

POLITECNICO DI TORINO

Corso di Laurea
in Ingegneria Matematica

Tesi di Laurea Magistrale

Modellizzazione della distribuzione della popolazione tra città su reti spaziali mediante la teoria cinetica dei sistemi multiagente



Relatori

prof. Andrea Tosin
prof. Nome Cognome

firma dei relatori

.....
.....

Candidato

Valerio Taralli

firma del candidato

.....

Anno Accademico 2025-2026

Ai miei genitori,
Elisabetta e Marco

INDICE

1	INTRODUZIONE	5
2	NOTE DI TEORIA DEI GRAFI	9
2.1	Definizioni miscellanee	9
2.2	Reti e città	10
2.3	Cenni sui dati	11
3	TEORIA CINETICA DEI SISTEMI MULTIAGENTE	15
3.1	Descrizione cinetica classica	15
3.1.1	Equazione di Boltzmann omogenea	15
3.1.2	Equazione di Boltzmann disomogenea	17
3.1.3	Equazione di tipo Boltzmann omogenea	17
3.2	Descrizione cinetica retale	17
3.2.1	Impostazione	17
3.2.2	Algoritmi d'interazione	18
3.2.3	Interazioni azione-reazione	18
3.3	Derivazione dell'equazioni cinetiche	19
3.3.1	Derivazione esatta	19
3.3.2	Interazioni tra città	21
3.3.3	Derivazione approssimata	23
3.4	Nesso discreto-continuo	23
3.4.1	Ipotesi semplificative	23
3.4.2	Analisi della IS₁	24
3.4.3	Analisi della IS₂	24
3.4.4	Analisi della IS₁ , IS₂ e IS₃	25
4	SIMULAZIONI	27
4.1	Metodo Monte Carlo	27
5	CONCLUSIONI	29
	APPENDICE	31
A	Codice	31



Lo studio della distribuzione della popolazione tra le città è stato un fenomeno già analizzato [sporadicamente] in passato. Il primo fu Auerbach [1] che all'inizio del 19-esimo secolo notò una caratteristica poi formalizzata successivamente da Zipf^f [14] quasi cinquantanni dopo, seppure inizialmente in un contesto linguistico; difatti la legge di Zipf è una legge empirica di forma

$$f \propto \frac{1}{r} \quad \text{ove} \begin{cases} f \text{ è la frequenza della parola,} \\ r \text{ è il suo rango nella classifica;} \end{cases} \quad (1)$$

per dare un esempio concreto, se la 10 parola più frequente compare 2000 volte allora esiste una costante C tale che $2000 \approx C/10$. La stessa legge descritta in (1) si può ritrovare in molt'altri contesti, tra i quali figura la distribuzione della popolazione tra le città se si in luogo di f la popolazione t [14, Fig. 9-2, p. 375] (t sta per *taglia*); tuttavia ciò è vero solo qualora $t \gg 1$, ossia nella coda della distribuzione, contrariamente al contesto linguistico originale di Zipf ove (1) vale sempre.

Più in generale la legge [empirica] di Zipf può essere descritta da una distribuzione di Pareto: sia T la variabile aleatoria legata alla taglia delle città e sia $f_T(t)$ la sua funzione di densità di probabilità², per la quale $f_T(t)dt$ indica il numero di città con $t \in (t, t+dt)$, allora si dice che T segue una distribuzione di Pareto se soddisfa

$$R(t) \approx \frac{1}{t^p} \quad \text{se } t \gg 1, \quad \text{ove } R(t) = \int_t^{+\infty} f_T(\tau) d\tau, \quad (2)$$

che altro non è che la funzione di ripartizione complementare di T , la quale rappresenta matematicamente il numero di città con popolazione maggiore o uguale a t ; si indica con $R(t)$ perché è analoga al rango nella (1) una volta che s'impone l'indice di Pareto p unitario.

Più recentemente [6, 7] si è scoperto che al di fuori della coda la $f_T(t)$ è ben fittata dalla distribuzione lognormale con funzione di distribuzione di probabilità

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \quad \text{ove } X \sim \text{Lognormale}(\mu, \sigma). \quad (3)$$

Come suggerisce il nome, la peculiarità di tale distribuzione è che vale $\log X \sim \mathcal{N}(\mu, \sigma)$ (Fig. 1.1) in cui $\mathcal{N}(\mu, \sigma)$ indica la distribuzione gaussiana.

¹ In realtà, come Zipf stesso ammette [14, p. 546], non fu il primo a scoprire tale legge; in ogni caso la popolarizzò a tal punto che oggi è conosciuta col suo nome.

² Affinché $f_T(t)$ esista bisogna rigorosamente anche supporre che sia assolutamente continua ma in questa introduzione non è necessario entrare nei dettagli, per i quali si rimanda alla § 3.

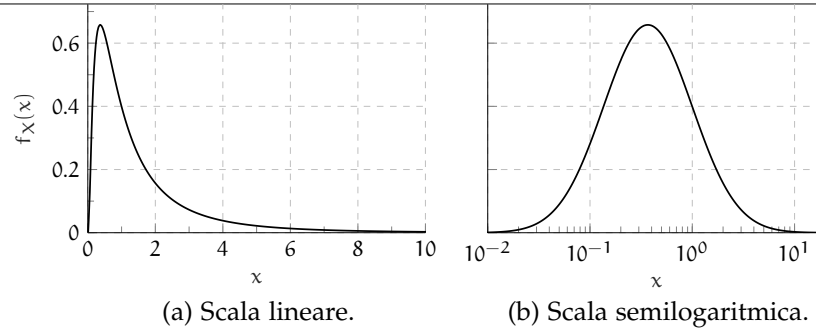


Figura 1.1: Funzione Lognormale(0,1).

Perdipiú da Gualandi *et al.* [7, 8] ci si è poi resi conto che per fittare l'intera distribuzione è sufficiente considerare una lognormale bimodale, vale a dire una combinazione convessa della (3)

$$f_T(t) = \xi f_T^1(t) + (1 - \xi) f_T^2(t), \quad (4)$$

ove $\xi \in [0, 1]$ è un parametro da fittare tanto quanto le medie e le varianze associate a $f_T^1(t)$ e $f_T^2(t)$ per un totale di 5 parametri.

Pertanto l'obbiettivo di questa tesi è il seguente: si vuole simulare attraverso la Teoria Cinetica dei Sistemi MultiAgente (TCSMA) le interazioni tra città su un grafo spaziale, tentando di riprodurre una distribuzione della popolazione con caratteristiche analoghe a quelle realmente misurate (principalmente corpo lognormale e coda di Pareto); una volta raggiunto questo scopo si potrà poi meditare sulle implicazioni della legge d'interazione riguardo al fenomeno dell'emigrazione, che è alla base delle interazioni tra città.

Si sottolinea che qui le città verranno invece considerate come agenti, caratterizzati dalla loro popolazione, che interagiscono attraverso la struttura sottostante di un grafo [spaziale]; tale descrizione, salvo il grafo, non è affatto dissimile a quella classica delle molecole di un gas caratterizzate dalla loro velocità e posizione.

Difatti, in letteratura, perlomeno quella a conoscenza dell'autore, non esistono articoli che trattano contemporaneamente le teorie cinetiche, la distribuzione della città e i grafi, nonostante la rappresentazione di queste su un grafo sia del tutto naturale; piú nel dettaglio gli articoli in letteratura

- ◇ o applicano la TCSMA alle città senza considerare una naturale topologia sottostante [8]³ o, in senso opposto, considera la struttura da grafo senza applicare la TCSMA[3];
- ◇ altri vedono i nodi piú come luoghi ove vivono e interagiscono gli agenti, creando cosí un modello descritto da un sistema di equazioni di Boltzmann elevate e accoppiate[10];
- ◇ infine l'articolo che piú si avvicina all'obbiettivo di questa tesi, applica tuttavia la sua teoria nel contesto delle reti sociali[12].

³ In particolare [8], nonostante sia un articolo molto interessante e che giunge a conclusioni altrettanto importanti, astrae eccessivamente le interazioni tra città oltre a non definire con sufficiente chiarezza che cosa s'intende per *dintorni*.

59 Questa lacuna è probabilmente attribuibile al fatto che gli sviluppi della
60 teoria dei grafi applicata alla TCSMA sono stati piuttosto recenti e prin-
61 cipalmente concentrati sulla prospettiva nodo-luogo anziché nodo-agente.
62 Dunque l'aspetto più innovativo di questo lavoro è mostrare che la pro-
63 spettiva nodo-agente è altrettanto valida quanto quella più comune e re-
64 cente del nodo-luogo; per il resto questa tesi dev'essere considerata come
65 un'esplorazione formale applicativa con pochi approfondimenti analitici.

