

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea in Fisica

Proprietà Statistiche di Network Dinamici

Relatore:
Prof. Armando Bazzani

Presentata da:
Marco Di Giovanni

Sessione II
Anno Accademico 2014/2015

Un giorno Alice arrivò a un
bivio sulla strada e vide lo
Stregatto sull'albero.
Che strada devo prendere?
chiese.
La risposta fu una domanda:
Dove vuoi andare?
Non lo so, rispose Alice.
Allora, - disse lo Stregatto - non
ha importanza.

Alice nel paese delle meraviglie.
Lewis Carroll

Sommario

L'obbiettivo di questa tesi è quello di studiare alcune proprietà statistiche di un random walk su network.

Dopo aver definito il concetto di network e di random walk su network, sono state studiate le caratteristiche dello stato stazionario di questo sistema, la loro dipendenza dalla topologia della rete e l'andamento del sistema verso l'equilibrio, con particolare interesse per la distribuzione delle fluttuazioni delle popolazioni sui differenti nodi, una volta raggiunto lo stato stazionario. In seguito, si è voluto osservare il comportamento del network sottoposto ad una forzatura costante, rappresentata da sorgenti e pozzi applicati in diversi nodi, e quindi la sua suscettività a perturbazioni esterne.

Tramite alcune simulazioni al computer, viene provato che una forzatura esterna modifica in modo diverso lo stato del network in base alla topologia di quest'ultimo. Dai risultati si è trovato quali sono i nodi che, una volta perturbati, sono in grado di cambiare ampiamente lo stato generale del sistema e quali lo influenzano in minima parte.

Indice

Sommario	III
Introduzione	1
1 Network e random walk	3
1.1 Network	3
1.2 Matrice di adiacenza	4
1.3 Random walk	6
1.4 Matrice delle probabilità di transizione	7
1.5 Dal processo di Markov alla Master Equation	9
1.6 Master Equation	10
1.7 Differenze tra la risoluzione della master equation e la simulazione di un random walk	11
2 Studio dello stato stazionario	13
2.1 Caratteristiche dello stato stazionario	13
2.2 Rilassamento all'equilibrio	14
2.3 Fluttuazioni all'equilibrio	16
2.3.1 Sistema caratterizzato da una sola grandezza	17
2.3.2 Sistema caratterizzato da n grandezze	18
2.4 Esempio con pochi nodi	23
3 Introduzione di una forzatura costante ed effetto della forzatura sullo stato stazionario	27
3.1 Descrizione della sorgente	27
3.2 Approccio perturbativo	31
4 Conclusioni	35
Appendice A: Runge-Kutta RK4	37
Appendice B: Suscettività generalizzata	39
Correlazione tra le fluttuazioni nel tempo	39
Decomposizione spettrale delle fluttuazioni	40

Suscettività generalizzata	41
Bibliografia	43

Introduzione

Lo studio dei network è molto importante nella fisica dei sistemi complessi, in quanto è stato provato che oggetti di questo tipo consentono di rappresentare in maniera efficace alcune interazioni tra individui, la cui natura può variare in base al problema.

I network sono strutture matematiche discrete utilizzate per modellizzare un'enorme varietà di fenomeni, non solo fisici, ma anche appartenenti ad altri ambiti, come la biologia, le scienze sociali, l'ecologia e l'economia.

Numerosi sono gli studi in questo campo per descrivere i comportamenti che si osservano quando molti individui, di qualsiasi genere, interagiscono tra loro in maniera complessa. [3]

La grande moltitudine di tipologie di grafi permette di rappresentare molteplici problemi, utilizzando gli stessi concetti di base.

L'applicazione di perturbazioni a network inizialmente in equilibrio è importante per studiare il comportamento di molti sistemi sottoposti a forzature di qualsiasi genere, come, ad esempio, la chiusura di una stazione ferroviaria, nel caso di un network di trasporto, oppure l'introduzione nell'ecosistema di una nuova specie animale, nell'ambito ecologico.

I risultati ottenuti dagli studi effettuati in passato hanno consentito lo sviluppo di questo ambito della fisica, permettendo la diffusione dei metodi utilizzati in altre scienze. [7]

Nel primo capitolo viene affrontata una descrizione generale dei network e delle loro rappresentazioni formali, tramite una matrice chiamata matrice di adiacenza. Vengono anche elencati diversi tipi di network, interessanti per le loro configurazioni e per le loro caratteristiche.

In seguito, viene definito un random walk come un particolare processo di Markov discreto. Utilizzando la matrice delle probabilità di transizione, si passa dalla descrizione di un processo di Markov (nel quale spazio e tempo sono discreti) alla master equation, ossia l'equazione che governa questi fenomeni nel caso in cui il tempo può essere considerato continuo. [4]

Nel secondo capitolo viene studiata la master equation. Per prima cosa si ricava lo stato stazionario, ossia lo stato in cui il sistema risulta essere in equilibrio. In seguito, si trova in che modo il sistema rilassa all'equilibrio,

ed il tempo caratteristico di questo rilassamento. [1]

Infine, particolare interesse viene dato alla distribuzione delle fluttuazioni delle popolazioni sui nodi quando il sistema è allo stato stazionario. Per ottenere una valida teoria, viene fatta una trattazione statistica, utilizzando il concetto di Entropia ed il Principio di Massima Entropia. Per maggiore chiarezza, prima è stato trattato il caso in cui il sistema è caratterizzato da una sola grandezza, poi il caso in cui le grandezze che descrivono il sistema sono molte (caso applicabile ad un network). [5]

Nel terzo capitolo viene studiato il comportamento del sistema se viene perturbato da alcune sorgenti e pozzi. Dopo aver definito formalmente una perturbazione di questo genere, viene osservato il modo in cui il sistema reagisce. Queste forzature costanti lo modificano in base alla sua topologia, ottenendo risultati diversi se vengono applicate in nodi differentemente sensibili. Viene trovato il nuovo stato stazionario e, tramite un approccio perturbativo, ne vengono descritte le proprietà, ottenendo in quali punti il sistema è più suscettibile.

Ogni studio precedentemente descritto è stato provato attraverso simulazioni al computer, effettuate utilizzando un programma scritto in C++ e root. In appendice A viene descritto il metodo con il quale viene risolto numericamente il sistema di equazioni differenziali che rappresenta la master equation (RK4).

In appendice B è riportata una descrizione formale della suscettività generalizzata, utilizzando i concetti di correlazione tra le fluttuazioni nel tempo e decomposizione spettrale delle fluttuazioni. [5]

Capitolo 1

Network e random walk

1.1 Network

Un network, o grafo, è un insieme di n elementi chiamati nodi, o vertici, collegati tra loro da linee chiamate link, o archi. Un network, quindi, è l'unione di un vettore, che rappresenta i nodi, e di un vettore di coppie, che rappresenta le connessioni tra i nodi. [2]

Ogni link connette due nodi tra loro, che sono chiamati gli estremi del link. Se i due estremi sono uguali, il link viene detto cappio. Se più link hanno gli stessi estremi, essi formano un multiarco.

In questa trattazione vengono affrontati solamente i network semplici, cioè i network senza cappi o multiarchi, quindi il vettore di coppie di nodi, che rappresenta i link, può essere pensato come una matrice, descritta in seguito.

Una caratteristica importante dei grafi è il minimo cammino medio che collega due nodi. Calcolando il numero di passi necessari per raggiungere un qualsiasi nodo partendo dal nodo fissato si trova questo valore, che è indice della connettività generale del network.

Infine, è possibile definire un'altra importante proprietà dei grafi, cioè quanto i nodi sono connessi tra loro, tramite il coefficiente di clustering. Il coefficiente di clustering locale riguarda un solo nodo di un network, mentre quello globale coinvolge l'intera rete. Questi coefficienti sono numeri compresi tra 0 ed 1, dove 0 rappresenta il caso in cui i collegamenti sono inesistenti, mentre 1 rappresenta il caso in cui il numero di collegamenti è massimo.

I link sono generalmente unidirezionali, quindi un nodo i può essere collegato con un nodo j senza che il nodo j sia collegato con il nodo i .

Esempi di network con link unidirezionali sono i network genetici, nei quali i geni sono connessi tra di loro in modo tale che l'attività di un gene influenza il comportamento di quelli vicini, oppure network che rappresentano catene alimentari, nei quali gli animali sono collegati tra loro se uno è la preda dell'altro.

Esistono anche situazioni nelle quali se un nodo i è collegato con un nodo j , allora il nodo j è obbligatoriamente collegato con il nodo i . In questo caso i link vengono chiamati bidirezionali.

Esempi di network con link bidirezionali sono i network di amicizie, dove le persone sono collegate tra loro se si conoscono reciprocamente, i network di trasporto, nei quali alcuni punti di interesse vengono collegati tramite strade, binari ferroviari oppure rotte aeree, o network telefonici, che collegano persone tramite antenne telefoniche.

1.2 Matrice di adiacenza

La topologia del network, ossia la presenza o l'assenza di un link tra ogni coppia di nodi (e la loro unidirezionalità o bidirezionalità), è descritta da una matrice booleana chiamata matrice di adiacenza A . L'elemento A_{ij} è 1 se esiste il link che collega il nodo j con il nodo i , altrimenti è 0. Se tutti i link sono bidirezionali, la matrice A è simmetrica.

Esistono diversi metodi per creare network interessanti tramite una matrice di adiacenza.

Si possono creare network nei quali ogni coppia di nodi ha una determinata probabilità di essere connessa, che sono chiamati Random Network ed hanno un breve minimo cammino medio, ma un coefficiente di clustering piccolo (Figura 1.1). [8]

Se si vuole studiare un network con una precisa topologia, si può creare la matrice di adiacenza in base alla natura del problema. Alcune delle configurazioni più utilizzate sono riportate in seguito.

Ad esempio, è utile studiare un network in cui ogni nodo è collegato solo con il precedente ed il successivo (l'ultimo è connesso anche con il primo), oppure con i due precedenti ed i due successivi.

Molto utilizzato per alcune particolari proprietà è anche quello che viene chiamato Small World Network, che consiste nel prendere uno dei due network descritti precedentemente e ricollegare alcuni link (generalmente pochi) scegliendo casualmente gli estremi (Figura 1.2). Questo tipo di configurazione è importante perché possiede un breve minimo cammino medio ma un grande coefficiente di clustering, quindi rappresenta fedelmente network di conoscenze tra persone (da qui il nome). [1]

Un'ulteriore tipologia di network sono i Network ad albero, cioè i network senza cicli (non esiste un percorso che permette di ritornare al nodo di partenza senza passare più volte sullo stesso link), tra i quali uno dei più importanti si crea aggiungendo sequenzialmente tutti i nodi, collegandoli ad un solo nodo già presente, con una probabilità dipendente dal numero di collegamenti effettuati precedentemente. Bisogna osservare che gli alberi

hanno, in generale, un lungo minimo cammino medio, ed un coefficiente di clustering nullo (Figura 1.3).

