



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Damián Kulich

Stochastické metody v krystalografii

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Prof. RNDr. Viktor Beneš, DrSc.

Studijní program: Matematika

Studijní obor: Obecná matematika

Praha 2020

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Chtěl bych poděkovat svému vedoucímu práce Prof. Rndr. Viktoru Benešovi, DrSc. za vedení práce a trpělivost. Chtěl bych poděkovat Mgr. Filipovi Seitlovi za rady k implementaci algoritmu a zpracování výsledků. Chtěl bych poděkovat svojí rodině za podporu při psaní práce a v celém mém dosavadním studiu.

Název práce: Stochastické metody v krystalografii

Autor: Damián Kulich

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Prof. RNDr. Viktor Beneš, DrSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V práci zavedeme mozaiky jako model mikrostruktury zrn v polykrystalickém materiálu. Dále představíme potřebné popisy trojrozměrných orientací pro kótování buněk. Představíme potřebnou teorii markovských řetězců abychom mohli používat MCMC algoritmy. Hlavním úkolem pak bude simulovat možná rozdělení misorientací mezi sousedními buňkami mozaiky. K tomu zavedeme parametrický stochastický model a ukážeme, že náhodný výběr z hledaného rozdělení lze simulovat pomocí MCMC metody. V závěrečné části diskutujeme výsledky simulací v závislosti na parametru a geometrii mozaiky.

Klíčová slova: mozaika, rozdělení misorientací, markovské Monte Carlo, krystalografie

Title: Stochastic Methods in Crystallography

Author: Damián Kulich

Department: Department of Probability and Mathematical Statistics

Supervisor: Prof. RNDr. Viktor Beneš, DrSc., Department of Probability and Mathematical Statistics

Abstract: First we define marked tessellations to use as a model for polycrystalline structure. Then we list the necessary descriptions of orientations to use as marks for the tessellation. We formulate the necessary theory of Markov chains, so that we can use MCMC algorithms. The main goal is to simulate possible distributions of misorientations between neighboring cells of a tessellation. For that we formulate a parametric stochastic model and show, that we can simulate from the target distribution using an MCMC method. In the final chapter, we discuss how the results depend on the parameter and geometry of the tessellation.

Keywords: tessellation, misorientation distribution, Markov chain Monte Carlo, crystallography

Obsah

| | |
|---|-----------|
| Úvod | 2 |
| 1 Teoretické základy | 3 |
| 1.1 Mozaiky v \mathbb{R}^3 | 3 |
| 1.2 Popis orientací v \mathbb{R}^3 | 4 |
| 1.3 Markovské řetězce, obecný stavový prostor | 6 |
| 1.4 Metody MCMC | 7 |
| 2 Stochastický model | 10 |
| 2.1 Model hustoty orientací buněk | 10 |
| 2.2 MCMC pro výběr z hustoty orientací | 11 |
| 3 Simulační studie | 13 |
| 3.1 Krystalografická interpretace | 13 |
| 3.2 Výpočetní programy | 13 |
| 3.3 Výsledky, závislost na θ | 14 |
| 3.3.1 $\theta = 0$ | 15 |
| 3.3.2 $\theta > 0$ | 15 |
| 3.3.3 $\theta < 0$ | 15 |
| 3.4 Výsledky, závislost na počtu sousedů | 16 |
| 3.5 Obrázky a tabulky | 17 |
| Závěr | 24 |
| Seznam použité literatury | 25 |

Úvod

Předmětem této práce jsou kótované mozaiky v \mathbb{R}^3 (Pawlas a kol., 2020), jejichž kótami jsou orientace kubických mřížek. Téma vychází z aplikace v materiálovém výzkumu, kde buňky jsou zrna polykrystalických materiálů (Morawiec, 2004), v nichž se sledují mřížky atomů. Zajímají nás možná rozdělení misorientací mezi kótami sousedních buněk mozaiky. Budeme se zaměřovat na extrémní rozdělení ve smyslu co největších, respektive co nejmenších misorientací. Volíme Stochastický model orientací buněk Gibbsova typu (Seitl a kol., 2020), který je třeba simulovat pomocí metod markovské Monte Carlo (Prokešová, 2020). Základy markovských řetězců s obecnou množinou stavů jsou dále vysvětleny v knize Meyn a Tweedie (1993).

Teoretickým výsledkem práce je volba modelu a ověření konvergence řetězce generovaného MCMC algoritmem, za pomoci Jones a kol. (2014). Z hlediska aplikace jde o vyšetření tvaru cílového rozdělení pro různé hodnoty parametru a v závislosti na tvaru vstupní mozaiky. Toho je dosaženo stochastickými simulacemi s pomocí markovského Monte Carlo.

Práce je členěna do tří kapitol. První z nich uvádí teoretické základy rozdělené do čtyř sekcí, první se zabývá definicí a vlastnostmi mozaiky, druhá popisuje orientace, třetí definuje markovské řetězce nad obecným stavovým prostorem a popisuje jejich vlastnosti a čtvrtá se týká MCMC metod a jejich konvergence. Druhá kapitola se zabývá konkrétním stochastickým modelem, který používáme, a popisuje jak z něj simulovat. Třetí kapitola nejprve popisuje krystalografickou interpretaci zavedené teorie, dále představuje použitý výpočetní program. Potom uvádí numerické výsledky simulací, které jsou rozděleny do dvou sekcí. Sekce 3.3 se zabývá závislostí cílového rozdělení na volbě parametru. Sekce 3.4 se zabývá tím, jak rozdělení závisí na množství stěn buněk mozaiky.

1. Teoretické základy

1.1 Mozaiky v \mathbb{R}^3

V trojrozměrném Eukleidovském prostoru \mathbb{R}^3 označíme $|\cdot|$ Eukleidovskou normu.

Definice 1 (Obecná poloha v \mathbb{R}^3). Říkáme, že n -tice bodů

$$\{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^3, n \in \mathbb{N}$$

je v **obecné poloze**, pokud je splněno následující:

- i. Žádné tři body neleží na přímce.
- ii. Žádné čtyři body neleží v rovině.
- iii. Žádných pět bodů neleží na sféře.

Definice 2 (Voronoiho mozaika v \mathbb{R}^3). Necht $\{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^3, n \in \mathbb{N}$ je v obecné poloze, pak

$$C_i = \{x \in \mathbb{R}^3; \forall k \neq i : |x - x_i| \leq |x - x_k|\}$$

nazveme **Voronoiho buňkou** příslušící x_i a množinu $V = \{C_1, C_2, \dots, C_n\}$ nazveme **Voronoiho mozaikou** v \mathbb{R}^3 .

Definice 3 (Laguerreho mozaika). Necht $y_i = (x_i, r_i) \in \mathbb{R}^3 \times \mathbb{R}^+, i = 1, \dots, n$ jsou dvojice, kde x_i jsou body v obecné poloze, pak

$$C_i = \{x \in \mathbb{R}^3; \forall k \neq i : \rho(x, y_i) \leq \rho(x, y_k)\},$$

kde

$$\rho(x, y_j) = |x - x_j|^2 - r_j,$$

nazveme **Laguerreho buňka** příslušící x_i a množinu $L = \{C_1, C_2, \dots, C_n\}$ nazveme **Laguerreho mozaikou** v \mathbb{R}^3 .

Je zřejmé, že Voronoiho mozaika je speciální případ Laguerreho mozaiky, kde pro každé i platí $r_i = 0$. Buňky Laguerreho i Voronoiho mozaiky jsou konvexní polygony, které vyplňují celý prostor, jejich vnitřky jsou po dvou disjunktní množiny.

Definice 4 (k -fasety buněk). Necht L je Laguerreho mozaika. Pro $k = 1, 2, 3, 4$, zdefinujeme **k -fasetu** pro různé buňky $C_{j_0}, \dots, C_{j_{4-k}} \in L$ jako $\bigcap_{i=0}^{4-k} C_{j_i}$, pokud je neprázdný. 0-fasetě říkáme vrchol, 1-fasetě hrana, 2-fasetě stěna a 3-faseta je celá příslušná buňka. Pokud mají dvě buňky společnou stěnu, řekneme, že **sousedí**. Relaci sousedění značíme pomocí \sim .

Definice 5 (Okno pozorování). Necht O je kompaktní, konvexní množina v \mathbb{R}^3 a necht L je Laguerreho mozaika. O nazveme **oknem pozorování**. Zavedeme $\hat{C}_i = C_i \cap O$ pro každé $C_i \in L$. Množinu $W = \{\hat{C}_1, \dots, \hat{C}_n\}$ nazveme **pozorovacím mozaikou v okně**.