66 Lo scritto sarà così suddiviso: nel secondo capitolo si affronteranno breve-
67 mente e [superficialmente] alcuni concetti della teoria dei grafi, soprattutto
68 quelli pertinenti alla topologia interurbana; quindi nel terzo la TCSMA è ap-
69 profondita prima da un punto di vista classico, e poi mediata dai grafi sia
70 esattamente che approssimativamente; infine negli ultimi due si illustreranno
71 i principali risultati concludendo con delle note finali e potenziali sviluppi
72 futuri.

73 Per il lettore interessato sarà anche presente un'appendice ove il codice di
74 Python dell'implementazione è spiegato a grandi linee.



In questo capitolo sono prima introdotti alcuni oggetti di base della teoria dei grafi, quindi si discute la natura della rete d'interesse, infine si descriveranno molto brevemente i dati usati.

2.1 DEFINIZIONI MISCELLANEE

Innanzitutto s'inizia dando la definizione di rete:

Definizione 2.1 (Grafo). Un grafo è formato dalla coppia $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, ove \mathcal{V} è l'insieme degli indici dei nodi mentre \mathcal{E} è l'insieme dei lati, vale a dire di coppie d'indici \mathcal{V} : due nodi $i, j \in \mathcal{V}$ sono connessi sse $(i, j) \in \mathcal{E}$.

Dall'insieme \mathcal{E} si può poi specializzare il concetto di grafo:

Definizione 2.2 (Grafo [in]diretto). Un grafo è indiretto sse dato $(i, j) \in \mathcal{E}$ allora $(j, i) \in \mathcal{E}$ e la direzione è trascurabile; altrimenti è detto diretto.

La trascurabilità della direzione sarà discussa poco dopo nella § 2.2, anche se è facilmente intuibile dall'esempio.

Definizione 2.3 (Matrice d'adiacenza unitaria e pesata). Sia $|\mathcal{V}|$ la cardinalità dell'insieme degli indici, ossia il numero d'indici totali, si definiscono la matrice d'adiacenza unitaria $A \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$ e pesata $W \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$ come:

$$a_{i,j} \equiv \begin{cases} 1 & \text{se } (i,j) \in \mathcal{E} \\ 0 & \text{altrimenti} \end{cases} \quad w_{i,j} \equiv \begin{cases} q_{i,j} & \text{se } (i,j) \in \mathcal{E} \\ 0 & \text{altrimenti} \end{cases} \quad (1)$$

ove $q_{i,j}$ è il peso¹ associato al lato (i,j) . Si noti che per i grafi indiretti ambo le matrici risultano simmetriche.

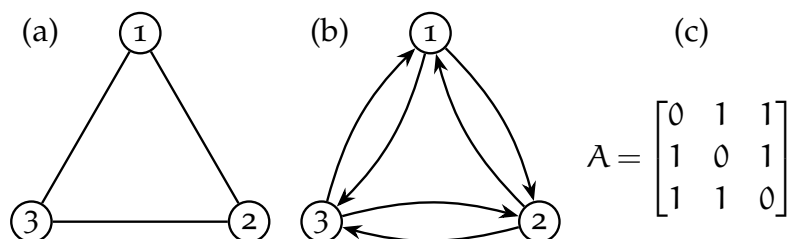


Figura 2.1: Esempio di un grafo indiretto, della sua equivalente forma diretta [simmetrica] e della loro [identica] matrice d'adiacenza.

¹ Per es. la matrice d'adiacenza unitaria può essere vista come pesata ponendo $q_{i,j} \equiv 1$.

Definizione 2.4 (Grado e Forza [entrante/uscente]). Dato un indice $i \in \mathcal{I}$, in un grafo indiretto non pesato si definisce grado la somma

$$k_i \equiv \sum_{j=1}^{|\mathcal{I}|} a_{i,j} = \sum_{j=1}^{|\mathcal{I}|} a_{j,i}; \quad (2)$$

d'altra parte in un grafo diretto non pesato la seconda equivalenza non è piú [necessariamente] valida, ragion per cui è necessario specializzare il grado in entrante e uscente

$$k_i^i \equiv \sum_{j=1}^{|\mathcal{I}|} a_{j,i} \quad \text{e} \quad k_i^o \equiv \sum_{j=1}^{|\mathcal{I}|} a_{i,j}, \quad (3)$$

rispettivamente. Se il grafo è invece pesato le definizioni sono le stesse di (2, 3) ma con $w_{i,j}/w_{j,i}$ in luogo di $a_{i,j}/a_{j,i}$.

Infine in questa trattazione vale la seguente fondamentale ipotesi:

Ipotesi 2.1. Il grafo \mathcal{G} è assunto statico: in altre parole, \mathcal{I} e \mathcal{E} sono costanti nel tempo.

In altre parole si è solo interessati a prevedere come la popolazione si distribuisce rispetto a una topologia prestabilita *a priori*; ovviamente si tratta di una forte semplificazione dato che nella realtà la topologia delle connessioni interurbane si è chiaramente coevoluta assieme allo sviluppo delle città stesse.

2.2 RETI E CITTÀ

Dopo aver studiato un po' di teoria sui grafi sorge successivamente il problema di come rappresentare mediante i grafi la moltitudine di collegamenti possibili tra le città

Difatti vi sono molti modi di rappresentare le città mediante i grafi: i primi si pongono a livello *intraurbano*, o considerando le strade come lati e le loro intersezioni come nodi [11, § 3.1.3.1 p. 17], o rappresentando la rete di trasporto di tram, di bus e della metro [11, § 3.2.1 p. 22]; altri si pongono piú propriamente a livello *interurbano* prendono singolarmente in considerazione varie reti dei trasporti (ferroviario [11, § 3.1.3.2 p. 17], navale [11, § 3.1.4 p. 19], aereo [11, § 3.1.2 p. 13], ecc.).

Tuttavia, da quanto detto nella § 1, è chiaro che per predire la distribuzione della popolazione tra città valgono le seguenti osservazioni:

1. tutte le rappresentazioni intraurbane vanno scartate perché sono troppo fini, oltre a considerare movimenti limitatamente a una sola città;
2. d'altra parte tutte quelle interurbane vanno considerate contemporaneamente e non singolarmente siccome ogn'individuo può scegliere diversi trasporti per muoversi.

Ecco perché la corretta rete da considerare è quella legata ai movimenti pendolari [11, § 3.1.3.3 p. 18; 5] tra città che mostrano olisticamente tutt'i possibili collegamenti tra le città a prescindere del trasporto scelto; inoltre essa mostra collegamenti realistici associati a movimenti quotidiani anziché straordinari (vacanze, visite mediche, ecc.).

Una volta fissata la rappresentazione è necessario comprendere il tipo di grafo con cui si ha a che fare. Eppure la risposta è molto semplice e immediata dopo una semplice osservazione: un ente, ovvero una città, può interagire con un secondo senza che questo interagisca a sua volta col primo; si è dunque di fronte a un grafo diretto ma simmetrico (2.2) perché il contesto richiede di considerare la direzione d'interazione.

Come nota finale si osserva che questo tipo di grafo è spaziale [5, 11], ovvero i suoi nodi occupano un punto nello spazio euclideo; oltre a ciò tale osservazione è puramente formale, ma servirà successivamente quando si definiranno le leggi d'interazione. Perdi più, contrariamente a quanto si possa pensare di primo inizialmente, tale grafo non è a invarianza di scala [2] proprio a causa della sua natura spaziale [4, 11]; ciò non esclude l'esistenza di nodi più centrali di altri, ma solo che il massimo grado di un nodo è limitato superiormente dalla natura spaziale del grafo.

2.3 CENNI SUI DATI

La matrice di pendolarismo costruita è quella fornita dall'ISTAT del 1999 [9] seguendo l'esempio di [5]. I dati sono di fatto un *file* di testo formato da una serie di righe di 29 numeri il cui significato è spiegato dalla Tab. 2.1.

Dato	Col. iniziale	Lunghezza
Provincia di partenza	1	3
Comune di partenza	4	3
Sesso	7	1
Mezzo di trasporto	8	1
Cond.professionale	9	1
Orario di uscita	10	1
Tempo di percorrenza	11	1
Provincia di arrivo	12	3
Comune di arrivo	15	3
Numero di persone	18	12

Tabella 2.1: Formattazione di una sola riga nei dati ISTAT²; le linee barrate corrispondono a dati trascurati.

I comuni considerati sono tutt'i quell'italiani (8100) nel 1999, quindi è necessario estrapolare da essi le matrici di pendolarismo delle singole regioni e parallelamente quella complessiva dell'Italia.

Si notano tuttavia due principali difetti dei dati contenuti nel documento `Pen_91It.txt`:

² si v. il documento `trapen91.txt` per maggiori informazioni.

1. in alcune linee come comune di destinazione è segnato il codice «022008» che però non è elencato nel documento `elencom91.xls` che riassume tutti gli 8100 comuni italiani nel 1991;
2. in molte linee come comune di destinazione sono segnati dei codici incompleti contenenti solo l'ultima metà (codice comune) ma non la prima (codice provincia), e questi sono in totale 8: «215», «229», «241», «216», «203», «224», «236», «246».

