

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE FÍSICA  
PÓS-GRADUAÇÃO EM FÍSICA

Tese de Doutorado

Padrões estruturados e campo aleatório  
em redes complexas.

**Felipe França Doria**

Tese apresentada ao Instituto de Física da  
Universidade Federal do Rio Grande do  
Sul, como parte dos créditos para a obten-  
ção do título de Doutor em Ciências.

Orientador: Prof. Dr. Rubem Erichsen Jr.  
Co-Orientador: Prof. Dr. David Dominguez

**Porto Alegre**  
**Maio de 2016**

*Não há coisa tão fácil como dar conselho, nem mais difícil do que sabê-lo dar. - Lope de Vega*

## **Agradecimentos**

Agradeço primeiramente aos meus orientadores, Rubem e David, pela paciência, sabedoria e amizade. Seus ensinamentos são de muita valia para mim.

Também agradeço a minha família, meu pai Luciano, minha mãe Mariluce e meu irmão Fernando. Sem eles, nada faria sentido.

Igualmente, agradeço a todos os meus amigos. Compartilhamos muitas experiências agradáveis juntos. Nossas conversas, risadas e camaradagem ao longo da minha vida foram fundamentais.

E por último, mas não menos importante. Gostaria de agradecer a Glaucia, pela presença, pelo respeito, e pelo amor. Amo você.

## Resumo

Este trabalho foca no estudo de duas redes complexas. A primeira é um modelo de Ising com campo aleatório. Este modelo segue uma distribuição de campo gaussiana e bimodal. Uma técnica de conectividade finita foi utilizada para resolvê-lo. Assim como um método de Monte Carlo foi aplicado para verificar os resultados. Há uma indicação em nossos resultados que para a distribuição gaussiana a transição de fase é sempre de segunda ordem. Para as distribuições bimodais há um ponto tricrítico, dependente do valor da conectividade  $c$ . Abaixo de um certo mínimo de  $c$ , só existe transição de segunda ordem.

A segunda é uma rede neural atratora métrica. Mais precisamente, estudamos a capacidade deste modelo para armazenar os padrões estruturados. Em particular, os padrões escolhidos foram retirados de impressões digitais, que apresentam algumas características locais. Os resultados mostram que quanto menor a atividade de padrões de impressões digitais, maior a relação de carga e a qualidade de recuperação. Uma teoria, também foi desenvolvido como uma função de cinco parâmetros: a relação de carga, a conectividade, o grau de densidade da rede, a relação de aleatoriedade e a correlação do padrão espacial.

**Palavras-chave:** Redes complexas; Sistemas desordenados; Modelo de Ising com campo aleatório; Conectividade finita; Rede neural atratora métrica; Reconhecimento de padrões; Impressões digitais.

## *Abstract*

This work focus on the study of two complex networks. The first one is a random field Ising model. This model follows a gaussian and bimodal distribution, for the random field. A finite connectivity technique was utilized to solve it. As well as a Monte Carlo method was applied to verify our results. There is an indication in our results that for a gaussian distribution the phase transition is always second-order. For the bimodal distribution there is a tricritical point, tha depends on the value of the connectivity  $c$ . Below a certain minimum  $c$ , there is only a second-order transition.

The second one is a metric attractor neural network. More precisely we study the ability of this model to learn structured patterns. In particular, the chosen patterns were taken from fingerprints, which present some local features. Our results show that the higher the load ratio and retrieval quality are the lower is the fingerprint patterns activity. A theoretical framework was also developed as a function of five parameters: the load ratio, the connectivity, the density degree of the network, the randomness ratio and the spatial pattern correlation.

**Keywords:**Complex networks; Disordered systems; Random field Ising model; Finite connectivity; Metric attractor neural network; Pattern recognition; Fingerprints.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>7</b>
<b>2</b>	<b>Modelo de Ising com campo aleatório em um grafo aleatório</b>	<b>10</b>
2.1	Introdução	10
2.2	Modelo	13
2.3	“Ansatz” de simetria de réplicas e a obtenção dos observáveis	19
2.4	Cálculo da energia livre	23
2.5	Resultados teóricos	28
2.6	Simulação Numérica	35
2.6.1	Método de Monte Carlo	35
2.6.2	Resultados numéricos	39
<b>3</b>	<b>Padrões estruturados e rede métrica atratora</b>	<b>42</b>
3.1	Introdução	42
3.2	O modelo	44
3.3	Experimentos de impressões digitais	47
3.3.1	Configurações topológicas e performance de recuperação	48
3.3.2	IDs esqueletizadas	50
3.3.3	IDs com atividade uniforme	51
3.3.4	Capacidade de Informação	52

---

3.3.5	Métrica de Boltzmann . . . . .	53
3.4	Teoria . . . . .	55
3.4.1	Divisão do campo em sinal e ruído . . . . .	56
3.4.2	Distribuição do campo . . . . .	59
3.4.3	Resultados teóricos . . . . .	64
<b>4</b>	<b>Conclusões e perspectivas</b>	<b>70</b>
4.1	Conclusões . . . . .	70
4.2	Perspectivas . . . . .	72
	<b>Referências</b>	<b>73</b>

# Capítulo 1

## Introdução

O estudo de redes complexas permeia diversos campos da ciência, indo da Neurobiologia até a Física Estatística [1, 2, 3, 4]. Um desafio da atualidade é descrever sistemas complexos de forma completa e precisa. Para isso, foram propostos modelos matemáticos que capturam suas características essenciais [5]. Contudo a arquitetura dessas redes pode variar indo de redes regulares a redes completamente aleatórias.

É possível fazer uma analogia ao imaginar  $n \gg 1$  botões espalhados no chão [6]. Se todos os botões fossem conectados por um fio em sequência, ou conectássemos todos os  $n$  botões entre si teríamos exemplos de redes regulares reais. De outro modo, ao pegar dois botões aleatoriamente e os conectar com um fio, ao repetir este procedimento  $c$  vezes (para um  $c$  grande, eventualmente serão selecionados botões já conectados), resultaria em um exemplo físico de um grafo aleatório com  $n$  vértices (sítios) e  $c$  arestas (conexões).

A forma que a topologia dos grafos aleatórios muda em função das conexões  $c$  foi estudada por Erdős e Rényi [7], sendo que para  $c$  pequeno o grafo tende a ser fragmentado e apresenta diversas aglomerações de sítios. A medida que  $c$  cresce, as aglomerações aumentam ligando-se com sítios isolados e outras aglomerações.

Assumindo que redes reais em geral seriam uma mistura entre esses dois extremos de



---

ordem de aleatoriedade, Watts e Strogatz [8, 9] propuseram um modelo simples chamado de “rede de pequenos mundos” que pode ser considerado um meio termo entre uma rede regular e um grafo aleatório. O modelo trata de uma rede regular cujas ligações locais possuem uma probabilidade  $\omega$  ( $0 \leq \omega \leq 1$ ) de serem trocadas por uma conexão aleatória. Esse parâmetro de aleatoriedade de uma rede de pequenos mundos proporciona duas propriedades: caminhos curtos entre sítios, que aumenta a velocidade de comunicação entre esses sítios e uma alta aglomeração (maior do que a de um grafo aleatório).

O Magnetismo utiliza o formalismo de redes complexas, para modelar sistemas magnéticos, como ferromagnetos [10], vidros de spin [11], entre outros [12, 13, 14]. Um dos modelos mais conhecidos e empregados para tal é o Modelo de Ising [15, 16]. Em geral o modelo é resolvido por meio da teoria de campo médio, que considera uma arquitetura de um grafo totalmente conectado, neste caso o  $c \rightarrow \infty$ . Porém, sistemas físicos reais possuem um  $c$  finito e a teoria de campo médio tem muitas dificuldades na resolução desses modelos. Para solucioná-los é aplicada a teoria das réplicas. Adicionando-se um campo local aleatório temos o modelo de Ising com campo aleatório, que é amplamente estudado no contexto da matéria condensada [12]. No capítulo 2 trataremos do modelo de Ising com conectividade finita com duas distribuições de campo aleatório em um grafo aleatório.

De forma semelhante, o modelo de Hopfield [17] estudado pela Neurociência computacional é isomórfico em relação ao modelo de Ising a uma temperatura  $T = 0$  [18]. Este modelo estuda neurônios que podem disparar ou não como os spins- $\frac{1}{2}$  encontrados no modelo de Ising. A função de troca entre os neurônios  $J_{ij}$  é escrita com pesos hebbianos, os neurônios que disparam juntos fortalecem suas conexões e neurônios que disparam fora de sincronia enfraquecem suas conexões [19]. Os pesos hebbianos geram atratores no espaço configuracional da energia e as entradas da rede convergem em direção a um padrão. A sobreposição desses padrões aprendidos e dos estados neuronais é equivalente à magnetização no modelo de Ising [20].

Recentemente arquiteturas de redes neurais atratoras mais estruturadas denominadas redes métricas foram estudadas, especialmente topologias de pequeno mundo [21, 22, 23]. Tais grafos são modelados por dois parâmetros, a conectividade e a aleatoriedade, e eles podem capturar a maior parte das características de uma grande gama de redes [8, 24]. A habilidade de recuperação dessas redes são comumente medidas pela sobreposição entre estados de neurônios e padrões armazenados, e o parâmetro de carga expresso como a razão entre os padrões armazenados e as ligações por nodo [25]. Padrões estruturados localmente seriam a evolução natural para uma rede neural atratora com topologia estruturada. No capítulo 3 estudamos Redes neurais atratoras métricas com padrões estruturados.

O objetivo desta tese é estudar duas arquiteturas de sistemas complexos. A primeira trata de um grafo aleatório. Neste caso vamos investigar um modelo de Ising com conectividade finita e duas distribuições de campo aleatório (bimodal e gaussiana), obter as curvas de transição de fase e o impacto que a conectividade finita tem nesta rede. A segunda trata de uma rede de pequeno mundo, na qual estudaremos uma rede neural atratora cujos padrões apresentam características globais e locais, onde fizemos experimentos e verificamos a habilidade de aprendizado da rede e adicionamos um parâmetro de aleatoriedade dos padrões. No capítulo 4 são expostas as conclusões e possíveis extensões do trabalho.

## Capítulo 2

# Modelo de Ising com campo aleatório em um grafo aleatório

### 2.1 Introdução

O modelo de Ising com campo aleatório (RFIM na sigla em inglês) atraiu interesse de pesquisadores desde sua introdução por Larkin [12] mais de quatro décadas atrás. Apesar de sua aparente simplicidade [13] este modelo apresenta uma riqueza em termos de comportamentos estático e dinâmico. Do ponto de vista experimental, o interesse surge do fato de que um fenômeno como um campo aleatório pode ser encontrado em uma variedade de cenários distintos, tais quais vórtices e supercondutores [14], sistemas desordenados de cério [26] e sistemas ferroelétricos [27].

Certamente, muito do interesse em RFIM de forma teórica e de forma experimental pode ser traçado para a proposição de Fishman e Aharony de que este modelo poderia ser gerado em sistemas físicos reais antiferromagnetos diluídos de Ising ao aplicarmos um campo magnético [28]. Na verdade, a realização dessa proposição pode ser observada em sistemas reais, tais quais  $Co_xZn_{1-x}F_2$  [29],  $Fe_xZn_{1-x}F_2$  [30] e  $Fe_xMg_{1-x}Cl_2$  [31],

o que permitiu a confirmação das previsões teóricas, como por exemplo, de que a fase ferromagnética (F) é suprimida pelo campo aleatório (RF) para uma dimensão  $d_c < 2$  [32]. Entretanto, ainda existem dúvidas em relação à ordem da transição de fase a baixas temperaturas e sobre a existência de um ponto tricrítico [33].

Resultados prévios relativos à ordem da transição de fase são bastante controversos. Para muitos pesquisadores, a teoria de campo médio convencional para RFIM indica que a ordem da transição de fase é bastante dependente do tipo de função de distribuição escolhida para o campo aleatório [34]. Por exemplo, utilizando uma distribuição gaussiana e o procedimento de réplica para tratar a desordem, Schneider e Pytte [35] perceberam que a transição entre a fase paramagnética (P) para uma fase ferromagnética (F) é sempre de segunda ordem. Se uma distribuição bimodal for utilizada para o campo aleatório, em baixas temperaturas, a transição de fase torna-se de primeira ordem com o aparecimento de um ponto tricrítico [34]. Na verdade, foi mostrado que, para a aproximação de campo médio convencional e campo zero, a transição de primeira ordem ocorreria quando a função de distribuição do campo aleatório fosse simétrica, com um mínimo, e a transição de segunda ordem ocorreria quando a função de distribuição fosse simétrica, com um máximo [34]. Trabalhos subsequentes indicaram que condições adicionais da distribuição de campo aleatório são necessárias para se obter uma transição de primeira ou segunda ordem [36, 37]. Por outro lado, um trabalho recente de Fytas e Martin-Mayor [38] nos mostra que uma transição ferro-para em um modelo com três dimensões do RFIM com interações de primeiros vizinhos é contínua para diferentes tipos de distribuição de campo aleatório.

Ainda assim, alguns dos resultados descritos acima [34, 36, 37, 38] foram obtidos utilizando uma aproximação de campo médio convencional, ou seja, para interações de longo alcance e conectividade infinita entre os spins. Poderíamos nos perguntar o que ocorreria em uma situação mais realista, isto é, para spins conectados com um número finito de outros spins. Em outras palavras, qual é o papel da conectividade em relação

à ordem da transição de fase? Resultados recentes em uma rede de Bethe com conectividade fixa e interações de curto alcance [39] mostraram que transições ferro-para são sempre contínuas para  $c \leq 3,0$  enquanto que para  $c \geq 4,0$  o ponto tricrítico aparece para baixas temperaturas, abaixo do qual a transição se torna descontínua. Alternativamente, podemos checar a robustez do resultado acima ao tratar de uma rede aleatória com uma conectividade finita (mas não fixa) e interações de longo alcance. Portanto, podemos esperar que alguns aspectos do sistema modelado pelo RFIM podem ser melhor descritos pela teoria de conectividade finita [40, 41], já que ela se aproxima mais da realidade.

Investigamos o RFIM usando a abordagem de conectividade finita [42, 43, 44, 45]. Nesse caso, uma rede aleatória é construída de tal forma que a conectividade entre dois sítios,  $i$  e  $j$ , segue uma distribuição bimodal com probabilidades  $\frac{c}{N}$  de estar presente e  $1 - \frac{c}{N}$  de estar ausente. Assim, podemos controlar a média da conectividade da rede ao ajustar o parâmetro real  $c$ . Para conexões existentes entre sítios,  $i$  e  $j$ , dois spins interagem por meio de uma ligação uniforme  $J_{ij}$ . Nessa rede controlada pela conectividade, um campo aleatório é aplicado e acoplado localmente aos spin. Recentemente, o método da cavidade foi utilizado para investigar a fase de vidro de spin em um RFIM [46]. Seguindo uma abordagem similar e usando o método das réplicas [11], obtivemos uma equação de ponto de sela para a distribuição do campo local efetivo. Isso permite calcular o parâmetro de ordem e a energia-livre. O foco do trabalho é verificar a ordem da transição de fase e a existência de um ponto tricrítico como função da conectividade. Para esse propósito, usamos distribuições de probabilidade gaussiana e bimodal para o campo aleatório.

Deve ser salientado que resultados experimentais para  $Fe_xMg_{1-x}Cl_2$  [31] apresentam um comportamento tipicamente de vidro de spin se  $x < 0,55$ . Entretanto, para valores maiores de  $x$ , este sistema se comportou como um RFIM com pontos tricrítico no diagrama de fase. Surpreendentemente, para  $x > 0,6$ , o ponto tricrítico desaparece e a transição de fase ferro-para torna-se de segunda ordem. Trabalhos teóricos mostra-

ram que o ponto tricrítico pode ser suprimido ao variarmos alguns parâmetros de uma distribuição de campo aleatório bimodal [47, 48]. Contudo, esses estudos ainda utilizam aproximação de campo médio convencional.

## 2.2 Modelo

O Hamiltoniano do modelo de Ising com campo aleatório pode ser escrito desta forma

$$H = - \sum_{i < j} J_{ij} \sigma_i \sigma_j - \sum_i \theta_i \sigma_i, \quad (2.1)$$

sendo  $J_{ij}$  o termo de troca,  $\theta_i$  o campo aleatório que atua na rede e  $\sigma_i$  os estados da rede.

Ao escrevermos o termo de troca como função de uma constante e uma quantidade referente à conectividade de um sítio  $J_{ij} = c_{ij}J$ , podemos reescrever o Hamiltoniano de Ising em um campo aleatório,

$$H = -J \sum_{i < j} c_{ij} \sigma_i \sigma_j - \sum_i \theta_i \sigma_i, \quad (2.2)$$

onde  $\sigma_i = \pm 1$  e  $c_{ij}$  é a conectividade que segue a distribuição de probabilidade

$$p_{c_{ij}} = \frac{c}{N} \delta_{c_{ij},1} + \left(1 - \frac{c}{N}\right) \delta_{c_{ij},0}. \quad (2.3)$$

O campo aleatório  $\theta_i$  pode seguir uma distribuição bimodal,

$$p_{\theta_i} = (1 - p) \delta(\theta_i - \theta_0) + p \delta(\theta_i + \theta_0), \quad (2.4)$$

ou uma distribuição gaussiana,

$$p_{\theta_i} = \frac{1}{\sqrt{2\pi\Delta^2}} \exp \left[ -\frac{(\theta_i - \theta_0)^2}{2\Delta^2} \right]. \quad (2.5)$$

As propriedades do sistema, quando em equilíbrio com um banho térmico de temperatura  $T = \beta^{-1}$  são derivadas da energia livre,

$$f(\beta) = - \lim_{N \rightarrow \infty} \frac{1}{\beta N} \langle \log Z \rangle_{\{\theta_i, c_{ij}\}}, \quad (2.6)$$

onde  $\langle \dots \rangle$  correspondem à média sobre a desordem,  $Z = \sum_{\sigma} e^{-\beta H}$  é a função de partição e o vetor de  $N$  coordenadas  $\sigma$  representa o estado do sistema. Para obter a média congelada (quenched), temos que  $\langle \log Z \rangle \neq \log \langle Z \rangle$ .

Tomando em conta que o cálculo da média de  $\log Z$  pode levar a dificuldades matemáticas, para contornar esse problema é utilizada a identidade de réplicas,

$$\langle \log(Z) \rangle = \lim_{n \rightarrow 0} \frac{1}{n} \log \langle Z^n \rangle. \quad (2.7)$$

Podemos escrever a função de partição,

$$Z = \sum_{\sigma} \exp \left\{ \frac{\beta}{c} \sum_{i < j} c_{ij} J \sigma_i \sigma_j + \beta \sum_i \theta_i \sigma_i \right\}. \quad (2.8)$$

Replicando a equação, obtemos

$$Z^n = \sum_{\sigma^1 \dots \sigma^n} e^{\beta \sum_{i,\alpha} \theta_i \sigma_i^\alpha} e^{\frac{\beta}{c} \sum_{i < j} c_{ij} J \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha}. \quad (2.9)$$

Aqui,  $\sigma^\alpha$  é um vetor  $N$  dimensional referente ao estado de  $N$  sítios na réplica  $\alpha$ .

Fazemos então a média sobre as variáveis aleatórias  $c_{ij}$ ,  $\theta_i$  e ficamos com

$$\langle Z^n \rangle = \sum_{\sigma^1 \dots \sigma^n} \left\langle e^{\beta \sum_{i,\alpha} \theta_i \sigma_i^\alpha} \right\rangle_{\{\theta_i\}} \prod_{i < j} \left\langle e^{\frac{\beta}{c} c_{ij} J \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha} \right\rangle_{c_{ij}}. \quad (2.10)$$

Ao realizarmos a média sobre a variável  $c_{ij}$ , temos

$$\langle Z^n \rangle = \sum_{\sigma^1 \dots \sigma^n} \left\langle e^{\beta \sum_{i,\alpha} \theta_i \sigma_i^\alpha} \right\rangle_{\{\theta_i\}} \prod_{i < j} \left[ 1 + \frac{c}{N} \left( e^{\frac{\beta}{c} J \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha} - 1 \right) \right]. \quad (2.11)$$

Tomando o limite para  $\frac{c}{N} \rightarrow 0$ , usamos a aproximação  $e^x \simeq 1 + x + \dots$  e a identidade  $\sum_{i < j} = \frac{1}{2} \sum_{i \neq j}$  para escrever a equação (2.11) desta maneira,

$$\langle Z^n \rangle = \sum_{\sigma^1 \dots \sigma^n} \left\langle \exp \left[ \beta \sum_{i, \alpha} \theta_i \sigma_i^\alpha + \frac{c}{2N} \sum_{i \neq j} \left( e^{\frac{\beta}{c} J \sigma_i \cdot \sigma_j} - 1 \right) \right] \right\rangle_{\{\theta_i\}}, \quad (2.12)$$

é importante salientar que utilizamos  $\sum_\alpha \sigma_i^\alpha \sigma_j^\alpha = \sigma_i \cdot \sigma_j$ , sendo  $\sigma_i$  um vetor  $n$  dimensional referente ao estado de  $n$  réplicas no sítio  $i$ .