Definice 6 (Kótovaná mozaika). *Nechť \mathbb{K} je libovolná neprázdná množina, nechť L je Laguerreho mozaika. Pak zavedeme pro každé $i = 1, \dots, n$ **kótovanou buňku** jako dvojici (C_i, k_i) , kde $k_i \in \mathbb{K}$. Dále pak nazveme **kótovanou Laguerreho mozaikou** množinu $\hat{L} = \{(C_1, k_1), \dots, (C_n, k_n)\}$. Nechť O je kompaktní množina v \mathbb{R}^3 . Zavedeme $\hat{C}_i = C_i \cap O$ pro každé $C_i \in L$. Množinu $\hat{W} = \{(\hat{C}_1, k_1), \dots, (\hat{C}_n, k_n)\}$ nazveme **pozorováním kótované mozaiky v okně**.*

1.2 Popis orientací v \mathbb{R}^3

Nechť b_{11}, b_{12}, b_{13} a b_{21}, b_{22}, b_{23} jsou ortonormální báze dvou různých souřadných systémů v \mathbb{R}^3 . Označme $\mathbf{B}_1 = (b_{11}|b_{12}|b_{13}), \mathbf{B}_2 = (b_{21}|b_{22}|b_{23})$ matice se sloupci, které jsou vektory bází.

Definice 7 (Matice orientace). *Nechť \mathbf{G} je reálná 3×3 matice taková, že*

$$\mathbf{G}^T \mathbf{G} = \mathbf{I},$$

*kde \mathbf{I} je jednotková matice. Pak \mathbf{G} nazveme **maticí orientace**.*

Definice 8 (Matice orientace mezi systémy souřadnic). \mathbf{G} nazveme **maticí orientace** systému (b_{21}, b_{22}, b_{23}) vůči (b_{11}, b_{12}, b_{13}) , pokud platí

$$\mathbf{B}_2 = \mathbf{G}\mathbf{B}_1$$

Říkáme, že systém \mathbf{B}_2 vznikl rotací systému \mathbf{B}_1 .

Poznámka. Rotace v \mathbb{R}^3 je zobrazení, které zachovává bod. Tento bod se nazývá střed rotace.

Prvek matice \mathbf{G} na pozici i, j značíme G_{ij} . Označíme-li úhel mezi b_{1j} a b_{2i} jako γ_{ij} , pak lze vyjádřit matici orientace jako

$$\mathbf{G} = \begin{pmatrix} \cos(\gamma_{11}) & \cos(\gamma_{12}) & \cos(\gamma_{13}) \\ \cos(\gamma_{21}) & \cos(\gamma_{22}) & \cos(\gamma_{23}) \\ \cos(\gamma_{31}) & \cos(\gamma_{32}) & \cos(\gamma_{33}) \end{pmatrix}$$

Věta 1. *Ke každé matici orientace \mathbf{G} existuje trojice úhlů $\varphi_1, \varphi_2 \in [0, 2\pi), \phi \in [0, \pi)$ taková, že $\mathbf{G} = \mathbf{G}_{\varphi_1} \mathbf{G}_{\phi} \mathbf{G}_{\varphi_2}$, kde*

$$\mathbf{G}_{\varphi_1} = \begin{pmatrix} \cos \varphi_1 & \sin \varphi_1 & 0 \\ -\sin \varphi_1 & \cos \varphi_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{G}_{\varphi_2} = \begin{pmatrix} \cos \varphi_2 & \sin \varphi_2 & 0 \\ -\sin \varphi_2 & \cos \varphi_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{G}_{\phi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix}$$

Navíc, pokud $G_{33} \neq 1$, potom jsou tyto úhly jednoznačné.

Důkaz. Budeme psát

$$\cos \varphi_i = c_i, \quad \sin \varphi_i = s_i, \quad \cos \phi = c, \quad \sin \phi = s$$

Pak tedy po roznásobení ortogonálních matic \mathbf{G}_{φ_1} , \mathbf{G}_ϕ , \mathbf{G}_{φ_2} dostaneme ortogonální matici

$$\mathbf{G} = \begin{pmatrix} c_1 c_2 - s_1 s_2 c & s_1 c_2 + c_1 s_2 c & s_2 s \\ -s_2 c_1 - s_1 c_2 c & -s_1 s_2 + c_1 c_2 c & c_2 s \\ s_1 s & -c_1 s & c \end{pmatrix}$$

Protože součin matic orientace zřejmě musí dát další matici orientace, každá trojice takovýchto úhlů popisuje orientaci. Navíc lze pomocí \mathbf{G} určit, že

$$\cos \phi = G_{33}$$

To jednoznačně určuje ϕ díky omezení hodnot ϕ na interval $[0, \pi)$. Dále zřejmě platí

$$\sqrt{1 - \cos^2 \phi} = \sqrt{\sin^2 \phi} = s$$

Tedy pro $G_{33} \neq 1$ lze psát

$$s_1 = \frac{G_{31}}{\sqrt{1 - c^2}}, \quad c_1 = -\frac{G_{32}}{\sqrt{1 - c^2}},$$

$$s_2 = \frac{G_{13}}{\sqrt{1 - c^2}}, \quad c_2 = \frac{G_{23}}{\sqrt{1 - c^2}},$$

což díky omezení $\varphi_1, \varphi_2 \in [0, 2\pi)$ dává jednoznačné vyjádření pro φ_1, φ_2 . Pro $c = 1$ je třeba najít jiné vyjádření, víme že to nastane pro $\phi = 0$. Můžeme psát

$$G_{11} = G_{22} = c_1 c_2 - s_1 s_2 = \cos(\varphi_1 + \varphi_2)$$

$$G_{12} = -G_{21} = s_1 c_2 + c_1 s_2 = \sin(\varphi_1 + \varphi_2)$$

Z toho plyne, že hledané úhly budou kterákoliv $\varphi_1, \varphi_2 \in [0, 2\pi)$, pro které toto platí. □

Trojice $(\varphi_1, \phi, \varphi_2)$ se nazývá **Eulerovy úhly**. Případu kde nastává nejednoznačnost vyjádření pomocí Eulerových úhlů se anglicky říká "gimbal lock" (Morawiec, 2004, str. 28). Pokud je v praktické aplikaci jednoznačnost potřeba, je možné pro tento případ zavést nějakou konvenci, která jednoznačnost zaručí, například $\varphi_1 = 0$.

Podle Eulerovy věty o rotaci (Morawiec, 2004, str. 1) lze každou rotaci v \mathbb{R}^3 vyjádřit jako rotaci která zachovává celou přímku. Ke každé rotaci existuje právě jedno takovéto vyjádření. Tuto rotaci lze popsat pomocí osy a úhlu následovně:

Definice 9 (Osa a úhel rotace). *Dvojici $(\hat{\mathbf{n}}, \omega)$, kde $\hat{\mathbf{n}} \in \mathbb{R}^3$ je jednotkový vektor a $\omega \in [-\pi, \pi)$ nazýváme **osou a úhlem rotace**.*

Vztah mezi maticí orientace a vyjádřením pomocí osy a úhlu je následovný (Karafiátová, 2017, str. 11):

$$\omega = \arccos \left(\frac{\text{tr}(G) - 1}{2} \right)$$

$$\hat{n}_1 = \frac{G_{32} - G_{23}}{2 \sin \omega}, \quad \hat{n}_2 = \frac{G_{13} - G_{31}}{2 \sin \omega}, \quad \hat{n}_3 = \frac{G_{21} - G_{12}}{2 \sin \omega}$$

Definice 10 (Misorientace). Pro dvě matice orientace $\mathbf{G}_1, \mathbf{G}_2$ definujeme **misorientaci** jako

$$\mathbf{M}_{12} = \mathbf{G}_2 \mathbf{G}_1^{-1}$$

Pro \mathbf{M}_{12} najdeme vyjádření pomocí osy a úhlu $(\hat{\mathbf{m}}(\mathbf{G}_1, \mathbf{G}_2), \alpha(\mathbf{G}_1, \mathbf{G}_2))$. Úhel $\alpha(\mathbf{G}_1, \mathbf{G}_2)$ nazveme úhlem misorientace pro orientace $\mathbf{G}_1, \mathbf{G}_2$.

Zjevně platí $\mathbf{M}_{12} = \mathbf{M}_{21}^{-1}$ a tedy $\alpha(\mathbf{G}_1, \mathbf{G}_2) = \alpha(\mathbf{G}_2, \mathbf{G}_1)$.

1.3 Markovské řetězce, obecný stavový prostor

Nechť $(\Omega, \mathcal{A}, \mathbb{P})$ je pravděpodobnostní prostor, nechť (T, \mathcal{T}) je měřitelný prostor se spočetně generovanou σ -algebrou \mathcal{T} .

