

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

Scuola di Scienze  
Dipartimento di Fisica e Astronomia  
Corso di Laurea in Fisica

# Proprietà statistiche dell'apprendimento nella rete di Hopfield diluita

Relatore:  
Prof. Armando Bazzani

Presentata da:  
Matteo Shqemza

Correlatore:  
Dott. Giulio Colombini

Anno Accademico 2021/2022



## Sommario

La rete di Hopfield è una tipologia di rete neurale che si comporta come una memoria associativa. La caratteristica più importante per questo tipo di rete è la capacità  $p_{max}$ , cioè quanti dati, chiamati ricordi, può apprendere prima che cambi comportamento. Infatti, per un numero grande di ricordi, avviene una transizione di fase e la rete smette di comportarsi come desiderato. La capacità di una rete di Hopfield è proporzionale al numero di neuroni  $N$  che la compongono,  $p_{max} = \alpha_c N$ , dove  $\alpha_c = 0.138$ .

Una variante importante di questo modello è la rete di Hopfield diluita. In questa rete i neuroni non sono tutti connessi tra loro ma l'esistenza di connessioni tra due neuroni è determinata in modo stocastico con probabilità  $\rho$  di esserci e  $1 - \rho$  di essere assente. Il grafo di una rete così definita è un grafo aleatorio di Erdős-Rényi.

Il lavoro qui presentato ha lo scopo di studiare le proprietà statistiche dell'apprendimento di questo tipo di reti neurali, specialmente l'andamento della capacità in funzione del parametro  $\rho$  e le connettività del grafo durante le transizioni di fase che avvengono nel network. Attraverso delle simulazioni, si è concluso che la capacità di una rete di Hopfield diluita  $p_{max}$  segue un andamento a potenza  $p_{max} = aN\rho^b + c$ , dove  $N$  è il numero di neuroni,  $a = (0.140 \pm 0.003)$ ,  $b = (0.49 \pm 0.03)$ , e  $c = (-11 \pm 2)$ . Dallo studio della connettività del grafo è emerso che la rete funge da memoria associativa finché il grafo del network risulta connesso.



# Indice

<b>Introduzione</b>	<b>1</b>
<b>1 Teoria dei grafi</b>	<b>3</b>
1.1 Grafo . . . . .	3
1.2 Coefficiente di Clustering . . . . .	5
1.3 Matrice laplaciana . . . . .	6
1.4 Grafi di Erdős-Rényi . . . . .	7
<b>2 Reti di Hopfield</b>	<b>8</b>
2.1 Cenni storici . . . . .	8
2.2 Il modello originale . . . . .	11
2.3 Reti stocastiche . . . . .	12
2.4 Capacità di una rete stocastica . . . . .	16
<b>3 Reti di Hopfield diluite</b>	<b>22</b>
3.1 Introduzione alle reti di Hopfield diluite . . . . .	22
3.2 Capacità delle reti debolmente diluite . . . . .	23
3.3 Capacità delle reti fortemente diluite . . . . .	23
<b>4 Risultati</b>	<b>25</b>
4.1 Simulazioni delle reti di Hopfield . . . . .	25
4.2 Simulazioni della rete di Hopfield diluito . . . . .	28
<b>5 Conclusioni</b>	<b>34</b>
<b>A Esistenza di una distribuzione stazionaria in un network di Hopfield stocastico</b>	<b>35</b>



# Introduzione

Il network di Hopfield [19] è uno dei primi modelli di rete neurale, cioè una rete che simula il funzionamento del cervello umano, in questo caso il funzionamento della memoria.

Il funzionamento di questa tipologia di reti neurali è spiegato come segue: dopo aver appreso delle informazioni, chiamate ricordi, se viene presentata alla rete una nuova informazione essa si ricondurrà al ricordo più simile. Per esempio, se la rete apprende immagini di felini quando incontrerà la foto di un gattino riconoscerà da sola che quello è un gatto. Questo tipo di processo è detto memoria associativa. La peculiarità di questo sistema è dato dal ricreare una memoria solamente attraverso il concetto di funzione energia ai cui minimi si trovano i ricordi immagazzinati. La caratteristica principale di una rete è la capacità, ovvero la quantità di ricordi che riesce a immagazzinare. Variando il numero di ricordi immagazzinati, avvengono delle transizioni di fase che modificano il comportamento della rete. Per studiare la capacità e le transizioni si utilizzano tecniche e concetti caratteristici della meccanica statistica, come il concetto di energia libera e di parametro d'ordine, ottenendo che la capacità è proporzionale al numero di neuroni[1] [2], le cellule del cervello.

In questo tipo di reti il grafo rappresentante il network, dove i nodi sono i neuroni e gli archi rappresentano le sinapsi, cioè le connessioni tra i neuroni, è completo, cioè ogni neurone è connesso con tutti gli altri. Questo, però, non è ciò che succede nel cervello umano, i neuroni sono connessi solo con una piccola frazione dei neuroni totali. Da tale fatto nasce la necessità di studiare reti non complete, chiamate reti diluite, in cui le connessioni tra i neuroni sono casuali, per capire come variano le proprietà principali, soprattutto la capacità. In letteratura è dimostrato che per reti molto diluite la capacità rimane proporzionale al numero di neuroni per la probabilità di connessione degli archi. Esiste, però, un dissenso riguardo al punto in cui avviene la transizione di fase: secondo alcuni la transizione avviene quando il grafo smette di essere connesso [21]; secondo altri essa avviene prima [3] [4].

Questo lavoro si affronta il problema di studiare la capacità delle reti di Hopfield diluite soffermandosi sulle transizioni di fase che avvengono nel network. Al fine di studiare queste transizioni è stata costruita una simulazione in C++ di una rete di Hopfield diluita. Successivamente è stato osservato il suo comportamento al variare dei ricordi

## Introduzione

---

memorizzati e analizzato il grafo, specialmente la matrice laplaciana, al fine di ottenere informazioni utili riguardo la connettività all'interno del network.

Uno studio di questo tipo può essere utile per capire meglio questo tipo di reti ed avere una miglior conoscenza del funzionamento della memoria nel cervello, nel modo in cui esso apprende tramite l'esperienza e soprattutto come i danni a livello delle sinapsi possono interferire con questo processo. Queste informazioni possono essere applicate anche in campo medico e biologico per comprendere l'effetto dell'invecchiamento e della morte cellulare sulla memoria e il cervello in generale.

La tesi è strutturata in cinque capitoli. Nel primo capitolo è presentata una breve introduzione alla teoria dei grafi e sono definiti strumenti utili per l'analisi dei network. Nel secondo capitolo è spiegato dettagliatamente il modello di Hopfield. Nel terzo capitolo è trattato approfonditamente il modello di Hopfield diluito. Nel quarto capitolo sono riportati i risultati ottenuti dalle simulazioni. Nel quinto capitolo, infine, sono esposte le conclusioni basate sui dati ottenuti.

# Capitolo 1

## Teoria dei grafi

Prima di addentrarsi nello studio delle reti di Hopfield è utile un'introduzione alla teoria dei grafi. In questo capitolo vengono presentati strumenti utili per studiare la connettività dei grafi. Questi strumenti verranno utilizzati successivamente nelle analisi sulla connettività delle reti di Hopfield diluite.

### 1.1 Grafo

Per prima cosa è utile fare un'introduzione ai grafi e descrivere le loro caratteristiche più importanti.

Un grafo è un insieme di elementi detti nodi, o vertici, connessi tra loro da elementi detti archi. Formalmente

**Definizione 1.1.1** (Grafo). Un grafo è definito dalla coppia ordinata  $(V, E)$  dove  $V = \{1, 2, 3, \dots, n\}$  è l'insieme di nodi ed  $E = \{(i, j) : i, j \in V; i \text{ è connesso a } j\}$  è l'insieme di archi.

Per avere una piena descrizione del grafico si definisce la matrice delle adiacenze.

**Definizione 1.1.2.** (Matrice delle adiacenze) Sia  $A = (A_{i,j})$  una matrice quadrata binaria di ordine  $n$ . Essa è detta matrice delle adiacenze se

$$A_{i,j} = \begin{cases} +1 & \text{se } i \text{ è connesso a } j \\ 0 & \text{altrimenti} \end{cases}. \quad (1.1)$$

Una importante tipologia di grafi sono i grafi non orientati.

**Definizione 1.1.3.** (Grafo non orientato) Un grafo è detto non orientato se la matrice delle adiacenze è simmetrica. In questo caso due nodi connessi tra loro sono detti adiacenti.

## 1. Teoria dei grafi

---

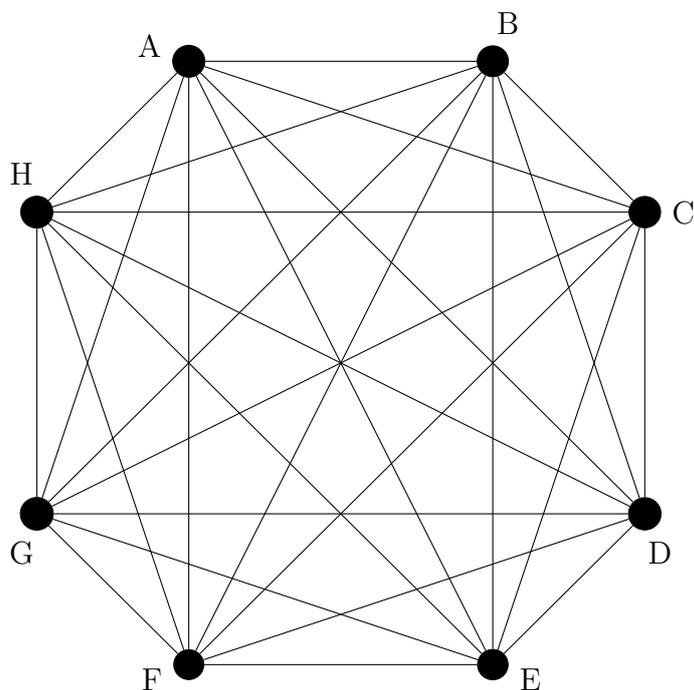


Figura 1.1: Esempio di grafo completo composto da 8 nodi e 64 archi.

Due tipi di grafi speciali sono il grafo nullo e il grafo completo.

**Definizione 1.1.4.** (Grafo nullo) Sia  $(V, E)$  un grafo, se  $E = \emptyset$  il grafo è detto nullo.

**Definizione 1.1.5.** (Grafo completo) Sia  $(V, E)$  un grafo, se tutti i nodi del grafo sono adiacenti tra loro il grafo è detto completo

Un'altra definizione utile è quella di sottografo.

**Definizione 1.1.6.** (Sottografo) Siano  $(V, E)$  e  $(V', E')$  con  $V' = \{x_1, \dots, x_k\}$  due grafi. Se  $V' \subseteq V$  e  $E' \subseteq E$  allora  $(V', E')$  è detto sottografo di  $(V, E)$ .

Soffermandosi sulle connessioni tra i nodi si possono definire il concetto di vicinato, grado e cammino.

**Definizione 1.1.7.** (Grado) Sia  $(V, E)$  un grafo, il grado del nodo  $i$  è definito come il numero di nodi adiacenti ad esso. Può essere espresso tramite la matrice dei nodi come

$$d_i = \sum_j A_{i,j}. \quad (1.2)$$

**Definizione 1.1.8.** (Cammino) Siano  $(V, E)$  un grafo e  $i, j \in V$  due nodi, un cammino di lunghezza  $k$  di estremi  $i$  e  $j$  è una sequenza di vertici  $(x_0, x_1, \dots, x_k)$  tale che  $x_0 = i$ ,

## 1. Teoria dei grafi

---

$x_k = j$ ,  $x_h \in V$  e  $(x_h, x_{h+1}) \in E$  per  $h = 0, 1, \dots, k-1$ . Un cammino è detto semplice se ogni arco è percorso una sola volta. Se per due nodi  $i, j \in V$  esiste un cammino di estremi  $i$  e  $j$ , i due vertici si dicono connessi.

Inoltre, si possono definire i concetti di ciclo, cappio, grafo connesso e componente connessa.

**Definizione 1.1.9.** (Ciclo) Un percorso in cui  $i = j$ , ovvero in cui i due estremi coincidono, si dice ciclo. Il ciclo è semplice se  $i = j$  è l'unico nodo che si ripete all'interno della sequenza.

**Definizione 1.1.10.** (Cappio) Un arco che ha due estremi coincidenti è detto cappio. Un cappio può essere visto anche come un ciclo di lunghezza 1.

**Definizione 1.1.11.** (Grafo connesso) Sia  $(V, E)$  un grafo, se ogni nodo  $i$  è connesso ad ogni nodo  $j$  per  $\forall i, j \in V$  il grafo è detto connesso.

**Definizione 1.1.12.** (Componente connessa) Una componente connessa di un grafo  $(V, E)$  è un sottografo  $(V', E')$  tale che ogni coppia di vertici  $i, j \in V'$  è connessa e ogni coppia di vertici  $i \in V'$  e  $j \notin V'$  non è connessa. Se il grafo è connesso esiste una sola componente connessa, il grafo stesso.

## 1.2 Coefficiente di Clustering

Il coefficiente di clustering è una misura di grado in cui i nodi di un grafo sono connessi tra loro. Si possono misurare due tipi di clustering, locale e globale. Questo coefficiente è utile per capire il modo in cui i neuroni all'interno della rete neurale sono interconnessi tra loro.

