



UNIVERSITÀ DI PISA

Dipartimento di Matematica
Corso di Laurea Triennale in Matematica

Tesi di Laurea

Il Modello Epidemiologico SIR sulle Reti Complesse

Relatore:
Dott. Fabio Durastante

Candidato:
Simmaco Di Lillo

Anno Accademico 2020/2021

Introduzione

Mi auguro solo che in una questione che riguarda così da vicino il bene dell'umanità, si decida con la piena consapevolezza che un po' di analisi e di calcolo possono fornire.

D. Bernoulli

Il tema della dinamica delle malattie infettive è stato ed è uno delle più ricche aree di applicazione della matematica in ambito biologico. Le epidemie hanno da sempre afflitto il genere umano. Da Tucidide e Lucrezio che ci hanno raccontato, con molta tragicità, dell'epidemia che nel 430 a.C. colpì la città di Atene; alla peste, chiamata morte nera, che terrorizzò l'Europa dal 1347 al 1352 causando 20 milioni di vittime [14] e la diminuzione di un terzo della popolazione del continente [1]. In epoche più recenti il colera, l'influenza spagnola, l'influenza suina e molte altre lungo i secoli fino ad arrivare ai giorni nostri con la pandemia di COVID-19.

La formulazione di modelli matematici per lo studio di questi fenomeni è utile a chiarire le ipotesi, le variabili e i parametri di interesse con il fine di valutare i possibili effetti della malattia sulla popolazione. Tali analisi possono indirizzare le decisioni politiche sulle strategie più opportune per affrontare, gestire e, nella migliore delle ipotesi, prevenire gli esiti di maggior impatto.

Storicamente, la prima trattazione matematica in ambito epidemiologico è attribuita a Daniel Bernoulli (1700-1782), il quale, servendosi di un'argomentazione teorica [3]¹ basata sulla crescita esponenziale, mostrò i vantaggi possibili ottenibili da una vaccinazione preventiva per il contenimento dell'epidemia di vaiolo che affliggeva l'Europa al volgere del diciottesimo secolo. In Europa il vaiolo era la prima causa di morte con 400.000 decessi annui [11] e aveva colpito pesantemente anche la famiglia del re Luigi XIV.

¹Traduzione e commento dell'articolo originale di D. Bernoulli.

Il paradigma di riferimento per la descrizione matematica delle epidemie è basato sull'utilizzo dei cosiddetti modelli compartmentali, ed è un'evoluzione del modello proposto nel 1927 dagli scienziati scozzesi William O. Kermack e Anderson G. McKendrick [6]. Per questo, nel primo capitolo, ci concentreremo sull'analisi di questo modello e del meccanismo con cui riesce a spiegare la rapida crescita e successiva decrescita del numero di infetti osservabili in alcune epidemie.

Per affrontare con maggior dettaglio lo sviluppo di una malattia è utile arricchire il modello di Kermack-McKendrick con delle informazioni relative alle modalità di trasmissione della malattia. Nel secondo capitolo ci concentreremo sull'estensione di questo modello alle *complex network*. Questo permette di aggiungere al modello una rappresentazione matematica della “rete dei contatti” tra i vettori del contagio. In particolare, ci focalizzeremo sul lavoro svolto da Kiss [7].

Nel terzo capitolo daremo infine una breve descrizione dei metodi per l'integrazione numerica di equazioni differenziali, ed effettueremo delle sperimentazioni numeriche su una rete reale per discutere delle difficoltà computazionali incontrate nel trattamento di questo problema su larga scala, e per validare i risultati discussi nei capitoli precedenti.

Tutti i codici utilizzati nella tesi sono scaricabili dal repository al link:

<https://github.com/simmaco99/Tesi>.

1 — Epidemiologia matematica e reti

1.1 — Il modello SIR

Il modello SIR è un modello compartmentale: la popolazione viene suddivisa in tre classi:

- S : i *suscettibili* ovvero individui che possono contrarre la malattia;
- I : gli *infetti* ovvero coloro che sono ammalati;
- R : i *rimossi* ovvero quelli tolti dalle classi precedenti perché completamente guariti (dunque immuni).

Tale modello si basa su alcune assunzioni.

- A1 Il numero della popolazione è costante nel tempo e verrà indicato con N , ovvero non si considerano nuove nascite o morti. Inoltre stiamo assumendo che N sia sufficientemente grande in modo da considerare la variazione in ogni classe continua e non discreta.
- A2 Esiste un fattore di contatto β . Tale rapporto indica il valor medio di contatti per infettivo per unità di tempo.
- A3 Gli infetti lasciano la classe I al tasso α per unità di tempo e migrano nella classe R , un individuo che entra nella classe R non uscirà da tale classe.

Da queste considerazioni segue che

$$\begin{aligned}\dot{S} &= -\beta SI, \\ \dot{I} &= \beta SI - \alpha I, \\ \dot{R} &= \alpha I.\end{aligned}\tag{1.1}$$

Poiché abbiamo assunto che $N = S + I + R$ sia costante, il sistema precedente risulta equivalente a

$$\begin{aligned}\dot{S} &= -\beta SI, \\ \dot{I} &= \beta SI - \alpha I.\end{aligned}\tag{1.2}$$

Studiamo cosa succede se introduciamo un piccolo numero di infetti in una popolazione di suscettibili, ovvero consideriamo il sistema (1.2) con le condizioni iniziali

$$\begin{aligned} I(0) &= I_0 > 0, \quad I_0 \ll N \text{ e} \\ S(0) &= S_0 = N - I_0. \end{aligned}$$

Da (1.2) osserviamo che $\dot{S} < 0$ per ogni tempo t , mentre $\dot{I} > 0$ se e solo se $\frac{\beta S}{\alpha} > 1$.

Definiamo

$$\mathcal{R}_0 = \frac{\beta S_0}{\alpha}$$

il *numero di riproduttività di base*. Esso rappresenta il numero di individui infettati all'interno di una popolazione di suscettibili.

Tale valore ci permette di descrivere la fase iniziale dell'epidemia:

- se $\mathcal{R}_0 < 1$ allora l'epidemia si estingue: sotto queste condizioni $\dot{I}(t) < 0$ per ogni tempo t ,
- se $\mathcal{R}_0 > 1$ allora I inizialmente aumenta e dunque l'epidemia ha inizio.

Studiamo ora gli equilibri del sistema autonomo di equazioni differenziali (1.2).

Sommando le due equazioni che definiscono tale sistema otteniamo

$$(S + I) = -\alpha I. \quad (1.3)$$

Ora $S + I$ è una funzione non negativa, decrescente dunque ammette un limite. Poiché la derivata di una funzione decrescente e limitata deve tendere a 0 si ha

$$I(t) \rightarrow 0.$$

Da queste due osservazioni si ha

$$S(t) \rightarrow S_\infty.$$

Integrando da 0 a $+\infty$ in (1.3) otteniamo

$$\alpha \int_0^{+\infty} I(t) dt = - \int_0^{+\infty} (S(t) + I(t)) dt = N - S_\infty.$$

In (1.2), dividendo per S e integrando da 0 a T otteniamo

$$\log \frac{S_0}{S_\infty} = \beta \int_0^{+\infty} I(t) dt = \frac{\beta}{\alpha} (N - S_\infty) = \mathcal{R}_0 \left(1 - \frac{S_\infty}{N} \right). \quad (1.4)$$

L'equazione (1.4) prende il nome di *relazione di dimensione finale* infatti fornisce una relazione tra il numero \mathcal{R}_0 e la dimensione dell'epidemia, ovvero il numero

di membri che sono stati infetti nel corso dell'epidemia: $N - S_\infty$.

Osservazione 1. Poiché il lato destro della (1.4) è finito lo è anche il lato sinistro: $S_\infty > 0$ ovvero finita l'epidemia esisteranno ancora degli individui suscettibili.

Proposizione 1.1.1. *La relazione finale (1.4) ha un'unica soluzione.*

Dimostrazione. Sia

$$g(x) = \log \frac{S_0}{x} - \mathcal{R}_0 \left(1 - \frac{x}{N}\right),$$

ora

$$\lim_{x \rightarrow 0^+} g(x) > 0 \text{ e } g(N) = \log \frac{S_0}{N} < 0,$$

mentre

$$g'(x) = -\frac{1}{x} + \frac{\mathcal{R}_0}{N} < 0 \Leftrightarrow x < \frac{N}{\mathcal{R}_0}.$$

- Se $\mathcal{R}_0 \leq 1$, allora $N < \frac{N}{\mathcal{R}_0}$ dunque g decresce da un valore positivo in 0^+ fino ad un valore negativo in N .

In questo caso $g(x)$ ammette un unico zero S_∞ con $S_\infty < N$.

- Se $\mathcal{R}_0 > 1$, allora $g(x)$ è monotona decrescente da un valore positivo in 0^+ fino al minimo in $\frac{N}{\mathcal{R}_0}$.

Ora

$$g\left(\frac{S_0}{\mathcal{R}_0}\right) = \log \mathcal{R}_0 - \mathcal{R}_0 + \frac{S_0}{N} \leq \log \mathcal{R}_0 \leq \mathcal{R}_0 + 1 < 0,$$

dove la penultima diseguaglianza deriva da $\log x < x - 1$ se $x > 0$.

Dunque $g(x)$ ha un unico zero in S_∞ con $S_\infty < \frac{N}{\mathcal{R}_0}$.

□

Andiamo ora a descrivere le orbite delle soluzioni nel piano (S, I) .
Dividendo per S l'equazione (1.2) e integrando tra 0 a t otteniamo

$$\log \frac{S_0}{S(t)} = \beta \int_0^{+\infty} I(t) dt = \frac{\beta}{\alpha} (N - S(t) - I(t)).$$

Per poter studiare il modello occorre stimare i parametri.

Il fattore β è di difficile stima: dipende dalla malattia in esame e, in particolar modo, da fattori sociali e comportamentali.

I valori di S_0 e S_∞ possono essere ricavati tramite test sierologici (misurazione della risposta immunitaria tramite analisi del sangue): da questi valori usando la (1.4) possiamo stimare \mathcal{R}_0 .

Questa stima, tuttavia, è a posteriori, cioè può essere ricavata solamente dopo che l'epidemia ha fatto il suo corso.

Un altro modo per stimare β può essere ricavato dalla seguente approssimazione

$$\dot{I} = (\beta N - \alpha) I,$$

da cui si ottiene che il numero degli infetti cresce esponenzialmente con un tasso di crescita

$$r = \beta N - \alpha = \alpha (\mathcal{R}_0 - 1),$$

ricavabile dall'incidenza della malattia all'inizio dell'epidemia.

Otteniamo così

$$\beta = \frac{r + \alpha}{N}.$$

Oltre ai parametri α e β due misure fondamentali nello studio dell'epidemia sono l'incidenza e la prevalenza. Usiamo la definizione data dal comitato di esperti dell'OMS [10].

La *prevalenza* indica il numero di individui malati in una determinata popolazione, senza nessuna distinzione tra nuovi e vecchi casi. La prevalenza può essere registrata in un determinato momento (prevalenza puntuale) o durante un determinato periodo di tempo (periodo di prevalenza). La “prevalenza puntiforme” è solitamente espressa come una frazione: il denominatore è il numero della popolazione.

L'*incidenza* è il numero di individui che si ammalano durante un dato periodo in una specifica popolazione. L'incidenza è solitamente espressa come frazione: il denominatore è il numero medio di persone nella popolazione.

Osservazione 2 (Immunizzazione). Se un gruppo di infetti viene introdotto in una popolazione, è possibile ridurre l'impatto dell'epidemia diminuendo \mathcal{R}_0 . Una strategia può essere tramite l'immunizzazione, il cui scopo è quello di trasferire membri della popolazione della classe S a quella R , così facendo viene ridotto il numero S_0 e dunque anche \mathcal{R}_0 .

Supponiamo che una frazione p della popolazione sia immunizzata: il numero dei suscettibili passa da S_0 a $S_0(1 - p)$.

Se inizialmente il numero di riproduzione di base era $\mathcal{R}_0 = \frac{\beta N}{\alpha}$, nella nuova situazione diventa $\mathcal{R}_0'' = \frac{\beta N(1-p)}{\alpha}$ da cui si ha che

$$\frac{\beta N(1-p)}{\alpha} < 1 \quad \Leftrightarrow \quad p > 1 - \frac{\alpha}{\beta N} = 1 - \frac{1}{\mathcal{R}_0}.$$

ciò ci fornisce una stima su p .

1.1.1 Un esempio

Analizziamo i dati della peste bubbonica del 1665-66 nel villaggio di Eyam in Inghilterra [12].

I membri del villaggio hanno annotato giorno per giorno il numero di decessi. Per appianare alcune significative variazioni giornaliera nel tasso di mortalità, Raggett [12] ha raccolto i dati con una cadenza di 15 giorni e mezzo a partire dal 18 Giugno del 1666 (Tabella 1.1).

Tabella 1.1: Popolazione di deceduti e rimossi.

Periodo (1666)	Deceduti	Rimossi (alla fine del periodo)
19 Giugno-3/4 Luglio	11.5	11.5
4/5 Luglio-19 Luglio	26.5	38
20 Luglio-3/4 Agosto	40.5	78.5
4/5 Agosto-19 Agosto	41.5	120
20 Agosto-3/4 Settembre	25	145
4/5 Settembre-19 Settembre	11	156
20 Settembre-4/5 Ottobre	11.5	167.5
5/6 Ottobre-20 Ottobre	10.5	178

Prendendo come periodo medio di infezione 11 giorni, possiamo stimare il numero degli infetti. Alla fine di ogni intervallo di tempo il numero di infetti è dato analizzando il diario dei decessi degli 11 giorni successivi.

Sfruttando la relazione

$$N = S(t) + I(t) + R(t),$$

otteniamo la Tabella 1.2.

Tabella 1.2: Numero di suscettibili ed infetti. Dati reali.

$S(0) = 254$ $I(0) = 7$ e $N = 261$		
Data (1666)	S	I
3/4 Luglio	235	14.5
19 Luglio	201	22
3/4 Agosto	153.5	29
19 Agosto	121	20
3/4 Settembre	108	8
19 Settembre	97	8
4/5 Ottobre	Sconosciuti	Sconosciuti
20 Ottobre	83	0

Ora sfruttando (1.4) otteniamo $\frac{\alpha}{\beta} \simeq 159$.

Simulando il modello (1.2) con condizioni iniziali $S(0) = 254$ e $I(0) = 7$ e utilizzando come parametri $\alpha = 2.73$ e $\beta = 0.0178$ troviamo i risultati in Tabella 1.3 e i grafici di S , I e R come funzioni del tempo (Figura 1.1(a)).

La Figura 1.1(b) mostra i punti reali insieme al ritratto di fase per il modello (1.2): notiamo che i dati effettivi sono notevolmente vicini alle previsioni ottenute.

Tabella 1.3: Numero di suscettibili ed infetti. Dati sperimentali.

$$S(0) = 254 \quad I(0) = 7 \quad N = 261$$

Data (1666)	S	I
3/4 Luglio	230	15
19 Luglio	190	26
3/4 Agosto	147	30
19 Agosto	115	24
3/4 Settembre	96	15
19 Settembre	86	9
4/5 Ottobre	81	4
20 Ottobre	78	2

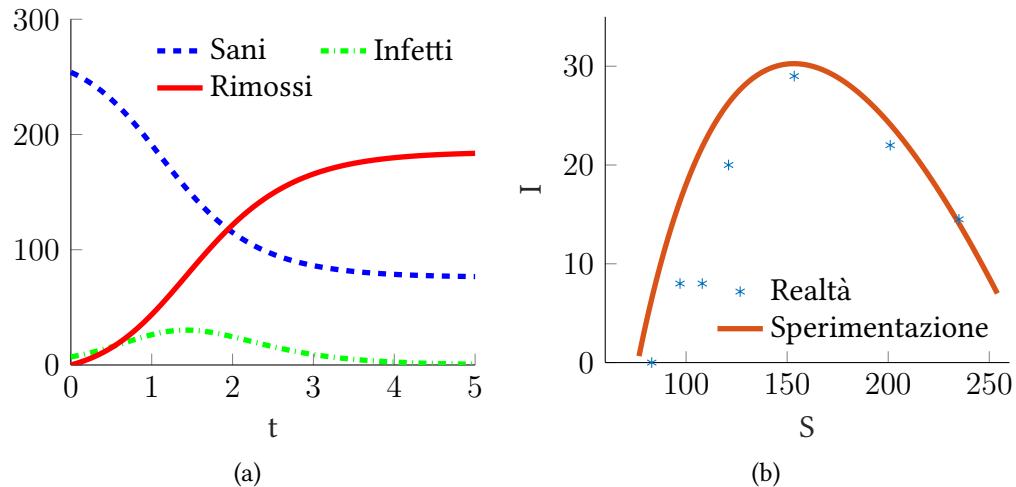


Figura 1.1: Popolazione delle singole classi in funzioni del tempo (a) e orbite nel piano (S, I) (b). Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, il problema di Cauchy (1.2) con condizioni iniziali $S(0) = 254$, $I(0) = 7$. Per quanto riguarda i parametri, abbiamo considerato $\alpha = 2.73$ e $\beta = 0.0178$.

1.2 — Reti complesse e grafi

Abbiamo trattato fino ad ora l'epidemia considerando la popolazione aggregata codificando l'intero meccanismo di contatto nel parametro β . Per avere un controllo maggiore dobbiamo estendere il modello affinché possa trattare direttamente le interazioni tra i vettori dell'epidemia.

