

# Modelagem de equilíbrio multifásico e Reconstrução Molecular aplicados ao processo de desasfaltação

Trabalho da disciplinas COQ712 - Termodinâmica de Soluções  
PEQ/COPPE/UFRJ

Aluno: Diego Telles Fernandes

Professores: Frederico W. Tavares & Iuri S. V. Segtovich

10 de Novembro de 2021

## Conteúdo

<b>1. Introdução</b>	<b>2</b>
<b>2. Desasfaltação</b>	<b>3</b>
<b>3. Reconstrução molecular</b>	<b>4</b>
3.1. Reconstrução Estocástica . . . . .	4
3.2. Partição Orientada a Medoides e Maximização entrópica . . . . .	6
<b>4. Representação molecular</b>	<b>8</b>
<b>5. Predição de propriedades</b>	<b>9</b>
<b>6. Equilíbrio Multifásico</b>	<b>10</b>
6.1. Distância ao Plano Tangente . . . . .	11
6.2. Análise de estabilidade . . . . .	14
6.3. Modelo ELLV . . . . .	17
<b>7. Resultados</b>	<b>22</b>
7.1. Predição de propriedades . . . . .	22
7.2. Equilíbrio de fases . . . . .	25
7.3. Reconstrução Molecular . . . . .	26
7.4. Desasfaltação . . . . .	26
<b>8. Conclusão</b>	<b>32</b>
<b>A. Estruturas do método de contribuição de grupos de Marrero e Gani</b>	<b>34</b>

## 1. Introdução

Este trabalho tem por objetivo modelar o equilíbrio de fases Líquido-Líquido-Vapor (ELLV) presente no processo de desasfaltação utilizando correntes de Resíduo de Vácuo (RV) que foram caracterizadas por Reconstrução Molecular.

O processo de desasfaltação pertence ao parque de refino do Brasil e tem por objetivo agregar valor ao RV separando-o em duas correntes através de uma extração líquido-líquido utilizando propano, ou GLP, como solvente. Os aspectos mais detalhados e a descrição do processo será apresentado na Seção 2.

A abordagem mais popular para caracterização de petróleos, e amplamente utilizada nos simuladores comerciais, é por pseudo-componente. Nesta abordagem, todas as propriedades de todas os pseudos são estimadas a partir de correlações que utilizam apenas duas informações da corrente, curva de destilação e densidade. Estas correlações foram elaboradas, em sua maioria, para petróleos e podem apresentar desvios grandes quando aplicadas a este corte mais pesado, o RV. Apesar de muito útil para o equilíbrio líquido-vapor, a caracterização por pseudo-componente apresenta algumas limitações para modelagem do equilíbrio líquido-líquido.

É um grande desafio modelar interações moleculares mais complexas, que dão origem a formação de duas fases líquidas, sem conhecer ao certo as moléculas presentes nas correntes. Para vencer-lo, este trabalho faz o uso da Reconstrução Molecular como método de caracterização para o RV, onde será proposta uma mistura de moléculas para representar esta corrente. Esta metodologia será melhor apresentada na Seção 3.

A Reconstrução Molecular está alicerçada numa representação das moléculas de forma objetiva e de fácil trato computacional, e também numa forma eficiente de estimação das propriedades das moléculas puras e de misturas, que neste caso se dá por contribuição de grupos. Estes aspectos serão discutidos nas Seções 4 e 5.

Uma vez definida as moléculas presentes no sistema, e o método de cálculo das propriedades, a próxima etapa para modelagem do processo de desasfaltação é a modelagem do equilíbrio multifásico. Esta etapa envolveu aplicação direta dos conhecimentos abordados na disciplina, e neste trabalho foi implementado o algoritmo proposto por Michelsen and Mollerup [2007] com algumas modificações que foram apresentadas na Seção 6.

Todas estas etapas envolvidas na construção deste modelos foram comparadas a dados experimentais, ou simuladores comerciais, afim de validar sua implementação. Após esta validação, o modelo foi utilizado para simulação da unidade de desasfaltação e seu desempenho foi comparado a dados experimentais levantados em células PVT. Esta discussão está presente na Seção 7. O modelo testado não apresentou desempenho satisfatório, a priori, pois não indicou a formação de duas fases líquidas. No entanto, a estratégia implementada parece promissora se uma forma melhor para estimar o parâmetro de interação binária da mistura for desenvolvida.

## 2. Desasfaltação

A desasfaltação é um processo de extração líquido-líquido que utiliza Propano, ou GLP, como solvente e tem por objetivo separar o Resíduo de Vácuo (RV) em duas correntes. Uma delas, que é a corrente insolúvel no solvente, é rica em moléculas asfálticas e é denominada como Resíduo Asfáltico (RASf) cuja principal utilidade é compor o Cimento Asfáltico de Petróleo (CAP). A outra corrente, a solúvel no solvente, é chamada de Óleo Desasfaltado (ODES) cujo destino é a fabricação de Óleos Lubrificantes ou carga de FCC.

Esta extração acontece numa coluna de pratos onde o líquido menos denso (ODES) sai pelo topo e o mais denso (RASf) pelo fundo. Em seguida, ambas as correntes são enviadas para suas seções de recuperação de solvente. Um diagrama esquemático de uma Unidade de Desasfaltação (UDASF) é apresentado na Figura 1.

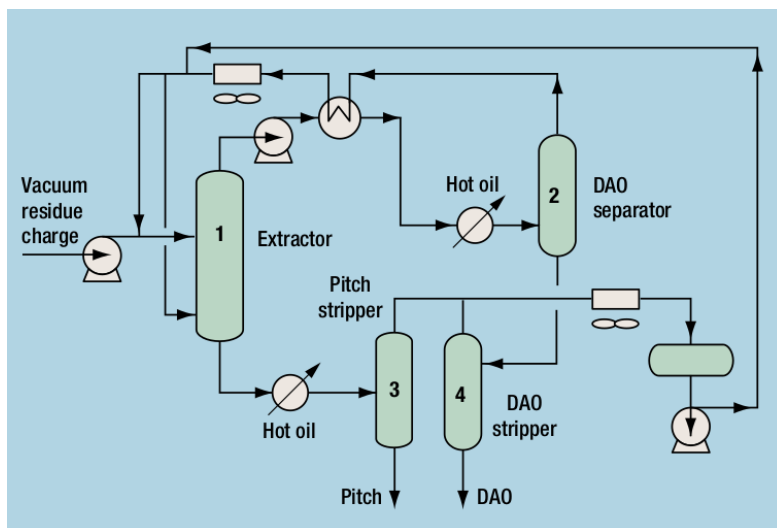


Figura 1: Fluxograma esquemático de uma UDASF (DAO = ODES e Pitch = RASf).  
Fonte: *Hydrocarbon Processing's Refining Processes 2006 Handbook*.

A torre extratora da unidade opera em pressão suficiente para garantir o solvente em fase líquida, de 30 a 45 kgf/cm<sup>2</sup> man, a uma temperatura que varia de 50 a 80°C. Apesar da torre conter vários pratos, a comparação do rendimento obtido em unidades industriais com o rendimento obtido em células PVT indicam que as unidades são bem representadas por um único estágio de equilíbrio.

A temperatura de extração influencia significativamente no rendimento de ODES e RASf da unidade, e a variação do rendimento com esta variável possui um comportamento particular neste processo. O rendimento de ODES decresce com o aumento da temperatura o que geralmente não é observado neste tipo de extração. Inclusive, neste quesito fica bastante claro a limitação da representação da corrente de Resíduo de Vácuo por pseudo-componentes. Quando um RV é representado desta forma, é difícil simular a formação de duas fases com rendimentos aceitáveis, e mesmo que isto aconteça, é bastante difícil

observar o mesmo comportamento dos rendimentos em função de temperatura e pressão que se observa nas unidades. Por isto, a Reconstrução Molecular do RV é recomendado para representação deste processo e esta metodologia será descrita na Seção 3 a seguir.

### 3. Reconstrução molecular

A reconstrução molecular é uma metodologia que propões uma mistura hipotética de moléculas que, segundo seus métodos de predição de propriedades, representa uma mistura real de hidrocarbonetos. As moléculas propostas certamente não estarão presentes na mistura, mas elas foram geradas contendo as principais estruturas de moléculas comumente presente neste corte (RV), e a mistura delas possui as mesmas propriedades da corrente experimental. Ou seja, a metodologia não pretende identificar as moléculas presentes numa determinada mistura, mas sim propor uma mistura hipotética que se comporta da mesma maneira. Para isto, esta metodologia é dividida em três etapas que serão apresentadas a seguir.

#### 3.1. Reconstrução Estocástica

Esta etapa foi inspirada no trabalho de Oliveira et al. [2013] e parte do princípio que alguns atributos estruturais presentes nas moléculas de uma determinada fração de petróleo seguem uma função de densidade de probabilidades (FDP) conhecida. Algumas modificações nas FDPs e seus parâmetros foi proposta por Maciel [2019] em sua dissertação de mestrado aqui na UFRJ, e esta contribuição está descrita nos Itens 4.3 e 5.5 do seu trabalho. Foi utilizada a proposta do Maciel [2019] e os atributos finais, as FDPs e seus parâmetros estão apresentadas de forma resumida na Tabela 1.

Tabela 1: Parâmetros e tipos das funções de densidade de probabilidade (FDP) associadas a cada atributo estrutural das moléculas de RV.

	Atributo estrutural	FDP	Parâmetros
1	Tipo de Molécula	Histograma	$\theta_1, \theta_2, \theta_3$
2	Comprimento de cadeias parafínicas	Chi-Quadrado	$\theta_4$
3	Número total de anéis	Chi-Quadrado	$\theta_5$
2	Comprimento de ramificações	Exponencial	$\theta_4$
5	Número de benzenos	Exponencial	$\theta_6$
6	Tipo de heterociclos	Histograma	$\theta_7, 1 - \theta_6, \theta_8$
7	Enxofre alifático	Histograma	$\theta_7$
8	Outro heteroátomo alifático	Histograma	$1 - ((1 - \theta_7)^2 + \theta_7)$
9	Heteroátomo alifático N ou O	Histograma	$\theta_8$
10	Número de núcleos	Exponencial	$\theta_9$

Importante observar que para representar o RV temos 10 FDPs e 9 parâmetros e que algumas FDPs compartilham os mesmos parâmetros. Realizar a Reconstrução Estocástica

é estimar estes 9 parâmetros tendo como função objetivo a comparação entre as propriedades calculadas pelo modelo e as experimentais fornecidas pelo usuário para uma corrente de RV desejada.

A metodologia consiste em gerar moléculas a partir de amostragens das FDPs seguindo um diagrama de construção. Cada etapa deste diagrama envolve a amostragem de uma das FDPs, e após o término de todas as etapas, uma molécula é montada. Amostragem de uma FDP implica em gerar um número aleatório a partir da forma funcional proposta. Ou seja, este é um procedimento que guarda uma certa “randomicidade”, e para que uma mistura de moléculas geradas a partir destas amostragens tenham de fato sua forma funcional proposta, uma grande quantidade de moléculas deve ser gerada. Assim, nesta etapa 5000 moléculas são geradas e misturadas com composição uniforme para compor uma mistura. As propriedades desta mistura são calculadas e comparadas com os valores fornecidos pelo usuário para orientar a estimação dos parâmetros das FDPs.