Definice 11 (Markovské jádro). Měřitelné zobrazení $P : T \times \mathcal{T} \rightarrow [0, 1]$ se nazývá **markovské jádro** na (T, \mathcal{T}) , pokud platí následující:

- i. pro každé $A \in \mathcal{T}$ je $P(\cdot, A)$ měřitelná funkce na T
- ii. pro každé $x \in T$ je $P(x, \cdot)$ pravděpodobnostní míra na \mathcal{T}

Definice 12 (Markovský řetězec). Nechť P je markovské jádro na (T, \mathcal{T}) a ρ je pravděpodobnostní rozdělení na \mathcal{T} . Pak náhodná posloupnost $\mathbf{X} = \{X_n, n \in \mathbb{N}_0\}$ se nazývá **markovský řetězec** s přechodovým jádrem P a počátečním rozdělením ρ , pokud splňuje

$$\begin{aligned} \mathbb{P}(X_0 \in A_0, \dots, X_n \in A_n) &= \\ &= \int_{A_0} \cdots \int_{A_{n-1}} P(y_{n-1}, A_n) dP(y_{n-2}, y_{n-1}) \cdots dP(y_0, y_1) d\rho(y_0) \end{aligned}$$

pro všechna $n \in \mathbb{N}_0, A_0, \dots, A_n \in \mathcal{T}$.

Je vidět, že markovské jádro popisuje pravděpodobnosti přechodu homogenního řetězce \mathbf{X} , pravděpodobnost přechodu ze stavu $x \in T$ do stavu v množině $A \in \mathcal{T}$ je $P(x, A)$.

Definice 13 (Přechodové jádro n -tého řádu). Zavedeme $P^0(x, A) = \delta_x(A)$. Definujeme **přechodové jádro n -tého řádu** induktivně jako

$$P^n(x, A) = \int_T P(y, A) dP^{n-1}(x, y)$$

Definice 14 (Limitní rozdělení markovského řetězce). Nechť π je rozdělení na (T, \mathcal{T}) . Říkáme, že π je **limitním rozdělením** markovského řetězce \mathbf{X} generovaného jádrem P , pokud

$$\lim_{n \rightarrow \infty} P^n(x, A) = \pi(A)$$

pro π -s.v. $x \in T$, pro všechna $A \in \mathcal{T}$.

Definice 15 (Stacionární rozdělení). Nechť π je rozdělení na (T, \mathcal{T}) . Řekneme, že π je **stacionární**, pokud

$$\pi(A) = \int_T P(x, A) d\pi(x), \quad \forall A \in \mathcal{T}.$$

Pokud existuje limitní rozdělení \mathbf{X} , pak je i stacionární (Prokešová, 2020, Tvrzení 5.2.4).

Definice 16 (Reverzibilní řetězec). *Nechť π je rozdělení na (T, \mathcal{T}) . Řekneme, že markovský řetězec s přechodovým jádrem P je **reverzibilní** vůči π , pokud*

$$\int_A P(x, B) d\pi(x) = \int_B P(x, A) d\pi(x)$$

pro každé $A, B \in \mathcal{T}$.

Toto lze také zapsat pomocí hustot jako

$$f(x)p(x, y) = f(y)p(y, x), \quad \forall x, y \in T,$$

kde $f(x)$ je hustota π a $p(x, y)$ je hustota markovského jádra, tedy

$$P(x, A) = \int_A p(x, y) d\nu(y),$$

kde ν je σ -konečná míra na (T, \mathcal{T}) . Reverzibilita je postačující podmínka stacionarity (Prokešová, 2020, Tvrzení 5.2.5).

Definice 17 (Nerzložitelnost markovského řetězce). *Nechť ρ je pravděpodobnostní míra na \mathcal{T} . Řekneme, že markovský řetězec \mathbf{X} je ρ -**nerozložitelný**, pokud pro každé $x \in T$ a $A \in \mathcal{T}$, pro které $\rho(A) > 0$, platí $P^n(x, A) > 0$ pro nějaké $n \in \mathbb{N}$.*

Definice 18 (Periodicita markovského řetězce). *Markovský řetězec \mathbf{X} se stacionárním rozdělením π se nazývá **periodický**, pokud existuje $q \in \mathbb{N}$, $q > 1$ a neprázdné disjunktní $A_0, \dots, A_{q-1}, A_q \in \mathcal{T}$, pro které $P(x, A_{i+1}) = 1$ pro každé $x \in A_i$, $i \in \{0, \dots, q-1\}$. V opačném případě se nazývá **neperiodický**.*

Definici periodicity bez předpokladu stacionarity lze najít v Meyn a Tweedie (1993). V této práci obecná definice nebude potřeba.

Věta 2 (Limitní rozdělení markovského řetězce). *Nechť \mathbf{X} je ρ -nerozložitelný neperiodický markovský řetězec se stacionárním rozdělením π . Pak je π i jeho limitním rozdělením (Prokešová, 2020, Věta 5.2.8.).*

1.4 Metody MCMC

Nechť π je cílové rozdělení s hustotou g vůči σ -konečné míře ν na měřitelném prostoru (T, \mathcal{T}) . Nechť $M = \{x \in T : g(x) > 0\}$. Markovské Monte Carlo (MCMC) je třída algoritmů, která využívá stacionárního rozdělení markovského řetězce pro simulaci náhodného výběru z cílového rozdělení. Jedním z takových algoritmů je takzvaný Gibbsův výběrový plán. Předpokládejme, že $T = \times_{i=1}^n T_i$ je součinný prostor a \mathcal{T} je příslušná součinnová σ -algebra. Nechť (Y_1, \dots, Y_n) je náhodný vektor s hodnotami v T a rozdělením π .

Algoritmus 1 (Gibbsův výběrový plán). *Nechť K je zvolený maximální počet kroků algoritmu.*

1. Zvolme počáteční stav $(y_1^{(0)}, \dots, y_n^{(0)}) \in M$, dále necht $t=0$
2. Simulujme $y_1^{(t+1)}$ z rozdělení $(Y_1|Y_2 = y_2^{(t)}, \dots, Y_n = y_n^{(t)})$
 Simulujme $y_2^{(t+1)}$ z rozdělení $(Y_2|Y_1 = y_1^{(t+1)}, Y_3 = y_3^{(t)}, \dots, Y_n = y_n^{(t)})$
 \vdots
 Simulujme $y_n^{(t+1)}$ z rozdělení $(Y_n|Y_1 = y_1^{(t+1)}, \dots, Y_{n-1} = y_{n-1}^{(t+1)})$
3. Pokud $t + 1 < K$, pak $t := t + 1$ a jdeme na 2. krok, jinak ukončíme algoritmus

Věta 3 (Stacionární rozdělení Gibbsova výběrového plánu). *Necht*

$$(\forall i \in \{1, \dots, n\} : g_i(x_i) > 0) \Rightarrow g(x) > 0, g_i : T_i \rightarrow R_+$$

kde g_i jsou marginální hustoty g a $x = (x_1, \dots, x_n)$. Potom π je stacionární rozdělení markovského řetězce generovaného kroky Algoritmu 1.

Věta 4 (Nerozložitelnost a neperiodicita Gibbsova výběrového plánu). *Necht* $M = \times_{i=1}^n M_i$ je součinnový prostor a $\nu = \otimes_{i=1}^n \nu_i$ je součinnová míra, $\nu_i(M_i) > 0$. Pak markovský řetězec generovaný Algoritmem 1 je ν -nerozložitelný a neperiodický.

Důkazy vět 3 a 4 viz (Prokešová, 2020, věty 4.1.3 a 5.3.1.). Z vět 2, 3, 4 plyne konvergence markovského řetězce z algoritmu 1 k limitnímu stacionárnímu rozdělení π .

Gibbsův výběrový plán předpokládá schopnost simulovat náhodný výběr ze všech plně podmíněných rozdělení. Pro případy, kdy to nejde, lze algoritmus modifikovat. Potřebujeme k tomu zavést následující pojem.

Definice 19 (Pravděpodobnost přijetí). *Necht* (S, \mathcal{S}, χ) je prostor s mírou a P je markovské jádro na S . Necht f je hustota cílového rozdělení na S . Necht $p(x, y)$ je hustota P vůči χ pro každé $x \in S$ a necht $P(x, \hat{M}) = 1$ pro $x \notin \hat{M}$, kde $\hat{M} = \{x \in S : f(x) > 0\}$. Hustotu p nazýváme **návrhovou hustotou**. Definujeme **pravděpodobnost přijetí** návrhu y , vycházíme-li ze stavu x , jako

$$H(x, y) = \begin{cases} \min \left\{ \frac{f(y)p(y,x)}{f(x)p(x,y)}, 1 \right\} \dots f(x)p(x,y) > 0 \\ 1 \dots \text{jinak} \end{cases} \quad (1.1)$$

Pravděpodobnost přijetí se využívá v tzv. Metropolisově-Hastingsově algoritmu (Prokešová, 2020, Algoritmus 4.2.1). My budeme využívat hybridní algoritmus, kde se krok Metropolisova-Hastingsova algoritmu vždy provede uvnitř kroku Gibbsova výběrového plánu.

Položme v definici 19 $S = T_i$, $P = Q$, $\chi = \nu_i$, $f = g_{i|i}$ postupně pro všechna $i = 1, \dots, n$. Zde $g_{i|i}$ značí hustotu podmíněného rozdělení $(Y_i|Y_1, \dots, Y_{i-1}, Y_{i+1}, \dots, Y_n)$. Dále označme $H_i(y_1, \dots, y_n, w) = H(y_i, w)$ pravděpodobnost přijetí návrhu w pro cílovou hustotu $g_{i|i}$.

