



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE  
Fakulta jaderná a fyzikálně inženýrská



# **Optimalizace algoritmu větví a mezí s aplikací na testování fázové stability vícesložkových směsí**

## **Optimization of the branch and bound algorithm with application for phase stability testing of multicomponent mixtures**

Diplomová práce

Autor: **Bc. Martin Jex**  
Vedoucí práce: **doc. Ing. Jiří Mikyška, Ph.D.**  
Akademický rok: 2020/2021

## ZADÁNÍ DIPLOMOVÉ PRÁCE

Student:	Bc. Martin Jex
Studijní program:	Aplikace přírodních věd
Studijní obor:	Aplikované matematicko-stochastické metody
Název práce (česky):	Optimalizace algoritmu větví a mezí s aplikací na testování fázové stability vícesložkových směsí
Název práce (anglicky):	Optimization of the branch and bound algorithm with application for phase stability testing of multicomponent mixtures

### Pokyny pro vypracování:

- 1) V návaznosti na Výzkumný úkol pokračujte ve vývoji algoritmu pro testování fázové stability vícesložkových směsí v proměnných VTN.
- 2) Pomocí konvexně-konkávních rozkladů příslušných veličin odvoďte horní a dolní odhady tlaku a chemických potenciálů na podsimplexech základního simplexu.
- 3) S využitím těchto odhadů navrhnete algoritmus testování fázové stability vícesložkových směsí a tento algoritmus implementujte.
- 4) Navržený algoritmus testujte na vhodných příkladech. Diskutujte zejména výpočetní náročnosti tohoto algoritmu ve srovnání s dříve navrženými metodami.
- 5) Shrňte svoje zkušenosti s využitím metod globální optimalizace pro testování stability vícesložkových směsí.

Doporučená literatura:

- 1) M. Locatelli, F. Schoen, Global Optimization - Theory, Algorithms, and Applications. SIAM Philadelphia, 2013.
- 2) R. Paulavičius, J. Žilinskas, Simplicial Global Optimization, Springer-Verlag New York, 2014.
- 3) B. Jaumard, Ch. Meyer, H. Tuy, Generalized Convex Multiplicative Programming via Quasiconcave Minimization. Journal of Global Optimization, 10, 1997, 229-256.
- 4) A. M. Oberman, The Convex Envelope is the Solution of a Nonlinear Obstacle Problem. In Proceedings of the American Mathematical Society, 135(6), June 2007, 1689-1694.
- 5) H. Tuy, Global Minimization of a Difference of 2 Convex Functions. Mathematical Programming Study, 30, 1987, 150-182.
- 6) H. Thi, T.P. Dinh, DC programming and DCA: thirty years of developments. Special Issue of Mathematical Programming, 169(1), 2018, 5-68.

Jméno a pracoviště vedoucího diplomové práce:

doc. Ing. Jiří Mikyška, Ph.D.

katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 120 00 Praha 2

Jméno a pracoviště konzultanta:

Datum zadání diplomové práce: 31.10.2020

Datum odevzdání diplomové práce: 3.5.2021

Doba platnosti zadání je dva roky od data zadání.

*Poděkování:*

Chtěl bych zde poděkovat svému školiteli doc. Ing. Jiřímu Mikyškovi, Ph.D. za ochotu, vstřícnost a odborné i lidské zázemí při vedení mé diplomové práce.

*Čestné prohlášení:*

Prohlašuji, že jsem tuto práci vypracoval samostatně a uvedl jsem všechnu použitou literaturu.

V Praze dne 5. května 2021

Martin Jex

*Název práce:*

**Optimalizace algoritmu větví a mezí s aplikací na testování fázové stability vícesložkových směsí**

*Autor:* Bc. Martin Jex

*Obor:* Aplikované matematicko-stochastické metody

*Druh práce:* Diplomová práce

*Vedoucí práce:* doc. Ing. Jiří Mikyška, Ph.D. Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 120 00 Praha 2

*Abstrakt:* Práce se zabývá otázkou stability směsi v proměnných VTN. Tento problém je řešen pomocí globální optimalizace funkce TPD (tangent plane distance). Globální optimalizace je provedena metodou větví a mezí, která je vylepšena oproti základnímu algoritmu o efektivnější prořezávání stromu vznikajícím touto metodou. Toto zlepšení je odvozeno z nutných podmínek extrému, které vedou na doplňující podmínky na tlak a chemické potenciály. Funkce popisující tyto podmínky nejsou konvexní, a tedy se v práci odvozuje jejich konvexně-konkávní rozklady.

*Klíčová slova:* fázová stabilita, globální optimalizace, konvexně-konkávní rozklad, metoda větví a mezí, směsi látek

*Title:*

**Optimization of the branch and bound algorithm with application for phase stability testing of multicomponent mixtures**

*Author:* Bc. Martin Jex

*Abstract:* This work examines the question of VTN phase stability testing. This problem is solved by global minimization of the TPD (tangent plane distance) function. The global optimization is performed using applying the branch and bound algorithm, which is improved, in comparison to its basic variant, by using a more effective pruning of the tree arising from the algorithm. This improvement is derived from the necessary conditions of an extremum, which leads to supplementary conditions for pressure and chemical potentials. Functions describing these conditions are not convex, therefore, in this work, we derive and apply its convex-concave decompositions.

*Key words:* phase stability, global optimisation, convex-concave split, branch and bound method, mixtures

# Obsah

Úvod	8
<b>1 Představení problému</b>	<b>9</b>
1.1 Fyzikální motivace	9
1.2 Matematická formulace	12
<b>2 Výpočetní algoritmus</b>	<b>13</b>
2.1 Algoritmus větví a mezí a jeho aplikace na naši situaci	13
2.2 Konvexní relaxace	16
2.3 Barycentrické souřadnice	16
2.4 Optimalizace v barycentrických souřadnicích	16
2.5 Dělení simplexu	17
2.6 Implementace algoritmu větví a mezí	18
<b>3 Efektivnější ořezávání stromu v algoritmu větví a mezí</b>	<b>19</b>
3.1 Nutná podmínka extrému	19
3.2 Ořezávání pomocí Pengovy-Robinsonovy stavové rovnice	20
3.2.1 Afinity transformace ve stavové rovnici	20
3.2.2 Ověřování předpokladů věty o konvexně-konkávním rozkladu a její použití	21
3.2.3 Konvexně-konkávní odhad Pengovy-Robinsonovy stavové rovnice	23
3.2.4 Ladění parametrů $\gamma_1, \gamma_2, \gamma_3$	24
3.2.5 Kritéria pro ořezávání simplexu na základě stavové rovnice	24
3.3 Ořezávání pomocí chemických potenciálů	25
3.3.1 Afinity transformace v chemických potenciálech	26
3.3.2 Ověřování předpokladů věty o konvexně-konkávním rozkladu a její použití	26
3.3.3 Konkávní a konvexní odhady chemických potenciálů	31
3.3.4 Ladění parametrů $\gamma$	32
3.3.5 Kritéria pro ořezávání simplexu na základě chemických potenciálů	33
3.4 Upravený algoritmus	34
<b>4 Příklady</b>	<b>35</b>
4.1 Směs metanu a propanu	36
4.2 Směs dekanu a oxidu uhličitého	37
<b>Závěr</b>	<b>47</b>

# Úvod

Fyzika je často motivací matematických problémů a tak je tomu i zde. V této práci se budeme zabývat testováním stability vícesložkové směsi, která je jedním z jejích základních problémů. Toto téma je velmi důležité v mnoha oblastech průmyslu (např. zkapaňování plynů a skladování směsí). Tato témata řeší užší odvětví fyziky jako je termodynamika nebo fyzikální chemie. Zajímavou otázkou je, jak bude daná látka, její složení, vypadat při určitých fyzikálních parametrech jako je např. teplota, tlak nebo přítomnost vnějších polí, tato otázka se řeší pomocí termodynamických potenciálů typicky Helmholtzovy volné energie nebo Gibbsova potenciálu. Tato veličina je funkcí stavových proměnných a koncentrací jednotlivých složek obsažených ve zkoumané směsi a je na ni kladena extrémální podmínka. Toto je důsledek druhého termodynamického zákona. Z termodynamického pohledu na věc je rovnovážný stav definován jako stav soustavy s maximální entropií  $S$ . Entropií se rozumí funkce vnitřní energie  $U$ , objemu  $V$  a počtu částic dílčích složek  $N_i$  systému  $S(U, V, N_1, \dots, N_k)$ . V řadě aplikací je nesrovnatelně výhodnější použít místo těchto proměnných jiné, pro daný problém přirozenější, jako například absolutní teplotu  $T$  nebo tlak  $P$ . V těchto situacích používáme právě už zmíněné potenciály Helmholtzův  $A(T, V, N_1, \dots, N_k)$  nebo Gibbsův  $G(T, p, N_1, \dots, N_k)$ , které jsou odvozeny od vnitřní energie  $U(S, V, N_1, \dots, N_k)$  pomocí Legendreovy transformace  $A = U - TS$  resp.  $G = U - TS - pV$  (teplota a tlak jsou definované jako derivace vnitřní energie podle odpovídající proměnné – entropie, objem). Výhodou použití těchto potenciálů je mimo jiné také možnost je jednoduše kombinovat se stavovou rovnicí, která je většinou psána v proměnných teplota  $T$ , objem  $V$  a látkové složení  $N_i$  tj.  $p = p(T, V, N_1, \dots, N_k)$ . Kromě volby potenciálu voleného podle relevantních stavových proměnných je druhou kritickou volbou výběr stavové rovnice média, které studujeme. Stavová rovnice je výraz svazující stavové proměnné tlak, látkové složení a objem. Fyzika nabízí v tomto směru nepřehledné množství kandidátů. Naše volba je diktována požadavkem dostatečné jednoduchosti výrazu pro tlak a fyzikální relevantností pro studované situace.

Používáme Pengovu-Robinsonovu stavovou rovnici v následujícím tvaru

$$p^{EOS}(T, \mathbf{c}) = \frac{\sum_{i=1}^N c_i RT}{1 - \sum_{i=1}^N b_i c_i} - \frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2}, \quad (1)$$

kde  $c_i = \frac{N_i}{V}$  je objemová koncentrace  $i$ -té složky a parametry  $a_{ij}$ ,  $b_i$  jsou definované následovně

$$a_{ij} = (1 - \delta_{i-j}) \sqrt{a_i a_j}, \quad (2)$$

$$a_i(T) = 0,45724 \frac{R^2 T_{i,crit}^2}{P_{i,crit}} \left[ 1 + m_i (1 - \sqrt{T_{i,r}}) \right]^2, \quad (3)$$

$$b_i = 0,0778 \frac{RT_{i,crit}}{P_{i,crit}}, \quad (4)$$

$$m_i = \begin{cases} 0,37464 + 1,54226\omega_i - 0,26992\omega_i^2 & \omega_i < 0,5; \\ 0,3796 + 1,485\omega_i - 0,1644\omega_i^2 + 0,01667\omega_i^3 & \omega_i \geq 0,5; \end{cases} \quad (5)$$

kde parametry  $T_{i,crit}$ ,  $P_{i,crit}$  jsou kritické teploty a tlaky  $i$ -té složky. Dále  $T_{i,r} = \frac{T}{T_{i,crit}}$  je redukovaná teplota,  $\omega_i$  je acentrický faktor,  $\delta_{i-j}$  je interakční koeficient a  $R$  je molární plynová konstanta. Pengovu-Robinsonovu rovnici používáme na vnitřku simplexu daném nerovnostmi

$$c_i > 0 \quad \text{a} \quad \sum_{i=1}^n b_i c_i < 1, \quad (6)$$

který pro nás bude přípustnou množinou.



# Kapitola 1

## Představení problému

Nejprve uvedeme fyzikální motivaci problému fázové stability vícesložkové směsi, poté přejdeme k matematické formulaci tohoto problému.

### 1.1 Fyzikální motivace

Zabýváme se směsí látek, u kterých jsou dané počáteční molární koncentrace  $\mathbf{c}^*$  a teplota  $T^*$ . Termodynamický rovnovážný stav tohoto systému je popsán Helmholtzovou energií  $A = A(V, T, N)$ , respektive hustotou Helmholtzovy energie  $a = \frac{A}{V}$  definovanou rovnicí

$$a(\mathbf{c}) = RT \sum_{i=1}^N c_i \ln \frac{c_i}{c_0} - RT \left( \sum_{i=1}^N c_i \right) \ln \left( 1 - \sum_{i=1}^N b_i c_i \right) - \frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{2\sqrt{2} \sum_{i=1}^N b_i c_i} \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right). \quad (1.1)$$

Pro popis naší situace jsou důležité parciální derivace podle proměnných  $c_i$ . Těmto parciálním derivacím hustoty Helmholtzovy energie (1.1) říkáme chemické potenciály a jsou definované následující rovnicí

$$\begin{aligned} \frac{\partial a}{\partial c_i}(\mathbf{c}) = \mu_i(\mathbf{c}) = & RT \left( \ln \frac{c_i}{c_0} + 1 \right) - RT \ln \left( 1 - \sum_{i=1}^N b_i c_i \right) + RT b_i \frac{\sum_{i=1}^N c_i}{1 - \sum_{i=1}^N b_i c_i} \\ & - \frac{\left( 2 \sum_{j=1}^N c_j a_{ij} \right) \left( \sum_{i=1}^N b_i c_i \right) - b_i \sum_{i,j=1}^N a_{ij} c_i c_j}{2\sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right) \\ & - \left( \frac{b_i \sum_{i,j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} \right). \end{aligned} \quad (1.2)$$

Někdy je vhodné použít vyjádření hustoty Helmholtzovy energie pomocí chemických potenciálů (1.2) a Pengovy-Robinsonovy stavové rovnice (1)

$$a(\mathbf{c}) = \sum_{i=1}^n c_i \frac{\partial a}{\partial c_i}(\mathbf{c}) - P^{EOS}(T, \mathbf{c}). \quad (1.3)$$

Naším cílem je rozhodnout, zda jsou všechny složky v jedné fázi (stabilní směs) nebo dojde k štěpení do více fází (nová vícefázová stabilní konfigurace). Tento problém se nazývá testování stability vícesložkové směsi v proměnných VTN. Zabýváme se situací, kde je termodynamika  $n$  složkové směsi popsána Pengovou-Robinsonovou stavovou rovnicí (1) a pomocí hustoty Helmholtzovy energie (1.1) a jejích partiálních derivací. Pro zkonstruování kritéria stability můžeme použít TPD funkci ve tvaru

$$TPD(\mathbf{c}, \mathbf{c}^*) = a(\mathbf{c}) - a(\mathbf{c}^*) - \sum_{i=1}^n \frac{\partial a}{\partial c_i}(\mathbf{c}^*)(c_i - c_i^*), \quad (1.4)$$

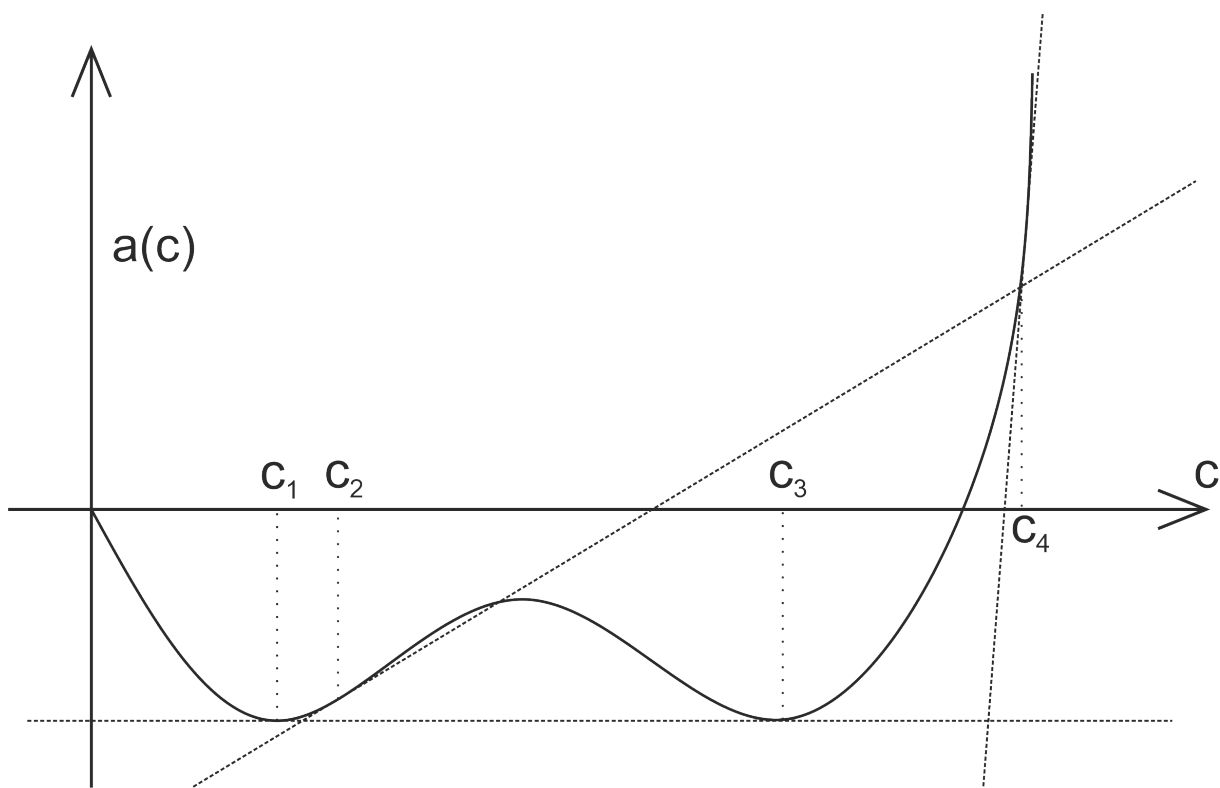
kde  $\mathbf{c}^*$  je vektor koncentrací testovaného stavu (stavu, u kterého rozhodujeme, zda je stabilní),  $\mathbf{c}$  je vektor koncentrací, funkce  $a(\mathbf{c})$  a funkce  $\frac{\partial a}{\partial c_i}(\mathbf{c})$  jsou definovány viz (1.1), (1.2). Kritérium pro fázovou stabilitu v proměnných VTN směsi bude formulováno pomocí následující věty.