O primeiro passo do diagrama, que é também a amostragem da FDP do primeiro atributo, diz respeito ao tipo de hidrocarboneto a ser gerado, e envolve uma escolha dentre as seguintes opções de moléculas: Parafínica, Naftênica, Aromática ou Aromática Polinuclear. Como a distribuição entre estas 4 categorias é bastante particular para cada tipo de óleo, esta FDP possui forma funcional de Histograma, onde a probabilidade de cada categoria é o próprio valor do parâmetro. Como o somatório das probabilidades é 1, temos 4 categorias e 3 parâmetros, assim a última categoria (Aromática Polinuclear) é calculada por diferença.

Para cada tipo de molécula um diagrama é seguido e o restante das decisões são tomadas, ou seja, as FDPs que cabem a cada tipo de molécula são amostradas e uma molécula é gerada. Por exemplo, para uma molécula parafínica somente mais uma FDP é necessária para defini-la completamente, que é a do atributo 2, onde se decide o tamanho da parafina. O isômero montado, com a quantidade de carbonos definida, é gerado de forma aleatória respeitando a regra eurística que haverá uma cadeia principal e todas as ramificações serão metílicas. Ou seja, a quantidade de ramificações da cadeia principal e a posição é definida randomicamente, mas ao final, contando cadeia principal e todas as ramificações, a molécula terá o tamanho definido anteriormente. O diagrama para construção de cada tipo de molécula e as regras eurísticas associadas a cada uma delas está descrita no trabalho do Maciel [2019].

Desta forma, cada molécula foi gerada a partir das FDPs e estão submetidas aos seus parâmetros. Ou seja, as propriedades calculadas para uma mistura hipotética estão diretamente ligadas aos parâmetros das FDPs que geram a mistura. De forma que, se utilizarmos os mesmos parâmetros para gerar várias misturas diferentes, estas misturas conterão moléculas diferentes entre si, mas terão praticamente as mesmas propriedades. As propriedades que estão atreladas aos parâmetros para o corte de RV, e que serão utilizadas na reconstrução estão apresentadas na Tabela 2.

Para estimação destes parâmetros, Maciel [2019] propõe um método bayesiano, e para melhor entendimento deste procedimento consultar seu trabalho. O importante é termos

Tabela 2: Propriedades experimentais fornecida para o corte de RV.

Dado experimental	Associação com atributo estrutural
Densidade	Número de benzenos / Tipo de Molécula
Análise Elementar (Teor de C, H, N, S, O)	Número de benzenos / quantidade de heteroátomos
SARA (Teor de Saturados, Aromáticos, Resinas e Asfaltenos)	Tipo de Molécula / Grupos funcionais
RMN Carbono e Hidrogênio	Tipo de Molécula / Grupos funcionais
Destilação Simulada	Comprimento de cadeias parafínicas / Número de benzenos / Número de Anéis

em mente que ao final da etapa de Reconstrução Estocástica teremos um conjunto de parâmetros que geram misturas de moléculas que possui propriedades semelhantes aos dados fornecidos pelo usuário.

Desta forma, a partir deste ponto, uma mistura é escolhida dentre várias para representar a corrente de RV. Esta mistura possui 5000 moléculas em composição equimolar e este número, apesar de necessário para garantir a representatividade do modelo, é muito grande para aplicação em simuladores de processos comerciais, ou mesmo para o cálculo de equilíbrio de fases deste trabalho. Assim uma maneira de reduzir o número de moléculas a uma quantidade factível será apresentada na Seção 3.2 a seguir.

### 3.2. Partição Orientada a Medoides e Maximização entrópica

Esta é uma metodologia de clusterização, onde um grupo de 100 moléculas dentre as 5000 será escolhido como representantes da mistura. Para isto uma matriz que computa a contribuição de cada molécula para cada propriedade, chamada Matriz  $g_{ij}$ , é definida e serve como referência para seleção dos clusters. Os procedimentos para montagem desta matriz e a definição de cada propriedade está descrita no trabalho do Maciel [2019].

O conceito de medoide está associado ao individuo de uma população que possui características mais próximas à media da população. Este conceito é aplicado a 100 grupos de moléculas que ao todo conterão as 5000 moléculas da mistura. A quantidade de moléculas em cada grupo e o seu representante é definido pelo próprio algoritmo.

Para que se possa avaliar a qualidade desta clusterização, as propriedades da mistura destas 100 é calculada e uma comparação com os dados experimentais e as propriedades

da mistura original de 5000 moléculas é feita. Para isto a composição destas 100 serão somada às composições de cada uma das outras 4900 moléculas que são representadas por estas 100. No entanto, esta é uma composição provisória só para avaliar a qualidade da clusterização.

A composição final da mistura é definida pelo método de Maximização Entrópica como proposto por Hudebine and Verstraete [2011] baseados na teoria da informação. Este é um método de otimização que pretende calcular a composição de cada molécula na mistura, mas não utiliza a própria composição como variável independente, caso o fizesse teríamos 99 parâmetros a serem estimados com poucas propriedades experimentais. Ao invés disso, o autor define a entropia de Shanon para uma mistura de molécula como função objetivo do problema, e utiliza multiplicadores de Lagrange como variáveis independentes que multiplicam as restrições do problema. Estas restrições são justamente a diferença entre as propriedades experimentais e as propriedades calculadas para as 100 moléculas a partir da Matriz  $g_{ij}$ . Esta metodologia, e a forma de calcular o  $g_i$  para cada propriedade  $j$ , estão descritos no artigo de Hudebine and Verstraete [2011] e também no trabalho do Maciel [2019]. A forma final da função objetivo pra o problema de otimização é:

$$H = \ln Z \quad (1)$$

e;

$$Z = \sum_{i=1}^N \exp\left(-\sum_{j=1}^J \lambda_j g_{ij}\right) \quad (2)$$

Onde,  $H$  é a entropia de Shannon a ser minimizada,  $N$  o número de moléculas,  $J$  o número de propriedades experimentais fornecidas,  $g_{ij}$  a contribuição da molécula  $i$  para a propriedade  $j$ , e  $\lambda_j$  o multiplicador de Lagrange associado a propriedade  $j$ .

Para solução deste problema de otimização foi utilizado o método de enxame de partículas, e a partir dos parâmetros otimizados a composição de cada molécula pode ser calculada da seguinte forma:

$$x_i = \frac{\exp\left(-\sum_{j=1}^J \lambda_j g_{ij}\right)}{Z} \quad (3)$$

Após as três etapas concluída temos uma mistura de 100 moléculas e sua composição que representa uma corrente de RV desejada. Ao longo dos passos do diagrama da etapa de Reconstrução Estocástica, estas moléculas foram sendo montadas e armazenadas segundo as metodologias de representação molecular que serão descritas na Seção 4 a seguir.

## 4. Representação molecular

A forma de representar as moléculas é um dos fundamentos desta metodologia e guia tanto a etapa de reconstrução molecular, quanto a etapa de predição de propriedades. A representação molecular é focada em moléculas de hidrocarbonetos presentes no petróleo e contempla apenas as principais estruturas orgânicas e os heteroátomos mais comuns e nas suas formas mais conhecidas.

Para isto, representamos cada molécula de duas formas, sendo a primeira delas uma expansão do vetor de Lumping Orientado a Estrutura (cuja sigla em inglês é SOL, Structure-Oriented Lumping). O vetor SOL foi originalmente proposto por Quann and Jaffe [1992] com 22 estruturas, onde cada uma delas representa uma estrutura química que compõe a molécula. Estas 22 estruturas estão muito bem descritas e apresentadas no artigo do Quann and Jaffe [1992].

Uma expansão desta representação foi proposta por Maciel [2019] e está descrita no Item 4.4 do seu trabalho. Esta expansão tem por objetivo detalhar melhor as estruturas e permitir uma fácil correlação entre o vetor de estruturas expandida (SOL Expandida) e outros métodos de contribuição de grupos. O vetor expandido passa de 22 para 35 estruturas e todos os vetores (uma para cada molécula) de uma mistura de hidrocarbonetos são empacotados numa única matriz. Esta Matriz SOL Expandida é uma das formas de representação das moléculas deste trabalho.

A segunda forma de representação molecular, que apesar de não ser inédita foi criada sem se embasar em nenhuma trabalho da literatura, é a representação por grafo. Neste formato cada molécula é entendida como um grafo, onde cada nó é diferenciado por tipo de átomo; Enxofre, Nitrogênio, Oxigênio, Carbono Alifático, Carbono Aromático, Carbono Naftênico, Carbono Oléfínico, Níquel e Vanádio. Hidrogênios não são representados. Além dos nós, as arestas representam as ligações e possuem informações sobre o tipo de ligação entre os nós; Ligação simples, dupla, tripla e ligações especiais com Heteroátomos ou Metais. Neste formato, cada molécula possui uma matriz da representação por grafo apelidada de Matriz de Ligações.

A primeira representação tem como vantagem clara a simplicidade. É fácil consultar informações sobre qualquer molécula a partir da Matriz SOL Expandida, mas não é possível diferenciar algumas estruturas básicas das moléculas. Por exemplo, a maioria dos isômeros terão vetores iguais na Matriz SOL Expandida.

A segunda representação tem como principal vantagem o grande detalhamento estrutural a respeito da molécula. Nesta representação cada ligação entre os átomos é detalhada e todas as moléculas são diferenciáveis, exceto os isômeros ópticos. A união destas duas formas de representar a molécula traz as vantagens de ambas as representações. Quando uma informação simples da molécula é requerida a Matriz SOL Expandida é consultada e quando uma informação detalhada da molécula é requerida a Matriz de Ligações é inspecionada.



Ambas as formas de representação são resultados do algoritmo de Reconstrução Molecular, bem como a composição de cada molécula na mistura. No software desenvolvido em C++ para este trabalho, a classe *Moleculas* lê os arquivos gerados pela reconstrução e monta a estrutura de dados responsável pelo armazenamento destas informações. Esta classe foi omitida deste trabalho por ser de propriedade intelectual da Petrobras, no entanto o arquivo *header* (*Moleculas.hpp*) apresenta o formato desta estrutura dos dados, o que já é suficiente para a compreensão do restante do algoritmo. O cálculo das propriedades dos componentes puros e da mistura a partir das moléculas será apresentado na Seção 5 a seguir.

## 5. Predição de propriedades

Neste trabalho foi utilizado o modelo de Peng-Robinson para o cálculo do equilíbrio de fases e predição das propriedades de puros e misturas que podem ser preditos por essa equação de estado. No entanto, algumas propriedades são necessárias para a utilização deste modelo. Sendo elas; a Temperatura Crítica ( $T_c$ ), Pressão Crítica ( $P_c$ ) e fator acêntrico ( $\omega$ ). Para misturas, um parâmetro de interação binária entre os componentes ( $k_{ij}$ ) é também requerido para uma melhor representação do sistema.