Il coefficiente di clustering locale  $C_i$  di un nodo  $i$  è dato dal numero di archi fra i nodi adiacenti ad esso diviso per il numero di archi potenziali fra loro.

**Definizione 1.2.1.** (Coefficiente di clustering locale) Il coefficiente di clustering locale per un grafo non orientato è definito come

$$C_i = \frac{2|\{(a, b) \in E : a, b \in N_i\}|}{d_i(d_i - 1)}, \quad (1.3)$$

Si può anche calcolare attraverso la matrice delle adiacenze

$$C_a = \frac{\sum_{b,c} A_{a,b}A_{b,c}A_{c,a}}{d_a(d_a - 1)} \quad (1.4)$$

Si può definire anche il coefficiente di clustering locale medio come

## 1. Teoria dei grafi

---

**Definizione 1.2.2.** (Coefficiente di clustering medio) Il coefficiente di clustering locale medio è definito come

$$\bar{C} = \frac{1}{n} \sum_{i=0}^n C_i. \quad (1.5)$$

Il coefficiente di clustering globale, invece, si basa sullo studio di triplette di nodi. Una tripletta è composta da tre nodi e 2 o 3 archi, rispettivamente tripletta aperta e chiusa.

**Definizione 1.2.3.** (Coefficiente di clustering globale) Il coefficiente è dato dal numero di triplette chiuse fratto il numero di triplette totali, aperte e chiuse. Per un grafo non orientato può essere espresso tramite la matrice delle adiacenze

$$C = \frac{\sum_{a,b,c} A_{a,b} A_{b,c} A_{c,a}}{\sum_i d_i (d_i - 1)} \quad (1.6)$$

### 1.3 Matrice laplaciana

La matrice laplaciana contiene al suo interno numerose informazioni riguardanti il grafo, in particolare il numero di componenti connesse che lo formano. Nello studio delle reti neurali è uno strumento utile per studiare la connettività della rete.

La matrice laplaciana è definita come

$$L_{ij} = d_i \delta_{i,j} - A_{i,j}, \quad (1.7)$$

dove  $A_{i,j}$  è la matrice delle adiacenze e  $d_i$  il grado del nodo.

La matrice è caratterizzata dalla presenza di elementi positivi unicamente sulla diagonale, mentre il resto della matrice è negativa o nulla. In caso di grafo non orientato essa risulta simmetrica.

L'analisi dello spettro della matrice laplaciana mostra la presenza di un autovalore nullo, rendendo la matrice non invertibile. Inoltre, la parte reale degli autovalori è sempre non negativa  $\text{Re}(\lambda) \geq 0$ .

Come accennato precedentemente, studiando gli autovalori della matrice laplaciana è possibile studiare la connettività del grafo. Infatti, il numero di componenti connesse è uguale alla dimensione del kernel della matrice, in altre parole il numero di componenti connesse è pari alla molteplicità algebrica dell'autovalore nullo. Da ciò deriva l'importanza del secondo autovalore più piccolo  $\lambda$ , chiamato connettività algebrica o autovalore di Fiedler, che determina se un grafo è connesso,  $\lambda > 0$ , oppure no,  $\lambda = 0$ . In caso il grafo sia completamente connesso si ha  $\lambda = N$ .

## 1.4 Grafi di Erdős-Rényi

Il grafo di una rete di Hopfield diluita è un grafo in cui le connessioni tra i nodi sono scelte in modo aleatorio. In teoria dei grafi, il modello più utilizzato per studiare questa tipologia di grafi è il modello di Erdős-Rényi [11] [10]. In questo modello sono studiati grafi aleatori  $(n, N)$  in cui gli archi vengono scelti in modo casuale. In modo indipendente da questi, Gilbert [13] propose un modello simile in cui il grafo è composto di  $n$  nodi e la probabilità di avere un arco tra i nodi  $i$  e  $j$  è uguale a  $\rho$ . Le due formulazioni, sebbene siano diverse, possono essere intercambiate in regime termodinamico, cioè per  $n \rightarrow \infty$ , se si studiano proprietà monotone del grafo. Per facilità dei calcoli si utilizza comunemente la formulazione di Gilbert.

Una proprietà importante dei grafi di Erdős-Rényi è la distribuzione di probabilità dei nodi, infatti la probabilità  $P_k$  che un nodo sia di grado  $d_i = k$  è data da

$$P_k = \binom{n-1}{k} \rho^k (1-\rho)^{n-1-k}, \quad (1.8)$$

la distribuzione è di Bernoulli.

In caso  $n \rightarrow \infty$  e  $n\rho$  costante, la distribuzione tende a una distribuzione di Poisson con probabilità  $P_k$  descritta da

$$P_k = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (1.9)$$

dove  $\lambda = n\rho$  è la media della distribuzione.

Nello studio dei grafi casuali è importante il concetto di componente gigante, ovvero una componente del grafo che contiene una frazione finita dei nodi.

**Definizione 1.4.1** (Componente gigante). Sia  $(n, N)$  un grafo di Erdős-Rényi, se il grafo ha una componente connessa  $(V, E)$  tale che  $V \approx n$ . In generale, se esiste una componente gigante essa è unica

Nei loro studi Erdős e Rényi [11] descrissero il comportamento di un grafo al variare di  $\rho$ . Dimostrarono che sia  $(V, N)$  un grafo e  $\rho$  la probabilità che un arco esista allora:

- se  $\rho < \frac{1}{n}$  il grafo è composto da componenti di dimensione  $\mathcal{O}(\log(n))$ ;
- se  $\rho > \frac{1}{n}$  il grafo contiene una componente gigante;
- se  $\rho > \frac{\log(n)}{n}$  il grafo risulta connesso.

# Capitolo 2

## Reti di Hopfield

In questo capitolo si analizza la rete di Hopfield e le sue principali caratteristiche. La prima sezione tratta un'introduzione storica all'argomento. La seconda e la terza espongono il modello e le sue tipologie più utilizzate, il network deterministico e quello stocastico. La quarta studia la caratteristica principale di una rete, ovvero la capacità, al variare dei parametri fondamentali del network.

### 2.1 Cenni storici

In origine lo studio delle reti neurali era focalizzato sullo studio delle interconnessioni reali che si formano tra i neuroni. Con il perfezionamento dei computer ci si rese conto che, nonostante i progressi ingegneristici, il cervello umano rimaneva molto più veloce nel risolvere azioni complicate, come il riconoscimento facciale, con un consumo minimo di energia. Questo, insieme ad altre capacità come la flessibilità di adattamento agli stimoli e la resistenza alla morte continua di neuroni, hanno spinto maggiormente nello studio di reti neurali e nella loro applicazione pratica.

Il cervello umano si occupa di molte funzioni all'interno del corpo, da quelle più semplici come raccogliere le informazioni sensoriali provenienti dall'esterno e coordinare i movimenti per arrivare a quelle più complesse come il linguaggio e i pensieri. Esso è formato da circa  $10^{11}$  cellule, chiamate neuroni, di forma ramificata e in continua comunicazione tra di loro e con il resto del corpo.

La maggior parte di queste diramazioni, chiamate dendriti, sono a corto raggio e captano i segnali provenienti dalle altre cellule; invece i rami a lungo raggio, chiamati assoni, trasmettono il segnale dal neurone alle altre cellule. All'estremità di questi rami ci sono le sinapsi, il punto in cui l'assone entra in comunicazione con le altre cellule. Un assone tipicamente comunica con svariate migliaia di altri neuroni.

La trasmissione del segnale nelle sinapsi è un complesso processo chimico e biologico in cui l'assone emette particolari composti chimici, chiamati neurotrasmettitori, che

## 2. Reti di Hopfield

---

vengono captati da specifici recettori della cellula destinataria. I neurotrasmettitori hanno la capacità di modificare il potenziale elettrico all'interno del neurone destinatario. Se il potenziale elettrico supera una certa soglia, diversa per ogni neurone, un impulso di durata e potenza fissati, chiamato potenziale d'azione, viene spedito lungo l'assone ai neuroni successivi. Un neurone in questo stato è detto eccitato. I neuroni comunicano anche con le altre cellule del corpo attraverso gli ormoni, come l'adrenalina, nel flusso sanguigno. Questi ormoni possono a loro volta influire sullo stato dei neuroni.

In figura 2.1 è mostrato una delle prime raffigurazioni di un neurone ad opera di Santiago Ramón y Cajal.

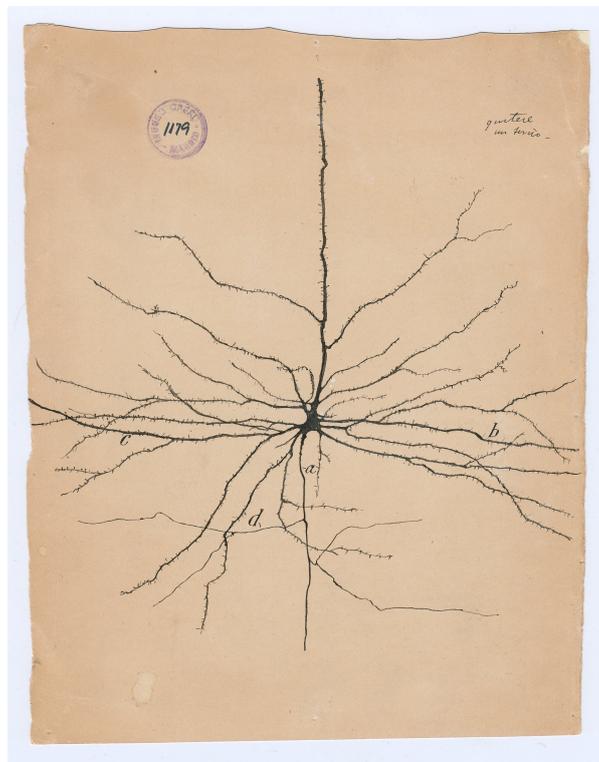


Figura 2.1: Santiago Ramón y Cajal, Neuroni piramidali della corteccia cerebrale, 1904 inchiostro e matita su carta, Crediti: Cajal Institute (CSIC), Madrid

È noto che lo stato del cervello è descritto dalla sua attività neurale, cioè da quali neuroni sono eccitati e da quali no. Questa attività varia con il passare del tempo in base al tipo di sinapsi tra i neuroni. Molte di queste sinapsi sono determinate dai geni e corrispondono ad azioni indispensabili per la vita, mentre alcune possono essere modificate, anche solo in parte, degli agenti esterni; infatti si è visto che l'esperienza e la memoria possono modificare o anche creare nuove sinapsi. Questo meccanismo è chiamato imprinting ed è indispensabile per l'apprendimento.

## 2. Reti di Hopfield

---

Il modello più famoso e studiato per l'imprinting fu proposto da Hebb [15] nel 1949. La regola di apprendimento da lui formulata dice che quando, in una particolare configurazione, due neuroni sono entrambi eccitati la sinapsi eccitatoria tra loro si rafforza mentre quella inibitoria si indebolisce. Quando, invece, uno è eccitato e l'altro no, la sinapsi eccitatoria si indebolisce e quella inibitoria si rafforza. Questo meccanismo permette di ricostruire l'attività neurale partendo da parte di essa.

Nel 1943 McCulloch e Pitts [22] proposero un modello semplice per spiegare questo processo. Esso si basa su una somma pesata dei segnali in ingresso paragonandoli con una soglia prefissata. L'evoluzione del neurone  $\sigma_i$  può essere riassunta come

$$\sigma_i(t+1) = \Theta \left( \sum_j w_{ij} \sigma_j(t) - \theta_i \right), \quad (2.1)$$

dove  $\sigma_j$  è lo stato dei neuroni in comunicazione con  $\sigma_i$  con i rispettivi pesi  $w_{ij}$ ,  $\theta_i$  è la soglia da superare e  $\Theta$  è la funzione a gradino di Heaviside. In questo modello quando il neurone è eccitato  $\sigma_i = 1$ , quando è a riposo  $\sigma_i = 0$ . I pesi  $w_{ij}$  rappresentano l'importanza della connessione tra l' $i$ -esimo neurone e il  $j$ -esimo, possono essere sia positivi che negativi a seconda se i neurotrasmettitori siano rispettivamente eccitatori o inibitori.

Questo primo modello, però, non racchiude al suo interno tutta la complessità riscontrata nel cervello, un neurone reale non ha una soglia così netta, ma ha una risposta adeguata all'intensità del segnale in entrata e la quantità di neurotrasmettitori emessi varia in maniera casuale.

Nonostante ciò, il modello di McCulloch e Pitts fu applicato anche in campi molto diversi dalla biologia come le scienze computazionali, sia in campo hardware che campo software.