Algoritmus 2 (Hybridní algoritmus). *Necht* K je zvolený maximální počet iterací algoritmu

1. Zvolme počáteční stav $(y_1^{(0)}, \dots, y_n^{(0)}) \in T$, $t := 0$
2. Vygenerujeme w z rozdělení $Q(y_1^{(t)}, \cdot)$ a s pravděpodobností $H_1(y_1^{(t)}, \dots, y_n^{(t)}, w)$ položíme $y_1^{(t+1)} := w$, jinak $y_1^{(t+1)} := y_1^{(t)}$
 Vygenerujeme w z rozdělení $Q(y_2^{(t)}, \cdot)$ a s pravděpodobností $H_2(y_1^{(t+1)}, y_2^{(t)}, \dots, y_n^{(t)}, w)$ položíme $y_2^{(t+1)} := w$, jinak $y_2^{(t+1)} := y_2^{(t)}$
 \vdots
 Vygenerujeme w z rozdělení $Q(y_n^{(t)}, \cdot)$ a s pravděpodobností $H_3(y_1^{(t+1)}, \dots, y_{n-1}^{(t+1)}, w)$ položíme $y_n^{(t+1)} := w$, jinak $y_n^{(t+1)} := y_n^{(t)}$
3. Pokud $t < K$, tak $t := t + 1$ a vracíme se na 2. krok, jinak ukončíme algoritmus.

Z článku Jones a kol. (2014) plyne, že konvergence základních algoritmů se přenáší i na hybridní algoritmus.

2. Stochastický model

V této kapitole zavedeme model pro náhodné orientace buněk v mozaice. Představíme kótovanou mozaiku, kde jako kóty použijeme trojice Eulerových úhlů. Tato mozaika bude deterministická, kóty budou náhodné. Sdružené rozdělení náhodných trojic Eulerových úhlů je dáno hustotou pravděpodobnosti f Gibbsova typu, kde funkce energie závisí na misorientaci sousedních buněk v mozaice. V druhé části kapitoly ověříme, že toto sdružené rozdělení lze simulovat pomocí MCMC a vyjádříme pravděpodobnost přijetí pro algoritmus 2.

2.1 Model hustoty orientací buněk

Nejprve je potřeba zavést model který budeme používat. Nechť (Ω, \mathcal{A}, P) je pravděpodobnostní prostor. Nechť $L = \{C_1, \dots, C_n\}$ je Laguerreho mozaika. Nechť

$$S = \{(\varphi_1, \phi, \varphi_2), \varphi_1, \varphi_2 \in [0, 2\pi), \phi \in [0, \phi)\} \quad (2.1)$$

a nechť \mathcal{S} je borelovská σ -algebra na S . Na (S, \mathcal{S}) dále definujeme míru μ pomocí předpisu

$$\mu(A) = \int_A |\cos \phi| d\phi d\varphi_1 d\varphi_2, \quad A \in \mathcal{S}$$

Zavedeme náhodný vektor orientací buněk jako

$$\mathbf{X} : (\Omega, \mathcal{A}) \rightarrow (S^n, \mathcal{S}^{\otimes n})$$

Tento náhodný vektor má rozdělení π s hustotou Gibbsova typu (Seitl a kol., 2020)

$$f(x_1, \dots, x_n) \propto \exp(\theta \varepsilon(x_1, \dots, x_n)), \quad (x_1, \dots, x_n) \in S^n \quad (2.2)$$

vzhledem k součinné míře $\mu^{\otimes n}$. Zde $\theta \in \mathbb{R}$ je libovolně volený parametr a ε je funkce energie. Pokud $\theta \geq 0$,

$$\varepsilon(x_1, \dots, x_n) = \sum_{j,i: C_i \sim C_j} \sin(\alpha(x_i, x_j)), \quad (2.3)$$

a pokud $\theta < 0$

$$\varepsilon(x_1, \dots, x_n) = \sum_{j,i: C_i \sim C_j} \sin(\alpha(x_i, x_j)) + \sum_{j,i: C_i \sim C_j} V_{hard}(C_i, C_j), \quad (2.4)$$

kde

$$V_{hard}(C_i, C_j) = \begin{cases} +\infty & \dots \alpha(C_i, C_j) < q, \\ 0 & \dots \alpha(C_i, C_j) \geq q \end{cases}$$

pro zvolenou minimální hodnotu misorientace q . Členu V_{hard} se říká **hardcore podmínka**.

Poznámka. Ač se Gibbsovo rozdělení a Gibbsův výběrový plán oba jmenují podle J. W. Gibbse, jde o dva zcela odlišné pojmy.

2.2 MCMC pro výběr z hustoty orientací

Protože nelze náhodný výběr z rozdělení π simulovat přímo, použijeme metodu MCMC. Rozdělení π bude cílovým rozdělením.

Věta 5. Pro hustotu (2.2) jsou předpoklady vět 3 a 4 splněny.

Důkaz. Protože f je ve tvaru exponenciály, zřejmě

$$f(x_1, \dots, x_n) > 0, \quad \forall (x_1, \dots, x_n) \in S^n.$$

Tím pádem jsou předpoklady věty 3 splněny. Zároveň potom $M = S^n$, je tedy v součinném tvaru, z definice je jasné

$$\nu_i(M_i) = \mu(M_i) = \mu(S) = 1 > 0$$

a jsou tedy splněny i předpoklady věty 4. □

Důsledek. Je-li π rozdělením náhodného vektoru \mathbf{X} s hustotou f , pak z vět 2 a 5 plyne, že π je limitním rozdělením MCMC algoritmu.

Použití Gibbsova výběrového plánu předpokládá možnost simulovat ze všech plně podmíněných rozdělení. Ověříme jak v případě (2.3) vypadají. Zavedeme

$$x_{-i} := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

Platí

$$\begin{aligned} f(x_1, \dots, x_n) &= \frac{\exp\{\theta\varepsilon(x_1, \dots, x_n)\}}{\int_{\mathbb{R}^n} \exp\{\theta\varepsilon(x_1, \dots, x_n)\} d(x_1, \dots, x_n)} \\ &= \frac{1}{c_{1:n}(\theta)} \exp\{\theta\varepsilon(x_1, \dots, x_n)\}, \\ f(x_{-i}) &:= \frac{\exp\{\theta\varepsilon(x_{-i})\}}{\int_{\mathbb{R}^{n-1}} \exp\{\theta\varepsilon(x_{-i})\} d(x_{-i})} = \frac{1}{c_{-i}(\theta)} \exp\{\theta\varepsilon(x_{-i})\}, \end{aligned}$$

kde $c_{1:n}$ a c_{-i} jsou příslušné konstanty, aby $f(x_1, \dots, x_n)$, $f(x_{-i})$ byly pravděpodobnostní hustoty. Pro hustotu podmíněného rozdělení $(X_i | X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ tedy platí

$$f_{i|i}(x_i | x_{-i}) = \frac{c_{-i}(\theta) \exp\{\theta\varepsilon(x_1, \dots, x_n)\}}{c_{1:n}(\theta) \exp\{\theta\varepsilon(x_{-i})\}} = c(\theta) \prod_{j: C_i \sim C_j} \exp\{\theta \sin(\alpha_{ij})\},$$

kde c je příslušná konstanta. Protože i tato rozdělení je obtížné simulovat, použijeme algoritmus 2. Zvolme návrhovou hustotu $p(x, y) = I[y \in S^n]$. Potom návrh hodnoty bude z rovnoměrného rozdělení na S^n . Pravděpodobnost přijetí má hodnotu

$$H_i(x, y) = \begin{cases} \frac{f_{i|i}(y|x_{-i})}{f_{i|i}(x_i|x_{-i})} \cdots f_{i|i}(x_i|x_{-i}) > 0 \\ 1 \dots \text{jinak} \end{cases} \quad (2.5)$$

Zřejmě platí

$$\frac{f_{i|i}(y|x_{-i})}{f_{i|i}(x_i|x_{-i})} = \exp \left\{ \theta \sum_{j: C_i \sim C_j} (\sin(\alpha(y, x_j)) - \sin(\alpha(x_i, x_j))) \right\}$$

To vede na tento speciální případ algoritmu 2:

Algoritmus 3. *Nechť K je zvolený maximální počet iterací algoritmu*

1. *Vygenerujeme počáteční stav $(x_1^{(0)}, \dots, x_n^{(0)})$ z rovnoměrného rozdělení na S^n ,
 $t := 0$*
2. *Vygenerujeme y z rovnoměrného rozdělení na S a s pravděpodobností
 $H_1(x_1^{(t)}, \dots, x_n^{(t)}, y)$ položíme $x_1^{(t+1)} := y$, jinak $x_1^{(t+1)} := x_1^{(t)}$
Vygenerujeme y z rovnoměrného rozdělení na S a s pravděpodobností
 $H_2(x_1^{(t+1)}, x_2^{(t)}, \dots, x_n^{(t)}, y)$ položíme $x_2^{(t+1)} := y$, jinak $x_2^{(t+1)} := x_2^{(t)}$
 \vdots
Vygenerujeme y z rovnoměrného rozdělení na S a s pravděpodobností
 $H_3(x_1^{(t+1)}, \dots, x_{n-1}^{(t+1)}, y)$ položíme $x_n^{(t+1)} := y$, jinak $x_n^{(t+1)} := x_n^{(t)}$*
3. *Pokud $t < K$, tak $t := t + 1$ a vracíme se na 2. krok, jinak ukončíme algoritmus.*

Pro případ (2.4) se navíc zamítají návrhy, pro které vyjde nějaká misorientace sousedních buněk menší než q .