Para as duas primeiras propriedades ( $T_c$  e  $P_c$ ), foi utilizado o método de contribuição de grupos proposto por Marrero and Gani [2001], mas somente a primeira ordem. Este método contempla 182 estruturas de primeira ordem, mas apenas 28 delas são suficiente para representar todas as estruturas e geradas pela Reconstrução Molecular. Estas estruturas são apresentadas no Apêndice A.

Como se pode observar, as estruturas possuem agrupamentos complexos de átomos e apenas uma consulta a Matriz SOL Expandida não é suficiente para diferenciar estes agrupamentos. Assim, uma inspeção mais detalhada da Matriz de Ligações é necessária e uma matriz contendo as moléculas nas linhas e as 28 estruturas nas colunas é montada para a mistura de hidrocarbonetos, esta matriz foi apelidada de Matriz de Grupos MG.

Para o fator acêntrico ( $\omega$ ) foi utilizado o método de contribuição de grupos proposto por Tahami et al. [2019]. Como os agrupamentos proposto por este método estão contidos nos agrupamentos proposto por Marrero and Gani [2001] a própria Matriz de Grupos MG produzida para o cálculo de  $T_c$  e  $P_c$  é utilizada para o cálculo de  $\omega$ .

A ultima propriedade necessária, o parâmetro de interação binária  $k_{ij}$ , foi calculado segundo o método de contribuição de grupos apresentado por Privat and Jaubert [2012] de nome PPR78. Mais uma vez, os agrupamentos proposto por este método estão contidos nos agrupamentos proposto por Marrero and Gani [2001] a própria Matriz de Grupos MG pode ser utilizada para o cálculo do  $k_{ij}$ .

Todos estes métodos estão implementados na Classe *MGProps* do algoritmo. Uma série de funções foi implementada para identificar as estruturas propostas por Marrero and Gani [2001] e diferencia-las umas das outras sem contar nenhum átomo mais de uma

vez, ou seja, um balaço para verificar se o numero de átomos contabilizados na Matriz de Grupos MG é de fato o número de átomos não hidrogênio contidos em cada molécula.

Mais uma vez, por questões de propriedade intelectual, o arquivo que detalha estas funções foi omitido do código, mas o arquivo *header* (MGProps.hpp) apresenta o formato desta estrutura dos dados, o que já é suficiente para a compreensão do restante do algoritmo. Todas as demais propriedades calculadas a partir da equação de estado (densidade, temperatura de ebulição e etc.), além de outras funções relativas às substâncias puras (fator de compressibilidade do puro, entalpia residual, entropia residual, energia livre residual e etc.) estão contidas na Classe *PRPropsPuros* disponível em anexo.

As propriedades da mistura (àquelas que dependem da composição) estão disponíveis na Classe *PropsMX* em anexo. Esta classe, além de conter as funções que calculam estas propriedades (como fator de compressibilidade e coeficientes de fugacidade da mistura), também contém as funções relativas ao equilíbrio de fases que serão discutidas na Seção 6 a seguir.

## 6. Equilíbrio Multifásico

O equilíbrio de fases parte das seguintes hipóteses: 1) homogeneidade de temperatura e pressão do sistema; 2) igualdade do potencial químico de cada componente em todas as fases; 3) o sistema em equilíbrio se encontra no mínimo global da energia livre de Gibbs. O que implica em:

$$\begin{cases} 1) & T^1 = T^2 = \dots = T^F \quad \& \quad P^1 = P^2 = \dots = P^F \\ 2) & \mu_i^1 = \mu_i^2 = \dots = \mu_i^F \\ 3) & \min(G) \end{cases}$$

Onde  $F$  é o numero total de fases em equilíbrio e o índice  $i$  representa um componente genérico do sistema.

A hipótese 2 é entendida como condição necessária para o cálculo do equilíbrio de fases, e é a partir dela que boa parte do equacionamento se inicia. Como a fugacidade é uma “escala” do potencial químico mais facilmente calculada, é conveniente transformar a igualdade proposta na hipótese 2 em igualdades de fugacidades, da seguinte forma:

$$f_i^1 = f_i^2 = \dots = f_i^F \quad (4)$$

A hipótese 3, que afirma que o sistema em equilíbrio está num mínimo global de energia livre, é entendida como condição suficiente do equilíbrio de fases. A resolução do pro-

blema proposto pela hipótese 2 nos leva naturalmente a uma condição de mínimo local de energia livre do sistema, no entanto, é preciso garantir que este mínimo seja também global. Para sistemas bifásicos, é sabido que só há um ponto de mínimo, e portanto, a condição necessária também é suficiente, ou seja, a resolução do problema proposto pela hipótese 2 é automaticamente um mínimo global e a condição suficiente está também atendida.

Para sistemas multifásicos uma verificação da estabilidade é necessária, ou seja, uma verificação se o ponto de mínimo encontrado é de fato um mínimo global. Para isto, serão apresentados as metodologias das seções 6.1 e 6.2 a seguir.

### 6.1. Distância ao Plano Tangente

A distância ao plano tangente (TPD) é uma equivalência geométrica a análise da variação da energia livre do sistema ( $\delta G$ ) em relação ao surgimento de nova fase hipotética em quantidade infinitesimal ( $\delta e$ ) com composição  $w$ . Então, segundo Michelsen and Mollerup [2007], esta razão entre  $\delta G$  e  $\delta e$  é a própria definição de TPD, dada por:

$$TPD = \frac{\delta G}{\delta e} = \sum_{i=1}^C w_i (\mu_i^w - \mu_i^z) \quad (5)$$

Onde  $C$  é o número de componentes,  $w_i$  é a composição de  $i$  na fase  $w$ ,  $\mu_i^w$  é o potencial químico do componente  $i$  na fase  $w$ ,  $\mu_i^z$  é o potencial químico do componente  $i$  na fase de referência  $z$ .

O calculo da TPD parte de uma fase de referência com composição conhecida ( $z$ ). Esta composição é um ponto da superfície de energia livre do sistema a partir do qual é construído um plano tangente. Outros pontos com quantidade infinitesimal ( $\delta e$ ) e composição proposta ( $w$ ) serão avaliados em relação a este plano tangente. Se o ponto de referência é um mínimo global, o valor de TPD calculada para qualquer outra fase e composição possível é sempre positivo, ou seja, estará acima do plano tangente. Em outras palavras, qualquer outra fase testada vai ter TPD positiva se o sistema é estável na única fase de referência ( $z$ ). Se o valor do TPD for zero para alguma fase testada, isto indica que estas fases podem estar em equilíbrio. Se alguma fase testada apresentar TPD negativo, indica que o sistema não encontra-se na sua forma mais estável na fase de referência.

Esta discussão fica um pouco mais clara quando aplicada a um sistema binário e bifásico. Para isto, Michelsen and Mollerup [2007] trouxe uma representação da energia livre do sistema como sendo a energia livre de Gibbs de mistura de dois componentes. A representação gráfica desta energia em função da composição de um dos componentes é apresentada na Figura 2, onde a linha trastejada representa a o plano tangente ao ponto de composição  $Z_1 = 0,13$ . Como se pode observar, tendo  $Z_1$  como referência, o valor da TPD para qualquer composição até o ponto A ( $Z_A$ ) tem como resultado valores positi-

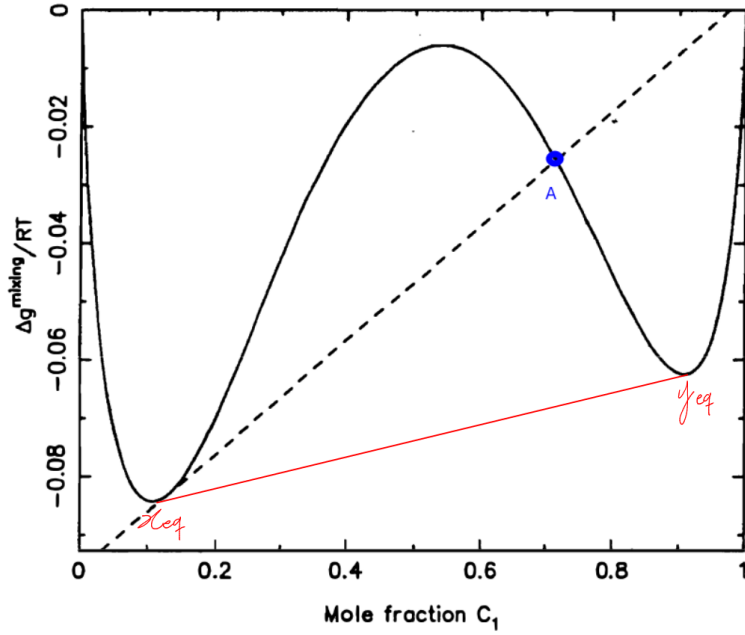


Figura 2: Energia Livre de Gibbs da mistura binária de CH<sub>4</sub> - H<sub>2</sub>S a 190 K e 4,56 MPa, indicando o plano tangente ao ponto de composição  $Z_1 = 0,13$ . (Fonte: Michelsen e Mollerup, 2007)

vos, estão acima do plano tangente. No entanto, qualquer composição onde  $Z > Z_A$  tem como resultado valores negativos, indicando que a mistura monofásica na composição de referência não é estável. Desta forma, o cálculo do equilíbrio de fases deve ser resolvido e o sistema deve ser bifásico com uma das fases com composição menor que  $Z_1$  e a outra fase com composição maior que  $Z_A$ .

A reta vermelha do gráfico indica o plano tangente das fases em equilíbrio de composições  $x_{eq}$  e  $y_{eq}$ , que será a solução do equilíbrio de fases. Ambas as fases tocam o mesmo plano tangente, e portanto a TPD delas é nula e nenhuma outra fase possível terá TPD negativa. Além disso, o sistema se encontra monofásico para qualquer composição  $z < x_{eq}$  ou  $z > y_{eq}$  e neste intervalo ( $x_{eq} < z < y_{eq}$ ) o sistema será bifásico com composição das fases constante  $x_{eq}$  e  $y_{eq}$ .