In entrambi i casi i dati relativi [a quelli che sono potenzialmente refusi] sono stati trascurati nella corrente tesi; il primo errore è già di per sé molto raro, mentre il secondo è eccessivamente ambiguo da risolvere come si può notare dalla linea successiva relativa al codice «215»:

00200714212215_____1,

ove _ indica uno spazio, per la quale, tralasciando il fatto che vi sono in totale 6 città italiane collo stesso codice comunale, se ci si limita alla regione del Piemonte (vale a dire la stessa del comune di partenza «002007») esistono in realtà ben due città condividenti lo stesso codice: «001215» e «004215», eppure entrambe non hanno lo stesso codice provinciale («002») del comune di partenza. Da questo piccolo estratto, quindi, non si può nemmeno ipotizzare che il codice provinciale sia lo stesso tra il comune di partenza e d'arrivo; è allora chiaro che sia impossibile scegliere la prima metà per completare il codice del comune di destinazione³.

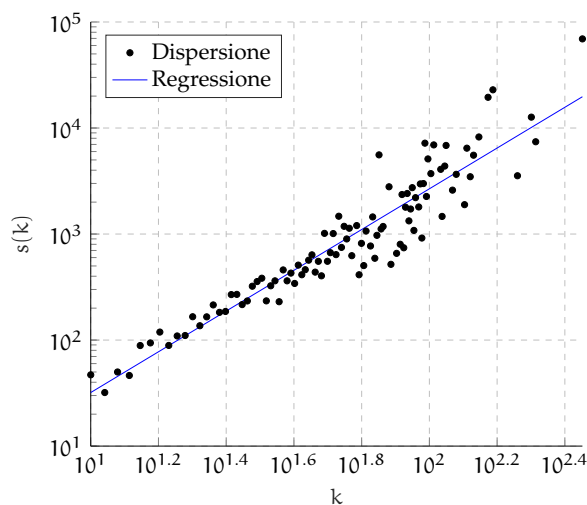


Figura 2.2: Forza contro grado $s(k)$ per la Sardegna.

Infine nelle Figg. 2.2 e 2.3 sono stati anche riprodotti i risultati di [5] colle seguenti osservazioni:

1. il coefficiente d'aggregazione medio $\langle C(k) \rangle = 0.453$ è quasi il doppio di quello riportato da [5, p. 911] (0.26);

³ Si potrebbe scegliere, tra tutt'i comuni collo stesso codice comunale, quello che minimizza la distanza geografica (ossia euleriana) siccome le coordinate dei punti rappresentativi di ogni comune sono disponibili nell'ISTAT; in ogni caso, come si vedrà nella § 4, sono correzioni del tutto innecessarie.

2. la Fig. 2.3c rispetto alla [5, Fig. 3b] presenta sia una forma leggermente diversa, seppure la tendenza a mezza luna sia la stessa, che limiti superiori e inferiori più estesi;
3. il coefficiente d'aggregazione pesato rappresentato nella Fig. 2.3g ha chiaramente una tendenza decrescente, anche se molto lenta, ragione per cui si può considerare comunque «circa costante» come afferma [5];
4. i primi due pesi più elevati, riportati nella Tab. 2.2, non corrispondono: ciò si può giustificare molto probabilmente come refuso da parte di [5] poiché i collegamenti «Cagliari-Sassari» e «Sassari-Olbia» sono molto distanti geograficamente⁴, per cui è ragionevole che i flussi siano bassi.

Connessione	Peso	Connessione	Peso
Cagliari-Quartu	14709	Cagliari-Sassari	13953
Cagliari-Selargius	7995	Sassari-Olbia	7246
Assemini-Selargius	4418	Cagliari-Assemini	4226
Porto Torres-Sassari	4149	Porto Torres-Sassari	3993
Cagliari-Capoterra	3865	Cagliari-Capoterra	3731
(a) Pesi maggiori correnti.		(b) Pesi maggiori di [5]	

Tabella 2.2: Confronto con [5] dei pesi maggiori.

Come si vedrà tra qualche capitolo, ma soprattutto perché sono davvero lievi, queste discrepanze sono irrilevanti per quanto concerne questo elaborato. Inoltre, l'unico risultato importante recuperato è la correlazione positiva tra la forza e il grado già mostrato nella Fig. 2.2.

⁴ Il primo equivale a percorrere l'intera lunghezza dell'isola ogni giorno.

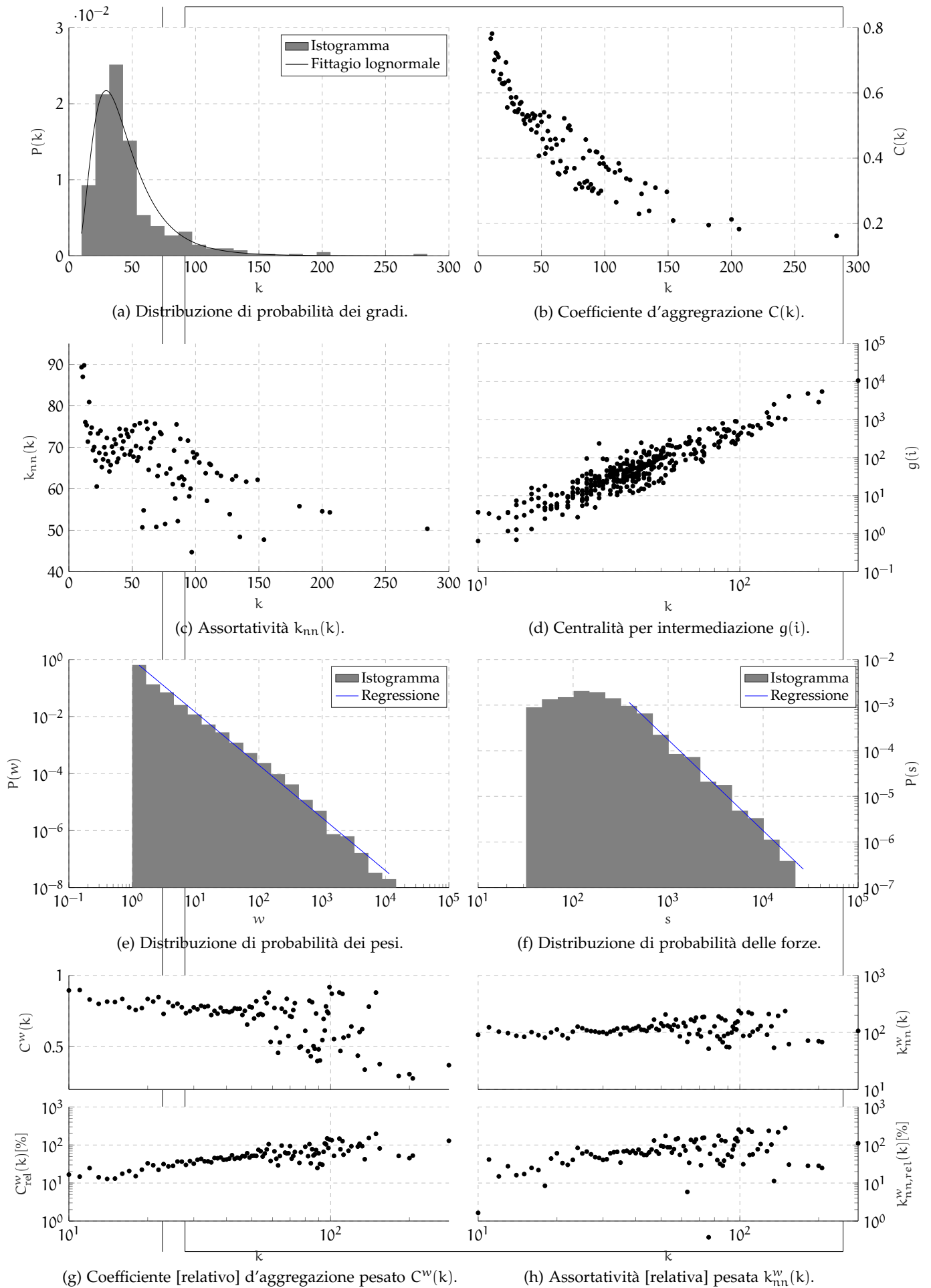


Figura 2.3: Riproduzione dei principali grafici di [5] [a cui si rimanda per maggiori dettagli sui vari coefficienti] relativi alla topologia della rete del pendolarismo sarda.

3

TEORIA CINETICA DEI SISTEMI MULTIAGENTE

In questo capitolo è discussa la Teoria Cinetica dei Sistemi MultiAgente (TCSMA) necessaria per ottenere i risultati nella § 4.

