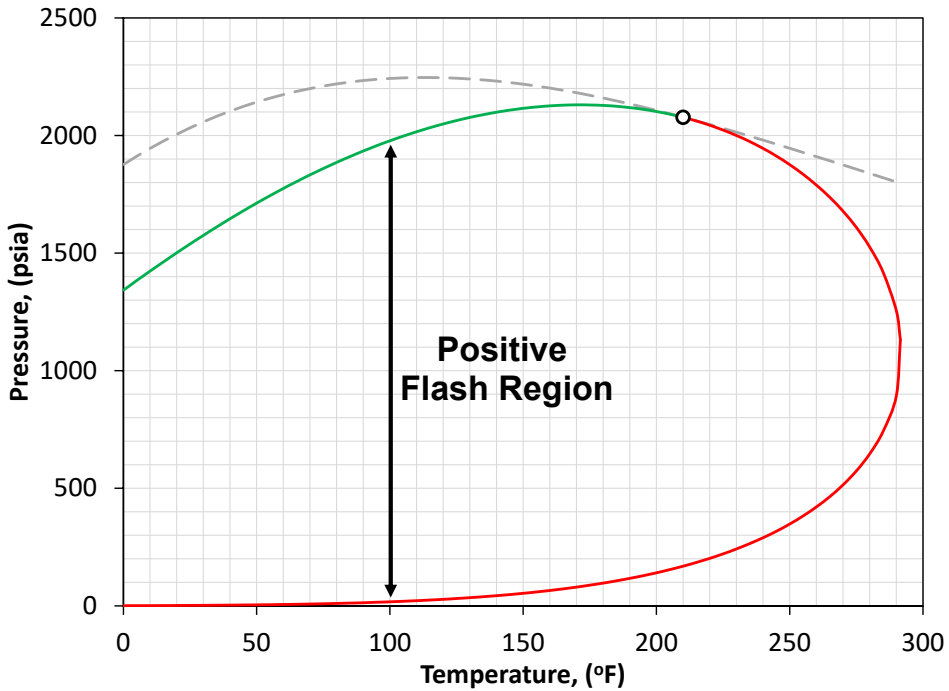
Figure 2.3: An example of a PT phase envelope (bubble point in green and dew point in red) for a uniform mixture with the EOS given in Table 2.1 and 2.2 as well as the convergence pressure as a function of temperature (grey dashed line) indicating the positive flash region for a temperature of $100°$F.

2.2.4 Negative Flash

In 1989 Whitson and Michelsen[3] published a paper describing the region of the flash calculation where the vapor fraction (F_V) was outside the range between 0 and 1, but within the calculated minimum and maximum vapor fractions from equation (2.26) [44]. This region of the flash calculation was defined as the *negative flash*. It was shown that the region gave thermodynamically consistent results. Furthermore, applications of the negative flash have resulted in practical methods like the fluid sampling methods proposed by Carlsen et al. [5]. An example of the negative flash region is shown in Figure 2.4 for a temperature of 100 °F for a uniform mixture with the EOS given in Table 2.1 and 2.2.

A special case of the negative flash is the transition from trivial solutions ($K_i = 1$) to the negative flash region. This pressure is called the *convergence pressure* and traces a line for different temperatures where the K-values tend to 1, while still yielding a negative flash solution. The convergence pressure line must go through the critical point of the given

---

[3]The authors note in the acknowledgements that: "The first author thanks Dr. Aaron Zick (ARC0 Oil and Gas Company) for introducing the idea of the "negative flash" in November 1985, and for many interesting discussions about the topic since then."

mixture and can considered a special case of the convergence pressure line where the flash calculation yields a positive flash solution. The convergence pressure line is shown as a grey dashed line in Figures 2.2 to 2.4.
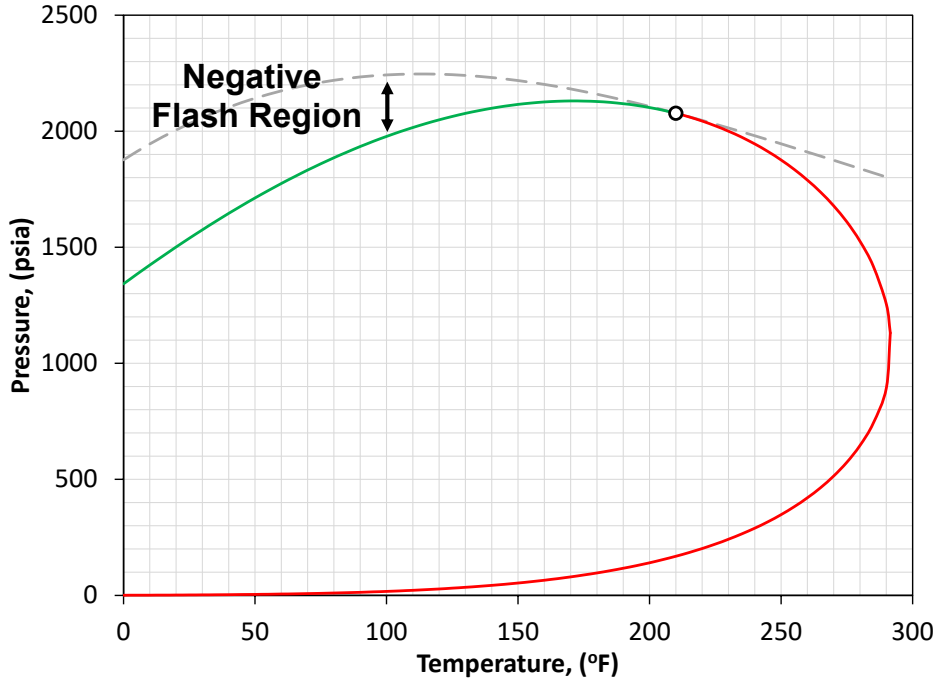
Figure 2.4: An example of a PT phase envelope (bubble point in green and dew point in red) for a uniform mixture with the EOS given in Table 2.1 and 2.2 as well as the convergence pressure as a function of temperature (grey dashed line) indicating the negative flash region for a temperature of $100°$F.

2.2.5 Ternary Diagram & Tie-Lines

For three component systems (ternary systems), visualizing the compositional space is made using a ternary diagram. The ternary diagram displays the quantity of each compositions along the three sides of an equilateral triangle. This type of plot is particularly useful for plotting the results of the flash calculation. An equivalent to the phase envelope, as discussed in section 2.2.2, can be drawn for a range of compositions z_i. The resulting liquid (x_i) and vapor (y_i) compositions for the two-phase flash can now be plotted on the ternary diagram with either a green color (representing liquid phase) or a red color (representing vapor phase). An example of this type of plot is shown in Figure 2.5.
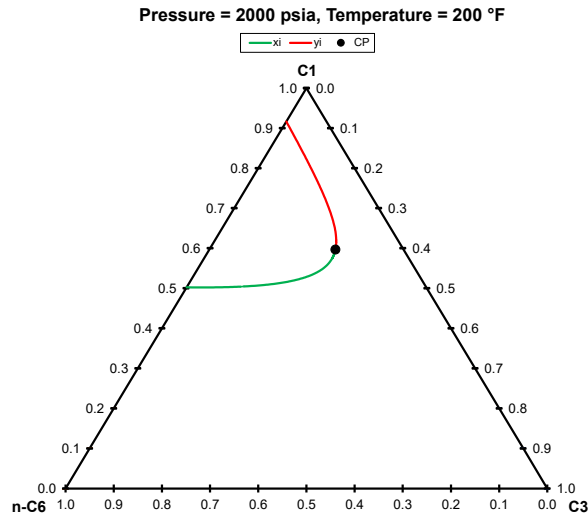
Figure 2.5: An example of the resulting phase compositions after a flash for a ternary system consisting of C_1, C_3 and n-C_6 at a temperature of $200°$F and pressure of 2000 psia using a PR EOS. The dew-point line is shown in red and the bubble-point line is shown in green with the critical point is shown as a black circle.
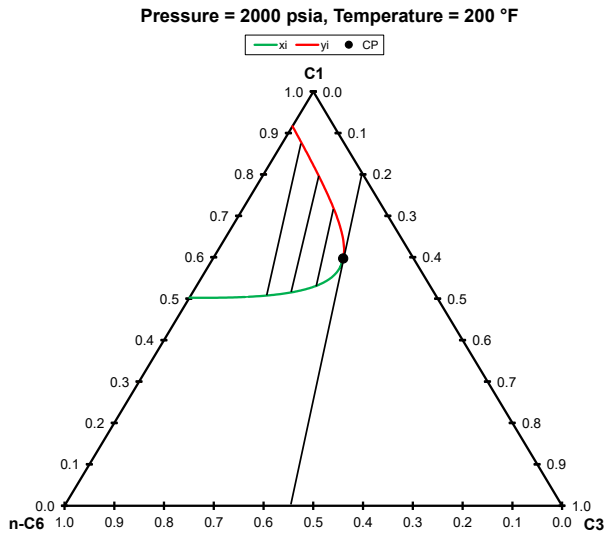
Figure 2.6: An example of the resulting phase compositions after a flash with tie-lines for a ternary system consisting of C_1, C_3 and n-C_6 at a temperature of $200°$F and pressure of 2000 psia using a PR EOS. The dew-point line is shown in red and the bubble-point line is shown in green with the critical point is shown as a black circle.

The line between the composition from the flash calculation (x_i and y_i) is referred to as the *tie-line* for the specific total composition (z_i). Examples of different tie-lines for different compositions are shown in Figure 2.6. The tie-line for the critical point (shown with a black circle) is tangent to the phase envelope at the critical point and is often referred to as the critical tie-line extension. An example of the critical tie-line extension is shown by a black line through the critical point in Figure 2.6.

2.3 Regression Methods

2.3.1 Newton's Method

The method commonly referred to as Newton's method, also known as the Newton-Raphson method, is a root finding algorithm that utilizes the slope (either analytical or approximated) of a root function $h(x)$ to determine the direction of the search. Traditional applications of Newton's method estimate the next step (x_{n+1}) by calculating the intercept of the slope at the current step (x_n) with the x-axis. Formally, the method can be described by equation (2.29) for a single variable root function.

$$x_{n+1} = x_n - \frac{h(x_n)}{h'(x_n)} \tag{2.29}$$

where $h'(x_n)$ is the derivative (i.e. the slope) of the root function at the current step (x_n). A visual example of the method is shown in Figure 2.7.

One significant drawback of Newton's method are cases where the derivative does not direct the search in the correct direction. As an example, in Figure 2.7, if the starting point is to the left of the local minima, the derivative would move the away from the true solution. For this reason, bounds are typically enforced such that if the solution leaves the relevant solution space, then the root finding algorithm should add a heuristic approach to converge notifying the user of the fact that the method has hit one of the bounds.

An equivalent approach to Newton's method is also available for multivariate objective functions. The formal definition of the multivariate version of Newton's method is given by

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{J}(\mathbf{x}_n)^{-1} \cdot \mathbf{h}(\mathbf{x}_n) \tag{2.30}$$

where $\mathbf{J}$ is the Jacobian matrix defined by:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_M}{\partial x_1} & \cdots & \frac{\partial h_M}{\partial x_N} \end{bmatrix} \tag{2.31}$$

and $M >= N$ is assumed for the system of linear equations to be solvable, and $\mathbf{h}(\mathbf{x})$ being the set of root functions.
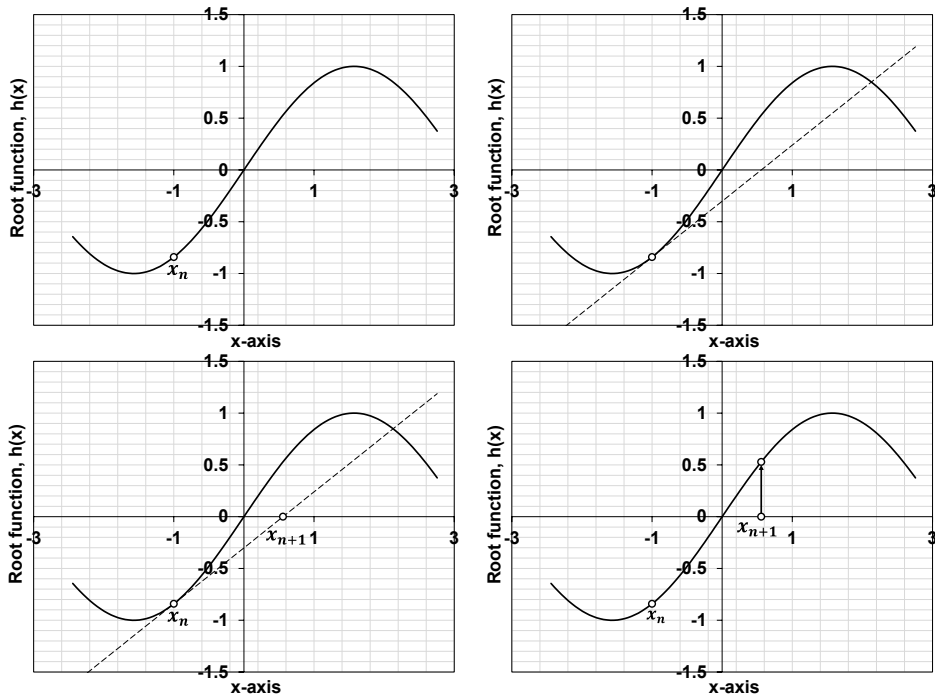
Figure 2.7: An example a step in the Newton method. The top-left figure shows the current position and value of the root function. The top-right figure shows the slope at the current position (black dashed line). The bottom-left figure shows the intercept between the slope and the x-axis. The bottom-right figure shows the new root function value at the new step.

2.3.2   Gradient Descent Method

Similar to Newton's method described in section 2.3.1, the gradient descent method utilizes the derivative (analytic or estimated) direct the search in the *right* direction and try to minimize its objective function. However, unlike Newton's method, gradient descent is not a root finding algorithm, but rather a minimization algorithm. Nevertheless, for the special case of sum of square (SSQ) minimization problems, the result of a global minimization and a root finding algorithm will coincide.

The basis of the gradient descent method is to calculate the gradient at a specific point ($\mathbf{x}$) which will point in the direction of the greatest accent of objective function in question at the specified point. By multiplying the gradient by -1, the product will by definition point in the direction of the steepest descent. A definition of the direction vector ($\vec{u}$) is given by

$$\vec{u} = -\nabla h(\mathbf{x}) \tag{2.32}$$

Normalizing the direction vector ($\vec{u}$) such that its 2-norm (magnitude) is unity yields a

(a) Calculate gradient at current step

(b) Calculate location of next step

(c) Calculate gradient at next step

(d) Calculate location of next step
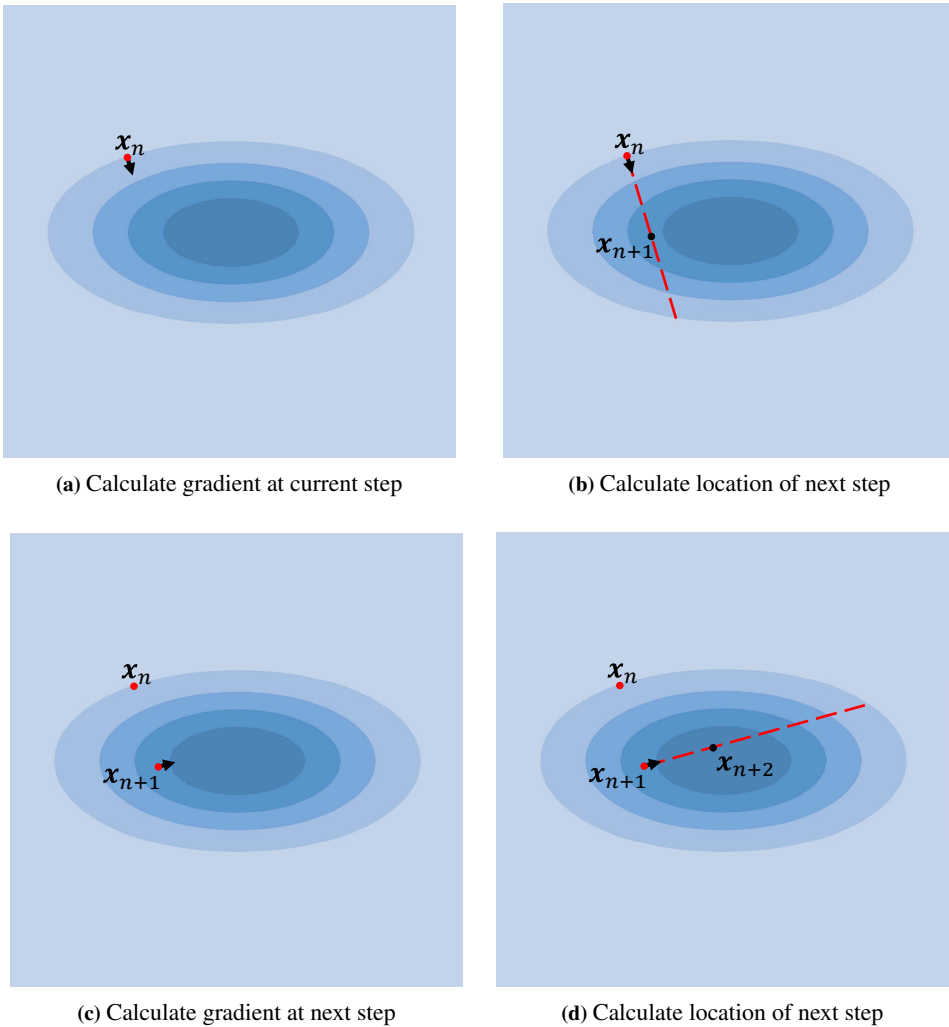
Figure 2.8: An example with two new estimated steps using the gradient descent method.

normalized direction vector ($\hat{u}$)

$$\hat{u} = -\frac{\nabla h(\mathbf{x})}{||\nabla h(\mathbf{x})||_2} \tag{2.33}$$

Utilizing the normalized direction vector ($\hat{u}$), a single variable (β) line search can be carried out

$$\beta^* = \underset{\beta \geq 0}{\operatorname{argmin}} \{h(\mathbf{x} + \beta\hat{u})\} \tag{2.34}$$

where β is a positive scalar value and β^* is the solution to the line search. The iterative method carried out the above methodology N times or until the distance between current and the previous step or the magnitude of the gradient vector is less than some threshold. The gradient descent method can be described by

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \beta^* \frac{\nabla h(\mathbf{x})}{||\nabla h(\mathbf{x})||_2} \tag{2.35}$$

An illustration of the methodology is shown in Figure 2.8.

2.3.3 Powell's Method

Powell's method was described in [30] as a restart procedure for the conjugate gradient method. The method is a local search procedure with implications as a module in wide range of search algorithms. A simplified version of Powell's method, referred to as coordinate descent, is described below.

Let $h(\mathbf{x})$ be the multivariate objective function and $\mathbf{x}$ be the set of N variables that make up the solution space. The essence of Powell's method is simply a combination of a heuristic for which variable in $\mathbf{x}$ to be used in a line search. The simplest heuristic is simply to choose each variable in $\mathbf{x}$ by turn until (a) the calculated value of of a sum of squares problem is less than a specific threshold, (b) some metric to determine that the distance traveled from the current step to the next is less than some threshold or (c) the maximum number of iterations is exceeded.

2.3.4 Pearson Correlation Coefficient

In statistics the concept of covariance is described by the measure of the joint variability of two stochastic variables. In other words, if the variability of the two stochastic variables tend to increase or decrease in union then the method is said to be positively covariant and if the sign of the change is different for the two variables then they are said to be negatively covariant. The definition of the covariance, σ, is given in equation (2.36).

$$\sigma(X, Y) = E[(X - E[X]) \cdot (Y - E[Y])] \tag{2.36}$$

where $E[\circ]$ is the expected value of some variable. For discrete variables, the covariance can be expressed as a function of measured data X_i and Y_i as in equation (2.37).

$$\sigma(X, Y) = \frac{1}{N} \sum_{i=1}^{N} (X_i - \mu_x)(Y_i - \mu_y) \tag{2.37a}$$

$$\sigma(X, Y) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{i=j}^{N} (X_i - X_j)(Y_i - Y_j) \tag{2.37b}$$

where μ is the mean value for the X or Y denoted by the subscript.

For two parameters that exhibit a linear relationship, the magnitude of the covariance will determine how strong the relationship between the two variables are. The idea of this method is to, by calculating the covariance of all parameters with the model function output, estimate the effect of the relationship between the parameters and the result. However, one major drawback to this method is that it only predicts linear, or at least a monotonic, relationship between the parameter and the solution of the model function.

As an example, take the covariance of a variable x and the function $f(x) = x^2$. If the range of x is strictly greater or less than 0 then the covariance will be non-zero. However if the range is between -1 and 1 then the covariance will, for a large number of measurements, show a covariance of ~ 0 (see Figure 2.9). The reason for this is that there is as much positive covariance as negative and the ability of this method to estimate the importance is lost. Nonetheless, if the range is "small" and the relationship is somewhat monotonic, then the calculated covariance can be a good estimate for the importance of a parameter.
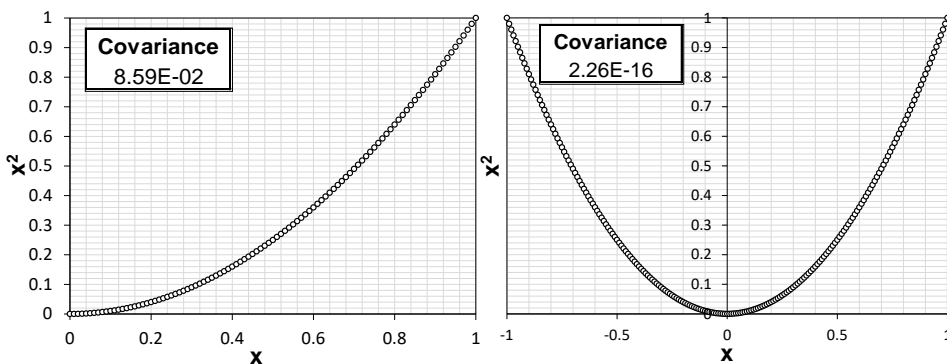
Figure 2.9: Plot of x^2 vs x and the calculated covariance with intervals of $0 \leq x \leq 1$ (left) and $-1 \leq x \leq 1$ (right).

A useful method to normalize the covariance with respect to units and be able to compare the results of different types of data is by calculating Pearson's product-moment correlation coefficient (PCC) [27]. This is shown in equation (2.38).

$$\rho(X, Y) = \frac{\sigma(X, Y)}{\sigma(X) \cdot \sigma(Y)} \tag{2.38}$$

In equation (2.38) $\sigma(\circ)$ is the standard deviation for some variable (e.g. X). One of the major features of the PCC method is that the magnitude will be normalized such that $|\rho| \leq 1$.

2.3.5 Random Forest Method

To be able to comprehend the random forest method, the underlying principles must be described in some detail. This will be described in the following paragraphs.

Decision Tree Method

The random forest algorithm is built of a range of different decision trees and it is therefore favorable to describe the mechanisms of these. The traditional discrete decision tree is composed of a set of features (F) which contain a set of attributes (A). For all the features, some combination of the attributes is collected with the output value of the specific case.

When generating the decision tree, the goal is to predict the outcome with as shallow a tree as possible to avoid over-fitting. This can be implemented by defining the $Gain$-function as in equation (2.39)

$$Gain(S, F) = Entropy(S) - \sum_{v \in Values(F)} \frac{|S_v|}{|S|} Entropy(S_v) \qquad (2.39)$$

where S is the system and F is the feature, $S_v = \{s \in S | F(s) = v\}$ is the system with attribute v and $|\circ|$ is the number of elements in the system. The entropy can be calculated as shown in equation (2.40)

$$Entropy(S) = \sum_{i=1}^{c} -p_i \cdot \log_2(p_i) \qquad (2.40)$$

where p_i is the probability of feature F of having attribute A and c is the total number of attributes.

A decision tree takes a dataset S and branches from a base node by the principle of maximum gain at each step. The termination of decision tree is based on the sub-dividing of the dataset until the gain of the final node of the decision tree (i.e. the terminal node) is equal to unity. This is equivalent to having a completely certain result. This description of the decision tree method was adapted from Michell [23] and a specific example of how a decision tree is built is shown in appendix D.

The final product of a decision tree is a function $h(\mathbf{x})$ that takes in a dataset x and trains the function such that a trained function $\hat{h}(\mathbf{x})$ that predicts the type output with a specific type of data. The main sets of user defined parameters in the decision tree method are the maximum depth of the tree and the number of branches per node (i.e. the width of the tree). The hyper-parameters depend on the problem at hand and need to be described based on the problem at hand.

Bagging Method

First described by Breiman in 1996 [4] the method of bagging tackles one of the main issues of the decision tree, namely overfitting of the data. The method for counteracting overfitting is simply by averaging over a number of decision trees with their own, smaller, set of training data. The averaging, also called voting, is shown in equation (2.41).

$$f(\mathbf{x}) = \frac{1}{J} \sum_{j=1}^{J} h_j(\mathbf{x}) \qquad (2.41)$$

In the above equation, $f(\mathbf{x})$ (typically referred to as *voting*) is the bagging method prediction function and J is the total number of trees used (i.e. the number of trees in the forest). Equation (2.41) assumes that the output is continuous, but this is not necessary and a discrete version can be found in the book by Zhang [47]. The topic of transferring from discrete to continuous problems is also described in detail, however this level of detail is not relevant for this work. There is also an additional hyper-parameter that comes with the bagging method and that is the number of decision trees.

Random Forest Method

Similar to the bagging method, the random forest also tries to predict the outcome of a system based on a set of decision trees. However, the traditional random forest methods deviate from the traditional bagging methods as described by Breiman [4] is that the random forest method utilizes only a random subset of the input variables to build the decision trees. This additional randomness aids in reducing overfitting and also has the added benefit of more accurately prediction a large set of systems. One concise description of the random forest method is given in equation (2.42)

$$f(\mathbf{x}) = \frac{1}{J} \sum_{j=1}^{J} h_j(\mathbf{x}_j) \qquad (2.42)$$

where $\mathbf{x}_j$ is a subset of of the entire input set $\mathbf{x}$. Similar to the bagging method, this extra degree of freedom introduces another hyper-parameter which is the size of the subset of input parameters.

Feature Importance

Based on the description of the random forest method and some dataset, it is possible to train the random forest function to the given data. Based on the trained random forest, an estimate of the feature importance can be made. The algorithm for calculating one such metric is detailed by Zhang [47].

Pro's and Con's of the Random Forest Method

Below is a list of some of the relevant pro's of the random forest method:

- The primary reason why the random forest method is useful is the fact that there is little to no risk of over-fitting the dataset, unlike other prominent methods like neural networks.

- It has been shown that the random forest algorithm can perform well with systems containing a large number of parameters [2].

- They have been shown to accurately predict irrelevant or null-effect variables [47]

However, there are some drawbacks to the method as described by Zhang et al.:

- They are not good at capturing relationships involving linear combinations of predictor variables;

- They are known to be unstable in the sense that if the data are perturbed slightly, the tree can change substantially;

- They are not as accurate as some of the more recently developed methods.

Example - Play Tennis Decision Tree

An example case from the book by Mitchell [23] is given in appendix D section D.2. This example goes through the details of how to build a decision tree for whether or not you are going to play tennis based on some data about the weather from previous days.

2.4 Software

The main software that have been utilized to generate the data for this work are summarized below.

2.4.1 PhazeComp

PhazeComp is a text-based PVT software developed by Zick Technologies that allows for a wide range of PVT calculations like vapor-liquid equilibrium calculations, standard PVT experiments like CCE, CVD, MSS and more. Another key feature of the software is the ability to do regression on the fluid model. The flexibility and accuracy of the text based software allows for a wide range of phase behavior computations and has been instrumental to the calculations used in this work, where data sets are automatically generated by Python code.

2.4.2 Python 3.7

Python is an object-oriented open source programming language. Python is one of the largest open source high-level programming languages and has a vast catalog of user developed libraries. The key libraries used in this work are summarized below:

- *numpy* v. 1.17.4: Numpy is a mathematics-based package which allows for easy handling of arrays and mathematical operations.

- *pandas* v. 0.25.3: Pandas is a data-structure library with a wide range of data management tools like making DataFrames (pandas specific object) which can easily be transported to and from Excel, numpy etc.

- *matplotlib* v. 3.1.2: Matplotlib is a plotting library which is designed to work in unison with pandas and numpy.

- *scikit-learn* v. 0.22.2: Scikit-learn or sklearn is a machine learning library with a wide range of different machine learning and artificial intelligence tools and methods.

- *python-ternary* v. 1.0.6: Plot-ternary is a plotting library which allows for plotting of data in ternary plots. A unique feature, which was utilized in this work, is the ability to plot ternary data as a heat-map or contour plot.

Another library that was used, but is not a public library is phasecomp-utils. This library was provided by Whitson AS and is a PhazeComp specific parser which has been an integral part of the supporting code for the data generation of this thesis.

2.4.3 PyCharm Community 2019.1.3

PyCharm Community is a free-to use Python environment. With options for running, writing and debugging Python code PyCharm allows for ease of use when developing Python code. There are also features like Git, which allows for more complex projects.

The Effect of Binary Interaction Parameters on Phase Behavior

"What is the use of a book without pictures or conversations?"

- Alice in Wonderland

In this chapter a study is made on how BIPs affect EOS phase behavior calculations, and particularly K-values, compositions and phase boundaries. First, binary systems are investigated, followed by a study of ternary systems. The aim is to first understand the simplest systems (binaries), and then the next simplest systems (ternaries). Understanding the simpler systems will hopefully help in understanding multi-component systems which can give significant insight to the effect on real petroleum systems. The study will investigate how the p-T-z relationship changes with BIPs ranging from negative to positive. A study on the effect of BIPs on the pressure dependence of equilibrium ratios (K-values) is also given for a range of temperatures and compositions. The study will also show the effect of BIPs on the phase envelope for a range of compositions as well as describe the shift in critical pressure and temperature and the expansion or contraction of the area of the PT phase envelope.

3.1   Binary System

In this section a range of values for the BIP between C_2 and C_7 will be applied. The effect on the critical pressure is given in section 3.1.1, and the effect on equilibrium ratios in section 3.1.2. The range of the BIP span from negative to positive values of various magnitudes (± 0.15).

3.1.1 Effect of BIP on the Critical Locus of a Binary

As described in section 2.2.2, any mixture's phase behavior can be described by the phase envelope for the specific mixture and given composition. Specifically, the pressure-temperature phase envelope can give a good indication of how the mixture will behave with respect to pressure and temperature. An example of the critical locus is shown in Figure 3.1 for a binary system consisting of a C_2 and C_7 mixture with the EOS from Table 4.4 and 4.6.
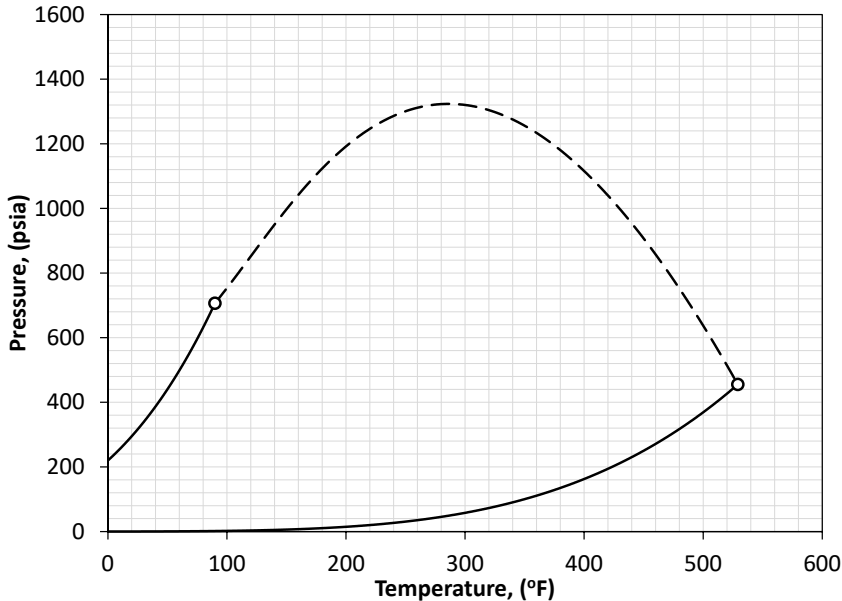
Figure 3.1: Example of critical locus for the binary system of with the EOS from Table 4.4 and 4.6. The solid lines are the vapor pressure lines for C_2 (left) and C_7 (right) and the dashed line is the critical locus for all mixtures ranging from 100% C_2 (left) to 100% C_7 (right).

The phase envelope shown in Figure 3.1 used PhazeComp to calculate the results of a constant composition expansion (CCE) test with injection at each stage. At the first stage of the CCE test the mixture is set to contain 100-ϵ % amount of C_2 and ϵ % amount of C_7. In this case ϵ is some small amount (e.g. 0.01%) followed by an increase of relative moles injected (RMI). The critical point was found for each composition yielding the temperature dependent critical locus. Based on the thermodynamics of binaries, the convergence pressure line for any binary composition is equal to the critical locus. This fact was used to more easily generate the critical locus as, for each stage of the CCE, the convergence pressure was calculated at the specific temperature instead of having to find the critical point. The convergence pressure algorithm in PhazeComp can also calculate the composition of the mixture at the convergence pressure by the negative flash. This convergence composition will be the same as the composition with its critical point at the convergence pressure.

This fact allows for a rigid bound on the possible temperature range for the convergence pressure based on the critical properties of the components in the binary system. As the composition approaches 100% of a single component, the limiting critical pressure must tend to the critical pressure of the specific component. This limits the temperature range that will yield a critical point (convergence pressures) to a temperature range of $T_{c,L} \leq T \leq T_{c,H}$ where $T_{c,L}$ is the light component critical temperature and $T_{c,H}$ is the heavy component critical temperature.

Including a non-zero BIP to the binary system and re-calculating the critical locus was the following step. This was done for a positive BIP with a magnitude ranging between 0 up to 0.145. Above this the critical locus either doesn't close or closes at a pressure higher than the maximum allowed pressure constraint in PhazeComp. The results for the BIPs of magnitudes 0, 0.05, 0.1 and 0.145 are shown in Figure 3.2.
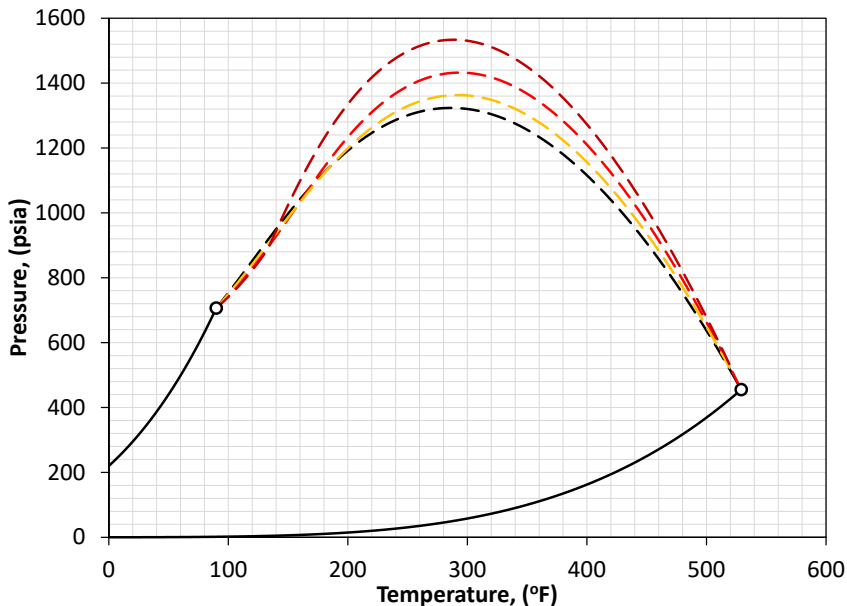
Figure 3.2: Effect of positive BIP on critical locus for C_2-C_7 mixture where BIP=0 (black), BIP=0.05 (orange), BIP=0.1 (red) and BIP=0.145 (dark red).
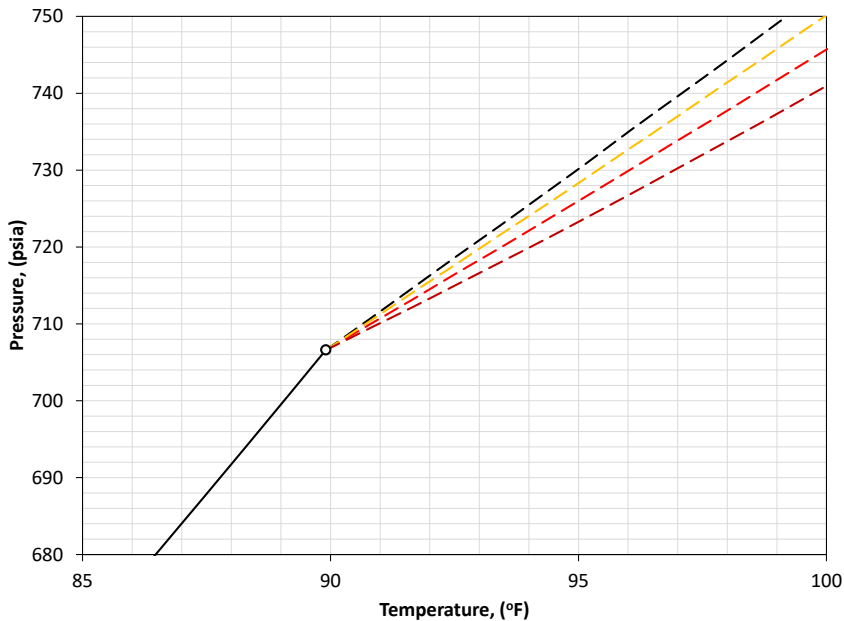
Figure 3.3: Effect of positive BIP on critical locus for C2-C7 mixture where BIP=0 (black), BIP=0.05 (yellow), BIP=0.1 (red) and BIP=0.145 (dark red) zoomed in near the light component critical point.

Zooming into the region near the critical temperature of the light component (C_2 in this case) shows that the behavior of the critical pressure tends to be divided into two parts. (1) For the mid- to high-temperature range the effect of the positive BIP is that the critical pressure is *increased*, and (2) for temperatures near the light component critical temperature, the convergence pressure tends to *decrease* with respect to the zero BIP case. This effect is also seen in a range of other binary pairs (see Appendix A) and seems to be a general trend.

On the other hand, negative BIP values ranging from 0 to -0.15 were also tested with results for the specific values of 0, -0.05, -0.1 and -0.15 shown in Figure 3.4.
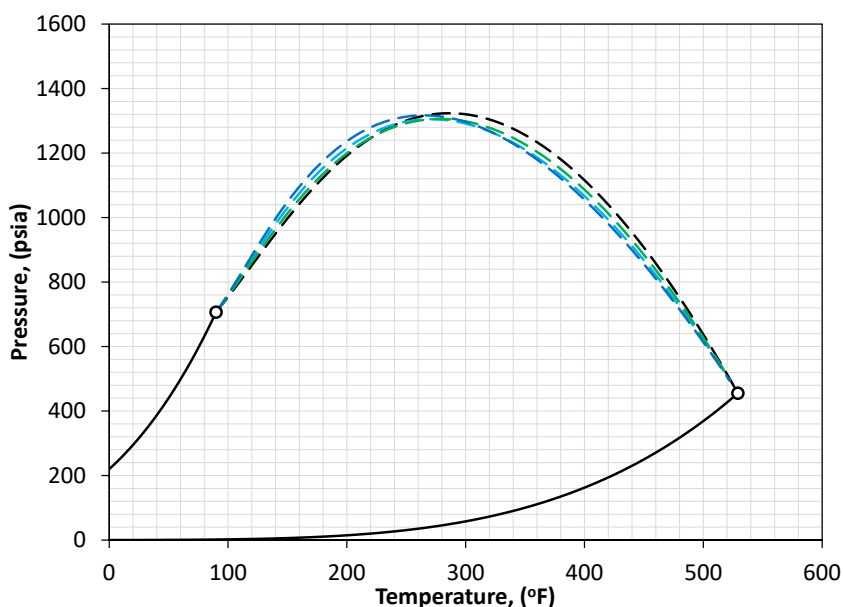
Figure 3.4: Effect of negative BIP on critical locus for C_2-C_7 mixture where BIP=0 (black), BIP=-0.05 (green), BIP=-0.1 (light blue) and BIP=0.15 (dark blue).

The general behavior of the negative BIP seems to also be divided into two regions: (1) for mid to high temperatures the critical pressure tends to *decrease* and (2) for low temperatures near the critical temperature of the light component the critical pressure tends to *increase*.

The overall effect of the BIP on the critical locus seems to depend on three main factors. (1) The sign of the BIP, (2) the magnitude of the BIP and (3) the components of the binary system. Figure 3.2 shows that the trend for positive BIPs increases the critical pressure above a certain temperature and decreases below that temperature. Similarly, Figure 3.4 shows that for negative BIPs the critical pressure decreases above a certain temperature and increases below that temperature. The magnitude amplifies the change in the critical pressure for both signs (e.g. larger magnitude results in larger deviation in critical pressure). In Appendix A, more binary pairs are shown and the relative magnitude of the effect for a given BIP tends to relate to the size of the molecule (e.g. the molecular weight).

Figure 3.5 shows the extremes of the positive BIP (value of 0.145) and negative BIP (value of -0.15), that were tested in this study, with respect to the zero BIP case. The figure shows the general trend of the effect of BIPs on the critical locus for binary pairs.
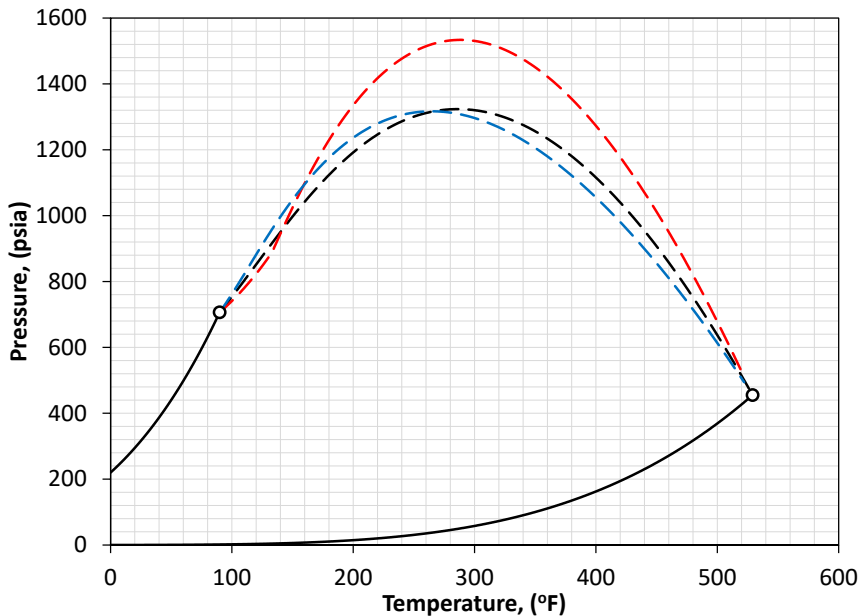
Figure 3.5: Comparison of positive and negative BIP on critical locus for C_2-C_7 mixture where BIP=0 (black), BIP=0.145 (dark red) and BIP=-0.15 (dark blue).

3.1.2 Effect of BIP on the K-values for Binary Systems

As previously mentioned, the critical and convergence pressures are uniquely connected for binary systems which allows the relationship between the critical locus and the shape of the pressure dependent equilibrium ratios (K-values). The relationship between a shift in critical pressure and the shape of the K-value plot is based on the fact that for any mixture with the convergence pressure equal to the critical pressure, the positive flash K-value plot will yield the entire shape (i.e. continuous positive flash K-values to the convergence pressure). This fact defines the relationship between the shift in critical pressure as one-to-one with the convergence pressure shift and equivalently the tip of the K-value curve.

The K-value plot was generated for the specific mixtures of C_2 and C_7 corresponding to two separate temperatures. The temperatures were chosen to display the effect of the BIPs in the two regions described in section 3.1.1. The two temperatures that are shown in the figures below are at 110°F and 300°F. The figures below show the full positive flash K-value plots with a zoomed plot of the high-pressure region for the low and intermediate pressures.

