

**Eduardo Serralta Hurtado de Menezes**

*Analizando diferenças ao simular sistemas de spins:  
o que é melhor banho térmico ou Metropolis?*

Porto Alegre

2012

**Eduardo Serralta Hurtado de Menezes**

***Analisando diferenças ao simular sistemas de spins:  
o que é melhor banho térmico ou Metropolis?***

Trabalho de conclusão de curso apresentado  
como parte das atividades para obtenção do  
título de bacharel em Física com ênfase em Pes-  
quisa Básica da Universidade Federal do Rio  
Grande do Sul - UFRGS, Instituto de Física

Orientador:

Prof. Dr. Roberto da Silva

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Porto Alegre

2012

*Este trabalho é dedicado à minha família  
e amigos cujo apoio  
foi fundamental para minha graduação.*

# *Agradecimentos*

Ao término deste trabalho, deixo aqui meus sinceros agradecimentos:

- a minha família;
- aos meus amigos;
- ao Prof. Dr. Roberto da Silva, por toda dedicação, paciência e estímulo em sua orientação;
- a todos os professores do Instituto de Física da UFRGS, por manter a excelência da Universidade;
- à Sociedade Brasileira por financiar meus estudos.

*“ O acaso é conceito mais fundamental que a causalidade.”*

Max Born.

# *Resumo*

Neste trabalho, estudam-se aspectos fundamentais de uma simulação de Monte Carlo. São apresentados conceitos de Mecânica Estatística e Dinâmica Estocástica fundamentais para o desenvolvimento desta ferramenta. Há uma breve explicação do Modelo de Ising por ser o escolhido para aplicação dos algoritmos. Descrevem-se os algoritmos de Monte Carlo de Metropolis e de Banho Térmico. Também é explicada a forma alternativa bastante comum que é na forma sequencial. Estudam-se as diferenças entre os dois algoritmos e suas respectivas alteações sequenciais usando a magnetização como parâmetro. Nos resultados obtidos, percebe-se a validade do Teorema de Peskun. Nota-se também que a alteração sequencial surte um efeito bastante positivo no caso do Algoritmo de Banho Térmico, aumentando sua acurácia, e diminuindo o tempo de convergência. No algoritmo de Metropolis, a alteração sequencial apresenta um comportamento mais complexo.

# *Abstract*

In this work, we studied fundamental aspects of Monte Carlo Markov Chain Methods. Moreover, some basic concepts of Statistical Mechanics and Stochastic Dynamics needed to understand the Monte Carlo Markov Chain Methods are presented. The Ising Model is explained and used to study the two most famous algorithms: Metropolis Algorithm and Gibbs Sampling. A very common alternative form of the algorithms is also presented, the sequential updating form. The difference between Metropolis and Gibbs Sampling is analyzed using the spontaneous magnetization. Peskun's theorem was verified from the data obtained. Two positive effects were noticed in the sequential form of Gibbs Sampling: it has higher accuracy and faster convergence than its random form. In the case of Metropolis algorithm, the sequential form has a more complex behavior and deserves more study.

# Lista de Figuras

6.1	Com a leitura sequencial no algoritmo de Metrópolis, esta configuração será repetida a cada dois passos de Monte Carlo . . . . .	18
6.2	(a): Evolução temporal da magnetização para realizações diferentes realizações usando Metropolis com leitura sequencial. . . . .	20
6.3	(b): A amplitude das flutuações para uso como critério de parada para o algoritmo que calcula a variância média, acurácia média e os tempos para quais as séries convergiram para o valor exato a menos dessa amplitude (precisão) . . . .	21
6.4	Distribuição acumulada dos tempos de convergência para as quatro diferentes formas de implementar o Ising aqui propostas e para $T = T_c/3$ . . . . .	25
6.5	Distribuição acumulada dos tempos de convergências para as quatro diferentes formas de implementar o Ising aqui propostas e para $T = T_c/2$ . . . . .	26
6.6	Distribuição acumulada dos tempos de convergência para as quatro diferentes formas de implementar o Ising aqui propostas e para $T = 2T_c/3$ . . . . .	27



# Lista de Tabelas

6.1	Tabela de erro médio da magnetização dos algoritmos . . . . .	22
6.2	Tabela da Média dos Desvios padrões da média de cada algoritmo, considerando apenas as séries que convergiram . . . . .	24

# Sumário

<b>1</b>	<b>Introdução</b>	<b>4</b>
<b>2</b>	<b>Metodologia</b>	<b>6</b>
<b>3</b>	<b>Desenvolvimento Teórico</b>	<b>7</b>
3.1	Mecânica Estatística . . . . .	7
3.2	Dinâmica Estocástica . . . . .	8
3.2.1	Processo estocástico de Markov . . . . .	9
3.2.2	Equação Mestra . . . . .	10
3.2.3	Balanco Detalhado e reversibilidade . . . . .	10
<b>4</b>	<b>Modelo de Ising</b>	<b>11</b>
<b>5</b>	<b>Algoritmo de Metropolis e Banho Térmico</b>	<b>13</b>
5.1	Algoritmo de Metropolis . . . . .	13
5.2	Algoritmo de Banho térmico . . . . .	14
<b>6</b>	<b>Análise dos Algoritmos e Resultados</b>	<b>16</b>
6.1	Aspectos iniciais dos programas . . . . .	16
6.2	Atualização Sequencial e Atualização Aleatória . . . . .	17
6.3	Comparação entre os Algoritmos . . . . .	19
6.3.1	Acurácia da Solução . . . . .	19
6.3.2	Desvio Padrão da Solução . . . . .	22
6.3.3	Tempo de Convergência . . . . .	24
<b>7</b>	<b>Considerações finais</b>	<b>28</b>

# Capítulo 1

## Introdução

Simulações de Monte Carlo com cadeias de Markov são uma ferramenta que ganha cada vez maior importância na Física. Ela vem se solidificando desde o advento da computação eletrônica até chegar aos clusters dos dias atuais, atividades que são amplamente utilizadas atualmente. O grande aumento da capacidade computacional permite cada vez mais realizar simulações e analisar problemas de extrema complexidade. Anteriormente insolúveis, tais problemas hoje são calculados com erros e tempos cada vez menores.

Com o crescente poder desta ferramenta, cunhou-se o termo **física computacional**, que é muito abrangente no contexto da física estatística. Ele pode referir-se a uma gama de diferentes aplicações que vão desde simulações de sistemas de longo alcance, aqui inclui-se por exemplo comportamento anômalo da água, simulação de enovelamento de proteínas, passando pela vastidão dos ditos sistemas de spins. As simulações também têm se mostrado úteis para explicar fenômenos socio-econômicos por meio de uma área relativamente nova denominada econofísica.

Nesta área frequentemente lida-se com equações extremamente complexas e/ou sistemas com grande número de graus de liberdade. O desenvolvimento das simulações de Monte Carlo para sistemas físicos é relativamente recente. Seu início é atribuído ao cérebre **algoritmo de Metropolis** desenvolvido em 1953 [10]<sup>1</sup>.

Hoje, há vários algoritmos baseados em geração de Cadeias de Markov com distribuição de probabilidades de estados gerados a partir de testes de Monte Carlo. O trabalho tem como objetivo estudar alguns conceitos relativos a convergência e a estabilidade de duas prescrições fortemente utilizadas para simular sistemas de spins: **o algoritmo de banho térmico** e **o algoritmo de Metropolis**.

Para isso, primeiramente, no capítulo 2, descreve-se muito brevemente o ambiente computacional que utilizamos para a realização do trabalho. Em seguida, analisa-se alguns aspectos teóricos para o desenvolvimento de tais algoritmos, isto é, delinea-se mesmo que de maneira básica, os conceitos de ergodicidade, equação mestra, condição do balanço detalhado e outros conceitos importantes relativo ao estudo. Isso é feito no capítulo 3, onde se descreve especificamente a prescrição de banho térmico, também conhecida como amostragem de Gibbs [4] e a prescrição de Metropolis [10]. No capítulo 4, é introduzido o modelo mais simples para explicar o magnetismo na matéria, o paradigmático modelo de Ising [7]. Neste modelo, serão aplicadas e analisadas nossas duas prescrições. O uso de algoritmos de Monte Carlo baseados

---

<sup>1</sup>O algoritmo de Metropolis está entre os 10 algoritmos mais utilizados do século passado, vide por exemplo: <http://www.uta.edu/faculty/rcli/TopTen/topten.pdf>

em cadeias de Markov é muitas vezes indiscriminado, ignorando possíveis nuances que cada algoritmo pode trazer ao sistema estudado. Por isso, fez-se análise de alguns aspectos importantes. A acurácia (diferença em valor absoluto entre a magnetização exata e a magnetização simulada) e a precisão (variância assintótica) bem como o tempo de convergência são estudados no capítulo 6.

Pelo menos quanto à forma e à análise em que se apresentam os resultados, acreditamos ter aspectos inéditos. A abordagem proposta neste trabalho pode ser de grande utilidade a futuros estudantes que venham a explorar esse tema, já que dificilmente se encontra na literatura dados sobre as taxas de convergência de cada algoritmo em função do número de passos.