Como a conectividade  $c$  é finita, não é possível expandir a exponencial interna. Para extrair as variáveis do somatório, introduzimos a identidade

$$1 = \sum_{\sigma} \delta_{\sigma \sigma_i} \equiv \sum_{\sigma} \prod_{\alpha=1}^n \delta_{\sigma^\alpha \sigma_i^\alpha}, \quad (2.13)$$

onde  $\delta_{\sigma \sigma_i} = 1$  para  $\sigma = \sigma_i$  e  $\delta_{\sigma \sigma_i} = 0$  para  $\sigma \neq \sigma_i$ .

Então, a Eq.(2.12) torna-se

$$\langle Z^n \rangle = \sum_{\sigma^1 \dots \sigma^n} \left\langle \exp \left[ \beta \sum_{i, \alpha} \theta_i \sigma_i^\alpha + \frac{c}{2N} \sum_{i \neq j} \sum_{\sigma, \tau} \delta_{\sigma, \sigma_i} \delta_{\tau, \sigma_j} \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right] \right\rangle_{\{\theta_i\}}. \quad (2.14)$$

Como o  $c$  é finito, não é possível truncar a expansão da exponencial interna o que conduz a uma quantidade infinita de parâmetros de ordem. Para contornar esse problema introduzimos uma função de ordem que apresenta a fração  $P(\sigma)$  de sítios onde a configuração  $\sigma$  acontece,

$$\int \prod_{\sigma} \left\{ dP(\sigma) \delta \left( P(\sigma) - \frac{1}{N} \sum_i \delta_{\sigma, \sigma_i} \right) \right\} = 1. \quad (2.15)$$

Introduzindo um parâmetro adicional  $\hat{P}(\sigma)$  a representação integral para a função  $\delta$ ,



$$\int \prod_{\sigma} [dP(\sigma) d\hat{P}(\sigma)] \exp \left\{ \sum_{\sigma} \hat{P}(\sigma) \left[ P(\sigma) - \frac{1}{N} \sum_i \delta_{\sigma, \sigma_i} \right] \right\} = 1. \quad (2.16)$$

Utilizando as Eqs.(2.15 e2.16) na Eq.(2.14), temos

$$\begin{aligned} \langle Z^n \rangle &= \sum_{\sigma_1 \dots \sigma_n} \left\langle \int \prod_{\sigma} [dP(\sigma) d\hat{P}(\sigma)] \exp \left[ \beta \sum_{i, \alpha} \theta_i \sigma_i^{\alpha} + \sum_{\sigma} \hat{P}(\sigma) P(\sigma) \right. \right. \\ &\quad \left. \left. - \frac{1}{N} \sum_{\sigma} \hat{P}(\sigma) \sum_i \delta_{\sigma, \sigma_i} + \frac{Nc}{2} \sum_{\sigma, \tau} P(\sigma) P(\tau) \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right] \right\rangle_{\theta_i}. \end{aligned} \quad (2.17)$$

O traço sobre as variáveis de spins replicadas e a média sobre as variáveis de campo aleatório  $\theta_i$  atuam apenas no primeiro e terceiro termo da exponencial externa, então,

$$\begin{aligned} \langle Z^n \rangle &= \int \prod_{\sigma} [dP(\sigma) d\hat{P}(\sigma)] \left\langle \sum_{\sigma_1 \dots \sigma_n} \exp \left[ \beta \sum_{i, \alpha} \theta_i \sigma_i^{\alpha} - \frac{1}{N} \sum_{\sigma} \hat{P}(\sigma) \sum_i \delta_{\sigma, \sigma_i} \right] \right\rangle_{\theta} \\ &\quad \times \exp \left\{ \sum_{\sigma} \hat{P}(\sigma) P(\sigma) + \frac{Nc}{2} \sum_{\sigma, \tau} P(\sigma) P(\tau) \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right\}. \end{aligned} \quad (2.18)$$

Passamos o somatório sobre os sítios  $i$  para fora da exponencial, obtemos um produto de  $N$  exponenciais,

$$\begin{aligned} \langle Z^n \rangle &= \int \prod_{\sigma} [dP(\sigma) d\hat{P}(\sigma)] \left( \sum_{\sigma} \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \exp \left[ - \frac{1}{N} \hat{P}(\sigma) \right] \right)^N \\ &\quad \times \exp \left\{ \sum_{\sigma} \hat{P}(\sigma) P(\sigma) + \frac{Nc}{2} \sum_{\sigma, \tau} P(\sigma) P(\tau) \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right\}, \end{aligned} \quad (2.19)$$

aplicando  $x = \exp[\log(x)]$  obtemos

$$\begin{aligned} \langle Z^n \rangle &= \int \prod_{\sigma} [dP(\sigma) d\hat{P}(\sigma)] \exp \left\{ N \log \sum_{\sigma} \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} - \frac{1}{N} \hat{P}(\sigma) \right] \right\rangle_{\theta} \right. \\ &\quad \left. + \sum_{\sigma} \hat{P}(\sigma) P(\sigma) + \frac{Nc}{2} \sum_{\sigma, \tau} P(\sigma) P(\tau) \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right\}. \end{aligned} \quad (2.20)$$

Fazemos então a transformação  $\hat{P}(\sigma) \rightarrow N\hat{P}(\sigma)$ , chegamos em

$$\begin{aligned} \langle Z^n \rangle &= \int \prod_{\sigma} [dP(\sigma) d\hat{P}(\sigma)] \exp \left\{ N \log \sum_{\sigma} \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} - \hat{P}(\sigma) \right] \right\rangle_{\theta} \right. \\ &\quad \left. + N \sum_{\sigma} \hat{P}(\sigma) P(\sigma) + \frac{Nc}{2} \sum_{\sigma, \tau} P(\sigma) P(\tau) \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right\}. \end{aligned} \quad (2.21)$$

Para  $N \rightarrow \infty$  podemos integrar esta equação por meio de ponto de sela e chegamos na energia livre ao calcular o extremo sobre  $P(\sigma)$  e  $\hat{P}(\sigma)$ ,

$$\begin{aligned} f(\beta) &= - \lim_{n \rightarrow 0} \frac{1}{n\beta} \text{Extr}_{P(\sigma), \hat{P}(\sigma)} \left\{ \log \sum_{\sigma} \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} - \hat{P}(\sigma) \right] \right\rangle_{\theta} + \sum_{\sigma} \hat{P}(\sigma) P(\sigma) \right. \\ &\quad \left. + \frac{c}{2} \sum_{\sigma, \tau} P(\sigma) P(\tau) \left( e^{\frac{\beta}{c} J \sigma \cdot \tau} - 1 \right) \right\}. \end{aligned} \quad (2.22)$$

As equações de ponto de sela podem ser obtidas ao derivar a equação Eq.(2.22) em relação às densidades  $P(\sigma)$  e  $\hat{P}(\sigma)$  e igualar a zero,

$$\frac{\partial f}{\partial \hat{P}(\sigma)} = \frac{\partial f}{\partial P(\sigma)} = 0; \quad (2.23)$$

o que nos leva a

$$P(\boldsymbol{\sigma}) - \frac{\left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} - \hat{P}(\boldsymbol{\sigma}) \right] \right\rangle_{\theta}}{\sum_{\boldsymbol{\sigma}} \left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} - \hat{P}(\boldsymbol{\sigma}) \right] \right\rangle_{\theta}} = 0, \quad (2.24)$$

e também a

$$\hat{P}(\boldsymbol{\sigma}) + c \sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \left( e^{\frac{\beta}{c} J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right) = 0. \quad (2.25)$$

Substituindo Eq.(2.24) na Eq.(2.25) podemos eliminar  $\hat{P}(\boldsymbol{\sigma})$ ,

$$P(\boldsymbol{\sigma}) = \frac{\left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} + c \sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \left( e^{\frac{\beta}{c} J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right) \right] \right\rangle_{\theta}}{\sum_{\boldsymbol{\tau}} \left\langle \exp \left[ \beta\theta \sum_{\alpha} \tau^{\alpha} + \sum_{\boldsymbol{\rho}} P(\boldsymbol{\rho}) \left( e^{\frac{\beta}{c} J \boldsymbol{\tau} \cdot \boldsymbol{\rho}} - 1 \right) \right] \right\rangle_{\theta}}. \quad (2.26)$$

Negligenciando o fator de normalização pois este é o mesmo para todos os  $P(\boldsymbol{\sigma})$ , temos

$$P(\boldsymbol{\sigma}) = \left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} + c \sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \left( e^{\frac{\beta}{c} J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right) \right] \right\rangle_{\theta}, \quad (2.27)$$

ao calcular o extremo sobre o  $P(\boldsymbol{\sigma})$ , a energia-livre é reescrita assim,

$$\begin{aligned} f(\beta) &= - \lim_{n \rightarrow 0} \frac{1}{\beta n} \text{Extr}_{P(\boldsymbol{\sigma})} \left\{ - \frac{c}{2} \sum_{\boldsymbol{\sigma} \boldsymbol{\tau}} P(\boldsymbol{\sigma}) P(\boldsymbol{\tau}) \left( e^{\beta J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right) \right. \\ &\quad \left. + \log \sum_{\boldsymbol{\sigma}} \left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} + c \sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \left( e^{\beta J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right) \right] \right\rangle_{\theta} \right\}. \quad (2.28) \end{aligned}$$

## 2.3 “Ansatz” de simetria de réplicas e a obtenção dos observáveis

Podemos assumir que a densidade  $P(\boldsymbol{\sigma})$  adquire uma certa forma. Uma forma na qual as permutações dos spins entre diferentes réplicas deixam  $P(\boldsymbol{\sigma})$  invariante. Estas condições são satisfeitas pelo “ansatz” de simetria de réplicas,

$$P(\boldsymbol{\sigma}) = \int dh W(h) \frac{e^{\beta h \sum_{\alpha} \sigma^{\alpha}}}{[2 \cosh(\beta h)]^n}. \quad (2.29)$$

Antes de introduzirmos o “ansatz” da Eq.(2.29) na equação de ponto de sela, Eq.(2.27), vamos expandir a exponencial da equação Eq.(2.27),

$$P(\boldsymbol{\sigma}) = \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \sum_{\boldsymbol{\tau}^1 \dots \boldsymbol{\tau}^k} P(\boldsymbol{\tau}^1) \dots P(\boldsymbol{\tau}^k) \times \prod_{\ell=1}^k e^{\frac{\beta J}{c} \boldsymbol{\sigma} \cdot \boldsymbol{\tau}^{\ell}}, \quad (2.30)$$

aplicando o “ansatz” de simetria de réplicas, Eq.(2.29), na equação expandida, Eq.(2.30), temos

$$P(\boldsymbol{\sigma}) = \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \sum_{\boldsymbol{\tau}^1 \dots \boldsymbol{\tau}^k} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\ \times \prod_{\ell=1}^k \left( \exp \left[ \frac{\beta J}{c} \boldsymbol{\sigma} \cdot \boldsymbol{\tau}^{\ell} + \beta h_{\ell} \sum_{\alpha} \tau_{\ell}^{\alpha} \right] \right), \quad (2.31)$$

então, ao calcular o traço  $\sum_{\boldsymbol{\tau}^1 \dots \boldsymbol{\tau}^n}$  e replicar a expressão, ficamos com

$$\begin{aligned}
P(\boldsymbol{\sigma}) = & \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\
& \times \prod_{\ell=1}^k \left[ \prod_{\alpha=1}^n \sum_{\tau_{\ell}^{\alpha}} \exp \left( \frac{\beta J}{c} \sigma^{\alpha} \tau_{\ell}^{\alpha} + \beta h_{\ell} \sum_{\alpha} \tau_{\ell}^{\alpha} \right) \right]. \tag{2.32}
\end{aligned}$$

Utilizando  $x^n = e^{n \log(x)} \approx 1 + n \log(x)$  escrevemos

$$\begin{aligned}
P(\boldsymbol{\sigma}) = & \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\
& \times \prod_{\ell=1}^k \left\{ \exp \left[ \sum_{\alpha} \log \left( \sum_{\tau_{\ell}^{\alpha}} \exp \left[ \frac{\beta J}{c} \sigma^{\alpha} \tau_{\ell}^{\alpha} + \beta h_{\ell} \sum_{\alpha} \tau_{\ell}^{\alpha} \right] \right) \right] \right\}, \tag{2.33}
\end{aligned}$$

como possuímos apenas dois estados de spins,  $\tau_{\ell}^{\alpha} = \pm 1$ , expandimos o somatório  $\sum_{\tau_{\ell}^{\alpha}}$  e obtemos

$$\begin{aligned}
P(\boldsymbol{\sigma}) = & \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\
& \times \prod_{\ell=1}^k \left\{ \exp \left[ \sum_{\alpha} \log \left( e^{\frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell}} + e^{-\frac{\beta J}{c} \sigma^{\alpha} - \beta h_{\ell}} \right) \right] \right\}, \tag{2.34}
\end{aligned}$$

por meio da função hiperbólica  $e^x + e^{-x} = 2 \cosh(x)$ , chegamos na equação,

$$\begin{aligned}
P(\boldsymbol{\sigma}) = & \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\
& \times \prod_{\ell=1}^k \left\{ \exp \left[ \sum_{\alpha} \log \left( 2 \cosh \left( \frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell} \right) \right) \right] \right\}, \tag{2.35}
\end{aligned}$$

utilizando a Eq.(2.13) e a identidade

$$\delta_{\sigma, \sigma^\alpha} = \frac{1}{2}(1 + \sigma\sigma^\alpha), \quad (2.36)$$

ficamos com

$$P(\boldsymbol{\sigma}) = \left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\ \times \prod_{\ell=1}^k \left\{ \exp \left[ \sum_{\alpha} \sum_{\sigma} \left( \frac{1}{2}(1 + \sigma\sigma^{\alpha}) \right) \log \left( 2 \cosh \left( \frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell} \right) \right) \right] \right\}. \quad (2.37)$$

Realizando o traço em  $\sigma = \pm 1$ ,

$$P(\boldsymbol{\sigma}) = \left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\ \times \prod_{\ell=1}^k \left\{ \exp \left[ \left( \frac{n}{2} + \frac{1}{2} \sum_{\alpha} \sigma^{\alpha} \right) \log \left[ 2 \cosh \left( \frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell} \right) \right] \right] \right. \\ \left. + \left( \frac{n}{2} - \frac{1}{2} \sum_{\alpha} \sigma^{\alpha} \right) \log \left( 2 \cosh \left( -\frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell} \right) \right) \right\}. \quad (2.38)$$

No limite para  $n \rightarrow 0$ , temos que

$$P(\boldsymbol{\sigma}) = \left\langle \exp \left[ \beta\theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\ \times \prod_{\ell=1}^k \exp \left\{ \left( \frac{1}{2} \sum_{\alpha} \sigma^{\alpha} \right) \log \left[ \frac{[\cosh(\frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell})]}{[\cosh(-\frac{\beta J}{c} \sigma^{\alpha} + \beta h_{\ell})]} \right] \right\}. \quad (2.39)$$

Utilizando relações conhecidas entre funções hiperbólicas, chegamos em

$$\begin{aligned}
P(\boldsymbol{\sigma}) &= \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} \frac{e^{-c} c^k}{k!} \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\
&\times \exp \left\{ \left( \sum_{\alpha} \sigma^{\alpha} \right) \sum_{\ell=1}^k \operatorname{arctanh} \left[ \tanh(\beta h_{\ell}) \tanh \left( \frac{\beta J_{\ell}}{c} \right) \right] \right\}. \quad (2.40)
\end{aligned}$$

Agora igualando o “ansatz”, Eq.(2.29) na Eq.(2.40),

$$\begin{aligned}
\int dh W(h) \frac{e^{\beta h \sum_{\alpha} \sigma^{\alpha}}}{[2 \cosh(\beta h)]^n} &= \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \right\rangle_{\theta} \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k \left\{ \frac{dh_{\ell} W(h_{\ell})}{[2 \cosh(\beta h_{\ell})]^n} \right\} \\
&\times \exp \left\{ \left( \sum_{\alpha} \sigma^{\alpha} \right) \sum_{\ell=1}^k \operatorname{arctanh} \left[ \tanh(\beta h_{\ell}) \tanh \left( \frac{\beta J}{c} \right) \right] \right\}, \quad (2.41)
\end{aligned}$$

onde  $P_k = \frac{e^{-c} c^k}{k!}$  que representa os pesos poissonianos com média  $c$ .

Agora, tomando o limite com  $n \rightarrow 0$ , ficamos com,

$$\begin{aligned}
\int dh W(h) e^{\beta h \sum_{\alpha} \sigma^{\alpha}} &= \int dx \left\langle \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) e^{\beta x \sum_{\alpha} \sigma^{\alpha}} \right. \\
&\times \left. \delta \left( x - \theta - \frac{1}{\beta} \sum_{\ell=1}^k \operatorname{arctanh} \left[ \tanh(\beta h_{\ell}) \tanh \left( \frac{\beta J}{c} \right) \right] \right) \right\rangle_{\theta}. \quad (2.42)
\end{aligned}$$

A partir deste ponto, comparando ambos os lados da Eq.2.42 e obtemos uma expressão para a distribuição de campos  $W(h)$  caso  $x = h$ , escrevemos

$$\begin{aligned}
W(h) &= \left\langle \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) \delta \left( h - \theta \right. \right. \\
&\left. \left. - \frac{1}{\beta} \sum_{\ell=1}^k \operatorname{arctanh} \left[ \tanh(\beta h_{\ell}) \tanh \left( \frac{\beta J}{c} \right) \right] \right) \right\rangle_{\theta}. \quad (2.43)
\end{aligned}$$

A Eq.(2.43) pode ser avaliada por meio de um algoritmo de “dinâmica de população” que será explicitado na Seção 2.6.

Depois de obter a distribuição de campos  $W(h)$  podemos por meio dela derivar a magnetização  $m$ ,

$$m = \sum_{\sigma^1 \dots \sigma^n} P(\boldsymbol{\sigma}) \sigma^1 = \int dh W(h) \tanh(\beta h). \quad (2.44)$$

Com a magnetização e a distribuição de campos locais, podemos tirar várias conclusões, mas para termos uma análise mais completa é necessária uma expressão mais detalhada para a energia livre.