From Figure 3.6 the same relationship as described in the previous section is found. The convergence pressure is decreased for the positive BIP in region 1 and increased for the negative BIP in region 1. The low pressure region shows an expansion of the K-value plot
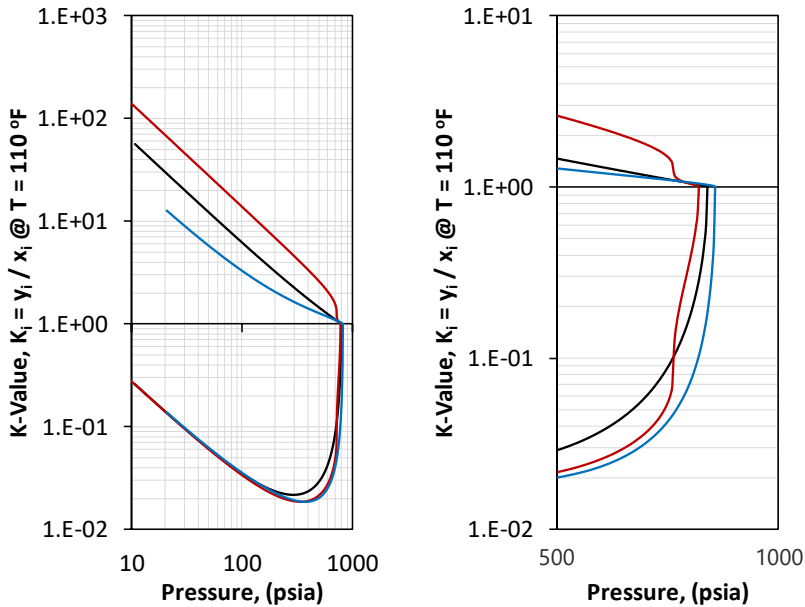
Figure 3.6: Comparison of the K-value plots for C_2-C_7 binary pair with BIP=0.1 (red), zero BIP (black) and BIP=-0.1 (blue) at a temperature of 110°F.

for the positive BIP and a contraction for the negative BIP. This effect is true for both the light and heavy component, yet the relative magnitude is larger for the lighter component. One note for the negative BIP K-value plot in Figure 3.6 is that the lower pressure at which the fluid form a single phase is increasing for the negative BIP and decreasing for the positive.

From Figure 3.7 the same relationship as in the previous section is also found for the intermediate temperature (i.e. region 2). The convergence pressure is increased for the positive BIP and decreased for the negative BIP. However, the low-pressure region shows an expansion of the K-value plot for the positive BIP and a contraction for the negative BIP. The low-pressure behavior therefore seems to be independent of temperature and only depends on the sign and magnitude of the BIP. This effect is true for both the light and heavy component, yet the relative magnitude of K-value change is larger for the lighter component.
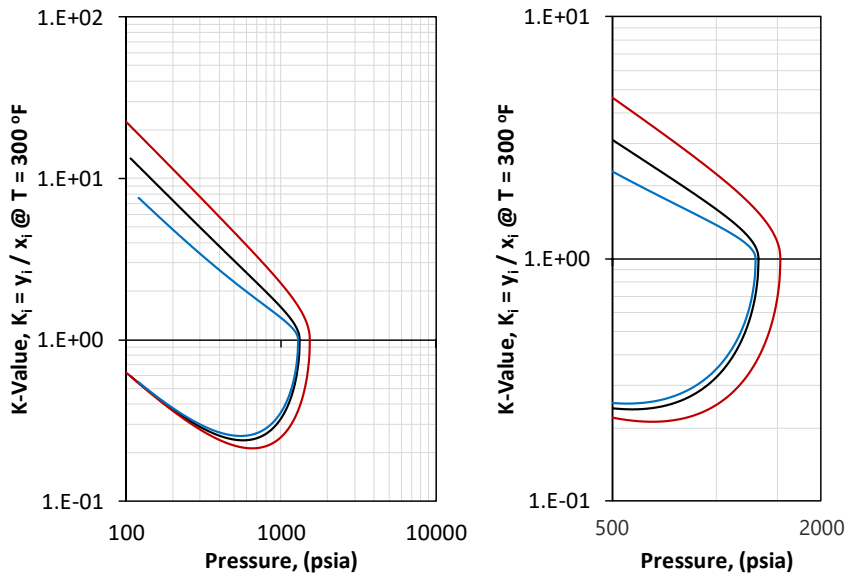
Figure 3.7: Comparison of the K-value plots for C_2-C_7 binary pair with BIP=0.1 (red), zero BIP (black) and BIP=-0.1 (blue) at a temperature of $300°$F.

3.2 Ternary System BIPs

In this section a range of values for BIPs between single carbon number (SCN) components C_2, C_5 and C_7 will be applied and the effects are presented in the form of the shift in the critical point and the impact on the pressure dependent equilibrium ratios. The range of the BIPs will span from negative to positive values with magnitudes ranging from 0 to 0.1.

3.2.1 Phase Envelope and Convergence Pressure

Similar to the binary system, it is possible to calculate the binary critical locus between each pair of binaries in the ternary system. The reason for making these calculations is that they are the limiting values for the critical point as the ternary compositions approaches any of the possible binary mixtures or a single component mixture. An example of such a plot given in Figure 3.8a for C_2, C_5 and C_7.

For any ternary mixture, it is possible to calculate the phase envelope and the corresponding temperature dependent convergence pressure line (shown by a light grey dashed line), as shown in Figure 3.8b for a mixture containing 50% C_2 and 25% of both C_5 and C_7. The range of the temperatures where the convergence pressure is calculated is limited by the software used (PhazeComp) and its ability to find a saturation pressure before a convergence pressure can be calculated. If a saturation pressure is not found, then the software

does not find the convergence pressure (which actually exists). Based on this limitation, the upper limit of the temperature range is the cricondenterm of the mixture.

(a)

(b)
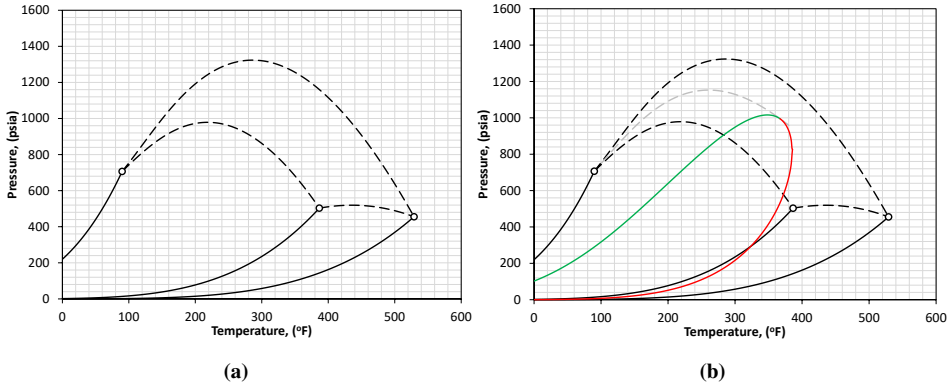
Figure 3.8: An example of a ternary system (C_2, C_5 and C_7) with the critical loci between all binaries (dashed lines). In (a) only the binary boundaries are shown while (b) gives an example of a phase envelope containing 50% C_2 and 25% C_5 and C_7.

Similar to section 3.1.1 the BIPs between the separate binaries ranges from positive to negative values. The corresponding shift in the phase envelope of the specific mixture for a positive BIP with the value of 0.1 for each binary is shown in Figure 3.9. The top left figure shows the shift of the phase envelope based on a BIP of 0.1 between C2 and C5, the top right figure shows the shift in the phase envelope based on a BIP of 0.1 between C2 and C7 and, finally, the bottom figure shows the shift of the phase envelope based on a BIP of 0.1 between C5 and C7. The original zero BIPs case is shown by a light grey plot behind the non-zero BIPs case in each sub-plot. The zero BIPs critical loci are shown with a black dashed line, whereas the new critical loci between the binaries with non-zero BIPs is shown in a dark red dashed line.

(a) BIP=0.1 between C_2 and C_5

(b) BIP=0.1 between C_2 and C_7

(c) BIP=0.1 between C_5 and C_7
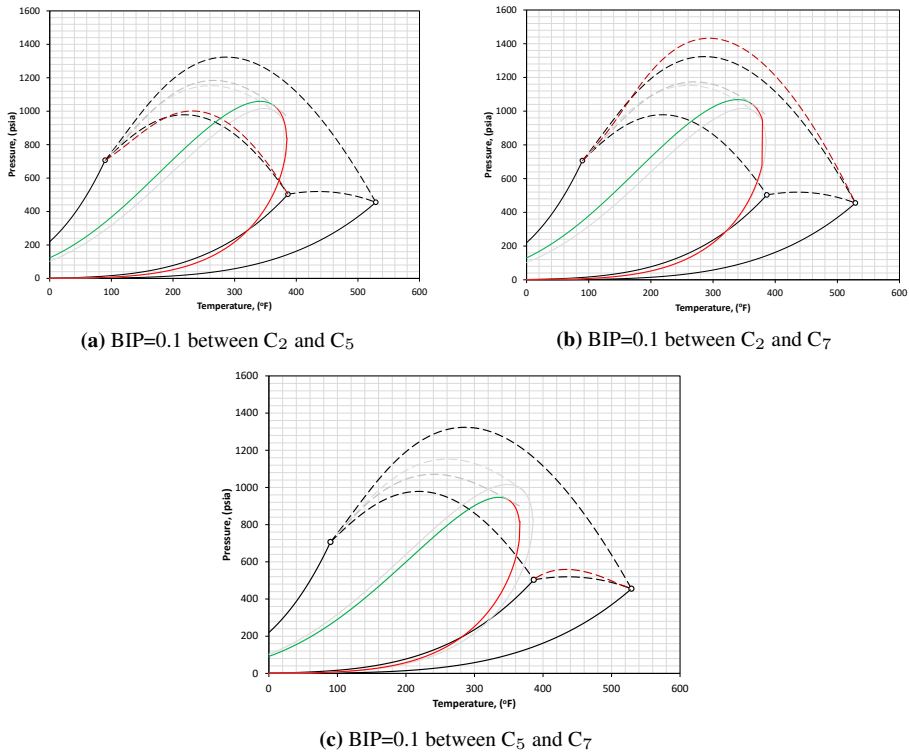
Figure 3.9: An example of a ternary system (C_2, C_5 and C_7) with the critical loci between all binaries as well as the phase envelope of a mixture containing 50% C_2 and 25% C_5 and C_7. The plot also contains the temperature dependent convergence pressure line (dashed grey).

One issue with the method of showing the shift in the phase envelope shown in Figure 3.9 is that it is dependent on the mixture composition used in the calculation. A method for describing the shift in shape and location of the phase envelope with respect to a zero BIPs case for all compositions with a constant set of BIPs was generated by calculating the following quantities:

$$\Delta PE = [\Delta p_c, \Delta T_c, \Delta A] \tag{3.1}$$

$$\Delta p_c = p_c(k_{ij} \neq 0) - p_c(k_{ij} = 0) \tag{3.2a}$$

$$\Delta T_c = T_c(k_{ij} \neq 0) - T_c(k_{ij} = 0) \tag{3.2b}$$

$$\Delta A = A(k_{ij} \neq 0) - A(k_{ij} = 0) \tag{3.2c}$$

Table 3.1: Full list of cases run for ternary plots showing the shift in the critical pressure as well as the change in area of the phase envelope. For results see Appendix B.

Case Name	BIP C2-C5	BIP C2-C7	BIP C5-C7
A-1	0.1	0	0
A-2	0	0.1	0
A-3	0	0	0.1
B-1	-0.1	0	0
B-2	0	-0.1	0
B-3	0	0	-0.1
C-1	0.1	0.1	0
C-2	0.1	0	0.1
C-3	0	0.1	0.1
D-1	-0.1	-0.1	0
D-2	-0.1	0	-0.1
D-3	0	-0.1	-0.1
E-1	0.1	-0.1	0
E-2	-0.1	0.1	0
F-1	0.1	0	-0.1
F-2	-0.1	0	0.1
G-1	0	0.1	-0.1
G-2	0	-0.1	0.1
H-1	0.1	0.1	0.1
H-2	-0.1	0.1	0.1
H-3	0.1	-0.1	0.1
H-4	0.1	0.1	-0.1
H-5	-0.1	-0.1	0.1
H-6	-0.1	0.1	-0.1
H-7	0.1	-0.1	-0.1
H-8	-0.1	-0.1	-0.1

where A is the enclosed area of the phase envelope for temperatures greater than $0°F$. This method was applied to mixtures containing C_2, C_5 and C_7 for values of the BIPs described in Table 3.1. An example of the ternary plots of the properties in equation (3.1) shown in Figure 3.10 for the case where the BIP between C_2 and C_5 is equal to 0.1, while the remaining set of BIPs are set to zero. The full set of different combinations of BIPs are given in Appendix B.

(a)

(b)

(c)
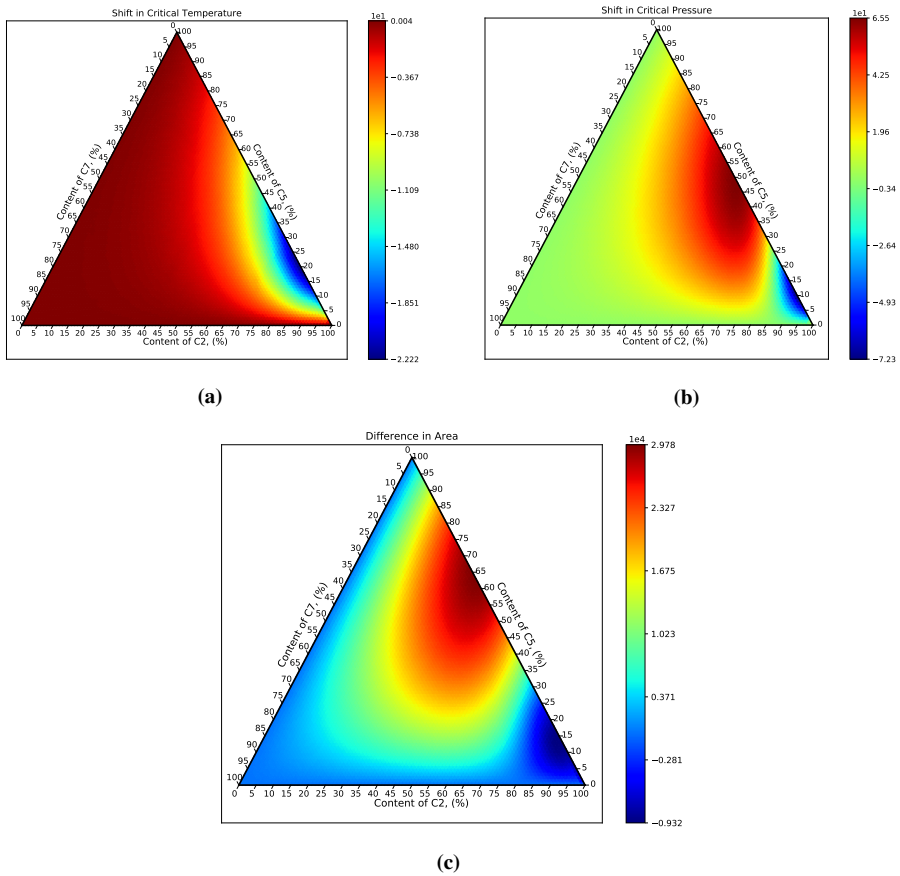
Figure 3.10: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2 and C_5 and zero BIPs for the C_2-C_7 and C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)

One important note for the difference in area (ΔA) is that the arbitrary cut-off point for the lower bound temperature will have an effect on the value for the difference in area. This is due to the fact that a shift in the of the phase envelope might result in the same area, yet due to the arbitrary cut of point at $0°$F the shift will exclude or add some of the shape, resulting in a reduction or increase in the estimated area. Therefore, the area difference plot is most useful for cases where the critical point is approximately constant, and the shape of the phase envelope is different (resulting in a change in the area). However, from the values of the ΔA for temperatures ranging from $0°$F to about $50°$F the difference between the saturation pressure (upper pressure) and the lower dew point pressure (lower pressure) is relatively small compared to the higher temperatures, which results in a small artificial shift because of the exclusion of the low temperature area.

3.2.2 Effect of BIPs on K-values for Ternary Systems

Similar to section 3.1.2 a mixture has to be used which has its saturation point at the critical point to calculate positive flash K-values for all pressures up to the point of convergence. However, unlike the case of binary systems, ternary and other multi-component systems don't have a single composition with this property, rather they have an infinite set of combinations along a critical tie-line or surface extension (as shown in Figure 2.5). Given this infinite range of compositions to choose from, the question becomes which composition should be used. The following section will describe a method that ties the initial feed composition (z_i^*) directly to a unique critical composition (z_{Ki}). The proposed procedure applies the negative flash calculation at the convergence pressure of the initial composition (i.e. $p_k(z_i^*)$) to yield the critical composition (z_{Ki}). K-value data are calculated for a range of temperatures with a variation of BIPs and compared to a zero BIPs case.

Convergence Pressure Composition
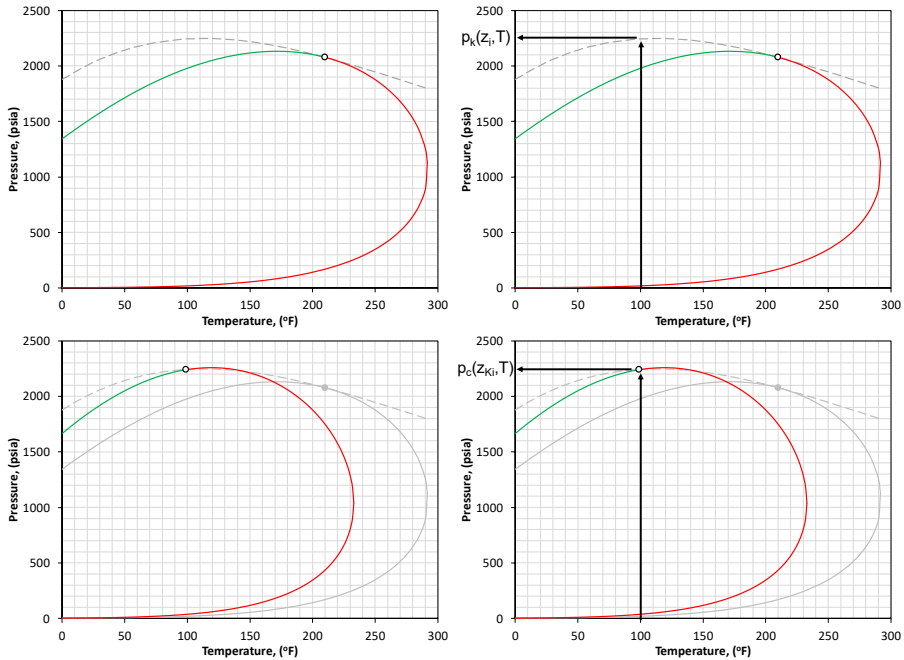
Figure 3.11: An example of convergence composition method for an initial equal molar amount (uniform) mixture with the EOS given in Tables 2.1 and 2.2 with a temperature of $100°$F.

As described briefly above, the goal of this method is to find a unique composition with the property of having its critical point at the same value as its convergence pressure based on some initial feed composition. First, let z_i^* be the initial feed composition. Using the initial composition at some temperature T^*, it is possible to calculate the convergence pressure of

the mixture at the specific temperature as $p_K(z_i^*, T^*)$. By solving the flash calculation at $(p_K(z_i^*, T^*), T^*)$ an equilibrium mixture for liquid (x_{eqi}) and vapor (y_{eqi}) is found by the negative flash. Because the pressure is calculated at the convergence pressure of the initial mixture, the theoretical equilibrium compositions will be equal (i.e. $x_{eqi} = y_{eqi}$). By defining a new composition z_{Ki} equal to the flashed composition (i.e. $z_{Ki} = x_{eqi} = y_{eqi}$) it is assured that the saturation pressure of the mixture will be equal to the critical pressure of the initial composition and because the composition of liquid and vapor are equal the saturation pressure point is also at the convergence (i.e. $K_i = 1$ for all components). Therefore, the proposed method retains the property of having a saturation pressure equal to its critical pressure which in turn defines the saturation pressure as the convergence pressure.

An added feature of the proposed method is that the new composition, which will from now on be referred to as the convergence composition (z_{Ki}), is defined by the initial composition (z_i^*). In fact, the convergence composition will be connected to the initial composition by the critical tie-line extension of the convergence composition at the convergence pressure of the initial fluid (z_i^*). An example is shown in Figure 3.11 where the top left figure shows the phase envelope of the initial composition (z_i^*) followed by the top right figure where a temperature (T^*) is chosen and a convergence pressure is calculated for the initial composition $(p_K(z_i^*, T^*))$. The bottom left figure shows the phase envelope of the new composition (i.e. the convergence composition z_{Ki}) with the old phase envelope shaded in a light grey color. Finally, the bottom right figure shows that the critical point of the convergence composition $(p_c(z_{Ki}, T^*))$ is equal to the initial convergence pressure $(p_K(z_i^*, T^*))$. See Appendix D for a ternary mixture example.

Table 3.2: Full list of cases run for ternary K-value plots showing the shift K-values for C2, C5 and C7 for a range of initial compositions (z_i^*). For results see Appendix B.

Case		Description
(1)	-	Iso-line starting at binary point [0.9,0.1,0]
(1.1)	-	Comp = [0.855,0.095,0.05]
(1.2)	-	Comp = [0.81,0.09,0.1]
(2)	-	Iso-line starting at binary point [0.55,0.45,0]
(2.1)	-	Comp = [0.5225,0.4275,0.05]
(2.2)	-	Comp = [0.495,0.405,0.1]
(3)	-	Comp = [0.2,0.6,0.2]

K-value Results

K-value plots for C_2-C_5-C_7 mixture were calculated at two temperatures (T=100°F and T=150°F) using the convergence composition method described in the previous section. One example is shown in Figure 3.12 where the mixture is composed of 85.5% C_2, 9.5% C_5 and 5% C_7 at 100°F. A list of all the cases that were calculated is given in Table 3.2. The results, similar to Figure 3.12, are given in appendix B for all cases in shown in Table 3.2.
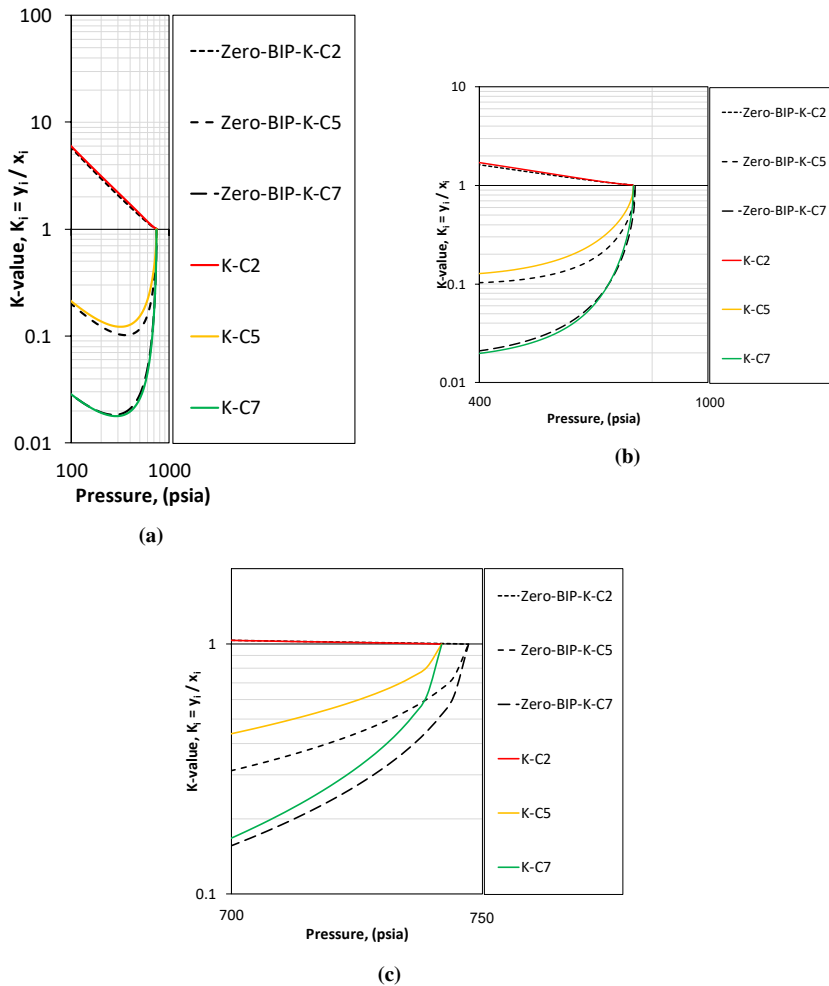
Figure 3.12: Example of K-value plot for a positive BIP = 0.1 between C_2 and C_5 for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.

As the composition is near the boundary binary composition of C_2-C_5, and based on the temperature and the results shown in Figure 3.9 the critical locus is reduced for a positive BIP between C_2-C_7. This is also seen in Figure 3.12 as the three component K-values tend to the convergence point at a lower pressure than the zero BIP case. It is also worth noting that the two lighter components (C_2 and C_5) both yield higher K-value results for low pressures while the heaviest component K-values are decreased.

There are also some additional complexities associated with ternary mixtures compared with binary systems. First, it should be noted that when applying the convergence composition methodology with and without non-zero BIPs the associated compositions at the convergence pressure will differ. This might lead to a somewhat inconsistent comparison when describing the phase behavior (i.e. the K-value plots). Second, as more BIPs are introduced with more components, the combinations of different effects from the separate BIPs make a detailed analysis harder to perform. An indication of this added complexity can be seen in Figure B.20 in Appendix B where three BIPs of different different signs are used.

In general, this type of K-value plot is useful for two main reasons. First and foremost, the data used in the regression approach in the following chapter is based on the K-value data. The reason for this is that it can be argued that the K-value data is the best metric for the fluid system phase behavior. Second, the K-value data is used for a specific type of consistency check, namely K-value crossing. K-value crossing is simply whether or not the K-values of two neighboring components cross at any point with the exception of the convergence pressure. This is believed to be non-physical behavior for hydrocarbons, but is not uncommon for non-hydrocarbon and hydrocarbon components (e.g. CO_2).

Chapter 4

Local Regression of BIPs and Selective Tuning Methodology

"First you guess. Don't laugh, this is the most important step. Then you compute the consequences. Compare the consequences to experience. If it disagrees with experience, the guess is wrong."

- Richard Feynman

In the following chapter, methods for regression of BIPs will be discussed and compared. First, in section 4.1, the regression problem at hand will be defined. In section 4.2, a simultaneous regression approach on all the BIPs using PhazeComp's solver is used to develop a base case description of a possible solution. Section 4.3 introduces a framework for a reduced regression problem, *selective tuning*, used to rank the importance of the model parameters. Sections 4.4.1 and 4.5 describe the methods for estimating "importance". The methods described in 4.5.1 and 4.5.2 are implemented for estimating the feature importance defined in section 4.3. Finally, in section 4.6, the results for two fluid systems different numbers of components are tested with the three proposed approaches.

4.1 Problem Statement

For this work it will be assumed that an EOS is to be tuned perfectly to some synthetic *true* solution with the exception of the BIPs. The task at hand is to find a BIP matrix ($\mathbf{K}$) that accurately predicts some set of experimental data from a synthetic EOS BIPs ($\mathbf{K}^*$). Given that there are N_e sets of PVT experiments, where each experiment has a set of data-points, then the set of observable data (Ω) can be defined by

$$\Omega = \left\{ \vec{\omega}_1, \vec{\omega}_2, ..., \vec{\omega}_{N_e} \right\} \tag{4.1}$$

where $\vec{\omega}_i$ is a set of measurements in a given experiment. As an example, let Ω be defined by a constant composition expansion (CCE), constant volume depletion (CVD) and multi-stage separator (MMS) test. Then $\vec{\omega}_1$ would be, for example, the volume of the PVT cell versus pressure used to estimate the saturation pressure of the fluid at a given temperature.

For the scope of this work, the only type of experiment used to describe the phase behavior will be the K-values at different pressures. However, the general problem description holds for any experiment type (e.g. CCE, CVD, DLE). For the specific type of conditions of the experiment(s) in Ω a supplementary set ($\mathcal{D}$) is calculated using the provided EOS and an initial set of BIPs ($\mathbf{K}^\circ$) as follows:

$$\mathcal{D} = \{\, \vec{d}_1, \vec{d}_2, ..., \vec{d}_{N_e} \,\} \tag{4.2}$$

Given an EOS with a specific set of BIPs, it is preferable to have a cost function (C) that can minimize the difference between the calculated ($\mathcal{D}$) and measured (Ω) data. In general, the objective of the regression approach can be described as

$$\mathbf{K}^* = \underset{\mathbf{K}}{\mathrm{argmin}}\{C(\Omega; \mathbf{K})\} \tag{4.3}$$

where $\mathbf{K}^*$ is the set of solution BIPs that is able to accurately model the measured data.

One specific type of cost function, called the root mean squared (RMS), will be used in this work and is defined by

$$C_{RMS}(\Omega, \mathbf{K}) = \sqrt{\frac{1}{N} \sum_{i=1}^{N_e} \langle \vec{r}_i, \vec{r}_i \rangle} \tag{4.4}$$

where $\langle \circ, \circ \rangle$ is the inner product[1] and $\vec{r}_i$ is the residual vector for experiment i defined by:

$$\vec{r}_i = \frac{\vec{d}_i(\mathbf{K}) - \vec{\omega}_i(\Omega)}{d_{ref}} \tag{4.5}$$

and N is the total number of measurements (data-points). The reference value ($d_{ref,i}$) depends on the type of RMS model. Two examples are (1) a relative RMS where $d_{ref,i} = \omega_i$ which weights small changes in the magnitude as much as large changes in the magnitude based on the observed value, and (2) a constant weighted RMS where $d_{ref,i}$ is constant, for example $d_{ref,i} = max\{\vec{\omega}_i\}$ which weights large changes in magnitude more than small changes in magnitude. In this work the constant type RMS with the maximum experiment value will be used as the reference as this is used in PhazeComp.

---

[1]For an example of how to calculate an inner product and the cost function, see appendix D

4.2 Simultaneous Tuning - Reference Case

Traditional solutions to the problem described in the previous section tend to be simultaneous tuning methods like multi-variate Newton's method described in section 2.3.1 or the gradient descent approach described in section 2.3.2. However, because of the $O(n^2)$ increase in number of BIPs with respect to the number of components, the computational time of simultaneous methods are not computationally feasible for systems with a large number of components, like real oil and gas reservoirs.

As an example, the PhazeComp solver shows, in section 4.6.2, that the computational time increases exponentially with respect to increasing number of components in the fluid model. This creates a constraint on the possible number of components that it is possible to use in the regression. Katz-Firoozabadi [18] and Yonous et al. [46] describe an approach where a subset of the BIPs have been found to have a good effect on specific fluid models. An approach that automatically determines which subset is relevant for the specific fluid model in question doesn't seem to exist.

This further expands the problem in the previous section to; finding a set of BIPs that describe the fluid system within computationally feasible time. This additional consideration of computational time is the basis for the proposed approach described in the following section.

4.3 Selective Tuning

The main objective of *selective tuning* is to reduce the problem given in equation (4.3) to a simplified system

$$\mathbf{K}_{opt}^* = \operatorname*{argmin}_{\mathbf{K}_{opt} \in \mathbf{K}} \{C(\Omega; \mathbf{K})\} \tag{4.6}$$

where $\mathbf{K}_{opt}$ is the subset of BIPs that yield the *simplest* possible solution based on the *importance* of each BIP. The definitions of *simplicity* and *importance* are given as follows:

A parameter solution ($\mathbf{K}_{opt}^*$) is *simple* if the number of parameters that deviate from some reference value is as small as possible while still yielding a valid solution to the system values, i.e. $\vec{d}_j = \vec{\omega}_j$. The reference parameters ($\mathbf{K}_0$) are arbitrary, yet the solution may be dependent on the reference values. It is therefore assumed that the reference parameters are valid and near the true solution or in the case of this work approach the correct solution. The validity of the solution is described by the physical constraints of the system while the nearness of the parameters ensures the solution will converge if a numerical scheme is applied to the deviating parameters (i.e. regression parameters).

The *importance* of a parameter is somewhat harder to describe as this is specific to the algorithm used to find the solution. However, a general description can be defined by the

ability of a single parameter to *simplify* the solution. As an example, if a parameter has little effect towards bringing the regression to a sufficient description of the system, then the parameter is said to be of little *importance*.

The general idea of selective tuning is rather straightforward; find some subset of the model parameters that yield a sufficient solution while minimizing the number of parameters that need to be tuned. The proposed method for estimating which parameters yield the simplest possible solution is by defining a metric for the parameter importance. Examples of possible metrics have been defined in sections 4.4.1, 4.5.1 and 4.5.2.

An iterative method is considered in tis work, where (1) the local importance of each BIP is found, then (2) a regression is applied to the n parameters with the highest importance, followed by (3) estimating the fit of the current set of parameters. If the fit of the regression made in step (3) is within the threshold, then the selective tuning process is finished. If the fit of the regression in step (3) is not within some threshold then revert back to step (1). This iterative version of the selective tuning approach is applied in this work. The methods used for estimating the local importance are the same as mentioned in the previous paragraph; the gradient methods, the PCC method and the random forest method. The relationship between this local version of the selective tuning process and the more general selective tuning, described in the beginning of this section, is based on the assumption that the local search tends to move the reference value ($\mathbf{K}_0$) closer to the true solution in an iterative manner.

4.4 Derivative Based Methods

In the following section a derivative based method is given that aims to describe the local importance of the BIPs based on the methodology stated in section 4.3. Specifically, the following section describes a methodology based on the gradient of the cost function to estimate the local parameter importance.

4.4.1 Gradient Method Approach

Similar to the method described in section 2.3.2, an estimate of the gradient of the cost function with respect to the BIPs can be calculated using a finite difference approximation with respect to an initial BIP matrix ($\mathbf{K}$) and a small perturbation for BIP shown by

$$\nabla C(\mathbf{K}) = \{\frac{\partial C}{\partial k_{ij}}\}_{i \neq j} \approx \{\frac{C(k_{ij} + \Delta k) - C(k_{ij} - \Delta k)}{2\Delta k}\}_{i \neq j} \qquad (4.7)$$

In addition to the gradient, a partial derivative matrix (∂C) is defined by

$$\partial C = \begin{bmatrix} 0 & \partial_{12}C & \partial_{13}C & \dots \\ \partial_{21}C & 0 & \partial_{23}C & \dots \\ \partial_{31}C & \partial_{32}C & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \qquad (4.8)$$

where $\partial_{ij}C$ is the partial derivative of the cost function with respect to k_{ij} as shown by

$$\partial_{ij}C = \frac{\partial C}{\partial k_{ij}} \tag{4.9}$$

As described in section 4.5 equation (4.17), one possible method for estimating the *importance* (ρ_{ij}) of BIP k_{ij} can be calculated by the partial derivative of the cost function with respect to k_{ij} divided by the 2-norm[2] of the cost function shown by

$$\rho_{ij}^{GD} = \frac{\partial_{ij}C}{||\nabla C||_2} \tag{4.10}$$

It is therefore possible to show that ∂C is proportional to the importance matrix (i.e. $\mathbf{\Sigma} \propto \partial C$).

4.5   Perturbation Approach

Unlike the gradient method of section 4.4.1, the perturbation methods described below are non-derivative based and belongs to the field of statistics. The aim of the following perturbation methods is to rank the importance of the current set of BIPs and extract the simplest set of parameters for the selective tuning process defined in section 4.3.

The main idea of the perturbation method is that, given a current set of BIPs ($\mathbf{K}$), an additional perturbation matrix ($\mathbf{\Delta}$) is added, and then the cost function (C_λ) is calculated at the perturbated set of BIPs as shown by

$$C_\lambda = C(\mathbf{K} + \mathbf{\Delta}_\lambda) \tag{4.11}$$

where

$$\mathbf{\Delta}_\lambda = \begin{bmatrix} 0 & \delta_{12}^{(\lambda)} & \delta_{13}^{(\lambda)} & \cdots \\ \delta_{21}^{(\lambda)} & 0 & \delta_{23}^{(\lambda)} & \cdots \\ \delta_{31}^{(\lambda)} & \delta_{32}^{(\lambda)} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{4.12}$$

and $\delta_{ij}^{(\lambda)}$ is some small perturbation with a given distribution (e.g. $\delta_{ij}^{(\lambda)} \sim \mathcal{N}(0, \sigma)$). Different sampling approaches may also be applied for the perturbations (e.g. latin hypercube), however these approaches were not implemented in this work. In equations (4.11) and (4.12) λ is the counting variable for each perturbation in the range $\lambda \in \{1, 2, ..., N_\lambda\}$ where N_λ is the total number of perturbations. In this work the perturbation deviation ($\delta^{(\lambda)}$) will be a uniform distribution within some range ($\pm|\Delta k|$)

---

[2]For an example of how to calculate the 2-norm see appendix D section D.5

Estimating the magnitude of the difference in cost function between the initial BIP matrix ($\mathbf{K}$) and the perturbed BIP matrix ($\mathbf{K} + \boldsymbol{\Delta}_\lambda$) for the total number of perturbations (N_λ) builds the database of change in cost as shown by

$$|\Delta \vec{C}| = \{|\Delta C_1|, |\Delta C_2|, ..., |\Delta C_{N_\lambda}|\} \tag{4.13}$$

where

$$|\Delta C_\lambda| = |C(\mathbf{K} + \boldsymbol{\Delta}_\lambda) - C(\mathbf{K})| \tag{4.14}$$

For convenience the set of perturbations ($\vec{\delta}_{ij}$) between components i and j is also stored in the following way

$$\vec{k}_{ij} = \{k_{ij} + \delta_{ij}^{(1)}, k_{ij} + \delta_{ij}^{(2)}, ..., k_{ij} + \delta_{ij}^{(N_\lambda)}\} \tag{4.15}$$

$$\vec{\delta}_{ij} = \{\delta_{ij}^{(1)}, \delta_{ij}^{(2)}, ..., \delta_{ij}^{(N_\lambda)}\} \tag{4.16}$$

Based on the change in magnitude of the cost function value for each perturbation ($|\Delta \vec{C}|$) as well as the size and magnitude of the perturbation ($\vec{\delta}_{ij}$) an *importance matrix* ($\boldsymbol{\Sigma}$) can be defined as

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0 & \rho_{12} & \rho_{13} & \cdots \\ \rho_{21} & 0 & \rho_{23} & \cdots \\ \rho_{31} & \rho_{32} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{4.17}$$

where ρ_{ij} is the local *importance* for the BIP k_{ij}.

4.5.1   Perturbation Methodology Using PCC

One method for estimating the local *importance* (ρ_{ij}) for BIP k_{ij} is by calculating the Pearson correlation coefficients as shown by

$$\rho_{ij}^{PCC} = \rho^{PCC}(\vec{k}_{ij}, \Delta \vec{C}) = \frac{\sigma(\vec{k}_{ij}, \Delta \vec{C})}{\sigma(\vec{k}_{ij}) \cdot \sigma(\Delta \vec{C})} \tag{4.18}$$

as described in section 2.3.4. Because the perturbation is local, it is possible to approximate the relationship between the two values of the PCC calculation as linear. This makes the effect of the PCC methodology useful to estimate the local *importance* of each BIP. Having defined the change in the cost function based only on the magnitude, the sign of

the PCC calculation yields a ranking of the most important set of BIPs and the direction which a regression technique should search similar to the gradient method. In general, for $\rho_{ij}^{PCC} > 0$ then the value of the BIP k_{ij} should decrease to yield a reduced cost function. Similarly, for $\rho_{ij}^{PCC} < 0$ then the value of the BIP k_{ij} should increase to yield a reduced cost function. An example of the PCC method is shown in section 2.3.4.

4.5.2 Perturbation Methodology Using Trained Random Forest

Another potential method to estimate the *importance* (ρ_{ij}) for k_{ij} is by using the set of perturbations ($\vec{\delta}_{ij}$) and the change in cost function ($\Delta\vec{C}$) to train a random forest as described in section 2.3.5. One feature of the random forest method is that it allows for an estimate of the *feature importance* for each parameter (see section 2.3.5 for more details). Having a trained random forest ($\hat{h}(\delta)$) will therefore be able to estimate the importance of each BIP shown by

$$\rho_{ij}^{RF} = \rho^{RF}(\vec{k}_{ij}, \Delta\vec{C}) = FI(\hat{h}(\delta)) \tag{4.19}$$

where FI is the feature importance of the trained random forest $\hat{h}(\delta)$ for the set of perturbations (δ) defined by

$$\delta = \{\vec{\delta}_{12}, \vec{\delta}_{13}, ..., \vec{\delta}_{21}, \vec{\delta}_{23}, ...\} \tag{4.20}$$

4.6 Results

In this section the selective tuning approach is applied to two sets of fluid mixtures. First, a description of the fluid systems will be given, followed by the specific approaches of the selective tuning method used in this work. A base case is defined using PhazeComp's solver on the entire set of BIPs. Following the base case results, the results for the three proposed approaches to the importance ranking are described for the first iteration, followed by the results of the regression scheme.

4.6.1 Case Description

In the following sections, the iterative selective tuning process was carried out for two cases. The first case (case 1) is defined by a mixture containing SCN components C_2, C_5, C_7 and C_{10}. The EOS used for case 1 is given in Table 4.1 and 4.2. The initial composition used was uniform and the temperature at which the experiments were carried out at 100 °F. The initial BIP matrix is given in Table 4.3 and contains zero BIPs with a bound on the regression variables between 0 and 0.2.

Table 4.1: Synthetic component properties, molecular weights and acentric factors for simple system (case 1).

Synthetic Component Properties					
Name	MW	Tc	Pc	AF	Vc
		(^{o}F)	(psia)		(m^3/kmol)
C2	30.07	89.91	706.6	0.099	0.146
C5	70.91	386.6	503.2	0.231	0.306
C7	95.63	529.0	455.3	0.274	0.387
C10	135.1	678.9	359.9	0.392	0.537

Table 4.2: Synthetic BIP matrix for simple system (case 1).

Synthetic BIP Matrix				
	C2	C5	C7	C10
C2	0	0.0076	0.0131	0.0232
C5	0.0076	0	0.0008	0.0044
C7	0.0131	0.0008	0	0.0015
C10	0.0232	0.0044	0.0015	0

The initial values for the BIP matrix was set to zero BIPs as shown in Table 4.3 and the bounds on the PhazeComp regression was set to 0 and 0.5 for all BIPs.

Table 4.3: Initial BIP matrix for simple system (case 1).

Initial BIP Matrix				
	C2	C5	C7	C10
C2	0	0.0000	0.0000	0.0000
C5	0.0000	0	0.0000	0.0000
C7	0.0000	0.0000	0	0.0000
C10	0.0000	0.0000	0.0000	0

The second case (case 2) is defined by a mixture containing SCN components C_1 to C_{10}. The EOS used for case 2 is given in Table 4.4 and 4.5. A uniform molar composition (equal molar amount) was used, and the experiments were carried out at a fixed temperature of 100 °F. The initial BIP matrix is given in Table 4.6 and contains zero BIPs with a bound on the regression variables between 0 and 0.2.

Table 4.4: Synthetic component properties, molecular weights and acentric factors for intermediate system (case 2).

Component Properties					
Name	**MW**	**Tc**	**Pc**	**AF**	**Vc**
		(°F)	(psia)		(m³/kmol)
C1	16.04	-116.7	667.0	0.011	0.099
C2	30.07	89.91	706.6	0.099	0.146
C3	44.10	206.0	616.1	0.152	0.200
C4	58.12	305.0	544.9	0.203	0.258
C5	70.91	386.6	503.2	0.231	0.306
C6	82.42	464.4	490.0	0.240	0.341
C7	95.63	529.0	455.3	0.274	0.387
C8	109.2	585.2	420.4	0.312	0.435
C9	122.2	634.5	388.1	0.352	0.485
C10	135.1	678.9	359.9	0.392	0.537

Table 4.5: Synthetic BIP matrix for complex system (case 2).

Synthetic BIP Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	**0**	0.0021	0.0069	0.0127	0.0175	0.0211	0.0254	0.0298	0.0343	0.0386
C2	0.0021	**0**	0.0014	0.0045	0.0076	0.0100	0.0131	0.0164	0.0198	0.0232
C3	0.0069	0.0014	**0**	0.0009	0.0025	0.0040	0.0060	0.0083	0.0108	0.0134
C4	0.0127	0.0045	0.0009	**0**	0.0004	0.0011	0.0023	0.0038	0.0056	0.0074
C5	0.0175	0.0076	0.0025	0.0004	**0**	0.0002	0.0008	0.0017	0.0030	0.0044
C6	0.0211	0.0100	0.0040	0.0011	0.0002	**0**	0.0002	0.0008	0.0017	0.0028
C7	0.0254	0.0131	0.0060	0.0023	0.0008	0.0002	**0**	0.0002	0.0007	0.0015
C8	0.0298	0.0164	0.0083	0.0038	0.0017	0.0008	0.0002	**0**	0.0002	0.0006
C9	0.0343	0.0198	0.0108	0.0056	0.0030	0.0017	0.0007	0.0002	**0**	0.0001
C10	0.0386	0.0232	0.0134	0.0074	0.0044	0.0028	0.0015	0.0006	0.0001	**0**

The initial values for the BIP matrix was set to zero BIPs as shown in Table 4.6 and the bounds on the PhazeComp regression was set to 0 and 0.5 for all BIPs.

Table 4.6: Initial BIP matrix for complex system (case 2).

Initial BIP Matrix									
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C2	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C3	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C4	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C5	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0000
C6	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000
C7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000
C8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000
C9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000
C10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0**

The methodology of selective tuning is similar for both cases and is given below:

1. Define the number of regression parameters (n), and maximum number of iterations (N_{max}).

2. Define the current BIP matrix.

3. Calculate the importance matrix for the current BIP matrix based on the specific method for estimation compared to the experimental data (Ω).