# Capítulo 2

## Metodologia

O trabalho foi feito primeiramente analisando o conteúdo teórico disponível sobre simulações de Monte Carlo em livros acadêmicos e artigos científicos. Todos os códigos construídos neste trabalho de conclusão, seja para fazer propriamente as simulações de Monte Carlo, seja para fazer a análise dos algoritmos, foi realizada por meio da implementação dos algoritmos estudados em linguagem Fortran 77-90 e compilados no cluster ada.if.ufrgs.br. Os algoritmos foram implementados pelo autor do trabalho em conjunto com seu orientador e com auxílio da bibliografia disponível. Também foram acrescentadas modificações aos algoritmos de forma a simplificar a análise. Os algoritmos foram rodados para diferentes sequências pseudo-aleatórias e os resultados obtidos foram comparados com a solução exata no limite termodinâmico [12]. Utilizamos como gerador de números pseudo-aleatórios o ran2.f do livro Numerical Recipes (para maiores informações <http://www.nr.com/>)[15]. Para as visualizações utilizamos o software origin 8.0 (para maiores informações <http://www.originlab.com/>).

# Capítulo 3

## Desenvolvimento Teórico

Neste capítulo, vamos apenas delinear alguns conceitos básicos importantes para entender a origem dos algoritmos utilizados para simular sistemas de spins, dentre os quais destacamos ergodicidade, processo estocástico, processo de Markov, equação mestra e condição de balanço detalhado. Caso o leitor se sinta desconfortável com a maneira superficial em que os tópicos são tratados recomenda-se os livros básicos para preencher as lacunas, que inclusive estão em português: Introdução a Física Estatística, Silvio Salinas, Editora EDUSP [1], e Dinâmica Estocástica e Irreversibilidade de T. Tomé e M. J. de Oliveira também editado pela EDUSP [2].

### 3.1 Mecânica Estatística

A mecânica estatística é a teoria cujo objeto de estudo são sistemas de muitas partículas, na verdade a mecânica estatística tem como objetivo principal dar explicações microscópicas, com uma teoria probabilística/combinatorial aos fenômenos termodinâmicos. Tais problemas têm como característica um número muito grande de graus de liberdade, isto resulta na impossibilidade de resolvê-los analiticamente. O exemplo padrão para demonstrar a utilidade da mecânica estatística considera a dinâmica de um gás monoatômico em um recipiente de aproximadamente 20 litros a temperatura ambiente. Tal sistema apresenta um número de partículas da ordem de  $N \sim 10^{23}$ .

Se supusermos que o recipiente do exemplo está isolado e tem paredes perfeitamente refletoras, ou seja, o sistema conserva energia, sabemos que o sistema irá evoluir seguindo uma trajetória particular no espaço de fase. São necessárias, portanto,  $6N$  equações e suas condições iniciais. Poderíamos a partir do resultado obtido das equações de movimento de cada partícula realizar a soma de quantidades físicas de interesse de todas partículas para obtermos as propriedades termodinâmicas do sistema. Esta resolução é, muitas vezes, impraticável, seja pela complexidade das equações, seja pelo tempo computacional necessário para algoritmos de dinâmica molecular.

A mecânica estatística de equilíbrio propõe uma abordagem alternativa a estes problemas. Ao invés de calcularmos o estado particular das partículas do sistema, as propriedades termodinâmicas são obtidas através de médias sobre todas as possibilidades de configurações que o sistema pode apresentar. O conjunto de sistemas de mesma configuração macroscópica com diferentes configurações microscópicas chama-se ensemble estatístico. O postulado que sus-

tenta a mecânica estatística clássica de equilíbrio supõe a equiprobabilidade dos microestados de mesma energia. Desta forma, é natural supor que, ao esperar um tempo razoável, todos os pontos do espaço de fase serão percorridos. Então se acredita que a grandeza termodinâmica  $O$  que medimos no laboratório através de uma média temporal em intervalo de tempo grande seja igual a média sobre todos os microestados acessíveis que o sistema possa assumir (ensemble). Em suma, para um sistema clássico,

$$\langle O \rangle_{tempo} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T O(t) dt,$$

a média temporal da grandeza  $O$  tomando-se um tempo arbitrariamente grande, deve ser igual a média no ensemble

$$\langle O \rangle_{ensemble} = \frac{\int O(p, q) f(p, q) dp dq}{\int f(p, q) dp dq},$$

onde  $f(p, q)$  descreve a densidade de pontos do espaço de fase. Dizer que estas duas médias são iguais, é algo realmente fundamental na mecânica estatística e denomina-se a tal suposição de **hipótese ergódica**.

Essa hipótese se baseia na conjectura que o sistema físico no equilíbrio se encontrará muito mais tempo nos macroestados em que há maior densidade de microestados e que se computarmos a média de todos estados possíveis, este valor não diferirá significativamente do valor médio sobre todos microestados que o sistema percorreu em um determinado período grande de tempo. A hipótese ergódica impõe que o sistema tenha a propriedade chamada **ergodicidade**, isto é todos os microestados do sistema devem poder ser acessados com o decorrer do tempo. Assim, ao realizar uma simulação de Monte Carlo que evolui um conjunto de spins para explicar fenômenos básicos de magnetismo, esperamos que a dinâmica de Markov que faça essa evolução gere configurações que satisfaçam a propriedade de ergodicidade. Na verdade, ergodicidade para as dinâmicas que regem a evolução destes sistemas é mais importante que a condição de balanço detalhado, esta tida como fundamental para se chegar na distribuição de equilíbrio dos sistemas. Não satisfazer ao balanço detalhado estritamente pode aumentar a velocidade de convergência como acontece no algoritmo de banho térmico, como se nota nos dados obtidos. Contudo, para entender como se estabelecem as dinâmicas que conduzem ao estado de equilíbrio, são necessários alguns conceitos básicos relacionados a dinâmica estocástica.

## 3.2 Dinâmica Estocástica

Chama-se de processo estocástico uma variável que assume valores aleatórios ao longo do tempo. Se tivermos uma equação de algum processo físico regido (em parte, ou totalmente) por uma variável aleatória, dizemos que temos uma dinâmica estocástica. Define-se realização, como uma particular sequência de valores assumidos pela variável aleatória. Supondo uma variável de uma dinâmica estocástica a tempo discreto, chamaremos  $x_j(t_i)$  os valores da variável  $x$  no instante  $t_i$ , com  $i = 1, \dots, n$  para a realização  $j = 1, \dots, m$ . Então a média temporal da variável de  $x$  é dada por:

$$\bar{x} = \frac{1}{n\Delta t} \sum_{i=1}^n x(t_i) \quad (3.1)$$

onde  $\Delta t$  expressa intervalos de tempo discretos supostos iguais, para os quais as transições ocorrem.

A média sobre  $m$  realizações é dada por:

$$\langle x(t_i) \rangle = \frac{1}{m} \sum_{j=1}^m x_j(t_i). \quad (3.2)$$

Se analisarmos o valor de  $x(t_{i+1})$ , ele estará de certa forma relacionado ao valor  $x(t_i)$ . Por exemplo, a posição de uma partícula em movimento browniano não deverá estar muito distante da posição anterior se o intervalo de tempo não for muito grande. Por isso, é fundamental definir a chamada função de autocorrelação de um processo estocástico e como ela varia com o tempo. A função de autocorrelação indica quanto o valor da variável aleatória depende do valor anterior. Podemos expressá-la por

$$\begin{aligned} C(t_i, t_j) &= \langle x(t_i)x(t_j) \rangle - \langle x(t_i) \rangle \langle x(t_j) \rangle \\ &= \frac{1}{m} \sum_{l=1}^m x_l(t_i)x_l(t_j) - \left[ \frac{1}{m} \sum_{l=1}^m x_l(t_i) \right] \left[ \sum_{l=1}^m x_l(t_j) \right] \end{aligned}$$

Quando a média da variável aleatória não depende do tempo e  $C(t_i, t_j)$  depende apenas da diferença  $|t_j - t_i|$ , o processo estocástico é dito estacionário. O comportamento da função de autocorrelação depende obviamente do processo analisado. Entretanto, quando nos referimos a processos estocásticos estacionários em física, como é o caso dos sistemas de spins, é muito corriqueiro, longe da transição de fase, observar um comportamento exponencial com  $|t_j - t_i|$  desses sistemas, o que pode ser expresso como:

$$C(t_i, t_j) \sim \exp(-|t_j - t_i|/\tau), \quad (3.3)$$

onde  $\tau$  define uma escala de tempo na qual a correlação é importante.

### 3.2.1 Processo estocástico de Markov

É natural pensar em um processo estocástico em que o valor de  $x(t_{i+1})$  depende de todos os valores anteriores a ele. Assim, possível é escrever isso como uma probabilidade condicional:

$$\Pr [x(t_{i+1})] = \Pr [x(t_{i+1}) | x(t_i), x(t_{i-1}), x(t_{i-2}), \dots, x(t_0)].$$

Uma classe de processos importantes na física são aqueles em que a probabilidade de um estado seguinte depende unicamente do estado atual. Estes são chamados **processos de Markov**. Um processo é chamado **markoviano** se sua probabilidade de transição de um estado para outro pode ser expressa pela relação

$$\Pr [x(t_{i+1})] = \Pr [x(t_{i+1}) | x(t_i)].$$

A sequência de valores gerados através deste processo é o que denominamos **cadeia de Markov**.

A importância dos processos de Markov se deve, principalmente, ao fato de vários processos serem markovianos em determinada escala de tempo. Este argumento é corroborado ao observar o comportamento exponencial de  $C(t_i, t_j)$  na eq. 3.3, que tende a zero quando  $|t_j - t_i|$  for substancialmente maior que  $\tau$ . O tratamento matemático de um processo de Markov é muito mais simples que um processo estocástico genérico cujas probabilidades de transição sejam correlacionadas com vários valores anteriores. A dinâmica dos processos de Markov é dada pela **equação mestra**.