**Věta 1** (VTN kritérium stability). *Mějme  $n$ -složkovou směs při fixní teplotě. Stav  $\mathbf{c}^*$  je stabilní právě tehdy, když pro každé  $\mathbf{c}$  z přípustné množiny platí, že  $TPD(\mathbf{c}, \mathbf{c}^*) \geq 0$ , kde  $TPD(\mathbf{c}, \mathbf{c}^*)$  je dána rovnicí (1.4).*

*Důkaz.* Viz [7]. □

Z této věty je zřejmé, že tento fyzikálně motivovaný problém, můžeme vyřešit pomocí nalezení globálního minima funkce, která se označuje jako TPD (tangent plane distance). Tuto funkci můžeme interpretovat jako rozdíl  $a(\mathbf{c})$  od hodnoty tečné nadroviny vedené k bodu  $\mathbf{c}^*$  v bodě  $\mathbf{c}$ . Na následujícím obrázku ukážeme 3 možnosti, které mohou nastat při testování fázové stability, těmi jsou stabilní stav, nestabilní stav a rovnovážný stav.

Obrázek 1.1: Ilustrace geometrického významu funkce TPD



Na obrázku 1.1 je vynesena hustota Helmholtzovy energie. Stavy  $c_1$  a  $c_3$  jsou v rovnováze a funkce TPD je v těchto bodech rovná nule  $TPD(c_1, c_3) = TPD(c_3, c_1) = 0$ . Stav  $c_2$  je nestabilní, protože například  $TPD(c_3, c_2) < 0$ . Naopak u stavu  $c_4$  vidíme, že pro jakékoliv  $c \neq c_4$  platí  $TPD(c, c_4) > 0$  a stav je tedy stabilní. Funkce  $a(c)$  je vynesena nepřerušovaně, čárkovaně jsou vyneseny tečny v daných bodech k funkci  $a(c)$  a tečkovaně jsou pro přehlednost zvýrazněny stavy na ose s hodnotou funkce ve stavu.

## 1.2 Matematická formulace

Jelikož potřebujeme k rozhodnutí o stabilitě systému podle Věty 1 ověřit nezápornost funkce TPD na celé přípustné množině, stačí nám nalézt globální minimum funkce TPD a pokud bude záporné, rozhodneme o nestabilitě systému (dojde ke štěpení do více fází). Naším matematickým problémem je tedy globální optimalizace nekonvexní funkce na konvexní množině

$$\begin{aligned} \min_{\mathbf{c}} \quad & TPD(\mathbf{c}, \mathbf{c}^*) \\ \text{za podmínek} \quad & c_i > 0; \quad i = 1, \dots, n \\ & \sum_{i=1}^n b_i c_i < 1. \end{aligned} \tag{1.5}$$

Přípustná množina určena podmínkami je vnitřek simplexu. Funkce TPD je nekonvexní, a tedy není zaručeno, že nalezením lokálního minima získáme dostatečnou informaci pro použití Věty 1. Pokud je lokální minimum záporné, k rozhodnutí o stabilitě to stačí, ale kladná lokální minima nám k rozhodnutí nestačí. Je nutné použít metody pro hledání globálního minima. To znamená, že naším úkolem je nalézt globální minimum funkce TPD na vnitřku simplexu s vrcholy  $\mathbf{c}^{(i)}$ . Za zmínku ještě stojí jedno zajímavé lokální minimum. Z tvaru TPD funkce je zřejmé, že platí  $TPD(\mathbf{c}^*, \mathbf{c}^*) = 0$ . Toto pozorování ukazuje, že globální minimum není nikdy kladné. Mohou nastat 2 možnosti, buď je globální minimum záporné nebo nulové. Pokud je globální minimum záporné, je systém nestabilní a dochází ke štěpení. Pokud je nulové a nabývá se ho pouze ve stavu  $\mathbf{c}^*$ , je systém stabilní a ke štěpení nedochází. Poslední možnost je, že globální minimum rovné nule se nabývá v bodě  $\mathbf{c} \neq \mathbf{c}^*$ . Pokud nastane tato možnost, jedná se o stav, který je energeticky ekvivalentní testovanému a je s ním v rovnováze.

## Kapitola 2

# Výpočetní algoritmus

V této kapitole představíme námi implementovaný algoritmus na nalezení globálního minima funkce TPD.

### 2.1 Algoritmus větví a mezí a jeho aplikace na naší situaci

Jedná se o algoritmus pro hledání globálního minima funkce na nějaké množině. Algoritmus v každém kroku rozdělí určitou část přípustné množiny na více menších a ponechá je na další prohledávání jen tehdy, pokud je možné, že v nich lze dosáhnout globálního optima. Vývoj tohoto postupu zaznamenáváme do stromu, kde je kořenem celá přípustná množina. Aktivní vrcholy stromu (nevyložené z algoritmu) reprezentují části přípustné množiny, na kterých algoritmus zatím nevyločil možnou globální optimalitu. O tom, zda si daný vrchol ponecháme, rozhodujeme na základě horního a dolního odhadu globálního optima na každé podmnožině. První horní odhad získáme lokální optimalizací funkce TPD na celé přípustné množině. Tento odhad poté aktualizujeme, pokud v průběhu algoritmu v jistém vrcholu obdržíme optimálnější řešení. Jako horní odhad je tedy použito zatím nejlepší nalezené lokální minimum funkce TPD. Pro dolní odhad je potřeba definovat podhodnocený problém, který budeme odvozovat pomocí následujících vět.

**Věta 2** (Konvexně-konkávní rozklad). *Necht'  $f$  je reálná funkce více proměnných, pro kterou platí  $f = gh$ . Necht' dále  $g$  a  $h$  jsou konvexní, nezáporné, reálné funkce více proměnných. Pak platí, že  $f = \frac{1}{2}(g+h)^2 - \frac{1}{2}(g^2 + h^2)$ . Navíc jsou funkce  $\frac{1}{2}(g+h)^2$  a  $\frac{1}{2}(g^2 + h^2)$  konvexní.*

*Důkaz.* Viz [2]. □

**Věta 3** (Jensenova nerovnost). *Necht'  $f$  je reálná, konvexní funkce na konvexní množině  $X$ , ( $n \in \mathbb{N}$ ) ( $\forall i \in \hat{n}$ ) ( $\mathbf{x}_i \in X$ ). Potom platí  $f\left(\sum_{i=1}^n \lambda_i \mathbf{x}_i\right) \leq \sum_{i=1}^n f(\lambda_i \mathbf{x}_i)$ , kde ( $\forall i \in \hat{n}$ ) ( $\lambda_i \in [0, 1]$ ) a  $\sum_{i=1}^n \lambda_i = 1$ . V případě, že je funkce  $f$  konkávní, platí opačná nerovnost.*

*Důkaz.* Viz [3]. □

Pomocí Věty 2 účelovou funkci rozdělíme na konvexní a konkávní část. Následnou aplikací Věty 3 na konkávní část dostaneme konvexní dolní odhad naší účelové funkce (tzv. konvexní relaxaci). Pokud nalezneme nějaké minimum této podhodnocené funkce, máme z následující věty zajištěno, že jde i o globální minimum.

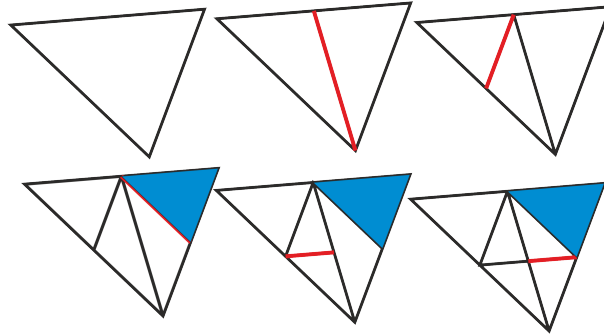
**Věta 4** (O lokálních minimech konvexní funkce). *Necht' je  $X \subset \mathbb{R}^n$  konvexní množina,  $f : X \rightarrow \mathbb{R}$  je konvexní funkce. Pak každé lokální minimum  $f$  na  $X$  je i globálním minimem  $f$  na  $X$ .*

Důkaz. Viz [1].

□

Jako poslední se musíme rozhodnout, jak budeme dělit naši přípustnou množinu. Naší přípustnou množinou je vnitřek simplexu, a tedy můžeme dělení provádět například rozdělením nejdelší hrany na půl. Tím získáme 2 nové simplexu. Na tomto obrázku ilustrujeme, jak algoritmus postupuje přípustnou

Obrázek 2.1: Ilustrace dělení simplexu ve 2D



množinou. Červeně je vyznačena hrana, která část simplexu rozdělila na dva. Modře je část simplexu, která byla vyloučena z algoritmu, protože algoritmus dokázal, že v této části nemůže být globální extrém.

Nejprve se budeme zabývat konstrukcí dolního odhadu hodnoty optimálního řešení na daném vnitřku simplexu s vrcholy  $\mathbf{c}^{(0)}, \dots, \mathbf{c}^{(N)}$  funkce TPD. Pro spodní odhad použijeme podhodnocenou TPD funkci, odvozenou stejným postupem jako v článku [8], ve tvaru:

$$TPD^{under} = a^{under}(\mathbf{c}) - a(\mathbf{c}^*) - \sum_{i=1}^N \frac{\partial a}{\partial c_i}(\mathbf{c}^*)(c_i - c_i^*). \quad (2.1)$$

K této funkci se dospěje následovně. O konvexitě funkce (1.4), jakožto funkce v proměnné  $\mathbf{c}$ , rozhoduje jen hustota Helmholtzovy energie (funkce (1.1)). Až na jeden člen je funkce (1.1) konvexní, důkaz tohoto tvrzení je v článku [11]. Můžeme tedy funkci (1.1) rozdělit na konvexní část

$$a_c(\mathbf{c}) = RT \sum_{i=1}^N c_i \ln \frac{c_i}{c_0} - RT \left( \sum_{i=1}^N c_i \right) \ln \left( 1 - \sum_{i=1}^N b_i c_i \right)$$

a nekonvexní část

$$a_{nc}(\mathbf{c}) = - \frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{2\sqrt{2} \sum_{i=1}^N b_i c_i} \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right),$$

kterou je potřeba ošetřit pomocí Věty 2 a Věty 3. Označme

$$f_1 = \frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{2\sqrt{2} \sum_{i=1}^N b_i c_i},$$

$$f_2 = \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right).$$

Předpokládáme, že kvadratická forma ve funkci  $f_1$  je pozitivně semi-definitní. Pokud by tomu tak nebylo, lze ji rozepsat jako rozdíl dvou pozitivně semidefinitních a postup analogicky použít na oba členy rozdílu. Funkce  $f_1$  a  $f_2$  jsou tedy nezáporné. Funkce  $f_1$  je konvexní. O této vlastnosti je možné rozhodnout z hessiánu. Druhá funkce je konkávní, a tedy ji upravíme, aby se hodila na použití Věty 2, které zajistíme vtáhnutím minusu a přičtením/odečtením kladné konstanty. Tímto se lišíme od článku [8], kde se funkce  $f_2$  označila za konvexní a to vedlo na jiný rozklad, který nesplňoval předpoklady Věty 2.

$$a_{nc} = -f_1 f_2 = f_1(C - f_2) - C f_1.$$

To zajistí, že funkce  $(C - f_2)$  je pro nějakou hodnotu  $C$  nezáporná. Nejmenší taková hodnota je

$$C = \ln\left(\frac{2 + \sqrt{2}}{2 - \sqrt{2}}\right),$$

což je nejvyšší možná hodnota  $f_2$  na naší přípustné množině. Funkce  $C f_1$  je konvexní, protože jsme ji přenásobili jen kladnou konstantou.

V našem případě je ještě rozklad z Věty 2 upraven o kladný škálující parametr  $\beta$ , který má fyzikální význam, protože funkce  $f_1$  a  $f_2$  mají různé jednotky. Potom za použití předchozí věty můžeme psát

$$a_{nc} = \frac{\beta}{2} \left( \frac{f_1}{\beta} + (C - f_2) \right)^2 - \frac{1}{2\beta} (f_1^2 + \beta^2 (C - f_2)^2) - C f_1,$$

kde poté využijeme Věty 3 na konkávní členy. Tento rozklad se liší od [8], kde byl člen  $f_2$  chybně označen za konvexní. což vedlo na rozdílný rozklad. Škálující parametr je zvolen následovně  $\beta = \frac{f_1(\mathbf{B})}{C - f_2(\mathbf{B})}$ , kde  $\mathbf{B}$  je těžiště simplexu. Hodnota parametru  $\beta$  byla zvolena tak, aby rozdíl mezi funkcemi  $a(\mathbf{c})$  a  $a^{\text{under}}$  byl co nejmenší. Toho je docíleno nalezením minima funkce

$$g_1 = -\frac{1}{2\beta} (f_1^2 + \beta^2 (C - f_2)^2) - C f_1$$

jakožto funkce  $\beta$ . To lze nalézt například derivováním podle  $\beta$ . Počítejme

$$\frac{\partial g_1}{\partial \beta} = -\frac{f_1^2 + \beta^2 (C - f_2)^2}{2\beta^2} + \frac{2\beta (C - f_2)^2}{2\beta} = -\frac{f_1^2}{2\beta^2} + \frac{(C - f_2)^2}{2}. \quad (2.2)$$

Tedy extrému se nabývá pro  $\beta = \frac{f_1}{C - f_2}$ . Druhá derivace

$$\frac{\partial^2 g_1}{\partial \beta^2} = \frac{f_1^2}{\beta^3}$$

je kladná a jedná se tedy o minimum. Označme část

$$g_1 = \frac{1}{2\beta} (f_1^2 + \beta^2 (C - f_2)^2) + C f_1$$

a konkávní část

$$g_2 = \frac{\beta}{2} \left( \frac{f_1}{\beta} + (C - f_2) \right)^2.$$

Platí tedy, že

$$a_{nc}(\mathbf{c}) = g_1(\mathbf{c}) - g_2(\mathbf{c}),$$

kde  $g_1(\mathbf{c})$  a  $g_2(\mathbf{c})$  jsou konvexní funkce. Dále odvodíme potřebný tvar  $a^{\text{under}}(\mathbf{c})$  pro konstrukci  $TPD^{\text{under}}$ .

$$a(\mathbf{c}) = a_c(\mathbf{c}) - g_1(\mathbf{c}) + g_2(\mathbf{c}) \geq a_c(\mathbf{c}) + g_2(\mathbf{c}) - \sum_{i=1}^{n+1} \alpha_i g_1(\mathbf{c}^{(i)}) = a^{\text{under}}(\mathbf{c}),$$

kde jsme v nerovnosti použili Větu 3 a  $\mathbf{c}^{(i)}$  jsou vrcholy simplexu. Protože  $a^{\text{under}}(\mathbf{c})$  a  $g_2(\mathbf{c})$  jsou konvexní funkce a  $\sum_{i=1}^{n+1} \alpha_i g_1(\mathbf{c}^{(i)})$  je afinní, je  $a^{\text{under}}$  konvexní funkce, která odhaduje  $a(\mathbf{c})$  zdola.