4. Determine the n most important BIPs from the importance matrix.

5. Regress the n most important BIPs based on the observed data and the cost function (RMS).

6. Calculate the total RMS after regression based on the observed data and the calculated data.

7. Either (a) the error is less than the threshold error (error $\leq \epsilon$), then the process is finished or (b) the number of iterations is equal to the maximum number of iterations, then the process is finished or (c) extract the regressed BIP matrix, set the current BIP matrix equal to BIP matrix after regression and continue to step 3.

To avoid getting stuck in a loop if the same subset of BIPs are chosen twice in a row, an additional heuristic was implemented. If this situation occurs, the number of BIPs in the subset (n) will be increased by one but restricted to be within some upper bound. The BIPs chosen for the subset are chosen from the ranking of importance from the selective tuning approach. As an example, take a mixture containing SCN components C_1 to C_{15}. The selective tuning approach might find that the BIP between C_1-C_{15} and C_1-C_{14} are ranked to the two most important BIPs. If this is the case for two iterations in a row, the third most important BIP will also be included in the subset of BIPs used in the regression (given that three BIPs are under the upper limit of allowed BIPs).

The basis for the proposed methodology was developed based on the traditional structure of EOS tuning described by Yonous et al. in [46]. There is also some precedence in the field of numerical methods to the proposed methodology. From the point of view of numerical methods, the proposed approach resembles a combination of the gradient descent method (see section 2.3.2) and a modified Powell's method (see section 2.3.3). However, the proposed method in its current form is, to the current knowledge of the author, a novel approach in the field of hydrocarbon fluid modeling.

For all the cases described in the following sections, the maximum number of iterations was set to 14 (starting at 0, 1, ... where step zero is the initial guess). The number of parameters that were tuned varies based on the case number (i.e. case 1 or 2) and is specified in the relevant sections.

4.6.2 Results - Base Case

The PVT software used in this work (PhazeComp) has an advanced regression solver. This solver will be used as a base case solution to the problem of BIP tuning. For both fluid model cases the result of a regression scheme where all the BIPs is used in the regression are given in Table 4.7 and 4.8. For reference, the time used by PhazeComp to solve the regression was recorded to be 0.31 s (case 1) and 3.05 s (case 2) respectively.

Table 4.7: BIP matrix solution for simple system (case 1) using Newton's method.

Final BIP Matrix				
	C2	C5	C7	C10
C2	0	0.0076	0.0131	0.0232
C5	0.0076	0	0.0008	0.0044
C7	0.0131	0.0008	0	0.0015
C10	0.0232	0.0044	0.0015	0

Table 4.8: BIP matrix solution for complex system (case 1) using Newtons method.

Final BIP Matrix										
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.0021	0.0069	0.0127	0.0175	0.0211	0.0254	0.0298	0.0343	0.0386
C2	0.0021	0	0.0014	0.0045	0.0076	0.0100	0.0131	0.0164	0.0198	0.0232
C3	0.0069	0.0014	0	0.0009	0.0025	0.0040	0.0060	0.0083	0.0108	0.0134
C4	0.0127	0.0045	0.0009	0	0.0004	0.0011	0.0023	0.0038	0.0056	0.0074
C5	0.0175	0.0076	0.0025	0.0004	0	0.0002	0.0008	0.0017	0.0030	0.0044
C6	0.0211	0.0100	0.0040	0.0011	0.0002	0	0.0002	0.0008	0.0017	0.0028
C7	0.0254	0.0131	0.0060	0.0023	0.0008	0.0002	0	0.0002	0.0007	0.0015
C8	0.0298	0.0164	0.0083	0.0038	0.0017	0.0008	0.0002	0	0.0002	0.0006
C9	0.0343	0.0198	0.0108	0.0056	0.0030	0.0017	0.0007	0.0002	0	0.0001
C10	0.0386	0.0232	0.0134	0.0074	0.0044	0.0028	0.0015	0.0006	0.0001	0

To estimate the increase in computational time for PhazeComp's regression algorithm with respect to increased number of components, a range of mixtures were tested containing SCN components ranging from C_1 to C_{20}. Starting from a mixture containing 5 SCN components ranging from C_1 to C_5 in an increasing manner (i.e. C_1, C_2, C_3, C_4 and C_5) represented by n=5, the average time for five runs was recorded versus the number of components. Adding a heavier component for each step (e.g. first C_1 to C_5 as n=5 then C_1 to C_6 as n=6 etc.) the same procedure was carried out for mixtures containing SCN components up to C_1 to C_{20} ($n = 20$). The plot of regression time versus the number of SCN components is shown in Figure 4.1 together with an exponential trend-line. By extrapolating the trend-line to a 40 and 50 component system, the time needed to solve the regression scheme with PhazeComp is estimated to take roughly 13 and 846 days respectively. The computational time for this size system is unfeasible for any regression scheme.
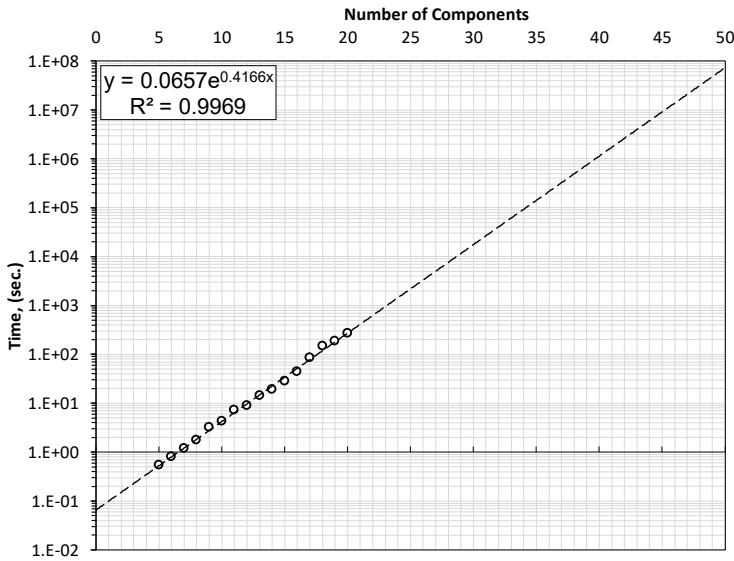
Figure 4.1: Plot displaying the average time for PhazeComp regression scheme for increasing number of SCN mixtures starting from a mixture of SCN compositions containing C_1 to C_5 (n=5) up towards C_1-C_{20} (n=20) as well as an exponential trend-line.

The reason for the exponential behavior can be attributed to a combination of the regression method used by PhazeComp, which is assumed to be a variation of a multivariate Newton's method, and the squared increase in the number of parameters in the regression scheme given by [3]

$$N_{bips} = \frac{N_{comps} \cdot (N_{comps} - 1)}{2} \qquad (4.21)$$

---

[3]A simple and visual derivation of this equation is given in Appendix D section D.4.

where N_{comps} is the number of components and N_{bips} is the number of BIPs.

4.6.3   Results - Gradient Method

Following the methodology described in section 4.6.1, application of the gradient method to estimate the local importance of every BIP for each iteration of the selective tuning was calculated. Examples of the importance matrix for the first iteration of the selective tuning process is shown for case 1 and 2 in Table 4.9 and 4.10.

Table 4.9: BIP importance matrix for the simple system (case 1) using gradient method for the first iteration.

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.0010	-0.0011	-0.8126
C5	-0.0010	0	-0.0014	-0.5424
C7	-0.0011	-0.0014	0	-0.2133
C10	-0.8126	-0.5424	-0.2133	0

Table 4.10: BIP importance matrix for the intermediate system (case 2) using the gradient method.

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.1179	-0.1344	-0.1370	-0.1370	-0.1447	-0.2244	-0.2114	-0.3130	-0.2955
C2	-0.1179	0	0.0000	0.0000	-0.0126	-0.0347	-0.0484	-0.0494	-0.1372	-0.1508
C3	-0.1344	0.0000	0	-0.0030	-0.0036	-0.0098	-0.0716	-0.0317	-0.0271	-0.1257
C4	-0.1370	0.0000	-0.0030	0	0.0312	0.0300	-0.0174	0.0253	-0.0569	-0.0743
C5	-0.1370	-0.0126	-0.0036	0.0312	0	0.0224	0.0168	0.0464	0.0702	0.0397
C6	-0.1447	-0.0347	-0.0098	0.0300	0.0224	0	0.0028	0.0858	0.1124	0.0767
C7	-0.2244	-0.0484	-0.0716	-0.0174	0.0168	0.0028	0	0.1310	0.1722	0.1732
C8	-0.2114	-0.0494	-0.0317	0.0253	0.0464	0.0858	0.1310	0	0.3248	0.3621
C9	-0.3130	-0.1372	-0.0271	-0.0569	0.0702	0.1124	0.1722	0.3248	0	0.4453
C10	-0.2955	-0.1508	-0.1257	-0.0743	0.0397	0.0767	0.1732	0.3621	0.4453	0

Based of the importance matrix, the BIPs with the largest magnitude of estimated importance were used in a PhazeComp regression run. The number of BIPs chosen for regression is specified beforehand and in the result shown in this work the range of important parameters is set to 2-4 for case 1 and 10-20 for case 2. The result of 14 iterations of the selective tuning approach is shown in Table 4.11 and 4.12.

Table 4.11: Final BIP matrix for the simple system (case 1) using the gradient method for the first iteration.

Final BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	**0**	0.0075	0.0132	0.0232
C5	0.0075	**0**	0.0008	0.0044
C7	0.0132	0.0008	**0**	0.0015
C10	0.0232	0.0044	0.0015	**0**

Table 4.12: Final BIP matrix for the intermediate system (case 2) using the gradient method.

Final BIP Matrix									
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	**0**	0.0160	0.0000	0.0041	0.0163	0.0169	0.0267	0.0368	0.0389	0.0310
C2	0.0160	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0628	0.0176	0.0016
C3	0.0000	0.0000	**0**	0.0000	0.0000	0.0011	0.0006	0.0001	0.0353	0.0000
C4	0.0041	0.0000	0.0000	**0**	0.0000	0.0008	0.0022	0.0015	0.0004	0.0133
C5	0.0163	0.0000	0.0000	0.0000	**0**	0.0000	0.0056	0.0021	0.0000	0.0044
C6	0.0169	0.0000	0.0011	0.0008	0.0000	**0**	0.0012	0.0017	0.0011	0.0020
C7	0.0267	0.0000	0.0006	0.0022	0.0056	0.0012	**0**	0.0001	0.0005	0.0007
C8	0.0368	0.0628	0.0001	0.0015	0.0021	0.0017	0.0001	**0**	0.0001	0.0000
C9	0.0389	0.0176	0.0353	0.0004	0.0000	0.0011	0.0005	0.0001	**0**	4.3E-05
C10	0.0310	0.0016	0.0000	0.0133	0.0044	0.0020	0.0007	0.0000	4.3E-05	**0**

A summary plot of the RMS versus the selective tuning iteration number is given (black solid line) in Figure 4.2 for case 1 and 2 respectively together with a random selection of important parameters (black dashed line) given as a reference.
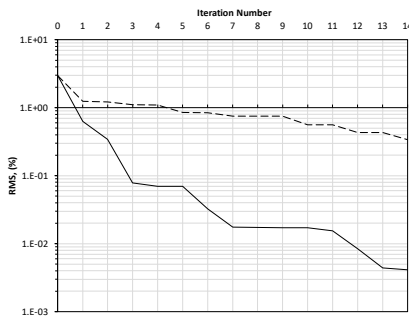
(a) Case 1 RMS results ($n = 100$).
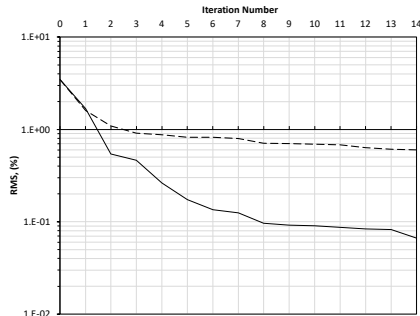
(b) Case 2 RMS results ($n = 100$).

Figure 4.2: Results for RMS versus selective tuning iteration number for case 1 (left) and 2 (right) for the gradient method (solid black line) and a random choice of important BIPs (dashed black line). (a) Case 1 mixture with 2 to 4 important BIPs while (b) case 2 mixture with 10 to 20 important BIPs.

As the simultaneous regression approach was shown not to be feasible with respect to run-time, an estimate of the run-time for the gradient based method with respect to the number of components of the system should also be performed. The results of SCN components ranging from C_1-C_5 to C_1-C_{24} for estimating the importance matrix of a single iteration are shown in Figure 4.3.
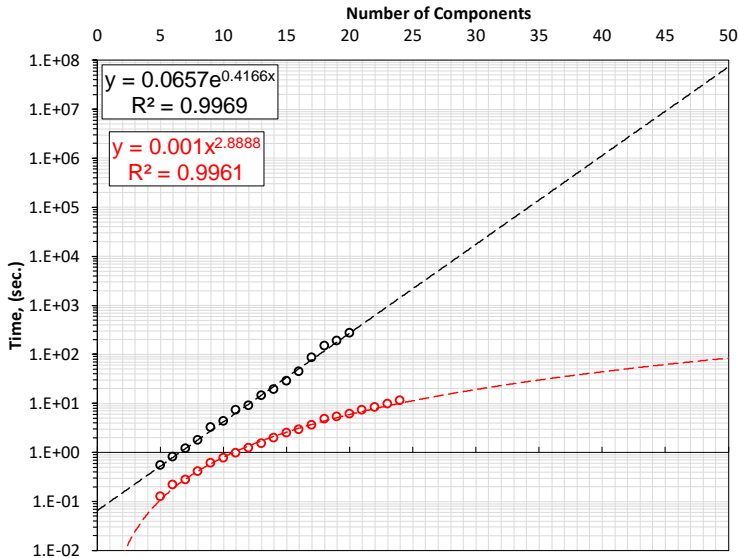
Figure 4.3: Plot displaying the average time for PhazeComp regression scheme (black points) and time to estimate importance using the gradient method (red points) for increasing number of SCN mixtures starting from C_1 to C_5 (n=5) up toward C_1-C_{20} (n=20) as well as trend-lines for both cases.

Assuming that the data will follow the power-law trendline displayed in Figure 4.3 the estimated time to calculate the importance matrix for a 40 and 50 component system for a single iteration of the selective tuning method is roughly 42.5 and 80.9 seconds respectively. This is significantly less than running the full regression scheme, however the number of iterations are assumed to range between 10 and 100, so the total regression time will range from 400 to 8000 seconds. Because the number of parameters used in the regression is relatively small, the time used for the regression can be neglected (e.g. 100 parameter regression takes between 10 to 100 seconds which is the most sub-parameters to be used). Furthermore, each iteration in the selective tuning method is possible to run in parallel in the sense that the 40-80 second range of each iteration can be run in parallel, reducing the time per iteration by a factor equal to the number of cores used to run in parallel. Assuming a range of cores between 5 (e.g. a personal computer) and 100 (e.g. a cluster) then it is estimated that the total regression time will range between 4 and 1600 seconds.

However, if the trendline does not hold for more complex systems and the method turn exponential assuming the same slope at $n = 24$, then the time per iteration will increase

significantly. An estimate using the slope at $n = 24$ yields a run-time of 1000 seconds for a single iteration for a 50-component system. With the number of iterations ranging from 10 to 100 the range of time for the entire selective tuning process will be between 1E4 and 1E5 seconds. Assuming that the process is run in parallel with a range of cores between 5 and 100 then the estimated total time is between 100 and 20,000 seconds or about 2 minutes to 5 and a half hours. This time range is still far below the simultaneous approach and is also within the realm of reasonable time spent on EOS fluid modeling.

4.6.4 Results - PCC Method

Following the methodology described in section 4.6.1, application of the PCC method to estimate the local importance of every BIP for each iteration of the selective tuning was calculated. Examples of the importance matrix for the first iteration of the selective tuning procedure is shown for case 1 and 2 in Table 4.13 and 4.14.

Table 4.13: BIP importance matrix for the simple system (case 1) using the PCC method.

Importance Matrix				
	C2	C5	C7	C10
C2	0	0.1079	-0.0347	-0.7872
C5	0.1079	0	0.1485	-0.5671
C7	-0.0347	0.1485	0	-0.2574
C10	-0.7872	-0.5671	-0.2574	0

Table 4.14: BIP importance matrix for the intermediate system (case 2) using the PCC method for the first iteration.

Importance Matrix										
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.1160	-0.1015	-0.2409	-0.1338	-0.3565	-0.1537	-0.2268	-0.2193	-0.3968
C2	-0.1160	0	0.0280	-0.0062	-0.0917	0.0276	-0.1539	-0.1219	-0.0157	-0.2636
C3	-0.1015	0.0280	0	-0.0273	0.0290	-0.2215	0.0621	0.0668	0.2037	-0.0880
C4	-0.2409	-0.0062	-0.0273	0	0.0111	0.0178	-0.0859	-0.0351	-0.2807	-0.0402
C5	-0.1338	-0.0917	0.0290	0.0111	0	0.1358	0.0919	0.1637	0.0209	0.1006
C6	-0.3565	0.0276	-0.2215	0.0178	0.1358	0	0.0347	0.0983	0.2158	-0.0058
C7	-0.1537	-0.1539	0.0621	-0.0859	0.0919	0.0347	0	0.1864	0.3138	0.1655
C8	-0.2268	-0.1219	0.0668	-0.0351	0.1637	0.0983	0.1864	0	0.2442	0.2122
C9	-0.2193	-0.0157	0.2037	-0.2807	0.0209	0.2158	0.3138	0.2442	0	0.4525
C10	-0.3968	-0.2636	-0.0880	-0.0402	0.1006	-0.0058	0.1655	0.2122	0.4525	0

Based of the importance matrix, the BIPs with the largest magnitude of the estimated importance were used in a PhazeComp regression run. The number of BIPs chosen for regression is specified beforehand and in the result shown in this work the range of important parameters is set to 2-4 for case 1 and 10-20 for case 2. The result of 14 iterations of

the selective tuning approach is shown in Table 4.15 and 4.16.

Table 4.15: Final BIP matrix for simple system (case 1) using the PCC method for the first iteration.

Final BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	**0**	0.0062	0.0135	0.0232
C5	0.0062	**0**	0.0000	0.0044
C7	0.0135	0.0000	**0**	0.0015
C10	0.0232	0.0044	0.0015	**0**

Table 4.16: Final BIP matrix for complex system (case 2) using the PCC method.

Final BIP Matrix									
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	**0**	0.0000	0.0000	0.0272	0.0378	0.0000	0.0374	0.0290	0.0263	0.0390
C2	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0144	0.0721	0.0066	0.0000
C3	0.0000	0.0000	**0**	0.0000	0.0193	0.0000	0.0060	0.0001	0.0000	0.0286
C4	0.0272	0.0000	0.0000	**0**	0.0000	0.0000	0.0083	0.0000	0.0135	0.0015
C5	0.0378	0.0000	0.0193	0.0000	**0**	0.0000	0.0065	0.0000	0.0000	0.0063
C6	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0008	0.0004	0.0000	0.0050
C7	0.0374	0.0144	0.0060	0.0083	0.0065	0.0008	**0**	0.0001	0.0002	0.0020
C8	0.0290	0.0721	0.0001	0.0000	0.0000	0.0004	0.0001	**0**	0.0012	0.0008
C9	0.0263	0.0066	0.0000	0.0135	0.0000	0.0000	0.0002	0.0012	**0**	0.0009
C10	0.0390	0.0000	0.0286	0.0015	0.0063	0.0050	0.0020	0.0008	0.0009	**0**

A summary plot of the RMS versus the selective tuning iteration number is given (black solid line) in Figure 4.4 for case 1 and 2 respectively together with a random selection of important parameters (black dashed line) given as a reference.

However, unlike the gradient method for estimating the importance matrix, the PCC method uses a dataset with n datapoints to estimate the importance. These datapoints were generated by perturbing around the current BIP matrix with a uniform distribution and a bound in the range $k_{ij} - dk < k_{ij} + \delta_{ij} < k_{ij} + dk$. For all cases in this work, the optimum value of dk was found to be 0.00001. The reasoning for the choice of dk was based on the assumption that as $dk \rightarrow 0$ then the correlation found by PCC will be linear. The lower bound for dk was found by trial and error where at smaller values the machine accuracy had a significant effect on the results.

A convergence study was also made for the number of perturbations (n) needed to yield representative solution for case 2. The results are given in appendix C in section C.2. The value used in the results above were calculated using $n = 100$ which, from the convergence study, is within the range of convergence. However, it is important to note that the number of datapoints needed increases with the iteration number and for larger iteration

(a) Case 1 RMS results (n=100).

(b) Case 2 RMS results (n=100).
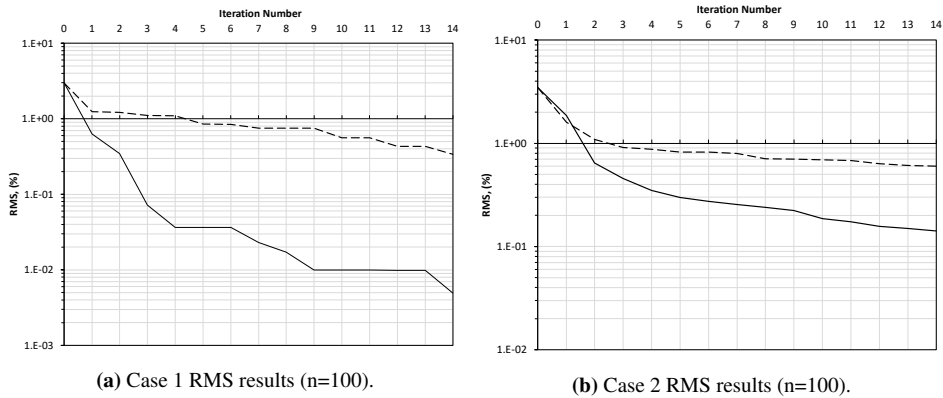
Figure 4.4: Results for RMS versus selective tuning iteration number for case 1 (left) and 2 (right) for the PCC method (solid black line) and a random choice of important BIPs (dashed black line). (a) Case 1 mixture with 2 to 4 important BIPs while (b) case 2 mixture with 10 to 20 important BIPs.

numbers, the deviation is more scattered for all cases in the convergence study (e.g. a bad choice in an early iteration will lead to a significant deviation at the later iterations).

4.6.5   Results - Random Forest Method

Following the methodology described in section 4.6.1, application of the random forest method to estimate the local importance of every BIP for each iteration of the selective tuning. Examples of the importance matrix for the first iteration of the selective tuning procedure is shown for case 1 and 2 in Table 4.17 and 4.18.

Table 4.17: BIP importance matrix for the simple system (case 1) using random forest method.

Importance Matrix				
	C2	**C5**	**C7**	**C10**
C2	**0**	0.0019	0.0016	0.0024
C5	0.0019	**0**	0.0021	0.3153
C7	0.0016	0.0021	**0**	0.6766
C10	0.0024	0.3153	0.6766	**0**

Based on the importance matrix, the BIPs with the largest magnitude of the estimated importance were used in a PhazeComp regression run. The number of BIPs chosen for regression is specified beforehand and in the result shown in this work the range of important parameters are set to 2-4 for case 1 and 10-20 for case 2. The result of 14 iterations of the selective tuning approach is shown in Table 4.19 and 4.20.

Table 4.18: BIP importance matrix for the intermediate system (case 2) using random forest method for the first iteration.

Importance Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	**0**	0.0095	0.0085	0.0105	0.0081	0.0195	0.0104	0.0109	0.0265	0.0258
C2	0.0095	**0**	0.0094	0.0078	0.0077	0.0091	0.0084	0.0098	0.0060	0.0074
C3	0.0085	0.0094	**0**	0.0120	0.0090	0.0090	0.0096	0.0087	0.0086	0.0178
C4	0.0105	0.0078	0.0120	**0**	0.0073	0.0076	0.0100	0.0123	0.0109	0.0149
C5	0.0081	0.0077	0.0090	0.0073	**0**	0.0128	0.0155	0.0253	0.0507	0.0881
C6	0.0195	0.0091	0.0090	0.0076	0.0128	**0**	0.0423	0.0349	0.0482	0.0689
C7	0.0104	0.0084	0.0096	0.0100	0.0155	0.0423	**0**	0.0748	0.0408	0.0345
C8	0.0109	0.0098	0.0087	0.0123	0.0253	0.0349	0.0748	**0**	0.0726	0.0492
C9	0.0265	0.0060	0.0086	0.0109	0.0507	0.0482	0.0408	0.0726	**0**	0.0086
C10	0.0258	0.0074	0.0178	0.0149	0.0881	0.0689	0.0345	0.0492	0.0086	**0**

Table 4.19: Final BIP matrix for simple system (case 1) using the random forest method for the first iteration.

Final BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	**0**	0.0060	0.0141	0.0232
C5	0.0060	**0**	0.0011	0.0044
C7	0.0141	0.0011	**0**	0.0015
C10	0.0232	0.0044	0.0015	**0**

Table 4.20: Final BIP matrix for complex system (case 2) using the random forest method.

Final BIP Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0989
C2	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C3	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
C4	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0063	0.0000
C5	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0000	0.0099
C6	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0000	0.0052
C7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000	0.0004
C8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	**0**	0.0000	0.0000
C9	0.0000	0.0000	0.0000	0.0063	0.0000	0.0000	0.0000	0.0000	**0**	0.0000
C10	0.0989	0.0000	0.0000	0.0000	0.0099	0.0052	0.0004	0.0000	0.0000	**0**

A summary plot of the RMS versus the selective tuning iteration number is given (black solid line) in Figure 4.5 for case 1 and 2 respectively together with a random selection of important parameters (black dashed line) given as a reference.

(a) Case 1 RMS results (n=100).

(b) Case 2 RMS results (n=100).
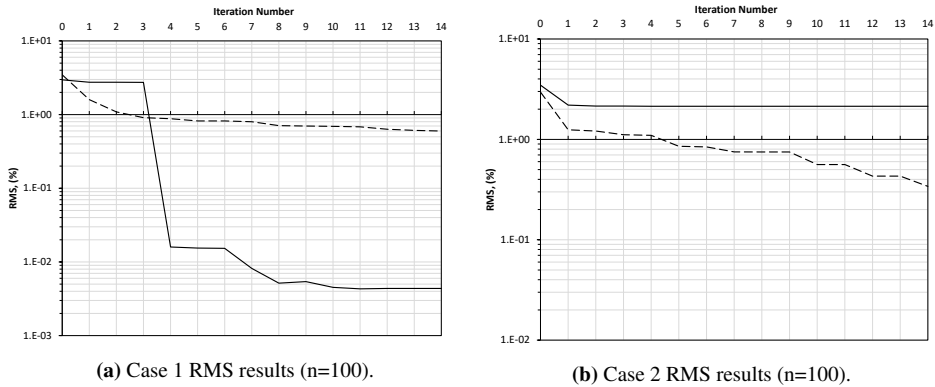
Figure 4.5: Results for RMS versus selective tuning iteration number for case 1 (left) and 2 (right) for the random forest method (solid black line) and a random choice of important BIPs (dashed black line). (a) Case 1 mixture with 2 to 4 important BIPs while (b) case 2 mixture with 10 to 20 important BIPs.

Similar to the PCC method, the random forest approach uses a dataset with n datapoints to estimate the importance. These datapoints were generated by perturbing around the current BIP matrix with a uniform distribution and a bound in the range $k_{ij} - dk < k_{ij} + \delta_{ij} < k_{ij} + dk$. For all cases in this work, the value of dk was set to 0.05 as one of the drawbacks of the random forest routine is that it does not correlate linear relationships well. This method is therefore the only approach which generates a more global than local estimation for the importance.

4.6.6 More Results

The results for each iteration of the gradient method and the first several iterations for the PCC and RF methods are summarized in appendix C section C.1. These results contain the following information for each iteration:

- Importance matrix
- BIP Matrix
- Final K-value data

5

Discussion

"Unforeseen surprises are the rule in science, not the exception. Remember: Stuff happens."

- Leonard Susskind

5.1   Discussion - Effects of BIPs on Phase Behavior

The results for the binary phase behavior given in section 3.1.1 and appendix A shows a similar trend in all cases. The difference in the phase behavior by introducing BIPs is divided into two regions. Region 1 is the temperature range near the light component critical temperature. There is a shift in the mixture's critical pressure depending (a) the sign of the BIP and (b) the magnitude of the BIP. The results for region 1 show a decrease in critical pressure for a positive BIP with a more dramatic shift for an increased magnitude of the BIP. Conversely, an increase in the mixture critical pressure is shown to be related to a negative BIP with an increased shift correlated to an increased magnitude of the BIP.

The second temperature region yields the opposite behavior of region 1, but is constrained to the same two variables as region 1 (i.e. BIP sign and magnitude). For a positive BIP in region 2 the critical pressure is increased, where the magnitude of the shift is related to the magnitude of the BIP. A negative BIP will have the opposite effect of a positive BIP, resulting in a decrease in the mixture critical pressure. The magnitude of the critical pressure shift is dependent on the magnitude of the BIP.

Given behavior of changing critical locus near the light (T_{cL}) and heavy (T_{cH}) component critical temperature, the non-zero BIP critical locus will cross the zero-BIP critical locus at some temperature in the range $T_{cL} < T < T_{cH}$. The magnitude of the maximum shift in critical pressure depends on the components in the binary as can be seen in the figures in Appendix A where, in the mixture with C_5-C_{10}, the negative BIP of -0.1 has a larger shift in the critical pressure than the positive BIP of 0.1. On the other hand, for the mix-

ture containing C_5-C_{15}, the effect is opposite. The apparent dependency is that for binary pair with increasing difference in molecular weight or true boiling point, the effect of the positive BIP shifts the critical pressure more.

Appendix A shows two cases where the critical locus does not follow the trends described in this work. However, the K-value data shows that for these two cases the results of the critical locus results yield trivial solutions ($K_i = 1$) in the flash calculation. Therefore, the difference in behavior can be neglected. The approach of using the K-value data to indicate trivial solutions in the results has shown to be useful in multiple situations and is a recommended diagnostic tool for all calculations where this is possible.

The effect of introducing a BIP on the pressure dependent equilibrium ratio plots can also be related to the shift in the critical locus. As described in section 3.2.2, the convergence pressure and the critical pressure are closely related for all compositions, but for binary pairs the relationship is even more coupled. In fact, the two properties, convergence pressure and critical pressure, are the same for all binary systems. The effect of the shift of the mixture critical pressure can therefore be directly correlated with the shift in the convergence pressure. This means that the relationships described above between the critical pressure and BIP sign, magnitude and binary compositions directly apply to the equilibrium ratio plots. The effect of a shift in the convergence pressure can be thought of shifting where the K-value plot converges to a point at high-pressures.

There is a second effect that a BIP has on the K-values for low pressures. The effect of the BIP on the low-pressure region is somewhat simpler to describe than the high-pressure region. The effect for binary mixtures is simply (1) for a positive BIP the K-value plot is expanded (i.e. the light K-value is increased and the heavy component is decreased) and (2) for a negative BIP the K-value plot is contracted (i.e. the light K-value is decreased and the heavy component is increased). The effect of the expansion or contraction is dependent on (a) the components in the binary system, (b) the sign of the BIP, (c) the magnitude of the BIP and (d) whether the component is the light or heavy component of the system. A general tendency of the low-pressure K-value shift is that the heavier component shift is significantly less than the lighter component. This can be seen in Figure 3.6, 3.7 and 3.12.

An important point to note is that the vapor pressure line is independent of the introduction of the BIP. This contradicts the common relation given in equation (2.28) for low-pressures (based on Raoult and Dalton's Law). The introduction of an apparent vapor pressure ($\hat{p}_v$) can be used by extrapolating the low-pressure K-value data to where the extrapolation intersects atmospheric pressure. The effect of introducing a BIP then becomes two-fold. (1) The apparent vapor pressure is increased / decreased dependent on the sign of the BIP and (2) the critical pressure is increased / decreased dependent on the sign of the BIP while maintaining the slope of the apparent vapor pressure line.

These two points describe the behavior discussed above. The shift in the low-pressure K-values shows that the apparent vapor pressure is increased for a positive BIP and decreased for a negative BIP. However, the component critical properties (solid circles on binary PT

diagrams) dictate the apparent vapor pressure must go through these points. This forces the apparent vapor pressure curve to change slope resulting in a change in the critical locus near the light component critical temperature. This also explains why there is little to no effect near the heavy component critical temperature, as the effect of the low-pressure K-values is significantly less for the heavy component. The relationship between the apparent vapor pressure and the critical locus is therefore key for describing the driving effects of the BIP on the phase behavior of the binary system.

Increasing the complexity of the analysis by the introduction of a third component to the mixture makes any generalized description difficult. However, as the ternary system tends to a reduced binary system it is possible to extrapolate some of the findings from the binary systems for the phase behavior of a ternary system based on the three binaries. There are two key aspects of the effect of BIPs on ternary systems that yield useful results; (1) critical point shifts and (2) K-value effects.

First, for the shift in the critical point there is a clear relationship between the "boundary condition" effects (i.e. the binary critical pressure shift behavior) and the ternary mixture critical pressure shift. The effects of the binary shift propagate into the ternary system as can be seen clearly in Figure 3.10. The decreased shift from the boundary condition is dependent on the compositional distance from the other boundaries as well as the sign and magnitude of the other boundary mixture BIPs. The data shown in Appendix B section B.1 has not yet been fully analyzed more utilized to its full potential. However, the approach of displaying the critical point shift with the two-phase area difference may help further study on the topic.

For the K-value plots of the ternary system, the same basic observations hold as for the binary systems. However, for the ternary mixtures, the fact that there are two BIPs effecting each component allows for a combination of effects on the high and low-pressure region of the K-value data. The increased flexibility is great for model tuning, but makes simple trends less likely. It is clear that the two regions (low- and high-pressure) behave in a specific manner, that allows for flexible shapes of the K-value plot. Because both high and low-pressure regions are affected by BIPs, both low and high-pressure K-value data are needed when tuning an EOS to fully describe the entire phase behavior.

5.2   Discussion - BIP Tuning

Systematic, consistent, and effective BIP tuning is a problem without a clear and prag-
matic solution. Even a set of reasonable guidelines is not available for how to tune BIPs.
This is partially due to the complexity of the problem at hand, but also because of the
lack of understanding on the effect that BIPs have on phase behavior. In the previous sec-
tion, the results of simple binary and ternary systems were discussed, and even for these
simple systems the underlying cause-and-effect is hard to describe. There were several
important observations made, such as the key shifts in the K-value data: (1) Change in the
convergence pressure and (2) change in the apparent vapor pressure. With this information
available, the task of tuning the BIPs to fit K-value data now has some basic guidelines.

From Figure 4.1 it is clear that the computational time of global BIP tuning is not feasi-
ble. This work tries to develop a more pragmatic approach by bringing the computational
time down significantly, but also attaining a solution that is within the possible accuracy of
laboratory data. From the K-value results given in appendix C section C.1 it is clear that
the numerical values are within a pragmatic range of what is needed to describe the phase
behavior for both the gradient and PCC methods.

The two methods that performed best in this work were the gradient based method and the
PCC based method. The gradient based method needs more computational time to give
any results, while the PCC approach performs decently with far shorter computational
time. This is seen in the sensitivity study performed in appendix C section C.2. However,
the gradient based method always tends to perform better (RMS vs iteration number) than
the PCC approach for the same computational time. The gradient based method is slower
but more accurate, while the PCC based method is faster but less accurate.

The random forest based method did not seem to accurately predict the importance of the
different BIPs as the method is set up in this work. The fact that the random forest method
is not a local search algorithm means that it is outside of its scope in this application.
There are a wide range of possibilities for applying the random forest differently to yield
possible solutions for the problem of BIP tuning, but these approaches tend to differ from
the general approach described in this work (i.e. selective tuning).

The overall results of the selective tuning procedure for the specific fluid systems described
in this work seem to have provided a reasonable approach. However, this is only one type
of fluid model where the BIPs are positive and monotonic. Further complexities will arise
as the synthetic fluid model BIPs become less structured.

Chapter 6

Conclusion

"*To a great mind, nothing is little.*" remarked Holmes, sententiously.

- Sherlock Holmes

6.1 Conclusions - Effects of BIPs on Phase Behavior

1. A consistent trend was found for the critical-point locus between two components with increasing temperature (a) a positive BIP will decrease the critical point line with respect to the zero-BIP curve followed by a sharp increase to a higher critical pressure than the zero-BIP curve; (b) a negative BIP will increase the critical point followed by a decrease in the critical point line to a lower pressure than the zero-BIP case.

2. For binary mixtures at all temperatures, (a) a positive BIP will increase the low pressure K-value data for the light component and will decrease the low pressure K-value data for the heavy component; (b) a negative BIP will decrease the low pressure K-value data for the light component and will increase the low pressure K-value data for the heavy component.

3. The apparent vapor pressure was found to be related to the shift in the critical pressure and is described.

4. For binary systems, large positive BIPs have a tendency to, at certain temperatures, increase the convergence pressure of the mixture in a seemingly non-physical manner.

5. A framework for displaying the shift and change in shape of the phase envelope of ternary systems was developed using a ternary plot heatmap for the shift in the critical point and difference in the phase envelope area.

6. For ternary mixtures the phase envelope shape and location is directly affected by the binary system critical locus boundaries.

6.2   Conclusions - BIP Tuning

1. Extrapolated values for the computational time to solve the regression problem of 40 to 50 component systems were shown to be infeasible using traditional regression techniques (e.g. PhazeComp's internal solver).

2. A novel approach to BIP regression was developed to significantly reduce the computational time of traditional (simultaneous) regression by selective tuning only a subset of BIPs to use as regression variables - *selective tuning*.

3. The gradient based method for selective tuning was found to handle 40 to 50 component mixtures in reasonable computational times.

4. A convergence study was carried out for the PCC based method used in the selective tuning process. The results shows that the PCC method can yield useful results for a sample size that is significantly smaller than the convergence sample size (see Appendix C section C.2).

5. The PCC based method for selective tuning was found to handle 40 to 50 component mixtures in reasonable computational times.

6. The random forest method for selective tuning, as implemented, did not work.

Chapter 7

Further Work

"*The measure of greatness*[1] *in a scientific idea is the extent to which it stimulates thought and opens up new lines of research.*"

- Paul Dirac

Based on the results of this thesis, the following points describe the proposed further work on the topics (a) the effects of BIPs on phase behavior and (b) the selective tuning approach for BIPs.

1. Develop a methodology for extrapolating the trends found in this work to quaternary systems.

2. Build a larger database of tested examples to confirm that the trends described in the discussion hold for a wider range of cases.

3. Using all types of PVT equilibrium data, instead of or together with K-value data.

4. Understanding the observed low pressure K-value behavior based on theoretical analysis.

5. Apply selective tuning of BIPs to any existing EOS.

6. Identify non-physical phase behavior caused by BIPs (e.g. crossing K-values).

---

[1]This quote is not meant to imply that this work is *greatness*, but rather that all research should strive to work on topics that will open new paths in science.

Bibliography

[1] Abudour, A.M., Mohammad, S.A., Robinson Jr, R.L., Gasem, K.A., 2017. Predicting pr eos binary interaction parameter using readily available molecular properties. Fluid Phase Equilibria 434, 130–140.

[2] Amaratunga, D., Cabrera, J., Lee, Y.S., 2008. Enriched random forests. Bioinformatics 24, 2010–2014.

[3] Benedict, M., Webb, G.B., Rubin, L.C., 1940. An empirical equation for thermodynamic properties of light hydrocarbons and their mixtures i. methane, ethane, propane and n-butane. The Journal of Chemical Physics 8, 334–345.

[4] Breiman, L., 1996. Bagging predictors. Machine learning 24, 123–140.

[5] Carlsen, M.L., Whitson, C.H., Alavian, A., Martinsen, S.Ø., Mydland, S., Singh, K., Younus, B., Yusra, I., et al., 2019. Fluid sampling in tight unconventionals, in: SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers.

[6] Chueh, P., Prausnitz, J., 1967. Vapor-liquid equilibria at high pressures: Calculation of partial molar volumes in nonpolar liquid mixtures. AIChE journal 13, 1099–1107.

[7] Gozalpour, F., Danesh, A., Todd, A., Tehrani, D., 1998. Rapid and robust phase equilibrium calculation to model fluids in reservoir and surface processing. Chemical Engineering Research and Design 76, 594–603.

[8] Hall, K.R., Yarborough, L., 1973. A new equation of state for z-factor calculations. Oil Gas J 71, 82.

[9] Haugen, K.B., Beckner, B.L., et al., 2011. Are reduced methods for eos calculations worth the effort?, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers.

[10] Hendriks, E., Van Bergen, A., 1992. Application of a reduction method to phase equilibria calculations. Fluid Phase Equilibria 74, 17–34.

[11] Hendriks, E.M., 1988. Reduction theorem for phase equilibrium problems. Industrial & engineering chemistry research 27, 1728–1732.

[12] Hoteit, H., Firoozabadi, A., 2006. Simple phase stability-testing algorithm in the reduction method. AIChE journal 52, 2909–2920.

[13] Jaubert, J.N., Mutelet, F., 2004. Vle predictions with the peng–robinson equation of state and temperature dependent kij calculated through a group contribution method. Fluid Phase Equilibria 224, 285–304.

[14] Jaubert, J.N., Privat, R., 2010. Relationship between the binary interaction parameters (kij) of the peng–robinson and those of the soave–redlich–kwong equations of state: Application to the definition of the pr2srk model. Fluid Phase Equilibria 295, 26–37.

[15] Jaubert, J.N., Privat, R., Mutelet, F., 2010. Predicting the phase equilibria of synthetic petroleum fluids with the ppr78 approach. AIChE journal 56, 3225–3235.

[16] Jaubert, J.N., Vitu, S., Mutelet, F., Corriou, J.P., 2005. Extension of the ppr78 model (predictive 1978, peng–robinson eos with temperature dependent kij calculated through a group contribution method) to systems containing aromatic compounds. Fluid Phase Equilibria 237, 193–211.

[17] Jensen, B.H., Fredenslund, A., 1987. A simplified flash procedure for multicomponent mixtures containing hydrocarbons and one non-hydrocarbon using two-parameter cubic equations of state. Industrial & engineering chemistry research 26, 2129–2134.

[18] Katz, D., Firoozabadi, A., et al., 1978. Predicting phase behavior of condensate/crude-oil systems using methane interaction coefficients. Journal of Petroleum Technology 30, 1–649.

[19] Li, Y., Johns, R.T., et al., 2005. Rapid flash calculations for compositional modeling, in: SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers.

[20] Michelsen, M., Yan, W., Stenby, E.H., et al., 2013. A comparative study of reduced-variables-based flash and conventional flash. SPE Journal 18, 952–959.

[21] Michelsen, M.L., 1986. Simplified flash calculations for cubic equations of state. Industrial & Engineering Chemistry Process Design and Development 25, 184–188.

[22] Michelsen, M.L., Mollerup, J., 2004. Thermodynamic modelling: fundamentals and computational aspects. Tie-Line Publications.

[23] Mitchell, T.M., et al., 1997. Machine learning. 1997. Burr Ridge, IL: McGraw Hill 45, 870–877.

[24] Muskat, M., McDowell, J., et al., 1949. An electrical computer for solving phase equilibrium problems. Journal of Petroleum Technology 1, 291–298.

[25] Nichita, D.V., Graciaa, A., 2011. A new reduction method for phase equilibrium calculations. Fluid Phase Equilibria 302, 226–233.

[26] Nida, E., 1991. Theories of translation. TTR: traduction, terminologie, rédaction 4, 19–32.

[27] Pearson, E.S., 1929. Some notes on sampling tests with two variables. Biometrika , 337–360.

[28] Péneloux, A., Rauzy, E., Fréze, R., 1982. A consistent correction for redlich-kwong-soave volumes. Fluid phase equilibria 8, 7–23.

[29] Peng, D.Y., Robinson, D.B., 1976. A new two-constant equation of state. Industrial & Engineering Chemistry Fundamentals 15, 59–64.

[30] Powell, M.J.D., 1977. Restart procedures for the conjugate gradient method. Mathematical programming 12, 241–254.

[31] Privat, R., Jaubert, J.N., Mutelet, F., 2008a. Addition of the nitrogen group to the ppr78 model (predictive 1978, peng robinson eos with temperature-dependent k ij calculated through a group contribution method). Industrial & engineering chemistry research 47, 2033–2048.

[32] Privat, R., Jaubert, J.N., Mutelet, F., 2008b. Use of the ppr78 model to predict new equilibrium data of binary systems involving hydrocarbons and nitrogen. comparison with other gceos. Industrial & engineering chemistry research 47, 7483–7489.

[33] Rachford Jr, H., Rice, J., et al., 1952. Procedure for use of electronic digital computers in calculating flash vaporization hydrocarbon equilibrium. Journal of Petroleum Technology 4, 19–3.

[34] Redlich, O., Kwong, J.N., 1949. On the thermodynamics of solutions. v. an equation of state. fugacities of gaseous solutions. Chemical reviews 44, 233–244.

[35] Robinson, D.B., Peng, D.Y., 1978. The characterization of the heptanes and heavier fractions for the GPA Peng-Robinson programs. Gas processors association.

[36] Soave, G., 1972. Equilibrium constants from a modified redlich-kwong equation of state. Chemical engineering science 27, 1197–1203.

[37] Standing, M., Katz, D., et al., 1944. Vapor-liquid equilibria of natural gas-crude oil systems. Transactions of the AIME 155, 232–245.

