

Bi-virus Epidemics over Large-scale Networks: Emergent Dynamics and Qualitative Analysis

Submitted in partial fulfillment of the requirements for
the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Augusto José R. A. Santos

B.S., Electrical and Computer Engineering, Instituto Superior Técnico,
Universidade de Lisboa
M.S., Electrical and Computer Engineering, Instituto Superior Técnico,
Universidade de Lisboa

Carnegie Mellon University
Pittsburgh, PA

September, 2014

I dedicate this Thesis to Benária –

*“...voa a gaivota com amor,
e é a razão pela qual vive,
foi o que aprendi,
com a melhor professora que já tive.*

Ivan Santos, Ode a Benária.”

– Thais, Ivan, and Agostinho, for tacitly teaching me that we can always find out a solution even when there is none.

Acknowledgements

I would like to thank:

Prof. José Moura: For his scientific mentorship commitment, overwhelmingly infectious enthusiasm, and humanistic concerns. While interacting with Prof. Moura I felt *silently my senses abandoning their defenses* to gently unconstrain my contrived *ab nihilo* tenets and utterly glimpse our scientific problems beyond trivial appearances. I feel inspired by his footprints of great scientific and humanistic deeds. I feel deeply grateful for everything. In sooth, if not for him, I shall say this Ph.D Thesis would have not come to terms. Professor José Moura was my advisor, and in writing those words, I am dawned by a blissful amazement, charm and honor.

Prof. João Xavier: For bringing forth my Ph.D application in the pristine stages of this venue. Also, I have profited from our discussions during this Ph.D.

Prof. Soummya Kar: For some quite helpful technical discussions.

The committee members: Prof. António Pacheco, Prof. Rohit Negi, Prof. Soummya Kar and Prof. João Pedro Gomes, for accepting to be part of the Ph.D Thesis defense and for reading this Ph.D Thesis.

Mr. Ivan Santos: For the craftsmanship manuscript illustrations provided throughout.

Mrs. Carolyn Patterson: For all the assistance provided, specially, by dedicating quite a while to assure an exceptional logistics to my office that helped me to overcome some asthma issues.

Mrs. Ana Mateus: For the bureaucratic assistance.

I acknowledge the excellent conditions I was provided with at Carnegie Mellon University.

This work was supported by the Fundação para a Ciência e a Tecnologia, Portugal, through a grant administered by the Carnegie Mellon Portugal Program. Partial support for this work was also provided by AFSOR and NSF.

Abstract

This Thesis studies bi-virus epidemics over large-scale networks. We set the rules of infection at the node level and determine the dynamical law governing the evolution of the fraction of infected nodes that naturally emerges in the limit of a large network. Then, we study the qualitative behavior of the fraction of infected nodes under the fluid limit dynamics to determine if and when the epidemics develops into a pandemic state, or leads to *natural selection*, with a dominant resilient virus strain. The Thesis is divided into two parts.

In the first part, we establish the fluid limit *macroscopic* dynamics of a multi-virus epidemics over classes of non-complete networks as the number of nodes grows large. We assume peer-to-peer random rules of infection in line with the Harris contact process. More specifically, The fluid limit ordinary differential equation dynamics is cast as the weak limit (in the number of nodes) of the fraction of infected nodes over time under the Skorokhod topology in the space of càdlàg sample paths. The *microscopic* model conforms to a Susceptible-Infectious-Susceptible model. A node is either infected or it is healthy and prone to receive infections. We prove the exact emergent dynamics for the class of complete-multipartite networks.

In the second part, we study the qualitative behavior of the fraction of infected nodes under the ordinary differential equation limiting dynamics obtained in the first part. Namely, we characterize the attractors – where the orbits of the differential equations converge to – and the corresponding basins of attraction. Due to the coupled nonlinear high-dimension nature of the mean field dynamics, there is no natural Lyapunov function to study their qualitative behavior. We establish their qualitative behavior, not by numerical simulations, but by bounding the epidemics dynamics for generic graph networks by the epidemics dynamics on two special regular networks – the inner and outer regular networks, for which we can carry out their qualitative analysis.

