

EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
TERMÉSZETTUDOMÁNYI KAR

---

Kovács Adrienn

MARKOV-LÁNC MONTE CARLO MÓDSZEREK ÉS  
ALKALMAZÁSAI GRÁFOKON

Szakdolgozat  
Alkalmazott matematikus MSc  
Sztochasztika szakirány

Témavezető:  
Backhausz Ágnes  
Valószínűségelméleti és Statisztika Tanszék



Budapest, 2019

# Köszönetnyilvánítás

Szeretném kifejezni köszönetemet témavezetőmnek, Backhausz Ágnesnek a sok segítségért, melyet a szakdolgozatban szereplő cikkek feldolgozásához és a szimuláció készítéséhez nyújtott.

# Tartalomjegyzék

Bevezetés . . . . .	4
<b>1. Markov-lánc</b>	<b>6</b>
<b>2. Markov-lánc Monte Carlo módszerek</b>	<b>11</b>
2.1. Metropolis–Hastings-algoritmus . . . . .	11
2.1.1. Szimmetrikus bázisú lánc . . . . .	11
2.1.2. Általános bázisú lánc . . . . .	12
2.2. Glauber-dinamika . . . . .	16
2.2.1. Ising-modell . . . . .	17
2.2.2. Potts-modell . . . . .	30
2.3. Swendsen–Wang-algoritmus . . . . .	32
2.3.1. Swendsen–Wang-algoritmus az Ising-modellen . . . . .	33
2.3.2. Swendsen–Wang-algoritmus a Potts-modellen . . . . .	35
2.4. Markov-lánc Monte Carlo módszerek alkalmazásai . . . . .	37
<b>3. Összefoglalás</b>	<b>39</b>
<b>4. Függelék</b>	<b>44</b>

## Bevezetés

Szakedolgozatom témája Markov-lánc Monte Carlo módszerek gráfokon. Az első fejezetben bemutatom, hogy mit értünk Markov-lánc alatt, illetve a témához kapcsolódó alapvető fogalmakat írom le. Fontos megemlíteni, hogy mi is az a Markov-lánc. A Markov-lánc egy olyan véletlen folyamat, amely Markov-tulajdonságú. Ez azt jelenti, hogy adott jelenbeli állapot mellett, a lánc jövőbeni állapotának eloszlása nem függ a múltbeliektől. Egyszerű példa az érmedobás, ha fejet dobunk elsőre, másodikra ugyanúgy 50 – 50 százalékkal dobhatunk írást vagy fejet egyaránt.

A második fejezetben három Markov-lánc Monte Carlo módszert ismertetek, először a Metropolis–Hastings-algoritmust, majd a Glauber-dinamikát, ezen belül az Ising-modellt és a Potts-modellt, valamint a Swendsen–Wang-algoritmus mutatom be. Szeretném pár mondatban összefoglalni ezeket a módszereket. A Monte Carlo módszer véletlenszerű mintavételen alapuló módszer, amelyet matematikai, fizikai és gazdasági számítások modellezésére használnak, Markov-lánc Monte Carlo módszerek adott valószínűségi eloszlásból való mintavételezéshez használt algoritmusokat foglalnak magukba, olyan Markov-láncot létrehozva, melynek az elérni kívánt eloszlás lesz a stacionárius eloszlása.

A Metropolis–Hastings-algortmushoz egy diszkrét eloszlásra és normális eloszlásra vonatkozóan programot is készítettem, majd a teljes variációs távolságot kiszámítva és ábrázolva látható, hogy a teljes variációs távolság mind a diszkrét eloszlás és mind a normális eloszlás esetén a mintanagyság növekedésével nullához tart, vagyis az így létrehozott láncok a kívánt eloszláshoz konvergálnak.

A Glauber-dinamikával tudjuk megadni többek között a fázisátmenetek dinamikáját, a fizikában a fázisátmenet a szilárd, folyadék és gáz állapot változást jelenti. Az egyik legegyszerűbb statisztikai modell, amely a fázisátmenet bemutatására szolgál, a kétdimenziós négyzetrács Ising-modell. Az Ising-modell fán való alkalmazását az információelméletben, genetikában és statisztikai fizikában használják. Ha adott egy  $T$  fa, és egy véletlen  $R$  bit, spin van a fa gyökerében, majd azt a bitet, például  $+1$  értéket a  $T$  fa élei  $\theta$  valószínűséggel továbbadják, vagy  $1 - \theta$  valószínűséggel nem adják tovább, azaz  $-1$ -et ad tovább a következő csúcsokba. Az alkalmazási területeken azt a modellt tanulmányozzák, amelyben a cél, hogy az  $R$ -et rekonstruálják azokból az értékekből, amelyek a fa  $n$ -edik szintjén megjelennek.

A Glauber-dinamika módszere a következő: egy halmazból egyenletesen véletlenül választunk egy csúcsot, és a csúcsnak új állapotot sorsolunk a halmazban fennmaradó csú-

csak állapotaira vett feltételes eloszlás szerint. A statisztikus fizikában a kristályszerkezetbe rendeződött atomok, ionok és molekulák közötti kapcsolatot az Ising-modell általánosítása, a Potts-modell írja le.

Az Ising-modellre vonatkozó Glauber-dinamika keverési ideje függ a gráf szerkezetétől. Dolgozatomban többféle gráf esetére vizsgáltam az Ising-modellt. Teljes gráfra vonatkozó Ising-modell keverési ideje kritikus hőmérséklet esetén nagyságrendekkel nagyobb lehet, mint magas illetve alacsony hőmérséklet esetén. A négyzetrács Ising-modellen az  $n$  csúcsú gráfon kritikus hőmérséklet esetén a keverési idő legfeljebb  $n^c$ . A fára vonatkozó Ising-modellen az  $n$  csúcsú gráfon a relaxációs idő polinomiális minden hőmérsékleti tartományban, és magas hőmérséklet esetén a keverési idő  $O(\log n)$  nagyságrendű. Teljes gráfra vonatkozó Potts-modell keverési ideje legalább három állapot esetén alacsony hőmérsékletnél nagyobb, mint magas illetve kritikus hőmérséklet esetén.

A Swendsen–Wang-algoritmus egy klaszterfrissítéses módszer, a Metropolis vagy "egy-spines" módszertől eltérően minden egyes lépés alatt  $n$  csúcsot frissítünk, ezért a Swendsen–Wang-algoritmus gyorsabban meg tudja közelíteni az egyensúlyi állapotot, mint az előzőekben említett algoritmusok keverési ideje. A Swendsen–Wang-algoritmus esetén a teljes gráfon az Ising-modellre és a Potts-modellre vonatkozó keverési időt vizsgáltam.

A második fejezet végén a Markov-lánc Monte Carlo módszerek fizikai és közgazdasági alkalmazásait is megemlítem.

# 1. fejezet

## Markov-lánc

Ezt a fejezetet az [1] könyv 1.,4. és 12. fejezete alapján írtam.

A Markov-lánc egy véletlen folyamat, amely az  $\Omega$  halmaz elemei között mozog a következő módon: ha  $x \in \Omega$  a lánc állapota egy adott időpontban, a következő állapot egy fix, csak  $x$ -től függő  $P(x, \cdot)$  valószínűségi eloszlásnak megfelelően választódik ki.

**1.1. Definíció.** Az  $(X_0, X_1, \dots)$  valószínűségi változók sorozatát az  $\Omega$  állapottéren  $P$  átmenet mátrixú Markov-láncnak nevezzük, ha minden  $x, y \in \Omega$ ,  $t \geq 1$  esetén minden

$$H_{t-1} = \bigcap_{s=0}^{t-1} \{X_s = x_s\} = \{X_{t-1} = x_{t-1}, \dots, X_1 = x_1, X_0 = x_0\}$$

*eseményre teljesül  $\mathbf{P}(H_{t-1} \cap \{X_t = x\}) > 0$ , és (a Markov-láncra) teljesül a következő:*

$$\mathbf{P}\{X_{t+1} = y | H_{t-1} \cap \{X_t = x\}\} = \mathbf{P}\{X_{t+1} = y | X_t = x\} = P(x, y). \quad (1.1)$$

Az (1.1) egyenletet Markov-tulajdonságnak nevezik.

A Markov-tulajdonság szerint, a  $t+1$  időponthoz tartozó állapot valószínűségi eloszlása csak az öt megelőző  $t$  időponthoz tartozó valószínűségi eloszlástól függ, és nem függ azoktól az állapotoktól, amelyen keresztül a folyamat eljutott a  $t$  időpontbeli állapotba. Ez azt jelenti, hogy a jelenlegi  $X_t = x$  érték ismeretében a Markov-lánc jövője, az  $X_{t+1} = y$  esemény független a múlttól, a  $H_{t-1}$  eseményektől.

A lánc  $P$  átmenetmátrixának  $x$ -edik sora a  $P(x, \cdot)$  eloszlása, annak a valószínűsége, hogy az  $x$  állapotból egy másik  $\Omega$ -beli állapotba kerülünk.

A  $P$  átmenetmátrix elemei az átmenetvalószínűségek. A  $P$  mátrix sztochasztikus, ha az elemei nemnegatívak, és minden sor összege 1. Az  $t$ -ed rendű átmenetvalószínűségek:  $P^t(i, j) = P(X_{t+m} = j | X_m = i)$ .

**1.2. Definíció.** Az  $i$  állapotból elérhető  $j$  ( $i \rightarrow j$ ), ha van olyan  $t \geq 0$ , hogy  $P^t(i, j) > 0$ . Ha  $i$  és  $j$  közlekednek ( $i \rightarrow j$  és  $j \rightarrow i$ ), akkor az állapotter osztályokra bomlik, ez azt jelenti, hogy bármelyik állapotból bármelyik állapotba el lehet jutni egy osztályon belül. A Markov-lánc irreducibilis, ha egyetlen osztályból áll.

**1.3. Definíció.** Az  $\{t > 0 : P^t(i, i) > 0\}$  halmaz legnagyobb közös osztója az  $i$  periódusa, jelölése  $d(i)$ . Ha a halmaz üres, akkor a periódust nem értelmezzük. A lánc aperiodikus, ha minden állapot periódusa 1. Ha a lánc nem aperiodikus, akkor periodikusnak hívjuk.

**1.4. Definíció.** Egy irreducibilis Markov-láncot ergodikusnak nevezünk, ha minden állapota aperiodikus.

**1.5. Definíció.** Tegyük fel, hogy egy  $\pi \in \Omega$  eloszlásra teljesül a következő egyenlet:

$$\pi(x)P(x, y) = P(y, x)\pi(y) \quad \forall x, y \in \Omega.$$

Ekkor a  $\pi$  stacionárius eloszlású és  $P$  átmenetmátrixú lánc megfordítható.

Legyen  $\Omega = \mathbb{Z}_n = \{0, 1, \dots, n-1\}$ , modulo  $n$  maradékainak osztálya. Tekintsük a következő átmenetmátrixot:

$$P(j, k) = \begin{cases} 1/2 & \text{ha } k \equiv j+1 \pmod{n}, \\ 1/2 & \text{ha } k \equiv j-1 \pmod{n}, \\ 0 & \text{egyébként.} \end{cases}$$

A kapcsolódó  $(X_t)$  Markov-láncot bolyongásnak, vagy véletlen sétának nevezzük az  $n$ -körön.

Az állapotokat úgy lehet elképzelni, mint egy körbe rendezve az egyenlő távolságra lévő pontokat.

Az  $n$ -körön vett bolyongás minden  $n \geq 1$ -re irreducibilis.

A Markov-lánc egy fontos típusának egy egyszerű esete az  $n$ -körön vett bolyongás.

Egy  $G = (V, E)$  gráf  $V$  csúcshalmazzal és  $E$  élhalmazzal, ahol  $E$  elemei a csúcsok rendezetlen párijai:  $E \subset \{\{x, y\} : x, y \in V, x \neq y\}$ . Gondolhatunk úgy  $V$ -re mint csúcsok halmazára, ahol két csúcs,  $x$  és  $y$  egy vonallal van összekötve akkor és csak akkor, ha  $\{x, y\}$  az élhalmaz egy eleme. Ha  $\{x, y\} \in E$ , azt mondjuk, hogy  $x$  az  $y$  szomszédja, és ezt  $x \sim y$ -nal jelöljük. Az  $x$  csúcs foka,  $\deg(x)$ , az  $x$  szomszédjainak száma. Adott egy  $G = (V, E)$

gráf, meghatározhatunk egy egyszerű véletlen sétát a  $G$  gráfon, amely egy  $V$  állapotterű Markov-lánc és átmenetmátrixa:

$$P(x, y) = \begin{cases} \frac{1}{\deg(x)} & \text{ha } y \sim x, \\ 0 & \text{egyébként.} \end{cases}$$

Így mondhatjuk, hogy amikor a lánc az  $x$  csúcsban van, megvizsgálja  $x$  összes szomszédját, véletlenszerűen egyenletesen kiválaszt egyet, és a kiválasztott csúcsba megy át.

**1.6. Definíció.** Legyen  $\pi$  egy eloszlás az  $\Omega$ -n. Ha teljesül a következő

$$\pi = \pi P, \tag{1.2}$$

akkor  $\pi$ -t a Markov-lánc stacionárius eloszlásának nevezzük (a  $P$  átmenetmátrixra nézve). Ha  $\pi$  stacionárius eloszlás és  $\mu_0 = \pi$  (a lánc stacionárius eloszlásból indul), akkor  $\mu_t = \pi$  minden  $t \geq 0$ -ra. Az (1.2) egyenletet írhatjuk elemenként, ekkor az ekvivalens formula

$$\pi(y) = \sum_{x \in \Omega} \pi(x) P(x, y) \quad \forall y \in \Omega.$$

Tekintsünk egy egyszerű bolyongást a  $G = (V, E)$  gráfon. Minden  $y \in V$  csúcsra

$$\sum_{x \in V} \deg(x) P(x, y) = \sum_{x \sim y} \frac{\deg(x)}{\deg(x)} = \deg(y).$$

Ahhoz, hogy egy valószínűséget kapjunk az  $\sum_{y \in V} \deg(y) = 2|E|$  mennyiséggel kell normalizálnunk. Így azt kapjuk, hogy a valószínűségi mérték

$$\pi(y) = \frac{\deg(y)}{2|E|} \quad \text{minden } y \in \Omega\text{-ra,}$$

amely a fokokkal arányos, és mindig egy stacionárius eloszlás a sétára vonatkozóan. A lánc megfordítható, mivel, hogy:

$$\pi(y) P(y, x) = \frac{\deg(y)}{2|E|} \frac{\mathbf{1}_{\{x \sim y\}}}{\deg(y)} = \frac{\mathbf{1}_{\{x \sim y\}}}{2|E|} = \pi(x) P(x, y).$$

Az előző kifejezésben az  $\mathbf{1}_A$  az  $A$  halmaz indikátor függvényét jelöli, ha  $a \in A$  akkor  $\mathbf{1}_A(a) = 1$ , egyébként  $\mathbf{1}_A(a) = 0$ .

Ahhoz, hogy a stacionárius eloszláshoz való konvergenciát meg tudjuk fogalmazni, szükség van eloszlások távolságának értelmezésére.



**1.7. Definíció.** A teljes variációs távolságot két  $\mu$  és  $\nu$  valószínűségi eloszlás között az  $\Omega$ -án a következő módon definiáljuk:

$$\|\mu - \nu\|_{TV} = \max_{A \subset \Omega} |\mu(A) - \nu(A)|.$$

A teljes variációs távolság definíció szerint  $\Omega$  összes részhalmazai fölötti maximum, ezért nem mindig kényelmes, célszerű ezzel becsülni a távolságot. Diszkrét esetben a következő állítás a teljes variációs távolságot egy az állapottér fölötti egyszerű összeggé, összeadássá, csökkenti.

**1.8. Állítás ([2] 4.2. Állítás).** Legyen  $\mu$  és  $\nu$  két valószínűségi eloszlás az  $\Omega$ -n. Ekkor

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Meghatározhatjuk a maximális távolságot  $P^t(x_0, \cdot)$  és  $\pi$  között,  $x_0 \in \Omega$ :

$$d(t) := \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{TV}.$$

Érdeemes bevezetni egy olyan paramétert, amely annak az idejét méri, hogy tetszőleges kezdeti állapotból indítva a Markov-lánc és a stacionárius eloszlás távolsága mikor lesz kicsi.