## 2.2 Konvexní relaxace

Jak už jsme naznačili dříve do algoritmu větví a mezí potřebujeme spodní odhad minima problému (1.5) na simplexu  $\mathbf{S}$ , který je konvexní (konvexní relaxace). Tento spodní odhad účelové funkce byl odvozen v předchozí kapitole a teď pouze shrneme, jak vypadá formulace tohoto problému.

$$\begin{aligned} \min_{\mathbf{c}} \quad & TPD^{\text{under}}(\mathbf{c}, \mathbf{c}^*) \\ \text{za podmíněk} \quad & \mathbf{c} \in \mathbf{S} \end{aligned} \tag{2.3}$$

Jak účelová funkce  $TPD^{\text{under}}$ , tak přípustná množina jsou konvexní a globální optimalizace lze díky Větě 4 řešit i lokálními metodami.

## 2.3 Barycentrické souřadnice

Při řešení úlohy se pohybujeme pouze na vnitřku simplexu, který je definovaný nerovnostmi  $c_i \geq 0$  a  $\sum_{i=1}^n b_i c_i < 1$ . Jakýkoliv bod uvnitř simplexu lze vyjádřit jako konvexní kombinaci vrcholů simplexu. Simplex má vždy  $n + 1$  vrcholů. V našem případě jsou počáteční vrcholy definované následovně  $\mathbf{c}^{(0)} = (0, \dots, 0)^T$  a  $\mathbf{c}^{(i)} = \frac{1}{b_i} \mathbf{e}_i$ , kde  $\mathbf{e}_i$  je  $i$ -tý vektor standardní báze  $\mathbb{R}^n$ . Maticově lze transformaci zapsat následujícím způsobem

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} c_1^{(0)} & c_1^{(1)} & \dots & c_1^{(n)} \\ c_2^{(0)} & c_2^{(1)} & \dots & c_2^{(n)} \\ \vdots & \vdots & & \vdots \\ c_n^{(0)} & c_n^{(1)} & \dots & c_n^{(n)} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{n+1} \end{pmatrix}.$$

Vektory  $\mathbf{c}$  a  $\boldsymbol{\alpha}$  mají sice jinou dimenzi, ale mají stejný počet stupňů volnosti. Pokud máme zadaných  $n$  složek vektoru  $\boldsymbol{\alpha}$ , můžeme dopočítat poslední složku jako doplněk do jedničky  $1 - \sum_{i=1; i \neq k}^n \alpha_i = \alpha_k$ . Práce v proměnných  $\boldsymbol{\alpha}$  na místo  $\mathbf{c}$  má pro naši práci několik příjemných vlastností. Například pohyb na simplexu je zaručen nezáporností  $\boldsymbol{\alpha}$  a jejich vysčítáním na jedničku.

## 2.4 Optimalizace v barycentrických souřadnicích

Protože je  $\mathbf{c}^*$  fixní vektor, nemusíme ho převádět do barycentrických souřadnic, a tedy stačí převést vektor  $\mathbf{c}$ . V proměnných  $\boldsymbol{\alpha}$  budeme využívat 2 optimalizačních úloh. Původní problém (1.5) bude



formulován následovně

$$\begin{aligned} \min_{\alpha} \quad & TPD \left( \sum_{i=1}^{n+1} \alpha_i \mathbf{c}^{(i)}, \mathbf{c}^* \right) \\ \text{za podmínek} \quad & \alpha_i > 0; i = 1, \dots, n \\ & \sum_{i=1}^{n+1} \alpha_i = 1, \end{aligned} \tag{2.4}$$

kde  $\mathbf{c}^{(i)}$  jsou vrcholy jistého podsimpllexu  $\mathbf{S}$ . A jeho konvexní relaxace (2.3) bude formulována následovně

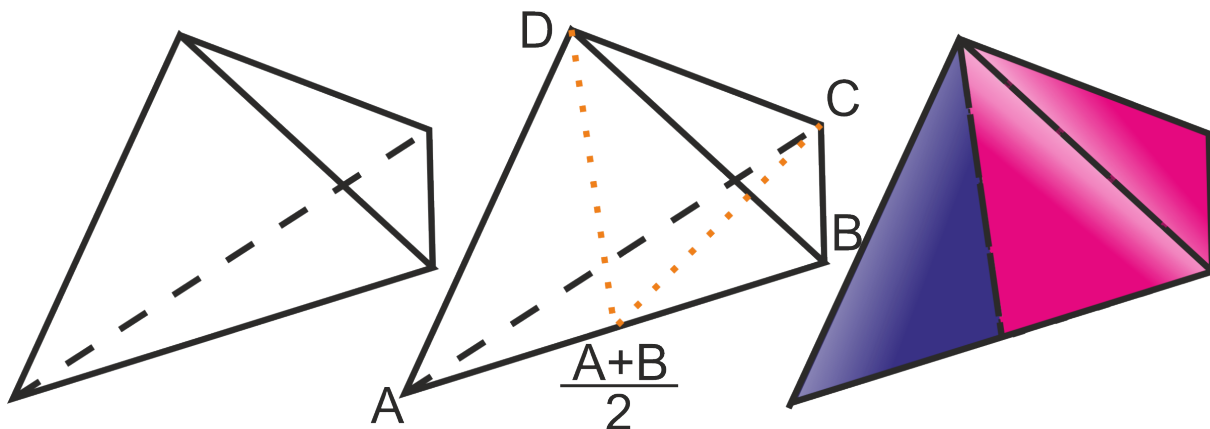
$$\begin{aligned} \min_{\alpha} \quad & TPD^{under} \left( \sum_{i=1}^{n+1} \alpha_i \mathbf{c}^{(i)}, \mathbf{c}^* \right) \\ \text{za podmínek} \quad & \alpha_i > 0; i = 1, \dots, n \\ & \sum_{i=1}^{n+1} \alpha_i = 1, \end{aligned} \tag{2.5}$$

kde  $\mathbf{c}^{(i)}$  jsou vrcholy jistého podsimpllexu  $\mathbf{S}$ . Tím jsme dokončili odvození problému, který chceme řešit a také jsme odvodili jeho konvexní dolní odhad a můžeme přejít k postupu na dělení simplexu.

## 2.5 Dělení simplexu

Pracujeme na vnitřku  $n$ -rozměrného simplexu. Zvolením některé z hran simplexu a nahrazením jednoho z vrcholů této hrany za jiný bod této hrany dostaneme znovu simplex. Takto můžeme konstruovat dělení simplexu v algoritmu. V každém kroku algoritmu takto rozdělíme simplex na dva menší a dále si ponecháme jen takové simplexu, v kterých ještě může účelová funkce původního problému nabývat globálního minima. Barycentrické souřadnice se nám starají o přípustnou množinu, protože je implicitně zadaná přes vrcholy simplexu. Pro ilustraci dělení uvedeme obrázek, jak k takovému dělení simplexu dochází ve 3D.

Obrázek 2.2: Dělení simplexu ve 3D



Jinak řečeno, mějme simplex  $\mathbf{S}$  s nejdelší hranou mezi vrcholy  $A$  a  $B$  z tohoto simplexu vytvoříme simplexu  $\mathbf{S}_1$  a  $\mathbf{S}_2$ .  $\mathbf{S}_1$  vznikne z  $\mathbf{S}$  nahrazením vrcholu  $A$  za vrchol  $\frac{A+B}{2}$  a analogicky vznikne  $\mathbf{S}_2$  nahrazením vrcholu  $B$  za vrchol  $\frac{A+B}{2}$ .

## 2.6 Implementace algoritmu větví a mezí

Nyní se konečně dostáváme k algoritmu větví a mezí [10]. Nejprve uvedeme základní algoritmus, který budeme dále v textu specifikovat/doplňovat. Algoritmus končí, pokud je horní odhad globálního optima účelové funkce **UB** dostatečně blízko spodnímu odhadu globálního optima účelové funkce **LB**. Množina aktivních simplexů  $S_a$  uchovává simplexu vznikající v algoritmu dělením počátečního a neobsahuje ty, v kterých už bylo dokázáno, že neobsahují globální optimum.

---

**Algoritmus 1:** Pseudoalgoritmus větví a mezí

---

**Result:** Vypiš **UB**

do  $S_a$  ulož celou přípustnou množinu (počáteční simplex);

**UB** =  $+\infty$ ;

**LB** =  $-\infty$ ;

**while** **UB** je blízko **LB** **do**

    vyber simplex  $S_a$  z množiny aktivních simplexů a odstraň ho z fronty;

    spočti původní problém (2.4) a optimum ulož do **UB1**;

**if** **UB1** < **UB** **then**

        | **UB1** =: **UB**;

**end**

    rozděl simplex **S** podle nějakého pravidla na **S1** a **S2** ;

    na simplexech **S1** a **S2** spočti podhodnocený problém (2.5) a optima ulož do **LB1** a **LB2** ;

**if** **LB1** < **UB** **then**

        | Přidej **S1** do  $S_a$ ;

**end**

**if** **LB2** < **UB** **then**

        | Přidej **S2** do  $S_a$ ;

**end**

    do **LB** ulož minimum z minim získaných spočtením podhodnoceného problému 2.5 ve vnitřku aktivních simplexů;

**end**

---

V algoritmu větví a mezí se standardně používá kritérium, zda simplex vyřadí z fronty, pokud je řešení konvexní relaxace větší nebo rovné našemu zatím nejmenšímu hornímu odhadu. V další kapitole odvodíme, jak strom můžeme prořezávat efektivněji.

## Kapitola 3

# Efektivnější ořezávání stromu v algoritmu větví a mezí

V této kapitole se budeme zabývat způsoby, jak efektivněji ořezávat strom vznikající v průběhu algoritmu, protože opakovaná optimalizace účelové funkce je výpočetně náročná a algoritmus je nedostatečně rychlý [9]. Pokud, ale využijeme nutných podmínek pro hledání optim, obdržíme další podmínky, které musí daný bod splňovat a to nám umožní odstraňovat neperspektivní simplex předtím, než bychom řešili samotnou konvexní relaxaci.

### 3.1 Nutná podmínka extrému

Ze základního kurzu matematické analýzy víme, že pokud má funkce více proměnných v daném bodě extrém, musí být všechny parciální derivace v daném bodě nulové. Připomeňme tvar funkce TPD

$$TPD(\mathbf{c}, \mathbf{c}^*) = \sum_{i=1}^n \frac{\partial a}{\partial c_i}(\mathbf{c}^*)(c_i^* - c_i) - (a(\mathbf{c}^*) - a(\mathbf{c})). \quad (3.1)$$

Počítejme parciální derivace funkce TPD podle proměnných  $c_i$  a položíme je rovné nule

$$\frac{\partial TPD(\mathbf{c}, \mathbf{c}^*)}{\partial c_i} = \frac{\partial a}{\partial c_i}(\mathbf{c}^*) - \frac{\partial a}{\partial c_i}(\mathbf{c}) \stackrel{!}{=} 0 \quad \forall i \in \{1, \dots, n\}. \quad (3.2)$$

Z rovnice (3.2) vidíme, že nutná podmínka extrému naší úlohy je, že chemické potenciály v testovaném stavu musí být rovné chemickému potenciálu v testujících stavech. Dále dosadíme vyjádření pro hustotu Helmholtzovy energie (1.3) do (1.4)

$$TPD(\mathbf{c}, \mathbf{c}^*) = \sum_{i=1}^n \frac{\partial a}{\partial c_i}(\mathbf{c}^*)(c_i^* - c_i) - \left( \sum_{i=1}^n c_i^* \frac{\partial a}{\partial c_i}(\mathbf{c}^*) - P^{EOS}(T, \mathbf{c}^*) - \sum_{i=1}^n c_i \frac{\partial a}{\partial c_i}(\mathbf{c}) + P^{EOS}(T, \mathbf{c}) \right). \quad (3.3)$$

Po dosažení nutných podmínek optimality naší úlohy z (3.2) do (3.3) získáme TPD funkci pro stacionární body  $\mathbf{c}$  ve tvaru

$$TPD(\mathbf{c}, \mathbf{c}^*) = P^{EOS}(T, \mathbf{c}^*) - P^{EOS}(T, \mathbf{c}). \quad (3.4)$$

Naším úkolem je nalézt globální minimum a jelikož ve stacionárních bodech hodnotu účelové funkce ovlivňuje pouze tlak v Pengově-Robinsonově rovnici, můžeme pomocí stavové rovnice stanovit další kritérium pro přípustnost stavu

$$P^{EOS}(T, \mathbf{c}^*) \leq P^{EOS}(T, \mathbf{c}). \quad (3.5)$$

Nejdůležitější výsledky kapitoly můžeme shrnout následujícím způsobem. Pokud má být pro nějaký stav TPD funkce záporná a zároveň být minimem, musí být potenciály jednotlivých složek rovny potenciálům jednotlivých složek v testovaném stavu a zároveň musí být tlak v tomto stavu větší nebo roven tlaku v testovaném stavu. Těchto poznatků budeme moci využít v algoritmu větví a mezí.

## 3.2 Ořezávání pomocí Pengovy-Robinsonovy stavové rovnice

Na základě výsledků předchozí kapitoly jsme dospěli k závěru, že můžeme využít stavovou rovnici pro rychlejší detekci neperspektivních částí simplexu. Přesněji danou část simplexu nemusíme dále prohledávat, pokud je v ní tlak menší než je v testovaném stavu  $\mathbf{c}^*$ . Jinými slovy, daný simplex  $\mathbf{S}$  vyloučíme z hledání, pokud platí

$$P^{EOS}(T, \mathbf{c}^*) \geq \max_{\mathbf{c} \in \mathbf{S}} P^{EOS}(T, \mathbf{c}). \quad (3.6)$$

Stavová rovnice není konkávní, a tedy by její globální maximalizace byla výpočetně náročná. V této kapitole budeme odvozovat horní odhad stavové rovnice, který budeme poté moci optimalizovat s menší výpočetní náročností. Jako první využijeme Věty 2, pomocí které odvodíme konkávní horní odhad, který lze optimalizovat jednou lokální optimalizací. Dále odvodíme tvar tečné nadroviny tohoto konkávního odhadu, z kterého budeme moci získat sice hrubší, ale výpočetně méně náročný, odhad chtěného maxima. Odhad maxima z nadroviny získáme porovnáním hodnot nadroviny v krajních bodech simplexu. Připomene tvar stavové rovnice

$$P^{EOS}(T, \mathbf{c}) = \frac{\left(\sum_{i=1}^N c_i\right) RT}{1 - \sum_{i=1}^N b_i c_i} - \frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left(\sum_{i=1}^N b_i c_i\right)^2}. \quad (3.7)$$

Jelikož se jedná o funkci více proměnných, budeme se snažit konvexitu ověřovat za použití následující věty.

**Věta 5** (O konvexitě funkce afinní transformace). *Necht' je  $\mathcal{X} \subset \mathbb{R}^n$  konvexní množina,  $h : \mathcal{X} \rightarrow \mathbb{R}$  je afinní funkce a  $g : \mathbb{R} \rightarrow \mathbb{R}$  je konvexní funkce. Pak  $f : \mathcal{X} \rightarrow \mathbb{R} : \mathbf{x} \mapsto g(h(\mathbf{x}))$  je konvexní funkce.*

*Důkaz.* Viz [2]. □

Nyní přejdeme k hornímu odhadu funkce (1). Jako první si uvědomíme na jakých výrazech funkce závisí. To nám ulehčí rozhodování o konvexitě a nezápornosti členů pro použití Věty 2, kde poté dostaneme aplikací Věty 3 horní odhad funkce (1).