Contents

Acknowledgements	iii
Abstract	v
1 Introduction	1
1.1 Part I – Emergent Dynamics	1
1.2 Part II – Qualitative Analysis of the Emergent Dynamics	6
I Emergent Dynamics	11
2 Background	13
2.1 Basic Definitions in Probability Theory	14
2.2 Skorokhod Topology	18
2.3 Weak Convergence and Tightness	24
2.4 Further Modes of Convergence	27
2.5 Summary of the Chapter	30
3 Problem Formulation	31
3.1 Stochastic Network Model	31
3.2 Pathwise Representation	39
3.3 Concluding Remarks	44
4 Single-virus over Complete-bipartite Networks	45
4.1 Martingale Vanishes in Probability and Weakly	47
4.2 Weak convergence	52
4.3 Concluding Remarks	58
5 Multivirus over Complete-Multipartite Networks	61
5.1 Single-virus over Complete-Multipartite Networks	61
5.2 Multivirus Case	64

5.3	Concluding Remarks	67
6	Epidemics over Regular Networks: Preliminary Results	69
6.1	Weak Convergence Framework	70
6.2	Epidemics over Regular Networks: Tightness	72
6.3	Epidemics over Regular Networks with Shuffling: F.D.D Convergence	79
6.4	Numerical Simulations	83
6.5	Concluding Remarks	85
II	Qualitative Analysis of the Emergent Dynamics	87
7	Model Analysis	89
7.1	E-networks	91
7.2	Dynamical Flow in E-networks	94
7.3	Concluding Remarks	97
8	Qualitative Analysis: Two-node E-networks	99
8.1	Single Virus Epidemics	99
8.2	Bi-virus Epidemics	104
8.3	Concluding Remarks	109
9	Qualitative Analysis: General E-networks	111
9.1	Single Virus Epidemics	111
9.2	Bi-virus Epidemics	116
9.3	Concluding Remarks	123
10	Conclusion and Future Work	125
10.1	Part I: Concluding Summary	125
10.2	Part I: Future Work	126
10.3	Part II: Concluding Summary	127
10.4	Part II: Future Work	128
A	Appendix	129

List of Figures

1.1	Evolution of the fraction of nodes at state 1. The blue (noisy) curve represents the actual stochastic evolution with 10% of nodes initially at state 1 and the red (boldfaced) curve represents the solution to the limiting dynamics $\dot{y}(t) = -(\mu + \gamma)y(t) + \gamma$ with initial condition $y(0) = 0.1$. The equilibrium is given by $y_{eq} = \gamma/(\mu + \gamma) = 2/(1 + 2) = 2/3$. . .	3
1.2	Informal summary of the goal of this Ph.D Thesis and future goal. This Ph.D concentrates on Large-scale complete-multipartite epidemics systems with preliminary results in systems with sparse topology, namely, regular.	6
2.1	Illustration of an M -valued stochastic process $(Z(t))$ as a function from the realization space Ω to $M \subset \mathbb{R}^{[0,T]}$. The real-valued function $(Z(\omega_1, t))$, respectively $(Z(\omega_2, t))$, is represented by the blue (darker color), respectively red (lighter color) sample path. The values assumed by the random variable $Z(s)$ at ω_1 and ω_2 are given by $Z(\omega_1, s)$ and $Z(\omega_2, s)$. 16	16
2.2	Illustration of the sequence f_n and the limiting function f . The sequence f_n does not converge to f , with respect to the uniform topology, as $\ f_n - f\ = 1$ for all n	19
2.3	Informal illustration of a concentration of a sequence of absolutely continuous probability measures μ_n about a dirac measure δ_{x^*} . Each function f_i represents the density associated with each probability measure μ_i	25
3.1	Complete-multipartite network representation. Nodes from the same island cannot transmit the infection amongst themselves. Nodes from an island can transmit the virus to nodes in neighboring islands. For instance, any node from island 5 can infect any node from islands 1 and 2. Nodes from island 1 cannot infect nodes from island 4 and vice-versa, as the islands 1 and 4 are not neighbors.	32
3.2	Illustration of the complete-multipartite network $G^{\mathbf{N}}$, with $\mathbf{N} = (3, 6, 5, 4, 8)$ and underlying graph G	33