## 2.4 Cálculo da energia livre

Retornemos agora à equação Eq.(2.28), podemos dividir a energia livre em dois termos,

$$f = f_1 + f_2, \quad (2.45)$$

sendo  $f_1$  o termo de energia e  $f_2$  o termo de entropia,

$$f_1 = \frac{c}{2\beta n} \sum_{\boldsymbol{\sigma} \boldsymbol{\tau}} P(\boldsymbol{\sigma}) P(\boldsymbol{\tau}) \left( e^{\beta J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right), \quad (2.46)$$

$$f_2 = -\frac{1}{\beta n} \log \sum_{\boldsymbol{\sigma}} \left\langle \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} + c \sum_{\boldsymbol{\tau}} P(\boldsymbol{\tau}) \left( e^{\beta J \boldsymbol{\sigma} \cdot \boldsymbol{\tau}} - 1 \right) \right] \right\rangle_{\theta}. \quad (2.47)$$

Para simplificar, vamos tratar primeiramente do termo  $f_1$ . Podemos escrever  $f_1$  da seguinte maneira



$$f_1 = \frac{c}{2\beta n} \left( f_{11} - 1 \right), \quad (2.48)$$

sendo que

$$f_{11} = \sum_{\sigma\tau} P(\sigma)P(\tau) \left( e^{\beta J \sigma \cdot \tau} \right). \quad (2.49)$$

Utilizando o “ansatz” Eq.(2.29) na Eq.(2.49), obtemos

$$f_{11} = \sum_{\sigma\tau} \int dh dh' W(h)W(h') \frac{e^{\beta h \sum_{\alpha} \sigma^{\alpha} + \beta h' \sum_{\alpha} \tau^{\alpha} + \beta J \sum_{\alpha} \sigma^{\alpha} \cdot \tau^{\alpha}}}{\left[ 4 \cosh(\beta h) \cosh(\beta h') \right]^n}. \quad (2.50)$$

Pelo fato das réplicas serem não-interagentes, o numerador da equação pode ser escrito como o produto de  $n$  exponenciais,

$$f_{11} = \int dh dh' W(h)W(h') \frac{\left[ \sum_{\sigma\tau} e^{\beta h \sigma + \beta h' \tau + \beta J \sigma \tau} \right]^n}{\left[ 4 \cosh(\beta h) \cosh(\beta h') \right]^n}. \quad (2.51)$$

A partir daí, é possível expandir os somatórios em  $\sigma$  e em  $\tau$ , sendo que os sítios possuem apenas dois estados,  $\sigma = \pm 1$  e  $\tau = \pm 1$ ,

$$f_{11} = \int dh dh' W(h)W(h') \left[ \frac{e^{\beta(h+h'+J)} + e^{\beta(h-h'-J)} + e^{\beta(-h+h'-J)} + e^{\beta(-h-h'+J)}}{4 \cosh(\beta h) \cosh(\beta h')} \right]^n. \quad (2.52)$$

Podemos reescrever a equação da seguinte forma

$$f_{11} = \int dh dh' W(h) W(h') \left[ \frac{2e^{\beta J} \cosh(\beta h + \beta h') + 2e^{-\beta J} \cosh(\beta h - \beta h')}{4 \cosh(\beta h) \cosh(\beta h')} \right]^n, \quad (2.53)$$

utilizando  $x^n = e^{n \log x} \approx 1 + n \log(x)$ , obtemos para  $n \rightarrow 0$  a expressão

$$f_{11} = 1 + n \int dh dh' W(h) W(h') \log \left( \frac{2e^{\beta J} \cosh(\beta h + \beta h') + 2e^{-\beta J} \cosh(\beta h - \beta h')}{4 \cosh(\beta h) \cosh(\beta h')} \right). \quad (2.54)$$

Inserindo a Eq.(2.54) na Eq.(2.48),

$$f_1 = \frac{c}{2\beta} \left[ \int dh dh' W(h) W(h') \log \left( \frac{2e^{\beta J} \cosh(\beta h + \beta h') + 2e^{-\beta J} \cosh(\beta h - \beta h')}{4 \cosh(\beta h) \cosh(\beta h')} \right) \right], \quad (2.55)$$

reescrevendo a equação por meio de relações conhecidas entre funções hiperbólicas chegamos na equação

$$\begin{aligned} f_1 &= \frac{c}{2\beta} \left[ \log \left( \cosh(\beta J) \right) \right] + \frac{c}{2\beta} \left[ \int dh dh' W(h) W(h') \right. \\ &\quad \times \left. \log \left( 1 + \tanh(\beta h) \tanh(\beta h') \tanh(\beta J) \right) \right]. \end{aligned} \quad (2.56)$$

Vamos escrever novamente o termo  $f_2$ ,

$$f_2 = -\frac{1}{\beta n} \log \left\langle \sum_{\sigma} \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} + c \sum_{\tau} P(\tau) \left( e^{\beta J \sigma \cdot \tau} - 1 \right) \right] \right\rangle_{\theta}, \quad (2.57)$$

expandindo a segunda exponencial e introduzindo o “ansatz” Eq.(2.29),

$$f_2 = -\frac{1}{\beta n} \log \left\langle \sum_{\sigma} \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k \frac{dh_{\ell} W(h_{\ell})}{\left[ 2 \cosh(\beta h_{\ell}) \right]^n} \right. \quad (2.58)$$

$$\left. \times \sum_{\tau^1 \dots \tau^k} \prod_{\ell=1}^k \left( e^{\beta h_{\ell} \sum_{\alpha} \tau_{\ell}^{\alpha} + \frac{\beta J}{c} \sum_{\alpha} \tau_{\ell}^{\alpha}} \right) \right\rangle_{\theta},$$

realizamos o traço sobre a variável  $\tau_{\ell}^{\alpha} = \pm 1$ ,

$$f_2 = -\frac{1}{\beta n} \log \left\langle \sum_{\sigma} \exp \left[ \beta \theta \sum_{\alpha} \sigma^{\alpha} \right] \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k \frac{dh_{\ell} W(h_{\ell})}{\left[ 2 \cosh(\beta h_{\ell}) \right]^n} \right. \quad (2.59)$$

$$\left. \times \prod_{\alpha} \prod_{\ell=1}^k \left( e^{\beta h_{\ell} + \frac{\beta J}{c} \sigma^{\alpha}} + e^{-\beta h_{\ell} - \frac{\beta J}{c} \sigma^{\alpha}} \right) \right\rangle_{\theta}.$$

Fazendo o traço sobre a variável  $\sigma^{\alpha}$ ,

$$f_2 = -\frac{1}{\beta n} \log \left\langle \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) \left\{ \frac{e^{\beta \theta} \prod_{\ell} \left[ 2 \cosh \left( \beta h_{\ell} + \frac{\beta J}{c} \right) \right]}{2 \cosh(\beta h_{\ell})} \right. \right. \quad (2.60)$$

$$\left. \left. + \frac{e^{-\beta \theta} \prod_{\ell} \left[ 2 \cosh \left( \beta h_{\ell} - \frac{\beta J}{c} \right) \right]}{2 \cosh(\beta h_{\ell})} \right\}^n \right\rangle_{\theta}.$$

Mais uma vez, por meio das relações entre funções hiperbólicas,

$$f_2 = -\frac{1}{\beta n} \log \left\langle \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) \left\{ e^{\beta \theta} \prod_{\ell} \cosh(\beta J) \left( 1 + \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right. \right. \quad (2.61)$$

$$\left. \left. + e^{-\beta \theta} \prod_{\ell} \cosh(\beta J) \left( 1 - \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right\}^n \right\rangle_{\theta}.$$

A partir daí, por meio das aproximações  $x^n = e^{n \log(x)}$ , para  $n \rightarrow 0$ ,

$$f_2 = -\frac{1}{\beta n} \log \left\langle 1 + n \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) \log \left\{ e^{\beta \theta} \prod_{\ell} \cosh(\beta J) \left( 1 + \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right. \right. \\ \left. \left. + e^{-\beta \theta} \prod_{\ell} \cosh(\beta J) \left( 1 - \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right\} \right\rangle_{\theta}, \quad (2.62)$$

o que nos leva à equação

$$f_2 = -\frac{1}{\beta} \log \left\langle \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) \left\{ \log \left[ \cosh(\beta J) \right]^k + \log \left[ e^{\beta \theta} \prod_{\ell} \left( 1 + \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right. \right. \right. \\ \left. \left. \left. + e^{-\beta \theta} \prod_{\ell} \left( 1 - \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right] \right\} \right\rangle_{\theta}. \quad (2.63)$$

O termo da interação das ligações  $\cosh(\beta J)$ , não depende da média sobre os termos de campos locais. Portanto, podemos tirá-lo da média e ao realizar os pesos poissonianos,

$$f_2 = -\frac{c}{\beta} \log \left[ \cosh(\beta J) \right] - \frac{1}{\beta n} \left\langle \sum_{k=0}^{\infty} P_k \int \prod_{\ell=1}^k dh_{\ell} W(h_{\ell}) \log \left\{ e^{\beta \theta} \prod_{\ell} \left( 1 + \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right. \right. \\ \left. \left. + e^{-\beta \theta} \prod_{\ell} \left( 1 - \tanh(\beta h_{\ell}) \tanh(\beta J) \right) \right\} \right\rangle_{\theta}. \quad (2.64)$$

E finalmente, ao substituirmos as equações Eq.(2.48) e Eq.(2.64) na Eq.(2.45), obtemos a energia livre,

$$\begin{aligned}
f(\beta) &= -\frac{c}{2\beta} \log \left[ \cosh(\beta J) \right] + \frac{c}{2\beta} \int dh dh' W(h) W(h') \log \left[ 1 + \tanh(\beta h) \tanh(\beta h') \tanh(\beta J) \right] \\
&- \frac{1}{\beta} \sum_k P_k \int \prod_l dh_\ell W(h_\ell) \left\langle \log \left\{ e^{\beta\theta} \prod_\ell \left( 1 + \tanh(\beta h_\ell) \tanh(\beta J) \right) \right. \right. \\
&+ \left. \left. e^{-\beta\theta} \prod_\ell \left( 1 - \tanh(\beta h_\ell) \tanh(\beta J) \right) \right\} \right\rangle_\theta.
\end{aligned} \tag{2.65}$$

Uma vez que temos a magnetização  $m$ , distribuição de campos  $W(h)$  e energia livre  $f$ , podemos traçar as curvas dos nossos resultados e observar o comportamento do modelo.

## 2.5 Resultados teóricos

Um resultado relevante deste trabalho é a observação de um ponto tricrítico na transição entre as fases ferromagnética (F) e paramagnética (P), e sua dependência com a conectividade  $c$ . Os observáveis que guiarão a discussão são a magnetização  $m$  e a energia-livre  $f$ . Para simplificar a discussão a seguir, foi fixado  $J = \frac{1}{c}$ ,  $p = 0,5$  para uma distribuição de campo aleatório bimodal e  $\theta_0 = 0$  para uma distribuição de campo aleatório gaussiana. A equação de ponto-de-sela, Eq.(2.43) é resolvida por um método iterativo descrito em [49, 50].

Inicialmente definimos uma população de campos locais reais. Tipicamente o tamanho da população (número de sítios) é de  $10^4$ . Um valor positivo é atribuído para cada campo como condição inicial. Isto ocorre para que uma magnetização positiva seja favorecida, mas poderia ter sido escolhida uma magnetização negativa. Então, em cada passo um número natural  $k$  é sorteado de acordo com uma distribuição de Poisson com média  $c$ , e  $k$  campos são aleatoriamente escolhidos da população de campos. Dado os  $k$  campos sorteados, o somatório que aparece no argumento da função  $\delta$  do lado direito

da Eq. (2.43) é calculado juntamente com o campo aleatório. Com a aplicação deste algoritimo a população de campo converge para para a distribuição de campos locais  $W(h)$ . Tipicamente, cada campo local é atualizado 100 vezes. Depois de alguns testes, foi verificado que é suficiente para alcançar a convergência. Como o principal resultado desse método numérico, é interessante checar como  $W(h)$  aparece para determinadas condições. Para ilustrar como a distribuição de campos se comporta de acordo com a variação de temperatura, foram escolhidos os parâmetros  $c = 4,0$  e  $\theta = 0,1$ , além disso um valor positivo é atribuído a todos os campos.

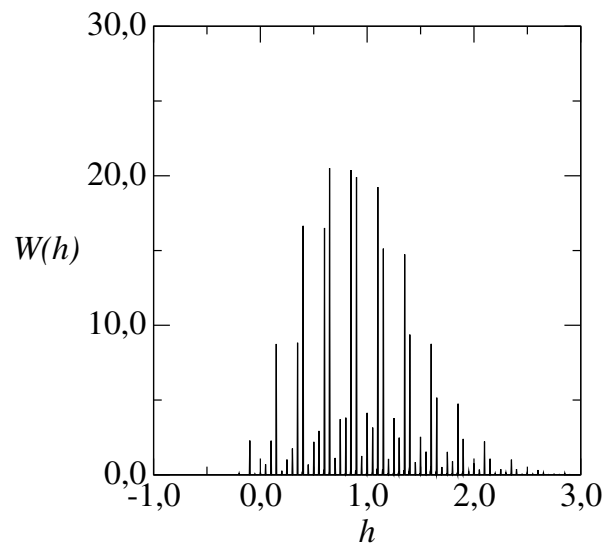


Figura 2.1: Distribuição para campo local com  $c = 4,0$ ,  $\theta = 0,1$  e  $T = 0,1$ .

Primeiramente a Fig.2.1 mostra que para o limite de temperaturas baixas,  $T = 0,1$ , o estado de equilíbrio é de fato ferromagnético uma vez que  $W(h)$  possui uma tendência positiva. Isto significa que a maior parte dos spins estão apontando para a mesma direção, pois para baixas temperaturas poucos *spins* trocam de direção. A distribuição de campos locais é discreta, o que reflete a conectividade de sítios (uma distribuição de Poisson) e a distribuição de campos aleatórios.

Em temperaturas intermediárias temos a Fig.2.2, onde  $T = 0,5$ . Podemos observar que o estado de equilíbrio permanece ferromagnético, mas devido a temperaturas maiores

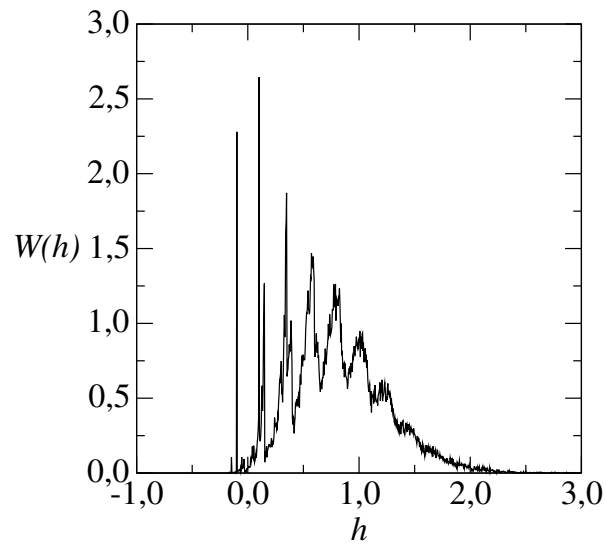


Figura 2.2: Distribuição para campo local com  $c = 4,0$ ,  $\theta = 0,1$  e  $T = 0,5$ .

a distribuição de campos torna-se contínua.

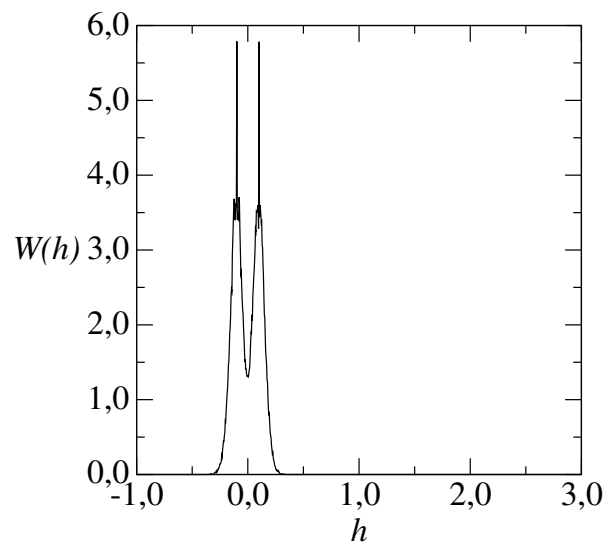


Figura 2.3: Distribuição para campo local com  $c = 4,0$ ,  $\theta = 0,1$  e  $T = 1,0$ .

Para altas temperaturas, como  $T = 1,0$ , a Fig.2.3 mostra que o estado de equilíbrio é paramagnético. Neste caso,  $W(h)$  é simetricamente distribuído em torno de  $h = 0$ , com deltas em  $h = \pm 0,1$  que são os dois possíveis valores do campo aleatório.

Foram traçados alguns gráficos da magnetização em função da amplitude  $\theta_0$  do campo aleatório. Neste caso é utilizada uma distribuição de campo bimodal onde foi fixado  $p = 0,5$ .

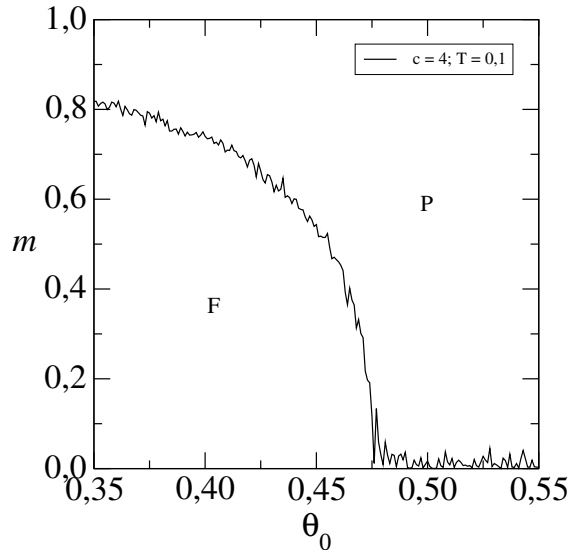


Figura 2.4: Magnetização  $m$  em função de  $\theta_0$  para  $p = 0,5$ ,  $c = 4,0$  e  $T = 0,1$ .

Podemos notar na Fig.2.4 que para conectividade  $c = 4,0$  com temperatura baixa,  $T = 0,1$ , a transição de fase  $F - P$  é contínua. Isso nos mostra que mesmo ao aumentar a temperatura, se  $c$  for muito pequeno, sempre teremos uma transição de fase de segunda ordem.

Fazendo  $c = 8$  e mantendo a temperatura baixa,  $T = 0,1$  na Fig.2.5 podemos perceber uma região de coexistência entre F e P, que caracteriza uma transição descontínua. A linha contínua (preta) é referente a o caso em que o sistema evolui ao aumentar o  $\theta_0$ , o ponto crítico para a espinodal neste caso ocorre na região aproximada de  $\theta_0 \sim 0,50$ . A linha tracejada (vermelha) mostra o caso em que o sistema evolui ao diminuir o  $\theta_0$  o ponto crítico para a espinodal neste caso ocorre na região aproximada de  $\theta_0 \sim 0,46$ .

Ao manter a conectividade  $c = 8,0$  e aumentar a temperatura para  $T = 0,6$  é observado na Fig.2.6 que a transição descontínua torna-se contínua. Este resultado indica a existência de um ponto tricrítico entre as temperaturas  $T = 0,1$  e  $T = 0,6$  para  $c = 8,0$ .



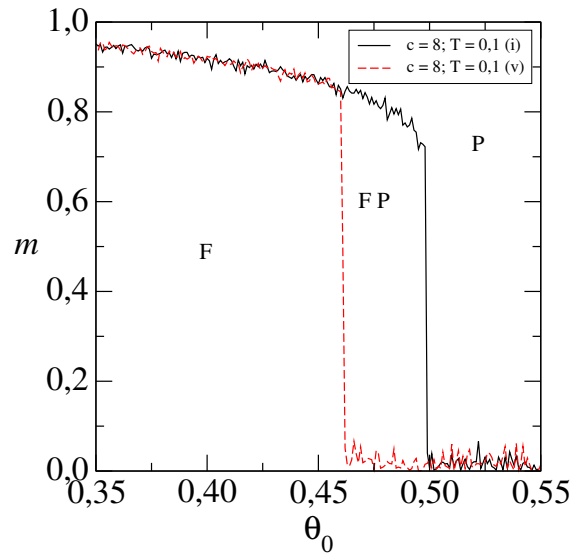


Figura 2.5: Magnetização  $m$  em função de  $\theta_0$  para  $p = 0,5$ ,  $c = 8,0$  e  $T = 0,1$ ; Linha sólida (preta) corresponde a evolução do sistema ao aumentar  $\theta_0$ ; Linha tracejada (vermelha) corresponde à evolução do sistema ao diminuir  $\theta_0$ .

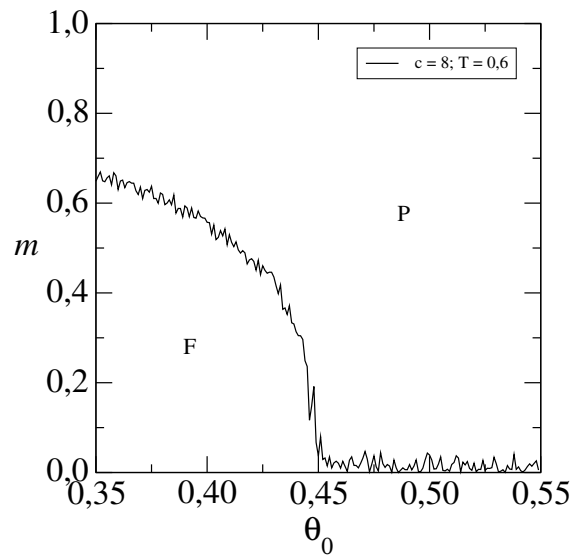


Figura 2.6: Magnetização  $m$  em função de  $\theta_0$  para  $p = 0,5$ ,  $c = 8,0$  e  $T = 0,6$ .