Da questo modello nacquero nuovi studi che, grazie all'aiuto di modelli fisici preesistenti, permisero di fare molti passi avanti. Nel 1954 Cragg e Temperley [8] riscrissero il modello di McCulloch e Pitts come una rete di spin magnetici con l'utilizzo del modello di Ising. In questo caso la memoria risiede nel fenomeno dell'isteresi nel dominio dei ricordi atteso per il sistema. Nel 1961 Caianiello [5] costruì una teoria statistica usando il formalismo della meccanica statistica e la regola d'apprendimento di Hebb. Nel 1982 Hopfield [19] aggiunse a questo modello una funzione energia e formulò una teoria dove i ricordi fungono da attrattori. Successivamente Hinton e Sejnowski [18] [17] e Peretto [23] aggiunsero casualità al modello di Hopfield utilizzando una dinamica stocastica in grado di compiere errori in analogia al concetto di temperatura in meccanica statistica. Infine, Amit et al. [1] [2] applicando metodi matematici appresi studiando il vetro di spin, delinearono le potenzialità e i limiti del modello di Hopfield.

Successivamente nacquero molte variazioni del modello di Hopfield, rompendo simmetrie, considerando particolari connessioni tra neuroni o aggiungendo la nozione di distanza. Una variazione importante è il modello di Hopfield diluito in cui non tutti i

## 2. Reti di Hopfield

---

neuroni sono connessi tra loro. Inizialmente il problema fu affrontato da Canning e Gardner [6] e Deridda, Gardner e Zippelius [9], per poi essere trattato in maniera rigorosa da Bovier e Gayraud [3] [4].

### 2.2 Il modello originale

La rete di Hopfield [19], chiamata anche rete neurale attrattiva, ideata nel 1982 da J. J. Hopfield, è un modello del comportamento dei neuroni nel cervello, focalizzandosi nei processi riguardanti la memoria associativa.

Per prima cosa, sia  $N$  il numero totale di neuroni. Ogni neurone è considerato come variabile binaria  $\sigma_i = \{\pm 1\}$  per indicare se il neurone è eccitato (+1) oppure a riposo (-1). Le sinapsi sono rappresentate attraverso dei pesi  $w_{ij}$ , quando non c'è connessione tra due neuroni  $w_{ij} = 0$ . Nel modello i pesi sono per definizione simmetrici e ogni neurone è connesso con tutti gli altri, anche se ciò non è riscontrabile nel cervello umano. Tutte le caratteristiche della rete sono determinate dai pesi.

Il modello originale è caratterizzato da una dinamica deterministica data da

$$\sigma_i(t+1) = \operatorname{sgn}\left(\sum_j w_{ij}\sigma_j\right), \quad (2.2)$$

dove  $\operatorname{sgn}(\cdot)$  è la funzione segno.

Avendo imposto la simmetria dei pesi  $w_{ij} = w_{ji}$ , possiamo definire una funzione energia

$$H(\sigma) = -\frac{1}{2} \sum_{ij} w_{ij}\sigma_i\sigma_j. \quad (2.3)$$

Si può notare che, utilizzando la dinamica deterministica 2.2, all'aumentare di  $t$  l'energia tende a diminuire. Infatti, ipotizzando il cambiamento dell' $i$ -esimo neurone secondo la 2.2 la variazione di energia d'energia è data da

$$\Delta H = 2\sigma_i \sum_j w_{ij}\sigma_j < 0. \quad (2.4)$$

Con il passare del tempo l'energia diminuisce finché non raggiunge un minimo locale. Questo tipo di evoluzione crea dei bacini d'attrazione che spingono la configurazione di neuroni  $\{\sigma\}$  verso il minimo dell'energia più vicino. Se i minimi locali corrispondono a un ricordo immagazzinato precedentemente nella rete questo processo può essere descritto come una memoria associativa, cioè dato un ricordo corrotto il network ricostruisce il ricordo originale. Ciò può essere ottenuto impostando i pesi secondo la regola d'apprendimento di Hebb [15]: sia  $p$  il numero di ricordi  $\xi^\mu$  con  $\mu = 1, 2, \dots, p$ , cioè configurazioni

## 2. Reti di Hopfield

---

di neuroni particolari che rappresentano le informazioni apprese dalla rete, i pesi sono determinati da

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

Il coefficiente  $\frac{1}{N}$  è una costante di normalizzazione.

Si nota che quando  $i = j$  modificando lo stato dell' $i$ -esimo neurone non c'è variazione di energia, poiché  $\sigma_i^2 = 1$ . Di conseguenza la 2.3 può essere riscritta come

$$H(\sigma) = -\frac{1}{2} \sum_{i \neq j} w_{ij} \sigma_i \sigma_j - \frac{1}{2} \sum_i w_{ii}. \quad (2.6)$$

Essendo l'ultimo termine una costante il suo valore non contribuisce all'andamento dell'energia e si può considerare nullo. Questo accorgimento rende il modello più vicino alla realtà cancellando le interazioni dei neuroni con se stessi, cosa che non avviene nel cervello umano. Per il resto della trattazione si considera  $w_{ii} = 0$ .

La rete di Hopfield così definita funge da memoria associativa con proprietà molto simili alla memoria umana e distanti dai meccanismi di memoria utilizzati oggi in informatica.

La proprietà più importante di questo tipo di reti è il numero massimo di ricordi che riescono ad immagazzinare, chiamata capacità. Attraverso uno studio qualitativo con simulazioni numeriche emerge che con un numero di ricordi piccoli questi sono minimi della funzione energia, ma non sono gli unici. Ovviamente  $-\xi^{\mu}$  è un minimo, per la simmetria dell'energia rispetto allo scambio  $\sigma_i \rightarrow -\sigma_i$  per ogni  $i$ ; ciò non è un problema dato che è possibile impostare dei parametri di controllo per invertire la configurazione dei neuroni nei casi opportuni. Anche le combinazioni lineari tra un numero dispari di ricordi sono minimi, chiamati minimi misti. Al crescere di  $p$  emergono altri minimi non correlati con alcun ricordo, chiamati minimi spuri. Infine quando il numero dei ricordi  $p$  è molto grande, i bacini d'attrazione diventano molto piccoli e la rete non si riconduce al ricordo corrispondente. Questo stato, simile al vetro di spin in meccanica statistica, si raggiunge quando  $p \approx 0.138N$ .

### 2.3 Reti stocastiche

Per rendere la rete neurale meno suscettibile ai minimi spuri si può aggiungere del rumore casuale alla rete, cioè una pseudo-temperatura. Come descritto da Hinton e Sejnowski [18] e Peretto [23], l'evoluzione della rete a partire da una configurazione nota di neuroni al tempo  $t$  è data dalla dinamica stocastica di Glauber [14] per gli spin per il modello di Ising [20] descritta come

## 2. Reti di Hopfield

---

$$\sigma_i(t+1) = \begin{cases} +1 & \text{con probabilità } p(h_i(t)) \\ -1 & \text{con probabilità } 1 - p(h_i(t)) \end{cases}, \quad (2.7)$$

e la probabilità  $p(h_i(t))$  è data

$$p(h_i) = \frac{1}{1 + e^{-2\beta h_i}}, \quad (2.8)$$

dove  $h_i(t) = \sum_j w_{ij}\sigma_j$  indica l'interazione del  $i$ -esimo neurone con gli altri e  $\beta = \frac{1}{T}$ .

L'equazione 2.8 è chiamata funzione logistica, mostrata in figura 2.2.  $T$  è una pseudo-temperatura che simula le fluttuazioni che possono avvenire a livello delle sinapsi, non corrisponde ad una temperatura fisica ma è solamente un parametro per la pendenza della funzione logistica vicino all'origine. Non essendo una grandezza fisica misurabile, la temperatura  $T$  è adimensionale e non necessita l'introduzione di un'eventuale costante di Boltzmann  $K_B$ . Azzerando la temperatura, l'equazione 2.8 si riduce alla funzione a gradino e la dinamica espressa dall'equazione 2.7 si riconduce alla dinamica deterministica dell'equazione 2.2.

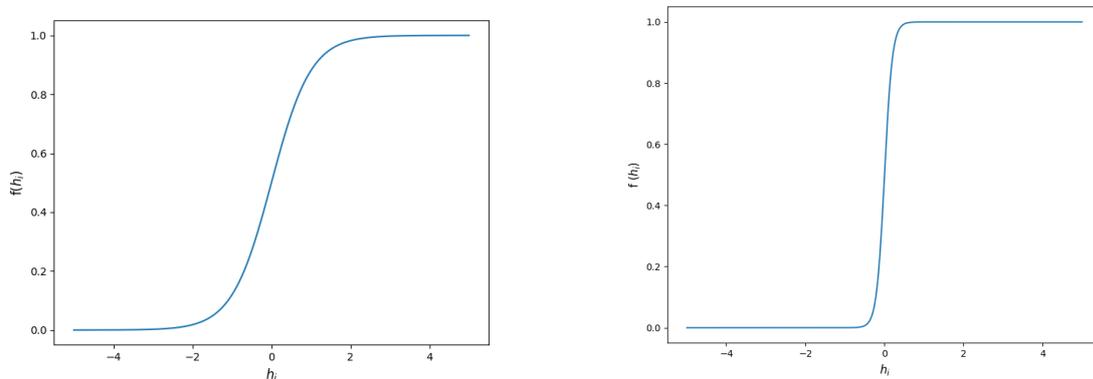


Figura 2.2: Funzione logistica. A sinistra  $\beta = 1$ , a destra  $\beta = 5$ . Si può notare come per  $\beta \rightarrow \infty$  la funzione tendi a una funzione a gradino.

L'utilizzo di una dinamica stocastica rende più facile lo studio della capacità della rete attraverso tecniche già sviluppate in meccanica statistica. Inoltre, la presenza di rumore rende la rete meno soggetta ad errori, permettendo al sistema di uscire dal bacino d'attrazione dei minimi spuri. Infatti spesso questi minimi sono maggiormente instabili rispetto ai ricordi.

Come per i sistemi termodinamici, la rete continuerà a evolvere, con il passare del tempo, senza fine. La rete non si fermerà mai in uno dei punto di stabilità, minimi della

## 2. Reti di Hopfield

---

funzione energia. Però esistono dei punti di equilibrio  $\{\langle\sigma_i\rangle\}$  in cui la media temporale delle grandezze di riferimento diventano indipendenti dal tempo.

Considerando il caso in cui la rete neurale ha interazioni uniformi, i pesi sono dati da

$$w_{ij} = w_{ji} = \frac{1}{N}\xi_i\xi_j = \frac{W}{N}. \quad (2.9)$$

In questo caso ci si può ricondurre al modello fisico dei ferromagnetici a range infinito ed è risolvibile attraverso la teoria di campo medio. La funzione energia 2.3 si riconduce all'hamiltoniana del modello di Ising

$$H(\sigma) = -\frac{NW}{2} \langle\sigma_i\rangle^2 + \frac{W}{2}, \quad (2.10)$$

con  $\langle\sigma_i\rangle$  è definito come

$$\langle\sigma_i\rangle = \frac{1}{N} \sum_i \sigma_i. \quad (2.11)$$

Dato che il parametro  $\langle\sigma_i\rangle$  è una media del valore dei singoli neuroni può essere espressa come

$$\langle\sigma_i\rangle = p(+1)(+1) + p(-1)(-1), \quad (2.12)$$

dove  $p(\pm 1)$  è la probabilità di trovare uno spin nello stato  $\sigma_i = \pm 1$  secondo la distribuzione di probabilità 2.8. Sostituendo si ottiene

$$\langle\sigma_i\rangle = \frac{1}{1 + e^{2\beta h_i}} - \frac{1}{1 + e^{-2\beta h_i}} = \tanh(\beta h_i), \quad (2.13)$$

dove  $\tanh(\cdot)$  è la funzione tangente iperbolica mostrata in figura 2.2.

L'interazione tra neuroni  $h_i$  per un solo ricordo è data da

$$h_i = \sum_j \frac{W}{N} \sigma_j = W \frac{1}{N} \sum_j \sigma_j = W \langle\sigma_i\rangle. \quad (2.14)$$

Sostituendo nella 2.13 si ottiene

$$\langle\sigma_i\rangle = \tanh \beta W \langle\sigma_i\rangle. \quad (2.15)$$

Questa equazione, nota come l'equazione di Curie-Weiss per il magnetismo, ha una sola soluzione  $\langle\sigma_i\rangle = 0$  per  $T = \frac{1}{\beta} \gg 1$  e tre soluzioni per  $T \ll 1$ . Il cambio di comportamento avviene quando  $\beta W = 1$ . Se  $W < 0$  esiste solo la soluzione nulla. Ciò si può vedere anche dal grafico in figura 2.3.