### 3.2.2 Equação Mestra

A equação mestra descreve a evolução temporal da probabilidade do sistema se encontrar em determinado estado microscópico. Sua construção é feita a partir de uma aproximação de primeira ordem. Suponhamos que  $T(w \rightarrow x)$  represente a densidade de probabilidade por tempo (taxa) de que o processo de Markov no estado  $w$  passe para o estado  $x$ , e que  $\Pr(x, t)$  seja a probabilidade do sistema estar no estado  $x$  no instante  $t$ . Então, a variação no tempo da probabilidade do sistema se encontrar em  $x$  é dada pela equação

$$\frac{d}{dt} \Pr(x, t) = \sum_w [\Pr(w, t)T(w \rightarrow x) - \Pr(x, t)T(x \rightarrow w)],$$

que é frequentemente chamada de equação mestra e por alguns de equação de Chapman-Kolmogorov. A interpretação dela pode ser memorizada da seguinte forma, a mudança no tempo da probabilidade de permanência no estado  $x$  é a diferença entre o que 'entra' e o que 'sai' do estado.

### 3.2.3 Balanço Detalhado e reversibilidade

A equação mestra rege a evolução temporal de qualquer processo de Markov, mas nem todo processo de Markov vai para um estado estacionário. A condição para um processo de Markov convergir para um estado estacionário é a exigência que  $d\Pr(x, t)/dt = 0$ . Nesta situação  $\Pr(x, t) = \pi(x)$  é dita a distribuição estacionária do sistema e neste caso dizemos que a condição de balanço é satisfeita, isto é

$$\sum_w [\pi(w)T(w \rightarrow x) - \pi(x)T(x \rightarrow w)] = 0 \quad (3.4)$$

Assim, sob esta condição, esperamos que para um tempo suficientemente longo  $\Pr(x, t) \rightarrow \pi(x)$ . Um estado estacionário não necessariamente estabelece **reversibilidade microscópica** do sistema, isto é

$$\pi(w)T(w \rightarrow x) = \pi(x)T(x \rightarrow w). \quad (3.5)$$

Estados estacionários que satisfazem a essa particular condição, conhecida como condição do balanço detalhado, são chamados de estados de equilíbrio e a distribuição estacionária  $\pi(x)$ , neste caso, é denominada como a distribuição de equilíbrio do sistema.

A condição do balanço detalhado 3.5 é mais restritiva que de estado estacionário 3.4. A condição 3.5 impõe que cada termo do somatório seja nulo. Ela satisfaz não apenas a condição de estado estacionário, como também diz que, no equilíbrio, o estado será tão mais populado quanto maior for a taxa de transição para ele. E será populado de forma inversamente proporcional a taxa de saída do estado. Não necessariamente existe uma única escolha que satisfaça ao balanço detalhado e que conduza ao estado de equilíbrio. Duas escolhas importantes, que definem taxas que realizam a evolução para a distribuição de equilíbrio de sistemas de spins são as prescrições de Metropolis e de banho térmico. Um sistema de spins particularmente importante onde se pode estudar tais evoluções é o modelo de Ising em duas dimensões, que foi o escolhido para esse trabalho pela sua simplicidade e existência de solução exata [12].

# Capítulo 4

## Modelo de Ising

Um dos temas de grande interesse na física é o magnetismo na matéria. Estudar propriedades macroscópicas, como a magnetização em um ferromagneto, a partir de interações microscópicas e sob quais condições estas propriedades se manifestam é de fundamental importância tanto do ponto de vista teórico como da aplicação tecnológica.

A interação dos spins foi calculada por W. Heisenberg [6] via o chamado hamiltoniano de Heisenberg sem a presença de campo externo:

$$\mathcal{H} = - \sum_{ij} J_{ij} \vec{S}_i \cdot \vec{S}_j,$$

na qual  $\vec{S}_i$  é o spin do  $i$ -ésimo elétron e  $J_{ij}$  representa a integral de troca do  $i$ -ésimo com o  $j$ -ésimo spin. Para uma melhor compreensão do que vem a ser uma integral de troca, pode se consultar por exemplo o livro Introduction to Quantum Mechanics de David J. Griffiths, da editora Prentice Hall [5]. As componentes  $S_i^x$ ,  $S_i^y$  e  $S_i^z$  são operadores que não comutam entre si. O tratamento via mecânica quântica não conduz a resultados exatos para propriedades termodinâmicas além de ser muito complicado de se estudar rigorosamente. Observa-se que, se  $J_{ij} > 0$ , os elétrons com alinhamento paralelo tem menor energia, e quando  $J_{ij} < 0$  o alinhamento anti-paralelo será privilegiado. Os materiais com  $J_{ij} > 0$  exibem comportamento ferromagnético enquanto os que tem  $J_{ij} < 0$  tem comportamento anti-ferromagnético.

Ernst Ising, orientado por W. Lenz, propôs tratar o hamiltoniano de Heisenberg utilizando algumas simplificações. Se o arranjo onde estão os átomos possuir uma grande anisotropia axial, ou um campo forte aplicado, podemos ignorar a contribuição de energia das componentes dos spins não paralelas a este eixo. Considerando a distância entre os spins vizinhos como sendo a mesma, então é razoável considerar que as integrais de troca tem um valor constante para os primeiros vizinhos em um reticulado. Supondo que a distância dos segundos vizinhos é grande,  $J$  torna-se nulo para os spins mais distantes. E. Ising também inseriu a interação dos spins com um campo magnético externo no seu modelo. Portanto podemos espessar o hamiltoniano do modelo de Ising, considerando  $N$  spins, através de:

$$\mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j + H \sum_{i=1}^N \sigma_i.$$

Agora, os  $\sigma_i$  não são mais operadores, mas representam a componente paralela ao eixo do campo magnético do  $i$ -ésimo spin, que podem assumir apenas dois valores  $\pm 1$  onde o valor  $\hbar/2$

foi incorporado a constante de acoplamento  $J$ . A  $\langle \cdot \rangle$  denota que a soma na equação deve ser realizada apenas para os primeiros vizinhos. O trunfo do modelo de Ising se deve a ele ser um modelo não-trivial com solução exata para campo  $H = 0$  em uma dimensão e em duas dimensões para uma rede quadrada [7] [12].

A solução em uma dimensão foi dada pelo próprio E. Ising [7], sua solução não prevê magnetização espontânea como podemos ver na sua expressão para a magnetização por spin:

$$m_{1d}(\beta, H) = \frac{\sinh(\beta H)}{[\sinh^2(\beta H) + e^{-4\beta J}]^{1/2}},$$

isto é  $m_{1d}(\beta, H = 0) = 0$ , onde  $\beta = 1/(k_B T)$  com  $k_B$  sendo a constante de Boltzmann.

Em duas dimensões, a solução foi dada por Lars Onsager [12]. A solução prevê magnetização espontânea, pois para  $H = 0$ , temos que a magnetização por spin é dada por

$$m_{2d}(T) = \left[ 1 - \sinh^{-4} \left( \frac{\ln(1 + \sqrt{2})T_c}{T} \right) \right]^{1/8}. \quad (4.1)$$

Os resultados de Onsager também determinam que o sistema apresenta uma transição de fase de segunda ordem (contínua) em

$$T_c = \frac{2J}{k_B \ln(1 + \sqrt{2})}.$$

Isto é, para  $T < T_c$ , o sistema apresenta magnetização espontânea enquanto que para  $T \geq T_c$  o sistema encontra-se completamente desordenado e portanto não apresenta magnetização espontânea. É comum se referir a este tipo de transição como ordem-desordem.

O modelo de Ising em 3 dimensões ainda não apresenta solução analítica, mas mesmo assim é possível provar a existência de magnetização espontânea através do argumento de Peierls [13] que se aplica para duas ou mais dimensões. No próximo capítulo, aplicam-se diferentes algoritmos para estudar a evolução temporal do modelo de Ising em 2 dimensões, bem como sua convergência ao estado de equilíbrio.

# Capítulo 5

## Algoritmo de Metropolis e Banho Térmico

Neste capítulo, exploram-se de maneira rápida duas dinâmicas importantes para evoluir o modelo de Ising. Na primeira seção, estuda-se o algoritmo de Metropolis. A seguir, na segunda seção, o algoritmo de banho térmico, também conhecido por amostragem de Gibbs, é estudado. Os dois modelos apresentados estão na versão de troca única de spin (do inglês: single spin flip).