Molti sistemi che si basano su interazioni tra soggetti sono rappresentabili e visualizzabili, in maniera naturale, come *network* dove i nodi rappresentano gli individui e gli archi le interazioni.

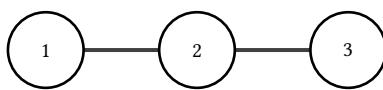
In questa sezione andremo a definire formalmente un grafo e alcune nozioni rilevanti.

Un *grafo* è una coppia $G = (V, E)$ dove:

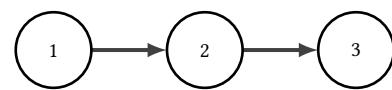
- V è detto insieme dei *vertici* o nodi;
- $E \subseteq V \times V$ è l'insieme degli *archi* o link.

Diremo che un grafo è *non orientato* se ogni arco appare in entrambe le direzioni, altrimenti diremo che è un grafo *diretto* o *orientato*.

Per visualizzare un grafo si può usare una rappresentazione come in Figura 1.2: cerchi, i vertici, connessi da frecce, nel caso di grafi diretti, oppure da segmenti, per grafi non orientati.



(a) Grafo non orientato



(b) Grafo orientato

Figura 1.2: Due esempi di grafi.

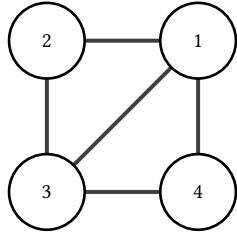
È possibile codificare un grafo, con un numero finito di nodi, attraverso una matrice detta matrice di adiacenza.

Supponiamo $|V| = N$ dunque esiste una biiezione $\phi : \{1, \dots, N\} \rightarrow E$.

La *matrice di adiacenza* di G è la matrice $H_G = (g_{ij})$ dove

$$g_{ij} = \begin{cases} 1 & \text{se } (\phi(i), \phi(j)) \in E \\ 0 & \text{altrimenti} \end{cases}.$$

Un esempio di matrice di adiacenza è riportata nella Figura 1.3.



$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Figura 1.3: Grafo e matrice di adiacenza del toast.

Un grafo G si dice *pesato* se ad ogni suo arco è associato un'etichetta. Formalmente, $G = (V, E)$ si dice grafo pesato se esiste una funzione $p : E \rightarrow T$ dove T è un insieme. Generalmente $T = \mathbb{R}$.

Dato un grafo $G = (V, E)$, un cammino P che connette i vertici $u, v \in V$ è una sequenza di vertici

$$P = (v_0, \dots, v_n)$$

dove $\forall i = 0, \dots, n - 1$ si ha $(v_i, v_{i+1}) \in E$, $v_0 = u$ e $v_n = v$.

Un grafo si dice *fortemente connesso* se per ogni coppia di nodi (u, v) esiste un cammino che li congiunge.

Definiamo il *diametro* di un grafo come la distanza massima tra due nodi, la distanza è il numero di archi di un cammino.

Poiché nello studio della diffusione di una popolazione con molti individui, è irrealistico sperare di conoscere l'esatta matrice di adiacenza è spesso utile supporre che la propagazione della malattia sia governata solamente da alcune misure qualitative. Ci focalizzeremo quindi sulle seguenti quantità.

Definiamo il *grado uscente (out-degree)* del nodo i come il numero di archi che partono dal nodo i . In termini della matrice di adiacenza

$$k_i^{out} = \sum_{j=1}^N g_{ij}.$$

Analogamente il *grado entrante (in-degree)* è il numero di archi che arrivano al nodo:

$$k_i^{in} = \sum_{j=1}^N g_{ji}.$$

Nel caso di grafi non orientati si ha $k_i^{out} = k_i^{in}$ che denoteremo, semplicemente, con k_i .

Il *grado medio* di un grafo non orientato è la quantità

$$\langle K \rangle = \frac{1}{N} \sum_{i=1}^N k_i.$$

Sia L il numero di gradi differenti e siano $\{d_1, \dots, d_L\}$ i possibili gradi.

Sia N_ℓ il numero di vertici con grado d_ℓ allora il *grado di distribuzione* del grado ℓ è la quantità

$$p_\ell = \frac{N_\ell}{N}.$$

Se tutti i vertici hanno lo stesso grado, diremo che il grafo è *regolare*.

2 — Il modello SIR *bottom-up* su una rete

In questo capitolo ci occuperemo di descrivere il modello SIR su di una rete sfruttando il modello scalare ed i concetti introdotti nella Sezione 1.2. Costruiremo un modello Markoviano SIR su una rete di N individui. Ogni individuo è in ogni istante temporale suscettibile, infetto o rimosso con una certa probabilità p . Come nel modello scalare del capitolo precedente, un individuo infetto può trasmettere il suo stato con la differenza che questo processo avviene unicamente per i nodi che sono suoi vicini nella rete.

In particolare, ci occuperemo di derivare le equazioni per lo stato di ogni singolo nodo della rete ottenendo un modello “*bottom-up*” esatto dell’epidemia. Successivamente cercheremo di ridurre il numero di equazioni del modello con il fine di mantenere l’esattezza e semplificare, dal punto di vista computazionale, il problema da integrare.

2.1 — Un primo esempio

Vediamo come è possibile applicare il modello *SIR* ad un semplice grafo con 3 nodi (Figura 2.1).

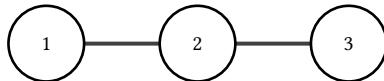


Figura 2.1: Grafo con 3 nodi.

Fissiamo la notazione:

- $\langle S_i \rangle(t)$ denota la probabilità che il nodo i sia suscettibile al tempo t ;
- $\langle I_i \rangle(t)$ denota la probabilità che il nodo i sia infetto al tempo t ;
- $\langle R_i \rangle(t)$ denota la probabilità che il nodo i sia rimosso al tempo t ;
- $(\langle I_i \rangle + \langle R_i \rangle)(t)$ denota la probabilità che il nodo i sia infetto o rimosso al tempo t ;

- $\langle S_i I_j \rangle(t)$ denota la probabilità che il nodo i sia sano e il nodo j sia infetto al tempo t .

In analogia a quanto fatto in 1.1 si ha che:

- ogni nodo i può diventare infetto se è suscettibile e almeno uno dei suoi vicini è infetto. Per ogni suo vicino infetto i diventa infetto con un tasso di τ ;
- una volta infettato, ogni nodo diventa rimosso con un tasso di γ indipendentemente dallo stato di qualsiasi altro nodo.

Passiamo ora allo studio del grafo 2.1.

- Poiché il nodo 1 ha come vicino solamente il nodo 2 la probabilità che possa infettarsi è $\langle S_1 I_2 \rangle$. $\langle S_1 \rangle$ diminuisce con un tasso di $\tau \langle S_1 I_2 \rangle$.
- Il nodo 2 ha due possibili fonti d'infezione: i nodi 1 e 2. Il suo tasso d'infezione è, dunque, $\tau (\langle S_2 I_1 \rangle + \langle S_2 I_3 \rangle)$.
- Il tasso d'infezione relativo al nodo 3 è, in analogia a quanto visto per il nodo 1, $\tau \langle S_3 I_2 \rangle$.

Mettendo insieme questi punti e ricordando il tasso di rimozione otteniamo il seguente sistema di equazioni differenziali:

$$\begin{aligned} \dot{\langle S_1 \rangle} &= -\tau \langle S_1 I_2 \rangle, & \dot{\langle I_1 \rangle} &= \tau \langle S_1 I_2 \rangle - \gamma \langle I_1 \rangle, \\ \dot{\langle S_2 \rangle} &= -\tau (\langle I_1 S_2 \rangle + \langle S_2 I_3 \rangle), & \dot{\langle I_2 \rangle} &= \tau (\langle I_1 S_2 \rangle + \langle S_2 I_3 \rangle) - \gamma \langle I_2 \rangle, \\ \dot{\langle S_3 \rangle} &= -\tau \langle I_2 S_3 \rangle, & \dot{\langle I_3 \rangle} &= \tau \langle I_2 S_3 \rangle - \gamma \langle I_3 \rangle, \\ \\ \dot{\langle R_1 \rangle} &= \gamma \langle I_1 \rangle, \\ \dot{\langle R_2 \rangle} &= \gamma \langle I_2 \rangle, \\ \dot{\langle R_3 \rangle} &= \gamma \langle I_3 \rangle. \end{aligned}$$

Analizzando attentamente il sistema possiamo notare che le equazioni per $\langle R_i \rangle$ possono essere rimosse infatti

$$\forall i \in \{1, 2, 3\} \quad \forall t \geq 0 \quad \langle R_i \rangle(t) = 1 - (\langle S_i \rangle + \langle I_i \rangle)(t).$$

Osserviamo, inoltre, che il sistema precedente non è chiuso: abbiamo altre quantità (ad esempio: $\langle S_1 I_2 \rangle$). Dobbiamo conoscere come esse evolvono nel tempo: servono

altre equazioni.

Usando argomenti simili a quanto fatto per i singoli nodi otteniamo il sistema

$$\begin{aligned}\dot{\langle S_1 I_2 \rangle} &= -\tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle S_1 I_2 \rangle, \\ \dot{\langle I_1 S_2 \rangle} &= -\tau \langle I_1 S_2 I_3 \rangle - (\tau + \gamma) \langle I_1 S_2 \rangle, \\ \dot{\langle S_2 I_3 \rangle} &= -\tau \langle I_1 S_2 I_3 \rangle - (\tau + \gamma) \langle S_2 I_3 \rangle, \\ \dot{\langle I_2 S_3 \rangle} &= -\tau \langle I_1 S_2 S_3 \rangle - (\tau + \gamma) \langle I_2 S_3 \rangle.\end{aligned}$$

Anche tale sistema non è chiuso: richiede ulteriori informazioni sulle triple.

Aggiungendo le equazioni necessarie otteniamo il sistema

$$\begin{aligned}\dot{\langle S_1 \rangle} &= -\tau \langle S_1 I_2 \rangle, \\ \dot{\langle I_1 \rangle} &= \tau \langle S_1 I_2 \rangle - \gamma \langle I_1 \rangle, \\ \dot{\langle S_2 \rangle} &= -\tau (\langle I_1 S_2 \rangle + \langle I_3 S_2 \rangle), \\ \dot{\langle I_2 \rangle} &= \tau (\langle I_1 S_2 \rangle + \langle I_3 S_2 \rangle) - \gamma \langle I_2 \rangle, \\ \dot{\langle S_3 \rangle} &= -\tau \langle I_2 S_3 \rangle, \\ \dot{\langle I_3 \rangle} &= \tau \langle I_2 S_3 \rangle - \gamma \langle I_3 \rangle, \\ \dot{\langle S_1 I_2 \rangle} &= \tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle S_1 I_2 \rangle, \\ \dot{\langle I_1 S_2 \rangle} &= -\tau \langle I_1 S_2 I_3 \rangle - (\tau + \gamma) \langle I_1 S_2 \rangle, \\ \dot{\langle I_2 S_3 \rangle} &= \tau \langle I_1 S_2 I_3 \rangle - (\tau + \gamma) \langle I_1 S_2 \rangle, \\ \dot{\langle S_2 I_3 \rangle} &= -\tau \langle I_1 S_2 I_3 \rangle - (\tau + \gamma) \langle I_1 S_2 \rangle, \\ \dot{\langle S_1 S_2 I_3 \rangle} &= -(\tau + \gamma) \langle S_1 S_2 I_3 \rangle, \\ \dot{\langle I_1 S_2 I_3 \rangle} &= -(2\tau + 2\gamma) \langle I_1 S_2 I_3 \rangle, \\ \dot{\langle I_1 S_2 S_3 \rangle} &= -(\tau + \gamma) \langle I_1 S_2 S_3 \rangle.\end{aligned}\tag{2.1}$$

Nella Figura 2.2 sono riportati i grafici dello stato dei nodi in funzione del tempo (a), (b), (c) e del grafico di prevalenza (d).

Per la sperimentazione, le condizioni iniziali sono state prese a partire da uno stato puro (1 infetto e 2, 3 sani):

$$\begin{aligned}\langle S_1 \rangle(0) &= 0, \\ \langle I_1 \rangle(0) &= 1, \\ \langle S_2 \rangle(0) &= 1, \\ \langle I_2 \rangle(0) &= 0, \\ \langle S_3 \rangle(0) &= 1, \\ \langle I_3 \rangle(0) &= 0.\end{aligned}\tag{2.2}$$

Abbiamo, inoltre, assunto l'indipendenza statistica delle condizioni iniziali (vedi [13]) ovvero

$$\langle A_i B_j \rangle(0) = \langle A_i \rangle(0) \langle B_j \rangle(0)$$

per $A, B \in \{S, I\}$ e $i, j = 1, 2, 3$.

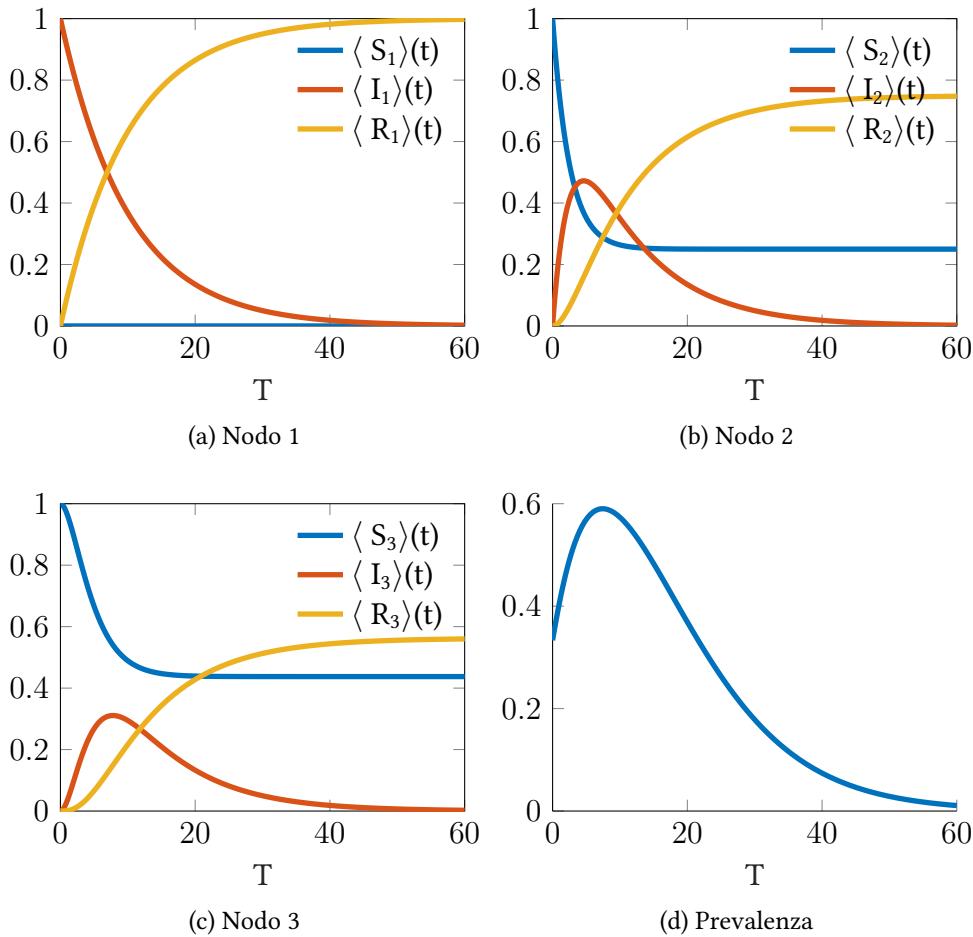


Figura 2.2: Divisione in classi nei singoli nodi (a)(b)(c) e grafico della prevalenza (d) per il grafo 2.1. Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, il problema di Cauchy (2.1) con condizioni iniziali di stati puri (2.2). Abbiamo, inoltre, supposto l'indipendenza statistica delle condizioni iniziali. Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

Poiché $\langle S_1 \rangle(0) = 0$ si ha che $\langle S_1 \rangle(t) = 0$ è soluzione del problema di Cauchy. Osserviamo, inoltre, che $\langle I_2 \rangle(t) \geq \langle I_3 \rangle(t)$ infatti il nodo 2 ha due strade per

infettarsi (di cui una inizialmente chiusa $\langle I_3 \rangle(0) = 0$) mentre il nodo 3 ne ha una sola (che inizialmente è chiusa).

2.2 — Chiusure

Come già osservato nell'esempio introduttivo, il modello bottom-up su una rete richiede un elevato numero di equazioni differenziali: le equazioni per i nodi dipendono dalle coppie, le coppie dalle triple e così via.

Tale approccio, al crescere del numero di nodi, risulterebbe computazionalmente intrattabile. Per risolvere questo problema dobbiamo trovare alcune semplificazioni che ci permettano di esprimere le coppie in termini dei singoli nodi, le triple in termini delle coppie e dei nodi e così via.

Se riusciamo a fare ciò, possiamo rompere la “cascata” nella quale ogni struttura dipende da tutte le strutture di ordine superiore.

Andiamo a presentare un approccio formale basato sul lavoro di Keeling [5] e van Baalen [15]. A tal fine introduciamo i *coefficienti di correlazione*.