O cálculo da TPD pode ser feito da forma apresentada na Equação 5, no entanto, o cálculo desta grandeza em função do coeficiente de fugacidade é mais simples e mais cômodo. Neste trabalho optei por utilizar a equação de estado de Peng-Robinson para todas as possíveis fases do sistema, líquida ou vapor, o que induz a um cálculo mais intuitivo do coeficiente de fugacidade de cada componente na mistura a partir dessa equação de estado. Assim, partindo da definição de fugacidade e tendo a condição de gás ideal como referência para ambas as fases, a diferença de potenciais químicos da Equação 5 pode ser escrita como:

$$\frac{\Delta\mu_i^{w-z}}{RT} = \ln \frac{f_i^w}{f_i^z} = \ln \frac{w_i \hat{\phi}_i^w}{z_i \hat{\phi}_i^z} \quad (6)$$

Onde  $\hat{\phi}_i^w$  é o coeficiente de fugacidade do componente  $i$  na fase  $w$  e, da mesma forma,  $\hat{\phi}_i^z$  é o coeficiente de fugacidade do componente  $i$  na fase  $z$ . Esta equação pode ser reescrita da seguinte forma:

$$\frac{\Delta\mu_i^{w-z}}{RT} = \ln w_i + \ln \hat{\phi}_i^w - h_i \quad (7)$$

Onde:

$$h_i = \ln z_i + \ln \hat{\phi}_i^z \quad (8)$$

Importante observar que o termo  $h_i$  só depende das informações referentes a fase de referência ( $z$ ) cujas propriedades são conhecidas e constantes, portanto,  $h_i$  é uma constante para cada componente. De maneira análoga, os dois primeiros termos da Equação 7 dependem de informações referentes a fase de teste  $w$ , cuja composição é desconhecida. Além disso, o próprio  $\hat{\phi}_i^w$  também depende da composição, o que implica num procedimento iterativo para o cálculo do TPD. Esta dedução foi desenvolvida e proposta por Michelsen and Mollerup [2007] e apresentada em aula durante o curso. Assim, o cálculo da TPD por substituições sucessivas pode ser feito a partir das seguintes equações:

$$K = -\ln \left( \sum_{i=1}^C Y_i \right) \quad (9)$$

Onde  $K$  é uma variável de estabilidade do sistema que se nula o sistema encontra-se no mínimo de energia livre. A contribuição de cada componente para a variável de estabilidade ( $Y_i$ ) é definida da seguinte forma:

$$Y_i = \exp(h_i - \ln(\hat{\phi}_i^w)) \quad (10)$$

Os termos que compõe  $Y_i$  já foram previamente apresentados, no entanto é importante lembrar que o cálculo de  $\hat{\phi}_i^w$  depende da composição da fase  $w$ , que ainda continua sendo desconhecida, mas pode ser calculada, segundo proposto por Michelsen and Mollerup [2007], da seguinte forma:

$$w_i = \frac{Y_i}{\sum_{i=1}^C Y_i} \quad (11)$$

Como a variável  $K$  contém, de forma implícita, o somatório das diferenças dos potências químicos de cada componente em ambas as fases, esta variável se relaciona com o próprio valor da TPD da seguinte forma:

$$TPD = K R T \quad (12)$$

Assim, a partir de uma estimativa inicial de composição da fase  $w$ , as Equações 8, 9, 10, 11 e 12 podem ser utilizadas para o cálculo da TPD por substituição sucessiva como apresentado no Algoritmo 1. No software desenvolvido em C++ para este trabalho, a função *tpd* da classe *PropsMX* realiza exatamente este procedimento.

Importante observar que a função *tpd*, apresentada no Algoritmo 1, além de retornar o valor da TPD para fase testada em relação a referência, também retorna o vetor de composição convergida da fase testada ( $w$ ) e o vetor dos coeficientes de fugacidade para cada componente na fase testada ( $\hat{\phi}^w$ ). Uma vez definida o conceito da TPD e sua forma de cálculo, fica ainda por se definir para quais fases possíveis  $w$  o sistema será testado. Este aspecto será discutido na Seção 6.2 a seguir.

## 6.2. Análise de estabilidade

A análise de estabilidade do sistema utiliza o valor da TPD discutida na seção anterior para indicar se o sistema encontra-se estável na fase de referência  $z$ , ou se existe outras possíveis fases mais estáveis ou em equilíbrio no sistema. Como o objetivo final deste trabalho é modelar o processos de desasfaltação, a fase de referência adotada é sempre a composição global da mistura de entrada na estado líquido. Para definir as possíveis fases a serem testadas foram utilizadas as boas praticas propostas por Michelsen and Mollerup [2007], e apresentadas em aula, válidas para o sistema estudado. Sendo elas:

- A mesma composição de referência na fase vapor
- Uma fase tipo líquido puro para cada componente

Como os resíduos de vácuo costuma ser representados por uma mistura diversa de hidrocarbonetos, a suposição de uma fase líquida para cada componente puro costuma indicar o aparecimento de muitas fases líquidas com tpd negativo. No entanto, a maioria delas resulta na mesma composição convergida, ou seja, convergem para mesma fase. Assim, todas as fases repetidas devem ser eliminadas. Tendo isso em mente, a função *AnáliseEstabilidade* da classe *PropsMX* realiza o procedimento apresentado no Algoritmo 2 para Análise de estabilidade do sistema.

---

**Algoritmo 1:** Cálculo da TPD

---

**Entrada:** z - Vetor composição da fase de referência z  
w - Vetor de estimativas iniciais da composição da fase w  
EFz - Variável booleana para o estado físico da fase z, 0=líqu e 1=vap  
EFw - Variável booleana para o estado físico da fase w, 0=líqu e 1=vap  
T - Temperatura do sistema em K  
P - Pressão do sistema em bar

função tpd (z, w, EFz, EFw, T, P)

$\hat{\phi}^z = f(z, EFz, T, P)$

**for**  $i=1$  até  $C$  **do**

$h_i = \ln z_i + \ln \hat{\phi}_i^z$       Equação 8

**end**

**while**  $SumErroY > tolerância$  **do**

$\hat{\phi}^w = f(w, EFw, T, P)$

$SumY = 0$

**for**  $i=1$  até  $C$  **do**

$Y_i = \exp(h_i - \ln(\hat{\phi}_i^w))$       Equação 10

$SumY += Y_i$

**end**

$SumErroY = 0$

**for**  $i=1$  até  $C$  **do**

$w_i^{new} = Y_i / SumY$       Equação 11

$SumErroY += (w_i - w_i^{new})^2$

$w_i = w_i^{new}$

**end**

**end**

$K = -\ln(SumY)$       Equação 9

$TPD = KRT$       Equação 12

retorna  $TPD, w, \hat{\phi}^w$

---

---

**Algoritmo 2:** Análise de Estabilidade

---

**Entrada:**  $z$  - Vetor composição da fase de referência  $z$

$T$  - Temperatura do sistema em K

$P$  - Pressão do sistema em bar

função *AnaliseEstabilidade* ( $z, T, P$ )

$vFases$  - vetor que contém as informações das fases com TPD negativas

$EFz = 0$

Testando composição  $z$  na fase vapor:

$w = z$

$EFw = 1$

$TPD, w, \hat{\phi}^w = tpd(z, w, EFz, EFw, T, P)$

**if**  $TPD < 0$  **then**

$vFases \leftarrow TPD, w, \hat{\phi}^w$

**end**

Testando fase tipo líquido puro para cada componente:

$EFw = 0$

**for**  $i=1$  até  $C$  **do**

**for**  $j=1$  até  $C$  **do**

**if**  $i == j$  **then**

$w_i = 1$

**else**

$w_i = 0$

**end**

**end**

$TPD, w, \hat{\phi}^w = tpd(z, w, EFz, EFw, T, P)$

**if**  $TPD < 0$  **then**

$vFases \leftarrow TPD, w, \hat{\phi}^w$

**end**

**end**

Remove fases repetidas

retorna  $vFases$

---



A partir do Algoritmo 2 é possível observar que a *função AnaliseEstabilidade* retorna um vetor de possíveis fases com TPD negativa contendo: o valor da TPD; a composição convergida de cada fase; e o coeficiente de fugacidade de cada componente nas fases. Desta forma, esta função possui ao menos duas utilidades, a primeira delas é ser executada antes do cálculo rigoroso de equilíbrio multifásico para indicar possíveis fases e suas estimativas iniciais de composição. A segunda utilidade é verificar se o resultado encontrado pelo cálculo rigoroso é de fato um mínimo global, ou seja, se o resultado do equilíbrio de fases rigoroso for submetido à *função AnaliseEstabilidade* e está retornar um vetor vazio, o resultado do equilíbrio de fases rigoroso é um mínimo global.

O cálculo rigoroso do equilíbrio multifásico e multicomponente parte do resultado da análise de estabilidade e será apresentado na Seção 6.3 a seguir.

### 6.3. Modelo ELLV

O cálculo rigoroso do equilíbrio multifásico, ou Equilíbrio Líquido-Líquido-Vapor (ELLV), para temperatura e pressão definidas se resume ao cálculo da fração de cada fase ( $\beta_j$ ) presente no sistema. Para isto, o algoritmo do ELLV foi implementado conforme proposto por Michelsen and Mollerup [2007].

Seu equacionamento parte da igualdade de potenciais químicos de cada componente em todas as fases, e do balanço material das fases e dos componentes, até chegar num sistema de equações de Rachford-Rice (uma para cada fase). Para ver o equacionamento de forma detalhada consultar Michelsen and Mollerup [2007], aqui trarei apenas as definições mais importantes para o desenvolvimento do modelo e algumas adaptações implementadas.

A primeira definição é a do fator de equilíbrio ( $K_i^j$ ) que tem como referência uma fase  $R$ , que é a fase mais abundante do sistema. Ao contrário do proposto por Michelsen and Mollerup [2007] para manter a fase de referência fixa, como sendo a ultima, deixarei a fase de referência dinâmica, sendo sempre a fase com maior fração ( $\beta_j$ ). Desta forma, o  $K_i^j$  é expresso como:

$$K_i^j = \frac{y_i^j}{y_i^R} = \frac{\hat{\phi}_i^R}{\hat{\phi}_i^j} \quad (13)$$

Sendo o  $K_i^j$  o fator de equilíbrio do componente  $i$  na fase  $j$ ;  $y_i^j$  a composição do componente  $i$  na fase  $j$ ;  $y_i^R$  a composição do componente  $i$  na fase de referência  $R$ ;  $\hat{\phi}_i^j$  o coeficiente de fugacidade do componente  $i$  na fase  $j$ ; e  $\hat{\phi}_i^R$  o coeficiente de fugacidade do componente  $i$  na fase de referência  $R$ . A composição do componente em uma fase qualquer pode ser calculada da seguinte forma:

$$y_i^j = \frac{z_i K_i^j}{1 + \sum_{l=1}^F \beta^l (K_i^l - 1)} \quad (14)$$

Onde  $F$  o numero total de fases. Optei por não diferenciar o calculo da composição da fase de referência das demais fases, pois o valor de  $K_i^R = 1$  o que faz com que esta fase não contribua com o somatório do denominador e nem altere o valor do numerador.

Michelsen and Mollerup [2007] apresenta uma função objetivo ( $Q(\beta)$ ) que transforma o sistema de equações de Rachford-Rice num problema de otimização multivariável onde deseja-se minimizar o valor da função de convergência  $Q(\beta)$ , que pode ser definida por:

$$Q(\beta) = \sum_{j=1}^F \beta^j - \sum_{i=1}^C z_i \ln E_i \quad (15)$$

Onde:

$$E_i = \sum_{k=1}^F \frac{\beta^k}{\hat{\phi}_i^k} \quad (16)$$

O vetor de gradientes ( $g^j$ ) da função objetivo  $Q(\beta)$  é dado por:

$$g^j = \frac{\partial Q}{\partial \beta^j} = 1 - \sum_{i=1}^C \frac{z_i}{E_i \hat{\phi}_i^j} \quad (17)$$

A matriz Hessiana ( $H^{jk}$ ) da função objetivo  $Q(\beta)$  é dado por:

$$H^{jk} = \frac{\partial g^j}{\partial \beta^k} = \sum_{i=1}^C \frac{z_i}{E_i^2 \hat{\phi}_i^j \hat{\phi}_i^k} \quad (18)$$

A partir do conjunto de equações definidas é possível resolver o problema de otimização seguindo as etapas propostas por Michelsen and Mollerup [2007] que apresentarei a seguir, com algumas adaptações.