3.3	Illustration of an infection. The infected (dark colored) node at island 1 activates three exponentially distributed random <i>clocks</i> , each dedicated to one neighbor island. Assuming that the clock dedicated to island 2 rings first, after a time $T_{12k}^c \sim \text{Exp}(\gamma_{12}^k)$, a node from island 2 is randomly picked to be infected. The clocks for infection keep triggering while the node is infected. Also, the infected node in island 1 heals after a time $T_{1k}^h \sim \text{Exp}(\mu_1^k)$	35
4.1	Summary of the proof to Theorem 25.	58
5.1	The plots represent the numerical evolution of the fractions of <i>x</i> -infected (in blue/solid) and <i>y</i> -infected (in red/dashed) nodes at each island 1 and 2. The boldfaced curves represent the solution of the limiting vector differential equation of a bi-viral epidemics in a bipartite network.	68
6.1	Illustration of two line networks with the same number of infected nodes (blue/colored nodes). The bottom line network shows a clustered configuration and therefore the only way to spread the infection is via the two links of contact between infected and healthy nodes. On the other hand, the top line network has more contact links between infected and healthy nodes. Therefore, and despite both networks having the same number of infected nodes, the rate to increase the fraction of infected nodes is grater in the top line network.	71
6.2	Illustration of the peer-to-peer diffusion model over a regular network with degree $d = 4$. Once a node is infected, it activates d clocks – one per neighbor – associated with infection, plus one clock for healing. For simplicity, we omitted the healing clock in the figure. All clocks are exponentially distributed and independent. The clocks for infection bear a rate of γ , while the healing clocks have rate μ	73
6.3	Illustration of the diffusion with shuffling. We omit the infection and healing clocks for simplicity. After a node is infected, the micro-state is permuted to display the new state $X_{\text{sh}}^N(t)$. Note that the intermediate state just after infection is not accounted by $X_{\text{sh}}^N(t)$	79
6.4	Evolution of the peer-to-peer epidemics with shuffling over cycle networks with different number of nodes: 100, 1000 and 4000, respectively. The noisy blue/darker color curves represent the evolution of the epidemics under the shuffling model, and the smooth red/lighter color curves represent the solution to the ODE (6.16) with the same initial condition. We assumed $\mu = 1$ and $\gamma = 2$	84
6.5	Evolution of the peer-to-peer epidemics without shuffling over cycle networks with different number of nodes: 100, 1000 and 4000, respectively. The noisy blue/darker color curves represent the evolution of the peer-to-peer epidemics driven by the random clocks, and the smooth red/lighter color curves represent the solution to the ODE (6.16) with the same initial condition. We assumed $\mu = 1$ and $\gamma = 2$	84

6.6	Evolution of the peer-to-peer epidemics without shuffling over cycle networks with different number of nodes: 100, 1000 and 4000, respectively. The noisy blue/darker color curves represent the evolution of the peer-to-peer epidemics driven by the random clocks, and the smooth red/lighter color curves represent the solution to the ODE (6.16) with the same initial condition. We assumed $\mu = 1$ and $\gamma = 0.8$	84
7.1	Graphical illustration on the left (blue/lighter color) of an e-network (all γ_{ij} sketched are positive). On the right (black/darker color) lies its support graph counterpart.	91
8.1	Phase space of the augmented dynamical system with $(x_i(t), y_i(t)) := (x_i, y_i)(t, \mathbf{z}_0)$, and $(\tilde{x}_i(t), \tilde{y}_i(t)) := (x_i, y_i)(t, \mathbf{z}(0))$. The red (dashed) curves represent the components of $(\tilde{x}_1(t), x_1(t), \tilde{y}_1(t), y_1(t), \tilde{x}_2(t), x_2(t), \tilde{y}_2(t), y_2(t))$. The fact that the red (dashed) curve cannot escape the blue (dark) region in finite time is equivalent to $x_i(t) \leq \tilde{x}_i(t)$ and $y_i(t) \geq \tilde{y}_i(t)$ for all time t	106
8.2	Population of blue (lighter color) in the center bipartite network is lower and upper bounded by the corresponding populations in the left and right bipartite networks, respectively. The same goes, in the other way around, for the red (darker) population. The symmetric configurations in the left and right bipartite networks induce well-known solutions that bound the qualitative behavior of the middle configuration for all time $t \geq 0$	108
9.1	Illustration of an orbit of the augmented system. The set B is invariant, which implies that if $\mathbf{y}(0) \geq \tilde{\mathbf{y}}(0)$ then $\mathbf{y}(t, \mathbf{y}(0)) \geq \mathbf{y}(t, \tilde{\mathbf{y}}(0))$, $\forall t \geq 0$. Whenever a component hits one of the 3 sides of a triangle, it returns inwards the corresponding triangle.	114
9.2	Phase space of the augmented dynamical system with $(x_i(t), y_i(t)) := (x_i, y_i)(t, \mathbf{z}_0)$, and $(\tilde{x}_i(t), \tilde{y}_i(t)) := (x_i, y_i)(t, \mathbf{z}(0))$ for $i = 1, \dots, M$. The red (dashed) curves represent the components of $(\mathbf{x}(t), \mathbf{y}(t), \tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t))$. The fact that the red (dashed) curve cannot escape the blue (dark) region in finite time is equivalent to $x_i(t) \leq \tilde{x}_i(t)$ and $y_i(t) \geq \tilde{y}_i(t)$ for all time t and $i = 1, \dots, M$	118
10.1	Fluid limit epidemics evolution over a 3 node e-network. The rates of virus x and y from node i to node j are given by γ_{ij}^x and γ_{ij}^y , respectively. Each curve in blue/lighter color and red/darker color represent the evolution of each $(x_i(t, \mathbf{x}_0, \mathbf{y}_0))$ and $(y_i(t, \mathbf{x}_0, \mathbf{y}_0))$, respectively (one curve per index i).	128