Siano $A, B \in \{S, I, R\}$ e $(i, j) \in E$ allora

$$C_{A_i B_j} = \frac{\langle A_i B_j \rangle}{\langle A_i \rangle \langle B_j \rangle}.$$

Tali coefficienti quantificano la propensione che due nodi adiacenti abbiano stato differente o identico.

Se $C_{A_i B_j} = 1$ allora possiamo, in modo equivalente, assumere A_i e B_j indipendenti.

Osserviamo che nel modello *SIR* gli stati non sono però indipendenti. I nodi infetti possono infettare i loro vicini, dunque, hanno una maggior probabilità di essere uniti ad altri nodi infetti. $C_{I_i I_j} \geq 1$. Diremo che I_i e I_j sono correlati positivamente.

Con medesime argomentazioni si arriva a dire che S_i e I_j sono correlati negativamente: $C_{S_i I_j} \leq 1$.

In un certo senso, sapere che j è infetto aumenta la nostra aspettativa che i sia infetto e diminuisce quella che sia sano.

2.2.1 Chiusura al livello delle coppie

Da quanto osservato nella parte introduttiva sulle chiusure, possiamo scrivere:

$$\langle A_i B_j \rangle = \langle A_i \rangle \langle B_j \rangle C_{A_i B_j} \text{ dove } A, B \in \{S, I, R\} \text{ e } (i, j) \in E.$$

Assumendo, in prima approssimazione, l'indipendenza a livello delle coppie abbiamo

$$\langle A_i B_j \rangle \approx \langle A_i \rangle \langle B_j \rangle.$$

Per enfatizzare che, in generale, non abbiamo identità esatte, quando andremo a risolvere un sistema ottenuto usando le chiusure denoteremo con $\langle X_i \rangle$ l'approssimazione di $\langle S_i \rangle$ e con $\langle Y_i \rangle$ quella di $\langle I_i \rangle$.

Presentiamo il modello bottom-up generale su una rete con N nodi senza loop e mostriamo come usando l'indipendenza a livello delle coppie si riesca ad ottenere un sistema di equazioni differenziali chiuse.

Sia $G = (g_{ij})$ la matrice di adiacenza del grafo G . Assumiamo che il tasso di trasmissione da i a j sia τg_{ij} e che il tasso di rimozione per il nodo i sia γ_i indipendentemente dallo stato di ogni altro nodo.

Dunque, le equazioni, per il sistema diventano

$$\begin{aligned}\dot{\langle S_i \rangle} &= -\tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle S_i I_j \rangle, \\ \dot{\langle I_i \rangle} &= \tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle S_i I_j \rangle - \gamma_i \langle I_i \rangle,\end{aligned}\tag{2.3}$$

dove ricordiamo che $\langle R_i \rangle = 1 - \langle S_i \rangle - \langle I_i \rangle$.

Usando l'indipendenza a livello delle coppie otteniamo un sistema chiuso:

$$\begin{aligned}\dot{\langle X_i \rangle} &= -\tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle X_i \rangle \langle Y_j \rangle, \\ \dot{\langle Y_i \rangle} &= \tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle X_i \rangle \langle Y_j \rangle - \gamma_i \langle Y_i \rangle.\end{aligned}\tag{2.4}$$

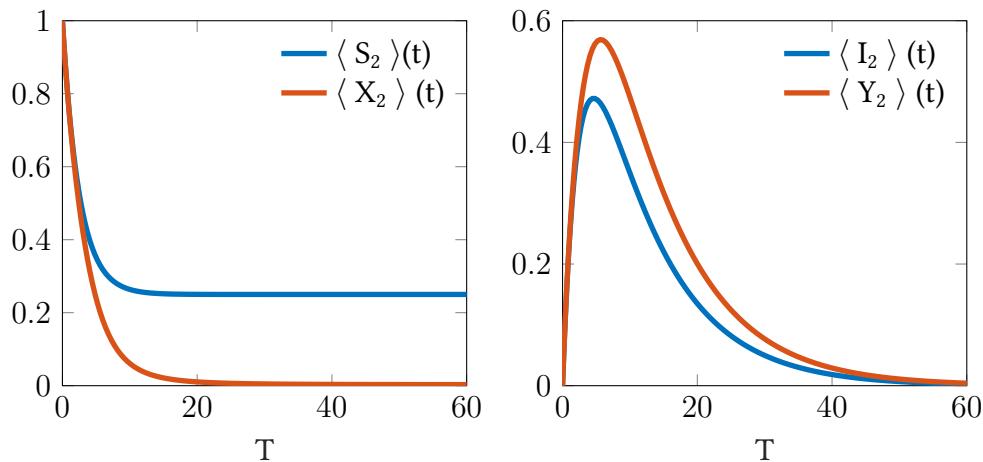
Possiamo scrivere il sistema precedente in forma vettoriale:

$$\begin{aligned}\dot{\langle X \rangle} &= -\tau G X * Y, \\ \dot{\langle Y \rangle} &= \tau G X * Y - \Gamma Y,\end{aligned}\tag{2.5}$$

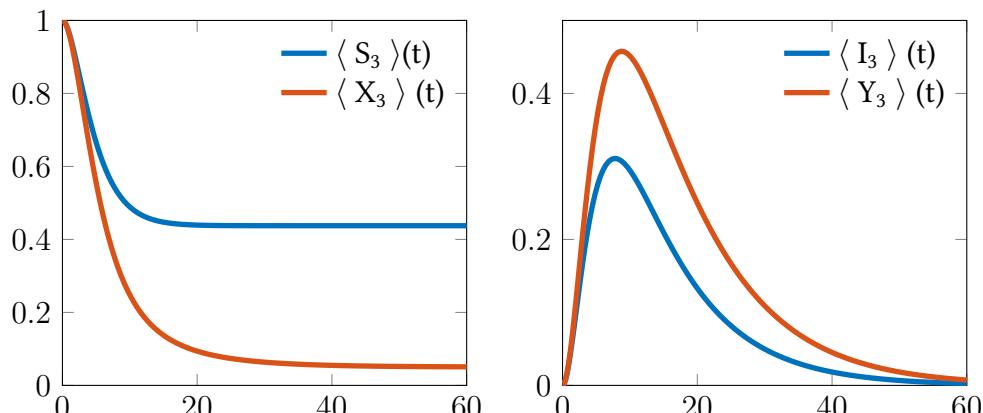
dove $\Gamma = \text{diag}(\gamma_1, \dots, \gamma_n)$ e il prodotto tra vettori è inteso elemento per elemento.

Nella Figura 2.3 possiamo confrontare le soluzioni del modello esatto con l'approssimazione ottenuta dall'indipendenza a livello delle coppie per il grafo 2.1. Per non appesantire i grafici abbiamo tracciato solamente $\langle S_i \rangle$ e $\langle I_i \rangle$ in funzione del tempo: $\langle R_i \rangle$ può essere ricavata. Dalla Figura 2.3 (d) possiamo notare che il

modello semplificato tende a sovrastimare la prevalenza della malattia. Tale sovrastima risulta sconveniente se volessimo usare queste predizioni per intervenire con delle politiche di contenimento. Nelle sezioni successive ci domanderemo se sia possibile trovare una rappresentazione con un numero minore di variabili che sia più accurata o, meglio ancora, esatta.



(a) Nodo 2



(b) Nodo 3

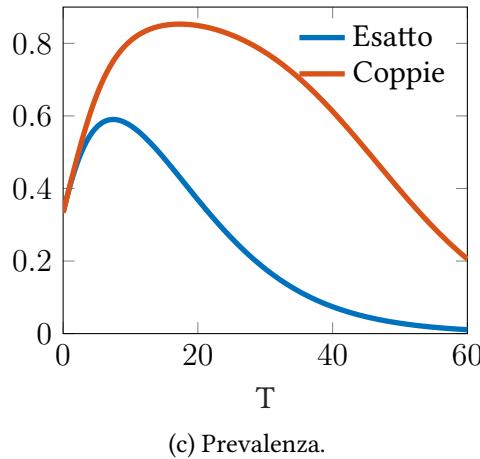


Figura 2.3: Divisione in classi dei singoli nodi (a)(b) e grafico della prevalenza (c). Non abbiamo riportato i grafici relativi al nodo 1 poiché in questo caso il modello esatto e il modello chiuso alle coppie coincidono. Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, il problema di Cauchy derivato dal modello esatto (2.1) (in blu) e quello approssimato ottenuto assumendo la chiusura alle coppie (2.4) (in rosso). Per entrambi i modelli le condizioni iniziali sono stati puri (2.2), inoltre, nel modello esatto abbiamo supposto l’indipendenza statistica di tali condizioni iniziali. Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

2.2.2 Chiusura al livello delle triple

In analogia a quanto fatto nel paragrafo precedente, cercheremo di sfruttare l’indipendenza a livello delle triple ovvero l’approssimazione

$$\langle A_i B_j C_k \rangle \approx \frac{\langle A_i B_j \rangle \langle B_j C_k \rangle}{\langle B_j \rangle}$$

per ottenere un modello chiuso (con un numero minore di equazioni rispetto a quello esatto). Tale modello ci fornirà un’approssimazione migliore di quello chiuso ottenuto assumendo l’indipendenza a livello delle coppie.

Come nel caso delle coppie, per rimarcare la non esattezza del modello useremo le lettere X, Y al posto di S e I .

Nel modello esatto, per i singoli nodi vale:

$$\begin{aligned}\langle \dot{S}_i \rangle &= -\tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle S_i I_j \rangle, \\ \langle \dot{I}_i \rangle &= -\tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle S_i I_j \rangle - \gamma \langle I_i \rangle.\end{aligned}$$

Mentre per le coppie

$$\langle \dot{S}_i S_j \rangle = \tau \sum_{\substack{k=1 \\ k \neq i}}^N g_{jk} \langle S_i S_j I_k \rangle - \tau \sum_{\substack{k=1 \\ k \neq j}}^N g_{ik} \langle I_k S_i I_j \rangle - \tau g_{ij} \langle S_i I_j \rangle - \gamma \langle S_i I_j \rangle.$$

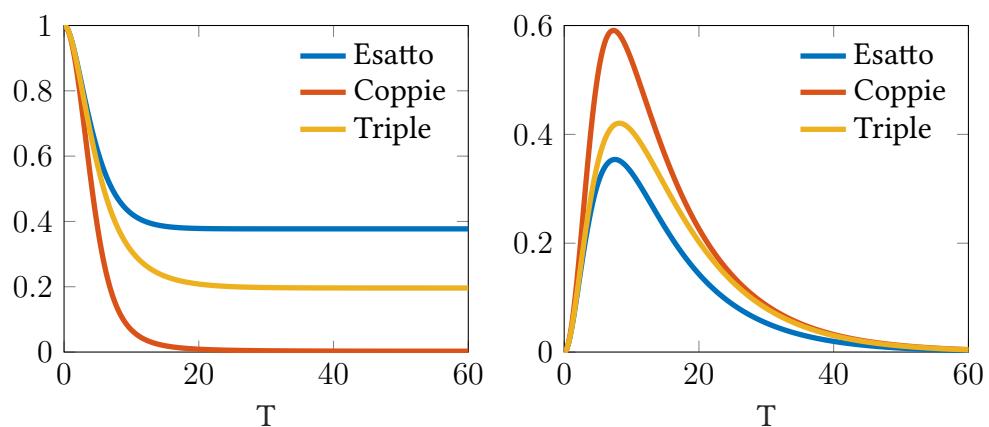
Dunque, usando la chiusura alle triple e osservando che vale

$$\langle \dot{S}_i S_j \rangle = -\tau \sum_{\substack{k=1 \\ k \neq i \\ k \neq j}}^N g_{ik} \langle I_k S_i S_j \rangle - \tau \sum_{\substack{k=1 \\ k \neq i}}^N g_{jk} \langle S_i S_j I_k \rangle,$$

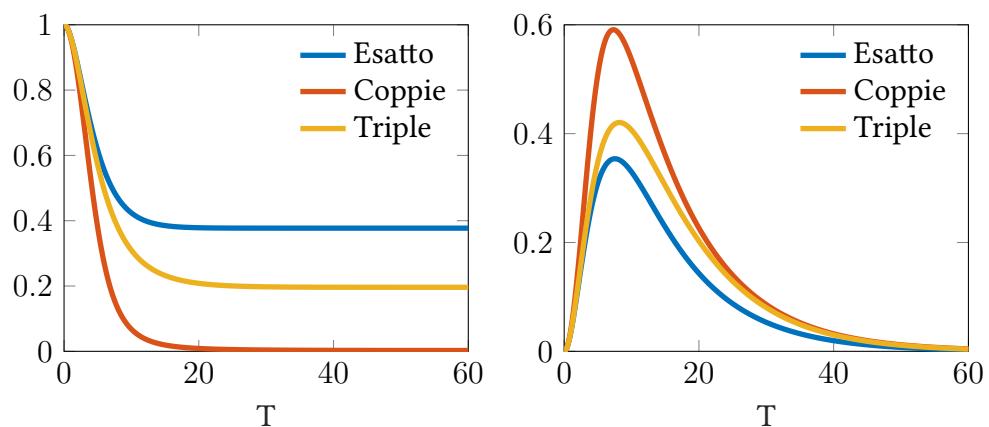
otteniamo il seguente sistema chiuso

$$\begin{aligned}\langle \dot{X}_i \rangle &= -\tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle X_i Y_j \rangle, \\ \langle \dot{Y}_i \rangle &= -\tau \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} \langle X_i Y_j \rangle - \gamma \langle I_i \rangle, \\ \langle \dot{X}_i Y_j \rangle &= -\tau \sum_{\substack{k=1 \\ k \neq i}}^N g_{jk} \frac{\langle X_i X_j \rangle \langle X_j Y_k \rangle}{\langle X_j \rangle} - \tau \sum_{\substack{k=1 \\ k \neq j}}^N g_{ik} \frac{\langle Y_k X_i \rangle \langle X_i Y_j \rangle}{\langle X_i \rangle} - \tau g_{ij} \langle X_i Y_j \rangle - \gamma \langle X_i Y_j \rangle, \\ \langle \dot{X}_i X_j \rangle &= -\tau \sum_{\substack{k=1 \\ k \neq j}}^N g_{ik} \frac{\langle Y_k X_i \rangle \langle X_i X_j \rangle}{\langle X_i \rangle} - \tau \sum_{\substack{k=1 \\ k \neq i}}^N g_{jk} \frac{\langle X_i X_j \rangle \langle X_j Y_k \rangle}{\langle X_j \rangle}.\end{aligned}\tag{2.6}$$

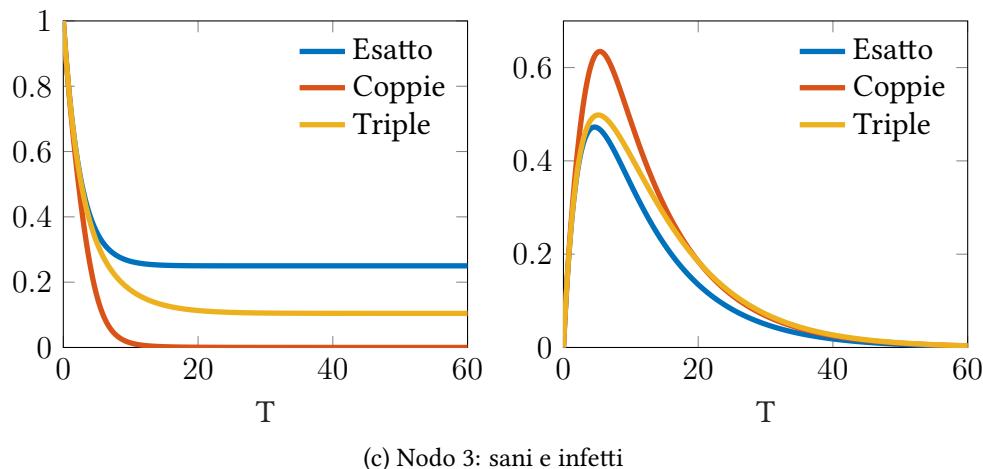
Andando a risolvere numericamente i 3 modelli osserviamo, in Figura 2.4, come il modello chiuso alle triple approssima in modo migliore, rispetto a quello chiuso alle coppie, il modello esatto.



(a) Nodo 1: sani e infetti



(b) Nodo 2: sani e infetti



(c) Nodo 3: sani e infetti

Figura 2.4: Divisione in classi dei singoli nodi per il grafo 2.9.

Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, il problema di Cauchy derivato dal modello esatto (2.3.1), dal modello chiuso alle coppie (2.4) e da quello alle triple (2.6). Per i 3 modelli le condizioni iniziali sono stati puri (2.9). Abbiamo, inoltre, supposto l'indipendenza statistica di tali condizioni iniziali. Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

Non abbiamo riportato i grafici relativi al nodo 4 poiché in questo caso i 3 modelli coincidevano (nei limiti della tolleranza della funzione `ode45`).

2.2.3 Approccio generale alla chiusura

Ci occuperemo di trovare un approccio generale alla chiusura che sfrutta le informazioni topologiche della rete [7].

Sia $G = (V, E)$ un grafo connesso.

- Diremo che $v \in V$ è un *cut-vertex* (vertice di taglio) se il grafo senza il nodo v risulta sconnesso.
- Diremo che $e \in E$ è un *ponte* se il grafo $G' = (V, E \setminus \{e\})$ risulta sconnesso.

Per un esempio di cut-vertex si guardi la Figura 2.5.