In esso si deriva prima la celeberrima equazione di Boltzmann a partire da una descrizione stocastica delle particelle di un gas; quindi si passa a generalizzare tali risultati per mediare le interazioni tramite una struttura topologica sottostante gli agenti, sia in modo esatto che approssimato; successivamente si descrivono le equazioni cinetiche nel caso corrente; infine si conclude analizzando le ipotesi semplificative che portano dalla presente teoria a quella classica di Boltzmann.

3.1 DESCRIZIONE CINETICA CLASSICA

3.1.1 Equazione di Boltzmann omogenea

Prima di procedere è necessario definire il tipo di gas che si sta considerando.

Ipotesi 3.1. Si consideri un gas composto da particelle che soddisfa le successive importanti ipotesi:

G₁ è uniformemente distribuito nello spazio cosicché in ogni punto la distribuzione statistica delle velocità è la stessa;

G₂ è rarefatto, da cui si può assumere che solo interazioni binarie possono aver luogo;

G₃ è composto da particelle indistinguibili e aventi ugual massa;

G₄ le collisioni sono elastiche per cui, unito alla G₂, porta alla conservazione dell'impulso e dell'energia:

$$\mathbf{v}' + \mathbf{v}'_* = \mathbf{v} + \mathbf{v}_*, \quad (\text{CI})$$

$$|\mathbf{v}'|^2 + |\mathbf{v}'_*|^2 = |\mathbf{v}|^2 + |\mathbf{v}_*|^2, \quad (\text{CE})$$

ove $\mathbf{v}', \mathbf{v}'_* \in \mathbb{R}^3$ sono le velocità poscollisionali che collidono colle velocità precollisionali $\mathbf{v}, \mathbf{v}_* \in \mathbb{R}^3$, delle due particelle interagenti¹.

Dalla G₄ si possono in realtà caratterizzare pienamente le velocità poscollisionali.

Proposizione 3.1. Esistono due funzioni $\hat{\mathbf{v}}, \hat{\mathbf{v}}_*: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ che soddisfanno (CI, CE) e tali che

$$\begin{aligned} \mathbf{v}' &= \hat{\mathbf{v}}(\mathbf{v}, \mathbf{v}_*) \equiv \mathbf{v} + [(\mathbf{v}_* - \mathbf{v}) \cdot \mathbf{n}] \mathbf{n}, \\ \mathbf{v}'_* &= \hat{\mathbf{v}}_*(\mathbf{v}, \mathbf{v}_*) \equiv \mathbf{v}_* + [(\mathbf{v} - \mathbf{v}_*) \cdot \mathbf{n}] \mathbf{n}, \end{aligned} \quad (1)$$

¹ In questo contesto non è necessario distinguere quale delle due sia quella interagente e quale quella ricevente, contrariamente a quello delle città.

ove $\mathbf{n} \in \mathbb{S}^2$ è un qualunque vettore appartenente alla sfera unitaria di \mathbb{R}^3 .

Dimostrazione. Assumendo l'ansatz

$$\mathbf{v}' = \mathbf{v} - \gamma \mathbf{n} \quad \text{e} \quad \mathbf{v}'_* = \mathbf{v}_* + \gamma \mathbf{n}, \quad (2)$$

in cui $\gamma \in \mathbb{R}$ è un parametro da determinare, si può notare che la (CI) è già verificata, mentre imponendo la (CE) si ha

$$\begin{aligned} |\mathbf{v}'|^2 + |\mathbf{v}'_*|^2 &= |\mathbf{v} - \gamma \mathbf{n}|^2 + |\mathbf{v}_* + \gamma \mathbf{n}|^2 \\ &= (\mathbf{v} - \gamma \mathbf{n}) \cdot (\mathbf{v} - \gamma \mathbf{n}) + (\mathbf{v}_* + \gamma \mathbf{n}) \cdot (\mathbf{v}_* + \gamma \mathbf{n}) \\ &= |\mathbf{v}|^2 - 2\gamma(\mathbf{v} \cdot \mathbf{n}) + \gamma^2 + |\mathbf{v}_*|^2 + 2\gamma(\mathbf{v}_* \cdot \mathbf{n}) + \gamma^2 \\ &= |\mathbf{v}|^2 + |\mathbf{v}_*|^2 - 2\gamma[(\mathbf{v}_* - \mathbf{v}) \cdot \mathbf{n} + 2\gamma] \\ &= |\mathbf{v}|^2 + |\mathbf{v}_*|^2 \implies \gamma[(\mathbf{v}_* - \mathbf{v}) \cdot \mathbf{n} + \gamma] = 0, \end{aligned}$$

ma escludendo il caso banale di $\gamma = 0$ (nessuna interazione) si ha

$$\gamma = \hat{\gamma}(\mathbf{v}, \mathbf{v}_*) \equiv (\mathbf{v} - \mathbf{v}_*) \cdot \mathbf{n},$$

per cui γ è in realtà una funzione delle velocità precollisionali; la (2) diventa allora

$$\mathbf{v}' = \mathbf{v} + [(\mathbf{v}_* - \mathbf{v}) \cdot \mathbf{n}] \mathbf{n} \quad \text{e} \quad \mathbf{v}'_* = \mathbf{v}_* + [(\mathbf{v} - \mathbf{v}_*) \cdot \mathbf{n}] \mathbf{n}. \quad (3)$$

□

Si noti come la precedente dimostrazione lasci arbitrario $\mathbf{n} \in \mathbb{S}^2$, seppure da un punto di vista sia ragionevole porlo parallelo al vettore passante per i centri delle particelle collidenti.

Per poter dedurre l'equazione di Boltzmann omogenea è però necessaria una descrizione statistica del precedente gas: sia $\mathbf{V}_t, \mathbf{V}_t^* \in \mathbb{R}^3$ le variabili aleatorie² delle velocità precollisionali, di cui le \mathbf{v} e \mathbf{v}_* precedenti sono due realizzazioni, al tempo t e similmente per $\mathbf{V}'_t, \mathbf{V}'_t^* \in \mathbb{R}^3$.

Le leggi d'interazione nella (1) diventano

$$\begin{aligned} \mathbf{V}' &= \hat{\mathbf{V}}(\mathbf{V}, \mathbf{V}_*) \equiv \mathbf{V} + [(\mathbf{V}_* - \mathbf{V}) \cdot \mathbf{n}] \mathbf{n}, \\ \mathbf{V}'_* &= \hat{\mathbf{V}}_*(\mathbf{V}, \mathbf{V}_*) \equiv \mathbf{V}_* + [(\mathbf{V} - \mathbf{V}_*) \cdot \mathbf{n}] \mathbf{n}, \end{aligned} \quad (4)$$

in cui anche \mathbf{n} risulta essere aleatoriamente distribuita; usualmente s'ipotizza $\mathbf{n} \sim \mathcal{U}(\mathbb{S}^2)$, vale a dire che non vi sono direzioni preferenziali essendo queste uniformemente distribuite sulla sfera unitaria.

Invece per modellare la dinamica basata su interazioni binarie si fissa un passo temporale $\Delta t > 0$, si campionano indipendentemente due agenti e si presuppone che questi interagiscano con una certa probabilità, ovvero

$$[\text{AR}] \begin{cases} \mathbf{V}_{t+\Delta t} = (1 - \Theta) \mathbf{V}_t + \Theta \mathbf{V}'_t, \\ \mathbf{V}_{t+\Delta t}^* = (1 - \Theta) \mathbf{V}_t^* + \Theta \mathbf{V}'_t^*. \end{cases}$$

ove

$$\Theta \sim \text{Bernoulli}(B((\mathbf{V}_t^* - \mathbf{V}_t) \cdot \mathbf{n}) \Delta t)$$

² Si noti come queste non dipendano dall'indice proprio per l'ipotesi G₃ d'indistinguibilità delle particelle.

è una variabile aleatoria, indipendente da V_t e V_t^* , che indica se l'interazione avviene o no; essa dipende dalla funzione $B: \mathbb{R} \rightarrow \mathbb{R}_+$, detta nucleo di collisione, che permette d'avere una maggiore espressività sulle collisioni molecolari le quali possono influenzare il tasso d'interazione tra le molecole. Per definizione di Θ deve valere $B((V_t^* - V_t) \cdot n) \Delta t \leq 1$ condizione soddisfacibile anche prendendo un passo temporale Δt adattivo.

Si noti come B dipenda da $(V_* - V) \cdot n$ cosa che permette d'interpretare n come la direzione lungo la quale le interazioni sono più frequenti.

3.1.2 Equazione di Boltzmann disomogenea

3.1.3 Equazione di tipo Boltzmann omogenea

3.2 DESCRIZIONE CINETICA RETALE

3.2.1 Impostazione

La popolazione degli agenti evolve a causa delle interazioni con altri agenti connessi. Seguendo la teoria cinetica collisionale, l'ipotesi fondamentale ipotizzata è che solo le interazioni binarie siano rilevanti: le interazioni fra tre o più agenti possono essere trascurate.