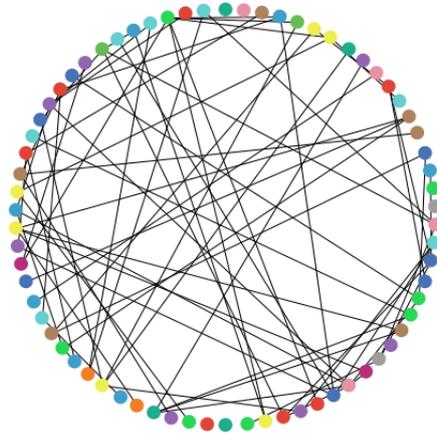


Figura 1.1: Illustrazione di un Random Network con 70 nodi e 70 link

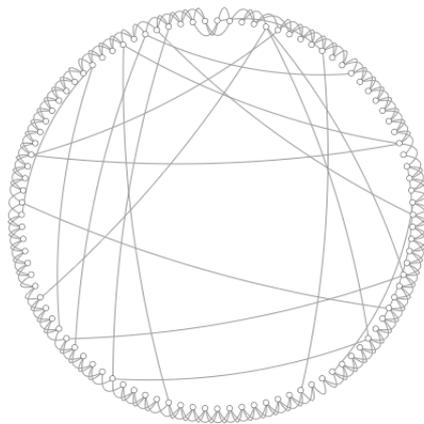


Figura 1.2: Illustrazione di uno Small World Network con 100 nodi

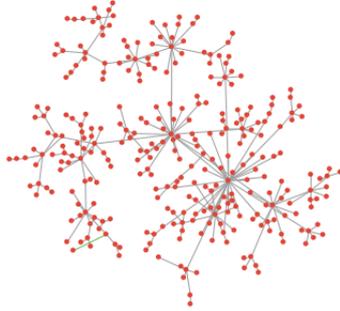


Figura 1.3: Illustrazione di un Network ad albero con 300 nodi

1.3 Random walk

Un random walk è un processo stocastico effettuato da uno o più individui, chiamati walkers, che partono da un punto qualsiasi fissato e prendono decisioni casuali sulla direzione e sulla lunghezza dei passi. Questo modello è utile per studiare il comportamento di molti processi in numerosi campi oltre la fisica, come la biologia, l'economia, la chimica e l'ecologia. [6]

Un random walk si svolge in un luogo che può essere scelto in base al problema da studiare. Si possono pensare random walks su rette (1 dimensione), piani (2 dimensioni) o nello spazio (3 dimensioni).

In questa trattazione verranno, invece, studiati i random walks su network, nei quali la direzione non è arbitraria e non viene definita la lunghezza dei passi.

Il percorso di un walker che si muove casualmente in un network è descritto da un vettore di K elementi, dove K è il numero di passi effettuati. Ogni elemento k_i rappresenta il nodo nel quale il walker si trova dopo aver effettuato i passi, perciò gli elementi del vettore sono numeri interi da 1 a N , dove N rappresenta il numero dei nodi del network. Non tutte le combinazioni degli elementi del vettore sono possibili, poichè non tutti i nodi sono generalmente connessi tra loro. Infatti, il walker che si trova su un qualsiasi nodo i potrà spostarsi solamente su un nodo j connesso al nodo i tramite un link.

Un random walk di questo genere è un processo di Markov discreto. [4] Dato un sistema il cui stato è descritto completamente da una variabile S , si ha un processo di Markov se lo stato del sistema nel futuro dipende solamente dallo stato presente ed è indipendente dal passato. Formalmente, la proprietà di Markov, nel caso in cui il processo sia discreto, è formulata nel seguente modo:

$$P(S_n = s_n | S_{n-1} = s_{n-1}, \dots, S_0 = s_0) = P(S_n = s_n | S_{n-1} = s_{n-1}) \quad (1.1)$$

dove $P(A|B)$ è la probabilità che si verifichi l'evento A , dato l'evento B , mentre S_i rappresenta lo stato del sistema dopo i passi.

Per quanto riguarda un random walk generale su network, lo stato del sistema è rappresentato da un vettore le cui componenti indicano il numero di walkers presenti su ogni nodo. Se c'è un solo walker nel network, questo vettore avrà tutte le componenti nulle tranne una in cui sarà presente un 1.

Ora è interessante studiare il caso in cui nel network sono distribuiti un numero N di walkers. Se n_i rappresenta la popolazione sul nodo i , allora vale

$$\sum_{i=1}^n n_i = N \quad (1.2)$$

Per semplicità al posto di n_i si preferisce considerare la frazione di walkers presente nel nodo, definita da

$$x_i = \frac{n_i}{N} \quad (1.3)$$

poiché queste quantità sono normalizzate a 1:

$$\sum_{i=1}^n x_i = 1 \quad (1.4)$$

Si nota che è sempre soddisfatta la relazione $0 \leq x_i \leq 1$.

1.4 Matrice delle probabilità di transizione

Per descrivere la dinamica del network viene definita la matrice delle probabilità di transizione P_{ij} . L'elemento P_{ij} rappresenta la probabilità che un walker vada nel nodo i dal nodo j , perciò vale $0 \leq P_{ij} \leq 1 \forall$ coppia i, j . Se l'elemento A_{ij} è nullo, quindi non esiste un link che collega il nodo j con il nodo i , allora anche l'elemento P_{ij} deve essere nullo. Infatti, la probabilità di passare da un nodo ad un altro non collegato con il primo deve essere nulla.

Inoltre la matrice P_{ij} deve essere stocastica, quindi

$$\sum_{i=1}^n P_{ij} = 1 \quad \forall j \quad (1.5)$$

Questo significa che la somma delle probabilità di andare in un qualsiasi nodo i partendo da un qualsiasi nodo j fissato deve essere unitaria. Perciò il numero di walkers viene conservato nel tempo.

Gli elementi di P_{ij} sulla diagonale ($i = j$) rappresentano la probabilità che

un walker rimanga nel nodo in cui è presente, senza muoversi.

Se la matrice P_{ij} dipende dallo stato della rete, quindi $P_{ij} = P_{ij}(x_1, \dots, x_n)$, significa che i walkers interagiscono tra loro e le probabilità di transizione sono influenzate dalla popolazione presente nei nodi. Un esempio sono i network con nodi con capacità limitata, in quanto tutti gli elementi della matrice P_{ij} con un i fissato devono essere nulli se la popolazione presente in quel nodo raggiunge la capacità massima.

Per creare una matrice P_{ij} ci sono diversi metodi che soddisfano le condizioni imposte in precedenza. Il metodo utilizzato calcola gli elementi P_{ij} utilizzando il grado dei nodi.

Si vuole, infatti, creare un network tale che la probabilità che un walker vada in un nodo i connesso al nodo j nel quale il walker si trova è identica a quella di andare in un qualsiasi altro nodo connesso al nodo di partenza. Nessun link deve essere preferito rispetto agli altri.

Viene definito grado di un nodo d_j il numero di nodi con cui è collegato, quindi il numero di link che partono da esso.

$$d_j = \sum_{i=1}^n A_{ij} \quad (1.6)$$

Gli elementi della matrice stocastica sono calcolati con la relazione:

$$P_{ij} = \frac{A_{ij}}{d_j} \quad (1.7)$$

Si può verificare l'effettiva stocasticità di P_{ij} .

Dimostrazione:

$$\sum_{j=1}^n P_{jk} = \sum_{j=1}^n \frac{A_{jk}}{d_k} = \frac{\sum_{j=1}^n A_{jk}}{d_k} = \frac{d_k}{d_k} = 1$$

Inoltre, poiché $A_{ij} = 0, 1$ e $d_i \geq 1 \forall i$ (nel network tutti i nodi sono collegati con almeno un altro nodo), vale la disuguaglianza $0 \leq P_{ij} \leq 1 \forall$ coppia i, j .

I network utilizzati hanno link bidirezionali. In questi network, se l'elemento $P_{ij} \neq 0$ allora anche l'elemento $P_{ji} \neq 0$, ma non è sempre vero che $P_{ij} = P_{ji}$ poiché, in generale $d_i \neq d_j$.

Sono state utilizzate P_{ij} con tutti gli elementi sulla diagonale nulli, in modo tale che i walkers siano obbligati a muoversi.

1.5 Dal processo di Markov alla Master Equation

Un processo di Markov su un network è un processo, con tempo e spazio discreti, descritto dall'equazione

$$p_i(t + \Delta t) = \sum_{j=1}^n P_{ij} p_j(t) \quad (1.8)$$

dove $p_i(t)$ è la probabilità di trovarsi nel nodo i al tempo t , mentre Δt è l'intervallo di tempo tra due istanti.

Per passare da un processo a tempo discreto ad uno a tempo continuo bisogna tenere conto del fatto che le probabilità di transizione variano nel tempo. Infatti diventa sempre più probabile, con il passare del tempo, attraversare un link e sempre meno probabile restare fermi. Per formalizzare questo concetto viene definita la matrice π_{ij} tale che valgano le seguenti proprietà:

$$P_{ij} = \pi_{ij} \Delta t \quad i \neq j \quad (1.9a)$$

$$P_{ij} = 1 - \pi_{ij} \Delta t \quad i = j \quad (1.9b)$$

La matrice π_{ij} rappresenta i rate di transizione tra tutte le coppie dei nodi del network. L'equazione (1.8) diventa:

$$p_i(t + \Delta t) = (1 - \pi_{ii} \Delta t) p_i(t) + \sum_{j \neq i} \pi_{ij} p_j(t) \Delta t$$

$$p_i(t + \Delta t) - p_i(t) = -\pi_{ii} \Delta t p_i(t) + \sum_{j \neq i} \pi_{ij} p_j(t) \Delta t$$

Ricordando che la matrice P_{ij} è stocastica, si ricava

$$1 = \sum_{i=1}^n P_{ij} = \sum_{i \neq j} (\pi_{ij} \Delta t) + 1 - \pi_{jj} \Delta t \quad \forall j$$

quindi

$$\pi_{jj} = \sum_{i \neq j} \pi_{ij} \quad \forall j$$

Sostituendo nella (1.8) si trova:

$$p_i(t + \Delta t) - p_i(t) = - \sum_{j \neq i} \pi_{ji} p_i(t) \Delta t + \sum_{j \neq i} \pi_{ij} p_j(t) \Delta t$$

$$\frac{p_i(t + \Delta t) - p_i(t)}{\Delta t} = - \sum_{j \neq i} \pi_{ji} p_i(t) + \sum_{j \neq i} \pi_{ij} p_j(t)$$

Ora è possibile effettuare il limite per $\Delta t \rightarrow 0$, per passare ad un processo con tempo continuo:

$$\dot{p}_i = \sum_{j \neq i} (\pi_{ij} p_j(t) - \pi_{ji} p_i(t)) \quad (1.10)$$

Questa equazione viene chiamata Master Equation e descrive i processi a tempo continuo e spazio discreto. [2]

1.6 Master Equation

I random walk su network sono governati dalla master equation:

$$\dot{p}_i = \sum_{j \neq i} (\pi_{ij} p_j - \pi_{ji} p_i) \quad (1.11)$$

Questa equazione descrive l'andamento della probabilità che un walker sia in un nodo in funzione della matrice π_{ij} e delle probabilità che il walker si trovi negli altri nodi nel passato. L'equazione, infatti, eguaglia la variazione della probabilità di essere in un nodo con la sommatoria, su ogni nodo, del rate di transizione da un qualsiasi nodo j al nodo studiato, moltiplicato per la probabilità di essere nel nodo j (flusso entrante), sottratto al rate di transizione dal nodo studiato al nodo j , moltiplicato per la probabilità di essere del nodo studiato nel passato (flusso uscente).