### 3.2.1 Afinní transformace ve stavové rovnici

Rovnice (1) závisí na vektoru koncentrací  $\mathbf{c}$  jen přes 3 výrazy. Pokud zavedeme následující substituce

$$x = \sum_{i=1}^N b_i c_i, \quad (3.8)$$

$$y = \sum_{i=1}^N c_i, \quad (3.9)$$

můžeme pro členy funkce obsahující jen jeden z těchto výrazů použít Větu 5 a místo konkávnosti funkce více proměnných ověřovat jen konkávnost funkce jedné proměnné. Poslední výraz je o něco složitější

$\sum_{i,j=1}^N a_{ij}c_i c_j$ . Jde o kvadratickou indefinitní formu danou symetrickou maticí  $\mathbb{A}$ . Z lineární algebry víme, že pro symetrickou matici vždy existuje diagonální tvar a její vlastní čísla jsou reálná a toho využijeme. Pokud je nějaké vlastní číslo nulové zahrneme ho do části s kladnými vlastními čísly  $\lambda$ .

$$\mathbb{A} = \mathbb{Q} \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & \iota_1 & & \\ & & & & \ddots & \\ & & & & & \iota_k \end{pmatrix} \mathbb{Q}^T = \mathbb{Q} \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \mathbb{Q}^T - \mathbb{Q} \begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & -\iota_1 & & \\ & & & & \ddots & \\ & & & & & -\iota_k \end{pmatrix} \mathbb{Q}^T. \quad (3.10)$$

Tedy indefinitní matici převedeme na rozdíl dvou pozitivně semidefinitních matic

$$\mathbb{A} = \mathbb{A}_c - \mathbb{A}_n. \quad (3.11)$$

Pokud je matice pozitivně definitní, což je běžné například u směsí uhlovodíků, tento rozpis nemusíme provádět (matice  $\mathbb{A}_n$  bude nulová). Tento rozpis by bylo možné použít i v případě podhodnocené TPD funkce (2.1). Těchto 2 substitucí a jednoho rozepsání budeme využívat v následující kapitole, kde budeme ověřovat předpoklady Věty 2 samostatně člen po členu.

### 3.2.2 Ověřování předpokladů věty o konvexně-konkávním rozkladu a její použití

Nejprve ověříme předpoklady u prvního členu stavové rovnice

$$\frac{\left(\sum_{i=1}^N c_i\right)RT}{1 - \sum_{i=1}^N b_i c_i}.$$

Zabývejme se čitatelem prvního členu. Snadno ověříme, že výraz  $\left(\sum_{i=1}^N c_i\right)RT$  je nezáporný ( $T$  je absolutní teplota (kladná konstanta) a koncentrace  $c$  jsou také nezáporné). Jde o lineární funkci v proměnných  $c_i$ , a tedy je konvexní. Alternativně můžeme použít substituci (3.9) společně s Větou 5. Tím převedeme problém na vyšetřování konvexity funkce  $y$  a dojdeme ke stejnému výsledku.

Dále se budeme zabývat druhou částí prvního členu, tím je výraz  $\frac{1}{1 - \sum_{i=1}^N b_i c_i}$ . Protože výraz diskutujeme jen

na přípustné množině definované v 6, vidíme, že je na dané množině jistě nezáporný. Použitím substituce (3.8) společně s Větou 5 se následná analýza zjednoduší a budeme ověřovat jen konvexitu funkce jedné proměnné  $f(x) = \frac{1}{1-x}$  pro  $x \in (0, 1)$ . Po vyjádření druhé derivace této funkce  $f^{(2)}(x) = \frac{2}{(1-x)^3}$  vidíme, že pro  $x \in (0, 1)$  je funkce jistě konvexní. Ověřili jsme předpoklady Věty 2 a můžeme ji použít na první člen stavové rovnice

$$\left(\sum_{i=1}^N c_i\right)RT \frac{1}{1 - \sum_{i=1}^N b_i c_i} = f_c^1 + f_{nc}^1, \quad (3.12)$$

kde jsou členy  $f_c^1$  a  $f_{nc}^1$  definované následovně

$$f_c^1 = \frac{\gamma_1}{2} \left( \frac{\left( \sum_{i=1}^N c_i \right) RT}{\gamma_1} + \frac{1}{1 - \sum_{i=1}^N b_i c_i} \right)^2,$$

$$f_{nc}^1 = -\frac{1}{2\gamma_1} \left( \left( \sum_{i=1}^N c_i \right) RT \right)^2 + \left( \frac{\gamma_1}{1 - \sum_{i=1}^N b_i c_i} \right)^2.$$

Člen  $f_c^1$  je konvexní a člen  $f_{nc}^1$  konkávní. Protože členy  $\left( \sum_{i=1}^N c_i \right) RT$  a  $\frac{1}{1 - \sum_{i=1}^N b_i c_i}$  mají různé jednotky, zavedli jsme nezáporný koeficient  $\gamma_1$ , který bude sloužit ke korekci jednotek. Jeho hodnotu budeme diskutovat na konci kapitoly. Nyní budeme podobně upravovat druhý člen v rovnici (1). Tím je

$$-\frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2}.$$

Výraz  $\sum_{i,j=1}^N a_{ij} c_i c_j$  je kvadratická forma. Tento výraz upravíme pro použití Věty 2 pomocí rozpisu (3.11) na rozdíl dvou pozitivně semidefinitních kvadratických forem.

Přesuneme se k výrazu  $\frac{1}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2}$ , znovu se díky (3.8) a Větě 5 můžeme zabývat jen konvexitou

funkce  $g(x) = \frac{1}{1 + 2x - x^2}$  pro  $x \in (0, 1)$ . Po napočítání druhé derivace  $g^{(2)}(x) = -\frac{2(3x^2 - 6x + 5)}{(x^2 - 2x - 1)^3}$  opět vidíme, že je druhá derivace nezáporná, a tedy je funkce  $g(x)$  konvexní pro  $x \in (0, 1)$ . Nyní můžeme přejít k aplikaci Věty 2 na druhý člen stavové rovnice

$$-\frac{\sum_{i,j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} = f_c^{2A_p} + f_{nc}^{2A_p} + f_c^{2A_n} + f_{nc}^{2A_n}, \quad (3.13)$$

kde jsou funkce  $f_c^{2A_p}$ ,  $f_{nc}^{2A_p}$ ,  $f_c^{2A_n}$  a  $f_{nc}^{2A_n}$  definované následovně

$$f_c^{2A_p} = \frac{1}{2\gamma_2} \left( \left( \mathbf{c}^T \mathbb{A}_p \mathbf{c} \right)^2 + \left( \frac{\gamma_2}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right),$$

$$f_{nc}^{2A_p} = -\frac{\gamma_2}{2} \left( \frac{\mathbf{c}^T \mathbb{A}_p \mathbf{c}}{\gamma_2} + \frac{1}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2,$$

$$f_c^{2A_n} = \frac{\gamma_3}{2} \left( \frac{\mathbf{c}^T A_n \mathbf{c}}{\gamma_3} + \frac{1}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2,$$

$$f_{nc}^{2A_n} = -\frac{1}{2\gamma_3} \left( (\mathbf{c}^T A_n \mathbf{c})^2 + \left( \frac{\gamma_3}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right).$$

Členy  $f_c^{2A_p}$  a  $f_c^{2A_n}$  jsou konvexní. Členy  $f_{nc}^{2A_p}$  a  $f_{nc}^{2A_n}$  jsou konkávní. Znovu mají členy  $\sum_{i,j=1}^N a_{ij} c_i c_j$  a  $\frac{1}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2}$  různé jednotky, a tedy bylo potřeba zavést nezáporné koeficienty  $\gamma_2$  a  $\gamma_3$ , které budou sloužit ke korekci jednotek. Jejich hodnoty budeme diskutovat na konci kapitoly.

### 3.2.3 Konvexně-konkávní odhad Pengovy-Robinsonovy stavové rovnice

Nyní už můžeme napsat konvexně konkávní rozklad funkce (1), na který budeme aplikovat Větu 3 a tím získáme horní odhad funkce (1). Nejprve napíšeme konvexní část

$$P_{convex}^{EOS}(T, \mathbf{c}) = f_c^1 + f_c^{2A_p} + f_c^{2A_n}. \quad (3.14)$$

Podobně napíšeme konkávní část

$$P_{concave}^{EOS}(T, \mathbf{c}) = f_{nc}^1 + f_{nc}^{2A_p} + f_{nc}^{2A_n}. \quad (3.15)$$

Jistě platí

$$P^{EOS}(T, \mathbf{c}) = P_{convex}^{EOS}(T, \mathbf{c}) + P_{concave}^{EOS}(T, \mathbf{c}).$$

V posledním kroku využijeme Jensenovy nerovnosti (Věta 3)

$$P^{EOS}(T, \mathbf{c}) = P_{convex}^{EOS}(T, \mathbf{c}) + P_{concave}^{EOS}(T, \mathbf{c}) \leq \sum_{i=1}^N \alpha_i P_{convex}^{EOS}(T, \mathbf{c}^{(i)}) + P_{concave}^{EOS}(T, \mathbf{c}).$$

Z tohoto tvaru je vidět, že bude výhodné používat barycentrické souřadnice. Pohybujeme se na simplexu, a tedy lze každý bod vyjádřit pomocí konvexní kombinace vrcholů ( $\mathbf{c}^{(i)}$ ) simplexu. Tímto jsme dokončili odvození horního odhadu stavové rovnice, který je konkávní

$$P_{upper}^{EOS}(T, \mathbf{c}) = \sum_{i=1}^N \alpha_i P_{convex}^{EOS}(T, \mathbf{c}^{(i)}) + P_{concave}^{EOS}(T, \mathbf{c}). \quad (3.16)$$

Ještě je potřeba uvést tvar tečné nadroviny ke středu zkoumaného simplexu  $\mathbf{c}_0$

$$P_{plane}^{EOS}(T, \mathbf{c}) = P_{upper}^{EOS}(T, \mathbf{c}_0) + \nabla P_{upper}^{EOS}(T, \mathbf{c}_0)(\mathbf{c} - \mathbf{c}_0). \quad (3.17)$$

### 3.2.4 Ladění parametrů $\gamma_1, \gamma_2, \gamma_3$

Obecně používáme vždy rozklad z Věty 2 typu

$$gh = \frac{\gamma}{2} \left( \frac{g}{\gamma} + h \right)^2 - \frac{1}{2\gamma} (g^2 + \gamma^2 h^2).$$

Situaci, kde je potřeba odhadnout druhý člen, jsme řešili v konstrukci podhodnocené účelové funkce v rovnici (2.2). Výhodná volba parametru v tomto případě byla  $\gamma = \frac{g(\mathbf{B})}{h(\mathbf{B})}$ , kde  $\mathbf{B}$  je těžiště simplexu. V případě, že bychom museli odhadnout první člen, bude výpočet totožný jako při ladění parametru v (2.2).

Počítejme tedy

$$\frac{\partial \frac{\gamma}{2} \left( \frac{g}{\gamma} + h \right)^2}{\partial \gamma} = \frac{g^2}{2\gamma^2} + \frac{2gh}{2\gamma} + \frac{h^2}{2} - \frac{2\gamma g}{2\gamma^2} \left( \frac{g}{\gamma} + h \right) = \frac{1}{2} \left( h^2 - \frac{g^2}{\gamma} \right).$$

Toto vede na totožnou volbu  $\gamma = \frac{g(\mathbf{B})}{h(\mathbf{B})}$ . Aplikací tohoto závěru na každou z odhadnutých funkcí zvlášť poté získáme volby

$$\gamma_1 = \frac{\left( \sum_{i=1}^N \mathbf{B}_i \right) RT}{1 - \sum_{i=1}^N b_i \mathbf{B}_i},$$

$$\gamma_2 = \frac{\mathbf{B}^T \mathbf{A}_p \mathbf{B}}{1 + 2 \sum_{i=1}^N b_i \mathbf{B}_i - \left( \sum_{i=1}^N b_i \mathbf{B}_i \right)^2},$$

$$\gamma_3 = \frac{\mathbf{B}^T \mathbf{A}_n \mathbf{B}}{1 + 2 \sum_{i=1}^N b_i \mathbf{B}_i - \left( \sum_{i=1}^N b_i \mathbf{B}_i \right)^2},$$

kde  $\mathbf{B}$  je těžiště simplexu  $\mathbf{S}$ .

### 3.2.5 Kritéria pro ořezávání simplexu na základě stavové rovnice

Jak jsme naznačili v úvodu této kapitoly, první možné kritérium lze volit následovně. Pokud pro simplex  $\mathbf{S}$  platí následující nerovnost

$$P^{EOS}(T, \mathbf{c}^*) \geq \max_{\mathbf{c} \in \mathbf{S}} P_{upper}^{EOS}(T, \mathbf{c}), \quad (3.18)$$

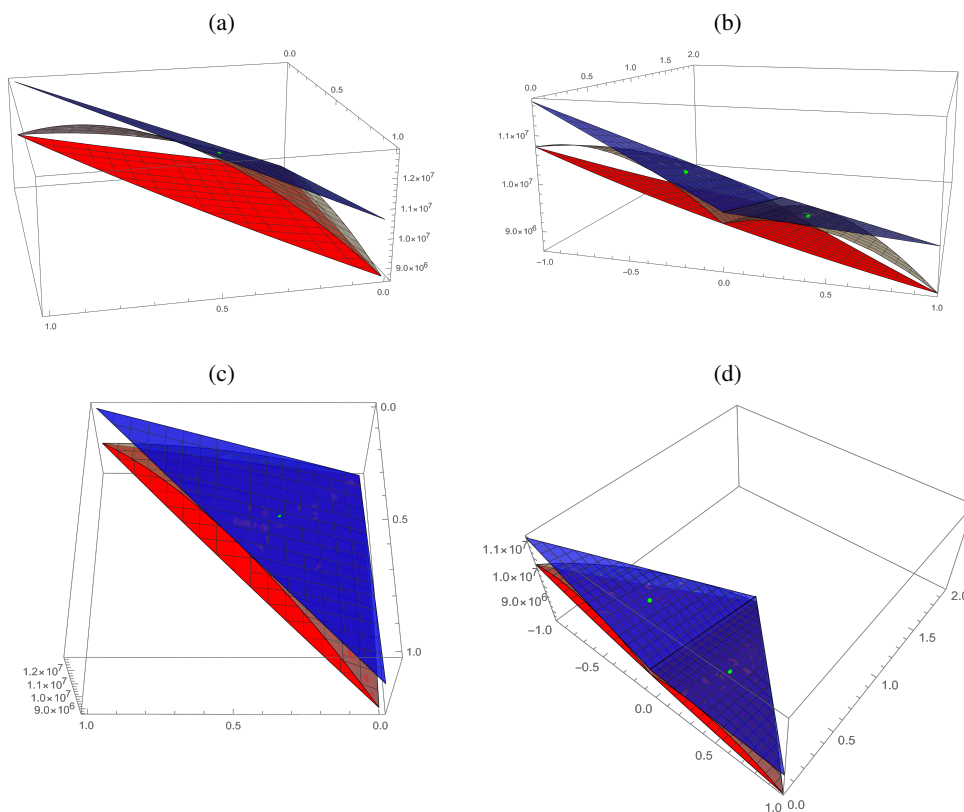
vyloučíme ho z přípustné množiny. Další možné kritérium formulujeme obdobně pomocí, tečné nadroviny. Pokud množinu vrcholů simplexu  $\mathbf{S}$  označíme jako  $\mathbf{S}_v$  a platí následující nerovnost

$$P^{EOS}(T, \mathbf{c}^*) \geq \max_{\mathbf{c} \in \mathbf{S}_v} P_{plane}^{EOS}(T, \mathbf{c}), \quad (3.19)$$

tak ho vyloučíme z přípustné množiny. Vyčíslení  $\max_{\mathbf{c} \in \mathbf{S}} P_{upper}^{EOS}(T, \mathbf{c})$  je získáno lokální optimalizací, zatímco  $\max_{\mathbf{c} \in \mathbf{S}_v} P_{plane}^{EOS}(T, \mathbf{c})$  je získáno jen vyčíslením konečného počtu hodnot funkce  $P_{plane}^{EOS}$  ve vrcholech simplexu  $\mathbf{S}$ .



Obrázek 3.1: Vizualizace Pengovy stavové rovnice



Na grafech a) a c) vidíme tentýž graf ze dvou různých úhlů na jednom simplexu, na grafech b) a d) vidíme grafy na stejném simplexu po jedné iteraci algoritmu vůči a) a c), červeně průběh tlaku z Pengovy-Robinsonovy stavové rovnice  $P^{EOS}(T, \mathbf{c})$  (1), hnědě horní odhad tlaku pomocí konvexně-konkávního rozkladu  $P_{upper}^{EOS}(T, \mathbf{c})$  (3.16), modře tečná nadrovina k hornímu odhadu tlaku v bodě vyznačeném zeleně (střed simplexu, kde se daný odhad konstruuje)  $P_{plane}^{EOS}(T, \mathbf{c})$  (3.17).