### 5.1 Algoritmo de Metropolis

O algoritmo de Metropolis é de fundamental importância, pois combinou os conceitos de cadeia de Markov com o método de Monte Carlo com amostragem por importância para estudar comportamentos de sistemas de grande número de graus de liberdade. A ideia fundamental nesse algoritmo é gerar uma cadeia de Markov de estados e calcular as propriedades físicas trocando complicadas médias configuracionais por simples médias aritméticas sobre os passos de Monte Carlo usando a propriedade de ergodicidade. O algoritmo de Metropolis define as taxas de transição entre estados de forma que a razão entre as taxas seja igual a razão entre as probabilidades dadas pela distribuição de equilíbrio térmico. Dessa forma, tenta-se gerar um processo de amostragem nos estados com maior probabilidade de ocorrência. Em resumo, a amostragem por importância obedecendo a distribuição de Boltzmann é feita por itermédio de uma cadeia de Markov de estados gerada com as taxas de transição ajustadas. O passo de Monte Carlo do algoritmo de Metropolis (de troca de um único spin) aplicado ao modelo de Ising para uma rede com  $N$  spins, sendo  $N = L^2$ , pode ser assim resumido:

#### Descrição

- 1 Sorteia-se um sítio da rede e inverte-se a orientação do spin;
- 2 Calcula-se a mudança de energia de tal troca  $\Delta E$ ;
- 3 Se a troca diminui ou não altera a energia, aceita-se a troca;
- 4 Se a troca aumenta a energia, sorteia-se um número pseudo-aleatório uniforme  $r$  entre  $[0, 1]$ , se  $r < e^{-\Delta E/k_B T}$ , aceita-se a troca, caso contrario rejeita-se.
- 5 Repete-se o mesmo procedimento de 1 à 4 outras  $N - 1$  vezes;

O algoritmo descrito apresenta ergodicidade uma vez que todos os estados do sistema são acessíveis a partir de qualquer estado. Outro aspecto importante é que o algoritmo, por construção, satisfaz ao balanço detalhado. Considerando que a probabilidade de ocupação de um sistema em contato com um reservatório térmico é dado pela distribuição de Boltzmann, a probabilidade do sistema estar no estado  $\zeta$  com energia  $E(\zeta)$  é dada pela distribuição canônica:

$$\Pr(\zeta) = \frac{e^{-\beta E(\zeta)}}{\sum_{\zeta^* \in \Omega} e^{-\beta E(\zeta^*)}}$$

onde  $\Omega$  denota o conjunto de todos os estados acessíveis ao sistema.

$$\frac{\Pr(\zeta)}{\Pr(\zeta^*)} = e^{-\frac{(E(\zeta) - E(\zeta^*))}{k_B T}}$$

Se o sistema se encontra no estado  $\zeta$  e sorteamos a mudança para  $\zeta^*$ , sua aceitação será  $\exp[-(E(\zeta^*) - E(\zeta))/k_B T]$  se a energia  $E(\zeta)$  for menor que  $E(\zeta^*)$ . Agora estando em  $\zeta^*$ , se sorteamos a mudança para o estado  $\zeta$  novamente, sua aceitação será 1. Portanto, o quociente entre as taxas de mudança é igual ao quociente das probabilidades, que é exatamente a condição de balanço detalhado 3.5.

## 5.2 Algoritmo de Banho térmico

Outra forma de simular o mesmo sistema de spins é com o algoritmo de banho térmico. Este algoritmo foi primeiramente proposto por S. Geman e D. Geman [4] em 1984. O algoritmo do banho térmico também se baseia em fazer uma amostragem por importância no espaço de fase. Este algoritmo é conceitualmente mais fácil de entender em comparação com o Metropolis. Ele determina uma probabilidade de orientação do spin baseado na orientação de seus vizinhos independente do seu estado atual.

Baseado na configuração dos spins vizinhos, calculam-se as energias correspondentes a cada estado possível de ser atingido (inclusive permanecer no mesmo). Como no caso do algoritmo de Metropolis, calculam-se as probabilidades de transição (permanência) para cada estado, utilizando a distribuição de Boltzmann. No modelo de Ising, a energia de interação depende apenas das interações com os primeiros vizinhos, então a probabilidade de orientação do spin depende apenas da orientação dos primeiros vizinhos. Definindo  $S_i$  como sendo a soma dos spins vizinhos ao spin lido  $\sigma_i$ , o passo de Monte Carlo deste algoritmo na sua forma de troca única de spin pode ser descrito da seguinte forma:

### Descrição

- 1 Sorteia-se um sítio da rede;
- 2 Lê-se a orientação dos spins vizinhos e calcula-se  $S_i$  ;
- 3 Define-se que a probabilidade do spin lido transicionar para  $\sigma_i = +1$  como
 
$$p(\cdot \rightarrow +) = \frac{e^{\beta JS}}{(e^{\beta JS} + e^{-\beta JS})};$$
- 4 Sorteia-se um número pseudo-aleatório  $r$  entre  $[0, 1]$ . Se  $r < p(\cdot \rightarrow +)$ ,  $\sigma_i := +1$  senão  $\sigma_i := -1$
- 5 Repete-se o mesmo procedimento de 1 à 4 outras  $N - 1$  vezes;

Obviamente, a probabilidade do spin se orientar para baixo  $p(\cdot \rightarrow -) = 1 - p(\cdot \rightarrow +)$ . Poderíamos alternativamente definir o passo 3 e 4 como função de  $p(\cdot \rightarrow -)$ , fazendo as devidas mudanças.

Novamente, observamos que esse algoritmo apresenta ergodicidade, uma vez que é possível chegar a qualquer configuração de spins alterando individualmente os spins.

Este algoritmo também satisfaz o balanço detalhado, pois

$$\frac{p(\cdot \rightarrow -)}{p(\cdot \rightarrow +)} = \exp(-2\beta JS) = \frac{\Pr[E(\sigma_i = -1)]}{\Pr[E(\sigma_i = +1)]},$$

que é equivalente à eq. 3.5.

Uma vez discutidos os algoritmos, no próximo capítulo apresentaremos uma análise quantitativa das principais diferenças entre eles.

# Capítulo 6

## Análise dos Algoritmos e Resultados

Nesta parte se faz uma análise e comparação de diversos pontos entre o algoritmo de banho térmico e Metrópolis, fazendo leituras sequencial ( que não satisfaz ao balanço detalhado) e aleatória ( que satisfaz ao balanço detalhado) e se apresentam alguns resultados comparativos exploratórios que tentam analisar o limite de validade do teorema de Peskun [14] para a variância assintótica. Também se explora como a forma de leitura influencia o tempo de convergência de cada algoritmo.

### 6.1 Aspectos iniciais dos programas

Para efeitos de comparação entre os dois algoritmos que estudamos, ambos foram implementados para uma única e mesma rede regular, quadrada, de lado  $L = 100$ , portanto com 10000 spins, pois o interesse aqui não é o estudo de efeitos de escala e mas uma comparação, ainda que empírica, da convergência e estabilidade (permanência nos estados de equilíbrio) das simulações. A rede com essas dimensões, sem considerações de simetria, comporta  $2^{10^4}$  possíveis configurações diferentes, o que equivale a ordem de grandeza de  $10^{3010}$ .

Os algoritmos foram implementados utilizando as chamadas condições periódicas de contorno. Essa condição prescreve que o vizinho do primeiro spin de uma linha (ou coluna) é igual ao último spin da linha (coluna) e vice-versa. Sabemos que essa condição gera simetrias frente a translações na rede. Há também simetria frente a inversão de todos de spins em campo nulo. Contudo, mesmo que tais considerações possam diminuir consideravelmente a quantidade de cálculos necessários ainda assim não seriam suficientes para realizar uma média das variáveis termodinâmicas em todo espaço de fase. Isto significa que não há, na prática, solução analítica para esta rede, mesmo que finita.

Desta forma, os resultados obtidos e que serão apresentados neste capítulo sempre foram comparados com o resultado exato obtido por Onsager no limite termodinâmico ( $L \rightarrow \infty$ ). Isso direcionou para um critério de parada (ainda que empírico) para observar as séries que convergiam para um determinado tempo decorrido de simulação como poderá ser visto e descrito ainda neste capítulo.

Cabe ressaltar que, caso diminuíssemos muito o tamanho a rede, poderíamos calcular os resultados exatos para redes finitas, contudo estes resultados não coincidem e até escondem efeitos fundamentais previstos nos cálculos teóricos no limite termodinâmico. As divergências que existem no limite termodinâmico não existem em redes finitas (vide pg.20 de [11]). Em

uma rede pequena essas curvas são mais suaves que redes grandes, como pode se ver na pg.79 de [8]. Por isso, foi escolhido um tamanho de rede que possibilitasse ter algum grau de confiabilidade no que se refere a tempo de convergência e estabilidade da solução no equilíbrio.

Certamente, com esse tamanho de rede, não será percorrida uma fração significativa do espaço de fases. Portanto, ressalta-se a importância de gerar uma dinâmica que corresponda a uma amostragem tendenciosa dos estados de equilíbrio. Como condição inicial, foram escolhidas configurações aleatórias. Cada ponto da rede tem cinquenta por cento de probabilidade de se encontrar orientado para cima ou para baixo. Tal condição inicial é chamada na literatura de condição de temperatura infinita. Assim, as simulações correspondem a relaxações que conduzem o sistema para o equilíbrio, isto é uma termalização, uma vez que o sistema inicialmente a uma temperatura infinita é resfriado rapidamente (em inglês: quenching) a temperatura de simulação  $T < T_c$ . O que estamos simulando é justamente esse trajeto em direção a uma solução estacionária, o que pode ser diferente para cada uma das dinâmicas.

Como apresentado no desenvolvimento teórico, o modelo de Ising tem como parâmetro crítico uma relação entre a constante de Boltzmann  $k_B$ , temperatura e o valor de acoplamento  $J$ . Nos resultados a seguir, todas as grandezas estão em um sistema de unidades em que  $J/k_B = 1$ . Desta forma, a única dependência do sistema físico para a transição de fase é da temperatura.