- i. Avalie a função objetivo  $Q$  para estimativa inicial de  $\beta$ , Equações 15 e 16.

- ii. Calcule o vetor de gradientes  $g^j$  e a matriz Hessiana  $H^{jk}$ , Equações 17 e 18. Resolva o problema de Newton para as correções de  $\beta$ ;  $H\Delta\beta + g = 0$ , que implica em  $\Delta\beta = -H^{-1}g^j$ . Comece com o valor do “modificador do passo”  $\alpha = 1$ .
- iii. Calcule os novos valores de fração das fases;  $\beta_{novo} = \beta_{antigo} + \alpha\Delta\beta$ . Se algum  $\beta_{novo}^j < 0$ , altere o valor de  $\alpha$  para que o menor  $\beta_{novo}^j$  seja **quase** zero.
- iv. Avalie a função objetivo  $Q_{novo}$  para o conjunto  $\beta_{novo}$ , se houver redução do valor da função objetivo;  $\beta = \beta_{novo}$ . Se um acréscimo da função objetivo for observado faça  $\alpha = \alpha/2$  e retorne ao item iii.
- v. Se alguma fase apresentou  $\beta^j = 0$  avalie o valor do gradiente desta fase  $j$ . Se  $g^j \geq 0$  remova esta fase permanentemente, se  $g^j < 0$  remova esta fase de forma temporária, posteriormente avalie o sistema com o retorno dessa fase.
- vi. Recalcule  $y_i^j$ ,  $K_i^j$  para o  $\beta$  convergido nos passos anteriores, Equações 13 e 14. Como  $y_i^j = f(K_i^j, \beta)$ , e por sua vez  $K_i^j = f(\phi_i^j)$ , que por sua vez  $\phi_i^j = f(y_i^j)$ , uma convergência neste passo é necessária.
- v. Reavalie a função objetivo  $Q_{novo}$  para  $y_i^j$ ,  $K_i^j$  e  $\beta$  convergidos. Se  $Q_{novo} = Q$  o sistema convergiu, se  $Q_{novo} < Q$  faça  $Q = Q_{novo}$  e retorne ao item ii. Se alguma fase foi removida de forma temporária na iteração anterior, retorne com esta fase ao sistema e retorne ao item ii.

A estimativa inicial das fases do sistema, bem como sua composição inicial é dada pela função *AnalyseEstabilidade*, e a estimativa inicial de  $\beta$  é uma distribuição uniforme entre as possíveis fases presentes no sistema, ou seja  $\beta^j = 1/F$ , sendo  $F$  o número de fases iniciais. Este procedimento foi implementado na função *FlashELLV* da classe *PropsMX* e está mais detalhadamente apresentada no Algoritmo 3.

Gostaria de destacar as alterações proposta para o algoritmo. Além da não diferenciação da fase de referência no calculo da composição e do fator de equilíbrio do sistema, como comentado anteriormente, alterei também o passo *iii* para que o novo beta fosse **quase** zero, e não diretamente zero como proposto originalmente. Algumas vezes, a solução do problema de Newton do passo *ii* nos leva a valores negativos de  $\beta$  em alguma fase, mas no próximo passo ela convergiria para um valor positivo próximo a zero. Para evitar que esta fase fosse removida do algoritmo antes mesmo de ter a oportunidade de existir, optei por modificar o valor de  $\alpha$  para que ele fosse 99,99% do valor que zeraria  $\beta$ . Se o sistema fica mais estável sem esta fase, isso acontecerá naturalmente no passo *iv*. Outra alteração implementada foi o passo *vi*, onde a composição, coeficiente de fugacidade e fator de equilíbrio dos componentes são calculados de forma interativa para os  $\beta$  convergidos anteriormente. A proposta original calcula a composição somente após a convergência completa de  $\beta$  mantendo os  $K_i^j$  fixos. Esta consideração pode até ser válida quando o sistema possui poucas fases e poucas moléculas, mas para sistemas mais complexos a variação do  $K_i^j$  ao longo do processo se mostrou bastante significativa e um calculo mais acurado desta variável ao longo da convergência do algoritmo foi requerida.

A partir do algoritmo proposto e implementado é possível avaliar o equilíbrio de fases do processo de desasfaltação para correntes de Resíduo de Vácuo que foram caracterizados por Reconstrução Molecular. A performance do modelo comparada aos dados experimentais será exibida na Seção 7 a seguir.

---

**Algoritmo 3:** Flash Multifásico

---

**Entrada:**  $z$  - Vetor composição da fase de referência  $z$   
 $T$  - Temperatura do sistema em K  
 $P$  - Pressão do sistema em bar

função FlashELLV ( $z, T, P$ )

$vFases$  - Vetor das fases do sistema

$Q$  - Variável que armazena o valor da função objetivo

$BQ$  - Variável que armazena o menor valor da função objetivo

$vFases = AnaliseEstabilidade(z, T, P)$

$F \leftarrow$  Tamanho de  $vFases$

$vFases.\beta = 1/F$

$Q \leftarrow$  Equações 15 e 16

**While**  $BQ \neq Q$

$BQ = Q$

$g \leftarrow$  Equação 17

$H \leftarrow$  Equação 18

$\alpha = 1$

$\Delta\beta = H^{-1} * g$

$\beta_{novo} = vFases.\beta + \alpha * \Delta\beta$

$\beta_{min} = 100$

**for**  $j=1$  até  $F$  **do**

**if**  $\beta_{novo}^j < 0$  &  $\beta_{novo}^j < \beta_{min}$  **then**

$\alpha = -\beta_{novo}^j / \Delta\beta^j$

$\beta_{min} = \beta_{novo}^j$

**end**

**end**

---

---



---

```

while  $BQ \geq Q$  do
  |  $\beta_{novo} = \beta + \alpha * \Delta\beta$ 
  |  $Q \leftarrow$  Equações 15 e 16
  |  $\alpha = \alpha/2$ 
end

 $vFases.\beta = \beta_{novo}$ 
for  $j=1$  até  $F$  do
  | if  $\beta^j == 0$  then
  |   | if  $g^j \geq 0$  then
  |   |   | Remove fase  $j$  definitivamente de  $vFases$ 
  |   | else
  |   |   | Transfere fase  $j$  de  $vFases$  para vetor de fases temporarias  $vTmp$ 
  |   | end
  | end
end

while  $SumYquad > tolerância$  do
  |  $y_{novo} \leftarrow$  Equação 14
  |  $vFases.\hat{\phi} = f(y_{novo})$ 
  |  $K_i^j \leftarrow$  Equação 13
  |  $SumYquad = (vFases.y - y_{novo})^2$ 
  |  $vFases.y = y_{novo}$ 
end

 $Q \leftarrow$  Equações 15 e 16
if Alguma fase removida na iteração anterior then
  | Retorna com a fase removida de  $vTmp$  para  $vFases$ 
end
 $F \leftarrow$  Tamanho de  $vFases$ 
end

retorna  $vFases$ 

```

---

## 7. Resultados

Nesta seção serão exibidos os resultados do desempenho do modelo sob alguns aspectos. O primeiro deles será em relação aos modelos de predição de propriedades por contribuição de grupos e a predição de propriedades de substâncias puras em geral. Em seguida, o desempenho do algoritmo do equilíbrio multifásico será avaliado quanto sua capacidade de prever o comportamento de algumas misturas conhecidas. Posteriormente, será avaliada a capacidade do método de Reconstrução Molecular para representar as propriedades desejadas de uma mistura, e também sua performance na predição de propriedades de validação não fornecidas como dado de entrada. Por fim, o desempenho do modelo completo como preditor do comportamento de unidades de desasfaltação será avaliado.

### 7.1. Predição de propriedades

A capacidade de predição de propriedades do modelo foi avaliada apenas para as propriedades necessárias para o cálculo do equilíbrio de fases utilizando a equação de estado de Peng Robinson, sendo elas; temperatura crítica ( $T_c$ ), pressão crítica ( $P_c$ ), fator acêntrico ( $\omega$ ) e o parâmetro de interação binária ( $k_{ij}$ ), que será avaliado em separado.

Foram selecionadas 21 moléculas cujos dados experimentais estavam disponíveis, que pudessem ser representadas pela representação molecular proposta e cujas estruturas pudessem ser avaliadas pelo algoritmo que contabiliza os grupos dos métodos de contribuição de grupos. O método de Tahami et al. [2019] foi utilizado para o cálculo do  $\omega$ , e o método de Marrero and Gani [2001] para  $T_c$ ,  $P_c$ . Além dessas propriedades, a temperatura normal de ebulição ( $T_b$ ) foi calculada pelo método de Marrero and Gani [2001] e também pela equação de estado de Peng Robinson. Todos estes dados estão apresentados na Tabela 3 e também nos gráficos da Figura 3.

Como se pode observar, o método de Marrero and Gani [2001] possui boa capacidade de predição das propriedades para a maioria das moléculas, mesmo estando somente com os grupos de primeira ordem implementados. Para moléculas mais pesadas, as propriedades começam a apresentar diferenças significativas entre os valores experimentais e os calculados, como se pode observar para o *N-heptacotane*, *pentatriacontylcyclohexane* e *pentatriacontylbenzene*. Estas moléculas estão justamente no meio da faixa de massa molar das moléculas geradas para representar o RV, que variam de 500 a 2000, e este desvio pode ser observado no final da curva da Figura 3a e no início da curva da Figura 3b.

Para o método de predição do fator acêntrico proposto por Tahami et al. [2019], Figura 3c, os valores estimados apresentam desvio significativo em toda faixa. No entanto, tanto para as propriedades críticas quanto para o  $\omega$ , nenhum outro método de contribuição de grupos testado apresentou estabilidade para a faixa de massa molar contida na mistura, e mesmo apresentando algum desvio os valores ainda são razoáveis. Avaliando a curva de ponto de ebulição, Figura 3d, observa-se que esta propriedade foi razoavelmente bem predita tanto utilizando diretamente o método de contribuição de grupos de Marrero and