[38] Starling, K., et al., 1966. A new approach for determining equation-of-state parameters using phase equilibria data. Society of Petroleum Engineers Journal 6, 363–371.

[39] Valderrama, J.O., Cisternas, L.A., et al., 1990. Binary interaction parameters in cubic equations of state for hydrogen—hydrocarbon mixtures. Chemical engineering science 45, 49–54.

[40] Vitu, S., Jaubert, J.N., Mutelet, F., 2006. Extension of the ppr78 model (predictive 1978, peng–robinson eos with temperature dependent kij calculated through a group contribution method) to systems containing naphtenic compounds. Fluid phase equilibria 243, 9–28.

[41] Vitu, S., Privat, R., Jaubert, J.N., Mutelet, F., 2008. Predicting the phase equilibria of co2+ hydrocarbon systems with the ppr78 model (pr eos and kij calculated through a group contribution method). The Journal of Supercritical Fluids 45, 1–26.

[42] Van der Waals, J., 1873. On the continuity of the gaseous and liquid states (doctoral dissertation). Universiteit Leiden .

[43] Whitson, C.H., Brulé, M.R., et al., 2000. Phase behavior. volume 20. Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers . . . .

[44] Whitson, C.H., Michelsen, M.L., 1989. The negative flash. Fluid Phase Equilibria 53, 51–71.

[45] Xu, X., Jaubert, J.N., Privat, R., Duchet-Suchaux, P., Bana-Mulero, F., 2015. Predicting binary-interaction parameters of cubic equations of state for petroleum fluids containing pseudo-components. Industrial & Engineering Chemistry Research 54, 2816–2824.

[46] Younus*, B., Whitson, C.H., Alavian, A., Carlsen, M.L., Martinsen, S.Ø., Singh, K., 2019. Field-wide equation of state model development, in: Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July 2019, Unconventional Resources Technology Conference (URTeC); Society of . . . . pp. 1031–1070.

[47] Zhang, C., Ma, Y., 2012. Ensemble machine learning: methods and applications. Springer.

Appendix A

Effect of BIPs on Binary Systems

"*I am a brain, Watson. The rest of me is a mere appendix.*"

- Sherlock Holmes

The following figures show the PT diagrams for all combinations of SCN components C_5, C_{10}, C_{15} and C_{20} with a the critical pressure locus with a zero BIP as well as a positive BIP with magnitude 0.1 and finally a negative BIP with the same magnitude (i.e. BIP = -0.1). A general note is that the lighter component vapor pressure line (solid black line) tends to be further towards the left compared to the heavier component.
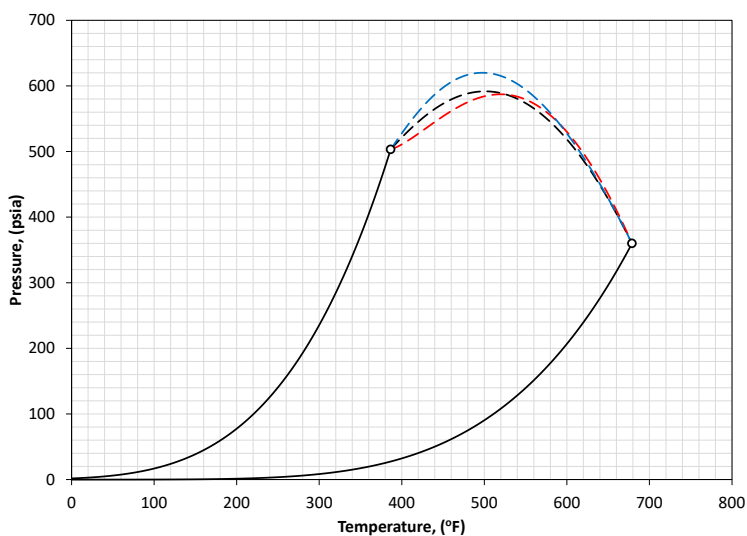
Figure A.1: Binary mixture containing C_5 and C_{10} critical pressure locus with zero BIP (black dashed line), BIP = 0.1 (red dashed line) and BIP = -0.1 (blue dashed line).
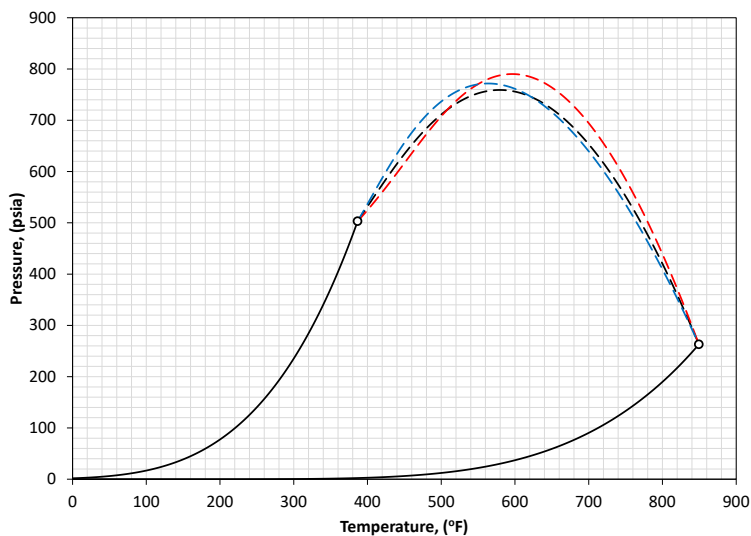
Figure A.2: Binary mixture containing C_5 and C_{15} critical pressure locus with zero BIP (black dashed line), BIP = 0.1 (red dashed line) and BIP = -0.1 (blue dashed line).
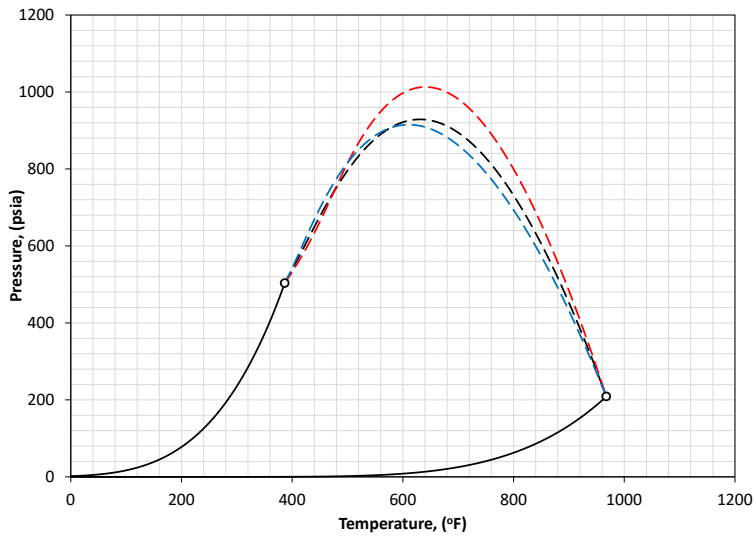
Figure A.3: Binary mixture containing C_5 and C_{20} critical pressure locus with zero BIP (black dashed line), BIP = 0.1 (red dashed line) and BIP = -0.1 (blue dashed line).
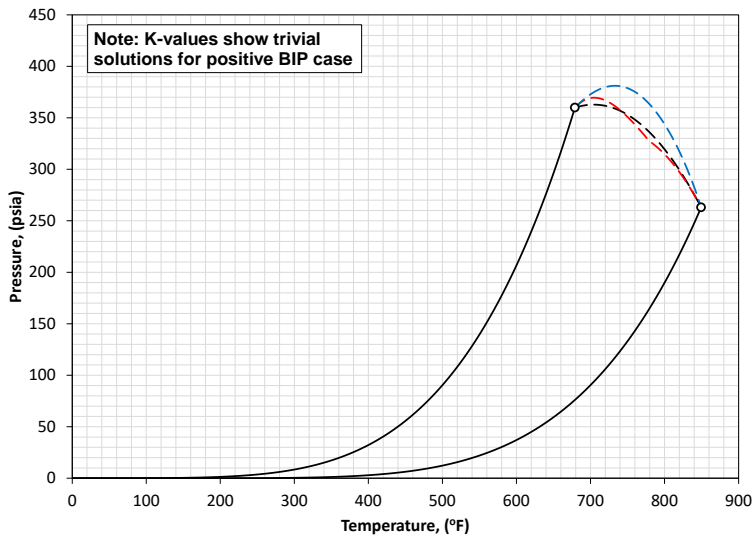
Figure A.4: Binary mixture containing C_{10} and C_{15} critical pressure locus with zero BIP (black dashed line), BIP = 0.1 (red dashed line) and BIP = -0.1 (blue dashed line).
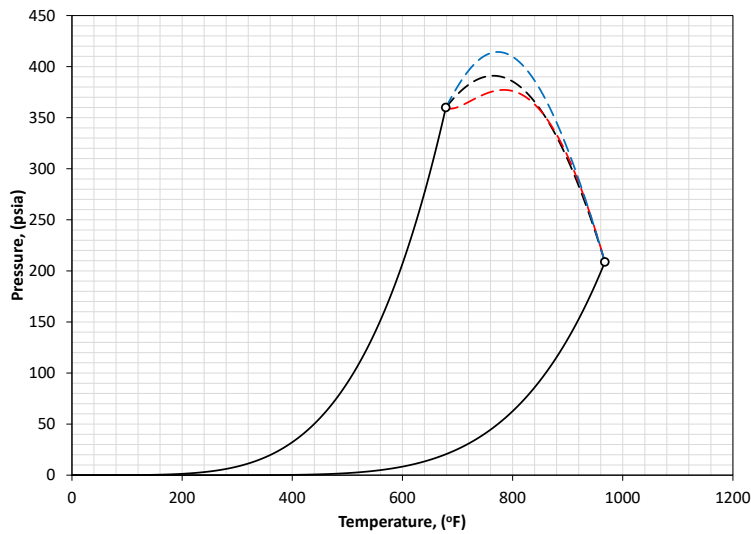
Figure A.5: Binary mixture containing C_{10} and C_{20} critical pressure locus with zero BIP (black dashed line), BIP = 0.1 (red dashed line) and BIP = -0.1 (blue dashed line).
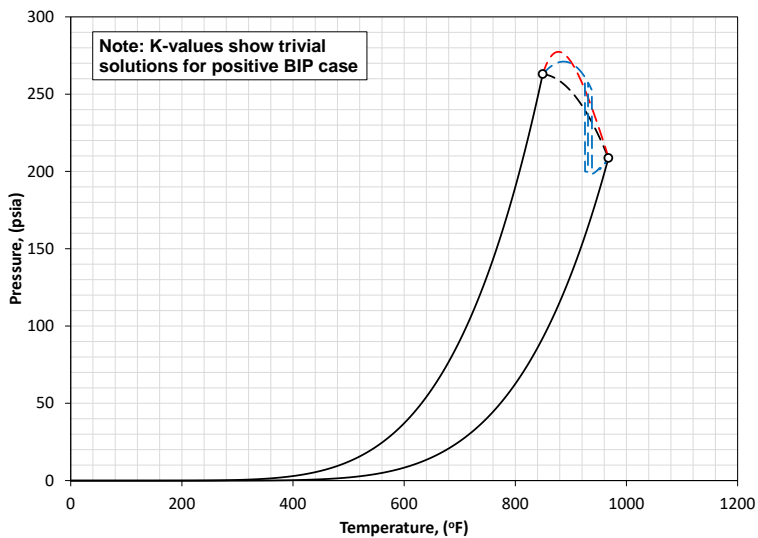
Figure A.6: Binary mixture containing C_{15} and C_{20} critical pressure locus with zero BIP (black dashed line), BIP = 0.1 (red dashed line) and BIP = -0.1 (blue dashed line).

Effect of BIPs on Ternary Systems

In the following sections (a) the PT shift (i.e. the critical point and phase envelope area) of the phase envelope for all cases in Table 3.1 are given and (b) the K-value plots from Table 3.2 are given.

It should be stated that when generating the results for section B.1 each set of three figures required a simulation time of roughly 1 hour. The range of results is not discussed in a great deal of detail in the discussion section, but is gathered in this appendix for potential further analysis for anyone to use.

B.1 PT Phase Behavior Shift

(a)

(b)

(c)
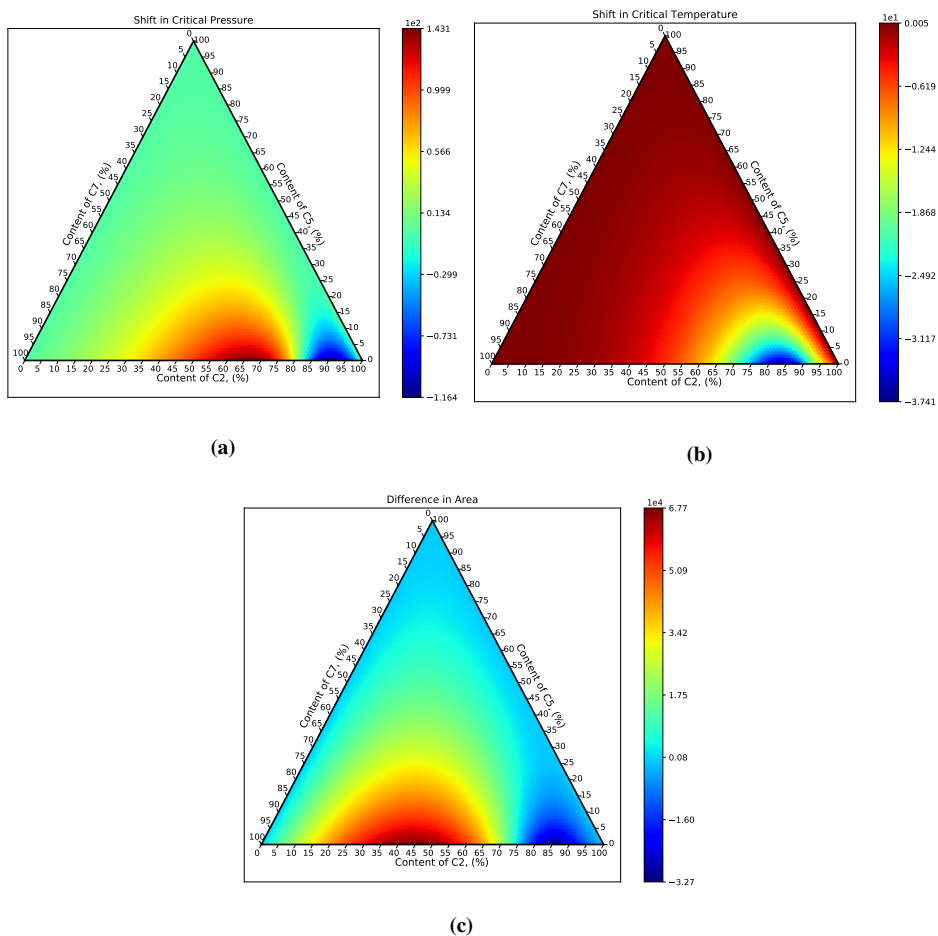
Figure B.1: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_7 and zero BIPs for the C_2-C_5 and C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)

(a)

(b)

(c)
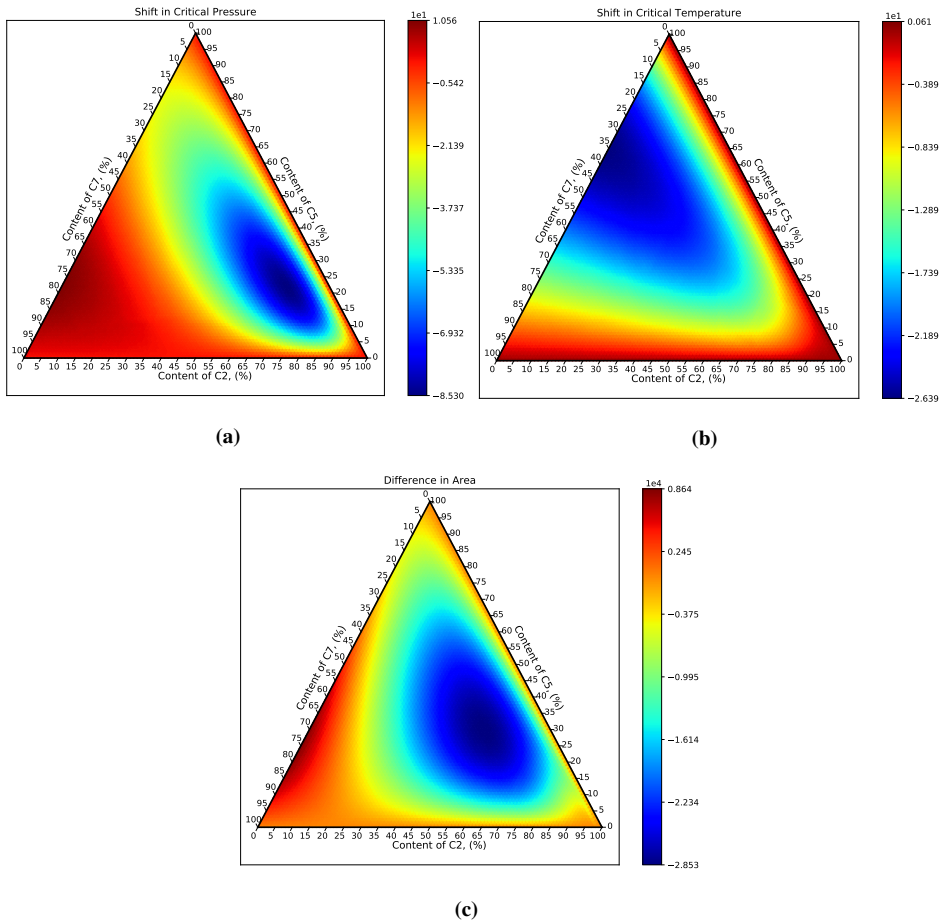
Figure B.2: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_5-C_7 and zero BIPs for the C_2-C_5 and C_2-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
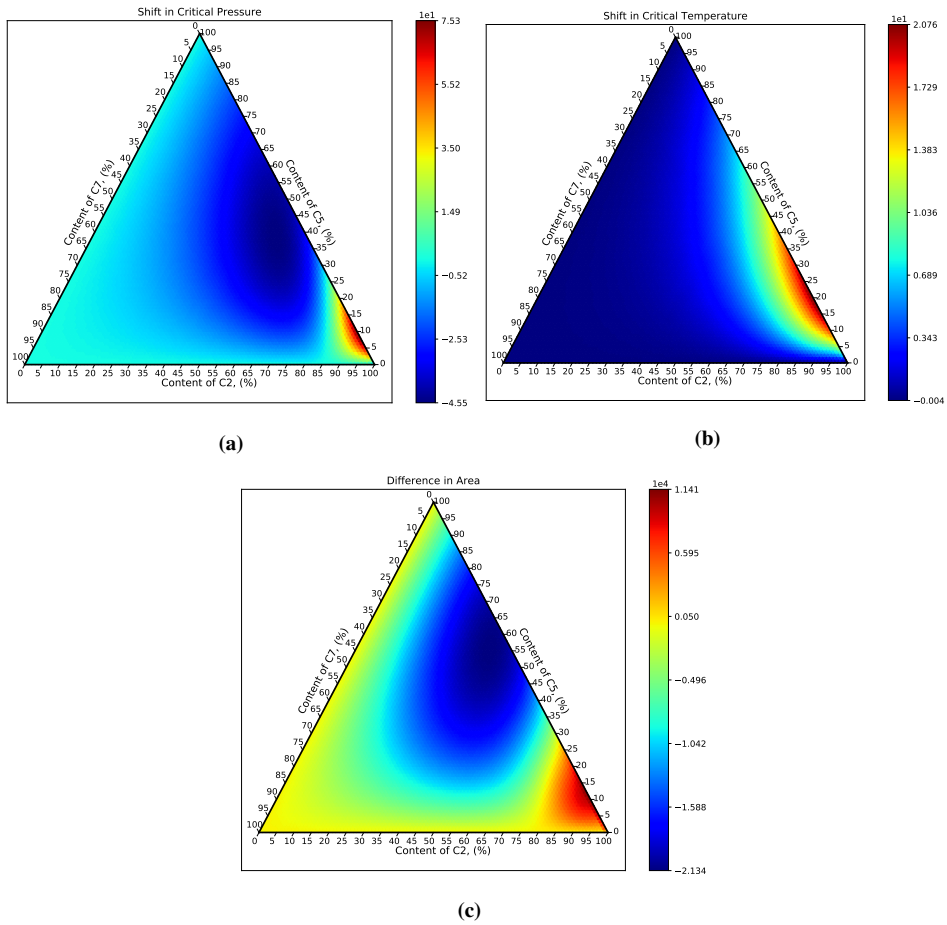
Figure B.3: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5 and zero BIPs for the C_2-C_7 and C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
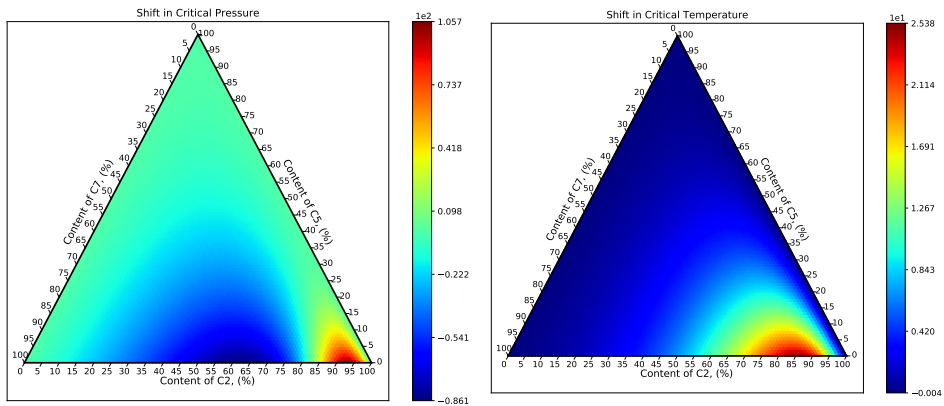
(a)

(b)
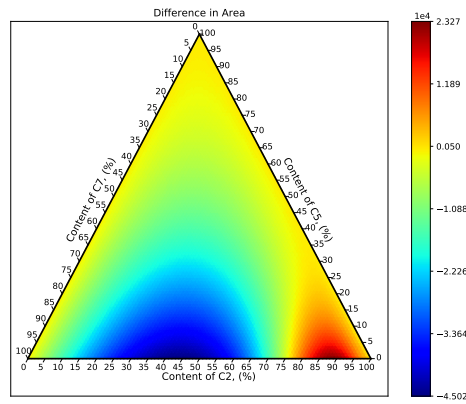
(c)

Figure B.4: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_7 and zero BIPs for the C_2-C_5 and C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
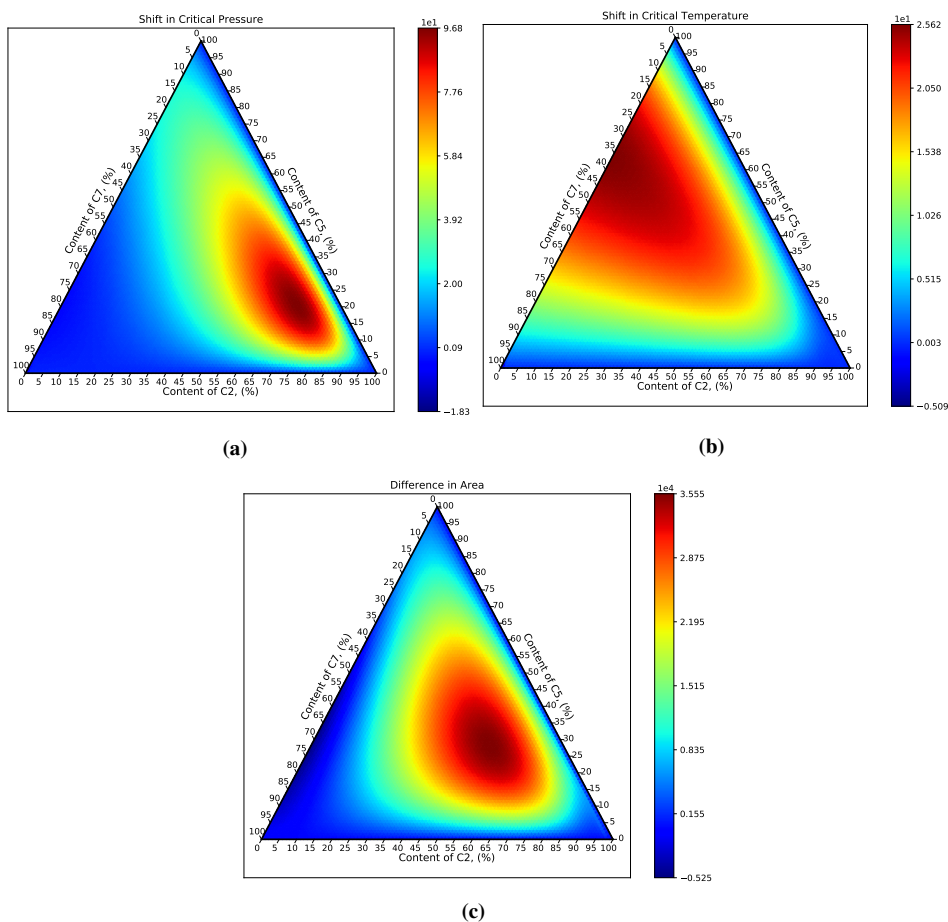
Figure B.5: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_5-C_7 and zero BIPs for the C_2-C_5 and C_2-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
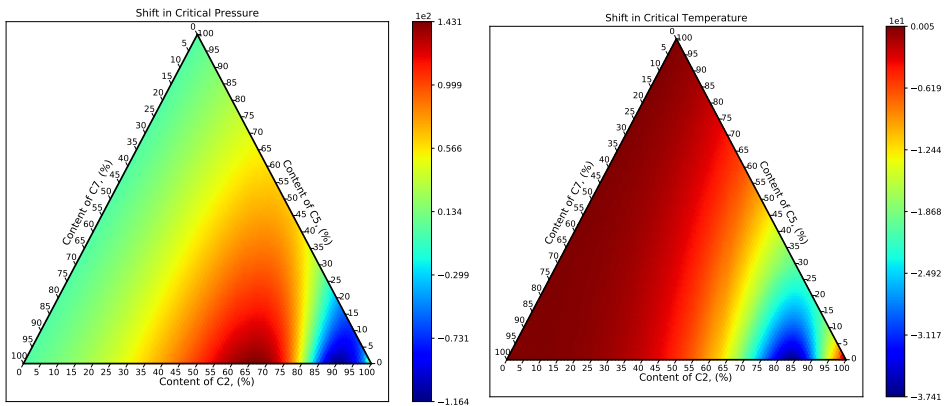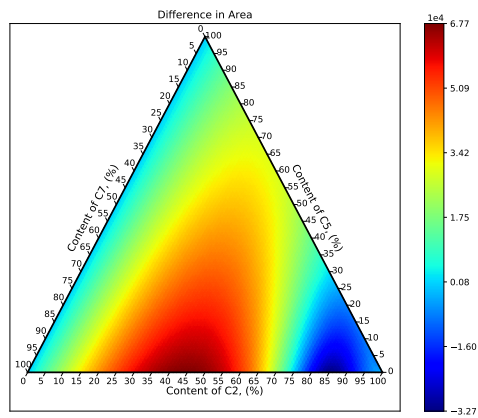
Figure B.6: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5 and C_2-C_7 and zero BIPs for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
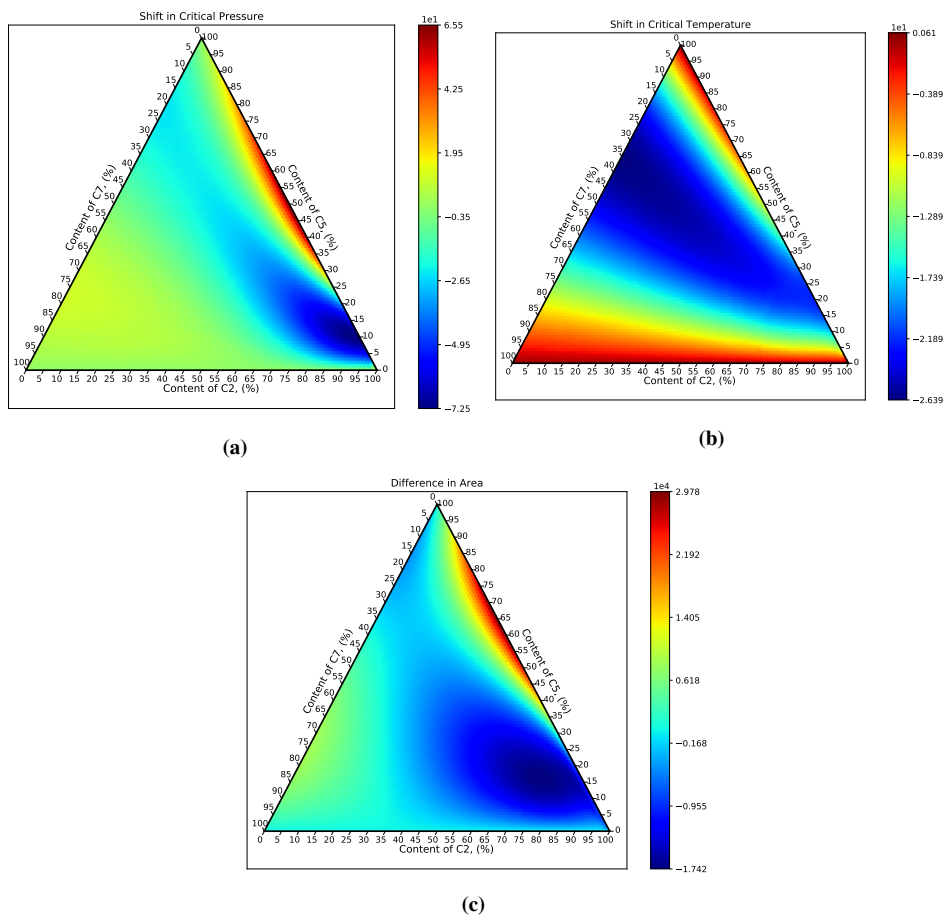
Figure B.7: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5 and C_5-C_7 and zero BIPs for the C_2-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)

(a)

(b)

(c)
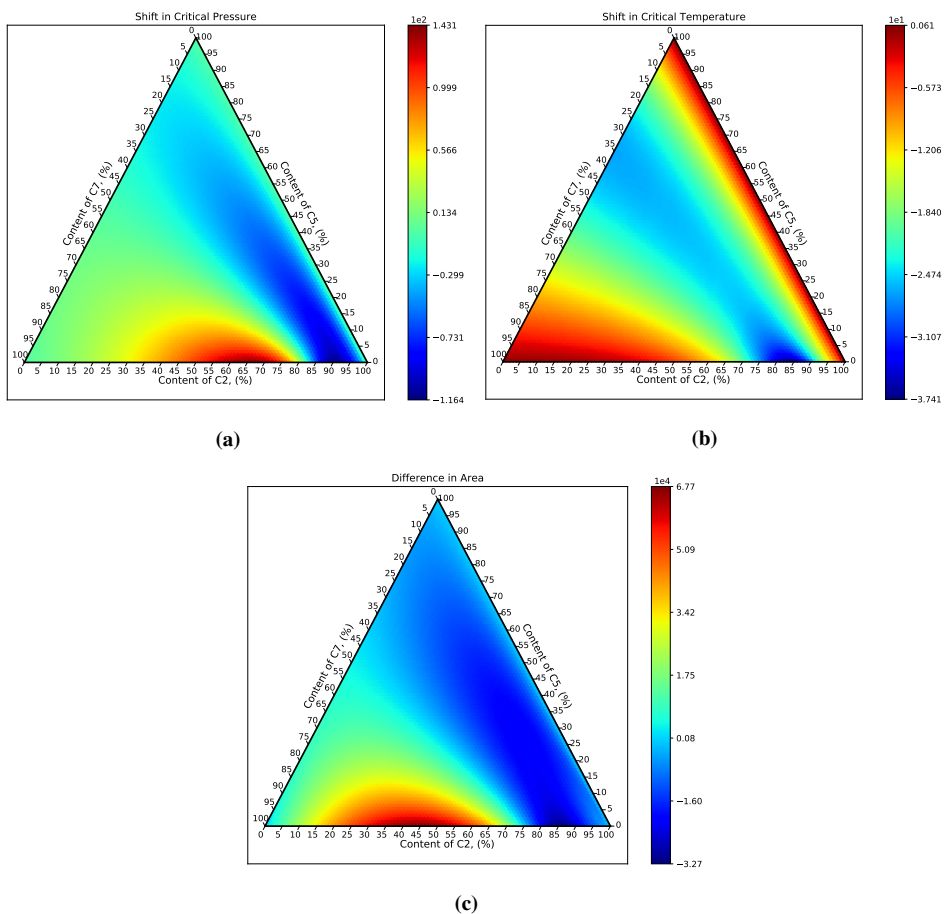
Figure B.8: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_7 and C_5-C_7 and zero BIPs for the C_2-C_5 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
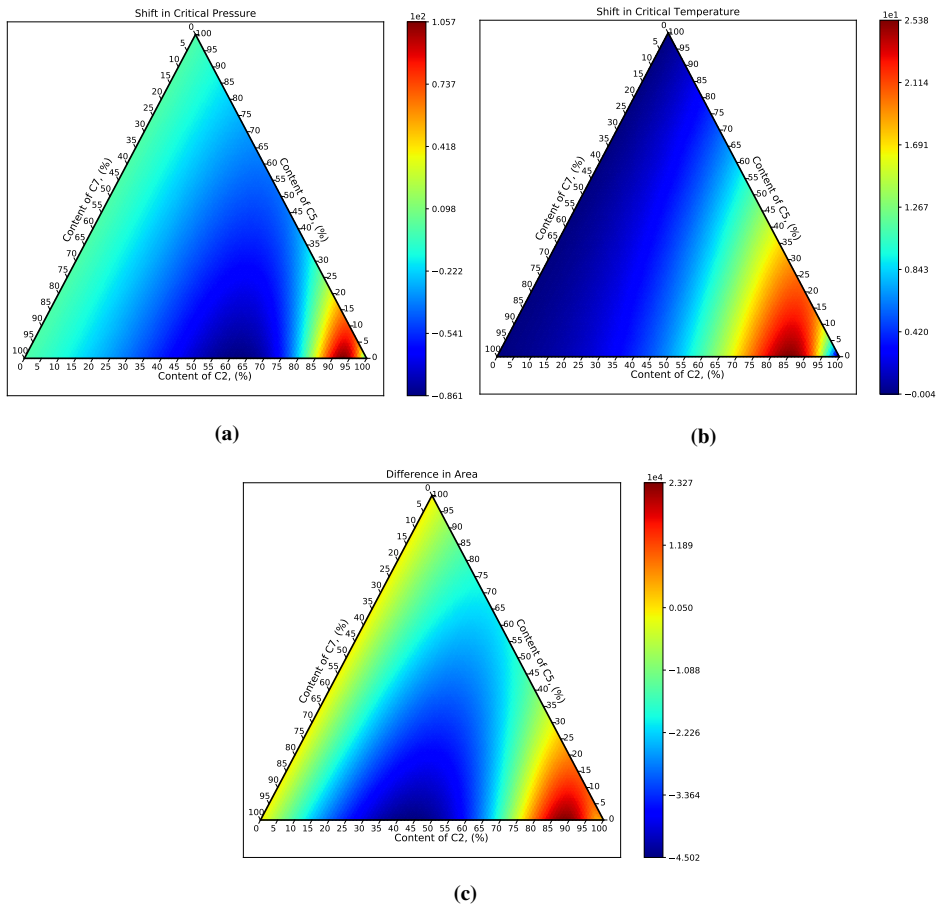
Figure B.9: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5 and C_2-C_7 and zero BIPs for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
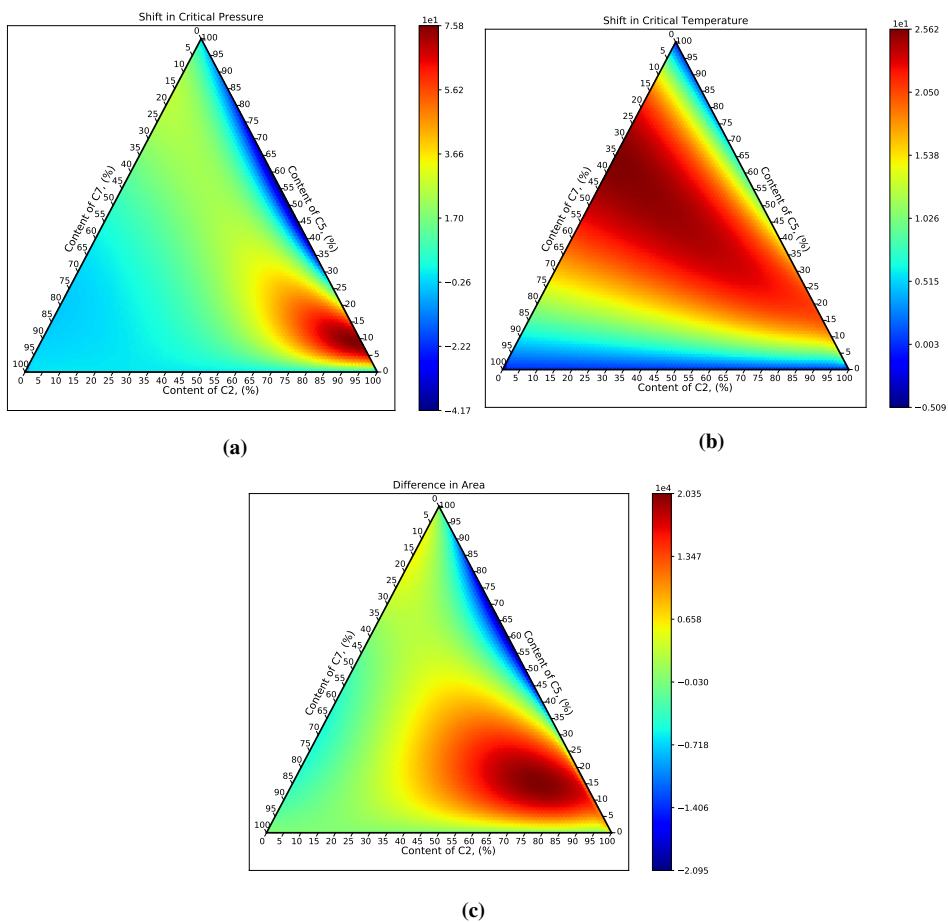
Figure B.10: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5 and C_5-C_7 and zero BIPs for the C_2-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
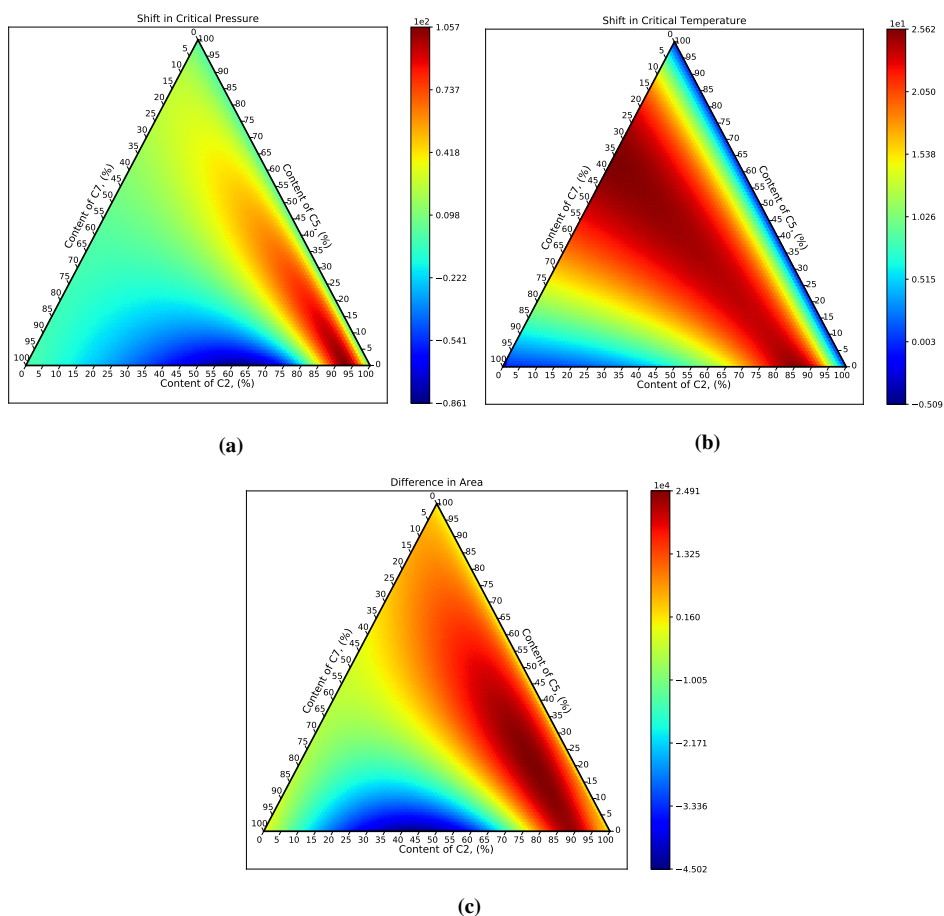
Figure B.11: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_7 and C_5-C_7 and zero BIPs for the C_2-C_5 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
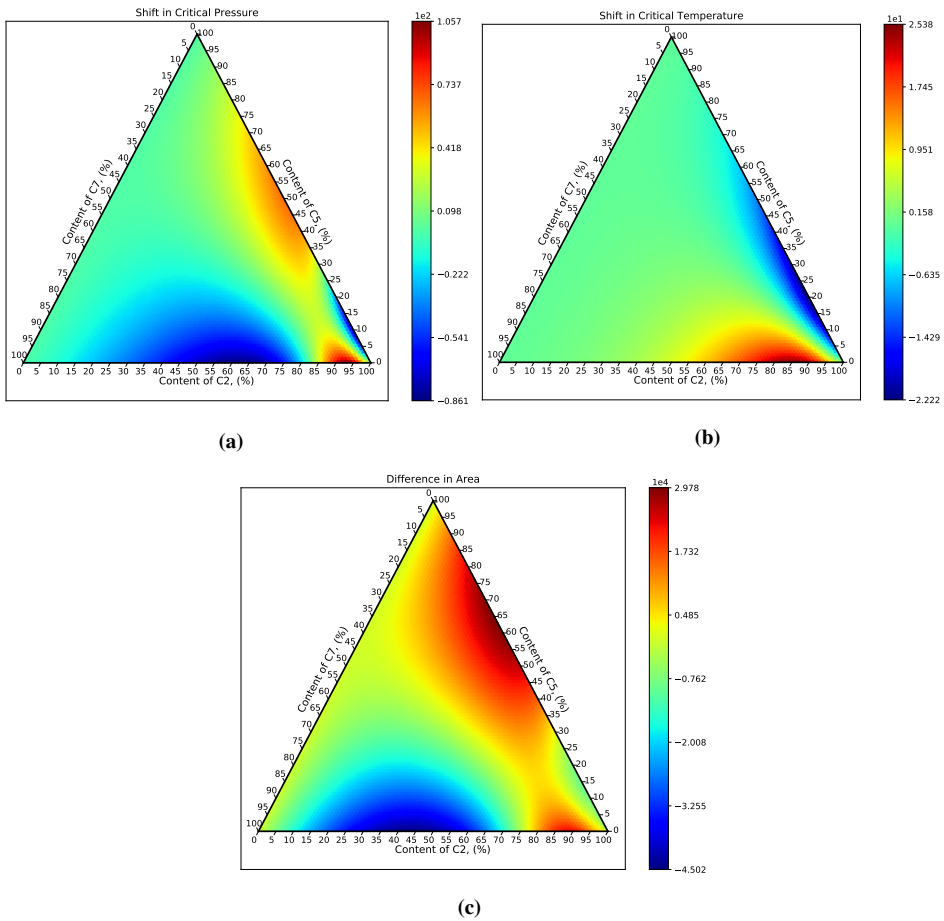
Figure B.12: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5 and a BIP of -0.1 C_2-C_7 and zero BIPs for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
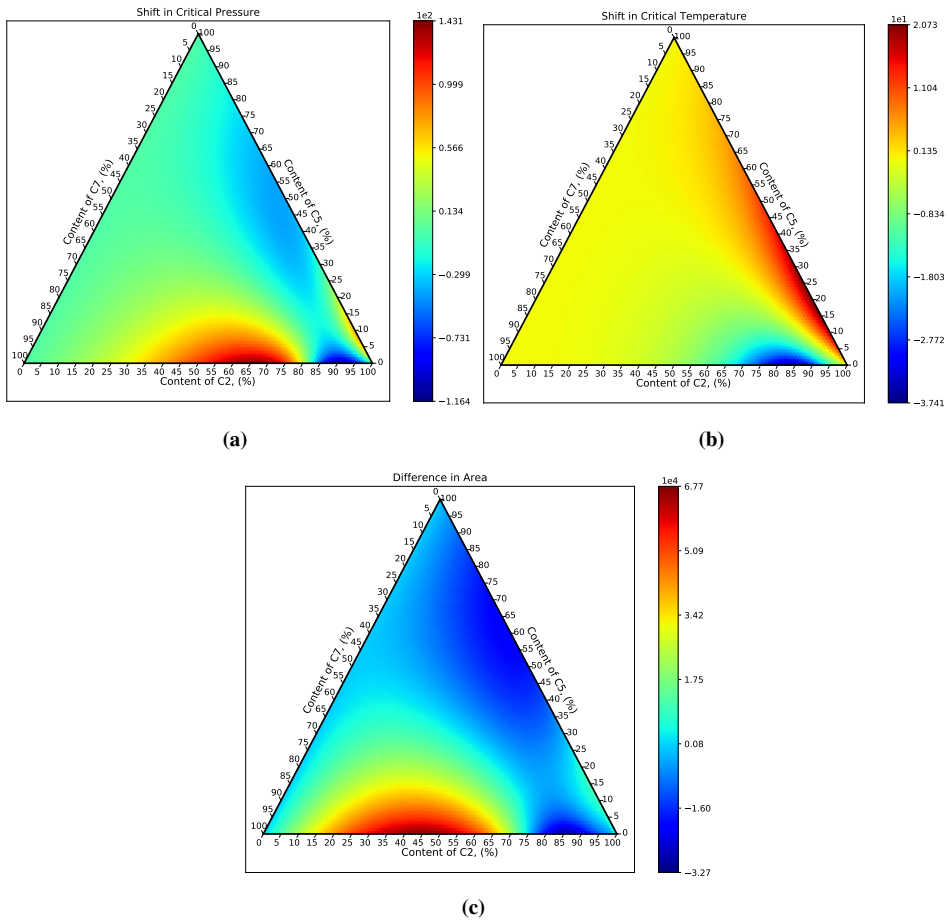
Figure B.13: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5 and a BIP of 0.1 C_2-C_7 and zero BIPs for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
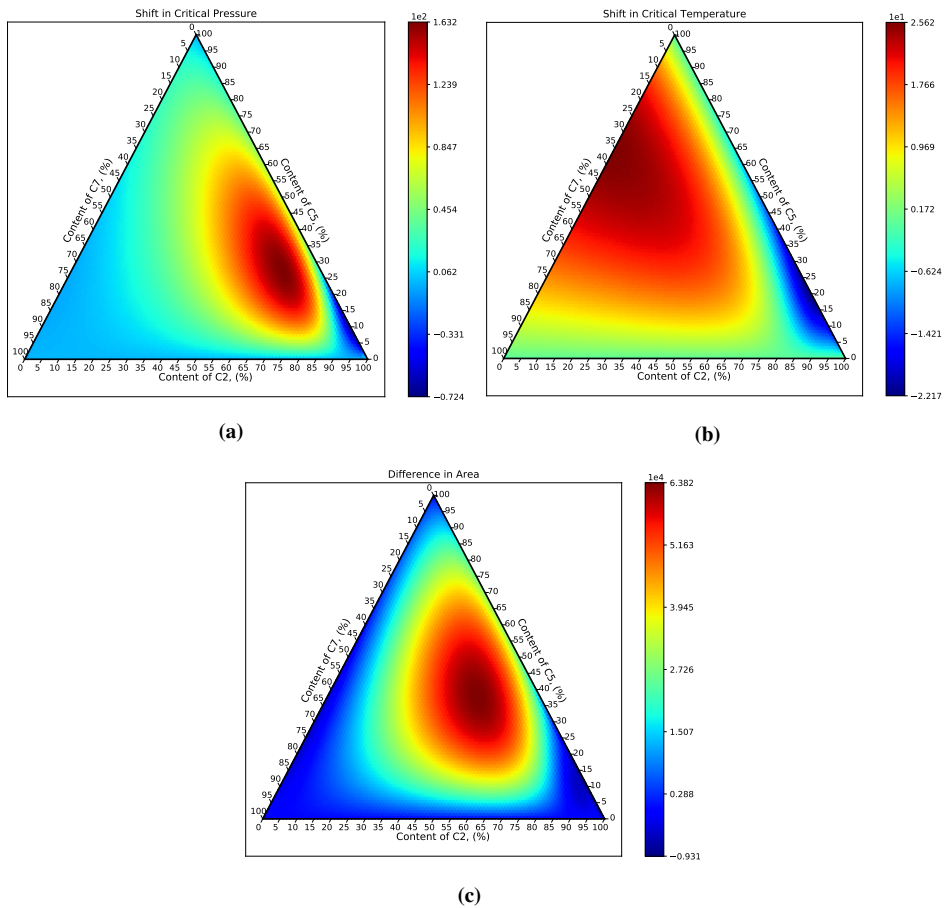
Figure B.14: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5 and a BIP of -0.1 C_5-C_7 and zero BIPs for the C_2-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
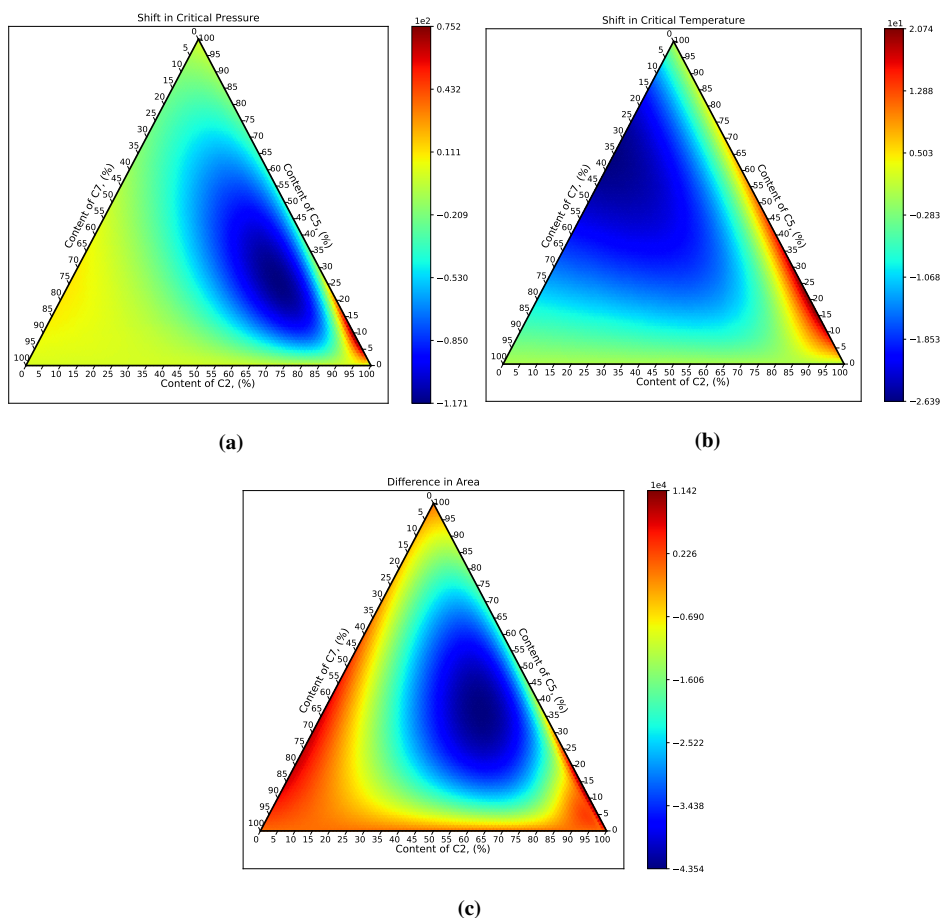
Figure B.15: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5 and a BIP of 0.1 C_5-C_7 and zero BIPs for the C_2-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)