## 6.2 Atualização Sequencial e Atualização Aleatória

Uma forma bastante comum de encontrarmos os dois algoritmos é a forma sequencial [16]. Na forma sequencial, não se sorteia o sítio alterado. Alternativamente, os testes são feitos percorrendo a rede de forma ordenada passando por todos os sítios em cada leitura da rede. Um exemplo de forma sequencial é começar testando um sítio e logo depois o sítio imediatamente a esquerda, e repetir o processo até testar toda linha. Depois repetir o processo para a linha imediatamente abaixo até percorrer todas linhas. Esta alteração é interessante pois permite a eliminação dos sorteios do sítio, tornando cada teste mais rápido. Outro aspecto desta forma de leitura da rede é a facilitação do processo de paralelização do programa. Entretanto, essa alteração tem seu custo. O algoritmo **não** apresenta **ergodicidade**, tampouco obedece ao **balanço detalhado**. Algumas configurações iniciais podem levar o algoritmo a passar sempre pela mesma sequência de estados. Um exemplo trivial de configuração inicial é representado pelo sistema  $2 \times 2$  da figura 6.1. Observe que ele não será ergódico se a leitura for sequencial e na ordem  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ , pois tal configuração se repetirá a cada 2 passos se for aplicada a prescrição de Metropolis.

O algoritmo assim implementado não satisfaz ao balanço detalhado uma vez que sempre é possível o sistema transicionar para uma configuração que difira apenas no spin a direita de um dado sítio, entretanto o sistema jamais transicionará para o estado anterior no próximo teste.

Embora a forma sequencial do algoritmo não tenha estes requisitos, vemos que o algoritmo obtém resultados bons de forma rápida [16]. Um estudo sobre o funcionamento da forma sequencial é encontrado em [9]. O trabalho apresenta argumentos com os quais se mostra que para convergir ao equilíbrio o algoritmo não necessita obedecer ao balanço detalhado em qualquer tempo, desde que a dinâmica seja regular. Ou seja, deve existir um número inteiro não



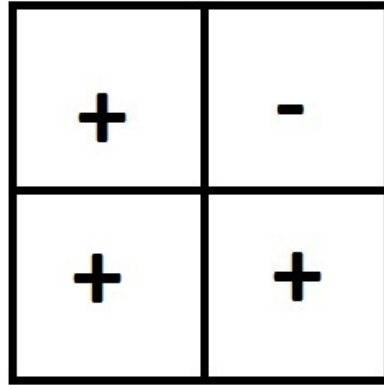


Figura 6.1: Com a leitura sequencial no algoritmo de Metrópolis, esta configuração será repetida a cada dois passos de Monte Carlo

nulo de testes no qual é possível acessar qualquer configuração independentemente do estado do sistema. Isto matematicamente pode ser dito da seguinte forma: a matriz de transição  $A$  entre os estados do problema (e aqui não nos restringimos apenas aos sistemas de spins) deve satisfazer à seguinte propriedade de regularidade na amostragem:

$$[A^m]_{ij} > 0 \text{ para todo } i, j \text{ para algum } m \text{ fixo}$$

onde  $[A^m]_{ij}$  denota a probabilidade de que o estado  $i$  transicione ao estado  $j$  em  $m$  passos.

Desta forma, a cadeia de Markov produzida em questão convergirá para uma cadeia com distribuição de probabilidades estacionária correspondente a distribuição definida pelo teste. Para o algoritmo de Metropolis na forma sequencial convergir sempre para a distribuição de Boltzmann basta subtrair um dado número entre zero e 1 de todas as probabilidades de transição. A transformação é descrita pela expressão:

$$A \rightarrow (1 - \alpha)A + \alpha\mathcal{I}.$$

Assim, tornam-se diferentes de 1 as transições que abaixam ou não alteram a energia do sistema. Também se alteram as taxas de transição que aumentam a energia. Esta alteração não leva o algoritmo à condição de balanço detalhado, mas é suficiente para gerar uma cadeia de Markov com distribuição de Boltzmann como distribuição estacionária. É interessante notar que esse número, embora tenha que diferir de zero, pode ser muito pequeno tornando o algoritmo quase igual ao Metropolis sequencial original. Neste trabalho, tal transformação não foi realizada, deixando a análise da influência de  $\alpha$  para estudos posteriores. Portanto, não há garantia teórica que o algoritmo de Metropolis sequencial deste trabalho terá uma distribuição estacionária. Embora, possamos inferir que em uma rede desta dimensão é muito difícil chegar a uma configuração que gere problemas de ergodicidade.

O algoritmo de banho térmico, por outro lado, se implementado na forma sequencial satisfaz a condição de regularidade sem que nenhuma alteração seja necessária. Portanto, não há aparentemente nenhuma desvantagem em realizá-lo assim.

Ainda com respeito ao balanço detalhado, é fácil ver que ele é satisfeito sempre tanto para o Metropolis quanto para o banho térmico se a escolha do spin é feita aleatoriamente e não de maneira sequencial. Suponha no caso do Metropolis por exemplo, com a dinâmica de unico spin trocado ( $\sigma_i \rightarrow -\sigma_i$ ) a probabilidade de se passar de uma energia  $E = -\sigma_i S + \Delta$  para  $\tilde{E} = \sigma_i S + \Delta$ , onde aqui  $S$  denota a soma dos primeiros vizinhos do spin  $\sigma_i$  e  $\Delta$  a energia de todos os spins fora dessa vizinhança, é dada por

$$w(E \rightarrow \tilde{E}) = w(\sigma_i \rightarrow -\sigma_i) = \min\{1, e^{-2\beta\sigma_i S}\}$$

e naturalmente  $w(\tilde{E} \rightarrow E) = w(-\sigma_i \rightarrow \sigma_i) + \sum_j w(\sigma_j \rightarrow \sigma_i) = \min\{1, e^{2\beta\sigma_i S}\}$ , pois a segunda soma é nula pois a troca é de um único spin. Isto implica que se  $w(E \rightarrow \tilde{E}) = 1$ , então  $w(\tilde{E} \rightarrow E) = e^{2\beta\sigma_i S}$  e se  $w(E \rightarrow \tilde{E}) = e^{-2\beta\sigma_i S}$ , então  $w(\tilde{E} \rightarrow E) = 1$ . Como as distribuições de equilíbrio de Boltzman são dadas por  $\Pr(E) = e^{-\beta(-\sigma_i S + \Delta)}/Z$  e  $\Pr(\tilde{E}) = e^{-\beta(\sigma_i S + \Delta)}/Z$  onde  $Z$  é a função de partição do problema temos que  $w(\tilde{E} \rightarrow E) \Pr(\tilde{E}) = w(E \rightarrow \tilde{E}) \Pr(E)$  sempre. Para o banho térmico, a demonstração é análoga.

## 6.3 Comparação entre os Algoritmos

O valor dado por uma simulação de Monte Carlo depende de vários fatores entre eles quantos pontos serão usados para realizar a média, de que instante da simulação será computada a média, da particular sequência de números aleatórios empregada. Neste trabalho, adotou-se como critério de avaliação dos algoritmo o valor absoluto da magnetização do sistema. Tal escolha se justifica por ser o parâmetro de ordem do sistema físico na teoria fenomenológica de Landau (vide pg.309 [1]) e por sua facilidade de cálculo.

A seguir, apresentam-se resultados quantitativos da comparação entre os algoritmos. As quantidades analisadas são: a **acurácia**, a **precisão** e o **tempo de convergência**. A **acurácia** é o quão próximo da solução exata é o valor dado pelo algoritmo. A **precisão** está associada a quanto varia da média os valores calculados pelo algoritmo. E, por final, o **tempo de convergência**, que é a quantidade de passos necessários para que o valor dado pelo algoritmo tenha um valor igual, com uma margem de erro (posteriormente explicitada), ao valor exato.

### 6.3.1 Acurácia da Solução

Um aspecto fundamental dos algoritmos é sua capacidade de calcular corretamente os valores de interesse do sistema físico. Uma qualidade da medida da quantidade física pode ser avaliada com a observação da acurácia. Para tanto, utilizou-se a expressão 4.1 que é a solução exata para o modelo em uma rede infinita. Define-se o erro médio  $e_m$  de acordo com a expressão 6.3.1:

$$e_m = \frac{\sum_{i=1}^W |\bar{m}_i - m_{2D}|}{W},$$

sendo  $W$  o número de realizações que convergiu para um valor e  $\bar{m}_i$  o valor médio da magnetização da  $i$ -ésima realização. A acurácia das soluções foi analisada apenas para as realizações que convergiram para um valor próximo ao dessa a expressão dentro de uma precisão  $\varepsilon$ . O critério para estabelecer o valor de  $\varepsilon$  é um problema neste caso. No entanto, quando se observa a evolução temporal de várias realizações 6.2, percebe-se que é possível determinar um critério visualmente. Primeiramente, evoluímos várias séries temporais da magnetização para uma dada temperatura de acordo com as dinâmicas de Metropolis e de banho térmico lendo a rede de forma sequencial e também aleatória. As figuras 6.2 e 6.3, ilustram essas evoluções para o algoritmo de Metropolis, com leitura sequencial para a temperatura  $T = T_c/2$ .