Chapter 1

Introduction

In this Thesis, we study bi-virus epidemics over large-scale networks. This Thesis is divided into two parts. In the first part, we determine the emergent dynamics, as an ordinary differential equation (ODE), describing the fraction of infected nodes over time of a bi-virus epidemics under a large number of interacting agents. In the second part, we study the qualitative behavior of such ODE's. The two parts comprising this Thesis may be read independently.

1.1 Part I – Emergent Dynamics

Many complex networked dynamical systems exhibit *emergent behavior* – well-structured macroscopic dynamics induced by simple, possibly random, local rules among interacting agents. Flocks of birds, ant colonies, beehives, brain neural networks, invasive tumor growth, and epidemics are all examples of large scale systems of interacting agents displaying complex adaptive functional behaviors. Under appropriate conditions, a flock of birds reaches consensus on its cruise velocity while each bird probes only its nearest neighbors dynamics without a preferred leader in the flock (refer to [1]). This gives rise to synchronized flocking flying formations. Ant colonies can design optimal trails to access sources of food even though no ant bears the cognitive ability to shape up the colony to its global behavior. Roughly, each scout-ant wanders around randomly tracking the leftover pheromone released by its scout peers. Reference [2] establishes the emergent dynamics of an idealized stochastic network model for ant colonies as the fluid limit dynamics of the network model (as the colony grows large). Seizure is an intricate outcome of the complex neural network dynamics of the brain. Reference [3] presents an overview of graphical dynamical models that have been applied to better understand the nature of seizures and bridge the microscopical electrical activity in the brain with the clinical observations of the phenomenon.

The challenge in studying such large scale dynamical systems lies in their **high-dimensionality** plus *sparse coupling* among the agents via their interactions. Together, these induce emergent behavior and, under this setting, comprehending the dynamics of individual agents does not shed much light into the holistic behavior of the large-scale networked system.

If we simply drop the coupling, then any weak law of large numbers will reflect the average behavior of each individual agent rather than an emergent global cooperative behavior. For instance, consider N **independent** – thus, uncoupled – agents whose state-vector

$$\mathbf{X}^N(t) := (X_1(t), \dots, X_N(t))$$

evolves as a jump Markov process over the state space $\mathcal{S}^N := \{0, 1\}^N$ as follows: once agent i is at state 0, $X_i^N(t) = 0$, the time to flip to 1 is independent of the state of the remaining agents, and it is exponentially distributed with rate γ ; similarly, once agent i is at state 1, $X_i^N(t) = 1$, the corresponding time to flip from 1 to 0 is exponential with rate μ . In this case, the state of each agent i evolves as a jump Markov process and, moreover, any *macroscopic* state construct

$$(f(X_1(t), \dots, X_N(t))),$$

is a Markov jump process, when

$$f : \{0, 1\}^N \rightarrow \mathbb{R}^M$$

bears appropriate measurability properties (we skip the details here) and we assume $M \ll N$. As an example, the fraction of agents at state 1,

$$\bar{Y}^N(t) = f(X_1(t), \dots, X_N(t)) = \sum_{i=1}^N X_i(t)/N,$$

is Markov. Even for large N , due to the independence assumption, a qualitative analysis over $(X^N(t))$ or

$$(f(X_1(t), \dots, X_N(t)))$$

becomes tractable; but this is achieved at the expense of loosing the coupling structure. Figure 1.1 illustrates that, as the number of nodes grows large, the limiting dynamics of $(\bar{Y}^N(t))$ is given by the linear first order ODE

$$\dot{y}(t) = -(\mu + \gamma)y(t) + \gamma,$$

as can be formally shown as a Corollary to the results developed in Chapter 4. Note that the global attractor equilibrium

$$y_{eq} = \frac{\gamma}{\gamma + \mu}$$

of the limit (in the number of nodes) dynamics also represents the expected limit (in time) fraction of time that any node i spends at state 1. In words, local information promptly maps onto the global behavior of the system and vice-versa, i.e., comprehending *the parts* is enough to plot *the whole*. This latter property should not be expected in coupled systems in general, where the topology of interactions plays a major role in the

