

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE FÍSICA

**Redes Neurais Atratoras com padrões que possuem atividade
em grafos aleatórios**

Alexandre Silveira

Dissertação de Mestrado realizada sob a orientação do Prof. Rubem Erichsen Junior, apresentado ao Instituto de Física da UFRGS em preenchimento parcial dos requisitos para a obtenção do grau de Mestre em Física.

Porto Alegre, 25 de julho de 2017

Sumário

1	Introdução	3
2	Redes Atratoras	6
2.1	A memória associativa e o modelo de Hopfield	6
2.1.1	Caso de um padrão	8
2.1.2	Múltiplos padrões	8
2.1.3	Capacidade da Rede de Hopfield	9
2.1.4	Energia do modelo	10
2.1.5	Apresentação alternativa do problema	10
2.2	Neurônios estocásticos	11
2.3	Modelo de Hopfield com campo médio	11
2.3.1	Regime $\alpha = 0$	11
2.3.2	Regime $\alpha > 0$	12
2.4	Redes Diluídas	13
2.5	Redes com atividade finita	14
3	Modelo	16
3.1	Grafos	16
3.1.1	Grafos Aleatórios	17
3.2	Rede atratora com atividade finita num grafo aleatório	17
3.2.1	Método de Réplicas	19
3.3	Cálculo das funções de ordem	20
3.3.1	Campo com vínculo	28
3.3.2	Energia livre por partícula com campo de vínculo	31
3.3.3	Expandindo a Energia livre	32
3.4	Cálculo das distribuições de campos locais em $T = 0$	34
4	Resultados	37
4.1	Cálculo da distribuição de campos locais de subrede	37
4.2	Comparando as regras de aprendizado	42
4.2.1	Diagramas de fase no plano (α, a)	42
4.2.2	Diagramas de fase no plano (α, T)	44
4.2.3	Diagramas de fase no plano (a, T)	51
5	Considerações finais e perspectivas	61

Agradecimentos

Primeiramente, gostaria de agradecer ao professor Rubem pela paciência e dedicação em me orientar.

Agradeço aos meus pais por todo apoio e carinho.

Aos colegas e amigos do curso de física.

Resumo

Com o avanço das técnicas analíticas, tem sido possível estudar redes neurais atratoras onde cada unidade de processamento é conectada com um número finito de vizinhos, sendo que esse número independe do tamanho do sistema. Aplicamos essas técnicas ao estudo de redes atratoras com padrões que possuem uma quebra de simetria sobre o número de bits ativos e quiscetes. O objetivo deste trabalho é estudar a capacidade da rede neural em armazenar padrões com atividade não nula, uma vez que a conectividade por neurônio é finita. Inicialmente, apresentamos os modelos predecessores de redes atratoras, como o modelo de Hopfield e os modelos de Amit, Gutfreund e Sompolinsky. Em tais modelos, o aprendizado é definido através de modificações sinápticas, inspiradas nas ideias de Hebb. Mostramos como é estimada a capacidade da rede. Mencionamos a introdução de uma função de energia para o sistema, que permite uma ligação com estudo de sistemas magnéticos através da mecânica estatística. Apresentamos também regras de aprendizado para lidar com padrões com atividade não nula. Num segundo momento, aplicamos o método de réplicas, utilizado para tratar sistemas desordenados, ao problema da rede atratora com conectividade e atividade dos padrões finitas. Utilizamos o formalismo de funções de ordem e fazemos uso do conceito de sub-redes, que permite particionar o grafo de acordo com os padrões a serem armazenados em cada neurônio. Obtemos, assim, uma função de ordem por sub-rede que contém toda informação sobre o estado do sistema. Aplicando o *ansatz* de simetria de réplicas, é possível derivar distribuições autoconsistentes dos campos locais para cada sub-rede. Tais distribuições passam a fornecer toda informação necessária para calcularmos os observáveis relevantes. As distribuições são calculadas numericamente a partir do método da dinâmica de populações. Em seguida, traçamos diagramas de fases para três regras de aprendizado. A partir desses, estimamos a capacidade, temperatura e atividade críticas. Observa-se a presença de fases de vidro de spin, transições descontínuas e pontos tricríticos.

Abstract

With the advance of analytical tools it has been possible to study attractor neural networks in which each processing unit is connected to a finite number of neighbours. Being that, the number of neighbours is independent of the size of the system. We apply these tools to the study of attractor networks in which the patterns have a broken symmetry with respect to the number of active and inactive bits. The objective of this work is to study the capacity of the neural network to store patterns with activity different from zero, being that the connectivity per neuron is finite. First, we present the predecessor models of attractor networks like the Hopfield and Amit, Gutfreund e Sompolinsky ones. In such models the learning is defined through synaptic modifications, inspired by Hebb's ideas. We show how to estimate the storage capacity of the network. We mention the introduction of a energy function for the system, which allows a link with the study of magnetic systems through statistical mechanics. We present learning rules to deal with patterns which have non zero activity. In the second part, we apply the replic method, utilized to deal with disordered systems to the problem of an attractor neural network with finite connectivity and activity. We utilized the formalism of order functions and the concept of sublattices, this concept allows to partition the graph according with the patterns to be stored in each neuron. This way, we obtain an order function per sublattice which contain all the information about the state of the system. Applying the replica symmetry *ansatz* it is possible to derive self-consistent distributions of the local fields per sublattice. Such distributions start to provide all the necessary information to calculate the relevant observables. These distributions are calculated numerically using the population dynamics method. Then, we draw phase diagrams for three learning rules. Using these, we estimate the storage capacity, the temperature and the critical activity. We observe the presence of spin glass phases, discontinuos phase transiotions and tricritical points.

Capítulo 1

Introdução

O cérebro humano é capaz de realizar muitas tarefas de forma mais eficiente que um computador. Uma dessas tarefas é o processamento visual. Suas unidades de processamento são os neurônios. Um neurônio é constituído de um corpo celular (soma), dendritos e um axônio. Uma representação está ilustrada na figura 1.1. O neurônio funciona como um mecanismo de recepção e transmissão de sinais elétricos. Sinais oriundos de outras células são recebidos pelo soma e pelos dendritos. A transmissão de sinais para outros neurônios é dada pelo axônio, através de uma sinapse, que é uma conexão entre o axônio de um neurônio e soma ou os dendritos de outro. Um neurônio pode possuir em média 10^4 sinapses. Sinapses podem ser excitatórias ou inibitórias. Sinapses excitatórias elevam o potencial elétrico do corpo do neurônio e sinapses inibitórias o reduzem. Após o potencial elétrico do soma exceder um certo limiar, um pulso elétrico é enviado ao axônio e propagado até outras células. Denomina-se esse estado do neurônio como ativo, ou disparando. Se o potencial do neurônio não excedeu seu limite e não há sinal sendo propagado através de seu axônio, denomina-se esse estado como quiescente.

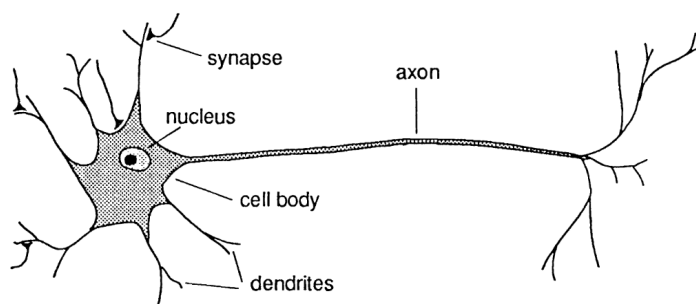


Figura 1.1: Representação de um neurônio. Figura retirada da referência [1]

Em 1943, McCulloch e Pitts (MP) [2] introduziram o modelo de neurônios formais, onde os neurônios são representados por unidades binárias que assumem valores $\sigma \in \{0, 1\}$. Quando $\sigma = 1$ o neurônio se encontra num estado ativo e $\sigma = 0$ significa que está quiescente. Similar ao comportamento de um neurônio real, o estado de uma unidade de processamento é o resultado das contribuições dos elementos adjacentes a ela:

$$\sigma_i(t+1) = \Theta\left(\sum_j \omega_{ij}\sigma_j(t) - U_i\right). \quad (1.1)$$

O tempo t é uma variável discreta. $\Theta(x)$ é a função de Heaviside. As variáveis ω_{ij} representam os pesos sinápticos e U_i representa o limiar de ativação do neurônio i . Valores positivos de ω_{ij} são análogos a sinapses excitatórias e negativos a inibitórias. Os autores demonstraram que tomando a escolha adequada de ω_{ij} e U_i , um neurônio formal pode representar uma função lógica, como OR ou AND. A partir desse modelo, surgem estudos de redes desses neurônios visando aplicá-las a tarefas computacionais. Uma dessas tarefas é a memória associativa, que consiste em armazenar e reconhecer padrões. Um padrão nesse contexto, é uma palavra composta por N bits, representados pela variável ξ_i . O valor que bit pode assumir está associado ao estado de ativação ou quiescência de um neurônio.

Em 1982, J. J. Hopfield [3] introduziu um modelo de rede constituída de neurônios formais de (MP), aplicado ao problema da memória associativa. Em seu modelo, as memórias ou padrões estão associados a atratores da dinâmica. As variáveis de neurônios em seu modelo assumem valores $\sigma \in \{-1, 1\}$, bem como os valores associados aos bits de um padrão $\xi \in \{-1, 1\}$. Baseando-se na ideia de Hebb [4], Hopfield introduz o aprendizado de forma local, recorrendo apenas a modificações sinápticas apropriadas. Fazendo a suposição de que as sinapses são simétricas, $J_{ij} = J_{ji}$, Hopfield introduz uma função de energia. Esse função permite uma analogia com sistemas magnéticos, como o modelo de Ising e também uma conexão com modelos de vidro de spin. Apesar do sucesso, o modelo de Hopfield possui vínculos que não existem em uma rede neural real, tais como: o sinal que um neurônio recebe de todos seus vizinhos é perfeito, não havendo atrasos ou variações de amplitude, logo, não há ruído sináptico; cada neurônio se conecta com todos neurônios da rede, porém, evidências biológicas mostram que no cérebro o número de sinapses por neurônio é finito; o bit de um certo padrão ξ_i^μ é uma variável aleatória, logo, bits ativos e quiescentes possuem a mesma probabilidade:

$$P(\pm 1) = \frac{1}{2}, \quad (1.2)$$

ou seja, em um padrão tem-se que 50% dos bits são ativos e 50% são quiescentes. Isso contrasta com evidências sobre a atividade neural no córtex, onde supostamente existem mais neurônios quiescentes do que ativos. Para contornar essa aleatoriedade, pode-se adicionar um *bias* à probabilidade do valor que um bit pode assumir:

$$P(\xi = \pm 1) = \frac{1}{2}(1 \pm a), \quad (1.3)$$

a variável $a \in \{-1, 1\}$ é denominada "atividade". Essa definição implica que a fração de bits ativos num padrão é dada por:

$$A = \frac{1}{2}(1 + a). \quad (1.4)$$

Logo, se $a = 1 (A = 1)$, existem apenas bits ativos, e se $a = -1 (A = 0)$, existem apenas bits quiescentes e não há atividade no padrão. O nível de atividade num padrão é medido por:

$$\frac{1}{N}(N_+ - N_-) = a, \quad (1.5)$$

onde N_+ representa o número de bits ativos e N_- o número de bits quiescentes. Pode-se também, motivar o estudo de padrões com *bias* traçando um paralelo com estudo de imagens, onde bits ativos podem ser associados a pixels pretos e bits quiescentes a pixels brancos, e levando em consideração que em imagens, geralmente, não há uma distribuição uniforme de pixels pretos e brancos [5].

A adição de ruído sináptico ao modelo foi feita por Amit, Gutfreund e Sompolinsky

(AGS). Os autores associaram o ruído sináptico à flutuações térmicas e resolveram o modelo de Hopfield, utilizando a teoria de campo médio. Na referência [6], (AGS) tratam do modelo de Hopfield com campo médio no regime em que o número de padrões armazenados é finito frente ao tamanho do sistema, e, na referência [7], estudam o caso em que o número de padrões escala de forma linear com o tamanho do sistema. O estudo de bits que possuem *bias* está contido na referência [8], nesse trabalho, regras sinápticas são introduzidas para otimizar o funcionamento da memória para padrões com atividade.

Redes neurais atratoras com conectividade finita trata-se de um assunto mais recente. Trabalhos precursores foram realizados nas chamadas redes diluídas, onde o número de conexões c escala com o tamanho do sistema da seguinte forma $c = \log(N)$. Apesar de existirem infinitas conexões, o número ainda é finito frente ao tamanho do sistema, ou seja: $\lim_{N \rightarrow \infty} \frac{c}{N} \rightarrow 0$. Derrida, Gardner e Zippelius (DGZ) [9] estudaram um modelo de rede diluída em que, diferentemente do modelo de Hopfield, as sinapses são assimétricas, $J_{ij} \neq J_{ji}$. Já no artigo de Watkin e Sherrington [10], os autores resolvem o modelo de rede neural diluída mantendo o vínculo de simetria das sinapses. O modelo de Hopfield com conectividade finita somente foi resolvido mais recentemente em 2003 por Coolen e Wemmenhove (CW) [11], pois as ferramentas para o tratamento analítico de sistemas onde c é finito e independe do tamanho do sistema, que envolvem funções de ordem, foram desenvolvidas apenas em 1997, num trabalho publicado por Monasson e Zechin (MZ) [12]. No trabalho de (CW), os autores diluem as conexões de forma aleatória, escolhendo uma probabilidade $\frac{c}{N}$ de existir uma conexão, mantendo c finito e independente do tamanho do sistema. Essa escolha de diluição caracteriza um grafo aleatório. Porém, existem trabalhos em outros tipo de redes, como redes sem escala [13] e redes de mundo pequeno [14].

O objetivo deste trabalho é estudar modelos de redes atratoras, onde a conectividade média de cada neurônio é finita e os padrões possuem atividade. Para tal, utilizaremos o formalismo desenvolvido por (CW) e aplicaremos as regras de aprendizado para padrões com atividade finita, desenvolvidas por (AGS). Desejamos estimar a capacidade de operação de três regras de aprendizado, bem como a tolerância das regras à elevação da atividade dos padrões.

O trabalho está organizado da seguinte forma: no capítulo 2, faremos uma revisão do modelo de Hopfield, explanaremos as regras de aprendizado como modificações sinápticas para um e múltiplos padrões e mostraremos uma estimativa para capacidade da rede para um certo número de padrões. Ainda nesse capítulo, faremos uma breve revisão dos modelos (AGS); No capítulo 3, faremos uma pequena apresentação sobre grafos e grafos aleatórios. Apresentaremos, em certo nível de detalhe, o método de réplicas. Aplicaremos o método de funções de ordem utilizado para sistemas com conectividade finita ao problema de uma rede neural atratora, do tipo Hopfield com padrões que possuem atividade. No capítulo 4, apresentaremos os resultados e dissertaremos acerca do efeito da atividade sobre a capacidade de funcionamento de três regras de aprendizado. No capítulo 5, apresentaremos uma discussão dos resultados e perspectivas futuras.

Capítulo 2

Redes Atratoras

Nesse capítulo, faremos uma breve introdução sobre modelos de redes neurais do tipo atratoras. Introduziremos a regra de aprendizado de Hebb e o modelo de rede de atratora de Hopfield, que traça uma ligação entre redes neurais e modelos de spins tratados com mecânica estatística. Explanaremos, brevemente, os trabalhos de (AGS) [6, 7], que resolvem o modelo de Hopfield na presença de ruído térmico, usando técnicas de campo médio desenvolvidas para tratar sistemas de vidros de spin. Será apresentado também, os resultados obtidos nas redes diluídas. Ainda nesse capítulo, mencionaremos o trabalho de (AGS) [8] que trata do modelo de Hopfield com padrões que possuem atividade finita.

2.1 A memória associativa e o modelo de Hopfield

O problema da memória associativa pode ser enunciado da seguinte forma. Considere um número de p padrões, ou memórias, que devem ser ensinados a memória, através de alguma regra de aprendizado. Uma vez que apresentamos um novo padrão a essa memória, ela deve ser capaz de reproduzir um estado que seja o mais próximo desse padrão, utilizando apenas os padrões previamente armazenados. Por exemplo, após uma etapa de treinamento, onde ensina-se a memória um conjunto de imagens, apresenta-se um fragmento da imagem, como na figura 2.1, e a partir disso a memória é capaz de recuperar a imagem completa.

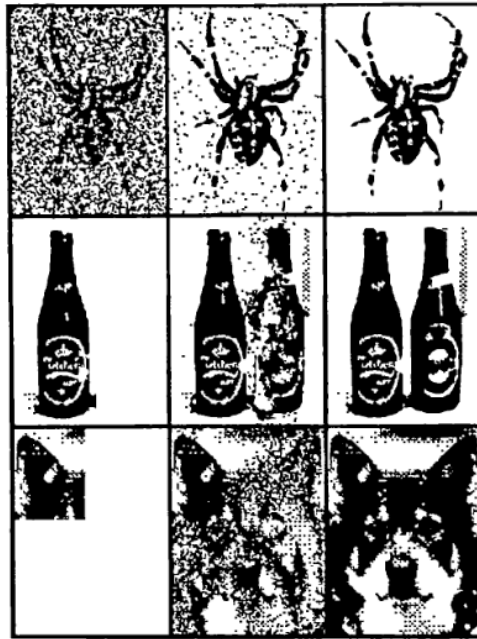


Figura 2.1: Figura retirada da referência [1]

Redes atratoras constituem de modelos teóricos, inspirados em redes neurais biológicas, que podem ser utilizados para executar a tarefa de memória associativa. Memórias ou padrões são pontos fixos (atratores) da dinâmica. As unidades de processamento, os neurônios, são versões simplificadas dos neurônios reais. Esses neurônios são interconectados de maneira a formar uma rede fechada. Sua estrutura de conexões contrasta com redes neurais em camadas, como o perceptron [15].

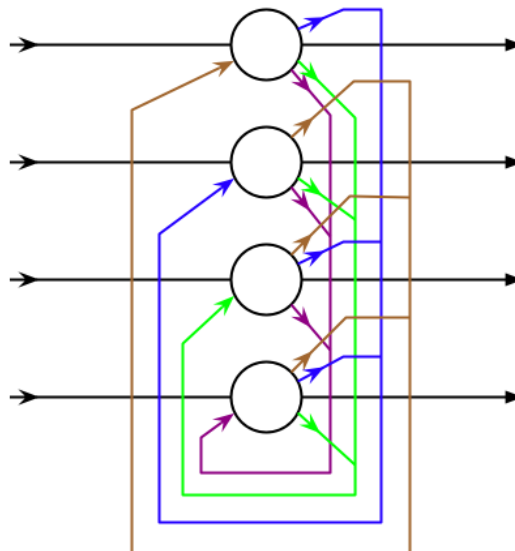


Figura 2.2: Rede de Hopfield, onde todos os neurônios estão completamente interconectados. Figura retirada da referência [16]

No modelo de Hopfield [3], os neurônios estão dispostos na rede de forma que cada um está completamente conectado com todos elementos da rede. Na figura 2.3, há um exemplo da

estrutura para o caso de quatro neurônios. As interações entre os neurônios devem ser simétricas, $J_{ij} = J_{ji}$, esse requerimento é necessário para que existam pontos de equilíbrio. O sistema é atualizado de forma assíncrona, cada neurônio é selecionado aleatoriamente e seu estado posterior é determinado através da equação:

$$\sigma_i = \text{sgn}\left(\sum_j J_{ij}\sigma_j\right), \quad (2.1)$$

onde

$$\text{sgn}(x) = \begin{cases} +1 & \text{se } x \geq 0 \\ -1 & \text{se } x < 0. \end{cases} \quad (2.2)$$

Hopfield utilizou como regra de aprendizado a ideia de Hebb [4], a qual afirma que a ativação simultânea de um par de neurônios acarreta num fortalecimento de suas sinapses. A regra de Hebb implica num aprendizado local, o estado de um neurônio depende apenas da interação com cada elemento a ele conectado pela sinapses J_{ij} .

2.1.1 Caso de um padrão

Partindo da ideia de introduzir aprendizado a partir das sinapses, apresentamos uma forma de motivar a regra sináptica utilizada por Hopfield. Considerando a equação 2.1, deseja-se que a rede possua um ponto estacionário $\sigma = \{\xi_1 \dots \xi_N\}$, que corresponda a uma memória. Uma possível escolha de J_{ij} que satisfaz essa condição é:

$$J_{ij} = \frac{1}{N}\xi_i\xi_j, \quad (2.3)$$

substituindo essa tentativa na equação 2.1, resulta em

$$\sigma_i = \text{sgn}(\xi_i), \quad (2.4)$$

de forma que o estado final que cada neurônio se encontra é o mesmo de um bit correspondente ξ_i da memória que se deseja inserir. Temos então, que a sinapse 2.3 garante que $\sigma = \{\xi_1 \dots \xi_N\}$ é um ponto fixo da dinâmica.

2.1.2 Múltiplos padrões

Uma extensão natural da regra sináptica para inserir muitos padrões é:

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu. \quad (2.5)$$

No entanto, essa sinapse pode acarretar instabilidades, caso o número de padrões inseridos exceda um certo limite. Para demonstrar isso, utilizamos a equação 2.1 e consideramos uma análise de sinal e ruído. Supondo que o sistema se encontra inicialmente no padrão: $\sigma = \{\xi_1^\mu \dots \xi_N^\mu\}$, o campo externo que atua sobre cada neurônio σ_i é dado por:

$$h_i^\mu = \frac{1}{N} \sum_{j=1}^N \sum_{\nu=1}^p \xi_i^\nu \xi_j^\nu \xi_j^\mu. \quad (2.6)$$

O requerimento de estabilidade do estado da rede no padrão μ é:

$$\xi_i^\mu = \text{sgn}(h_i^\mu). \quad (2.7)$$

Separamos o campo em duas contribuições

$$h_i^\mu = \xi_i^\mu + \delta_i, \quad (2.8)$$

o termo ξ_i^μ é denominado termo de sinal e $\delta_i = \frac{1}{N} \sum_{j=1}^N \sum_{\substack{\nu \\ \nu \neq \mu}}^p \xi_i^\nu \xi_j^\nu \xi_j^\mu$ corresponde ao termo do ruído provocado pela existência dos outros $p - 1$ padrões. Se δ_i for suficientemente pequeno, o termo de ruído não afeta o sinal de cada bit ξ_i^μ , mantendo a rede estável no padrão μ . No caso de padrões aleatórios, δ_i é apenas a soma de $(N - 1)(p - 1)$ variáveis aleatórias que assumem valores -1 ou 1 com probabilidade $p = \frac{1}{2}$. A média dessa soma é $\langle \delta_i \rangle = 0$ e a variância $s^2 = \langle \delta_i^2 \rangle - \langle \delta_i \rangle^2 = \frac{p}{N}$. Logo, se p for suficientemente pequeno em relação a N , o valor de δ_i será aproximadamente zero e o termo de sinal dominará o campo h_i^μ , tornando o estado da rede no padrão ξ_i^μ estável.

2.1.3 Capacidade da Rede de Hopfield

Faremos aqui uma pequena estimativa para quais valores de p o modelo de Hopfield consegue operar como memória associativa. Definindo a quantidade:

$$C_i^\mu = -\xi_i^\mu \frac{1}{N} \sum_{j=1}^N \sum_{\substack{\nu \\ \nu \neq \mu}}^p \xi_i^\nu \xi_j^\nu \xi_j^\mu, \quad (2.9)$$

Se $C_i^\mu > 1$ o termo de sinal do padrão ξ_i^μ e o termo de ruído δ_i possuem sinais opostos, isso também indica que o valor do termo de ruído $\delta_i > 1$, logo o termo de ruído domina o campo local e o bit ξ_i^μ encontra-se numa configuração instável. De forma que se $C_i^\mu > 1$ o bit i possui uma probabilidade de estar errado, ou seja

$$P_{\text{erro}} = P(C_i^\mu > 1) \quad (2.10)$$

O termo δ_i é regido por uma distribuição binomial que possui média nula e variância p/N . No limite $Np \gg 1$ esta distribuição pode ser aproximada por uma distribuição Gaussiana com as mesmas médias e variâncias.

$$P(C_i, \langle C_i \rangle = 0, s = \sqrt{p/N}) = \frac{1}{\sqrt{2s^2\pi}} \exp\left(-\frac{(C_i - \langle C_i \rangle)^2}{2s^2}\right) \quad (2.11)$$

P_{error}	p_{max}/N
0.001	0.105
0.0036	0.138
0.01	0.185
0.05	0.37
0.1	0.61

Figura 2.3: Tabela de valores de p_{max} para o grau de tolerância de erros . Figura retirada da referência [1]

A probabilidade de erro P_{erro} é a integral sobre a área da Gaussiana que corresponde à valores de $C_i^\mu > 1$

$$P_{erro} = \int_1^\infty dC_i \frac{1}{\sqrt{2\sigma\pi}} \exp\left(-\frac{(C_i - \langle C_i \rangle)^2}{2\sigma^2}\right) \quad (2.12)$$

$$= \frac{1}{2} \left[1 - \operatorname{erf}\left(\frac{1}{\sqrt{2\sigma^2}}\right)\right] = \frac{1}{2} \left[1 - \operatorname{erf}\left(\sqrt{N/2p}\right)\right] \quad (2.13)$$

Isto implica, que, caso se deseje que a probabilidade de erro de um bit seja inferior a $P_{erro} < 0.036$, o valor máximo de padrões que permite essa escolha de erro é $p \leq 0.138N$. A partir dessa análise, podemos estimar que a capacidade de operação da rede está relacionado à fração do número de padrões inseridos pelo número de sítios da rede:

$$\alpha = \frac{p}{N}. \quad (2.14)$$

2.1.4 Energia do modelo

Talvez a maior contribuição do artigo publicado por Hopfield seja a introdução de uma função energia. Hopfield argumenta que caso $J_{ij} = J_{ji}$, e a dinâmica dos neurônios seja regida pela equação 2.1, existe uma função que é decrescente sobre a mudança de $\sigma_i \rightarrow -\sigma_i$. Essa função é definida por:

$$E = -\frac{1}{2} \sum_{i,j} \sum_{\mu} \xi_i^\mu \xi_j^\mu \sigma_i \sigma_j. \quad (2.15)$$

A introdução dessa função de energia permite uma analogia direta com modelos como o de Ising e traz o problema de memória associativa a uma classe de problemas que pode ser estudado com mecânica estatística.