### 3.3 Ořezávání pomocí chemických potenciálů

V kapitole 3.1, kde jsme se zabývali tvary nutných podmínek extrému, jsme jako další možnost na vylepšení algoritmu větví a mezí uvedli použití chemických potenciálů, kterými se budeme zabývat v této kapitole. Z rovnice (3.2) máme nutné podmínky pro globální minima. Přesněji máme  $n$  podmínek, které je možné zapsat vektorově

$$\nabla a(\mathbf{c}^*) = \nabla a(\mathbf{c}). \quad (3.20)$$

Jinými slovy chemické potenciály v testovaném stavu musí být rovny potenciálům v testujícím stavu. Chemické potenciály opět nejsou konvexní nebo konkávni funkcí, a tedy jsme nuceni odvodit odhady, které už konvexní/konkávni budou. Tedy podmínka (3.20) bude zeslabena na to, že každý chemický potenciál v testovaném stavu bude ležet v intervalu vymezeném horním a dolním odhadem. Pokud pro simplex  $\mathbf{S}$  neplatí následující nerovnost

$$\min_{\mathbf{c} \in \mathbf{S}} \mu_{under}(\mathbf{c}) \leq \mu(\mathbf{c}^*) \leq \max_{\mathbf{c} \in \mathbf{S}} \mu_{upper}(\mathbf{c}), \quad (3.21)$$

vyloučíme ho z algoritmu. Připomene tvar chemických potenciálů

$$\begin{aligned} \frac{\partial a}{\partial c_i}(\mathbf{c}) = \mu_i(\mathbf{c}) = & RT \left( \ln \frac{c_i}{c_0} + 1 \right) - RT \ln \left( 1 - \sum_{i=1}^N b_i c_i \right) + RT b_i \frac{\sum_{i=1}^N c_i}{1 - \sum_{i=1}^N b_i c_i} \\ & - \frac{\left( 2 \sum_{j=1}^N c_j a_{ij} \right) \left( \sum_{i=1}^N b_i c_i \right) - b_i \sum_{i,j=1}^N a_{ij} c_i c_j}{2 \sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right) \\ & - \left( \frac{b_i \sum_{i,j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left( \sum_{i=1}^N b_i c_i \right)^2} \right). \end{aligned} \quad (3.22)$$

Jelikož je chemický potenciál definován pro každou složku, máme celkem  $2n$  nerovností, kde  $n$  je počet látek ve směsi. Analogicky jako jsme prováděli konstrukci horního odhadu funkce (1) nyní provedeme pro chemické potenciály. První dva kroky jsou totožné. Nejdříve si uvědomíme na jakých výrazech závisí chemické potenciály. V druhém ověříme konvexitu pro použití Věty 2. Protože potřebujeme horní i dolní odhad chemických potenciálů (1.2), budeme Větu 3 aplikovat dvakrát na každou složku.

### 3.3.1 Afinní transformace v chemických potenciálech

Podobně jako u stavové rovnice, zaleží potenciály na vektoru  $\mathbf{c}$  jen přes několik výrazů. Opět budeme využívat několika substitucí, které nám společně s Větou 5 ulehčí ověřování předpokladů. Navíc od stavové rovnice přibude jen jedna, ale pro přehlednost je uvedeme znovu. Tyto substituce jsou následující:

$$x = \sum_{i=1}^N b_i c_i, \quad (3.23)$$

$$y = \sum_{i=1}^N c_i, \quad (3.24)$$

$$z_i = \sum_{j=1}^N c_j a_{ij}. \quad (3.25)$$

Dále znovu využijeme rozklad indefinitní matice na rozdíl pozitivně semidefinitních (3.11)

$$\mathbb{A} = \mathbb{A}_c - \mathbb{A}_n. \quad (3.26)$$

### 3.3.2 Ověřování předpokladů věty o konvexně-konkávním rozkladu a její použití

První člen v rovnici chemických potenciálů (1.2) je funkce pouze jedné proměnné

$$f_{i,nc}^1 = RT \left( \ln \frac{c_i}{c_0} \right), \quad (3.27)$$

je jistě konkávní jakožto logaritmus, a tedy ho můžeme zahrnout do konkávní části rozkladu. Druhý člen v rovnici chemických potenciálů (1.2)

$$f_{i,c}^2 = -RT \ln \left( 1 - \sum_{i=1}^N b_i c_i \right) \quad (3.28)$$

je opět zřejmý, jelikož po použití substituce (3.23) a Věty 5 jde o záporně vzatý logaritmus, tedy jde o konvexní funkci a zahrneme ho do konvexní části rozkladu. U prvních dvou členů jsme se tedy obešli bez Věty 2. Třetí člen rovnice chemických potenciálů (1.2) už není konkávní/konvexní, a tedy budeme muset provést rozklad. Splnění předpokladů Věty 2 je lehce nahlédnutelné. Člen  $\sum_{i=1}^N c_i$  je nezáporný a lineární, tedy i konvexní a díky (3.23) a Větě 5 se místo konvexity  $\frac{1}{1 - \sum_{i=1}^N b_i c_i}$  můžeme zabývat jen funkcí

$\frac{1}{1-x}$ . Funkce je na daném intervalu zřejmě kladná. Konvexitu ověříme spočtením druhé derivace, která je rovna  $\frac{2}{(1-x)^3}$ . Tato derivace je na daném intervalu kladná. Ověřili jsme předpoklady Věty 2 a použijeme ji

$$RT b_i \frac{\sum_{i=1}^N c_i}{1 - \sum_{i=1}^N b_i c_i} = f_{i,c}^3 + f_{i,nc}^3, \quad (3.29)$$

kde je

$$f_{i,c}^3 = \frac{RT b_i \gamma_1}{2} \left( \frac{\sum_{i=1}^N c_i}{\gamma_1} + \frac{1}{1 - \sum_{i=1}^N b_i c_i} \right)^2,$$

$$f_{i,nc}^3 = -\frac{RT b_i}{2\gamma_1} \left( \left( \sum_{i=1}^N c_i \right)^2 + \left( \frac{\gamma_1}{1 - \sum_{i=1}^N b_i c_i} \right)^2 \right),$$

první člen je konvexní a druhý konkávní a zahrneme je do příslušných částí rozkladu. U dalšího členu chemických potenciálů (1.2) nejdříve upravíme první člen rozdílu a využijeme vlastnosti logaritmu podílu

$$-\frac{\left( \sum_{j=1}^N c_j a_{ij} \right)}{\sqrt{2} \sum_{i=1}^N b_i c_i} \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right) =$$

$$\frac{\left( \sum_{j=1}^N c_j a_{ij} \right)}{\sqrt{2} \sum_{i=1}^N b_i c_i} \ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right) - \frac{\left( \sum_{j=1}^N c_j a_{ij} \right)}{\sqrt{2} \sum_{i=1}^N b_i c_i} \ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right). \quad (3.30)$$

Zabývejme se nejdříve prvním členem na pravé straně (3.30). Funkce  $\sum_{j=1}^N c_j a_{ij}$  je afinní nezáporná funkce,

a tedy i konvexní. Dále se budeme zabývat členem  $\frac{\ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\sum_{i=1}^N b_i c_i}$ , který je záporný, a tedy musíme vy-

tknout mínus před výraz. Opět využijeme (3.23) a Větu 5 a budeme se zabývat jen konvexitou funkce jedné proměnné na intervalu  $(0, 1)$ . O konvexitě funkce  $-\frac{\ln(1+(1-\sqrt{2})x)}{x}$  rozhodneme s pomocí druhé derivace, která je kladná a funkce je tedy konvexní. Ověřili jsme předpoklady a můžeme přejít k aplikaci Věty 2

$$\frac{\left(\sum_{j=1}^N c_j a_{ij}\right)}{\sqrt{2} \sum_{i=1}^N b_i c_i} \ln\left(1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i\right) = f_{i,c}^4 + f_{i,nc}^4, \quad (3.31)$$

kde

$$f_{i,c}^4 = \frac{1}{2\sqrt{2}\gamma_2} \left( \left( \sum_{j=1}^N c_j a_{ij} \right)^2 + \left( \frac{\gamma_2 \ln\left(1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i\right)}{\sum_{i=1}^N b_i c_i} \right)^2 \right),$$

$$f_{i,nc}^4 = -\frac{\gamma_2}{2\sqrt{2}} \left( \frac{\sum_{j=1}^N c_j a_{ij}}{\gamma_2} - \frac{\ln\left(1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i\right)}{\sum_{i=1}^N b_i c_i} \right)^2.$$

Zde je první člen konvexní a druhý konkávní a zahrneme je do příslušných částí rozkladu. Druhý člen z pravé strany rovnice (3.30) dopadne velice podobně. Funkci  $\sum_{j=1}^N c_j a_{ij}$  jsme již komentovali a funkce  $\frac{\ln(1+(1+\sqrt{2})\sum_{i=1}^N b_i c_i)}{\sum_{i=1}^N b_i c_i}$  je kladná, a tedy opět diskutujeme její konvexitu. Po použití substituce (3.23) se zabý-

váme konvexitou funkce  $\frac{\ln(1+(1+\sqrt{2})x)}{x}$  na intervalu  $(0, 1)$ . Druhá derivace je kladná, a tedy je člen konvexní a opět můžeme aplikovat Větu 2

$$-\frac{\left(\sum_{j=1}^N c_j a_{ij}\right)}{\sqrt{2} \sum_{i=1}^N b_i c_i} \ln\left(1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i\right) = f_{i,c}^5 + f_{i,nc}^5, \quad (3.32)$$

kde

$$f_{i,c}^5 = -\frac{1}{2\sqrt{2}\gamma_3} \left( \left( \sum_{j=1}^N c_j a_{ij} \right)^2 + \left( \frac{\gamma_3 \ln\left(1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i\right)}{\sum_{i=1}^N b_i c_i} \right)^2 \right),$$

$$f_{i,nc}^5 = \frac{\gamma_3}{2\sqrt{2}} \left( \frac{\sum_{i=1}^N c_j a_{ij}}{\gamma_3} - \frac{\ln\left(1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i\right)}{\sum_{i=1}^N b_i c_i} \right)^2.$$

Zde je první člen konvexní a druhý konkávní a zahrneme je do příslušných částí rozkladu. Další člen chemických potenciálů (1.2) opět rozepíšeme jako rozdíl logaritmů

$$\begin{aligned} & \frac{b_i \sum_{i,j=1}^N a_{ij}c_i c_j}{2\sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( \frac{1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i}{1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i} \right) = \\ & \frac{b_i \sum_{i,j=1}^N a_{ij}c_i c_j}{2\sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right) - \frac{b_i \sum_{i,j=1}^N a_{ij}c_i c_j}{2\sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right). \end{aligned} \quad (3.33)$$

Opět se zabýváme prvním členem na pravé straně (3.33). Člen obsahuje totožnou kvadratickou formu jako v rovnici (3.13), a tedy ho ošetříme stejně, pomocí rozpisu na rozdíl kvadratických forem. Zbylá část členu se po provedení substituce (3.23) dostane do tvaru  $\frac{\ln(1+(1+\sqrt{2})x)}{x^2}$ . Funkce je opět kladná a její druhá derivace je opět kladná, a tedy se jedná o konvexní funkci a přejdeme k aplikaci Věty 2

$$\frac{b_i \sum_{i,j=1}^N a_{ij}c_i c_j}{2\sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right) = f_{i,c}^{6A_p} + f_{i,nc}^{6A_p} + f_{i,c}^{6A_n} + f_{i,nc}^{6A_n}, \quad (3.34)$$

kde jsou funkce  $f_{i,c}^{6A_p}$ ,  $f_{i,nc}^{6A_p}$ ,  $f_{i,c}^{6A_n}$  a  $f_{i,nc}^{6A_n}$  definované následovně

$$\begin{aligned} f_{i,c}^{6A_p} &= \frac{b_i \gamma_4}{4\sqrt{2}} \left( \frac{\mathbf{c}^T \mathbb{A}_p \mathbf{c}}{\gamma_4} + \left( \frac{\ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right), \\ f_{i,nc}^{6A_p} &= -\frac{b_i}{4\sqrt{2}\gamma_4} \left( (\mathbf{c}^T \mathbb{A}_p \mathbf{c})^2 + \left( \frac{\ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right), \\ f_{i,nc}^{6A_n} &= -\frac{b_i \gamma_5}{4\sqrt{2}} \left( \frac{\mathbf{c}^T \mathbb{A}_n \mathbf{c}}{\gamma_5} + \left( \frac{\ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right), \\ f_{i,c}^{6A_n} &= \frac{b_i}{4\sqrt{2}\gamma_5} \left( (\mathbf{c}^T \mathbb{A}_n \mathbf{c})^2 + \left( \frac{\ln \left( 1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right), \end{aligned}$$

kde  $f_{i,c}^{6A_p}$ ,  $f_{i,c}^{6A_n}$  jsou konvexní a  $f_{i,nc}^{6A_p}$ ,  $f_{i,nc}^{6A_n}$  jsou konkávní a zahrneme je do příslušných částí rozkladu.

Dále analogicky upravíme druhý výraz na pravé straně rovnice (3.33). První člen je opět kvadratická indefinitní forma, kterou ošetříme totožně jako v rovnici (3.13). Zbylá část členu se po provedení substituce (3.23) dostane do tvaru  $-\frac{\ln(1+(1-\sqrt{2})x)}{x^2}$ , což je kladná funkce na intervalu (0, 1). Druhá derivace je opět kladná, a tedy se jedná o konvexní funkci. Aplikujeme Větu 2 a napíšeme výraz ve tvaru

$$-\frac{b_i \sum_{i,j=1}^N a_{ij}c_i c_j}{2\sqrt{2} \left( \sum_{i=1}^N b_i c_i \right)^2} \ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right) = f_{i,c}^{7A_p} + f_{i,nc}^{7A_p} + f_{i,c}^{7A_n} + f_{i,nc}^{7A_n}, \quad (3.35)$$

kde jsou funkce  $f_{i,c}^{7A_p}$ ,  $f_{i,nc}^{7A_p}$ ,  $f_{i,c}^{7A_n}$  a  $f_{i,nc}^{7A_n}$  definované následovně

$$f_{i,c}^{7A_p} = \frac{b_i \gamma_6}{4\sqrt{2}} \left( \frac{\mathbf{c}^T \mathbb{A}_p \mathbf{c}}{\gamma_6} + \left( -\frac{\ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right),$$

$$f_{i,nc}^{7A_p} = -\frac{b_i}{4\sqrt{2}\gamma_6} \left( (\mathbf{c}^T \mathbb{A}_p \mathbf{c})^2 + \left( -\gamma_6 \frac{\ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right),$$

$$f_{i,nc}^{7A_n} = -\frac{b_i \gamma_7}{4\sqrt{2}} \left( \frac{\mathbf{c}^T \mathbb{A}_n \mathbf{c}}{\gamma_7} + \left( -\frac{\ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right),$$

$$f_{i,c}^{7A_n} = \frac{b_i}{4\sqrt{2}\gamma_7} \left( (\mathbf{c}^T \mathbb{A}_n \mathbf{c})^2 + \left( -\gamma_7 \frac{\ln \left( 1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i c_i \right)}{\left( \sum_{i=1}^N b_i c_i \right)^2} \right)^2 \right),$$

kde  $f_{i,c}^{7A_p}$ ,  $f_{i,c}^{7A_n}$  jsou konvexní a  $f_{i,nc}^{7A_p}$ ,  $f_{i,nc}^{7A_n}$  jsou konkávní a zahrneme je do příslušných částí rozkladu. Nakonec rozebereme poslední výraz v rovnici pro chemické potenciály (1.2)

$$\left( \frac{b_i \sum_{i,j=1}^N a_{ij}c_i c_j}{-1 - 2 \sum_{i=1}^N b_i c_i + \left( \sum_{i=1}^N b_i c_i \right)^2} \right). \quad (3.36)$$

Jako první člen do rozkladu opět využijeme kvadratickou formu, kterou jsme již diskutovali a druhý člen má tvar  $\left( \frac{1}{-1 - 2 \sum_{i=1}^N b_i c_i + \left( \sum_{i=1}^N b_i c_i \right)^2} \right)$ , tento výraz je záporný, a tedy vytkneme mínus. Po použití substituce