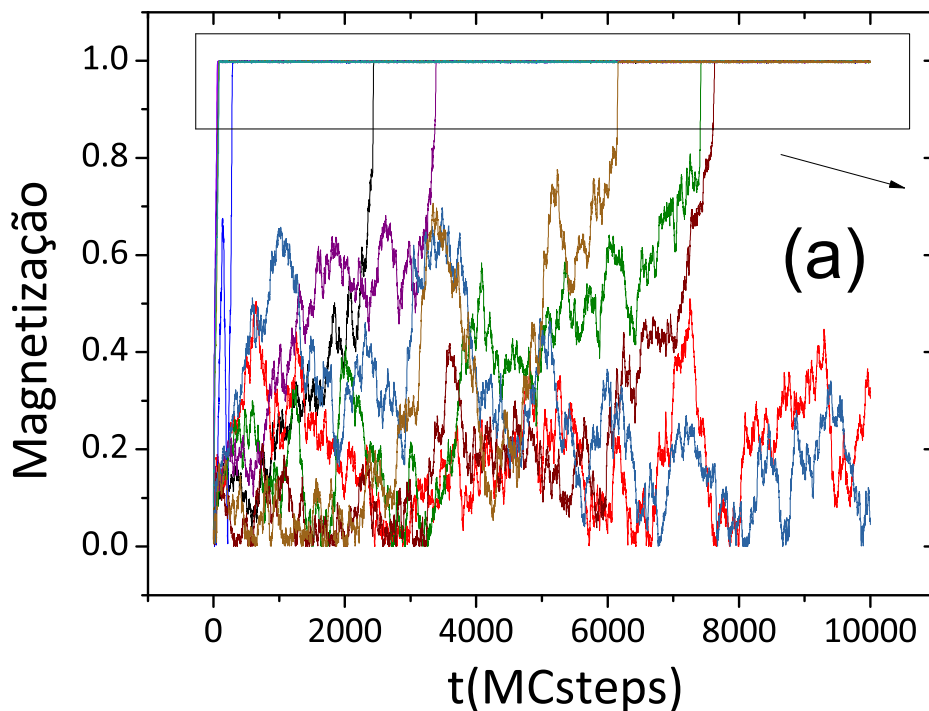


Figura 6.2: (a): Evolução temporal da magnetização para realizações diferentes realizações usando Metropolis com leitura sequencial.

Na figura 6.3 (b), mostra-se a região do retângulo da figura 6.2 (a) ampliada. Analisando as flutuações térmicas, determinou-se a margem de erro  $\varepsilon$  como entrada para nosso algoritmo auxiliar. Neste caso, de forma segura,  $\varepsilon$  é dado por 0.01. Um algoritmo auxiliar computa que a magnetização convergiu para o magnetização de equilíbrio se, em valor absoluto, o valor da magnetização exata( dada por 4.1) menos a média da magnetização no equilíbrio é um número menor que  $\varepsilon$ . O número de realizações que convergem para um dado valor depende do número de passos, este ponto será abordado posteriormente.

Para fins de análise de acurácia das soluções, foram computadas médias temporais da magnetização para cada semente aleatória em cada instante posterior a 2000 passos até 10000,

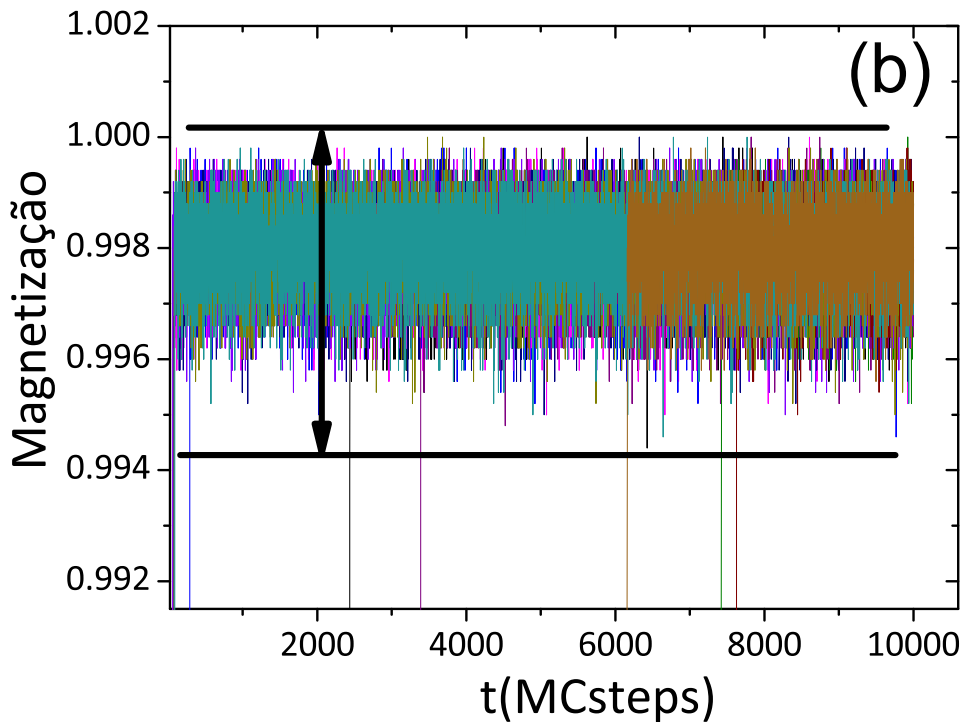


Figura 6.3: (b): A amplitude das flutuações para uso como critério de parada para o algoritmo que calcula a variância média, acurácia média e os tempos para quais as séries convergiram para o valor exato a menos dessa amplitude (precisão)

apenas para as séries que convergiram. O mesmo procedimento foi realizado para instantes posteriores de 5000 passos até 10000. Este procedimento foi repetido para as seguintes temperaturas:  $T_c/3$ ,  $T_c/2$  e  $2T_c/3$ . Os valores determinados para cada temperatura foram  $\varepsilon = 0.004$ , 0.01 e 0.04 respectivamente. Realizou-se 30 diferentes séries com sementes diferentes para cada temperatura para cada algoritmo.

A tabela 6.1 apresenta o erro médio da magnetização (dado por 6.3.1) obtida pelos algoritmos de banho térmico e Metropolis aleatórios e suas respectivas versões sequenciais. A primeira coisa que se nota é que o erro médio é baixo. Todos algoritmos atingem o valor teórico com uma margem de erro bastante baixa. Nota-se também que usar apenas os pontos finais de uma simulação que já tenha convergido não melhora o resultado, pois o uso de maior número de pontos é mais importante neste caso.

Observa-se que usar os 5000 passos finais das simulações que convergiram é pior que usar os 8000 passos finais, neste caso específico. Embora os últimos 5000 passos de uma simulação tenham, em geral, um desvio padrão menor (observe a tabela 6.2) e mais séries que convergiram, seu erro médio é maior.

Tabela de Erro Médio da Magnetização ( $\times 10^{-7}$ )				
Acima de 2000 passos de Monte Carlo				
	Metropolis	Banho Térmico	Metropolis Sequencial	Banho térmico Sequencial
$T_c/3$	17,87	06,65	12,76	09,09
$T_c/2$	83,71	63,96	73,30	61,72
$2T_c/3$	322,17	317,68	297,74	322,23
Acima de 5000 passos de Monte Carlo				
	Metropolis	Banho Térmico	Metropolis Sequencial	Banho Térmico Sequencial
$T_c/3$	19,02	14,43	12,76	11,23
$T_c/2$	112,47	131,61	90,66	88,57
$2T_c/3$	381,01	481,02	348,49	348,46

Tabela 6.1: Tabela de erro médio da magnetização dos algoritmos

Outro aspecto importante é que a alteração sequencial parece melhorar a acurácia no caso do algoritmo de Metropolis mesmo tendo a ergodicidade quebrada. A alteração sequencial no algoritmo de banho térmico parece ter um efeito significativo positivo, embora o erro médio ainda seja maior em alguns casos.

### 6.3.2 Desvio Padrão da Solução

Outra quantidade fundamental na análise da qualidade da solução é a precisão do valor obtido na simulação. Tal quantidade é importante por ser costumeiro tomar a média de apenas uma parte dos valores de uma sequência que convergiu. Cabe ressaltar que essa média de um valor terá associada a ela um desvio padrão da média  $\sigma_m$  dado por 6.1 quando aplicado à magnetização por partícula. É bastante claro que o erro associado a uma medida será tanto menor quanto maior a quantidade de estados amostrados.

$$\sigma_m = \left( \frac{\sum_{i=1}^N (m_i - \bar{m})^2}{N(N-1)} \right)^{1/2} = \frac{\sigma}{\sqrt{N}} \quad (6.1)$$

com  $\bar{m} = (1/N) \sum_{i=1}^N m_i$  a magnetização média tomada sobre  $N$  passos de Monte Carlo.

Peskun em [14] demonstrou um teorema que associa a variância de uma medida através de uma simulação de Monte Carlo a sua taxa de transição de estados na cadeia de Markov. O teorema de Peskun pode ser enunciado da seguinte forma:

**Teorema 7.1:** *Sejam  $P_1$  e  $P_2$  duas matrizes irredutíveis de transição que satisfazem a condição de balanço detalhado para a mesma distribuição de probabilidade de estados  $\pi$ . Se  $P_2 \leq P_1$  isto é, se todas as probabilidades de transição, entre quaisquer par de estados diferentes determinado por  $P_2$  são menores que os correspondentes de  $P_1$ , então o valor da variância assintótica  $\sigma(f, \pi, P_1)$  da quantidade física  $f$ , estimada por  $I_1 = (1/N) \sum_{t=1}^N f(X_t^{(1)})$  associada a cadeia de Markov  $X_t^{(1)}$  evoluído de acordo com  $P_1$  é menor que a variância assintótica  $\sigma(f, \pi, P_2)$  da mesma quantidade física associada a cadeia de Markov evoluída  $X_t^{(2)}$  por  $P_2$ , isto é  $I_2 = (1/N) \sum_{t=1}^N f(X_t^{(2)})$*

**Demonstração:** Vide Peskun [14].

O teorema de Perskun assim mostra que o desvio padrão assintótico, isto é no limite de  $N \rightarrow \infty$ , de uma quantidade física  $m$  estimada por uma cadeia de Markov de estados

$$\sigma(m, \pi, P) = \lim_{N \rightarrow \infty} \left( \frac{\sum_{i=1}^N (m_i - \bar{m})^2}{(N-1)} \right)^{1/2}$$

é uma propriedade associada à matriz de transição do algoritmo. Uma vez que a variância assintótica é tanto menor quanto maior os elementos da matriz de transição, a maior precisão será alcançada maximizando os elementos da matriz de transição, na construção de um algoritmo que obedeça ao balanço detalhado.