2.1.5 Apresentação alternativa do problema

Baseado na ideia de uma função de energia, podemos introduzir o problema de memória associativa da seguinte forma: Seja uma rede de N neurônios de 2 estados, sendo que estados ativos são representados por $\sigma_i = 1$ e quiescentes por $\sigma_i = -1$. Dado que um número de p padrões $\xi^\mu = \{\xi_1^\mu \dots \xi_N^\mu\}$ é armazenado na memória, com $\xi_i^\mu \in \{-1, 1\}$ e $\mu = 1, \dots, p$. Deseja-se que, a partir de uma configuração inicial $\sigma = \{\sigma_1 \dots \sigma_N\}$ próxima a um dos p padrões, o sistema seja capaz de evoluir de forma a recuperar completamente esse padrão. Isso pode ser expresso da seguinte forma: considere a quantidade

$$d_H = \frac{1}{2N} \sum_{i=1}^N (\sigma_i - \xi_i^\mu)^2 = (1 - m^\mu), \quad (2.16)$$

que mede a distância de cada bit σ_i , que corresponde ao estado atual do sistema, se encontra de um bit ξ_i^μ de um certo padrão. O parâmetro de ordem $m^\mu = \frac{1}{N} \sum_{i=1}^N \sigma_i \xi_i^\mu$, conhecido como *overlap*, mede a correlação do estado da rede com o padrão $\{\xi_1^\mu \dots \xi_N^\mu\}$. A tarefa da rede deve ser a de minimizar a distância d_H , o que é o mesmo que maximizar o *overlap* m^μ . Logo, podemos escrever essa tarefa como um problema de minimizar uma certa função custo:

$$H = -\frac{1}{2N} \sum_{\mu=1}^p (m^\mu)^2, \quad (2.17)$$

substituindo a definição do *overlap* e expandindo os produtos, obtemos a mesma expressão de energia introduzida por Hopfield.

2.2 Neurônios estocásticos

Em redes neurais reais, o potencial de ativação dos neurônios é variável e as sinapses possuem atrasos. Esses efeitos podem ser entendidos como ruído sináptico e, numa formulação de mecânica estatística, representados por flutuações térmicas, onde a dinâmica não é deterministicamente regida pelo conjunto de equações 2.1, mas sim por uma dinâmica estocástica [17], em que atribuímos uma densidade de probabilidade:

$$P(\pm\sigma_i) = \frac{1}{1 + e^{\mp 2\beta h_i \sigma_i}}, \quad (2.18)$$

a um dado estado σ_i do neurônio. A variável $\beta = \frac{1}{T}$ controla o nível de ruído. Nessa dinâmica estocástica, calcula-se o valor médio do estado do neurônio:

$$\langle \sigma_i \rangle = \sum_{\sigma_i = \pm 1} P(\sigma_i) \sigma_i = \tanh(\beta h_i). \quad (2.19)$$

Sendo que $\beta \rightarrow \infty$, a dinâmica torna-se determinística novamente, e o estado de cada neurônio fica determinado pela equação 2.1.

2.3 Modelo de Hopfield com campo médio

A resolução do modelo de Hopfield utilizando o método de campo médio se deve a (AGS). Nos artigos [6] [7], são estudados, respectivamente, os modelos no regime de baixa alocação de padrões $p \ll N$ ($\alpha = 0$) e no regime de saturação $p = \alpha N$.

2.3.1 Regime $\alpha = 0$

O primeiro caso resolvido foi no regime $N \rightarrow \infty$ e p finito. Os resultados obtidos são que abaixo de uma temperatura $T = T_c = 1$, o sistema passa de uma dinâmica ergódica, uma fase paramagnética (P), onde não existe ordenamento global do sistema com nenhum dos p padrões, para um fase de quebra de ergodicidade, em que existem $2p$ estados estáveis. Cada p desses estados está relacionado com ordenamento macroscópico da rede com um dos p padrões. O fator 2 é devido ao fato que os estados $\xi = \{\xi_1^\mu \dots \xi_N^\mu\}$ e $\xi = \{-\xi_1^\mu \dots -\xi_N^\mu\}$ são energeticamente degenerados. Denominam-se estados em que o sistema está correlacionado unicamente com um padrão de estados de recuperação (R), ou também como de Mattis [18]. No intervalo de $0.46 < T < 1$, os estados (R) são os únicos estáveis, abaixo de $T = 0.46$ estados de misturas aparecem e seu número aumenta a medida que T decresce até $T = 0$. Estados de mistura são aqueles em que o sistema está correlacionado com mais de um padrão. Misturas podem ser tanto simétricas ou assimétricas. No capítulo 4 faremos uma classificação mais completa de estados de mistura.

2.3.2 Regime $\alpha > 0$

Quando o número de padrões escala de forma $p = \alpha N$, a informação sobre o estado do sistema fica determinado em um diagrama de fases no plano (α, T) . Os parâmetros de ordem que conferem a informação sobre a fase termodinâmica desse sistema são calculados a partir da energia livre por partícula:

$$f = - \lim_{N \rightarrow \infty} \frac{1}{\beta N} \langle \langle \log Z \rangle \rangle, \quad (2.20)$$

onde $\langle \langle \dots \rangle \rangle$ representa a média sobre todas as possíveis realizações de $\xi = \{\xi_1^\mu \dots \xi_N^\mu\}$. O cálculo da média $\langle \langle \log(Z) \rangle \rangle$, requer o uso do método de réplicas, originalmente utilizado para tratar sistemas de vidro de spin, como os trabalhos de Sherrington-Kirkpatrick (SK) [19] e Edwards-Anderson (EA) [20]. De forma resumida, o método consiste em utilizar uma das seguintes identidades:

$$\langle \langle \log(Z) \rangle \rangle = \lim_{n \rightarrow 0} \frac{\langle \langle Z^n \rangle \rangle - 1}{n}, \quad (2.21)$$

ou

$$\langle \langle \log(Z) \rangle \rangle = \lim_{n \rightarrow 0} \frac{\log \langle \langle Z^n \rangle \rangle}{n} \quad (2.22)$$

e assumir que $n \in \mathbb{N}$, de maneira a executar o produto de n funções de partição $Z^n = \prod_{\rho=1}^n Z^\rho$ e tratá-las como elementos independentes. Após tomar o traço na função de partição, e obter uma expressão fechada para f , o que usualmente implica em trocar a ordem que executamos os limites $\lim_{N \rightarrow \infty}$ e $n \rightarrow 0$, tomando primeiramente o limite termodinâmico. Em seguida continua-se analiticamente f para o domínio de $n \in \mathbb{C}$, e assim toma-se o $n \rightarrow 0$ [21]. Vidros de spin são sistemas caracterizados por possuírem desordem e frustração. Desordem significa que as interações entre os pares de spins não são uniformes, mas sim variáveis aleatórias distribuídas de acordo com alguma função de probabilidade $P(J_{ij})$. Frustração é devido à incapacidade dos spins satisfazerem todas suas interações com a vizinhança. A fase de vidro de spin (VS) é caracterizada pelo parâmetro de ordem de (EA):

$$q = \left\langle \frac{1}{N} \sum_{i=1}^N \langle \sigma_i \rangle^2 \right\rangle_J, \quad (2.23)$$

que mede o grau de atividade local dos spins da rede. O termo $\langle \dots \rangle_J$ indica que é necessário tomar a média sobre todas realizações de J . Numa fase VS, não há ordenamento global, porém cada spin se comporta como se sua orientação estivesse *congelada* numa certa direção. Os resultados obtidos por Amit estão expostos na figura 2.4.

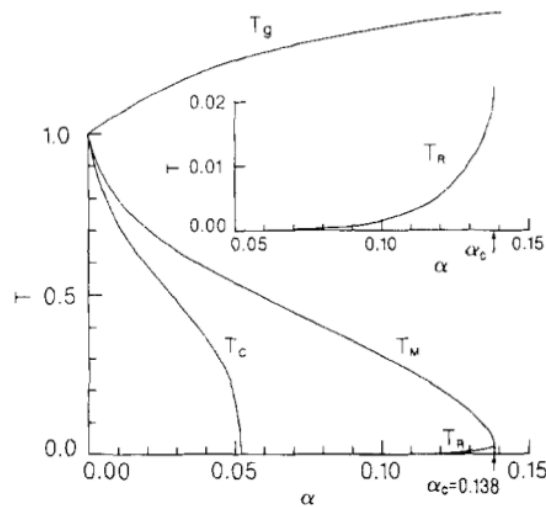


Figura 2.4: Diagrama de fases do modelo de Hopfield com $p = \alpha N$. Figura retirada da referência [6]

Em $T = 0$ e abaixo de $\alpha_c = 0.138$, o sistema funciona efetivamente como memória associativa e possui um *overlap* com um único padrão da ordem $m^\mu \approx 0.97$. Em α_c o sistema possui uma transição de fase descontínua, passando de uma fase com $m^\mu \approx 0.97$ para $m^\mu = 0$. Se $\alpha \rightarrow 0$, o *overlap* alcança seu valor máximo $m = 1$. Acima da linha T_g o sistema encontra-se numa fase P, em T_g temos uma transição de fase contínua P-VS. No intervalo de $T_M < T < T_g$ o sistema encontra-se numa fase VS, com $q \neq 0$ e $m^\mu = 0$. Em T_M o sistema sofre uma transição descontínua e estados R aparecem, sendo metaestáveis no intervalo $T_c < T < T_M$. Em T_c há uma transição descontínua onde o sistema entra numa fase em que estados R são estáveis. Para $\alpha \leq 0.03$, existem estados de misturas simétricas. Abaixo da linha de Almeida-Thouless T_R [22], a solução de simetria de réplicas é instável.

2.4 Redes Diluídas

Utilizando a teoria de simetria de réplicas Watkin e Sherrington [10] resolvem o modelo de Hopfield com interações simétricas numa rede diluída, onde o número de conexões por neurônio c escala com N , de forma que $\lim_{N \rightarrow \infty} \frac{c}{N} \rightarrow 0$. Nesse trabalho os autores verificam que para $T \rightarrow 0$ existem soluções de $m^\mu \neq 0$ apenas para valores de $\alpha < \frac{2}{\pi} = 0,634$, logo há um aumento da capacidade na rede diluída.

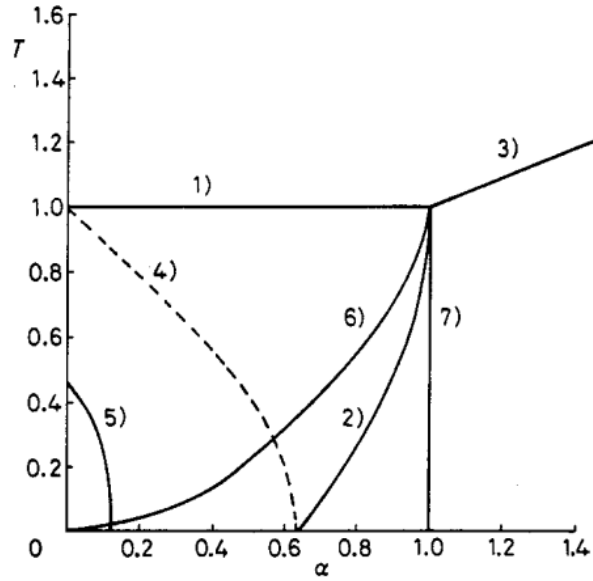


Figura 2.5: Diagrama de fases no plano (α, T) para rede diluída. Figura retirada de [10].

O resultado completo do modelo se apresenta no plano (α, T) . A linha de transição 1) separa soluções R de P, 2) representa a fronteira entre R e VS obtida utilizando a solução de simetria de réplicas, 3) separa as fases VS de P, a linha tracejada 4) representa a fronteira entre as fases R e VS no modelo de DGZ [9] onde as sinapses são assimétricas, abaixo de 5) existem estados de mistura, a linha 6) representa a linha de Almeida-Thouless [22] e a linha 7) representa a fronteira entre a fase R e VS utilizando a solução de quebra de simetria de réplicas.

2.5 Redes com atividade finita

Consideramos padrões que possuem atividade, de forma que cada componente de um certo padrão, ξ_i^μ , possui uma probabilidade:

$$P(\xi_i^\mu) = \frac{1+a}{2} \delta_{\xi_i^\mu, +1} + \frac{1-a}{2} \delta_{\xi_i^\mu, -1}, \quad (2.24)$$

onde $a \in \{-1, 1\}$.

Isso implica que a atividade média do padrão é

$$\langle \xi_i^\mu \rangle_{\xi^\mu} = a, \quad (2.25)$$

e, também que a correlação entre padrões é

$$\langle \xi_i^\mu \xi_i^\nu \rangle_{\xi^\mu} = \delta^{\mu\nu} + a^2(1 - \delta^{\mu\nu}). \quad (2.26)$$

Logo, embora dois padrões distintos: $\xi^1 = (\xi_1^1, \dots, \xi_N^1)$ e $\xi^2 = (\xi_1^2, \dots, \xi_N^2)$ sejam estatisticamente independentes obtidos, eles são correlacionados. Uma consequência disso é que o termo de ruído da equação 2.8 não tende a zero, tornando instáveis os estados R. Faz-se necessário

então uma modificação na regra de aprendizado. Na referência [8], (AGS) sugerem a seguinte modificação:

$$J_{ij} = \sum_{\mu=1}^p (\xi_i^\mu - a)(\xi_j^\mu - a), \quad (2.27)$$

que pode ser motivada por uma análise de sinal e ruído. Suponha que a rede se encontra num dos p padrões $\{\sigma_i = \xi_i^\mu\}$ e não há existência de ruído térmico $T = 0$. O campo externo que um neurônio sente de todos seus vizinhos numa rede totalmente conectada é:

$$h_i^\mu = \sum_j J_{ij} \xi_j^\mu = \xi_i^\mu [(1 - \xi_i^\mu a)(1 - a^2) + \delta_i], \quad (2.28)$$

onde $\delta_i = \frac{1}{N} \sum_{j \neq i}^N \sum_{\nu \neq \mu}^p \xi_i^\mu (\xi_i^\nu - a)(\xi_j^\nu - a) \xi_j^\mu$ representa o termo de ruído que os $(p-1)$ padrões fazem sobre o estado do neurônio i . Reescrevendo esse termo da seguinte forma:

$$\delta_i = \frac{1}{N} \sum_{\nu \neq \mu}^p \xi_i^\mu (\xi_i^\nu - a) \sum_{j \neq i}^N (\xi_j^\nu - a) \xi_j^\mu. \quad (2.29)$$

E notando que no limite $N \rightarrow \infty$:

$$\sum_{j \neq i}^N (\xi_j^\nu - a) \xi_j^\mu \rightarrow 0. \quad (2.30)$$

Conclui-se que com a nova regra, o termo de ruído tende a zero, e o estado da rede $\{\sigma_i = \xi_i^\mu\}$ é estável.

A segunda regra de aprendizado sugerida por (AGS) faz uso da regra sináptica 2.27, em conjunto com um campo externo que vincula a atividade média dos neurônios, de forma que esses possuam a mesma atividade dos padrões:

$$\sum_{i=1}^N \sigma_i = Na. \quad (2.31)$$

Na referência [8], mostra-se que com a adição desse campo de vínculo há um aumento na capacidade crítica em função da atividade $\alpha_c(a)$. No capítulo seguinte, mostraremos uma forma de adicionar esse vínculo a sistemas com conectividade finita.

Capítulo 3

Modelo

Nesse capítulo, faremos uma breve introdução sobre grafos e derivaremos a distribuição do número de vizinhos de um sítio (grau do sítio), no caso de um grafo aleatório. Introduziremos o método de réplicas, amplamente aplicado ao estudo de sistemas desordenados, como vidros de spin (SK) [19]. Em seguida, aplicaremos o formalismo de réplicas para sistemas com conectividade finita, que consiste em funções de ordem, desenvolvido inicialmente no contexto de problemas de satisfatibilidade (MZ) [12] e aplicado ao problema de redes neurais com conectividade finita por (CW) [11]. Aplicaremos esse método ao problema de uma rede neural atratora num grafo aleatório com atividade finita para os padrões.

3.1 Grafos

Nessa seção, apresentamos, brevemente, alguns conceitos e definições sobre grafos.

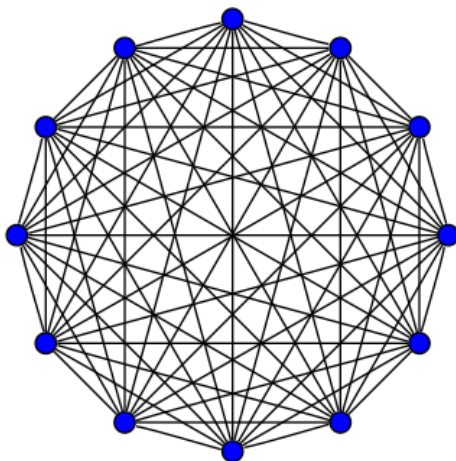


Figura 3.1: Grafo completo K_{12} que contém 66 arestas. Figura retirada de [23].

Grafos consistem de um sistema de pontos (nós, sítios, vértices) e linhas (conexões, ligações, arestas) entre um par de nodos. Matematicamente, um grafo pode ser definido como [24, 25] um par de conjuntos $G = (V, E)$, onde V consiste no conjunto de todos vértices do grafo G e E consiste no conjunto de todas as arestas do grafo, onde uma aresta representa uma ligação entre um

par de vértices (v, w) . Um grafo pode ser representado por sua matriz de adjacência $\mathbf{C} = \{c_{ij}\}$, onde as variáveis c_{ij} representam uma conexão se $c_{ij} = 1$, e a ausência de uma conexão, se $c_{ij} = 0$. Uma maneira de classificar o grafo é segundo a distribuição de probabilidade de seu grau $P(k)$, onde o grau k de um certo vértice mede o número de sítios vizinhos. Na figura 3.1, exemplificamos um grafo regular completo, onde cada vértice possui o mesmo número $k = 12$ de vizinhos.

3.1.1 Grafos Aleatórios

O estudo de grafos aleatórios foi iniciado por Paul Erdős e Alfréd Rényi (ER) [26]. Existem 2 modelos de grafos aleatórios. No primeiro modelo $G(N, M)$, o número de arestas é fixo e designa-se uma distribuição de probabilidade para obter-se um grafo de todos os possíveis grafos com N vertices e M arestas

$$P(G) = \frac{1}{\binom{\frac{N(N-1)}{2}}{M}}, \quad (3.1)$$

onde $\frac{N(N-1)}{2}$ mede o número de pares de sítios no grafo. Outra forma de obter um grafo aleatório é determinando uma probabilidade p de haver uma aresta entre dois nodos e uma probabilidade $(1-p)$ desta aresta não existir. Pode-se definir p de forma que no limite $N \rightarrow \infty$ a conectividade média por sítio c seja fixa. Adotando a seguinte definição:

$$P(c_{ij}) = \frac{c}{N} \delta_{c_{ij},1} + \left(1 - \frac{c}{N}\right) \delta_{c_{ij},0}, \quad (3.2)$$

temos que a conectividade média por sítio fica:

$$\frac{1}{N} \sum_{i,j} c_{ij} = c. \quad (3.3)$$

A probabilidade de um sítio qualquer possuir um grau k é determinada por:

$$P(k) = \frac{N!}{(N-k)!k!} \left(\frac{c}{N}\right)^k \left(1 - \frac{c}{N}\right)^{N-k}. \quad (3.4)$$

Tomando o limite de $N \rightarrow \infty$ obtemos o seguinte:

$$P(k) = \frac{N!}{(N-k)!k!} \left(\frac{c}{N}\right)^k \left(1 - \frac{c}{N}\right)^{N-k} \approx \frac{\exp(-c)c^k}{k!}. \quad (3.5)$$

Isso mostra que, no limite $N \rightarrow \infty$, a distribuição do grau de um sítio de um grafo do tipo (ER) se transforma numa distribuição Poissoniana de média c .

3.2 Rede atratora com atividade finita num grafo aleatório

Nessa seção, aplicamos método de réplicas para sistemas com conectividade finita a uma rede atratora na presença de ruído térmico. O estudo de mecânica estatística de sistemas com conectividade finita no contexto de simetria de réplicas, possui aplicabilidade em muitos

campos, alguns como: criptografia [27, 28], correção e detecção de erros em códigos [29, 30], problemas de otimização [12, 31], vidros de spin [32, 14, 33] e redes neurais [11, 13, 34, 35]. Trabalhos preliminares, utilizando a solução de simetria de réplicas em sistemas com conectividade finita [33], apesar de extraírem com sucesso as propriedades do sistema, possuíam uma complexidade analítica notável. Em alguns trabalhos posteriores, ferramentas mais sofisticadas e intuitivas foram desenvolvidas [12].

Faremos uso de um conjunto de p padrões $\xi^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)$ como memórias da rede. No caso tratado nesse trabalho, a atividade dos padrões é finita e as variáveis $\xi_i^\mu \in \{-1, 1\}$ possuem uma distribuição de probabilidade dada por:

$$P(\xi_i^\mu) = \frac{1+a}{2} \delta_{\xi_i^\mu, +1} + \frac{1-a}{2} \delta_{\xi_i^\mu, -1}. \quad (3.6)$$

Os neurônios da rede são representados por variáveis de spin do tipo Ising: $\sigma_i \in \{-1, 1\}$. Se $\sigma_i = 1$, o neurônio está ativo e se $\sigma_i = -1$, está quiescente. As variáveis c_{ij} representam conexões entre pares de neurônios, e a matriz de conectividade da rede $\mathbf{C} = \{c_{ij}\}$, é desenhada a partir da distribuição:

$$P(c_{ij}) = \frac{c}{N} \delta_{c_{ij}, 1} + \left(1 - \frac{c}{N}\right) \delta_{c_{ij}, 0}, \quad (3.7)$$

de forma que a conectividade média por neurônio seja fixa:

$$c = \frac{1}{N} \sum_{i,j} c_{ij}. \quad (3.8)$$

Como estamos trabalhando com padrões que possuem atividade finita a , é interessante introduzirmos as regras de aprendizados propostas por Amit [8]. É importante lembrar que essas regras foram estudadas num modelo de rede totalmente conectada. No entanto, até o presente momento, não foram testadas numa arquitetura de rede com conectividade finita do tipo (ER).

Utilizando a regra de Hebb modificada, que introduzimos na seção 2.5 do capítulo 2, definimos o Hamiltoniano do sistema:

$$H = -\frac{1}{2} \sum_{i,j} \frac{c_{ij}}{c} \sum_{\mu=1}^p (\xi_i^\mu - ab)(\xi_j^\mu - ab) \sigma_i \sigma_j, \quad (3.9)$$

Introduzimos uma variável $b \in \{0, 1\}$. Isso se mostrará útil no momento de implementar os cálculos numericamente. Então, se $b = 0$, temos a regra de Hebb original. Caso $b = 1$, temos a versão modificada da regra de Hebb. O somatório $\sum_{i,j}$ representa uma soma sobre todos os pares de neurônios da rede e a fração $\frac{1}{2}$ indica que a interação entre os pares é simétrica, $J_{ij} = J_{ji}$.

Consideramos a rede neural como um sistema estocástico, em função da presença do ruído sináptico. Para simular a natureza estocástica do problema, consideraremos uma rede neural em equilíbrio com um reservatório de ruído térmico à temperatura T . Logo, a quantidade de relevância será a energia livre por partícula:

$$f(\beta, \mathbf{C}, \mathbf{X}) = - \lim_{N \rightarrow \infty} \frac{1}{\beta N} \log(Z(\mathbf{C}, \mathbf{X})). \quad (3.10)$$

Na equação 3.10, a energia livre por partícula é calculada sobre uma realização da matriz de conectividade \mathbf{C} e sobre uma realização $\mathbf{X} = \{\xi^\mu\}_{\mu=1 \dots p}$ dos p padrões. A matriz \mathbf{X} corresponde

à matriz dos p padrões inseridos pela regra de aprendizado. Não é interessante obter a energia livre e os parâmetros de ordem para apenas uma realização de \mathbf{C} e \mathbf{X} . Estamos interessados numa quantidade que seja independente da escolha das variáveis congeladas (*quenched*) \mathbf{C} e \mathbf{X} . Logo, aproveitando-se do fato que f é uma quantidade autome diante, ou seja [36]:

$$\frac{\langle f^2 \rangle - \langle f \rangle^2}{\langle f \rangle^2} \rightarrow 0, \quad (3.11)$$

calculamos então a média da energia livre sobre \mathbf{C} e \mathbf{X}

$$f = - \lim_{N \rightarrow \infty} \frac{1}{\beta N} \langle \log(Z(\mathbf{C}, \boldsymbol{\xi})) \rangle_{\mathbf{C}, \mathbf{X}}, \quad (3.12)$$

em que $\langle \dots \rangle_{\mathbf{X}} = \prod_{\mu, i} P(\xi_i^\mu) f(\{\xi_i^\mu\}_{i=1 \dots N}^{\mu=1 \dots p})$ corresponde à média sobre todas realizações desses p padrões e $\langle \dots \rangle_{\mathbf{C}} = \prod_{i < j} \sum_{c_{ij}=0,1} P(c_{ij}) f(c_{ij})$ equivale à média sobre todos os grafos gerados pela distribuição 3.7. A função de partição é dada por

$$Z = \sum_{\boldsymbol{\sigma}} \exp(-\beta H(\boldsymbol{\sigma}; \mathbf{C}, \mathbf{X})), \quad (3.13)$$

onde o traço $\sum_{\boldsymbol{\sigma}}$ é tomado sobre todas as 2^N possíveis configurações de spin. As variáveis $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_N)$ são conhecidas como *annealed* e possuem liberdade de flutuar de valor de acordo com o nível de ruído térmico $\beta = \frac{1}{T}$.