(3.23) a Věty 5 diskutujeme konvexitu výrazu  $\left( \frac{1}{1+2x-(x)^2} \right)$ , po výpočtu druhé derivace znovu potvrzujeme konvexitu. Ověřili jsme předpoklady Věty 2 a můžeme přejít k její aplikaci

$$-\left(\frac{b_i \sum_{j=1}^N a_{ij} c_i c_j}{1 + 2 \sum_{i=1}^N b_i c_i - \left(\sum_{i=1}^N b_i c_i\right)^2}\right) = f_{i,c}^{8A_p} + f_{i,nc}^{8A_p} + f_{i,c}^{8A_n} + f_{i,nc}^{8A_n}, \quad (3.37)$$

kde jsou funkce  $f_{i,nc}^{8A_p}$ ,  $f_{i,c}^{8A_p}$ ,  $f_{i,c}^{8A_n}$  a  $f_{i,nc}^{8A_n}$  definované následovně

$$f_{i,nc}^{8A_p} = -\frac{b_i \gamma_8}{2} \left( \frac{\mathbf{c}^T \mathbb{A}_p \mathbf{c}}{\gamma_8} + \left( \frac{1}{1 + 2 \sum_{i=1}^N b_i c_i - \left(\sum_{i=1}^N b_i c_i\right)^2} \right)^2 \right),$$

$$f_{i,c}^{8A_p} = \frac{b_i}{2\gamma_8} \left( \left( \mathbf{c}^T \mathbb{A}_p \mathbf{c} \right)^2 + \left( \frac{\gamma_8}{1 + 2 \sum_{i=1}^N b_i c_i - \left(\sum_{i=1}^N b_i c_i\right)^2} \right)^2 \right),$$

$$f_{i,c}^{8A_n} = \frac{b_i \gamma_9}{2} \left( \frac{\mathbf{c}^T \mathbb{A}_n \mathbf{c}}{\gamma_9} + \left( \frac{1}{1 + 2 \sum_{i=1}^N b_i c_i - \left(\sum_{i=1}^N b_i c_i\right)^2} \right)^2 \right),$$

$$f_{i,nc}^{8A_n} = -\frac{b_i}{2\gamma_9} \left( \left( \mathbf{c}^T \mathbb{A}_n \mathbf{c} \right)^2 + \left( \frac{\gamma_9}{1 + 2 \sum_{i=1}^N b_i c_i - \left(\sum_{i=1}^N b_i c_i\right)^2} \right)^2 \right),$$

kde  $f_{i,c}^{8A_p}$ ,  $f_{i,c}^{8A_n}$  jsou konvexní a  $f_{i,nc}^{8A_p}$ ,  $f_{i,nc}^{8A_n}$  jsou konkávní a zahrneme je do příslušných částí rozkladu. Tímto jsme dokončili rozklad potenciálů a můžeme definovat tvary našich horních a dolních odhadů potenciálů pomocí Věty 3.

### 3.3.3 Konkávní a konvexní odhady chemických potenciálů

Uvedeme rozpis chemických potenciálů (1.2) do konvexní a do konkávní části. Nejdříve konvexní část

$$\mu_i^{convex}(\mathbf{c}) = f_{i,c}^2 + f_{i,c}^3 + f_{i,c}^4 + f_{i,c}^5 + f_{i,c}^{6A_p} + f_{i,c}^{6A_n} + f_{i,c}^{7A_p} + f_{i,c}^{7A_n} + f_{i,c}^{8A_p} + f_{i,c}^{8A_n}. \quad (3.38)$$

Podobně napíšeme konkávní část

$$\mu_i^{concave}(\mathbf{c}) = f_{i,nc}^1 + f_{i,nc}^3 + f_{i,nc}^4 + f_{i,nc}^5 + f_{i,nc}^{6A_p} + f_{i,nc}^{6A_n} + f_{i,nc}^{7A_p} + f_{i,nc}^{7A_n} + f_{i,nc}^{8A_p} + f_{i,nc}^{8A_n}. \quad (3.39)$$

Jistě platí

$$\mu_i(\mathbf{c}) = \mu_i^{convex}(\mathbf{c}) + \mu_i^{concave}(\mathbf{c}).$$

V posledním kroku využijeme Jensenovy nerovnosti (Věta 3) v obou směrech

$$\mu_i(\mathbf{c}) = \mu_i^{convex}(\mathbf{c}) + \mu_i^{concave}(\mathbf{c}) \leq \sum_{i=1}^N \alpha_i \mu_i^{convex}(\mathbf{c}^{(i)}) + \mu_i^{concave}(\mathbf{c}),$$

$$\mu_i(\mathbf{c}) = \mu_i^{\text{convex}}(\mathbf{c}) + \mu_i^{\text{concave}}(\mathbf{c}) \geq \mu_i^{\text{convex}}(\mathbf{c}) + \sum_{i=1}^N \alpha_i \mu_i^{\text{concave}}(\mathbf{c}^{(i)}).$$

Z tohoto tvaru je vidět, že bude výhodné používat barycentrické souřadnice. Pohybujeme se na simplexu, a tedy lze každý bod vyjádřit pomocí konvexní kombinace vrcholů ( $\mathbf{c}^{(i)}$ ) simplexu. Tímto jsme dokončili odvození konkávního horního a konvexního dolního odhadu chemických potenciálů na simplexu  $\mathbf{S}$ .

$$\mu_i^{\text{upper}}(\mathbf{c}) = \sum_{i=1}^N \alpha_i \mu_i^{\text{convex}}(\mathbf{c}^{(i)}) + \mu_i^{\text{concave}}(\mathbf{c}), \quad (3.40)$$

$$\mu_i^{\text{lower}}(\mathbf{c}) = \mu_i^{\text{convex}}(\mathbf{c}) + \sum_{i=1}^N \alpha_i \mu_i^{\text{concave}}(\mathbf{c}^{(i)}). \quad (3.41)$$

Ještě je potřeba uvést tvar tečné nadroviny ke středu zkoumaného simplexu  $\mathbf{c}_0$

$$\mu_{i,\text{plane}}^{\text{upper}}(\mathbf{c}) = \mu_i^{\text{upper}}(\mathbf{c}_0) + \nabla \mu_i^{\text{upper}}(\mathbf{c}_0)(\mathbf{c} - \mathbf{c}_0), \quad (3.42)$$

$$\mu_{i,\text{plane}}^{\text{lower}}(\mathbf{c}) = \mu_i^{\text{lower}}(\mathbf{c}_0) + \nabla \mu_i^{\text{lower}}(\mathbf{c}_0)(\mathbf{c} - \mathbf{c}_0). \quad (3.43)$$

### 3.3.4 Ladění parametru $\gamma$

Obdobně jako u stavové rovnice vždy upravujeme buď konkávní nebo konvexní část rozkladu. Z předchozích kapitol jsme došli k závěru, že v obou případech je prospěšné volit parameter  $\gamma$  jako podíl funkcí v rozkladu braných uprostřed daného simplexu přesněji, jestliže odhadujeme funkci  $f = gh$ , potom  $\gamma = \frac{g(\mathbf{B})}{h(\mathbf{B})}$ , kde  $\mathbf{B}$  je těžiště simplexu.

Aplikací tohoto závěru na každou z odhadnutých funkcí zvlášť poté získáme volby

$$\gamma_1 = \frac{\sum_{i=1}^N \mathbf{B}_i}{1 - \sum_{i=1}^N b_i \mathbf{B}_i},$$

$$\gamma_2 = \frac{\sum_{j=1}^N \mathbf{B}_j a_{ij}}{\frac{\ln\left(1 + (1 - \sqrt{2}) \sum_{i=1}^N b_i \mathbf{B}_i\right)}{\sum_{i=1}^N b_i \mathbf{B}_i}},$$

$$\gamma_3 = \frac{\sum_{j=1}^N \mathbf{B}_j a_{ij}}{\frac{\ln\left(1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i \mathbf{B}_i\right)}{\sum_{i=1}^N b_i \mathbf{B}_i}},$$

$$\gamma_4 = \frac{\mathbf{B}^T \mathbb{A}_p \mathbf{B}}{\frac{\ln\left(1 + (1 + \sqrt{2}) \sum_{i=1}^N b_i \mathbf{B}_i\right)}{\left(\sum_{i=1}^N b_i \mathbf{B}_i\right)^2}},$$



$$\gamma_5 = \frac{\mathbf{B}^T \mathbf{A}_n \mathbf{B}}{\frac{\ln\left(1+(1+\sqrt{2})\sum_{i=1}^N b_i \mathbf{B}_i\right)}{\left(\sum_{i=1}^N b_i \mathbf{B}_i\right)^2}},$$

$$\gamma_6 = \frac{\mathbf{B}^T \mathbf{A}_p \mathbf{B}}{\frac{\ln\left(1+(1-\sqrt{2})\sum_{i=1}^N b_i \mathbf{B}_i\right)}{\left(\sum_{i=1}^N b_i \mathbf{B}_i\right)^2}},$$

$$\gamma_7 = \frac{\mathbf{B}^T \mathbf{A}_n \mathbf{B}}{\frac{\ln\left(1+(1-\sqrt{2})\sum_{i=1}^N b_i \mathbf{B}_i\right)}{\left(\sum_{i=1}^N b_i \mathbf{B}_i\right)^2}},$$

$$\gamma_8 = \frac{\mathbf{B}^T \mathbf{A}_p \mathbf{B}}{\frac{1}{1+2\sum_{i=1}^N b_i \mathbf{B}_i - \left(\sum_{i=1}^N b_i \mathbf{B}_i\right)^2}},$$

$$\gamma_9 = \frac{\mathbf{B}^T \mathbf{A}_n \mathbf{B}}{\frac{1}{1+2\sum_{i=1}^N b_i \mathbf{B}_i - \left(\sum_{i=1}^N b_i \mathbf{B}_i\right)^2}},$$

kde  $\mathbf{B}$  je těžiště simplexu.

### 3.3.5 Kritéria pro ořezávání simplexu na základě chemických potenciálů

Jak jsme naznačili v úvodu této kapitoly, první možné kritérium lze volit následovně. Pokud pro simplex  $\mathbf{S}$  platí jedna z následujících nerovností

$$\mu(\mathbf{c}^*) \geq \max_{\mathbf{c} \in \mathbf{S}} \mu_i^{upper}(\mathbf{c}), \quad (3.44)$$

$$\mu(\mathbf{c}^*) \leq \min_{\mathbf{c} \in \mathbf{S}} \mu_i^{lower}(\mathbf{c}), \quad (3.45)$$

vyloučíme ho z přípustné množiny. Další možné kritérium formulujeme obdobně pomocí téčné nadroviny. Pokud množinu vrcholů simplexu  $\mathbf{S}$  označíme jako  $\mathbf{S}_e$  a platí nějaká z následujících nerovností

$$\mu(\mathbf{c}^*) \geq \max_{\mathbf{c} \in \mathbf{S}_e} \mu_{i,plane}^{upper}(\mathbf{c}), \quad (3.46)$$

$$\mu(\mathbf{c}^*) \leq \min_{\mathbf{c} \in \mathbf{S}_e} \mu_{i,plane}^{lower}(\mathbf{c}), \quad (3.47)$$

tak ho vyloučíme z přípustné množiny. Vyčíslení maxima v (3.44), respektive minima v (3.45) je získáno lokální optimalizací, zatímco maxima v (3.46), respektive minima v (3.47) jsou získána jen vyčíslením konečného počtu hodnot funkce  $\mu_{i,plane}^{upper}(\mathbf{c})$ , respektive  $\mu_{i,plane}^{lower}(\mathbf{c})$  ve vrcholech simplexu  $\mathbf{S}$ .

### 3.4 Upravený algoritmus

V další kapitole budeme zkoumat vliv podmínek z kapitol 3.2.5 a 3.3.5 na výpočetní složitost algoritmu větví a mezí. Předtím, než budeme řešit podhodnocený problém (2.5), nejdříve ověříme, že na příslušném simplexu je splněna nějaká z podmínek z kapitol 3.2.5 a 3.3.5. Pokud tomu tak není, vyloučíme tento simplex.

---

**Algoritmus 2:** Pseudoalgoritmus větví a mezí

---

**Result:** Vypiš **UB**  
do  $S_a$  ulož celou přípustnou množinu (počáteční simplex);  
**UB** =  $+\infty$ ;  
**LB** =  $-\infty$ ;  
**while** **UB** je blízko **LB** **do**  
    vyber simplex  $S_a$  z množiny aktivních simplexů a odstraň ho z fronty;  
    spočti původní problém (2.4) a optimum ulož do **UB1**;  
    **if** **UB1** < **UB** **then**  
        | **UB1** =: **UB**;  
    **end**  
    rozděl simplex **S** podle nějakého pravidla na **S1** a **S2** ;  
    **if** na **S1** je splněna nutná podmínka **then**  
        | spočti podhodnocený problém (2.5) na simplexech a optimum ulož do **LB1**;  
    **else**  
        | **LB1** =  $\infty$   
    **end**  
    **if** na **S2** je splněna nutná podmínka **then**  
        | spočti podhodnocený problém (2.5) na simplexech a optimum ulož do **LB2**;  
    **else**  
        | **LB2** =  $\infty$   
    **end**  
    **if** **LB1** < **UB** **then**  
        | Přidej **S1** do  $S_a$ ;  
    **end**  
    **if** **LB2** < **UB** **then**  
        | Přidej **S2** do  $S_a$ ;  
    **end**  
    do **LB** ulož minimum z minim získaných spočtením podhodnoceného problému 2.5 ve vnitřku aktivních simplexů;  
**end**

---

Nutná podmínka zde může být jakákoliv kombinace z určená rovnicemi (3.18), (3.19), (3.44), (3.45), (3.46) nebo (3.47). My se omezíme na 4 varianty následovně

- (a) Nutná podmínka: (3.18)
- (b) Nutná podmínka: (3.19)
- (c) Nutná podmínka: (3.44) a (3.45)
- (d) Nutná podmínka: (3.46) a (3.47).

## Kapitola 4

### Příklady

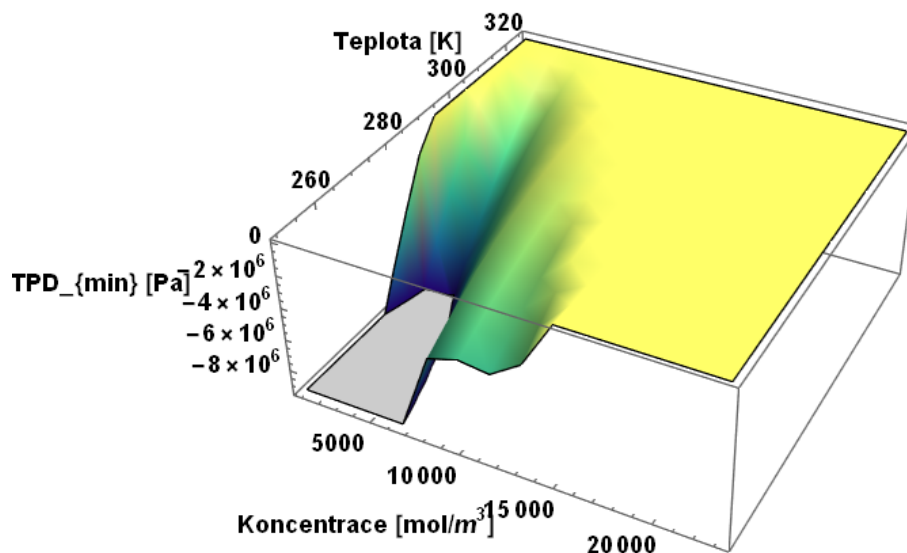
V této kapitole budeme postupy odvozené v práci aplikovat na příklady z literatury. V tabulce 4.1 jsou vyneseny fyzikální parametry použitých látek. Jako zastavovací kritérium jsme použili  $\frac{|UB-LB|}{|UB|+1} < 10^{-5}$ . Základní algoritmus porovnáváme se 4 verzemi upraveného algoritmu. V prvních dvou případech budeme využívat nutnou podmínku na tlak, tedy (3.18) a následně (3.19). V dalších dvou budeme využívat chemických potenciálů, tedy buď rovnic (3.44) a (3.45) nebo (3.46) a (3.47). Algoritmy budeme porovnávat na základě počtu iterací potřebných k nalezení dostatečně přesného výsledku a také času, který k výpočtu byl potřeba. Algoritmy jsme implementovali v programu Wolfram Mathematica. Na hledání lokálního minima byla použita metoda vnitřního bodu založené na technice prediktor-korektor [17]. Jako počáteční odhad řešení používáme těžiště aktivního simplexu. Výpočet byl proveden na počítači s procesorem AMD Ryzen 5 3600XT 3,8GHz a operační paměť 32GB.