**1.9. Definíció.** A keverési időt az alábbi módon definiáljuk:

$$t_{MIX}(\varepsilon) := \min \{t : d(t) \leq \varepsilon\}$$

és

$$t_{MIX} := t_{MIX}(1/4).$$

A keverési idő a lépések száma a teljes variációs távolság és a stacionaritás között. Véges állapotterű irreducibilis, aperiodikus Markov-lánc  $t$  időbeli eloszlása a  $\pi$  stacionárius eloszláshoz konvergál, ahogy  $t \rightarrow \infty$ , a lánc kezdő állapotától eltekintve. Fontos kérdés az, hogy a  $t$ -nek milyen nagynak kell lennie ahhoz, hogy a lánc  $t$  időbeli eloszlása megközelítőleg  $\pi$  legyen.

**1.10. Definíció.** Az  $(X_0, X_1, \dots)$   $P$  átmenetmátrixú Markov-lánc a véges  $\Omega$  állapottéren. A Markov-lánc  $x \in \Omega$ -ra vonatkozó elérési ideje  $\tau_x := \min \{t \geq 0 : X_t = x\}$ , az első idő, amikor a lánc az  $x$  állapotba kerül.

A következő tétel mondja ki, hogy az irreducibilis, aperiodikus Markov-láncok konvergálnak a hozzájuk tartozó stacionárius eloszlásokhoz.

**1.11. Tétel.** (Konvergenciatétel)[[2] 4.9. Tétel] Tegyük fel, hogy  $P$  irreducibilis és aperiodikus,  $\pi$  stacionárius eloszlással. Ekkor léteznek  $\alpha \in (0, 1)$  konstansok és  $C > 0$ , hogy

$$d(t) = \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{TV} \leq C\alpha^t.$$

**1.12. Definíció.** Egy  $P$  megfordítható átmenetmátrix sajátértékei:  $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_{|\Omega|} \geq -1$ .

Legyen

$$\lambda_* := \max \{|\lambda| : \lambda \text{ a } P \text{ sajátértéke, } \lambda \neq 1\}.$$

A  $\gamma_* := 1 - \lambda_*$  különbséget az abszolút spektrális résnek nevezzük.

Egy megfordítható lánc spektrális részét a következő módon definiálhatjuk  $\gamma := 1 - \lambda_2$ .

**1.13. Definíció.** Egy megfordítható Markov-lánc relaxációs ideje,  $t_{rel}$ , a  $\gamma_*$  abszolút spektrális réssel a következő:

$$t_{rel} := \frac{1}{\gamma_*}.$$

**1.14. Definíció.** A  $\mu$  és  $\nu$  két valószínűségi eloszlás csatolása,  $(X, Y)$  tetszőleges ugyanazon a valószínűségi mezőn lévő valószínűségi változók együttes eloszlása, ahol  $X$  eloszlása  $\mu$ ,  $Y$  eloszlása pedig  $\nu$ . Így, egy  $(X, Y)$  csatolásra  $P(X = x) = \mu(x)$  és  $P(Y = y) = \nu(y)$  teljesül.

## 2. fejezet

# Markov-lánc Monte Carlo módszerek

Ezt a fejezetet az [1] könyv 3. fejezete, [3], [4], [6] és [7] cikkek és [5] alapján írtam.

Adott egy  $\pi$  valószínűségi eloszlás az  $\Omega$ -n, találhatunk-e olyan átmenetmátrixot, amelynek  $\pi$  a stacionárius eloszlása. Az egyik megközelítés, hogy Markov-láncot használjunk mintavételezéshez. Tegyük fel, hogy  $(X_t)$  egy lánc az  $\Omega$  állapottéren és stacionárius eloszlása az  $\Omega$ -án egyenletes. A konvergenciatétel szerint  $X_t$  megközelítőleg egyenletes eloszlású, ha  $t$  nagy. Adott valószínűségi eloszlásból vett mintavételezés ezen módszerét Markov-lánc Monte Carlo módszernek nevezzük.

### 2.1. Metropolis–Hastings-algoritmus

Adott tetszőleges lánc az  $\Omega$  állapottéren és egy tetszőleges stacionárius eloszlás, lehet-e a láncot úgy módosítani, hogy az új lánc stacionárius eloszlása  $\pi$  legyen? A Metropolis–Hastings-algoritmus teljesíti ezt. A Metropolis–Hastings-algoritmus egy bolyongáson alapuló mintavételezés, ezzel a módszerrel létrehozott minták eloszlása a kívánt eloszláshoz tart. Azt a módszert, amikor a Markov-láncot szimmetrikus eloszlásból generáljuk, Metropolis–Hastings-algoritmusnak nevezzük, ez a Metropolis–Hastings-algoritmus speciális esete. A Markov-láncot generálhatjuk általános, nem feltétlenül szimmetrikus eloszlásból, ebben az esetben Metropolis-Hastings-algoritmusról beszélünk.

#### 2.1.1. Szimmetrikus bázisú lánc

Tegyük fel, hogy  $\Psi$  egy szimmetrikus átmenetmátrix. Ebben az esetben,  $\Psi$  megfordítható az  $\Omega$ -n vett egyenletes eloszlásra vonatkozóan. Az új lánc a következőképpen alakul ki.

Amikor a lánc az  $x$  állapotban van, a  $\Psi(x, \cdot)$  eloszlásból mozgás generálódik. Az új állapot  $y$ , ami  $a(x, y)$  valószínűséggel elfogadott, így a lánc új állapota  $y$ , és a fennmaradó  $1 - a(x, y)$  valószínűséggel a lánc az  $x$  állapotban marad. Az új lánc átmenetmátrixa:

$$P(x, y) = \begin{cases} \Psi(x, y)a(x, y) & \text{ha } y \neq x, \\ 1 - \sum_{z:z \neq x} \Psi(x, z)a(x, z) & \text{ha } y = x. \end{cases}$$

Az  $a(x, y)$  elfogadási valószínűség  $(\pi(y)/\pi(x)) \wedge 1$ -gyel egyenlő. Ezzel a Metropolis-lánc egy  $\pi$  valószínűségre és egy  $\Psi$  szimmetrikus átmenetmátrixra a következő módon definiálható:

$$P(x, y) = \begin{cases} \Psi(x, y) \left[ 1 \wedge \frac{\pi(y)}{\pi(x)} \right] & \text{ha } y \neq x, \\ 1 - \sum_{z:z \neq x} \Psi(x, z) \left[ 1 \wedge \frac{\pi(z)}{\pi(x)} \right] & \text{ha } y = x. \end{cases}$$

### 2.1.2. Általános bázisú lánc

A Metropolis–Hastings-láncot akkor is definiálhatjuk, ha a kezdeti átmenetmátrix nem szimmetrikus. Egy  $\Psi$  általános irreducibilis átmenetmátrixra és az  $\Omega$ -án egy tetszőleges  $\pi$  valószínűségeloszlásra az új lánc a következőképpen jön létre. Amennyiben  $x$  állapotban van a lánc,  $\Psi(x, \cdot)$ -ből generálunk egy  $y$  állapotot, és

$$\frac{\pi(y)\Psi(y, x)}{\pi(x)\Psi(x, y)} \wedge 1$$

valószínűséggel az  $y$  állapotba kerülünk, és komplementer valószínűséggel az  $x$  állapotban marad a lánc. A lánc  $P$  átmenetmátrixa a következő:

$$P(x, y) = \begin{cases} \Psi(x, y) \left[ \frac{\pi(y)\Psi(y, x)}{\pi(x)\Psi(x, y)} \wedge 1 \right] & \text{ha } y \neq x, \\ 1 - \sum_{z:z \neq x} \Psi(x, z) \left[ \frac{\pi(z)\Psi(z, x)}{\pi(x)\Psi(x, z)} \wedge 1 \right] & \text{ha } y = x. \end{cases}$$

A Metropolis–Hastings-algoritmus az  $\Omega$  állapottéren létrehoz egy Markov-láncot, úgy, hogy adott egy  $G = (\Omega, E)$  gráf, és ezen akarunk egy  $\pi$  eloszlást megközelíteni. Ha az irreducibilis Markov-lánc átmenetvalószínűsége  $P$ , akkor a Metropolis–Hastings-lánc a következő módon jön létre:

- ha  $X_n = v_i$ , akkor a  $P(i, \cdot)$  szerint  $v_i$  szomszédjai közül egyenletesen választjuk  $v_j$ -t,
- a

$$\min \left( 1, \frac{\pi(v_j)P(j, i)}{\pi(v_i)P(i, j)} \right) = \min \left( 1, \frac{\pi(v_j) \deg(v_i)}{\pi(v_i) \deg(v_j)} \right)$$

valószínűséggel elfogadjuk  $v_j$ -t,

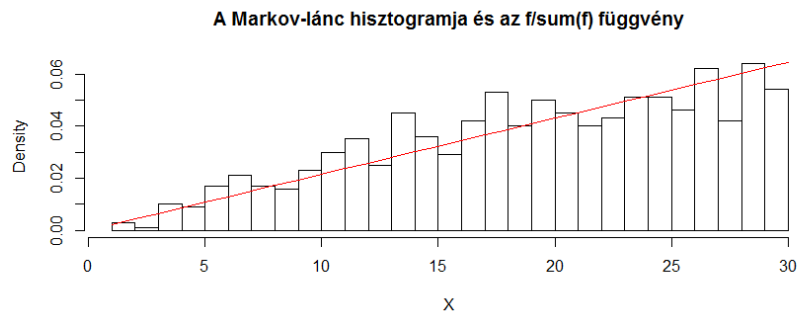
- 

$$1 - \min \left( 1, \frac{\pi(v_j)P(j, i)}{\pi(v_i)P(i, j)} \right) = 1 - \min \left( 1, \frac{\pi(v_j) \deg(v_i)}{\pi(v_i) \deg(v_j)} \right)$$

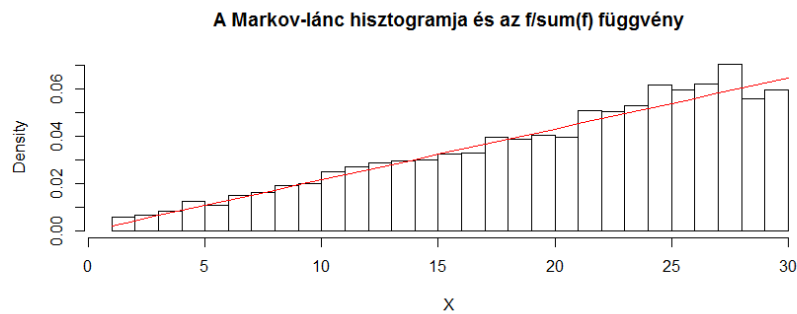
valószínűséggel az új állapot  $v_i$  marad,

- ha elfogadtuk  $v_j$ -t, akkor az új állapot  $v_j$  lesz,  $X_{n+1} = v_j$ , egyébként  $X_{n+1} = v_i$ , vagyis maradunk.

A Metropolis–Hastings-algoritmus által létrehozott Markov-lánc stacionárius eloszlása  $\pi$ , vagyis az 1.11 konvergenciatételnek megfelelően a lánc eloszlása konvergál a stacionárius eloszláshoz.

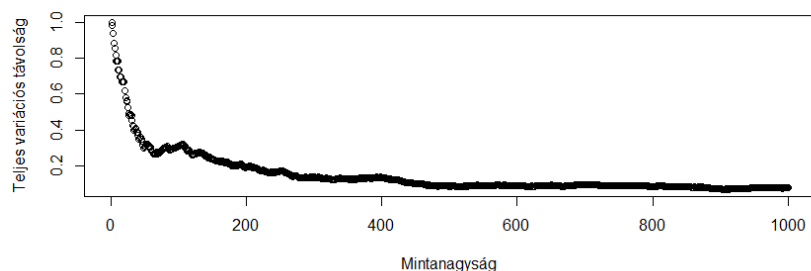


2.1. ábra.

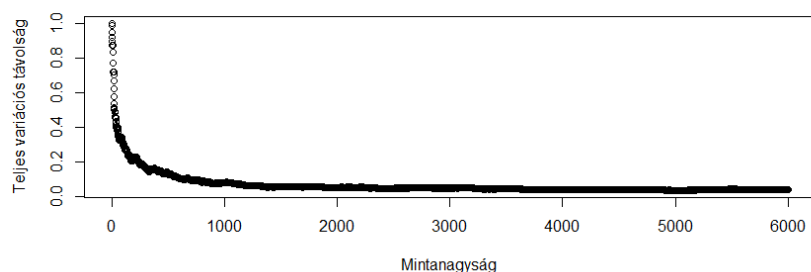


2.2. ábra.

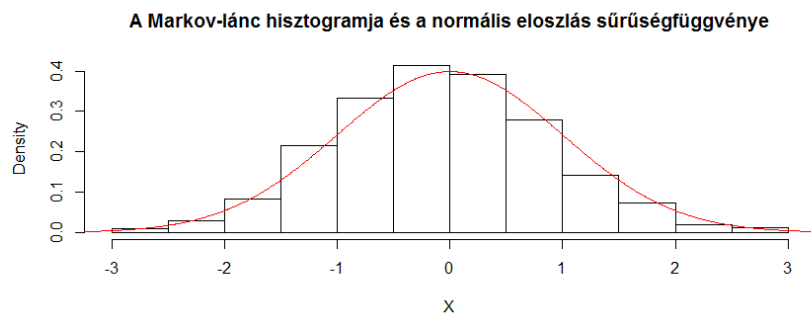
A 2.1 ábrán a Metropolis–Hastings-algoritmus által egy diszkrét eloszlásból szimulált 1000 mintanagyságú Markov-lánc hisztogramja és az  $f/\text{sum}(f)$  függvény látható, ahol az  $f = 1, \dots, 30$ . A diszkrét eloszlásból generált Markov-lánc azt jelenti, hogy az  $f/\text{sum}(f)$  függvény által megadott eloszlásból lett a lánc generálva, így 30 állapot lehetséges, azaz



2.3. ábra. Teljes variációs távolság 1000 mintanagyság esetén



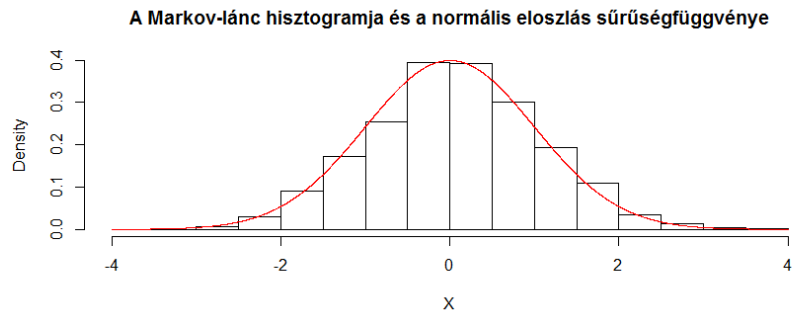
2.4. ábra. Teljes variációs távolság 6000 mintanagyság esetén



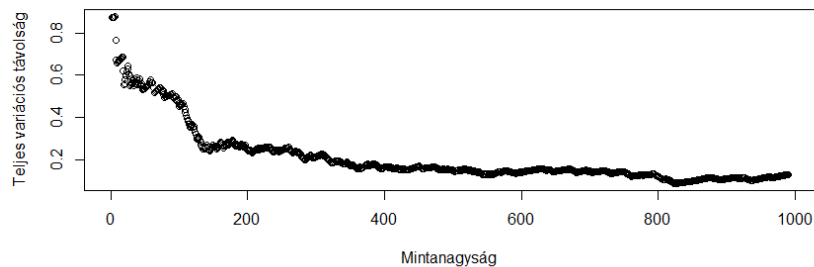
2.5. ábra.