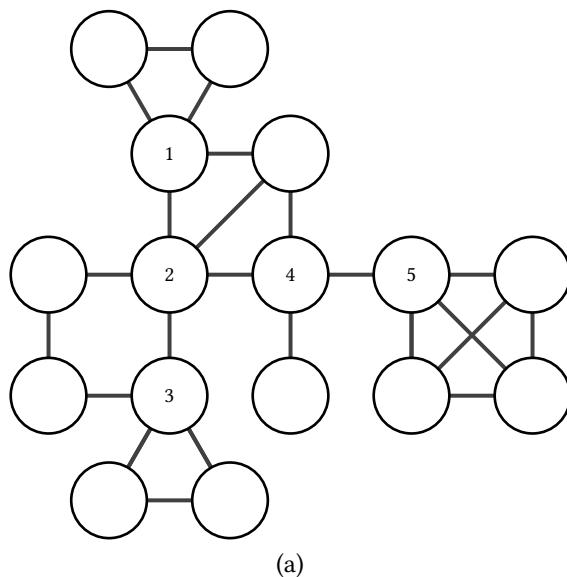
Prima di presentare il teorema cardine del capitolo ricordiamo però, la seguente definizione.

Siano A, B due eventi di uno spazio di probabilità con $\mathbb{P}(B) > 0$. Si dice *probabilità condizionale* di A dato B la quantità

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \quad (2.7)$$

Teorema 2.2.1. *Sia $G = (V, E)$ un grafo e $F = \{v_1, \dots, v_k\}$ un sottoinsieme connesso di vertici e sia v_i un suo cut-vertex. Poniamo*

$$F_1 = \{v_1, \dots, v_{i-1}\} \text{ e } F_2 = \{v_{i+1}, \dots, v_k\}.$$



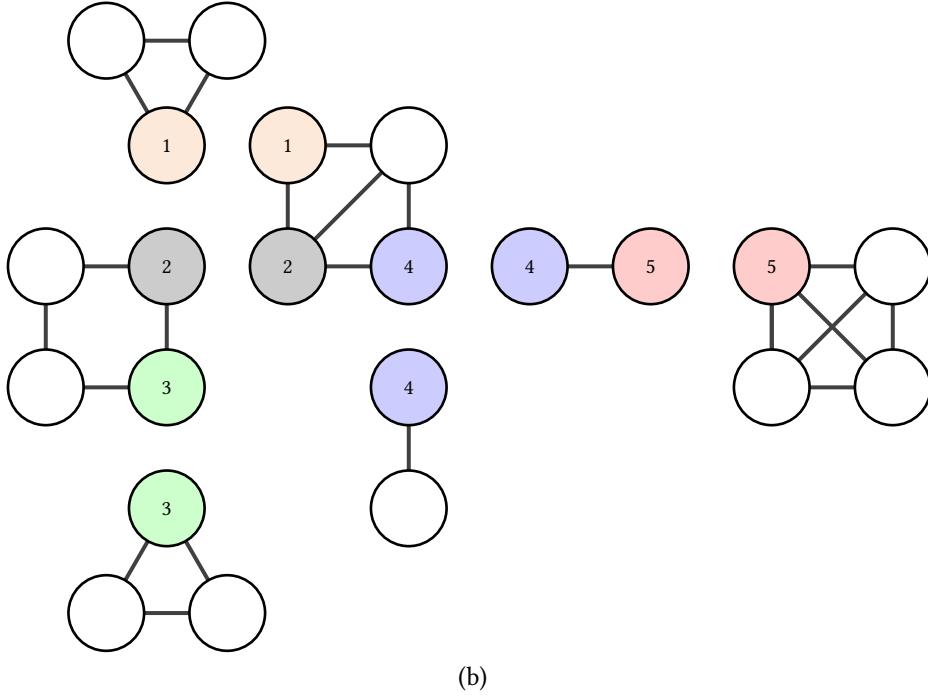


Figura 2.5: (a) Un esempio di rete con 5 cut-vertices e (b) la rete equivalente dopo la decomposizioni in sottoreti.

Se ogni cammino che connette un nodo in F_1 ad uno in F_2 passa da v_i allora:

$$\langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} \rangle = \langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} \rangle \langle S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} \rangle \quad (2.8)$$

dove $Z \in \{S, I, R\}$.

Dimostrazione. Se $\langle S_{v_i} \rangle = 0$ allora l'uguaglianza (2.8) risulta banalmente vera.

Sia $\langle S_{v_i} \rangle \neq 0$. Utilizzando la definizione di probabilità condizionale (2.7)

$$\langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} \rangle = \langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} | S_{v_i} \rangle \langle S_{v_i} \rangle.$$

Notiamo che

$$\langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} | S_{v_i} \rangle = \langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} | S_{v_i} \rangle \langle S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} | S_{v_i} \rangle,$$

infatti, ogni percorso da F_1 a F_2 deve passare attraverso v_i . Ora v_i è suscettibile dunque la trasmissione non può avvenire tra un nodo in F_1 ed uno in F_2 . Per tale motivo i due eventi presenti nel membro di sinistra sono indipendenti.

Se riapplichiamo la definizione di probabilità condizionale otteniamo

$$\langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} | S_{v_i} \rangle = \frac{\langle Z_{v_1} \dots Z_{v_{i-1}} S_{v_i} \rangle}{\langle S_{v_i} \rangle},$$

$$\langle S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} | S_{v_i} \rangle = \frac{\langle S_{v_i} Z_{v_{i+1}} \dots Z_{v_k} \rangle}{\langle S_{v_i} \rangle}$$

da cui la tesi. □

Grazie a questo teorema siamo pronti per presentare l'algoritmo generale per le chiusure.

1. Si trovano tutti i cut-vertex di G e si denotino con $C = \{v_{i_1}, \dots, v_{i_L}\}$. Tale procedimento può essere fatto con un costo in tempo di $O(|E| + |V|)$ utilizzando una visita in profondità (DFS).
2. Si divide la rete originale in sottoreti connesse a due a due scollegate. Tale procedura conduce ad una famiglia di sottoreti distinte G_1, \dots, G_P . Le sottoreti vengono create in modo che i cut-vertex siano mantenuti in tutte le sottoreti generate. Si veda la Figura 2.5.
3. Per ogni nodo i delle sottoreti, si ha

$$\begin{aligned} \dot{\langle S_i \rangle} &= -\tau \sum_j g_{ij} \langle S_i I_j \rangle, \\ \dot{\langle I_i \rangle} &= \tau \sum_j g_{ij} \langle S_i I_j \rangle - \gamma \langle I_i \rangle, \\ \dot{\langle R_i \rangle} &= 1 - \langle S_i \rangle - \langle I_i \rangle. \end{aligned}$$

Si possono trovare equazioni simili per le derivate di tutte le coppie che sorgono in queste equazioni. Queste coppie dipendono dalle triple. A loro volta le triple dipendono dalle quadruple. Si forma, così, una gerarchia di equazioni.

4. Nella gerarchia che si verrà a creare, se appare un termine composto da vertici di sottoreti diverse allora in esso è presente un cut-vertex suscettibile. Usando il Teorema (2.2.1) è possibile esprimere questo termine usando termini più semplici.

2.2.4 Un primo esempio. Continuazione

Concludiamo l'esempio, che abbiamo iniziato a inizio capitolo, applicando l'algoritmo appena presentato alla rete in Figura 2.1.

Come visto nel modello completo, per i singoli nodi si ha

$$\begin{aligned}\langle \dot{S}_1 \rangle &= -\tau \langle S_1 I_2 \rangle, & \langle \dot{I}_1 \rangle &= \tau \langle S_1 I_2 \rangle - \gamma \langle I_1 \rangle, \\ \langle \dot{S}_2 \rangle &= -\tau (\langle I_1 S_2 \rangle + \langle I_3 S_2 \rangle), & \langle \dot{I}_2 \rangle &= \tau (\langle I_1 S_2 \rangle + \langle I_3 S_2 \rangle) - \gamma \langle I_2 \rangle, \\ \langle \dot{S}_3 \rangle &= -\tau \langle I_2 S_3 \rangle, & \langle \dot{I}_3 \rangle &= \tau \langle I_2 S_3 \rangle - \gamma \langle I_3 \rangle.\end{aligned}$$

Nelle precedenti equazioni appaiono delle coppie: scriviamo le equazioni anche per loro.

$$\langle \dot{S}_1 I_2 \rangle = \tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle S_1 I_2 \rangle = \frac{\langle S_1 S_2 \rangle \langle S_2 I_3 \rangle}{\langle S_2 \rangle} - (\tau + \gamma) \langle S_1 I_2 \rangle$$

infatti dal Teorema 2.2.1 applicato al cut-vertex 2 si ha

$$\langle S_1 S_2 I_3 \rangle = \frac{\langle S_1 S_2 \rangle \langle S_2 I_3 \rangle}{\langle S_2 \rangle}.$$

In maniera analoga otteniamo

$$\begin{aligned}\langle \dot{I}_1 S_2 \rangle &= -\tau \frac{\langle I_1 S_2 \rangle \langle S_2 I_3 \rangle}{\langle S_2 \rangle} - (\tau + \gamma) \langle I_1 S_2 \rangle, \\ \langle \dot{S}_2 I_3 \rangle &= -\tau \frac{\langle I_1 S_2 \rangle \langle S_2 I_3 \rangle}{\langle S_2 \rangle} - (\tau + \gamma) \langle S_2 I_3 \rangle, \\ \langle \dot{I}_2 S_3 \rangle &= \tau \frac{\langle I_1 S_2 \rangle \langle S_2 S_3 \rangle}{\langle S_2 \rangle} - (\tau + \gamma) \langle I_2 S_3 \rangle.\end{aligned}$$

Le precedenti equazioni introducono altre due coppie. Aggiungendole otteniamo un modello chiuso:

$$\begin{aligned}\langle \dot{S}_1 S_2 \rangle &= -\tau \frac{\langle S_1 S_2 \rangle \langle S_2 I_3 \rangle}{\langle S_2 \rangle}, \\ \langle \dot{S}_2 S_3 \rangle &= -\tau \frac{\langle I_1 S_2 \rangle \langle S_2 S_3 \rangle}{\langle S_2 \rangle}.\end{aligned}$$

In questo esempio, non abbiamo ottenuto un notevole vantaggio, infatti abbiamo un sistema di 12 equazioni mentre il modello esatto (2.1) ne ha 13 equazioni. Nell'esempio, che vedremo successivamente, potremmo apprezzare meglio il risparmio computazionale.

Andando a risolvere numericamente il sistema e tracciando le soluzioni otteniamo la Figura 2.6. Possiamo notare, come previsto dalla teoria, che tali grafici

sono identici a quelli ottenuti nel modello esteso 2.2. Inoltre, se andiamo a calcolare l'errore assoluto tra le soluzioni (Figura 2.7) otteniamo che tali errori sono dell'ordine della tolleranza dell'integratore cosa che non accade confrontando il modello esatto con quello chiuso alle coppie (Figura 2.8).

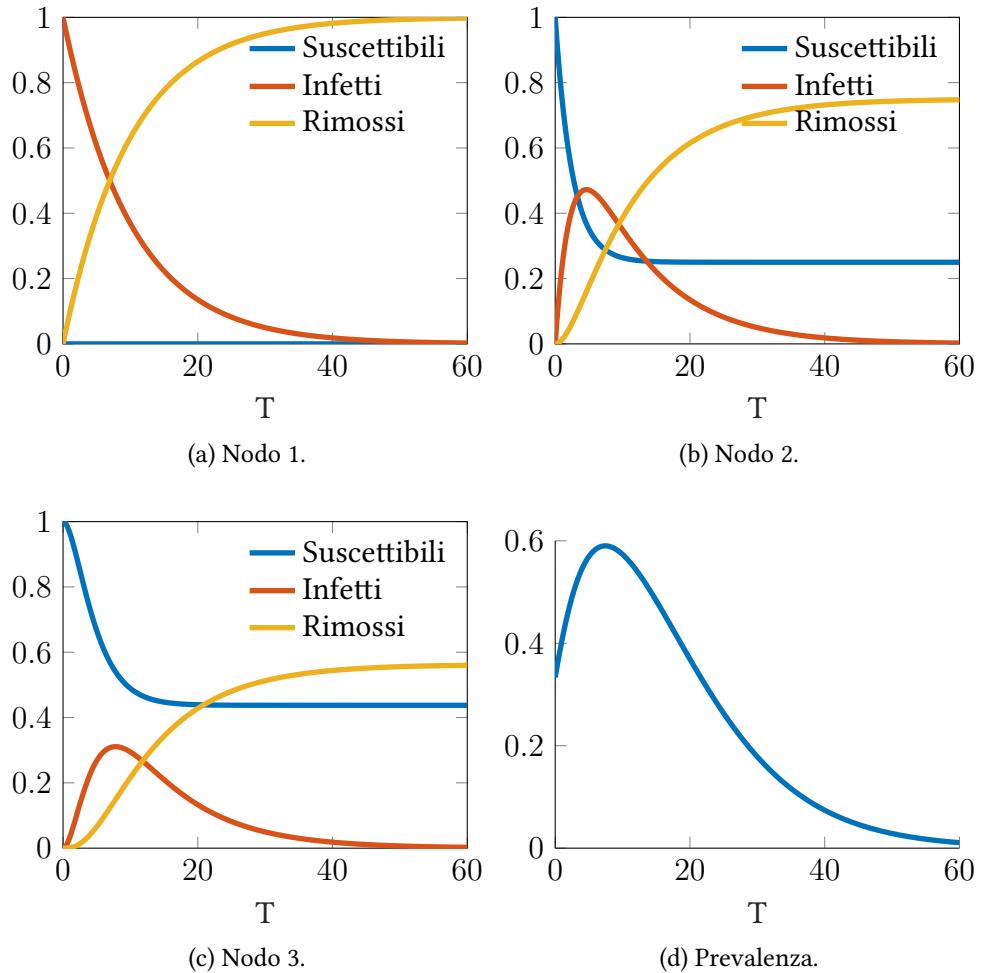


Figura 2.6: Divisione in classi nei singoli nodi (a)(b)(c) e grafico della prevalenza (d) per il grafo 2.1. Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, il problema di Cauchy ottenuto usando il Teorema 2.2.1 con condizioni iniziali di stati puri (2.2). Abbiamo inoltre supposto l'indipendenza statistica di tali condizioni iniziali.

Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

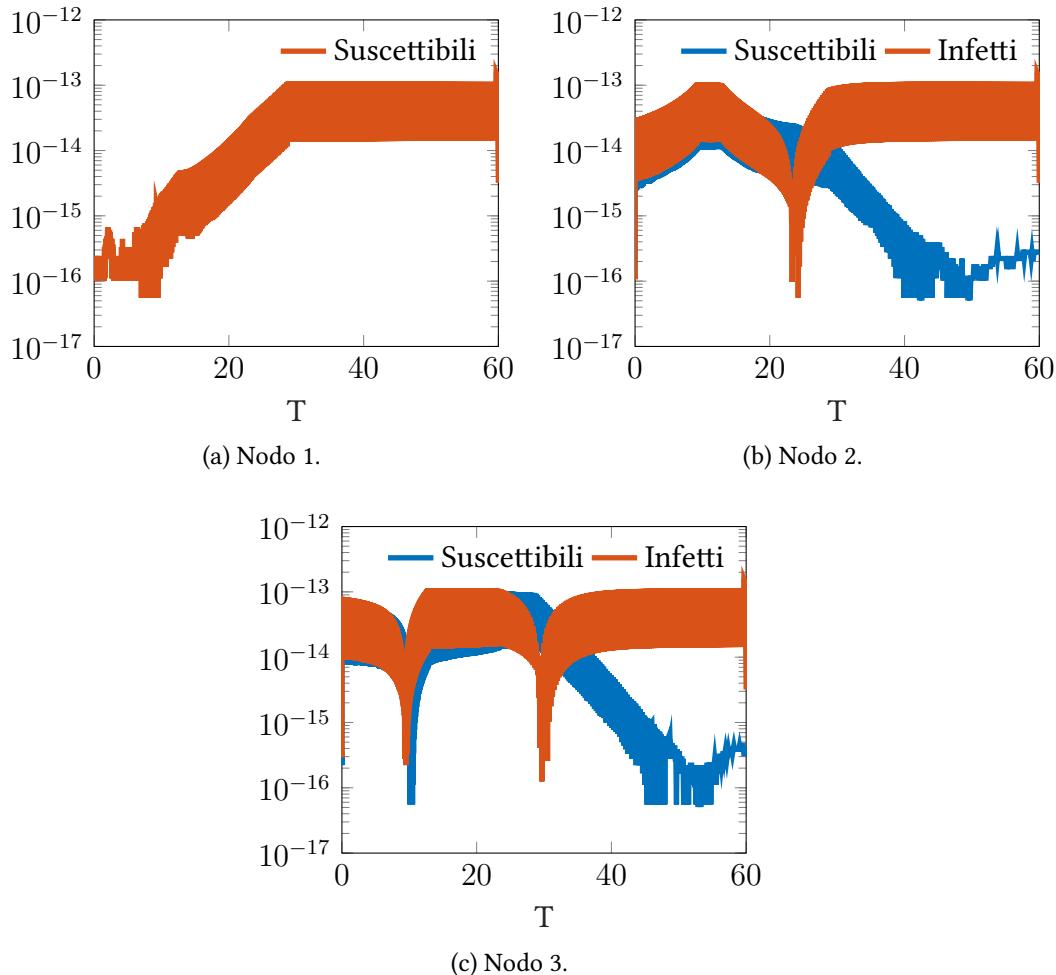


Figura 2.7: Errore assoluto (in scala logaritmica) tra la soluzione del problema di Cauchy (2.1) e quello ottenuto utilizzando il Teorema 2.2.1.