Tabulka 4.1: Fyzikální parametry použitých látek

Látka	Kritická teplota (K)	Kritický tlak (Pa)	Acentrický faktor
Metan ( $C_1$ )	190,40	$46,00 \cdot 10^5$	0,011
Propan ( $C_3$ )	369,80	$42,50 \cdot 10^5$	0,1530
Oxid uhličitý	304,14	$73,75 \cdot 10^5$	0,2390
Dekan ( $C_{10}$ )	617,7	$73,21 \cdot 10^5$	0,489

## 4.1 Směs metanu a propanu

Obrázek 4.1: Globální minimum funkce TPD v prvním příkladu



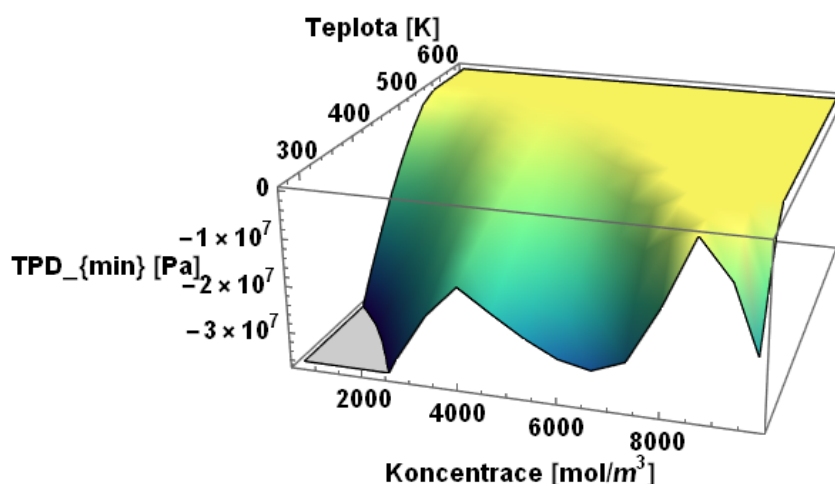
V prvním příkladu studujeme fázovou stabilitu metan-propanové směsi. Molární zlomky použitých látek jsou  $z_{C_3}^* = 0,452587$  a  $z_{C_1}^* = 0,547413$ . Interakční koeficient je  $\delta_{C_1-C_3} = 0,0365$  a fyzikální parametry jsou v tabulce 4.1. Stabilita směsi byla testována na mřížce  $15 \times 15$  pro teploty v rozsahu  $T = (250 \text{ K}; 650 \text{ K})$  a všechny přípustné koncentrace. V grafu 4.1 vidíme globální minimum funkce TPD pro tyto teploty a koncentrace. V druhém grafu vidíme časovou výpočetní náročnost a ve třetím potřebný počet iterací algoritmu. Při použití algoritmu za využití horního konvexně-konkávního odhadu Pengovy stavové rovnice se počet iterací zmenšil v průměru o 14,9 %, ale výpočetní čas se zhoršil v průměru o 16,7 %. Při použití tečné nadrovinu k tomuto hornímu odhadu se počet iterací zmenšil o 14,1 %, ale nyní došlo i ke zlepšení z hlediska výpočetního času, který se zlepšil o 16,8 %. Ke zlepšení dochází především v oblasti fázového přechodu na straně, kde jsou stavy stabilní. V grafech 4.3 a 4.4 jsou vyneseny trojice grafů. V prvním z trojice je celkový počet časů/iterací, v druhém je změna, o kolik se tyto veličiny změnily při použití kritéria (3.18) a na třetím změna za použití kritéria (3.19). Při používání chemických potenciálů už k vylepšení nedošlo ani v jedné alternativě algoritmu. Při používání konvexně-konkávních odhadů a lokální minimalizace se čas potřebný k výpočtu zvětšil v průměru o 312,12 %, což není tak překvapivé z důvodu řešení dalších lokálních minimalizací v každé iteraci. Ani z hlediska počtu iterací nedošlo k výraznému poklesu potřebných iterací, ten klesl o 3,97 %. Při použití alternativy s tečnou nadrovinou nyní nedošlo k ušetření času, ten se zhoršil o 41,65 %. Počet potřebných iterací opět klesl oproti základnímu algoritmu o 3,59 %. Opět uvedeme grafy 4.5 a 4.6 porovnávající výpočetní složitost oproti základnímu algoritmu. V tabulce 4.2 jsou uvedeny výsledky z tohoto příkladu.

Tabulka 4.2: Výsledky z prvního příkladu, zlepšení/zhoršení je vztaženo k základnímu algoritmu.

Nutná podmínka v algoritmu	a)	b)	c)	d)
průměrné ušetření času	-16,7 %	16,8 %	-312,12 %	-41,65 %
průměrné ušetření počtu iterací	14,9 %	14,1 %	3,97 %	3,59 %

## 4.2 Směs dekanu a oxidu uhličitého

Obrázek 4.2: Globální minimum funkce TPD v druhém příkladu



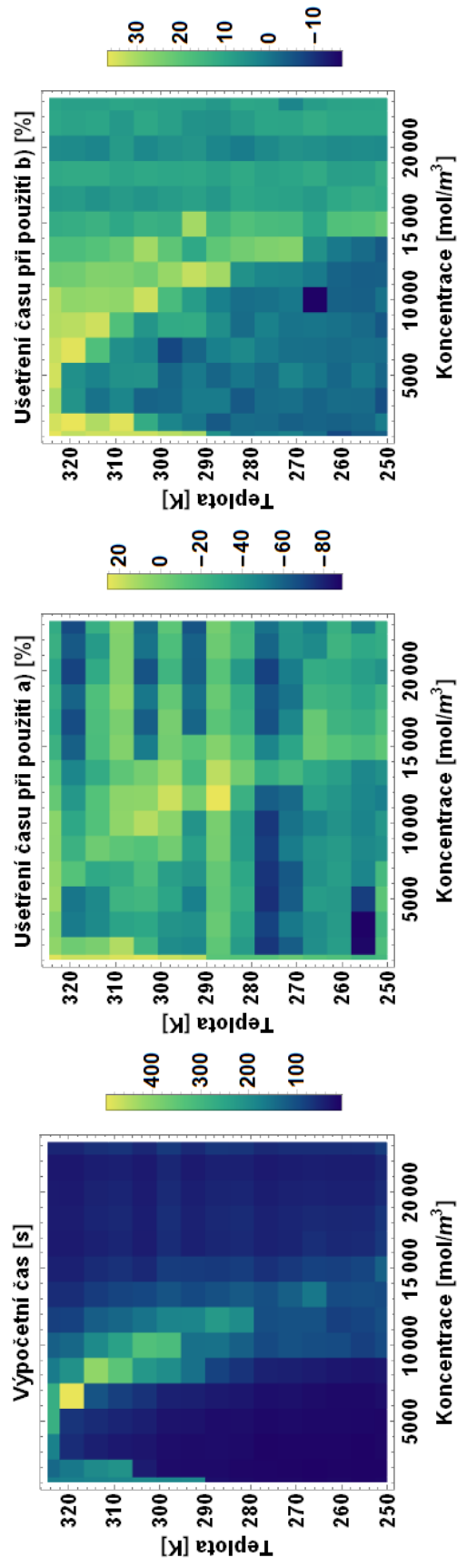
V druhém příkladu studujeme fázovou stabilitu směsi oxidu uhličitého a dekanu. Molární zlomky použitých látek jsou  $z_{CO_2}^* = 0,547413$  a  $z_{nC_{10}}^* = 0,452587$ . Interakční koeficient je  $\delta_{nC_{10}-CO_2} = 0,15$  a fyzikální parametry jsou v tabulce 4.1. Stabilita směsi byla testována na mřížce  $15 \times 15$  pro teploty v rozsahu  $T = (250 \text{ K}; 650 \text{ K})$  a všechny přípustné koncentrace. Upravený algoritmus vždy porovnáme se základním. V grafu 4.2 vidíme globální minimum funkce TPD pro tyto teploty a koncentrace. V druhém grafu vidíme časovou výpočetní náročnost a ve třetím potřebný počet iterací algoritmu. Při použití algoritmu za využití horního konvexně-konkávního odhadu Pengovy stavové rovnice se počet iterací zmenšil v průměru o 10,15 % a výpočetní čas se zhoršil v průměru o 18,88 %. Při použití tečné nadroviny k tomuto hornímu odhadu se počet iterací zmenšil o 9,36 % a znovu došlo i ke zlepšení z hlediska výpočetního času, který se zlepšil o 11,57 %. Ke zlepšení dochází především v oblasti fázového přechodu na straně, kde jsou stavy stabilní. V grafech 4.7 a 4.8 jsou vyneseny trojice grafů. V prvním z trojice je celkový počet časů/iterací, v druhém je změna, o kolik se tyto veličiny změnily při použití kritéria (3.18) a ve třetím změna za použití kritéria (3.19). Při používání chemických potenciálů už k vylepšení nedošlo ani v jedné alternativě algoritmu. Při používání konvexně-konkávních odhadů a lokální minimalizace se čas potřebný k výpočtu zvětšil v průměru o 342,12 %, což není tak překvapivé z důvodu řešení dalších lokálních minimalizací v každé iteraci. Ani z hlediska počtu iterací nedošlo k výraznému poklesu potřebných iterací, ten klesl o 6,31 %. Při použití alternativy s tečnou nadrovinou nyní nedošlo k ušetření

času, ten se zhoršil o 36,89 %. Počet potřebných iterací opět klesl oproti základnímu algoritmu o 4,39 %. Opět uvedeme grafy 4.9 a 4.10 porovnávající výpočetní složitost oproti základnímu algoritmu. V tabulce 4.3 jsou uvedeny výsledky z tohoto příkladu.

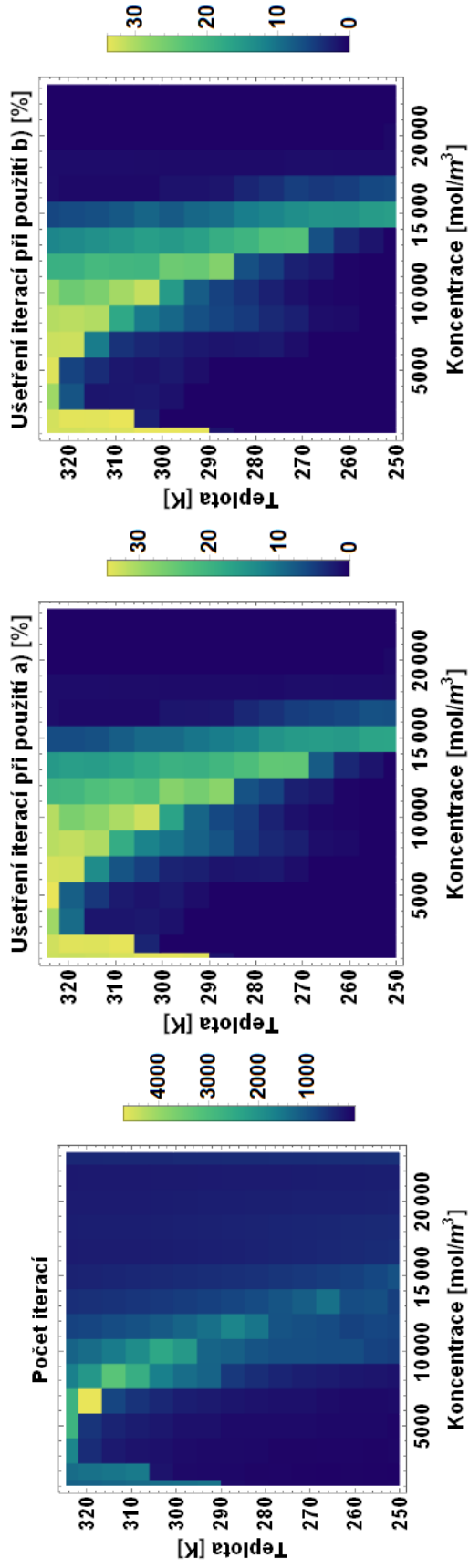
Tabulka 4.3: Výsledky z druhého příkladu, zlepšení/zhoršení je vztaženo k základnímu algoritmu.

Nutná podmínka v algoritmu	a)	b)	c)	d)
průměrné ušetření času	-18,88 %	11,57 %	-342,12 %	-36,89 %
průměrné ušetření počtu iterací	10,15 %	9,36 %	6,31 %	4,39 %

Obrázek 4.3: Srovnání výpočetního času v prvním příkladu přístupu využívajícím a) a b)

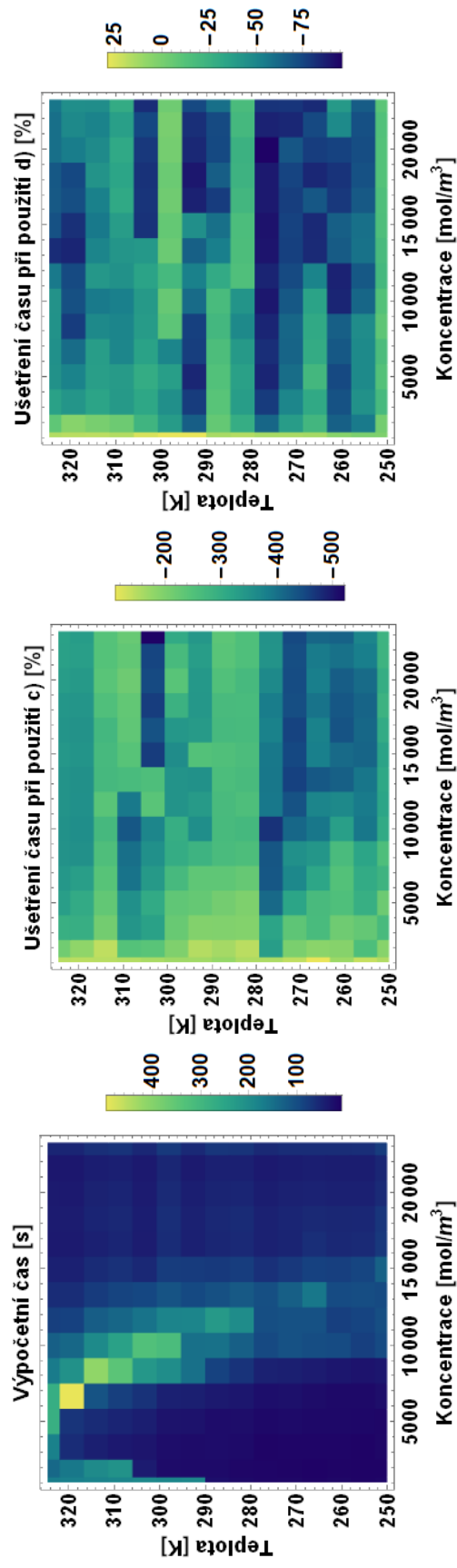


Obrázek 4.4: Srovnání počtu iterací v prvním příkladu přístupem využívajícím a) a b)