Dopodiché, sia $X \in \mathcal{I}$ la posizione di un agente sul grafo $\mathcal{G} = (\mathcal{I}, \mathcal{E})$, ove \mathcal{I} è l'insieme dei vertici mentre \mathcal{E} dei lati di \mathcal{G} . Si assume che il grafico sia statico, ovvero che le connessioni tra agenti non varia nel tempo.

Si consideri, allora, un generico agente rappresentativo, il cui stato microscopico è descritto dal processo stocastico $(X, S_t)_{t \geq 0}$; la funzione $S_t: \Omega \rightarrow \mathcal{P}$ è una variabile aleatoria da uno spazio astratto Ω allo spazio delle popolazioni \mathcal{P} e indica la popolazione dell'agente al tempo $t \geq 0$. Tale variabile aleatoria evolve nel tempo per le interazioni binarie con altri agenti mediate dalle connessioni descritte da \mathcal{E} , definendo così un processo stocastico $\{S_t, t \in [0, +\infty)\}$.

Nel complesso si descrive statisticamente lo stato microscopico X, S_t dell'agente mediante una probabilità di misura $f = f(x, s, t)$, discreta in $x \in \mathcal{I}$ e continua in $s \in \mathcal{P}$. Pertanto si può dare alla f la seguente forma

$$f(x, s, t) = \frac{1}{N} \sum_{i \in \mathcal{I}} f_i(s, t) \otimes \delta(x - i), \quad (6)$$

ove $N \equiv |\mathcal{I}|$ è il numero totale d'agenti/vertici del grafo mentre $\delta(\cdot)$ denota la delta di Dirac centrata all'origine; d'altra parte

$$f_i = f_i(s, t): \mathcal{P} \times [0, +\infty) \rightarrow \mathbb{R}_+$$

è la densità di probabilità della taglia S_t dell'agente $X = i$.

Logicamente si richiede

$$\int_{\mathcal{P}} f_i(s, t) ds = 1, \quad \forall t \geq 0, \forall i \in \mathcal{I},$$

che implica coerentemente

$$\int_{\mathcal{I}} \int_{\mathcal{P}} f(x, s, t) ds dx = 1 \quad \forall t \geq 0$$

3.2.2 Algoritmi d'interazione

Un algoritmo d'interazione è una regola che descrive come gli agenti interagiscono a coppie e modificano di conseguenza il loro stato nel tempo; nel dettaglio, in un dato passo temporale $\Delta t > 0$ si assume che un agente $(X, S_t) \in \mathcal{I} \times \mathcal{P}$ cambi la sua popolazione a $S_{t+\Delta t} \in \mathcal{P}$ a seguito di un'interazione con un altro agente $(X^*, S_t^*) \in \mathcal{I} \times \mathcal{P}$ secondo il successivo schema

$$S_{t+\Delta t} = (1 - \Theta)S_t + \Theta S_t', \quad (7)$$

ove $\Theta \in [0, 1]$ è una variabile aleatoria che tiene in considerazione qualora l'interazione tra i due agenti effettivamente si manifesti ($\Theta = 1$) o no ($\Theta = 0$); d'altro canto $S_t' \in \mathcal{P}$ è la nuova popolazione ottenuta dall'agente (X, V_t) in seguito a un'interazione avvenuta.

Con maggiore dettaglio si pone

$$\Theta \sim \text{Bernoulli}(A(X, X^*), \Delta t), \quad (8)$$

il che significa che la probabilità che un'interazione avvenga è proporzionale al passo temporale d'interazione Δt mediante un nucleo d'interazione $A(X, X^*) = 1$, che contiene le informazioni sui lati del grafo, e quindi alle connessioni tra gli agenti, ponendo

$$A(X, X^*) = \begin{cases} 1 & \text{se } (X, X^*) \in \mathcal{E}, \\ 0 & \text{se } (X, X^*) \notin \mathcal{E}, \end{cases}$$

dove la coppia ordinata (X, X^*) denota il lato dal vertice X al vertice X^* ; per coerenza è necessario imporre $\Delta t \leq 1$ che impone un limite superiore al massimo passo temporale ammissibile, seppure tale condizione sia molto facile da verificare nella pratica.

La popolazione postinterazione è una variabile aleatoria $S_t' : \Omega \rightarrow \mathcal{P}$ dipendente in generale dagli stati preinterazione V_t, V_t^* degli agenti integranti:

$$V_t'(\omega) = \Psi(S_t(\omega), S_t^*(\omega), \omega), \quad \omega \in \Omega, \quad (9)$$

in cui $\Psi : \mathcal{P}^2 \times \Omega \rightarrow \mathcal{P}$ è funzione nota potenzialmente stocastica.

3.2.3 Interazioni azione-reazione

Nel contesto delle città è chiaro che qualora una città-nodo interagisca con un'altra entrambe debbano variare il loro stato.

Sia S la città interagente e S^* quella subente, allora in un grafo diretto si può distinguere il senso d'interazione: quella *in avanti* avviene sse $(S, S^*) \in \mathcal{E}$ mentre quella *in indietro* sse $(S^*, S) \in \mathcal{E}$; tuttavia questa distinzione è inutile in questo caso di grafo diretto con matrice d'adiacenza simmetrica. Pertanto l'agente (X^*, S_t^*) aggiorna la sua popolazione attraverso una regola analoga a quella dell'agente (X, S_t) :

$$S_{t+\Delta t}^* = (1 - \Theta)S_t^* + \Theta S_t'^*, \quad (10)$$

ove si osserva che la Θ è la stessa della (7) la cui legge dipende da $A(X, X^*)$ ma non da $A(X^*, X)$; tale dettaglio è da tenere in considerazione nel caso

in cui la matrice di adiacenza non sia simmetrica, ma in questo caso di simmetria non è rilevante. L'opinione postinterazione

$$S'_t(\omega) = \Psi_*(S_t^*(\omega), S_t(\omega), \omega), \quad \omega \in \Omega,$$

è definita mediante una funzione $\Psi_*: \mathcal{P}^2 \times \Omega \rightarrow \mathcal{P}$ potenzialmente diversa da Ψ . Questo tipo d'interazione, prendendo come riferimento [12, § 2.2.1] sarà identificato come azione-reazione, e riassunto dall'algoritmo

$$[\text{AR}] \begin{cases} S_{t+\Delta t} = (1-\Theta)S_t + \Theta S'_t, \\ S_{t+\Delta t}^* = (1-\Theta)S_t^* + \Theta S_t^{*'} \end{cases}$$

Si conclude questa sezione osservando che gli agenti (X, S_t) , (X^*, S_t^*) sono campionati casualmente e uniformemente a ogni passo temporale.

3.3 DERIVAZIONE DELL'EQUAZIONI CINETICHE

3.3.1 Derivazione esatta

Una descrizione cinetica dell'algoritmo [AR] coincide con dell'equazioni d'evoluzione per le distribuzioni di probabilità f_i delle opinioni degli agenti; per derivarle si procede mediante un metodo classico nella teoria dei sistemi multiagente [12, 13].

Sia $\Phi \equiv \Phi(x, s): \mathcal{I} \times \mathcal{P} \rightarrow \mathbb{R}$ un osservabile arbitrario (funzione *test*), cioè una quantità che si può calcolare sapendo lo stato microscopico di un generico agente rappresentativo del sistema. Allora dalla prima equazione in [AR], valutando il valore atteso dell'osservabile postinterazione rispetto agli indici e alla popolazione a tempo fissato, si ha §

$$\begin{aligned} \mathbb{E}[\Phi(X, S_{t+\Delta t})] &= \mathbb{E}[\mathbb{E}[\Phi(X, (1-\Theta)S_t + \Theta\Psi(S_t, S_t^*, \omega))A(X, X_*)\Delta t | X, X_*]] \\ &= \mathbb{E}[\mathbb{E}[\Phi(X, S_t)(1-A(X, X_*)\Delta t) \\ &\quad + \Phi(X, \Psi(S_t, S_t^*, \omega))A(X, X_*)\Delta t]], \end{aligned}$$

da cui, riordinando i termini e dividendo ambo i membri per Δt , si deduce

$$\frac{\mathbb{E}[\Phi(X, S_{t+\Delta t})] - \mathbb{E}[\Phi(X, S_t)]}{\Delta t} = \mathbb{E}[A(X, X_*)(\Phi(X, \Psi(S_t, S_t^*, \omega)) - \Phi(X, S_t))],$$

laddove, prendendo il limite $\Delta t \rightarrow 0^+$, si ricava formalmente

$$\frac{d\mathbb{E}[\Phi(X, S_t)]}{dt} = \mathbb{E}[A(X, X^*)(\Phi(X, \Psi(S_t, S_t^*, \omega)) - \Phi(X, S_t))]. \quad (11)$$

Si può ricavare una simile equazione ripetendo i precedenti passaggi ma colla seconda equazione di [AR], da cui

$$\frac{d\mathbb{E}[\Phi(X^*, S_t^*)]}{dt} = \mathbb{E}[A(X, X^*)(\Phi(X^*, \Psi_*(S_t^*, S_t, \omega)) - \Phi(X^*, S_t^*))]. \quad (12)$$