In maniera più rigorosa si può espandere la tangente iperbolica con la serie di Taylor attorno a  $\langle\sigma_i\rangle = 0$  ottenendo

## 2. Reti di Hopfield

---

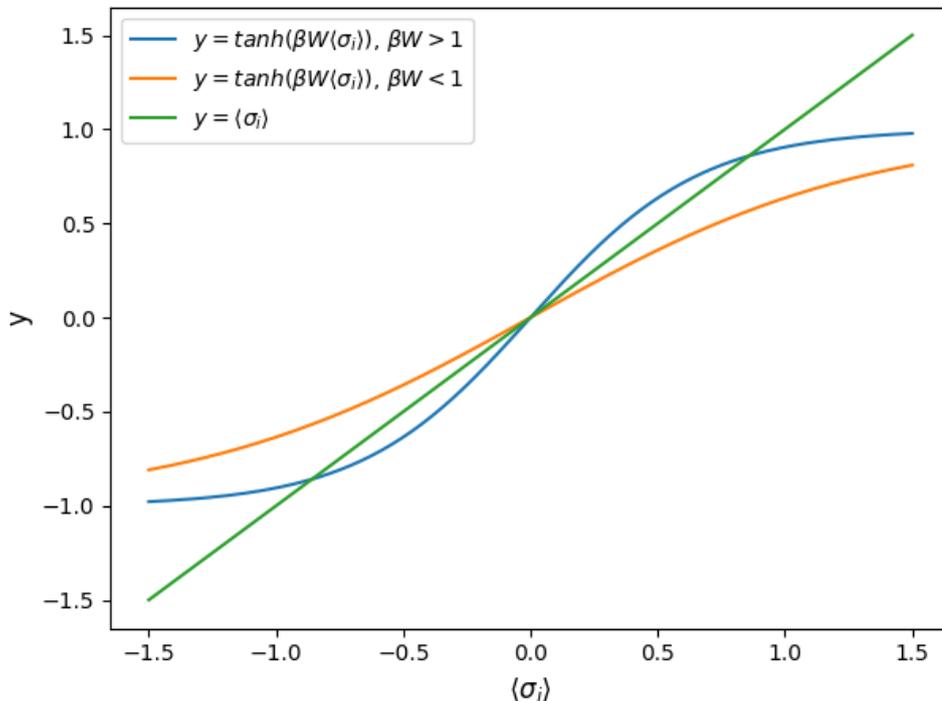


Figura 2.3: Soluzione grafiche dell'equazione di Curie-Weiss, cioè  $\langle \sigma_i \rangle = \tanh \beta W \langle \sigma_i \rangle$ . In blu e arancione le tangenti iperboliche rispettivamente con  $\beta W > 1$  e  $\beta W < 1$ , in verde la retta  $y = \langle \sigma_i \rangle$ . Le soluzioni sono date dai punti d'intersezione della tangente iperbolica con la retta. Per  $\beta W < 1$  l'unica soluzione è  $\langle \sigma_i \rangle = 0$ , per  $\beta W > 1$  si riscontrano due soluzioni aggiuntive.

$$\langle \sigma_i \rangle = \beta W \langle \sigma_i \rangle - \frac{1}{3}(\beta W)^3 \langle \sigma_i \rangle^3 + \mathcal{O}(\langle \sigma_i \rangle^5). \quad (2.16)$$

Trascurando i termini di ordine maggiore al terzo, la soluzione non nulla può essere scritta come

$$\langle \sigma_i \rangle^2 = \frac{3(\beta W - 1)}{(\beta W)^3}. \quad (2.17)$$

Per  $\beta W \approx 1$ , il numeratore domina sul denominatore;  $\langle \sigma_i \rangle$  si può scrivere come

$$\langle \sigma_i \rangle = \pm \sqrt{3(\beta W - 1)}. \quad (2.18)$$

## 2. Reti di Hopfield

---

Questa equazione è valida fintanto che  $\beta W > 1$ . Definendo  $T_c = \frac{1}{\beta_c} = W$  la temperatura critica in cui avviene il cambio di comportamento, la 2.18 può essere riscritta come

$$\langle \sigma_i \rangle = \pm \sqrt{\frac{3(T_c - T)}{T_c}}. \quad (2.19)$$

Il brusco cambio di comportamento di  $\langle \sigma_i \rangle$  per  $T = T_c$  è indice di una transizione di fase. Infatti, superata la temperatura critica,  $T < T_c$ , i minimi corrispondenti ai ricordi smettono di essere attrattori e la rete non funge da una memoria associativa. Ogni punto di minimo ha la propria temperatura critica per cui, superandola, questo smette di essere un attrattore. Di solito la temperatura critica dei ricordi  $T_c$  è maggiore di quella dei minimi spuri e misti  $T_s$ . Per un buon funzionamento della rete di Hopfield la temperatura  $T$  deve essere compresa tra la temperatura critica dei minimi spuri e misti e la temperatura critica della rete,  $T_s < T < T_c$ .

### 2.4 Capacità di una rete stocastica

Studiare la capacità di questo tipo di reti presenta notevoli difficoltà. Successivamente è esposta una trattazione qualitativa ad opera di Geszti [12] descritta in [16].

Per prima cosa, considerando il limite termodinamico  $N \rightarrow \infty$  e  $p \rightarrow \infty$ , dove  $N$  è il numero di neuroni e  $p$  è il numero di ricordi memorizzati dalla rete, si definisce un nuovo parametro  $\alpha = \frac{p}{N}$  e il parametro d'ordine  $m^\mu$ , il quale indica la sovrapposizione tra la configurazione  $\{\sigma\}$  e il ricordo  $\xi^\mu$ .

$$m^\mu(\sigma) = \frac{1}{N} \sum_i \xi_i^\mu \langle \sigma_i \rangle. \quad (2.20)$$

Considerando il sistema simile al primo ricordo, si definisce la grandezza  $r$  che indica la media quadrata della sovrapposizione della configurazione  $\{\sigma\}$  con i ricordi diversi dal primo.

$$r = \frac{1}{\alpha} \sum_{\mu \neq 1} (m^\mu)^2. \quad (2.21)$$

Il coefficiente  $\frac{1}{\alpha} = \frac{N}{p}$  permette di cancellare l'andamento  $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$  di  $m^\mu$  per  $\mu \neq 1$ .

Utilizzando la teoria di campo medio e la 2.15, si ottiene

$$m^\nu = \frac{1}{N} \sum_i \xi_i^\nu \tanh \left( \beta \sum_\mu \xi_i^\mu m^\mu \right). \quad (2.22)$$

## 2. Reti di Hopfield

---

Per  $\nu \neq 1$  l'argomento della tangente iperbolica si può separare in due termini

$$m^\nu = \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[ \beta \left( m^1 + \xi_i^\nu \xi_i^1 m^\nu + \sum_{\mu \neq i, \nu} \xi_i^\mu \xi_i^1 m^\mu \right) \right]. \quad (2.23)$$

Si è potuto portare  $\xi_i^1$  fuori da  $\tanh$  sfruttando la proprietà  $\tanh(\pm x) = \pm \tanh(x)$ .

Per ipotesi il secondo termine è molto minore del primo, che è circa 1, quindi possiamo espanderlo come

$$m^\nu = \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[ \beta \left( m^1 + \sum_{\mu \neq i, \nu} \xi_i^\mu \xi_i^1 m^\mu \right) \right] + \frac{\beta}{N} \sum_i \left\{ 1 - \tanh^2 \left[ \beta \left( m^1 + \sum_{\mu \neq i, \nu} \xi_i^\mu \xi_i^1 m^\mu \right) \right] \right\} \quad (2.24)$$

Assumiamo che le piccole sovrapposizioni  $m^\mu$  con  $\mu \neq 1$  siano variabili stocastiche indipendenti con media nulla e varianza  $\alpha r p^{-1}$ . Nella seconda parte di 2.24,  $\xi_i^\mu \xi_i^1$  sono variabili stocastiche e indipendenti da  $m^\mu$ , quindi applicando il teorema del limite centrale, il secondo termine diventa la media del “rumore” gaussiano  $\sum_{\mu \neq i, \nu} \xi_i^\mu \xi_i^1 m^\mu$  con varianza  $\alpha r$ . La 2.24 si riduce a

$$m^\nu = \frac{1}{N} \sum_i \xi_i^\nu \xi_i^1 \tanh \left[ \beta \left( m^1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m^\mu \right) \right] + \beta m^\nu + \beta q m^\nu \quad (2.25)$$

esprimibile anche come

$$m^\nu = \frac{\sum_i \xi_i^\nu \xi_i^1 \tanh \left[ \beta \left( m^1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m^\mu \right) \right]}{N[1 - \beta(1 - q)]}, \quad (2.26)$$

dove

$$q = \int \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \tanh^2[\beta(m^1 + \sqrt{\alpha r}z)]. \quad (2.27)$$

Per stimare  $r$  si procede in questo modo. Per prima cosa si ricava  $(m^\nu)^2$  da 2.26 ottenendo

$$(m^\mu)^2 = \frac{1}{N^2[1 - \beta(1 - q)]^2} \sum_{i,j} \xi_i^\nu \xi_i^1 \xi_j^\nu \xi_j^1 \times \tanh \left[ \beta \left( m^1 + \sum_{\mu \neq 1, \nu} \xi_i^\mu \xi_i^1 m^\mu \right) \right] \times \tanh \left[ \beta \left( m^1 + \sum_{\mu \neq 1, \nu} \xi_j^\mu \xi_j^1 m^\mu \right) \right]. \quad (2.28)$$

## 2. Reti di Hopfield

---

Successivamente si media sui ricordi. Dato che il  $\nu$ -esimo ricordo non compare all'interno di  $\tanh$ ,  $\xi_i^\nu \xi_i^1 \xi_j^\nu \xi_j^1$  e le due tangenti iperboliche possono essere trattate separatamente. Gli unici termini che rimangono della sommatoria sono quelli per cui  $i = j$ . In questo caso la media della tangente iperbolica restituisce  $q$ . Il risultato finale risulta indipendente da  $\nu$  ed è

$$r = \frac{q}{[1 - \beta(1 - q)]^2}. \quad (2.29)$$

Rimane ora da ricavare anche  $m^1$ . Ripercorrendo gli stessi passaggi a partire dalla [2.22](#) si ottiene facilmente

$$m^1 = \int \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \tanh[\beta(m^1 + \sqrt{ar}z)]. \quad (2.30)$$

Le equazioni [2.30](#), [2.29](#) e [2.27](#) si possono risolvere simultaneamente. Generalmente ciò viene fatto numericamente, mentre nel limite  $T \rightarrow 0$ , i.e.  $\beta \rightarrow \infty$ , si può trovare una soluzione analitica usando le identità

$$\int \frac{dz}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \{1 - \tanh^2[\beta(a + bz)]\} = \sqrt{\frac{2}{\pi}} \frac{1}{b\beta} e^{-\frac{a^2}{2b^2}}, \quad (2.31)$$

e

$$\int \frac{dz}{\sqrt{2\pi}} e^{\frac{1}{2}z^2} \tanh[\beta(a + bz)] = \operatorname{erf}\left(\frac{a}{\sqrt{2}b}\right), \quad (2.32)$$

dove  $\operatorname{erf}(\cdot)$  è la funzione degli errori, mostrata in [figura 2.4](#), definita come

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (2.33)$$

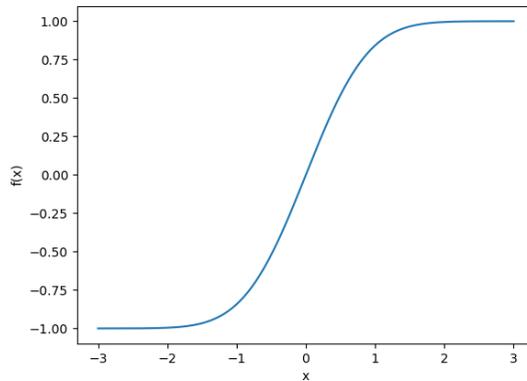


Figura 2.4: Funzione degli errori

## 2. Reti di Hopfield

---

Le tre equazioni diventano

$$C = \beta(1 - q) = \sqrt{\frac{2}{\pi\alpha r}} e^{-\frac{m^1}{2\alpha r}}, \quad (2.34)$$

$$r = \frac{1}{(1 - C)^2}, \quad (2.35)$$

$$m^1 = \operatorname{erf}\left(\frac{m^1}{\sqrt{2\alpha r}}\right). \quad (2.36)$$

Facendo un cambio di variabile  $x = \frac{m^1}{\sqrt{2\alpha r}}$  si ottiene

$$x \left( \sqrt{2\alpha} + \frac{2}{\sqrt{\pi}} e^{-x^2} \right) = \operatorname{erf}(x). \quad (2.37)$$

La 2.37 è facilmente risolvibile graficamente come mostrato in figura 2.5. Si può chiaramente vedere che per  $\alpha > \alpha_c$  esiste solo la soluzione nulla. Da analisi numeriche emerge che  $\alpha_c = 0.138$ .

Una trattazione più rigorosa è data in [7]. Ridefinendo alcune grandezze come la sovrapposizione con il  $\mu$ -esimo ricordo  $m^\mu$  e la funzione energia  $H$  e introducendo nuove grandezze derivate dalla meccanica statistica, come la probabilità di Gibbs e l'energia libera di Helmholtz, si può giungere allo stesso risultato.