30 elemű az állapottér, és az egyes állapotok valószínűségei a sorszámukkal arányosak. Az elfogadási valószínűség  $y/x$ , ahol  $y \in \{1, \dots, 30\}$ .

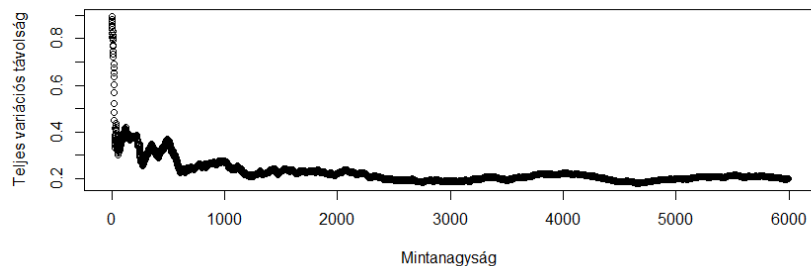
A 2.2 ábrán a Metropolis-Hastings algoritmus által egy diszkrét eloszlásból szimulált 6000 mintanagyságú Markov-lánc hisztogramja és az  $f/\text{sum}(f)$  függvény látható, ahol az  $f = 1, \dots, 30$ . Ebben az esetben is a diszkrét eloszlásból generált Markov-lánc azt jelenti,



2.6. ábra.



2.7. ábra. Teljes variációs távolság 1000 mintanagyság esetén



2.8. ábra. Teljes variációs távolság 6000 mintanagyság esetén

hogy az  $f/\text{sum}(f)$  függvény által megadott eloszlásból lett a lánc generálva, így 30 állapot lehetséges, azaz 30 elemű az állapottér. Az elfogadási valószínűség ebben az esetben is  $y/x$ , ahol  $y \in \{1, \dots, 30\}$ . A Markov-lánc eloszlása és az eredeti eloszlás közötti teljes variációs távolság az 1.11 konvergenciatételnek megfelelően a mintanagyság növelésével csökken, amely a 2.3 és a 2.4 ábrákon látható, azaz a lánc eloszlása tart az eredeti eloszláshoz.

A 2.5 ábrán a Metropolis-Hastings által  $N(0, 1)$  eloszlásból szimulált 1000 mintanagyságú Markov-lánc hisztogramja és a  $N(0, 1)$  eloszlás sűrűségfüggvénye látható. Az elfogadási valószínűség  $\min(1, dnorm(y, 0, 1) / dnorm(x, 0, 1))$ . A 2.6 ábrán a Metropolis-Hastings által  $N(0, 1)$  eloszlásból szimulált 6000 mintanagyságú Markov-lánc hisztogramja és a  $N(0, 1)$  eloszlás sűrűségfüggvénye látható. A Markov-lánc eloszlása és az eredeti eloszlás közötti teljes variációs távolság az 1.11 konvergenciatételnek megfelelően a mintanagyság növelésével csökken, amely a 2.7 és a 2.8 ábrákon látható, azaz a lánc eloszlása tart az eredeti eloszláshoz.

## 2.2. Glauber-dinamika

Legyen  $V$  egy gráf csúcshalmaza és  $S$  véges halmaz. Az  $S^V$  elemeit konfigurációknak nevezzük. Elképzelhetünk egy konfigurációt úgy mint, hogy az  $S$  elemeivel felcímkézzük a csúcsokat. Adott egy  $\pi$  valószínűségi eloszlás a konfigurációk terén. A  $\pi$ -re vett Glauber-dinamika egy Markov-lánc az  $\Omega$  állapottéren, melynek a stacionárius eloszlása  $\pi$ . Ezt a láncot gyakran nevezik Gibbs-mintavételezésnek. Általánosan, legyenek  $V$  és  $S$  véges halmazok, és tegyük fel, hogy  $\Omega$  az  $S^V$  részhalmaza. Legyen  $\pi$  egy valószínűségi eloszlás, melynek értelmezési tartománya  $\Omega$ .

A Glauber-lánc az  $x$  állapotból a következőképp mozog: a  $V$  halmazból egyetlenesen véletlenül választunk egy  $v$  csúcsot, és egy új állapotot választunk  $v$ -nek a  $\pi$  mérték szerint véletlenül, de feltételezve, hogy az állapotok halmaza egyenlő  $x$ -szel minden  $v$ -től eltérő csúcson.

Legyen  $x \in \Omega$  és  $v \in V$ , ekkor

$$\Omega(x, v) = \{y \in \Omega : y(w) = x(w) \ \forall w \neq v\}$$

legyen az olyan állapotok halmaza, amik  $v$ -t kivéve mindenol megegyeznek  $x$ -szel. Az átmenetvalószínűségek

$$\pi^{x,v}(y) = \pi(y|\Omega(x, v)) = \begin{cases} \frac{\pi(y)}{\pi(\Omega(x, v))} & \text{ha } y \in \Omega(x, v), \\ 0 & \text{ha } y \notin \Omega(x, v), \end{cases}$$

ahol  $\pi^{x,v}$  az  $\Omega(x, v)$  halmazon értelmezett feltételes eloszlás. Az  $x$  konfiguráció frissítésének menete, hogy egyetlenesen véletlenül veszünk egy  $v$  csúcsot, és  $\pi^{x,v}$ -nek megfelelő új konfigurációt választunk. A Glauber-dinamika szerint a  $\pi$  eloszlás mindig stacionárius és megfordítható.



### 2.2.1. Ising-modell

A 2.2.1. alfejezet [4] cikk alapján készült. Legyen  $G = (V, E)$  véges gráf. A spin rendszer egy valószínűségi eloszlás az  $\Omega = \{-1, 1\}^V$  állapotterén, ahol a  $G = (V, E)$  gráf csúcshalmaza  $V$ . Az  $\Omega := \{-1, 1\}^V$  állapotter elemeit konfigurációknak nevezzük, és  $\sigma \in \Omega$ -ra, a  $\sigma(v)$  értéket a  $v$  csúchoz tartozó spinnek nevezzük.

Az Ising-modell esetében a gráf csúcaiba  $+1$  vagy  $-1$  spint helyezünk.

A valószínűségeloszlást az energiafüggvény segítségével adjuk meg. Egy  $\sigma \in \{-1, 1\}^V$  konfiguráció legközelebbi szomszéd energiája,  $H(\sigma)$  a következő módon definiálható:

$$H(\sigma) := - \sum_{v,w \in V, v \sim w} J(v, w) \sigma(v) \sigma(w), \quad (2.1)$$

ahol  $w \sim v$  azt jelenti, hogy  $\{w, v\} \in E$ . A  $J(v, w)$  paraméterek a csúcsok közötti kapcsolat erősségét mérik, mindig feltesszük, hogy  $J(v, w) \equiv J$ , ahol  $J$  pozitív konstans.

Ha  $J(v, w) > 0$ , akkor a modell ferromágneses. A ferromágneses Ising-modellben a szomszédos csúcsok nagyobb valószínűséggel kapnak azonos spint.

Ha  $J(v, w) < 0$ , akkor a modell antiferromágneses. Az antiferromágneses modellben a szomszédos csúcsok nagyobb valószínűséggel kapnak különböző spint. Ha  $J(v, w) = 0$ , akkor nincs kapcsolat a spinek között.

Adott  $\beta \geq 0$ -ra a  $G$  gráfon az Ising-modell a  $\beta$  paraméterrel az a  $\mu$  valószínűségi mérték az  $\Omega$ -án, amely a  $H$  energiára vonatkozó Gibbs-eloszlással adható meg, és amely a következőképpen definiálható:

$$\mu(\sigma) = \frac{1}{Z(\beta)} e^{-\beta H(\sigma)}, \quad (2.2)$$

ahol  $Z(\beta) := \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}$  a partíciós függvény, a normalizáló konstans, amely szükséges ahhoz, hogy a  $\mu$  valószínűségi eloszlást létrehozzuk. A  $\beta$  paraméter a rendszer inverz hőmérséklete, ha  $\beta \geq 0$  akkor a modell ferromágneses, egyébként anti-ferromágneses.

A valószínűségi eloszlás az előző (2.1) és (2.2) képletekből adódóan a következő:

$$\mu(\sigma) = Z(\beta)^{-1} \exp \left( \beta \sum_{v,w \in V, v \sim w} J \sigma(v) \sigma(w) \right). \quad (2.3)$$

A  $\beta$  paraméter fizikailag a hőmérséklet inverzét jelenti, és a  $H$  energia függvény hatását méri a valószínűségi eloszláson. Végtelen hőmérsékletnél, ennek megfelelően  $\beta = 0$  esetén, a  $H$  energia függvény nem játszik szerepet és  $\mu$  egyenletes eloszlás  $\Omega$ -án. Ez azt jelenti,

hogy független azonos eloszlás szerint sorsoljuk ki a spinek értékét, azaz  $1/2$  valószínűséggel  $+1$  és  $1/2$  valószínűséggel  $-1$  lesz a spinek értéke. Ebben az esetben nincs összefüggés a különböző csúcsokon vett spinek között és a  $\{\sigma(v)\}_{v \in V}$  valószínűségi változók függetlenek.

Alacsony hőmérsékletnél  $\beta > 1$  esetén vagy minden spin pozitív, vagy minden spin negatív, ekkor az energiafüggvény negatív értéket ad. Ez azt jelenti, hogy az alacsonyabb energiájú konfigurációk valószínűbbek. Amint  $\beta > 0$  növekszik,  $\mu$  torzítása szintén növekszik az alacsonyabb energiájú konfigurációk irányába.

A  $\beta_c(q)$  a kritikus hőmérséklet inverze, a kritikus hőmérséklet azt a legmagasabb hőmérsékletet jelenti, amelyre lehet nemnulla mágnesezettség. Ebben a pontban a rendszer egy úgynevezett rendezett-rendezetlen átmenetbe kerül, amit fázisátmenetnek nevezünk. A fázisátmenet kifejezés leggyakrabban a szilárd, folyékony, gáz és plazma fázisok, állapotok közötti átmenetet jelenti.

Most nézzük részletesebben a Glauber-dinamikát a (2.3) eloszlásra. Ez a lánc úgy fejlődik, hogy a  $V$  halmazból egyenletesen véletlenül választunk egy  $v$  csúcsot, és a  $v$  csúcsához tartozó spin értéket  $\mu$  eloszlás szerint frissítjük, vagyis új konfiguráció keletkezik, a  $\{\eta \in \Omega : \eta(w) = \sigma(v), w \neq v\}$  halmazra vonatkozó feltételes eloszlás szerint.

Más szavakkal, ha a  $v$  csúcsot választottuk ki, az új konfiguráció  $\sigma$ -val megegyezik mindenhol kivéve  $v$  lehetőségét, és  $v$ -nél a spin  $+1$

$$p(\sigma, v) := \frac{e^{\beta S(\sigma, v)}}{e^{\beta S(\sigma, v)} + e^{-\beta S(\sigma, v)}} = \frac{1 + \tanh(\beta S(\sigma, v))}{2}.$$

valószínűséggel, ahol az  $S(\sigma, v) := J \sum_{w: w \sim v} \sigma(w)$ .

Így, erre a láncra vonatkozó átmenetmátrix a következőképpen adható meg:

$$P(\sigma, \sigma') = \frac{1}{|V|} \sum_{w \in V} \frac{e^{\beta \sigma'(w) S(\sigma, w)}}{e^{\beta \sigma'(w) S(\sigma, w)} + e^{-\beta \sigma'(w) S(\sigma, w)}} \cdot \mathbf{1}_{\{\sigma'(w) = \sigma(w) \ \forall w \neq v\}}.$$

Nyilvánvalóan, a  $v$  csúcs új spinjének eloszlása csak a  $v$  szomszédjainak jelenlegi spinjétől függ. Így könnyen belátható, hogy  $(X_t)$ , a Glauber-dinamika egy változata  $\sigma$  kezdő állapottal a  $\mu$  mértékre nézve megfordítható (2.2)-ben.

A keverési idő a  $\beta$ -tól függően változhat, ha  $\beta$  kicsi, a lánc keverési ideje gyors, míg ha  $\beta$  nagy, a lánc lassan fog konvergálni a stacionárius eloszláshoz.

A következőkben, a Glauber-dinamikára az  $(X_t)_{t=0}^{\infty}$  jelölést használjuk, az  $(X_t)$  a Glauber-dinamika egy verziója  $\sigma$  kezdő állapottal. A  $\mathbf{P}_{\sigma}$  és  $\mathbf{E}_{\sigma}$  jelöléseket használjuk a valószínűségi mértékre és a hozzá kapcsolódó operátorhoz, amikor  $X_0 = \sigma$ .

Adott egy  $G_n = (V_n, E_n)$  gráfok sorozata, és  $\mu_n$ -t használjuk az Ising-mértékhez és  $(X_t^n)$  a Glauber-dinamikához a  $G_n$  gráf sorozaton. A Glauber-dinamika lánc stacionárius eloszlástól vett legrosszabb esetű (maximális, legnagyobb) távolsága  $t$  lépés után:

$$d_n(t) := \max_{\sigma \in \Omega_n} \|\mathbf{P}_\sigma(X_t^n \in \cdot) - \mu_n\|_{TV}, \quad (2.4)$$

ahol  $\|\mu - \nu\|_{TV}$  jelöli a teljes variációs távolságot a  $\mu$  és  $\nu$  valószínűségi mértékek között.

A  $t_{MIX}(n)$  keverési időt a következő módon definiálhatjuk:

$$t_{MIX(\varepsilon)}(n) := \min \{t : d_n(t) \leq \varepsilon\}, \quad (2.5)$$

ahol  $\varepsilon \in (0, 1)$  és

$$t_{MIX}(n) := t_{MIX(1/4)}(n) \quad (2.6)$$

Ergodikus Markov-láncokra vonatkozó 1.11 konvergenciatétel szerint  $t_{MIX}(n)$  véges minden rögzített  $n$ -re,  $d_n(t) \rightarrow 0$  ahogy  $t \rightarrow \infty$ . Mindazonáltal,  $t_{MIX}(n)$  általában  $n$ -nel tart a végtelenbe. A Glauber-dinamika egy cutoff-ot mutat  $\{t_n\}$ -nél  $\{w_n\}$  ablakkal ha  $w_n = o(t_n)$  és

$$\lim_{\gamma \rightarrow \infty} \liminf_{n \rightarrow \infty} d_n(t_n - \gamma w_n) = 1,$$

$$\lim_{\gamma \rightarrow \infty} \limsup_{n \rightarrow \infty} d_n(t_n + \gamma w_n) = 0.$$

A cutoff jelenség a Markov-lánc stacionárius eloszlásától vett teljes variációs távolság hirtelen csökkenését jelenti, amely a  $w_n$  ablak mérettel következik be.

## Az Ising-modell a teljes gráfon

Legyen  $K_n$  az  $n$  csúcsú teljes gráf, olyan gráf, amelynek minden csúcsa össze van kötve minden más csúccsal. Tekintsük az Ising-modellt a  $K_n$  teljes gráfon. A csúcshalmaz  $V_n = \{1, 2, \dots, n\}$ , és az  $\mathcal{E}_n$  élhalmaz minden  $\{i, j\}$  párt tartalmaz  $1 \leq i < j \leq n$ . Legyen a  $J$  paraméter  $1/n$ , ebben az esetben a  $\mu$  Ising-mérték a  $\{-1, 1\}^n$  állapotéren:

$$\mu(\sigma) = \mu_n(\sigma) = \frac{1}{Z(\beta)} \exp\left(\frac{\beta}{n} \sum_{1 \leq i < j \leq n} \sigma(i) \sigma(j)\right).$$

**2.1. Tétel ( [4] 1. Tétel).** *Tegyük fel, hogy  $\beta < 1$ . A  $K_n$ -en az Ising-modellre vonatkozó Glauber-dinamikának van cutoff jelensége a  $t_n = [2(1 - \beta)]^{-1} n \log n$ -nél  $n$  ablak mérettel.*

A  $\beta = 1$  kritikus esetben a Glauber-dinamika keverési idejének rendje  $n^{3/2}$ .