(a)

(b)

(c)
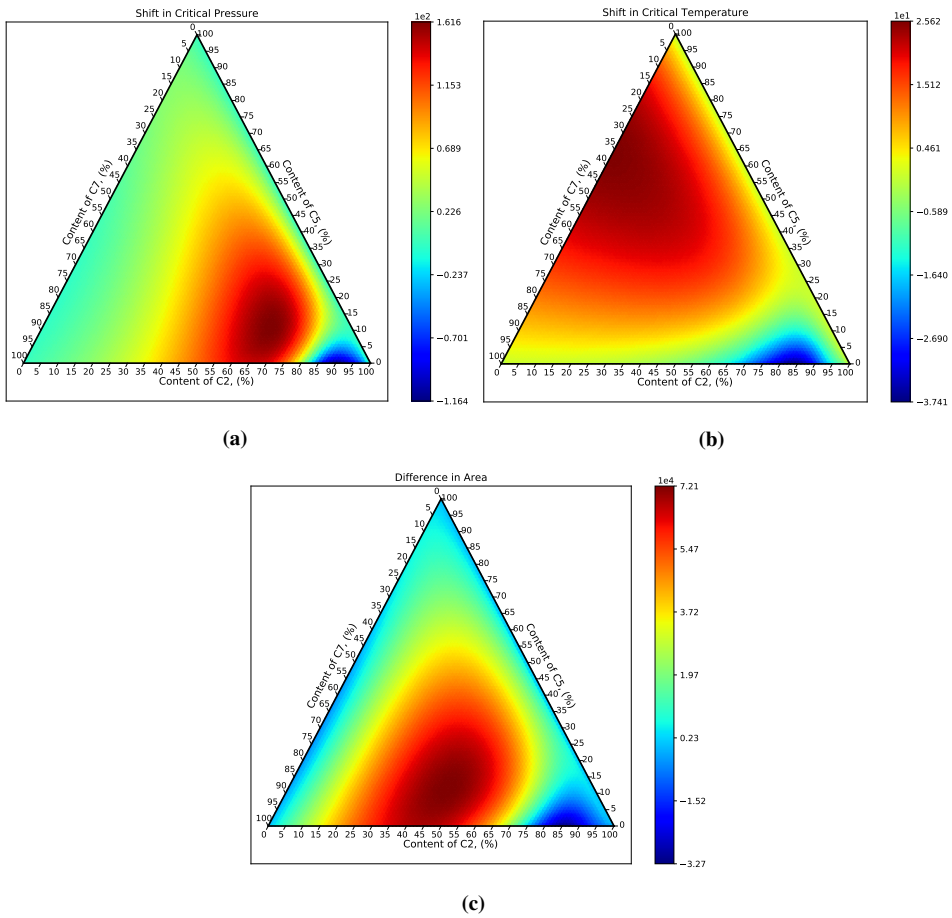
Figure B.16: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_7 and a BIP of -0.1 C_5-C_7 and zero BIPs for the C_2-C_5 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
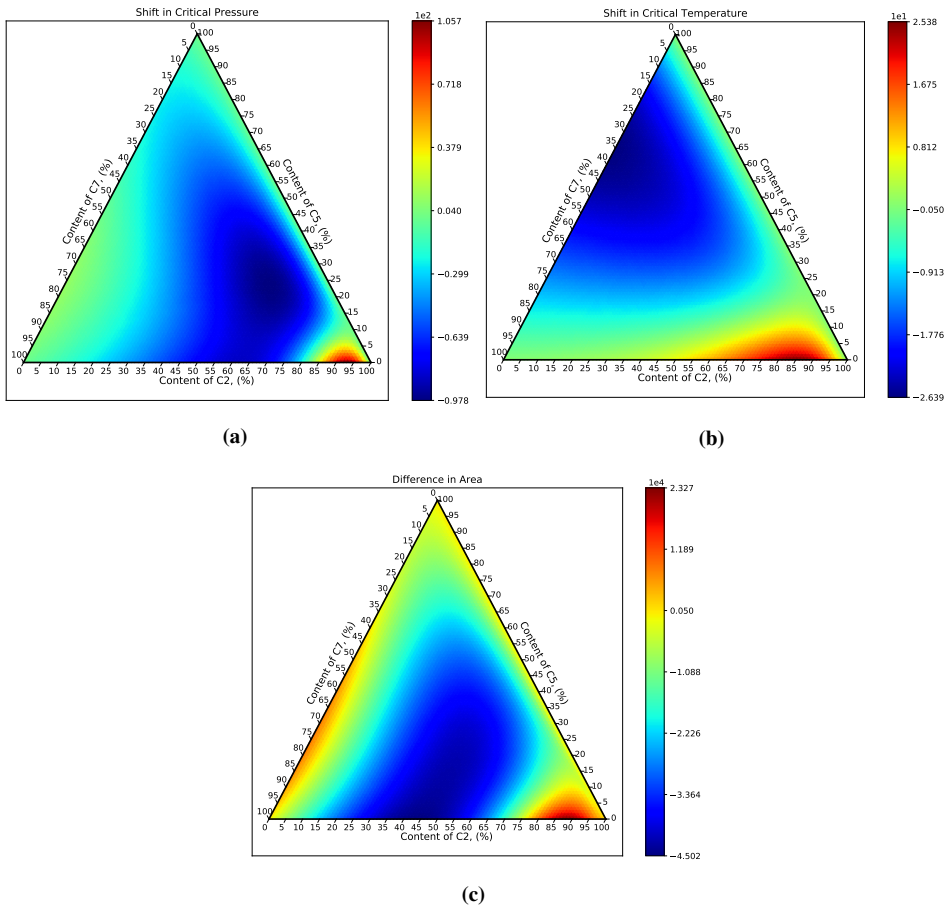
Figure B.17: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_7 and a BIP of 0.1 C_5-C_7 and zero BIPs for the C_2-C_5 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
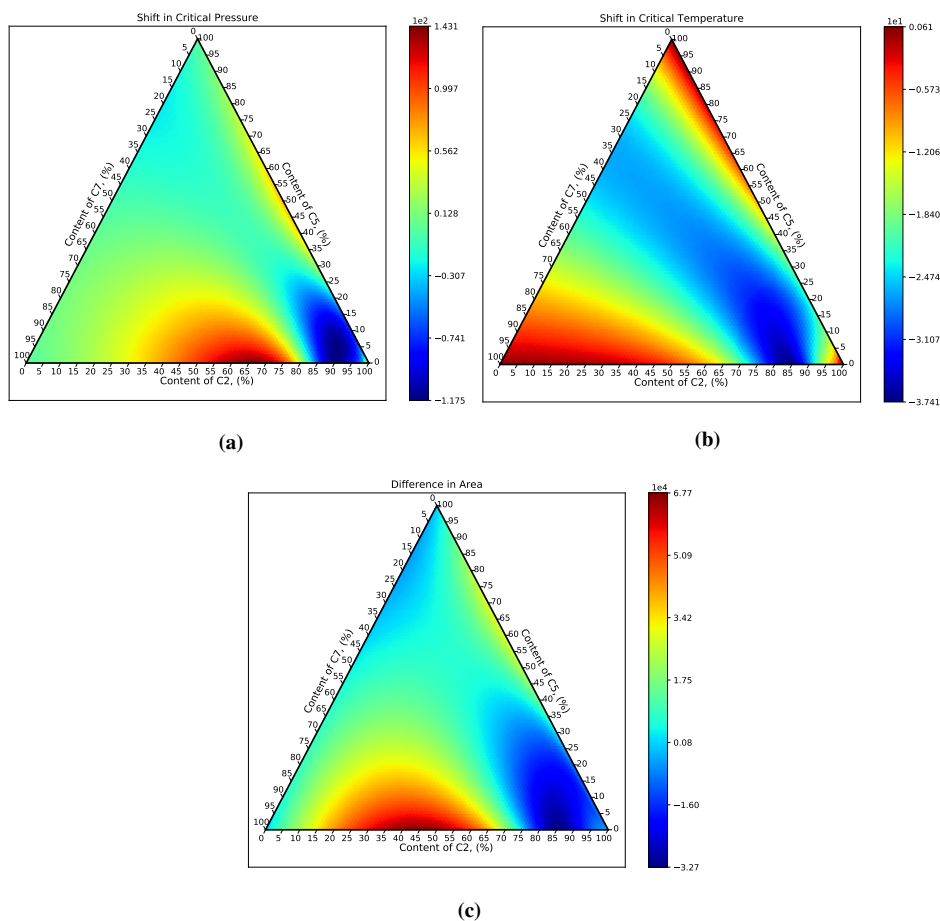
Figure B.18: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5, a BIP of 0.1 C_2-C_7 and a BIPs of 0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
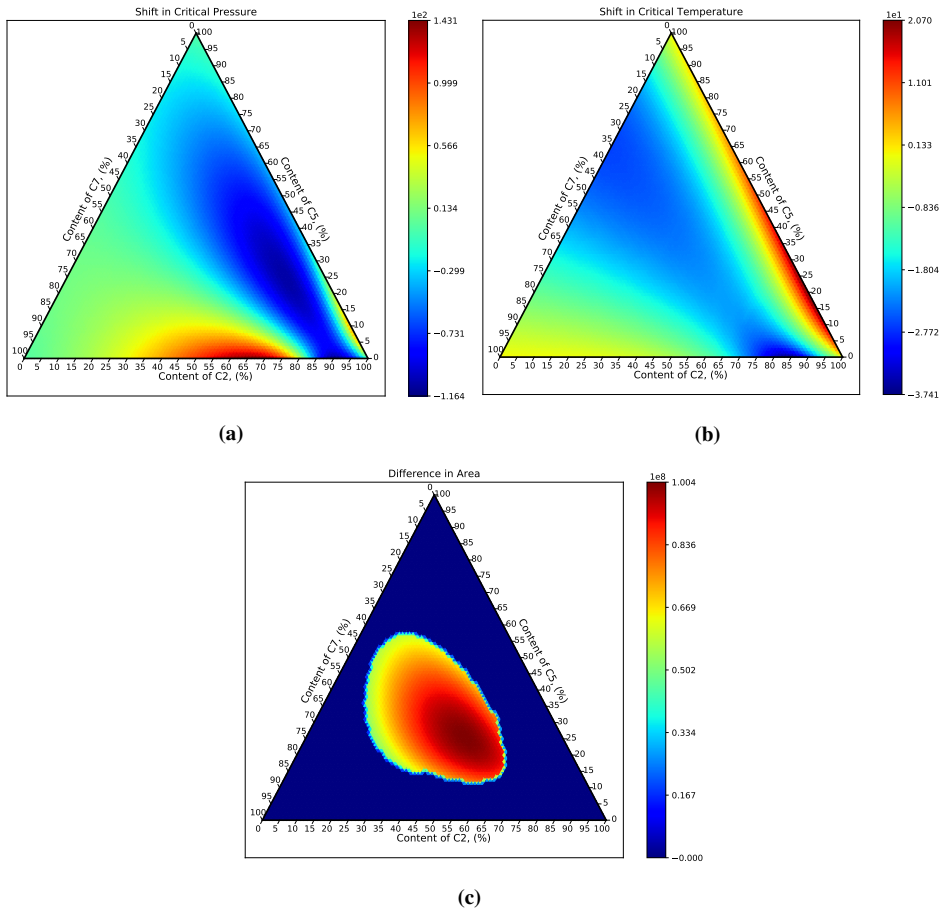
Figure B.19: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5, a BIP of 0.1 C_2-C_7 and a BIPs of 0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
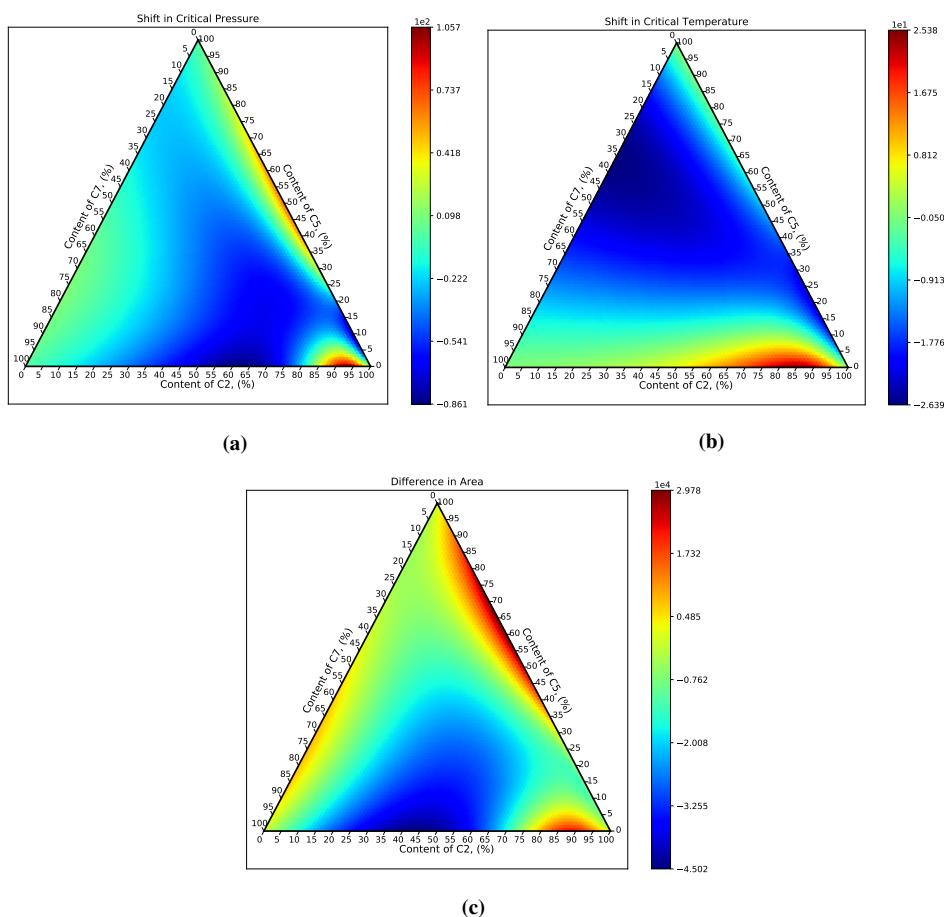
Figure B.20: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5, a BIP of -0.1 C_2-C_7 and a BIPs of 0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
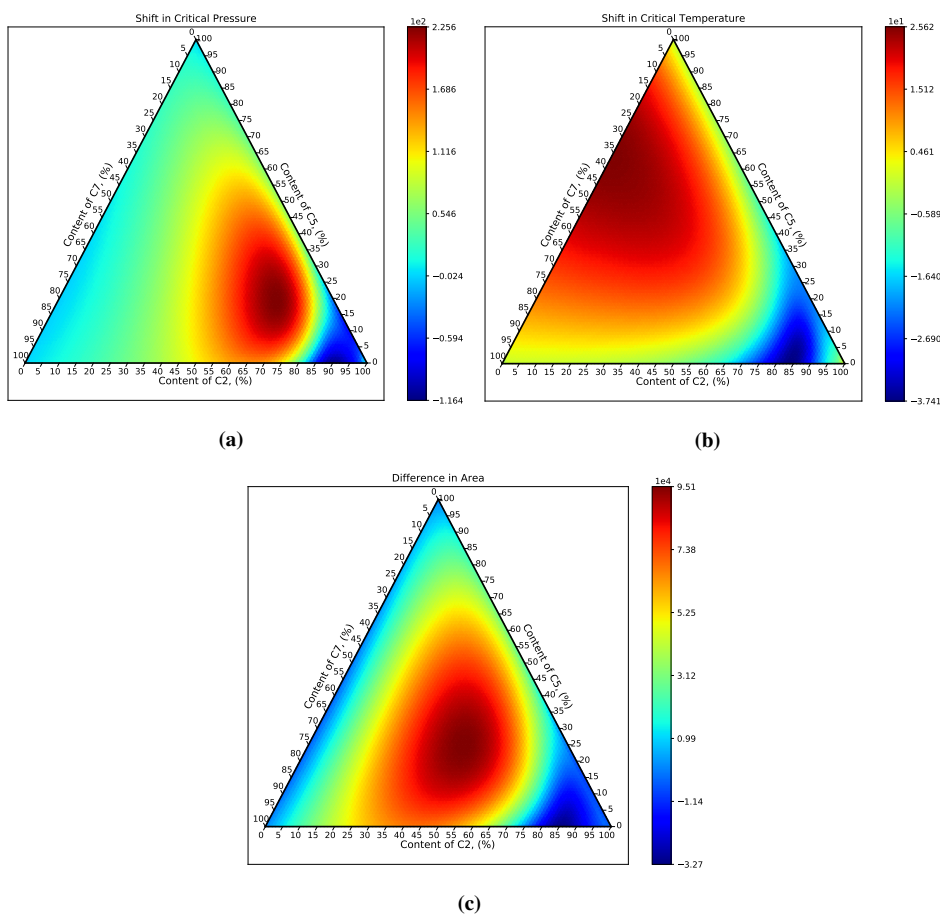
Figure B.21: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5, a BIP of 0.1 C_2-C_7 and a BIPs of -0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
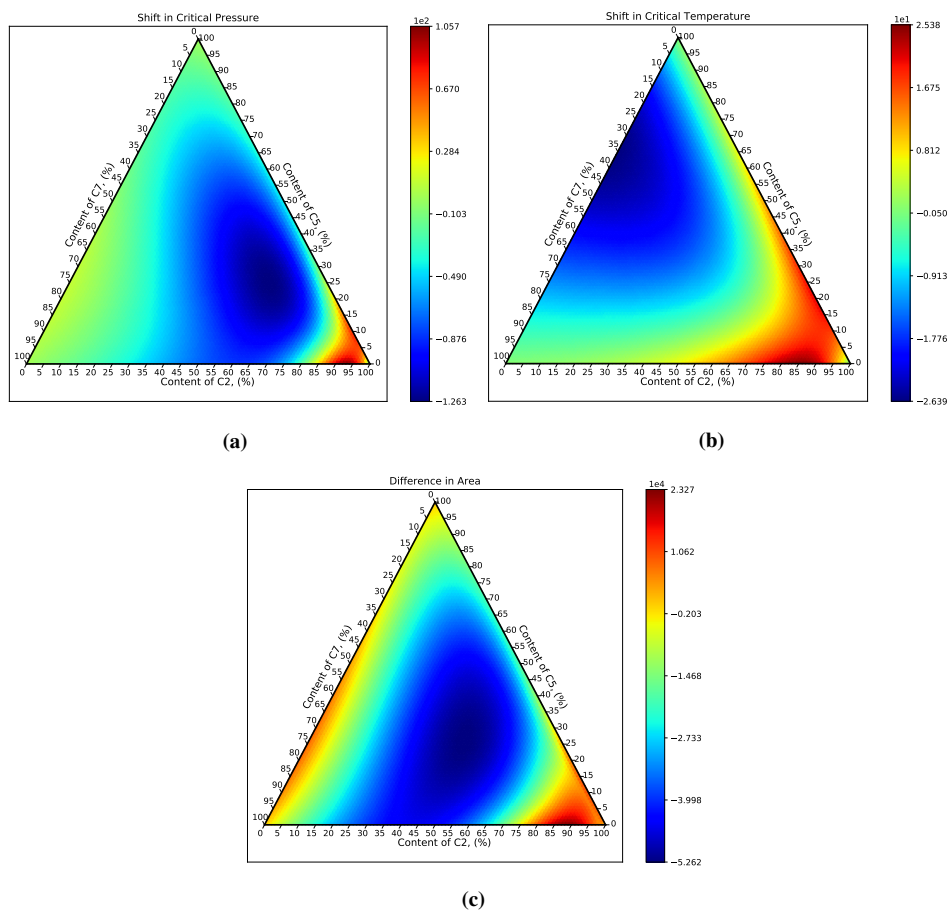
Figure B.22: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5, a BIP of -0.1 C_2-C_7 and a BIPs of 0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)

(a)

(b)

(c)
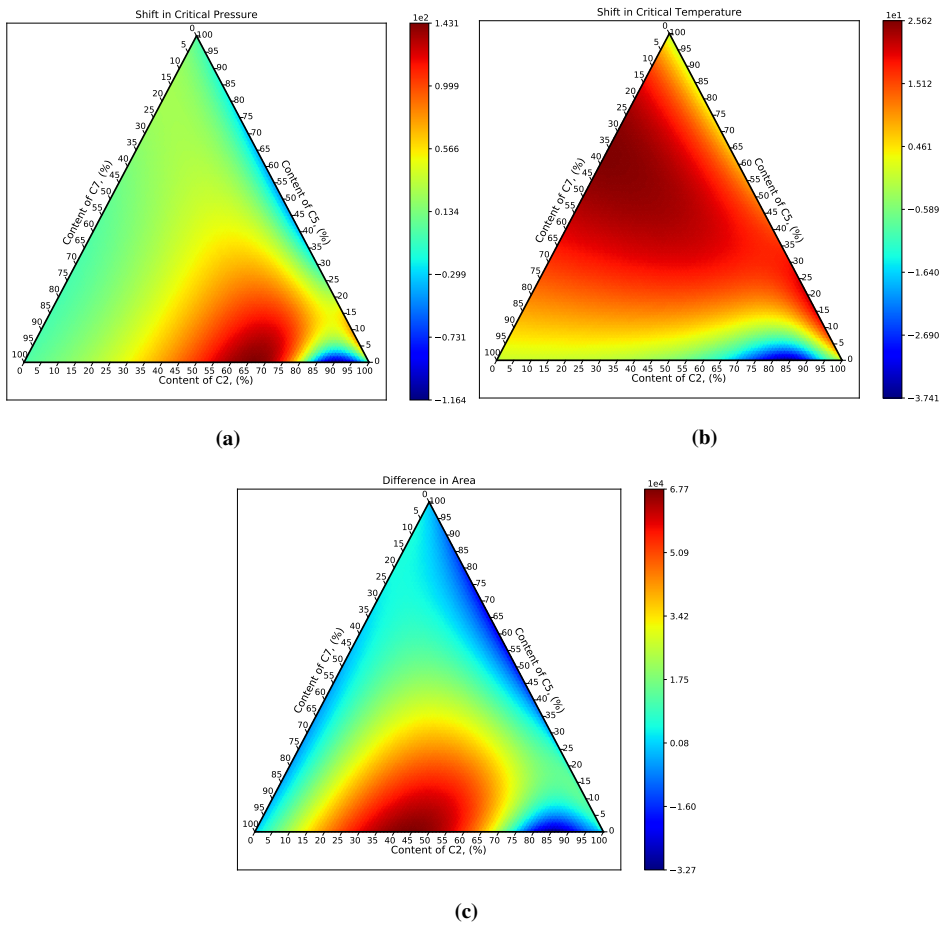
Figure B.23: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5, a BIP of 0.1 C_2-C_7 and a BIPs of -0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
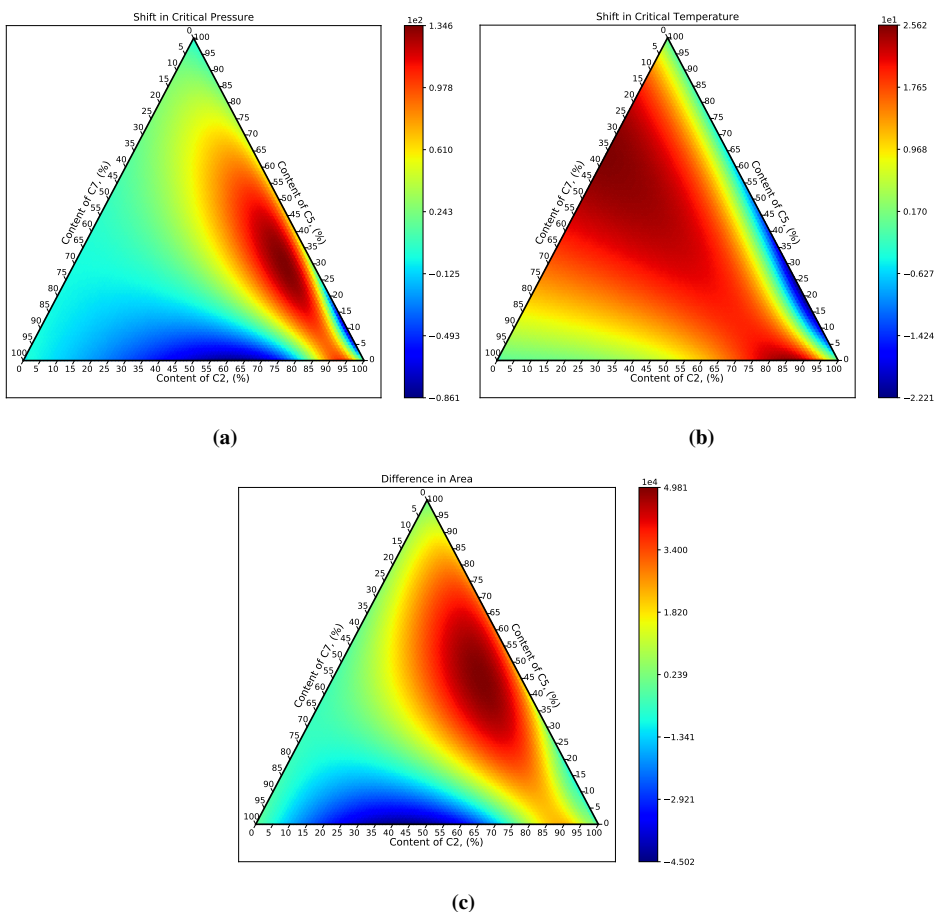
Figure B.24: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of 0.1 between C_2-C_5, a BIP of -0.1 C_2-C_7 and a BIPs of -0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)
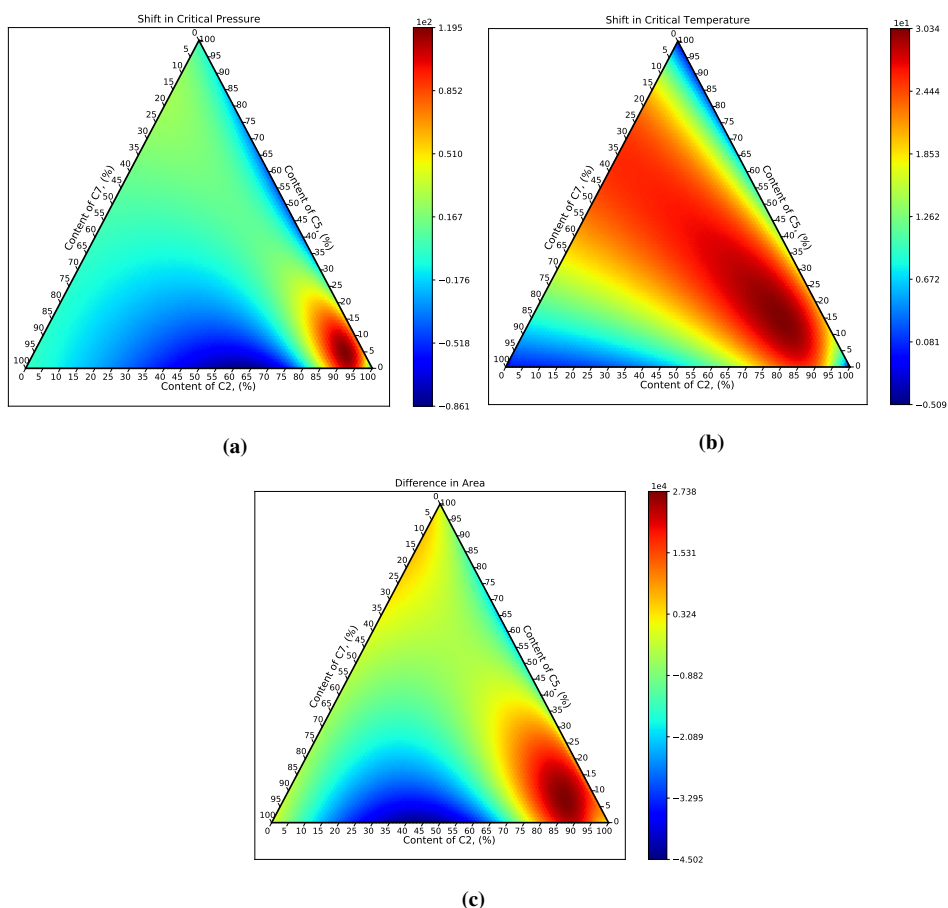
Figure B.25: Shift in critical point and difference in phase envelope area for C_2-C_5-C_7 mixture with a BIP of -0.1 between C_2-C_5, a BIP of -0.1 C_2-C_7 and a BIPs of -0.1 for the C_5-C_7 binaries. (a) shift in critical temperature, (b) shift in critical pressure and (c) difference between area of phase envelope with non-zero BIPs and with zero BIPs as defined in equation (3.2)

B.2 Ternary K-value Plots

(a)

(b)