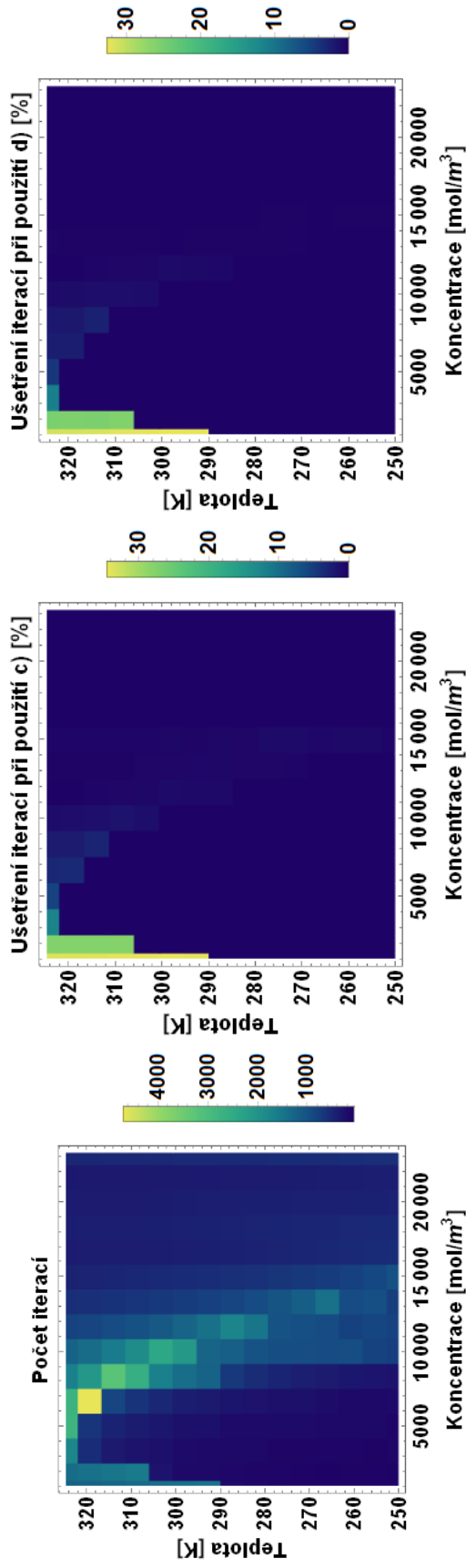




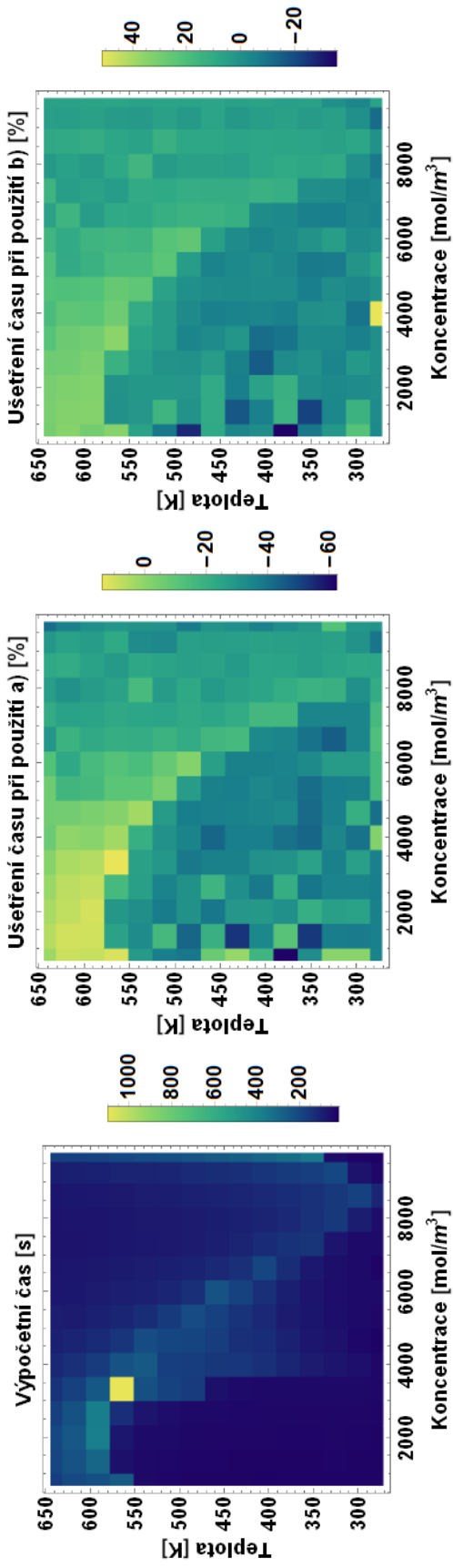
Obrázek 4.5: Srovnání výpočetního času v prvním příkladu přístupu využívajícím c) a d)



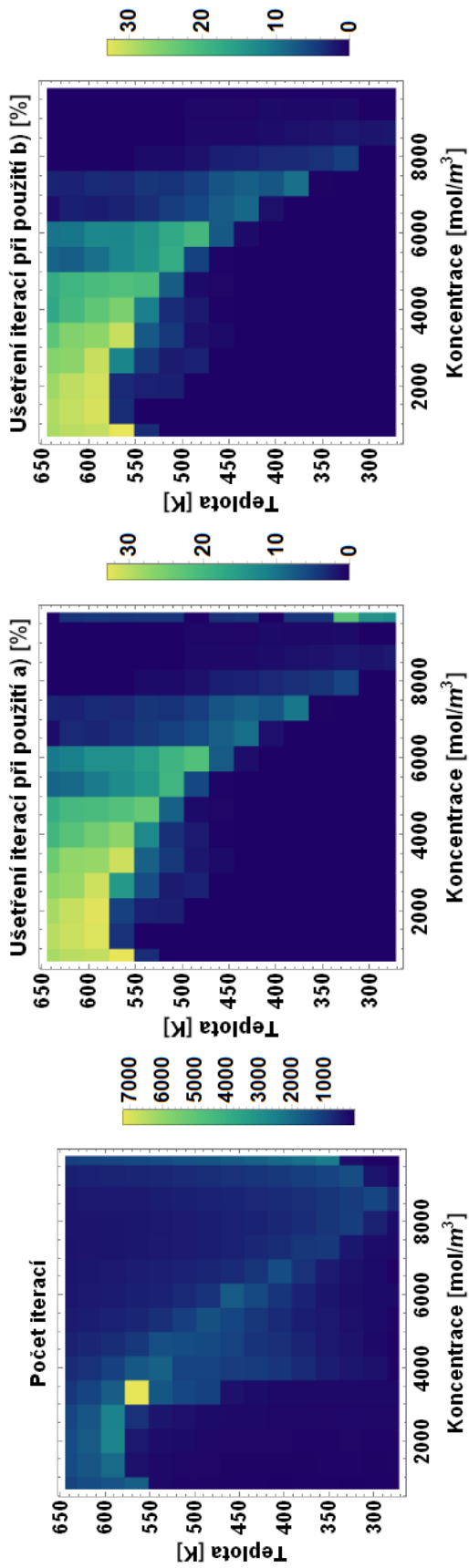
Obrázek 4.6: Srovnání počtu iterací v prvním příkladu přístupu využívajícím c) a d)



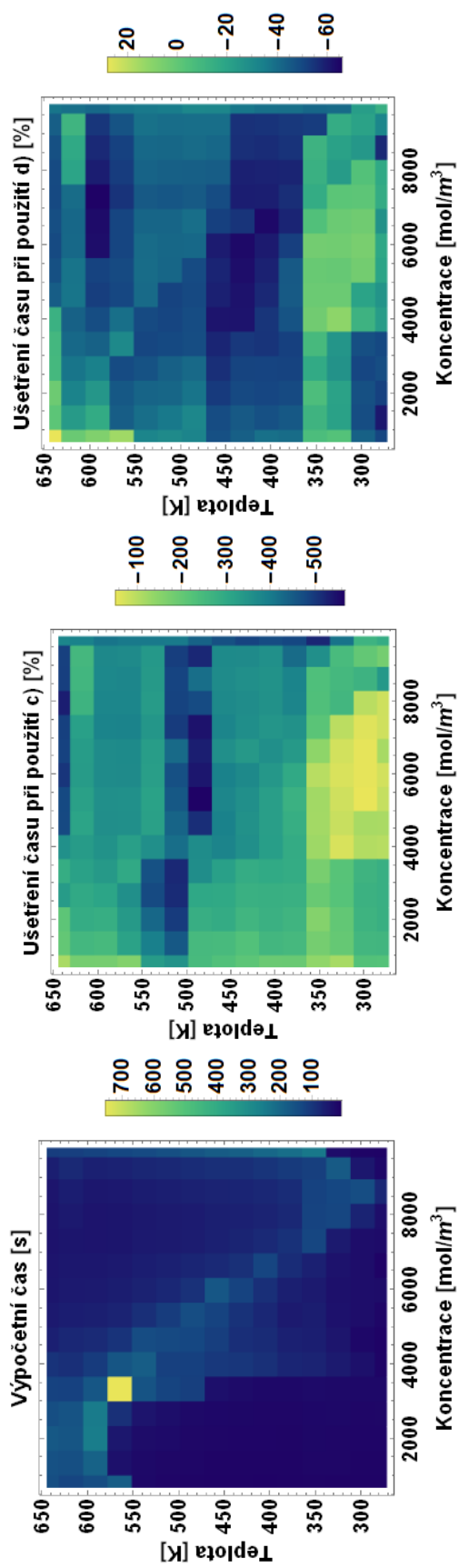
Obrázek 4.7: Srovnání výpočetního času v druhém příkladu přístupu využívajícím a) a b)



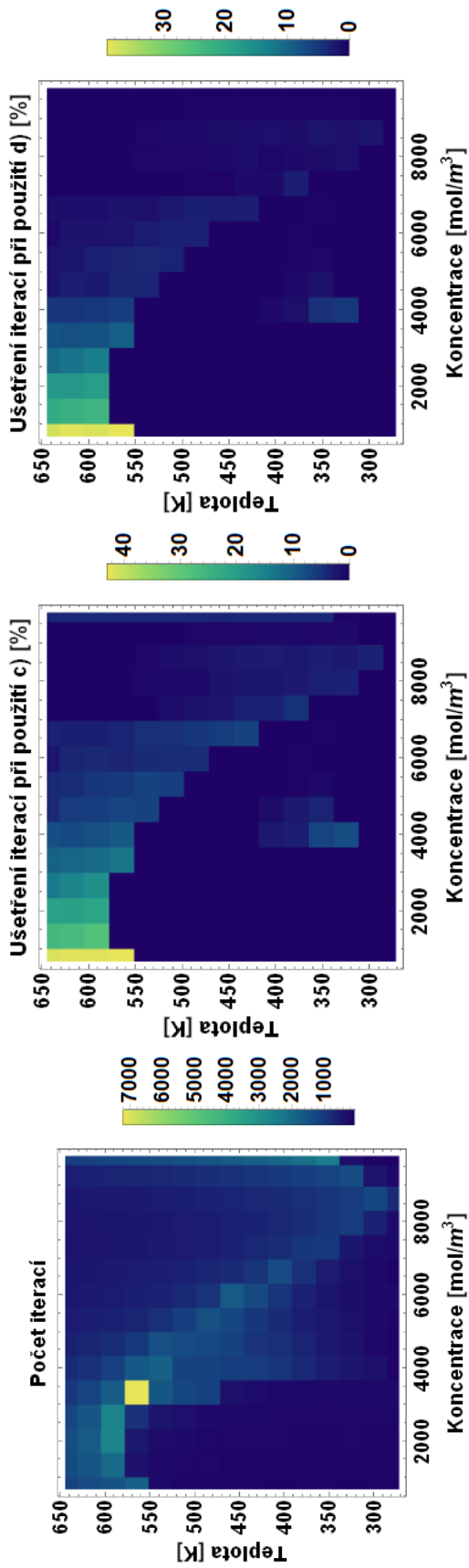
Obrázek 4.8: Srovnání počtu iterací v druhém příkladu přístupem využívajícím a) a b)



Obrázek 4.9: Srovnání výpočetního času v druhém příkladu průstupu využívajícím c) a d)



Obrázek 4.10: Srovnání počtu iterací v druhém příkladu přístupem využívajícím c) a d)



# Závěr

V práci jsme se zabývali efektivnějším prořezáváním stromu vznikajícím v algoritmu větví a mezí při aplikaci na testování fázové stability. Při použití Pengovy stavové rovnice, přesněji při aplikaci podmínky (3.19) došlo ke zlepšení z hlediska průměrného výpočetního času a počtu potřebných iterací. Jinými slovy odhad tlaku získaný z tečné nadroviny k hornímu konvexně-konkávnímu odhadu je dostatečně kvalitní a v algoritmu větví a mezí se vyvaruje vyhodnocování účelové funkce na části neperspektivních simplexů. Samotný přístup s využitím horního konvexně-konkávního odhadu tlaku (3.18) ke zlepšení nestačil, protože na získání jeho optima potřebujeme lokální optimalizaci a následně musíme znovu získat optimum dolního konvexně-konkávního odhadu funkce TPD. Důvodem je, že opakovaně v jedné iteraci algoritmu provádíme na místo jedné optimalizace dvě a počet iterací, které tímto postupem ušetříme, není dostačující ke zlepšení průměrného výpočetního času, i když byl odhad kvalitnější (větší počet ušetřených iterací) oproti přístupu s použitím rovnice (3.19). Při použití chemických potenciálů, podmínky (3.44)-(3.47), nedošlo k výraznému zlepšení, což napovídá tomu, že konvexně-konkávní odhad není dostatečně kvalitní. To je způsobeno počtem členů, na které používáme Větu 2, který je u chemických potenciálů (1.2) vyšší než u Pengovy-Robinsonovy rovnice (1), což vede na hrubší odhad. Do budoucna se můžeme zabývat jinými odhady než odhady konstruovanými pomocí Věty 2. Ideální by bylo použít konvexní obal grafu, který je ale náročný na konstrukci. My jsme se snažili přiblížit k tomuto konvexnímu obalu pomocí Věty 2, která při použití na Pengovu-Robinsonovu rovnici (1) vedla k výraznému zlepšení v oblasti fázového přechodu. V práci byl také upraven konvexně-konkávní rozklad z článku [8] funkce TPD (1.4) pro splnění předpokladů Věty 2.

Tato diplomová práce byla podpořena z projektu SGS20/184/OHK4/3T/14 (Vývoj a aplikace pokročilých metod matematického modelování přírodních a průmyslových procesů s využitím vysoce výkonné výpočetní techniky).

# Literatura

- [1] Hartman, P. *On functions representable as a difference of convex functions*. Pacific Journal of Mathematics, 9, 707-713, 1959.
- [2] Lachout, P.; Dupačová, J. *Úvod do optimalizace*. Matfypress, První vydání, Praha 2011
- [3] Jensen, J. L. W. V. Acta Math. 30 (1906), 175–193. *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*. Acta Math, 30, 175–193, 1906.
- [4] Peng, D. Y.; Robinson, D. B. *A New Two-Constant Equation of State*. Industrial & Engineering Chemistry Fundamentals, 15, 59–64, 1976.
- [5] Stryjek, R.; Vera, J. H. *PRSV: An improved peng-Robinson equation of state for pure compounds and mixtures*. The Canadian Journal of Chemical Engineering, 64 (2), 323–333, 1986.
- [6] Stryjek, R.; Vera, J. H. *PRSV2: A cubic equation of state for accurate vapor–liquid equilibria calculations*. The Canadian Journal of Chemical Engineering, 64 (5), 820–826, 1986.
- [7] Mikyška, J.; Firoozabadi A. *Investigation of Mixture Stability at Given Volume, temperature and number of moles*. Fluid Phase Equilibria, 321, 1-9, 2012.
- [8] Smejkal T.; Mikyška J. *VTN-phase stability testing using the Branch and Bound strategy and the convex-concave splitting of Helmholtz free energy density*. Fluid Phase Equilibria, 504, 112323, 2020.
- [9] Jex M. *Testování fázové stability vícesložkových směsí s využitím metod globální optimalizace*. Výzkumný úkol, Katedra matematiky FJFI 2020.
- [10] Locatelli M.; Schoen F. *Global optimization - Theory, Algorithms, and Applications*. SIAM Philadelphia, 2013.
- [11] Jüngel, A., Mikyška, J., and Zamponi, N. *Existence analysis of a single phase flow mixture with van der Waals pressure*. SIAM Journal on Mathematical Analysis, 50(1): 1367–1395, 2018.
- [12] Paulavičius R., Žilinskas J. *Springer-Verlag New York*. Springer-Verlag New York, 2014.
- [13] Jaumard B., Meyer Ch., Tuy H. *Generalized Convex Multiplicative Programming via Quasiconcave Minimization*. Journal of Global Optimization, 10, 229-256, 1997.
- [14] Oberman A. M. *The Convex Envelope is the Solution of a Nonlinear Obstacle Problem*. In Proceedings of the American Mathematical Society, 135(6), 1689-1694, 2007.
- [15] Tuy H. *Global Minimization of a Difference of 2 Convex Functions*. Mathematical Programming Study, 30, 150-182, 1987.



- [16] Thi H., Dinh T.P. *DC programming and DCA: thirty years of developments*. Special Issue of Mathematical Programming, 169(1), 5-68, 2018.
- [17] Mehrotra, S. *On the Implementation of a Primal-Dual Interior Point Method*. SIAM J. Optimization, 2, 575-601, 1992.