**2.2. Tétel ([4] 2. Tétel).** *Ha  $\beta = 1$ , akkor léteznek  $C_1, C_2 > 0$  konstansok, úgy, hogy a  $K_n$ -en az Ising-modellre vonatkozó Glauber-dinamikára:*

$$C_1 n^{3/2} \leq t_{MIX}(n) \leq C_2 n^{3/2}.$$

Végezetül tekintsük az alacsony hőmérsékletű esetet, amely  $\beta > 1$ -nek felel meg. A következőkhöz meg kell említenünk a normalizált mágnesezettséget, egy  $S$  függvényt, ami  $\sigma$  konfigurációkon definiálva a következő:

$$S(\sigma) := \frac{1}{n} \sum_{i=1}^n \sigma(i).$$

Mágnesezettség alatt a spinek átlagát értjük. Szintén definiáljuk az  $\Omega^+$  állapotok halmazát nemnegatív magnetizációval:

$$\Omega^+ := \{\omega \in X : S(\sigma) \geq 0\}.$$

Ezek azok a konfigurációk, amikben több  $+1$  értékű spin van, mint  $-1$  értékű spin. A keverési idő  $n \log n$  rendű, ha a lánc az  $\Omega^+$  halmazra van megszorítva. Pontosabban, a megszorított dinamika a következő módon fejlődik az  $\Omega^+$ -on: generáljunk egy jelöltet, amely  $\eta$ -ba megy a szokásos Glauber-dinamikának megfelelően. Ha  $S(\eta) \geq 0$ , akkor  $\eta$ -t elfogadjuk mint az új állapotot, amíg ha  $S(\eta) < 0$ , akkor helyette  $-\eta$ -ba megy. Az  $\eta$  kiválasztott állapotot akkor fogadjuk el, ha  $\eta$ -ban a spinek átlaga nemnegatív, ha az  $\eta$  állapotban a spinek átlaga negatív, akkor a  $-\eta$  állapotot fogadjuk el, ekkor a spinek átlaga nemnegatív lesz. Ez szemléletesen azt jelenti, hogy, ha egy olyan állapotból indulunk el, amiben a spinek átlaga nemnegatív, azaz több vagy egyenlő a pozitív spinek száma mint a negatív spinek száma, ez az állapot megfelelhet a jég állapotnak. Ebből egy olyan állapotba kerülünk, amiben több vagy egyenlő a pozitív spinek száma mint a negatív spinek száma. Vagyis a jég állapotból nem kerülhetünk át víz állapotba, ami megfelel annak, hogy a spinek átlaga negatív, vagyis több a negatív spinek száma, mint a pozitív spinek száma. Ez azt jelenti, hogy nem jön létre fázisámenet. Mindig hasonló állapotban maradunk.

**2.3. Tétel ([4] 3. Tétel).** *Ha  $\beta > 1$ , akkor léteznek  $C_3(\beta), C_4(\beta) > 0$   $\beta$ -től függő konstansok, úgy, hogy a  $K_n$ -en az Ising-modellre vonatkozó megszorított Glauber-dinamikára,*

$$C_3(\beta) n \log n \leq t_{MIX}(n) \leq C_4(\beta) n \log n.$$

A 2.1, 2.2 és 2.3 tételekből az látszik, hogy nagyobb hőmérséklet esetén kisebb a keverési idő, a kritikus hőmérséklet környékén a keverési időben nagyságrendi változás lép fel.

### Az Ising-modell a $\mathbb{Z}^2$ négyzetrácson

Ez a rész [6] cikk alapján készült. Tekintsük az Ising-modellt az eddigi teljes gráf helyett a négyzetrácson. A klasszikus Ising modell egy  $\Lambda$  véges dobozon a  $\mathbb{Z}^2$  négyzetrácson külső mezők nélkül az alábbi módon definiálható. Az  $\Omega = \{\pm 1\}^\Lambda$  lehetséges konfigurációk halmaza, ahol egyes konfigurációk a  $\Lambda$ -ban a csúcsokhoz plusz/mínusz spinek egy hozzárendelésének felelnek meg. Azt a valószínűséget, hogy a rendszer egy  $\sigma \in \Omega$  konfigurációban van, a Gibbs-eloszlás adja meg:

$$\mu_\Lambda(\sigma) = \frac{1}{Z(\beta)} \exp\left(\beta \sum_{u \sim v} \sigma(u) \sigma(v)\right),$$

ahol a  $Z$  partíciófüggvény egy normáló konstans. Ez a definíció a végtelen  $\mathbb{Z}^2$  négyzetrácsra vonatkozik, ahol van egy kritikus pont  $\beta_c = \frac{1}{2} \log(1 + \sqrt{2})$  úgy, hogy a Gibbs-eloszlás egyértelmű akkor és csak akkor, ha  $\beta \leq \beta_c$ . Tekintsük a  $\mathbb{Z}^2$ -en a kritikus Ising-modellre vonatkozó Glauber-dinamika spektrális részét. Amikor a dinamika kritikusan lelassul, akkor kritikus Ising-modellről beszélhetünk.

A  $\Lambda \subset V$  határát, amelyet  $\partial\Lambda$ -val jelölünk, úgy definiálhatjuk, mint a  $V \setminus \Lambda$  halmaz  $\Lambda$ -val szomszédos csúcsait, és  $\tau \in \{\pm 1\}^{\partial\Lambda}$  a határfeltétel. Tekintsünk egy  $\Lambda$  véges dobozt, a határfeltétel azon dobozon kívül eső csúcsok spinjeinek értékére utal, amely csúcsoknak van szomszédos csúcsa a dobozon belül. A  $+$  határfeltétel azt jelenti, hogy azon csúcsok spinjeinek az értéke  $+1$ , amelyek a dobozon kívül esnek, de van szomszédjuk a dobozon belül. A  $-$  határfeltétel azt jelenti, hogy azon csúcsok spinjeinek az értéke  $-1$ , amelyek a dobozon kívül esnek, de van szomszédjuk a dobozon belül. A szabad határfeltétel esetén nincs határfeltétel meghatározva. Tekintsünk egy  $\Lambda$  véges dobozt  $+$  határfeltétellel és tekintsük a Glauber-dinamikát, úgy, hogy a dobozban lévő csúcsoknak nincs állapota, vagyis nincs a csúcsoknak spin értéke. Kiválasztunk egy dobozon belüli csúcsot, amely szomszédja a dobozon kívüli csúcs és a Glauber-dinamika szerint meghatározzuk az állapotát. A kiválasztott csúcsnak nagy valószínűséggel  $+1$  lesz a spin értéke, mivel a szomszédja is  $+1$ .

Magas és közepes hőmérséklet esetén,  $+$  határfeltétel mellett a doboz közepén lévő csúcs a Glauber-dinamika szerint nagyobb valószínűséggel kap  $-1$  értékű spint, mint alacsony hőmérséklet esetén. Alacsony hőmérséklet esetén,  $+$  határfeltétel mellett a doboz

közepén lévő csúcs a Glauber-dinamika szerint nagyobb valószínűséggel kap  $+1$  értékű spint mint  $-1$  értékűt.

Az Ising-modellre vonatkozó folytonos idejű Glauber-dinamika Gibbs-eloszlása a következő generátorral adható meg:

$$(Lf)(\sigma) = \sum_{x \in \Lambda} c(x, \sigma) (f(\sigma^x) - f(\sigma)),$$

ahol  $\sigma^x$  az a  $\sigma$  konfiguráció, amelynél  $x$ -ben a spin értéke az ellentettje. A Glauber-dinamika  $c(x, \sigma)$  átmenet rátái:

$$c(x, \sigma) = \left( 1 + \exp \left( -2\beta\sigma(x) \sum_{y \sim x} \sigma(y) \right) \right)^{-1}.$$

Az így megadott Glauber-dinamika ekvivalens az egy rátájú Poisson órák szerinti spinek frissítésével, vagyis véletlen, exponenciális időközönként a kiválasztott csúcs spinjét kitöröljük és a szomszédos csúcsainak spinjétől függően új spint helyezünk a kiválasztott csúcsba. A dinamika  $\mathbb{Z}^2$ -en kritikus lelassulást mutat: magas hőmérsékletnél a spektrális rés inverze  $O(1)$ , a kritikus  $\beta_c$ -nél polinomiális, alacsony hőmérsékletnél exponenciális.

**2.4. Tétel ([6] 1. Tétel).** *Tekintsük a kritikus Ising-modell egy  $n$  oldalhosszú  $\Lambda \subset \mathbb{Z}^2$  véges dobozon, az inverz hőmérséklet  $\beta_c = \frac{1}{2} \log(1 + \sqrt{2})$ . Legyen  $\text{gap}_\Lambda^\tau$  a spektrális rés a megfelelő folytonos idejű Glauber-dinamika generátorában egy  $\tau$  tetszőleges, rögzített határfeltétel mellett. Ekkor létezik egy abszolút konstans  $C > 0$  ( $\Lambda$ -tól,  $\tau$ -tól független) úgy, hogy  $(\text{gap}_\Lambda^\tau)^{-1} \leq n^C$ .*

Könnyen belátható, hogy  $t_{MIX} \leq \text{gap}^{-1} \log \frac{e}{\mu_{min}}$ , ahol  $\mu_{min} = \min_{x \in \Omega} \mu(x)$ . Ebből és a 2.4 tételből következik az alábbi következmény.

**2.5. Következmény ([6] 2.1. Következmény).** *Tekintsük a kritikus Ising-modell egy  $n$  oldalhosszú  $\Lambda \subset \mathbb{Z}^2$  véges dobozon, az inverz hőmérséklet  $\beta_c = \frac{1}{2} \log(1 + \sqrt{2})$ . Legyen  $t_{MIX}^\tau$  a megfelelő folytonos idejű Glauber-dinamika legrosszabb esetű teljes variációs keverési ideje, egy  $\tau$  tetszőleges, rögzített határfeltétel mellett. Ekkor létezik egy abszolút konstans  $C > 0$  ( $\Lambda$ -tól és  $\tau$ -tól független) úgy, hogy  $t_{MIX}^\tau \leq n^C$ .*

### Az Ising-modell a $T_r^{(b)}$ fán

Ez a rész [7] cikk alapján készült. Legyen  $G = (V, E)$  egy végtelen gráf  $\Delta$  maximális fokkal. Legyen  $\rho$  egy megkülönböztetett csúcs és  $G_r = (V_r, E_r)$  jelölje a feszített gráfot, ahol  $V_r =$

$\{v \in V : \text{dist}(\rho, v) \leq r\}$ . Legyen  $n_r$  a csúcsok száma  $G_r$ -ben. Nézzük azt az esetet amikor  $G = T = (V, E)$   $b$ -reguláris fa, ekkor  $T_r^{(b)} = (V_r, E_r)$  jelöli az  $r$ -szintű,  $b$ -reguláris fát. A  $G_r$  gráfon az Ising-modellben  $\beta$  inverz hőmérsékletnél minden  $\sigma \in \{-1, 1\}^{V_r}$  konfigurációhoz az alábbi valószínűség rendelhető hozzá:

$$\mu(\sigma) = Z(\beta)^{-1} \exp \left( \beta \sum_{\{v,w\} \in E_r} J(v,w) \sigma(v) \sigma(w) + \beta \sum_{v \in V_r} h_v \sigma(v) \right),$$

ahol  $Z(\beta)$  egy normáló konstans. A  $h$  a modell külső mezejének létezésére utal. A  $h_v$  azt is megmutatja, hogy a  $v$  spinje milyen kapcsolatban van a külső mezővel. Ha  $h = 0$ , akkor a modellnek nincs külső mezeje, így nincs külső hatás. Ha  $h > 0$ , a külső mező spinjei pozitívak, akkor a belső mező spinjei, amelyek a külső mező spinjeivel kapcsolatban vannak nagyobb valószínűséggel pozitív értékűek lesznek. Ha  $h < 0$ , a külső mező spinjei negatívak, akkor a belső mező spinjei, amelyek a külső mező spinjeivel kapcsolatban vannak nagyobb valószínűséggel negatív értékűek lesznek. Ha  $\beta = 0$  és  $h = 0$ , akkor  $\mu$  egyenletes eloszlás  $\Omega$ -án. Ez azt jelenti, hogy független azonos eloszlás szerint sorsoljuk ki a spinek értékét, azaz  $1/2$  valószínűséggel  $+1$  és  $1/2$  valószínűséggel  $-1$  lesz a spinek értéke. Ha van külső mező, azaz  $h < 0$  vagy  $h > 0$ , akkor nem független azonos eloszlás szerint sorsoljuk ki a spinek értékét, mivel a külső mező hatással van a belső mezőre.

Egy  $G = (V, E)$  irányítatlan gráfon definiáljunk egy konfigurációt, mint az  $A^V$  egy  $\sigma$  eleme, ahol  $A$  véges halmaz, az Ising-modell esetén  $A = \{+1, -1\}$ , és minden egyes élre  $(v, w) \in E$ , vegyünk egy  $\alpha_{vw} : A \times A \rightarrow \mathbb{R}_+$  súlyfüggvényt.

A Gibbs-eloszlás minden  $\sigma$  konfigurációhoz egy  $\prod_{\{v,w\} \in E} \alpha(\sigma_v, \sigma_w)$  valószínűségi hányadost rendel hozzá. Az Ising-modell esetén  $\alpha_{vw}(\sigma_v, \sigma_w) = e^{\beta \sigma_v \sigma_w}$ .

Egy véges gráfon az Ising-modellre vett Glauber-dinamika megadható a következő generátorral:

$$(\mathcal{L}(f))(\sigma) = \sum_{v \in V} \left( \sum_{a \in A} K[\sigma \rightarrow \sigma_v^a] (f(\sigma_v^a) - f(\sigma)) \right),$$

ahol  $\sigma_v^a$  olyan konfiguráció, amelyre

$$\sigma_v^a(w) = \begin{cases} a & \text{ha } w = v \\ \sigma(w) & \text{ha } w \neq v, \end{cases}$$

és

$$K[\sigma \rightarrow \sigma_v^a] = \frac{\prod_{w:(w,v) \in E} \alpha_{vw}(a, \sigma_w)}{\sum_{a' \in A} \left( \prod_{w:(w,v) \in E} \alpha_{vw}(a', \sigma_w) \right)}.$$

Az  $f$  a konfigurációkon értelmezett függvény, a  $K$  átmenetvalószínűség azt mondja meg, hogy melyik konfigurációból melyik konfigurációba mennyi valószínűséggel megyünk. Az alapján, hogy várhatóan mennyi változást generál az  $f$  a Glauber-dinamikának egy lépését végrehajtva, az összes csúcshoz egy 1-rátájú Poisson óra tartozik. Ezek a Poisson órák függetlenek egymástól. Tegyük fel, hogy  $v$ -ben az óra  $t$ -kor csöng és  $t$  előtt a konfiguráció  $\sigma$  volt. Ekkor  $t$ -kor a  $v$ -ben lévő spint egy véletlen új spinre cseréljük a fennmaradó spinekre vett feltételes Gibbs-eloszlás szerint:

$$\frac{\mathbf{P}[\sigma'(v) = i | \sigma]}{\mathbf{P}[\sigma'(v) = j | \sigma]} = \prod_{w: \{v,w\} \in E} \frac{\alpha_{vw}(i, \sigma(w))}{\alpha_{vw}(j, \sigma(w))}.$$

Ha  $G_r = T_r^{(b)}$  és  $\varepsilon = (1 + e^{2\beta})^{-1}$  rögzített, akkor a módszer a következő. Válasszunk egy véletlen  $\pm 1$  spint egyenletesen a fa gyökeréhez, vizsgáljuk át a fát felülről lefelé, rendeljük a  $v$  csúcshoz egy spint a  $v$  csúcs szülőjének a spinjével egyenlően  $1 - \varepsilon$  valószínűségével és ellenkezőleg  $\varepsilon$  valószínűséggel. Az Ising-modellre vonatkozó Glauber-dinamika az új  $\sigma'(v)$  spint a következő módon választja ki:

$$\frac{P(\sigma'(v) = +1 | \sigma)}{P(\sigma'(v) = -1 | \sigma)} = \exp \left( 2\beta \sum_{w: \{w,v\} \in E_r} \sigma(w) \right).$$

Az Ising-modell a  $b$ -rendű fán három különböző tartománnyal rendelkezik. A magas hőmérsékletű tartományban, ahol  $1 - 2\varepsilon < 1/b$ , a végtelen fán egy egyértelmű Gibbs-mérték van, és a gyökérben a spin várható értéke,  $\sigma_\rho$  bármilyen  $\sigma_{\partial T_r^{(b)}}$  határfeltétel mellett  $r$ -ben exponenciálisan csökken. Az  $\partial T_r^{(b)}$  a gyökértől  $r$  távolságra lévő csúcsok. Magas hőmérséklet esetén a gyökér spinjének várható értékének exponenciális csökkenése azt jelenti, hogy, ha a  $+1$  határfeltételt tekintjük, vagyis a határon az összes spin  $+1$ , akkor ahogy a gyökér felé haladunk egyre kevésbé lesznek valószínűbbek a  $+1$  értékű spinek. A határtól a gyökér felé tartva a  $+1$  értékű spinek valószínűsége exponenciálisan csökken, ekkor a spinek várható értékei is exponenciálisan csökkennek.