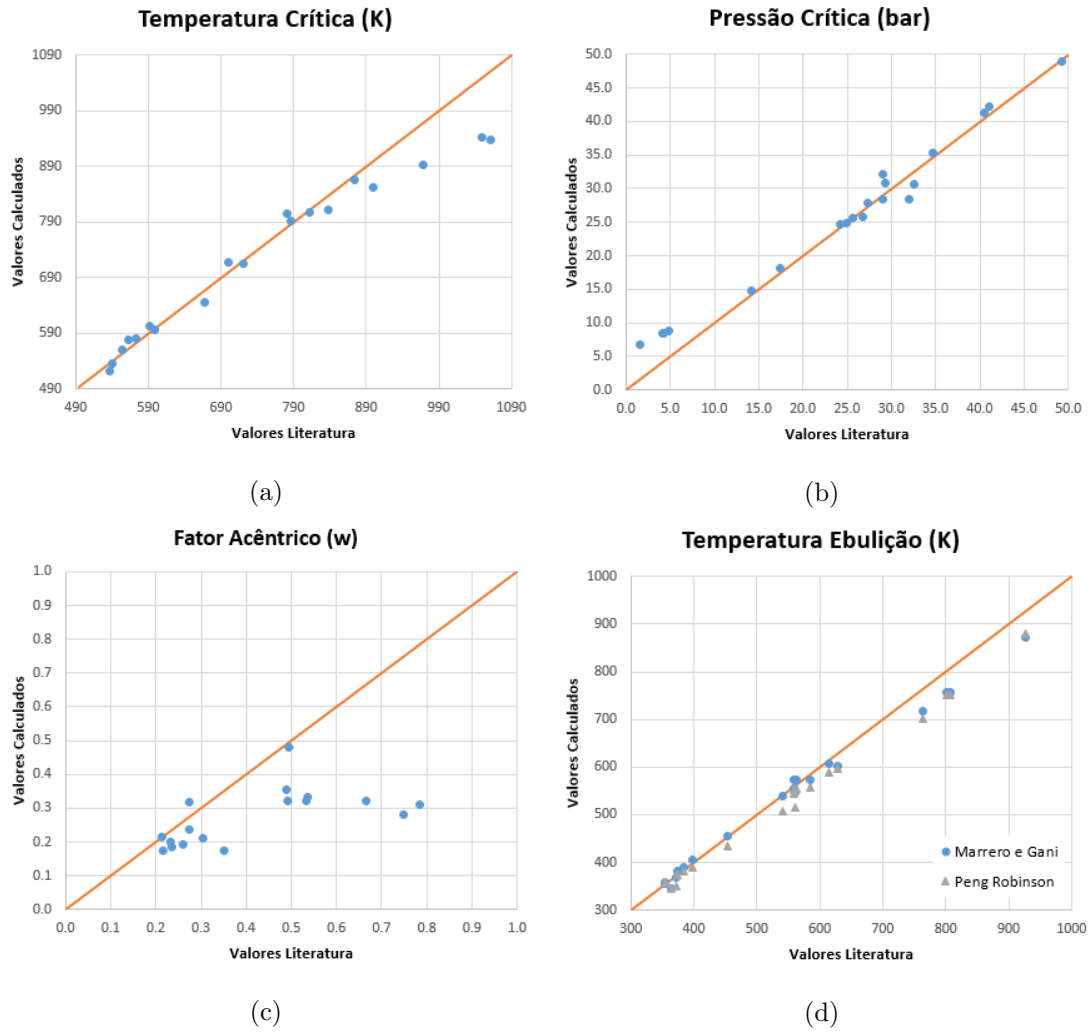


Figura 3: Comparação dos valores experimentais com o calculado pelos métodos de contribuição de grupos para as seguintes propriedades: (a) Temperatura crítica por Marrero and Gani [2001] ; (b) Pressão crítica por Marrero and Gani [2001]; (c) Fator acêntrico por Tahami et al. [2019]; (d) Temperatura normal de ebulição por Marrero and Gani [2001] (azul) e por Peng Robinson usando  $T_c$ ,  $P_c$  e  $w$  calculados pelos métodos de contribuição de grupos.

Tabela 3: Propriedades experimentais e calculadas.

Experimentais								
n <sup>o</sup>	Molécula	formula	MMw	$\omega$	$T_c[K]$	$P_c[bar]$	$T_b[K]$	fonte
0	anthracene	C14H10	178.2	0.4890	873.0	29.0	615.2	yaws
1	1,3,8-trimethylnaphthalene	C13H14	170.3	0.4920	786.8	26.7	558.2	yaws
2	dibenzofuran	C12H8O	168.2	0.2750	837.8	32.0	557.9	yaws
3	dibenzopyrrole	C12H9N	167.2	0.4940	899.0	32.6	627.9	yaws
4	1-tridecanal	C13H26O	198.3	0.7850	700.0	17.4	540.2	yaws
5	4,6-dimethyl-1-ethylnaphthalene	C14H16	184.3	0.5360	811.4	24.2	585.4	yaws
6	trans-1,3-dimethylcyclohexane	C8H16	112.2	0.2340	598.0	29.4	397.6	yaws
7	butylcyclohexane	C10H20	140.3	0.2740	667.0	25.7	454.1	yaws
8	2-butylnaphthalene	C14H16	184.3	0.5330	781.0	25.0	562.2	yaws
9	pentatriacontylbenzene	C41H76	569.1	-	1048.4	4.3	801.2	yaws
10	pentatriacontylcyclohexane	C41H82	575.1	-	1060.8	4.1	807.2	yaws
11	pentatriacontane	C35H72	493.0	-	968.3	4.8	763.2	yaws
12	hexadecane	C16H34	226.4	0.7470	720.6	14.2	560.0	yaws
13	N-heptacontane	C70H142	983.9	-	7326.5	1.6	926.2	yaws
14	benzene	C6H6	78.1	0.2150	562.1	49.2	353.2	petrosim
15	toluene	C7H8	92.1	0.2596	591.8	41.0	383.8	petrosim
16	n-heptane	C7H16	100.2	0.3498	540.2	27.4	371.6	petrosim
17	cyclohexene	C6H12	84.2	0.2133	553.2	40.5	353.9	petrosim
18	methyl-cyclohexane	C7H14	98.2	0.2330	572.1	34.8	374.1	petrosim
19	2,3-dimethylpentane	C7H16	100.2	0.3050	537.3	29.1	362.9	petrosim
20	propanol	C3H8O	60.1	0.6650	508.3	47.6	355.4	petrosim

Calculadas								
n <sup>o</sup>	Molécula	formula	MMw	$T_b^{PR}[K]$	$T_b[K]$	$T_c[K]$	$P_c[bar]$	$\omega$
0	anthracene	C14H10	178.2	588.1	607	866.1	32.05	0.3557
1	1,3,8-trimethylnaphthalene	C13H14	170.3	543.5	555.9	790.9	25.83	0.3206
2	dibenzofuran	C12H8O	168.2	552.8	573.9	812.2	28.49	0.3191
3	dibenzopyrrole	C12H9N	167.2	596.1	601.5	852	30.62	0.4807
4	1-tridecanal	C13H26O	198.3	508.5	538.5	717	18.14	0.3095
5	4,6-dimethyl-1-ethylnaphthalene	C14H16	184.3	558.1	572.8	806.7	24.73	0.333
6	trans-1,3-dimethylcyclohexane	C8H16	112.2	389.3	405.4	597.2	30.75	0.185
7	butylcyclohexane	C10H20	140.3	434.3	454.3	644.8	25.69	0.2371
8	2-butylnaphthalene	C14H16	184.3	555.5	572.3	805.3	24.77	0.32
9	pentatriacontylbenzene	C41H76	569.1	752.5	757.6	942.1	8.51	0.603
10	pentatriacontylcyclohexane	C41H82	575.1	750.3	755.7	937.2	8.38	0.6111
11	pentatriacontane	C35H72	493.0	703	718.7	892.7	8.82	0.5141
12	hexadecane	C16H34	226.4	515.6	547.3	716.4	14.85	0.2826
13	N-heptacontane	C70H142	983.9	879.2	871.7	1050.9	6.86	0.9004
14	benzene	C6H6	78.1	358.9	359	578.5	48.95	0.1734
15	toluene	C7H8	92.1	382.5	389.2	604.1	42.18	0.1918
16	n-heptane	C7H16	100.2	351.7	369.8	536.2	27.83	0.1725
17	cyclohexene	C6H12	84.2	358.3	355.5	560.5	41.29	0.2158
18	methyl-cyclohexane	C7H14	98.2	374.3	381.8	579.6	35.32	0.2004
19	2,3-dimethylpentane	C7H16	100.2	345.3	344.2	521.2	28.33	0.2127
20	propanol	C3H8O	60.1	321	322.8	500.8	56.41	0.3235



Gani [2001], quanto calculados por Peng Robinson utilizando as propriedades críticas e fator acêntrico estimados, mesmo para moléculas de alta massa molar, o que indica que, mesmo com algum desvio nas propriedades estimadas, o modelo ainda consegue prever as propriedades dos puros de forma satisfatória.

## 7.2. Equilíbrio de fases

O algoritmo proposto neste trabalho foi implementando conforme apresentado nos arquivos em anexo, e antes de avaliar o desempenho da desasfaltação, é válido testar esta implementação diante de dados de ELLV em comparação com simuladores comerciais. Para isto, foi feita uma simulação no Petro-SIM contendo *benzeno*, *n-hexano* e *água*. Como o objetivo é avaliar a implementação do modelo, os valores de  $T_c$ ,  $P_c$ ,  $\omega$  e  $k_{ij}$  do simulador foram alimentados diretamente no código. O simulador fixa o valor de  $k_{ij}$  em 0,5 entre a água e qualquer outro hidrocarboneto, para o binário *benzeno* e *n-hexano* o  $k_{ij}$  utilizado é o meso do simulador de  $7.10^{-3}$ . O restante das propriedades dos puros estão na Tabela 3. O resultado para este equilíbrio multifásico está apresentado na Tabela 4.

Tabela 4: Comparação entre o algoritmo implementado e o simulador comercial Petro-SIM para a mistura *benzeno*, *n-hexano* e *água* a 1 atm e com composição global de 0,3000; 0,3000; 0,4000, respectivamente.

	Petro-SIM			Este trabalho		
Temperatura [°C]	63,2			63,9		
fase	vapor	líquida	aquosa	vapor	líquida	aquosa
fração da fase	0,5000	0,2129	0,2871	0,5000	0,2060	0,2941
	<i>composição</i>					
<i>benzeno</i>	0,3634	0,5558	0,0000	0,3719	0,5538	0,0000
<i>n-hexano</i>	0,4117	0,4422	0,0000	0,4170	0,4443	0,0000
<i>água</i>	0,2249	0,0020	1,0000	0,2111	0,0019	1,0000

É possível observar que, mesmo com alguma diferença entre os resultados, o desempenho é bem similar. Uma diferença de 0,7°C foi observada para encontrar a mesma fração vaporizada da mistura (0,5) com algumas diferenças na composição. No entanto, o algoritmo implementado indicou a formação das 3 fases do sistema, o que indica que este foi bem implementado e está apto a prever o equilíbrio multifásico.

Para avaliar equilíbrio de fases com as propriedades estimadas pelos métodos de contribuição de grupos, foi calculado o envelope de fases para o sistema *benzeno* e *n-hexano*. Estes dados foram comparados com a curva gerada pelo simulador comercial Petro-SIM e também com dados experimentais disponibilizados pela universidade de Dortmund em seu site de banco de dados que estão apresentados da Figura 4. A comparação entre as propriedades dos componentes puros envolvidos e seus dados experimentais já foram apresentadas na Tabela 3. O valor do  $k_{ij}$  gerado para esta mistura foi de  $6,564.10^{-3}$  enquanto que o valor presente no Petro-SIM é de  $7.10^{-3}$ .

Como se pode observar, os dados gerados pelo Petro-SIM são bem aderentes aos dados experimentais. Os dados gerados com as propriedades dos puros estimadas possui um erro significativo em relação ao dados experimentais. O valor do  $k_{ij}$  para este binário foi bem estimado, mas mesmo uma pequena diferença entre os valores ainda geram uma mudança de comportamento da curva. Ou seja, o equilíbrio de fases está bem implementado no algoritmo proposto e os métodos de contribuição de grupos estima de forma satisfatória as propriedades, mesmo que contenham algum desvio, é uma forma interessante de estimar propriedades de moléculas desconhecidas.