3.2.1 Método de Réplicas

Para tomar a média de $\langle \log(x) \rangle$ é necessário o uso do método de réplicas. Que consiste em escrever $\log(x)$ como:

$$\log(x) = \lim_{n \rightarrow 0} \frac{x^n - 1}{n}. \quad (3.14)$$

Isto pode ser derivado a partir de:

$$x^n = \exp(\log(x^n)) = \exp(n \log(x)), \quad (3.15)$$

quando $n \rightarrow 0$

$$x^n \approx 1 + n \log(x) + O \Rightarrow \log(x) = \lim_{n \rightarrow 0} \frac{x^n - 1}{n}. \quad (3.16)$$

Sabendo-se que:

$$\log(1 + ny) \approx ny, \quad (3.17)$$

tomando-se $y = \langle \log(x) \rangle$,

$$\log(1 + n \langle \log(x) \rangle) = n \langle \log(x) \rangle, \quad (3.18)$$

utilizando-se 3.14

$$\lim_{n \rightarrow 0} \frac{1}{n} \log(1 + n \frac{\langle x^n \rangle - 1}{n}) = \langle \log(x) \rangle, \quad (3.19)$$

logo:

$$\langle \log(x) \rangle = \lim_{n \rightarrow 0} \frac{1}{n} \log(\langle x^n \rangle). \quad (3.20)$$

A partir dessa identidade, podemos reescrever a energia livre na seguinte forma:

$$f = - \lim_{N \rightarrow \infty} \frac{1}{\beta N} \lim_{n \rightarrow 0} \frac{1}{n} \log \langle Z^n(\mathbf{C}, \mathbf{X}) \rangle_{\mathbf{C}, \mathbf{X}}. \quad (3.21)$$

Ao calcular $\langle Z^n(\mathbf{C}, \mathbf{X}) \rangle_{\mathbf{C}, \mathbf{X}}$, é necessário assumir que $n \in \mathbb{N}$, embora na identidade derivada acima, o limite seja tomado sobre $n \in \mathbb{R}$. Se promovemos n para o conjunto dos naturais, podemos tomar o produto da função de partição de forma:

$$\langle Z^n(\mathbf{C}, \mathbf{X}) \rangle_{\mathbf{C}, \mathbf{X}} = \langle \sum_{\sigma^1} \exp(-\beta H(\sigma^1; \mathbf{C}, \mathbf{X})) \cdots \sum_{\sigma^n} \exp(-\beta H(\sigma^n; \mathbf{C}, \mathbf{X})) \rangle_{\mathbf{C}, \mathbf{X}}, \quad (3.22)$$

porém, ao tomar o limite $n \rightarrow 0$, continua-se analiticamente a função para $n \in \mathbb{C}$. Usualmente, em trabalhos que fazem uso do método de réplicas, após algumas manipulações na média da função de partição replicada, chega-se a uma expressão da energia livre no formato:

$$f = - \lim_{N \rightarrow \infty} \frac{1}{\beta N} \lim_{n \rightarrow 0} \frac{1}{n} \log \int dP(\sigma) e^{N\phi(P(\sigma))}, \quad (3.23)$$

onde $P(\sigma)$ corresponde a alguma função ou parâmetro de ordem. Como no limite termodinâmico a integral é dominada pelo extremo da função $\phi(P(\sigma))$, faz-se uso do método de ponto de sela para calcular a integral. No entanto, é necessário permutar a ordem dos limites, tomando primeiramente $N \rightarrow \infty$. No artigo publicado por Van Hemmen e Palmer [37], esses garantem a comutatividade desses limites no modelo de (SK) [19]. Na derivação a seguir, faremos uso dessa troca de limites e utilizaremos o método derivado na referência [12]. Existe, no entanto, outro método [38], que obtem os mesmos resultados, porém tomando primeiro o limite de réplicas e depois o limite termodinâmico.

3.3 Cálculo das funções de ordem

Trabalhamos, inicialmente, calculando as médias sobre a função de partição replicada

$$\langle Z^n \rangle_{\mathbf{C}, \mathbf{X}} = \langle \prod_{\alpha=1}^n Z_\alpha \rangle = \langle \prod_{\alpha=1}^n \sum_{\sigma^\alpha} e^{-\beta H(\sigma^\alpha)} \rangle, \quad (3.24)$$

onde, por simplicidade removeremos os subíndices das médias. O Hamiltoniano replicado fica

$$H(\sigma^\alpha) = -\frac{1}{2} \sum_{i,j} \frac{c_{ij}}{c} \sum_{\mu=1}^p (\xi_i^\mu - ab)(\xi_j^\mu - ab) \sigma_i^\alpha \sigma_j^\alpha. \quad (3.25)$$

Tomando inicialmente a média sobre as realizações de \mathbf{C}

$$\begin{aligned}
 \langle Z^n \rangle &= \left\langle \sum_{\boldsymbol{\sigma}^1 \dots \boldsymbol{\sigma}^n} \exp \left[\frac{\beta}{2} \sum_{i \neq j} \frac{c_{ij}}{c} \sum_{\mu=1}^p (\xi_i^\mu - ab)(\xi_j^\mu - ab) \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha \right] \right\rangle \\
 &= \sum_{\boldsymbol{\sigma}^1 \dots \boldsymbol{\sigma}^n} \prod_{i < j} \sum_{c_{ij}=0,1} P(c_{ij}) \langle \exp \left[\beta \frac{c_{ij}}{c} (\boldsymbol{\gamma}_i \cdot \boldsymbol{\gamma}_j) \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha \right] \rangle \\
 &= \sum_{\boldsymbol{\sigma}^1 \dots \boldsymbol{\sigma}^n} \prod_{i < j} \left[1 + \frac{c}{N} \langle [e^{\frac{\beta}{c} (\boldsymbol{\gamma}_i \cdot \boldsymbol{\gamma}_j) \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha} - 1] \rangle \right],
 \end{aligned} \tag{3.26}$$

o vetor $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_N)$ pertence ao espaço de 2^N configurações de spin, $\boldsymbol{\xi}_i = (\xi_i^1, \dots, \xi_i^p)$ corresponde vetor no espaço de p padrões e $\boldsymbol{\gamma}_i = (\xi_i^1 - ab, \dots, \xi_i^p - ab)$ ao vetor $\boldsymbol{\xi}_i$ transformado para incluir a regra de Hebb modificada.

Como no limite termodinâmico $\frac{c}{N} \rightarrow 0$, fazemos uso da expansão $\exp(x) \approx 1 + x$. Termos de maior ordem foram desprezados. Reescrevemos o termo $\langle 1 + \frac{c}{N} [\dots] \rangle$ na forma:

$$\langle Z^n \rangle = \sum_{\boldsymbol{\sigma}^1 \dots \boldsymbol{\sigma}^n} \exp \left[\frac{c}{2N} \sum_{i,j} \langle [e^{\frac{\beta}{c} (\boldsymbol{\gamma}_i \cdot \boldsymbol{\gamma}_j) \sum_{\alpha} \sigma_i^\alpha \sigma_j^\alpha} - 1] \rangle \right]. \tag{3.27}$$

Em seguida, removemos as variáveis σ_i^α de dentro da exponencial interna fazendo uso da identidade [11, 34]:

$$\sum_{\boldsymbol{\sigma}} \left[\prod_{\alpha=1}^n \delta_{\sigma_\alpha, \sigma_i^\alpha} \right] = \sum_{\boldsymbol{\sigma}} \delta_{\boldsymbol{\sigma}, \boldsymbol{\sigma}_i} = 1, \tag{3.28}$$

onde $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)$ é um vetor de spin no espaço de n réplicas e $\boldsymbol{\sigma}_i$ equivale a alguma configuração nesse espaço para o i -ésimo spin. A função de partição replicada fica:

$$\langle Z^n \rangle = \prod_{\alpha=1}^n \sum_{\boldsymbol{\sigma}^\alpha} \exp \left[\frac{c}{2N} \sum_{i,j} \langle \sum_{\boldsymbol{\sigma}, \boldsymbol{\sigma}'} [e^{\frac{\beta}{c} (\boldsymbol{\gamma}_i \cdot \boldsymbol{\gamma}_j) \sum_{\alpha} \sigma_\alpha \sigma'_\alpha} - 1] \delta_{\boldsymbol{\sigma}, \boldsymbol{\sigma}_i} \delta_{\boldsymbol{\sigma}', \boldsymbol{\sigma}_j} \rangle \right]. \tag{3.29}$$

O traço sobre os spins pode ser reescrito na forma:

$$\prod_{\alpha=1}^n \sum_{\boldsymbol{\sigma}^\alpha} = \prod_{i=1}^N \sum_{\boldsymbol{\sigma}_i}, \tag{3.30}$$

com $\boldsymbol{\sigma}_i = (\sigma_i^1, \dots, \sigma_i^n)$. Logo,

$$\langle Z^n \rangle = \prod_{i=1}^N \sum_{\boldsymbol{\sigma}_i} \exp \left[\frac{c}{2N} \sum_{i,j} \langle \sum_{\boldsymbol{\sigma}, \boldsymbol{\sigma}'} [e^{\frac{\beta}{c} (\boldsymbol{\gamma}_i \cdot \boldsymbol{\gamma}_j) (\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}')} - 1] \delta_{\boldsymbol{\sigma}, \boldsymbol{\sigma}_i} \delta_{\boldsymbol{\sigma}', \boldsymbol{\sigma}_j} \rangle \right]. \tag{3.31}$$

Para tomar a média sobre os padrões $\langle \dots \rangle_{\boldsymbol{\xi}_i, \boldsymbol{\xi}_j}$, faremos uso do conceito de sub-redes [39]. A sub-rede é definida como:

$$I_{\boldsymbol{\xi}} = \{i | \boldsymbol{\xi}_i = \boldsymbol{\xi}\}. \tag{3.32}$$

Isto significa que, caso ξ_i for um vetor de p padrões que corresponde ao vetor ξ , ele pertence à sub-rede ξ e I_ξ representa o conjunto de todos os neurônios que pertencem a esse sub-rede, onde existem 2^p sub-redes. O módulo da sub-rede $|I_\xi|$ mede o número de sítios que pertencem a sub-rede ξ e a probabilidade de um sítio pertencer a esta sub-rede é:

$$p_\xi = \lim_{N \rightarrow \infty} \frac{|I_\xi|}{N}. \quad (3.33)$$

Médias de quantidades sobre as sub-redes são abreviadas como $\sum_\xi f(\xi) = \langle f(\xi) \rangle_\xi$.

Um exemplo do conceito de sub-rede para o caso de $p = 2$. O número de possíveis sub-redes é $2^2 = 4$, que são:

$$\begin{aligned} \xi^1 &= (1, 1) \\ \xi^2 &= (-1, 1) \\ \xi^3 &= (1, -1) \\ \xi^4 &= (-1, -1). \end{aligned}$$

Seja i um spin da rede, se $(\xi_i^1 = 1, \xi_i^2 = 1)$, esse spin pertence à sub-rede 1. Reescrevendo o somatório $\sum_{i=1}^N = \sum_\xi \sum_{i \in I_\xi}$, de forma a agrupar os sítios que pertencem a mesma sub-rede e inserindo adequadamente $N^{-1} N \frac{|I_\xi|}{|I_\xi|} = 1$ na equação 3.31, temos:

$$\langle Z^n \rangle = \prod_{i=1}^N \sum_{\sigma_i} \exp \left[\frac{Nc}{2} \sum_{\xi, \xi'} \sum_{\sigma, \sigma'} [e^{\frac{\beta}{c} (\gamma \cdot \gamma') (\sigma \cdot \sigma')} - 1] \sum_{i \in I_\xi} \frac{|I_\xi|}{N |I_\xi|} \delta_{\sigma, \sigma_i} \sum_{j \in I_{\xi'}} \frac{|I_{\xi'}|}{N |I_{\xi'}|} \delta_{\sigma', \sigma_j} \right]. \quad (3.34)$$

Definindo a função de ordem

$$P_\xi(\sigma) = \frac{1}{|I_\xi|} \sum_{i \in I_\xi} \delta_{\sigma, \sigma_i}, \quad (3.35)$$

que representa a distribuição de probabilidade de um vetor de spin de réplica σ na sub-rede ξ , ou seja, é a fração de vetores de spins de réplica σ_i que pertence ao conjunto da sub-rede I_ξ estão no estado σ .

Utilizando a seguinte identidade

$$1 = \int_{-\infty}^{\infty} \prod_{\xi, \sigma} \left\{ dP_\xi(\sigma) \delta \left[P_\xi(\sigma) - \frac{1}{|I_\xi|} \sum_{i \in I_\xi} \delta_{\sigma, \sigma_i} \right] \right\} \quad (3.36)$$

e escrevendo a δ de Dirac em sua representação integral

$$1 = \int \left[\prod_{\xi, \sigma} dP_\xi(\sigma) d\hat{P}_\xi(\sigma) \right] \exp \left\{ i \sum_{\xi, \sigma} \hat{P}_\xi(\sigma) \left[P_\xi(\sigma) - \frac{1}{|I_\xi|} \sum_{i \in I_\xi} \delta_{\sigma, \sigma_i} \right] \right\}, \quad (3.37)$$

é possível escrever a função de partição em termos das funções de ordem $[P_\xi(\sigma), \hat{P}_\xi(\sigma)]$

$$\begin{aligned} \langle Z^n \rangle = & \prod_{i=1}^N \sum_{\sigma_i} \int \left[\prod_{\xi, \sigma} dP_{\xi}(\sigma) d\hat{P}_{\xi}(\sigma) \right] \exp \left\{ i \sum_{\xi, \sigma} \hat{P}_{\xi}(\sigma) P_{\xi}(\sigma) - i \sum_{\xi, i \in I_{\xi}} \sum_{\sigma} \frac{1}{|I_{\xi}|} \hat{P}_{\xi}(\sigma) \delta_{\sigma, \sigma_i} \right. \\ & \left. + \sum_{\xi, \xi'} \sum_{\sigma, \sigma'} \frac{Nc}{2} [e^{\frac{\beta}{c}(\gamma \cdot \gamma')(\sigma \cdot \sigma')} - 1] p_{\xi} p_{\xi'} P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\}. \end{aligned} \quad (3.38)$$

Como resultado dessas manipulações, o traço sobre as configurações de spin atua somente no seguinte termo:

$$\prod_{i=1}^N \sum_{\sigma_i} \exp \left\{ -i \sum_{\xi, i \in I_{\xi}} \sum_{\sigma} \frac{1}{|I_{\xi}|} \hat{P}_{\xi}(\sigma) \delta_{\sigma, \sigma_i} \right\}. \quad (3.39)$$

Aplicando a transformação $\hat{P}_{\xi}(\sigma) \rightarrow |I_{\xi}| \hat{P}_{\xi}(\sigma)$ e fazendo um pouco de álgebra obtemos:

$$\prod_{i=1}^N \sum_{\sigma_i} \exp \left\{ -i \sum_{\xi, i \in I_{\xi}} \sum_{\sigma} \frac{1}{|I_{\xi}|} \hat{P}_{\xi}(\sigma) \delta_{\sigma, \sigma_i} \right\} = \exp \left\{ N \left\langle \log \left[\sum_{\sigma} e^{-i \hat{P}_{\xi}(\sigma)} \right] \right\rangle_{\xi} \right\}. \quad (3.40)$$

Substituindo na energia livre, temos:

$$f = - \lim_{n \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{\beta N n} \log \int \left[\prod_{\xi, \sigma} dP_{\xi}(\sigma) d\hat{P}_{\xi}(\sigma) \right] \exp[N \Psi(P_{\xi}(\sigma), \hat{P}_{\xi}(\sigma))], \quad (3.41)$$

sendo que

$$\begin{aligned} \Psi(P_{\xi}(\sigma), \hat{P}_{\xi}(\sigma)) = & \left\{ \left\langle i \sum_{\sigma} \hat{P}_{\xi}(\sigma) P_{\xi}(\sigma) \right\rangle_{\xi} \right. \\ & + \left\langle \left\langle \sum_{\sigma, \sigma'} \frac{c}{2} [e^{\frac{\beta}{c}(\gamma \cdot \gamma')(\sigma \cdot \sigma')} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \\ & \left. + \left\langle \log \left[\sum_{\sigma} e^{-i \hat{P}_{\xi}(\sigma)} \right] \right\rangle_{\xi} \right\}. \end{aligned} \quad (3.42)$$

No limite termodinâmico, $N \rightarrow \infty$, a integral da equação 3.41 é dominada pelo extremo de $\Psi(P_{\xi}(\sigma), \hat{P}_{\xi}(\sigma))$ com relação à $P_{\xi}(\sigma), \hat{P}_{\xi}(\sigma)$. Assumindo que é possível trocar a ordem dos limites de réplica e termodinâmico, a energia livre resulta em:

$$\begin{aligned} f = & - \lim_{n \rightarrow 0} \frac{1}{\beta n} \text{extr}_{\{P_{\xi}(\sigma), \hat{P}_{\xi}(\sigma)\}} \left\{ i \left\langle \sum_{\sigma} \hat{P}_{\xi}(\sigma) P_{\xi}(\sigma) \right\rangle_{\xi} \right. \\ & + \left\langle \left\langle \sum_{\sigma, \sigma'} \frac{c}{2} [e^{\frac{\beta}{c}(\gamma \cdot \gamma')(\sigma \cdot \sigma')} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \\ & \left. + \left\langle \log \left[\sum_{\sigma} e^{-i \hat{P}_{\xi}(\sigma)} \right] \right\rangle_{\xi} \right\}. \end{aligned} \quad (3.43)$$

Derivando f com relação a $\hat{P}_{\xi}(\sigma)$ obtemos:

$$P_{\xi}(\sigma) = \frac{\exp[-i \hat{P}_{\xi}(\sigma)]}{\sum_{\sigma} \exp[-i \hat{P}_{\xi}(\sigma)]}. \quad (3.44)$$

Derivando f com relação a $P_{\xi}(\sigma)$

$$\hat{P}_{\xi}(\sigma) = ic \left\langle \sum_{\sigma'} [\exp(\frac{\beta}{c}(\gamma \cdot \gamma')(\sigma \cdot \sigma')) - 1] P_{\xi'}(\sigma') \right\rangle_{\xi}, \quad (3.45)$$

onde utilizamos que

$$\frac{\delta P_{\xi}(\sigma)}{\delta P_{\xi'}(\sigma')} = \delta(\sigma - \sigma') \delta(\xi - \xi'). \quad (3.46)$$

Substituindo 3.45 em 3.44, obtemos uma equação para $P_{\xi}(\sigma)$:

$$P_{\xi}(\sigma) = \frac{\exp \left[c \left\langle \sum_{\sigma'} [\exp(\frac{\beta}{c}(\gamma \cdot \gamma')\sigma \cdot \sigma') - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} \right]}{\sum_{\sigma''} \exp \left[c \left\langle \sum_{\sigma'} [\exp(\frac{\beta}{c}(\gamma \cdot \gamma')\sigma' \cdot \sigma'') - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} \right]}. \quad (3.47)$$

Utilizando a equação 3.45, podemos eliminar $\hat{P}_{\xi}(\sigma)$ e reescrever f como:

$$f = \lim_{n \rightarrow 0} \frac{c}{2\beta n} \left\langle \left\langle \sum_{\sigma, \sigma'} [\exp(\frac{\beta}{c}(\gamma \cdot \gamma')\sigma \cdot \sigma') - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle_{\xi, \xi'} \right\rangle - \frac{1}{\beta n} \left\langle \log \left[\sum_{\sigma} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{(\frac{\beta}{c}(\gamma \cdot \gamma')\sigma \cdot \sigma')} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} \right\} \right] \right\rangle_{\xi}. \quad (3.48)$$

Para tomar o limite $n \rightarrow 0$ em f , necessitamos da função de ordem $P_{\xi}(\sigma)$, que deve ser obtida a partir da expressão 3.47. Pela equação 3.47, $P_{\xi}(\sigma)$ é obtida de forma autoconsistente. Então, necessitamos de um *ansatz*. Utilizamos nesse trabalho o *ansatz* de simetria de réplicas. A suposição de simetria de réplicas, no contexto de redes neurais atratoras [11, 13, 34, 35], implica que $P_{\xi}(\sigma)$ é invariante frente a permutação das componentes σ^{α} , logo tem-se que:

$$P_{\xi}(\sigma) = \int_{-\infty}^{\infty} dh W_{\xi}(h) \prod_{\alpha=1}^n \frac{e^{\beta h \sigma^{\alpha}}}{2 \cosh(\beta h)}. \quad (3.49)$$

Podemos motivar o *ansatz* da seguinte forma

$$P_{\xi}(\sigma) = \int_{-\infty}^{\infty} dh W_{\xi}(h) P_{\xi}(h|\sigma), \quad (3.50)$$

onde $P_{\xi}(h|\sigma)$ é a probabilidade de um certo vetor σ sob a ação de um campo externo local h_{ξ} por cada sub-rede. Como as réplicas são independentes, tem-se:

$$P_{\xi}(h|\sigma) = \prod_{\alpha=1}^n P_{\xi}(h|\sigma^{\alpha}). \quad (3.51)$$

A distribuição de probabilidade de equilíbrio de cada réplica sob ação de um campo médio $h = \beta^{-1} \arctan(\langle \sigma^{\alpha} \rangle_{\beta})$, fica:

$$P(h|\sigma^{\alpha}) = \frac{e^{\beta h \sigma^{\alpha}}}{\sum_{\sigma^{\alpha}=\pm 1} e^{\beta h \sigma^{\alpha}}} = \frac{e^{\beta h \sigma^{\alpha}}}{2 \cosh(\beta h)}, \quad (3.52)$$

como estamos tratando de uma rede aleatória, em que tomamos a média sobre todos os possíveis \mathcal{C} , a natureza do campo médio também depende de cada realização de \mathcal{C} . Logo, é necessário integrar sobre as possíveis distribuições de campos locais por sub-rede $W_{\xi}(h)$. Obtemos finalmente

$$P_{\xi}(\boldsymbol{\sigma}) = \int_{-\infty}^{\infty} dh W_{\xi}(h) \prod_{\alpha=1}^n \frac{e^{\beta h \sigma^{\alpha}}}{2 \cosh(\beta h)}, \quad (3.53)$$

o *ansatz* de simetria de réplicas.

Usando a suposição de simetria de réplicas, podemos calcular os observáveis:

A magnetização de sub-rede

$$m_{\xi} = \sum_{\boldsymbol{\sigma}} P_{\xi}(\boldsymbol{\sigma}) \sigma^{\alpha}, \quad (3.54)$$

o *overlap*

$$m^{\mu} = \langle \xi^{\mu} m_{\xi} \rangle_{\xi} \quad (3.55)$$

e o parâmetro de ordem de (EA)

$$q = \langle \sum_{\boldsymbol{\sigma}} P_{\xi}(\boldsymbol{\sigma}) \sigma^{\alpha} \sigma^{\gamma} \rangle_{\xi}. \quad (3.56)$$

A seguir aplicaremos o *ansatz*:

$$P_{\xi}(\boldsymbol{\sigma}) = \int_{-\infty}^{\infty} dh W_{\xi}(h) \frac{e^{\beta h \sum_{\alpha} \sigma^{\alpha}}}{[2 \cosh(\beta h)]^n}, \quad (3.57)$$

na equação 3.47 e faremos uma expansão, de forma a obter uma expressão para $W_{\xi}(h)$. Veremos que essa distribuição, de fato depende da conectividade média, e também da distribuição de grau do grafo aleatório no $\lim_{N \rightarrow \infty}$.