È importante osservare che la master equation descrive il comportamento di un sistema con il tempo continuo, ma con lo spazio discreto, infatti, è presente la derivata rispetto al tempo della popolazione dei nodi. La discretezza del network impone che le distanze tra i vari nodi non possano essere modificate per poter essere considerate infinitesime.

Conviene definire la matrice Laplaciana:

$$L_{ij} = \delta_{ij} - P_{ij} \quad (1.12)$$

Quindi, ricordando (1.9), questa matrice può essere scritta come

$$L_{ij} = \pi_{ij} \Delta t \quad i = j \quad (1.13a)$$

$$L_{ij} = -\pi_{ij} \Delta t \quad i \neq j \quad (1.13b)$$

In questo modo, per $\Delta t = 1$, la master equation diventa:

$$\dot{p}_i = - \sum_{j=1}^n L_{ij} p_j \quad (1.14)$$

Dimostrazione:

$$\begin{aligned}\dot{p}_k &= \sum_{j \neq i} (\pi_{kj} p_j - \pi_{jk} p_k) = - \sum_{j \neq i} L_{kj} p_j - p_k \sum_{j \neq i} P_{jk} = \\ &= - \sum_{j \neq i} L_{kj} p_j - p_k = - \sum_{j \neq i} L_{kj} p_j - L_{kk} p_k = - \sum_{j=1}^n L_{kj} p_j\end{aligned}$$

ricordando che

$$\begin{aligned}\sum_{j \neq i} P_{jk} &= 1 \quad \forall k \\ L_{kk} &= 1 \quad \forall k\end{aligned}$$

1.7 Differenze tra la risoluzione della master equation e la simulazione di un random walk

Quando si vuole simulare un random walk su network, bisogna considerare che questo è un processo discreto nel tempo. Ad ogni istante i walkers prendono contemporaneamente una decisione casuale, basandosi solamente sulla matrice P_{ij} , non sullo stato del sistema.

Simulando questo fenomeno con un numero elevato di walker si ottiene un andamento delle popolazioni dei nodi nel tempo, con il quale è possibile studiare fluttuazioni rispetto ai valori medi.

È possibile approssimare l'andamento delle popolazioni sui nodi simulato precedentemente tramite la master equation, considerando il tempo continuo. L'equazione (1.14) è un sistema di equazioni differenziali in p_i .

La soluzione può essere trovata anche analiticamente, ma, per network con un gran numero di nodi, i conti si complicano, quindi è stato deciso di utilizzare un algoritmo numerico per trovare la soluzione approssimata del sistema. Il metodo utilizzato è RK4, l'algoritmo di Runge-Kutta al quarto ordine (descritto in appendice A).

Per risolvere il sistema di equazioni differenziali è necessaria la condizione iniziale del sistema, chiamata $p_i^{in} = p_i(0)$, che rappresenta il valore delle popolazioni su ogni nodo prima di iniziare l'evoluzione del sistema.

I risultati ottenuti con la master equation saranno simili a quelli ottenuti attraverso la simulazione per quanto riguarda la soluzione stazionaria, mentre saranno meno compatibili nello studio dei transienti, poiché l'approssimazione funziona meglio se le grandezze p_i variano poco nell'intervallo di tempo Δt utilizzato per rendere il processo continuo.

Quindi, il primo procedimento permette di studiare molto efficacemente lo stato stazionario ed il tempo di rilassamento di un network, mentre il secondo metodo permette di studiare le fluttuazioni delle popolazioni rispetto alla media.

Capitolo 2

Studio dello stato stazionario

2.1 Caratteristiche dello stato stazionario

Data la master equation descritta precedentemente

$$\dot{p} + Lp = 0 \quad (2.1)$$

esiste sempre una soluzione stazionaria p^s , unica e attrattiva, perciò, date condizioni iniziali qualsiasi, normalizzate a 1 ($\sum_{i=1}^n p_i^{in} = 1$), dopo un determinato tempo, dipendente dalla topologia del network, il sistema si troverà all'equilibrio.

La soluzione stazionaria corrisponde all'autovettore con autovalore nullo della matrice Laplaciana L , quindi dipende solo dalla forma del network, mentre è indipendente da altri fattori, come la popolazione iniziale.

Il sistema rilassa verso l'equilibrio con un andamento esponenziale, la cui scala di tempo è data dall'inverso del più piccolo autovalore non nullo, in valore assoluto.

Per dimostrare che il determinante di L è nullo, bisogna ricordare la seguente proprietà:

$$\sum_{j=1}^n L_{jk} = 0$$

Questo vuol dire che

$$L^T \bar{1} = 0$$

dove $\bar{1}$ è un vettore le cui componenti sono tutte 1. Quindi L^T ha un autovettore con autovalore nullo, ossia esiste una soluzione non triviale dell'equazione agli autovalori di L^T :

$$L^T v_i = v_i \lambda_i$$

dove v_i è un generico autovettore con autovalore nullo ($\lambda_i = 0$).

Poiché L ed L^T hanno gli stessi autovalori, allora anche L ha un autovalore

nullo. Quindi il determinante di L , che è dato dal prodotto degli autovalori, è nullo.

Scegliendo $P_{ij} = \frac{A_{ij}}{d_j}$ la soluzione stazionaria sarà proporzionale al grado del nodo corrispondente: $p_j^s \propto d_j$.

Ponendo $c = \frac{p_j^s}{d_j}$ una costante indipendente dal nodo, considerando lo stato stazionario e dividendo entrambi i membri della master equation per d_i si ottiene:

$$0 = \dot{p}_i = \sum_{j \neq i} (\pi_{ij} p_j^s - \pi_{ji} p_i^s)$$

quindi

$$\frac{\sum_{j \neq i} \pi_{ij} p_j^s}{d_i} = \frac{\sum_{j \neq i} \frac{d_i}{d_j} p_j^s}{d_i} = \frac{c \sum_{j \neq i} 1}{d_i} = c$$

$$\frac{\sum_{j \neq i} \pi_{ji} p_i^s}{d_i} = \frac{p_i^s \sum_{j \neq i} \pi_{ji}}{d_i} = \frac{p_i^s}{d_i} = c$$

Poiché deve valere la relazione

$$\sum_{i=1} p_i^s = 1 \quad (2.2)$$

la soluzione all'equilibrio sarà

$$p_i^s = \frac{d_i}{d_{tot}} \quad (2.3)$$

dove

$$d_{tot} = \sum_{i=1}^n d_i \quad (2.4)$$

2.2 Rilassamento all'equilibrio

La master equation (2.1) è un sistema di equazioni lineari, quindi la soluzione si può scrivere come somma di esponenziali:

$$p(t) = \sum_{i=1}^n c_i e^{-\lambda_i t} v_i \quad (2.5)$$

dove c_i sono coefficienti costanti e v_i sono gli autovettori con autovalore corrispondente λ_i della matrice L .

La matrice Laplaciana è costruita in modo tale da avere tutti gli autovalori positivi tranne uno nullo, quindi tutti i termini della sommatoria corrispondenti ad autovettori con autovalore diverso da zero, per tempi lunghi,

daranno un contributo nullo. Per questo motivo rimane solo il termine che corrisponde all'autovettore con autovalore nullo, che rappresenta la soluzione all'equilibrio.

Il tempo di rilassamento del sistema è dato dall'autovalore più piccolo tra quelli positivi, poiché questo corrisponderà al termine della sommatoria che tende a zero più lentamente.

Vengono ora mostrati i risultati della master equation, calcolati con il metodo Runge-Kutta RK4, per un Random Network di 100 nodi. All'inizio la popolazione è distribuita equamente in tutti i nodi. Si può notare come il sistema evolve verso l'equilibrio come descritto precedentemente (Figura 2.1). In ogni grafico rappresentato, il tempo è misurato in unità di tempo arbitrarie.

Osservando in particolare l'andamento della somma delle differenze della popolazione sui nodi rispetto all'equilibrio (Figura 2.2) si può calcolare il tempo di rilassamento del sistema, che coincide con il minimo autovalore non nullo della matrice Laplaciana utilizzata.

Eseguendo un fit lineare dei dati ottenuti, scartando alcuni dati iniziali poiché il sistema, in quel momento, è influenzato non solo dall'autovettore con autovalore minore, ma anche dagli altri, si ottiene una buona correlazione lineare.

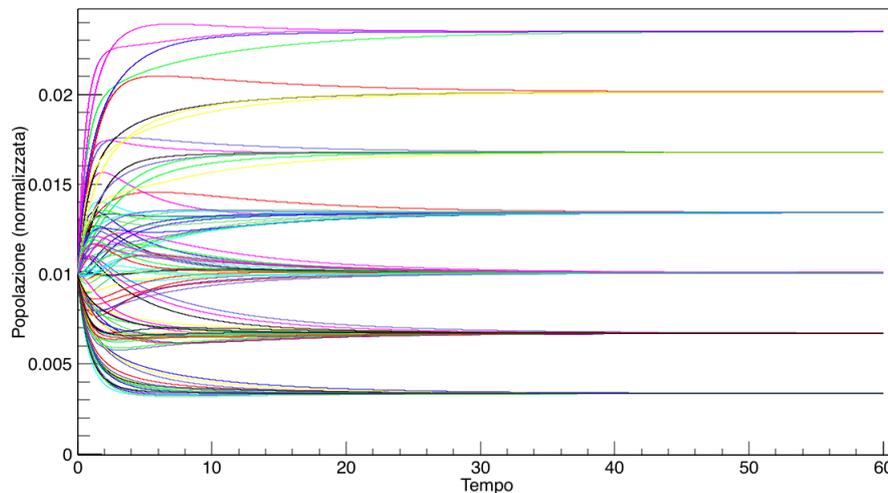


Figura 2.1: Risoluzione della master equation in un Random Network con 100 nodi e popolazione iniziale distribuita equamente su ogni nodo. Ogni curva colorata rappresenta l'andamento della popolazione in un diverso nodo del network.