3. Simulační studie

3.1 Krystalografická interpretace

Polykrystalické materiály ve fyzice kovů mají mikrostrukturu mozaik, jejichž buňky se nazývají zrna. Každému zrnu je přiřazena orientace jeho atomové mřížky vůči souřadnému systému vzorku. V této práci uvažujeme krychlovou mřížku, jejíž orientace popisujeme Eulerovými úhly. Tento popis je ovšem nejednoznačný vůči symetriím mřížky, kterých je v tomto případě 24. Různé trojice Eulerových úhlů tak mohou reprezentovat stejnou orientaci mřížky.

Cílem simulačních studií je ukázat tvar rozdělení misorientací ve formě histogramu, získaného ze simulací z hustoty Eulerových úhlů $f(x)$ (viz vzorec (2.2)) a následného výpočtu nejmenší misorientace přes všechny symetrie mřížek každých dvou sousedních zrn. V literatuře se pro tuto nejmenší misorientaci někdy zavádí pojem disorientace, my ale tento pojem neužíváme. Simulace jsou provedeny přes pevnou vstupní mozaiku.

Zaměříme se na extrémní případy možných rozdělení misorientací ve smyslu polohy histogramu. Intuitivně lze očekávat, že hodnota parametru θ bude mít výrazný vliv na tato rozdělení. Pro velké záporné hodnoty budou misorientace na dolní mezi dané hardcore podmínkou. Pro hodnoty velkého kladného parametru se naopak budou posouvat k maximální hodnotě misorientace, která je vzhledem ke krychlové mřížce zhruba 1.08 rad (Morawiec, 2004). Tvar tohoto histogramu jsme v literatuře nenalezli, přitom byl hlavní motivací práce. Výsledky simulací jsou diskutovány v sekci 3.3.

Další úkol je zabývat se tím, nakolik prováděné simulace závisí na vstupní mozaice. Toto je diskutováno v sekci 3.4 pro speciální případ mozaik s velkým resp. malým počtem stěn v zrnech.

3.2 Výpočetní programy

Potřebný program jsem napsal v jazyce C++. Soubory `source.cpp`, `initialize.cpp`, `initialize.h` jsem napsal sám, soubor `misorientation.cpp` jsem převzal z článku Pawlas a kol. (2020) a zbytek souborů je z otevřené knihovny Voro++ (Rycroft, 2009), která je určená na generování Voronoiho mozaik a práci s nimi.

`initialize.cpp` obsahuje dvě možné inicializační procedury. `init` vygeneruje sama Voronoiho mozaiku v s příslušným počtem buněk zatímco `init2` mozaiku načte z přiloženého souboru `input.txt`. Ten musí obsahovat řádek pro každou buňku v mozaice, na tomto řádku musí být nejprve číslo řádku, které slouží jako identifikátor buňky, dále počet jejích sousedů a identifikátory těchto sousedů. Mezi těmito inicializačními procedurami se přepíná tak, že se přepíšou na příslušném místě v kódu. Obě inicializační procedury dále každé buňce přiřadí

náhodnou trojici Eulerových úhlů a pomocí funkce `get_misorientation`, která je převzata z `misorientation.cpp`, se spočítají nejmenší misorientace mezi dvojicemi sousedních buněk. Tyto informace jsou uloženy do matice, kde každý řádek reprezentuje sousedy dané buňky, u každého souseda je uloženo jeho identifikační číslo a příslušná misorientace. V prvním sloupci matice je navíc ke každé buňce uložena její orientace v Eulerových úhlech.

Tato matice obsahuje všechny informace potřebné pro průběh algoritmu 3. Po uplynutí daného počtu iterací jsou výsledky vypsány do souboru `output.txt`. Tento soubor tedy v každém řádku obsahuje misorientace mezi sousedními buňkami v dané iteraci. Program dále vypisuje do konzole číslo právě prováděné úpravy orientace buňky, aby bylo patrné že simulace probíhají bez problému. Na konci do konzole vypíše kolik návrhů z algoritmu 3 bylo přijato.

Vstupní parametry lze upravit modifikací konstant ve zdrojovém kódu, `c` je počet buněk, `theta` je hodnota parametru θ , `q` je minimální hodnota misorientace pro hardcore podmínku, `maxIt` je maximální počet iterací algoritmu a `outstep` je po kolika iteracích program vypíše současné misorientace do souboru.

Doba průběhu programu roste s počtem iterací a s počtem buněk v mozaice. S počtem iterací roste lineárně, zvyšování počtu buněk zvedá počet sousedících buněk rychleji.

3.3 Výsledky, závislost na θ

Misorientace náhodných dvojic orientací s kubickými symetriemi mají takzvané Mackenzieho rozdělení. Pro hustotu tohoto rozdělení je charakteristický modus přibližně 0.8, přičemž na intervalu kde je rostoucí je konvexní a na intervalu kde je klesající je konkávní. Střední hodnota tohoto rozdělení je přibližně 0.710 (Morawiec, 2004, str.121). Uvažujeme-li Voronoiho mozaiku, pak za předpokladu, že orientace buněk budou voleny rovnoměrně náhodně, se ukazuje, že misorientace aproximují Mackenzieho rozdělení velice dobře. Příklad rozdělení misorientací 1000 buněk je vidět na obrázku 3.1. Toto rozdělení je počátečním rozdělením pro markovský řetězec generovaný algoritmem 3.

Konvergenci algoritmu můžeme zkoumat pomocí výběrového průměru a výběrového rozptylu, viz tabulky 3.1 a 3.2. Dalším důležitým měřítkem je takzvaná diskrepance histogramu (Seitl a kol., 2020). Ta je definovaná jako

$$d = \sum_{i=1}^k \left| \frac{l_i}{N_1} - \frac{m_i}{N_2} \right|$$

kde k je počet tříd histogramu a l_i , m_i jsou četnosti prvků v dané třídě pro porovnávané histogramy a N_1 , N_2 jsou celkové počty prvků v jednotlivých histogramech. Zřejmě platí $d \in [0,2]$. V této práci jsou diskrepance počítané pro histogram o deseti rovnoměrných třídách na oboru hodnot misorientace, to nutně nemusí přesně odpovídat obrázkům. Výsledné diskrepance jsou v tabulce 3.3.

Další zajímavý výsledek, který dostaneme ze simulace je takzvaný poměr přijetí p_a , definovaný jako poměr přijatých návrhů a celkových návrhů v chodu algoritmu 3. Tyto poměry vidíme v tabulce 3.4.

3.3.1 $\theta = 0$

Zřejmě pro volbu $\theta = 0$ bude v průběhu algoritmu každý návrh přijat. Protože návrhy jsou brány z rovnoměrného rozdělení orientací, v každé iteraci algoritmu dostaneme výběr z Mackenzieho rozdělení a toto rozdělení tedy bude i cílovým rozdělením.

3.3.2 $\theta > 0$

Podívejme se na $\theta = 20$. Z obrázku 3.2 je vidět, že už po prvních 20000 iteracích se histogram značně změní, rozdělení se posune směrem k maximální hodnotě misorientace. Napříč dalším průběhem simulace se již histogram výrazně nemění, je to patrné z diskrepance, viz tabulka 3.3. Z tabulek 3.1 a 3.2 vidíme, že i výběrový průměr a rozptyl se nejvíce změní během prvních 20000 iterací a pak zůstávají víceméně stabilní. Zdá se tedy, že během prvních 20000 iterací rozdělení zkonverguje blízko k limitnímu.

Pro $\theta = 10$ je situace velice podobná, evidentní rozdíl je, že se rozdělení méně liší od Mackenzieho rozdělení. Lze tak usoudit z tvaru histogramu (obrázek 3.3), menší diskrepance na prvních 20000 iterací a také z toho že výběrový průměr je menší (tabulka 3.1).

Pro $\theta = 1$ je situace složitější v tom, že diskrepance už od začátku osciluje, nedojde k tak výraznému zmenšení (tabulka 3.3). Histogramy také vypadají velice podobně, viz obrázek 3.4. Rozdělení je pro tuto hodnotu pravděpodobně tak blízko Mackenzieho rozdělení, že pomocí diskrepance se toho nedá mnoho zjistit. Je ovšem vidět z tabulky 3.1, že výběrový průměr se už během prvních 20000 iterací lehce zvýšil a zůstal tak v každém dalším měření.