Desconsideramos o denominador de 3.47, pois esse é apenas uma constante de normalização. Expandindo a exponencial externa temos:

$$P_{\xi}(\boldsymbol{\sigma}) \sim \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \sum_{\boldsymbol{\sigma}'} \exp\left(\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}') \boldsymbol{\sigma} \cdot \boldsymbol{\sigma}'\right) P_{\xi'}(\boldsymbol{\sigma}') \right\rangle \right\rangle_{\xi'}^k. \quad (3.58)$$

Fazendo a mudança de variáveis $\xi' \rightarrow \zeta$, $\boldsymbol{\sigma}' \rightarrow \boldsymbol{\tau}$ e $\boldsymbol{\gamma}' \rightarrow \boldsymbol{\eta}$, onde $\boldsymbol{\eta} = (\zeta^1 - a, \dots, \zeta^p - a)$

$$P_{\xi}(\boldsymbol{\sigma}) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \sum_{\boldsymbol{\tau}^1 \dots \boldsymbol{\tau}^k} \left[P_{\zeta^1}(\boldsymbol{\tau}^1) \dots P_{\zeta^k}(\boldsymbol{\tau}^k) \right] \left[\prod_{l=1}^k \exp\left(\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \boldsymbol{\sigma} \cdot \boldsymbol{\tau}^l\right) \right] \right\rangle \right\rangle_{\zeta^1 \dots \zeta^k}. \quad (3.59)$$

Substituindo o *ansatz* de simetria de réplicas 3.57

$$P_{\xi}(\boldsymbol{\sigma}) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \int \left[\prod_{l=1}^k dh^l W_{\zeta^l}(h^l) \right] \left[\prod_{l=1}^k \frac{1}{[2 \cosh(\beta h^l)]^n} \right] \right. \right. \\ \left. \left. \times \sum_{\boldsymbol{\tau}^1 \dots \boldsymbol{\tau}^k} \prod_{l=1}^k \left[e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \sum_{\alpha} \sigma_{\alpha} \tau_{\alpha}^l + \beta h^l \sum_{\alpha} \tau_{\alpha}^l} \right] \right\rangle \right\rangle_{\zeta^1 \dots \zeta^k}. \quad (3.60)$$

Colocando num formato mais conveniente para tomar o traço sobre τ_α^l

$$P_\xi(\boldsymbol{\sigma}) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \int \left[\prod_{l=1}^k dh^l W_{\boldsymbol{\zeta}^l}(h^l) \right] \left[\prod_{l=1}^k \frac{1}{[2 \cosh(\beta h^l)]^n} \right] \right. \right. \\ \left. \left. \times \prod_{l=1}^k \exp \left[\sum_{\alpha} \log \sum_{\tau_\alpha^l = \pm 1} e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \sigma_\alpha \tau_\alpha^l + \beta h^l \tau_\alpha^l} \right] \right\rangle \right\rangle_{\boldsymbol{\zeta}^1 \dots \boldsymbol{\zeta}^k}. \quad (3.61)$$

Separando o último termo da direita da equação acima

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \prod_{l=1}^k \exp \left[\sum_{\alpha} \log \sum_{\tau_\alpha^l = \pm 1} e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \sigma_\alpha \tau_\alpha^l + \beta h^l \tau_\alpha^l} \right]. \quad (3.62)$$

Tomando o traço sobre τ_α^l

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \prod_{l=1}^k \exp \left[\sum_{\alpha} \log [2 \cosh(\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \sigma_\alpha + \beta h^l)] \right]. \quad (3.63)$$

Utilizando $\sum_{\sigma = \pm 1} \delta_{\sigma, \sigma_\alpha}$ para remover σ_α de dentro do cosh

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \prod_{l=1}^k \exp \left[\sum_{\alpha} \sum_{\sigma = \pm 1} \delta_{\sigma, \sigma_\alpha} \log [2 \cosh(\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \sigma + \beta h^l)] \right]. \quad (3.64)$$

Reescrevendo a delta como

$$\delta_{\sigma, \sigma_\alpha} = \frac{1}{2} (1 + \sigma \sigma_\alpha), \quad (3.65)$$

temos

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \prod_{l=1}^k \exp \left[\sum_{\alpha} \sum_{\sigma = \pm 1} \frac{1}{2} (1 + \sigma \sigma_\alpha) \log [2 \cosh(\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l) \sigma + \beta h^l)] \right]. \quad (3.66)$$

Tomando o traço $\sum_{\sigma = \pm 1}$

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \prod_{l=1}^k \exp \left[\left(\frac{1}{2} n + \sum_{\alpha} \frac{1}{2} \sigma_\alpha \right) \log [2 \cosh(\beta h^l + \frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))] \right. \\ \left. + \left(\frac{1}{2} n - \sum_{\alpha} \frac{1}{2} \sigma_\alpha \right) \log [2 \cosh(\beta h^l - \frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))] \right]. \quad (3.67)$$

No lim $n \rightarrow 0$, pode-se reescrever

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \prod_{l=1}^k \exp \left[\sum_{\alpha} \frac{1}{2} \sigma_\alpha \log \frac{2 \cosh(\beta h^l + \frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))}{2 \cosh(\beta h^l - \frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))} \right]. \quad (3.68)$$

O termo de dentro do log pode ser reescrito como

$$\frac{2 \cosh(\beta h^l + \frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))}{2 \cosh(\beta h^l - \frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))} = \frac{1 + \tanh(\beta h^l) \tanh(\frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))}{1 - \tanh(\beta h^l) \tanh(\frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))}, \quad (3.69)$$

fazendo uso da identidade

$$\cosh(x + \lambda y) = [1 + \lambda \tanh(x) \tanh(y)] \cosh(x) \cosh(y) \quad (3.70)$$

e utilizando também a seguinte identidade

$$\operatorname{atanh}(x) = \frac{1}{2} \log\left(\frac{1+x}{1-x}\right), \quad (3.71)$$

tem-se como resultado

$$I(\boldsymbol{\sigma}, \mathbf{h}) = \exp \left[\left(\sum_{\alpha} \sigma_{\alpha} \right) \sum_{l=1}^k \operatorname{atanh}[\tanh(\beta h^l) \tanh(\frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))] \right]. \quad (3.72)$$

Colocando novamente na equação 3.61, ficamos com:

$$P_{\boldsymbol{\xi}}(\boldsymbol{\sigma}) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \int [\prod_{l=1}^k dh^l W_{\zeta^l}(h^l)] \right. \right. \\ \left. \left. \times \exp \left[\left(\sum_{\alpha} \sigma_{\alpha} \right) \sum_{l=1}^k \operatorname{atanh}[\tanh(\beta h^l) \tanh(\frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))] \right] \right\rangle \right\rangle_{\zeta^1 \dots \zeta^k}. \quad (3.73)$$

Reescrevendo a equação acima usando a delta de Dirac

$$P_{\boldsymbol{\xi}}(\boldsymbol{\sigma}) = \int dx \left\{ \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \int [\prod_{l=1}^k dh^l W_{\zeta^l}(h^l)] \right. \right. \right. \\ \left. \left. \times \delta \left[x - \beta^{-1} \sum_{l=1}^k \operatorname{atanh}[\tanh(\beta h^l) \tanh(\frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))] \right] \right\rangle \right\rangle_{\zeta^1 \dots \zeta^k} \left. \right\} e^{x\beta \sum_{\alpha} \sigma_{\alpha}}. \quad (3.74)$$

Como no lim $n \rightarrow 0$, $P_{\boldsymbol{\xi}}(\boldsymbol{\sigma}) = \int dh W_{\boldsymbol{\xi}}(h) e^{\beta h \sum_{\alpha} \sigma_{\alpha}}$, identifica-se $W_{\boldsymbol{\xi}}(h)$ como

$$W_{\boldsymbol{\xi}}(h) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \left\langle \int [\prod_{l=1}^k dh^l W_{\zeta^l}(h^l)] \right. \right. \\ \left. \left. \times \delta \left[h - \beta^{-1} \sum_{l=1}^k \operatorname{atanh}[\tanh(\beta h^l) \tanh(\frac{\beta}{c}(\boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l))] \right] \right\rangle \right\rangle_{\zeta^1 \dots \zeta^k}. \quad (3.75)$$

Analisando a equação 3.75, vemos que a distribuição de campos locais é uma relação autoconsistente que depende da média sobre todas as sub-redes. Nessa expressão observa-se, também, a distribuição Poissoniana $P(k) = \frac{c^k e^{-c}}{k!}$, a qual está associada à distribuição de grau de um sítio de um grafo do tipo (ER) no limite $N \rightarrow \infty$.

3.3.1 Campo com vínculo

Nessa seção, adicionaremos o seguinte vínculo $\sum_i \langle \sigma_i \rangle_\beta = Na$ sobre a atividade dos neurônios da rede. Definindo $\gamma = ((\xi^\mu - a), \dots, (\xi^p - a))$ e $\sigma = (\sigma^1, \dots, \sigma^n)$, tomamos novamente a energia livre após o cálculo da integral pelo método do ponto de sela

$$\begin{aligned}
 f = & - \lim_{n \rightarrow 0} \frac{1}{\beta n} \text{extr}_{\{P_\xi(\sigma), \hat{P}_\xi(\sigma)\}} \left\{ i \left\langle \sum_{\sigma} \hat{P}_\xi(\sigma) P_\xi(\sigma) \right\rangle_{\xi} \right. \\
 & + \left\langle \left\langle \sum_{\sigma, \sigma'} \frac{c}{2} [e^{\frac{\beta}{c} (\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_\xi(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \\
 & \left. + \left\langle \log \left[\sum_{\sigma} e^{-i \hat{P}_\xi(\sigma)} \right] \right\rangle_{\xi} \right\}. \tag{3.76}
 \end{aligned}$$

Calculamos novamente $\delta \hat{P}_\xi(\sigma)$ e obtemos:

$$P_\xi(\sigma) = \frac{e^{-i \hat{P}_\xi(\sigma)}}{\sum_{\sigma} e^{-i \hat{P}_\xi(\sigma)}}. \tag{3.77}$$

No entanto, reescrevemos essa equação

$$\begin{aligned}
 \log P_\xi(\sigma) &= -i \hat{P}_\xi(\sigma) - \log \sum_{\sigma} e^{-i \hat{P}_\xi(\sigma)} \\
 i \hat{P}_\xi(\sigma) &= -\log P_\xi(\sigma) - \log \sum_{\sigma} e^{-i \hat{P}_\xi(\sigma)} \tag{3.78}
 \end{aligned}$$

e substituímos em 3.76. Após algumas manipulações e impondo que a distribuição seja normalizada $\sum_{\sigma} P_\xi(\sigma) = 1$, obtemos:

$$\begin{aligned}
 f = & - \lim_{n \rightarrow 0} \frac{1}{\beta n} \text{extr}_{\{P_\xi(\sigma)\}} \left\{ \left\langle \left\langle \sum_{\sigma, \sigma'} \frac{c}{2} [e^{\frac{\beta}{c} (\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_\xi(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \right. \\
 & \left. - \left\langle \sum_{\sigma} P_\xi(\sigma) \log P_\xi(\sigma) \right\rangle_{\xi} \right\} \tag{3.79}
 \end{aligned}$$

O próximo passo para obter a energia livre é extremar o funcional

$$\begin{aligned}
 \phi(P_\xi(\sigma)) &= \frac{c}{2} \left\langle \left\langle \sum_{\sigma, \sigma'} [e^{\frac{\beta}{c} (\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_\xi(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \\
 & - \left\langle \sum_{\sigma} P_\xi(\sigma) \log P_\xi(\sigma) \right\rangle_{\xi}. \tag{3.80}
 \end{aligned}$$

No entanto, queremos adicionar o seguinte vínculo

$$\left\langle \sum_{\sigma} P_{\xi}(\sigma) \sigma^{\alpha} \right\rangle_{\xi} = a. \quad (3.81)$$

Por conveniência, utilizaremos

$$\sum_{\alpha} \left\langle \sum_{\sigma} P_{\xi}(\sigma) \sigma^{\alpha} \right\rangle_{\xi} = na, \quad (3.82)$$

que corresponde ao mesmo que $\sum_{\alpha} \sigma^{\alpha} = na$. Desejamos garantir também o vínculo $\sum_{\sigma} P_{\xi}(\sigma) = 1$.

Para tanto, ao invés de extremizarmos $\phi(P_{\xi}(\sigma))$, introduzimos o funcional

$$L(P_{\xi}(\sigma); \lambda_1, \lambda_2) = \phi(P_{\xi}(\sigma)) - \lambda_1 \left(\sum_{\alpha} \left\langle \sum_{\sigma} P_{\xi}(\sigma) \sigma^{\alpha} \right\rangle_{\xi} - na \right) - \lambda_2 \left(\sum_{\sigma} P_{\xi}(\sigma) - 1 \right). \quad (3.83)$$

Extremando $L(P_{\xi}(\sigma); \lambda_1, \lambda_2)$ com relação a $P_{\xi}(\sigma)$, temos

$$\begin{aligned} \frac{\partial \phi}{\partial P_{\xi}(\sigma)} &= \lambda_1 \sum_{\alpha} \sigma^{\alpha} + \lambda_2, \\ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \log P_{\xi}(\sigma) - 1 &= \lambda_1 \sum_{\alpha} \sigma^{\alpha} + \lambda_2, \end{aligned}$$

assim obtemos

$$P_{\xi}(\sigma) = \frac{\exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \lambda_1 \sum_{\alpha} \sigma^{\alpha} \right\}}{e^{\lambda_2 + 1}}. \quad (3.84)$$

Tomando o traço sobre ambos lados da igualdade, encontramos λ_2

$$\lambda_2 = \log \left\{ \sum_{\sigma} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \lambda_1 \sum_{\alpha} \sigma^{\alpha} \right\} \right\} - 1, \quad (3.85)$$

resultando em

$$P_{\xi}(\sigma) = \frac{\exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \lambda_1 \sum_{\alpha} \sigma^{\alpha} \right\}}{\sum_{\sigma''} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma'' \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \lambda_1 \sum_{\alpha} \sigma''^{\alpha} \right\}}. \quad (3.86)$$

Usando o ansatz de simetria de réplica 3.57 e expandindo 3.86, de forma semelhante ao que foi feito na seção anterior, chegamos a uma expressão para distribuição de campos locais de sub-rede

$$W_{\xi}(h) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi^l}(h^l) \right] \times \delta \left[h - \beta^{-1} \lambda_1 - \beta^{-1} \sum_{l=1}^k \operatorname{atanh}[\tanh(\beta h^l) \tanh \left(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l \right)] \right] \right\rangle_{\xi^1 \dots \xi^k}. \quad (3.87)$$

Para inferir sobre a forma de λ_1 , substituímos o ansatz 3.57 na equação 3.81 e tomamos o $\lim_{n \rightarrow 0}$

$$\left\langle \int dh W_{\xi}(h) \tanh(\beta h) \right\rangle_{\xi} = a. \quad (3.88)$$

Caso os campos locais sejam nulos $W_{\xi^l}(h^l) = \delta(h^l) \implies W_{\xi}(h) = \delta(h - \beta^{-1} \lambda_1)$, logo

$$\left\langle \int dh \delta(h - \beta^{-1} \lambda_1) \tanh(\beta h) \right\rangle_{\xi} = a,$$

$$\tanh(\beta \beta^{-1} \lambda_1) = a.$$

Transformando $\beta^{-1} \lambda_1 = h_0$

$$\tanh(\beta h_0) = a,$$

$$h_0 = \beta^{-1} \operatorname{atanh}(a). \quad (3.89)$$

A distribuição de campos locais de sub-rede fica então,

$$W_{\xi}(h) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi^l}(h^l) \right] \times \delta \left[h - \beta^{-1} \operatorname{atanh}(a) - \beta^{-1} \sum_{l=1}^k \operatorname{atanh}[\tanh(\beta h^l) \tanh \left(\frac{\beta}{c} \sum_{\mu=1}^p (\xi^{\mu} - a)(\zeta^{\mu} - a) \right)] \right] \right\rangle_{\xi^1 \dots \xi^k}. \quad (3.90)$$

3.3.2 Energia livre por partícula com campo de vínculo

Uma vez que adicionamos um vínculo sobre a atividade dos neurônios, a função de energia livre deve apresentar uma modificação. Espera-se o aparecimento de um termo proporcional a $h_0 a$. Isso é análogo à adição de um vínculo de volume sobre o sistema, no caso vínculo sobre a atividade a , por uma pressão P , aqui sendo dada pelo campo de vínculo h_0 .

Tomando a energia livre

$$f = - \lim_{n \rightarrow 0} \frac{1}{\beta n} \left\{ \left\langle \left\langle \sum_{\sigma, \sigma'} \frac{c}{2} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} - \left\langle \sum_{\sigma} P_{\xi}(\sigma) \log P_{\xi}(\sigma) \right\rangle_{\xi} \right\} \quad (3.91)$$

e a equação da função de ordem na presença do campo de vínculo, derivada na seção anterior

$$P_{\xi}(\sigma) = \frac{\exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma^{\alpha} \right\}}{\sum_{\sigma''} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma'' \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma''^{\alpha} \right\}}, \quad (3.92)$$

definimos a constante de normalização de $P_{\xi}(\sigma)$ como:

$$\mathcal{N} = \sum_{\sigma''} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma'' \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma''^{\alpha} \right\}. \quad (3.93)$$

Reescrevendo a função de ordem

$$P_{\xi}(\sigma) = \frac{1}{\mathcal{N}} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma^{\alpha} \right\}, \quad (3.94)$$

tomando o log em ambos lados da igualdade

$$\log P_{\xi}(\sigma) = c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma^{\alpha} - \log \mathcal{N}, \quad (3.95)$$

substituindo $\log P_{\xi}(\sigma)$ na equação 3.91

$$f = - \lim_{n \rightarrow 0} \frac{1}{\beta n} \left\{ - \frac{c}{2} \left\langle \left\langle \sum_{\sigma, \sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} + \beta h_0 \sum_{\alpha} \left\langle \sum_{\sigma} P_{\xi}(\sigma) \sigma^{\alpha} \right\rangle_{\xi} + \left\langle \sum_{\sigma} P_{\xi}(\sigma) \log \mathcal{N} \right\rangle_{\xi} \right\} \quad (3.96)$$

e sabendo que $\sum_{\alpha} \langle \sum_{\sigma} P_{\xi}(\sigma) \sigma^{\alpha} \rangle_{\xi} = na$ e $\sum_{\sigma} P_{\xi}(\sigma) = 1$, tem-se

$$f = -\lim_{n \rightarrow 0} \frac{1}{\beta n} \left\{ -\frac{c}{2} \left\langle \left\langle \sum_{\sigma, \sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \right. \\ \left. + \beta h_0 na + \left\langle \log \left[\sum_{\sigma''} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma'' \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma''^{\alpha} \right\} \right] \right\rangle_{\xi} \right\}, \quad (3.97)$$

logo,

$$f = \lim_{n \rightarrow 0} \frac{c}{2\beta n} \left\langle \left\langle \sum_{\sigma, \sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'} \\ - \frac{1}{\beta n} \left\langle \log \left[\sum_{\sigma''} \exp \left\{ c \left\langle \sum_{\sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma'' \cdot \sigma'} - 1] P_{\xi'}(\sigma') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma''^{\alpha} \right\} \right] \right\rangle_{\xi} - h_0 a. \quad (3.98)$$

3.3.3 Expandindo a Energia livre

De posse da equação da energia livre 3.98 e fazendo uso do ansatz de simetria de réplicas, calculamos a energia livre numa forma que permite tomar o $\lim_{n \rightarrow 0}$. Separamos a energia livre em três termos:

$$f = f_1 + f_2 - h_0 a. \quad (3.99)$$

Trabalhando inicialmente com f_1

$$f_1 = \frac{c}{2\beta n} \left\langle \left\langle \sum_{\sigma, \sigma'} [e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma \cdot \sigma'} - 1] P_{\xi}(\sigma) P_{\xi'}(\sigma') \right\rangle \right\rangle_{\xi, \xi'}, \quad (3.100)$$

substituindo o ansatz de simetria de réplicas 3.20 e fazendo um pouco de algebra, obtém-se:

$$f_1 = \frac{c}{2\beta n} \left\langle \left\langle \int dh W_{\xi}(h) dh' W_{\xi'}(h') \right. \right. \\ \left. \left. \times \left\{ \left[\frac{1}{[4 \cosh(\beta h) \cosh(\beta h')]^n} \prod_{\alpha=1}^n \sum_{\sigma_{\alpha}, \sigma'_{\alpha}} e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma_{\alpha} \sigma'_{\alpha} + \beta h \sigma_{\alpha} + \beta h' \sigma'_{\alpha}} \right] - 1 \right\} \right\rangle \right\rangle_{\xi, \xi'}. \quad (3.101)$$

Pode-se demonstrar que o traço sobre as variáveis $\sum_{\sigma_{\alpha}, \sigma'_{\alpha}}$ fica:

$$\sum_{\sigma_{\alpha}, \sigma'_{\alpha}} e^{\frac{\beta}{c}(\gamma \cdot \gamma') \sigma_{\alpha} \sigma'_{\alpha} + \beta h \sigma_{\alpha} + \beta h' \sigma'_{\alpha}} = \\ 4 \left[1 + \tanh(\beta h) \tanh(\beta h') \tanh\left(\frac{\beta}{c}(\gamma \cdot \gamma')\right) \right] \cosh(\beta h) \cosh(\beta h') \cosh\left(\frac{\beta}{c}(\gamma \cdot \gamma')\right). \quad (3.102)$$

Aplicando esse resultado, obtemos

$$f_1 = \frac{c}{2\beta n} \left\langle \left\langle \int dh W_{\xi}(h) dh' W_{\xi'}(h') \right. \right. \quad (3.103)$$

$$\left. \left. \times \left\{ \cosh^n \left(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}' \right) \left[1 + \tanh(\beta h) \tanh(\beta h') \tanh \left(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}' \right) \right]^n - 1 \right\} \right\rangle_{\xi, \xi'}.$$

Utilizando a identidade 3.14, no lim $n \rightarrow 0$, resulta em

$$f_1 = \frac{c}{2\beta} \left\langle \left\langle \int dh W_{\xi}(h) dh' W_{\xi'}(h') \right. \right. \quad (3.104)$$

$$\left. \left. \times \log \left[\cosh \left(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}' \right) \left[1 + \tanh(\beta h) \tanh(\beta h') \tanh \left(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}' \right) \right] \right] \right\rangle_{\xi, \xi'}.$$

O termo f_2 é um pouco mais laborioso

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{\boldsymbol{\sigma}''} \exp \left\{ c \left\langle \sum_{\boldsymbol{\sigma}'} \left[e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}') \boldsymbol{\sigma}'' \cdot \boldsymbol{\sigma}'} - 1 \right] P_{\xi'}(\boldsymbol{\sigma}') \right\rangle_{\xi'} - \beta h_0 \sum_{\alpha} \sigma''^{\alpha} \right\} \right] \right\rangle_{\xi}. \quad (3.105)$$

Reescrevendo

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{\boldsymbol{\sigma}''} e^{-c} e^{c \left\langle \sum_{\boldsymbol{\sigma}'} e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}') \boldsymbol{\sigma}'' \cdot \boldsymbol{\sigma}'} P_{\xi'}(\boldsymbol{\sigma}') \right\rangle_{\xi'} e^{-\beta h_0 \sum_{\alpha} \sigma''^{\alpha}}} \right] \right\rangle_{\xi} \quad (3.106)$$

e usando a expansão da exponencial

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{\boldsymbol{\sigma}''} \sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \sum_{\boldsymbol{\sigma}'} e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}') \boldsymbol{\sigma}'' \cdot \boldsymbol{\sigma}'} P_{\xi'}(\boldsymbol{\sigma}') \right\rangle_{\xi'}^k e^{-\beta h_0 \sum_{\alpha} \sigma''^{\alpha}} \right] \right\rangle_{\xi}. \quad (3.107)$$

Trocando $\boldsymbol{\sigma}' \rightarrow \boldsymbol{\tau}$ e $\boldsymbol{\sigma}'' \rightarrow \boldsymbol{\sigma}$ e reescrevendo o termo $\langle \dots \rangle_{\xi'}^k$

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{\boldsymbol{\sigma}} \sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \prod_{\alpha=1}^n \prod_{l=1}^k \sum_{\tau_{\alpha}^l = \pm 1} e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}')_{\sigma_{\alpha} \tau_{\alpha}^l} P_{\xi'_l}(\boldsymbol{\tau}^l)} \right\rangle_{\xi'_1 \dots \xi'_k} e^{-\beta h_0 \sum_{\alpha} \sigma^{\alpha}} \right] \right\rangle_{\xi} \quad (3.108)$$

e substituindo o ansatz de simetria de réplicas 3.57

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{\boldsymbol{\sigma}} \sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi'_l}(h^l) \frac{1}{[2 \cosh(\beta h^l)]^n} \right. \right. \right. \quad (3.109)$$

$$\left. \left. \times \prod_{\alpha=1}^n \prod_{l=1}^k \sum_{\tau_{\alpha}^l = \pm 1} e^{\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}')_{\sigma_{\alpha} \tau_{\alpha}^l + \beta h^l \tau_{\alpha}^l} \right\rangle_{\xi'_1 \dots \xi'_k} e^{-\beta h_0 \sum_{\alpha} \sigma^{\alpha}} \right] \right\rangle_{\xi}.$$

Tomando o traço sobre τ_{α}^l , rearranjando a expressão utilizando a comutatividade da soma e levando em consideração que o termo $e^{-\sum_{\alpha} \beta h_0 \sigma^{\alpha}}$ é uma constante sobre as médias $\langle \dots \rangle_{\xi'_1 \dots \xi'_k}$

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi'_l}(h^l) \frac{1}{[2 \cosh(\beta h^l)]^n} \right. \right. \right. \quad (3.110)$$

$$\left. \left. \times \prod_{\alpha=1}^n \sum_{\sigma_{\alpha}} e^{-\beta h_0 \sigma^{\alpha}} \prod_{l=1}^k 2 \cosh \left[\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}')_{\sigma_{\alpha}} + \beta h^l \right] \right\rangle_{\xi'_1 \dots \xi'_k} \right] \right\rangle_{\xi}$$

Executando o traço \sum_{σ_α}

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi'_l}(h^l) \frac{1}{[2 \cosh(\beta h^l)]^n} \right] \right. \right. \right. \\ \left. \left. \left. \times \prod_{\alpha=1}^n \left\{ e^{-\beta h_0} \prod_{l=1}^k 2 \cosh \left[\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l) + \beta h^l \right] + e^{\beta h_0} \prod_{l=1}^k 2 \cosh \left[\frac{\beta}{c} (\boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l) - \beta h^l \right] \right\} \right] \right\rangle_{\xi'_1 \dots \xi'_k} \right\rangle_{\xi} \quad (3.111)$$

Utilizando a identidade 3.70 e simplificando a expressão, ficamos com

$$f_2 = -\frac{1}{\beta n} \left\langle \log \left[\sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi'_l}(h^l) \right] \right. \right. \\ \left. \left. \times \left\{ \sum_{\lambda=\pm} e^{-\lambda \beta h_0} \prod_{l=1}^k [1 + \lambda \tanh(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l) \tanh(\beta h^l)] \cosh(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l) \right\}^n \right] \right\rangle_{\xi'_1 \dots \xi'_k} \right\rangle_{\xi} \quad (3.112)$$

Tomando o limite $n \rightarrow 0$ e utilizando a identidade de réplicas 3.20,

$$f_2 = -\frac{1}{\beta} \sum_{k=0}^{\infty} e^{-c} \frac{c^k}{k!} \left\langle \left\langle \int \left[\prod_{l=1}^k dh^l W_{\xi'_l}(h^l) \right] \right. \right. \\ \left. \left. \times \log \left[\sum_{\lambda=\pm} e^{-\lambda \beta h_0} \prod_{l=1}^k [1 + \lambda \tanh(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l) \tanh(\beta h^l)] \cosh(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l) \right] \right\rangle_{\xi'_1 \dots \xi'_k} \right\rangle_{\xi}. \quad (3.113)$$

3.4 Cálculo das distribuições de campos locais em $T = 0$

Nessa seção, é demonstrado como calcular as distribuições de campos locais de sub-rede e os observáveis no caso de $T = 0$. Como $T = 0$, $\beta \rightarrow \infty$. Para encontrar a distribuição de campos locais por sub-rede é necessário expandir o campo efetivo

$$U = \beta^{-1} \sum_{l=1}^k \text{atanh}[\tanh(\beta h^l) \tanh(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)]. \quad (3.114)$$

Reescrevendo a expressão 3.114 como

$$U = \beta^{-1} \sum_{l=1}^k \frac{1}{2} \log \left[\frac{e^{\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l} + e^{-\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l}}{e^{\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l} + e^{-\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l}} \right], \quad (3.115)$$

e considerando cada possível caso de h^l e $\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l$:

$$h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l > 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l > 0 \quad (3.116)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &< 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &> 0 \end{aligned} \quad (3.117)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &> 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &< 0 \end{aligned} \quad (3.118)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &< 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &< 0 \end{aligned} \quad (3.119)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &= 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &> 0 \end{aligned} \quad (3.120)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &= 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &< 0 \end{aligned} \quad (3.121)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &> 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &= 0 \end{aligned} \quad (3.122)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &< 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &= 0 \end{aligned} \quad (3.123)$$

$$\begin{aligned} h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &= 0 \\ h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l &= 0 \end{aligned} \quad (3.124)$$

Como exemplo, calculamos o valor do campo efetivo 3.115 para o caso 3.116

$$U = \sum_{l=1}^k \frac{1}{2} \beta^{-1} \log \left[\frac{e^{\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l} (1 + e^{-2(\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)})}{e^{\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l} (1 + e^{-2(\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)})} \right] \quad (3.125)$$

$$U = \sum_{l=1}^k \frac{1}{2} \beta^{-1} \left[\log e^{\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l} (1 + e^{-2(\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)}) - \log e^{\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l} (1 + e^{-2(\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)}) \right] \quad (3.126)$$

$$U = \sum_{l=1}^k \frac{1}{2} \beta^{-1} \left[\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l - \beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l + \log (1 + e^{-2(\beta h^l + \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)}) - \log (1 + e^{-2(\beta h^l - \frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l)}) \right], \quad (3.127)$$

como $\beta \rightarrow \infty$ os dois últimos termos vão a zero, logo

$$U = \sum_{l=1}^k \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l. \quad (3.128)$$

Realizando esse processo para cada caso obtemos a seguinte expressão

$$U = \frac{1}{2} \sum_{l=1}^k (|h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l| - |h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l|). \quad (3.129)$$

Por meio da equação 3.129, podemos calcular a distribuições de sub-rede em $T = 0$

$$W_{\boldsymbol{\xi}}(h) = \sum_{k=0}^{\infty} \frac{c^k e^{-c}}{k!} \left\langle \int \left[\prod_{l=1}^k dh^l W_{\boldsymbol{\xi}^l}(h^l) \right] \times \delta \left[h - \frac{1}{2} \sum_{l=1}^k (|h^l + \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l| - |h^l - \frac{1}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\eta}^l|) \right] \right\rangle_{\boldsymbol{\xi}^1 \dots \boldsymbol{\xi}^k}. \quad (3.130)$$

Em seguida, calculam-se os observáveis. No limite $\lim_{\beta \rightarrow \infty}$ os observáveis ficam:

$$m^{\mu} = \langle \xi^{\mu} \int dh W_{\boldsymbol{\xi}}(h) \operatorname{sgn}(h) \rangle_{\boldsymbol{\xi}} \quad (3.131)$$

e

$$q = \langle \int dh W_{\boldsymbol{\xi}}(h) \operatorname{sgn}^2(h) \rangle_{\boldsymbol{\xi}} = 1. \quad (3.132)$$

No capítulo seguinte apresentamos o procedimento para calcular numericamente as distribuições de campos através do método de dinâmica de populações.