Per prima cosa si ridefinisce il parametro d'ordine  $m^\mu$

$$m^\mu = \frac{1}{N} \sum_i \xi_i^\mu \sigma_i. \quad (2.38)$$

L'energia del sistema si può scrivere

$$H(\sigma) = -\frac{1}{2} \sum_{ij} \left( \frac{1}{N} \sum_\mu \xi_i^\mu \xi_j^\mu \right) \sigma_i \sigma_j = -\frac{N}{2} \sum_\mu \left( \sum_i \xi_i^\mu \sigma_i \right)^2. \quad (2.39)$$

La 2.39 diventa

$$H = -\frac{N}{2} \sum_\mu (m^\mu)^2. \quad (2.40)$$

Si definisce la funzione di partizione  $Z$

$$Z = \sum_\sigma e^{-\beta H(\sigma)}, \quad (2.41)$$

la probabilità di trovare il sistema in una certa configurazione come

## 2. Reti di Hopfield

---

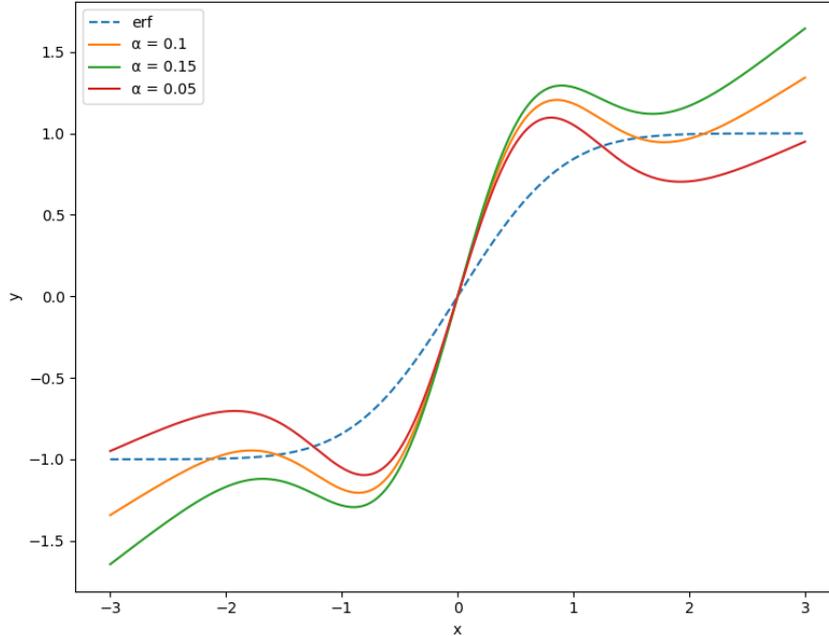


Figura 2.5: Soluzione grafica della 2.37. Le soluzioni sono date dall'intersezione della funzione degli errori, curva tratteggiata, e la funzione  $x \left( \sqrt{2\alpha} + \frac{2}{\sqrt{\pi}} e^{-x^2} \right)$ , curve continue. Sono state disegnate 3 curve continue con diversi valori di  $\alpha$ , cioè  $\alpha = 0.15$ ,  $\alpha = 0.10$ ,  $\alpha = 0.05$ . Si nota che per  $\alpha > \alpha_c$  c'è solo la soluzione nulla.

$$p(\sigma) = \frac{1}{Z} e^{-\beta H(\sigma)} \quad (2.42)$$

e l'energia libera di Helmholtz

$$F = -\frac{1}{\beta} \ln Z. \quad (2.43)$$

Studiando queste quantità al variare di  $\alpha$  e di  $T$  e utilizzando le tecniche apprese nello studio del vetro di spin quali la teoria di campo medio, la replica symmetry (RS) e la replica symmetry breaking (RSB) si trova la capacità della rete e il limiti per cui il network funge ancora da memoria associativa.

Da queste analisi emergono delle transizioni di fase del secondo ordine dipendenti dal parametro  $\alpha$  e dalla temperatura  $T$ . Si possono distinguere quattro fasi diverse: fase paramagnetica, magnetica, ferromagnetica e vetro di spin. In figura 2.6 è mostrato il diagramma di fase della rete di Hopfield. Nella fase paramagnetica la temperatura è

## 2. Reti di Hopfield

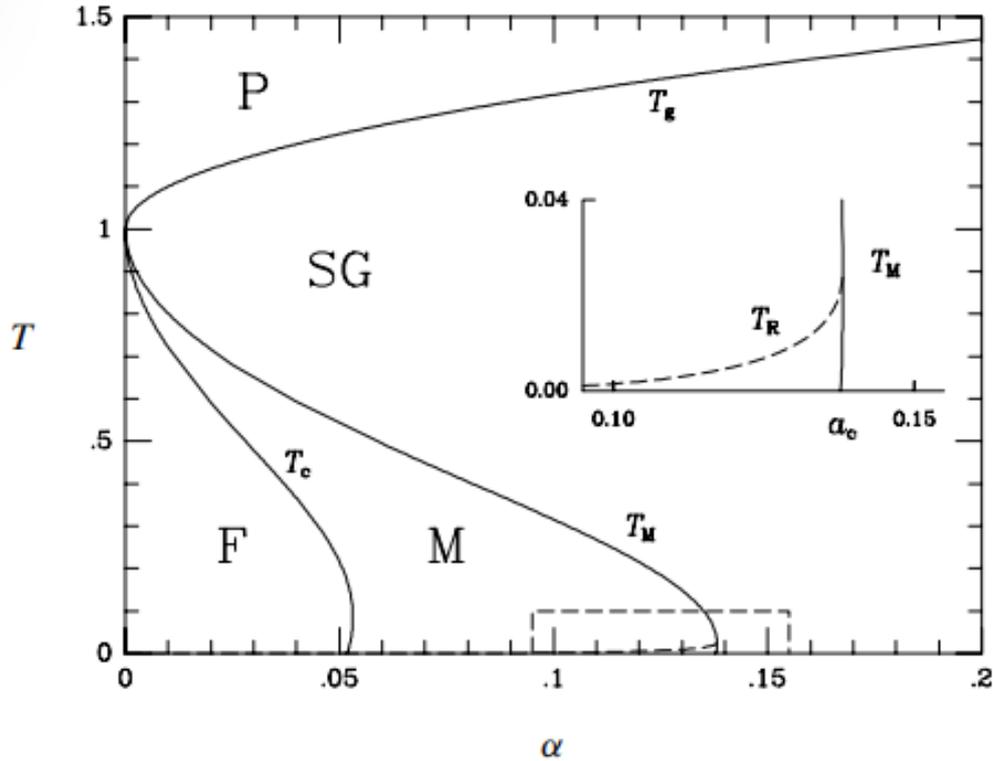


Figura 2.6: Diagramma di fase della rete di Hopfield. P: fase paramagnetica,  $m = q = 0$ . SG: fase vetro di spin,  $m = 0, q > 0$ . F: fase ferromagnetica,  $m \neq 0, q > 0$ . M: fase magnetica. La linea tratteggiata rappresenta l'instabilità AT. Crediti:[7]

troppo alta per avere dei punti di equilibrio. Nella fase vetro di spin i bacini d'attrazione dei minimi sono piccoli e fitti, la rete funge da memoria solo se l'input è molto vicino al ricordo. La rete in queste due fasi smette di essere una memoria associativa, l'unica configurazione d'equilibrio è  $m = 0$ . Nella fase ferromagnetica, invece, i ricordi sono solo quelli che sono minimi dell'energia. Nella fase magnetica, invece, i ricordi sono minimi ma possono emergere minimi spuri non collegati a quest'ultimi. La rete in queste due fasi può essere usata come memoria associativa, ci sono varie configurazioni d'equilibrio  $m \neq 0$  oltre a quella banale. La transizione tra la fase magnetica e la fase vetro di spin, per temperature non troppo alte, avviene quando  $\alpha$  supera il valore critico  $\alpha_c = 0.138$ .

Per temperature prossime allo zero la replica symmetry smette di essere affidabile e avviene una transizione di fase verso il vetro di spin prima che  $\alpha$  raggiunga  $\alpha_c$ . Questo fenomeno descritto dalla RSB è chiamato instabilità AT.

# Capitolo 3

## Reti di Hopfield diluite

In questo capitolo si analizza una particolare variazione della rete di Hopfield, la rete di Hopfield diluita. In questo caso i neuroni non sono tutti connessi tra loro ma le connessioni vengono generate in modo casuale.

### 3.1 Introduzione alle reti di Hopfield diluite

La rete fino ad adesso presentata necessita di una connessione totale tra i neuroni, considerando il network come un grafo dove i neuroni sono i nodi e le sinapsi gli archi esso risulta completo. Ciò, però, non è riscontrabile nel cervello dove i neuroni più interconnessi al massimo formano  $10^5$  sinapsi rispetto alla loro totalità di circa  $10^{11}$ . Questo fatto spinge a creare un modello in cui l' $i$ -esimo neurone è connesso solo ad una frazione, scelta in modo casuale, degli  $N$  neuroni. Questo problema viene chiamato in letteratura rete di Hopfield diluita.

Una prima descrizione di questo problema fu data da Derrida, Gardner e Zippelius [9] nel 1987. Essi partendo dalla rete di Hopfield modificarono i pesi  $w_{ij}$  in modo da ottenere

$$w_{ij} = \frac{1}{N\rho} \sum_{\mu} A_{ij} \xi_i^{\mu} \xi_j^{\mu}, \quad (3.1)$$

dove  $A_{ij}$  è una variabile casuale indipendente uniformemente distribuita con probabilità  $\rho$  di avere  $A_{ij} = 1$  e probabilità  $1 - \rho$  di avere  $A_{i,j} = 0$ . La variabile  $A_{i,j}$  così definita segue una distribuzione di Bernoulli. Essa indica la presenza di una connessione tra  $i$ -esimo neurone e il  $j$ -esimo.

$$A_{ij} = \begin{cases} +1 & \text{se } i \text{ e } j \text{ sono connessi} \\ 0 & \text{altrimenti} \end{cases}. \quad (3.2)$$

### 3. Reti di Hopfield diluite

---

La simmetria dei pesi non è richiesta per il funzionamento del modello. Nonostante ciò per il resto della trattazione si considerano  $A_{ij} = A_{ji}$ .

Con i nuovi pesi 3.1 la funzione energia 2.3 diventa

$$H = -\frac{1}{2N\rho} \sum_{i,j} \sum_{\mu} A_{i,j} \sigma_i \sigma_j \xi_i^{\mu} \xi_j^{\mu}. \quad (3.3)$$

### 3.2 Capacità delle reti debolmente diluite

In caso di una diluizione piccola  $\rho \approx 1$ , questo modello è chiamato in letteratura rete di Hopfield debolmente diluita, il sistema si può analizzare matematicamente e si ottiene un riscaldamento della temperatura. Questo si può vedere qualitativamente in [16] considerando un numero di ricordi molto inferiore al numero di neuroni.

Infatti analizzando l'interazione dell' $i$ -esimo neurone  $h_i$  si ottiene

$$h_i = \sum_j A_{ij} w_{ij} \sigma_j. \quad (3.4)$$

Avendo rimosso poche sinapsi rimangono tanti termini nella sommatoria e si può utilizzare la teoria di campo medio ottenendo

$$h_i = A \sum_j w_{ij} \langle \sigma_i \rangle \quad (3.5)$$

dove  $A$  indica la media degli  $A_{ij}$ .

La 3.5 risulta un'equazione di campo medio con una riscalatura della temperatura  $T$ .

Per un numero di ricordi  $p$  grande la situazione si complica, ma questa proprietà continua ad essere verificata.

### 3.3 Capacità delle reti fortemente diluite

Per diluizioni maggiori  $\rho \ll 1$ , invece, i primi a studiare la rete di Hopfield fortemente diluita in modo rigoroso furono Bovier e Gayraud, sia a temperatura zero [3] che in regime termodinamico [4]. Essi dimostrarono che la capacità di una rete è proporzionale al numero di sinapsi, in questo caso  $p \propto \rho N$ , fintanto che  $\rho > \sqrt{\frac{\ln N}{N}}$ . Questo risultato si dimostra in accordo con gli studi precedenti di Peretto e Niez [24] che dimostravano questa proporzionalità per piccole diluizioni,  $\rho \approx 1$ .

### 3. Reti di Hopfield diluite

---

Considerando il grafo del network, la matrice composta dalle variabili  $(A_{i,j})$  è la matrice delle adiacenze. Data la natura stocastica del modello il grafo è un grafo di Erdős-Rényi  $G(N, \rho)$ , dove  $N$  è il numero di nodi e  $\rho$  è la probabilità che due nodi siano adiacenti. Basandosi sul comportamento della componente connessa del grafo di Erdős-Rényi in regime termodinamico si possono supporre alcune proprietà del network. Infatti per  $\rho < \frac{1}{N}$  il grafo è composto da tante componenti connesse finite, per  $\rho > \frac{1}{N}$  il grafo presenta una componente gigante e per  $\rho = \frac{\ln(N)}{N}$  il grafo risulta connesso. Intuitivamente, si potrebbe ipotizzare che il network funga da memoria associativa fintanto che il grafo risulta connesso e i neuroni riescono a comunicare tra loro, ciò avviene per  $\rho > \frac{\ln(N)}{N}$ . Però Bovier e Gayraud proposero una congettura [3]: la transizione di fase avvenga per  $\rho \approx \sqrt{\frac{1}{N}}$ .

Successivamente Löwe e Vermet [21] dimostrarono che la capacità della rete è proporzionale al numero di sinapsi,  $p \propto \rho N$ , finché il grafo del network risulta connesso, ovvero  $\rho > \frac{\log N}{N}$ .

Questo teorema è in forte disaccordo con la congettura di Bovier e Gayraud e, nel corso della trattazione, si cercherà di studiare queste transizioni tramite delle simulazioni.

# Capitolo 4

## Risultati

In questo capitolo vengono analizzati il comportamento e le proprietà statistiche delle reti di Hopfield e reti di Hopfield diluite soffermandosi specialmente sulla capacità e sulle transizioni di fase. Per fare ciò sono state fatte varie simulazioni in C++.