Osservando che le coppie (X, S_t) e (X^*, S_t^*) fanno riferimento a un agente rappresentativo generico del sistema, vale

$$\mathbb{E}[\Phi(X, S_t)] = \mathbb{E}[\Phi(X^*, S_t^*)]$$

cosicché, sommando (11, 12), si ha

$$\frac{dE[\Phi(X, S_t)]}{dt} = \frac{1}{2} E[A(X, X^*) (\Phi(X, \Psi(S_t, S_t^*, \omega)) + \Phi(X^*, \Psi_*(S_t^*, S_t, \omega)) - \Phi(X, S_t) - \Phi(X^*, S_t^*))],$$

ed espandendo la definizione della media si arriva a

$$\frac{d}{dt} \int_{\mathcal{J}} \int_{\mathcal{P}} \Phi f dv dx = \int_{\mathcal{J}^2} \int_{\mathcal{P}^2} A(x, x_*) \frac{\langle \Phi' + \Phi'_* - \Phi - \Phi_* \rangle}{2} f f_* ds ds_* dx dx_*, \quad (13)$$

ove [per brevità] si sono omessi gli argomenti dell'osservabile e della distribuzione (6):

$$f \equiv f(x, s, t), \quad \Phi \equiv \Phi(x, s), \quad \Phi_* \equiv \Phi(x_*, s_*), \quad \Phi' \equiv \Phi(x, s'), \quad \Phi'_* \equiv \Phi(x_*, s'_*),$$

si è inoltre imposto

$$s' = \Psi(s, s_*, \omega) \quad s'_* = \Psi_*(s_*, s, \omega), \quad (14)$$

mentre $\langle \cdot \rangle$ indica il valore atteso rispetto alla potenziale stocasticità delle funzioni Ψ e Ψ_* .

Si noti che la (13) è valida per ogni funzione *test* Φ per cui è un'equazione debole per la distribuzione f . (commento su Fokker-Planck {S})

Si osservi anche che la (13) è scritta sotto l'ipotesi della propagazione del caos: ogni due potenziali agenti interagenti sono tra di loro campionati indipendentemente. Questa assunzione è classicamente usata, per es. nella teorica cinetica di Boltzmann, per ottenere un'equazione chiusa per la distribuzione f di una particella, siccome permette di fattorizzare la distribuzione di probabilità congiunta $g(x, x_*, s, s_*, t)$ degli agenti interagenti nel prodotto $f(x, s, t) f(x_*, s_*, t)$.

Dalla (13) con una scelta adeguata della funzione *test* Ψ , è possibile recuperare un sistema di equazioni debole per le f_i . Sia $\Psi(x, s) = \phi_i(x) \varphi(s)$, dove $\phi_i: \mathcal{J} \rightarrow \mathbb{R}$ è tale che $\phi_i(i) = 1$ mentre $\phi_i(x) = 0$ per ogni $x \in \mathcal{J} \setminus \{i\}$ e $\varphi: \mathcal{O} \rightarrow \mathbb{R}$ è arbitrario. Allora usando la (6) dentro la (13) si ricava

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{P}} \varphi f_i ds &= \frac{1}{2N} \sum_{j \in \mathcal{J}} A(i, j) \int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle f_i f_j^* ds ds_* \\ &\quad + \frac{1}{2N} \sum_{j \in \mathcal{J}} A(j, i) \int_{\mathcal{P}^2} \langle \varphi'_* - \varphi_* \rangle f_j f_i^* ds ds_*, \quad \forall i \in \mathcal{J}, \end{aligned} \quad (15)$$

ove gli argomenti di tutte le funzioni sono stati sottintesi, vale a dire

$$f_i \equiv f_i(s, t) \quad \forall i \in \mathcal{J}, \quad \varphi \equiv \varphi(s), \quad \varphi_* \equiv \varphi(s_*) \quad \text{e} \quad \varphi' \equiv \varphi(s').$$

Tal'equazione si può anche derivare, sempre sotto l'ipotesi della propagazione del caos, dalla gerarchia tipo BBGKY (v. se aggiungere il riferimento {S}). In aggiunta si può convertire in forma matriciale introducendo la distribuzione vettoriale $\mathbf{f} \equiv (f_i(s, t))_{i \in \mathcal{J}}$ e la matrice $\mathbf{M} \equiv (A(i, j))_{i, j \in \mathcal{J}} \in \mathbb{R}^{N \times N}$:

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{P}} \varphi \mathbf{f} ds &= \frac{1}{2N} \int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle \mathbf{f} \odot \mathbf{M} \mathbf{f}_* ds ds_* \\ &\quad + \frac{1}{2N} \int_{\mathcal{P}^2} \langle \varphi'_* - \varphi_* \rangle \mathbf{M}^\top \mathbf{f} \odot \mathbf{f}_* ds ds_*, \end{aligned} \quad (16)$$

ove \odot indica il prodotto di Hadamard e \mathbf{M}^\top al trasposto di \mathbf{M} . Si noti che \mathbf{M} altro non è che la matrice d'adiacenza di \mathcal{G} .

3.3.2 Interazioni tra città

Si può ora approfondire il tipo d'interazioni ipotizzate tra città su grafi.

Innanzitutto, è chiaro che l'interazione d'interesse sia di tipo «azione-reazione» descritta da [AR]: infatti, se una città interagisce con un'altra, scambiando popolazione, entrambi variano il proprio stato ma non è detto che la seconda interagisca a sua volta colla prima.

Tuttavia, se una città può interagire con un'altra, allora è sempre possibile l'opposto; dunque il grafo in questione è diretto ma con struttura indiretta, ovvero la sua matrice d'adiacenza è simmetrica; a livello matematico, ciò implica che la matrice d'adiacenza \mathbf{M} è simmetrica.

Gli stati postinterazione (14) prendono come riferimento leggi d'interazioni lineari

$$\begin{cases} S'_t = pS_t + qS_t^*, \\ S_t^{*'} = p_*S_t + q_*S_t^*, \end{cases} \quad (17)$$

le quali, specializzate, presentano invece la seguente forma:

$$\begin{cases} s' = s(1 - E(s, s_*) + \gamma) \\ s_*' = s_* + sI(s, s_*) \end{cases} \quad (18)$$

ove s e s_* sono le città interagente e subente rispettivamente, $E(s, s_*)$ e $I(s, s_*)$ sono rispettivamente i tassi di emigrazione e immigrazione, mentre γ rappresenta fluttuazioni stocastiche da definire; rispetto alle leggi d'interazioni lineari (17) le (18) soddisfano

$$\begin{aligned} p(s, s_*) &\equiv s[1 - E(s, s_*) + \gamma] & e & & p_*(s, s_*) &\equiv s_* \\ q(s, s_*) &\equiv 0, & e & & q_*(s, s_*) &\equiv sI(s, s_*), \end{aligned} \quad (19)$$

rispettivamente per la prima e seconda legge. Ovviamente questo scambio deve conservare [in media] la popolazione totale da cui

$$\begin{aligned} s + s_* &= \langle s + s_* \rangle = \langle s' + s_*' \rangle \\ &= s - sE(s, s_*) + s_* + sI(s, s_*) \implies E(s, s_*) = I(s, s_*), \end{aligned} \quad (20)$$

ciò ha senso perché l'emigrazione e l'immigrazione sono fenomeni relativi (invertendo s ed s_* sarebbe l'opposto). La scelta di $E(s, s_*)$ dipende da come si vuole modellizzare il fenomeno dell'immigrazione, e in questo manoscritto si è modificata la [8, (2.2), § 2, p. 223] mediante la [8, (4.5), § 4, p. 228]:

$$E(s, s_*) \equiv \lambda \frac{(s_*/s)^\alpha}{1 + (s_*/s)^\alpha}, \quad (21)$$

che in essenza è una funzione di Hill di ordine α , in cui v'è un tasso di emigrazione maggiore verso città con popolazione relativa, data dal rapporto s_*/s , maggiore; gli unici due parametri presenti, invece, presentano il seguente significato:

- ◇ $\lambda \in (0, 1)$ rappresenta l'attrattività dei poli, ossia la frazione che le città più popolose riescono al massimo ad attrarre in un'interazione;
- ◇ $\alpha \in \mathbb{R}^+$ indica la rapidità d'emigrazione e influenza quanto rapidamente il rapporto s_*/s raggiunge la massima attrattività λ .

In poche parole la (21) descrive la tendenza degli individui di aggregarsi per i più svariati motivi: lavoro, sicurezza, famiglia, eccetera.

In questo caso si hanno quindi interazioni non simmetriche, poiché dalla (19) $p \neq q$ e $q \neq p$, e non lineari, a causa della (21).