Na Fig.2.7 temos as curvas de energia-livre em função da amplitude  $\theta_0$  do campo aleatório. Observa-se que para  $c = 4,0$  e  $T = 0,1$  (linha pontilhada azul) e para  $c = 8,0$   $T = 0,6$  (linha sólida verde), a transição de fase é contínua. As curvas da energia-livre

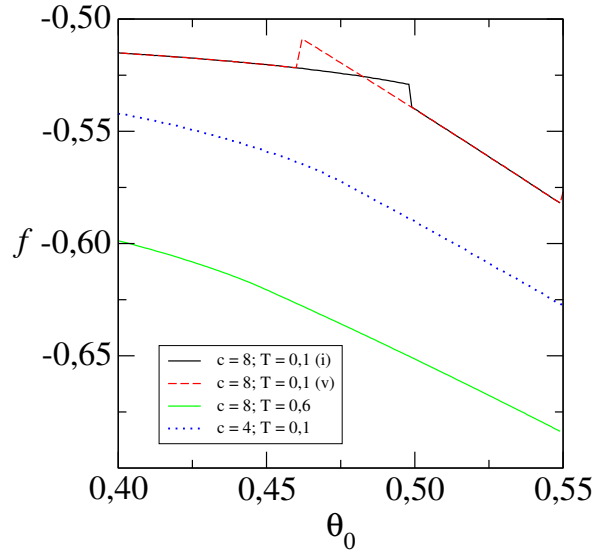


Figura 2.7: Energia livre  $f$  em função de  $\theta_0$  para  $p = 0,5$ . Linha sólida (preta) e linha tracejada (vermelha) correspondem a  $c = 8,0$  e  $T = 0,1$  aumentando e diminuindo o  $\theta_0$  respectivamente; Linha pontilhada (azul) a  $c = 4,0$  e  $T = 0,1$ ; Linha sólida (verde) corresponde a  $c = 8,0$  e  $T = 0,6$ .

são iguais ao aumentarmos ou diminuirmos o  $\theta_0$ . O mesmo não ocorre para  $c = 8,0$  e  $T = 0,1$ , onde o cruzamento entre a energia-livre para um sistema que evolui ao aumentar o  $\theta_0$  e a energia-livre para um sistema que evolui ao diminuir o  $\theta_0$  nos fornece o  $\theta_0$  crítico  $\theta_0 \approx 0,47$ .

O comportamento geral de uma rede de conectividade finita sobre o efeito de um campo aleatório bimodal pode ser resumido pelos diagramas de fases de  $T$  contra  $\theta_0$  mostrados na Fig. 2.8 (*esquerda*). Como discutido, para  $c = 4,0$ , a transição de fase é sempre contínua. Quando a conectividade aumenta para  $c = 8,0$ , a transição F-P permanece contínua para altas temperaturas, mas torna-se descontínua em um regime de baixas temperaturas. Existe um ponto tricrítico em  $T_{tri} \sim 0,45$  e  $\theta_{0,tri} \sim 4,7$ . Em ambos os casos, contínuo e descontínuo, a linha de transição termodinâmica incide perpendicularmente no eixo horizontal  $T = 0$ . Isto é, não foi observada uma reentrância. Além disso é interessante salientar que o tamanho da fase F é fracamente afetada pela conectividade. Isto é justificado pelo fato de que, no modelo atual, os acoplamentos são

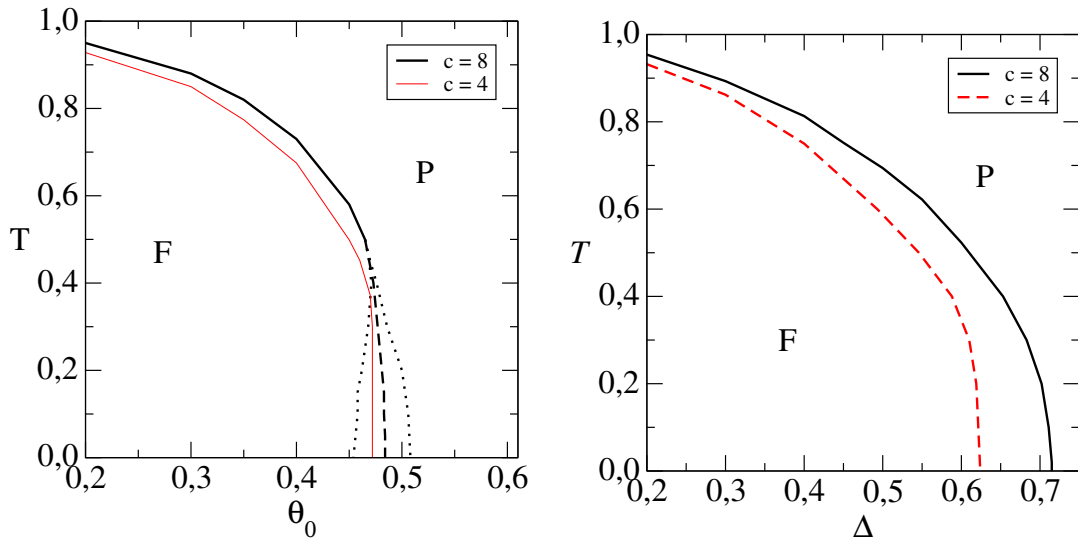


Figura 2.8: Esquerda: Diagramas de fases  $T$  vs.  $\theta_0$  para uma distribuição bimodal do campo aleatório. Linha fina (vermelha), sólida, transição F-P contínua para  $c = 4,0$ . Linhas grossas (pretas) corresponde a  $c = 8,0$ : linha sólida, transição F-P contínua; linha tracejada, transição F-P descontínua; linha pontilhada, espinodais da transição descontínua. Direita: Diagrama de fase  $T$  vs.  $\Delta$  para distribuição gaussiana do campo aleatório. Linha tracejada (vermelha) e linha sólida (preta): Transição contínua F-P para  $c = 4,0$  e  $c = 8,0$  respectivamente.

normalizados com  $c$ . Para completar, o diagrama de fases para uma distribuição gaussiana com  $c = 4,0$  e  $c = 8,0$  é mostrado na Fig.2.8 (*direita*). Ao contrário da distribuição bimodal, a transição é sempre contínua. Os resultados apresentados demonstram que a natureza da transição F-P em um ferromagneto com campo aleatório não depende apenas do tipo de distribuição do campo aleatório, mas também da conectividade.

Mais detalhes na dependência da conectividade são dados na Fig. 2.9, onde a temperatura do ponto tricrítico é mostrada como função de  $c$ , para a distribuição bimodal do campo aleatório. Isto indica que a temperatura tricrítica decai com  $c$ , e é suprimida para  $c < c_{min} = 5,0$ . Para este leque de conectividade, a transição F-P é sempre contínua mesmo para uma distribuição bimodal, ou seja, usando uma distribuição de campo aleatório simétrica com um mínimo.

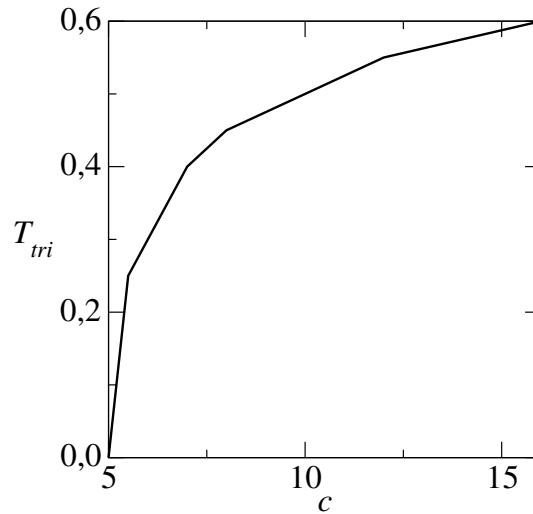


Figura 2.9: Ponto tricrítico como função da conectividade  $c$ , para uma distribuição bimodal do campo aleatório.

## 2.6 Simulação Numérica

Foram realizadas simulações numéricas para verificar a validade dos resultados teóricos.

### 2.6.1 Método de Monte Carlo

O objetivo da simulação de Monte Carlo de um sistema térmico é o cálculo do valor esperado  $\langle Q \rangle$  de uma quantidade observável  $Q$ , tal como a energia interna de um gás, ou a magnetização em um modelo magnético [51].

Idealmente, calcula-se a média da quantidade sobre todos os estados  $\mu$  do sistema, pesando-os com sua própria probabilidade de Boltzmann

$$\langle Q \rangle = \frac{\sum_{\mu} Q_{\mu} e^{-\beta E_{\mu}}}{\sum_{\mu} e^{-\beta E_{\mu}}}, \quad (2.66)$$

onde  $E_{\mu}$  corresponde à energia no estado  $\mu$ . No entanto, para grandes sistemas não é possível fazer a média sobre todos os estados. Por isso é feita a média sobre um subgrupo

deles, apesar de introduzir imprecisões aos cálculos. O método seleciona um subgrupo aleatoriamente por meio de uma distribuição de probabilidade  $p_\mu$ . Ao escolher  $M$  estados  $\{\mu_1 \cdots \mu_M\}$ , a melhor estimativa para a quantidade  $Q$  é dada por

$$Q_M = \frac{\sum_{i=1}^M Q_{\mu_i} p_{\mu_i}^{-1} e^{-\beta E_{\mu_i}}}{\sum_{j=1}^M p_{\mu_j}^{-1} e^{-\beta E_{\mu_j}}}. \quad (2.67)$$

$Q_M$  é conhecido como o estimador de  $Q$ , para  $M \rightarrow \infty$ , obtêm-se  $Q_M = \langle Q \rangle$ . Para obter um estimador preciso de  $\langle Q \rangle$  é necessário optar por uma distribuição de probabilidade de forma cuidadosa, caso contrário a soma sobre  $\mu_i$  pode ser dominada por uma pequena parcela de estados, com a contribuição da maioria sendo praticamente negligenciável. Por outro lado, se for possível descobrir quais estados contribuem mais para a soma da média e selecionar apenas esses estados, seria obtido um bom estimador.

Para fazer uma amostragem de importância, escolhe-se uma amostra de acordo com a distribuição de Boltzmann. Com  $p_\mu = e^{-\beta E_\mu} / \sum_\mu e^{-\beta E_\mu}$ . O estimador torna-se

$$Q_M = \frac{1}{M} \sum_i^M Q_{\mu_i}. \quad (2.68)$$

Para que o conjunto de estados seja gerado com uma distribuição de Boltzmann, é utilizado um processo de Markov, que é um mecanismo que dado um estado inicial  $\mu$ , gera um novo estado  $\nu$  de maneira aleatória de forma que o estado gerado é diferente, mesmo se for selecionado  $\mu$  como estado inicial novamente.

A probabilidade de gerar um estado  $\nu$  dado  $\mu$  é chamada de probabilidade de transição  $P(\mu \rightarrow \nu)$  e esta não pode variar no tempo e deve depender apenas das propriedades dos estados  $\mu$  e  $\nu$ . Além de satisfazer ao vínculo

$$\sum_\nu P(\mu \rightarrow \nu) = 1. \quad (2.69)$$

Na simulação de Monte Carlo, é utilizado o processo de Markov repetidamente para gerar

uma cadeia de Markov de estados. Quando o processo é repetido por vezes o suficiente o sistema chega ao equilíbrio e os estados que aparecem possuem probabilidades dadas pela distribuição de Boltzmann. Para que isso seja possível são adicionadas mais duas condições para o processo de Markov, a ergodicidade e o balanço detalhado.

A ergodicidade é o requerimento de que deve ser possível para o processo de Markov alcançar qualquer estado do sistema de qualquer outro estado, se realizarmos esse processo por tempo o suficiente.

O balanço detalhado garante que a distribuição após o sistema entrar em equilíbrio seja a distribuição de Boltzmann. Ao considerar que a taxa em que o sistema faz a transição de qualquer e para qualquer estado  $\mu$  deve ser igual a

$$\sum_{\nu} p_{\mu} P(\mu \rightarrow \nu) = \sum_{\nu} p_{\nu} P(\nu \rightarrow \mu), \quad (2.70)$$

tiramos a soma em  $\nu$  de ambos os lados da equação e chegamos na equação de balanço detalhado.

$$p_{\mu} P(\mu \rightarrow \nu) = p_{\nu} P(\nu \rightarrow \mu). \quad (2.71)$$

Dado que o objetivo é que no equilíbrio o sistema possua uma distribuição de Boltzmann, a equação de balanço detalhado diz que a probabilidade de transição deve obedecer

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{p_{\nu}}{p_{\mu}} = e^{-\beta(E_{\nu} - E_{\mu})}, \quad (2.72)$$

a Eq.(2.69) e a Eq.(2.72) são vínculos escolhidos para que no equilíbrio o processo de Markov seja uma distribuição de Boltzmann.

Contudo, é difícil saber qual processo de Markov possui a transição de probabilidade apropriada e encontrar isto pode ser um procedimento longo de tentativa e erro. Para evitar este método empírico é então escolhida uma probabilidade de transição  $P(\mu \rightarrow \mu)$

diferente de zero. Ao fazer  $\nu = \mu$  na Eq.(2.72) obtêm-se  $1 = 1$ , o que mostra que o balanço detalhado sempre é satisfeito para  $P(\mu \rightarrow \mu)$ . Isto significa que é possível ajustar o valor de qualquer  $P(\mu \rightarrow \nu)$  e manter a Eq.(2.69) satisfeita ao compensar com um ajuste igual e contrário de  $P(\mu \rightarrow \mu)$ , contanto que  $P(\mu \rightarrow \mu)$  mantenha-se entre 0 e 1. Para que a Eq.(2.72) também mantenha-se satisfeita, simultaneamente é feita uma mudança em  $P(\nu \rightarrow \mu)$ , para que a razão entre as duas probabilidades permaneça constante. Então separa-se a probabilidade de transição em dois termos

$$P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu)A(\mu \rightarrow \nu), \quad (2.73)$$

a quantidade  $g(\mu \rightarrow \nu)$  é a probabilidade selecionada que dado um estado inicial  $\mu$  gera um estado  $\nu$ , e  $A(\mu \rightarrow \nu)$  é a razão de aceitação, que tem a liberdade de variar entre 0 e 1 e representa a aceitação da mudança do estado  $\mu$  para  $\nu$  em uma fração de vezes. Isso permite que  $g(\mu \rightarrow \nu)$  tenha qualquer valor, já que o vínculo da Eq.(2.72) apenas fixa a razão

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{g(\mu \rightarrow \nu)A(\mu \rightarrow \nu)}{g(\nu \rightarrow \mu)A(\nu \rightarrow \mu)}. \quad (2.74)$$

Existem muitos algoritmos de Monte Carlo e cada um utiliza uma forma diferente na obtenção do  $g(\mu \rightarrow \nu)$  e do  $A(\mu \rightarrow \nu)$ . Para a resolução do modelo de Ising, um dos métodos mais utilizado é o algoritmo de Metropolis [51]. Neste caso, a probabilidade selecionada  $g(\mu \rightarrow \nu)$  dos possíveis estados de  $\nu$  é escolhida como igual. Todos os outros estados possuem a probabilidade selecionada nula. Se o sistema simulado tem  $N$  spins, então existem  $N$  estados possíveis de  $\nu$  que podem ser alcançados do estado  $\mu$ . Portanto existem  $N$  probabilidades selecionadas que toma o valor de

$$g(\mu \rightarrow \nu) = \frac{1}{N}. \quad (2.75)$$

Ao aplicar a Eq.(2.75) na Eq.(2.74) chegamos em

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta(E_\nu - E_\mu)}, \quad (2.76)$$

onde o termo  $e^{-\beta(E_\nu - E_\mu)}$  é um peso de Boltzmann e a energia  $E_\nu$  é maior do que a energia  $E_\mu$ .

Como as razões de aceitação estão entre os valores 0 e 1, fazemos a maior delas igual a 1 e a outra toma qualquer valor necessário para que a Eq.(2.76) seja respeitada, portanto

$$A(\mu \rightarrow \nu) = \begin{cases} e^{-\beta(E_\nu - E_\mu)} & \text{se } E_\nu - E_\mu > 0; \\ 1 & \text{para outros casos.} \end{cases} \quad (2.77)$$

O que significa que a transição sempre é aceita para estados com energia menor ou igual ao estado atual. Caso o estado possua energia maior, é utilizado um peso de Boltzmann para decidir se a transição é aceita ou não.

## 2.6.2 Resultados numéricos

Esta seção utiliza um algoritmo de Metropolis para redes de tamanho  $N = 256$  até  $N = 4096$  adaptado de [52].

Além do procedimento padrão do método de Metropolis para atualizar o estado do sistema, foram criadas  $L$  réplicas não interagentes, cada uma com diferentes temperaturas  $T_i$ . Para os resultados teóricos obtidos nesse trabalho, as curvas de coexistência a baixas temperaturas são sensíveis à amplitude do campo aleatório. Por consequência, para que cada uma das  $L$  cópias evoluam em paralelo, foram atribuídos diferentes valores de campo aleatório  $\theta_i$  a uma temperatura fixa.

Essas  $L$  cópias existem para evitar mínimos locais do espaço de configurações da energia, ou seja, cada sistema possui diferentes temperaturas fixas. Ao evoluir essas cópias paralelamente, variando os valores dos campos aleatórios  $\theta_i$ , é possível que uma



das cópias entre em um mínimo local, mas outra com um campo aleatório ligeiramente diferente não fique presa por este mínimo. Eventualmente, os estados entre duas réplicas são trocados seguindo uma distribuição de Boltzmann. O sistema com  $\theta_i$  maior atravessa o mínimo local e ao fazê-lo devolve o estado para a réplica de menor campo aleatório.

A figura 2.10 (a) (topo) mostra o resultado da razão de  $m$  vs.  $\theta_0$  com  $c = 4,0$  para redes de tamanho  $N = 256$  e  $N = 512$ . Apesar de uma pequena discrepância, os resultados numéricos concordam com a natureza contínua da transição. Na figura 2.10 (a) (baixo) temos  $c = 8,0$  e  $N = 256$ ,  $N = 512$  e  $N = 1024$ , temos uma curva compatível com a transição descontínua, temos um  $\theta_{0c} \approx 0,48$ . Na figura 2.10 (b), temos uma distribuição gaussiana para  $c = 8,0$ ,  $N = 256$  e  $N = 512$  e como esperado a transição de fase é contínua.

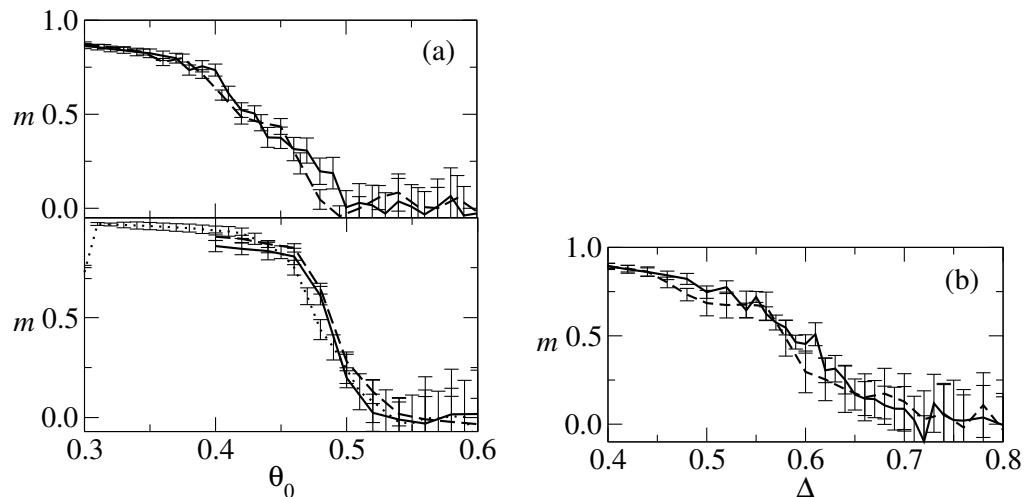


Figura 2.10: (a) Topo: Magnetização  $m$  em função de  $\theta_0$ , para  $T = 0,1$  constante, com distribuição bimodal do campo aleatório,  $c = 4,0$ ; tamanhos das redes são  $N = 256$  (linha tracejada) e  $N = 512$  (linha sólida). (a) Baixo:  $c = 8,0$ ; tamanho da rede é  $N = 256$  (linha pontilhada),  $N = 512$  (linha tracejada) e  $N = 1024$  (linha sólida). (b) Magnetização  $m$  em função de  $\Delta$ , a  $T = 0,1$  constante, com distribuição gaussiana do campo aleatório, para  $c = 8,0$ ; tamanhos das redes são  $N = 256$  (linha tracejada) e  $N = 512$  (linha sólida).