### 7.3. Reconstrução Molecular

Uma vez avaliado o desempenho da predição de propriedades e da implementação do algoritmo de equilíbrio multifásico, é importante avaliar a Reconstrução Molecular quanto a capacidade de prever as propriedades fornecidas da mistura, e também algumas propriedades de validação não fornecidas. Para isto, foram reconstruídas as propriedades de três RVs que aqui serão identificados apenas como A, B e C, devido ao sigilo das informações. As propriedades destas correntes estão apresentadas juntamente com o resultado da caracterização destes RVs com 100 moléculas nas Tabelas 5, 6 e 7.

Como se pode observar, as propriedades fornecidas foram bem representadas após a reconstrução. As propriedades de validação variaram sempre em torno do mesmo percentual de erro para todos os RVs. A RMNH do está intimamente ligada a regras eurísticas assumidas durante a proposição do diagrama de construção das moléculas, mas os valores apresentados parecem coerentes com os valores experimentais. A destilação e densidade apresentam erros dentro da faixa dos valores esperados pelos métodos de contribuição de grupos (ainda mais se mudarmos a escala de temperatura para Kelvin), além de serem propriedades de fácil correção caso se deseje. Além disso, a curva de destilação aqui foi calculada apenas pelo ordenamento dos pontos de ebulição dos componentes puros (o que é prática comum na caracterização de petróleo), e não por modelo de equilíbrio de fases. Desta forma, apesar de haver algumas diferenças entre as propriedades experimentais e calculadas, os RVs estão bem caracterizados e as propriedades da mistura proposta representam de forma satisfatória os dados experimentais.

### 7.4. Desasfaltação

Uma vez caracterizado os RVs e implementado o modelo de equilíbrio multifásico, é chegada a hora de avaliar a aplicação destas ferramentas na representação do processo de desasfaltação. Para isto, foram levantados dados experimentais do equilíbrio de fases destes três RVs em células PVT. Os RVs foram alimentados no algoritmo proposto e todos eles sem a presença de solvente formaram uma mistura homogênea em fase líquida. Quando misturado com o propano nenhum deles apresentou a formação de duas fases líquidas. Com GLP somente, o RV-A indicou a formação de duas fases, o que pode indicar que a formação de duas fases líquidas está eminente no modelo.

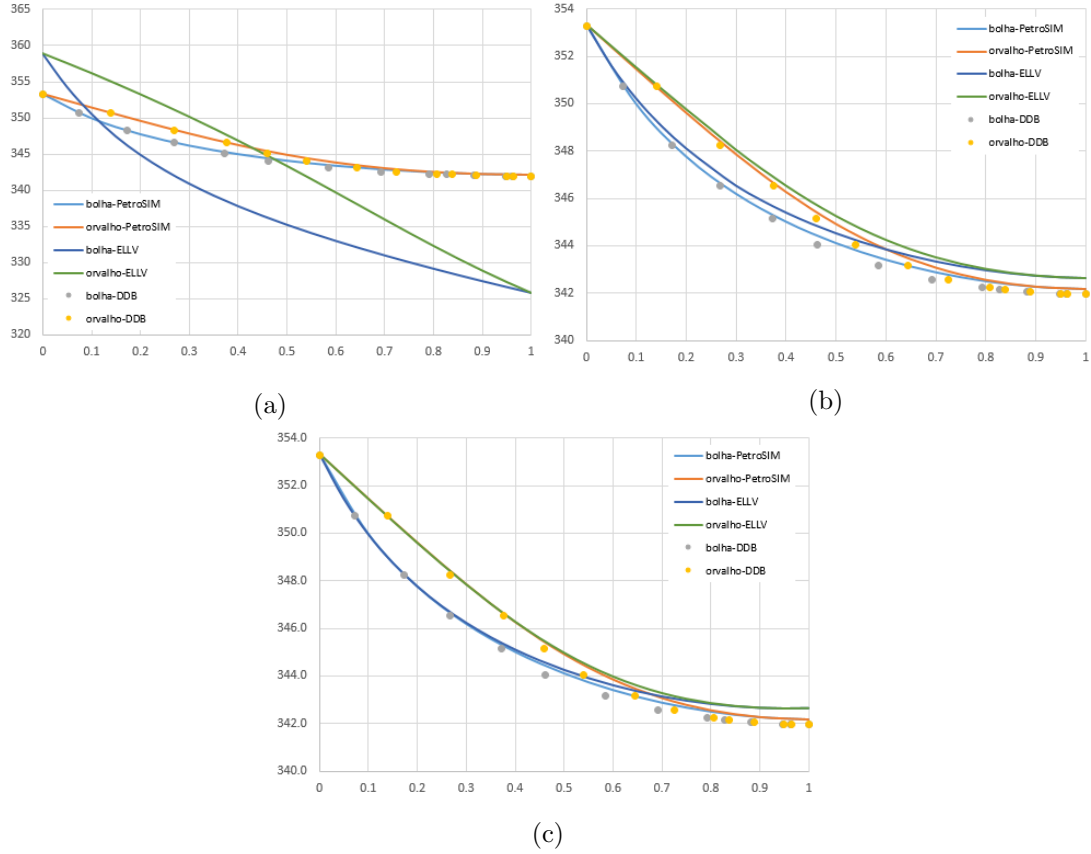


Figura 4: Comparação entre envelope de fases gerado pelo modelo, os dados gerados no Petro-SIM e os dados experimentais do banco de dados da universidade de Dortmund (DDB), sendo: (a) curva do modelo gerada com propriedades estimadas pelos métodos de contribuição de grupos ( $k_{ij} = 6,564.10^{-3}$ ); (b) curva do modelo gerada com as mesmas propriedades dos puros do Petro-SIM e  $k_{ij}$  estimado por contribuição de grupos ( $k_{ij} = 6,564.10^{-3}$ ); (c) curva do modelo gerada com as mesmas propriedades dos puros e  $k_{ij}$  do Petro-SIM ( $k_{ij} = 7.10^{-3}$ )

Tabela 5: Propriedades do RV-A e comparação com resultado da caracterização.

Propriedades	Experimental	Reconstruído	Erro
<b>Propriedades Fornecidas</b>			
MMW	786,4	786,4	0,0%
<b>Análise Elementar</b>			
Carbono (%m/m)	87,21	87,21	0,0%
Hidrogênio (%m/m)	10,64	10,64	0,0%
Enxofre (%m/m)	1,28	1,28	0,0%
Nitrogênio (%m/m)	0,86	0,86	0,0%
Oxigênio (%m/m)	0,01	0,01	0,3%
<b>Análise SARA</b>			
Saturados (%m/m)	12	12	0,0%
Aromáticos (%m/m)	37	37	0,0%
Resinas (%m/m)	32	32	0,0%
Asfaltenos (%m/m)	18	18	0,0%
<b>Ressonância Magnética de Carbono (RMNC)</b>			
Carbono saturado (%m/m)	69,9	69,9	0,0%
Carbono insaturado (%m/m)	30,1	30,1	0,0%
<b>Propriedades de Validação</b>			
<b>Ressonância Magnética de Hidrogênio (RMNH)</b>			
Hidrogênio aromático (%m/m)	6,4	4,9	23,0%
Hidrogênio $\alpha$ (%m/m)	13,4	7,8	42,0%
Hidrogênio $\beta$ (%m/m)	61,2	75,6	23,5%
Hidrogênio $\gamma$ (%m/m)	19,0	11,7	38,3%
<b>Destilação Simulada (°C)</b>			
5(%m)	456	464	1,8%
10(%m)	527	490	7,1%
30(%m)	596	521	12,5%
50(%m)	646	606	6,2%
70(%m)	702	753	7,3%
<b>Densidade</b>			
densidade (20/4)	1,0090	1,1385	12,8%

Tabela 6: Propriedades do RV-B e comparação com resultado da caracterização.

Propriedades	Experimental	Reconstruído	Erro
<b>Propriedades Fornecidas</b>			
MMW	717,9	717,9	0,0%
<b>Análise Elementar</b>			
Carbono (%m/m)	86,70	86,70	0,0%
Hidrogênio (%m/m)	11,40	11,40	0,0%
Enxofre (%m/m)	1,10	1,10	0,0%
Nitrogênio (%m/m)	0,80	0,27	66,7%
Oxigênio (%m/m)	0,00	0,53	-
<b>Análise SARA</b>			
Saturados (%m/m)	19	19	0,0%
Aromáticos (%m/m)	40	40	0,0%
Resinas (%m/m)	34	34	0,0%
Asfaltenos (%m/m)	7	7	0,0%
<b>Ressonância Magnética de Carbono (RMNC)</b>			
Carbono saturado (%m/m)	77,6	78,0	0,5%
Carbono insaturado (%m/m)	22,4	22,0	1,6%
<b>Propriedades de Validação</b>			
<b>Ressonância Magnética de Hidrogênio (RMNH)</b>			
Hidrogênio aromático (%m/m)	4,8	2,5	48,5%
Hidrogênio $\alpha$ (%m/m)	9,5	6,8	28,2%
Hidrogênio $\beta$ (%m/m)	64,9	69,4	7,0%
Hidrogênio $\gamma$ (%m/m)	20,6	21,3	3,2%
<b>Destilação Simulada (°C)</b>			
5(%m)	490	439	10,6%
10(%m)	525	490	6,6%
30(%m)	576	493	14,3%
50(%m)	620	563	9,3%
70(%m)	676	654	3,3%
<b>Densidade</b>			
densidade (20/4)	0,9770	1,0701	9,5%

Tabela 7: Propriedades do RV-C e comparação com resultado da caracterização.

Propriedades	Experimental	Reconstruído	Erro
<b>Propriedades Fornecidas</b>			
MMW	737,7	737,7	0,0%
<b>Análise Elementar</b>			
Carbono (%m/m)	87,10	87,10	0,0%
Hidrogênio (%m/m)	11,50	11,50	0,0%
Enxofre (%m/m)	0,59	0,59	0,0%
Nitrogênio (%m/m)	0,78	0,78	0,0%
Oxigênio (%m/m)	0,03	0,03	0,0%
<b>Análise SARA</b>			
Saturados (%m/m)	-	33	-
Aromáticos (%m/m)	-	27	-
Resinas (%m/m)	-	37	-
Asfaltenos (%m/m)	2	2	-
<b>Ressonância Magnética de Carbono (RMNC)</b>			
Carbono saturado (%m/m)	75,6	75,6	0,0%
Carbono insaturado (%m/m)	24,4	24,4	0,0%
<b>Propriedades de Validação</b>			
<b>Ressonância Magnética de Hidrogênio (RMNH)</b>			
Hidrogênio aromático (%m/m)	5,3	5,7	8,0%
Hidrogênio $\alpha$ (%m/m)	9,9	6,9	29,9%
Hidrogênio $\beta$ (%m/m)	66,0	63,9	3,2%
Hidrogênio $\gamma$ (%m/m)	18,8	23,4	24,6%
<b>Destilação Simulada (°C)</b>			
5(%m)	505	449	11,2%
10(%m)	526	455	13,4%
30(%m)	579	464	19,8%
50(%m)	628	546	13,0%
70(%m)	690	735	6,6%
<b>Densidade</b>			
densidade (20/4)	0,9847	1,0808	9,5%

Tabela 8: Comparação dos resultados obtidos para rendimentos do equilíbrio Líquido-Líquido entre os RVs e solventes.