Per ottenere i grafici abbiamo risolto numericamente, con una tolleranza di $1e-12$ i due problemi di Cauchy con condizioni iniziali di stati puri (2.2). Abbiamo inoltre supposto l'indipendenza statistica di tali condizioni iniziali.

Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

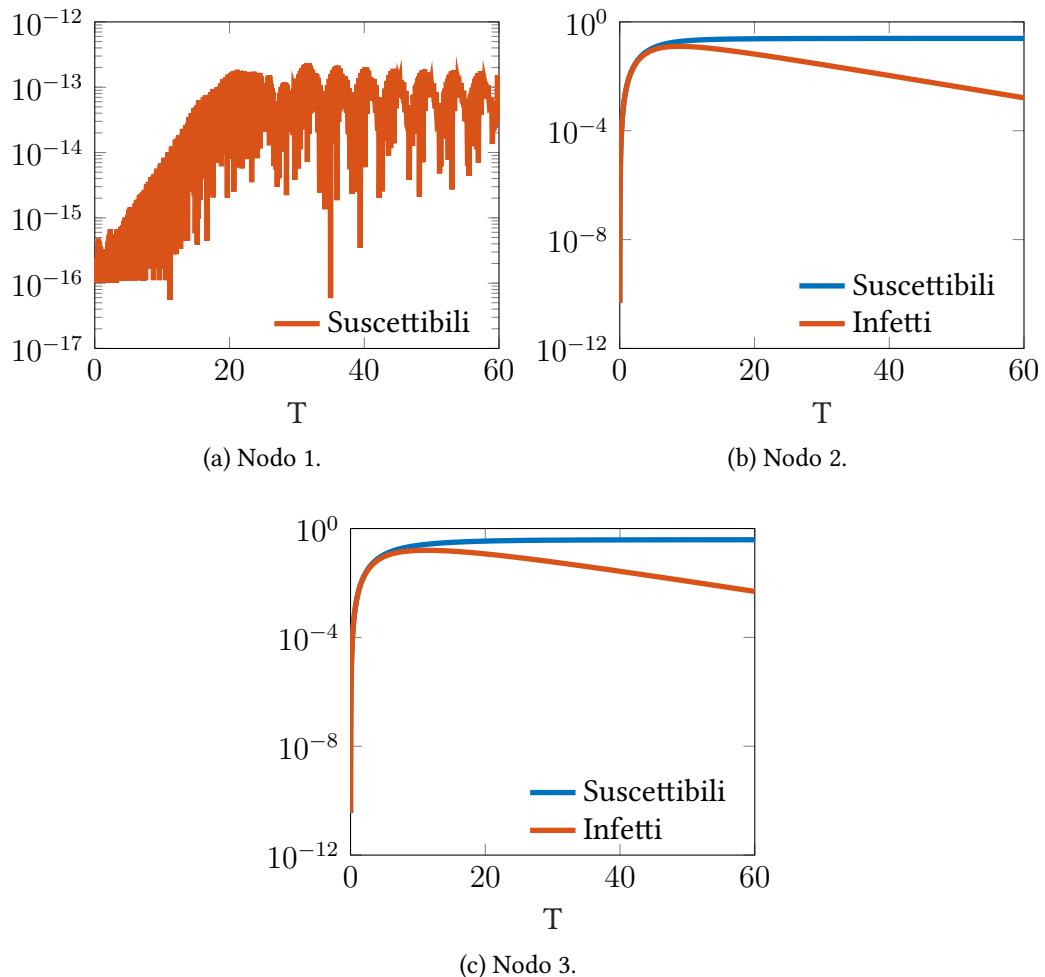


Figura 2.8: Errore assoluto (in scala logaritmica) tra la soluzione del problema di Cauchy (2.1) e quello utilizzando la chiusura alle coppie.

Per ottenere i grafici abbiamo risolto numericamente, con una tolleranza di $1e-12$ i due problemi di Cauchy con condizioni iniziali di stati puri (2.2). Abbiamo inoltre supposto l'indipendenza statistica di tali condizioni iniziali.

Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

2.3 — Un esempio completo

In questa sezione andremo a ricavare le equazioni per la rete in Figura 2.9 sia usando il modello esteso che attraverso l'uso dei cut-vertex.

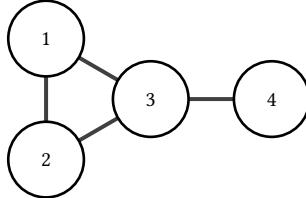


Figura 2.9: Il network lollipop.

2.3.1 Modello esteso

Le equazioni per i singoli nodi sono:

$$\begin{aligned}\dot{\langle S_1 \rangle} &= -\tau (\langle S_1 I_3 \rangle + \langle S_1 I_2 \rangle), \\ \dot{\langle I_1 \rangle} &= \tau (\langle S_1 I_2 \rangle + \langle S_1 I_3 \rangle) - \gamma \langle I_1 \rangle, \\ \dot{\langle S_2 \rangle} &= -\tau (\langle I_1 S_2 \rangle + \langle S_2 I_3 \rangle), \\ \dot{\langle I_2 \rangle} &= \tau (\langle I_1 S_2 \rangle + \langle S_2 I_3 \rangle) - \gamma \langle I_2 \rangle, \\ \dot{\langle S_3 \rangle} &= -\tau (\langle I_1 S_3 \rangle + \langle I_2 S_3 \rangle + \langle S_3 I_4 \rangle), \\ \dot{\langle I_3 \rangle} &= \tau (\langle I_1 S_3 \rangle + \langle I_2 S_3 \rangle + \langle S_3 I_4 \rangle) - \gamma \langle I_3 \rangle, \\ \dot{\langle S_4 \rangle} &= -\tau \langle I_3 S_4 \rangle, \\ \dot{\langle I_4 \rangle} &= \tau \langle I_3 S_4 \rangle - \gamma \langle I_4 \rangle.\end{aligned}$$

Tali equazioni dipendono da alcune coppie: tutte le disposizioni di archi con un nodo suscettibile ed un uno infetto. Abbiamo bisogno di equazioni addizionali:

$$\begin{aligned}\dot{\langle S_1 I_2 \rangle} &= \tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle S_1 I_2 \rangle - \tau \langle S_1 I_2 I_3 \rangle, \\ \dot{\langle S_1 I_3 \rangle} &= \tau (\langle S_1 I_2 S_3 \rangle + \langle S_1 S_3 I_4 \rangle) - (\tau + \gamma) \langle S_1 I_3 \rangle - \tau \langle S_1 I_2 I_3 \rangle, \\ \dot{\langle I_1 S_2 \rangle} &= \tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle I_1 S_2 \rangle - \tau \langle I_1 S_2 I_3 \rangle, \\ \dot{\langle S_2 I_3 \rangle} &= \tau (\langle I_1 S_2 S_3 \rangle + \langle S_2 S_3 I_4 \rangle) - (\tau + \gamma) \langle S_2 I_3 \rangle - \tau \langle I_1 S_2 I_3 \rangle, \\ \dot{\langle I_1 S_3 \rangle} &= -\tau (\langle I_1 I_2 S_3 \rangle \langle I_1 S_3 I_4 \rangle) - (\tau + \gamma) \langle I_1 S_3 \rangle + \tau \langle S_1 I_2 S_3 \rangle, \\ \dot{\langle I_2 S_3 \rangle} &= -\tau (\langle I_1 I_2 S_3 \rangle + \langle I_2 S_3 I_4 \rangle) - (\tau + \gamma) \langle I_2 S_3 \rangle + \tau \langle I_1 S_2 S_3 \rangle, \\ \dot{\langle S_3 I_4 \rangle} &= -\tau (\langle I_1 S_3 I_4 \rangle + \langle I_2 S_3 I_4 \rangle) - (\tau + \gamma) \langle S_3 I_4 \rangle, \\ \dot{\langle I_3 S_4 \rangle} &= \tau (\langle I_1 S_3 S_4 \rangle + \langle I_2 S_3 S_4 \rangle) - (\tau + \gamma) \langle I_3 S_4 \rangle.\end{aligned}$$

A questo punto, però necessitiamo anche delle triple:

$$\begin{aligned}\langle S_1 \dot{S}_2 I_3 \rangle &= \tau \langle S_1 S_2 S_3 I_4 \rangle - (2\tau + \gamma) \langle S_1 S_2 I_3 \rangle, \\ \langle S_1 \dot{I}_2 I_3 \rangle &= \tau (\langle S_1 S_2 I_3 \rangle + \langle S_1 I_2 S_3 I_4 \rangle) - 2(\tau + \gamma) \langle S_1 I_2 I_3 \rangle + \tau \langle S_1 I_2 S_3 \rangle, \\ \langle S_1 \dot{I}_2 S_3 \rangle &= -\tau \langle S_1 I_2 S_3 I_4 \rangle - (2\tau + \gamma) \langle S_1 I_2 S_3 \rangle, \\ \langle I_1 \dot{S}_2 I_3 \rangle &= \tau \langle S_1 S_2 I_3 \rangle - 2(\tau + \gamma) \langle I_1 S_2 I_3 \rangle + \tau \langle I_1 S_2 S_3 \rangle + \tau \langle I_1 S_2 S_3 I_4 \rangle, \\ \langle I_1 \dot{S}_2 S_3 \rangle &= -\tau \langle I_1 S_2 S_3 I_4 \rangle - (2\tau + \gamma) \langle I_1 S_2 S_3 \rangle, \\ \langle I_1 \dot{I}_2 S_3 \rangle &= \tau (\langle S_1 I_2 S_3 \rangle + \langle I_1 S_2 S_3 \rangle - \langle I_1 I_2 S_3 I_4 \rangle) - 2(\tau + \gamma) \langle I_1 I_2 S_3 \rangle.\end{aligned}$$

Mancano:

$$\begin{aligned}\langle I_1 S_3 I_4 \rangle &= -2(\tau + \gamma) \langle I_1 S_3 I_4 \rangle + \tau \langle S_1 I_2 S_3 I_4 \rangle - \tau \langle I_1 I_2 S_3 I_4 \rangle, \\ \langle I_1 S_3 S_4 \rangle &= -(\tau + \gamma) \langle I_1 S_3 S_4 \rangle + \tau \langle S_1 I_2 S_3 S_4 \rangle - \tau \langle I_1 I_2 S_3 S_4 \rangle, \\ \langle S_2 S_3 I_4 \rangle &= -(\tau + \gamma) \langle S_2 S_3 I_4 \rangle - 2\tau \langle I_1 S_2 S_3 I_4 \rangle, \\ \langle I_2 S_3 I_4 \rangle &= -2(\tau + \gamma) \langle I_2 S_3 I_4 \rangle + \tau \langle I_1 S_2 S_3 I_4 \rangle - \tau \langle I_1 I_2 S_3 I_4 \rangle, \\ \langle I_2 S_3 S_4 \rangle &= -(\tau + \gamma) \langle I_2 S_3 S_4 \rangle + \tau \langle I_1 S_2 S_3 S_4 \rangle - \tau \langle I_1 I_2 S_3 S_4 \rangle, \\ \langle S_1 S_3 I_4 \rangle &= -2\tau \langle S_1 I_2 S_3 I_4 \rangle - (\tau + \gamma) \langle S_1 S_3 I_4 \rangle,\end{aligned}$$

e le quadruple:

$$\begin{aligned}\langle I_1 I_2 S_3 I_4 \rangle &= -3(\tau + \gamma) \langle I_1 I_2 S_3 I_4 \rangle + \tau (\langle S_1 I_2 S_3 I_4 \rangle + \langle I_1 S_2 S_3 I_4 \rangle), \\ \langle S_1 I_2 S_3 I_4 \rangle &= -(3\tau + 2\gamma) \langle S_1 I_2 S_3 I_4 \rangle, \\ \langle I_1 S_2 S_3 I_4 \rangle &= -(3\tau + 2\gamma) \langle I_1 S_2 S_3 I_4 \rangle, \\ \langle I_1 I_2 S_3 S_4 \rangle &= \tau (\langle S_1 I_2 S_3 S_4 \rangle + \langle I_1 S_2 S_3 S_4 \rangle) - 2(\tau + \gamma) \langle I_1 I_2 S_3 S_4 \rangle, \\ \langle I_1 S_2 S_3 S_4 \rangle &= -(2\tau + \gamma) \langle I_1 S_2 S_3 S_4 \rangle, \\ \langle S_1 I_2 S_3 S_4 \rangle &= -(2\tau + \gamma) \langle S_1 I_2 S_3 S_4 \rangle, \\ \langle S_1 S_2 S_3 I_4 \rangle &= -(\tau + \gamma) \langle S_1 S_2 S_3 I_4 \rangle.\end{aligned}$$

Abbiamo così ottenuto un sistema chiuso di 35 equazioni.

2.3.2 Modello chiuso

Vediamo come possiamo sfruttare il Teorema 2.2.1 per ridurre il numero di equazioni.

La rete lollipop possiede un solo cut-vertex: il vertice 3. Decomponiamo il grafo in un triangolo ed in un singolo arco (Figura 2.10).

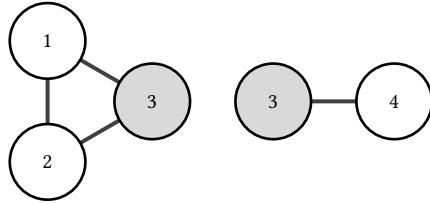


Figura 2.10: Decomposizione della rete lollipop.

Per i singoli nodi otteniamo le stesse equazioni viste in precedenza:

$$\begin{aligned}\dot{\langle S_1 \rangle} &= -\tau (\langle S_1 I_3 \rangle + \langle S_1 I_2 \rangle), \\ \dot{\langle I_1 \rangle} &= \tau (\langle S_1 I_2 \rangle + \langle S_1 I_3 \rangle) - \gamma \langle I_1 \rangle, \\ \dot{\langle S_2 \rangle} &= -\tau (\langle I_1 S_2 \rangle + \langle S_2 I_3 \rangle), \\ \dot{\langle I_2 \rangle} &= \tau (\langle I_1 S_2 \rangle + \langle S_2 I_3 \rangle) - \gamma \langle I_2 \rangle \\ \dot{\langle S_3 \rangle} &= -\tau (\langle I_1 S_3 \rangle + \langle I_2 S_3 \rangle + \langle S_3 I_4 \rangle), \\ \dot{\langle I_3 \rangle} &= \tau (\langle I_1 S_3 \rangle + \langle I_2 S_3 \rangle + \langle S_3 I_4 \rangle) - \gamma \langle I_3 \rangle, \\ \dot{\langle S_4 \rangle} &= -\tau \langle I_3 S_4 \rangle, \\ \dot{\langle I_4 \rangle} &= \tau \langle I_3 S_4 \rangle - \gamma \langle I_4 \rangle.\end{aligned}$$

Andiamo a ricavarci le equazioni per le coppie.

Esaminiamo l'equazione per $\langle S_1 I_3 \rangle$.

$$\begin{aligned}\dot{\langle S_1 I_3 \rangle} &= \tau (\langle S_1 I_2 S_3 \rangle + \langle S_1 S_3 I_4 \rangle) - (\tau + \gamma) \langle S_1 I_3 \rangle - \tau \langle S_1 I_2 I_3 \rangle = \\ &= \tau \left(\langle S_1 I_2 S_3 \rangle + \frac{\langle S_1 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - (\tau + \gamma) \langle S_1 I_3 \rangle - \tau \langle S_1 I_2 S_3 \rangle\end{aligned}$$

dove abbiamo usato il Teorema 2.2.1 per sostituire $\langle S_1 S_3 I_4 \rangle$.