(c)
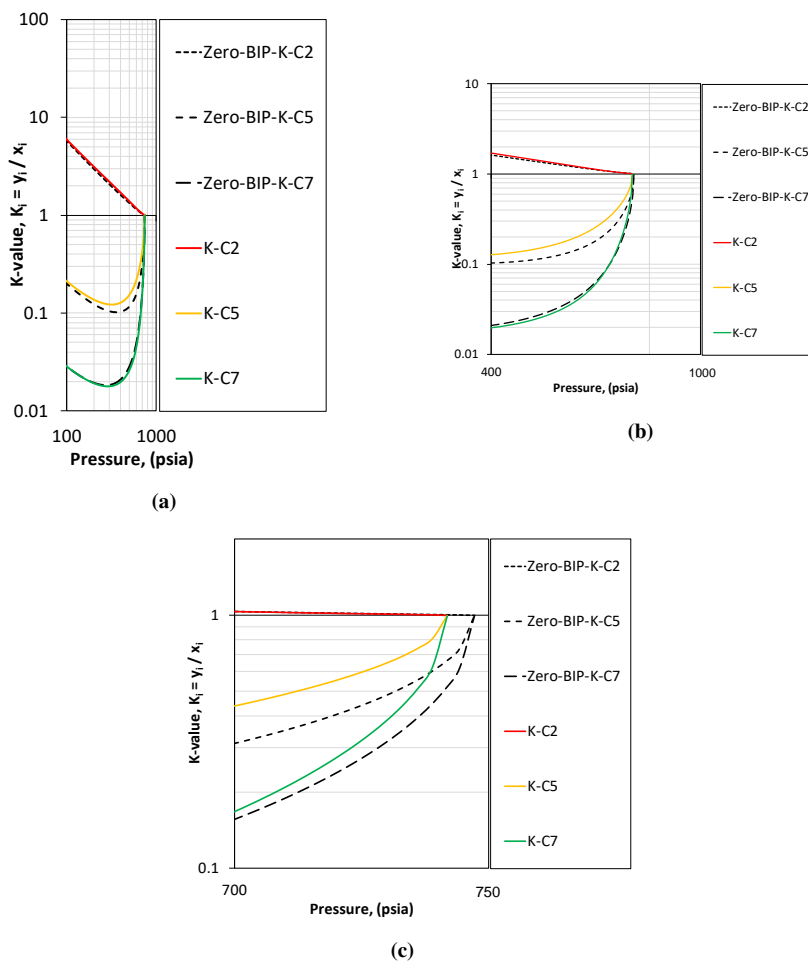
Figure B.26: K-value plot for case 1.1 described in Table 3.2 at 100°F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
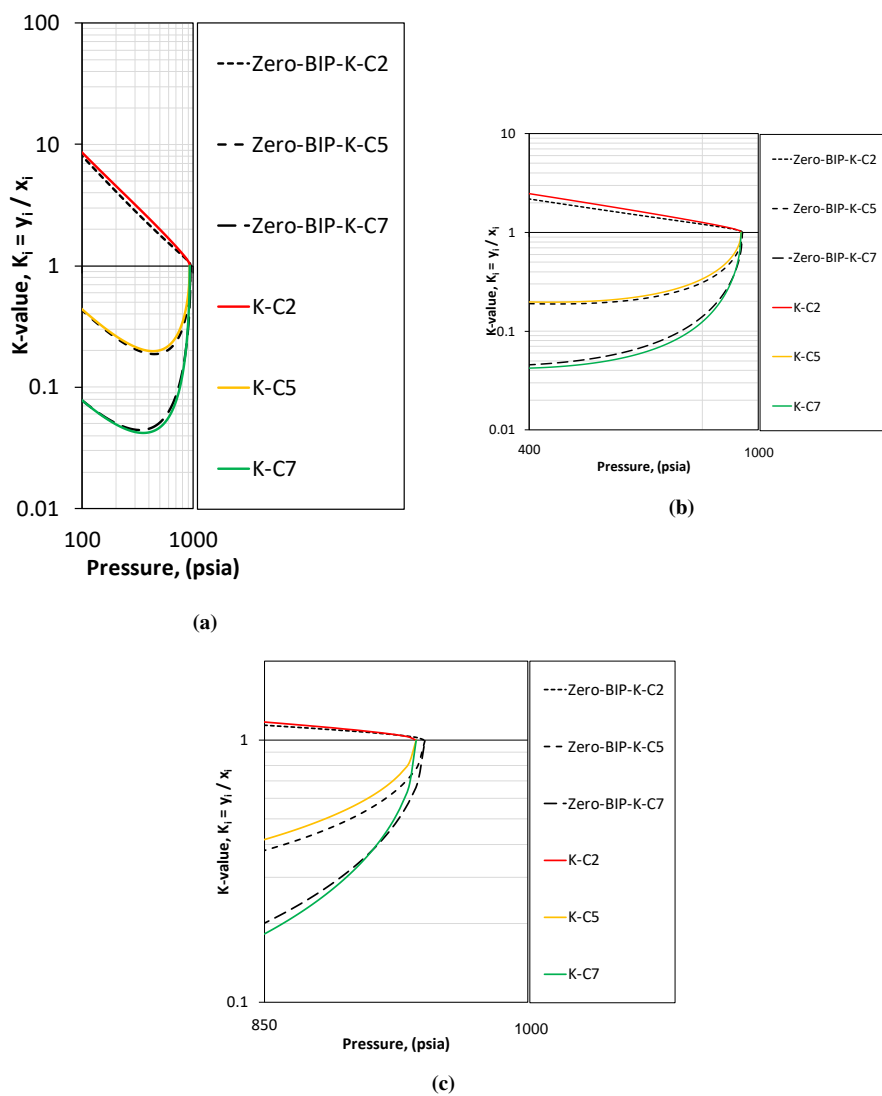
Figure B.27: K-value plot for case 1.1 described in Table 3.2 at 150°F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
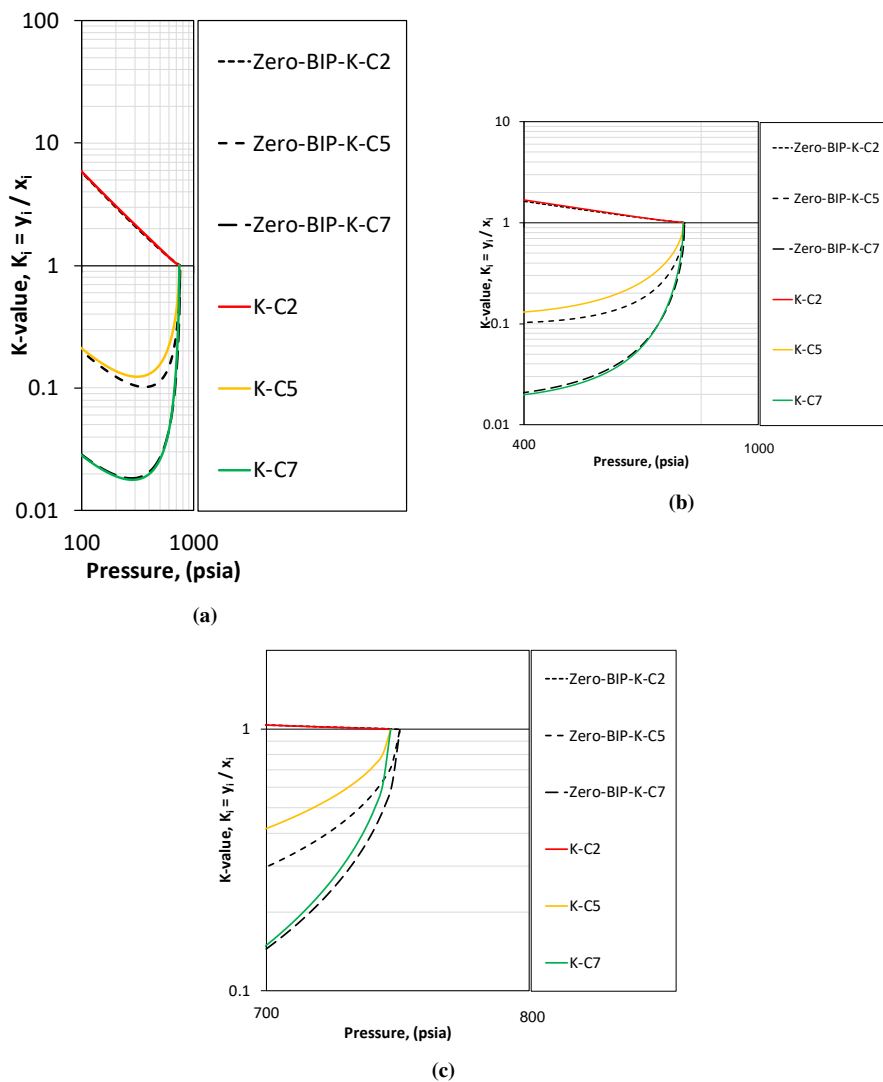
Figure B.28: K-value plot for case 1.2 described in Table 3.2 at 100°F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
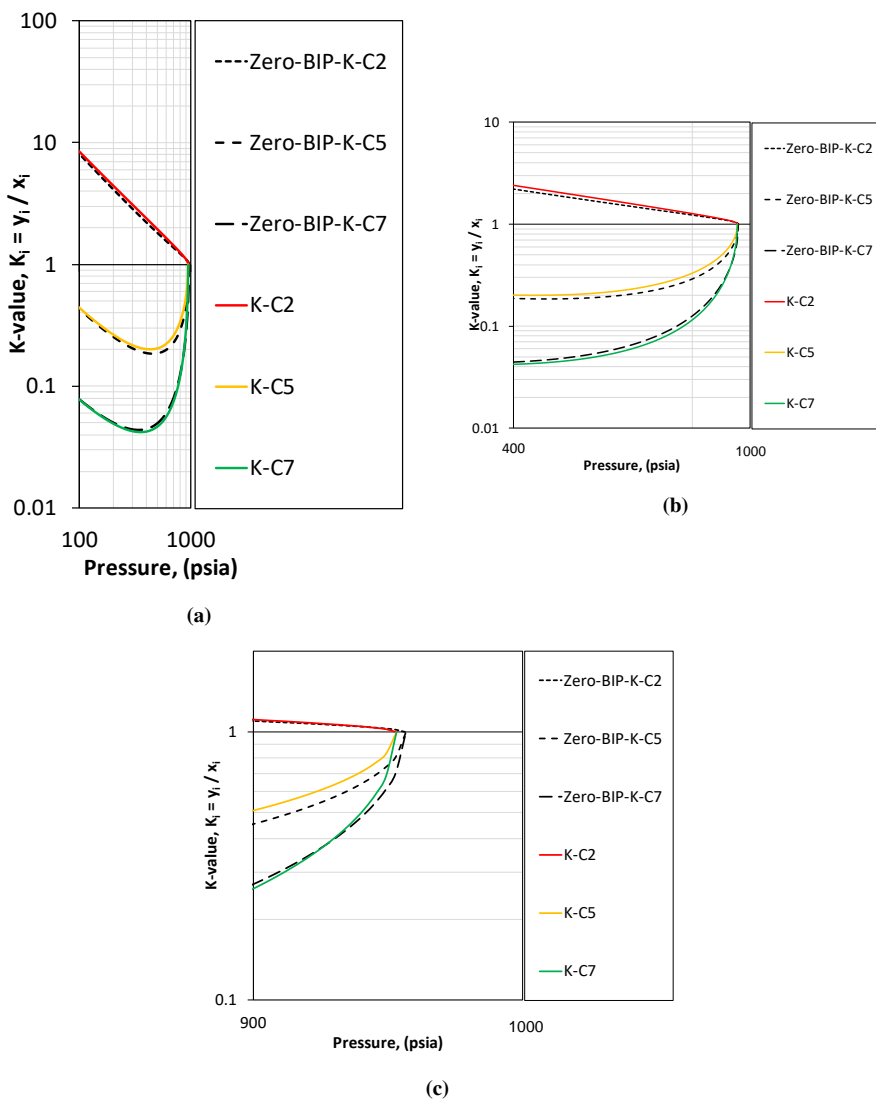
Figure B.29: K-value plot for case 1.2 described in Table 3.2 at $150°$F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
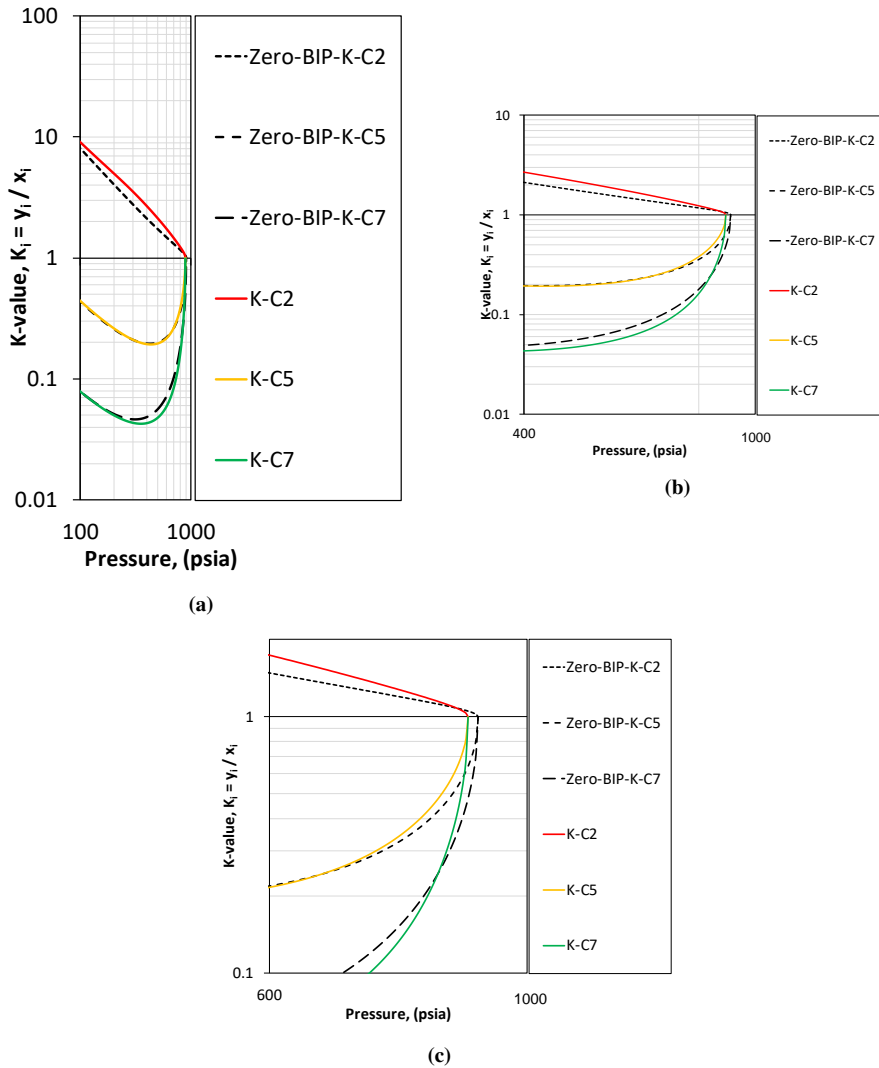
Figure B.30: K-value plot for case 2.1 described in Table 3.2 at $150°F$ for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
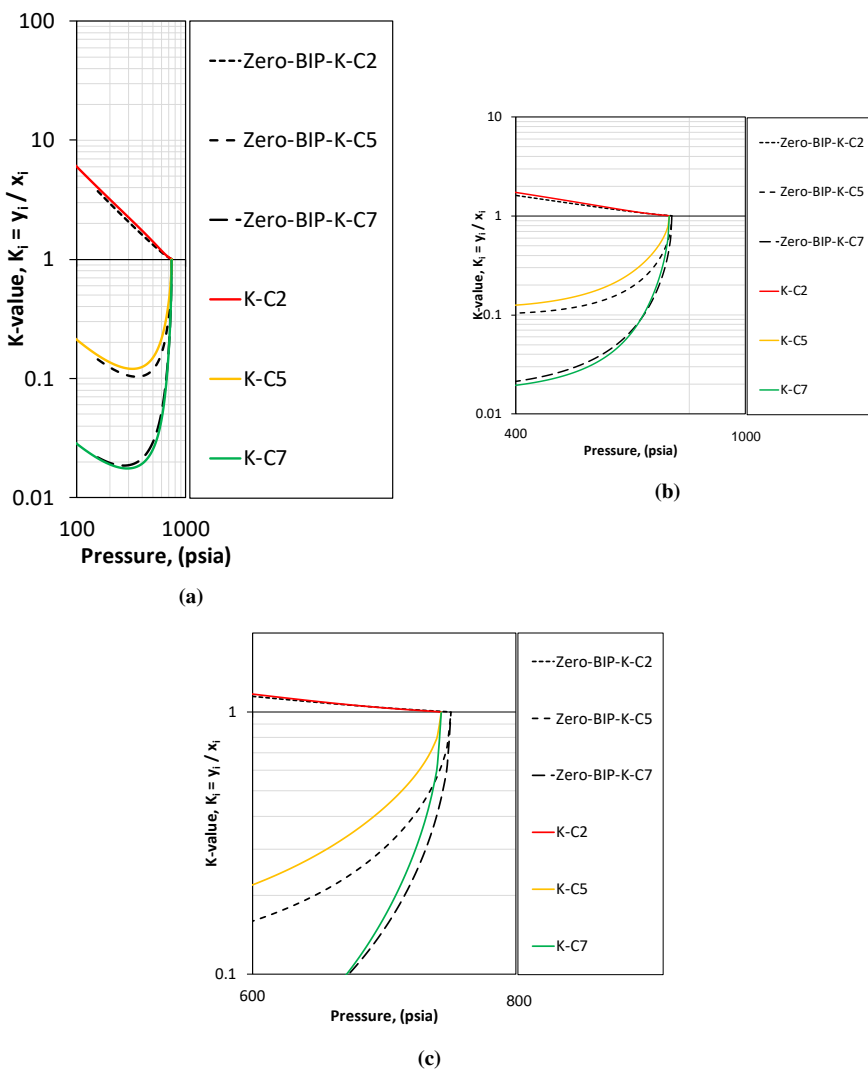
Figure B.31: K-value plot for case 2.2 described in Table 3.2 at $100°$F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
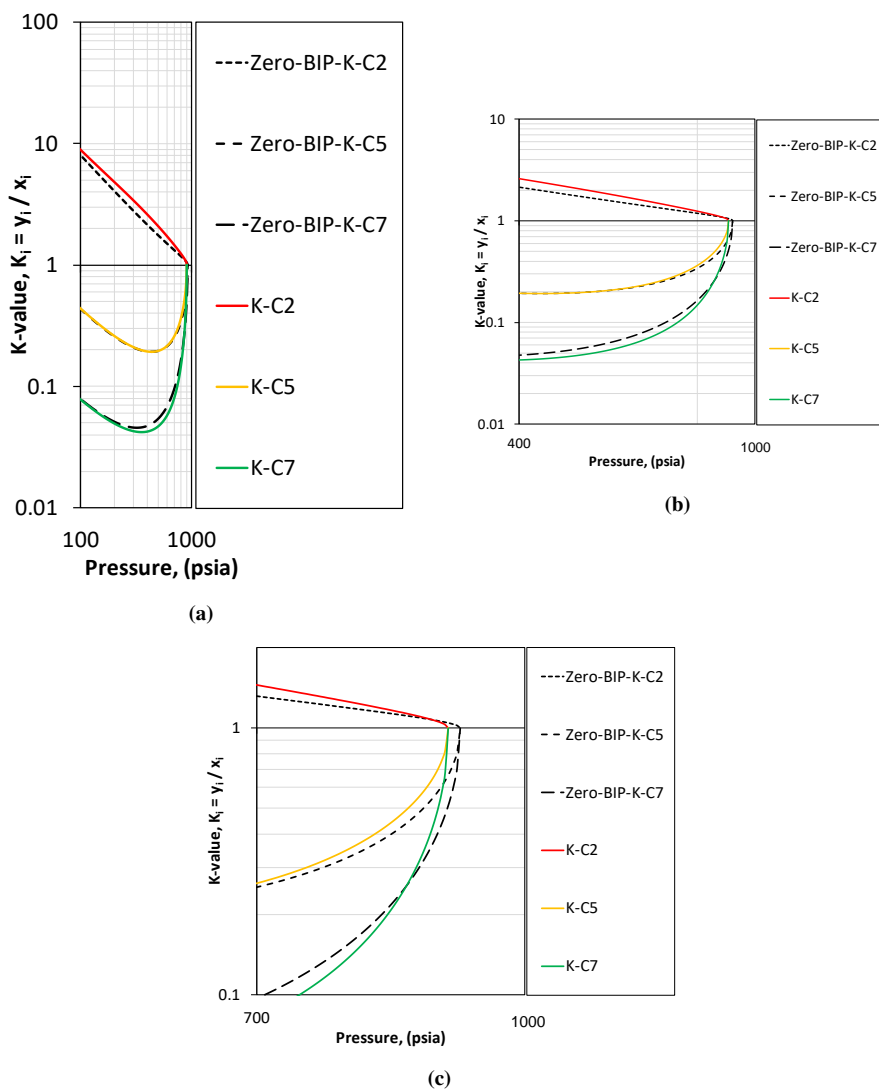
Figure B.32: K-value plot for case 2.2 described in Table 3.2 at $150°$F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
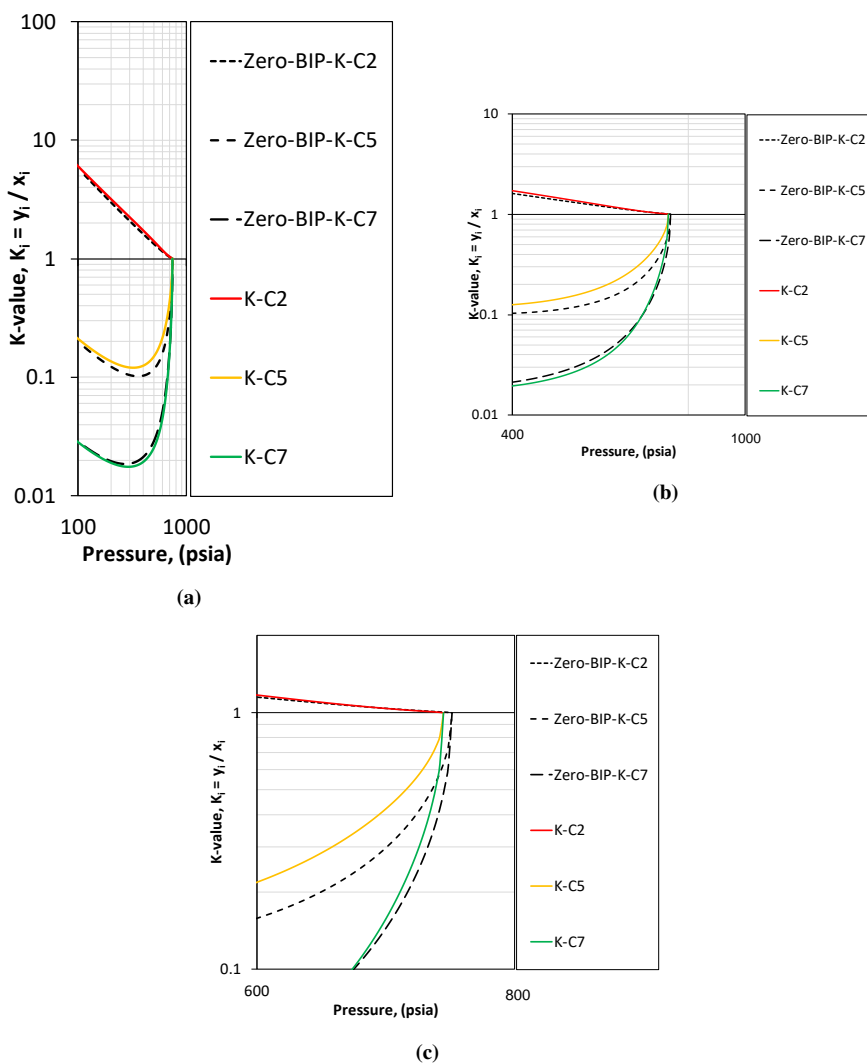
Figure B.33: K-value plot for case 3 described in Table 3.2 at $100°$F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.
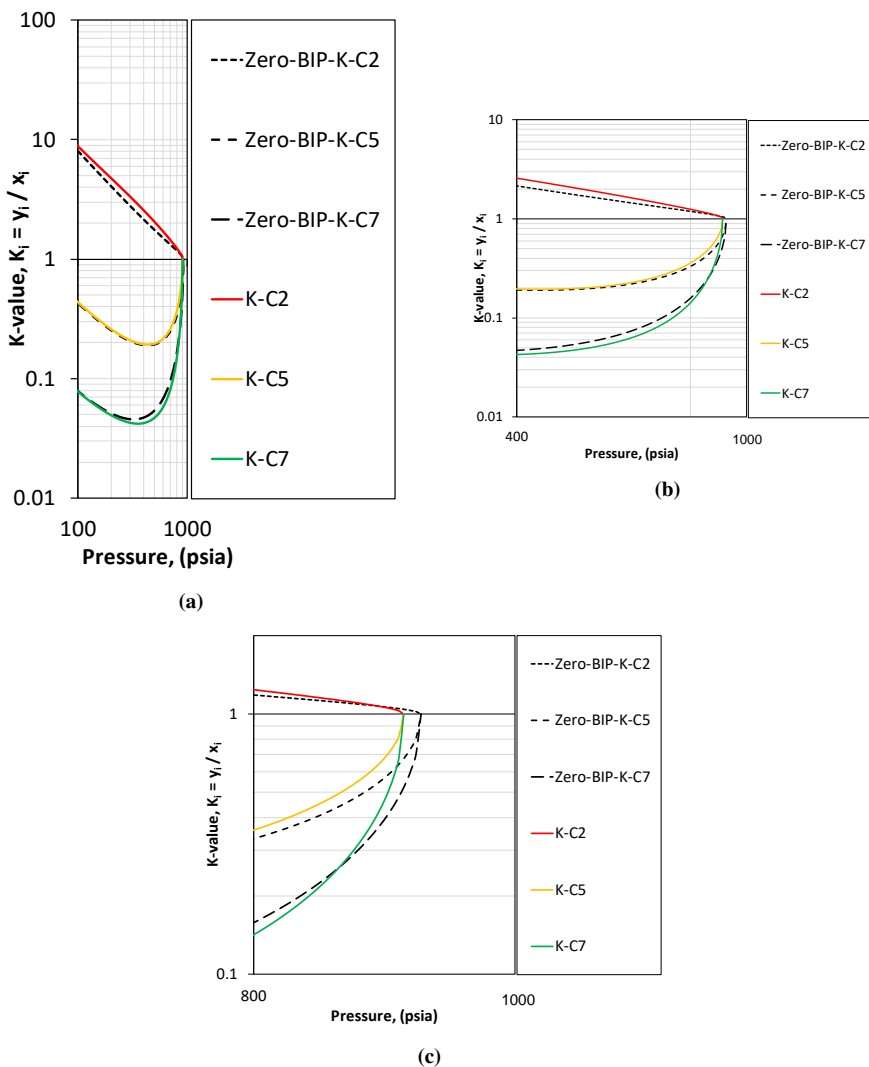
Figure B.34: K-value plot for case 3 described in Table 3.2 at $150°$F for a ternary system containing C_2 (red), C_5 (orange) and C_7 (green) with the EOS described in Table 4.4 and 4.6 as the zero-BIP case (black lines). (a) Full pressure range plot, (b) intermediate zoom at high pressure and (c) enhanced zoom near the convergence pressure.

C

Detailed Results for Regression

C.1   Selective Tuning Results

C.1.1   Gradient Method

Case 1 Fluid System
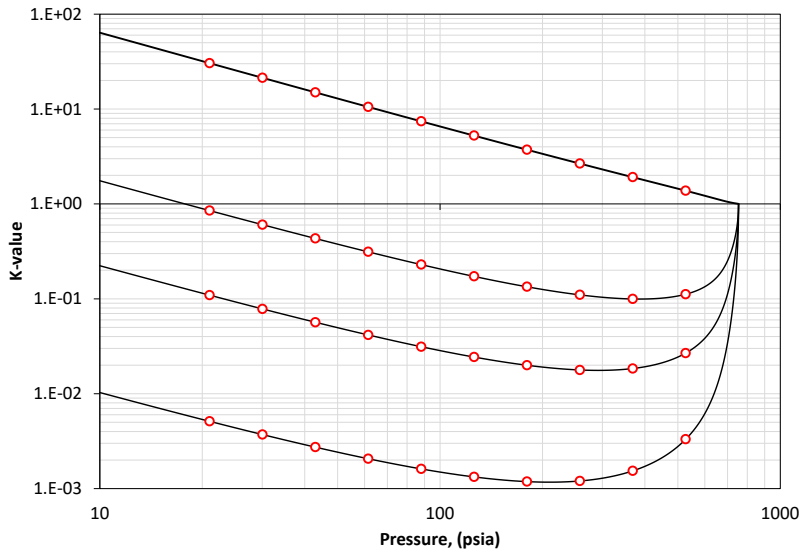
Figure C.1: K-value data after 14 iterations of the selective tuning approach with the gradient based method. The resulting K-value estimates of the selective tuning method are shown as red circles while the detailed synthetic K-value results are given as a solid black line.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.000	0.000
C5	0.000	0	0.000	0.000
C7	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	-0.001	0.000	-0.733
C5	-0.001	0	-0.002	-0.625
C7	0.000	-0.002	0	-0.269
C10	-0.733	-0.625	-0.269	0

Table C.1: Current BIP matrix and importance matrix for iteration 0 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.000	0.023
C5	0.000	0	0.000	0.004
C7	0.000	0.000	0	0.000
C10	0.023	0.004	0.000	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	-0.022	-0.023	-0.025
C5	-0.022	0	-0.001	-0.012
C7	-0.023	-0.001	0	-0.999
C10	-0.025	-0.012	-0.999	0

Table C.2: Current BIP matrix and importance matrix for iteration 1 of the gradient based method with fluid system 1.

Current BIP Matrix				
	C2	C5	C7	C10
C2	0	0.000	0.000	0.023
C5	0.000	0	0.000	0.004
C7	0.000	0.000	0	0.002
C10	0.023	0.004	0.002	0

Importance Matrix				
	C2	C5	C7	C10
C2	0	0.000	-0.752	-0.368
C5	0.000	0	-0.009	-0.469
C7	-0.752	-0.009	0	-0.281
C10	-0.368	-0.469	-0.281	0

Table C.3: Current BIP matrix and importance matrix for iteration 2 of the gradient based method with fluid system 1.

Current BIP Matrix				
	C2	C5	C7	C10
C2	0	0.000	0.015	0.023
C5	0.000	0	0.000	0.004
C7	0.015	0.000	0	0.002
C10	0.023	0.004	0.002	0

Importance Matrix				
	C2	C5	C7	C10
C2	0	-0.078	-0.004	0.149
C5	-0.078	0	-0.040	0.008
C7	-0.004	-0.040	0	0.985
C10	0.149	0.008	0.985	0

Table C.4: Current BIP matrix and importance matrix for iteration 3 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.015	0.023
C5	0.000	0	0.000	0.004
C7	0.015	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	-0.197	-0.367	-0.011
C5	-0.197	0	-0.240	0.721
C7	-0.367	-0.240	0	-0.499
C10	-0.011	0.721	-0.499	0

Table C.5: Current BIP matrix and importance matrix for iteration 4 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.015	0.023
C5	0.000	0	0.000	0.004
C7	0.015	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	-0.735	-0.305	0.335
C5	-0.735	0	-0.150	0.125
C7	-0.305	-0.150	0	-0.464
C10	0.335	0.125	-0.464	0

Table C.6: Current BIP matrix and importance matrix for iteration 5 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.006	0.015	0.023
C5	0.006	0	0.000	0.004
C7	0.015	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.039	0.111	0.051
C5	0.039	0	-0.014	0.991
C7	0.111	-0.014	0	-0.045
C10	0.051	0.991	-0.045	0

Table C.7: Current BIP matrix and importance matrix for iteration 6 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.006	0.014	0.023
C5	0.006	0	0.000	0.004
C7	0.014	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	-0.358	0.425	-0.616
C5	-0.358	0	-0.078	0.059
C7	0.425	-0.078	0	-0.550
C10	-0.616	0.059	-0.550	0

Table C.8: Current BIP matrix and importance matrix for iteration 7 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.006	0.014	0.023
C5	0.006	0	0.000	0.004
C7	0.014	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.080	0.508	-0.601
C5	-0.080	0	-0.091	0.126
C7	0.508	-0.091	0	0.592
C10	-0.601	0.126	0.592	0

Table C.9: Current BIP matrix and importance matrix for iteration 8 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.006	0.014	0.023
C5	0.006	0	0.000	0.004
C7	0.014	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.087	0.513	-0.627
C5	-0.087	0	-0.167	-0.050
C7	0.513	-0.167	0	0.553
C10	-0.627	-0.050	0.553	0

Table C.10: Current BIP matrix and importance matrix for iteration 9 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.006	0.014	0.023
C5	0.006	0	0.000	0.004
C7	0.014	0.000	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.087	0.513	-0.627
C5	-0.087	0	-0.167	-0.050
C7	0.513	-0.167	0	0.553
C10	-0.627	-0.050	0.553	0

Table C.11: Current BIP matrix and importance matrix for iteration 10 of the gradient based method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.006	0.013	0.023
C5	0.006	0	0.001	0.004
C7	0.013	0.001	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.099	0.003	0.089
C5	-0.099	0	-0.053	0.983
C7	0.003	-0.053	0	-0.117
C10	0.089	0.983	-0.117	0

Table C.12: Current BIP matrix and importance matrix for iteration 11 of the gradient based method with fluid system 1.

Current BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.007	0.013	0.023
C5	0.007	0	0.001	0.004
C7	0.013	0.001	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	-0.076	0.888	-0.446
C5	-0.076	0	-0.022	0.023
C7	0.888	-0.022	0	0.079
C10	-0.446	0.023	0.079	0

Table C.13: Current BIP matrix and importance matrix for iteration 12 of the gradient based method with fluid system 1.

Current BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.008	0.013	0.023
C5	0.008	0	0.001	0.004
C7	0.013	0.001	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.001	0.076	-0.106
C5	0.001	0	0.050	0.937
C7	0.076	0.050	0	-0.320
C10	-0.106	0.937	-0.320	0

Table C.14: Current BIP matrix and importance matrix for iteration 13 of the gradient based method with fluid system 1.

Current BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.008	0.013	0.023
C5	0.008	0	0.001	0.004
C7	0.013	0.001	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.174	0.437	-0.399
C5	0.174	0	0.583	-0.326
C7	0.437	0.583	0	-0.416
C10	-0.399	-0.326	-0.416	0

Table C.15: Current BIP matrix and importance matrix for iteration 14 of the gradient based method with fluid system 1.

Case 2 Fluid System
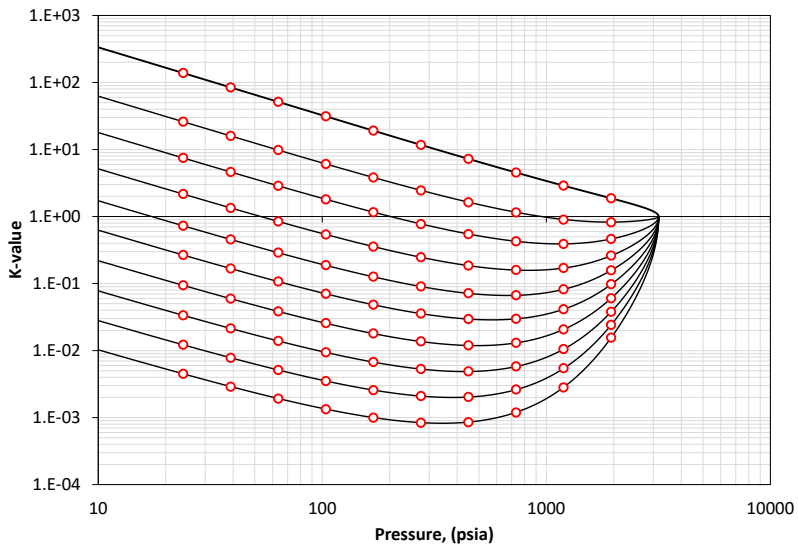
Figure C.2: K-value data after 14 iterations of the selective tuning approach with the gradient based method. The resulting K-value estimates of the selective tuning method are shown as red circles while the detailed synthetic K-value results are given as a solid black line.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.136	-0.157	-0.159	-0.159	-0.164	-0.221	-0.203	-0.332	-0.313
C2	-0.136	0	0.000	0.000	-0.006	-0.022	-0.025	0.006	-0.082	-0.078
C3	-0.157	0.000	0	-0.002	0.000	-0.005	-0.044	0.012	0.044	-0.049
C4	-0.159	0.000	-0.002	0	0.034	0.033	-0.003	0.048	0.017	-0.002
C5	-0.159	-0.006	0.000	0.034	0	0.027	0.023	0.066	0.110	0.077
C6	-0.164	-0.022	-0.005	0.033	0.027	0	0.006	0.094	0.140	0.111
C7	-0.221	-0.025	-0.044	-0.003	0.023	0.006	0	0.133	0.191	0.189
C8	-0.203	0.006	0.012	0.048	0.066	0.094	0.133	0	0.316	0.348
C9	-0.332	-0.082	0.044	0.017	0.110	0.140	0.191	0.316	0	0.425
C10	-0.313	-0.078	-0.049	-0.002	0.077	0.111	0.189	0.348	0.425	0

Table C.16: Current BIP matrix and importance matrix for iteration 0 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.017	0.031	0.034	0.034	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.017	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.031	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.031	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.002	0.005	0.003	0.003	0.003	0.003	0.003	-0.004	0.001
C2	0.002	0	-0.001	-0.025	-0.028	-0.069	-0.128	-0.205	-0.259	-0.310
C3	0.005	-0.001	0	-0.023	-0.123	-0.123	-0.123	-0.247	-0.251	-0.251
C4	0.003	-0.025	-0.023	0	-0.042	-0.041	-0.041	-0.175	-0.179	-0.180
C5	0.003	-0.028	-0.123	-0.042	0	-0.037	-0.047	-0.086	-0.126	-0.129
C6	0.003	-0.069	-0.123	-0.041	-0.037	0	-0.011	0.097	0.037	0.014
C7	0.003	-0.128	-0.123	-0.041	-0.047	-0.011	0	0.124	0.154	0.146
C8	0.003	-0.205	-0.247	-0.175	-0.086	0.097	0.124	0	0.303	0.321
C9	-0.004	-0.259	-0.251	-0.179	-0.126	0.037	0.154	0.303	0	0.359
C10	0.001	-0.310	-0.251	-0.180	-0.129	0.014	0.146	0.321	0.359	0

Table C.17: Current BIP matrix and importance matrix for iteration 1 of the gradient based method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.017	0.031	0.034	0.034	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.022	0.010
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.013
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.017	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.031	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.034	0.063	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.034	0.018	0.022	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.031	0.000	0.010	0.013	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.030	-0.042	-0.042	-0.041	-0.047	-0.046	-0.148	-0.143	-0.147
C2	-0.030	0	0.000	0.000	-0.001	-0.003	-0.004	-0.007	-0.012	0.003
C3	-0.042	0.000	0	0.000	-0.007	-0.004	-0.022	-0.027	-0.038	0.006
C4	-0.042	0.000	0.000	0	-0.054	-0.042	-0.054	-0.005	0.028	0.005
C5	-0.041	-0.001	-0.007	-0.054	0	-0.179	-0.169	-0.141	-0.222	-0.228
C6	-0.047	-0.003	-0.004	-0.042	-0.179	0	-0.151	-0.069	-0.075	-0.115
C7	-0.046	-0.004	-0.022	-0.054	-0.169	-0.151	0	0.192	0.215	0.210
C8	-0.148	-0.007	-0.027	-0.005	-0.141	-0.069	0.192	0	0.413	0.413
C9	-0.143	-0.012	-0.038	0.028	-0.222	-0.075	0.215	0.413	0	0.466
C10	-0.147	0.003	0.006	0.005	-0.228	-0.115	0.210	0.413	0.466	0

Table C.18: Current BIP matrix and importance matrix for iteration 2 of the gradient based method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.017	0.031	0.034	0.034	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.022	0.010
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.013
C5	0.000	0.000	0.000	0.000	0	0.009	0.008	0.000	0.000	0.000
C6	0.017	0.000	0.000	0.000	0.009	0	0.000	0.000	0.000	0.000
C7	0.031	0.000	0.000	0.000	0.008	0.000	0	0.000	0.000	0.000
C8	0.034	0.063	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.034	0.018	0.022	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.031	0.000	0.010	0.013	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.168	-0.226	-0.265	-0.270	-0.266	-0.280	-0.276	-0.282	-0.364
C2	-0.168	0	-0.001	0.000	0.002	-0.002	-0.005	-0.010	-0.015	0.004
C3	-0.226	-0.001	0	-0.009	-0.013	-0.024	-0.028	0.017	0.078	0.013
C4	-0.265	0.000	-0.009	0	0.025	0.017	0.020	0.048	0.098	0.012
C5	-0.270	0.002	-0.013	0.025	0	0.040	-0.005	0.005	0.087	-0.003
C6	-0.266	-0.002	-0.024	0.017	0.040	0	-0.040	-0.054	-0.044	-0.065
C7	-0.280	-0.005	-0.028	0.020	-0.005	-0.040	0	0.046	0.109	0.066
C8	-0.276	-0.010	0.017	0.048	0.005	-0.054	0.046	0	0.209	0.242
C9	-0.282	-0.015	0.078	0.098	0.087	-0.044	0.109	0.209	0	0.423
C10	-0.364	0.004	0.013	0.012	-0.003	-0.065	0.066	0.242	0.423	0

Table C.19: Current BIP matrix and importance matrix for iteration 3 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.035	0.035	0.032
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.022	0.010
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.013
C5	0.026	0.000	0.000	0.000	0	0.009	0.008	0.000	0.000	0.000
C6	0.020	0.000	0.000	0.000	0.009	0	0.000	0.000	0.000	0.000
C7	0.027	0.000	0.000	0.000	0.008	0.000	0	0.000	0.000	0.000
C8	0.035	0.063	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.035	0.018	0.022	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.032	0.000	0.010	0.013	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.004	0.001	0.002	-0.010	-0.001	0.002	-0.001	-0.003
C2	0.000	0	-0.001	0.000	-0.003	-0.007	-0.013	-0.021	-0.027	-0.038
C3	0.004	-0.029	0	-0.001	-0.001	-0.029	-0.060	-0.094	-0.108	-0.127
C4	0.001	0.000	-0.001	0	-0.001	-0.033	0.002	-0.001	-0.172	-0.178
C5	0.002	-0.003	-0.001	-0.001	0	-0.062	-0.095	-0.207	-0.239	-0.300
C6	-0.010	-0.007	-0.029	-0.033	-0.062	0	-0.210	-0.298	-0.419	-0.425
C7	-0.001	-0.013	-0.060	0.002	-0.095	-0.210	0	-0.091	-0.061	-0.052
C8	0.002	-0.021	-0.094	-0.001	-0.207	-0.298	-0.091	0	0.081	0.223
C9	-0.001	-0.027	-0.108	-0.172	-0.239	-0.419	-0.061	0.081	0	0.361
C10	-0.003	-0.038	-0.127	-0.178	-0.300	-0.425	-0.052	0.223	0.361	0

Table C.20: Current BIP matrix and importance matrix for iteration 4 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.035	0.035	0.032
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.022	0.010
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.015
C5	0.026	0.000	0.000	0.000	0	0.009	0.008	0.000	0.002	0.001
C6	0.020	0.000	0.000	0.000	0.009	0	0.003	0.003	0.000	0.000
C7	0.027	0.000	0.000	0.000	0.008	0.003	0	0.000	0.000	0.000
C8	0.035	0.063	0.000	0.000	0.000	0.003	0.000	0	0.000	0.000
C9	0.035	0.018	0.022	0.000	0.002	0.000	0.000	0.000	0	0.000
C10	0.032	0.000	0.010	0.015	0.001	0.000	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.018	-0.111	-0.132	-0.132	-0.156	-0.148	-0.132	-0.218	-0.201
C2	-0.018	0	-0.017	-0.019	-0.021	-0.037	-0.054	-0.069	-0.087	-0.071
C3	-0.111	-0.017	0	0.016	-0.059	-0.096	-0.079	-0.101	-0.195	-0.271
C4	-0.132	-0.019	0.016	0	0.019	-0.059	-0.096	-0.063	-0.026	-0.083
C5	-0.132	-0.021	-0.059	0.019	0	-0.058	-0.009	0.043	0.043	-0.031
C6	-0.156	-0.037	-0.096	-0.059	-0.058	0	0.030	0.031	0.070	-0.174
C7	-0.148	-0.054	-0.079	-0.096	-0.009	0.030	0	0.113	0.196	0.172
C8	-0.132	-0.069	-0.101	-0.063	0.043	0.031	0.113	0	0.348	0.371
C9	-0.218	-0.087	-0.195	-0.026	0.043	0.070	0.196	0.348	0	0.478
C10	-0.201	-0.071	-0.271	-0.083	-0.031	-0.174	0.172	0.371	0.478	0

Table C.21: Current BIP matrix and importance matrix for iteration 5 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.035	0.042	0.030
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.015
C5	0.026	0.000	0.000	0.000	0	0.009	0.008	0.000	0.002	0.001
C6	0.020	0.000	0.000	0.000	0.009	0	0.003	0.003	0.000	0.001
C7	0.027	0.000	0.000	0.000	0.008	0.003	0	0.000	0.000	0.001
C8	0.035	0.063	0.000	0.000	0.000	0.003	0.000	0	0.000	0.000
C9	0.042	0.018	0.036	0.000	0.002	0.000	0.000	0.000	0	0.000
C10	0.030	0.000	0.000	0.015	0.001	0.001	0.001	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.006	-0.012	-0.016	-0.026	-0.147	-0.010	-0.070	0.050	0.060
C2	-0.006	0	-0.006	-0.003	-0.021	-0.037	-0.070	-0.105	-0.144	-0.174
C3	-0.012	-0.006	0	-0.007	0.013	0.013	0.032	0.028	0.028	0.019
C4	-0.016	-0.003	-0.007	0	-0.187	-0.244	-0.204	-0.353	-0.343	-0.352
C5	-0.026	-0.021	0.013	-0.187	0	0.006	0.033	-0.052	-0.107	-0.294
C6	-0.147	-0.037	0.013	-0.244	0.006	0	0.348	0.283	0.165	0.087
C7	-0.010	-0.070	0.032	-0.204	0.033	0.348	0	0.093	0.074	-0.060
C8	-0.070	-0.105	0.028	-0.353	-0.052	0.283	0.093	0	0.099	0.020
C9	0.050	-0.144	0.028	-0.343	-0.107	0.165	0.074	0.099	0	0.160
C10	0.060	-0.174	0.019	-0.352	-0.294	0.087	-0.060	0.020	0.160	0

Table C.22: Current BIP matrix and importance matrix for iteration 6 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.035	0.042	0.030
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.001	0.000	0.000	0.015
C5	0.026	0.000	0.000	0.000	0	0.009	0.008	0.000	0.002	0.001
C6	0.020	0.000	0.000	0.000	0.009	0	0.001	0.003	0.000	0.001
C7	0.027	0.000	0.000	0.001	0.008	0.001	0	0.000	0.000	0.001
C8	0.035	0.063	0.000	0.000	0.000	0.003	0.000	0	0.000	0.000
C9	0.042	0.018	0.036	0.000	0.002	0.000	0.000	0.000	0	0.000
C10	0.030	0.002	0.000	0.015	0.001	0.001	0.001	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.129	-0.134	-0.132	-0.111	-0.321	-0.089	-0.333	0.025	-0.036
C2	-0.129	0	-0.003	-0.010	-0.009	0.018	-0.001	0.005	0.004	0.039
C3	-0.134	-0.003	0	-0.048	-0.110	-0.101	-0.047	-0.060	0.013	-0.018
C4	-0.132	-0.010	-0.048	0	0.026	0.079	0.085	0.258	0.139	0.132
C5	-0.111	-0.009	-0.110	0.026	0	0.276	0.209	0.174	0.073	0.000
C6	-0.321	0.018	-0.101	0.079	0.276	0	0.049	-0.122	-0.144	-0.203
C7	-0.089	-0.001	-0.047	0.085	0.209	0.049	0	-0.129	-0.280	-0.434
C8	-0.333	0.005	-0.060	0.258	0.174	-0.122	-0.129	0	-0.092	-0.130
C9	0.025	0.004	0.013	0.139	0.073	-0.144	-0.280	-0.092	0	-0.134
C10	-0.036	0.039	-0.018	0.132	0.000	-0.203	-0.434	-0.130	-0.134	0

Table C.23: Current BIP matrix and importance matrix for iteration 7 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.037	0.042	0.030
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.001	0.000	0.000	0.015
C5	0.026	0.000	0.000	0.000	0	0.000	0.007	0.003	0.002	0.001
C6	0.020	0.000	0.000	0.000	0.000	0	0.001	0.003	0.001	0.001
C7	0.027	0.000	0.000	0.001	0.007	0.001	0	0.000	0.000	0.001
C8	0.037	0.063	0.000	0.000	0.003	0.003	0.000	0	0.000	0.000
C9	0.042	0.018	0.036	0.000	0.002	0.001	0.000	0.000	0	0.000
C10	0.030	0.002	0.000	0.015	0.001	0.001	0.001	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.028	-0.028	-0.117	-0.019	-0.059	0.004	-0.044	0.202	0.147
C2	-0.028	0	0.000	-0.012	-0.007	0.012	0.025	-0.026	-0.045	-0.020
C3	-0.028	0.000	0	0.000	-0.028	-0.092	-0.131	-0.008	-0.131	-0.077
C4	-0.117	-0.012	0.000	0	-0.004	-0.004	0.041	0.269	0.088	0.140
C5	-0.019	-0.007	-0.028	-0.004	0	-0.069	0.049	0.055	-0.328	-0.331
C6	-0.059	0.012	-0.092	-0.004	-0.069	0	0.099	0.190	-0.066	-0.039
C7	0.004	0.025	-0.131	0.041	0.049	0.099	0	0.342	-0.103	-0.152
C8	-0.044	-0.026	-0.008	0.269	0.055	0.190	0.342	0	0.156	0.063
C9	0.202	-0.045	-0.131	0.088	-0.328	-0.066	-0.103	0.156	0	-0.550
C10	0.147	-0.020	-0.077	0.140	-0.331	-0.039	-0.152	0.063	-0.550	0

Table C.24: Current BIP matrix and importance matrix for iteration 8 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.037	0.040	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.001	0.000	0.000	0.015
C5	0.026	0.000	0.000	0.000	0	0.000	0.007	0.003	0.000	0.003
C6	0.020	0.000	0.000	0.000	0.000	0	0.001	0.003	0.001	0.001
C7	0.027	0.000	0.000	0.001	0.007	0.001	0	0.000	0.000	0.001
C8	0.037	0.063	0.000	0.000	0.003	0.003	0.000	0	0.000	0.000
C9	0.040	0.018	0.036	0.000	0.000	0.001	0.000	0.000	0	0.000
C10	0.031	0.002	0.000	0.015	0.003	0.001	0.001	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.064	-0.084	-0.084	-0.020	-0.016	0.129	-0.110	-0.035	0.020
C2	-0.064	0	-0.020	-0.005	-0.009	0.000	0.018	-0.033	-0.048	-0.006
C3	-0.084	-0.020	0	-0.129	-0.075	-0.075	-0.174	-0.065	-0.065	-0.042
C4	-0.084	-0.005	-0.129	0	0.002	0.000	0.199	0.143	0.171	0.230
C5	-0.020	-0.009	-0.075	0.002	0	0.090	0.025	0.109	-0.057	-0.146
C6	-0.016	0.000	-0.075	0.000	0.090	0	0.103	0.005	-0.174	-0.246
C7	0.129	0.018	-0.174	0.199	0.025	0.103	0	0.465	0.192	0.248
C8	-0.110	-0.033	-0.065	0.143	0.109	0.005	0.465	0	0.351	0.379
C9	-0.035	-0.048	-0.065	0.171	-0.057	-0.174	0.192	0.351	0	-0.085
C10	0.020	-0.006	-0.042	0.230	-0.146	-0.246	0.248	0.379	-0.085	0

**Table C.25:** Current BIP matrix and importance matrix for iteration 9 of the gradient based method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.037	0.040	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.000	0.001	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.002	0.000	0.000	0.014
C5	0.026	0.000	0.000	0.000	0	0.000	0.007	0.003	0.000	0.003
C6	0.020	0.000	0.000	0.000	0.000	0	0.001	0.003	0.000	0.002
C7	0.027	0.000	0.001	0.002	0.007	0.001	0	0.000	0.001	0.000
C8	0.037	0.063	0.000	0.000	0.003	0.003	0.000	0	0.000	0.000
C9	0.040	0.018	0.036	0.000	0.000	0.000	0.001	0.000	0	0.000
C10	0.031	0.002	0.000	0.014	0.003	0.002	0.000	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.050	-0.082	-0.099	-0.107	-0.093	-0.044	0.129	-0.011	-0.072
C2	-0.050	0	-0.014	-0.034	-0.040	0.011	0.005	-0.006	-0.068	-0.093
C3	-0.082	-0.014	0	-0.089	-0.102	-0.225	-0.058	0.007	0.034	0.032
C4	-0.099	-0.034	-0.089	0	0.291	0.181	0.092	0.215	0.063	0.028
C5	-0.107	-0.040	-0.102	0.291	0	0.104	0.344	0.260	0.310	-0.112
C6	-0.093	0.011	-0.225	0.181	0.104	0	0.077	0.163	0.083	-0.196
C7	-0.044	0.005	-0.058	0.092	0.344	0.077	0	0.439	-0.070	0.014
C8	0.129	-0.006	0.007	0.215	0.260	0.163	0.439	0	-0.096	-0.052
C9	-0.011	-0.068	0.034	0.063	0.310	0.083	-0.070	-0.096	0	-0.276
C10	-0.072	-0.093	0.032	0.028	-0.112	-0.196	0.014	-0.052	-0.276	0