Na forma aleatória, podemos definir a transição de estados como o sorteio de uma mudança de posição do spin multiplicada por uma taxa de aceitação. Assim, a probabilidade de sorteio de cada sítio é igual nos dois algoritmos. A aceitação do novo estado correspondente a transição ( $\sigma_i \rightarrow -\sigma_i$ ) será dado pelas expressões 6.2 e 6.3 para o algoritmo de Metropolis e banho térmico, respectivamente:

$$w_M(\sigma_i \rightarrow -\sigma_i) = \min \left\{ 1, e^{-\left(\frac{2JS}{k_B T}\right)} \right\} \quad (6.2)$$

e

$$w_{BT}(\sigma_i \rightarrow -\sigma_i) = \frac{1}{1 + e^{\frac{2JS}{k_B T}}} \quad (6.3)$$

onde  $S$  denota a soma das variáveis de spins dos vizinhos do spin lido:  $\sigma_i$ .

Nas expressões 6.2 e 6.3 para qualquer  $S$  possível, vemos que aceitação do algoritmo de Metropolis é a máxima. Pois a condição

$$e^{-\left(\frac{2JS}{k_B T}\right)} > \frac{1}{1 + e^{\frac{2JS}{k_B T}}} \quad (6.4)$$

uma vez que

$$\frac{e^{\left(\frac{2JS}{k_B T}\right)}}{\left(1 + e^{\frac{2JS}{k_B T}}\right)} < 1 \quad (6.5)$$

pois

$$e^{\pm\left(\frac{2JS}{k_B T}\right)} > 0. \quad (6.6)$$

Também

<b>Média dos desvio padrões da média (<math>\times 10^{-5}</math>)</b>				
<b>Acima de 2000 passos de Monte Carlo</b>				
	Metropolis	Banho Térmico	Metropolis Sequencial	Banho térmico Sequencial
$T_c/3$	10,33	10,34	10,36	10,34
$T_c/2$	67,01	67,25	66,95	67,16
$2T_c/3$	207,97	208,78	208,71	207,45
<b>Acima de 5000 passos de Monte Carlo</b>				
	Metropolis	Banho Térmico	Metropolis Sequencial	Banho Térmico Sequencial
$T_c/3$	10,26	10,33	10,36	10,30
$T_c/2$	67,02	67,20	67,12	67,18
$2T_c/3$	208,23	208,73	208,57	206,94

Tabela 6.2: Tabela da Média dos Desvios padrões da média de cada algoritmo, considerando apenas as séries que convergiram

$$\frac{1}{1 + e^{\frac{2JS}{k_B T}}} < 1, \quad (6.7)$$

para qualquer  $S$ . Contudo, nos casos em que a atualização é sequencial, o teorema de Peskun nada diz. Assim é necessário traçar um comparativo. Para verificar quanto a precisão de um algoritmo é melhor que a do outro, foi construída a tabela 6.2. Nesta tabela, apresentam-se os valores médios dos desvios padrões da média de cada algoritmo nas mesmas condições da tabela anterior. Só foram computadas os desvios padrões para as séries convergiram que de acordo com o critério de acurácia anteriormente comentado em detalhes.

Nela, podemos constatar que os algoritmos obedecem ao teorema de Peskun. Cabe ressaltar também que para a forma sequencial o Teorema de Peskun não se aplica. A tabela 6.2 mostra que em geral o desvio padrão de Metropolis na forma sequencial é maior que no algoritmo de banho térmico. Este ponto ainda merece uma maior investigação numérica e não explou-se tão enfaticamente neste trabalho.

Finalmente, o apresentamos o ponto que carece de maiores discussões na literatura: o tempo de convergência. Na próxima seção, apresentam-se dados que podem ser bastante elucidativos aos usuários destas prescrições.

### 6.3.3 Tempo de Convergência

Uma das características de um bom algoritmo é o tempo de convergência da sua solução, ou seja, quantos passos de Monte Carlo são necessários para a cadeia de Markov gerada pelo algoritmo atingir a distribuição estacionária esperada. Um estudo baseado nos autovalores da matriz de transição dos dois algoritmos é encontrado em [3]. No trabalho, o autor conclui que a velocidade de convergência no modelo de Ising é melhor para o algoritmo de Metropolis para baixas temperaturas. Também conclui que a velocidade de convergência do algoritmo de Metropolis é pior que a do algoritmo de banho térmico a altas temperaturas. Contudo, consideramos trabalho bastante inconclusivo e com pequena faixa de aplicabilidade. Ainda

sobre esse trabalho, é de difícil comparação com resultados práticos em simulações Monte Carlo.

Para verificar a velocidade de convergência, novamente se recorreu às simulações. Assim, 10000 diferentes realizações para cada uma das três temperaturas  $T = T_c/3$ ,  $T = T_c/2$  e  $T = 2T_c/3$  foram analisadas para cada algoritmo e sua respectiva forma sequencial. Neste caso, o algoritmo auxiliar fornece o passo onde os critérios de paradas (mesmos da acurácia) são satisfeitos. As distribuições acumuladas dos tempos de convergência de cada caso são apresentadas no gráfico 6.4 para  $T = T_c/3$ , no gráfico 6.5 para  $T = T_c/2$ , e no gráfico 6.6 para  $T = 2T_c/3$ .

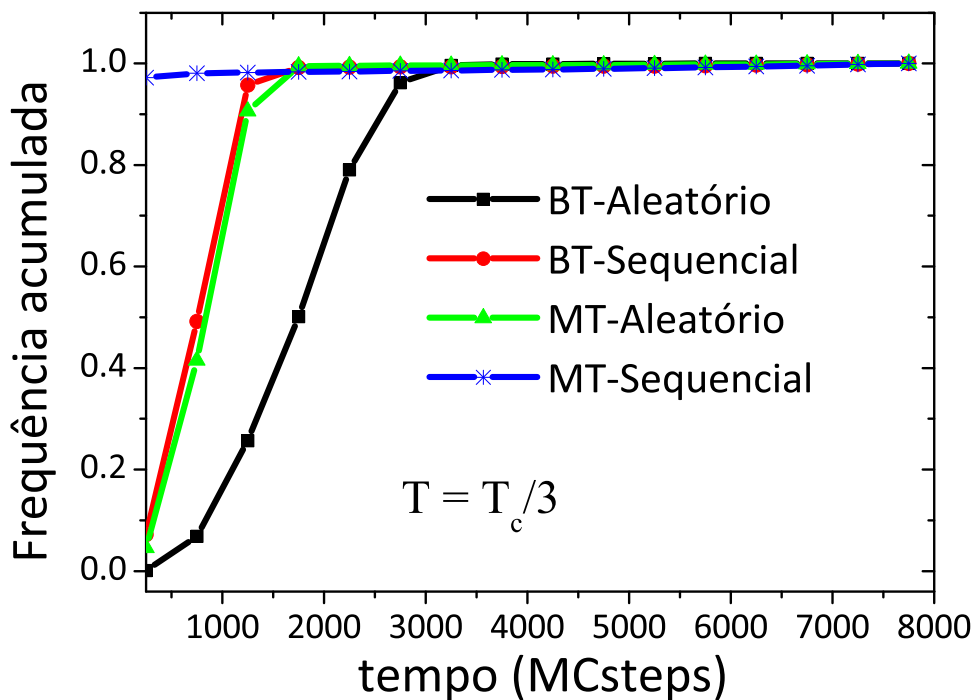


Figura 6.4: Distribuição acumulada dos tempos de convergência para as quatro diferentes formas de implementar o Ising aqui propostas e para  $T = T_c/3$

Ao observar estes gráficos, alguns pontos importantes devem ser analisados. Primeiramente a temperaturas mais altas ( $T = T_c/2$  e  $T = 2T_c/3$ ) no Metropolis sequencial, onde justamente temos um problema de ergodicidade, um fenômeno interessante é observado. Quando as séries convergem, elas convergem rapidamente mas há uma boa parcela delas que persiste a tempos muito longos (mais que  $8 \cdot 10^3$  passos de Monte Carlo). Isso pode ser bastante perigoso para quem faz simulações Monte Carlo, principalmente quando se faz programas em que simplesmente se espera um tempo razoável sem nenhuma verificação formal sobre a convergência e tomam-se as médias arbitrariamente. Neste caso, resultados inverossímeis podem surgir. Contudo, para a temperatura mais baixa estudada  $T = T_c/3$ , se verifica um excelente desempenho para Metropolis sequencial.