Con considerazioni analoghe otteniamo

$$\begin{aligned}\langle \dot{S_1 I_2} \rangle &= \tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle S_1 I_2 \rangle - \tau \langle S_1 I_2 I_3 \rangle, \\ \langle \dot{I_1 S_2} \rangle &= \tau \langle S_1 S_2 I_3 \rangle - (\tau + \gamma) \langle I_1 S_2 \rangle - \tau \langle I_1 S_2 I_3 \rangle, \\ \langle \dot{S_2 I_3} \rangle &= \tau \left(\langle I_1 S_2 S_3 \rangle + \frac{\langle S_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - (\tau + \gamma) \langle S_2 I_3 \rangle - \tau \langle I_1 S_2 I_3 \rangle, \\ \langle \dot{I_1 S_3} \rangle &= -\tau \left(\langle I_1 I_2 S_3 \rangle + \frac{\langle I_1 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - (\tau + \gamma) \langle I_1 S_3 \rangle + \tau \langle S_1 I_2 S_3 \rangle, \\ \langle \dot{I_2 S_3} \rangle &= -\tau \left(\langle I_1 I_2 S_3 \rangle + \frac{\langle I_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - (\tau + \gamma) \langle I_2 S_3 \rangle + \tau \langle I_1 S_2 S_3 \rangle, \\ \langle \dot{S_3 I_4} \rangle &= -\tau \left(\frac{\langle I_1 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} + \frac{\langle I_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - (\tau + \gamma) \langle S_3 I_4 \rangle, \\ \langle \dot{I_3 S_4} \rangle &= \tau \left(\frac{\langle I_1 S_3 \rangle \langle S_3 S_4 \rangle}{\langle S_3 \rangle} + \frac{\langle I_2 S_3 \rangle \langle S_3 S_4 \rangle}{\langle S_3 \rangle} \right) - (\tau + \gamma) \langle I_3 S_4 \rangle.\end{aligned}$$

Da queste equazioni emergono nuove coppie, dunque necessitiamo anche di

$$\begin{aligned}\langle \dot{S_1 S_3} \rangle &= -\tau \frac{\langle S_1 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} - 2\tau \langle S_1 I_2 S_3 \rangle, \\ \langle \dot{S_2 S_3} \rangle &= -\tau \frac{\langle S_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} - 2\tau \langle I_1 S_2 S_3 \rangle, \\ \langle \dot{S_3 S_4} \rangle &= -\tau \frac{\langle I_1 S_3 \rangle \langle S_3 S_4 \rangle}{\langle S_3 \rangle} - \tau \frac{\langle I_2 S_3 \rangle \langle S_3 S_4 \rangle}{\langle S_3 \rangle}.\end{aligned}$$

Le equazioni per le coppie introducono alcune triple. Aggiungiamole.

$$\begin{aligned}\langle \dot{S_1 S_2 I_3} \rangle &= \tau \frac{\langle S_1 S_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} - (2\tau + \gamma) \langle S_1 S_2 I_3 \rangle, \\ \langle \dot{S_1 I_2 I_3} \rangle &= \tau \left(\langle S_1 S_2 I_3 \rangle + \frac{\langle S_1 I_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - 2(\tau + \gamma) \langle S_1 I_2 I_3 \rangle + \\ &\quad + \tau \langle S_1 I_2 S_3 \rangle, \\ \langle \dot{S_1 I_2 S_3} \rangle &= -\tau \frac{\langle S_1 I_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} - (2\tau + \gamma) \langle S_1 I_2 S_3 \rangle, \\ \langle \dot{I_1 S_2 I_3} \rangle &= \tau \langle S_1 S_2 I_3 \rangle - 2(\tau + \gamma) \langle I_1 S_2 I_3 \rangle + \tau \langle I_1 S_2 S_3 \rangle + \\ &\quad + \tau \frac{\langle I_1 S_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle}, \\ \langle \dot{I_1 S_2 S_3} \rangle &= -\tau \frac{\langle I_1 S_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} - (2\tau + \gamma) \langle I_1 S_2 S_3 \rangle,\end{aligned}$$

$$\begin{aligned}\langle I_1 \dot{I}_2 S_3 \rangle = & -\tau \left(\langle S_1 I_2 S_3 \rangle + \langle I_1 S_2 S_3 \rangle - \frac{\langle I_1 I_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle} \right) - \\ & - 2(\tau + \gamma) \langle I_1 I_2 S_3 \rangle.\end{aligned}$$

Scrivendo queste equazioni appare una nuova tripla. Completiamo il sistema aggiungendo l'equazione mancante:

$$\langle S_1 \dot{S}_2 S_3 \rangle = -\tau \frac{\langle S_1 S_2 S_3 \rangle \langle S_3 I_4 \rangle}{\langle S_3 \rangle}.$$

Siamo riusciti a ridurre il numero di equazioni da 35 a 27.

Andiamo a risolvere numericamente i due sistemi con condizioni iniziali a partire da uno stato puro (4 certamente infetto e 1, 2, 3 certamente sani):

$$\begin{aligned}\langle S_1 \rangle(0) &= 1, \\ \langle I_1 \rangle(0) &= 0, \\ \langle S_2 \rangle(0) &= 1, \\ \langle I_2 \rangle(0) &= 0, \\ \langle S_3 \rangle(0) &= 1, \\ \langle I_3 \rangle(0) &= 0, \\ \langle S_4 \rangle(0) &= 0, \\ \langle I_4 \rangle(0) &= 1.\end{aligned}\tag{2.9}$$

Nella Figura 2.11 sono riportati i grafici dello stato dei nodi in funzione del tempo sia per il modello esteso che per quello chiuso. Come ci aspettavamo dalla teoria i grafici ottenuti dai due modelli coincidono su ogni nodo.

Per marcire ciò abbiamo calcolato gli errori assoluti tra le due soluzioni (Figura 2.12) ottenendo che tali errori sono dell'ordine della tolleranza dell'integratore. Facendo la medesima cosa tra il modello esatto e quello chiuso alle coppie otteniamo che gli errori sono molto maggiori (Figura 2.13).

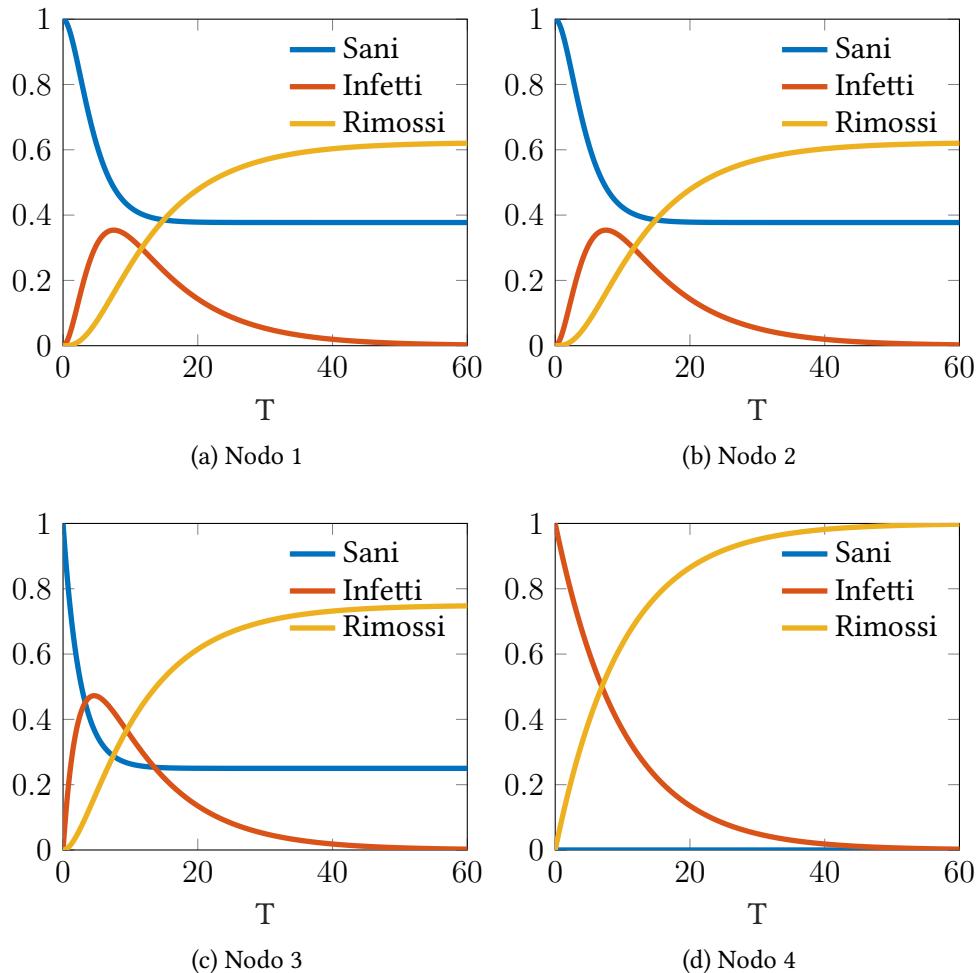


Figura 2.11: Divisione in classe dei singoli nodi per il grafo 2.9.

Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, i due problemi di Cauchy con condizioni iniziali di stati puri (2.9). Abbiamo inoltre supposto l'indipendenza statistica di tali condizioni iniziali.

Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

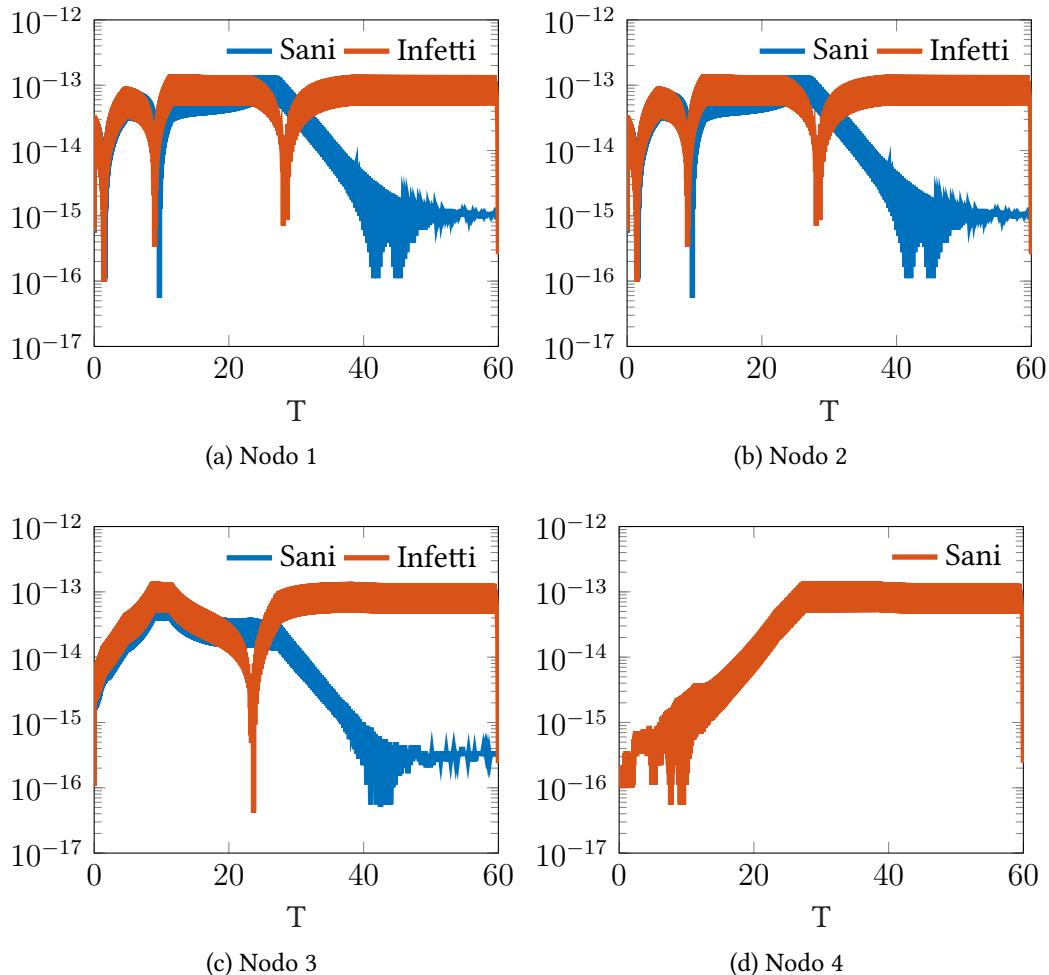


Figura 2.12: Errore assoluto (in scala logaritmica) tra la soluzione del problema esteso e quello chiuso con cut-vertex.

Per ottenere i grafici abbiamo risolto numericamente, con una tolleranza di $1e-12$ i due problemi di Cauchy con condizioni iniziali di stati puri (2.9). Abbiamo inoltre supposto l'indipendenza statistica delle condizioni iniziali.

Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

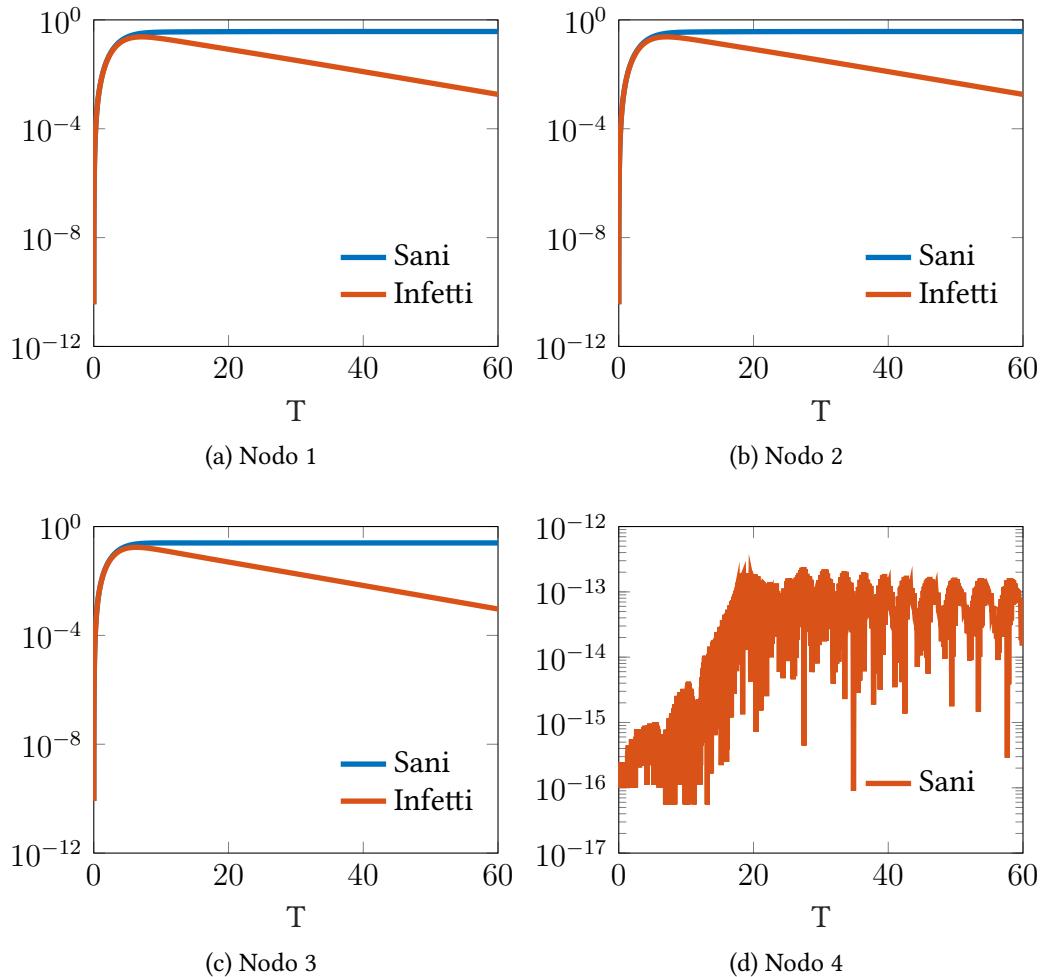


Figura 2.13: Errore assoluto (in scala logaritmica) tra la soluzione del problema esteso e quello chiuso alle coppie.

Per ottenere i grafici abbiamo risolto numericamente, con una tolleranza di $1e-12$ i due problemi di Cauchy con condizioni iniziali di stati puri (2.9). Abbiamo inoltre supposto l'indipendenza statistica di tali condizioni iniziali.

Per la sperimentazione abbiamo usato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

3 — Integrazione numerica di ODE

In questo capitolo presenteremo alcuni metodi utilizzabili per l'integrazione di ODE. Per la parte teorica abbiamo fatto riferimento a [9, Capitoli 5-8]. Sia $f(t, u)$ la funzione

$$f : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$$

continua in t e lipschitziana in y .

Consideriamo il problema al valore iniziale per un sistema di n equazioni ordinarie del primo ordine

$$\begin{cases} u' = f(t, u) \text{ per } t > t_0 \\ u(t_0) = u_0 \end{cases} \quad (3.1)$$

Sotto tali condizioni il problema di Cauchy (3.1) ammette una ed una sola soluzione $u(t)$, per $t \geq t_0$.

3.1 — Metodi d'integrazione ad un passo

I metodi d'integrazione ad un passo sono una classe molto utilizzata per il calcolo numerico della soluzione di (3.1). Tali metodi approssimano i valori della soluzione in corrispondenza di un insieme discreto di valori t_0, t_1, \dots ; posto $U^0 = u_0$ si calcolano U^1, U^2, \dots in modo che

$$U^n \approx u(t_n).$$

In generale un metodo ad un passo ha la forma

$$U^{n+1} = U^n + h_{n+1} \Phi(t_n, U^n, h_{n+1}, f).$$

Tale metodo è caratterizzato dal fatto che il valore approssimato U^{n+1} è ottenuto a partire da un solo valore precedente U^n .

Sono detti metodi *impliciti* quelli che richiedono di risolvere, ad ogni passo, un sistema di equazioni per ricavare U^{n+1} .

L'*errore locale di troncamento*

$$\sigma(n) = u(t_n) - U^n$$

è l'errore dovuto all'impiego del valore approssimato fornito dal modello.

3.1.1 Eulero esplicito

Fissato un numero k di intervalli temporali, poniamo $t_n = nk$.

Noto U^n per calcolare U^{n+1} possiamo approssimare il limite che definisce $u'(t_n)$:

$$\lim_{h \rightarrow 0^+} \frac{u(t_n + h) - u(t_n)}{h}$$

con la differenza finita *in avanti*

$$D_+ U^n = \frac{U^{n+1} - U^n}{k}.$$

Tale approssimazione ci permette di ottenere

$$f(U^n) = \frac{U^{n+1} - U^n}{k} \text{ per } n = 0, 1, \dots$$

Utilizzando la seguente formula

$$U^{n+1} = kf(U^n) + U^n,$$

possiamo calcolare esplicitamente U^{n+1} .