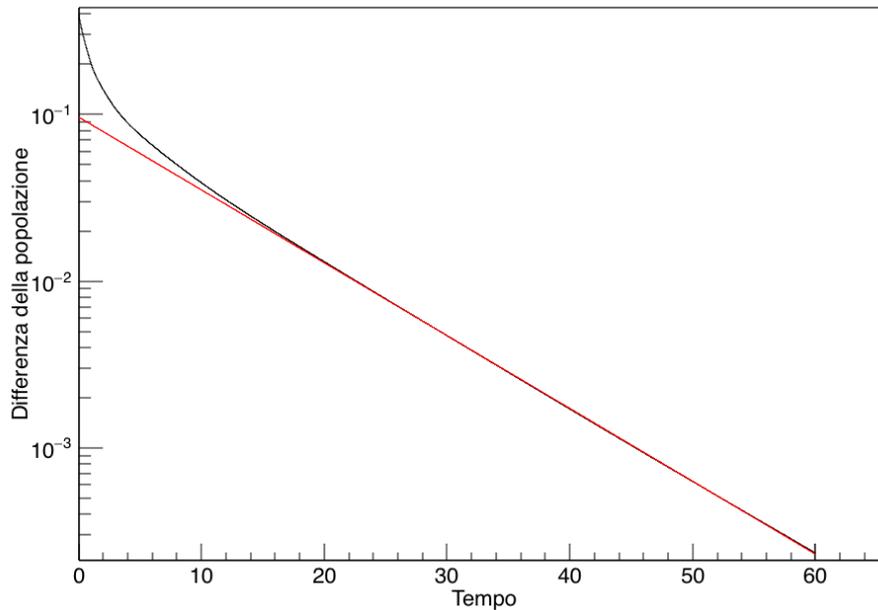


Figura 2.2: Andamento della differenza tra la popolazione dei nodi nel tempo rispetto alla soluzione all'equilibrio, nello stesso network descritto nella figura precedente. La retta rossa rappresenta il fit della curva. Il coefficiente di correlazione lineare calcolato risulta $R=-0.996$. L'equazione della retta è $\ln(\Delta p) = -2.34 - 0.10t$

2.3 Fluttuazioni all'equilibrio

Nel caso di un random walk su network, una popolazione di N walkers si muove negli n nodi del network fino a raggiungere l'equilibrio. In questo stato, le popolazioni dei singoli nodi sono simili al loro valore medio, ma è importante tenere conto delle piccole deviazioni che si verificano, chiamate fluttuazioni. Lo studio delle fluttuazioni consiste nel trovare la distribuzione delle probabilità di queste deviazioni rispetto al valore medio.

Se il numero di walkers è piccolo, le fluttuazioni diventano così importanti che non è più possibile considerarle piccole deviazioni dal valore medio. In questo caso non si può effettuare la seguente trattazione, che si basa sull'ipotesi che N sia abbastanza grande da poter pensare questo sistema come un sistema termodinamico, nel quale le grandezze subiscono fluttuazioni piccole rispetto ai loro valori medi. [5]

2.3.1 Sistema caratterizzato da una sola grandezza

Inizialmente è preferibile affrontare il caso in cui un sistema sia descritto da una singola variabile macroscopica x , traslata in modo tale che $\bar{x} = 0$. Questa variabile termodinamica, quando il sistema è in equilibrio, fluttua attorno al suo valore medio. Per trovare la distribuzione di queste fluttuazioni si definisce la probabilità $w(x)$ che la variabile x si trovi nell'intervallo tra x e $x + dx$, utilizzando il concetto di entropia di Boltzmann $S(x)$:

$$w(x) = ce^{S(x)} \quad (2.6)$$

dove c è una costante.

Il Principio di Massima Entropia afferma che un sistema isolato evolve verso lo stato in cui la sua entropia è massima. In questo stato il sistema è in equilibrio ($x = \bar{x} = 0$).

Valgono le relazioni:

$$\left. \frac{\partial S}{\partial x} \right|_{x=0} = 0$$
$$\left. \frac{\partial^2 S}{\partial x^2} \right|_{x=0} < 0$$

Sviluppando in serie di potenze fino al secondo ordine si trova:

$$S(x) = S_0 - \frac{\beta}{2}x^2 \quad (2.7)$$

con la costante $\beta > 0$.

L'equazione (2.6) diventa

$$w(x) = N_{norm} e^{-\frac{\beta}{2}x^2} \quad (2.8)$$

dove N_{norm} è la costante di normalizzazione tale che

$$\int_{-\infty}^{+\infty} w(x) dx = 1$$

Risolvendo l'integrale si ottiene

$$N_{norm} = \sqrt{\frac{\beta}{2\pi}}$$

La distribuzione della grandezza x è quindi una gaussiana:

$$w(x) = \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta}{2}x^2} \quad (2.9)$$

Calcolando il momento secondo della distribuzione si trova

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2 w(x) dx = \frac{1}{\beta} \quad (2.10)$$

Sostituendo nell'equazione (2.9), la distribuzione diventa

$$w(x) = \frac{1}{\sqrt{2\pi\langle x^2 \rangle}} e^{-\frac{x^2}{2\langle x^2 \rangle}} \quad (2.11)$$

Si osserva che, più è piccolo $\langle x^2 \rangle$, più la curva è stretta ed il massimo è elevato, in quanto abbiamo imposto la normalizzazione costante ad 1.

Concludendo, un sistema isolato descritto da una variabile x evolve nel tempo fino ad uno stato di equilibrio che massimizza l'entropia del sistema stesso. In questo stato, la variabile x fluttua e i suoi valori sono distribuiti secondo una gaussiana centrata nel valore medio della variabile.

2.3.2 Sistema caratterizzato da n grandezze

Quando un sistema è caratterizzato da più grandezze termodinamiche, si può ripetere un ragionamento simile al precedente.

Nel caso di un random walk su network, il sistema è descritto da n variabili x_1, \dots, x_n tali che x_i rappresenta la popolazione di walkers presente nel nodo i . Se il numero di walkers è grande, le variabili x_i possono essere considerate variabili termodinamiche e la trattazione seguente può essere effettuata. In questo caso, il limite termodinamico ci permette di considerare i valori delle fluttuazioni relative che vogliamo studiare abbastanza piccoli rispetto al valore medio delle variabili x_i . Infatti:

$$\begin{aligned} \bar{x}_i &\propto N \\ \sigma_i(x_i) &\propto \sqrt{N} \end{aligned}$$

Si può definire l'entropia totale del sistema $S(x_1, \dots, x_n)$ come una funzione di tutte e solo le variabili x_1, \dots, x_n .

Anche in questo caso vale il principio di massima entropia: un sistema isolato evolve verso lo stato in cui la sua entropia è massima. Il sistema, perciò, muterà in modo tale da assumere determinati valori delle grandezze x_1, \dots, x_n che massimizzano la funzione $S(x_1, \dots, x_n)$. È comodo, in questo caso, traslare le variabili in modo tale che $\bar{x}_1 = \dots = \bar{x}_n = 0$. Questa operazione non modifica il comportamento del sistema.

Definendo $w(x_1, \dots, x_n)$ la probabilità che il sistema si trovi nello stato in cui le sue variabili assumano i valori x_1, \dots, x_n , questa funzione è determinata dalla formula precedentemente citata

$$w(x_1, \dots, x_n) = N_{norm} e^{S(x_1, \dots, x_n)} \quad (2.12)$$

Sviluppando $S(x_1, \dots, x_n) - S_0$ con Taylor fino al secondo ordine si ottiene

$$S(x_1, \dots, x_n) - S_0 = -\frac{1}{2} \sum_{i,k=1}^n \beta_{ik} x_i x_k \quad (2.13)$$

con $\beta_{ik} = \beta_{ki}$.

L'equazione (2.12) diventa

$$w(x_1, \dots, x_n) = N_{norm} e^{-\frac{1}{2} \sum_{i,k=1}^n \beta_{ik} x_i x_k} \quad (2.14)$$

N_{norm} è una costante di normalizzazione, in questo caso calcolata ponendo

$$\int w dx_1 \dots dx_n = 1$$

Risolvendo l'integrale si ottiene

$$N_{norm} = \frac{\sqrt{\beta}}{(2\pi)^{n/2}}$$

e sostituendo

$$w(x_1, \dots, x_n) = \frac{\sqrt{\beta}}{(2\pi)^{n/2}} e^{-\frac{1}{2} \sum_{i,k=1}^n \beta_{ik} x_i x_k} \quad (2.15)$$

È utile introdurre le grandezze termodinamicamente reciproche con x_i :

$$X_i = -\frac{\partial S}{\partial x_i} = \beta_{ik} x_k \quad (2.16)$$

Queste grandezze sono importanti perchè la reciprocità è bilaterale:

$$x_i = -\frac{\partial S}{\partial X_i}$$

Calcolando i valori medi dei prodotti, si trova [5]

$$\langle x_i X_k \rangle = \delta_{ik} \quad (2.17)$$

Inserendo la definizione (2.16), si ottengono le seguenti relazioni tra β_{ik} e la matrice di covarianza delle grandezze x_i e delle X_i :

$$\langle x_i x_k \rangle = \beta_{ik}^{-1}$$

$$\langle X_i X_k \rangle = \beta_{ik}$$

Nel caso di un network, le popolazioni sui nodi non sono indipendenti, ma la popolazione su un nodo influenza la popolazione sui nodi collegati ad esso, soprattutto durante il rilassamento del sistema, prima di raggiungere lo stato stazionario. All'equilibrio, però, le popolazioni possono essere considerate indipendenti in quanto il numero di walker in un nodo non influenza in modo consistente il numero di walker nei nodi vicini.

Quindi, se x_i e x_j sono indipendenti, $\langle x_i x_j \rangle = \bar{x}_i \bar{x}_j = 0$. Questo significa che tutti i termini della matrice β_{ij} fuori dalla diagonale sono nulli.

Se si hanno solamente due variabili e queste sono indipendenti, la distribuzione cercata è

$$w_{12} = \text{cost} \cdot e^{-\frac{1}{2}\beta_{11}x_1^2 - \beta_{12}x_1x_2 - \frac{1}{2}\beta_{22}x_2^2}$$

Poiché $\beta_{12}^{-1} = 0$, allora $\beta_{12} = 0$, quindi:

$$w_{12} = \text{cost} \cdot e^{-\frac{1}{2}\beta_{11}x_1^2 - \frac{1}{2}\beta_{22}x_2^2}$$

Questa distribuzione è il prodotto di due gaussiane indipendenti tra loro. Si può effettuare un ragionamento simile nel caso in cui ci siano n variabili, cercando la distribuzione di una coppia ij . La distribuzione w_{ij} si ottiene integrando $w(x_1, \dots, x_n)$ su tutte le x_k con $k \neq i, j$, perciò

$$w_{ij} = \text{cost} \cdot e^{-\frac{1}{2}\beta'_{ii}x_i^2 - \beta'_{ij}x_i x_j - \frac{1}{2}\beta'_{jj}x_j^2}$$

Se le due variabili sono indipendenti, $\beta'_{ij} = 0$, quindi $\beta'_{ij} = 0$.