A közepes hőmérsékletű tartományban, ahol  $1/b < 1 - 2\varepsilon < 1/\sqrt{b}$ , a fent leírt exponenciális csökkenés még tart a tipikus határfeltételekkel, de vannak bizonyos kivételes határfeltételek, úgy mint a csak pozitív spinekből álló határ, következésképp, végtelen sok Gibbs-mérték van a végtelen fán. Ilyenkor három extrémális Gibbs-mérték van, a csak pozitív, csak negatív határfeltételhez tartozók, illetve a szabad határfeltételhez tartozó. Minden más Gibbs-mérték előáll ezek konvex kombinációjaként.



Az alacsony hőmérsékletű tartományban, ahol  $1 - 2\varepsilon > 1/\sqrt{b}$ , erősebb a szomszédos csúcsok közötti kölcsönhatás, így a tipikus határfeltételek torzíthatnak a spin várható értékében a  $\sigma_\rho$  gyökérben. Alacsony hőmérséklet esetén a határtól a gyökér felé tartva a spinek várható értékei polinomiálisan csökkennek, vagyis a határtól távolabbi csúcsok spinjei nagyobb valószínűséggel egyeznek meg a határon lévőkével, mint magasabb hőmérséklet esetén. Ennek megfelelően változik a relaxációs idő is, amint azt a következő tétel mutatja.

**2.6. Tétel ([7] 1.4. Tétel).** *Tekintsük az Ising-modellt a  $T_r = T_r^{(b)}$   $r$ -szintű  $b$ -reguláris fán. Legyen  $\varepsilon = (1 + e^{2\beta})^{-1}$ . A Glauber-dinamika  $\tau_2$  relaxációs idejének a  $T_r^{(b)}$ -n a következő korlátjai lehetnek.*

1. *A relaxációs idő polinomiális minden hőmérsékleten:  $\tau_2 = n_r^{O(\log(1/\varepsilon))}$ . Továbbá, az alábbi határérték:*

$$\lim_{r \rightarrow \infty} \frac{\log \left( \tau_2 \left( T_r^{(b)}, \beta \right) \right)}{\log n_r}$$

*létezik.*

2. *Alacsony hőmérsékletű tartomány:*

- *Ha  $1 - 2\varepsilon \geq 1/\sqrt{b}$ , akkor  $\sup_r \tau_2(T_r) = \infty$ . Valójában,  $\tau_2(T_r) = \Omega \left( n_r^{\log_b(b(1-2\varepsilon)^2)} \right)$  amikor  $1 - 2\varepsilon > 1/\sqrt{b}$  és  $\tau_2(T_r) = \Omega(\log n_r)$  amikor  $1 - 2\varepsilon = 1/\sqrt{b}$ .*
- *Valamint,  $\tau_2$  tart a végtelenbe, amint  $\varepsilon$  tart 0-hoz:  $\tau_2(T_r) = n_r^{\Omega(\log(1/\varepsilon))}$ .*

3. *Közepes és magas hőmérsékletű tartományok: Ha  $1 - 2\varepsilon < 1/\sqrt{b}$ , akkor a relaxációs idő egyenletesen korlátos:  $\tau_2 = O(1)$ . Továbbá ez az eredmény minden  $\{h(v)\}_{v \in T_r}$  külső mezőn érvényes.*

**Bizonyítás.** A 2.6 Tétel bizonyítása [7] cikk alapján magas hőmérsékletű tartomány esetén. A bizonyítás során blokkdinamikát és csatolást alkalmazunk. A blokkdinamika a következő módszert jelenti: tekintsük a  $T = T_r^{(b)}$  fát, amely egy nagyobb  $r + 2h$  magasságú  $b$ -reguláris  $T_*$  fa része, ahol a  $T$  fa  $\rho$  gyökere a  $T_*$  fa  $h$  szintjén van. A  $T_*$  minden egyes  $v$  csúcsához tekintsünk olyan  $h$  magasságú részfát, amelynek  $v$  a gyökere. Egy blokk definíció szerint  $T$  egy ilyen részfával való metszete. Minden egyes blokknak van egy egy rátájú Poisson órája és amikor az óra csöng a blokkhoz tartozó összes csúcs spinjét kitöröljük, és a  $T$  fa fennmaradó részének azon spinjeitől függően, amelyek a blokkal szomszédosak,

a Gibbs-eloszlás szerint új spineket helyezünk a blokkhoz tartozó csúcsokba. Ebben az esetben véletlen időközönként választunk ki egy blokkot és frissítjük a blokkhoz tartozó csúcsok állapotát. Minden egyes blokkhoz tartozó Poisson óra független egymástól. Az ehhez kapcsolódó diszkrét idejű dinamika, a blokkdinamika minden egyes lépésénél véletlenszerűen válasszunk ki egy blokkot, a blokkhoz tartozó csúcsok összes spinjét töröljük ki, és a  $T$  fa fennmaradó részének spinjeire vett feltételes Gibbs-eloszlásnak megfelelően tegyünk új spineket a csúcsokba. Ebben az esetben nem véletlen időközönként választunk ki egy blokkot, hanem lépésenként, fix időközönként. A csatolás azt jelenti, hogy ha tekintünk két különböző konfigurációt, elindítjuk a Glauber-dinamikát és minden egyes lépésben ugyanazt a csúcsot frissítjük, úgy, hogy a végeredmény a lehető legnagyobb valószínűséggel megegyezzen. A konfigurációkon alkalmazzunk súlyozott Hamming-metrikát,

$$d(\sigma, \eta) = \sum_v \lambda^{|v|} \mathbf{1}_{(\sigma_v \neq \eta_v)},$$

ahol  $|v|$  a gyökér  $v$  csúcstól vett távolságát jelöli.

Két konfiguráció távolságát, úgy kapjuk meg, hogy megszámláljuk, hogy hány olyan  $v$  van, ahol a  $\sigma$  és  $\eta$  különböző és  $\lambda^{|v|}$ -el súlyozva összeadjuk. Legyen  $\theta = 1 - 2\varepsilon$  és  $\lambda = 1/\sqrt{b}$ ,  $b\lambda\theta < 1$  és  $\theta < \lambda$ . Két különböző  $\sigma$  és  $\eta$  konfigurációból elindulva, a csatolás mindig az azonos blokkot választja  $\sigma$ -ban és  $\eta$ -ban és a két blokk között csatolást választunk, amely minimalizálja  $d(\sigma', \eta')$ -t. Olyan csatolást használunk, hogy minden konfiguráció párra, amely egy spinnel különbözik, a blokkdinamika egy lépését alkalmazva csökkeni fog a két konfiguráció közti távolság várható értéke. Az eredetileg különböző konfigurációkra ezt a módszert alkalmazva, a konfigurációk tartani fognak egymáshoz. Legyen  $v$  egy olyan csúcs, amelyre  $\sigma_v \neq \eta_v$ , azaz az  $\sigma$  és  $\eta$  konfigurációk különbözőek a  $v$  csúcsban. Ekkor  $d(\sigma, \eta) = \lambda^{|v|}$ . Jelölje  $B$  a kiválasztott blokkot és  $\sigma', \eta'$  legyenek a blokkdinamika egy lépése utáni konfigurációk.

Az alábbi lemmát bizonyítás nélkül felhasználjuk.

**2.7. Lemma.** *Legyen  $T$  egy véges fa és legyenek  $v \neq w$  csúcsok  $T$ -ben. Legyen  $\{\beta_e \geq 0\}_{e \in E(T)}$  ferromágneses hatások a  $T$ -n, és legyen  $\{-\infty < H(u) < \infty\}_{u \in V(T)}$  egy külső mező a  $T$  csúcsain. Tekintsük a következő feltételes Gibbs-mértékeket:*

$\mu_{+,H}$  : Gibbs-mérték  $H$  külső mezővel feltéve, hogy  $\sigma_v = 1$ .

$\mu_{-,H}$  : Gibbs-mérték  $H$  külső mezővel feltéve, hogy  $\sigma_v = -1$ .

*Ekkor az  $\mu_{+,H}[\sigma_w] - \mu_{-,H}[\sigma_w]$  függvény maximuma  $H \equiv 0$  esetén adódik.*

Az 2.7 Lemmában a  $\{\beta_e \geq 0\}_{e \in E(T)}$  azt jelenti, hogy az inverz hőmérséklet élről élre

változik. Négy esetet tekinthetünk.

**1. Eset** Ha a  $B$  blokk nem tartalmazza sem a  $v$  csúcst sem a  $v$  csúcs bármelyik szomszéd csúcsát, akkor a  $B$  határán lévő spinek megegyeznek, nem keletkezik különbség, de nem is szűnik meg, és így  $d(\sigma', \eta') = d(\sigma, \eta)$ .

**2. Eset** Ha a  $B$  blokk tartalmazza a  $v$  csúcst, akkor a  $v$ -ben lévő különbség eltűnik,  $\sigma' = \eta'$ , és  $d(\sigma', \eta') = d(\sigma, \eta) - \lambda^{|v|}$ .  $h$  olyan blokk van, amely a  $v$  csúcs  $h$  darab őseinek megfelelő a  $v$  felett lévő  $1, 2, \dots, h$  generációknál, azaz egyes blokkok olyan részfák, amelyek gyökerei a  $v$  csúcs ősei, és így tartalmazzák  $v$ -t. Azaz ha a  $B$  blokk olyan részfa, amely tartalmazza a  $v$  csúcst, az eredetileg különböző  $\sigma$  és  $\eta$  konfigurációkból indulva, egy lépés blokkdinamika alkalmazása után az új  $\sigma'$  és  $\eta'$  konfigurációk megegyeznek.

**3. Eset** Ha a  $B$  blokk olyan részfa, amely gyökere a  $v$  csúcs egyik gyereke, akkor a feltételes valószínűségek a  $B$  külső határai által adottak, amelyek nem ugyanolyanok, ugyanis az egyik blokk fölött  $+1$  van, a másik blokk fölött  $-1$  van. Azaz az egyik blokk gyökerének szülőjénél  $+1$  értékű spin van, a másik blokk gyökerének szülőjénél  $-1$  értékű spin van. Mindkét blokk leveleinek szomszédjának ugyanazok a határkonfigurációi vannak. Amikor a blokkon tekintjük a folyamatot, akkor a határkonfiguráció hatásakor a külső mező megváltozásával lehet számolni. Ez azt jelenti, hogy  $\sigma$ -nak és  $\eta$ -nak ugyanaz a határkonfigurációja minden határ csúcson kivéve  $v$ -t. A 2.7 Lemmából adódik, hogy a változás felső becsléséhez elég azt az esetet tekinteni, amikor nincs külső mező. Tekintsük azt, amikor az egyik blokk gyökerének van  $+1$  szomszédja és a másik blokk gyökerének  $-1$  szomszédja van, és nincs külső mező vagy határfeltétel. Azután a blokkot feltöltjük felülről lefelé, minden egyes él  $\theta$  valószínűséggel továbbadja a spin értékét, azaz a jelenlegi csúcs spinje megegyezik a csúcs szülőjének spinjével, és  $1 - \theta$  valószínűséggel elfelejti a spin értékét, vagyis a jelenlegi csúcs spinje egy új véletlenszerű,  $1/2$  valószínűséggel  $+1$ , illetve  $1/2$  valószínűséggel  $-1$  értékű spin. Csatoljuk ezeket a választásokat  $\sigma$  és  $\eta$  megfelelő éleihez és azt kapjuk, hogy a  $\sigma'$  és  $\eta'$  között kialakuló távolság várható értéke meg fog egyezni a  $v$  spinjeit megkapó csoport, a  $v$ -t tartalmazó klaszter várható súlyával, ami  $\sum_j \lambda^{|v|+j} b^j \theta^j \leq \lambda^{|v|} / (1 - b\lambda\theta)$ . Itt  $\theta^j$  annak a valószínűsége, hogy a  $j$ -szinten egy csúcsban különbség van a spinek között, és  $b^j$  a  $j$ . szinten lévő csúcsok száma,  $\lambda^{|v|+j}$  a súly.  $b$  olyan blokk van, ami  $v$  gyerekeinek felel meg, azaz olyan blokkok, amelyek gyökerei a  $v$  csúcs gyerekei.

**4. Eset** Ha a  $B$  blokk gyökere a  $v$  csúcs őse, pontosabban a  $h + 1$  generáció  $v$  felett, ekkor a blokkhoz nem tartozik hozzá a  $v$  csúcs, akkor a feltételes valószínűségek nem ugyanazok, ugyanis az egyik blokkban van egy csúcs akinek a szomszédja  $v$ , amely már nem tartozik

a blokkhoz  $+1$  értékű, és a másik blokkban van egy csúcs akinek a szomszédja  $v$ , amely már nem tartozik a blokkhoz  $-1$  értékű. Pontosan egy ilyen blokk van. Ez a különbség a két blokk között legfeljebb annyi változást generálhat a konfigurációk közötti távolságban, mint amennyi a blokkhoz tartozó  $w$  csúcs klaszterének a mérete. Ha az egyik blokknál a  $v$  csúcsban  $+1$  van a másik blokknál pedig  $-1$  van a  $v$  csúcsban akkor a blokkok konfigurációja közötti távolság a  $v$  csúcs azon  $\theta$  klaszterétől függ, amely a  $v$  őseire vonatkozik. Azaz  $v$  őseire visszamenőleg mennyi csúcson van az egyik blokkban  $+1$ , és a másik blokkban  $-1$ , és visszamenőleg az utolsó  $w$  csúcs is továbbadhatja az eltérést,  $\theta$  valószínűséggel az egyik blokkban  $+1$ -et, a másik blokkban pedig  $-1$ -et, és annál nagyobb lesz a két blokk konfigurációja közötti távolság.

Alkalmazzuk a 2.7 Lemmát, hogy megmutassuk, hogy a távolság várható értéke a  $w$   $\theta$  klaszterének méretétől függ. Egy csúcs  $\theta$  klasztere azt a csúcshalmazt jelenti, hogy ha tekintjük a csúcs állapotát és a csúcsból kiinduló éleken sorsolunk függetlenül egy  $\theta$  és  $1-\theta$  valószínűséget, és  $\theta$  valószínűséggel továbbadjuk azt a spin értéket ami a csúcsban van a következő csúcsba, akkor mely csúcsokba jut el a csúcs állapota. A  $v$  csúcs klaszterének várható súlya a következő módon számolható ki:

$$\sum_w \theta^{|v|-|w|} \sum_j \lambda^{|w|+j} b^j \theta^j = \frac{\sum_w \lambda^{|w|} \theta^{|v|-|w|}}{1 - b\lambda\theta} = \frac{\lambda^{|v|}}{(1 - \theta\lambda^{-1})(1 - b\lambda\theta)}.$$

Az előző számolásban, ha a  $v$  csúcs klaszterét tekintjük, akkor azt, hogy mennyi annak a valószínűsége, hogy a  $w$ , amely a  $v$  őse még benne van a  $v$  klaszterében a  $\theta^{|v|-|w|}$  adja meg. A  $\lambda^{|w|+j}$  a  $|w|+j$  szinten lévő csúcsához tartozó súly,  $b^j$  a  $|w|+j$  szinten lévő csúcsok száma  $w$  leszármazottai közül, és  $\theta^j$  annak a valószínűsége, hogy a spin érték tovább adódott a  $|w|+j$  szintig, azaz ezen a szinten lévő csúcs benne van a klaszterben, ha  $w$  benne volt. Geometriai sorral felülről becsülve kapjuk meg az eredményt, véges sok  $w$  van, de ebben az esetben végtelen sort tekintünk.