3.1.2 Eulero implicito

L'idea che sta dietro a questo metodo è simile a quella del metodo precedente. Approssimiamo il limite che definisce $u'(t_{n+1})$ con la differenza finita *all'indietro*

$$D_- U^{n+1} = \frac{U^{n+1} - U^n}{k}.$$

Dunque otteniamo

$$\frac{U^{n+1} - U^n}{k} = f(U^{n+1}),$$

ovvero

$$U^{n+1} = U^n + kf(U^{n+1}). \quad (3.2)$$

Per calcolare U^{n+1} usando la (3.2) occorre cercare uno zero della funzione

$$g(x) = x - kf(x) - U^n$$

che, in generale, non è lineare. Un'approssimazione di tale zero può essere trovata utilizzando un metodo di Newton.

3.1.3 Serie di Taylor

Il metodo di Eulero esplicito può essere derivato usando uno sviluppo di $u(t_{n+1})$ in serie di Taylor centrato in $u(t_n)$ troncato al secondo termine:

$$u(t_{n+1}) = u(t_n) + ku'(t_n) + O(k^2)$$

sostituendo $u'(t_n)$ con $f(u(t_n), t_n)$.

Un metodo di ordine maggiore può essere derivato mantenendo più termine nella serie di Taylor: se tronchiamo lo sviluppo dopo i primi $p+1$ termini

$$u(t_{n+1}) \approx u(t_n) + ku'(t_n) + \frac{1}{2}k^2u''(t_n) + \cdots + \frac{1}{p!}k^p u^{(p)}(t_n)$$

otteniamo un metodo di ordine p .

Il problema, di tale metodo, è che avendo a disposizione solamente

$$u'(t) = f(u(t), t)$$

dobbiamo calcolare le derivate superiori per differenziazione ripetuta della funzione f . Per esempio, possiamo calcolare

$$\begin{aligned} u''(t) &= \frac{\partial}{\partial u} f(u(t), t) u'(t) + \frac{\partial}{\partial t} f(u(t), t) = \\ &= \frac{\partial}{\partial u} f(u(t), t) f(u(t), t) + \frac{\partial}{\partial t} f(u(t), t). \end{aligned}$$

Tale metodo può originare un elevato numero di espressioni differenti che devono essere ricavate e calcolate per ogni equazione. Per questo motivo tale metodo nella pratica risulta essere scarsamente utilizzato.

3.1.4 Runge-Kutta

I metodi di Runge-Kutta evitano gli inconvenienti emersi dall'impiego delle formule di Taylor: le derivate successive vengono approssimate calcolando la funzione f in opportuni istanti nell'intervallo (t_n, t_{n+1}) .

Per presentare l'idea alla base di questa classe di metodi mostriamo come ottenere un metodo del secondo ordine.

Poniamo

$$\phi(t, u; h) = c_1 f(t, u) + c_2 f(t + ha_2, u + hb_2 f(t, u))$$

dove i coefficienti a_2 , b_{21} , c_1 e c_2 sono da determinare in modo che il metodo abbia ordine 2. Confrontando gli sviluppi di Taylor della soluzione calcolata e di quella esatta otteniamo le condizioni

$$\begin{aligned} c_1 + c_2 &= 1 \\ c_1 a_1 &= \frac{1}{2} \\ c_2 b_2 &= \frac{1}{2}. \end{aligned} \tag{3.3}$$

Per calcolare U^{n+1} possiamo far ricorso alla formula

$$U^{n+1} = U^n + k\phi(t_n, U^n; k)$$

Il procedimento esposto in precedenza può essere esteso in modo da utilizzare un qualsiasi numero m di valutazione della funzione f . In questo modo otterremo il metodo generale di Runge-Kutta esplicito a m stadi:

$$\begin{aligned} h_1(t, u; k) &= f(u, t) \\ h_r(t, u; k) &= f\left(t + ka_r, u + h \sum_{s=1}^{r-1} b_{r,s} h_s\right) \text{ per } r = 2, \dots, m \\ \phi(t, u; k) &= \sum_{r=1}^m c_r h_r(t, u) \\ U^{n+1} &= U^n + h\phi(t_n, U^n; k) \end{aligned} \tag{3.4}$$

Un modo usuale di rappresentare tale metodo è sotto forma di *tableau di Butcher* come fatto in Figura 3.1.

La maggior parte dei software per la risoluzione di ODE non utilizza un passo temporale fisso ma, piuttosto, lo regola durante il processo d'integrazione in modo da avere un controllo sull'errore locale di troncamento.

Un modo utilizzato per stimare l'errore è quello di confrontare le soluzioni con due metodi diversi: uno di ordine p ed un altro di ordine $p + 1$ (o maggiore). In questo modo, se il passo è sufficientemente piccolo, il metodo di ordine maggiore genera un'approssimazione migliore dell'altro. La differenza tra le due soluzioni fornisce quindi una stima sull'errore del metodo di ordine inferiore.

Possiamo implementare questa strategia anche nei metodi di Runge-Kutta. Le soluzioni dei due metodi vengono prese scegliendo semplicemente due vettori c nel *tableau* in Figura 3.1; in questo modo, sfruttando le medesime valutazioni della dinamica f del sistema, si generano le due soluzioni .

Questa versione del metodo di Runge-Kutta viene implementata nella funzione `ode45` di MATLAB di cui riportiamo il tableau (Figura 3.2), si veda [2] per i dettagli e la derivazione.

0			
a_2	$b_{2,1}$		
\vdots	\vdots	\ddots	
a_m	$b_{m,1}$	\dots	$b_{m,m-1}$
	c_1	\dots	\dots
			c_m

Figura 3.1: Tableau del metodo generale di Runge-Kutta esplicito a m stadi

3.2 — Assoluta stabilità

Consideriamo il problema test

$$\begin{cases} u(t) = \lambda u(t) \\ u(t_0) = 1 \end{cases} \quad (3.5)$$

dove $Re(\lambda) < 0$.

Di tale problema conosciamo la soluzione esatta:

$$u(t) = e^{\lambda t}.$$

Diremo che un metodo numerico è *assolutamente stabile* se

$$\lim_{n \rightarrow +\infty} |u_n| = 0$$

Chiameremo *regione di assoluta stabilità* del metodo numerico

$$\mathcal{A} = \{z = \lambda k \in \mathbb{C} : |u_n| \rightarrow 0 \text{ per } t_n \rightarrow +\infty\}$$

Diremo che un metodo è \mathcal{A} -stabile se

$$\mathcal{A} \supseteq \mathbb{C}^-$$

ovvero se per $Re(\lambda) < 0$ la condizione di assoluta stabilità è verificata per ogni valore di k .

Il metodo di Eulero esplicito applicato al problema test (3.5) produce

$$U^{n+1} = (1 + \lambda k)U^n.$$

Per ricorsione su n otteniamo

$$U^n = (1 + \lambda k)^n$$

	0						
	$\frac{1}{5}$	$\frac{1}{5}$					
	$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$				
	$\frac{4}{5}$	$\frac{44}{45}$	$-\frac{56}{15}$	$\frac{32}{9}$			
	$\frac{8}{9}$	$\frac{19372}{6561}$	$-\frac{25360}{2187}$	$\frac{64448}{6561}$	$-\frac{212}{729}$		
	1	$\frac{9017}{3168}$	$-\frac{355}{33}$	$\frac{46732}{5247}$	$\frac{49}{176}$	$-\frac{5103}{18656}$	
	1	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$
		$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$
		$\frac{5179}{57600}$	0	$\frac{7571}{16695}$	$\frac{393}{640}$	$-\frac{92097}{339200}$	$\frac{187}{2100}$
							$\frac{1}{40}$

Figura 3.2: Tableau del metodo di Runge-Kutta implementato dalla funzione `ode45` di MATLAB

dunque il metodo è assolutamente stabile se $|1 + \lambda k| < 1$. La regione di assoluta stabilità è il disco (nel piano complesso) di raggio 1 centrato in $(-1, 0)$.

Il metodo di Eulero implicito applicato al problema test (3.5) produce

$$U^{n+1} = \frac{1}{1 - \lambda k} U^n.$$

Per ricorsione su n otteniamo

$$U^n = \frac{1}{(1 - \lambda k)^n}$$

dunque il metodo è assolutamente stabile se $|1 - \lambda k| > 1$. La regione di assoluta stabilità è la parte esterna (nel piano complesso) del disco di raggio 1 centrato in $(1, 0)$.

Le due regioni di assoluta stabilità sono riportate per confronto nella Figura 3.3. Osserviamo, in particolare, che il metodo di Eulero implicito è un metodo \mathcal{A} -stabile, poiché l'intero \mathbb{C}^- è contenuto in \mathcal{A} .

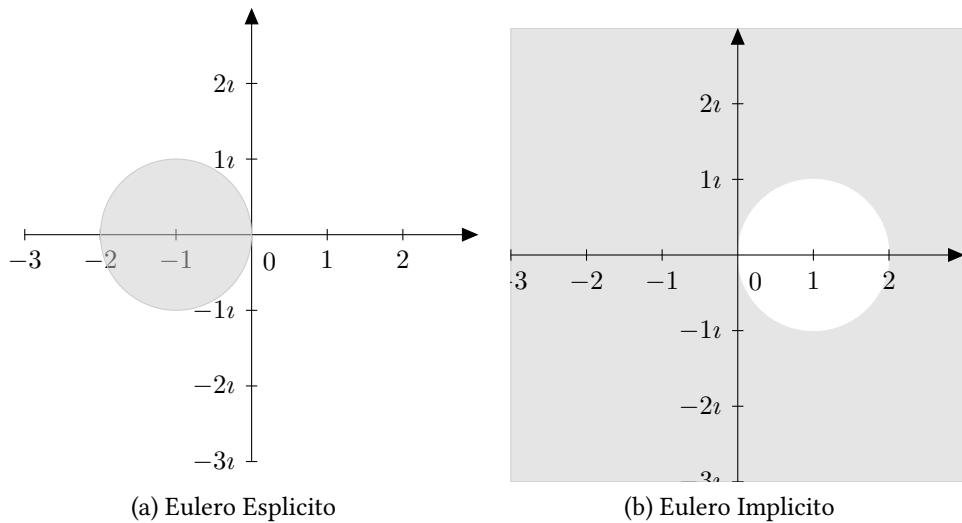


Figura 3.3: Regioni di assolutà stabilità \mathcal{A} dei metodi di Eulero Implicito ed Esplicito.

3.3 — Stiffness

Dato il problema (3.1) definiamo il quoziente di stiffness nel punto $(\bar{t}, u(\bar{t}))$ come il rapporto

$$\frac{\max |\lambda_i|}{\min |\lambda_i|}$$

dove λ_i sono gli autovalori della matrice Jacobiana del sistema calcolata nel punto $(\bar{t}, u(\bar{t}))$.

Diremo che un problema è stiff se tale rapporto è molto più grande di 1.

Tale definizione formale presenta diverse limitazioni e non sempre caratterizza accuratamente i sistemi stiff. Riportiamo una definizione di carattere più euristico dovuta a J. Lambert [8] e qui tradotta in italiano:

“Un sistema di equazioni differenziali ordinarie si dice stiff se, approssimandolo con uno schema numerico che presenti una regione di assoluta stabilità limitata, esso, per ogni dato iniziale, obbliga lo schema numerico ad impiegare un passo di discretizzazione molto più piccolo di quello realmente necessario per descrivere ragionevolmente l’andamento della soluzione esatta”.

3.4 — Esempio su una rete reale

Concludiamo la tesi con un'applicazione del modello SIR su una rete reale: la rete stradale del Minnesota (Figura 3.4). In questa rete i nodi rappresentano gli incroci tra diverse stradi riportate come archi. In particolare, ci concentriamo su alcuni

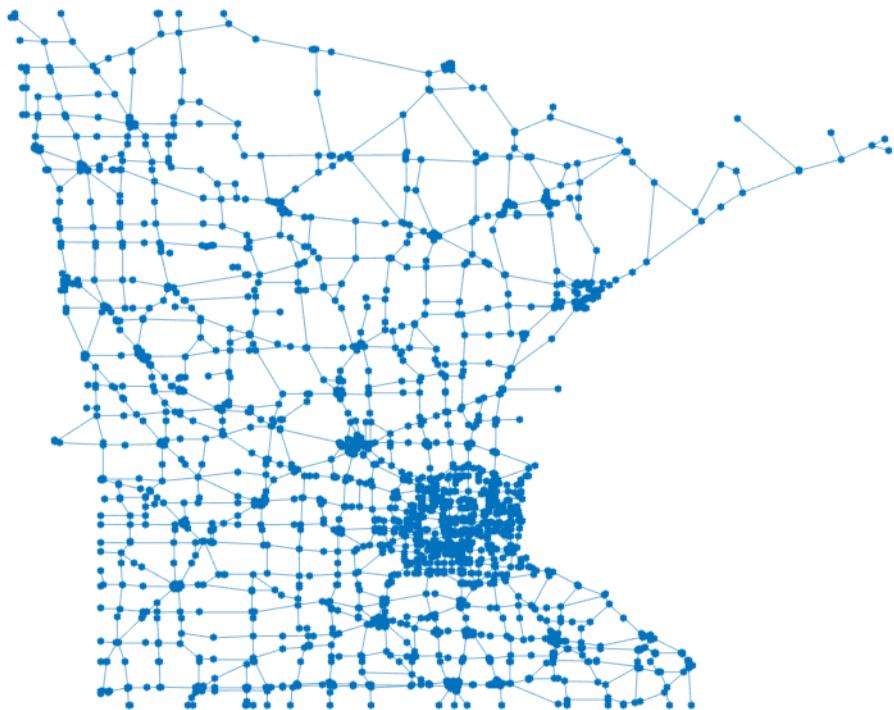


Figura 3.4: Grafico della rete stradale del Minnesota. I dati provengono dalla *SuiteSparse Matrix Collection* [4].

aspetti numerici concernenti l'integrazione del sistema di ODE risultante.

Un primo quesito su cui abbiamo indagato è come la scelta dei tassi d'infezione e di recupero influenzassero la scelta dei passi temporali (*time-step*) nei due metodi. Il time-step deve essere scelto sufficientemente piccolo in modo da garantire il controllo dell'errore locale di troncamento. In particolare, se le condizioni sulla stabilità ci costringono ad utilizzare un passo molto più piccolo di quanto indicato dall'errore di troncamento locale allora il metodo che abbiamo scelto non è ottimale per il problema. Ciò accade, ad esempio, se proviamo ad utilizzare un metodo esplicito per risolvere un problema stiff.

In Figura 3.5 si può osservare come varia il numero di time-step utilizzati dalla funzione `ode45` e `ode15s` per diversi valori dei parametri τ e γ . Per tutti le

combinazioni dei parametri la funzione `ode15s` impiega molte meno iterazioni: il modello, dunque, è stiff. Un'altra conferma di ciò ci viene data anche calcolando il rapporto di stiffness (Figura 3.6) nelle condizioni iniziali.

Questo ci suggerisce di utilizzare metodi di tipo implicito per integrare i sistemi di ode (2.4).

Un altro aspetto su cui abbiamo indagato è il legame tra il grado del nodo in cui “inizia” l’epidemia e il suo sviluppo. Nella rete in esame ogni nodo ha come grado un numero compreso tra 1 e 5 dunque abbiamo risolto 5 problemi di Cauchy facendo partire l’epidemia ogni volta da un nodo di grado diverso. In Figura 3.7 si osserva come al crescere del grado del nodo inizialmente infetto si velocizzi il tempo per arrivare al picco. Notiamo, inoltre, che il picco rimane pressappoco uguale sia che si scelga un nodo di grado 1 che uno di grado 5.

L’ultimo quesito su cui abbiamo indagato è come l’immunizzazione cambi le curve di prevalenza. Per simulare l’immunizzazione di una percentuale x della popolazione abbiamo considerato come condizioni iniziale un solo nodo infetto con probabilità $1 - x$ e rimosso con probabilità x mentre gli altri nodi sani con probabilità x e rimossi con probabilità $1 - x$. In Figura 3.8 è possibile notare come una maggior immunizzazione riduca il picco e lo sposti nel tempo.