A figura 2.11 (a) mostra mais evidências da descontinuidade da transição de fases  $F - P$  para baixas temperaturas provendo as curvas de histerese para a razão de  $m$  por

$\theta_0$ . Podemos comparar esses resultados com o da Fig.2.5. A figura 2.11(b) mostra as curvas espinodais superiores e inferiores em função de  $T$  e podemos comparar com os resultados da figura 2.8.

As simulações estão de acordo com os resultados teóricos. Para conectividade pequena  $c < 5,0$ , independe se a distribuição de campos seja bimodal ou gaussiana, a transição de fase  $F - P$  é sempre contínua, mesmo em temperaturas baixas, como podemos comparar a linha de  $c = 4,0$  na Fig.2.8 (esquerda) e a Fig.2.10 (a) topo. Para  $c \geq 5,0$  em uma distribuição gaussiana, a transição  $F - P$  ainda é contínua como podemos comparar por meio da curva de  $c = 8,0$  na Fig.2.8 (direita) e a Fig.2.10 (b). Para  $c \geq 5,0$  em uma distribuição bimodal, encontramos na teoria e no experimento uma transição  $F - P$  descontínua como podemos comparar para a curva de  $c = 8,0$  na Fig.2.8 (esquerda) e a Fig.2.10 (a) (baixo). As espinodais teóricas e experimentais também estão de acordo com valores teóricos ambas com  $\theta_0$  aproximadamente entre 0,46 e 0,51, como pode ser comparado para as curvas pontilhadas na Fig.2.11 (direita) e na Fig.2.8 (esquerda).

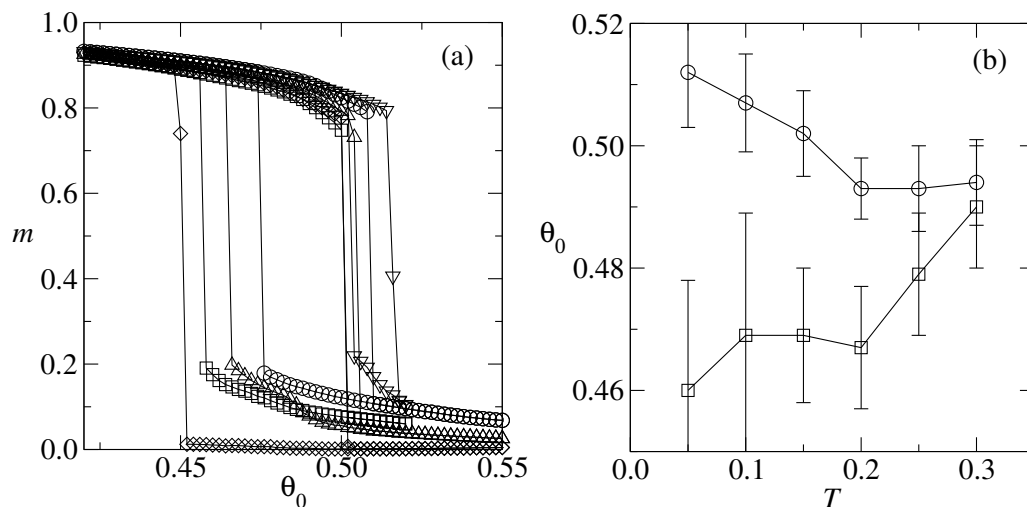


Figura 2.11: Resultados da simulação para redes com  $N = 4096$ ,  $c = 8,0$ , bimodal, com  $p = 0,5$ . (a) Magnetização  $m$  em função da amplitude do campo aleatório  $\theta_0$ . Os dados mostrados corresponde a 5 realizações de desordem, para  $\theta_0$  crescente e decrescente. (b)  $\theta_0$ -espinodais em função da temperatura  $T$ . Cada ponto corresponde a uma média sobre 10 realizações de desordem.

## Capítulo 3

# Recuperação de padrões estruturados utilizando uma rede métrica atratora: aplicação no reconhecimento de impressões digitais

### 3.1 Introdução

Investigar a recuperação de padrões por meio de uma Rede Neural Atratora (ANN como sua sigla em inglês) tem sido um dos objetivos da comunidade científica já há algum tempo. Principalmente ao tratar de conectividade uniforme, como por exemplo redes completamente conectadas [53] e redes aleatórias [25].

Entretanto, ao trabalharmos com sistemas neuronais biológicos reais e suas contra-

partes artificiais, sabemos que esses sistemas são redes métricas e, portanto, a distância entre os sítios não é uniforme.

Além disso, apesar de demonstrar resultados relativamente satisfatórios na recuperação de padrões aleatórios uniformes, a recuperação de padrões que possuem estados distribuídos de forma não uniforme no espaço, não é boa para redes atratoras.

Para lidar com problemas que levam em conta padrões estruturados, como classificação biométrica, são introduzidas as Redes Neurais Atratoras Métricas (MANN como sua sigla em inglês) [54]. Essas redes possuem conexões organizadas localmente com algumas ligações "atalho" de longo alcance.

Um problema típico de classificação biométrica é o do reconhecimento de impressões digitais. Levando em consideração a detecção e a medição de características, tais como pontos singulares, direções ou bifurcações de cumes, é possível identificar o dono de uma impressão digital (ID).

Algumas Redes neurais são aplicadas em tarefas como classificação, reconhecimento e verificação dessas impressões digitais [55, 56, 57]. Inclusive, essas redes também são utilizadas na identificação de impressões digitais [58, 59].

Os contornos da superfície que definem um objeto físico são conhecidos como região de interesse (ROI). Idealmente, todo o contorno de uma ID seria sua ROI. Porém, por causa de recursos computacionais, essas redes atratoras, consideram uma região de interesse (ROI) menor do que deveriam. É possível aumentar a ROI para seu tamanho ideal se utilizarmos uma rede métrica diluída, que permite cobrir praticamente toda a ID.

Trabalhos recentes demonstraram um sucesso razoável na recuperação de padrões de impressões digitais por meio de uma MANN [60]. Neste modelo, um parâmetro equivalente ao da probabilidade de reconexão de uma rede de pequeno mundo, é conhecido como aleatoriedade topológica  $\omega$ . E é obtido por meio da razão entre as conexões de longo alcance sobre toda a conectividade. Este parâmetro possui um papel importante na recuperação de padrões.

Em geral, ao considerarmos uma estrutura local de padrões, tentamos responder a uma pergunta: "o papel da conectividade da rede é sempre relevante?".

Apesar de termos algumas respostas [61, 62, 63, 64, 65, 66, 67], elas são baseadas em situações específicas. Para respondermos de uma maneira mais geral utilizamos uma MANN e introduzimos um parâmetro que mede a localidade dos padrões  $\psi$ .

Ao resolvermos o modelo, obtemos a sobreposição entre os padrões  $m$  e os estados neuronais, como função da capacidade de carga ao variarmos  $\omega$  e  $\psi$ .

É importante salientar que trabalhos anteriores consideravam a atividade média dos padrões como  $a = 0,23$ , por tratar de um valor semelhante a estruturas reais das IDs.

Entretanto, cada ID possui sua atividade típica. Por isso, foram realizados novos experimentos com IDs, mas variando a atividade média dos padrões, indo de  $a \sim 0,05$  (código esparsos) até  $a \sim 0,48$ .

Além disso, foi desenvolvida uma teoria levando em conta uma média na topologia (portanto os detalhes da métrica não serão capturados).

## 3.2 O modelo

O estado da rede em um dado tempo  $t$  é definido por um conjunto de  $N$  variáveis de neurônios binários e independentes  $\boldsymbol{\tau}^t = \{\tau_i^t \in \{0,1\}; i = 1, \dots, N\}$ , onde 1 e  $-1$  representam, respectivamente, estados ativos e inativos. A rede tem o objetivo de recuperar  $P$  padrões independentes  $\{\boldsymbol{\eta}^\mu, \mu = 1, \dots, P\}$  que foram armazenados por um processo de aprendizado. Isto significa que um estado recuperado estável satisfaz  $\boldsymbol{\tau}^t = \boldsymbol{\eta}^\mu$ , para um tempo  $t$  grande o suficiente.

Cada padrão,  $\boldsymbol{\eta}^\mu = \{\eta_i^\mu \in \pm 1; i = 1, \dots, N\}$  é caracterizado pela atividade média  $a^\mu = \langle \eta^\mu \rangle \equiv \sum_i \frac{\eta_i^\mu}{N}$  e pela correlação espacial  $\psi$ .

Os acoplamentos sinápticos entre os neurônios  $i$  e  $j$  são dados por uma matriz de

interação fatorada em uma matriz de topologia  $c_{ij}$  e nos pesos de aprendizagem  $W_{ij}$ ,

$$J_{ij} \equiv c_{ij}W_{ij}. \quad (3.1)$$

A topologia, definida pela matriz,  $c_{ij} \in \{0,1\}$ , com  $c_{ii} = 0$ , divide as conexões entre locais e aleatórias. As ligações locais conectam cada neurônio com seus  $c_\ell$  vizinhos mais próximos. As ligações aleatórias,  $c_r$ , são distribuída ao longo da rede. Portanto a conectividade total da rede é

$$c = c_\ell + c_r. \quad (3.2)$$

Desta forma, a topologia da rede é caracterizada por dois parâmetros: a razão de conectividade

$$\gamma = \frac{c}{N}, \quad (3.3)$$

e a razão de aleatoriedade,

$$\omega = \frac{c_r}{c}. \quad (3.4)$$

Por conveniência, a seguir utilizaremos variáveis normalizadas,  $\sigma \equiv \frac{\tau - q}{\sqrt{Q}}$  e  $\xi \equiv \frac{\eta - a}{\sqrt{A}}$ , sendo que  $q \equiv \langle \tau \rangle$  é a atividade média dos neurônios,  $a \equiv \langle \eta \rangle$  a atividade média dos padrões,  $Q \equiv Var(\tau)$  a variância dos neurônios e  $A \equiv Var(\eta)$  a variância dos padrões.

Por meio dessas variáveis normalizadas, a recuperação de padrões é alcançada por

meio de uma dinâmica determinística,

$$\sigma_i^{t+1} = g(h_i^t - \theta_i^t, q_i^t), \quad (3.5)$$

$$h_i^t \equiv \frac{1}{c} \sum_j J_{ij} \sigma_j^t, i = 1, \dots, N,$$

onde  $h_i^t$  corresponde ao campo local no neurônio  $i$  no tempo  $t$ , e  $\theta_i$  é o limiar de disparo. A função ganho é definida da seguinte forma

$$g(x, y) \equiv \frac{\Theta(x) - y}{\sqrt{y(1 - y)}}, \quad (3.6)$$

sendo  $\Theta(x)$  uma função degrau.

A matriz peso é atualizada de acordo com a regra de Hebb,

$$W_{ij}^\mu = W_{ij}^{\mu-1} + \xi_i^\mu \xi_j^\mu, \quad (3.7)$$

os pesos começam em  $W_{ij}^0 = 0$  e após  $P$  passos de aprendizado eles alcançam o valor  $W_{ij} = \sum_\mu \xi_i^\mu \xi_j^\mu$ .

Para recuperarmos padrões com baixa atividade,  $a < 0,5$ , é necessário utilizar uma estratégia de limiar adaptativo, caso contrário, a atividade pode ser maior (menor) do que a atividade do padrão sempre que o limiar for pequeno (grande). Na literatura de código esparsos [68] foi desenvolvido o limiar

$$\theta_i^t = \begin{cases} \frac{1-2a}{2\sqrt{A}} & \text{se } q_i^t < 0,5; \\ -\frac{1-2a}{2\sqrt{A}} & \text{se } q_i^t > 0,5. \end{cases} \quad (3.8)$$

Também definimos um parâmetro  $\zeta$ , com um valor empírico de  $\zeta = 0,7$  e trocamos  $\theta_0$  por  $\zeta\theta_0$  ( $\frac{1}{\zeta}\theta_0$ ) para os primeiros (últimos) passos de tempo, digamos  $t < 20$  ( $t \geq 20$ ), para

que a rede evite ser presa em mínimos espúrios.

Para avaliarmos a habilidade de recuperação da rede utilizamos dois parâmetros obtidos por meio das equações de dinâmica, a sobreposição de padrões sobre os estados da rede que a partir de agora chamaremos apenas de “sobreposição” e a atividade neural média definida pelas expressões

$$m^t \equiv \frac{1}{N} \sum_i^N \xi_i \sigma_i^t, \tag{3.9}$$

$$q^t \equiv \frac{1}{N} \sum_i^N \tau_i^t.$$

Quando o padrão é recuperado, temos que  $m \sim 1$  e  $q \sim a$ . Além da sobreposição dos estados pelos padrões e da atividade neural, também nos interessamos pela relação de carga,  $\alpha \equiv \frac{P}{N}$ , que é a razão entre os padrões armazenados e o tamanho da rede.

### 3.3 Experimentos de impressões digitais

Os experimentos utilizaram IDs coletadas da Competição Internacional de Verificação de Impressões Digitais, *FVC2000*, *FVC2002* e *FVC2004* [69].

Esses padrões de ID foram pré-processados usando o *MATLAB* [70] e por meio dos seguintes passos: Realce, binarização, operações morfológicas e recorte de imagens.

Por meio da *Short Time Fourier Transform* (STFT) é realizado o realce da imagem. Em seguida é utilizada a função *im2bw* para se obter uma imagem binarizada. Depois, por meio da função *bwmorph*, são feitas operações morfológicas, tais como *Thin*, que remove pixels de objetos e os afina em linhas ou anéis, *Close*, que realiza um fecho morfológico, *Clean* e *Spur*, que removem respectivamente pixels isolados e espúrios. Finalmente a imagem é cortada de uma resolução original de  $388 \times 374$  pixels mantendo a



maior parte da informação da ID para uma resolução final de  $263 \times 340$  pixels. Podemos ver exemplos na Fig.3.1.



Figura 3.1: (Da esquerda para direita) ID em tons de cinza. ID realçada e binarizada ( $a \sim 0,23$ ). ID dilatada ( $a \sim 0,48$ ). ID esqueletizada ( $a \sim 0,08$ ).

### 3.3.1 Configurações topológicas e performance de recuperação

Utilizando uma rede de pequeno-mundo, foram realizados vários testes na configuração das topologias. Cada topologia representa como os sítios estão conectados. Os parâmetros que as definem são a conectividade  $c$  e a razão de aleatoriedade  $\omega$ . Ao aumentarmos  $c$ , teremos mais conexões entre os sítios, sempre respeitando a forma topológica escolhida. Para  $\omega = 0,0$  as ligações são sempre locais. Ao aumentar o  $\omega$  existirão conexões aleatórias de longo alcance entre os sítios. Foi considerada uma condição periódica de contorno.

As topologias utilizadas neste trabalho, são representadas esquematicamente na Fig.3.2. Da esquerda para a direita são: Anel, Cruz e X. Sendo  $c = 4$   $\omega = 0,0$  para todos os painéis.

A diluição da rede é um fator importante que determina o número de padrões que a rede é capaz de armazenar ( $P \equiv \alpha c$ ), sem contar que, computacionalmente o tempo de aprendizado/recuperação é da ordem de  $O(N \times c \times P)$ .

Os estados estacionários  $m^*$  alcançados pela rede foram estudados como função da relação de carga  $\alpha$ , para diferentes valores do parâmetro topológico  $\omega$ .

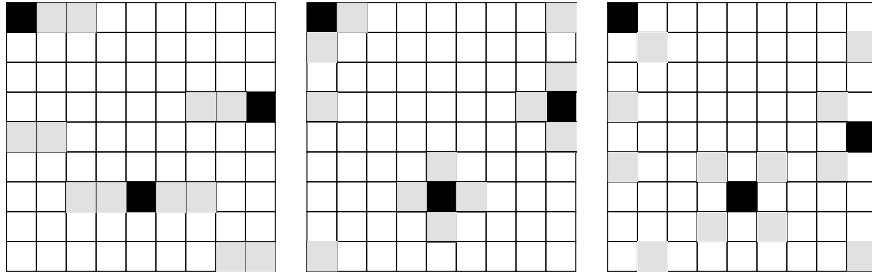


Figura 3.2: Topologias (preto = sítio de referência, cinza = sítios vizinhos): Anel (Esquerda), Cruz (meio), X (direita).  $c = 4, \omega = 0.0$ .

Os resultados das simulações são traçados para uma rede com  $N = 340 \times 263 = 89420$  (que corresponde à resolução espacial de uma imagem de uma ID), e  $c = 200$ . A rede começa com uma condição inicial sem ruído com  $m^0 = 1,0$  e evolui para o valor de  $t = 100$  passos no tempo, que foi encontrado empiricamente como sendo um tempo longo o suficiente. Um exemplo da recuperação estacionária das sobreposições  $m^*$  é mostrado na Fig.3.3. Nesta figura,  $m^*$  é representada com símbolos e as linhas sólidas são correspondentes a uma média para os diferentes valores de  $\omega$ .

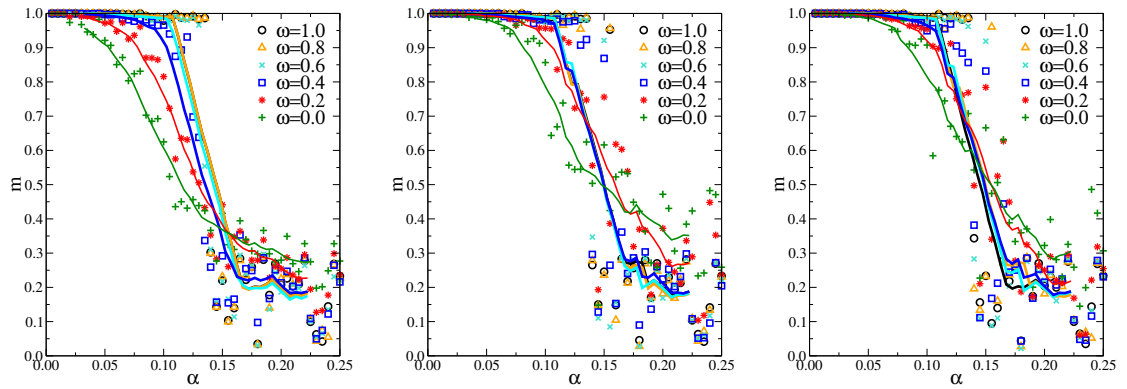


Figura 3.3: Capacidade de recuperação de impressões digitais sem ruído para diferentes valores do parâmetro de aleatoriedade  $\omega$ , com  $N = 89420$ ,  $c = 200$ , condição inicial  $m^0 = 1,0$  e atividade  $a = 0,23$ . Anel(esquerda), Cruz(meio), X(direita). Sendo  $m^*$  representada com símbolos e as linhas sólidas são correspondentes a uma média suave para os diferentes valores de  $\omega$ .

É possível definir o valor crítico  $\alpha_c$  tal que a sobreposição cai abaixo de  $m_c \sim 0,5$ . No painel da direita para a topologia X com  $\omega = 1,0$ , é possível observar que a rede

possui uma relação de carga crítica de  $\alpha_c = \frac{P}{c} \sim 0,13$ . A rede diminui sua performance para pequenos valores de  $\omega$ , alcançando  $\alpha_c \sim 0,09$  para  $\omega = 0,0$ . Entretanto, para  $\omega = 0,4$  a performance de recuperação da rede é equivalente a de uma rede com  $\omega = 1,0$ , com exceção do custo das conexões. A topologia Cruz no painel central mostra um comportamento similar. A topologia de Anel, mostra uma piora na recuperação para  $\omega < 1,0$ , como ilustrado no painel da esquerda.

### 3.3.2 IDs esqueletizadas

No aprendizado de IDs esqueletizadas (com baixa atividade  $a \sim 0,085$ ,  $\theta_0 = 1,49$ ) a rede armazena mais padrões devido à escassez de atividade, e por isso ocorre a ativação intensa de um conjunto relativamente pequeno de neurônios [71].

Os estados estacionários  $m^*$  alcançados pela rede foram estudados em função da relação de carga  $\alpha$ , para valores diferentes do parâmetro topológico  $\omega$ .

Os resultados de simulação são traçados para uma rede com  $N = 89420$  e  $c = 200$ . A rede possui uma condição inicial sem ruído de  $m^0 = 1,0$ , e é atualizada para  $t = 100$  passos no tempo.