Z těchto tří příkladů lze usoudit, že pro $\theta > 0$ je střední hodnota výsledného rozdělení vyšší než u Mackenzieho rozdělení, a že čím větší θ je, tím větší je tento rozdíl. I tvarově se víc odlišují rozdělení s větším θ . Rozdíly v poměru přijetí jsou způsobeny tím, že při pravděpodobnosti přijetí 1 se jedná o Mackenzieho rozdělení, čili zřejmě čím je výsledné rozdělení dál od Mackenzieho, tím menší pravděpodobnost přijetí bude mít.

3.3.3 $\theta < 0$

Nyní se podíváme na příklad pro záporné θ . Zde musíme zavést minimální misorientaci pro hardcore podmínku, nechť $q = 0.018$. Podívejme se na $\theta = -20$. Je vidět (tabulka 3.1, obrázek 3.5), že se rozdělení velice rychle posune k dolní hranici misorientace, kde zůstane a téměř se nemění.

Pro $\theta = 10$ vidíme z obrázku 3.6, že se hodnoty opět velice rychle posunou k minimální hodnotě. Ze středních hodnot (tabulka 3.1) je ale vidět, že se rozdělení se drží na vyšší střední hodnotě než v případě pro $\theta = 20$. Navíc je z porovnání s obrázkem 3.5 vidět jiný tvar rozdělení

Pro $\theta = -1$ se rozdělení téměř měnit nebude, bude oscilovat velice blízko Mackenzieho rozdělení (obrázek 3.7, tabulka 3.3). Z tabulky 3.1 je ale vidět, že střední hodnota se lehce zmenší a zůstane tak.

V tabulce 3.4 vidíme poměry přijetí pro záporná θ . Opět je vidět, že pro rozdělení, která jsou dál od Mackenzieho rozdělení je pravděpodobnost přijetí menší. Porovnáme-li tyto výsledky a výsledky z tabulek 3.1, 3.3, zdá se, že pro záporné hodnoty θ se rozdělení liší od Mackenzieho víc než pro kladné θ se stejnou absolutní hodnotou.

3.4 Výsledky, závislost na počtu sousedů

Rozdělení misorientací ve Voronoiho mozaice s rovnoměrně náhodnými orientacemi Mackenzieho rozdělení pouze aproximuje, neboť existují závislosti mezi misorientacemi. Otázka je jak je toto rozdělení ovlivněno volbou jiné Laguerreho mozaiky, protože to se promítne do jiného průměrného počtu stěn a tedy i do jiných závislostí mezi misorientacemi. V této sekci používám dva speciální případy Laguerreho mozaiky převzaté z článku Stoyan a kol. (2020). Jedna má 366 buněk a průměrný počet sousedů 12, druhá 347 buněk a průměrný počet sousedů 18.

Pro mozaiku s průměrným počtem sousedů vidíme z obrázku 3.8 a tabulek 3.5, 3.6 velice málo rozdílů oproti Voronoiho mozaice se stejným θ .

Podíváme-li se na mozaiku s průměrným počtem stěn 18, z obrázku 3.9 a tabulek 3.5, 3.6 opět není příliš mnoho změn.

Nakonec se můžeme podívat na diskrepance mezi těmito dvěma mozaikami a Voronoiho mozaikou, tabulka 3.8. I zde je vidět, že rozdíly nejsou příliš velké. Avšak zdá se, že mezi mozaikou s průměrným počtem 18 sousedů a Voronoiho mozaikou se drží konzistentní malý rozdíl. Dá se tedy tvrdit, že rozdělení je na průměrném počtu sousedů závislé, ale jen málo.

3.5 Obrázky a tabulky

| iterace | 0 | 20000 | 40000 | 60000 | 80000 | 100000 |
|----------|-------|-------|-------|-------|-------|--------|
| θ | | | | | | |
| 20 | 0.703 | 0.875 | 0.877 | 0.886 | 0.880 | 0.881 |
| 10 | 0.713 | 0.826 | 0.835 | 0.826 | 0.828 | 0.834 |
| 1 | 0.716 | 0.728 | 0.733 | 0.748 | 0.746 | 0.732 |
| -1 | 0.707 | 0.670 | 0.671 | 0.642 | 0.665 | 0.662 |
| -10 | 0.705 | 0.078 | 0.086 | 0.074 | 0.070 | 0.066 |
| -20 | 0.710 | 0.068 | 0.060 | 0.062 | 0.061 | 0.060 |

Tabulka 3.1: V prvním sloupci jsou hodnoty parametru θ . Na prvním řádku je počet iterací. Ve vnitřku tabulky jsou výběrové průměry misorientací pro příslušnou hodnotu θ a počet iterací.

| iterace | 0 | 20000 | 40000 | 60000 | 80000 | 100000 |
|----------|-------|-------|-------|-------|-------|--------|
| θ | | | | | | |
| 20 | 0.038 | 0.015 | 0.016 | 0.015 | 0.016 | 0.017 |
| 10 | 0.037 | 0.022 | 0.020 | 0.024 | 0.016 | 0.020 |
| 1 | 0.037 | 0.031 | 0.035 | 0.034 | 0.035 | 0.034 |
| -1 | 0.045 | 0.040 | 0.043 | 0.043 | 0.045 | 0.042 |
| -10 | 0.041 | 0.001 | 0.003 | 0.001 | 0.001 | 0.001 |
| -20 | 0.040 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

Tabulka 3.2: V prvním sloupci jsou hodnoty parametru θ . Na prvním řádku je počet iterací. Ve vnitřku tabulky jsou výběrové rozptyly misorientací pro příslušnou hodnotu θ a počet iterací.

| iterace po tisících θ | 0-20 | 20-40 | 40-60 | 60-80 | 80-100 |
|---------------------------------|-------|-------|-------|-------|--------|
| 20 | 0.724 | 0.068 | 0.086 | 0.083 | 0.079 |
| 10 | 0.460 | 0.128 | 0.147 | 0.128 | 0.113 |
| 1 | 0.217 | 0.140 | 0.143 | 0.088 | 0.121 |
| -1 | 0.237 | 0.101 | 0.101 | 0.179 | 0.195 |
| -10 | 1.964 | 0.155 | 0.152 | 0.072 | 0.126 |
| -20 | 1.980 | 0.291 | 0.046 | 0.046 | 0.084 |

Tabulka 3.3: V prvním sloupci jsou hodnoty parametru θ . Na prvním řádku jsou rozmezí iterací. Ve vnitřku tabulky jsou diskrepance mezi příslušnými iteracemi pro příslušné θ .

| θ | 20 | 10 | 1 | -1 | -10 | -20 |
|----------|-------|------|------|------|-------|--------|
| p_a | 0.006 | 0.05 | 0.76 | 0.67 | 0.001 | 0.0005 |

Tabulka 3.4: V prvním řádku jsou hodnoty parametru θ , ve druhém poměry přijetí pro příslušné θ .

| iterace | 0 | 20000 | 40000 | 60000 | 80000 | 100000 |
|---------------------|-------|-------|-------|-------|-------|--------|
| prům. počet sousedů | | | | | | |
| 12 | 0.708 | 0.871 | 0.870 | 0.870 | 0.874 | 0.871 |
| 18 | 0.710 | 0.863 | 0.864 | 0.864 | 0.863 | 0.867 |

Tabulka 3.5: V prvním sloupci jsou průměrné počty stěn. Na prvním řádku je počet iterací. Ve vnitřku tabulky jsou výběrové průměry misorientací pro příslušný počet sousedů a počet iterací. Pro $\theta = 20$.

| iterace | 0 | 20000 | 40000 | 60000 | 80000 | 100000 |
|---------------------|-------|-------|-------|-------|-------|--------|
| prům. počet sousedů | | | | | | |
| 12 | 0.039 | 0.019 | 0.019 | 0.018 | 0.017 | 0.018 |
| 18 | 0.039 | 0.025 | 0.023 | 0.023 | 0.023 | 0.023 |

Tabulka 3.6: V prvním sloupci jsou průměrné počty sousedů. Na prvním řádku je počet iterací. Ve vnitřku tabulky jsou výběrové rozptyly misorientací pro příslušný počet sousedů a počet iterací. Pro $\theta = 20$.

| iterace po tisících | 0-20 | 20-40 | 40-60 | 60-80 | 80-100 |
|---------------------|-------|-------|-------|-------|--------|
| prům. počet sousedů | | | | | |
| 12 | 0.698 | 0.074 | 0.046 | 0.078 | 0.105 |
| 18 | 0.741 | 0.074 | 0.060 | 0.038 | 0.055 |