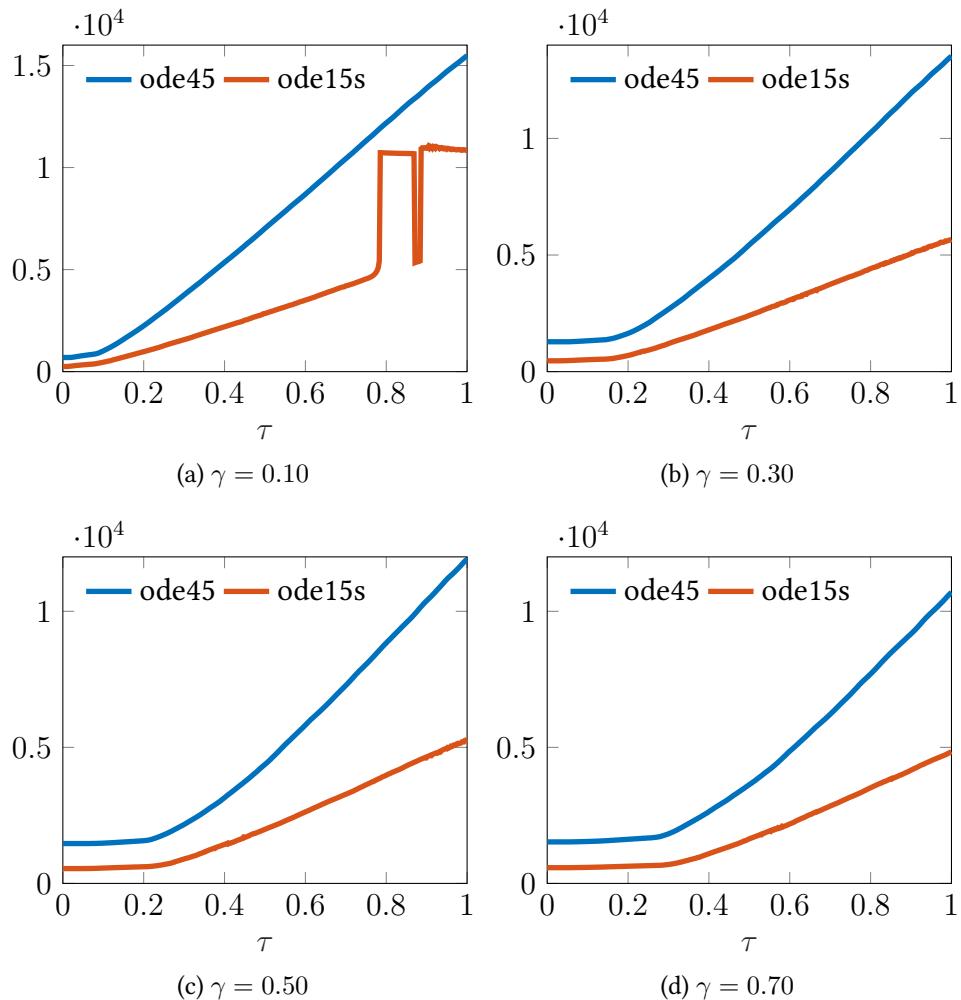


Figura 3.5: Numero di time-step per `ode45` e `ode15s` sulla rete stradale del Minnesota al variare dei parametri.

Per ottenere i grafici abbiamo risolto numericamente, usando gli integratori `ode45` e `ode15s`, il problema chiuso alle coppie con condizioni iniziali di stati puri (1 sicuramente infetto, gli altri nodi sicuramente sani) per la rete 3.4.

Per la sperimentazione abbiamo usato l'intervallo temporale [0 30].

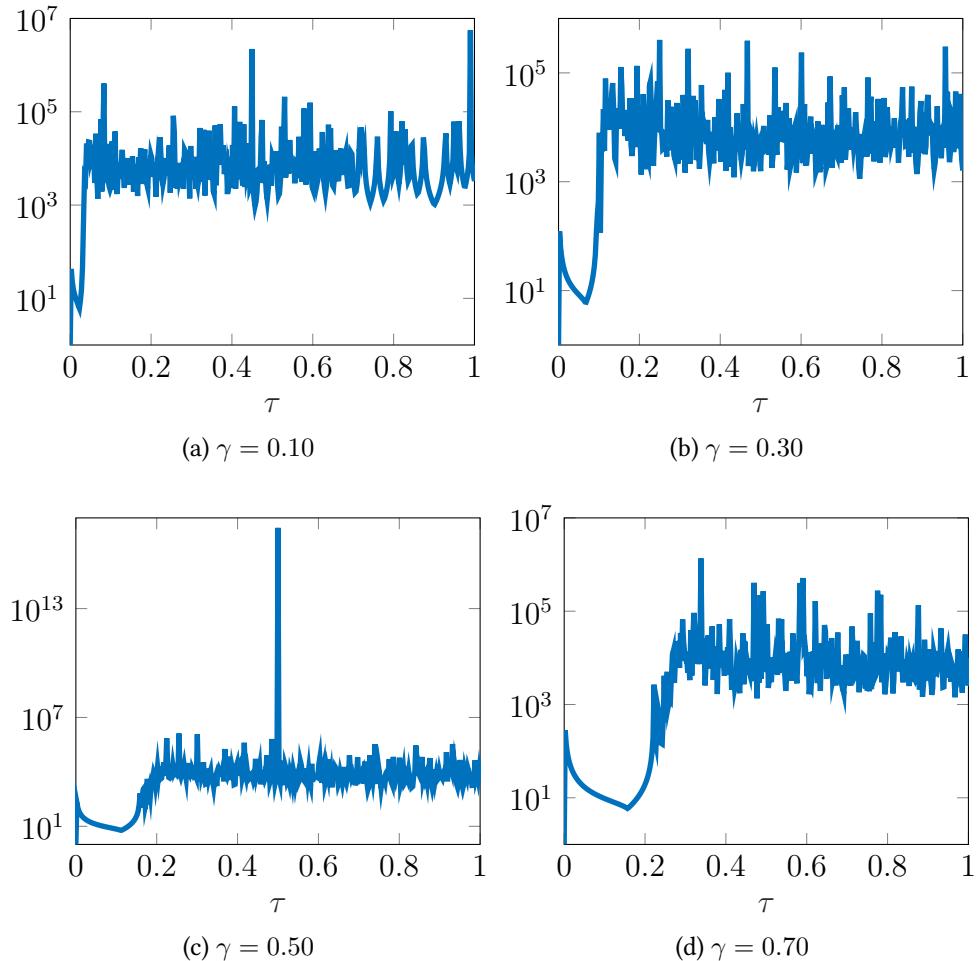


Figura 3.6: Stiff ratio calcolato alle condizioni iniziali: stati puri (1 sicuramente infetto, gli altri nodi sicuramente sani) al variare dei parametri τ e γ .

Per ottenere i grafici abbiamo calcolato il rapporto tra l'autovalore di modulo massimo e quello minimo della matrice Jacobiana, calcolato alle condizioni iniziali, del modello chiuso alle coppie.

Per la sperimentazione abbiamo calcolato numericamente gli autovalori della matrice Jacobiana usando il comando `eig` di MATLAB.

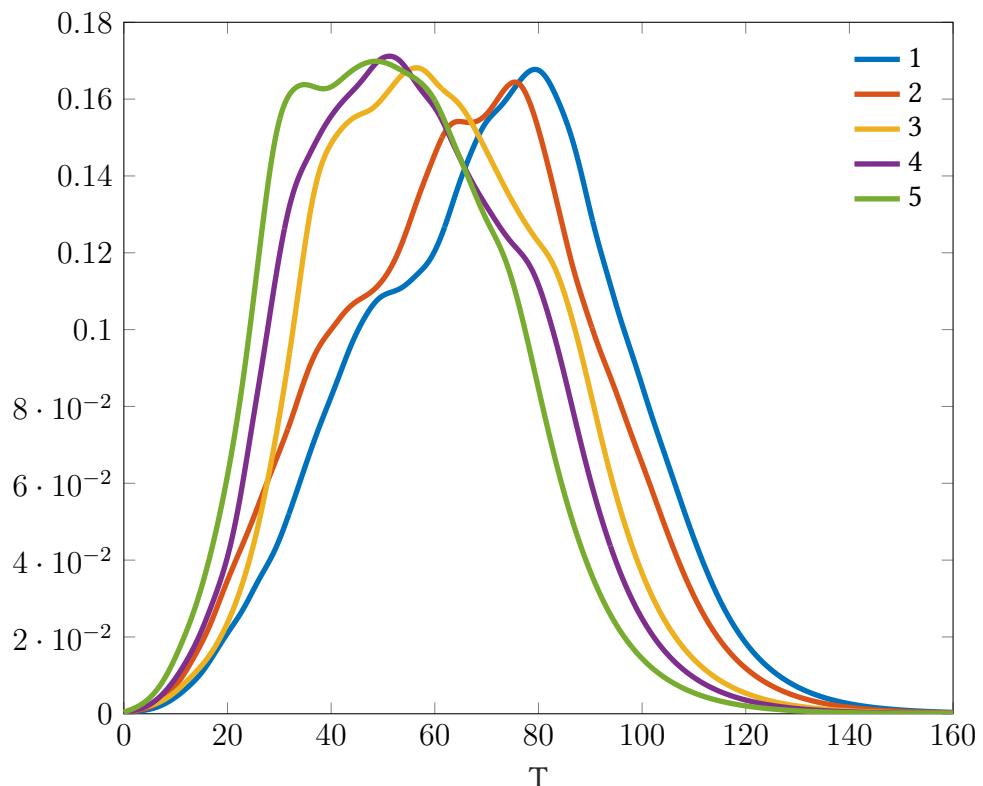


Figura 3.7: Grafico della prevalenza al variare del grado del nodo inizialmente infetto.

Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, cinque problemi di Cauchy ottenuti dal modello chiuso alle coppie.

Per le condizioni iniziali abbiamo preso ogni volta un nodo certamente infetto (con grado diverso e tutti gli altri nodi certamente sani).

Per la sperimentazione abbiamo utilizzato come parametri $\tau = 0.3$ e $\gamma = 0.1$.

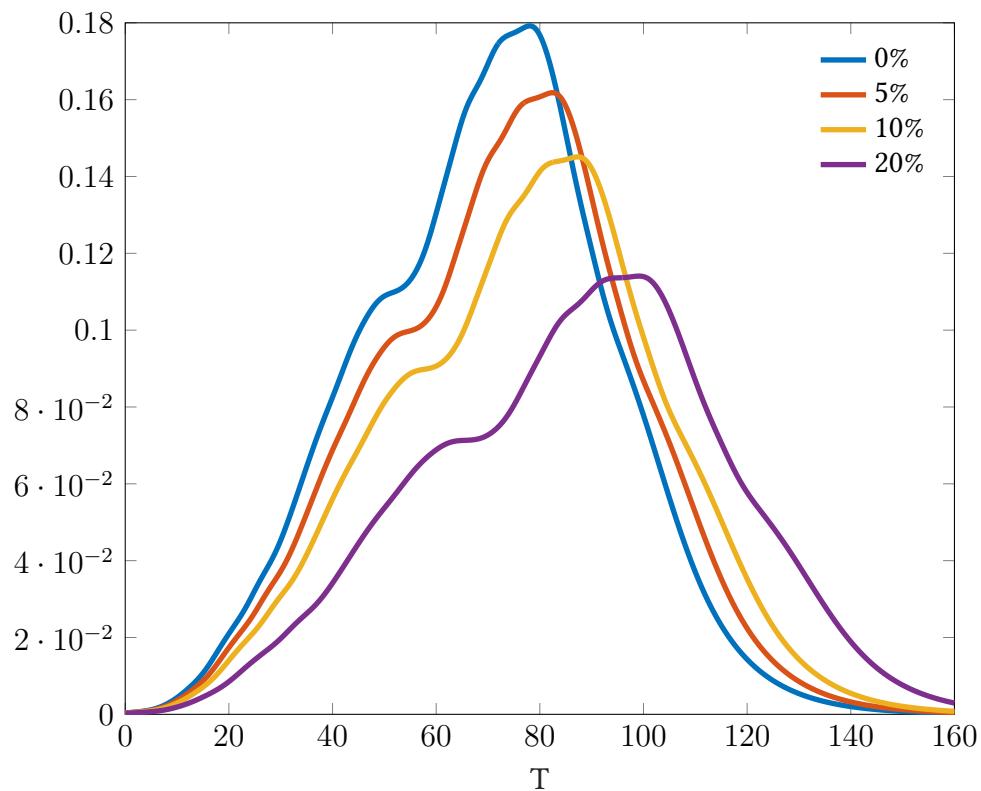


Figura 3.8: Grafico della prevalenza al variare della percentuale d'immunizzazione al tempo iniziale.

Per ottenere i grafici abbiamo risolto numericamente, usando MATLAB, quattro problemi di Cauchy ottenuti dal modello chiuso alle coppie utilizzando come parametri $\tau = 0.3$ e $\gamma = 0.1$.

Bibliografia

- [1] Suzanne Austin Alchon. *A pest in the land: new world epidemics in a global perspective*. UNM Press, 2003.
- [2] Ryuichi Ashino, Michihiro Nagase, and Rémi Vaillancourt. Behind and beyond the matlab ode suite. *Computers & Mathematics With Applications - COMPUT MATH APPL*, 40, 02 2000.
- [3] Daniel Bernoulli and Sally Blower. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Reviews in medical virology*, 14(5):275, 2004.
- [4] Timothy A. Davis and Yifan Hu. The University of Florida Sparse Matrix Collection. *ACM Trans. Math. Softw.*, 38(1), December 2011.
- [5] Matthew James Keeling. *The ecology and evolution of spatial host-parasite systems*. PhD thesis, University of Warwick, 1995.
- [6] William Ogilvy Kermack, A. G. McKendrick, and Gilbert Thomas Walker. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721, 1927.
- [7] Istvan Z. Kiss, Charles G. Morris, Fanni Sélley, Péter L. Simon, and Robert R. Wilkinson. Exact deterministic representation of Markovian *SIR* epidemics on networks with and without loops. *J. Math. Biol.*, 70(3):437–464, 2015.
- [8] John Denholm Lambert et al. *Numerical methods for ordinary differential systems*, volume 146. Wiley New York, 1991.
- [9] Randall LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems (Classics in Applied Mathematics Classics in Applied Mathemat)*. Society for Industrial and Applied Mathematics, USA, 2007.

- [10] World Health Organization et al. *Expert Committee on Healt Statistics [meeting held in Geneva from 29 September to 4 October 1958]: Sixth report*. World Health Organization, 1959.
- [11] Stanley A. Plotkin, Edward A. Mortimer, and Walter A. Orenstein. *Vaccines*. W.B. Saunders, Philadelphia, 1999.
- [12] GF Raggett. A stochastic model of the Eyam plague. *Journal of Applied Statistics*, 9(2):212–225, 1982.
- [13] K. J. Sharkey, I. Z. Kiss, R. R. Wilkinson, and P. L. Simon. Exact equations for SIR epidemics on tree graphs. *Bull. Math. Biol.*, 77(4):614–645, 2015.
- [14] Stefan Cunha Ujvari. *Storia delle epidemie*. Odoya, 2020.
- [15] Minus Van Baalen. Pair approximations for different spatial geometries. *The geometry of ecological interactions: simplifying spatial complexity*, 742:359–387, 2000.

Elenco delle figure

1.1	Analisi dei dati della peste di Eyam	10
1.2	Due esempi di grafi.	11
1.3	Grafo e matrice di adiacenza del toast.	12
2.1	Grafo con 3 nodi.	14
2.2	Sperimentazione in MATLAB relativo al grafo 2.1	17
2.3	Confronto tra modello esatto e chiuso alle coppie per 2.1	22
2.4	Confronto tra il modello esatto, chiuso alle coppie e alle triple per il grafo 2.9	25
2.5	Cut-vertex e decomposizione in sottoreti	27
2.6	Sperimentazione in MATLAB relativo al grafo 2.1	30
2.7	Errori assoluti relativi al grafo 2.1 tra modello esatto e cut-vertex	31
2.8	Errori assoluti relativi al grafo 2.1 tra modello esatto e chiuso alle coppie	32
2.9	Lollipop	33
2.10	Decomposizione della rete lollipop.	35
2.11	Sperimentazione in MATLAB relativo al grafo 2.9	38
2.12	Errori assoluti relativi al grafo 2.9 tra il modello esatto e con cut-vertex	39
2.13	Errori assoluti relativi al grafo 2.9 tra il modello esatto e quello chiuso alle coppie	40
3.1	Tableau del metodo generale di Runge-Kutta esplicito a m stadi	45
3.2	Tableau del metodo di Runge-Kutta implementato dalla funzione <code>ode45</code> di MATLAB	46
3.3	Regioni di assoluta stabilità.	47
3.4	Grafico della rete stradale del Minnesota.	48
3.5	Numero di time-step per <code>ode45</code> e <code>ode15s</code> sulla rete stradale del Minnesota al variare dei parametri	50
3.6	Stiff ratio alle condizioni iniziali sulla rete stradale del Minnesota al variare dei parametri	51
3.7	Grafico della prevalenza al variare del grado del nodo inizialmente infetto.	52

- 3.8 Grafico della prevalenza al variare della percentuale d'immunizzazione. 53

Elenco delle tabelle

- | | | |
|-----|---|----|
| 1.1 | Popolazione di deceduti e rimossi. | 9 |
| 1.2 | Numero di suscettibili ed infetti. Dati reali. | 9 |
| 1.3 | Numero di suscettibili ed infetti. Dati sperimentali. | 10 |

Indice

Introduzione	2
1 Epidemiologia matematica e reti	4
1.1 Il modello SIR	4
1.1.1 Un esempio	9
1.2 Reti complesse e grafi	11
2 Il modello SIR <i>bottom-up</i> su una rete	14
2.1 Un primo esempio	14
2.2 Chiusure	19
2.2.1 Chiusura al livello delle coppie	19
2.2.2 Chiusura al livello delle triple	22
2.2.3 Approccio generale alla chiusura	26
2.2.4 Un primo esempio. Continuazione	29
2.3 Un esempio completo	33
2.3.1 Modello esteso	33
2.3.2 Modello chiuso	35
3 Integrazione numerica di ODE	41
3.1 Metodi d'integrazione ad un passo	41
3.1.1 Eulero esplicito	42
3.1.2 Eulero implicito	42
3.1.3 Serie di Taylor	43
3.1.4 Runge-Kutta	43
3.2 Assoluta stabilità	45
3.3 Stiffness	47
3.4 Esempio su una rete reale	48
Bibliografia	55
Elenco delle figure	58
Elenco delle tabelle	58