Capítulo 4

Resultados

Nesse capítulo, introduzimos um método numérico para calcular a distribuição de campos locais de sub-rede 3.75, a partir disso podemos calcular os observáveis relevantes m^μ e q em função dos parâmetros do sistema: temperatura T , número de padrões p , atividade dos padrões a e conectividade média c . De posse dos observáveis é possível traçar diagramas de fase nos planos (α, a) , (α, T) , (a, T) . No plano (α, a) , podemos estimar a capacidade da memória $\alpha(a)$, na ausência de ruído térmico. No plano (α, T) , verificamos a capacidade $\alpha(T)$ para um valor fixo de atividade. No plano (a, T) , investigamos para um valor fixo de α , qual o valor de atividade máxima $a(T)$, na presença de ruído térmico. Os diagramas de fase foram realizados para cada uma das seguintes regras de aprendizado:

Regra de Hebb original

A regra sináptica utilizada originalmente no modelo de Hopfield [3],

$$J_{ij} = \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu. \quad (4.1)$$

Regra de Hebb modificada

A modificação da regra de Hebb original para padrões com atividade finita, introduzido na referência [8],

$$J_{ij} = \sum_{\mu=1}^p (\xi_i^\mu - a)(\xi_j^\mu - a). \quad (4.2)$$

Regra de Hebb modificada com campo de vínculo

Regra de Hebb modificada incluindo um campo que vincula a atividade dos neurônios,

$$\frac{1}{N} \sum_{i=1}^N \sigma_i = a. \quad (4.3)$$

4.1 Cálculo da distribuição de campos locais de subrede

Para calcular os observáveis relevantes m^μ e q , é necessário calcular a distribuição de campos locais para cada sub-rede $W_\xi(h)$, através da equação 3.75. Este cálculo é feito

numericamente através de uma dinâmica de populações. A partir de uma condição inicial, cria-se uma população de campos $\{h_{\xi}^1 \dots h_{\xi}^{\mathcal{N}}\}$ para cada uma das 2^p sub-redes ξ . A variável h_{ξ} representa o campo local de uma determinada subrede ξ . Inicialmente, selecionamos uma sub-rede ξ que desejamos atualizar. Escolhemos um número k de uma distribuição Poissoniana com média c .

Selecionamos uma outra sub-rede ξ' com probabilidade:

$$p_{\xi'} = \prod_{\mu=1} \frac{1}{2} (1 + \xi'^{\mu} a), \quad (4.4)$$

onde a representa a atividade dos padrões. A seguir, seleciona-se aleatoriamente um dos \mathcal{N} elementos da população de campos $\{h_{\xi'}^1 \dots h_{\xi'}^{\mathcal{N}}\}$ e calcula-se a quantidade

$$U_l = \arctan[\tanh(\beta h_{\xi'}^l) \tanh(\frac{\beta}{c} \boldsymbol{\gamma} \cdot \boldsymbol{\gamma}'_l)], \quad (4.5)$$

onde $\boldsymbol{\gamma}' = (\xi'^1 - ab, \dots, \xi'^p - ab)$ corresponde à sub-rede sorteada e $\boldsymbol{\gamma} = (\xi^1 - ab, \dots, \xi^p - ab)$ à sub-rede inicialmente selecionada. Após k execuções calcula-se

$$U = \beta^{-1} \sum_{l=1}^k U_l \quad (4.6)$$

e atribui-se, aleatoriamente, a um dos \mathcal{N} elementos da população $\{h_{\xi}^1 \dots h_{\xi}^{\mathcal{N}}\}$ o valor calculado na equação 4.6. Esse processo é realizado novamente para cada sub-rede e iterado até todas as distribuições de campos atingirem um estado estacionário. Nas simulações realizadas, escolheu-se usualmente um número de $200\mathcal{N}$ iterações, mas essa escolha não é definitiva e pode sofrer ajustes. A condição inicial das populações de campos é escolhida de forma que o campo local de cada sub-rede corresponda a um dos padrões:

$$h_{\xi} = \xi^p. \quad (4.7)$$

Essa estratégia é adotada para iniciar o sistema mais próximo de uma das 2^p soluções de recuperação. Um exemplo de condição inicial das populações para o caso de $p = 2$ está esboçado na figura 4.1.

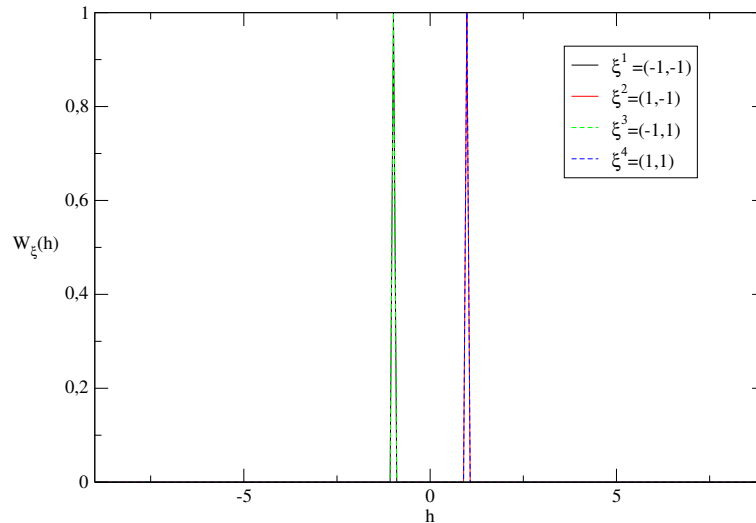


Figura 4.1: Distribuição de campos de sub-rede. A condição inicial escolhida é colocar os campos locais no primeiro padrão.

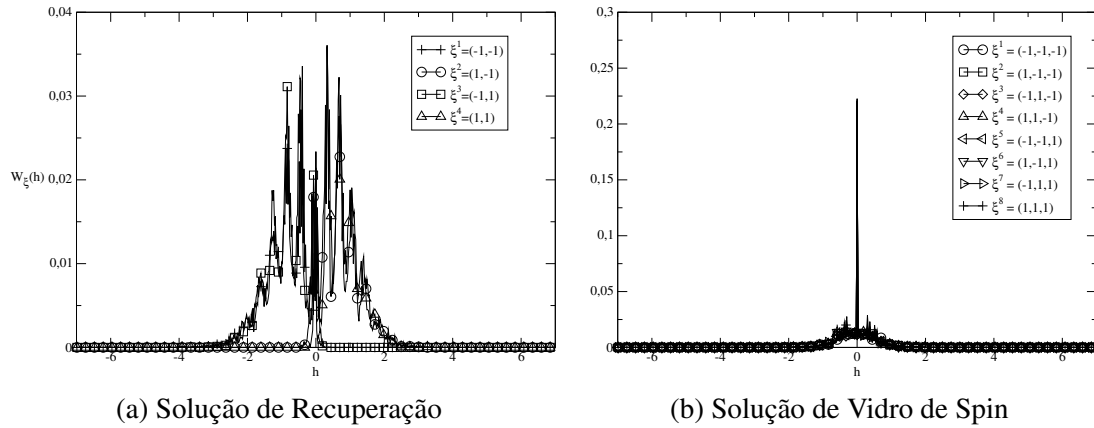


Figura 4.2: Distribuição de campos locais por sub-rede para uma atividade e temperaturas fixas em $a = 0.1$ e $T = 0.2$. Figura 4.2a corresponde a uma fase R com $p = 2$ e $c = 5$. Figura 4.2b corresponde a uma fase VS com $p = 3$ e $c = 2$.

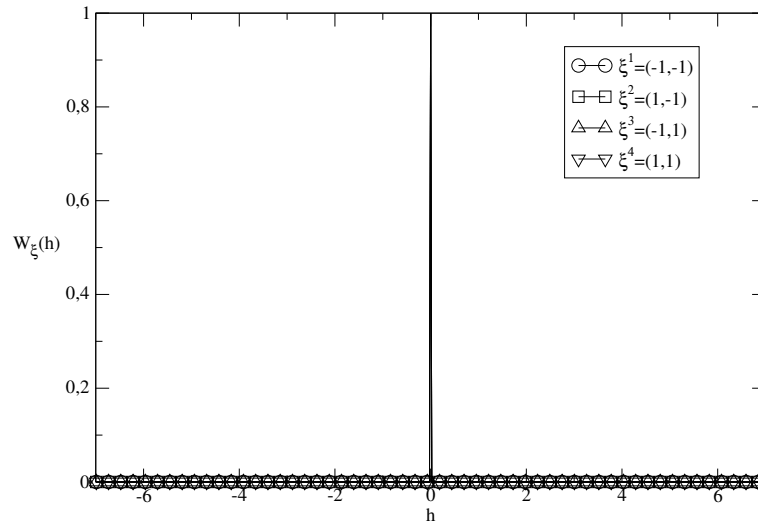


Figura 4.3: Solução paramagnética da distribuição de campos locais de sub-rede para uma atividade fixa em $a = 0.1$ e temperatura em $T = 1.1$ com $p = 2$ e $c = 5$.

Nas figuras 4.2a, 4.2b e 4.3, temos respectivamente o perfil da distribuição de campos locais por sub-rede em uma fase de recuperação de um padrão (R), fase de vidro de spin (VS) e fase paramagnética (P).

Para verificar qual tipo de fase o sistema se encontra calculamos os observáveis

$$m^\mu = \langle \xi^\mu \int dh W_\xi(h) \tanh(\beta h) \rangle_\xi \quad (4.8)$$

e

$$q = \langle \int dh W_\xi(h) \tanh^2(\beta h) \rangle_\xi. \quad (4.9)$$

O observável m^μ mede a correlação entre o estado dos neurônios da rede com um dos p padrões inseridos na memória, e o observável q é o parâmetro de ordem de vidro de spin.

Soluções do tipo $\mathbf{m} = (m^1, \dots, m^p) = (0, \dots, 0)$ e $q = 0$ caracterizam uma fase (P). A fase de vidro de spin (VS) é definida como $\mathbf{m} = (m^1, \dots, m^p) = (0, \dots, 0)$ e $q \neq 0$. Fases do tipo $\mathbf{m} = (m^1, \dots, m^p) \neq (0, \dots, 0)$ e $q \neq 0$ são divididas de acordo com a simetria de $\mathbf{m} = (m^1, \dots, m^p)$.

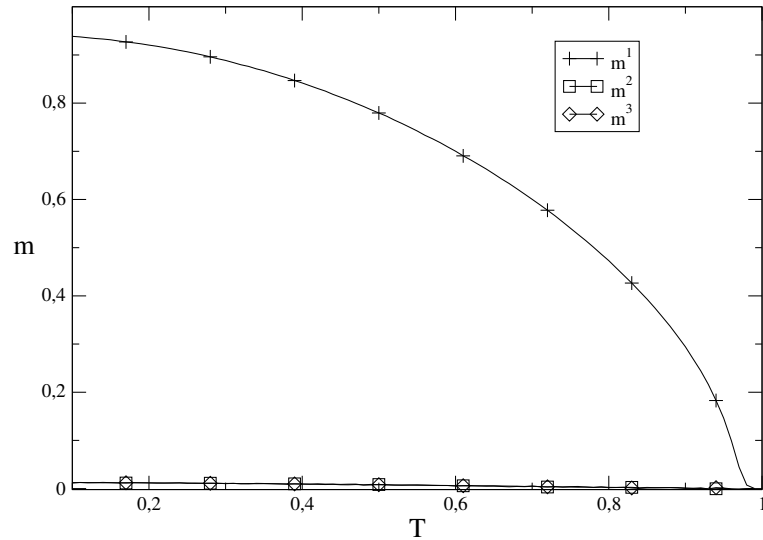


Figura 4.4: Transição R-P. As variáveis m^1 , m^2 e m^3 representam respectivamente os *overlaps* da rede com os padrões 1, 2 e 3.

Os estados mais interessantes são os de recuperação de apenas um padrão (R), por exemplo: $\mathbf{m} = (m^1, m^2, m^2, m^2, m^2, m^2)$, onde $|m^1| > |m^2|$. Apresentamos um exemplo da fase R na figura 4.4. No presente trabalho, esse estado seria o mesmo, caso se permutasse $m^1 \rightarrow m^2$ ou $m^1 \rightarrow m^3$. Ou seja, desde que *overlap* de um padrão se sobreponha em valor a todos outros, esse é considerado um estado de recuperação, não havendo distinção em que ordem os padrões foram recuperados. Estados da forma $\mathbf{m} = m(1, \dots, 1)$ são denominados simétricos (ou de simetria completa) (S). Nesses estados, o valor do *overlap* é igual para todos padrões, de forma que a rede não faz distinção entre as memórias. Uma analogia com esse estado seria a seguinte: Implementa-se numa memória três padrões, representados pelas letras A , B e C . Nesse estado de simetria completa, a memória não é capaz de detectar diferença entre os três caracteres. Na figura 4.5, mostramos exemplo de transição de fase R-S.

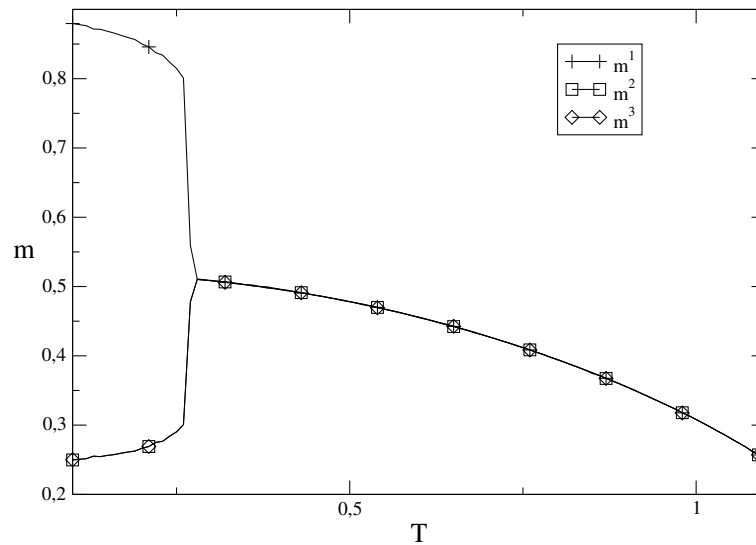


Figura 4.5: Transição descontínua R- S_3 .

Misturas assimétricas do tipo $\mathbf{m} = (m^1, m^1, m^1, m^2, m^2, m^2)$, e suas permutações sobre m_1 e m_2 , são representadas por S_{n_1, n_2} , n_1 corresponde ao número padrões com valor m^1 e n_2 ao número de m^2 padrões. Isso significa que a memória recupera $n = n_1 + n_2$ padrões e faz distinção sobre n_1 e n_2 . Na figura 4.6, mostramos um exemplo de uma transição fase R- $S_{2,2}$.

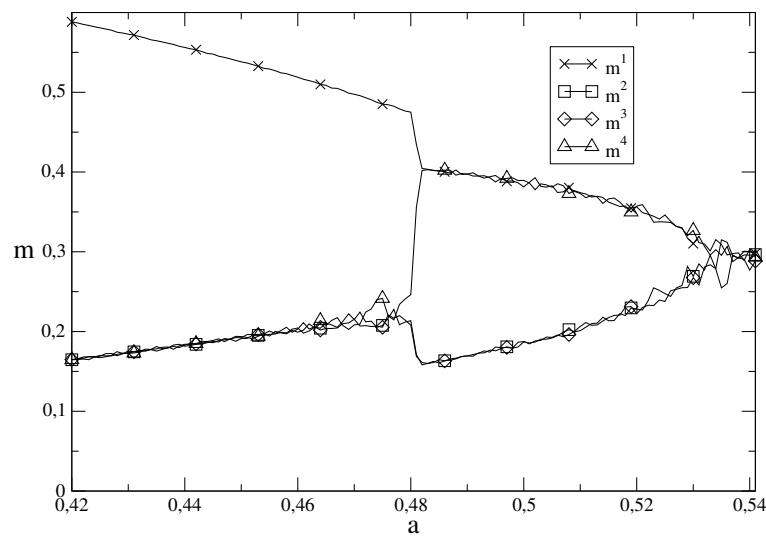
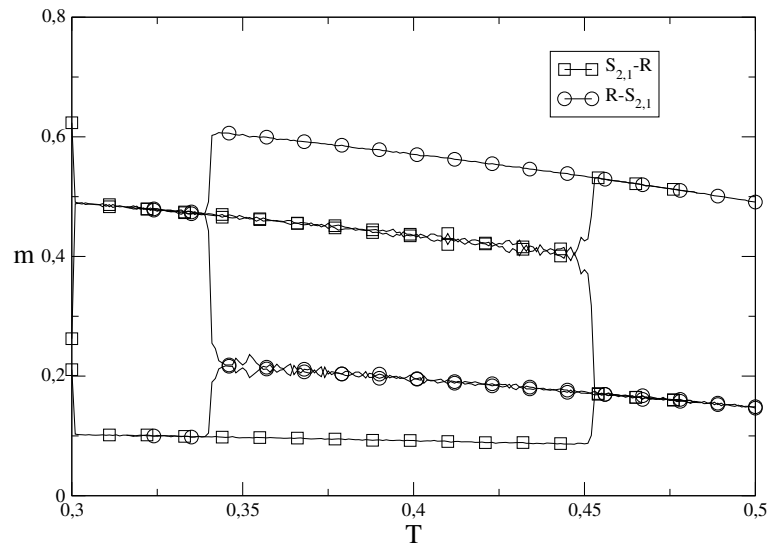


Figura 4.6: Transição de fase R para $S_{2,2}$.

Usualmente, em trabalhos de redes neurais, o interesse é focado na fase R e procura-se delimitar a fronteira dessa fase com outras. Nesse trabalho, a procura por pontos de transição de fase é feita a partir do cálculo dos parâmetros de ordem m^μ e q , sem fazer uso da energia livre. Se a transição é descontínua, procuramos a espinodal da fase R para outras fases. Uma espinodal é o limite em que uma determinada fase seja metaestável, numa transição descontínua. Um exemplo de espinodal é apresetado na figura 4.7.

Figura 4.7: Espinodais $S_{2,1}-R$ e $R-S_{2,1}$.

4.2 Comparando as regras de aprendizado

Nesta seção, comparamos as três regras de aprendizado nos planos (α, a) , (α, T) e (a, T) , de forma a mostrar as propriedades dessas regras. No plano (α, a) , como $T = 0$, o campo de vínculo desaparece. As linhas presentes nos diagramas dos planos (α, a) e (α, T) servem apenas para guiar os olhos; os pontos marcados nos diagramas desses planos representam as soluções calculadas, referentes aos parâmetros de ordem. Procurou-se trabalhar apenas em planos com conectividade média (c) fixa, seguindo as referências [11, 34].

4.2.1 Diagramas de fase no plano (α, a)

Primeiramente, analisando o plano (α, a) , comparamos as regras de Hebb e a versão modificada. Na figuras 4.8 e 4.9, traçamos curvas de m^μ vs. a em valores de $p = 2$ até $p = 7$, com as regras de Hebb e Hebb modificada. A regra de Hebb apresenta uma transição de fase contínua em $p = 2$, e transições descontínuas em $p = 3$ a $p = 7$. Uma vez na fase S, o comportamento de m^μ em função da atividade é linear. Pode-se observar pela figura 4.9, que regra de Hebb modificada possui apenas transições de fase contínuas. Em contraste com a regra de Hebb, na fase simétrica, a regra modificada apresenta um comportamento similar a uma função quadrática.

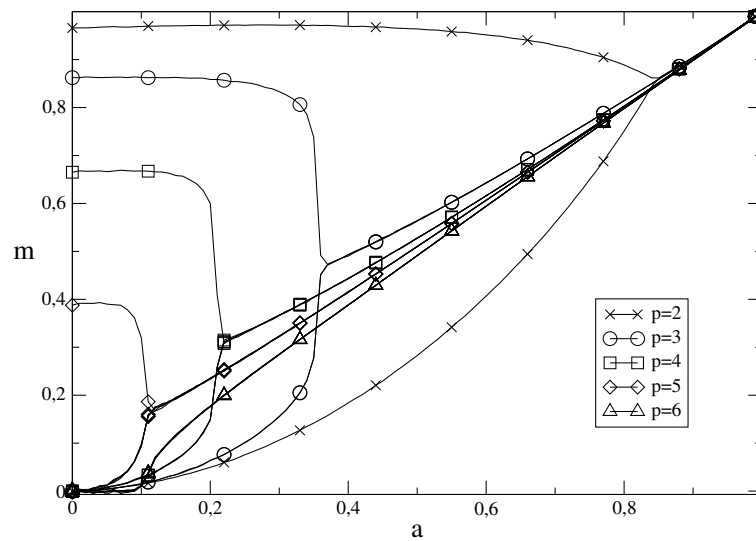


Figura 4.8: *Overlaps* m^μ vs. a de $p = 2$ a $p = 7$ com a regra de Hebb em $T = 0$ e $c = 7$.

A partir dessas medidas, traçamos os diagramas de fase, mostrados na figura 4.11. Em $p = 2$, a regra de Hebb opera melhor, com um valor de atividade crítica $a_c = 0,84$ na transição R-S. Porém, para $p \geq 3$, o valor a_c dessa regra começa a decair, e a partir de $p = 6$ ($\alpha \approx 0,857$), a regra de Hebb encontra-se numa fase VS. A regra de Hebb modificada apresenta uma pequena melhora no que diz respeito à atividade crítica para valores de $p \geq 3$, tendo valor máximo de $a_c = 0,42$ para $p = 3$ ($\alpha \approx 0,429$). Novamente em $p = 6$ ($\alpha \approx 0,857$), essa regra também encontra-se numa fase VS.

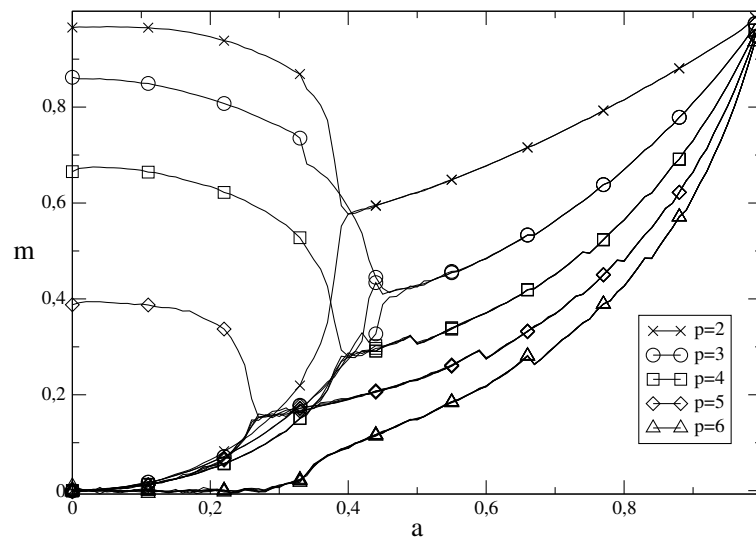


Figura 4.9: *Overlaps* m^μ vs. a de $p = 2$ a $p = 7$ com a regra de Hebb modificada em $T = 0$ e $c = 7$.

Realizando novamente medidas com regra de Hebb modificada, figura 4.10, porém com conectividade média $c = 10$, para valores de $p = 3$ ($\alpha = 0,3$), $p = 4$ ($\alpha = 0,4$) e $p = 5$ ($\alpha = 0,5$), verifica-se que o valor que possui o maior $a_c = 0,477$ na transição R-S, é $p = 4$ ($\alpha = 0,4$). Isso indica que o α que possui o maior a_c tende a diminuir à medida que c é elevado. No entanto, observou-se uma melhora no valor de a_c . Vale observar que não desenhamos o diagrama de

fases completo para $c = 10$, pois, à medida que nos aproximamos de $\alpha = 1 (p = c)$, o número de sub-redes escala na ordem de 2^{10} , e o tempo computacional para calcular os observáveis torna-se muito elevado.

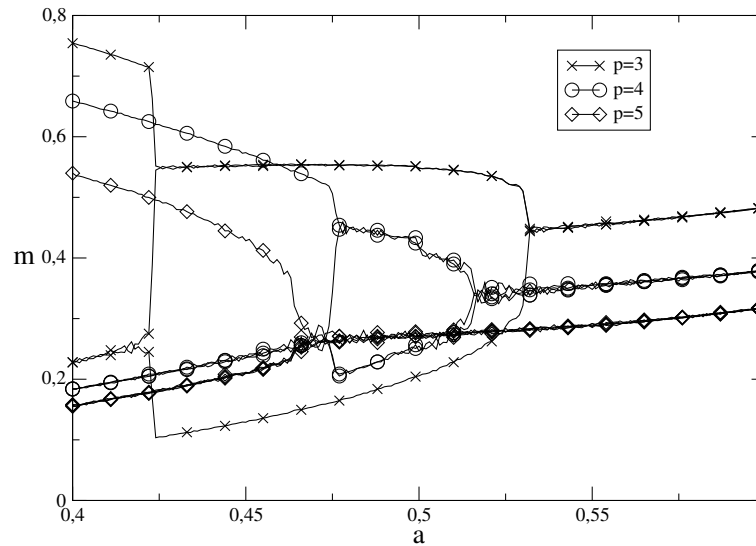


Figura 4.10: *Overlaps* m^μ vs. a para valores de $p = 3$, $p = 4$ e $p = 5$, com Regra de Hebb modificada em $T = 0$ e com $c = 10$.