Non manca che caratterizzare il tipo di perturbazione γ per avere uno stato postinterazione fisicamente sensato; difatti, è chiaro che rigorosamente $\mathcal{P} \equiv \mathbb{N}$ ma è più agevole supporre $\mathcal{P} \equiv \mathbb{R}^+$ per poi approssimare per eccesso o difetto il numero intero effettivo; pertanto, dalla (18), si ha

$$s' > 0 \implies \gamma > E(s, s_*) - 1 \quad (22)$$

mentre s'_* è per definizione sempre positivo. La scelta di $>$ anziché \geq nella (22) è ben fondata: di fatto si sta supponendo che le fluttuazioni non possono annullare la popolazione di una città; difatti qualora $\gamma = E(s, s_*) - 1$ si avrebbe $s' = 0$ dalla (18), situazione che si vuole evitare³ dato che nella (21) compare il rapporto tra popolazioni delle città interagenti.

Perdipiù, le fluttuazioni rappresentano a grandi linee quei fenomeni complessivi di nascita e di morte che vengono considerati ma non direttamente modellati; pertanto γ deve soddisfare le seguenti due caratteristiche:

- F1 possono assumere sia valori positivi che negativi, ma non minori del vincolo imposto da (22);
- F2 la media è scelta arbitrariamente posta allo zero, ossia $\langle \gamma \rangle = 0$;
- F3 seppure non vi siano limiti superiori per l'entità della fluttuazione, è chiaro che più grande questa è meno è probabile.

Con queste si possono allora analizzare alcune distribuzioni continue:

- ◇ la distribuzione normale non soddisfa la F1 poiché può assumere valori reali arbitrari con probabilità non nulla;
- ◇ la distribuzione uniforme è adeguata solo per intervalli finiti e diventa degenerare quando un suo estremo diverge, per cui non soddisfa né F2 né F3, mentre F1 si;
- ◇ la distribuzione esponenziale è quella più promettente perché riflette sia F2 (dopo un'opportuna traslazione dei valori campionati) che F3, ma sfortunatamente non F1 perché il valore estremo $\gamma = E(s, s_*) - 1$ ha probabilità non nulla [anzi massima] d'essere campionato;
- ◇ l'unica distribuzione che soddisfa tutt'e tre le caratteristiche ricercate è proprio la distribuzione gamma.

Si consideri allora una distribuzione gamma avente la seguente funzione di densità di probabilità:

$$f(x) = \frac{1}{\theta \Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\theta}\right), \quad (23)$$

ove α e θ sono i parametri rispettivamente di forma e di scala, mentre $\Gamma(\alpha) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ è la funzione gamma:

$$\Gamma(\alpha) \equiv \int_0^{+\infty} y^{\alpha-1} e^{-y} dy. \quad (24)$$

³ Ciò non significa che il modello non possa modellare lo spopolamento di una città, poiché la sua taglia può arbitrariamente avvicinarsi a zero [ma mai esserne uguale], raggiungendolo solo *a posteriori* dopo l'approssimazione dai numeri reali a quelli interi.

437 Si scelga $\hat{\gamma} \sim \text{Gamma}(\alpha, \theta)$, che soddisfa per definizione la F3, e s'imponga

$$\langle \hat{\gamma} \rangle = 1 - E(s, s_*) = \alpha \theta \quad \text{e} \quad \langle \hat{\gamma}^2 \rangle = \sigma^2 = \alpha \theta^2, \quad (25)$$

438 con $\sigma \in \mathbb{R}^+$ equivalente alla deviazione *standard* mentre σ^2 alla varianza, da
439 cui

$$\alpha = \frac{(1 - E(s, s_*))^2}{\sigma^2} \quad \text{e} \quad \theta = \frac{\sigma^2}{1 - E(s, s_*)}. \quad (26)$$

440 Con tale scelta dei parametri è possibile soddisfare la F2 semplicemente tra-
441 slando i valori campionanti della $\hat{\gamma}$ di $-\langle \hat{\gamma} \rangle$, ossia si considera la distribuzio-
442 ne $\gamma \sim \hat{\gamma} - \langle \hat{\gamma} \rangle$:

$$\langle \gamma \rangle = \langle \hat{\gamma} - \langle \hat{\gamma} \rangle \rangle = \langle \hat{\gamma} \rangle - \langle \hat{\gamma} \rangle = 0.$$

443 D'altra parte la F1 necessità di salvaguardarsi dai casi degeneri della distri-
444 buzione gamma: essa infatti se $\alpha = 1$ diventa un'esponenziale di parametro
445 θ , mentre se $\alpha < 1$ diverge all'origine; per avere quindi una probabilità nulla
446 di campionare l'origine [e quindi $-\langle \hat{\gamma} \rangle$ dopo la traslazione] è necessario porre

$$\alpha > 1 \implies \frac{(1 - E(s, s_*))^2}{\sigma^2} > 1,$$

447 ma nel caso peggiore $1 - E(s, s_*) = 1 - \lambda$ da cui

$$\frac{(1 - \lambda)^2}{\sigma^2} > 1 \implies \sigma^2 < (1 - \lambda)^2 \implies \sigma < |1 - \lambda| = 1 - \lambda, \quad (27)$$

448 siccome $E(s, s_*) \in (0, \lambda) \quad \forall s, s_* \in \mathcal{P}$ e $\lambda \in (0, 1)$. La (27) implica quindi che non
449 è possibile avere una varianza arbitraria per poter soddisfare la F1, ma che
450 questa è limitata superiormente dall'attrattività dei poli: più è grande λ più
451 piccola è la varianza, e viceversa.

452 3.3.3 Derivazione approssimata

453 3.4 NESSO DISCRETO-CONTINUO

454 È d'interesse esplorare il legame presente tra la (15) coll'equazione classica
455 di Boltzmann {S}.

456 3.4.1 Ipotesi semplificative

457 A questo scopo si possono fare tre principali ipotesi semplificative da appli-
458 care alla (15):

459 IS1 Si presuppone che il grafo sia completamente connesso e quindi
460 che la matrice d'adiacenza sia unitaria $A \equiv I$.

461 IS2 Si assume che gli agenti siano indistinguibili:

$$f_i(s, t) = f_j(s, t) = f(s, t) \quad \forall i, j \in \mathcal{I}, \quad (28)$$

462 IS3 S'ipotizza che le interazioni siano simmetriche:

$$s' = \Psi(s, s_*) = \Psi_*(s_*, s), \quad \text{ove } s_* = \Psi_*(s, s_*). \quad (29)$$

3.4.2 Analisi della IS1

Con tal'ipotesi la (15) diventa

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi f_i ds = \frac{1}{2N} \left[\sum_{j \in \mathcal{J}} \int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle f_i f_j^* ds ds_* + \sum_{j \in \mathcal{J}} \int_{\mathcal{P}^2} \langle \varphi'_* - \varphi_* \rangle f_j f_i^* ds ds_* \right],$$

e valutando la distribuzione marginale della (6) rispetto agli indici

$$F(s, t) \equiv \int_{\mathcal{J}} f(x, s, t) dx = \frac{1}{N} \sum_{i \in \mathcal{J}} f_i(s, t) \otimes \int_{\mathcal{J}} \delta(x - i) dx = \frac{1}{N} \sum_{i \in \mathcal{J}} f_i(s, t), \quad (30)$$

che corrisponde a una media tra le distribuzioni dei singoli agenti, si ricava

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi f_i ds = \frac{1}{2} \left[\int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle f_i F^* ds ds_* + \int_{\mathcal{P}^2} \langle \varphi'_* - \varphi_* \rangle F f_i^* ds ds_* \right];$$

mediando ora rispetto a tutti gli agenti, si ha

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi F ds = \frac{1}{2} \int_{\mathcal{P}^2} \langle \varphi' - \varphi + \varphi'_* - \varphi_* \rangle F F^* ds ds_*, \quad (31)$$

la quale è formalmente analoga a quella classica di Boltzmann [8]. Ciò significa che con tal'ipotesi semplificativa, nonostante gli agenti siano distinti, questi si possono vedere come indistinguibili purché si consideri la distribuzione media (30).

Tale risultato è anche confermato a livello pratico nell'algoritmo 1, illustrato nel paragrafo a venire, ove una matrice d'adiacenza unitaria porta ad avere un algoritmo del tutto analogo a quello classico; pertanto l'unica distribuzione che può calcolare 1 è proprio quella media F.