### 4.1 Simulazioni delle reti di Hopfield

La simulazione è costruita da  $N$  neuroni in comunicazione tra loro con i pesi determinati dalla regola di Hebb. La dinamica avviene in maniera asincrona, a turno viene scelto un neurone in modo casuale secondo una distribuzione di Monte Carlo da evolvere con probabilità di transizione, cioè la probabilità di invertire un neurone, data da

$$p(-\sigma_i|\sigma_i) = \frac{1}{1 + e^{\beta\Delta H}} \quad (4.1)$$

dove  $\Delta H = 2h_i\sigma_i$  è la differenza tra l'energia della configurazione iniziale e quella della configurazione con l' $i$ -esimo neurone invertito. La probabilità di transizione racchiude al suo interno tutte le caratteristiche della dinamica.

Essendo più formali, questo algoritmo è un processo markoviano sullo spazio delle configurazioni  $\sigma$  con probabilità di transizione  $p(\sigma'|\sigma)$  di passare dalla configurazione  $\{\sigma\}$  alla configurazione  $\{\sigma'\}$ . In caso  $p(\sigma'|\sigma) = 0$  se cambiano 2 o più neuroni. Questo processo rilassa a una distribuzione stazionaria. Maggiori informazioni sono in [Appendice A](#).

In questa rete vengono immagazzinati un numero di ricordi  $p$  generati in modo casuale, in modo tale da avere una probabilità  $\frac{1}{2}$  che  $\xi_i^\mu$  sia  $+1$  e  $\frac{1}{2}$  che sia  $-1$ .

Per verificare il funzionamento del network è stato scelto casualmente un ricordo per poi essere corrotto. Il processo di corruzione avviene utilizzando una funzione che inverte casualmente  $\frac{1}{5}$  dei neuroni della configurazione del ricordo scelto. Successivamente, applicando la dinamica di Glauber [4.1](#) con temperatura fissata  $T = 0.2$ , si controlla se,

## 4. Risultati

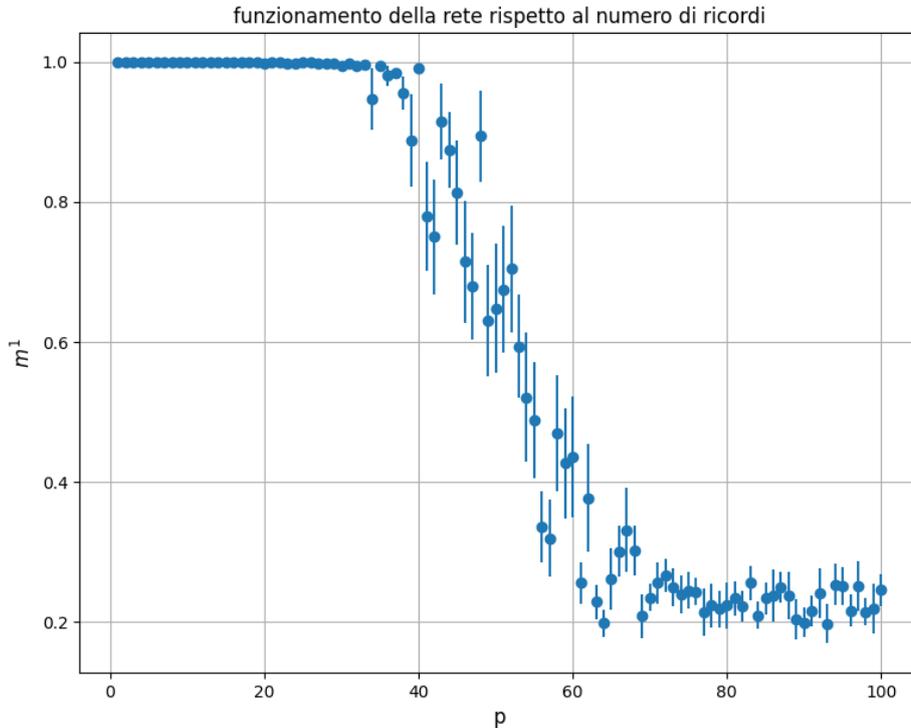


Figura 4.1: Andamento del parametro d'ordine  $m^1$ , dove con  $\xi^1$  si indica il ricordo che è stato corrotto, all'aumentare del numero di ricordi  $p$ . Il network è composto da 400 neuroni ed ha una temperatura  $T = 0.2$ . I dati sperimentali, in blu, sono stati ottenuti tramite cinque simulazioni ricavandone successivamente media e deviazione standard. Per  $p < 40$  le incertezze sui dati sono trascurabili. Nel grafico si possono notare le transizioni di fase. Per  $p$  piccoli il sistema si riconduce con precisione al ricordo iniziale, la rete è nella fase ferromagnetica. Per  $p > 40$  iniziano a crearsi delle fluttuazioni e la configurazione finale non combacia perfettamente con il ricordo di partenza, la rete è nella fase magnetica. Per  $p > 50$  il sistema non si riconduce più ad alcun ricordo, la rete è nella fase di vetro di spin.

dopo un numero prestabilito di iterazioni, la configurazione si riconduce al ricordo di partenza. È stato preso come riferimento  $50N$  iterazioni. Ciò si può verificare controllando il parametro d'ordine  $m^1$ , definito come nella 2.38, dove  $\xi_1$  si indica il ricordo corrotto; se  $m^1 \approx 1$  la rete riconduce il dato corrotto al ricordo precedentemente memorizzato, la rete così fatta risulta una memoria associativa. Si può facilmente osservare che, raggiunto un certo valore di  $p$ , la rete smette di comportarsi come una memoria associativa e il sistema entra nella fase vetro di spin. Questo processo è mostrato in figura 4.1.

## 4. Risultati

Lo stesso procedimento è stato fatto variando il numero di neuroni  $N$  lasciando inalterati gli altri parametri come la temperatura  $T$  e la probabilità di corruzione. Analizzando il valore di  $p_{max}$ , cioè il numero di ricordi immagazzinati per cui il network entra nella fase vetro di spin, in funzione di  $N$ , si nota che c'è un andamento lineare  $p_{max} = \alpha_c N + q_0$  e, facendo un fit, si ottiene  $\alpha_c = (0.124 \pm 0.008)$  e  $q_0 = (-1 \pm 5)$ ; il  $\tilde{\chi}^2$  del fit è uguale a 2.38.

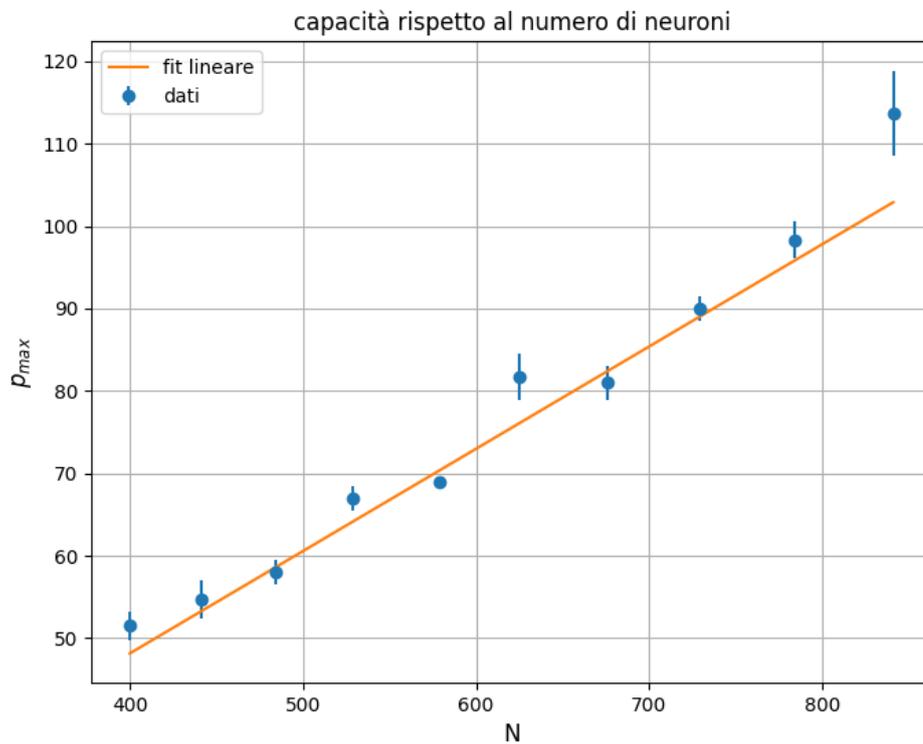


Figura 4.2: Andamento della capacità di una rete di Hopfield  $p_{max}$  in funzione del numero di neuroni  $N$ . La temperatura delle rete è fissata a  $T = 0.2$  e probabilità di corruzione  $\frac{1}{5}$ . I dati sperimentali, in blu, sono stati ottenuti tramite cinque simulazioni ricavandone successivamente media e deviazione standard, mentre il fit lineare  $p_{max} = \alpha_c N + q_0$ , in arancione, è stato ottenuto tramite il metodo dei minimi quadrati. I parametri ottenuti sono  $\alpha_c = (0.124 \pm 0.008)$  e  $q_0 = (-1 \pm 5)$ . il  $\tilde{\chi}^2$  è 2.38.

Questo risultato è compatibile con il valore di  $\alpha_c = 0.138$  ricavato dalla 2.37. Il  $\tilde{\chi}^2$  è maggiore di 1, ciò può essere dovuto alla natura stocastica e discreta del sistema e ai pochi dati presi in considerazione. Il risultato del fit è mostrato in figura 4.2.

## 4.2 Simulazioni della rete di Hopfield diluita

Come per la rete di Hopfield, anche per la versione diluita è stata creata una simulazione per verificarne le proprietà.

La simulazione è costruita da  $N$  neuroni, le sinapsi sono scelte casualmente tramite una generazione di Montecarlo uniforme con probabilità  $\rho$  che due neuroni siano connessi e  $1 - \rho$  che due neuroni siano sconnessi. L'evoluzione avviene in modo asincrono e la probabilità di transizione è data da 4.1.

Nella rete vengono immagazzinati un numero  $p$  di ricordi generati in modo casuale attraverso un processo di Monte Carlo, in modo tale che la probabilità che  $\xi_i^\mu = \pm 1$  sia rispettivamente  $\frac{1}{2}$ . Per diluire il network si usa una funzione che in maniera aleatoria associa ai pesi  $w_{ij}$  il valore ricavato attraverso la regola di Hebb oppure il valore nullo in modo tale da avere una probabilità  $\rho$  di avere una sinapsi tra il neurone  $i$  e  $j$ .

Come nel capitolo precedente, si sceglie casualmente un ricordo e si corrompe invertendo dei neuroni con probabilità di corruzione  $\frac{1}{5}$ . Successivamente, applicando la dinamica di Glauber 4.1 con temperatura fissata  $T = 0.2$ , si controlla se, dopo un numero prestabilito di iterazioni, la configurazione si riconduce al ricordo di partenza. È stato preso come riferimento 50  $N$  iterazioni. Per verificare ciò si controlla il parametro d'ordine  $m^1$ , definito come nella 2.38, dove  $\xi_1$  indica il ricordo corrotto; se  $m^1 \approx 1$  la rete si comporta come una memoria associativa.

In figura 4.3 è mostrato l'andamento del parametro d'ordine  $m^1$  in funzione di  $\rho$  per  $p = 9$ . Si può osservare che per un dato  $p$ , al variare di  $\rho$ , il network funziona come memoria associativa finché la probabilità di connessione non è inferiore a un valore critico  $\rho_c$ . Superata questa soglia la rete non si riconduce più al ricordo di partenza. Questo fenomeno è mostrato in figura 4.3.

Per studiare questa transizione sono stati misurati in modo qualitativo la grandezza dei bacini di attrazione dei ricordi. Per fare ciò si è controllato per diverse probabilità di corruzione se la rete si riconduce al ricordo di partenza, cioè se  $m^1 \approx 1$ . Dalle simulazioni si osserva che la grandezza dei bacini rimane costante finché non si raggiunge la transizione di fase della rete e smettono di esserci attrattori; si nota, inoltre, che questa transizione avviene in modo molto brusco. Una possibile causa di questo andamento a gradino può essere trovata nel modo in cui i ricordi vengono generati. Infatti, essi vengono generati in maniera uniforme e, quindi, risultano distribuiti uniformemente nello spazio delle configurazioni. Variando il modo in cui questi vengono generati, per esempio creando dei ricordi maggiormente ravvicinati rispetto agli altri, si osserva che per quelli più vicini tra loro la transizione di fase avviene prima rispetto agli altri. Per creare dei ricordi ravvicinati si è scelto casualmente uno, è stato corrotto e il prodotto finale è stato memorizzato nella rete come ricordo aggiuntivo. Questi andamenti sono mostrati in figura 4.4.