Na Fig.3.4,  $m^*$  é representada com símbolos e as linhas sólidas representam a média suavizada das curvas para os diferentes  $\omega$ . No painel da direita, para a topologia X com  $\omega = 1,0$ , pode ser verificado que a relação de carga crítica é de  $\alpha_c = \frac{P}{c} \sim 0,29$ . A rede diminui sua performance para baixos valores de  $\omega$ , alcançando  $\alpha_c \sim 0,125$  para  $\omega = 0,0$ . Entretanto, para  $\omega = 0,6$  a performance de recuperação da rede é equivalente à rede aleatória com  $\omega = 1,0$ , economizando o custo de ligações.

A topologia Cruz no painel do meio piora sua performance para valores  $\omega < 1,0$  comparando com a topologia X. A topologia de Anel tem uma performance similar à topologia de Cruz.

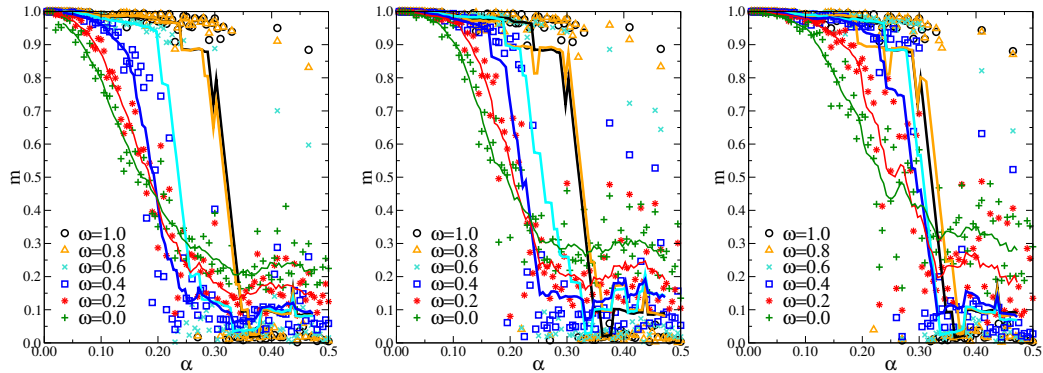


Figura 3.4: Capacidade de recuperação de impressões digitais sem ruído esqueletizadas para diferentes valores do parâmetro de aleatoriedade  $\omega$ , com  $N = 89420$ ,  $c = 200$ , condição inicial  $m^0 = 1,0$  e atividade  $a = 0,08$ . Anel(esquerda), Cruz(meio), X(direita). Sendo  $m^*$  representada com símbolos e as linhas sólidas são correspondentes a uma média suave para os diferentes valores de  $\omega$ .

### 3.3.3 IDs com atividade uniforme

A Fig.3.5 mostra a performance de recuperação da rede para padrões de atividade uniforme, ou seja, padrões com  $a \sim 0,5$ . Mais precisamente, no experimento utilizamos  $a = 0,48$ .

Neste caso, não existe uma tendência na atividade, em contraste aos padrões estudados anteriormente. Não é necessária estratégia para o limiar na recuperação de padrões de atividade uniforme, portanto  $\theta_0 = 0,0$  [53].

Novamente as três topologias são estudadas. Como esperado, a performance de recuperação é menor quando comparada com padrões de atividade tendenciosas ( $a = 0,23$ ) e padrões mais escassos ( $a = 0,08$ ). A relação de carga crítica  $\alpha_c \sim 0,07$  é menor do que o  $\alpha_c$  obtido nos casos anteriores. É importante salientar que para baixos valores de  $\omega$ , a rede recupera melhor padrões de atividade uniforme.

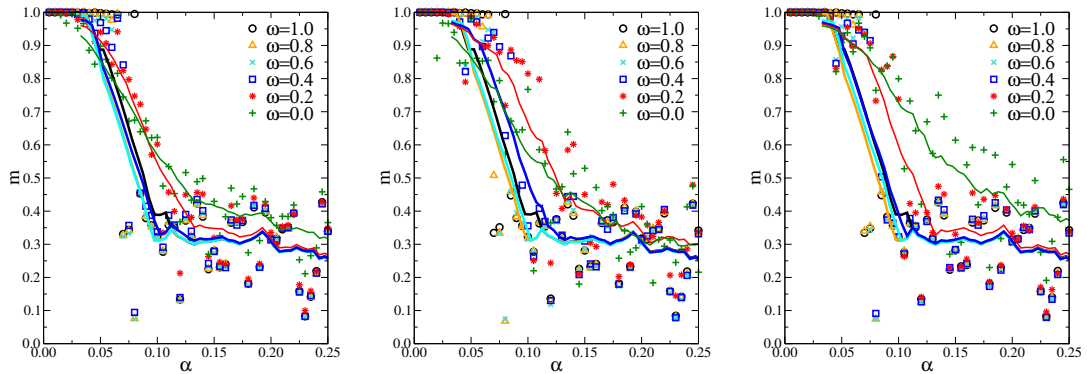


Figura 3.5: Capacidade de recuperação de impressões digitais dilatadas ruidosas uniformes para diferentes valores do parâmetro de aleatoriedade  $\omega$ , com  $N = 89420$ ,  $c = 200$ ,  $a = 0,48$ ,  $\theta_0 = 0,0$  e condição inicial  $m^0 = 1,0$ . Anel(esquerda), Cruz(meio), X(direita). Sendo  $m^*$  representada com símbolos e as linhas sólidas são correspondentes a uma média suave para os diferentes valores de  $\omega$ .

### 3.3.4 Capacidade de Informação

Dos resultados das sessões anteriores, é possível supor que armazenar padrões com baixa atividade possui uma grande vantagem em relação a padrões uniformes. Entretanto, a informação de um único padrão com baixa atividade é menor do que a informação de um padrão uniforme, portanto é necessário calcular a capacidade de informação, que leva em conta tanto a capacidade e a recuperação em uma única função escalar. Assim, é possível estudar como a capacidade de informação  $i$  se comporta com a atividade  $a$ .

A informação foi calculada em uma topologia de Anel usando  $i = \alpha_0 M(m, \alpha_0)$ , onde  $\alpha_0$  é a relação de carga ótima, e  $M(m, \alpha_0)$  é a informação mútua entre o neurônio e o padrão recuperado com valor máximo de  $i$ .

A informação mútua  $M(m, \alpha_0)$  é uma ferramenta útil na teoria da informação e corresponde a quantidade de informação que uma variável contém em relação a outra e que é recuperada [72]. Em geral, calcula-se,  $M$  por meio de  $M((m, \alpha_0) = S(\sigma) - S(\sigma|\xi)$ , onde  $S$  é a entropia de Shannon. Porém, como queremos a informação mútua com o valor máximo de  $i$ , a entropia  $S(m|\alpha_0)$  torna-se zero, e ficamos apenas com  $S(m)$ . Utilizamos

Atividade	Carga	Entropia	Informação
0,08	0,34	0,402	0,1367
0,23	0,11	0,778	0,0856
0,48	0,045	0,999	0,0449

Tabela 3.1: Atividade de padrões vs. Capacidade de informação.

o  $S(m)$  obtido em [73]

$$S(m) = -a \ln(a) - (1 - a) \ln(1 - a). \quad (3.10)$$

Como a transição é abrupta, é estimado que a entropia  $S$  de um padrão recuperado (onde  $m \approx 1$ ), possui erros de 5%, isto é,  $m(\alpha_c) = 0,95$ , para o qual definimos uma relação de carga crítica  $\alpha_c$ . Então,  $i = \alpha_c S(m)$ .

Na Tabela 3.1, mostramos o valor da carga crítica  $\alpha_c$ , a entropia  $S$  e a capacidade de informação  $i$  para cada valor da atividade  $a$ .

Apesar do aumento da capacidade de informação com a diminuição da atividade não é isso o que ocorre para a capacidade de carga, mesmo assim é observada uma melhoria moderada em  $i$  para baixas atividades. O preço a pagar é a resolução de cada padrão.

### 3.3.5 Métrica de Boltzmann

Fizemos mais um experimento, para ver se o comportamento observado nas Figs.3.3, 3.4, 3.5 estão de acordo. Foi utilizada uma distribuição de probabilidade de conexões representada pela Fig.3.6, que nomeamos de Topologia de Boltzmann, uma vez que essa distribuição pode ser escrita da seguinte forma:

$$p_{c_{ij}} = \frac{1}{Z} \exp\left(-\frac{D_{ij}}{T}\right), \quad (3.11)$$

onde o parâmetro de temperatura  $T$  está intimamente conectado ao parâmetro  $\omega$  dos

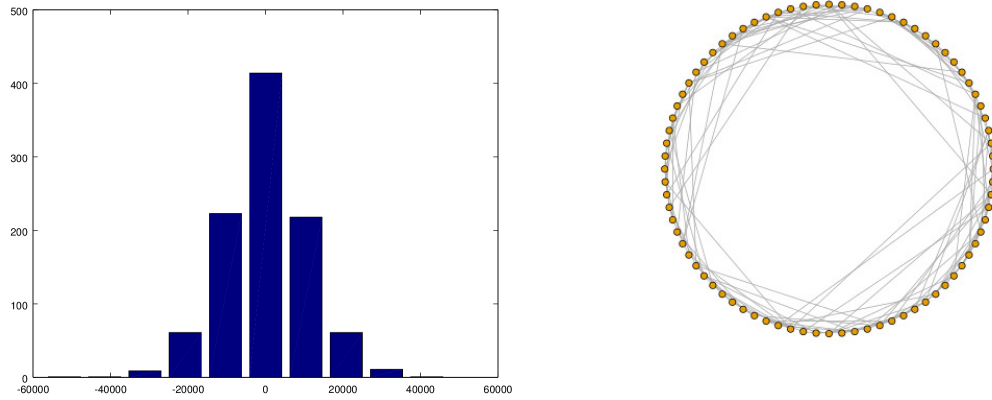


Figura 3.6: Grafo de Boltzmann. Esquerda: Histograma da distância das conexões. Direita: Um grafo típico para  $N = 81$ ,  $c = 4$ ,  $T = 10$

outros experimentos e representa o nível de ruído na vizinhança. Para  $T$  pequeno, a maioria das conexões ligam neurônios locais, enquanto que para  $T$  grande, conexões de longo alcance dominam. A distância em um Anel fechado é  $D_{ij} = (i - j) \bmod \frac{N}{2}$ , e a normalização constante é de  $Z = T \sinh\left(\frac{N}{2T}\right)$ .

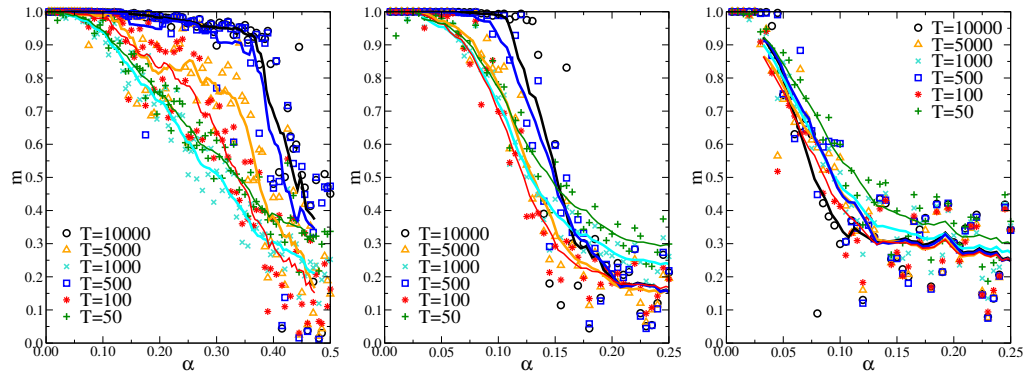


Figura 3.7: Capacidade de recuperação de IDs para uma topologia de Boltzmann e diferentes valores do ruído  $T$ , com  $N = 89420$ ,  $c = 200$ ,  $m^0 = 1,0$ . Esquerda:  $a = 0,08$ . Meio:  $a = 0,23$ , Direita:  $a = 0,48$ . Sendo  $m^*$  representada com símbolos e as linhas sólidas são correspondentes a uma média suave para os diferentes valores de  $T$ .

Calculamos a sobreposição como função da relação de carga, para a métrica de Boltzmann, e a representamos na Fig.3.7, com os mesmos valores da atividade de padrões

das Figs.3.3, 3.4, 3.5. Podemos comparar as curvas e verificar que a forma da topologia de Boltzmann é similar às topologias estudadas anteriormente ao trocarmos a razão de aleatoriedade  $\omega$  pelo ruído  $T$ .

### 3.4 Teoria

Na seção anterior fizemos um tratamento experimental de uma MANN para armazenar e recuperar padrões de IDs. Vários tipos de atividade foram consideradas, indo de código esparsos até padrões uniformes.

Os experimentos demonstraram que em uma MANN para uma dada atividade típica, quanto maior o  $\omega$ , melhor é a performance da rede. Contudo, é difícil de assegurar que isto seja verdade para  $\omega \sim 1$ , pois na Fig.3.4 existe um comportamento não monótono, isto é, existe um valor ótimo de  $\omega_0 \sim 0,8$  no qual a capacidade de carga é máxima. Tal  $\omega_0$  parece depender da atividade média e dos detalhes particulares da topologia considerada. Ao aumentar a atividade típica de padrões de IDs, a capacidade de carga  $\alpha_c$  diminui. Além disso, o espalhamento entre as curvas da sobreposição contra a relação de carga,  $m \times \alpha$ , aumenta da direção de  $\omega = 0$  para  $\omega = 1$ , quando a atividade aumenta. Também é observada uma remanência depois do valor crítico do ruído.

Tentaremos explicar esses comportamentos e suas principais características na abordagem teórica. Sendo capaz de explicar um sistema de IDs utilizando a média de uma topologia métrica (dependente apenas de  $\omega$  [72]), a média da atividade de padrões, e um novo parâmetro que foi introduzido para medir o caráter local de cada ID.

Este último, chamamos de  $\psi$ , ele quantifica a correlação espacial dos pixels da ID. Para um dado padrão,  $\{\xi_i \equiv \xi_i^\mu\}$ ,  $\psi$  é estatisticamente definido de acordo com

$$\xi_i \doteq \psi \xi_i^r + (1 - \psi) \xi_i^\ell, \quad (3.12)$$



onde  $\xi_i^r$  corresponde a natureza aleatória dos padrões, e  $\xi_i^\ell$  representa a distribuição local dos pixels nos padrões.

É interessante notar que a Eq.(3.12) define os padrões de ID como a soma global de uma parte aleatória e uma parte local, no sentido de distribuição de probabilidade. Mais precisamente, os padrões aleatórios,  $\{\xi_i^r\}_i^N$  são um conjunto de variáveis aleatórias identicamente distribuídas e independentes (iidrv),  $\langle \xi_i^r \xi_j^r \rangle \equiv \delta_{ij}$ . Os padrões locais,  $\{\xi_i^\ell\}_i^N$  são correlacionados localmente,  $\langle \xi_i^\ell \xi_j^\ell \rangle \equiv 1$ , se  $\{j \in n_i\}$ , onde  $n_i$  é a vizinhança do iésimo pixel, caso contrário  $\langle \xi_i^\ell \xi_j^\ell \rangle \equiv 0$ . No caso de pixels independentes,  $\psi = 1$ . Para pixels completamente localizados,  $\psi = 0$ . Para correlações intermediárias temos  $0 < \psi < 1$ . O valor de  $\psi$  de uma ID não é medido diretamente. Ele será determinado por um ajuste entre teoria e experimentos.

O experimento real que estamos tentando modelar não possui tamanho infinito, mas utiliza valores grandes e finitos de uma rede de tamanho,  $N \sim 89420$ , e uma conectividade,  $c = 200$ , assim como  $\frac{N}{c} \sim 447$ . Portanto a teoria é parametrizada pelos seguintes parâmetros:  $\alpha \equiv \frac{P}{c}$ ,  $\beta_c \equiv \frac{1}{c}$ ,  $\gamma \equiv \frac{c}{N}$ . Temos uma boa aproximação para o limite de pequenos  $\alpha$ ,  $\beta_c$ ,  $\gamma$ , o que pode ser facilmente generalizado para outros grupos de padrões estruturados com características similares.

### 3.4.1 Divisão do campo em sinal e ruído

A performance da recuperação de padrões em uma MANN pode ser medida por meio de duas relações: a sobreposição  $m$ , dada pela Eq.(3.2), que corresponde a sobreposição dos padrões armazenados em relação aos estados estacionários da dinâmica neural, e a capacidade  $\alpha$ .

Para recuperar padrões, é necessária uma relação de recorrência para a sobreposição,  $m^{t+1} = f(m^t)$ . Relação esta que deve conter a regra de atualização dos neurônios, obtida na Eq.(3.5). Porém, para obter esta dependência é necessário levar em consideração o

caráter local da matriz topológica e dos padrões.

Ao fazer  $g(h_i^t - \theta_i^t, q_i^t) = \text{sign}(h_i^t)$ , facilita a operação dos padrões binários  $\xi_i^\mu \in \pm 1$ . Podemos escrever a dinâmica para a sobreposição usando a Eq.(3.2) e obtemos

$$m^{t+1} = \langle \text{sign}(\hat{h}_i^t) \rangle, \quad (3.13)$$

$$\hat{h}_i^t \equiv \xi_i h_i^t.$$

Vamos considerar o campo pós-sináptico, Eq.(3.5), atuando no neurônio  $i$  em um dado tempo  $t$ , digamos  $\sigma_i \equiv \sigma_i^t$ . Substituindo a matriz de adjacência, Eq.(3.1), e os pesos de aprendizagem, Eq.(3.7) podemos reescrever o campo local como

$$\hat{h}_i = \frac{\xi_i}{c} \sum_{\mu} \xi_i^{\mu} \sum_j c_{ij} \xi_j^{\mu} \sigma_j. \quad (3.14)$$

Caso a rede esteja recuperando um dado padrão, digamos  $\xi_i^1 \equiv \xi_i$ , então é possível reescrever o campo em termos de um sinal  $m_i$  e do ruído de interferência  $\Omega_i$ ,

$$\hat{h}_i = m_i + \Omega_i, \quad (3.15)$$

sendo a parte de sinal

$$m_i \equiv \frac{1}{c} \sum_j c_{ij} \xi_j \sigma_j, \quad (3.16)$$

e a parte da interferência

$$\Omega_i \equiv \xi_i \sum_{\mu>1} \xi_i^\mu m_i^\mu, \quad (3.17)$$

$$m_i^\mu \equiv \frac{1}{c} \sum_j c_{ij} \xi_j^\mu \sigma_j.$$

O termo de sinal tem o papel de uma sobreposição local, dependendo da matriz de conectividade. Esta por sua vez pode ser escrita como a soma de termos locais e aleatórios [72]:

$$c_{ij} \doteq \omega c_{ij}^\rho + (1 - \omega) c_{ij}^\lambda, \quad (3.18)$$

onde  $c_{ij}^\rho$  são as conexões aleatórias, uniformemente distribuídas entre os acoplamentos dos neurônios  $i, j$ , e  $c_{ij}^\lambda$  são as conexões locais, que representam qualquer das métricas descritas na Fig.3.2. O alcance da rede é medido por meio de  $\omega$ .