In questo caso, i coefficienti β' sono in generale differenti dai coefficienti β in quanto vengono calcolati attraverso l'integrazione descritta precedentemente. Ciononostante, se la matrice β_{ij} è diagonale, lo sarà anche la matrice β'_{ij} . La distribuzione delle fluttuazioni di due variabili w_{ij} si separa nel prodotto di due distribuzioni Gaussiane indipendenti.

Per potere confrontare le diverse distribuzioni delle fluttuazioni nei diversi nodi, è necessario considerare una nuova grandezza scalata

$$x' = \frac{x}{\sigma} \tag{2.18}$$

ricordando che

$$\sigma^2 = \langle x^2 \rangle - \bar{x}^2 = \langle x^2 \rangle \tag{2.19}$$

In questo modo, se x era una grandezza distribuita con media nulla e varianza σ , la nuova variabile avrà una distribuzione con media nulla e varianza unitaria:

$$\bar{x}' = 0 \quad \sigma' = 1 \tag{2.20}$$

Dimostrazione:

$$\begin{aligned}\overline{x'} &= \frac{\sum_{t=0}^T x'}{T} = \frac{\sum_{t=0}^T \frac{x}{\sigma}}{T} = \frac{\sum_{t=0}^T x}{T\sigma} = \frac{\overline{x}}{\sigma} = 0 \\ \sigma' &= \sqrt{\frac{\sum_{t=0}^T (x'^2)}{T}} = \sqrt{\frac{\sum_{t=0}^T (x^2)}{T\sigma^2}} = \frac{\sqrt{\frac{\sum_{t=0}^T (x^2)}{T}}}{\sigma} = \frac{\sqrt{\langle x^2 \rangle}}{\sigma} = 1\end{aligned}$$

La distribuzione reale non è una gaussiana, poichè le variabili x_1, \dots, x_n , non traslate, sono discrete e non negative. Verranno quindi indicate con m_1, \dots, m_n .

Per network con un piccolo numero di nodi, la distribuzione della popolazione in ogni nodo si comporta come una binomiale:

$$w(m_i) = \binom{N}{m_i} p_i^{m_i} (1 - p_i)^{N - m_i} \quad (2.21)$$

dove N è il numero totale di walkers e p_i è la probabilità che un walker sia sul nodo i (la soluzione della master equation all'equilibrio), quindi $p_i = \frac{d_i}{d_{tot}}$, come dimostrato precedentemente.

Se il numero dei walkers è elevato, la distribuzione tende ad una Poissoniana, con $\mu_i = Np_i$ mantenuto costante quando si effettua il limite:

$$w(m_i) = \frac{\mu_i^{m_i}}{m_i!} e^{-\mu_i} \quad (2.22)$$

Dimostrazione:

Ricordando l'approssimazione di Stirling per k elevati

$$k! \simeq \left(\frac{k}{e}\right)^k$$

ed i seguenti limiti:

$$\begin{aligned}\lim_{N \rightarrow +\infty} \left(1 - \frac{\mu_i}{N}\right)^{\mu_i} &= 1 \\ \lim_{N \rightarrow +\infty} \left(1 - \frac{\mu_i}{N}\right)^N &= e^{-\mu_i}\end{aligned}$$

Si ricava, sostituendo nella distribuzione binomiale:

$$\begin{aligned}
 w(m_i) &= \binom{N}{m_i} p_i^{m_i} (1 - p_i)^{N - m_i} = \frac{N!}{m_i! (N - m_i)!} \left(\frac{\mu_i}{N} \right)^{m_i} \left(1 - \frac{\mu_i}{N} \right)^{N - m_i} = \\
 &= \frac{N^N e^{-N}}{m_i! (N - m_i)^{N - m_i} e^{-N + m_i}} \left(\frac{\mu_i}{N} \right)^{m_i} \left(1 - \frac{\mu_i}{N} \right)^{N - m_i} \\
 &= \frac{1}{\left(1 - \frac{m_i}{N} \right)^{N - m_i} e^{m_i}} \frac{\mu_i^{m_i}}{m_i!} \left(1 - \frac{\mu_i}{N} \right)^{N - m_i} = \frac{\mu_i^{m_i}}{m_i!} e^{-\mu_i}
 \end{aligned}$$

Questo significa che una buona approssimazione di una distribuzione binomiale, se N è elevato, è data dalla distribuzione Poissoniana (Figura 2.3). [4]

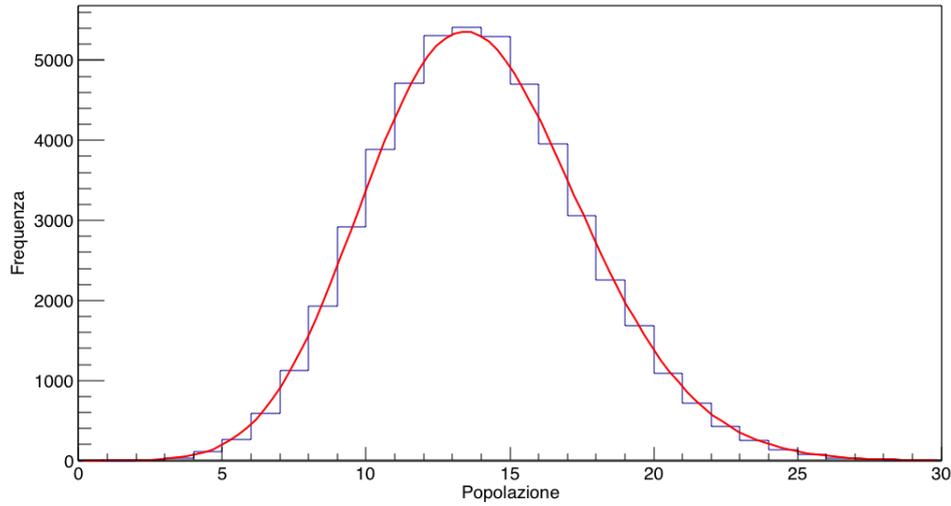


Figura 2.3: Distribuzione della popolazione di un nodo di grado 4 nel Random Network descritto precedentemente (Capitolo 2). Il numero di walkers è 1000, ed ognuno ha la possibilità di fare 50000 passi. Per raffigurare la distribuzione si sono scartati i primi passi, durante i quali il network ancora non ha raggiunto l'equilibrio. La curva rossa rappresenta il fit poissoniano dell'istogramma, la cui media $\mu = 13.42$, che corrisponde al valore della frazione della popolazione all'equilibrio, moltiplicato per il numero di walkers (vedi Figura 2.1).

2.4 Esempio con pochi nodi

Viene in seguito riportato un esempio di un network con pochi nodi. In questo caso è possibile osservare le matrici descritte precedentemente e la soluzione della master equation. Interessante è anche simulare un random walk sul network plottando in un grafico i valori ottenuti.

Il network seguente (Figura 2.4) ha 4 nodi collegati casualmente nel modo descritto dalla matrice A . Sono riportati in seguito anche il grado dei nodi, la matrice P e la matrice Laplaciana.

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad d = (3 \quad 2 \quad 2 \quad 1)$$

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 & 1 \\ 1/3 & 0 & 1/2 & 0 \\ 1/3 & 1/2 & 0 & 0 \\ 1/3 & 0 & 0 & 0 \end{pmatrix} \quad L = \begin{pmatrix} 1 & -1/2 & -1/2 & -1 \\ -1/3 & 1 & -1/2 & 0 \\ -1/3 & -1/2 & 1 & 0 \\ -1/3 & 0 & 0 & 1 \end{pmatrix}$$

Tramite appositi algoritmi sono stati calcolati gli autovalori della matrice Laplaciana e la soluzione all'equilibrio

$$\lambda = (0 \quad 0.771 \quad 1.729 \quad 1.5) \quad p^s = (3/8 \quad 2/8 \quad 2/8 \quad 1/8)$$

Le due figure riportate in seguito rappresentano l'andamento della popolazione dei nodi nel tempo. Ogni curva rappresenta un nodo. I colori differenti servono per distinguere i diversi nodi.

La Figura 2.5 mostra la soluzione della master equation con il metodo di Runge-Kutta al quarto ordine. L'algoritmo è stato iterato dal tempo iniziale $t_0 = 0$, per 1000 volte con un incremento $\Delta t = 0.01$ ognuna, arrivando quindi al tempo $t_f = 10$.

La Figura 2.6 mostra una simulazione di random walk su network, con 1000 walkers che effettuano 100 passi ciascuno.

La Figura 2.7 verifica il rilassamento esponenziale del sistema con tempo caratteristico coincidente con il più piccolo autovalore non nullo, calcolato precedentemente.

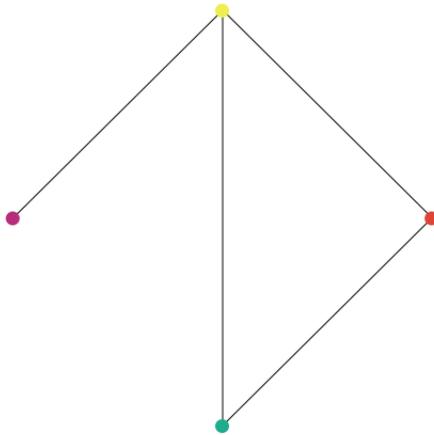


Figura 2.4: Illustrazione del network

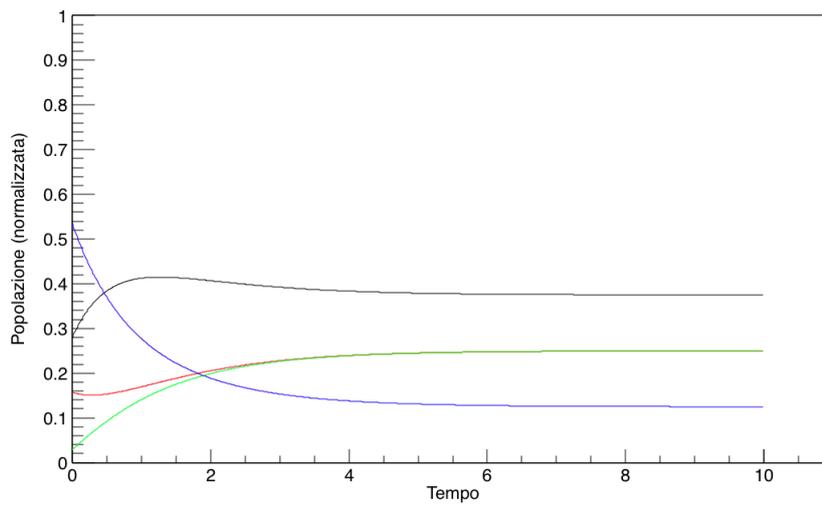


Figura 2.5: Andamento delle popolazioni dei nodi calcolato con RK4. Ogni curva rappresenta la popolazione di un nodo diverso.