Table C.26: Current BIP matrix and importance matrix for iteration 10 of the gradient based method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.037	0.040	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.001	0.001	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.001	0.002	0.000	0.000	0.014
C5	0.026	0.000	0.000	0.000	0	0.000	0.006	0.003	0.001	0.003
C6	0.020	0.000	0.001	0.001	0.000	0	0.001	0.003	0.000	0.002
C7	0.027	0.000	0.001	0.002	0.006	0.001	0	0.000	0.001	0.000
C8	0.037	0.063	0.000	0.000	0.003	0.003	0.000	0	0.000	0.000
C9	0.040	0.018	0.036	0.000	0.001	0.000	0.001	0.000	0	0.000
C10	0.031	0.002	0.000	0.014	0.003	0.002	0.000	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.057	-0.139	-0.110	0.002	0.007	0.029	-0.075	-0.398	-0.012
C2	-0.057	0	0.000	-0.006	-0.009	0.005	-0.021	0.067	-0.041	-0.070
C3	-0.139	0.000	0	0.000	-0.126	-0.055	-0.068	0.192	0.063	0.147
C4	-0.110	-0.006	0.000	0	-0.051	-0.097	-0.081	0.175	0.041	0.213
C5	0.002	-0.009	-0.126	-0.051	0	-0.090	-0.037	0.083	-0.177	-0.158
C6	0.007	0.005	-0.055	-0.097	-0.090	0	0.134	0.186	-0.154	-0.032
C7	0.029	-0.021	-0.068	-0.081	-0.037	0.134	0	0.079	-0.438	-0.482
C8	-0.075	0.067	0.192	0.175	0.083	0.186	0.079	0	-0.065	0.032
C9	-0.398	-0.041	0.063	0.041	-0.177	-0.154	-0.438	-0.065	0	-0.142
C10	-0.012	-0.070	0.147	0.213	-0.158	-0.032	-0.482	0.032	-0.142	0

Table C.27: Current BIP matrix and importance matrix for iteration 11 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.037	0.040	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.001	0.001	0.000	0.036	0.000
C4	0.000	0.000	0.000	0	0.000	0.001	0.002	0.002	0.000	0.013
C5	0.026	0.000	0.000	0.000	0	0.000	0.006	0.003	0.000	0.004
C6	0.020	0.000	0.001	0.001	0.000	0	0.001	0.002	0.001	0.002
C7	0.027	0.000	0.001	0.002	0.006	0.001	0	0.000	0.001	0.001
C8	0.037	0.063	0.000	0.002	0.003	0.002	0.000	0	0.000	0.000
C9	0.040	0.018	0.036	0.000	0.000	0.001	0.001	0.000	0	0.000
C10	0.031	0.002	0.000	0.013	0.004	0.002	0.001	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.046	-0.101	-0.149	-0.149	-0.149	0.025	-0.129	0.060	0.089
C2	-0.046	0	-0.055	-0.046	-0.001	0.000	0.007	-0.062	-0.099	-0.018
C3	-0.101	-0.055	0	0.050	0.139	0.012	-0.062	0.230	0.161	0.241
C4	-0.149	-0.046	0.050	0	0.116	0.055	0.165	0.113	0.140	0.029
C5	-0.149	-0.001	0.139	0.116	0	0.120	0.166	-0.008	0.150	0.021
C6	-0.149	0.000	0.012	0.055	0.120	0	0.079	0.004	0.002	0.046
C7	0.025	0.007	-0.062	0.165	0.166	0.079	0	-0.121	0.188	-0.154
C8	-0.129	-0.062	0.230	0.113	-0.008	0.004	-0.121	0	-0.495	-0.421
C9	0.060	-0.099	0.161	0.140	0.150	0.002	0.188	-0.495	0	-0.253
C10	0.089	-0.018	0.241	0.029	0.021	0.046	-0.154	-0.421	-0.253	0

Table C.28: Current BIP matrix and importance matrix for iteration 12 of the gradient based method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.026	0.020	0.027	0.037	0.040	0.031
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.001	0.001	0.000	0.035	0.000
C4	0.000	0.000	0.000	0	0.000	0.001	0.002	0.002	0.000	0.013
C5	0.026	0.000	0.000	0.000	0	0.000	0.006	0.003	0.000	0.004
C6	0.020	0.000	0.001	0.001	0.000	0	0.001	0.002	0.001	0.002
C7	0.027	0.000	0.001	0.002	0.006	0.001	0	0.000	0.001	0.001
C8	0.037	0.063	0.000	0.002	0.003	0.002	0.000	0	0.000	0.000
C9	0.040	0.018	0.035	0.000	0.000	0.001	0.001	0.000	0	0.000
C10	0.031	0.002	0.000	0.013	0.004	0.002	0.001	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.190	-0.248	-0.306	-0.252	-0.252	-0.046	0.045	-0.342	0.054
C2	-0.190	0	-0.058	-0.024	-0.003	0.024	0.009	0.077	-0.128	-0.012
C3	-0.248	-0.058	0	0.058	0.058	-0.091	-0.066	0.064	-0.012	0.104
C4	-0.306	-0.024	0.058	0	-0.111	-0.111	-0.106	0.124	0.186	0.059
C5	-0.252	-0.003	0.058	-0.111	0	-0.029	0.047	-0.244	-0.078	-0.252
C6	-0.252	0.024	-0.091	-0.111	-0.029	0	0.021	0.001	0.101	0.126
C7	-0.046	0.009	-0.066	-0.106	0.047	0.021	0	0.071	0.043	0.133
C8	0.045	0.077	0.064	0.124	-0.244	0.001	0.071	0	-0.117	-0.015
C9	-0.342	-0.128	-0.012	0.186	-0.078	0.101	0.043	-0.117	0	0.457
C10	0.054	-0.012	0.104	0.059	-0.252	0.126	0.133	-0.015	0.457	0

Table C.29: Current BIP matrix and importance matrix for iteration 13 of the gradient based method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.016	0.000	0.004	0.016	0.017	0.027	0.037	0.039	0.031
C2	0.016	0	0.000	0.000	0.000	0.000	0.000	0.063	0.018	0.002
C3	0.000	0.000	0	0.000	0.000	0.001	0.001	0.000	0.035	0.000
C4	0.004	0.000	0.000	0	0.000	0.001	0.002	0.002	0.000	0.013
C5	0.016	0.000	0.000	0.000	0	0.000	0.006	0.002	0.000	0.004
C6	0.017	0.000	0.001	0.001	0.000	0	0.001	0.002	0.001	0.002
C7	0.027	0.000	0.001	0.002	0.006	0.001	0	0.000	0.001	0.001
C8	0.037	0.063	0.000	0.002	0.002	0.002	0.000	0	0.000	0.000
C9	0.039	0.018	0.035	0.000	0.000	0.001	0.001	0.000	0	0.000
C10	0.031	0.002	0.000	0.013	0.004	0.002	0.001	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.043	0.043	-0.079	-0.110	-0.207	-0.027	0.048	-0.034	-0.016
C2	0.043	0	0.017	0.017	0.021	-0.035	0.031	-0.059	-0.074	-0.141
C3	0.043	0.017	0	0.015	0.015	-0.036	-0.046	-0.119	0.010	-0.206
C4	-0.079	0.017	0.015	0	-0.015	-0.015	0.119	-0.026	0.155	-0.015
C5	-0.110	0.021	0.015	-0.015	0	0.056	0.125	0.035	0.120	-0.021
C6	-0.207	-0.035	-0.036	-0.015	0.056	0	0.153	-0.108	0.158	-0.173
C7	-0.027	0.031	-0.046	0.119	0.125	0.153	0	-0.344	-0.004	-0.166
C8	0.048	-0.059	-0.119	-0.026	0.035	-0.108	-0.344	0	-0.404	-0.576
C9	-0.034	-0.074	0.010	0.155	0.120	0.158	-0.004	-0.404	0	-0.180
C10	-0.016	-0.141	-0.206	-0.015	-0.021	-0.173	-0.166	-0.576	-0.180	0

Table C.30: Current BIP matrix and importance matrix for iteration 14 of the gradient based method with fluid system 2.

C.1.2   PCC Method

Case 1 Fluid System
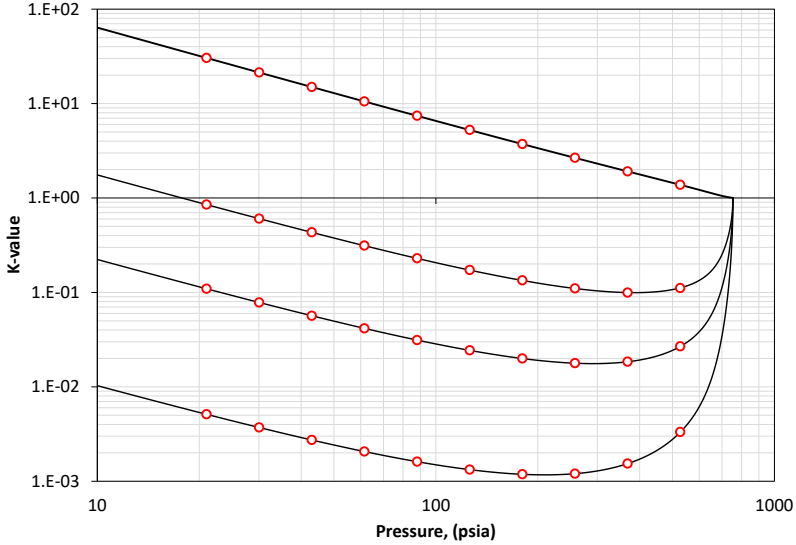
Figure C.3: K-value data after 14 iterations of the selective tuning approach with the PCC based method. The resulting K-value estimates of the selective tuning method are shown as red circles while the detailed synthetic K-value results are given as a solid black line.

Current BIP Matrix				
	C2	C5	C7	C10
C2	0	0.000	0.000	0.000
C5	0.000	0	0.000	0.000
C7	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0

Importance Matrix				
	C2	C5	C7	C10
C2	0	-0.059	-0.011	-0.671
C5	-0.059	0	0.220	-0.637
C7	-0.011	0.220	0	-0.434
C10	-0.671	-0.637	-0.434	0

Table C.31: Current BIP matrix and importance matrix for iteration 0 of the PCC method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.000	0.000	0.023
C5	0.000	0	0.000	0.004
C7	0.000	0.000	0	0.000
C10	0.023	0.004	0.000	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.009	-0.155	0.065
C5	-0.009	0	-0.186	0.009
C7	-0.155	-0.186	0	-0.997
C10	0.065	0.009	-0.997	0

Table C.32: Current BIP matrix and importance matrix for iteration 1 of the PCC method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.000	0.000	0.023
C5	0.000	0	0.010	0.004
C7	0.000	0.010	0	0.002
C10	0.023	0.004	0.002	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.081	-0.436	-0.093
C5	-0.081	0	0.072	0.894
C7	-0.436	0.072	0	0.071
C10	-0.093	0.894	0.071	0

Table C.33: Current BIP matrix and importance matrix for iteration 2 of the PCC method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.000	0.014	0.023
C5	0.000	0	0.010	0.004
C7	0.014	0.010	0	0.002
C10	0.023	0.004	0.002	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	-0.029	0.066	-0.060
C5	-0.029	0	-0.033	0.234
C7	0.066	-0.033	0	0.971
C10	-0.060	0.234	0.971	0

Table C.34: Current BIP matrix and importance matrix for iteration 3 of the PCC method with fluid system 1.

Current BIP Matrix			
	C2	C5	C7	C10
C2	0	0.000	0.014	0.023
C5	0.000	0	0.010	0.004
C7	0.014	0.010	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix			
	C2	C5	C7	C10
C2	0	0.120	-0.109	-0.228
C5	0.120	0	-0.076	0.494
C7	-0.109	-0.076	0	0.816
C10	-0.228	0.494	0.816	0

Table C.35: Current BIP matrix and importance matrix for iteration 4 of the PCC method with fluid system 1.

Current BIP Matrix				
	C2	C5	C7	C10
C2	0	0.000	0.014	0.023
C5	0.000	0	0.010	0.004
C7	0.014	0.010	0	0.001
C10	0.023	0.004	0.001	0

Importance Matrix				
	C2	C5	C7	C10
C2	0	-0.073	-0.010	-0.298
C5	-0.073	0	0.076	0.442
C7	-0.010	0.076	0	0.787
C10	-0.298	0.442	0.787	0

Table C.36: Current BIP matrix and importance matrix for iteration 5 of the PCC method with fluid system 1.

Case 2 Fluid Model
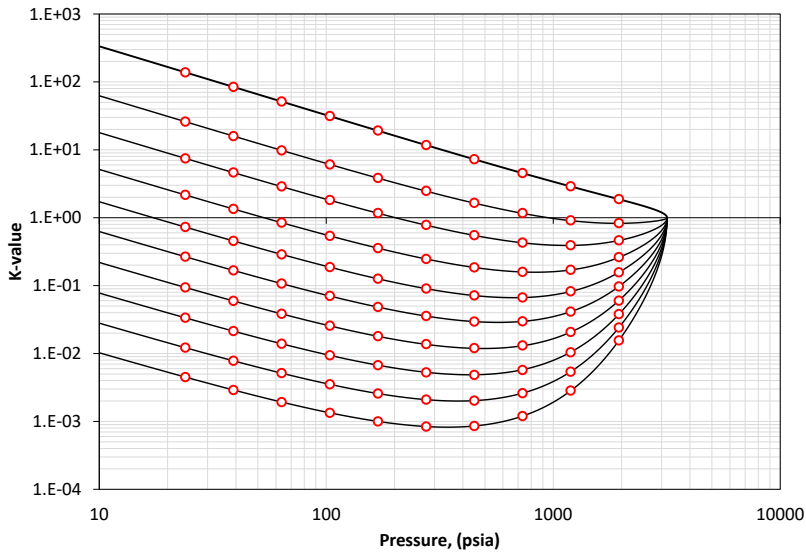
Figure C.4: K-value data after 14 iterations of the selective tuning approach with the PCC based method. The resulting K-value estimates of the selective tuning method are shown as red circles while the detailed synthetic K-value results are given as a solid black line.

Current BIP Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	0.009	-0.031	-0.199	-0.078	-0.333	-0.203	-0.188	-0.385	-0.407
C2	0.009	0	-0.009	-0.123	0.039	-0.125	-0.180	-0.091	0.049	-0.254
C3	-0.031	-0.009	0	0.040	-0.014	0.070	0.201	-0.123	0.012	0.021
C4	-0.199	-0.123	0.040	0	-0.057	0.140	-0.024	-0.069	-0.058	0.135
C5	-0.078	0.039	-0.014	-0.057	0	0.003	-0.097	0.035	0.125	0.121
C6	-0.333	-0.125	0.070	0.140	0.003	0	-0.011	-0.061	0.125	-0.005
C7	-0.203	-0.180	0.201	-0.024	-0.097	-0.011	0	0.278	0.169	0.310
C8	-0.188	-0.091	-0.123	-0.069	0.035	-0.061	0.278	0	0.385	0.429
C9	-0.385	0.049	0.012	-0.058	0.125	0.125	0.169	0.385	0	0.385
C10	-0.407	-0.254	0.021	0.135	0.121	-0.005	0.310	0.429	0.385	0

Table C.37: Current BIP matrix and importance matrix for iteration 0 of the PCC method with fluid system 2.

Current BIP Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	0.000	0.000	0.000	0.000	0.051	0.036	0.000	0.031	0.041
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.060
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.051	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.036	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.031	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.041	0.060	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	0.131	-0.024	0.007	-0.116	-0.035	-0.016	-0.248	0.089	-0.007
C2	0.131	0	0.021	0.018	-0.072	-0.104	-0.196	0.116	-0.003	-0.032
C3	-0.024	0.021	0	-0.032	-0.105	0.025	-0.216	-0.189	-0.467	-0.549
C4	0.007	0.018	-0.032	0	-0.213	-0.204	-0.103	-0.060	-0.291	-0.422
C5	-0.116	-0.072	-0.105	-0.213	0	-0.115	-0.036	-0.026	0.016	-0.217
C6	-0.035	-0.104	0.025	-0.204	-0.115	0	0.034	-0.151	0.096	-0.071
C7	-0.016	-0.196	-0.216	-0.103	-0.036	0.034	0	0.060	0.037	0.065
C8	-0.248	0.116	-0.189	-0.060	-0.026	-0.151	0.060	0	0.216	0.434
C9	0.089	-0.003	-0.467	-0.291	0.016	0.096	0.037	0.216	0	0.160
C10	-0.007	-0.032	-0.549	-0.422	-0.217	-0.071	0.065	0.434	0.160	0

Table C.38: Current BIP matrix and importance matrix for iteration 1 of the PCC method with fluid system 2.

Current BIP Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	0.000	0.000	0.000	0.000	0.051	0.036	0.019	0.031	0.041
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.060
C3	0.000	0.000	0	0.000	0.000	0.000	0.034	0.000	0.002	0.014
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.017	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.007
C6	0.051	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.036	0.000	0.034	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.031	0.000	0.002	0.017	0.000	0.000	0.000	0.000	0	0.000
C10	0.041	0.060	0.014	0.000	0.007	0.000	0.000	0.000	0.000	0

Importance Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	-0.111	0.069	-0.135	0.028	0.011	0.229	0.091	-0.047	0.073
C2	-0.111	0	0.008	-0.115	-0.340	0.189	0.139	-0.073	0.104	0.206
C3	0.069	0.008	0	-0.017	-0.065	-0.015	-0.109	0.152	0.245	-0.139
C4	-0.135	-0.115	-0.017	0	-0.162	-0.137	-0.038	-0.079	-0.103	0.117
C5	0.028	-0.340	-0.065	-0.162	0	-0.120	0.105	0.169	-0.049	0.170
C6	0.011	0.189	-0.015	-0.137	-0.120	0	-0.461	-0.317	-0.503	-0.445
C7	0.229	0.139	-0.109	-0.038	0.105	-0.461	0	-0.228	0.079	-0.056
C8	0.091	-0.073	0.152	-0.079	0.169	-0.317	-0.228	0	0.140	0.122
C9	-0.047	0.104	0.245	-0.103	-0.049	-0.503	0.079	0.140	0	0.309
C10	0.073	0.206	-0.139	0.117	0.170	-0.445	-0.056	0.122	0.309	0

Table C.39: Current BIP matrix and importance matrix for iteration 2 of the PCC method with fluid system 2.

Current BIP Matrix										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	0.000	0.000	0.000	0.000	0.051	0.038	0.019	0.031	0.041
C2	0.000	0	0.000	0.000	0.058	0.000	0.000	0.000	0.000	0.054
C3	0.000	0.000	0	0.000	0.000	0.000	0.034	0.000	0.003	0.014
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.017	0.000
C5	0.000	0.058	0.000	0.000	0	0.000	0.000	0.000	0.000	0.007
C6	0.051	0.000	0.000	0.000	0.000	0	0.005	0.002	0.000	0.002
C7	0.038	0.000	0.034	0.000	0.000	0.005	0	0.001	0.000	0.000
C8	0.019	0.000	0.000	0.000	0.000	0.002	0.001	0	0.000	0.000
C9	0.031	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0	0.000
C10	0.041	0.054	0.014	0.000	0.007	0.002	0.000	0.000	0.000	0

| **Importance Matrix**										
	C1	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**
C1	0	-0.119	-0.095	-0.118	-0.125	0.170	0.233	-0.353	-0.140	0.068
C2	-0.119	0	-0.114	0.089	0.150	-0.099	0.002	0.043	-0.136	-0.061
C3	-0.095	-0.114	0	0.223	0.227	-0.048	0.135	0.137	0.070	-0.104
C4	-0.118	0.089	0.223	0	-0.078	-0.179	-0.179	-0.274	-0.205	-0.241
C5	-0.125	0.150	0.227	-0.078	0	-0.216	0.030	-0.257	-0.317	-0.395
C6	0.170	-0.099	-0.048	-0.179	-0.216	0	0.031	0.224	0.155	0.161
C7	0.233	0.002	0.135	-0.179	0.030	0.031	0	0.066	0.011	-0.019
C8	-0.353	0.043	0.137	-0.274	-0.257	0.224	0.066	0	0.087	0.172
C9	-0.140	-0.136	0.070	-0.205	-0.317	0.155	0.011	0.087	0	0.042
C10	0.068	-0.061	-0.104	-0.241	-0.395	0.161	-0.019	0.172	0.042	0

Table C.40: Current BIP matrix and importance matrix for iteration 3 of the PCC method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.051	0.036	0.021	0.031	0.041
C2	0.000	0	0.000	0.000	0.058	0.000	0.000	0.000	0.000	0.054
C3	0.000	0.000	0	0.000	0.000	0.000	0.034	0.000	0.003	0.014
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.001	0.017	0.000
C5	0.000	0.058	0.000	0.000	0	0.000	0.000	0.000	0.000	0.007
C6	0.051	0.000	0.000	0.000	0.000	0	0.005	0.002	0.000	0.002
C7	0.036	0.000	0.034	0.000	0.000	0.005	0	0.001	0.000	0.000
C8	0.021	0.000	0.000	0.001	0.000	0.002	0.001	0	0.000	0.000
C9	0.031	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0	0.000
C10	0.041	0.054	0.014	0.000	0.007	0.002	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	-0.200	-0.054	-0.014	-0.026	0.202	0.070	0.067	0.085	0.066
C2	-0.200	0	0.099	0.093	0.145	-0.041	-0.091	-0.053	-0.157	0.070
C3	-0.054	0.099	0	0.060	0.052	0.106	0.281	0.130	0.112	-0.155
C4	-0.014	0.093	0.060	0	0.246	-0.148	0.102	0.118	0.184	0.249
C5	-0.026	0.145	0.052	0.246	0	-0.031	0.318	0.261	-0.022	0.056
C6	0.202	-0.041	0.106	-0.148	-0.031	0	0.070	0.138	0.231	0.433
C7	0.070	-0.091	0.281	0.102	0.318	0.070	0	0.158	-0.153	-0.241
C8	0.067	-0.053	0.130	0.118	0.261	0.138	0.158	0	-0.328	-0.339
C9	0.085	-0.157	0.112	0.184	-0.022	0.231	-0.153	-0.328	0	-0.388
C10	0.066	0.070	-0.155	0.249	0.056	0.433	-0.241	-0.339	-0.388	0

Table C.41: Current BIP matrix and importance matrix for iteration 4 of the PCC method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.051	0.036	0.021	0.031	0.041
C2	0.000	0	0.000	0.000	0.058	0.000	0.000	0.000	0.000	0.054
C3	0.000	0.000	0	0.000	0.000	0.000	0.033	0.000	0.003	0.014
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.001	0.017	0.000
C5	0.000	0.058	0.000	0.000	0	0.000	0.000	0.000	0.000	0.007
C6	0.051	0.000	0.000	0.000	0.000	0	0.005	0.002	0.000	0.002
C7	0.036	0.000	0.033	0.000	0.000	0.005	0	0.001	0.000	0.000
C8	0.021	0.000	0.000	0.001	0.000	0.002	0.001	0	0.000	0.000
C9	0.031	0.000	0.003	0.017	0.000	0.000	0.000	0.000	0	0.000
C10	0.041	0.054	0.014	0.000	0.007	0.002	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.064	-0.125	0.004	-0.055	0.385	0.001	-0.177	-0.231	-0.075
C2	0.064	0	-0.031	0.141	-0.078	-0.076	-0.232	0.032	-0.016	0.009
C3	-0.125	-0.031	0	0.102	0.056	-0.023	-0.166	-0.187	-0.205	-0.374
C4	0.004	0.141	0.102	0	0.203	-0.279	0.182	0.009	0.204	-0.013
C5	-0.055	-0.078	0.056	0.203	0	-0.251	0.434	-0.192	0.041	0.048
C6	0.385	-0.076	-0.023	-0.279	-0.251	0	0.252	0.130	0.089	0.190
C7	0.001	-0.232	-0.166	0.182	0.434	0.252	0	0.315	0.121	0.232
C8	-0.177	0.032	-0.187	0.009	-0.192	0.130	0.315	0	-0.069	-0.044
C9	-0.231	-0.016	-0.205	0.204	0.041	0.089	0.121	-0.069	0	-0.100
C10	-0.075	0.009	-0.374	-0.013	0.048	0.190	0.232	-0.044	-0.100	0

Table C.42: Current BIP matrix and importance matrix for iteration 5 of the PCC method with fluid system 2.

C.1.3   Random Forest Method

Case 1 Fluid Model
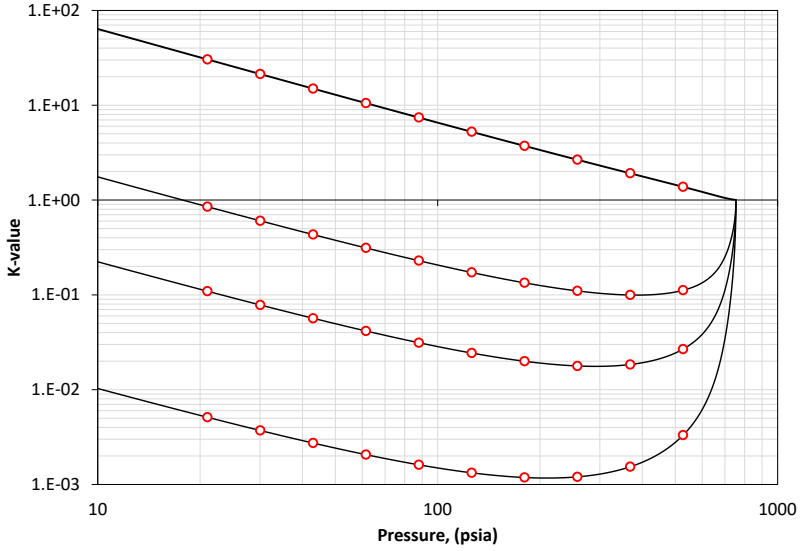
Figure C.5: K-value data after 14 iterations of the selective tuning approach with the RF based method. The resulting K-value estimates of the selective tuning method are shown as red circles while the detailed synthetic K-value results are given as a solid black line.

Current BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.000	0.000
C5	0.000	0	0.000	0.000
C7	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0

Importance Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.001	0.001	0.001
C5	0.001	0	0.001	0.195
C7	0.001	0.001	0	0.801
C10	0.001	0.195	0.801	0

Table C.43: Current BIP matrix and importance matrix for iteration 0 of the RF method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.000	0.000
C5	0.000	0	0.000	0.004
C7	0.000	0.000	0	0.001
C10	0.000	0.004	0.001	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.001	0.001	0.001
C5	0.001	0	0.001	0.212
C7	0.001	0.001	0	0.784
C10	0.001	0.212	0.784	0

Table C.44: Current BIP matrix and importance matrix for iteration 1 of the RF method with fluid system 1.

Current BIP Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.000	0.000
C5	0.000	0	0.000	0.004
C7	0.000	0.000	0	0.001
C10	0.000	0.004	0.001	0

Importance Matrix			
	C2	**C5**	**C7**	**C10**
C2	0	0.001	0.001	0.001
C5	0.001	0	0.001	0.256
C7	0.001	0.001	0	0.740
C10	0.001	0.256	0.740	0

Table C.45: Current BIP matrix and importance matrix for iteration 2 of the RF method with fluid system 1.

Current BIP Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.000	0.012	0.000
C5	0.000	0	0.000	0.004
C7	0.012	0.000	0	0.001
C10	0.000	0.004	0.001	0

Importance Matrix				
	C2	**C5**	**C7**	**C10**
C2	0	0.001	0.001	0.001
C5	0.001	0	0.001	0.218
C7	0.001	0.001	0	0.778
C10	0.001	0.218	0.778	0

Table C.46: Current BIP matrix and importance matrix for iteration 3 of the RF method with fluid system 1.

Case 2 Fluid Model
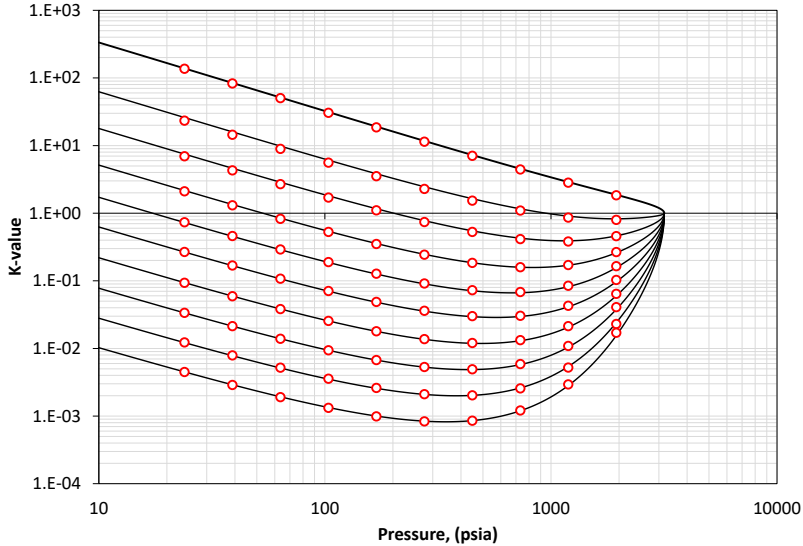
Figure C.6: K-value data after 14 iterations of the selective tuning approach with the RF based method. The resulting K-value estimates of the selective tuning method are shown as red circles while the detailed synthetic K-value results are given as a solid black line.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000
C7	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix

|	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.016	0.012	0.009	0.012	0.008	0.019	0.016	0.031	0.018
C2	0.016	0	0.010	0.007	0.010	0.008	0.009	0.012	0.010	0.008
C3	0.012	0.010	0	0.006	0.010	0.009	0.011	0.008	0.008	0.016
C4	0.009	0.007	0.006	0	0.010	0.009	0.011	0.017	0.013	0.014
C5	0.012	0.010	0.010	0.010	0	0.015	0.014	0.022	0.023	0.040
C6	0.008	0.008	0.009	0.009	0.015	0	0.014	0.043	0.056	0.070
C7	0.019	0.009	0.011	0.011	0.014	0.014	0	0.081	0.093	0.080
C8	0.016	0.012	0.008	0.017	0.022	0.043	0.081	0	0.037	0.045
C9	0.031	0.010	0.008	0.013	0.023	0.056	0.093	0.037	0	0.008
C10	0.018	0.008	0.016	0.014	0.040	0.070	0.080	0.045	0.008	0

Table C.47: Current BIP matrix and importance matrix for iteration 0 of the RF method with fluid system 2.

Current BIP Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.103	0.000
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000
C6	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.005	0.000
C7	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.103	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.011	0.007	0.008	0.009	0.008	0.010	0.010	0.009	0.012
C2	0.011	0	0.009	0.009	0.009	0.006	0.009	0.008	0.008	0.008
C3	0.007	0.009	0	0.010	0.011	0.008	0.009	0.012	0.007	0.006
C4	0.008	0.009	0.010	0	0.007	0.008	0.009	0.009	0.008	0.015
C5	0.009	0.009	0.011	0.007	0	0.009	0.018	0.030	0.046	0.038
C6	0.008	0.006	0.008	0.008	0.009	0	0.022	0.065	0.086	0.061
C7	0.010	0.009	0.009	0.009	0.018	0.022	0	0.120	0.100	0.062
C8	0.010	0.008	0.012	0.009	0.030	0.065	0.120	0	0.021	0.035
C9	0.009	0.008	0.007	0.008	0.046	0.086	0.100	0.021	0	0.018
C10	0.012	0.008	0.006	0.015	0.038	0.061	0.062	0.035	0.018	0

Table C.48: Current BIP matrix and importance matrix for iteration 1 of the RF method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.103	0.000
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.011	0.000
C6	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.005	0.000
C7	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.103	0.000	0.000	0.000	0.011	0.005	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.015	0.007	0.011	0.015	0.009	0.012	0.008	0.007	0.011
C2	0.015	0	0.007	0.008	0.008	0.008	0.008	0.009	0.007	0.010
C3	0.007	0.007	0	0.007	0.009	0.008	0.008	0.009	0.007	0.008
C4	0.011	0.008	0.007	0	0.008	0.008	0.009	0.008	0.012	0.025
C5	0.015	0.008	0.009	0.008	0	0.009	0.018	0.019	0.037	0.083
C6	0.009	0.008	0.008	0.008	0.009	0	0.039	0.071	0.030	0.099
C7	0.012	0.008	0.008	0.009	0.018	0.039	0	0.078	0.034	0.122
C8	0.008	0.009	0.009	0.008	0.019	0.071	0.078	0	0.025	0.017
C9	0.007	0.007	0.007	0.012	0.037	0.030	0.034	0.025	0	0.035
C10	0.011	0.010	0.008	0.025	0.083	0.099	0.122	0.017	0.035	0

Table C.49: Current BIP matrix and importance matrix for iteration 2 of the RF method with fluid system 2.

Current BIP Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.103	0.000
C2	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C4	0.000	0.000	0.000	0	0.000	0.000	0.000	0.000	0.000	0.000
C5	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000	0.011	0.000
C6	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.005	0.000
C7	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000	0.000
C8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0	0.000	0.000
C9	0.103	0.000	0.000	0.000	0.011	0.005	0.000	0.000	0	0.000
C10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0

Importance Matrix									
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	0.013	0.008	0.009	0.008	0.008	0.010	0.010	0.011	0.014
C2	0.013	0	0.008	0.008	0.011	0.007	0.008	0.007	0.006	0.007
C3	0.008	0.008	0	0.009	0.009	0.007	0.007	0.011	0.008	0.008
C4	0.009	0.008	0.009	0	0.012	0.008	0.010	0.010	0.011	0.045
C5	0.008	0.011	0.009	0.012	0	0.011	0.014	0.035	0.035	0.079
C6	0.008	0.007	0.007	0.008	0.011	0	0.016	0.034	0.044	0.065
C7	0.010	0.008	0.007	0.010	0.014	0.016	0	0.131	0.047	0.111
C8	0.010	0.007	0.011	0.010	0.035	0.034	0.131	0	0.022	0.029
C9	0.011	0.006	0.008	0.011	0.035	0.044	0.047	0.022	0	0.019
C10	0.014	0.007	0.008	0.045	0.079	0.065	0.111	0.029	0.019	0

**Table C.50:** Current BIP matrix and importance matrix for iteration 3 of the RF method with fluid system 2.

C.2 PCC Convergence Study Results

In the following figures, the results for 14 iteration of the selective tuning regression for a mixture of SCN C_1 to C_{10}. The initial composition used is uniform and the temperature is set to $100°$F with a synthetic model with Chueh-Prausnitz BIPs. The observed data used K-value data for each component at 10 pressure points which are logarithmically distributed. Each iteration (i.e. each figure) also displays the RMS for the gradient method (solid red line) as described in section 4.4.1 as well as an average RMS for for 20 cases with random choice of most important BIPs used in the regression. The number of important BIPs in all cases in this section was set to 10.
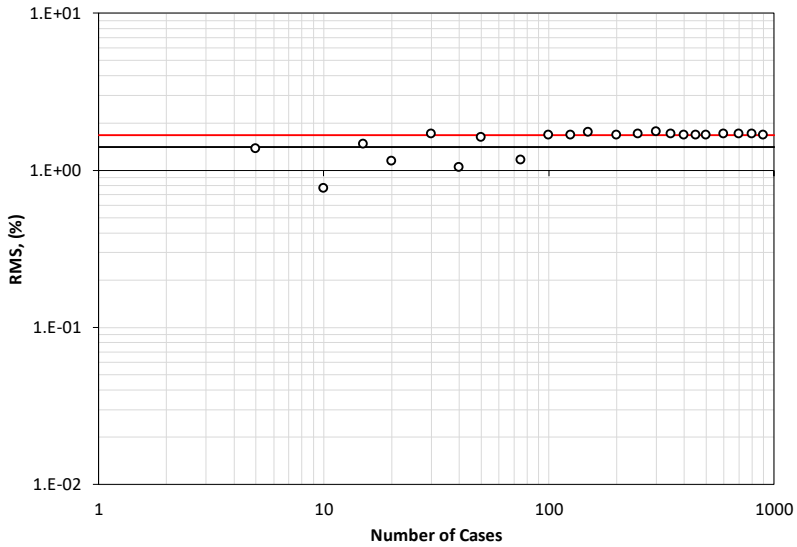
Figure C.7: The 1st iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).
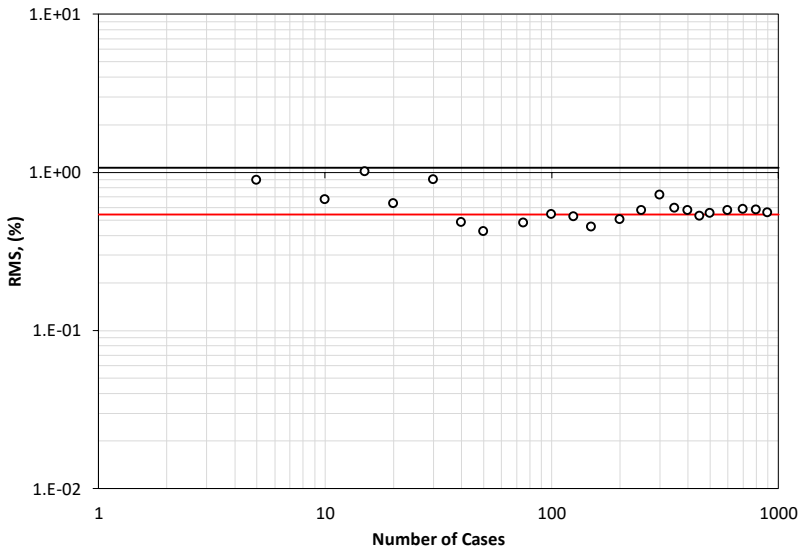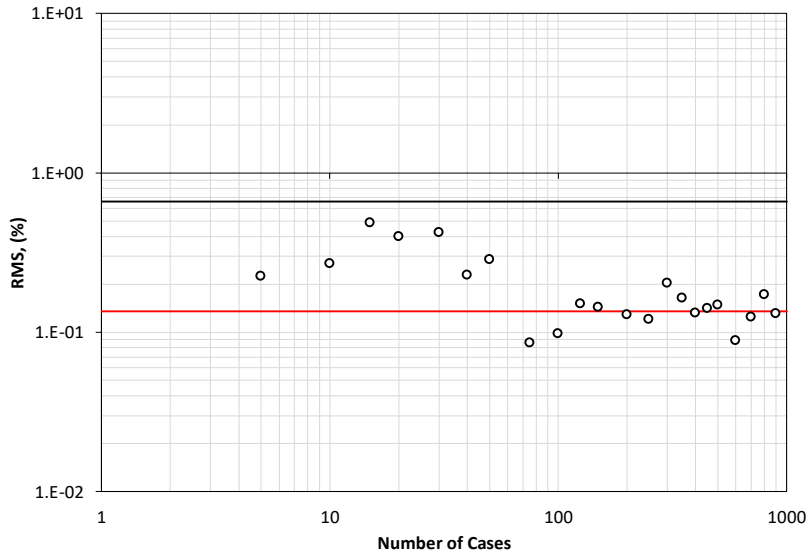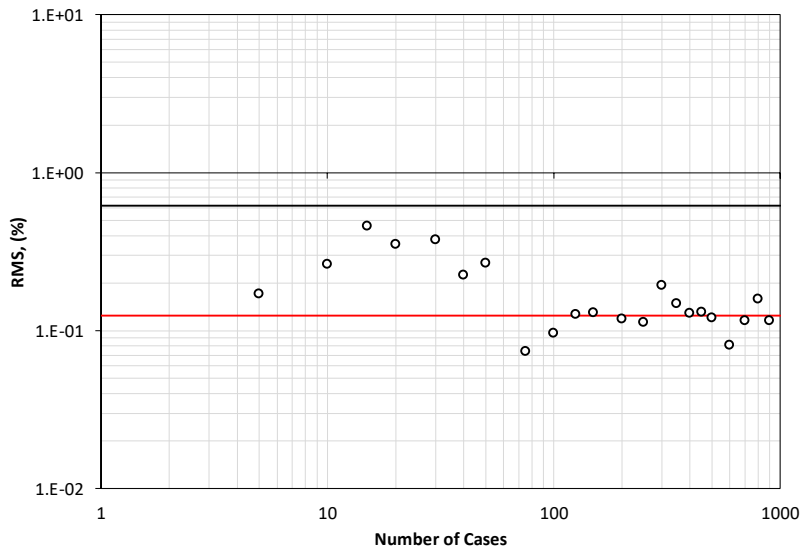
Figure C.8: The 2nd iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).
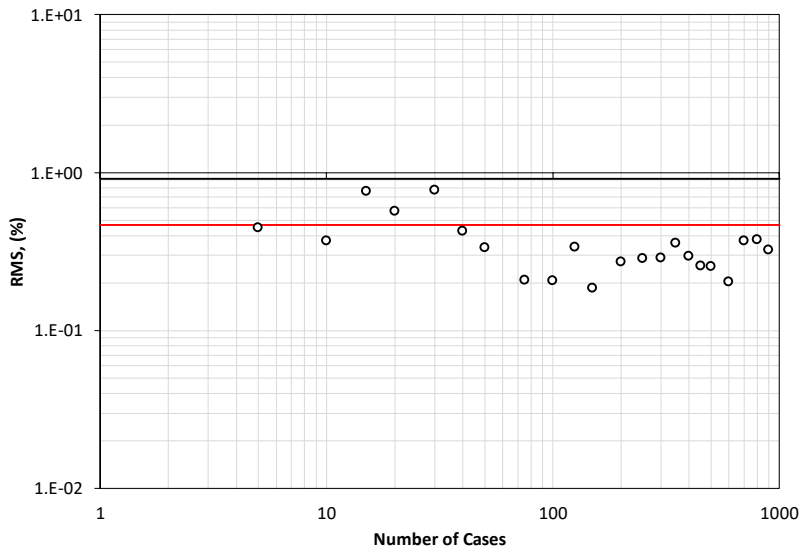
Figure C.9: The 3rd iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).
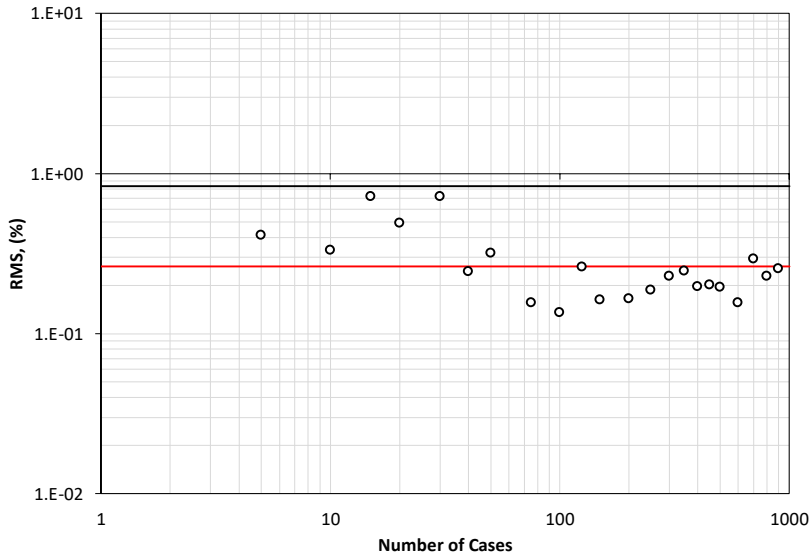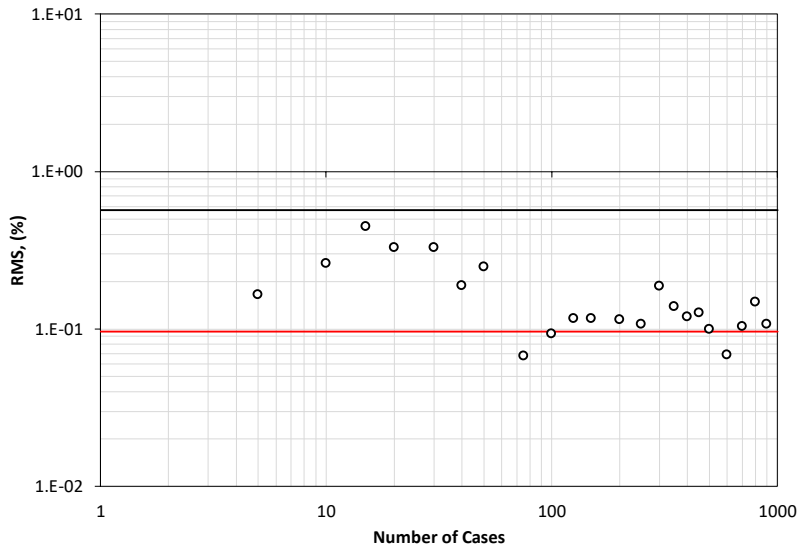
Figure C.10: The 4th iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).
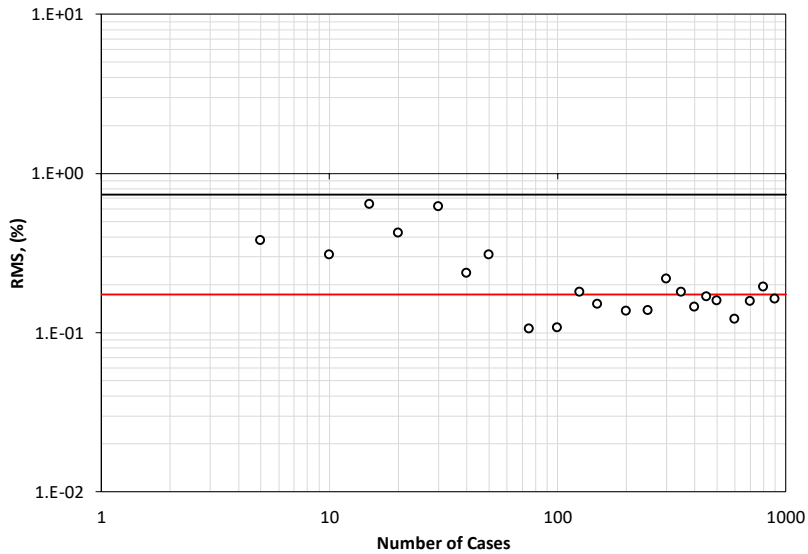
Figure C.11: The 5th iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).
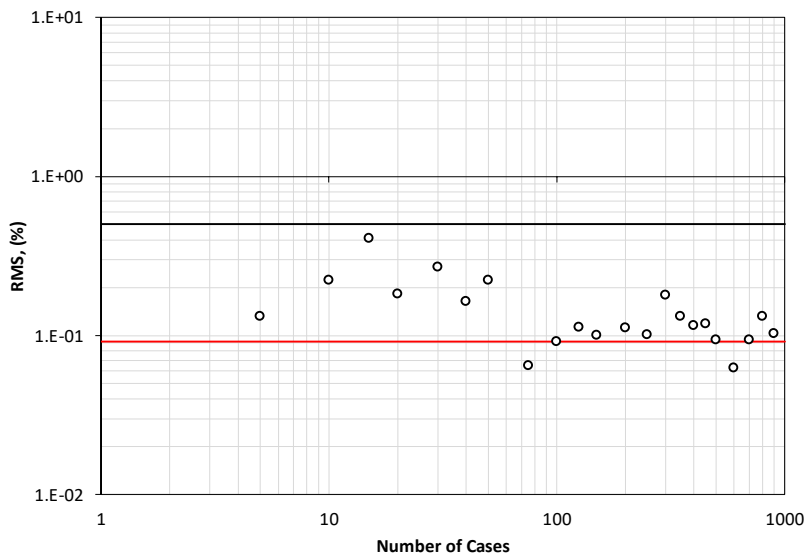
Figure C.12: The 6th iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).