A sobreposição dos sítios, Eq.(3.16) se torna

$$m_i = \omega m^\rho + (1 - \omega) m^\lambda, \quad (3.19)$$

onde a sobreposição com topologia aleatória é escrito desta forma,

$$m^\rho = \frac{1}{c} \sum_j c_{ij}^\rho \xi_j \sigma_j, \quad (3.20)$$

e a sobreposição com topologia local é,

$$m^\lambda = \frac{1}{c} \sum_j c_{ij}^\lambda \xi_j \sigma_j. \quad (3.21)$$

Uma vez que os padrões também se dividem em termos locais e aleatórios, Eq.(3.12),

cada termo nas Eqs.(3.20 e 3.21) tornam-se dois

$$m_i = \psi[\omega m_i^{r\rho} + (1 - \omega)m_i^{r\lambda}] + (1 - \psi)[\omega m_i^{\ell\rho} + (1 - \omega)m_i^{\ell\lambda}], \quad (3.22)$$

onde para os padrões aleatórios temos

$$m^{r\rho} = \frac{1}{c} \sum_j c_{ij}^\rho \xi_j^r \sigma_j, \quad (3.23)$$

$$m^{r\lambda} = \frac{1}{c} \sum_j c_{ij}^\lambda \xi_j^r \sigma_j, \quad (3.24)$$

e para os padrões locais,

$$m^{\ell\rho} = \frac{1}{c} \sum_j c_{ij}^\rho \xi_j^\ell \sigma_j, \quad (3.25)$$

$$m^{\ell\lambda} = \frac{1}{c} \sum_j c_{ij}^\lambda \xi_j^\ell \sigma_j. \quad (3.26)$$

### 3.4.2 Distribuição do campo

Foi considerada uma aproximação gaussiana para todos os termos nos campos locais:

$$\hat{h}_i \doteq \aleph[\langle \hat{h}_i \rangle, V_{\hat{h}}] \doteq \langle \hat{h}_i \rangle + Z_{\hat{h}} \sqrt{V_{\hat{h}}}, \quad (3.27)$$

sendo  $V_{\hat{h}} \equiv Var(\hat{h}_i)$  a variância de  $\hat{h}_i$  e  $Z_{\hat{h}} \doteq \aleph(0,1)$  uma variável gaussiana normalizada. Com a Eq.(3.27) e de acordo com a Eq.(3.15) podemos calcular a média do campo local ao calcularmos a média do termo de sinal, Eq.(3.22). Esta média converge, para o limite de  $c$  grande, para uma única variável,

$$\langle \hat{h}_i \rangle = m, \quad (3.28)$$

sendo  $m$  a sobreposição global, como definida na Eq.(3.13).

O campo local,  $\hat{h}_i$ , possui duas partes, como podemos ver na Eq.(3.15) e, portanto, sua variância também possui duas partes:

$$V_{\hat{h}} = V_{m_i} + V_{\Omega}. \quad (3.29)$$

O primeiro termo é a variância da sobreposição de sítio  $m_i$ , o segundo termo é a variância do ruído de interferência,  $\Omega$ . O termo de ruído possui uma forma proposta por Amit [53]

$$V_{\Omega} = \alpha r, \quad (3.30)$$

onde  $\alpha$  é a relação de carga e  $r$  é o termo de retroalimentação, que também é dividido em termos de retroalimentação para padrões aleatórios  $r_r$  e termos de retroalimentação para padrões locais  $r_\ell$ :

$$r = \psi r_r + (1 - \psi) r_\ell. \quad (3.31)$$

Cada um desses termos pode ser dividido em topologia aleatória e topologia métrica:

$$r = \psi[\omega r_{r\rho} + (1 - \omega)r_{r\lambda}] + (1 - \psi)[\omega r_{\ell\rho} + (1 - \omega)r_{\ell\lambda}]. \quad (3.32)$$

Os valores de retroalimentação são avaliados da seguinte maneira:

$$\begin{aligned} r_{r\rho} &= 1, \\ r_{r\lambda} &= 1 + b \cdot m^2, \\ r_{\ell\rho} &= 1 + d, \\ r_{\ell\lambda} &= 1 + b \cdot m^2. \end{aligned} \quad (3.33)$$

Os verdadeiros valores de cada  $r$  são funções complexas, as quais são recursivas em  $m$  e todos os tipos de retroalimentação  $r$ .

Vamos dar uma olhada no que cada  $r$  quer dizer. Começamos com  $r_{r\rho} = 1$ , pois a matriz topológica  $c_{ij}^\rho$  é extremamente diluída e os padrões aleatórios  $\xi^r$  são distribuídos independentemente. Portanto não existe realimentação presente neste caso [74]. A seguir,  $r_{\ell\rho} = 1 + d$ , precisa ser maior do que  $r_{r\rho}$  devido à dependência do sítio dos padrões. O fator  $r_{r\lambda} = r_{\ell\lambda} = 1 + b \cdot m^2$ , por causa da conectividade local. Mesmo em casos bastante diluídos, a retroalimentação é intensa, de forma similar a uma rede com todos os seus elementos conectados.

O que implica que existe uma dependência no valor absoluto de  $m$  que é simplificado do valor conhecido de topologias métricas [72]. Temos que  $r_{\ell\lambda}$  deve ser maior do que  $r_{\ell\rho}$  devido a alta correlação da topologia métrica, exceto para pequenas sobreposições quando alguma memória está presente na dinâmica neuronal, semelhante ao efeito de remanência no ferromagnetismo.

Neste trabalho, os fatores  $b$  e  $d$  são ajustados com os experimentos e demonstram uma dependência no grau de densidade  $\gamma$ . Porém, para termos o menor grau de liberdade no parâmetro ajustável, utilizamos uma função simples  $d(a,b)$ , que satisfaz o vínculo  $d \leq b$ , para todos os valores trabalhados da atividade. Ademais, procuramos por um valor empírico  $b(\gamma)$  para o  $\gamma = 200/89420$  fixo utilizado nos experimentos.

Encontramos que as funções:

$$d = 2ab, \tag{3.34}$$

$$b = 10,$$

se encaixam nos resultados experimentais da seção 3.3.

Agora, consideremos a variância da sobreposição de sítios. Uma vez que todos os

quatro tipos de sobreposição de sítio são independentes, podemos escrever a variância  $V_{m_i}$  como uma soma

$$V_{m_i} = \psi[\omega V_{r\rho} + (1 - \omega)V_{r\lambda}] + (1 - \psi)[\omega V_{\ell\rho} + (1 - \omega)V_{\ell\lambda}], \quad (3.35)$$

onde cada uma das quatro variâncias são calculadas das Eqs.(3.23,3.24,3.25,3.26) e obtemos

$$V_{r\rho} = V_{r\lambda} = \beta_c, \quad (3.36)$$

$$V_{\ell\rho} = V_{\ell\lambda} = \beta_c + o - m^2,$$

até a primeira ordem em  $O(\beta_c \equiv \frac{1}{c})$ .

O termo  $o \equiv o_\ell$  que aparece na equação anterior corresponde a função de correlação neuronal  $o \equiv \langle \sigma_i \sigma_j \rangle_\ell$  para os padrões locais, isto é,  $\psi = 0$ . Esta é uma função da atividade  $a$  e da sobreposição  $m$ , que deve estar próxima de  $o = 1$  para uma atividade de  $a = 1/2$  e por volta de  $o = m^2$  para código esparsos,  $a \rightarrow 0$ . Escolhemos uma interpolação linear capaz de capturar este comportamento:

$$o \equiv \langle \sigma_i \sigma_j \rangle_\ell = 2a(1 - m^2) + m^2. \quad (3.37)$$

O vínculo  $0 \leq o \leq 1$  é satisfeito. Para padrões aleatórios a função correlação entre estados neuronais é  $o_r = o \equiv \langle \sigma_i \sigma_j \rangle_r = 0$ .

Obtemos então

$$V_{\hat{h}} = \psi\beta_c + (1 - \psi)(\beta_c + o - m^2) + \alpha r, \quad (3.38)$$

e também

$$r = \psi[\omega + (1 - \omega)(1 + bm^2)] + (1 - \psi)[\omega(1 + 2ab) + (1 - \omega)(1 + bm^2)]. \quad (3.39)$$

Na Eq.(3.13), temos uma média discreta sobre os estados ( $N^{-1} \sum_i$ ), como consideramos uma aproximação gaussiana para todos os termos do campo local, também podemos utilizar uma distribuição gaussiana e a média passa a ser contínua, temos:

$$m = \int_{-\infty}^{+\infty} \text{sign}(\hat{h}) d\phi_{(Z_{\hat{h}})}, \quad (3.40)$$

sendo  $d\phi_{Z_{\hat{h}}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z_{\hat{h}}^2}{2}} dZ_{\hat{h}}$ , aplicando a Eq.(3.27) obtemos:

$$m = \int_{-\infty}^{+\infty} \text{sign} \left( \langle \hat{h} \rangle + Z_{\hat{h}} \sqrt{V_{\hat{h}}} \right) d\phi_{(Z_{\hat{h}})}. \quad (3.41)$$

Podemos resolver esta integral por partes fazendo:

$$dv = d\phi_{(Z_{\hat{h}})} \quad (3.42)$$

$$u = \text{sign} \left( \hat{h} + Z_{\hat{h}} \sqrt{V_{\hat{h}}} \right),$$

deste modo, escrevemos:

$$m = \phi_{Z_{\hat{h}}} \text{sign} \left( \langle \hat{h} \rangle + Z_{\hat{h}} \sqrt{V_{\hat{h}}} \right) \Big|_0^1 - \int \phi_{(Z_{\hat{h}})} d \text{sign} \left( \langle \hat{h} \rangle + Z_{\hat{h}} \sqrt{V_{\hat{h}}} \right), \quad (3.43)$$

isto leva a:

$$m = 1 - 2 \int_{-\frac{\langle \hat{h} \rangle}{\sqrt{V_{\hat{h}}}}^{+\infty} \phi_{(Z_{\hat{h}})} d(Z_{\hat{h}} \sqrt{V_{\hat{h}}}) \delta_{\langle \hat{h} \rangle + Z_{\hat{h}} \sqrt{V_{\hat{h}}}}, \quad (3.44)$$



usando a identidade  $\delta(xy) = \frac{1}{x}\delta(y)$ , a equação fica assim:

$$m = 1 - 2\phi\left(-\frac{\langle\hat{h}\rangle}{\sqrt{V_h}}\right), \quad (3.45)$$

podemos escrever isso como:

$$m = \operatorname{erf}\left(\frac{\langle\hat{h}\rangle}{\sqrt{2V_h}}\right) \quad (3.46)$$

e usando a Eq.(3.28) chegamos em:

$$m = \operatorname{erf}\left(\frac{m}{\sqrt{2V_h}}\right). \quad (3.47)$$

Obtendo assim a relação de recorrência da sobreposição, dependendo apenas da média e da variância do campo local.

### 3.4.3 Resultados teóricos

Resolvendo as equações recursivas Eqs.(3.37, 3.38, 3.39, 3.47) podemos então comparar as previsões teóricas com os experimentos de IDs. O único parâmetro ajustável é o  $\psi$ , que estima a localidade espacial dos padrões nas amostras aprendidas. As amostras consideradas possuem padrões de tamanho  $N = 89420$ , e a rede utiliza a conectividade média  $c = 200$  por neurônio. De acordo com o pré-processamento, treinamos a rede com diferentes valores de atividade de padrão:  $a = 0,08; 0,23; 0,48$ , que vai de código esparsos ( $a \rightarrow 0$ ), até padrões uniformes ( $a \sim \frac{1}{2}$ ).

Considerando inicialmente uma atividade média  $a = 0,23$ , procuramos por um valor de  $\psi$  que represente de forma ótima as curvas da sobreposição  $m$  pela relação de carga  $\alpha$  na Fig.3.3, para os diversos valores de  $\omega$ .

Alguns valores de  $\psi$  são retratados na Fig. 3.8. Quando a relação de carga,  $\alpha$ , é baixa

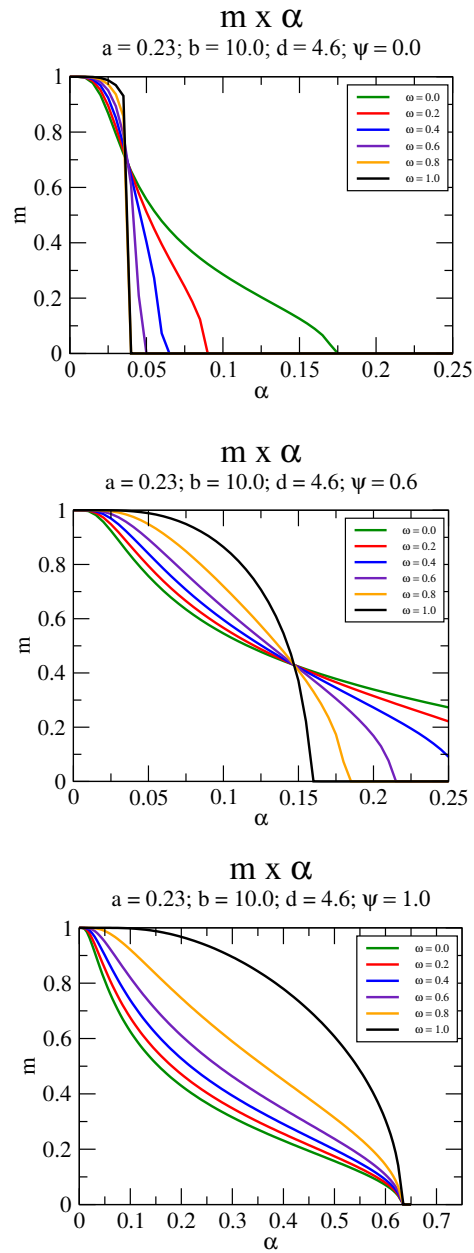


Figura 3.8: Sobreposição em função da razão de carga para diferentes valores do parâmetro de aleatoriedade topológica  $\omega$ , com  $c = 200$ , variando o parâmetro de aleatoriedade dos padrões  $\psi$ . Topo:  $\psi = 0,0$ ; Centro:  $\psi = 0,6$ ; Inferior:  $\psi = 1,0$ .

as curvas possuem uma sobreposição maior para conexões de longo alcance ( $\omega \sim 1$ ) do que para curto alcance ( $\omega \sim 0$ ), caso  $\alpha$  seja alto, verificamos o comportamento contrário.

Percebe-se que  $\psi = 0,6$  é o que melhor descreve os experimentos na Fig.3.3. As transições são sempre contínuas, mas seu declive é mais acentuado para longo alcance do que para curto alcance. O cruzamento entre as curvas de longo e curto alcance acontece em cerca de  $m \sim 0,45$  e  $a \sim 0,15$ .

Podemos compreender estes comportamentos por meio das equações recursivas. O termo de sinal na Eq.(3.47) não depende nem de  $\omega$  nem de  $\psi$ . O ruído do campo local na Eq.(3.38) possui um termo proporcional a  $(1 - \psi)$  que decresce com  $m$ , e um termo de retroalimentação que depende tanto do  $\psi$  quanto do  $\omega$ .

Como podemos ver na Eq.(3.39),  $r$  possui um ruído maior quando  $\omega \sim 1$  do que para  $\omega \sim 0$  se  $m$  for pequeno, o inverso ocorre para  $m$  grande. Portanto existe um cruzamento de  $m$  na direção oposta de  $r$ .

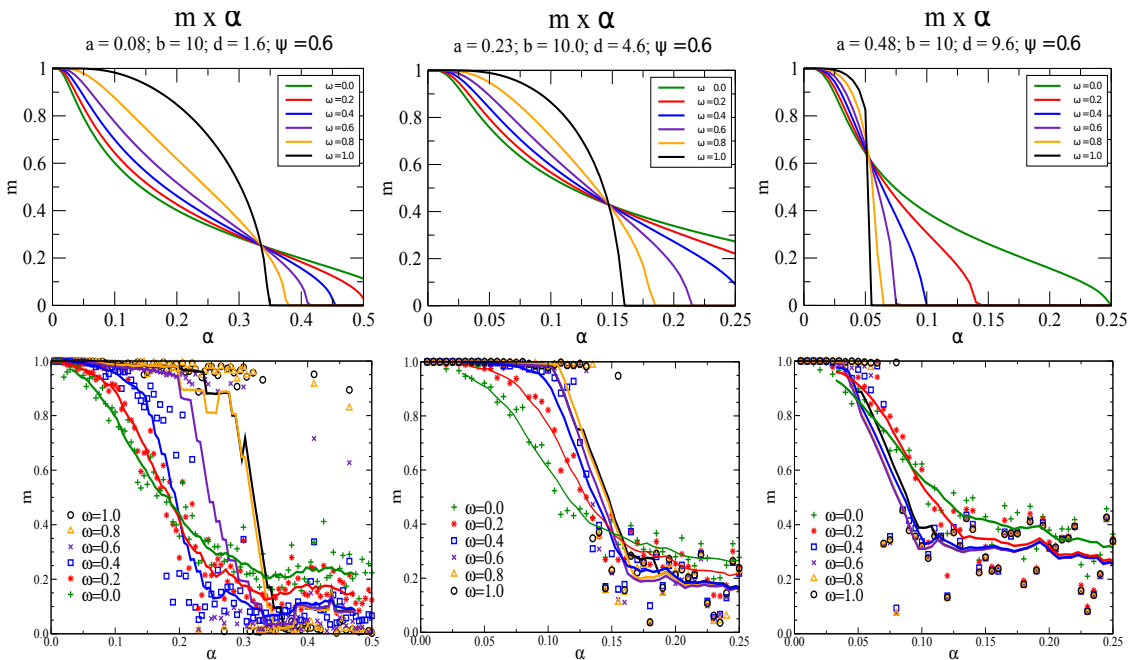


Figura 3.9: Sobreposição em função da razão de carga para diferentes valores do parâmetro de aleatoriedade topológica  $\omega$  e diferentes atividade:  $a = 0,08$  (esquerda),  $a = 0,23$  (centro) e  $a = 0,48$  (direita). Painéis superiores são os resultados teóricos com  $c = 200$ ,  $\psi = 0,6$ . Painéis inferiores, padrões de recuperação de impressões digitais sem ruído  $m^0 = 1,0$ . Topologia de anel com  $N = 89420$ ,  $c = 200$ ,  $\theta_0 = 1,49$  (esquerda),  $\theta_0 = 0,66$  (meio),  $\theta_0 = 0,0$  (direita).

Mantivemos constante o valor do parâmetro de localidade  $\psi = 0,6$  para todos os valores de atividade de padrões  $a$ , e calculamos as curvas  $m \times \alpha$ . O gráfico de  $m(\alpha)$  ao variar o alcance das ligações  $\omega$  é desenhado na Fig.3.9, para três valores de atividade  $a$ .

No painel esquerdo é apresentado o resultado teórico para  $a = 0,08$ , que concorda com os experimentos da Fig.3.4. A sobreposição para as topologias  $\omega \sim 1$  é maior do que para  $\omega \sim 0$ , para quase todos os intervalos de atividade, exceto quando  $a \geq 0,35$ . Neste caso, a sobreposição para  $\omega < 1$  possui um efeito de remanência. O cruzamento entre essas curvas acontece por volta de  $m \sim 0,25$ . Também temos uma comparação quantitativa para todos os valores da relação de carga. Por exemplo  $\alpha = 0,15$ , quando  $m(\omega = 1) \sim 0,95$  e  $m(\omega = 0) \sim 0,5$ , mostrando um espalhamento "gordo" entre as curvas. Para  $\omega = 1,0$  a relação de carga crítica é  $\alpha \simeq 0,64$ , que corresponde com os resultados esperados na literatura [25]

No painel central, apresentamos as curvas teóricas para  $a = 0,23$ , que são comparadas com os experimentos da Fig.3.3. É interessante notar que o espalhamento entre as curvas para  $\omega = 1,0$  e  $\omega = 0$  decresce enquanto a atividade  $a$  aumenta.

No painel da direita, temos os resultados teóricos que correspondem à Fig.3.5. O espalhamento praticamente desaparece para  $a = 0,48$ . A única diferença entre os experimentos e a teoria é uma cauda em  $m(\alpha)$  para  $\omega \sim 1$ , ou seja, neste caso não foi prevista a remanência verificada nos experimentos. Essas caudas são devido aos efeitos de tamanho finito ou aos ruídos reais que não são capturados pela teoria.

Estimamos que a diferença entre a teoria para  $c, N \rightarrow \infty$  e os experimentos para  $c$  e  $N$  finitos por meio do desvio padrão do ruído, de acordo com os termos negligenciados na Eq.3.35:  $\delta \approx O\left(\frac{1}{\sqrt{c}}\right)$ . Para  $c = 200$ , temos  $\delta(m) \approx 0,07$  nas regiões onde a sobreposição é pequena.

Finalmente, calculamos a correlação espacial para o conjunto de dados de IDs. Essas

Atividade	0,0844	0,1605	0,2257	0,3012	0,4817
Correlação	0,2012	0,5401	0,6400	0,7034	0,74034

Tabela 3.2: Atividade de padrões vs. Correlação espacial.

IDs foram estimadas ao representarmos Eq.3.12 assim

$$\psi = \langle \xi_i \xi_j \rangle_r \doteq \frac{1}{PN} \sum_{i,\mu} \xi_i^\mu \xi_{i+1}^\mu. \quad (3.48)$$

Os valores são dados pela Tabela 3.2 e por meio dela é possível notar que para esse conjunto específico de estruturas de padrões quanto mais esparsa o código é, menor é a correlação.