## 4. Risultati

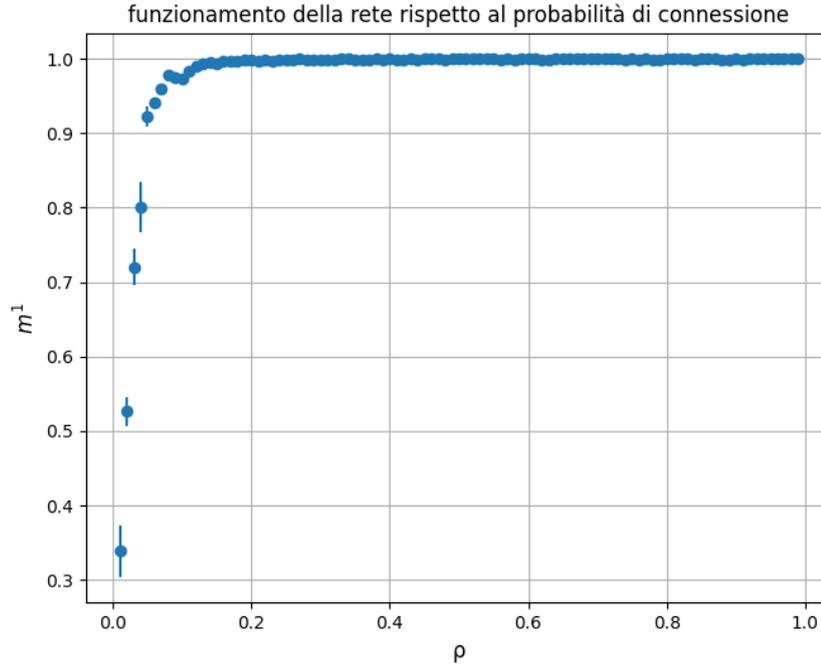


Figura 4.3: Andamento del parametro di sovrapposizione  $m^1$ , dove con  $\xi_1$  si indica il ricordo che è stato corrotto, in funzione della probabilità di connessione  $\rho$ . Il network è composto da  $N = 625$  neuroni, ha una temperatura  $T = 0.2$  e sono stati immagazzinati  $p = 9$  ricordi. I dati sperimentali, in blu, sono stati ottenuti tramite cinque simulazioni ricavandone successivamente media e deviazione standard. Si può notare che per  $\rho = 0.04$  avviene una transizione di fase e la rete non si riconduce più al ricordo.

Per studiare la capacità del network  $p_{max}$  in funzione della probabilità di connessione  $\rho$ , si è analizzata la capacità della rete al variare della probabilità lasciando invariati il numero di neuroni  $N$ , la temperatura  $T = 0.2$  e la probabilità di corruzione del ricordo. Analizzando questi dati si osserva un andamento a potenza. Facendo un fit con un andamento a potenza  $(aN)\rho^b + c$  si ottiene  $a = (0.140 \pm 0.003)$ ,  $b = (0.49 \pm 0.03)$  e  $c = (-11 \pm 2)$ ; il  $\chi^2$  del fit è uguale a 0.82. In figura 4.5 è mostrato il risultato del fit. Si può notare che per  $\rho = 1$  si ottiene la dipendenza della capacità in funzione dei neuroni per una rete di Hopfield classica,  $p_{max} = \alpha_c N$ . Questo risultato non è previsto dalla letteratura scientifica e le possibili cause di questo andamento sono, ipoteticamente, da riscontrare in un effetto di saturazione della rete dovuto al numero finito di archi.

Per capire questa transizione sono state studiate anche alcune grandezze del grafo del network, come la matrice laplaciana e il coefficiente di clustering globale nel momento in cui il network transisce dalla fase magnetica alla fase vetro di spin. Studiando il

## 4. Risultati

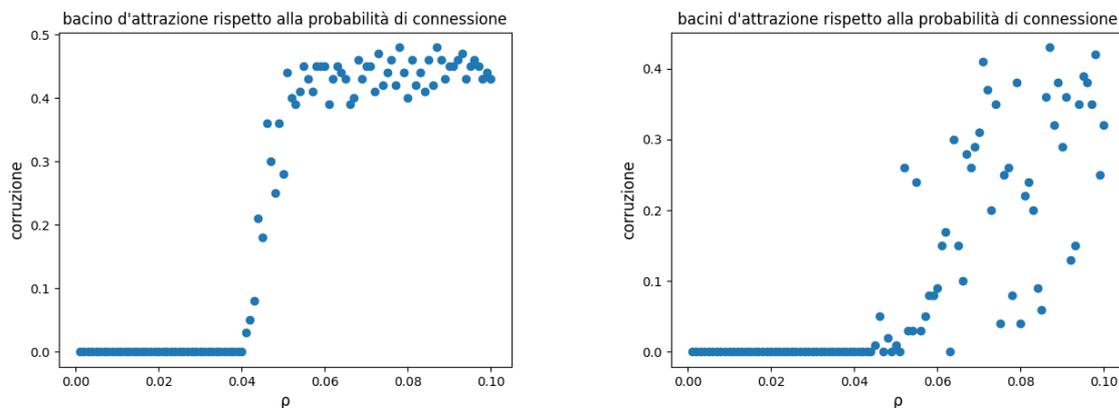


Figura 4.4: Andamento dei bacini in funzione della corruzione. Il network è composta da  $N = 625$  neuroni, ha una temperatura di  $T = 0.2$  e sono stati memorizzati  $p = 9$  ricordi. La grandezza dei bacini è stata misurata qualitativamente variando la probabilità di corruzione e verificando il funzionamento della rete con il parametro  $m^1$ . In blu sono segnati i valori della probabilità di corruzione per i quali la rete non si riconduce più al ricordo di partenza  $m^1 < 1$  ed avviene la transizione di fase. A destra i ricordi sono generati in modo uniforme; a sinistra, invece, sono stati generati in modo tale da avere un gruppo di 4 ricordi più vicini e i rimanenti 5 ricordi più distanti. Si nota che in caso di ricordi ravvicinati essi sono più instabili rispetto a quelli più distanti. Infatti, per ricordi distanziati la transizione di fase avviene quasi nello stesso momento; mentre, per ricordi ravvicinati la transizione di fase avviene in momenti diversi per i vari ricordi.

coefficiente di clustering in funzione di  $\rho$  si nota un andamento lineare caratteristico dei grafi di Erdős-Rényi. In figura 4.6 è mostrato questo andamento.

Studiando la matrice laplaciana del network specialmente l'autovalore di Fiedler, cioè il secondo autovalore più piccolo della matrice, in funzione di  $\rho$  e  $m^1$  si possono ottenere informazioni sulla transizione di fase. Osservando i dati si nota che, al diminuire del numero di ricordi  $p$ , il  $\rho_c$  in cui avviene il passaggio di fase è sempre più vicino a  $\rho = \frac{\log(N)}{N}$ , punto in cui il grafo del network smette di essere connesso, tanto che per  $p = 1$  si supera questa soglia. Questo fenomeno è mostrato in figura 4.7.

## 4. Risultati

---

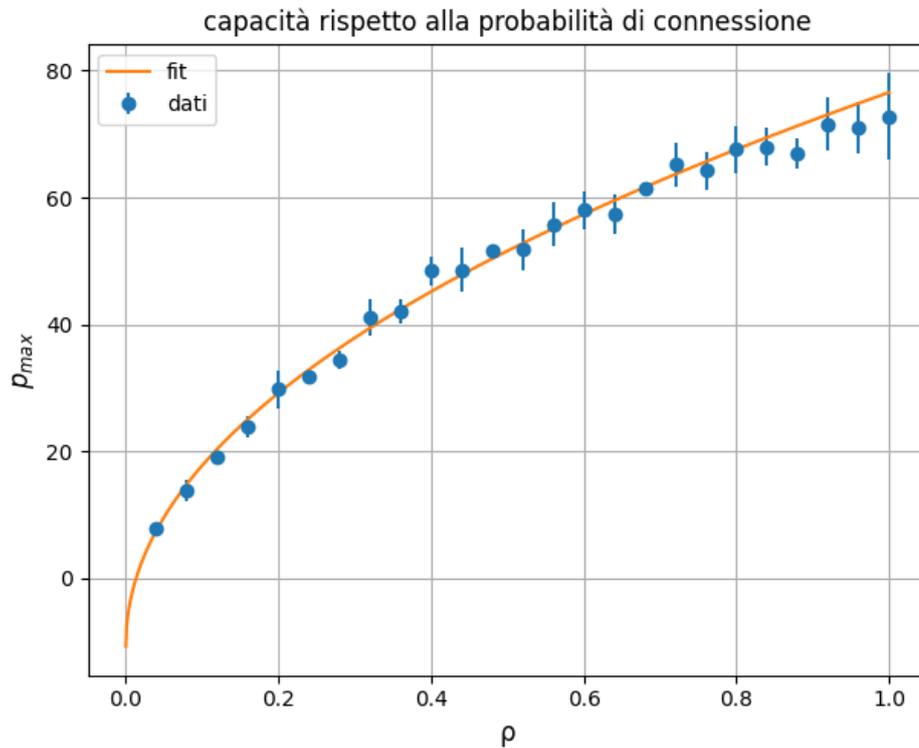


Figura 4.5: Andamento della capacità della rete  $p_{max}$  in funzione della probabilità di connessione  $\rho$ . La rete utilizzata per ottenere i dati è formata da  $N = 625$  neuroni, opera a temperatura  $T = 0.2$  e la probabilità di corruzione è  $\frac{1}{5}$ . I dati sperimentali, in blu, sono stati ottenuti tramite cinque simulazione ricavandone successivamente, media e deviazione standard, mentre il fit con una curva a potenza  $(aN)\rho^b + c$ , in arancione, è stato ricavato attraverso il metodo dei minimi quadrati. Sono stati ottenuti i seguenti parametri  $a = (0.140 \pm 0.003)$ ,  $b = (0.49 \pm 0.03)$  e  $c = (-11 \pm 2)$ . Il  $\tilde{\chi}^2$  è 0.82.

## 4. Risultati

---

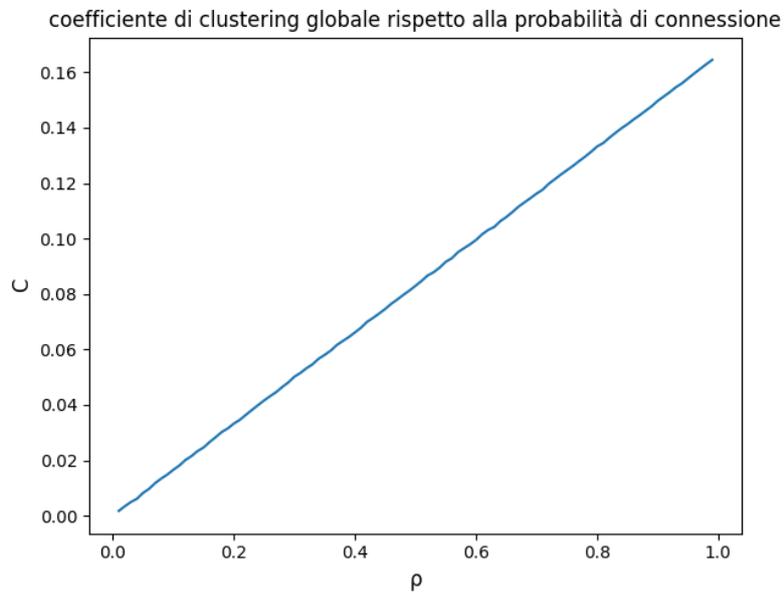


Figura 4.6: Andamento del coefficiente di clustering  $C$  del network in funzione di  $\rho$ . Si può notare un andamento lineare tipico dei grafici di Erdős-Rényi. Gli errori sono trascurabili.

## 4. Risultati

---

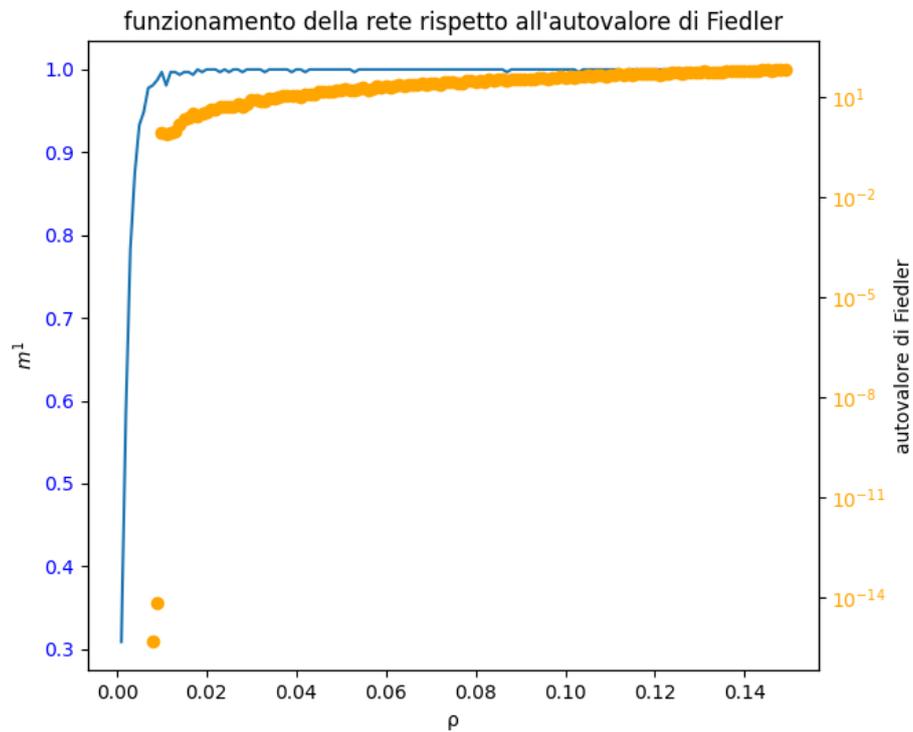


Figura 4.7: Andamento del parametro d'ordine  $m^1$  (in blu) e del secondo autovalore della matrice laplaciana, detto anche autovalore di Fiedler, (in arancione) in funzione della probabilità di connessione  $\rho$ . La rete utilizzata è formata da  $N = 625$  neuroni, lavora a temperatura  $T = 0.2$ , la probabilità di corruzione è  $\frac{1}{5}$  e ha immagazzinato  $p = 1$  ricordi. Si nota che la transizione di fase avviene dopo che il network perde la connessione.