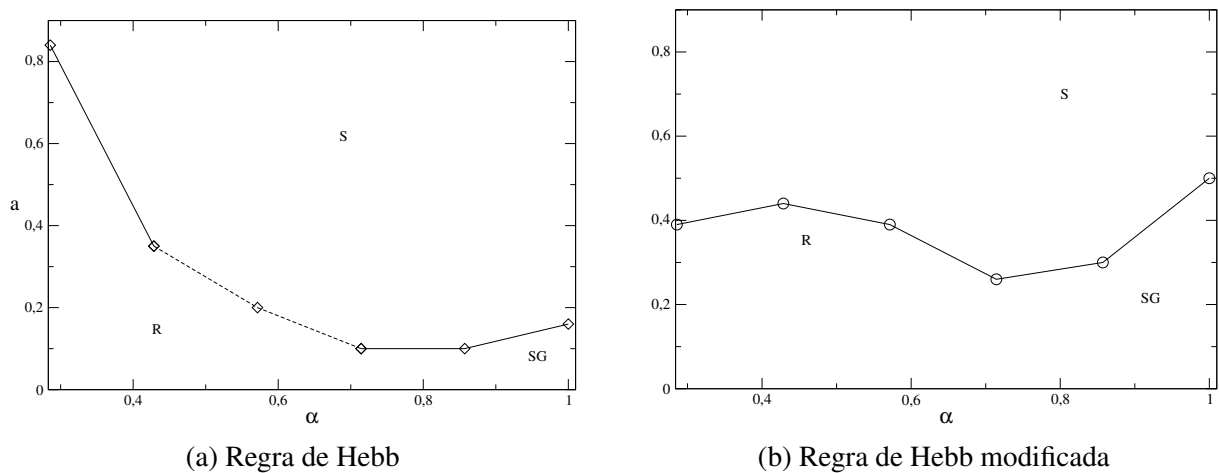


Figura 4.11: Diagramas no plano (α, a) com $c = 7$. Utilizando as regras de Hebb e a regra de Hebb modificada. Linhas cheias representam transições de fase contínuas e linhas tracejadas representam transições de fase descontínuas.

4.2.2 Diagramas de fase no plano (α, T)

No plano (α, T) , fixamos atividade e a conectividade médias. Traçamos curvas a p fixo de m^μ vs. T , identificando a temperatura crítica para cada α . Inicialmente, fixamos a atividade em $a = 0, 1$ e a conectividade em $c = 7$, na figura 4.16, desenhamos os diagramas de fase para cada regra de aprendizado.

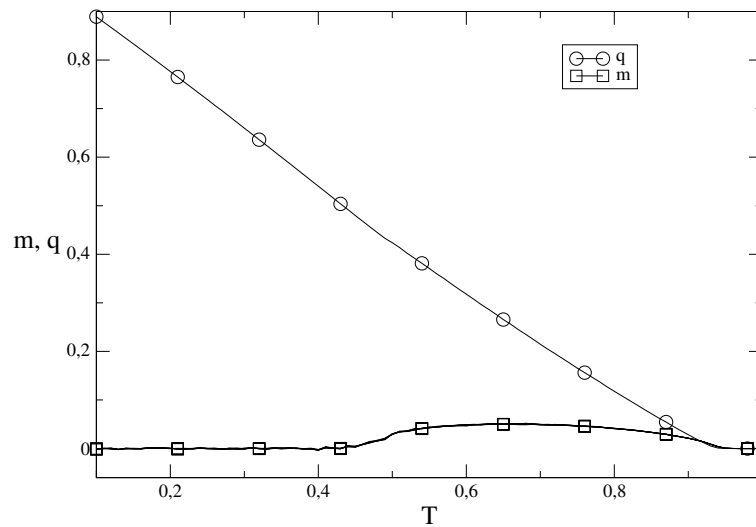


Figura 4.12: Regra de Hebb. Observáveis q e m^μ vs. T , $c = 7$, $p = 7$ e $a = 0, 1$. Transições de fase VS-S seguido de S-P.

No diagrama de fases da regra de Hebb, figura 4.16a, a fase de R está presente em valores de $p < 5$. À partir de $p > 5$ ($\alpha \approx 0, 714$), não há fase R, mesmo a T baixo. À medida que T é elevado a fase R dá lugar a fase S e para T suficientemente alto aparece uma fase P. Em $p = 7$ observa-se, figura 4.12, que existe uma transição VS-S, seguido de uma transição de S-P.

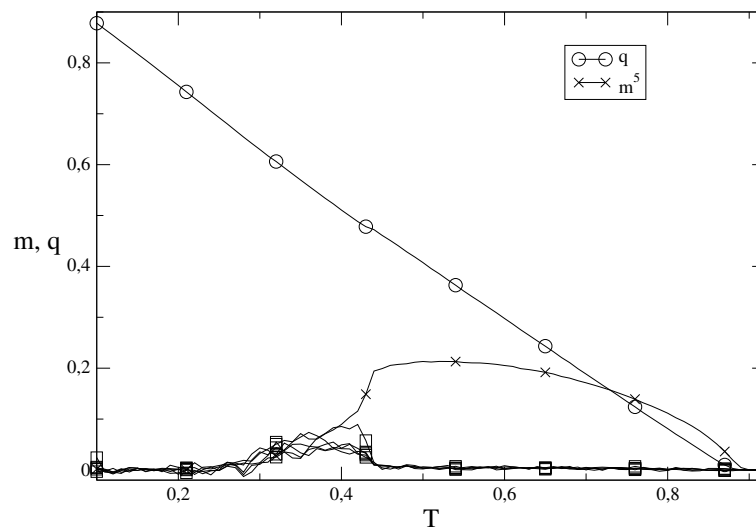


Figura 4.13: Regra de Hebb modificada. Observáveis q e m^μ vs. T , $c = 7$, $p = 6$ e $a = 0, 1$. Em $T = 0, 28$ existe uma transição de fase VS-S, para $T = 0, 36$ há uma transição S-R e finalmente uma transição R-P em $T = 0, 9$.

Na regra de Hebb modificada, figura 4.16b, a fase R está presente até $p = 6$ ($\alpha \approx 0, 857$). Observando a figura 4.13, nota-se que o aumento de T é necessário para o aparecimento de uma fase R. O sistema inicia de uma fase de VS à T baixo, onde há presença apenas de ordenamento local, elevando T uma fase S aparece em seguida para um T suficiente temos a fase R em que o sistema possui um ordenamento global com um padrão.

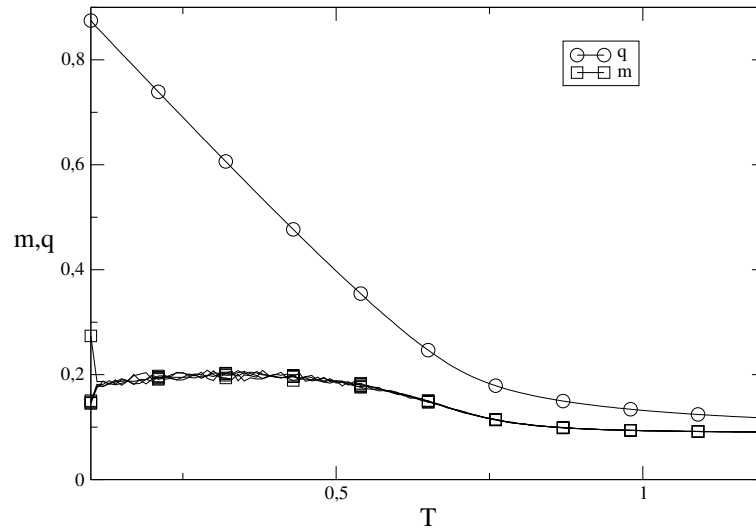


Figura 4.14: Regra de Hebb modificada com campo de vínculo. Observáveis q e m^μ vs. T , $c = 7$, $p = 7$ e $a = 0,3$. Para valores de T elevado, q e m^μ se aproximam de a^2 .

Na regra de Hebb modificada com campo de vínculo, ao elevar T , os parâmetros de ordem se aproximam de:

$$m^\mu = \langle \xi^\mu \langle \sigma \rangle_\beta \rangle_\xi \rightarrow a^2. \quad (4.10)$$

e

$$q = \langle \langle \sigma \rangle_\beta^2 \rangle_\xi \rightarrow a^2. \quad (4.11)$$

Isso pode ser justificado pois a atividade dos padrões é fixa em a , logo $\langle \xi^\mu \rangle_\xi = a$ e também devido à presença do campo de vínculo à atividade dos neurônios é $\langle \sigma \rangle_\beta = a$. Na figura 4.14 temos uma ilustração dessa propriedade.

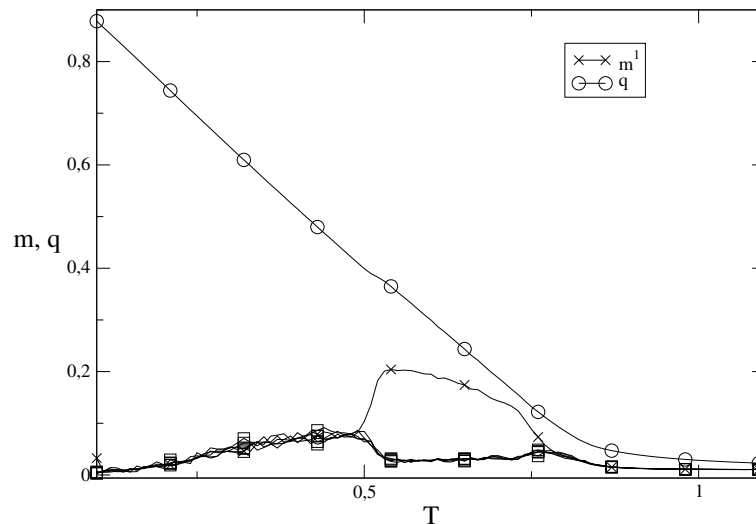
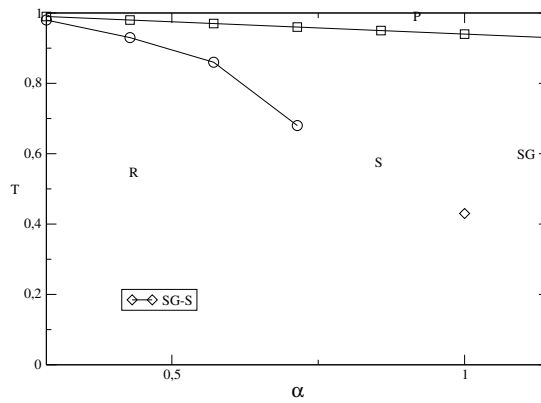


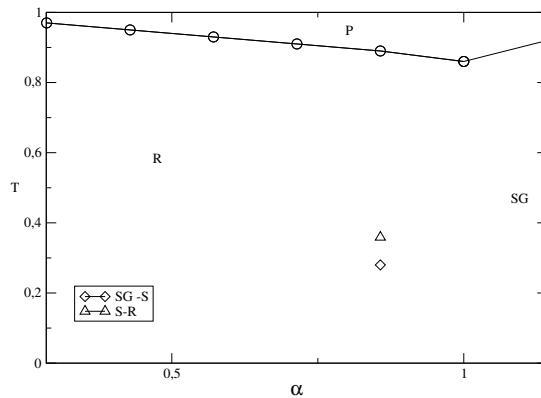
Figura 4.15: Regra de Hebb modificada com campo de vínculo. Observáveis q e m^μ vs. T , $c = 7$, $p = 7$ e $a = 0,1$. Em $T = 0,5$, temos uma transição S-R, em seguida em $T = 0,78$ o sistema retorna para a fase S.

A regra com campo de vínculo também possui uma fase R até $\alpha \approx 0,857$. Na figura 4.15, em

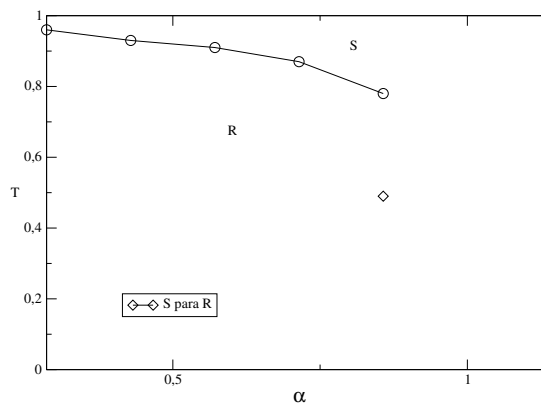
$p = 6(\alpha \approx 0,857)$, verificamos a presença de uma reentrância: Em baixas temperaturas, há uma fase S, para um T suficiente, o sistema favorece uma fase R e se elevado demais T retorna-se a uma fase S. Nota-se que com o campo de vínculo, a região da fase R em $p = 6$ é inferior a região de R em $p = 6$ da regra sem vínculo.



(a) Regra de Hebb



(b) Regra de Hebb modificada



(c) Regra de Hebb modificada com campo de vínculo

Figura 4.16: Diagramas de fase no plano (α, T) . Regras de Hebb 4.16a, Regra de Hebb modificada 4.16b e Regra de Hebb modificada campo de vínculo 4.16c. Conectividade média $c = 7$ e atividade $a = 0.1$.

Observando os diagramas de fase das regras modificada e com vínculo, verificamos que quando $p \sim c$, é necessário um certo nível de ruído térmico para auxiliar fases termodinâmicas

com maior ordenamento global. A regra de aprendizado de Hebb, 4.16a, apresenta valores de α e T_c inferiores as outras duas regras de aprendizado.

Efeitos do aumento da atividade

Analisamos o efeito do aumento da atividade sobre as regras de aprendizado, realizando diagramas de fases em cada regra de aprendizado, mantendo a conectividade média fixa em $c = 7$. Na regra de Hebb, figura 4.17b, o efeito do aumento de atividade acarretou numa diminuição da região da fase R, bem como a ausência de fases de VS. Quando $a = 0.3$, pode-se notar que a temperatura crítica T_c das transição S-P é elevada com o aumento de p .

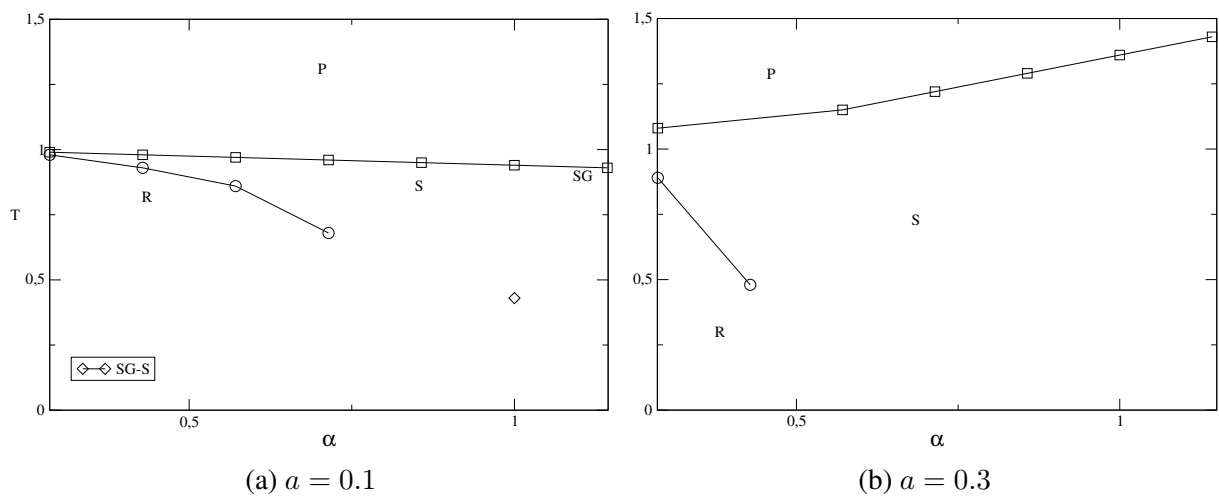


Figura 4.17: Diagramas com Regra de Hebb, $c = 7$.

Na figura 4.18, comparamos a regra de Hebb modificada para valores de $a = 0, 1$ e $a = 0, 3$. O primeiro aspecto notável é que a fase R começa a perder lugar para a fase de simetria S. Podemos fazer uma comparação com o trabalho de Amit [8], onde o autor, fazendo uso da regra modificada, a partir de uma análise de sinal e ruído, demonstra que a capacidade crítica numa rede totalmente conectada e com atividade finita, à $T = 0$, é $\alpha_c = (1 - |a|)^2 \alpha_c(a = 0)$, onde $\alpha_c(a = 0) = 0, 138$ [6]. Se calcularmos as capacidades críticas para as respectivas atividades $a = 0, 1$ e $a = 0, 3$, em seu modelo, obtemos $\alpha_c(a = 0, 1) \approx 0, 112$ e $\alpha_c(a = 0, 3) \approx 0, 068$. Em comparação, em nosso modelo com conectividade finita e ruído térmico presente, se estimarmos, à partir dos diagramas da figura 4.18, em quais valores de α , a regra de Hebb modificada ainda mantém a memória operante, obtemos $\alpha(a = 0, 1) \approx 0, 857$ e $\alpha(a = 0, 3) \approx 0, 714$.

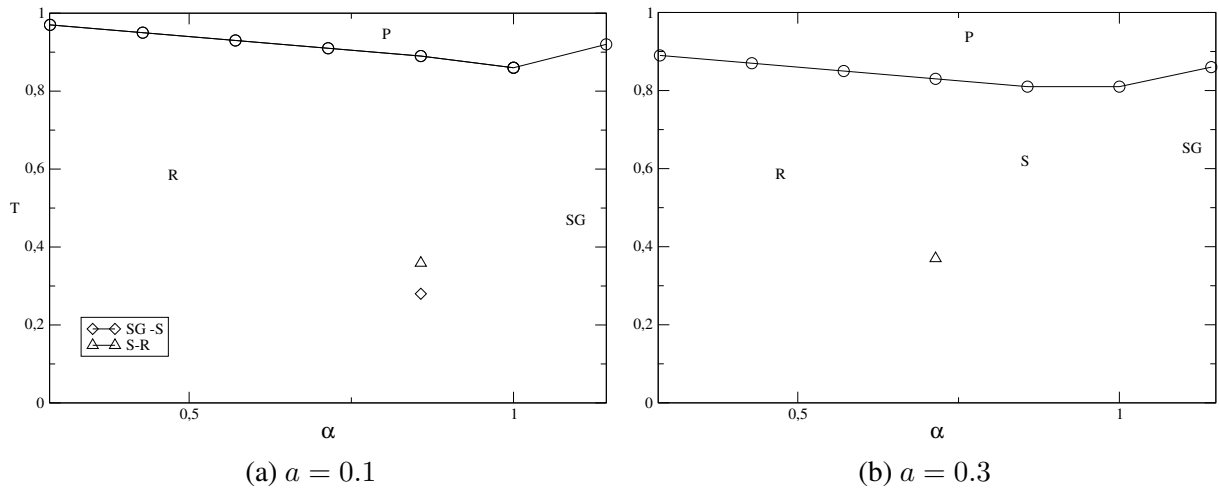


Figura 4.18: Diagramas com Regra de Hebb modificada, $c = 7$.

Comparando a regra de Hebb modificada com campo de vínculo, com dois valores distintos de atividade, figura 4.19, verifica-se que o efeito do aumento de a torna a fase R presente apenas em $p \leq 4$. Logo, a adição do vínculo, acarretou numa redução da fase R, em contraste com a regra modificada sem vínculo que possui uma fase R para valores de $p \leq 5$.

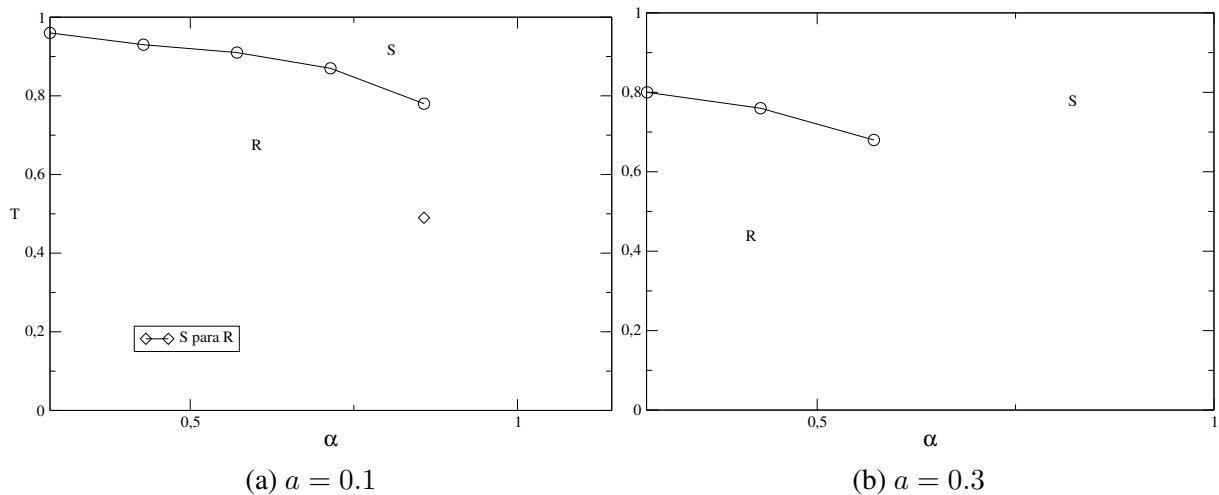


Figura 4.19: Diagramas com Regra de Hebb modificada com campo de vínculo, $c = 7$.

Os diagramas 4.17, 4.18 e 4.19 mostram que todas regras operam como memória associativa na presença de padrões com atividade finita, sendo que das três regras de aprendizado, a que mostra maior capacidade $\alpha(a)$ é a regra de Hebb modificada sem campo de vínculo.

Efeitos do aumento da conectividade média

Semelhante aos trabalhos de redes neurais atratoras em grafos aleatórios de Coolen e Pérez Castillo [11, 34], testamos o efeito da mudança de conectividade média sobre as três regras de aprendizado. Observamos que o aumento de c eleva a T_c das transições R-P e VS-P, figura 4.20.

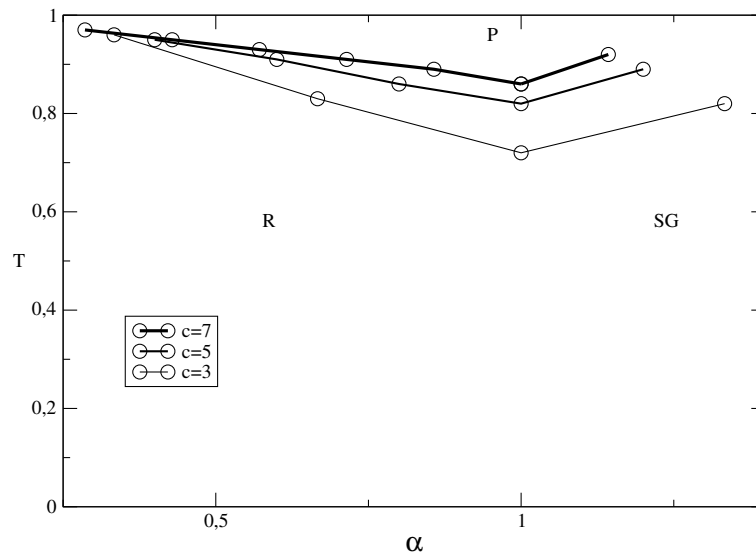


Figura 4.20: Diagramas de fase com a regra de Hebb modificada no plano (α, T) para valores $c = 3, c = 5, c = 7$ e $a = 0.1$.

Na regra de Hebb original, figura 4.21, a capacidade em $c = 5$ é de $\alpha(c = 5) = 0.6$. Ao elevar a conectividade para $c = 7$ temos $\alpha(c = 7) \approx 0.714$. No entanto, a T_c da transição R-S sofre um decréscimo em $p = 5$ ($\alpha = 0, 714$), figura 4.21b.

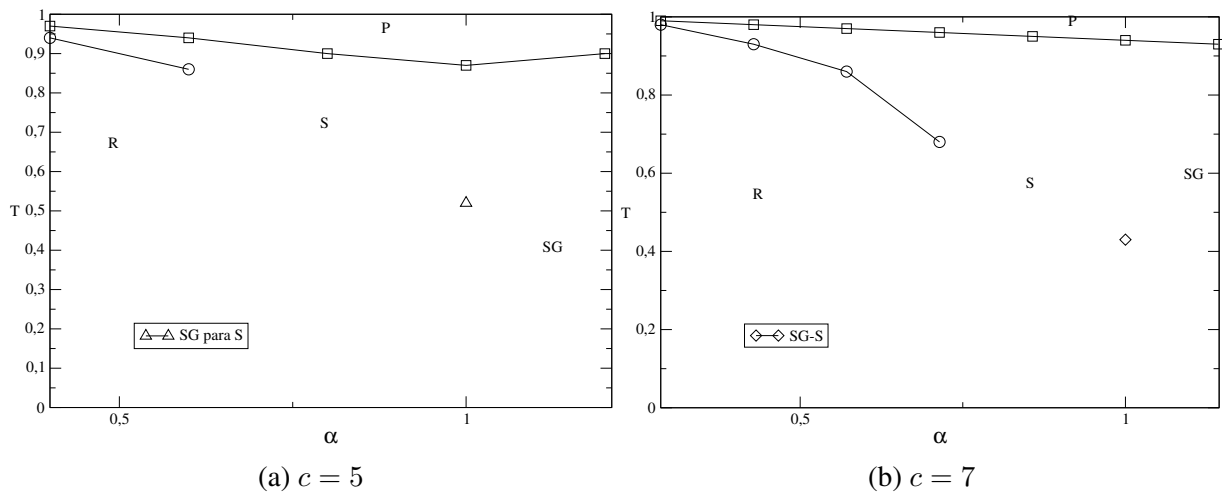


Figura 4.21: Diagramas com regra de Hebb para conectividades $c = 5$ e $c = 7$ com atividade $a = 0.1$

Na regra de Hebb modificada, figura 4.22, o aumento de c ocasiona, quando $p \sim c$, o aparecimento de transições de fase de VS-S e S-R a baixo T . Algo semelhante ocorre com a regra na presença de campo de vínculo, figura 4.23, porém a transição é do tipo S-R.