Solvente	T	P	RSO	Experimental		Simulado	
	[°C]	[kgf/cm²g]		ODES [%]	RASF [%]	ODES [%]	RASF [%]
RV-A							
Propano	50	30	4	20,1	79,9	66,0	34,0
Propano	50	30	7	19,3	80,7	73,0	27,0
Propano	50	30	9	21,6	78,4	76,1	23,9
Propano	70	30	4	13,3	86,7	39,0	61,0
Propano	70	30	7	18,0	82,0	46,4	53,6
Propano	70	30	9	25,8	74,2	50,0	50,0
GLP	50	30	4	35,3	64,7	39,1	60,9
GLP	70	30	4	38,9	61,1	37,0	63,0
GLP	70	30	7	46,3	53,7	46,2	53,8
GLP	70	30	9	47,7	52,3	50,5	49,5
RV-B							
Propano	70	30	4	19,3	80,7	56,9	43,1
Propano	70	30	7	26,7	73,3	65,6	34,4
Propano	70	30	9	32,1	67,9	70,1	29,9
GLP	70	30	4	67,7	32,3	53,0	48,0
GLP	70	30	9	68,3	31,7	71,0	29,0
RV-C							
Propano	60	45	6	29,3	70,8	85,8	14,2
GLP	60	30	5	43,7	56,3	63,3	36,7
GLP	60	35	5	46,6	53,4	63,3	36,7
GLP	65	35	5	41,5	58,5	64,6	35,4
GLP	70	30	5	37,0	63,0	64,6	35,4

Com o intuito de avaliar a variação de rendimento com as demais variáveis do processo e vendo que o simulador atribui o valor contante de 0,5 para o  $k_{ij}$  entre água e qualquer outro hidrocarboneto, e que isto implica na formação automática de duas fases líquido, adicionei o valor de 0,2 ao  $k_{ij}$  entre solvente e RV. Após esta manipulação com intuítos meramente acadêmicos, o resultado de rendimentos da desasfaltação está apresentado na Tabela 8.

Apesar dos rendimentos simulados para o propano ficarem bem distantes dos valores experimentais, os valores calculados para o GLP ficaram razoavelmente satisfatórios. Mas acima de tudo, as tendencias dos rendimentos ficaram coerentes o que indica que o modelo captura o comportamento do processo. Ao que parece, uma forma mais eficiente de estimar os valores de  $k_{ij}$  pode resultar num modelo com performance de rendimentos suficientes para o uso industrial. No entanto, este estudo fica para o trabalho os próximos trabalhos.

## 8. Conclusão

Este trabalho teve a pretensão de modelar o processo de desasfaltação, e o equilíbrio de fases envolvido nesta extração, utilizando a Reconstrução Molecular para estimação das moléculas presentes na corrente de Resíduo de Vácuo, que é a carga da unidade. Para isto, métodos de contribuição de grupos foram utilizados para estimar as propriedades das moléculas e da mistura e um algoritmo de equilíbrio multifásico foi implementado utilizando a equação de estado de Peng Robinson.

O modelo de propriedades por contribuição de grupos possibilita a estimativa automatizada de propriedades de moléculas desconhecidas. Desta forma, mesmo apresentando algum desvio em relação aos valores experimentais, ele atende de forma satisfatória a predição de propriedades para este fim.

A metodologia de Reconstrução Molecular representa de forma satisfatória as propriedades dos Resíduos de de Vácuo, tanto para propriedades fornecida quanto para propriedades de validação, e erro encontrado nas de validação são da mesma ordem de grandeza dos erros dos próprios métodos de contribuição de grupos.

O algoritmo de equilíbrio multifásico foi implementado e reproduziu resultados semelhantes ao simulador comercial Petro-SIM, o que indica sucesso na sua implementação. Algumas modificações foram sugeridas para melhorar a convergência do método proposto por Michelsen and Mollerup [2007], como discutido na Seção 6.3, e algumas sugestões para futuras alteração foram identificadas nesta mesma seção. Apesar de algumas limitações de convergência, a implementação atendeu de forma satisfatória a este trabalho convergindo em todas as simulações.

A representação da unidade de desasfaltação através do modelo proposto não apresentou o desempenho satisfatório. Quando o RV foi misturado ao solvente, com os parâmetros  $k_{ij}$  estimados por contribuição de grupos, não houve a formação de duas fases líquidas, como acontece na unidade. Observando que, no simulador comercial, o parâmetro  $k_{ij}$  entre água e qualquer hidrocarboneto é sempre de 0,5, e que isto implica na formação automática de duas fases líquidas, foi acrescido o valor de 0,2 ao  $k_{ij}$  entre as moléculas do RV e as moléculas de solvente. Após esta intervenção uma segunda fase líquida é observada ao misturar solvente e RV, e a variação dos rendimentos de ODES e RASF em relação as variáveis mais importantes do processo, como a Temperatura e Pressão da extração e a Razão Solvente Óleo, obedecem as tendências observadas na unidade.

Em suma, os modelos de predição de propriedades dos puros, a Reconstrução Molecular e implementação do algoritmo de equilíbrio multifásico apresentou desempenho satisfatório. O modelo de Peng-Robinson e/ou a estimativa do  $k_{ij}$  não apresentou um resultado satisfatório e um método mais promissor para a estimativa deste parâmetro entre RV e solvente parece promissor para tornar esta ferramenta aderente o suficiente para se tornar aplicável em processos industriais.


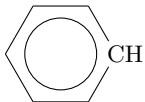
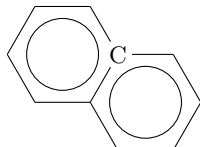
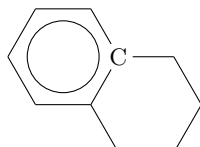
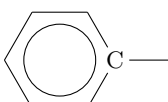

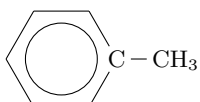
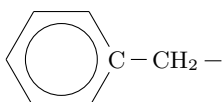
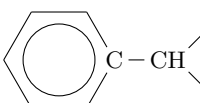
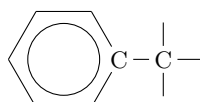


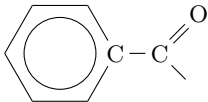
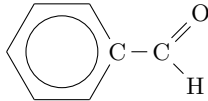
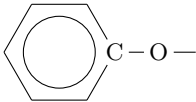
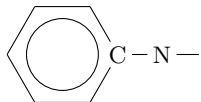
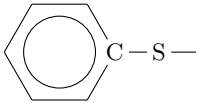
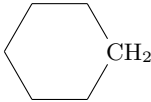
## Referências

- Michael L. Michelsen and Jorgen M. Mollerup. *Thermodynamic Models: Fundamentals & Computational Aspects*. Tie-Line Publications, 2 edition, 2007. ISBN 87-989964-3-4.
- L. Pereira de Oliveira, A. Trujillo Vazquez, J. J. Verstraete, and M. Kolb. Molecular reconstruction of petroleum fractions: Application to vacuum residues from different origins. *Energy & Fuels*, 27:3622–3641, 2013.
- Helton Siqueira Maciel. Molecular reconstruction of heavy petroleum fractions, 2019.
- D. Hudebine and J. J. Verstraete. Reconstruction of petroleum feedstocks by entropy maximization. application to fcc gasolines. *Oil & Gas Science and Technology - Rev. IFP Energies nouvelles*, 66(2):437–460, 2011.
- R. J. Quann and S. B. Jaffe. Structure-oriented lumping: Describing the chemistry of complex hydrocarbon mixtures. *Industrial & Engineering Chemistry Research*, 31: 2483–2497, 1992.
- Jorge Marrero and Rafiqul Gani. Group-contribution based estimation of pure component properties. *Fluid Phase Equilibria*, pages 183–207, 2001.
- Saba Tahami, Habib Ghasemitabar, and Kamyar Movagharnejad. Estimation of the acentric factor of organic compounds via a new group contribution method. *Fluid Phase Equilibria*, 499, 2019.
- Romain Privat and Jean-Noël Jaubert. Addition of the sulfhydryl group ( sh) to the ppr78 model: Estimation of missing group-interaction parameters for systems containing mercaptans and carbon dioxide or nitrogen or methane, from newly published data. *Fluid Phase Equilibria*, 334:197–203, 2012.

## A. Estruturas do método de contribuição de grupos de Marrero e Gani

Neste apêndice são apresentadas as estruturas do método de contribuição de grupos que se aplicam às moléculas geradas pela Reconstrução Molecular, sendo eles:

1. Carbono alifático primário;	$CH_3$	$CH_3 -$
2. Carbono alifático secundário;	$CH_2$	$-CH_2-$
3. Carbono alifático terciário;	$CH$	$-CH$ 
4. Carbono aromático;	$aCH$	
5. Carbono aromático fundido com aromático;	$aC_{fa}$	
6. Carbono aromático fundido com naftênico;	$aC_{fn}$	
7. Carbono aromático substituído genérico;	$aC_{qq}$	
8. Nitrogênio piridínico;	$aN$	
9. Carbono aromático ligado a $CH_3$ ;	$aCCH_3$	
10. Carbono aromático ligado a $CH_2$ ;	$aCCH_2$	
11. Carbono aromático ligado a $CH$ ;	$aCCH$	
12. Carbono aromático ligado a $C$ ;	$aCC$	

13. Hidroxila;	$OH$	$-OH$
14. Cetona ligada a $CH_3$ ;	$CH_3CO$	$CH_3-C \begin{array}{l} \nearrow O \\ \searrow \end{array}$
15. Cetona ligada a $CH_2$ ;	$CH_2CO$	$-CH_2-C \begin{array}{l} \nearrow O \\ \searrow \end{array}$
16. Cetona ligada a $CH$ ;	$CHCO$	$\begin{array}{l} \diagup \\ \diagdown \end{array} CH-C \begin{array}{l} \nearrow O \\ \searrow \end{array}$
17. Cetona ligada a Carbono aromático;	$aCCO$	
18. Aldeído;	$COH$	$\begin{array}{l} \nearrow O \\ -C \\ \searrow H \end{array}$
19. Aldeído ligado a Carbono aromático;	$aCalCO$	
20. Éter ligado a Carbono aromático;	$aCO$	
21. Amina ligada a $CH_2$ ;	$NH_2CH_2$	$NH_2-CH_2-$
22. Amina ligada a Carbono aromático;	$aCN$	
23. Mercapta ligada a $CH_2$ ;	$SHCH_2$	$SH-CH_2-$
24. Mercapta ligada a átomo genérico;	$SH$	$SH-$
25. Amina ligada a átomo genérico;	$NH_2$	$NH_2-$
26. Enxofre ligada a Carbono aromático;	$aCS$	
27. Carbono naftênico secundário;	$nCH_2$	
28. Carbono naftênico terciário;	$nCH$	