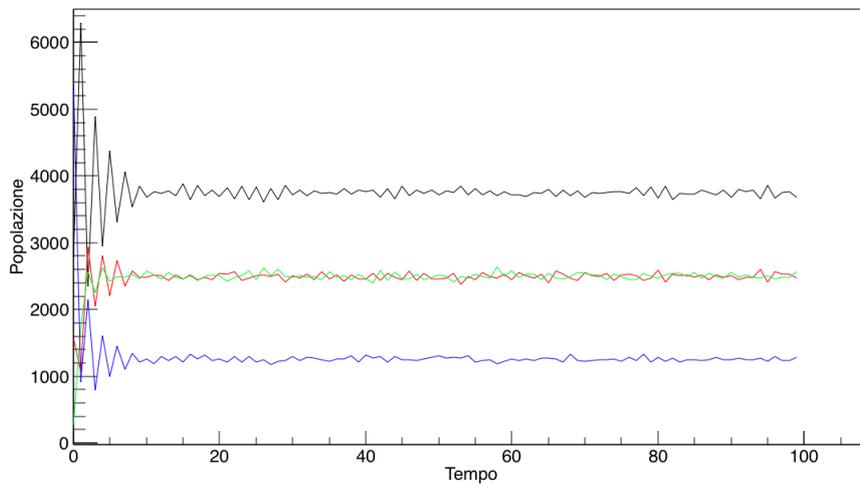


Figura 2.6: Simulazione del Random Walk di 10000 walkers sul network. Ognuno effettua 100 passi.

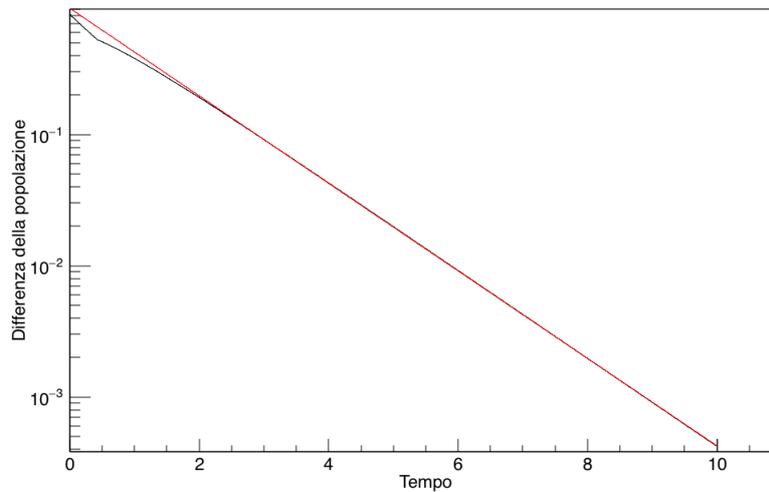


Figura 2.7: Andamento della somma delle differenze tra le popolazioni dei nodi e l'equilibrio (in valore assoluto). La retta rossa rappresenta il fit lineare, effettuato scartando i dati ottenuti durante la fase di transizione. Il coefficiente di correlazione calcolato risulta $R=-0.9998$. L'equazione della retta è: $\ln(\Delta p) = -0.09 - 0.76t$. Si può notare che la pendenza della retta coincide approssimativamente con il più piccolo autovalore non nullo della matrice Laplaciana.

Capitolo 3

Introduzione di una forzatura costante ed effetto della forzatura sullo stato stazionario

3.1 Descrizione della sorgente

Si vogliono ora studiare gli effetti di una forzatura sul network. In questo caso si prende come forzatura una sorgente ed un pozzo. La sorgente è caratterizzata da un rate s_0 tale che vale l'equazione

$$\dot{p}_0 = s_0 - \sum_{j=1}^n \pi_{j0} p_0 \quad (3.1)$$

dove p_0 è la popolazione della sorgente e π_{j0} è il rate di transizione tra la sorgente ed un nodo j . Consideriamo il caso in cui la sorgente è collegata con alcuni nodi e divide equamente tra questi la sua potenza. Perciò chiamiamo $\pi_{j0} = \pi_0 = \frac{1}{n_{sorg}} \quad \forall j$ dove n_{sorg} è il numero di nodi collegati alla sorgente. L'equazione diventa

$$\dot{p}_0 = s_0 - p_0 \quad (3.2)$$

Dimostrazione:

$$\dot{p}_0 = s_0 - \sum_{j=1}^n \pi_{j0} p_0 = s_0 - p_0 \sum_{j=1}^n \pi_{j0} = s_0 - p_0 \frac{1}{n_{sorg}} \sum_{j=1}^{n_{sorg}} 1 = s_0 - p_0$$

dove si è tenuto conto che $\sum_{j=1}^n \pi_{j0} = 1$.

La soluzione di questa equazione differenziale è un esponenziale della forma

$$p_0(t) = s_0 - (s_0 + p_0^{in})e^{-t} \quad (3.3)$$

considerando $p_0^{in} = p_0(0)$. Si può verificare che all'equilibrio (t grande) $p_0^s = s_0$.

Contemporaneamente viene aggiunto un pozzo collegato con tutti i nodi. La master equation viene modificata e diventa:

$$\dot{p}_k = \sum_{j=0}^n (\pi_{kj} p_j - \pi_{jk} p_k) \quad (3.4)$$

La nuova matrice π_{ij} utilizzata è una matrice quadrata $(n+1) \times (n+1)$. All'equilibrio, l'equazione per p_0 diventa

$$\sum_{j=0}^n \pi_{0j} p_j^s = \sum_{j=0}^n \pi_{j0} p_0^s = s_0$$

Se imponiamo che tutti gli elementi π_{0j} siano uguali $\forall j$, quindi che il pozzo prenda da ogni nodo del network i walker con lo stesso rate, allora, chiamando $\pi_{0j} = \epsilon$ l'equazione diventa

$$\epsilon \sum_{j=0}^n p_j^s = s_0$$

Inoltre

$$1 + p_0^{in} = \sum_{j=0}^n p_j^s = \sum_{j=1}^n p_j^s + p_0^s = \sum_{j=1}^n p_j^s + s_0$$

Quindi

$$\sum_{j=1}^n p_j^s = 1 + p_0^{in} - s_0 \quad (3.5)$$

$$\epsilon(1 + p_0^{in} - s_0) = s_0 \quad (3.6)$$

Si ottengono le seguenti relazioni tra ϵ ed s_0 :

$$s_0 = \frac{1 + p_0^{in}}{1 + \epsilon^{-1}}$$

$$\epsilon = \frac{s_0}{1 + p_0^{in} - s_0}$$

Questo significa che, siccome $0 \leq \epsilon \leq 1$ allora $0 \leq s_0 \leq \frac{1+p_0^{in}}{2}$. Quindi il rate di transizione deve essere minore della metà del numero di walkers presenti in tutto il sistema (sorgente compresa), altrimenti in un solo istante alcuni walkers dovrebbero entrare nel pozzo e istantaneamente uscire dalla sorgente, e questo è impossibile.

Se $\epsilon \rightarrow 0$, $s_0 \rightarrow 0$ la sorgente è spenta ed il sistema è imperturbato.

Se $\epsilon \rightarrow 1$, $s_0 \rightarrow \frac{1+p_0^{in}}{2}$ la sorgente ed i pozzi sono così potenti da rendere ininfluenti le altre connessioni del network. La soluzione stazionaria sarà data da una una popolazione p^{lim} diversa da zero solo nei nodi in cui agisce la sorgente, tale che

$$p^{lim} = \frac{1 + p_0^{in}}{2n_{sorg}} \quad (3.7)$$

Infatti, se $\pi_{0j} = 1$ allora $\sum_{i=1}^n \pi_{ij} = 0 \quad \forall j$ per la stocasticità di π_{ij} . La master equation all'equilibrio diventa

$$0 = \dot{p}_k = \sum_{j=0}^n (\pi_{kj} p_j^s - \pi_{jk} p_k^s) = \pi_{k0} p_0^s - p_k^s$$

Quindi

$$p_k^s = \frac{1 + p_0^{in}}{2n_{sorg}} \quad (3.8)$$

indipendente dal nodo k .

Dimostrazione:

$$p_k^s = \pi_{k0} p_0^s = \frac{1}{n_{sorg}} s_0 = \frac{1 + p_0^{in}}{2n_{sorg}}$$

La nuova matrice P_{ij} si costruisce, quindi, con il seguente metodo.

Gli elementi $P_{0j} = \epsilon \quad \forall j \neq 0$, come descritto in precedenza, dove ϵ può essere ricavato a partire da s_0 ,

Gli elementi $P_{j0} = \frac{1}{n_{sorg}}$, tutti uguali tra loro.

La restante matrice $n \times n$ viene trovata partendo dalla matrice di adiacenza A_{ij} , ed il grado dei nodi d_j , prima di inserire sorgenti e pozzi. Il nuovo elemento $P_{ij} = \frac{1-\epsilon}{d_j}$.

In questo modo la nuova matrice P_{ij} è stocastica:

$$\sum_{i=0}^n P_{ij} = 1 \quad \forall j \quad (3.9)$$

Dimostrazione:

$$\sum_{i=0}^n P_{ij} = \epsilon + \sum_{i=1}^n \frac{1-\epsilon}{d_j} = \epsilon + \frac{1-\epsilon}{d_j} \sum_{i=1}^{d_j} 1 = \epsilon + 1 - \epsilon = 1$$

La nuova matrice Laplaciana è

$$L_{ij}^{ex} = \delta_{ij} - P_{ij} \quad (3.10)$$

Il nuovo sistema di equazioni differenziali da risolvere per trovare il comportamento del sistema è

$$\dot{p} + L^{ex} p = 0 \quad (3.11)$$

Ora è interessante studiare il caso in cui la sorgente non influenza lo stato stazionario. Perciò si impone $p'^s = p^s$, ricordando che p^s è la soluzione stazionaria della master equation del sistema imperturbato

$$\dot{p}_i = \sum_{j=1}^n (\pi_{ij} p_j - \pi_{ji} p_i)$$

mentre p'^s è la soluzione stazionaria della master equation del sistema perturbato

$$\dot{p}_i = \sum_{j=0}^n (\pi_{ij} p_j - \pi_{ji} p_i)$$

Si trova che la condizione coincide con il bilancio dettagliato per le sorgenti/pozzi

$$\pi_{i0} p_0^s = \pi_{0i} p_i^s$$

Dimostrazione:

$$\dot{p}_i = \sum_{j=0}^n (\pi_{ij} p_j - \pi_{ji} p_i) = \dot{p}_i = \sum_{j=1}^n (\pi_{ij} p_j - \pi_{ji} p_i) + \pi_{i0} p_0 - \pi_{0i} p_i$$

Allo stato stazionario:

$$\dot{p}_i = \sum_{j=1}^n (\pi_{ij} p_j^s - \pi_{ji} p_i^s) = 0$$

Quindi

$$\pi_{i0} p_0 - \pi_{0i} p_i = 0$$

Nel caso in cui la sorgente sia del tipo descritto precedentemente, per verificare quella condizione è necessario, per prima cosa, che tutti i nodi siano collegati alla sorgente, quindi $\pi_{i0} = \frac{1}{n}$. In più deve valere

$$s_0 = p_0^{in}$$

$$s_0 = \epsilon$$

Dimostrazione:

$$\pi_{i0} p_0^s = \frac{1}{n} s_0$$

$$\pi_{0i} p_i^s = \epsilon \frac{d_i}{d_{tot}}$$

Ricordando

$$\epsilon = \frac{s_0}{1 + p_0^{in} - s_0}$$

si trova

$$\frac{1}{n} s_0 = \frac{s_0}{1 + p_0^{in} - s_0} \frac{d_i}{d_{tot}}$$

$$s_0 = 1 + p_0^{in} - n \frac{d_j}{d_{tot}}$$

Infatti s_0 non deve dipendere dal nodo, d_j deve essere uguale $\forall j$, perciò $d_{tot} = n d_j$.

$$s_0 = p_0^{in}$$

Inoltre

$$s_0 = \frac{1 + p_0^{in}}{1 + \epsilon^{-1}}$$

Quindi

$$p_0^{in} = \frac{1 + p_0^{in}}{1 + \epsilon^{-1}}$$

Risolvendo per ϵ si trova

$$\epsilon = p_0^{in}$$

Questo è il valore di ϵ da impostare se si vuole una sorgente che non influenzi lo stato stazionario del network, ricordandosi di applicare la sorgente a tutti i nodi del network.