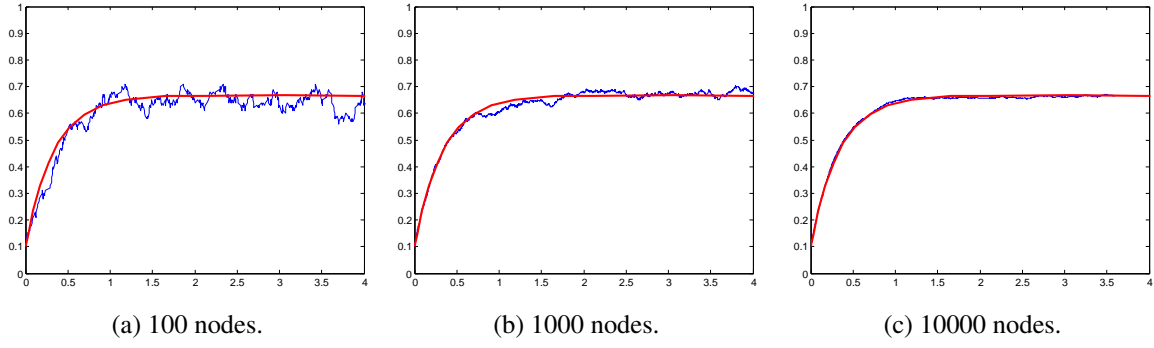


Figure 1.1: Evolution of the fraction of nodes at state 1. The blue (noisy) curve represents the actual stochastic evolution with 10% of nodes initially at state 1 and the red (boldfaced) curve represents the solution to the limiting dynamics $\dot{y}(t) = -(\mu + \gamma)y(t) + \gamma$ with initial condition $y(0) = 0.1$. The equilibrium is given by $y_{eq} = \gamma/(\mu + \gamma) = 2/(1 + 2) = 2/3$.

ultimate macroscopic dynamics of the system.

On the other side of the spectrum, under a dense coupling among the agents, e.g., assuming a *complete* topology of interactions – any agent evenly affects any other agent in the cloud, e.g., an agent i switches to state 1 with a rate that is proportional to the number of agents at state 1 – low-dimensional *macroscopic* state-variables may still be Markov, even though the state of each individual agent is no longer Markov. The fraction of nodes at state 1,

$$f(X_1(t), \dots, X_N(t)) = \sum_{i=1}^N X_i(t)/N,$$

is Markov. The Markov property at the macroscopic level leads to a dynamical system structure that is often tractable. The emergent behavior is framed as the fluid limit dynamics of a global state variable

$$(\mathbf{Y}(t)) := (f(X_1(t), \dots, X_N(t)))$$

of interest. For example, reference [4] considers a multiclass flow of packets over a complete network with finite capacity nodes. It defines the macroscopic state variable

$$(\mathbf{Y}^N(t)) = (Y_1^N(t), \dots, Y_L^N(t))$$

that collects the fraction of nodes $Y_i^N(t)$ with a particular distribution i of packets over the different classes. Reference [4] proves that the empirical distribution $(\mathbf{Y}^N(t))$ converges weakly, with respect to the Skorokhod topology on the space of sample paths, to the solution of a vector ordinary differential equation.

In real life complex networked dynamical systems, as the ones referred to in the first paragraph of this Section, the underlying nature of the topology of interactions is neither complete nor uncoupled. When the agents are coupled under an arbitrary sparse topology, the evolution of the macroscopic state variables is intricately tied to the high-dimensional *microscopic* state $(\mathbf{X}^N(t))$ of the system, that is, neither the state

$X_i(t)$ of each agent i is Markov nor the macroscopic low-dimensional states

$$(f(X_1(t), \dots, X_N(t)))$$

are Markov, in general. Determining the emergent dynamics or, in other words, establishing the functional weak law of large numbers under an arbitrary coupling topology of the agents is challenging. Reference [5] proposes to consider the impact of the topology on the diffusion of a virus in a network, but to overcome the coupling difficulty that arises with non complete networks, the authors in [5] replace the exact transition rates of the microstate process $(\mathbf{X}(t))$ by their average to establish their N -intertwined model. If the states of the nodes were independent processes (a very strong assumption), the resulting N -intertwined model would be an exact model to describe the dynamics of the likelihood of infection of each node as pointed out by the authors.