Az eseteket figyelembe véve, és a 4 esetben kapott konfigurációk közti távolság változásait összeadva, és figyelembe véve, hogy hány blokk van és mennyi a valószínűsége, hogy milyen blokkot választunk, kapjuk meg a távolság változásának várható értékét:

$$\mathbf{E} \left( d(\sigma', \eta') - d(\sigma, \eta) \right) \leq \left( \frac{b\lambda^{|v|}}{1 - b\lambda\theta} + \frac{\lambda^{|v|}}{(1 - \theta\lambda^{-1})(1 - b\lambda\theta)} - h\lambda^{|v|} \right) \frac{1}{n + h - 1}.$$

Ha a  $h$  blokk magasság egy jelentősen nagy konstans,  $n$  a fában lévő csúcsok száma, akkor az első két tag nullához tart, amint  $h \rightarrow \infty$ , és a fennmaradó  $-h\lambda^{|v|}/(n + h - 1)$ -ra

valamilyen pozitív  $c$  konstansra kapjuk, hogy

$$\mathbf{E} \left( d \left( \sigma', \eta' \right) - d \left( \sigma, \eta \right) \right) \leq \frac{-c\lambda^{|\nu|}}{n} \leq \frac{-c}{n} d \left( \sigma, \eta \right) \quad (2.7)$$

$$\mathbf{E} \left( d \left( \sigma', \eta' \right) \right) \leq \left( 1 - \frac{c}{n} \right) d \left( \sigma, \eta \right). \quad (2.8)$$

Jegyezzük meg, hogy  $\max d \left( \sigma, \eta \right) = \sum_{j \leq r} b^j \lambda^j \leq \sqrt{n}$ . A csatolással [18] azt kapjuk, hogy a blokkdinamikára a keverési idő legfeljebb  $O(n \log n)$ .

A (2.7) eredményt ismerve a diszkrét idejű blokkdinamika spektrális részére is tudunk eredményeket megadni. A csatolás minden egyes lépése várható értékben  $(1 - c/n)$ -szeres csökkenést jelent a konfigurációk távolsága között, ekkor a blokkdinamika spektrális rése legalább  $c/n$ . [19] Legyen  $\lambda_2$  abszolút értékben a második legnagyobb sajátérték és  $f$  egy sajátvektor  $\lambda_2$ -höz. Legyen  $M = \sup_{\sigma, \eta} |f \left( \sigma \right) - f \left( \eta \right)| / d \left( \sigma, \eta \right)$  és jelölje  $\mathbf{P}$  az átmenetoperátort, azaz  $\lambda_2 f = \mathbf{P} f$ . Ekkor

$$\begin{aligned} |\lambda_2| M &= \sup_{\sigma, \eta} \frac{|\mathbf{P} f \left( \sigma \right) - \mathbf{P} f \left( \eta \right)|}{d \left( \sigma, \eta \right)} \quad (2.9) \\ &\leq \sup_{\sigma, \eta} \sum_{\sigma', \eta'} \mathbf{P} \left[ \left( \sigma, \eta \right) \rightarrow \left( \sigma', \eta' \right) \right] \frac{|f \left( \sigma' \right) - f \left( \eta' \right)|}{d \left( \sigma', \eta' \right)} \frac{d \left( \sigma', \eta' \right)}{d \left( \sigma, \eta \right)} \\ &\leq \sup_{\sigma, \eta} \sum_{\sigma', \eta'} \mathbf{P} \left[ \left( \sigma, \eta \right) \rightarrow \left( \sigma', \eta' \right) \right] M \frac{d \left( \sigma', \eta' \right)}{d \left( \sigma, \eta \right)} \\ &= M \sup_{\sigma, \eta} \frac{\mathbf{E} \left[ d \left( \sigma', \eta' \right) \right]}{d \left( \sigma, \eta \right)} \leq \left( 1 - c/n \right) M. \end{aligned}$$

A (2.9) egyenletben az utolsó lépés a (2.7) képletből következik.

Így  $|\lambda_2| M \leq (1 - c/n) M$ , ahonnan a diszkrét idejű blokkdinamika relaxációs ideje legfeljebb  $O(n)$ .

A folytonos idejű blokkdinamika relaxációs idejére a következő eredményt kapjuk. Folytonos idejű Glauber-dinamika esetén minden csúcshoz tartozik egy egy rátájú Poisson-folyamat, ami exponenciális időközönként újra sorsolja a csúcs állapotát. Ha független Poisson-folyamatokat tekintünk és egyesítjük az eseményeket, akkor olyan Poisson-folyamatot kapunk, melynek rátája az eredeti ráták összege. Mivel  $n$  csúcs van, ezért  $n$  rátájú Poisson-folyamatunk van, ami azt jelenti, hogy egységnyi idő alatt a történések eloszlása  $n$  paraméterű Poisson, azaz várhatóan  $n$  lépés történik egységnyi idő alatt. Míg diszkrét időben egységnyi idő alatt egy lépés történik egy csúcsnál. A folytonos idejű dinamika  $n$ -szer gyorsabb mint a diszkrét idejű dinamika. Ez azért igaz, mert a diszkrét

idejű dinamika átmenetmátrixa  $M = \mathbf{I} + \frac{1}{n}\mathcal{L}$ , ahol  $\mathbf{I}$  a  $2^n$  dimenziós egységmátrix, és  $\mathcal{L}$  a folytonos idejű Glauber-dinamika generátora. Ekkor

$$\tau_2(\text{blokkdinamika}) = O(1).$$

Az egy csúcsra vonatkozó blokkdinamika relaxációs ideje korlátos. Minden egyes blokkot frissíthetünk úgy, hogy a blokkon belül konstans számú egy csúcsra vonatkozó frissítést csinálunk és minden egyes fa csúcs megszámlálható számú blokkhoz tartozik. Egy blokk-dinamika lépés, ha  $n$  csúcs van, akkor  $n$  csúcsonkénti lépésnek felel meg, ekkor a relaxációs idő konstansszorosára változhat. Az egy csúcsra vonatkozó Glauber-dinamika relaxációs ideje  $O(1)$ .

□

## 2.2.2. Potts-modell

A 2.2.2. alfejezet [3] cikk alapján készült. Következőkben teljes gráfon vett Potts-modellt tekintünk. A Potts-eloszlás egy valószínűségi mérték a  $\Sigma_n = Q^V$ -án, ahol  $Q = \{1, \dots, q\}$  és  $V = \{1, \dots, n\}$ ,  $n \geq 1$  és  $\beta \geq 0$ , a következőképpen definiálható

$$\mu_n(\sigma) = Z_{\beta,n}^{-1} \exp \left\{ (\beta/n) \sum_{u,v \in V} \mathbb{I}_{\sigma(u)=\sigma(v)} \right\},$$

ahol  $\sigma \in \Sigma_n$  és  $Z_{\beta,n}$  normáló konstans.

A klasszikus Ising-modell, amikor  $q = 2$ , most azt az esetet nézzük, amikor  $q \geq 3$ . A Potts-modellre alkalmazott diszkrét idejű Glauber-dinamikát jelöljük  $(\sigma_t)_{t \geq 0}$ -val.  $\sigma_0$ -ból indulva, minden lépésnél egyenletesen választunk egy  $u \in V$  csúcsot és legyen

$$\sigma_{t+1}(v) = \begin{cases} \sigma_t(v) & \text{ha } v \neq u \\ k & \mu_n(\sigma(u) = k | \sigma(w) = \sigma_t(w) \forall w \neq u) \text{ valószínűséggel ha } v = u. \end{cases}$$

A  $\beta_c(q)$  a kritikus hőmérséklet inverze, ugyanazt jelenti mint az Ising-modell esetén. A Markov-folyamat átmenet valószínűségét  $P_n$ -nel jelöljük, és  $\mathbb{P}_{\sigma_0}$ -val a hozzá tartozó valószínűségi mértéket.

A teljes variációs távolságot és a keverési időt Potts-modell esetében az Ising-modellnél definiált (2.4),(2.5) és (2.6) képletekkel lehet megadni.

Ha bármilyen rögzített  $\varepsilon \in (0, 1)$  esetén

$$w_\varepsilon(n) = t_{MIX(\varepsilon)}(n) - t_{MIX(1-\varepsilon)}(n) = o(t_{MIX(1/4)}(n)), \text{ ahogy } n \rightarrow \infty,$$

akkor azt modjuk, hogy a Markov-láncok családja bemutatja a cutoff jelenséget, ami leír egy erős csökkenést a teljes variációs távolságon az 1-től 0-ig zárt intervallumon. A kisebb rendű idő intervallumot mint a  $t_{MIX}(n)$  cutoff window-nak jelöljük.

A cutoff akkor és csak akkor fordul elő ha  $t_{MIX(\delta)}(n)/t_{MIX(\varepsilon)}(n) \rightarrow 1$  ahogy  $n \rightarrow \infty$  minden rögzített  $\delta, \varepsilon \in (0, 1)$  esetén.

A Potts-model dinamikája kritikusan lelassul egy  $\beta_s(q) > 0$  inverz hőmérsékleten. Ezen dinamikai határt a következő összefüggés adja meg:

$$\beta_s(q) = \sup \left\{ \beta \geq 0 : \left( 1 + (q-1) e^{2\beta \frac{1-qx}{q-1}} \right)^{-1} - x \neq 0 \text{ minden } x \in (1/q, 1) \text{-ra} \right\}.$$

Nem úgy mint az Ising-modell esetén, amikor  $\beta_s(2) = \beta_c(2) = 1$ , a dinamikai átmenet  $q \geq 3$ -ra szigorúan magasabb hőmérsékleten jelenik meg, mint a statikus fázis átmenete, átalakulása, mivel  $\beta_s(q) < \beta_c(q)$ .

A Glauber-dinamika Potts-modellre való alkalmazása során a  $q \geq 3$  esetben, megmutatható, hogy egy  $\beta_s(q)$  inverz hőmérsékletnél kritikusan lelassul és  $\beta_s(q)$  szigorúan alacsonyabb lesz, mint a kritikus  $\beta_c(q)$  a termodinamikai határ egyértelműségére vonatkozóan. A  $\beta < \beta_s(q)$  tartományban a gyors keverés  $O(n \log n)$  lépés alatt és a dinamika cutoffot mutat  $O(n)$  ablak mérettel, azaz a Markov-lánc stacionárius eloszlásához való konvergenciája felgyorsul.

**2.8. Tétel ([3] 1. Tétel).** *Legyen  $q \geq 3$  egész szám. Ha  $\beta < \beta_s(q)$ , akkor a Glauber-dinamika a  $q$ -állapotú Potts-modellre a teljes gráfon cutoffot mutat a*

$$t_{MIX}(n) = \alpha_1(\beta, q) n \log n$$

$$\text{keverési időben, } w_\varepsilon(n) = O_\varepsilon(n) \text{ cutoff ablakkal, ahol } \alpha_1(\beta, q) = [2(1 - 2\beta/q)]^{-1}.$$

Elemezzük a keverési idő rendjét amint  $\beta(n) \rightarrow \beta_s(q)$  amint  $n \rightarrow \infty$ ,

$$\beta(n) = \beta_s(q) - \xi(n),$$

ahol  $\xi(n) \rightarrow 0$ , amint  $n \rightarrow \infty$ . A keverési idő aszimptotikája függeni fog attól, hogy  $\xi$  csökkenése milyen gyors, de a cutoff csak akkor figyelhető meg, ha elég lassú a csökkenés. Ezt a következő tétel fogalmazza meg.

**2.9. Tétel ([3] 2. Tétel).** *Legyen  $q \geq 3$  egész szám. Az  $\beta(n) = \beta_s(q) - \xi(n)$  egyenlettel megadott  $\beta(n)$ -el kapjuk az:*

1. Ha  $\lim_{n \rightarrow \infty} n^{2/3} \xi(n) = \infty$ , akkor a Glauber-dinamikának a teljes gráfon van cutoff-ja  $t_{MIX}(n)$  keverési idővel és  $w_\varepsilon(n)$  cutoff ablakkal:

$$t_{MIX}(n) = \alpha_1(\beta(n), q) n \log n + \alpha_2(q) n / \sqrt{\xi(n)},$$

$$w_\varepsilon(n) = O_\varepsilon \left( n + \sqrt{n/\xi(n)^{5/2}} \right),$$

ahol  $\alpha_2(q)$  egy pozitív konstans és  $\alpha_1$  a 2.8 Tétel által definiált konstans.

2. Ha  $0 \leq \liminf_{n \rightarrow \infty} n^{2/3} \xi(n) \leq \limsup_{n \rightarrow \infty} n^{2/3} \xi(n) < \infty$ , akkor a dinamika a teljes gráfon nem mutat cutoff-ot és a keverési ideje:

$$t_{MIX(\varepsilon)}(n) = \Theta_\varepsilon(n^{4/3}).$$

A 2.9 Tétel 2. részét különösképpen a kritikus  $\beta = \beta_s(q)$  esetén alkalmazzák, ahol a keverési idő  $n^{4/3}$  rendű,  $n^{-2/3}$  rendű skálázási ablakkal, ellentétben a kritikus Ising-modell  $n^{3/2}$  rendű keverési idejével,  $\sqrt{n}$  ablakkal. A  $\beta_s(q)$  felett a keverési idő exponenciálisan nagy  $n$ -ben.

**2.10. Tétel ([3] 3. Tétel).** Legyen  $q \geq 3$  egész, és rögzített  $\beta > \beta_s(q)$ . Minden  $0 < \varepsilon < 1$ -hoz létezik  $C_1, C_2 > 0$ , úgy, hogy minden  $n$ -re,

$$t_{MIX(\varepsilon)}(n) \geq C_1 \exp(C_2 n).$$

A teljes gráfon vett Potts-modellre vonatkozó Glauber-dinamika keverési idejének teljes elemzését adják a 2.8, 2.9 és 2.10 Tételek. A 2.8 Tétel az  $\beta < \beta_s(q)$  esetet, a 2.9 Tétel az  $\beta(n) = \beta_s(q) - \xi(n)$  esetet és az 2.10 Tétel az  $\beta > \beta_s(q)$  esetet tekinti a keverési idő vizsgálatánál, így az összes lehetőséget megnéztük.

## 2.3. Swendsen–Wang-algoritmus

A statisztikai fizikában az egyik fő cél az, hogy megértsük a rendezetlen rendszerek időbeli fejlődését, és egy jellegzetes példa a klasszikus spin rendszerek, úgy mint az Ising-modell tanulmányozása. Az állapot tér általában nagy,  $2^n$  nagyságrendű, ahol  $n$  a csúcsok száma. Lokális dinamika esetén nagyon lassan közelít a folyamat az egyensúlyi állapot irányába, mivel egyszerre mindig egy spin frissül. Nem lokális, vagyis globális dinamika gyorsabban meg tudja közelíteni az egyensúlyi állapotot, mivel az állapottéren nagy ugrások megengedettek, a szomszédos kapcsolatok viszonylag korlátlanok. Az előző fejezetben láttuk, hogy



már a blokkdinamika is gyorsabb eredményt ad, mint a spinek egyesével való frissítése. A véletlenszerű klaszter modell szerepet játszott egy ilyen egyszerű és előnyös rendszer kifejlesztésében, amelyet Swendsen és Wang javasolt. Swendsen és Wang [33] 1987-ben bemutattak egy új Monte-Carlo algoritmust az Ising-modellre, ami nagymértékben eltér a Metropolis vagy "egy-spines" módszertől, ez egy klaszter algoritmus. A klaszter frissítéses algoritmus a legsikeresebb globális frissítéses módszer az alkalmazásokban. Ezek a módszerek egy lépésben globálisan frissítik a változókat, miközben a sztenderd lokális módszerek egyszerre egy változóval dolgoznak. Adott spinek  $s_i$  konfigurációjából indulva, a Swendsen–Wang-algoritmus folyamata a következő:

- Vizsgáljuk meg az összes  $s_i, s_j$  állapotot, ha  $s_i = s_j$ , akkor az  $i, j$  helyek között hozzunk létre élt  $p = 1 - \exp(-2\beta)$  valószínűséggel, egyébként ne hozzunk létre élt.
- Legyenek klaszterek az élek által összekapcsolt pontok halmazai.
- Minden egyes klaszternek az értéke véletlenszerűen  $\pm 1$ , a többitől függetlenül.
- Az összes él törlődik és az új spin konfiguráció létrejön, és az algoritmus újraindul az első lépéssel.