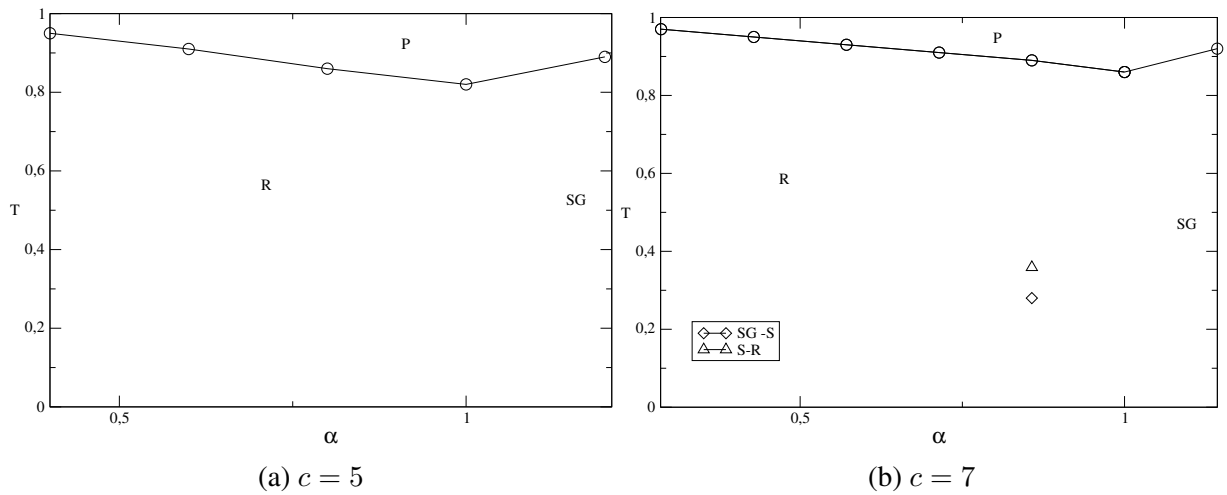


Figura 4.22: Diagramas com regra de Hebb modificada para conectividades $c = 5$ e $c = 7$ com atividade $a = 0.1$

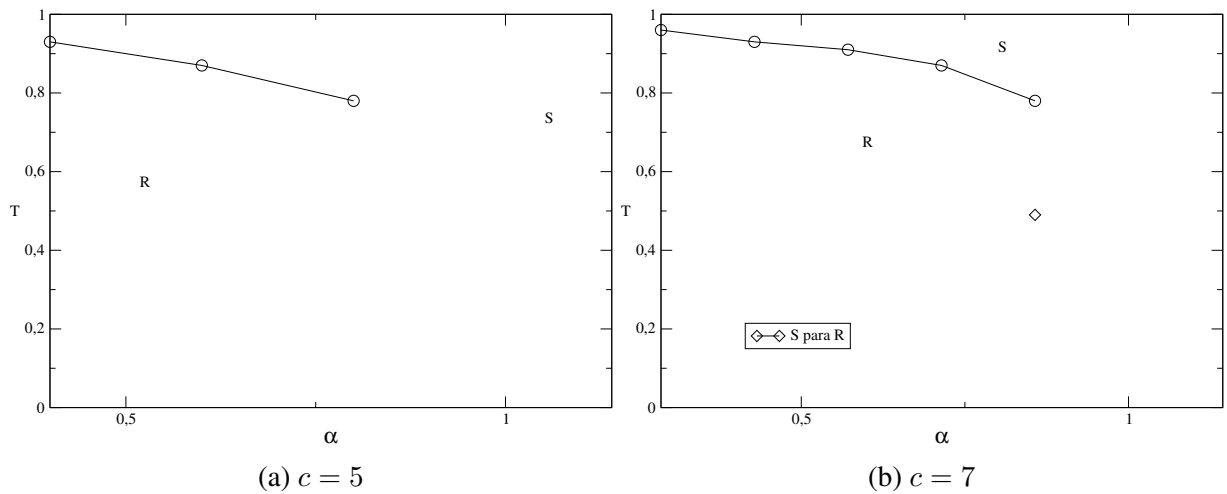


Figura 4.23: Diagramas com regra de Hebb modificada com campo de vínculo para conectividades $c = 5$ e $c = 7$ com atividade $a = 0.1$

A elevação de c aumenta a fase R, no entanto, como observamos nos diagramas de fase 4.21, 4.22 e 4.23, esse aumento introduz estados espúrios de VS e S.

4.2.3 Diagramas de fase no plano (a, T)

No plano (a, T) , fixamos o número de padrões e a conectividade média para cada diagrama. Nos diagramas descritos abaixo, iniciamos utilizando uma conectividade média $c = 10$ e um número $p = 3$ de padrões. Determina-se os pontos de transição de fase traçando curvas de m^μ e q em função de a e T .

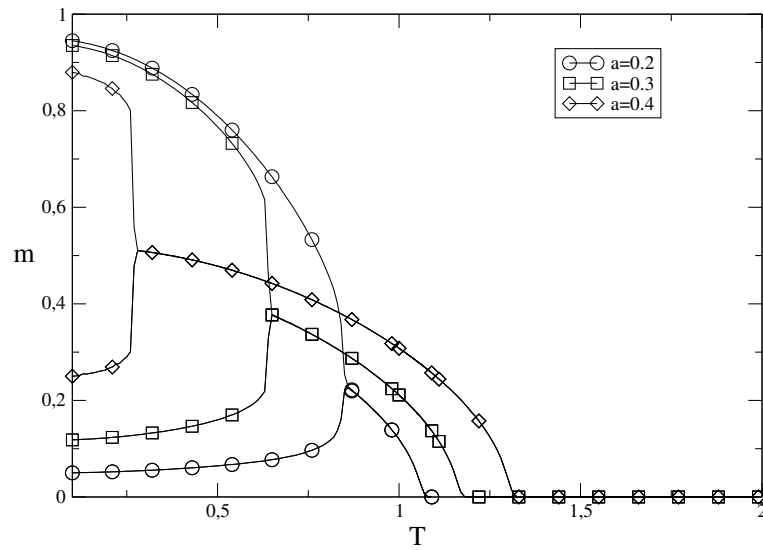


Figura 4.24: Observável m^μ vs. T utilizando a regra de Hebb, $p = 3$, conectividade $c = 10$ para valores de $a = 0.2$, $a = 0.3$ e $a = 0.4$

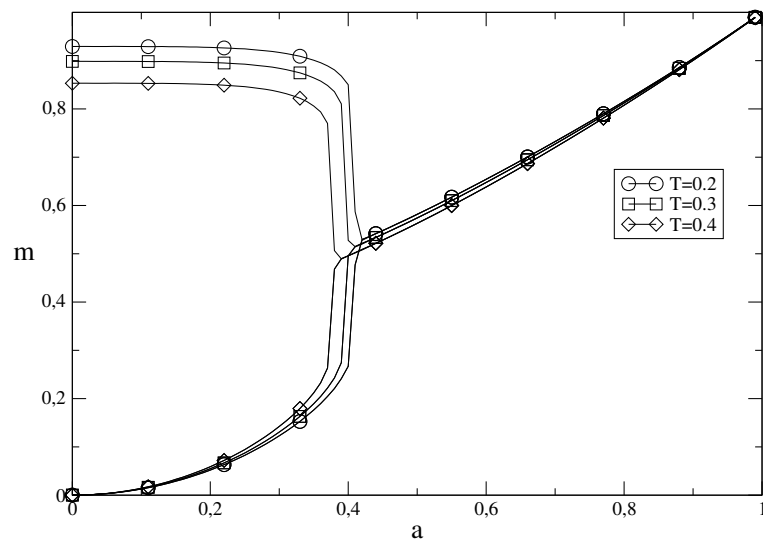
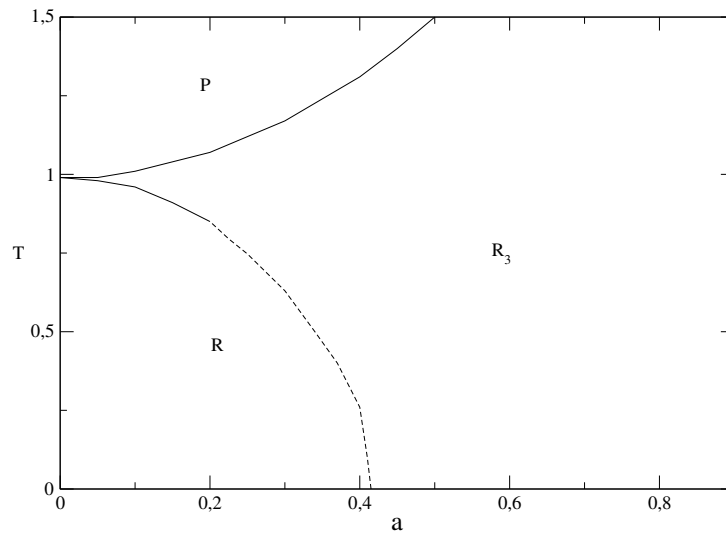
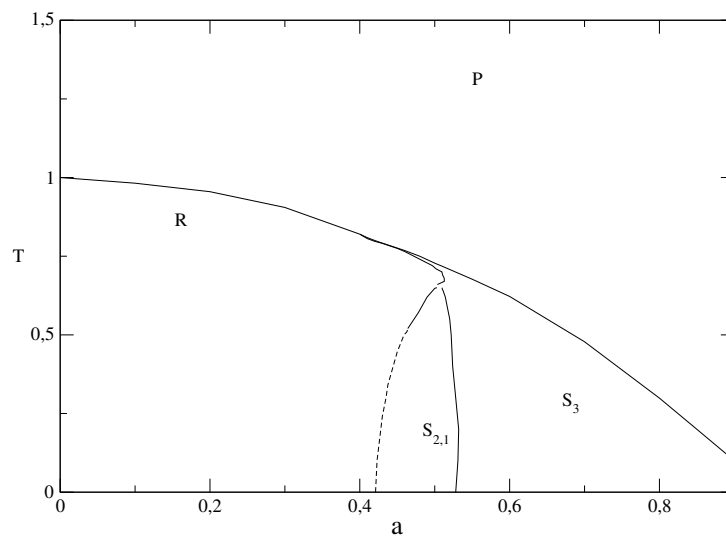


Figura 4.25: Observável m^μ vs. a utilizando a regra de Hebb com $p = 3$, conectividade $c = 10$ e valores de $T = 0.2$, $T = 0.3$ e $T = 0.4$

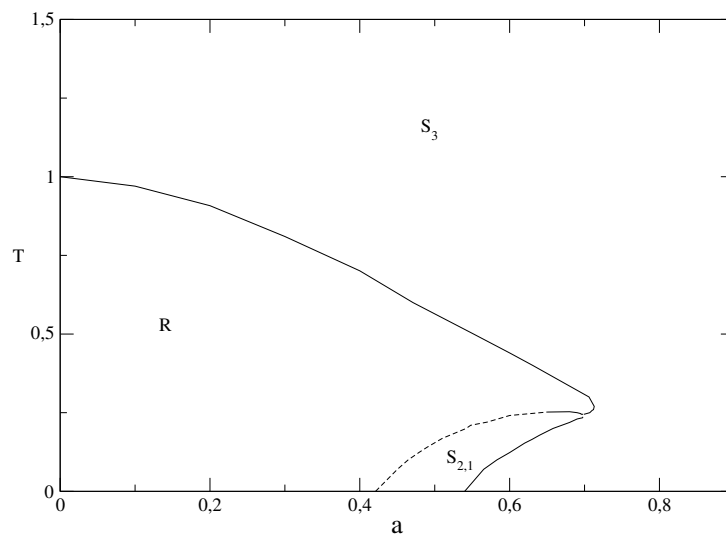
Na regra de Hebb, figuras 4.24 e 4.25, a transição R-S apresenta descontinuidades. Numa curva em $a = 0, 2$ fixo, existe uma transição de fase contínua em $T = 0, 84$, realizando mais medidas em curvas de atividade fixa com valores $a < 0, 2$, verificamos que as transições de fase tornam-se contínuas. Logo, marcamos o ponto $(a = 0, 2, T = 0, 84)$ como sendo o divisor entre transições descontínuas e contínuas, esse ponto é conhecido como ponto tricrítico, figura 4.26a.



(a) Regra de Hebb



(b) Regra de Hebb modificada



(c) Regra de Hebb modificada com campo de vínculo

Figura 4.26: Diagramas de fase no plano (a, T) . Regras de Hebb 4.26a, Regra de Hebb modificada 4.26b e Regra de Hebb modificada campo de vínculo 4.26c. Conectividade média $c = 10$ e número de padrões $p = 3$. Linhas tracejadas indicam transições de fase descontínuas e linhas cheias indicam transições contínuas.

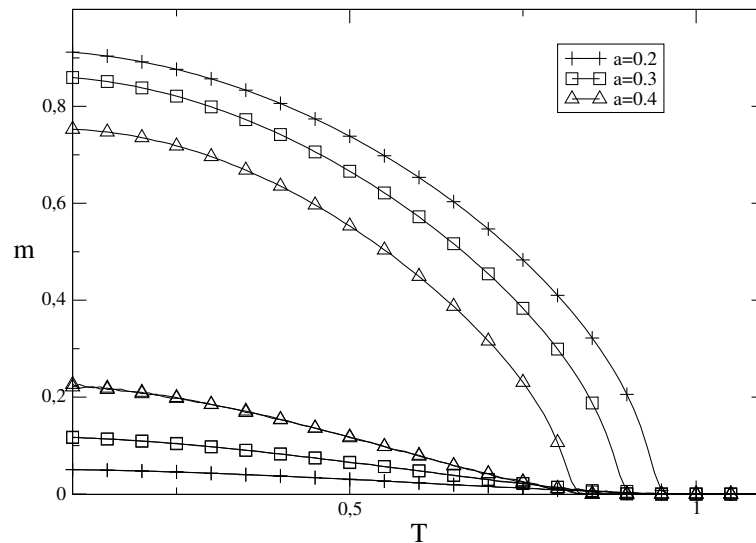


Figura 4.27: Observável m^μ vs. T utilizando a regra de Hebb modificada, $p = 3$, conectividade $c = 10$ para valores de $a = 0.2$, $a = 0.3$ e $a = 0.4$

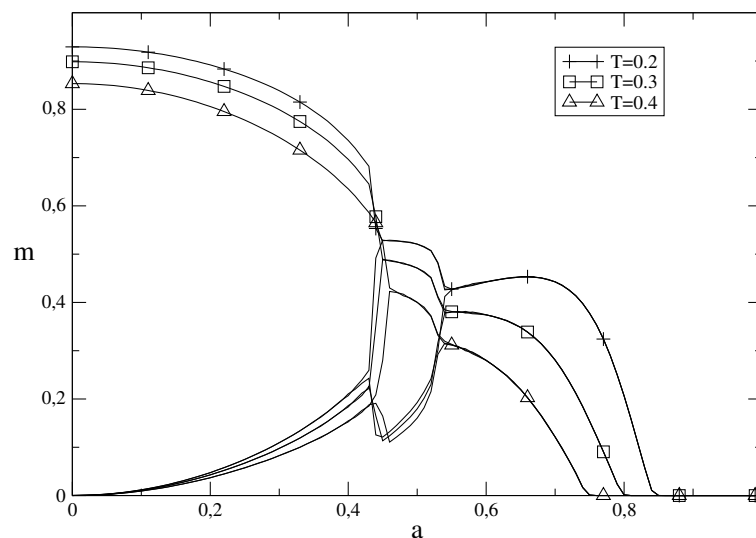


Figura 4.28: Observável m^μ vs. a utilizando a regra de Hebb modificada com $p = 3$, conectividade $c = 10$ e valores de $T = 0.2$, $T = 0.3$ e $T = 0.4$

Nas figuras 4.27 e 4.28, traçamos curvas de m^μ , utilizando a regra de Hebb modificada. Observamos a presença de uma fase assimétrica do tipo $S_{2,1}$, onde os *overlaps* m^1 e m^2 possuem o mesmo valor. O diagrama de fase dessa regra está representado na figura 4.26b. As transições de fase R- $S_{2,1}$ possuem descontinuidades em certos intervalos de a e T . Exibimos, na figura 4.29, em curvas de m^μ vs. T , alguns exemplos dessas transições descontínuas, para valores fixos de atividade $a = 0,43$ e $a = 0,45$. Quando o nível de atividade é relativamente alto, o ruído térmico aparenta colaborar para o sistema entrar numa fase de recuperação R. O ponto ($a = 0,464, T = 0,52$), na espínodal da transição R- $S_{2,1}$, é um ponto tricrítico.

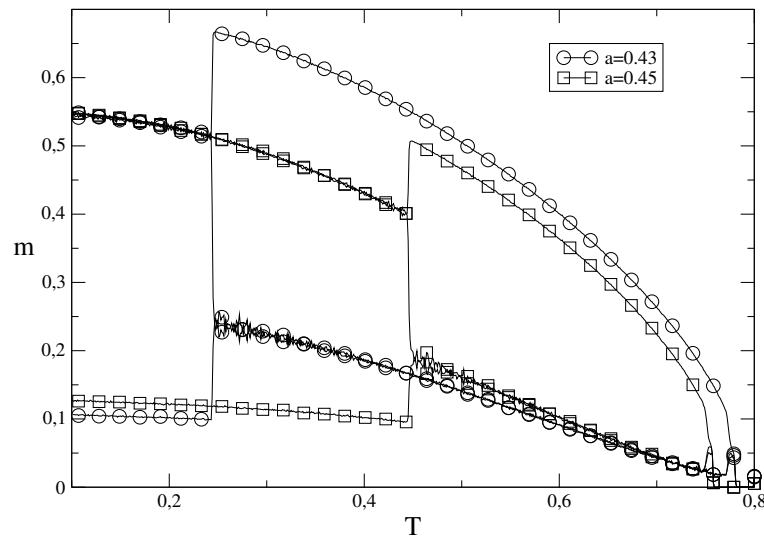


Figura 4.29: Observável m^μ vs. T utilizando a regra de Hebb modificada, $p = 3$, $c = 10$. Mostrando transições de fase descontínuas R-S_{2,1}.

Curvas do *overlap* em função de a e T usando a regra de Hebb modificada com o campo de vínculo são apresentadas nas figuras 4.30 e 4.31. O campo mantém o valor da atividade dos neurônios finita, como consequência, o observável q não será nulo, nem mesmo a altas temperaturas, esse efeito está mostrado na figura 4.33. Novamente, verificamos a presença de um ponto tricrítico na espinodal R-S_{2,1} localizado em $(a = 0,65, T = 0,252)$.

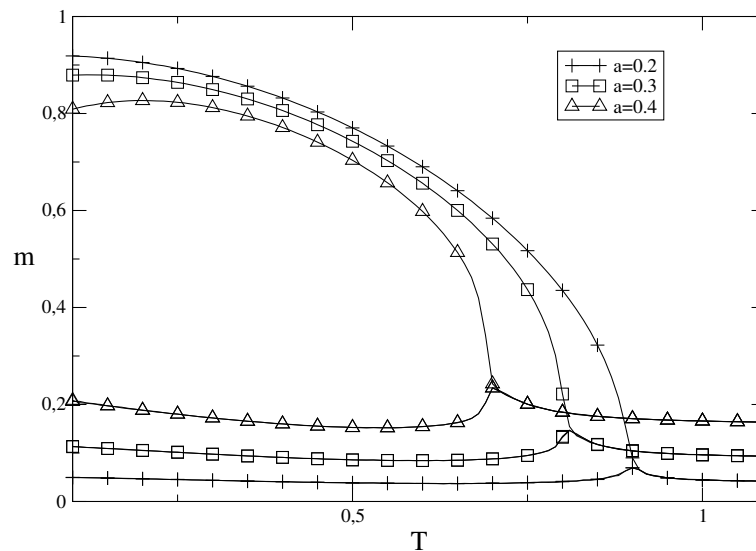


Figura 4.30: Observável m^μ vs. T utilizando a regra de Hebb modificada com campo de vínculo, $p = 3$, $c = 10$ para valores de $a = 0.2$, $a = 0.3$ e $a = 0.4$

A figura 4.31 mostra um gráfico comparativo de m^μ vs. a em três valores fixos de T . Em $T = 0,3$, observa-se o maior valor de a_c na transição de fase R-S. Comparando os diagramas de fase da regra de Hebb modificada 4.26b e a versão com campo de vínculo 4.26c, percebe-se que o campo de vínculo torna a rede melhor operante em valores de atividade elevada, porém, observando o diagrama 4.26c, vemos que quando a atividade é suficientemente elevada, o nível de T tolerável diminui. Na figura 4.34, desenhamos curvas de m^μ vs. a para cada regra de

aprendizado, numa temperatura fixa em $T = 0.4$. Essa figura evidencia que o campo de vínculo auxilia mais a memória em regimes com maior atividade sobre os padrões.

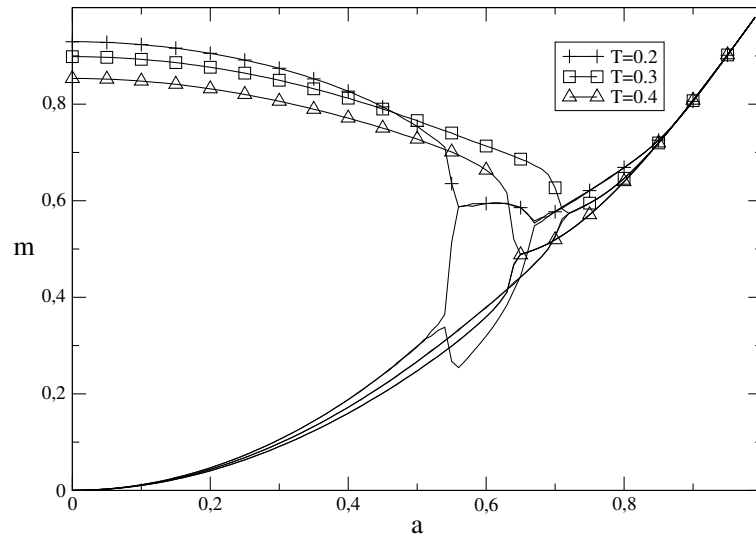


Figura 4.31: Observável m^μ vs. a utilizando a regra de Hebb modificada com campo de vínculo, $p = 3$, $c = 10$ e valores de $T = 0.2$, $T = 0.3$ e $T = 0.4$

Observando a figura 4.32, verificamos que, a baixas temperaturas, o sistema encontra-se numa fase S, à medida que T é elevada, o sistema passa por uma fase $S_{2,1}$, seguido de uma fase de R, e por fim retorna a S. Isso caracteriza um fenômeno de reentrância, que pode ser observado no diagrama de fase, na figura 4.26c.

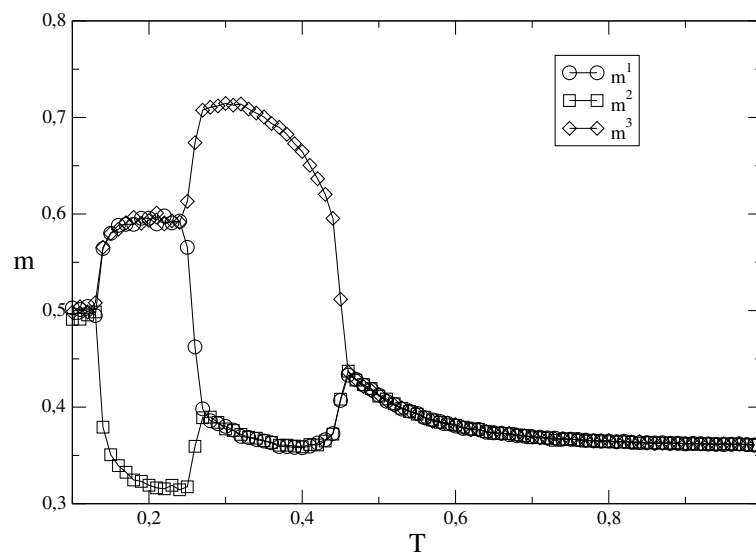


Figura 4.32: Observável m^μ vs. T , atividade $a = 0.6$, conectividade média $c = 10$. Regra de Hebb modificada com campo de vínculo.

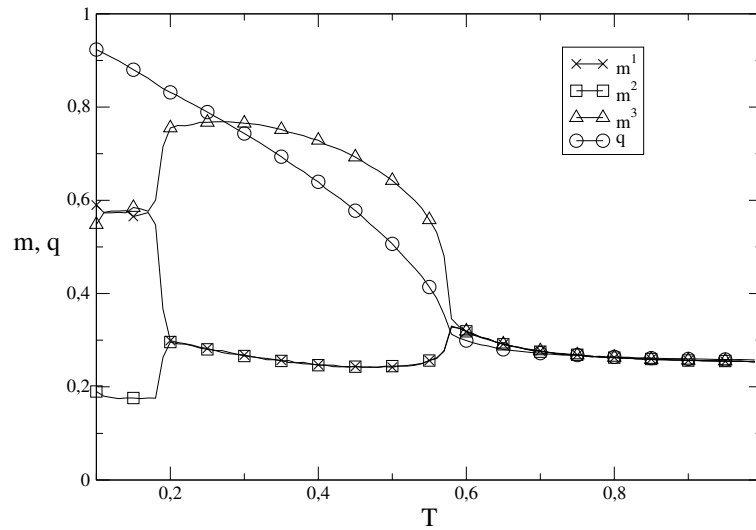


Figura 4.33: Observáveis m^μ e q vs. T , atividade $a = 0.5$, conectividade média $c = 10$. Regra de Hebb modificada com campo de vínculo.

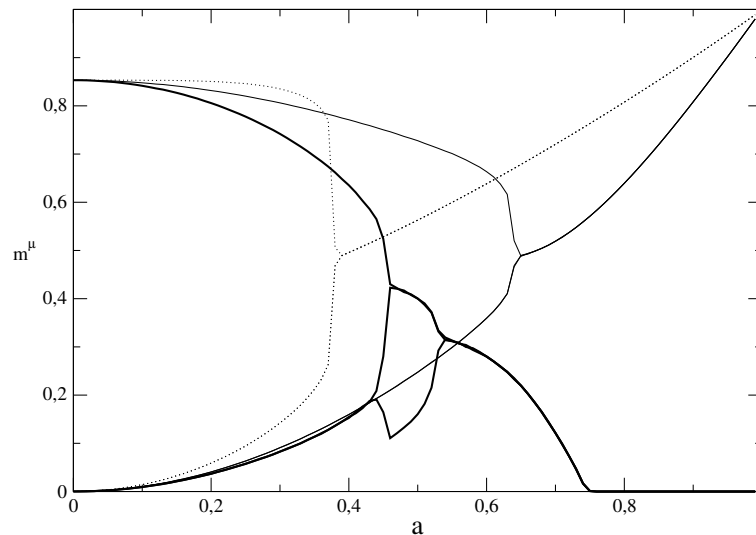


Figura 4.34: Observável m^μ vs. a , temperatura $T = 0.4$, conectividade média $c = 10$. Linhas pontilhadas representam a regra de Hebb; linhas cheias e grossas representam a regra de Hebb modificada; linhas cheias finas representam a regra com campo de vínculo.