3.2 Approccio perturbativo

Per studiare l'effetto di una perturbazione sul network si utilizza un approccio perturbativo, nel quale la nuova matrice Laplaciana può essere pensata come la vecchia matrice Laplaciana sommata ad una perturbazione ΔL . Il nuovo stato stazionario è il prodotto diretto tra il vecchio stato stazionario e lo stato p_0 della sorgente.

Dalla definizione della matrice Laplaciana si può notare che è sempre possibile scrivere

$$L = DG \tag{3.12}$$

dove D è una matrice simmetrica e G è una matrice diagonale.

Infatti:

$$L_{ij} = \delta_{ij} - P_{ij} = \delta_{ij} - \frac{A_{ij}}{d_j} = (\delta_{ik} d_i - A_{ik}) \frac{\delta_{kj}}{d_j}$$

Quindi

$$D = \delta_{ik} d_i - A_{ik}$$

è simmetrica poiché sia la matrice identità δ_{ik} sia la matrice di adiacenza A_{ik} sono simmetriche.

La matrice

$$G = \frac{\delta_{kj}}{d_j}$$

è diagonale poiché la matrice δ_{kj} è diagonale e rappresenta la metrica.

Rispetto alla metrica G si può definire il prodotto scalare

$$u \cdot v = \sum_{ij} u_i G_{ij} v_j$$

Si suppone inoltre che L sia autoaggiunta rispetto al prodotto scalare:

$$u \cdot Lv = Lu \cdot v$$

La matrice L , perciò, è simmetrica rispetto alla metrica G , cioè

$$L^T G = GL$$

Si verifica che, dati due autovettori di L con autovalori diversi ($\lambda_1 \neq \lambda_2$) il loro prodotto scalare è nullo, perciò sono ortogonali.

$$\lambda_1 v_1 \cdot v_2 = Lv_1 \cdot v_2 = DGv_1 Gv_2 = Gv_1 DGv_2 = Gv_1 Lv_2 = Gv_1 \lambda_2 v_2 = \lambda v_2 \cdot v_1$$

$$v_1 \cdot v_2 = 0$$

Una volta trovata una base ortonormale di autovettori di L , per trovare la nuova soluzione stazionaria bisogna risolvere l'equazione perturbativa

$$(L_0 + \Delta L) \sum_{j=0}^n c_j v_j = 0 \quad (3.13)$$

dove si è scomposta la soluzione nelle componenti rispetto alla base appena trovata

$$v = \sum_{j=0}^n c_j v_j$$

con i rispettivi coefficienti c_j .

Poiché la base è un insieme di autovettori di L , l'equazione (3.13) diventa

$$L_0 \sum_{j=0}^n c_j v_j + \Delta L \sum_{j=0}^n c_j v_j = \sum_{j=1}^n \lambda_j c_j v_j + \sum_{j=0}^n c_j \Delta L v_j = 0$$

Ora è necessario proiettare l'equazione in un qualsiasi vettore di base v_i fissato, ricavando

$$\lambda_i c_i + \sum_{j=0}^n v_i \cdot c_j \Delta L v_j = 0 \quad (3.14)$$

$$\sum_{j=0}^n v_0 \cdot \Delta L v_j c_j = 0 \quad (3.15)$$

ricordandosi che i v_i sono ortonormali, quindi

$$v_i \cdot v_j = 0 \quad i \neq j$$

$$v_i \cdot v_i = 1$$

Se è un caso non degenero, si utilizza la prima equazione, approssimata al primo ordine rispetto agli autovettori v_i , per trovare i coefficienti c_i , considerando $c_0 = 1$ e $c_j = 0$:

$$c_i = -\frac{\Delta L_{i0}}{\lambda_i} \quad (3.16)$$

dove

$$\Delta L_{ij} = v_i \cdot \Delta L v_j$$

Si osserva che il massimo cambiamento delle componenti corrisponde ad una perturbazione che agisce sull'autovettore con autovalore minore (non nullo). Per avere una soluzione molto diversa dalla soluzione del sistema imperturbato bisogna, quindi, creare sorgenti e pozzi tali che $\Delta L_{i0} \neq 0$ e che λ_i sia l'autovalore più piccolo in valore assoluto della vecchia matrice Laplaciana. Questo metodo funziona per soluzioni che possono essere considerate perturbazioni dello stato stazionario. Però, dato l'autovalore minimo λ_{min} , esiste un valore di soglia della perturbazione $\|\Delta L\|$ tale che, per Laplaciane perturbate maggiormente, la soluzione non può più essere considerata una perturbazione dello stato stazionario ed il metodo perturbativo risulta inefficace. Questo avviene quando $\|\Delta L\| \simeq \lambda_{min}$.

Viene in seguito riportata la soluzione della master equation di un Random Network con 20 nodi. Ogni curva rappresenta la popolazione di un nodo.

Nella figura Figura 3.2 è riportato il caso in cui la sorgente viene applicata ad un nodo di importanza marginale. Si può notare come il sistema non cambia di molto la sua configurazione all'equilibrio.

Nella figura Figura 3.2 è mostrato lo stesso network, al quale viene applicata una sorgente identica alla precedente, questa volta ad un nodo molto suscettibile. Si nota come lo stato del sistema, dopo la perturbazione, è molto differente da quello imperturbato all'equilibrio. Questo è dovuto dal fatto che l'autovettore corrispondente al più piccolo autovalore non nullo della matrice Laplaciana ha una componente corrispondente al nodo suscettibile molto più grande della componente corrispondente al nodo ininfluenza.

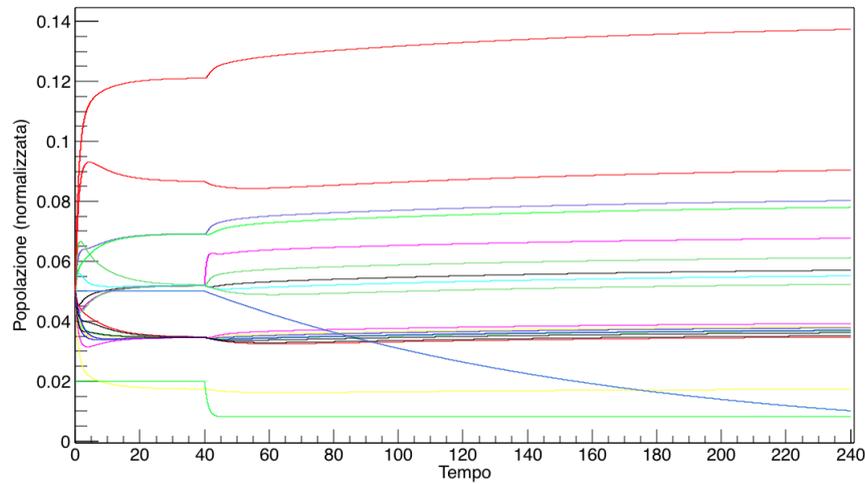


Figura 3.1: Andamento delle popolazioni dei nodi calcolato con RK4 di un Random Network di 20 nodi. Al tempo $t = 40$ viene applicata una forzatura ad un nodo poco suscettibile. Lo stato del sistema non varia molto rispetto allo stato del sistema imperturbato.

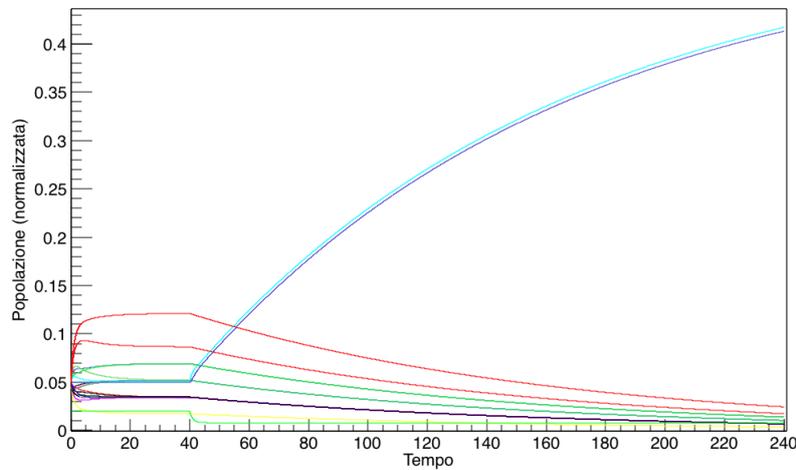


Figura 3.2: Andamento delle popolazioni dei nodi dello stesso network della figura precedente. Al tempo $t = 40$ viene applicata una forzatura ad un nodo molto suscettibile (trovato osservando autovalori e autovettori della matrice Laplaciana). Lo stato del sistema varia molto rispetto allo stato del sistema imperturbato.

Capitolo 4

Conclusioni

Il lavoro effettuato in questa tesi ha confermato le previsioni teoriche formulate in precedenza.

Tramite un apposito programma scritto in C++ e root, si è potuto verificare l'effettivo comportamento di alcuni walkers, che effettuano un random walk su network. I comportamenti simulati rispecchiano le previsioni teoriche, confermando i principali risultati aspettati.

Si è potuto verificare che le soluzioni della master equation sono una buona approssimazione del random walk, poiché i grafici che raffigurano i due fenomeni sono compatibili, ricordando comunque che il primo rappresenta un fenomeno continuo, mentre il secondo è un processo discreto. Nella simulazione riguardante il network con 4 nodi (Capitolo 2.4) si può perfino verificare direttamente la somiglianza.