A statisztikai fizikai modellek szimulációjában fontos szerepe van a Markov-láncok keverési idejének. Különösen érdekes megfigyelni, hogy ezekben a modellekben a fázis átmenetek hogyan nyilvánulnak meg a keverési idővel kapcsolatban. A következő fejezetben a Swendsen–Wang-algoritmusra vonatkozó keverési időt fogjuk vizsgálni.

### 2.3.1. Swendsen–Wang-algoritmus az Ising-modellen

A 2.3.1. alfejezet [23] cikk alapján készült. A következőkben az Ising-modellre vonatkozó Swendsen–Wang-algoritmussal foglalkozunk az  $n$  csúcsú teljes gráfon. A teljes gráfon a kritikus Ising-modell esetén a Swendsen–Wang-dinamikával minden egyes lépésnél  $n$  csúcsot frissítünk, így a keverési idő gyorsabb, mint a Glauber-dinamika esetén. Egy  $G = (V, E)$  gráfon a Swendsen–Wang-algoritmus stacionárius eloszlása, amely az Ising-modellként is ismert. Ez Boltzmann–Gibbs-eloszlásként is ismert valószínűségi mérték, az  $\Omega = \{1, -1\}^V$  halmazon, ahol minden egyes  $\sigma \in \Omega$  valószínűsége

$$\mathbf{P}(\sigma) = \frac{e^{\beta \sum_{(u,v) \in E} \sigma(u)\sigma(v)}}{Z(G)},$$

ahol  $\beta \in [0, \infty]$  egy paraméter, amelyet úgy ismerünk, mint az inverz hőmérséklet, és a normáló konstans  $Z(G)$  a következő módon definiált:

$$Z(G) = \sum_{\sigma \in \Omega} e^{\beta \sum_{(u,v) \in E} \sigma(u)\sigma(v)}.$$

$\sigma \in \Omega$ -ra. Legyen  $G^+(\sigma)$  a  $G$  azon csúcsai által kifeszített gráf, amelyek  $\sigma$  által 1-gyel jelöltek és hasonlóan legyen  $G^-(\sigma)$  a  $G$  csúcsai által kifeszített gráf, amelyek  $\sigma$  által  $-1$ -gyel jelöltek. Egy  $G$  gráfon a Swendsen–Wang-dinamika a  $p \in [0, 1]$  perkolációs paraméterrel egy Markov-lánc  $\Omega$ -án. Adott a lánc  $\sigma_t$  aktuális állapota, a következő  $\sigma_{t+1}$  állapotot az alábbi két lépéses procedúra határozza meg:

- Hajtsunk végre a  $G^+(\sigma_t)$  és a  $G^-(\sigma_t)$  gráfon külön-külön egy független  $p$ -élű perkolációt. Ami azt jelenti, hogy  $G^+(\sigma_t)$  és  $G^-(\sigma_t)$  minden egyes élét tartsuk meg  $p$  valószínűséggel és töröljük  $1 - p$  valószínűséggel, minden éltől függetlenül. Ennek megfelelően nevezzük a kapott gráfokat  $G_p^+$ -nek és  $G_p^-$ -nek.
- Ahhoz, hogy megkapjuk  $\sigma_{t+1}$ -et,  $G_p^+$ -nek és  $G_p^-$ -nek minden egyes csatlakoztatott  $\mathcal{C}$  összefüggőségi komponensében, a  $\mathcal{C}$  minden csúcsához  $1/2$ -ed valószínűséggel rendeljük hozzá  $1$ -et és  $1/2$ -ed valószínűséggel rendeljük hozzá  $-1$ -et, minden egyes komponensről függetlenül.

A Swendsen–Wang-algoritmus egyértelműen egy aperiodikus és irreducibilis Markov-lánc. Bármilyen  $\sigma_0$  konfigurációból indulva,  $t$  frissítés után kapott  $\sigma_t$  törvényszerűen eloszlásban konvergál a stacionárius Ising-mértékhez. A következő tétel minden hőmérsékletre megadja a keverési idő rendjét.

**2.11. Tétel ([23] 2.1. Tétel).** *Tekintsük a Swendsen–Wang-algoritmust az  $n$  csúcsú  $K_n$  teljes gráfon,  $p = \frac{c}{n}$  perkolációs paraméterrel, ahol  $c$  egy  $n$ -től független konstans. Ekkor,*

- ha  $c > 2$ , akkor  $T_{mix} = \Theta(\log n)$ ,
- ha  $c = 2$ , akkor  $T_{mix} = \Theta(n^{1/4})$ ,
- ha  $c < 2$ , akkor  $T_{mix} = \Theta(1)$ .

Azaz a keverési idő alacsony hőmérsékleten, amikor  $\beta > \beta_c$ ,  $\Theta(\log n)$ , kritikus hőmérsékleten, amikor  $\beta = \beta_c$ ,  $\Theta(n^{1/4})$  és magas hőmérsékleten, amikor  $\beta < \beta_c$ ,  $\Theta(1)$ . Amíg az  $n$  csúcsú teljes gráfon a kritikus Ising-modellre a Glauber-dinamika keverési ideje  $\Theta(n^{3/2})$ ,

addig a Swendsen–Wang-dinamika keverési ideje kritikus Ising-modell esetén legfeljebb  $\Theta(n^{5/4})$ , mivel minden egyes lépésnél  $n$  csúcsot frissítünk, ebben az esetben a Swendsen–Wang-dinamika  $n^{1/4}$ -el gyorsabb, mint a Glauber-dinamika.

### 2.3.2. Swendsen–Wang-algoritmus a Potts-modellen

A 2.3.2. alfejezet [22] cikk alapján készült. A következőkben a  $q$ -állapotú ferromágneses Potts-moddellel foglalkozunk az  $n$  csúcsú teljes gráfon. A következő definícióban  $q = 2$  az Ising-modellnek és  $q \geq 3$  a Potts-modellnek feleltethető meg. Egy  $G = (V, E)$  gráfra a modell konfigurációi spinek csúcsokhoz való  $\sigma : V \rightarrow [q]$  hozzárendelése, a konfigurációk halmazát jelölje  $\Omega$ . A modell a  $\beta > 0$ -val paraméterezett, amit úgy ismerünk mint az inverz hőmérséklet. Egy  $\sigma \in \Omega$  konfigurációra legyen  $m(\sigma)$   $E$ -ben az élek száma, amelyek  $\sigma$ -ban egyszínűek és legyen a súlyuk  $w(\sigma) = \exp(\beta m(\sigma))$ . Ekkor a  $\mu$  Gibbs-eloszlás a következőképpen definiált:  $\sigma \in \Omega$ -ra  $\mu(\sigma) = w(\sigma) / Z(\beta)$ , ahol  $Z(\beta) = \sum_{\sigma \in \Omega} w(\sigma)$  a normáló konstans. Az  $n$  csúcsú  $G = (V, E)$  teljes gráfon a  $B > 0$  paraméterhez tartozó Gibbs-eloszlás a következő:

$$\mu(\sigma) = \frac{1}{Z(\beta)} (1 - B/n)^{-m(\sigma)}.$$

Ha  $n$  nagy, akkor  $\beta = -\ln(1 - B/n) \sim B/n$ . A  $B$  paraméterre a kritikus pontok  $\mathfrak{B}_u < \mathfrak{B}_0 < \mathfrak{B}_{rc}$  és

$$\mathfrak{B}_u = \sup \left\{ B \geq 0 \mid \frac{B - z}{B + (q - 1)z} \neq e^{-z} \text{ minden } z > 0\text{-ra} \right\} = \min_{z \geq 0} \left\{ z + \frac{qz}{e^z - 1} \right\},$$

$$\mathfrak{B}_0 = \frac{2(q - 1) \ln(q - 1)}{q - 2}, \mathfrak{B}_{rc} = q.$$

$\mathfrak{B}_0$  ekvivalens az előző fejezetben szereplő  $\beta_c$ -vel és  $\mathfrak{B}_u$  ekvivalens az előző fejezetben szereplő  $\beta_s$ -sel, a  $z = B(qx - 1) / (q - 1)$  paraméterezés alatt. Ahogy az előző részben is láttuk, a Glauber-dinamika egy klasszikus eszköz a Gibbs-eloszlás tanulmányozására. Ez lokális átmenetekkel rendelkező Markov-láncok családja, amelyben egy véletlenszerűen kiválasztott csúcsban frissül a konfiguráció és amely úgy van kialakítva, hogy a stacionárius eloszlás a Gibbs-eloszlás legyen. Egy lokális Markov-lánc korlátja, úgy mint a Glauber-dinamika, az, hogy alacsony hőmérsékletnél, amikor a  $B$  nagy tipikusan lassul a konvergencia. A Swendsen–Wang-algoritmus egy jóval kifinomultabb Markov-lánc, amely a Potts-modell véletlen klaszter reprezentációt használja fel, hogy potenciálisan le tudja

küzdeni az egyszerű Glauber-dinamikát akadályozó szűk keresztmetszeteket. A Swendsen–Wang-algoritmus egy  $(X_t)$  Markov-lánc, amelynek  $X_t \rightarrow X_{t+1}$  átmenetei a következők. Egy  $X_t \in \Omega$  konfigurációból:

- Legyen  $M$  az egyszínű élek halmaza  $X_t$ -ben.
- Perkolációs lépés: minden egyes  $e \in M$  élre,  $B/n$  valószínűséggel, a többitől függetlenül, tartsuk meg,  $1 - B/n$  valószínűséggel töröljük ki. Jelölje  $M'$  a fennmaradó egyszínű élek halmazát.
- Színező lépés: a  $(V, M')$  gráfban, függetlenül minden egyes csatlakoztatott összefüggőségi komponenshez, válasszunk  $[q]$ -ből egyenletesen véletlenszerűen egy színt, és ebben a komponensben minden csúcshoz rendeljük hozzá a kiválasztott színt. Jelölje  $X_{t+1}$  az eredményül kapott spin konfigurációt.

Egy általános tény, hogy a lánc megfordítható a  $\mu$  Gibbs-eloszlásra vonatkozóan, és konvergál is ahhoz. Minket a lánc  $T_{mix}$  keverési ideje érdekel, amely a legrosszabb kezdő állapottól a  $\mu$  eloszlás  $1/4$ -ed teljes variációs távolságával megtett lépések számával definiált. A ferromágneses Potts-modellre vonatkozó Swendsen–Wang-algoritmusra megmutatható, hogy a keverési idő a rendezett/rendezetlen fázis átmenet pontnál  $\exp(n^{\Omega(1)})$  rendű lehet. A teljes gráfra Gore és Jerrum [21] megmutatta, hogy a  $B = \mathfrak{B}_0$  kritikus pontnál a keverési idő  $\exp(\Omega(\sqrt{n}))$ . Most a teljes gráfon a ferromágneses Potts-modellre vonatkozó Swendsen–Wang-algoritmus keverési idejét vizsgáljuk. Korábban a teljes gráfon a ferromágneses Potts-modellre vonatkozó Glauber-dinamika keverési idejére kapott eredmények:  $\Theta(n \log n)$ , ha  $B < \mathfrak{B}_u$ ,  $\exp(\Omega(n))$ , ha  $B > \mathfrak{B}_u$ , és  $\Theta(n^{4/3})$ , ha  $B = \mathfrak{B}_u$ . A fő eredményt a teljes gráfon a Swendsen–Wang-dinamika keverési idejére, ha a  $B$  paraméter egy  $n$ -től független konstans a következő tétel adja meg.

**2.12. Tétel ([22] 1. Tétel).** *Minden  $q \geq 3$  egészre, az  $n$  csúcsú teljes gráfon a Swendsen–Wang-algoritmus  $T_{mix}$  keverési idejére teljesül:*

- Minden  $B < \mathfrak{B}_u$ -ra  $T_{mix} = \Theta(1)$ .
- $B = \mathfrak{B}_u$ -ra  $T_{mix} = \Theta(n^{1/3})$ .
- Minden  $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$ ,  $T_{mix} = \exp(n^{\Omega(1)})$ .
- Minden  $B \geq \mathfrak{B}_{rc}$ -re  $T_{mix} = \Theta(\log n)$ .

Azaz a keverési idő alacsony hőmérsékleten  $\Theta(\log n)$ , kritikus hőmérsékleten  $\Theta(n^{1/3})$  és magas hőmérsékleten  $\Theta(1)$ . A keverési idő alacsony hőmérsékleten és magas hőmérsékleten a Potts-modellre és az Ising-modellre a Swendsen–Wang-dinamika keverési ideje megegyezik, a kritikus hőmérsékletnél Ising-modell esetén a keverési idő  $\Theta(n^{1/4})$ , Potts-modell esetén a keverési idő  $\Theta(n^{1/3})$ . Amíg az  $n$  csúcsú teljes gráfon a Potts-modellre a kritikus pontban a Glauber-dinamika keverési ideje  $\Theta(n^{4/3})$ , addig a Swendsen–Wang-dinamika keverési ideje Potts-modell esetén a kritikus pontban  $\Theta(n^{1/3})$ .

## 2.4. Markov-lánc Monte Carlo módszerek alkalmazásai

A Monte Carlo szimuláció véletlenszerű objektumok vagy folyamatok generálása számítógép segítségével. Ezek a tárgyak és folyamatok felmerülhetnek a valós életben egy rendszer modellezésének részeként, mint például egy komplex úthálózat, a neutronok szállítása, vagy a tőzsdei piac fejlődése. Sok esetben a Monte Carlo technikákat determinisztikus problémák megoldására alkalmazzák. Ebben az esetben egy adott valószínűségi eloszlásból történő véletlenszerű mintavételezést jelent. A mintavételezésnél a cél az, hogy véletlenszerű objektumokról azok számos realizációját megfigyelve információt nyerjünk. Egy példa erre a szimulációs modellezés, ahol a véletlenszerű folyamat utánozza a valódi élet számos rendszerét, úgy mint a termelési hálózat és távközlési hálózat. A Monte Carlo technikák bizonyítottan nagyon hatékonyak bizonyultak a véletlenszerű struktúrák és gráfok tulajdonságainak tanulmányozásában a statisztikai fizikában, a valószínűségelméletben és a számítástudományban. A ferromágnesesség klasszikus modelljei az Ising-modell és a Potts-modell ezeknek a véletlenszerű struktúráknak a példái. A Monte Carlo technikák szintén kulcs szerepet játszanak a perkolációs elmélet tanulmányozásában, amely a valószínűségelmélet és a statisztikai fizika metszéspontja. A [27] cikk az Ising-modell részvénytiacon való alkalmazását mutatja be. Az Ising-modell használatának hosszú hagyománya van a társadalmi kapcsolatok és szervezetek reprezentálásában. A statisztikai modellek között az Ising-modell fontos szerepet játszik, hogy megértsük a valós pénzügyi piac tényszerű adatainak eredetét. A közgazdasági modellek nagy csoportját hozzá lehet rendelni az Ising-modell számos változatához, hogy figyelembe tudjuk venni azt, hogy a társadalmi hatások milyen szerepet játszanak az egyéni döntések során. A gazdaságfizikában (Econophysics) a kutatók az Ising-modellt a részvénytiacon a vélemények reprezentálására használják. Az egydimenziós Ising-modellt hatékonyan lehet használni a kereskedő döntések bemutatására a vásárlások és eladások során, amely meghatározza a piac általános

viselkedését. Az Ising-modellben az események vagy spinek  $\pm 1$  értéket vehetnek fel, a mágnesezettség a spinek átlagát jelenti. A fizikában az  $m$  mágnesezettség az árbevételnek felel meg. Ha a vásárlók száma nagyobb mint az eladók száma, azaz  $m$  pozitív, az ár egyértelműen emelkedik. Másrészt, ha az eladók száma több,  $m$  negatív az ár pedig csökken. Az egy ügynökre jutó átlagos pénzmennyiséget a gazdaság  $T$  hőmérsékletének, fűtöttségének nevezzük. Amikor  $T$  alacsony a másolási tendencia nagy, és a  $T$  növekedésével a kereskedők önállóan döntenek, és egy meghatározott hőmérsékleten van egy teljes piaci rendellenesség, majd a piac lassan stabilizálódik. A [28] cikk azt mutatja be, hogy a Potts-modell általánosításával modellezhetjük a relatívan komplexen összefüggő adatokat. Ehhez a 2000-es év népsűrűség és éjszakai fény adatokat használnak. Az utóbbi gyakorlatilag olyan éjszakai fotók a földről, ahol a települések mesterséges fényét rögzítik. Számos tanulmány utal éjszakai fényminták és a régiók gazdasági fejlődésére. Az éjszakai fények intenzitása és sűrűsége objektív gazdasági fejlettségi indexnek tekinthető, különösen az alacsony statisztikai rendszerrel rendelkező országok esetén használható ez a módszer.