Figure C.13: The 7th iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).

Figure C.14: The 8th iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).

Figure C.15: The 9th iteration of the selective tuning for the number of cases used to estimate the PCC ranging from 5 to 900. Also, the RMS value for the gradient method (red solid line) and an average random choice of important BIPs (solid black line).

Appendix D

Example Calculations

In the following sections examples of the following topics are given: RMS calculations for multiple experiments, building a decision tree, ternary convergence pressure composition, proof between number of BIPs and number of components and how to calculate 2-norm.

D.1   RMS Calculation

In this section, an example will be given for how to calculate the RMS for multiple experiments. The experiments that are used in this section are given in Table D.1. The main equation used to calculate the RMS is given in equation (4.4) in chapter 4 section 4.1.

	Experiment 1		Experiment 2	
Case	**Experimental**	**Calculated**	**Experimental**	**Calculated**
1	0.023	0.000	0.003	0.000
2	0.123	0.100	0.014	0.010
3	0.245	0.200	0.032	0.040
4	0.325	0.300	0.122	0.090
5	0.394	0.400	0.170	0.160
6	0.471	0.500	0.245	0.250
7	0.640	0.600		
8	0.748	0.700		
9	0.792	0.800		
10	0.892	0.900		

Experiment & Calculated Data

Table D.1: Example data for two sets of experiments with corresponding calculated data.

The first step of calculating the RMS of the multiple experiments is by calculating the RMS of the individual RMS values of the individual experiments.

Calculation for experiment 1:

$$d_{ref} = \max_i \{\omega_i\} = 0.892 \tag{D.1a}$$

$$\vec{r}_{exp.1} = [\frac{(0.023 - 0.0)}{d_{ref}}, \frac{(0.123 - 0.1)}{d_{ref}}, ..., \frac{(0.892 - 0.9)}{d_{ref}}] \tag{D.1b}$$

$$\langle \vec{r}_{exp.1}, \vec{r}_{exp.1} \rangle = (\frac{0.023 - 0.0}{d_{ref}})^2 + (\frac{0.123 - 0.1}{d_{ref}})^2 + ... + (\frac{0.892 - 0.9}{d_{ref}})^2 \tag{D.1c}$$

$$N = N_{exp.1} = 10 \tag{D.1d}$$

$$RMS_{exp.1} = \sqrt{\frac{1}{N} \langle \vec{r}_{exp.1}, \vec{r}_{exp.1} \rangle} \cdot 100 = 3.3\% \tag{D.1e}$$

Calculation for experiment 2:

$$d_{ref} = \max_i \{\omega_i\} = 0.245 \tag{D.2a}$$

$$\vec{r}_{exp.2} = [\frac{(0.003 - 0.0)}{d_{ref}}, \frac{(0.014 - 0.01)}{d_{ref}}, ..., \frac{(0.245 - 0.25)}{d_{ref}}] \tag{D.2b}$$

$$\langle \vec{r}_{exp.2}, \vec{r}_{exp.2} \rangle = (\frac{0.003 - 0.0}{d_{ref}})^2 + (\frac{0.014 - 0.01}{d_{ref}})^2 + ... + (\frac{0.245 - 0.25}{d_{ref}})^2 \tag{D.2c}$$

$$N = N_{exp.2} = 5 \tag{D.2d}$$

$$RMS_{exp.2} = \sqrt{\frac{1}{N} \langle \vec{r}_{exp.2}, \vec{r}_{exp.2} \rangle} \cdot 100 = 6.4\% \tag{D.2e}$$

The second step is to combine the RMS values of the individual experiments to yield the total RMS.

$$N = N_{exp.1} + N_{exp.2} = 10 + 5 \tag{D.3a}$$

$$RMS_{tot} = \sqrt{\frac{1}{N} (\langle \vec{r}_{exp.1}, \vec{r}_{exp.1} \rangle + \langle \vec{r}_{exp.2}, \vec{r}_{exp.2} \rangle)} \cdot 100 = 4.6\% \tag{D.3b}$$

Overview of RMS for Experiments		
Experiment 1	**Experiment 2**	**Experiment 1 & 2**
(%)	(%)	(%)
3.3	6.4	4.6

Table D.2: RMS values for experiment 1, 2 and the combination of experiments for the data given in Table D.1.

D.2 Decision Tree Example

This section will describe an example of how to build a decision tree based on an adaptation from the book by Mitchell [23] based on estimating whether or not to play tennis based on certain features. The features in this example is the outlook, temperature, humidity and the wind. The attributes associated with, for example, the outlook is sunny, overcast and rain. The dataset is given in Table D.3.

Table D.3: Dataset from Mitchell [23] for estimating if the conditions are acceptable to play tennis.

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

When generating the decision tree, the goal is to predict the outcome with as shallow a tree as possible to avoid over-fitting. This can be implemented by defining the Gain-function as in equation (2.39). The entropy can be calculated as shown in equation (2.40).

As an example, let S be the total system given in Table D.3 and the attribute be either yes

or no. Based on this knowledge we define S such that $S=[9+,5-]$ meaning that there are 9 results that yield yes and 5 values that yield no. The entropy of the system can then be calculated using equation (2.40) and yields:

$$Entropy(S) = -\frac{9}{14} \cdot \log_2(\frac{9}{14}) - \frac{5}{14} \cdot \log_2(\frac{5}{14}) = 0.940 \qquad \text{(D.4)}$$

Similarly the entropy for the attributes "Weak" and "Strong" can be calculated by defining the attribute system $S_{Weak}=[6+,2-]$ which means that for the subset with the attribute "weak" there are 6 results that yield yes and 2 that yield *no*. Similarly, the attribute system $S_{Strong}=[3+,3-]$ can be defined, which means that for the subset with the attribute "strong" there are 3 results that yield yes and 3 that yield *no*.

$$Entropy(S_{weak}) = -\frac{6}{8} \cdot \log_2(\frac{6}{8}) - \frac{2}{8} \cdot \log_2(\frac{2}{8}) = 0.811 \qquad \text{(D.5)}$$

$$Entropy(S_{strong}) = -\frac{3}{6} \cdot \log_2(\frac{3}{6}) - \frac{3}{6} \cdot \log_2(\frac{3}{6}) = 1 \qquad \text{(D.6)}$$

By combining the results in (D.4), (D.5) and (D.6) together with equation (2.39) the gain of the "wind" feature can be calculated to be:

$$Gain(S, wind) = 0.940 - \frac{8}{14} \cdot 0.811 - \frac{6}{14} \cdot 1 = 0.048 \qquad \text{(D.7)}$$

Similarly, it is possible to calculate the gain for all the features. The results for this are given by

Table D.4: Gain for all features from the example by Mitchell [23].

Feature	Gain
Outlook	0.246
Humidity	0.151
Wind	0.048
Tempearture	0.029

From the results given in Table D.4 it is possible to determine the feature that yields the most knowledge from the highest valued gain. The methodology detailed above is relevant for the first node in the decision tree, and the same procedure can be applied recursively for each layer in the tree with a subset of the data [23] as shown in Table D.7. One exception to this algorithm occurs when the results are all of one type (e.g. either all yes or no). This defines the terminal node of the decision tree.
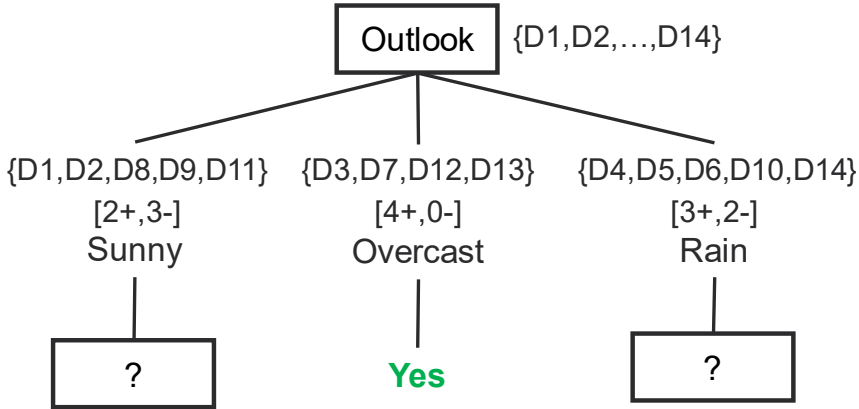
Figure D.1: Schematic of first node in decision tree based on the data from Mitchell [23] showing the sub-dividing of the original dataset.

D.3   Convergence Composition

The convergence pressure composition methodology is described in chapter 3 section 3.2.2, however, to make the approach more clear an example is given here for a ternary system with SCN components C_1, C_3 and C_6. The approach can be summarized by (1) starting with an initial composition (z_i^*) at a pressure (p^*) and temperature (T^*) then (2) the convergence pressure of the initial composition is calculated ($p_K(z_i^*, T^*)$) followed by (3) applying the negative flash to yield equilibrium composition for the vapor (y_{eqi}) and liquid (x_{eqi}) which are equal and finally (4) define the convergence pressure composition ($z_{Ki} = y_{eqi} = x_{eqi}$) which will yield a critical point at the convergence pressure of the initial composition ($p_c(z_{Ki}, T^*) = p_K(z_i^*, T^*)$).

This procedure was carried out for the ternary mixture containing SCN C_1, C_3 and C_6 and the compositional phase envelope for a temperature of 100°C at the convergence (1986.95 psia) and initial pressure (1000 psia) in Figure D.2. The critical tie-line extension at the critical point for the convergence pressure is also shown in the figure. This shows, graphically, the connection between initial composition and the convergence pressure composition as being connected via the critical tie-line extension.
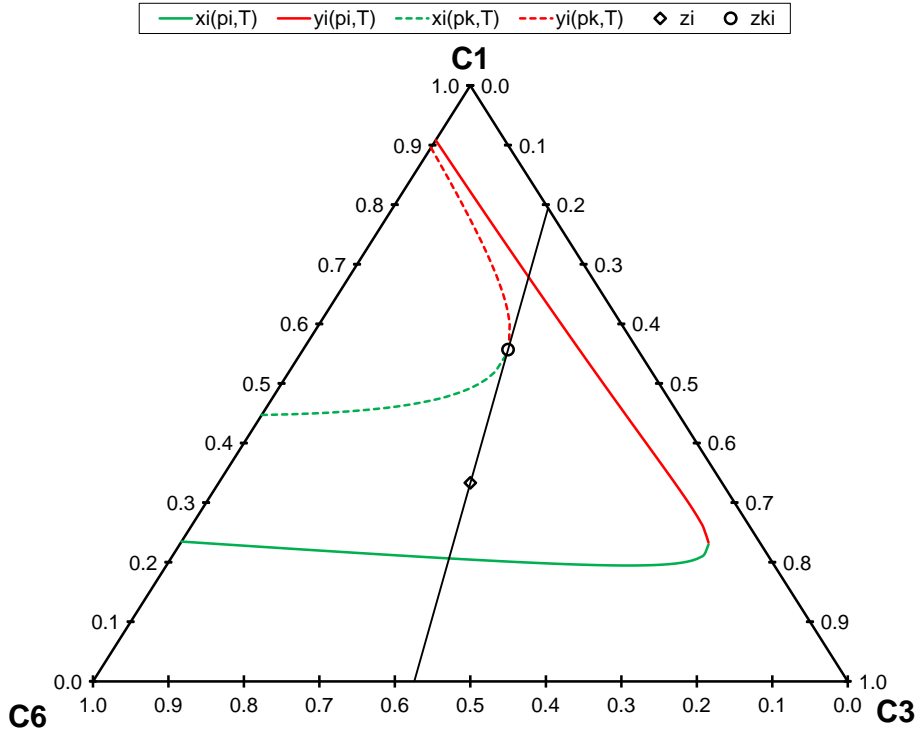
Figure D.2: Example of ternary compositional phase envelope for initial pressure (solid line) and convergence pressure (dashed line) as well as the inital composition (square symbol) and the convergence pressure composition (circle symbol).

D.4 Number of BIPs Derivation

A simple and visual proof of the equation for the number of BIPs with respect to the number of components is given below. First make a square with N_{comps} boxes and we start by filling all the boxes ($N_{filled} = N_{comps} \cdot N_{comps}$). Then, remove the diagonal, which contains as many boxes as the number of components ($N_{filled} = N_{comps} \cdot N_{comps} - N_{comps}$). Then remove the symmetric upper triangle by dividing by two ($N_{filled} = \frac{N_{comps} \cdot N_{comps} - N_{comps}}{2}$) and finally factoring out the common factor of N_{comps} resulting in

$$N_{filled} = \frac{N_{comps}(N_{comps} - 1)}{2} \tag{D.8}$$

From the figure below it is clear that the number of filled boxes is equal to the number of BIPs. This concludes the proof.

(a) Fill square.

(b) Remove diagonal.

(c) Remove upper triangle

(d) Factor the 5.

Figure D.3: Example of filling proof for $N_{comps} = 5$.

D.5 Two-Norm Example Calculation

Let $\vec{u} = [1, 4, 2, -2]$ be some arbitrary vector. A general description of how to calculate the 2-norm of any vector $\vec{v} = [v_1, v_2, ..., v_n]$ is given by

$$||\vec{v}||_2 = \sqrt{\sum_{i=1}^{n} v_i^2} \tag{D.9}$$

Applying equation (D.9) to the specific vector $\vec{u}$ yields

$$||\vec{u}||_2 = \sqrt{(1)^2 + (4)^2 + (2)^2 + (-2)^2} = \sqrt{25} = 5 \tag{D.10}$$

Appendix E

Code Reference

The code used to generate the data for chapters 3 and 4 can be found in a *github* repository at: `https://github.com/MarkusHays/Master-Thesis`

PhazeComp Solver K-value Regression Study

This chapter of the appendix is a collection of work done on the ability of PhazeComp to do regression of BIPs with synthetic K-value data. The range of results presented are meant to show when PhazeComp is able to accurately solve for the K-values or not and what information is needed (e.g. K-value data) to accurately predict the phase behavior

F.1 Gridding Structure

Withing this study three types of pressure point gridding was used to generate the K-value data. The three types of pressure point gridding are (a) linear gridding (equal spacing), (b) logarithmic gridding (equal spacing on a logarithmic axis) and (c) geometric gridding. The equations used to generate the pressure data are given for (a) to (c) are given equations (F.1) to (F.3).

$$x_k = x_0 + k\Delta x \tag{F.1a}$$

$$\Delta k = \frac{x_{N-1} - x_0}{N - 1} \tag{F.1b}$$

$$x_k = x_0 \cdot 10^{k\Delta x} \tag{F.2a}$$

$$\Delta k = \frac{\log(x_{N-1}/x_0)}{N - 1} \tag{F.2b}$$

$$x_k = x_0 + (x_{N-1} - x_0)(1 - r^k) \tag{F.3}$$

Below is a range of different cases where (i) the number of pressure points are varying, (b) the pressure gridding structure is different and (iii) for the case of geometric gridding the placement of the pressure points (r-value) are changed. For the figures where there are five components, the fluid system used is a C_1-C_8-C_{20}-C_{30}-C_{40} single carbon number mixture at 100°C.
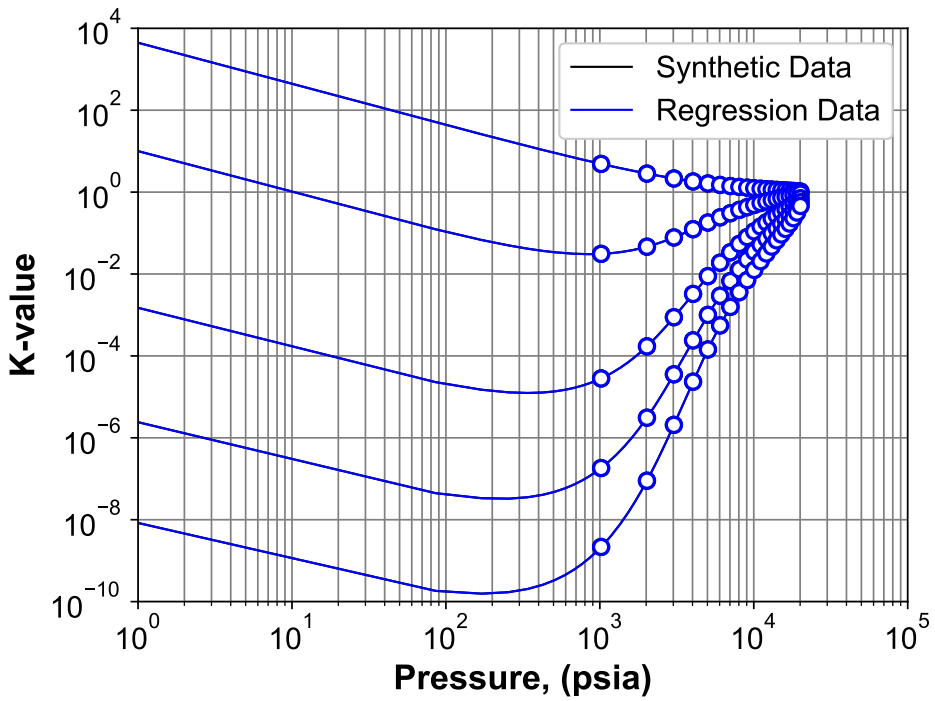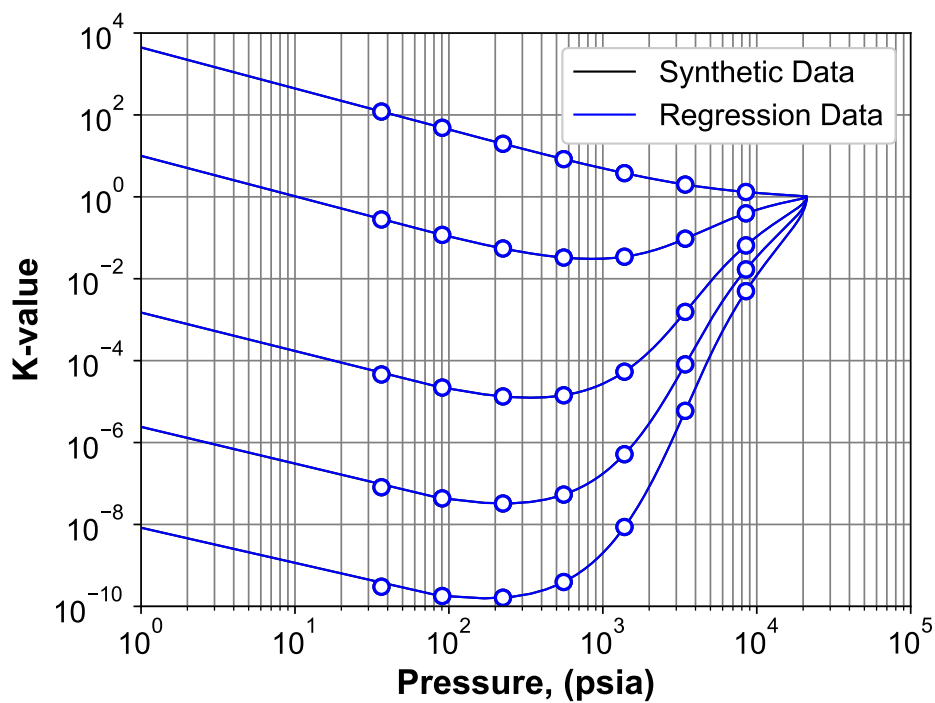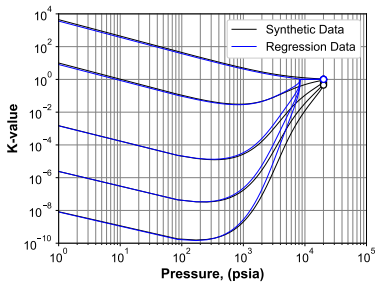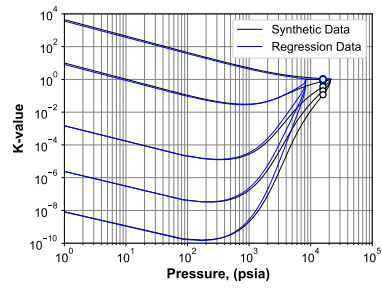

Figure F.1: Example of linear gridding.

Figure F.2: Example of logarithmic gridding.

F.2 Geometric Gridding Sensitivity

Below are three figures containing a sensitivity study of geometric gridding the r-value (i.e. the pressure range) and the number of pressure points.
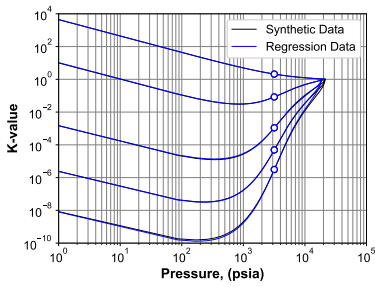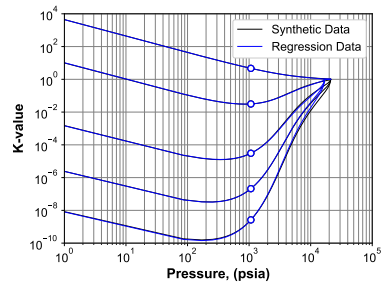
(a) r=0.05

(b) r=0.25
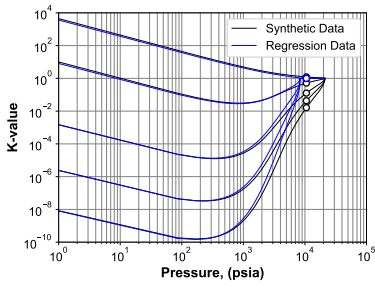
(c) r=0.50
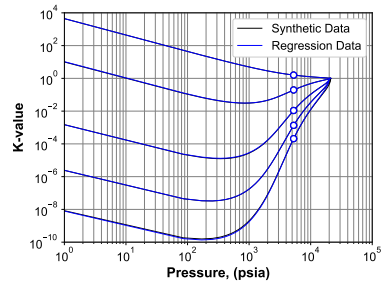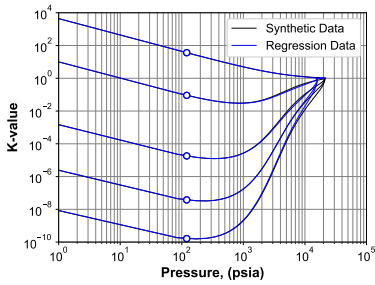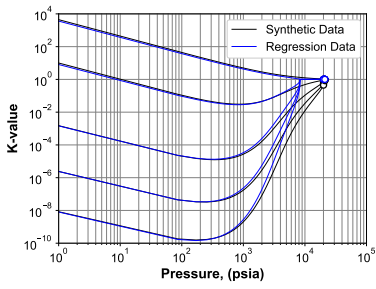
(d) r=0.75
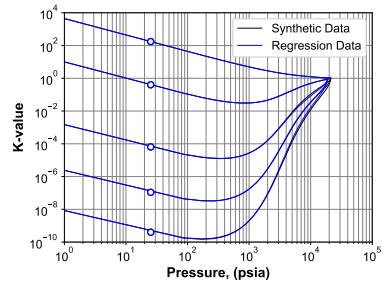
(e) r=0.85

(f) r=0.95

(g) r=0.995

(h) r=0.9995

Figure F.3: One pressure point geometric gridding with varying r-value.
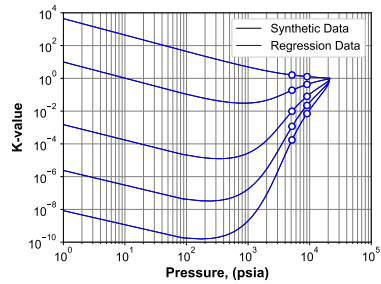
(a) r=0.05
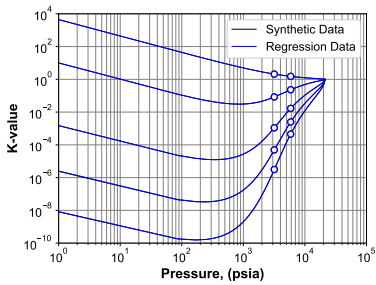
(b) r=0.25
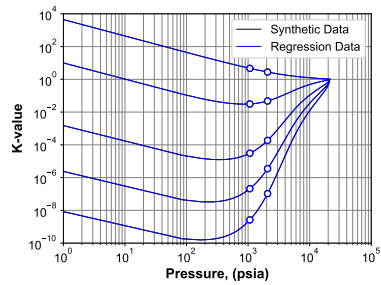
(c) r=0.50
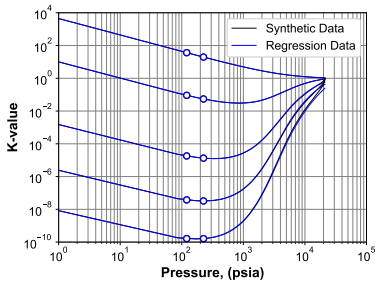
(d) r=0.75

(e) r=0.85

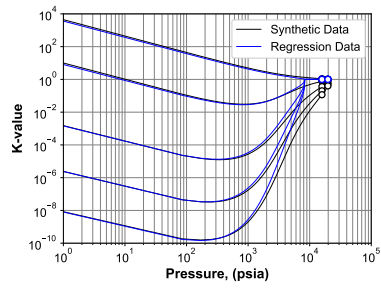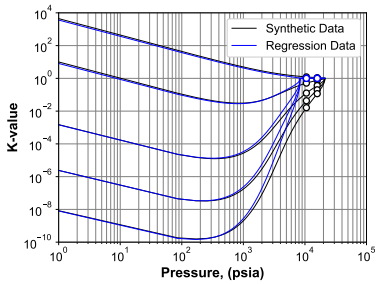**(f)** r=0.95

(g) r=0.995

(h) r=0.9995

Figure F.4: Two pressure points geometric gridding with varying r-value.
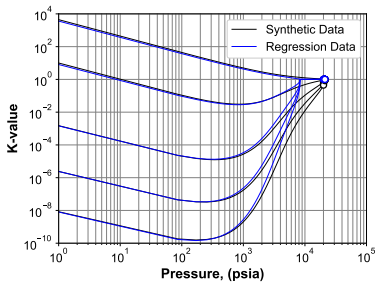
(a) r=0.05
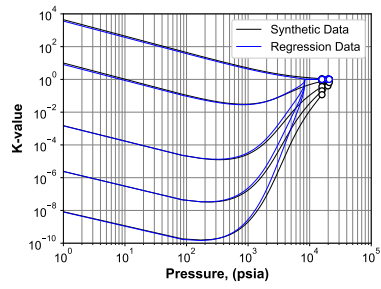
(b) r=0.25
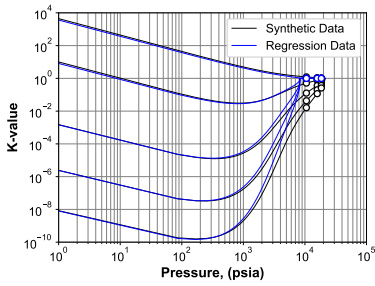
(c) r=0.50
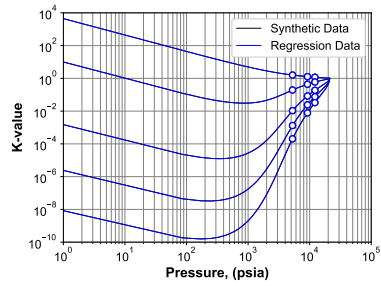
(d) r=0.75
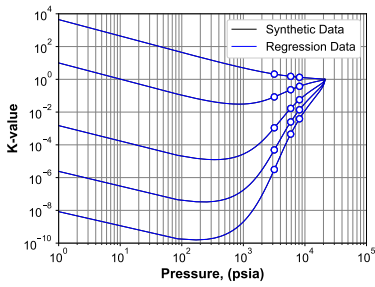
(e) r=0.85
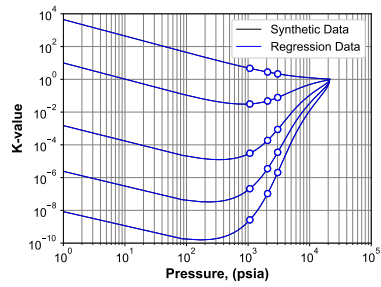
(f) r=0.95
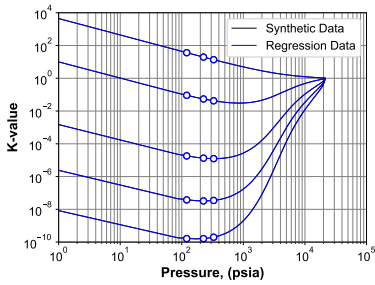
(g) r=0.995
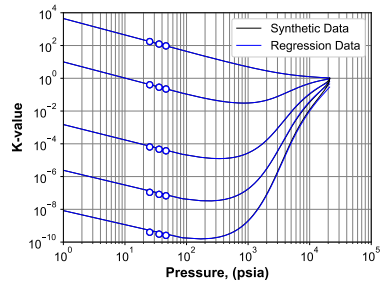
(h) r=0.9995

Figure F.5: Three pressure points geometric gridding with varying r-value.

F.3 Strange Case of Low-Pressure Crossing
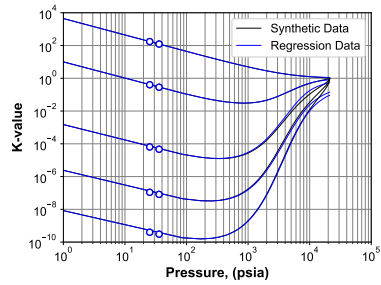
A unforeseen result of the flexibility of the effect of BIPs on the phase behavior is given below. The following two figures show a regression of 20 pressure points for a C_1-C_{15} SCN mixture at 100 °C K-value data. The interesting effect of the two regression solutions is shown in the low pressure region as both regression solutions fit a wide range of pressure data. However, as the pressure tends to 1 psia the K-value data of several components (specifically C_5 and C_6) starts deviating from the linear trend. The only difference in the regression is the range of the search, where in Figure F.6 the range is strictly positive and in Figure F.7 the lower range is set to -0.1.
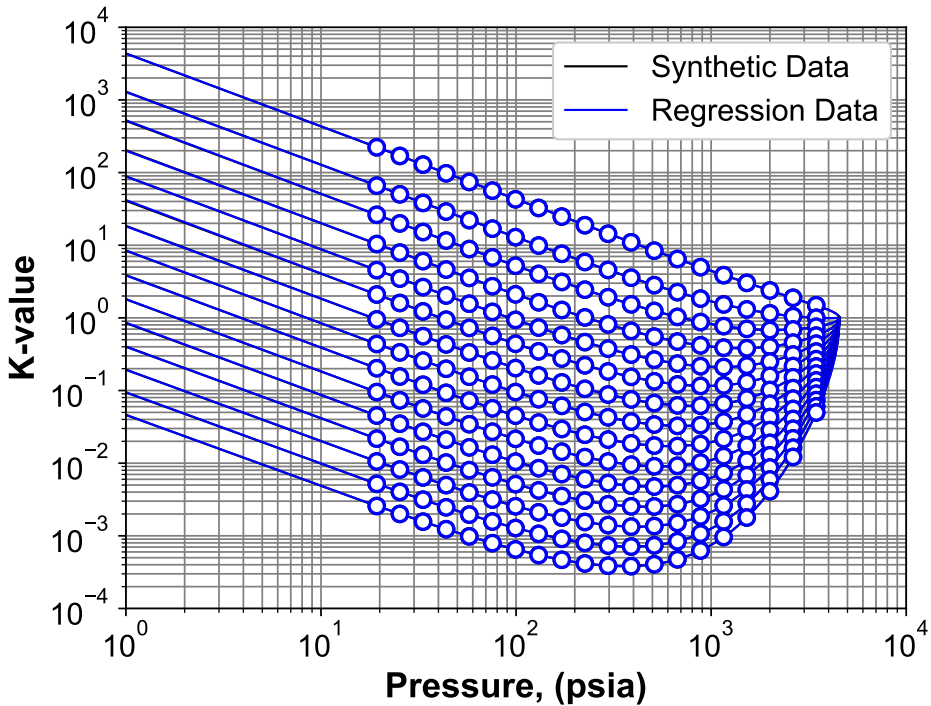
Figure F.6: Final results of BIP regression with 20 pressure points for a C_1 to C_{15} SCN mixture at 100 °C with a BIP ranging from 0 to 0.5.
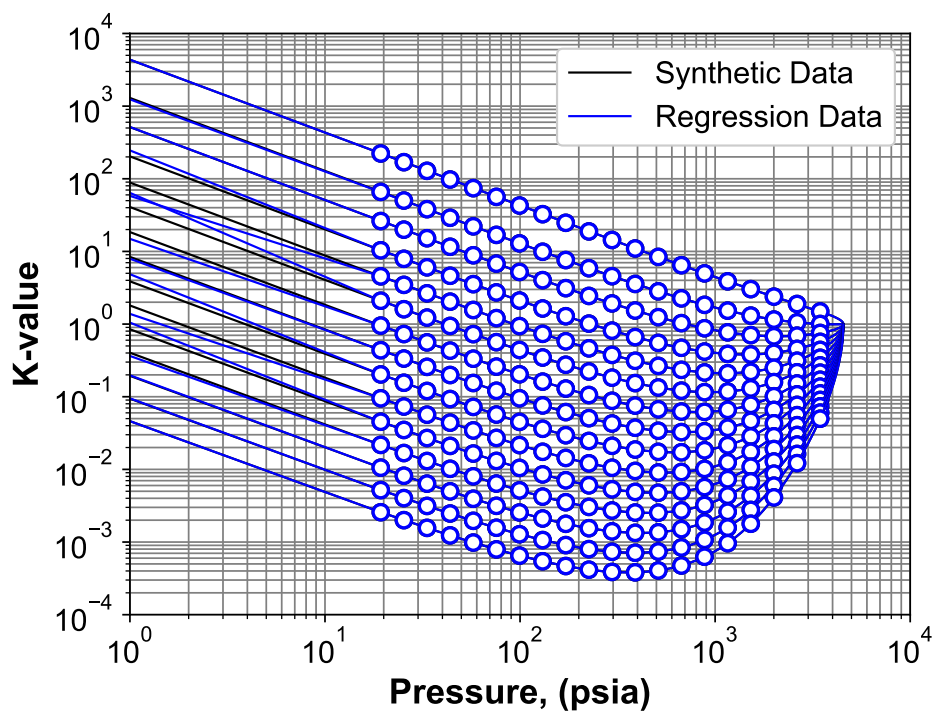
Figure F.7: Final results of BIP regression with 20 pressure points for a C_1 to C_{15} SCN mixture at 100 °C with a BIP ranging from -0.1 to 0.5.

The resulting equation of state models from the regression in Figure F.6 and F.7 are given in Table F.3 and F.4. For the associated PhazeComp file and the respective output file, the reader is able to retrieve the data at the GitHub website linked to in the previous chapter.

Table F.1: Component properties for all components used in this section (i.e. SCN C_1 to C_{15}).

Component Name	MW	T_c (°C)	p_c (psia)	ω	V_c
C1	16.04	-82.59	667.0	0.011	0.099
C2	30.07	32.17	706.6	0.099	0.146
C3	44.10	96.68	616.1	0.152	0.200
C4	58.12	151.6	544.9	0.203	0.258
C5	70.91	197.0	503.2	0.231	0.306
C6	82.42	240.2	490.0	0.240	0.341
C7	95.63	276.1	455.3	0.274	0.387
C8	109.24	307.3	420.4	0.312	0.435
C9	122.22	334.7	388.1	0.352	0.485
C10	135.14	359.4	359.9	0.392	0.537
C11	147.98	381.7	335.3	0.432	0.588
C12	160.73	402.1	313.7	0.471	0.640
C13	173.38	420.8	294.7	0.506	0.691
C14	185.92	438.1	277.9	0.543	0.742
C15	198.36	454.1	263.0	0.579	0.792

Table F.2: Synthetic BIPs (Chueh-Prausnitz) for all components used in this section (i.e. SCN C_1 to C_{15}).

Binary Interaction Parameters

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
C1	0	0.002	0.007	0.013	0.018	0.021	0.025	0.030	0.034	0.039	0.043	0.047	0.050	0.054	0.057
C2	0.002	0	0.001	0.005	0.008	0.010	0.013	0.016	0.020	0.023	0.027	0.030	0.033	0.036	0.039
C3	0.007	0.001	0	0.001	0.002	0.004	0.006	0.008	0.011	0.013	0.016	0.018	0.021	0.023	0.026
C4	0.013	0.005	0.001	0	0.000	0.001	0.002	0.004	0.006	0.007	0.009	0.011	0.013	0.015	0.017
C5	0.018	0.008	0.002	0.000	0	0.000	0.001	0.002	0.003	0.004	0.006	0.008	0.009	0.011	0.012
C6	0.021	0.010	0.004	0.001	0.000	0	0.000	0.001	0.002	0.003	0.004	0.005	0.007	0.008	0.010
C7	0.025	0.013	0.006	0.002	0.001	0.000	0	0.000	0.001	0.001	0.002	0.004	0.005	0.006	0.007
C8	0.030	0.016	0.008	0.004	0.002	0.001	0.000	0	0.000	0.001	0.001	0.002	0.003	0.004	0.005
C9	0.034	0.020	0.011	0.006	0.003	0.002	0.001	0.000	0	0.000	0.001	0.001	0.002	0.003	0.003
C10	0.039	0.023	0.013	0.007	0.004	0.003	0.001	0.001	0.000	0	0.000	0.000	0.001	0.001	0.002
C11	0.043	0.027	0.016	0.009	0.006	0.004	0.002	0.001	0.001	0.000	0	0.000	0.000	0.001	0.001
C12	0.047	0.030	0.018	0.011	0.008	0.005	0.004	0.002	0.001	0.000	0.000	0	0.000	0.000	0.001
C13	0.050	0.033	0.021	0.013	0.009	0.007	0.005	0.003	0.002	0.001	0.000	0.000	0	0.000	0.000
C14	0.054	0.036	0.023	0.015	0.011	0.008	0.006	0.004	0.003	0.001	0.001	0.000	0.000	0	0.000
C15	0.057	0.039	0.026	0.017	0.012	0.010	0.007	0.005	0.003	0.002	0.001	0.001	0.000	0.000	0

Table F.3: Final results of the regression BIPs with a range on the BIPs between 0 and 0.5 for all components used in this section (i.e. SCN C_1 to C_{15}).

Binary Interaction Parameters															
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
C1	0	0.002	0.006	0.013	0.018	0.020	0.026	0.029	0.035	0.039	0.042	0.047	0.052	0.053	0.058
C2	0.002	0	0.014	0.000	0.000	0.022	0.000	0.031	0.010	0.018	0.044	0.029	0.000	0.066	0.032
C3	0.006	0.014	0	0.000	0.000	0.004	0.012	0.000	0.014	0.021	0.000	0.023	0.041	0.000	0.033
C4	0.013	0.000	0.000	0	0.003	0.000	0.002	0.003	0.009	0.003	0.012	0.008	0.018	0.018	0.013
C5	0.018	0.000	0.000	0.003	0	0.000	0.001	0.004	0.000	0.004	0.011	0.007	0.000	0.019	0.011
C6	0.020	0.022	0.004	0.000	0.000	0	0.000	0.000	0.002	0.006	0.000	0.007	0.011	0.000	0.014
C7	0.026	0.000	0.012	0.002	0.001	0.000	0	0.000	0.001	0.000	0.004	0.004	0.000	0.012	0.005
C8	0.029	0.031	0.000	0.003	0.004	0.000	0.000	0	0.000	0.001	0.000	0.001	0.007	0.000	0.006
C9	0.035	0.010	0.014	0.009	0.000	0.002	0.001	0.000	0	0.000	0.000	0.002	0.001	0.004	0.002
C10	0.039	0.018	0.021	0.003	0.004	0.006	0.000	0.001	0.000	0	0.001	0.000	0.000	0.002	0.002
C11	0.042	0.044	0.000	0.012	0.011	0.000	0.004	0.000	0.000	0.001	0	0.000	0.000	0.000	0.002
C12	0.047	0.029	0.023	0.008	0.007	0.007	0.004	0.001	0.002	0.000	0.000	0	0.001	0.000	0.000
C13	0.052	0.000	0.041	0.018	0.000	0.011	0.000	0.007	0.001	0.000	0.000	0.001	0	0.000	0.000
C14	0.053	0.066	0.000	0.018	0.019	0.000	0.012	0.000	0.004	0.002	0.000	0.000	0.000	0	0.000
C15	0.058	0.032	0.033	0.013	0.011	0.014	0.005	0.006	0.002	0.002	0.002	0.000	0.000	0.000	0

Table F.4: Final results of the regression BIPs with a range on the BIPs between -0.1 and 0.5 for all components used in this section (i.e. SCN C_1 to C_{15}).

Binary Interaction Parameters															
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
C1	0	0.002	0.005	0.014	0.015	0.026	0.020	0.034	0.031	0.043	0.037	0.051	0.050	0.052	0.059
C2	0.002	0	0.026	-0.012	0.058	-0.092	0.121	-0.067	0.090	-0.065	0.138	-0.058	0.043	0.073	0.020
C3	0.005	0.026	0	0.000	-0.053	0.110	-0.100	0.089	-0.049	0.072	-0.051	0.080	-0.011	0.030	0.026
C4	0.014	-0.012	0.000	0	0.019	-0.033	0.040	-0.027	0.018	0.023	-0.034	0.046	0.047	-0.070	0.062
C5	0.015	0.058	-0.053	0.019	0	0.000	0.002	-0.006	0.028	-0.058	0.115	-0.100	0.001	0.142	-0.068
C6	0.026	-0.092	0.110	-0.033	0.000	0	-0.003	0.015	-0.029	0.064	-0.096	0.113	-0.007	-0.100	0.084
C7	0.020	0.121	-0.100	0.040	0.002	-0.003	0	-0.005	0.016	-0.031	0.057	-0.058	0.023	0.047	-0.024
C8	0.034	-0.067	0.089	-0.027	-0.006	0.015	-0.005	0	-0.003	0.011	-0.018	0.025	-0.019	0.023	-0.004
C9	0.031	0.090	-0.049	0.018	0.028	-0.029	0.016	-0.003	0	-0.002	0.002	-0.004	0.030	-0.049	0.031
C10	0.043	-0.065	0.072	0.023	-0.058	0.064	-0.031	0.011	-0.002	0	0.002	0.002	-0.026	0.055	-0.027
C11	0.037	0.138	-0.051	-0.034	0.115	-0.096	0.057	-0.018	0.002	0.002	0	-0.002	0.016	-0.034	0.022
C12	0.051	-0.058	0.080	0.046	-0.100	0.113	-0.058	0.025	-0.004	0.002	-0.002	0	-0.003	0.013	-0.008
C13	0.050	0.043	-0.011	0.047	0.001	-0.007	0.023	-0.019	0.030	-0.026	0.016	-0.003	0	-0.002	0.002
C14	0.052	0.073	0.030	-0.070	0.142	-0.100	0.047	0.023	-0.049	0.055	-0.034	0.013	-0.002	0	0.000
C15	0.059	0.020	0.026	0.062	-0.068	0.084	-0.024	-0.004	0.031	-0.027	0.022	-0.008	0.002	0.000	0

Markus Hays Nielsen

The Effect of Binary Interaction Parameters on Phase Behavior

NTNU
Norwegian University of
Science and Technology