É interessante observar que o valor obtido por um ajuste simples de  $\psi$  para  $a = 0,23$  da Fig.3.8, para os experimentos, temos que  $\psi \approx 0,6$ , que é bem próximo do valor experimental de  $\psi = 0,64$ . Ao invés de usarmos um valor fixo, recalculamos as curvas teóricas com os  $\psi$  experimentais. Tanto o caso de  $a = 0,23$  quanto o caso de  $a = 0,48$  possuem  $\psi$  semelhante ao valor ajustado, mas o caso em que  $a = 0,08$  possui um  $\psi = 0,20$  que é bastante diferente do fixado, mas ajusta a curva de forma mais satisfatória. A Fig.3.10 ilustra essa situação.

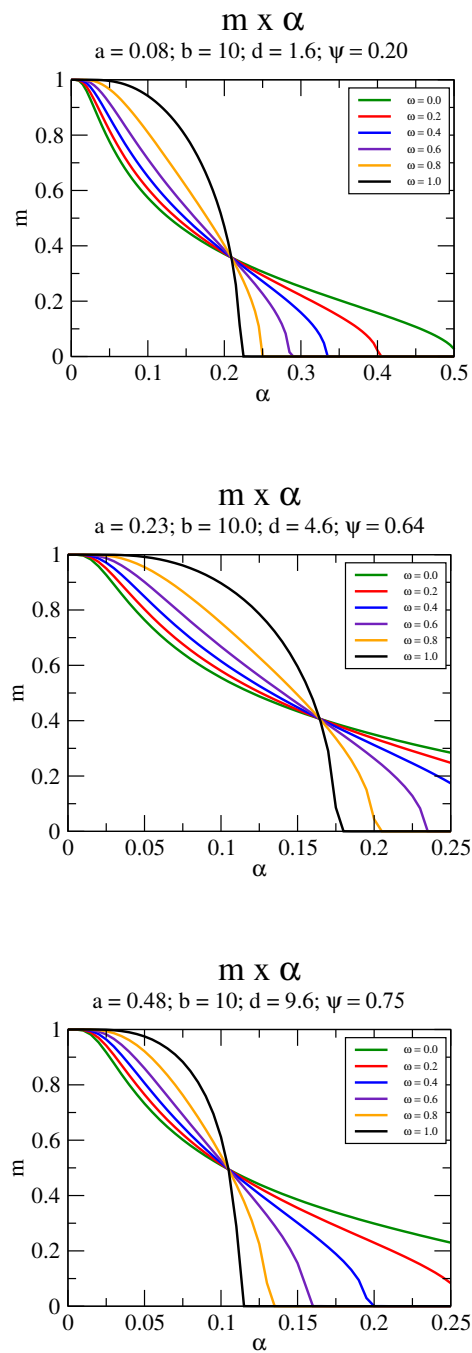


Figura 3.10: Curvas teóricas da sobreposição em função da razão de carga para diferentes valores do parâmetro de aleatoriedade topológica  $\omega$  e diferentes atividade:  $a = 0,08$  (topo),  $a = 0,23$  (centro) e  $a = 0,48$  (inferior), com suas respectivas correlações espaciais:  $\psi = 0,20$ ,  $\psi = 0,64$ ,  $\psi = 0,75$ .

## Capítulo 4

# Conclusões e perspectivas

### 4.1 Conclusões

Nesta tese investigamos duas redes complexas. Uma rede complexa de conectividade finita, sem estrutura. Onde foi estudada a transição  $F - P$ . E uma rede complexa estruturada. Onde estudamos o armazenamento e recuperação de padrões complexos.

No primeiro caso, o objetivo principal foi investigar a natureza da transição  $F - P$  para um Modelo de Ising com um campo aleatório em função da conectividade finita  $c$  que segue uma distribuição de probabilidade de um grafo aleatório com uma distribuição de campo aleatório tanto bimodal quanto gaussiana.

Os resultados mostram que a natureza da transição de fase depende da distribuição de campos aleatórios. Para uma distribuição gaussiana a transição é sempre de segunda ordem. Para uma distribuição bimodal, um ponto tricrítico aparece. Neste caso, sua localização é fortemente dependente do  $c$ . Na realidade, a temperatura do ponto tricrítico diminui quando  $c$  diminui até alcançar um valor mínimo da conectividade  $c_{min}$ . Em suma, os resultados afirmam que acima de um certo valor da conectividade  $c_{min}$  existe o surgimento de um ponto tricrítico para uma distribuição bimodal. Isto confirma as

conclusões de Aharony [34] utilizando conectividade infinita. Ainda mais importante, temos que para pequenos valores de conectividade ( $c < c_{min}$ ) a transição  $F - P$  é sempre contínua, independentemente do tipo de distribuição de campo. Assim, os resultados para  $c$  grande com uma distribuição bimodal do campo aleatório confirmam a existência de um ponto tricrítico, que foi relatado anteriormente por uma rede de Bethe [39, 75]. As diferenças no valor do  $c_{min}$  podem ser creditadas a diferenças nas arquiteturas de curto e longo alcance. Como mencionado anteriormente, Fytas e Martin-Mayor [38], apoiados por simulações numéricas, encontraram que a transição  $F - P$  para um RFIM com interações de primeiros vizinhos é de segunda ordem. O que entra em desacordo com os resultados obtidos, já que a transição de segunda ordem foi encontrada apenas para  $c < 5,0$ . Possivelmente, o número de coordenação efetivo em uma rede aleatória será diferente do de uma rede  $3D$  de primeiros vizinhos. A concordância entre os resultados analíticos e simulados pelo método de Monte Carlo validam os resultados obtidos. Os resultados deste trabalho foram publicados na revista internacional *Physica A* [76]

Para o segundo caso procuramos recuperar padrões em uma rede neural atratora com métrica de pequeno mundo. Neste caso foi utilizada a ideia de que padrões neurais no cérebro, assim como padrões no mundo real são estruturados (espacial e/ou temporalmente). Portanto foi estudado o processamento de informação por meio de padrões estruturados, de uma perspectiva neurobiológica. As habilidades da rede métrica atratora e sua performance foi testada por padrões estruturados com diferentes níveis de atividade.

Impressões digitais foram os padrões escolhidos, em razão de apresentarem estruturas espaciais únicas. Uma teoria foi desenvolvida como função de cinco parâmetros: a relação de carga, a conectividade finita, o grau de densidade da rede, a razão de aleatoriedade, e a correlação do padrão espacial. Este último sendo conectado com uma nova variável neural dinâmica, a correlação espacial neural. As simulações concordaram com a teoria para o parâmetro de localidade  $\psi = 0,6$ , que estima a estrutura espacial local dos padrões nas



impressões digitais armazenadas. Quer seja para atividades de padrões esparsa ( $a \ll 1$ ) ou para atividades de padrões uniformes ( $a \approx 1/2$ ), a teoria está de acordo com os resultados experimentais. Os resultados deste trabalho foram publicados na revista internacional *Physica A* [77].

## 4.2 Perspectivas

O próximo passo natural para o modelo de Ising com campo aleatório em um grafo aleatório é a adição de desordem nas interações. Neste caso a função de troca  $J$  deixa de ser constante e será necessário realizar uma média configuracional sobre esta variável também. Outra modificação é uma distribuição híbrida bimodal e gaussiana entre os campos aleatórios.

Para a rede neural atratora métrica com padrões estruturados será obtida de forma mais rigorosa uma função de correlação entre os neurônios locais. Será adicionada a métrica na teoria. Este modelo pode ser utilizado para identificação de padrões em redes sociais.

# Referências

- [1] A.T. Winfree. *The geometry of biological time*. Springer, 1980.
- [2] Y. Kuramoto. Chemical oscillations, waves, and turbulence. *Journal of Applied Mathematics and Mechanics*, 66(7):296–296, 1986.
- [3] Kurt Wiesenfeld, Pere Colet, and Steven H. Strogatz. Frequency locking in josephson arrays: Connection with the kuramoto model. *Phys. Rev. E*, 57:1563–1569, Feb 1998.
- [4] P. A. Gillespie. Fractals and chaos in geology and geophysics. *Geological Journal*, 28(2):214–215, 1993.
- [5] S.H. Strogatz. Exploring complex networks. *Nature*, 410:268–276, 2001.
- [6] S. Kauffman. *At home in the universe*. Oxford university press, 1996.
- [7] P. Erdős and A. Rényi. On the evolution of random graphs. In *Publication of the mathematical institute of hungarian academy of sciences*, pages 17–61, 1960.
- [8] D.J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.
- [9] D.J. Watts. *Small worlds*. Princeton university press, 1999.
- [10] T. Nattermann. Theory of the Random Field Ising Model. *Spin Glasses And Random Fields.*, 12:277–298, December 1997.

- 
- [11] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, 35:1792–1796, Dec 1975.
- [12] A.I. Larkin. Effect of inhomogeneities on the structure of the mixed state of superconductors. *Sov. Phys. JETP*, 31(4):784, 1970.
- [13] D.P. Belanger and A.P. Young. The random field ising model. *Journal of Magnetism and Magnetic Materials*, 100(1):272 – 291, 1991.
- [14] Pierre Le Doussal. Novel phases of vortices in superconductors. *International Journal of Modern Physics B*, 24(20n21):3855–3914, 2010.
- [15] Martin Niss. History of the lenz-ising model 1920–1950: From ferromagnetic to cooperative phenomena. *Archive for History of Exact Sciences*, 59(3):267–318, 2005.
- [16] Martin Niss. History of the lenz–ising model 1950–1965: from irrelevance to relevance. *Archive for History of Exact Sciences*, 63(3):243–287, 2008.
- [17] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.
- [18] R. Rojas. *Neural Networks*. Springer, 1996.
- [19] Donald Hebb. The organization of behavior. In *Neurocomputing*. MIT Press., 1949.
- [20] G. Tkacik, E. Schneidman, M. J. I. Berry, and W. Bialek. Ising models for networks of real neurons. *eprint arXiv:q-bio/0611072*, November 2006.
- [21] C. Li and G. Chen. Stability of a neural network model with small-world connections. 68(5):052901, November 2003.
- [22] Patrick N. McGraw and Michael Menzinger. Topology and computational performance of attractor neural networks. *Phys. Rev. E*, 68:047102, Oct 2003.

- [23] David R. C. Dominguez, Kostadin Koroutchev, Eduardo Serrano, and Francisco de Borja Rodríguez Ortiz. Information and topology in attractor neural networks. *Neural Computation*, 19(4):956–973, 2007.
- [24] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, Jan 2002.
- [25] John Hertz, Richard G. Palmer, and Anders S. Krogh. *Introduction to the Theory of Neural Computation*. Perseus Publishing, 1st edition, 1991.
- [26] Mesoscopic magnetic states in metallic alloys with strong electronic correlations: A percolative scenario for  $\text{Cu}_{1-x}\text{Ni}_x$ , author = Marcano, N. and Gómez Sal, J. C. and Espeso, J. I. and De Teresa, J. M. and Algarabel, P. A. and Paulsen, C. and Iglesias, J. R., journal = *Phys. Rev. Lett.*, volume = 98, issue = 16, pages = 166406, numpages = 4, year = 2007, month = Apr, publisher = American Physical Society, doi = 10.1103/PhysRevLett.98.166406, url = <http://link.aps.org/doi/10.1103/PhysRevLett.98.166406>.
- [27] Wolfgang Kleemann. Random fields in relaxor ferroelectrics - a jubilee review. *Journal of Advanced Dielectrics*, 02(02):1241001, 2012.
- [28] Shmuel Fishman and Amnon Aharony. Random field effects in disordered anisotropic antiferromagnets. *Journal of Physics C: Solid State Physics*, 12(18):L729, 1979.
- [29] H. Yoshizawa, R. A. Cowley, G. Shirane, R. J. Birgeneau, H. J. Guggenheim, and H. Ikeda. Random-field effects in two- and three-dimensional ising antiferromagnets. *Phys. Rev. Lett.*, 48:438–441, Feb 1982.
- [30] DP Belanger, AR King, and V Jaccarino. Random-field critical behavior of a d=3 ising system: Neutron scattering studies of  $\text{Fe}_{0.6}\text{Zn}_{0.4}\text{F}_2$ . *Physical Review B*, 31(7):4538, 1985.

- 
- [31] A Peter Young. *Spin glasses and random fields*, volume 12. World Scientific, 1997.
- [32] Yoseph Imry and Shang-keng Ma. Random-field instability of the ordered state of continuous symmetry. *Phys. Rev. Lett.*, 35:1399–1401, Nov 1975.
- [33] Vik S Dotsenko. On the nature of the phase transition in the three-dimensional random field ising model. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(09):P09005, 2007.
- [34] Amnon Aharony. Tricritical points in systems with random fields. *Phys. Rev. B*, 18:3318–3327, Oct 1978.
- [35] T. Schneider and E. Pytte. Random-field instability of the ferromagnetic state. *Phys. Rev. B*, 15:1519–1522, Feb 1977.
- [36] David Andelman. First- and second-order phase transitions with random fields at low temperatures. *Phys. Rev. B*, 27:3079–3080, Mar 1983.
- [37] Serge Galam and Joseph L. Birman. Random-field distributions and tricritical points. *Phys. Rev. B*, 28:5322–5322, Nov 1983.
- [38] Nikolaos G. Fytas and Víctor Martín-Mayor. Universality in the three-dimensional random-field ising model. *Phys. Rev. Lett.*, 110:227201, May 2013.
- [39] Yoshiki Matsuda, Hidetoshi Nishimori, Lenka Zdeborová, and Florent Krzakala. Random-field  $p$ -spin-glass model on regular random graphs. *Journal of Physics A: Mathematical and Theoretical*, 44(18):185002, 2011.
- [40] R. Bruinsma. Random-field ising model on a bethe lattice. *Phys. Rev. B*, 30:289–299, Jul 1984.
- [41] R.A. Barrio, G.G. Naumis, and C. Wang. Bethe lattice studies in glasses. *Journal of Non-Crystalline Solids*, 182(1-2):22–26, 1995. cited By 2.

- 
- [42] I. Kanter and H. Sompolinsky. Mean-field theory of spin-glasses with finite coordination number. *Phys. Rev. Lett.*, 58:164–167, Jan 1987.
- [43] L Viana and A J Bray. Phase diagrams for dilute spin glasses. *Journal of Physics C: Solid State Physics*, 18(15):3037, 1985.
- [44] M. Mézard and G. Parisi. Mean-field theory of randomly frustrated systems with finite connectivity. *EPL*, 3(10):1067–1074, 1987. cited By 35.
- [45] Rémi Monasson. Optimization problems and replica symmetry breaking in finite connectivity spin glasses. *Journal of Physics A: Mathematical and General*, 31(2):513, 1998.
- [46] Florent Krzakala, Federico Ricci-Tersenghi, and Lenka Zdeborová. Elusive spin-glass phase in the random field ising model. *Phys. Rev. Lett.*, 104:207208, May 2010.
- [47] Nuno Crokidakis and Fernando D Nobre. Destruction of first-order phase transition in a random-field ising model. *Journal of Physics: Condensed Matter*, 20(14):145211, 2008.
- [48] Octavio R Salmon, Nuno Crokidakis, and Fernando D Nobre. Multicritical behavior in a random-field ising model under a continuous-field probability distribution. *Journal of Physics: Condensed Matter*, 21(5):056005, 2009.
- [49] I. Pérez Castillo and N.S. Skantzos. The little-hopfield model on a sparse random graph. *Journal of Physics A: Mathematical and General*, 37(39):9087–9099, 2004. cited By 18.
- [50] M. Mézard and G. Parisi. The bethe lattice spin glass revisited. *The European Physical Journal B - Condensed Matter and Complex Systems*, 20(2):217–233, 2001.
- [51] M.E.J. Newman and Barkema G.T. *Monte Carlo Methods in Statistical Physics*. Oxford university press, 1999.

- 
- [52] E. Marinari and G. Parisi. Simulated tempering: A new monte carlo scheme. *EPL (Europhysics Letters)*, 19(6):451–458, 1992.
- [53] Daniel J. Amit. *Modeling Brain Function&Mdash;the World of Attractor Neural Networks*. Cambridge University Press, New York, NY, USA, 1989.
- [54] Mario González, David Dominguez, and Francisco B. Rodríguez. Block attractor in spatially organized neural networks. *Neurocomputing*, 72(16–18):3795 – 3801, 2009. Financial EngineeringComputational and Ambient Intelligence (IWANN 2007).
- [55] S. Bernard, N. Boujemaa, D. Vitale, and C. Bricot. Fingerprint classification using kohonen topologic map. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 3, pages 230–233 vol.3, 2001.
- [56] Pierre Baldi and Yves Chauvin. Neural networks for fingerprint recognition. *Neural Comput.*, 5(3):402–418, 1993.
- [57] M.M. Rashid and A.K. Hossaim. Fingerprint verification system using artificial neural network. *Information Technology*, 5(6), 2006.
- [58] Christopher Hillar, Jascha Sohl-Dickstein, and Kilian Koepsell. Efficient and optimal binary hopfield associative memory storage using minimum probability flow. *arXiv preprint arXiv:1204.2916*, 2012.
- [59] K.N. Mutter, Z.M. Jafri, and A.A. Aziz. Automatic fingerprint identification using gray hopfield neural network improved by run-length encoding. In *Computer Graphics, Imaging and Visualisation, 2008. CGIV '08. Fifth International Conference on*, pages 205 –210, aug. 2008.
- [60] Mario González, David Dominguez, Francisco B Rodríguez, and Angel Sanchez. Retrieval of noisy fingerprint patterns using metric attractor networks. *International journal of neural systems*, 24(07), 2014.

- 
- [61] R. Picard. Structured patterns from random fields. In *Proc. Asilomar Conf. Signals, Syst. and Comp.*, volume 9, pages 749–756., 1992.
- [62] Patricio Pérez and Dino Salinas. Storage of structured patterns in a neural network. *Phys. Rev. E*, 50:4182–4186, Nov 1994.
- [63] Michael Biehl, Reimer Kühn, and Ion-Olimpiu Stamatescu. Learning structured data from unspecific reinforcement. *Journal of Physics A: Mathematical and General*, 33(39):6843, 2000.
- [64] R Monasson. Properties of neural networks storing spatially correlated patterns. *Journal of Physics A: Mathematical and General*, 25(13):3701, 1992.
- [65] A. Singh. Detecting weak but hierarchically-structured patterns in networks. In *Journal of Machine Learning Research: Workshop and Conference Proceedings*, pages 1011–1015, 2010.
- [66] Shinsuke Koyama. Storage capacity of two-dimensional neural networks. *Phys. Rev. E*, 65:016124, Dec 2001.
- [67] Hidetoshi Nishimori, W. Whyte, and D. Sherrington. Finite-dimensional neural networks storing structured patterns. *Phys. Rev. E*, 51:3628–3642, Apr 1995.
- [68] David Dominguez, Mario González, Francisco B. Rodríguez, Eduardo Serrano, R. Erichsen Jr., and W.K. Theumann. Structured information in sparse-code metric neural networks. *Physica A: Statistical Mechanics and its Applications*, 391(3):799 – 808, 2012.
- [69] Maio Maltoni Cappelli, D. Maio, D. Maltoni, J. L. Wayman, and A. K. Jain. FVC2004: Third fingerprint verification competition. In *in Proceedings of the First International Conference on Biometric Authentication*, pages 1–7, 2004.
- [70] MATLAB and Statistics Toolbox. *The MathWorks, Inc.*, 2012.



- 
- [71] Bruno A Olshausen and David J Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4):481–487, 2004.
- [72] David Dominguez, Mario González, Eduardo Serrano, and Francisco B. Rodríguez. Structured information in small-world neural networks. *Phys. Rev. E*, 79(2):021909, Feb 2009.
- [73] D. R. C. Dominguez and D. Bollé. Self-control in sparsely coded networks. *Phys. Rev. Lett.*, 80:2961–2964, Mar 1998.
- [74] B Derrida, E Gardner, and A Zippelius. An exactly solvable asymmetric neural network model. *EPL (Europhysics Letters)*, 4(2):167, 1987.
- [75] O E Wohlman and C Domb. Random field ising model on the bethe lattice. *Journal of Physics A: Mathematical and General*, 17(11):2247, 1984.
- [76] F.F. Doria, R. Erichsen Jr., D. Dominguez, Mario González, and S.G. Magalhaes. Random field ising model in a random graph. *Physica A: Statistical Mechanics and its Applications*, 422:58 – 65, 2015.
- [77] F.F. Doria, R. Erichsen Jr., Mario González, Francisco B. Rodríguez, Ángel Sánchez, and David Dominguez. Structured patterns retrieval using a metric attractor network: Application to fingerprint recognition. *Physica A: Statistical Mechanics and its Applications*, 457:424 – 436, 2016.