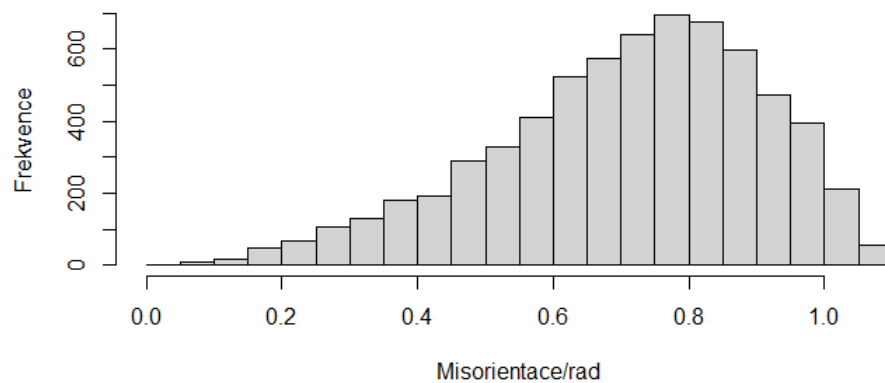
Tabulka 3.7: V prvním sloupci jsou průměrné počty sousedů. Na prvním řádku jsou rozmezí iterací. Ve vnitřku tabulky jsou diskrepance mezi příslušnými iteracemi pro příslušný počet sousedů. Pro $\theta = 20$.

| iterace po tisících | 0 | 20 | 40 | 60 | 80 |
|---------------------|-------|-------|-------|-------|-------|
| porovnávané mozaiky | | | | | |
| 12 sous./18 sous. | 0.059 | 0.099 | 0.139 | 0.115 | 0.071 |
| 12 sous./Voro. | 0.106 | 0.064 | 0.070 | 0.104 | 0.078 |
| 18 sous./Voro. | 0.105 | 0.125 | 0.120 | 0.125 | 0.130 |

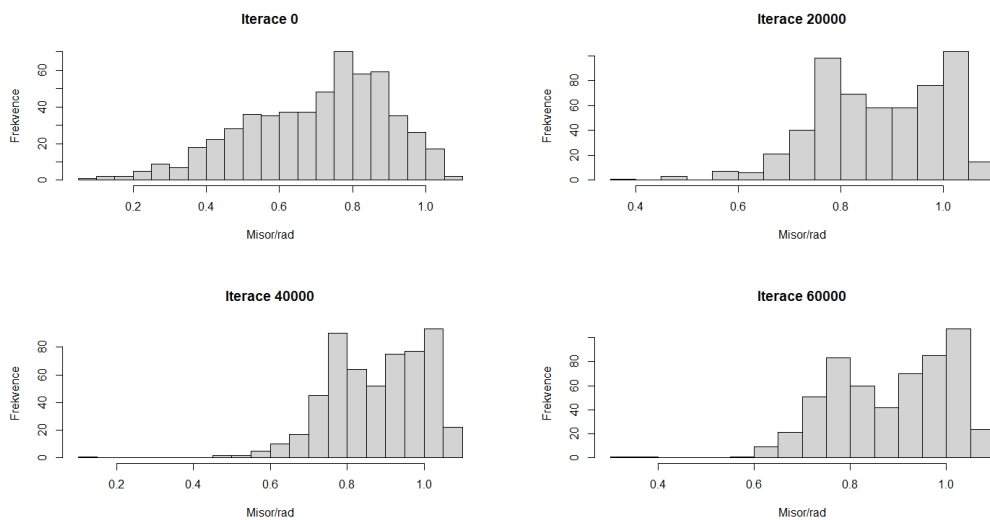
Tabulka 3.8: Diskrepance mezi histogramy pro mozaiky s průměrným počtem 12 a 18 sousedů mezi sebou a s Voronoiho mozaikou z 3.3. V prvním řádku jsou uvedeny iterace. Pro $\theta = 20$.

| prům. počet sousedů | 12 | 18 |
|---------------------|-------|-------|
| p_a | 0.013 | 0.003 |

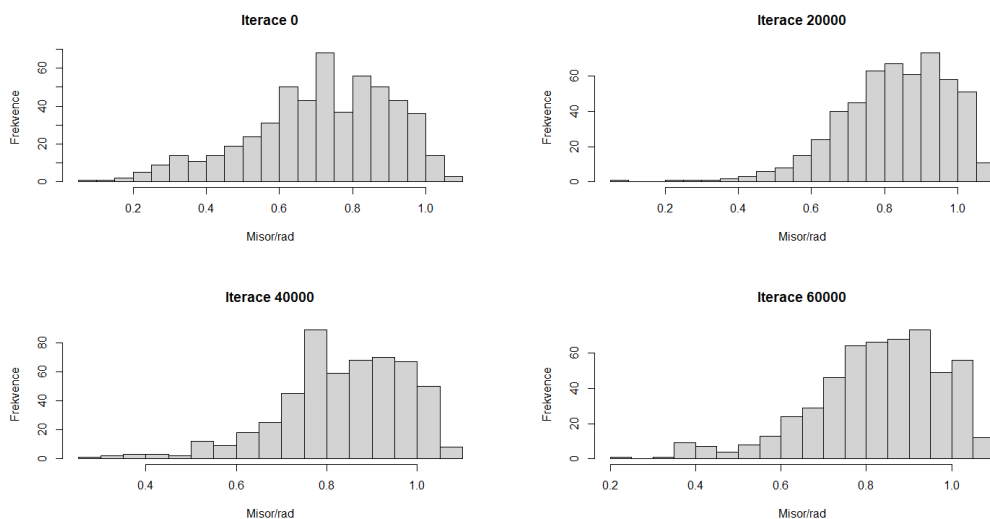
Tabulka 3.9: V prvním řádku jsou průměrné počty sousedů, ve druhém poměry přijetí pro příslušný počet sousedů. Pro $\theta = 20$.



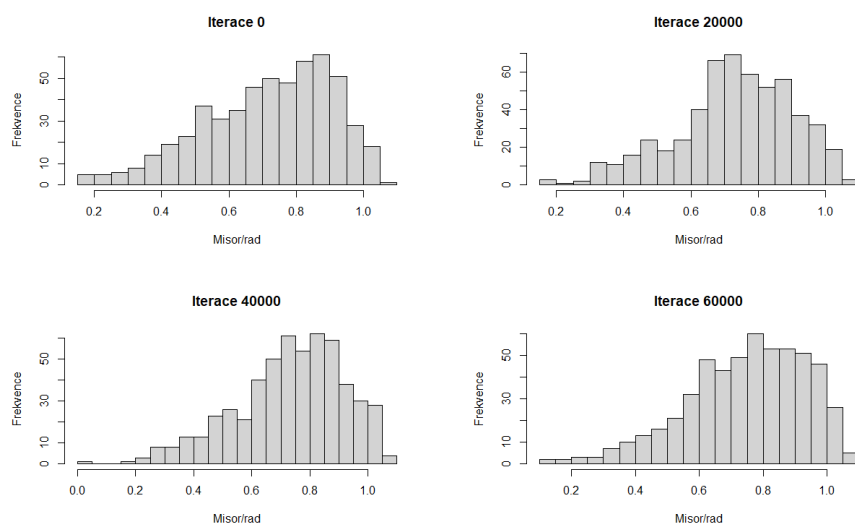
Obrázek 3.1: Aproximace Mackenzieho rozdělení pro 1000 buněk



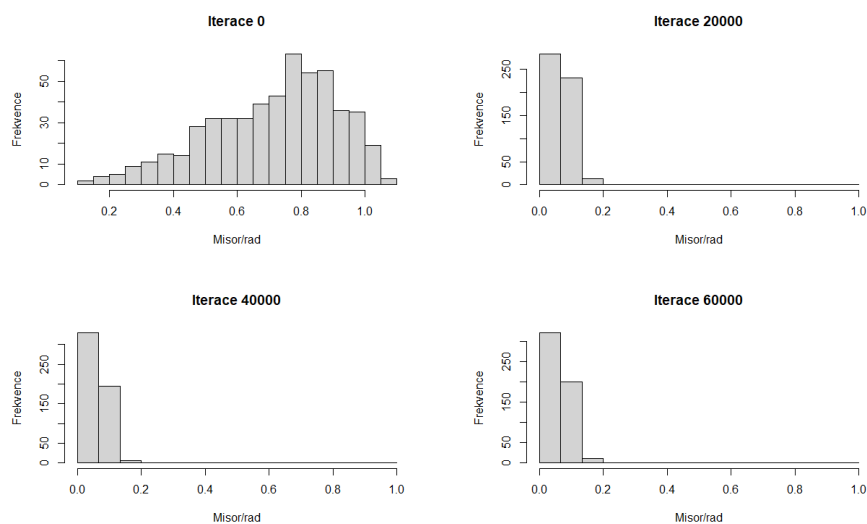
Obrázek 3.2: Hodnota parametru $\theta = 20$. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