A statisztikai fizikában és matematikában a perkolációs elmélet egy véletlen gráfon az egymással kapcsolatban lévő klaszterek viselkedését írja le. Tegyük fel, hogy valamilyen porózus anyagra folyadékot öntünk, lejut-e a folyadék a lyukakon, és eléri az anyag alját? Ezt a fizikai kérdést matematikailag modellezzük csúcsok  $n \times n \times n$  három dimenziós hálózataként, ahol az egyes szomszédok közötti élek vagy kötések  $p$  valószínűséggel nyitva vannak (átengedve a folyadékot), vagy  $1 - p$  valószínűséggel zárva vannak, és feltesszük, hogy az élek függetlenek. Egy adott  $p$  esetében mekkora a valószínűsége annak, hogy létezik egy útvonal, amelynek mindegyik éle nyitott a tetejétől az aljáig? Nagyszámú  $n$  esetén ez a viselkedés különösen fontos. Ezt a feladatot kötés perkolációnak hívják, amelyet Broadbent és Hammersley (1957) vezették be a matematikai irodalomba. A komplex hálózatokat, úgy mint a távközlési hálózatok, számítógépes hálózatok, biológiai hálózatok és közösségi hálózatokat tekinthetjük úgy, mint egyes elemeket, amelyeket csúcsok reprezentálnak és élek jelenítik meg az elemek közötti kapcsolatokat. A hálózat elemzésében a hálózat strukturális méretét a perkolációs elmélet alkalmazásával tanulmányozzák. Amikor csúcsok kritikus részét eltávolítják a hálózat apró klaszterekre bomlik. Ez az előbb említett perkoláció jelensége. Az Ising-modell és a él perkoláció közötti kapcsolat meglehetősen erős, a véletlenszerű klaszter modell néven ismert általánosított perkoláción keresztül kapcsolódnak egymáshoz.

## 3. fejezet

# Összefoglalás

Szeretném összefoglalni a dolgozatomban bemutatott eredményeket. A Metropolis–Hastings-algoritmushoz készített programhoz tartozó ábrákból azt a következtetést lehet levonni, hogy az algoritmus által létrehozott Markov-láncok eloszlása mind a diszkrét eloszlás és mind a folytonos eloszlás esetén konvergálnak az elérni kívánt eloszláshoz.

Az Ising-modellre vonatkozó Glauber-dinamika esetén a keverési idő függ a gráf szerkezetétől. Teljes gráfra vonatkozóan a három hőmérsékleti tartomány esetén a keverési időre kapott eredmények a következők: magas hőmérsékletnél  $\beta < 1$  esetén  $t_n = [2(1 - \beta)]^{-1} n \log n$ , kritikus hőmérsékletnél  $\beta = 1$  esetén  $C_1 n^{3/2} \leq t_n \leq C_2 n^{3/2}$  és alacsony hőmérsékletnél  $\beta > 1$  esetén  $C_3(\beta) n \log n \leq t_n \leq C_4(\beta) n \log n$ . Látható, hogy teljes gráfon az Ising-modell esetén kritikus hőmérsékletnél nagyobb a keverési idő, mint alacsony és magas hőmérséklet esetén. A négyzetrácsra vonatkozóan kritikus hőmérsékletnél a keverési idő  $t_n \leq n^c$ .

A fára vonatkozó Ising-modell relaxációs ideje a 2.6 tétel szerint polinomiális minden hőmérsékleti tartományban. A magas hőmérsékleti tartományra vonatkozó bizonyításból kapjuk, hogy a keverési idő magas hőmérséklet esetén  $O(\log n)$  nagyságrendű és  $n$  csúcsra vonatkozó diszkrét idejű blokkdinamika relaxációs ideje legfeljebb  $O(n)$ , így az egy csúcsra vonatkozó Glauber-dinamika relaxációs ideje  $O(1)$ .

A teljes gráfra vonatkozó legalább három állapotú Potts-modell keverési idejére kapott eredmények három hőmérsékleti tartomány esetén: magas hőmérséklet esetén  $t_n = \alpha_1(\beta, q) n \log n$ , kritikus hőmérséklet esetén  $t_n = \Theta_\varepsilon(n^{4/3})$  és alacsony hőmérsékletnél  $t_n \leq C_1 \exp(C_2 n)$ . Teljes gráfon a Potts-modell keverési ideje alacsony hőmérsékletnél nagyobb, mint magas és kritikus hőmérséklet esetén.

Teljes gráfon az Ising-modellre a Swendsen–Wang-algoritmus keverési ideje alacsony

hőmérsékleten  $\Theta(\log n)$ , kritikus hőmérsékleten  $\Theta(n^{1/4})$  és magas hőmérsékleten  $\Theta(1)$ . Teljes gráfon a Potts-modellre a Swendsen–Wang-algoritmus keverési ideje alacsony hőmérsékleten  $\Theta(\log n)$ , kritikus hőmérsékleten  $\Theta(n^{1/3})$  és magas hőmérsékleten  $\Theta(1)$ .

Az előzőekben összefoglaltakból látható, hogy az  $n$  csúcsú teljes gráfon a kritikus Ising-modellre a Glauber-dinamika keverési ideje  $\Theta(n^{3/2})$ , addig a Swendsen–Wang-dinamika keverési ideje kritikus Ising-modell esetén  $\Theta(n^{1/4})$ , ebben az esetben a Swendsen–Wang-dinamika gyorsabb, mint a Glauber-dinamika. Amíg az  $n$  csúcsú teljes gráfon a Potts-modellre a kritikus pontban a Glauber-dinamika keverési ideje  $\Theta(n^{4/3})$ , addig a Swendsen–Wang-dinamika keverési ideje Potts-modell esetén a kritikus pontban  $\Theta(n^{1/3})$ , ebben az esetben is a Swendsen–Wang-dinamika gyorsabb, mint a Glauber-dinamika.



# Irodalomjegyzék

- [1] David A. Levin, Yuval Peres, Elizabeth L. Wilmer, *Markov chains and mixing times*, 1st edition, American Mathematical Society, 2008, 37-270. MR2466937
- [2] David A. Levin, Yuval Peres, Elizabeth L. Wilmer, *Markov chains and mixing times*, second edition of [MR2466937], American Mathematical Society, Providence, RI, 2017, 37-390. MR3726904
- [3] P. Cuff, J. Ding, O. Louidor, E. Lubetzky, Y. Peres, and A. Sly, *Glauber dynamics for the mean-field Potts model*, J. Stat. Phys. **149** (2012), no. 3, 432–477, 1-8. MR2992796
- [4] D. A. Levin, M. J. Luczak and Y. Peres, *Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability*, Probab. Theory Related Fields **146** (2010), no. 1-2, 223–265. MR2550363
- [5] Csizsár Villő, *Diszkrét és folytonos paraméterű Markov láncok*
- [6] E. Lubetzky and A. Sly, *Critical Ising on the square lattice mixes in polynomial time*, Comm. Math. Phys. **313** (2012), no. 3, 815–836. MR2945623
- [7] N. Berger et al., *Glauber dynamics on trees and hyperbolic graphs*, Probab. Theory Related Fields **131** (2005), no. 3, 311–340. MR2123248
- [8] Nathanael Berestycki, *Mixing Times of Markov Chains: Techniques and Examples*, University of Cambridge, 2016
- [9] Geoff Nicholls, *MSc MT15. Further Statistical Methods: MCMC, Lecture 5-6: Markov chains; Metropolis Hastings MCMC*, University of Oxford, 2015, <http://www.stats.ox.ac.uk/~nicholls/MScMCMC15/L5MScMCMC15.pdf>
- [10] <http://hu.m.wikipedia.org/wiki/Monte-Carlo-műsdszer>

- [11] [https://en.m.wikipedia.org/wiki/Ising\\_model](https://en.m.wikipedia.org/wiki/Ising_model)
- [12] [https://en.m.wikipedia.org/wiki/Roy\\_J.\\_Glauber](https://en.m.wikipedia.org/wiki/Roy_J._Glauber)
- [13] [https://vallalatipenzugyek.blog.hu/2017/12/29/egyszeru\\_monte-carlo\\_szimulacio\\_excelben](https://vallalatipenzugyek.blog.hu/2017/12/29/egyszeru_monte-carlo_szimulacio_excelben)
- [14] [https://en.m.wikipedia.org/wiki/Markov\\_chain\\_Monte\\_Carlo](https://en.m.wikipedia.org/wiki/Markov_chain_Monte_Carlo)
- [15] W. Evans et al., *Broadcasting on trees and the Ising model*, Ann. Appl. Probab. **10** (2000), no. 2, 410–433. MR1768240
- [16] [https://en.m.wikipedia.org/wiki/Crystal\\_structure](https://en.m.wikipedia.org/wiki/Crystal_structure)
- [17] [https://en.m.wikipedia.org/wiki/Potts\\_model](https://en.m.wikipedia.org/wiki/Potts_model)
- [18] R. Bubly, M. Dyer, *Path coupling: A technique for proving rapid mixing in Markov chains*, In Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS), 1997, 223-231.
- [19] M.-F. Chen, *Trilogy of couplings and general formulas for lower bound of spectral gap*, in *Probability towards 2000 (New York, 1995)*, 123–136, Lect. Notes Stat., 128, Springer, New York. MR1632643
- [20] F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, in *Lectures on probability theory and statistics (Saint-Flour, 1997)*, 93–191, Lecture Notes in Math., 1717, Springer, Berlin. MR1746301
- [21] V. K. Gore and M. R. Jerrum. *The Swendsen-Wang process does not always mix rapidly*. J. Statist. Phys. **97** (1999), no. 1-2, 67–86. MR1733467
- [22] A. Galanis, D. Štefankovič and E. Vigoda, *Swendsen-Wang algorithm on the mean-field Potts model*, Random Structures Algorithms **54** (2019), no. 1, 82–147. MR3884616
- [23] Long, Yun; Nachmias, Asaf; Ning, Weiyang; Peres, Yuval. *A power law of order 1/4 for critical mean field Swendsen-Wang dynamics*. Mem. Amer. Math. Soc. **232** (2014), no. 1092, vi+84 pp. MR3243141

- [24] E. Luijten *Introduction to Cluster Monte Carlo Algorithms*, Lect. Notes Phys. 703, 13–38 (2006), Department of Materials Science and Engineering, Frederick Seitz Materials Research Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, U.S.A.
- [25] Kari Rummukainen *Cluster algorithms*, University of Helsinki, Department of physics, <https://www.mv.helsinki.fi/home/rummukai/simu/cluster.pdf>
- [26] Dirk P. Kroese, Tim Brereton, Thomas Taimre, Zdravko I. Botev *Why the Monte Carlo Method is so important today*, Journal WIREs Computational Statistics, Volume 6 Issue 6, November 2014, (Pages 386-392)
- [27] Lisha Damodaran and K. M. Udayanandan *Dynamics of stock market, using Ising model*. Hyperion International Journal of Econophysics & New Economy, 11 (2018), 1, 03, 1-8
- [28] Gabriell Máté *A Potts Model for Night Light and Human Population*. Kézirat. arXiv:1501.04229[physics.soc-ph], Institute for Theoretical Physics, Heidelberg University, Philosophenweg 19, Heidelberg, Germany 2015
- [29] [https://en.wikipedia.org/wiki/Percolation\\_theory](https://en.wikipedia.org/wiki/Percolation_theory)
- [30] [https://en.wikipedia.org/wiki/Network\\_science](https://en.wikipedia.org/wiki/Network_science)
- [31] G. Grimmett, *Percolation*, Springer-Verlag, New York, 1989. MR0995460
- [32] G. Grimmett, *The random-cluster model*, Grundlehren der Mathematischen Wissenschaften, 333, Springer-Verlag, Berlin, 2006. MR2243761
- [33] Swendsen, R. H., and Wang, J.-S. *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett., (1987), 58(2):86-88.

## 4. fejezet

# Függelék

```
m<-30
  f<-1:m
  n<-1000; X<-rep(1,n); X[1]<-1
  track<-c(X[1]-runif(1),X[1]+runif(1))
  for (t in 1:(n-1))
    x<-X[t]
    y<-ceiling(m*runif(1))
    a<-min(1,f[y]/f[x])
    U<-runif(1)
    if (U<=a) X[t+1]<-y else X[t+1]<-x
  track<-c(track,x)
  plot(tail(track,100),1:min(100,length(track)),
  main='',type='l',ylab='Trace')
  abline(v=x,lty=3,col="gray")
  hist(X,freq=FALSE,breaks=c(1:30), main="A Markov-lánc hisztogramja és
  az f/sum(f) függvény"); lines(1:m,f/sum(f), col="red")
  dtv<-function(t) 1/2*sum(abs((f/sum(f))-(hist(X[1:t],
  breaks=c(0:30),plot=FALSE)$density)))
  plot(sapply(1:1000,dtv), xlab="Mintanagyság",
  ylab="Teljes variációs távolság")
  f<-function(x)
  dnorm(x,0,1)
  n<-6000; X<-rep(1,n); X[1]<-0
```

```

track<-c(-1,1)
for (t in 1:(n-1))
x<-X[t]
y<-X[t]+rnorm(1,0,1)
a<-min(1,f(y)/f(x))
U<-runif(1)
if (U<=a) X[t+1]<-y else X[t+1]<-x
track<-c(track,x)
plot(tail(track,100),1:min(100,length(track)),
main="",type='l',ylab='Trace')
abline(v=x,lty=3,col="gray")
a<-seq(from=-4,to=4.5,by=0.0011)
b=dnorm(a,mean=0,sd=1)
hist(X,freq=FALSE,
main="A Markov-lánc hisztogramja és a
normális eloszlás sűrűségfüggvénye");lines(a,b, col="red")
w$y<-b
w$x<-a
xv<-c( -3.0, -2.5, -2.0, -1.5, -1.0, -0.5,
0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0 )
yv<-rep(NA,length(xv))
for(k in 1:length(xv)) xi<-xv[k]
xa<-w$x[max(which(w$x<=xi))]
xb<-w$x[min(which(xi<=w$x))]
ya<-w$y[max(which(w$x<=xi))]
yb<-w$y[min(which(xi<=w$x))]
yi<-if(xa==xb) ya else ((xb-xi)*ya+(xi-xa)*yb)/(xb-xa)
yv[k]<-yi plot(a,b)
points(xv,yv,col="red",pch=20)
dtv<-function(t)
1/2*sum(abs(yv-(hist(X[1:t],breaks=c(hist(X)$breaks),
plot=FALSE)$density)))
plot(sapply(1:6000,dtv),xlab="Mintanagyság",
ylab="Teljes variációs távolság")

```