I grafici in scala logaritmica delle differenze delle popolazioni rispetto allo stato stazionario nel tempo provano che il rilassamento verso l'equilibrio è esponenziale, governato dal più piccolo autovalore non nullo (gli autovalori delle matrici sono stati calcolati con appositi programmi di calcolo). La perfetta correlazione lineare fornisce valori attendibili riguardanti i dati cercati. La soluzione della master equation tramite l'algoritmo di Runge-Kutta verifica che la popolazione allo stato stazionario è proporzionale al grado del nodo, come previsto nel Capitolo 2. Dai rispettivi grafici si può notare che, all'equilibrio, la popolazione dei diversi nodi con lo stesso grado tende allo stesso valore.

I grafici nel Capitolo 3 verificano che in un generico network, alcuni nodi sono più suscettibili alle perturbazioni di altri. Infatti, perturbando diversi punti del network con la stessa sorgente, si verifica che il sistema si comporta in maniera differente. Il suo stato, dopo aver applicato la perturbazione ad un nodo poco suscettibile, è simile allo stato stazionario del sistema imperturbato, mentre dopo aver applicato la perturbazione ad un nodo molto suscettibile, si nota una forte divergenza.

Appendice A: Runge-Kutta RK4

Metodi di Runge-Kutta

I metodi di Runge-Kutta sono un insieme di metodi iterativi utilizzati per approssimare le soluzioni di equazioni differenziali ordinarie con condizioni iniziali fissate (problemi di Cauchy):

$$\begin{cases} \dot{x} = f(x, t) \\ x(t_0) = x_0 \end{cases} \quad (4.1)$$

dove $x(t) : \mathbb{R} \rightarrow \mathbb{R}^n$ ed $f(x, t) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$.

Il metodo utilizzato per risolvere la Master Equation è RK4, cioè il metodo di Runge-Kutta al quarto ordine.

Il metodo consiste nell'eseguire la seguente funzione per calcolare tutti i valori di $x(t_0 + h)$, dove h è piccolo e rappresenta l'incremento del tempo tra le iterazioni ($h > 0$)

$$x(t_0 + h) = x(t_0) + \frac{h}{6}(f(x(t_0), t_0) + 2f(y, t_0 + \frac{h}{2}) + 2f(z, t_0 + \frac{h}{2}) + f(w, t_0 + h)) \quad (4.2)$$

dove

$$y = x(t_0) + \frac{h}{2}f(x(t_0), t_0) \quad (4.3a)$$

$$z = x(t_0) + \frac{h}{2}f(y, t_0 + \frac{h}{2}) \quad (4.3b)$$

$$w = x(t_0) + hf(z, t_0 + \frac{h}{2}) \quad (4.3c)$$

Si può notare che questo metodo utilizza una approssimazione delle derivate fino al quarto ordine.

Iterando questo metodo si trova il valore numerico di $x(t)$ fino ad $x(t_f)$ dove t_f è il tempo finale, tale che $t_f - t_0 = kh$, con k il numero di iterazioni effettuate.

Per avere valori numerici precisi è necessario scegliere un valore dell'incremento h piccolo, poiché l'algoritmo commette un errore proporzionale ad h^4 .

Appendice B: Suscettività generalizzata

Questo capitolo riguarda una trattazione formale della suscettività generalizzata, utilizzando il concetto di correlazione tra le fluttuazioni nel tempo e di decomposizione spettrale delle fluttuazioni. [5]

Correlazione tra le fluttuazioni nel tempo

Considerato un sistema il cui stato può essere descritto completamente da una variabile termodinamica fluttuante $x(t)$, si può notare che i diversi valori di $x(t)$ sono correlati tra loro. Quindi i valori di $x(t')$ non sono indipendenti dai valori $x(t)$, con $t' \neq t$, ma vengono influenzati tramite la funzione di correlazione temporale

$$\varphi(t' - t) = \langle x(t)x(t') \rangle \quad (4.4)$$

che dipende solo da $t' - t$ e quindi può essere scritta

$$\varphi(t) = \langle x(0)x(t) \rangle \quad (4.5)$$

Per $t \rightarrow \infty$, la correlazione scompare, quindi $\varphi \rightarrow 0$.

Per definizione, $\varphi(t)$ è pari, quindi

$$\varphi(-t) = \varphi(t) \quad (4.6)$$

Se la grandezza x è lontano dall'origine (rispetto alla fluttuazione media), il sistema non è all'equilibrio. Col passare del tempo il sistema tenderà all'equilibrio con andamento descritto dall'equazione

$$\dot{x} = \dot{x}(x) \quad (4.7)$$

dove si nota che il comportamento dipende solo dallo stato del sistema. Considerando solo il termine lineare, troviamo

$$\dot{x} = -\lambda x \quad (4.8)$$

con $\lambda > 0$ costante. Il sistema è in equilibrio se $x = \bar{x}$. Viene traslato in modo che $\bar{x} = 0$, quindi

$$\left. \frac{dx}{dt} \right|_{x=0} = 0$$

La grandezza $\frac{1}{\lambda}$ rappresenta il tempo di rilassamento del sistema.

Introducendo la grandezza $\xi_x(t)$, che rappresenta il valore medio di x al tempo t se al tempo $t = 0$ aveva il valore x , si scrive la funzione di correlazione nel seguente modo:

$$\varphi(t) = \langle x \xi_x(t) \rangle \quad (4.9)$$

La media è fatta rispetto alle possibili x al tempo iniziale.

Dall'equazione (4.8) si trova

$$\dot{\xi}_x = -\lambda \xi_x \quad t > 0 \quad (4.10)$$

Risolvendo l'equazione sapendo che, per definizione, $\xi_x(0) = x$, si trova

$$\xi(t) = x e^{-\lambda t} \quad (4.11)$$

Quindi l'equazione (4.9) diventa

$$\varphi(t) = \langle x^2 \rangle e^{-\lambda t} \quad t > 0 \quad (4.12)$$

Perciò, ricordando (2.10), si ricava

$$\varphi(t) = \frac{1}{\beta} e^{-\lambda|t|} \quad \forall t \quad (4.13)$$

poiché $\varphi(t)$ è pari.

Decomposizione spettrale delle fluttuazioni

Per decomporre spettralmente la grandezza fluttuante $x(t)$ si utilizza lo sviluppo di Fourier:

$$x_\omega = \int_{-\infty}^{+\infty} x(t) e^{i\omega t} dt \quad (4.14)$$

$$x(t) = \int_{-\infty}^{+\infty} x_\omega e^{-i\omega t} \frac{d\omega}{2\pi} \quad (4.15)$$

La funzione di correlazione diventa

$$\varphi(t' - t) = \langle x(t')x(t) \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle x_{\omega}x_{\omega'} \rangle e^{-i(\omega t + \omega' t')} \frac{d\omega d\omega'}{(2\pi)^2} \quad (4.16)$$

Poiché $\varphi(t' - t)$ deve essere funzione solo della differenza tra i tempi $t' - t$, bisogna imporre

$$\langle x_{\omega}x_{\omega'} \rangle = 2\pi(x^2)_{\omega}\delta(\omega + \omega') \quad (4.17)$$

Questa è considerata la definizione di $(x^2)_{\omega}$. Si può notare che $(x^2)_{\omega}$ è sempre reale.

L'equazione diventa

$$\varphi(t) = \int_{-\infty}^{+\infty} (x^2)_{\omega} e^{-i\omega t} \frac{d\omega}{2\pi} \quad (4.18)$$

Al tempo $t = 0$, la correlazione corrisponde al momento secondo di x , quindi

$$\varphi(0) = \int_{-\infty}^{+\infty} (x^2)_{\omega} \frac{d\omega}{2\pi} = \langle x^2 \rangle \quad (4.19)$$

Inoltre

$$(x^2)_{\omega} = \int_{-\infty}^{+\infty} \varphi(t) e^{i\omega t} dt \quad (4.20)$$

Sostituendo

$$\varphi(t) = \frac{1}{\beta} e^{-\lambda|t|} \quad (4.21)$$

Si trova

$$(x^2)_{\omega} = \frac{2\lambda}{\beta(\omega^2 + \lambda^2)} \quad (4.22)$$

Suscettività generalizzata

Poiché è impossibile trovare una formula sempre valida equivalente alla (4.22), valida per le fluttuazioni quasi stazionarie, è necessario trovare un metodo per legare le proprietà delle fluttuazioni alle proprietà delle forze esterne. Si può scrivere l'equazione

$$\bar{x}(t) = \hat{\alpha}f = \int_0^{\infty} \alpha(\tau) f(t - \tau) d\tau \quad (4.23)$$

dove \bar{x} è il valore medio della grandezza perturbata, che è diverso da zero in quanto il sistema non è all'equilibrio, $\hat{\alpha}$ è un operatore lineare integrale ed f è la forza perturbatrice generalizzata.

Si può osservare che il valore di \bar{x} dipende solo dai valori di f riferiti a tempi antecedenti. Quindi, \bar{x} è la risposta del sistema alla perturbazione.

Riducendo la perturbazione tramite lo sviluppo in serie di Fourier, si ottiene una relazione tra le componenti della forza e la risposta del sistema

$$\bar{x}_\omega = \alpha(\omega) f_\omega \quad (4.24)$$

dove $\alpha(\omega)$ è chiamata suscettività generalizzata

$$\alpha(\omega) = \int_0^{+\infty} \alpha(t) e^{i\omega t} dt \quad (4.25)$$

Questa funzione determina il comportamento del sistema quando viene perturbato, poiché fluttuazioni di x si possono esprimere in funzione della suscettività generalizzata. Poiché $\alpha(\omega)$ è complessa

$$\alpha(\omega) = \alpha'(\omega) + i\alpha''(\omega) \quad (4.26)$$

Dalla definizione

$$\alpha(-\omega) = \alpha^*(\omega) \quad (4.27)$$

Quindi

$$\alpha'(-\omega) = \alpha'(\omega) \quad \alpha''(-\omega) = -\alpha''(\omega) \quad (4.28)$$

Perciò $\alpha'(\omega)$ è pari, mentre $\alpha''(\omega)$ è dispari.

Bibliografia

- [1] A. Barabási, *Linked: The New Science of Networks*, Perseus Books Group, 2002.
- [2] A. Barabási, M. E. J. Newman, D. J. Watts, *The Structure and Dynamics of Networks*, Princeton University Press, 2006.
- [3] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford Finance Series, 2007
- [4] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. 1*, Wiley, 1968.
- [5] Lev D Landau, Evenij M. Lifshits, *Fisica Teorica V: Fisica Statistica*. Editori Riuniti, 1978.
- [6] L. Lovász, *Random Walks on Graphs: A Survey*, Keszthely, 1993
- [7] M. E. J. Newman, *The structure and function of complex networks*, 2003.
- [8] https://en.wikipedia.org/wiki/Random_graph