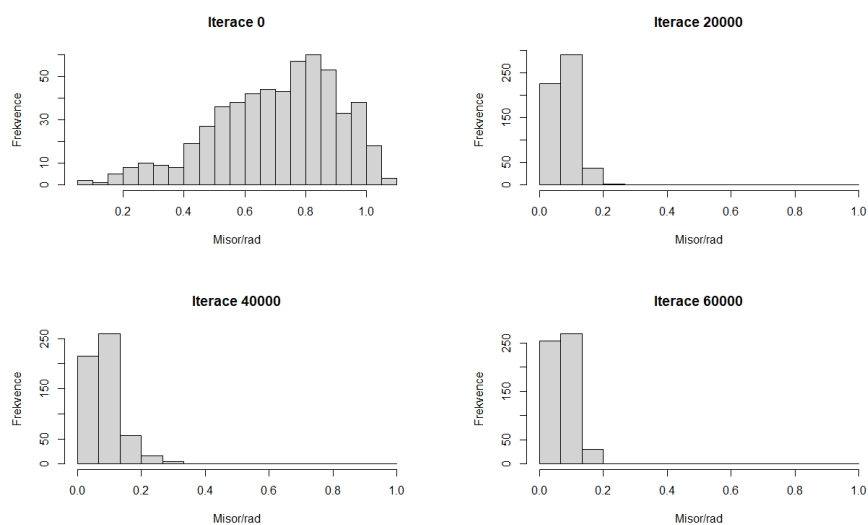
Obrázek 3.3: Hodnota parametru $\theta = 10$. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



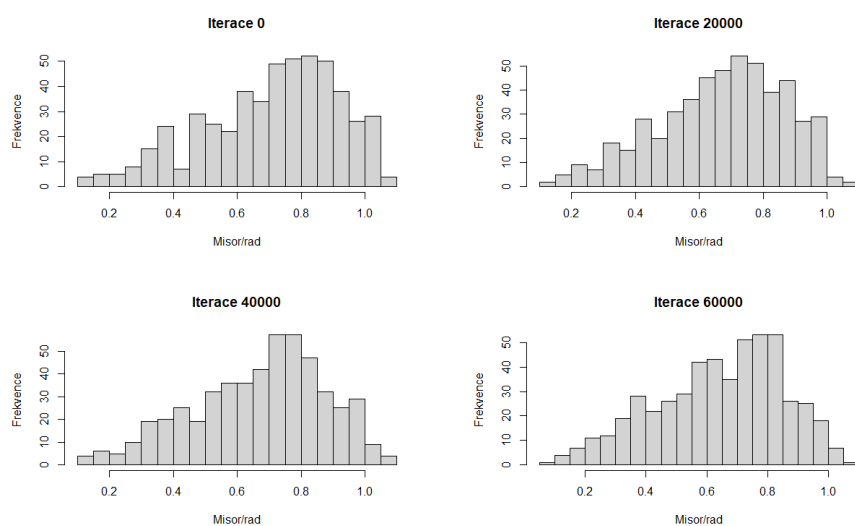
Obrázek 3.4: Hodnota parametru $\theta = 1$. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



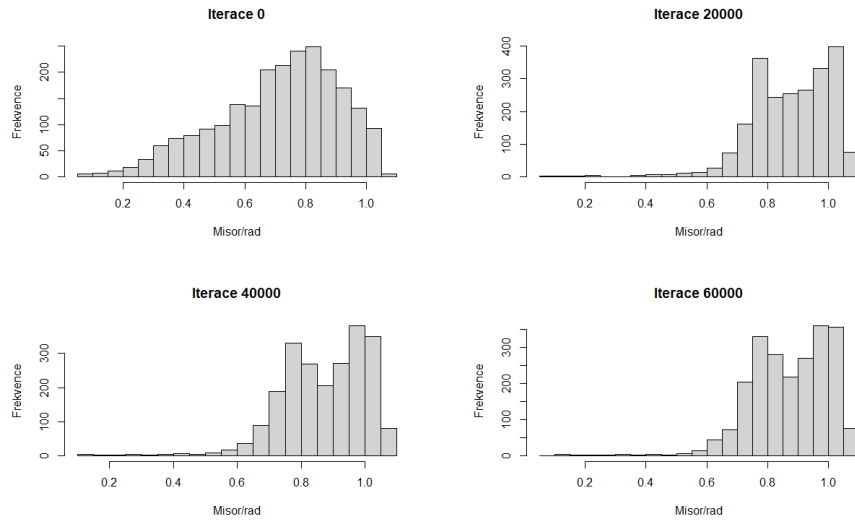
Obrázek 3.5: Hodnota parametru $\theta = -20$. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



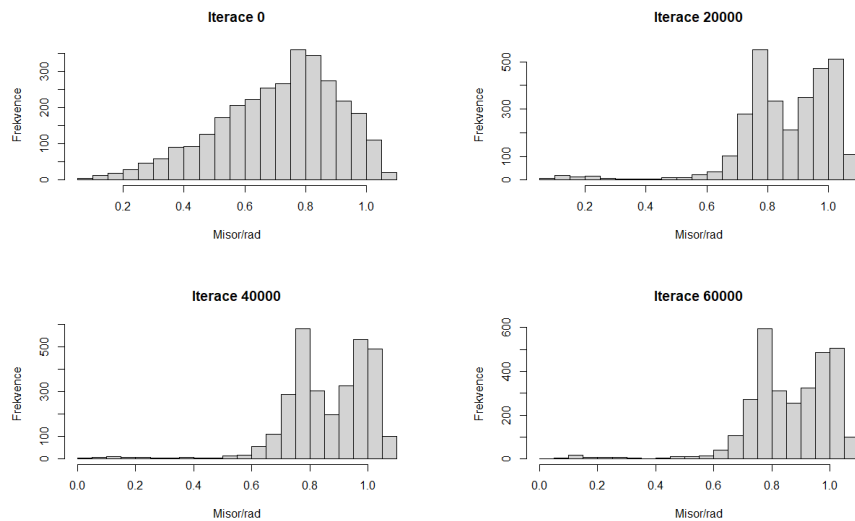
Obrázek 3.6: Hodnota parametru $\theta = -10$. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



Obrázek 3.7: Hodnota parametru $\theta = -1$. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



Obrázek 3.8: Hodnota parametru $\theta = 20$, průměrný počet sousedů je 12. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.



Obrázek 3.9: Hodnota parametru $\theta = 20$, průměrný počet sousedů je 18. Histogramy misorientací během jednoho chodu MCMC řetězce po počtu iterací 0 (výchozí stav), 20000, 40000, 60000.

Závěr

Pomocí MCMC algoritmu se povedlo ukázat, že parametr θ z rozdělení (2.2) má zásadní vliv na tvar rozdělení misorientací. Čím větší je jeho absolutní hodnota, tím dále je od Mackenzieho rozdělení. Pro kladné hodnoty se střední hodnota zvyšuje, pro záporné se snižuje. Pro hodnoty parametru se stejnou absolutní hodnotou platí, že záporná hodnota směřuje k větším rozdílům od Mackenzieho rozdělení, normovaná diskrepance se blíží maximální teoretické hodnotě 2.

Dále se povedlo ukázat, že ač průměrný počet stěn mozaiky lehce ovlivňuje rozdělení misorientací, je tento vliv velice malý až zanedbatelný. Počty stěn ale ovlivňují pravděpodobnost přijetí návrhu a rychlost konvergence algoritmu, jež jsou nižší u složitějších prostorových závislostí, tedy v případě průměrného počtu stěn rovného 18.

Seznam použité literatury

- JONES, G., ROBERTS, G. a ROSENTHAL, J. (2014). Convergence of conditional Metropolis-Hastings samplers. *Advances in Applied Probability*, **46**, 422–445.
- KARAFIÁTOVÁ, I. (2017). Statistika směřových dat s využitím v krystalografii. Bakalářská práce, MFF Univerzity Karlovy v Praze.
- MEYN, S. P. a TWEEDIE, R. L. (1993). *Markov Chain and Stochastic Stability*. Springer-Verlag, New York. ISBN 978-1-4471-3267-7.
- MORAWIEC, A. (2004). *Orientations and Rotations*. Springer-Verlag, New York. ISBN 978-3-642-07386-1.
- PAWLAS, Z., KARAFIÁTOVÁ, I. a HELLER, L. (2020). Random tessellations marked with crystallographic orientations. *Spatial Statistics*, **39**(100469).
- PROKEŠOVÁ, M. (2020). Metody MCMC. poznámky k přednášce, MFF Univerzity Karlovy v Praze <https://www2.karlin.mff.cuni.cz/~proke-sov/2020/mcmcLN.pdf>.
- RYCROFT, C. (2009). Voro++: a three-dimensional Voronoi cell library in C++. *Chaos*, **19**(041111).
- SEITL, F., PETRICH, L., KRILL III, C. E., SCHMIDT, V., STANĚK, J. a BENEŠ, V. (2020). Exploration of Gibbs-Laguerre tessellations for three-dimensional stochastic modeling. *Methodology and Computing in Applied Probability*. <https://doi.org/10.1007/s11009-019-09757-x>.
- STOYAN, D., SEITL, F. a BENEŠ, V. (2020). Dependent radius marks of Laguerre tessellations: A case study. *Australian and New Zealand Journal of Statistics*, *in print*.