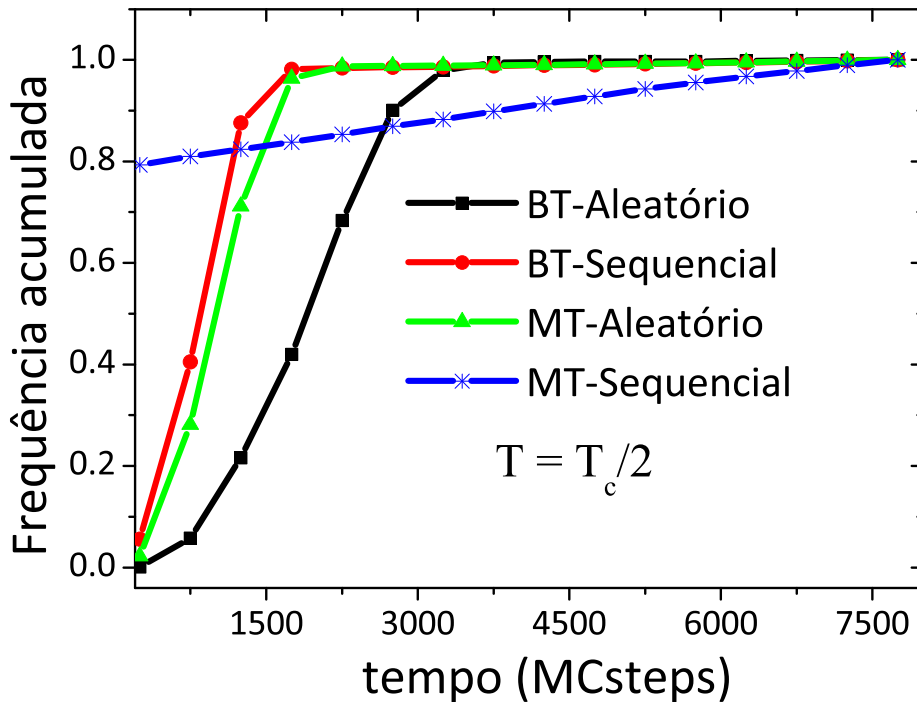


Figura 6.5: Distribuição acumulada dos tempos de convergências para as quatro diferentes formas de implementar o Ising aqui propostas e para  $T = T_c/2$

Outra conclusão bastante importante vem da análise do banho térmico sequencial versus o Metropolis aleatório. Observa-se, nesta comparação, um ligeiro ganho em fazer o banho térmico sequencial, que não satisfaz o balanço detalhado, contra um que satisfaz, que é o Metropolis aleatório. Entretanto, este ganho vai se tornando inexpressivo conforme a temperatura é baixada.

É importante também salientar que para qualquer temperatura estudada qualquer uma das três prescrições banho térmico e Metropolis na forma aleatória ou sequencial têm praticamente 100% das séries convergidas em aproximadamente  $N > N_{\min} = 6000$  passos de Monte Carlo. Portanto, 6000 passos parece ser uma boa estimativa para um tempo de convergência seguro a ser adotado em simulações. Cabe alertar para ser um pouco mais cuidadoso com o caso do Metropolis sequencial.

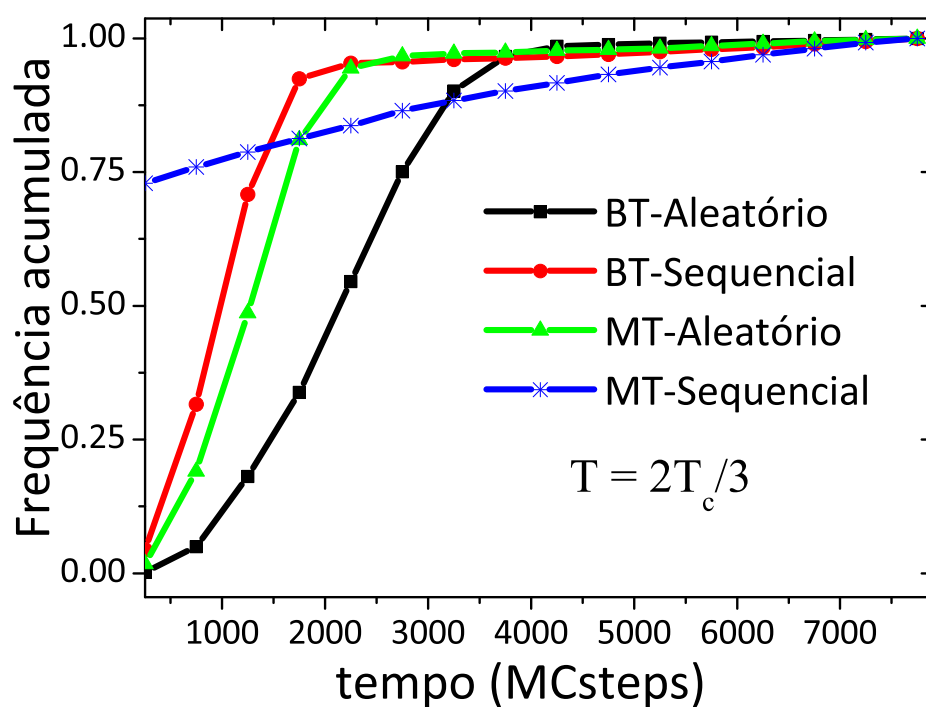


Figura 6.6: Distribuição acumulada dos tempos de convergência para as quatro diferentes formas de implementar o Ising aqui propostas e para  $T = 2T_c/3$

# Capítulo 7

## Considerações finais

Na comparação direta entre os dois algoritmos, não há resposta clara sobre qual algoritmo é melhor para simular o modelo de Ising. O algoritmo de Metropolis, com leitura aleatória, tem uma convergência mais rápida e um desvio padrão assintótico menor que o de banho térmico. Entretanto, a acurácia do algoritmo de banho térmico é melhor.

Sobre a alteração sequencial no algoritmo de Metropolis também não se pode garantir que é melhor ou pior que a forma aleatória. A forma aleatória preenche todos pre-requisitos teóricos embora tenha taxa de convergência maior para uma grande quantidade de passos, a forma aleatória apresenta uma acurácia pior que sua forma sequencial. A forma sequencial não apresenta ergodicidade tampouco obedece ao princípio do balanço detalhado, no entanto tem uma taxa de convergência muito superior aos outros algoritmos com uma quantidade pequena de passos de Monte Carlo. Porém, deve-se salientar que caso a série não convirja em poucos passos, a convergência se dá a um tempo muito longo, o que é indesejável para tal tipo de simulação.

No caso do algoritmo de banho térmico, pelos dados apresentados na tabela de erro médio e nos gráficos de tempo de convergência, conclui-se que a forma mais conveniente de implementação é a forma sequencial, uma vez que os resultados tem, na maioria dos casos, melhor acurácia e menor tempo de convergência. Embora a forma sequencial do algoritmo de banho térmico não obedeça ao balanço detalhado, a convergência a uma cadeia de Markov com a distribuição estacionária de Boltzmann é respaldada por apresentar uma amostragem regular. Este algoritmo, com leitura sequencial, não apresenta o problema das séries de longo tempo de convergência do Metropolis sequencial, é mais rápido que o Metropolis aleatório, tem a melhor acurácia e baixo desvio padrão assintótico. Ainda que o algoritmo de Metropolis aleatório tenha menor desvio padrão assintótico, a diferença é pequena, portanto não é um problema relevante. Por esses motivos, entre os algoritmos estudados, recomenda-se a escolha do banho térmico sequencial.

Os resultados apresentados não são definitivos embora deem uma boa noção do comportamento das soluções. Uma quantidade maior de estatísticas bem como um acompanhamento mais detalhado das quantidades apresentadas em função da temperatura são necessários. O estudo desses algoritmos deve ser continuado e sua influência sobre as ditas simulações dependentes do tempo e considerando também análise de outras quantidades físicas, tais como a energia e susceptibilidade magnética. A determinação dos critérios de convergência em função de alguma quantidade termodinâmica como a susceptibilidade magnética e a influência da condição

inicial na dinâmica do sistemas são temas que podem ser abordados em estudos futuros.

# Referências Bibliográficas

- [1] S.R. De Azevedo Salinas. *Introdução a Física Estatística Vol. 09*. EDUSP, 1997.
- [2] M.J. DE OLIVEIRA. *Dinâmica Estocástica e Irreversibilidade Vol. 35*. EDUSP, 2001.
- [3] A. Frigessi, P. Di Stefano, C. R. Hwang, and S. J. Sheu. Convergence rate of the Gibbs sampler, the Metropolis algorithm and other single-site updating dynamics. *Journal of the Royal Statistical Society*, 55:205–219, 1993.
- [4] Stuart Geman and Donald Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-6(6):721–741, November 1984.
- [5] D.J. Griffiths. *Introduction to quantum mechanics*. Pearson Prentice Hall, 2005.
- [6] W. Heisenberg. Zur Theorie des Ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 49(9):619–636, September 1928.
- [7] Ernst Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):253–258, February 1925.
- [8] D. P. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Statistical Physics - 2nd Edition*. September 2005.
- [9] V. I. Manousiouthakis and M. W. Deem. Strict detailed balance is unnecessary in Monte Carlo simulation. , 110:2753–2756, February 1999.
- [10] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. , 21:1087–1092, June 1953.
- [11] M. E. J. Newman and G. T. Barkema. *Monte Carlo methods in statistical physics*. 1999.
- [12] L. Onsager. Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition. *Physical Review*, 65:117–149, February 1944.
- [13] R. Peierls and M. Born. Statistical theory of adsorption with interaction between the adsorbed atoms. *Proceedings of the Cambridge Philosophical Society*, 32:471, 1936.
- [14] P. H. Peskun. Optimum monte-carlo sampling using markov chains. *Biometrika*, 60(3):pp. 607–612, 1973.

- 
- [15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in FORTRAN. The art of scientific computing*. 1992.
- [16] R. Ren and G. Orkoulas. Acceleration of Markov chain Monte Carlo simulations through sequential updating. , 124(6):064109, February 2006.