# Capitolo 5

## Conclusioni

Questo lavoro si occupa dello studio delle proprietà statistiche delle reti di Hopfield diluite, specialmente della capacità delle reti e le transizioni di fase.

Per fare ciò sono state fatte delle simulazioni in C++ osservando il comportamento della rete al variare del numero di ricordi memorizzati e al variare delle connessioni tra le sinapsi, analizzando la connettività del grafo e l'andamento della capacità rispetto alla probabilità di connessione.

Per la rete di Hopfield le analisi sulla capacità sono compatibili con la letteratura, ottenendo una dipendenza lineare rispetto al numero di neuroni. Invece per la rete di Hopfield diluita si è osservata una dipendenza non lineare dal numero di neuroni per la probabilità di connessione. Dalle osservazioni è emerso che la capacità  $p_{max}$  segue un andamento a potenza  $p_{max} = aN\rho^b + c$ , dove  $N$  è il numero di neuroni,  $\rho$  probabilità di connessione,  $a = (0.140 \pm 0.003)$ ,  $b = (0.49 \pm 0.03)$  e  $c = (-11 \pm 2)$ . Una prima ipotesi per spiegare questo andamento è data dalla saturazione dello spazio delle configurazioni. Essendo uno spazio finito, man mano che si riducono gli archi si va a interferire con i bacini d'attrazione fino al punto che non sono più utilizzabili.

Dalle analisi della matrice laplaciana si nota che la rete funge da memoria associativa fintanto che il network è connesso, modificando in modo brusco il proprio comportamento nel momento in cui il grafo perde la connessione. Questo comportamento risulta in contrasto con la congettura di Bovier e Gayraud dato che la rete funge da memoria associativa ben oltre la soglia di  $\rho = \sqrt{\frac{1}{N}}$ .

Questo lavoro è solo un primo passo verso la comprensione del problema. Ci sono ancora molti aspetti oscuri di questo modello, a partire dall'andamento da legge a potenza, che potranno essere esplorati in futuro.

# Appendice A

## Esistenza di una distribuzione stazionaria in un network di Hopfield stocastico

Un network di Hopfield stocastico è composto da  $N$  neuroni e con i pesi determinati dalla regola di Hebb 2.5. Scegliendo un'evoluzione asincrona, a turno viene scelto secondo una distribuzione di Monte Carlo un neurone da evolvere con probabilità di transizione, cioè la probabilità di invertire un neurone, data da

$$W(-\sigma_i|\sigma_i) = \frac{1}{1 + e^{\beta\Delta H}} \quad (\text{A.1})$$

dove  $\Delta H = 2h_i\sigma_i$  è la differenza tra l'energia della configurazione iniziale e quella della configurazione con l' $i$ -esimo neurone invertito. La probabilità di transizione racchiude al suo interno tutte le caratteristiche della dinamica.

Essendo più formali, questo algoritmo è un processo markoviano sullo spazio delle configurazioni con probabilità  $W(\sigma'|\sigma)$  di passare dalla configurazione  $\{\sigma\}$  alla configurazione  $\{\sigma'\}$ . In caso  $W(\sigma'|\sigma) = 0$  se cambiano 2 o più neuroni. Generalmente questi processi possono creare comportamenti ciclici o caotici. In caso la probabilità di avere una configurazione  $p(\sigma)$  tende per il tempo  $t$  che va a infinito a una distribuzione stazionaria allora la rete tende a una configurazione di equilibrio e non ci sono correnti di probabilità tra due configurazioni differenti.

Un processo stocastico tende a una distribuzione stazionaria  $p_\infty(\sigma)$  se è ergodico, cioè esplora tutto lo spazio delle configurazioni, oppure la dinamica obbedisce al bilancio dettagliato

$$p(\sigma)W(\sigma'|\sigma) = p(\sigma')W(\sigma|\sigma'), \quad (\text{A.2})$$

## A. Esistenza di una distribuzione stazionaria in un network di Hopfield stocastico

---

cioè il sistema ha la stessa probabilità di passare dalla configurazione  $\{\sigma\}$  alla configurazione  $\{\sigma'\}$  e viceversa. Se esiste, la distribuzione stazionaria è unica.

Per questo modello, dato che le interazioni sono simmetriche, vale la relazione di bilancio dettagliato e si può dimostrare [7] che la distribuzione di probabilità stazionaria tende alla distribuzione di Maxwell-Boltzmann

$$p_\infty(\sigma) = Ae^{-\beta H(\sigma)}. \quad (\text{A.3})$$

Il coefficiente  $A$  è una costante di normalizzazione e  $H(\sigma)$  è la funzione energia della configurazione  $\{\sigma\}$ .

In realtà il sistema non è completamente ergodico. Infatti, per passare da un minimo all'altro, il sistema impiega un tempo  $t$  proporzionale a  $e^N$ . Utilizzando come unità di misura del tempo il clock interno di un computer, già per  $N = 400$  occorre aspettare  $t = e^{400} = 5.210^{173}$  clock; anche avendo a disposizione il computer più performante al mondo, un clock è circa  $10^{-9}s$ , si dovrebbe aspettare un tempo maggiore dell'età dell'universo.

Per  $N \rightarrow \infty$  il sistema rimane bloccato nel bacino d'attrazione del minimo della funzione energia tanto che se può considerare bloccato nel bacino di attrazione del minimo con conseguente rottura di ergodicità.

Lo stesso ragionamento può essere fatto per il network di Hopfield diluito.

# Bibliografia

- [1] D. J. Amit, H. Gutfreund e H. Sompolinsky. «Spin-glass models of neural networks». In: *Phys. Rev. A* 32 (2 ago. 1985), pp. 1007–1018. DOI: [10.1103/PhysRevA.32.1007](https://doi.org/10.1103/PhysRevA.32.1007).
- [2] D. J. Amit, H. Gutfreund e H. Sompolinsky. «Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks». In: *Phys. Rev. Lett.* 55 (14 set. 1985), pp. 1530–1533. DOI: [10.1103/PhysRevLett.55.1530](https://doi.org/10.1103/PhysRevLett.55.1530).
- [3] A. Bovier e V. Gayrard. «Rigorous bounds on the storage capacity of the dilute Hopfield model». In: *Journal of statistical physics* 69.3 (1992), pp. 597–627. DOI: [10.1007/BF01050427](https://doi.org/10.1007/BF01050427).
- [4] A. Bovier e V. Gayrard. «Rigorous results on the thermodynamics of the dilute Hopfield model». In: *Journal of statistical physics* 72.1 (1993), pp. 79–112. DOI: [10.1007/BF01048041](https://doi.org/10.1007/BF01048041).
- [5] E.R. Caianiello. «Outline of a theory of thought-processes and thinking machines». In: *Journal of Theoretical Biology* 1.2 (1961), pp. 204–235. DOI: [10.1016/0022-5193\(61\)90046-7](https://doi.org/10.1016/0022-5193(61)90046-7).
- [6] A. Canning e E. Gardner. «Partially connected models of neural networks». In: *Journal of Physics A: Mathematical and General* 21.15 (ago. 1988), pp. 3275–3284. DOI: [10.1088/0305-4470/21/15/016](https://doi.org/10.1088/0305-4470/21/15/016).
- [7] A.C.C. Coolen, R. Kuehn e P. Sollich. *Theory of Neural Information Processing Systems*. Oxford University Press, set. 2005. ISBN: 978-0198530244.
- [8] B.G. Cragg e H.N.V. Temperley. «The organisation of neurones: A co-operative analogy». In: *Electroencephalography and Clinical Neurophysiology* 6 (1954), pp. 85–92. DOI: [10.1016/0013-4694\(54\)90008-5](https://doi.org/10.1016/0013-4694(54)90008-5).
- [9] B. Derrida, E. Gardner e A Zippelius. «An Exactly Solvable Asymmetric Neural Network Model». In: *Europhysics Letters (EPL)* 4.2 (lug. 1987), pp. 167–173. DOI: [10.1209/0295-5075/4/2/007](https://doi.org/10.1209/0295-5075/4/2/007).
- [10] P. Erdős e A Rényi. «On Random Graphs I». In: *Publicationes Mathematicae* 6 (1959), pp. 290–297.

## Bibliografia

---

- [11] P. Erdős e A Rényi. «On the Evolution of Random Graphs». In: *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* 5.1 (1960), pp. 17–61.
- [12] T. Geszti. *Physical Models of Neural Networks*. Singapore: World Scientific, 1990. ISBN: 9789810200121.
- [13] E. N. Gilbert. «Random Graphs». In: *The Annals of Mathematical Statistics* 30.4 (1959), pp. 1141–1144. DOI: [10.1214/aoms/1177706098](https://doi.org/10.1214/aoms/1177706098).
- [14] R. J. Glauber. «Time-dependent statistics of the Ising model». In: *Journal of Mathematical Physics* 4.2 (feb. 1963), pp. 294–307. DOI: [10.1063/1.1703954](https://doi.org/10.1063/1.1703954).
- [15] D. O. Hebb. *The Organization of Behavior*. John Wiley & Sons (New York), 1949. ISBN: 9780471367277.
- [16] J. Hertz, Anders Krogh e Richard Palmer. *Introduction to the Theory of Neural Computation*. Taylor & Francis group, mar. 2018. ISBN: 9780429499661. DOI: [10.1201/9780429499661](https://doi.org/10.1201/9780429499661).
- [17] G. E. Hinton e T. J. Sejnowski. «Learning and relearning in Boltzmann machines». In: *Parallel distributed processing: Explorations in the microstructure of cognition* 1.282-317 (1986), p. 2.
- [18] G. E. Hinton e T. J. Sejnowski. «Optimal perceptual inference». In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. Vol. 448. Citeseer. 1983, pp. 448–453.
- [19] J. J. Hopfield. «Neural networks and physical systems with emergent collective computational abilities.» In: *Proceedings of the National Academy of Sciences* 79.8 (1982), pp. 2554–2558. DOI: [10.1073/pnas.79.8.2554](https://doi.org/10.1073/pnas.79.8.2554).
- [20] E. Ising. «Beitrag zur Theorie des Ferromagnetismus». In: *Zeitschrift für Physik* 31 (feb. 1925), pp. 253–258. DOI: [10.1007/BF02980577](https://doi.org/10.1007/BF02980577).
- [21] M. Löwe e F. Vermet. «The hopfield model on a sparse erdős-renyi graph». In: *Journal of Statistical Physics* 143.1 (2011), pp. 205–214. DOI: [10.1007/s10955-011-0167-1](https://doi.org/10.1007/s10955-011-0167-1).
- [22] W. S. McCulloch e W. Pitts. «A Logical Calculus of the Ideas Immanent in Nervous Activity». In: *The Bulletin of Mathematical biophysics* 5 (1943), pp. 115–133. DOI: [10.1007/BF02478259](https://doi.org/10.1007/BF02478259).
- [23] P. Peretto. «Collective properties of neural networks: a statistical physics approach». In: *Biological Cybernetics* 50 (1984), pp. 51–62. DOI: [10.1007/BF00317939](https://doi.org/10.1007/BF00317939).
- [24] P. Peretto e J.J. Niez. «Long term memory storage capacity of multiconnected neural networks». In: *Biological Cybernetics* 54.1 (1986), pp. 53–63. DOI: [10.1007/BF00337115](https://doi.org/10.1007/BF00337115).

# Ringraziamenti

Ringrazio il prof. Armando Bazzani per la sua disponibilità, nonostante i suoi mille impegni, e per il suo aiuto nella realizzazione di questo lavoro. Ringrazio il dott. Giulio Colombini per l'aiuto che mi ha dato nel approfondire questo argomento, le reti di Hopfield, e nel cercare di interpretare i risultati sperimentali che ottenevo. È stato molto bello e divertente realizzare questo lavoro con voi e spero di poter in futuro ripetere questa esperienza.

Ringrazio anche tutti i miei amici, Morgan, per la sua disponibilità e per i suoi consigli su come affrontare i vari esami, il Gruppo Eccezionale, per la compagnia durante le pause tra le lezioni e le bevute per festeggiare gli esami svolti, R1s1Ko!1!1!1!1! e Rodolfo, per bellissime feste all'area di sosta e per le serate davanti alla chiesa di Rio, i coinquilini presenti e passati e tutti i compagni del corso. Vi ringrazio per tutto l'affetto che mi avete dato.

Una menzione d'onore va ad Antonio per avermi sopportato e aiutato a scrivere i codici sia in quel bellissimo linguaggio che è C++ sia in quell'orribile linguaggio che è Python. Ti ringrazio anche per tutto il supporto che mi hai dato e per le virgole che hai aggiunto alla tesi.