3.4.3 Analisi della IS2

La previa discussione suggerisce di studiare anche il caso in cui gli agenti siano effettivamente indistinguibili; tuttavia, prima di affrontarlo assieme alla prima ipotesi risulta interessante analizzare tale ipotesi isolatamente. Pertanto la (15) diventa

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{P}} \varphi f ds &= \frac{1}{2N} \sum_{j \in \mathcal{J}} A(i, j) \int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle f f^* ds ds_* \\ &\quad + \frac{1}{2N} \sum_{j \in \mathcal{J}} A(j, i) \int_{\mathcal{P}^2} \langle \varphi'_* - \varphi_* \rangle f f^* ds ds_*, \end{aligned}$$

che sommata su tutti gli indici porta a

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi f ds = \frac{1}{2N^2} \sum_{i, j \in \mathcal{J}} A(i, j) \int_{\mathcal{P}^2} \langle \varphi' - \varphi + \varphi'_* - \varphi_* \rangle f f^* ds ds_*,$$

e definendo $L \equiv \sum_{i, j \in \mathcal{J}} A(i, j)$ si arriva a

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi f ds = \frac{L}{N^2} \left[\frac{1}{2} \sum_{i, j \in \mathcal{J}} A(i, j) \int_{\mathcal{P}^2} \langle \varphi' - \varphi + \varphi'_* - \varphi_* \rangle f f^* ds ds_* \right]. \quad (32)$$

In questo contesto il rapporto $L/N^2 \in [0,1]$ rappresenta topologicamente simile è la rete a una completamente connessa⁴; d'altra parte l'equazione è analoga a quella classica di Boltzmann **(S)**.

Dunque l'indistinguibilità degli agenti ha una notevole conseguenza sulla (15), riassumendo l'effetto complessivo del grafo al solo coefficiente L/N^2 che quindi ne rappresenta gli ultimi bagliori prima di una sua totale scomparsa per la IS1.

3.4.4 Analisi della IS1, IS2 e IS3

Visto che vale IS1 si può partire dalla (31) nella quale la distribuzione media (30) diventa per IS2

$$F(s,t) = \frac{1}{N} \sum_{i \in \mathcal{I}} f_i(s,t) \stackrel{2^\circ}{=} \frac{1}{N} \sum_{i \in \mathcal{I}} f(s,t) = f(s,t),$$

ossia la F coincide con quella di tutti gli agenti⁵, essendo questi, appunto, indistinguibili.

In tal modo la (31) diventa

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi f ds = \frac{1}{2} \int_{\mathcal{P}^2} \langle \varphi' - \varphi + \varphi'_* - \varphi_* \rangle f f^* ds ds_*,$$

che unita all'IS3 porta all'equivalenza (mediante il cambio di variabili $s_* = s$ e $s = s_*$)

$$\int_{\mathcal{P}^2} \langle \varphi'_* - \varphi_* \rangle f f_* ds ds_* = \int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle f f^* ds ds_*,$$

e quindi a

$$\frac{d}{dt} \int_{\mathcal{P}} \varphi f ds = \int_{\mathcal{P}^2} \langle \varphi' - \varphi \rangle f f^* ds ds_*,$$

che equivale alla formula classica di Boltzmann con interazioni simmetriche **(S)** usata classicamente per modellizzare la distribuzione dell'energia cinetica tra una popolazione di particelle di un gas.

⁴ Difatti L è interpretabile come il numero di lati presenti in un grafo diretto e che ha come limite superiore proprio N^2 , ossia il numero totale di coppie [e quindi lati] dati N nodi.

⁵ Si noti anche che la perdita della dipendenza della distribuzione media dal numero di nodi N è coerente colla situazione in cui $N \rightarrow \infty$, condizione fondamentale analoga a casi classici come lo studio del gas nel quale $N \gg 1$ ben approssima il limite.



Algorithm 1 Algoritmo di Monte Carlo per equazioni di tipo su un grafo

Require: adjacency matrix \mathbf{M} ; initial state $V_0 \in \mathcal{O}^N$; time step $\Delta t > 0$; final time $T > 0$

```
1:  $\tilde{V} \leftarrow V_0$ 
2:  $t \leftarrow 1$ 
3: for  $t < T$  do
4:    $\langle \varphi \rangle(t) \leftarrow \frac{1}{N} \sum_{i=1}^N \varphi(\tilde{V}(i))$ 
5:    $V \leftarrow \tilde{V}$ 
6:    $P \leftarrow$  random permutation of  $\{1, \dots, N\}$ 
7:    $p_1 \leftarrow (P(1), \dots, P(N/2))$ 
8:    $p_2 \leftarrow (P(N/2+1), \dots, P(N))$ 
9:    $i \leftarrow 1$ 
10:  for  $i < N/2$  do
11:     $\Theta \sim \text{Bernoulli}(B(p_1(i), p_2(i))\Delta t)$ 
12:     $\tilde{V}(p_1(i)) \leftarrow V(p_1(i))(1-\Theta) + \Psi(V(p_1(i)), V(p_2(i)))\Theta$ 
13:     $\tilde{V}(p_2(i)) \leftarrow V(p_2(i))(1-\Theta) + \Psi_*(V(p_2(i)), V(p_1(i)))\Theta$ 
14:     $i \leftarrow i+1$ 
15:  end for
16:   $t \leftarrow t + \Delta t$ 
17: end for
```

4

SIMULAZIONI

4.1 METODO MONTE CARLO



504

5

CONCLUSIONI

505

Maggiori sviluppi teorici, specialmente per ricavare un'equazione di Fokker-Planck per la distribuzione stazionaria.

506

507

Maggiori sviluppi pratici per abbandonare l'ipotesi semplificativa della rete statica.

508

509

Applicare questa teoria anche a reti europee o internazionali.



510 APPENDICE

511 A CODICE



512

ELENCO DELLE FIGURE

513	Figura 1.1	Funzione Lognormale(0,1).	6
514	Figura 2.1	Esempio di un grafo indiretto, della sua equivalen-	
515		te forma diretta [simmetrica] e della loro [identica]	
516		matrice d'adiacenza.	9
517	Figura 2.2	Forza contro grado $s(k)$ per la Sardegna.	12
518	Figura 2.3	Riproduzione dei principali grafici di [5] [a cui si ri-	
519		manda per maggiori dettagli sui vari coefficienti] re-	
520		lativi alla topologia della rete del pendolarismo sarda.	14



521

ELENCO DELLE TABELLE

522

Tabella 2.1 Formattazione di una sola riga nei dati ISTAT¹; le
523 linee barrate corrispondono a dati trascurati. 11

524

Tabella 2.2 Confronto con [5] dei pesi maggiori. 13



BIBLIOGRAFIA

- [1] Auerbach, Felix. «Das Gesetz der Bevölkerungskonzentration [The law of population concentration]». In: *Petermanns Geographische Mitteilungen* 59 (1913), pp. 74–76.
- [2] Barabási, Albert-László and Albert, Réka and Jeong, Hawoong. «Mean-field theory for scale-free random networks». In: *Physica A: Statistical Mechanics and its Applications* 272.1-2 (1999), pp. 173–187.
- [3] Berliant, Marcus and Watanabe, Axel H. «A scale-free transportation network explains the city-size distribution». In: *Quantitative Economics* 9.3 (2018), pp. 1419–1451.
- [4] Broido, Anna D and Clauset, Aaron. «Scale-free networks are rare». In: *Nature communications* 10.1 (2019), p. 1017.
- [5] De Montis, Andrea and Barthélemy, Marc and Chessa, Alessandro and Vespignani, Alessandro. «The structure of interurban traffic: a weighted network analysis». In: *Environment and Planning B: Planning and Design* 34.5 (2007), pp. 905–924.
- [6] Eeckhout, Jan. «Gibrat’s law for (all) cities». In: *American Economic Review* 94.5 (2004), pp. 1429–1451.
- [7] Gualandi, Stefano and Toscani, Giuseppe. «Human behavior and lognormal distribution. A kinetic description». In: *Mathematical Models and Methods in Applied Sciences* 29.04 (2019), pp. 717–753.
- [8] Gualandi, Stefano and Toscani, Giuseppe. «Size distribution of cities: A kinetic explanation». In: *Physica A: Statistical Mechanics and its Applications* 524 (2019), pp. 221–234.
- [9] ISTAT. *Matrici di contiguità, distanza e pendolarismo*. 2023. URL: <https://www.istat.it/notizia/matrici-di-contiguita-distanza-e-pendolarismo/> (visitato il giorno 12/01/2026).
- [10] Loy, Nadia and Tosin, Andrea. «A viral load-based model for epidemic spread on spatial networks». In: *arXiv preprint arXiv:2104.12107* (2021).
- [11] Marc Barthélemy. «Spatial networks». In: *Physics Reports* 499.1 (2011), pp. 1–101. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2010.11.002>. URL: <https://www.sciencedirect.com/science/article/pii/S037015731000308X>.
- [12] Nurisso, Marco and Raviola, Matteo and Tosin, Andrea. «Network-based kinetic models: Emergence of a statistical description of the graph topology». In: *European Journal of Applied Mathematics* (2024), pp. 1–22. DOI: [10.1017/S0956792524000020](https://doi.org/10.1017/S0956792524000020).
- [13] Pareschi, Lorenzo and Toscani, Giuseppe. *Interacting multiagent systems: kinetic equations and Monte Carlo methods*. OUP Oxford, 2013.
- [14] Zipf, George Kingsley. *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio books, 2016.