Efeito do aumento da conectividade média

Examinamos o efeito do aumento da conectividade média no plano (a, T) . Na figura 4.35, mantendo o número de padrões fixo em $p = 3$, desenhamos curvas de m^μ vs. a para valores de conectividade média $c = 5$, $c = 10$, $c = 30$ e $c = 50$. Observa-se que a medida que c é aumentado, o valor de a_c da fase de recuperação R é elevado. Altas conectividades médias melhoram o *overlap* m^μ , isso acarreta numa melhor recuperação do padrão. Uma analogia pode ser feita entre um padrão ensinado à rede e uma imagem. Quanto maior o valor de m^μ , melhor a resolução da imagem recuperada pela memória. Em contrapartida, o aumento de c eleva a região ocupada pela fase $S_{2,1}$ e faz também que as transições de fase de tornam-se descontínuas. Na

figura 4.36 traçamos diagramas para as três regras utilizando conectividades médias de $c = 7$ e $c = 10$.

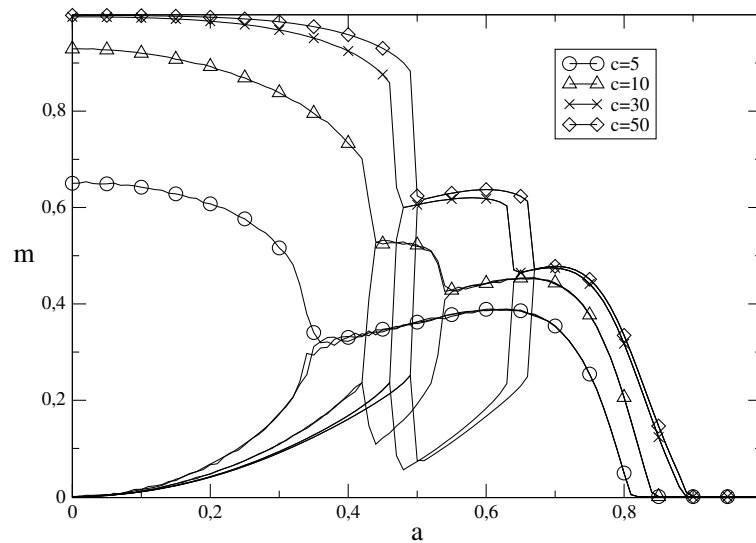


Figura 4.35: Observável m^μ vs. a para $T = 0.2$ em valores de $c = 5$, $c = 10$ e $c = 30$. Regra de Hebb modificada.

Nos diagramas representados na figura 4.36, na regra de Hebb, o valor de a_c na transição de fase R-S em $c = 7$ é $a_c = 0,35$, aumentado a conectividade para $c = 10$, o valor sobe para $a_c = 0,4$. O ponto tricrítico em $c = 7$ é localizado em $(a = 0,3, T = 0,45)$, já em $c = 10$ ele muda para $(a = 0,464, T = 0,52)$. A regra de Hebb modificada em $c = 7$ não apresenta transições de fase descontínuas. O mesmo ocorre com a regra modificada na presença do vínculo. Observando o diagrama de fases da regra com vínculo, em $c = 7$, verificamos que a fase $S_{2,1}$ é quase inexistente.

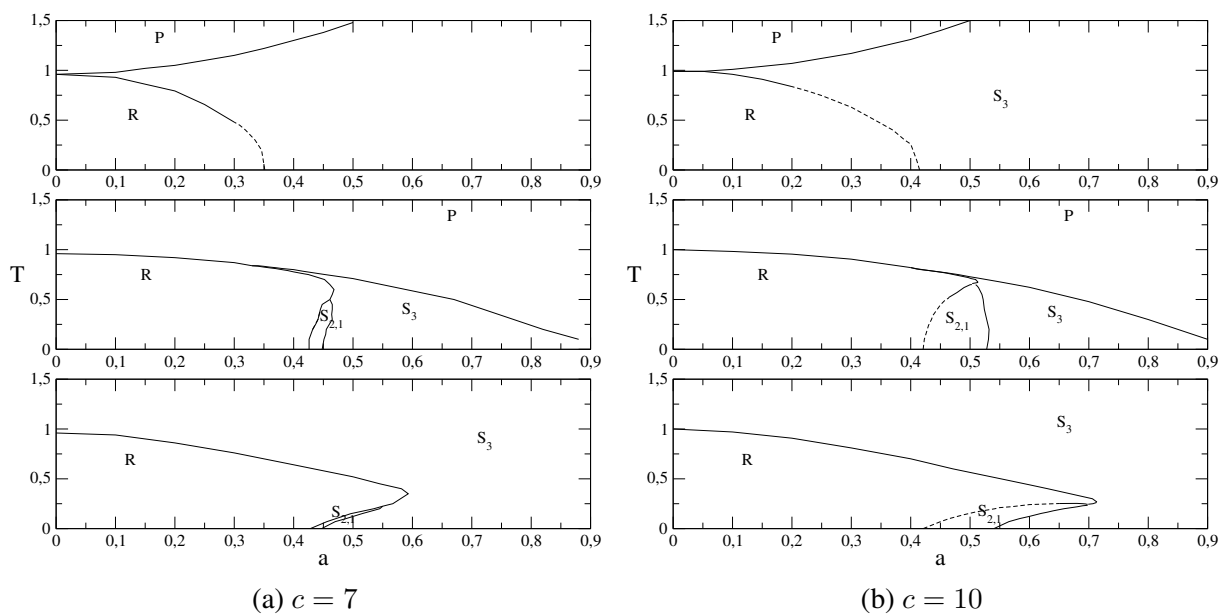


Figura 4.36: Diagramas de fase das respectivas regras: Hebb, Hebb modificada e com campo de vínculo, na figura 4.36a, $c = 7$, e na figura 4.26 $c = 10$, todos diagramas realizados com $p = 3$.

Efeito de um número diferente de padrões

Comparamos também o efeito que um número diferente de padrões pode acarretar. Na figura 4.37, traçamos diagramas usando as três regras, com $c = 10$, em valores de $p = 3$ e $p = 4$. Quando utilizamos três padrões, só existe um possível estado assimétrico $S_{2,1}$. No caso de $p = 4$, os possíveis estados assimétricos são $S_{2,2}$ e $S_{3,1}$. Sendo que $S_{2,2}$ é representado por $m^1 = m^2 = m$ e $m^3 = m^4 = m'$. Já o estado $S_{3,1}$, corresponde a $m^1 = m^2 = m^3 = m$ e $m^4 = m'$, incluindo as possíveis permutações. No diagrama da figura 4.37 que corresponde a regra de Hebb modificada com $p = 4$, notamos apenas a presença da fase $S_{2,2}$. Porém, examinando a figura 4.38, onde elevamos a conectividade média até $c = 13.333$, verificamos o aparecimento da fase $S_{3,1}$. Isso indica que o aumento da conectividade média favorece o aparecimento de outras fases de assimetria.

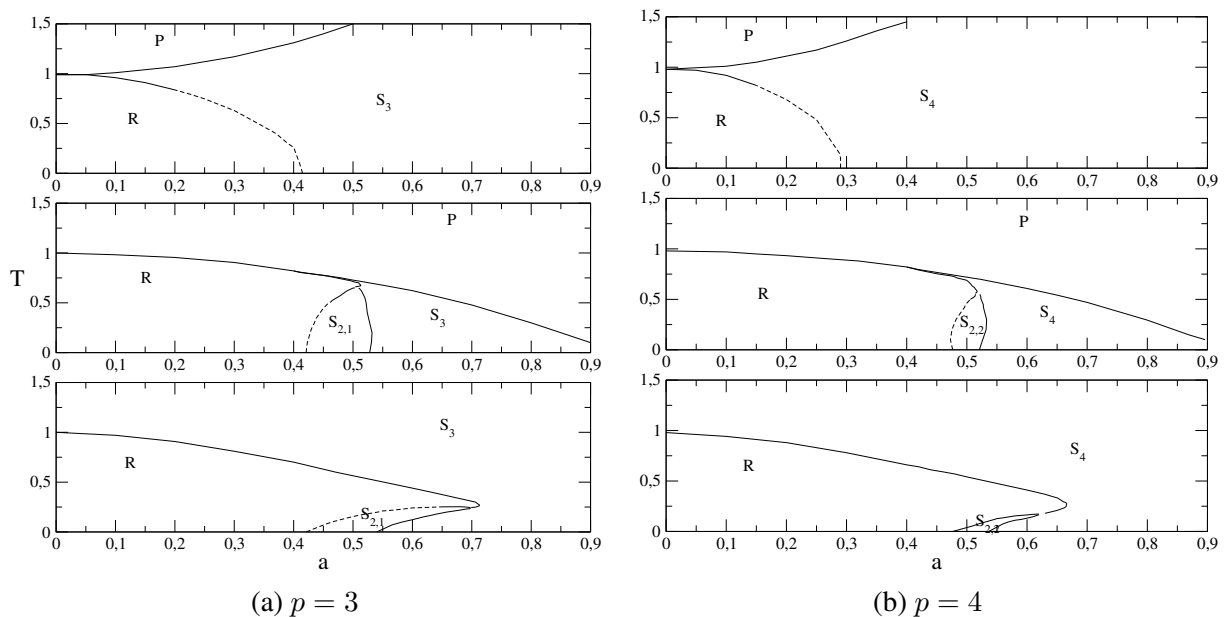


Figura 4.37: Diagramas de fase das respectivas regras: Hebb, Hebb modificada e com campo de vínculo, na figura 4.37a, $p = 3$, e na figura 4.37b $p = 4$, todos diagramas realizados com $c = 10$.

Notamos que a mudança do número de padrões não afeta a topologia dos diagramas de fase em cada regra. Inicialmente, isso era uma dúvida, motivada pelo fato de que com três padrões o valor da interação entre dois neurônios não pode ser zero. Porém, com quatro padrões podemos ter algo como:

$$\xi \cdot \xi' = (+1, +1, +1, +1) \cdot (+1, +1, -1, -1) = 0. \quad (4.12)$$

Isso acarretaria, caso a fosse suficientemente elevado para favorecer essas interações, numa decorrelação entre os neurônios, esperando assim uma região de fase paramagnética dominante nos diagramas de fase.

No diagrama da regra de Hebb, o valor de a_c decai em $p = 4$, bem como observado no plano (α, a) , no diagrama da figura 4.11. Na regra de Hebb modificada, em $p = 4$, a fase assimétrica $S_{2,2}$ possui uma região menor do que a da fase $S_{2,1}$, em $p = 3$. No diagrama da regra com campo de vínculo, em $p = 4$, não foi possível identificar de forma clara a natureza da transição R- $S_{2,2}$, logo, desenhou-se a linha dessa transição como contínua.

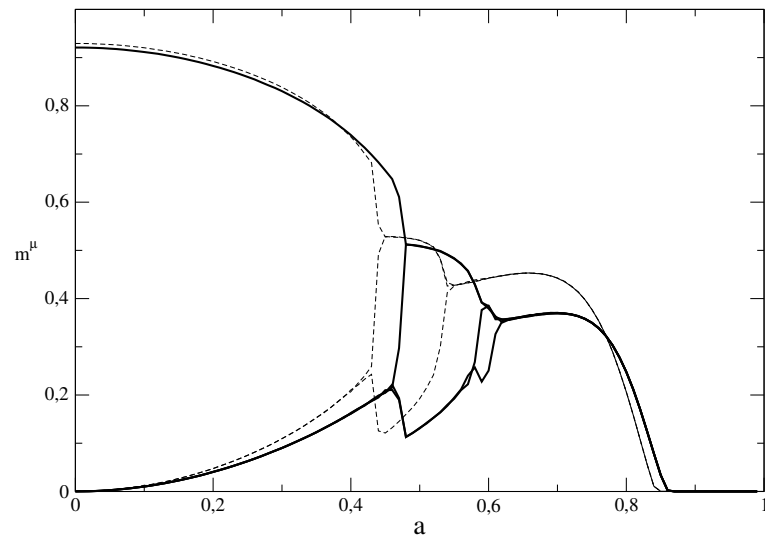


Figura 4.38: Regra de Hebb modificada. Observável m^μ vs. a , $T = 0.2$. Linhas pontilhadas correspondem a $p = 3$ e $c = 10$. Linhas contínuas a $p = 4$ e $c = 13.333$, ambos com mesmo $\alpha = 0.3$.

Capítulo 5

Considerações finais e perspectivas

Neste trabalho, estudamos o comportamento de uma rede neural atratora com um número finito de conexões por sítio quando os padrões possuem uma atividade não nula. Os objetivos desse trabalho foram estudar as propriedades de três regras de aprendizado, sujeitas às condições de conectividade e atividade finitas, e averiguar as seguintes propriedades: a atividade crítica $a_c(\alpha, T = 0)$ na ausência de ruído térmico; a capacidade em função da temperatura $\alpha_c(T)$ e atividade dos padrões $\alpha_c(a)$; a atividade crítica em função da temperatura $a_c(T)$; o efeito da conectividade sobre a capacidade α_c e a atividade crítica a_c ; o efeito do número de padrões sobre a a_c .

No Capítulo 2, foi feita uma introdução de modelos de redes atradoras. Apresentamos a regra de Hebb e o modelo determinístico de Hopfield. Em seguida, mostramos o modelo estocástico de Hopfield, estudado por (AGS), nos regimes $\alpha = 0$ e $\alpha = \frac{p}{N}$, e explanamos também as regras de aprendizado adaptadas para otimizar a capacidade da memória na presença de atividade $a \neq 0$.

No Capítulo 3, desenvolvemos os cálculos para uma rede atratora num grafo aleatório, com padrões de atividade finita. Para tal, utilizamos o formalismo de simetria de réplicas para sistemas com conectividade finita, que faz uso de funções de ordem. Utilizamos o conceito de sub-redes para particionar os sítios de acordo com os padrões que se deseja inserir. A partir da introdução de funções de ordem, calculou-se energia livre através do método de ponto de sela. A equação de recorrência para as funções de ordem por sub-rede foi obtida através do cálculo do extremo da função de energia livre. Para prosseguir com os cálculos, fizemos uso do *ansatz* de simetria de réplicas, que é uma função da distribuição de campos locais. Em seguida obtivemos uma expressão analítica para as distribuições de campos locais por sub-rede $W_\xi(h)$, a partir da qual extraímos toda informação sobre o sistema, e calculamos os observáveis que permitem identificar a fase termodinâmica. Foi feita adição de um vínculo sobre as funções de ordem de forma que a atividade dos neurônios seja limitada à atividade dos padrões. Novamente, obtivemos uma expressão para $W_\xi(h)$ na presença desse vínculo. Ainda nesse capítulo derivamos uma expressão para $W_\xi(h)$ na ausência de ruído térmico, $T = 0$.

No Capítulo 4, explanamos como é implementado numericamente o cálculo das $W_\xi(h)$. Mostramos exemplos do formato das $W_\xi(h)$ após atingirem um estado estacionário em cada uma das fases macroscópicas: paramagnética (P), recuperação de um padrão (R), fases simétricas (S), assimétricas S_{n_1, n_2} e fase de vidro de spin (VS). Estudamos três regras de aprendizado: a regra de Hebb original, a versão modificada por (AGS) para otimizar a rede na presença de padrões com atividade e uma regra que inclui a versão modificada e um campo que vincula a atividade dos neurônios à dos padrões. Dividimos os resultados nos planos de fase: (α, a) , (α, T) e (a, T) .

No plano (α, a) , verificamos que a capacidade da regra de Hebb e a versão modificada é similar. Contudo a regra de Hebb modificada possui valores de a_c mais elevados à medida que o número de padrões é elevado. Elevando a conectividade média na regra modificada para $c = 10$, verificamos que isso acarreta diminuição de α com maior valor de a_c .

No plano (α, T) , fixando $c = 7$, as regras apresentam fases desordenadas a baixas temperaturas. Aparentemente, isso ocorre quando $p \sim c$. Na regra com o campo de vínculo, verificamos a presença de transições reentrantes. As três regras sofrem perda na região de fase R com o aumento de a . A regra modificada mostra-se a mais resistente ao aumento de a . A vantagem da regra modificada sobre a original de Hebb pode ser argumentada em torno da análise de sinal e ruído derivada na seção 2.5. Devido ao aumento de p e a presença de padrões correlacionados, o termo de ruído da regra original torna-se suficientemente grande para tornar os estados R instáveis. O campo de vínculo, apesar de igualar a atividade dos neurônios a dos padrões, mostra um desempenho inferior ao da regra modificada sem vínculo. Vemos que a elevação de c favorece a fase (R) nas três regras, sendo que as regras modificada e com campo de vínculo mostram o mesmo ganho na fase R. Porém as T_c na região de $p \sim c$ sofrem uma perda maior nas regras de Hebb original e com vínculo. O aumento da conectividade também possui alguma relação com o aparecimento de transições de fases menos ordenadas para fases mais ordenadas, em baixos valores de T . No que tange à capacidade das regras, a versão modificada sem vínculo mostra os melhores resultados nesse plano.

No plano (a, T) , fixando $c = 10$, analisamos as três regras de aprendizado. Verificamos que, na regra de Hebb a transição de fase R-S é descontínua dentro de um intervalo de a e T , apresentando um ponto tricrítico em $(0,2,0,84)$. Nas regra de Hebb modificada e com campo de vínculo, existem transições descontínuas R- $S_{2,1}$. Verificamos também a presença de pontos tricríticos nessas regras, localizados respectivamente em $(0,464,0,52)$ e $(0,65,0,252)$ na espinodal R- $S_{2,1}$. Na regra com vínculo, verificamos uma forte reentrância no diagrama de fases. Aparentemente, quando o nível de atividade é suficientemente elevado, e portanto, os padrões estão suficientemente correlacionados, é necessário uma certa quantidade de ruído térmico T para auxiliar o sistema a sair das fases S e $S_{2,1}$ e entrar na fase R. A regra com campo de vínculo possui valores de a_c superior ao das outras regras, porém, em comparação com a regra modificada sem vínculo, apresenta menor resistência à elevação de T . O incremento de c aumenta o valor de m^μ , auxilia no valor de a_c da transição R-S e R- S_{n_1, n_2} , torna as transições de fase descontínuas e amplia a região da fase $S_{2,1}$. Ao examinar os efeitos de um número distinto de padrões, não identificamos grandes mudanças. Na regra de Hebb, como esperado, o aumento de p acarretou em a_c menor. Nas regras modificada e com campo de vínculo, encontram-se outras fases de assimetria em $p = 4$, tais como $S_{2,2}$ e $S_{3,1}$. No entanto, nessas regras não há grandes modificações nos diagrama de fases.

Como perspectivas futuras, podemos introduzir o cálculo do conteúdo de informação em cada fase, estudar a fase de vidro de spin utilizando a teoria de quebra de simetria de réplicas, testar outras regras de aprendizado para padrões com atividade finita e estudar redes atratoras com padrões de atividade finita em outros tipos de grafos, como sem escala e pequeno mundo.

Referências Bibliográficas

- [1] HERTZ, J.; KROGH, A.; PALMER, R. *Introduction to the theory of neural computation*. Addison-Wesley Computation and Neural Systems Series. Avalon Publishing, 1991.
- [2] MCCULLOCH, W. S.; PITTS, W. A. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, v. 5, p. 99–115, 1943.
- [3] HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci USA*, v. 79, p. 2554–2558, 1982.
- [4] HEBB, D. O. *The organization of behavior*. Wiley, New York, 1949.
- [5] AMIT, D. *Modeling brain function: The world of attractor neural networks*. Cambridge University Press, 1992.
- [6] D. J. AMIT, H. G.; SOMPOLINSKY, H. Statistical mechanics of neural networks near saturation. *Ann. Phys.*, v. 173, p. 30–67, 1987.
- [7] D. J. AMIT, H. G.; SOMPOLINSKY, H. Spin-glass models of neural networks. *Phys. Rev. A*, v. 32, p. 1007–1018, 1985.
- [8] D. J. AMIT, H. G.; SOMPOLINSKY, H. Information storage in neural networks with low levels of activity. *Phys. Rev. A*, v. 35, p. 2293–2303, 1987.
- [9] B. DERRIDA, E. G.; ZIPPELIUS, A. An exactly solvable asymmetric neural network model. *Europhys. Lett.*, v. 4, p. 167–173, 1987.
- [10] WATKIN, T. L. H.; SHERRINGTON, D. A neural network with low symmetric connectivity. *Europhys. Lett.*, v. 14, p. 791–796, 1991.
- [11] WEMMENHOVE, B.; COOLEN, A. C. C. Finite connectivity attractor neural networks. *J. Phys. A: Math. Gen.*, v. 36, p. 9617–9633, 2003.
- [12] MONASSON, R.; ZECCHINA, R. Statistical mechanics of the random k-satisfiability model. *Phys. Rev. E*, v. 56, p. 1357–, 1997.
- [13] Pérez Castillo, I.; Wemmenhove, B.; Hatchett, J. P. L.; Coolen, A. C. C.; Skantzos, N. S.; Nikolettopoulos, T. Analytic solution of attractor neural networks on scale-free graphs. *Journal of Physics A Mathematical General*, v. 37, p. 8789–8799, Sept. 2004.
- [14] Nikolettopoulos, T.; Coolen, A. C. C.; Pérez Castillo, I.; Skantzos, N. S.; Hatchett, J. P. L.; Wemmenhove, B. Replicated transfer matrix analysis of Ising spin models on 'small world' lattices. *Journal of Physics A Mathematical General*, v. 37, p. 6455–6475, June 2004.
- [15] ROSENBLATT, F. *Principles of neurodynamics*. Spartan, Washington, DC, 1961.

- [16] WIKIMEDIA. A hopfield net with four nodes., 2006. File: File:Hopfield-net.png.
- [17] PERETTO, P. Collective properties of neural networks: A statistical physics approach. *Biol. Cybern.*, v. 50, p. 51–62, 1984.
- [18] Mattis, D. C. Solvable spin systems with random interactions. *Physics Letters A*, Amsterdam, v. 56, p. 421–422, Apr. 1976.
- [19] KIRKPATRICK, S.; SHERRINGTON, D. Infinite-ranged models of spin-glasses. *Phys. Rev. B*, v. 17, p. 4384–4403, 1978.
- [20] EDWARDS, S. F.; ANDERSON, P. W. Theory of spin glasses. *J. Phys. F: Met. Phys.*, v. 5, p. 965–975, 1975.
- [21] Castellani, T.; Cavagna, A. Spin-glass theory for pedestrians. *Journal of Statistical Mechanics: Theory and Experiment*, v. 5, p. 05012, May 2005.
- [22] DE ALMEIDA, J. R. L.; THOULESS, D. J. Stability of the sherrington-kirkpatrick solution of a spin glass model. *J. Phys. A: Math. Gen.*, v. 11, p. 983–990, 1978.
- [23] WIKIMEDIA. 11-simplex graph, 2010. File: File:11-simplex graph.svg.
- [24] J. L. GROSS, J. Y.; ZHANG, P. *Handbook of graph theory, second edition*. CRC Press, 2014.
- [25] BOLLOBÁS, B. *Modern graph theory*. Springer New York, 1998.
- [26] ERDŐS, P.; RÉNYI, A. *On random graphs. i*. 1959. v. 6.
- [27] Kabashima, Y.; Murayama, T.; Saad, D. Typical Performance of Gallager-Type Error-Correcting Codes. *Physical Review Letters*, Woodbury, v. 84, p. 1355–1358, Feb. 2000.
- [28] Skantzos, N. S.; Saad, D.; Kabashima, Y. Analysis of common attacks in public-key cryptosystems based on low-density parity-check codes. *Phys. Rev. E*, v. 68, n. 5, p. 056125, Nov. 2003.
- [29] KABASHIMA, Y.; SAAD, D. Statistical mechanics of low-density parity-check codes. *J. Phys. A: Math. Gen.*, v. 37, p. R1–R43, 2004.
- [30] Skantzos, N. S.; van Mourik, J.; Saad, D.; Kabashima, Y. Average and reliability error exponents in low-density parity-check codes. *Journal of Physics A Mathematical General*, v. 36, p. 11131–11141, Oct. 2003.
- [31] Cocco, S.; Monasson, R. Trajectories in Phase Diagrams, Growth Processes, and Computational Complexity: How Search Algorithms Solve the 3-Satisfiability Problem. *Physical Review Letters*, Woodbury, v. 86, p. 1654–1657, Feb. 2001.
- [32] Mézard, M.; Parisi, G. The Bethe lattice spin glass revisited. *European Physical Journal B*, v. 20, p. 217–233, Mar. 2001.
- [33] VIANA, L.; BRAY, A. J. Phase diagrams for dilute spin glasses. *J. Phys. C: Solid State Phys.*, v. 18, p. 3037, 1985.
- [34] Pérez Castillo, I.; Skantzos, N. S. The Little Hopfield model on a sparse random graph. *Journal of Physics A Mathematical General*, v. 37, p. 9087–9099, Oct. 2004.

-
- [35] Wemmenhove, B.; Skantzos, N. S. Slowly evolving random graphs II: adaptive geometry in finite-connectivity Hopfield models. *Journal of Physics A Mathematical General*, v. 37, p. 7843–7858, Aug. 2004.
- [36] E. ORLANDINI, M. C. T.; WHITTINGTON, S. G. Self-averaging in the statistical mechanics of some lattice models. *J. Phys. A: Math. Gen.*, v. 35, p. 4219, 2002.
- [37] VAN HEMMEN, J. L.; PALMER, R. G. The replica method and solvable spin glass model. *J. Phys. A: Math. Gen.*, v. 12, p. 563–580, 1979.
- [38] KABASHIMA, Y.; SAAD, D. Statistical mechanics of error-correcting codes. *Europhys. Lett.*, v. 45, p. 97–103, 1999.
- [39] VAN HEMMEN, J. L.; KÜHN, R. Nonlinear neural networks. *Phys. Rev. Lett.*, v. 57, p. 913–916, Aug 1986.