In part I of the Thesis, we go beyond the complete network model to establish the **exact** fluid limit dynamics of a multi-virus epidemics over the class of complete-multipartite networks; we do not impose simplifying assumptions or approximations. This establishes the emergent dynamics of the epidemics system under a topology whose connectivity is neither completely dense nor uncoupled. Moreover, and as we refer next, large-scale bi-virus epidemics over complete-multipartite networks is far from being an artificial problem and finds motivation in real life epidemics.

Multipartite structures find motivation in real life epidemics. A network is multipartite if the nodes can be partitioned into groups (or partites) where within each group nodes are not connected, but there may be connections between nodes of different groups. For example, the dengue disease is spread between humans and *Aedes Aegypti* mosquitoes according to a bipartite structure, refer to [6]. Indeed, a healthy human gets infected once bitten by an infected mosquito. A healthy mosquito gets infected once it bites an infected human being. The disease does not spread among humans nor does it spread among mosquitoes, i.e., only cross-species infection is allowed. Therefore, the underlying network of infections has two partites connected together: one comprised of mosquitoes and the other of human beings.

Bi-virus dynamics epidemics model the virus pandemic versus vaccination. In many real life settings, once an individual is infected, it cannot be vaccinated¹. If the individual is vaccinated, it is not prone (ideally) to be infected. Thus, the population can be partitioned into these disjoint sets: infected individuals; not infected and not vaccinated (susceptible); and vaccinated (and thus, not susceptible) individuals. We refer to this constraint as the *exclusion principle*, which is an important assumption embedded in the stochastic network microscopic model of diffusion studied in Part I. The exclusion principle couples the dynamics of virus spread and vaccination and leads to a competition for space in the networked system between the virus and vaccines. The following assertions suggest that vaccines may spread across the agents of a network similarly to a virus:

¹In fact, a vaccine is comprised of biological micro-agents that resemble the virus and aim to train or stimulate the immune system against the virus, which turns out to be ineffective if the individual is already infected. For more details, refer to [7].

- While a virus spreads across the population according to the underlying network of contacts, references [8], [9] suggest that vaccines obey similar dynamics: the social network of influences plays a major role on the decision of parents vaccinating their children.

- Recent research results in malaria vaccines (e.g., [10]) suggest that infection inhibition via mosquitoes bites is possible. In other words, the vaccines would spread across the population via similar dynamics as the main *virus*². In this case, one recovers the bi-virus epidemics setting in the network where the vaccine may be cast as another virus strain spreading similarly to the main virus with mutual inhibition between the strains.

Therefore, the virus-vaccine problem may be abstracted as two strains of virus competing for space – in line with the exclusion principle – in the network. The multipartite structure is a natural generalization of cross-species spread for dengue like diseases across more than two species. The weak limit (as the number of nodes goes to infinity) formalizes the real life large scale epidemics. The emergent behavior is the fluid limit dynamics of the fraction of infected nodes over time. Namely, we show that, when the number of agents goes to infinity in a sense to be defined in Chapter 4, the fraction of infected agents converges weakly to the solution of a set of nonlinear ordinary differential equations, which describes the dynamics of spread of virus and vaccines across the large-scale complete-multipartite network.

In Chapter 6, we present our current results on epidemics over regular networks, to obtain the weak functional law of large numbers associated with the evolution of the fraction of infected nodes over time as the regular network grows large. These results are a first step into this direction of deriving methodologies to explain formally the emergent dynamics on sparser topologies. As referred before, in this setting the macroscopic state-variables are not Markov and new approaches are required that can bypass the lack of classical dynamical systems tools in a non-Markov framework. Fig. 1.2 summarizes informally our ideas.

Contribution of Part I: We formally establish, from the peer-to-peer rules of infection, the limiting dynamics of multi-virus epidemics over the class of complete-multipartite networks. These results have been partially reported in reference [11]. Namely, we prove that the càdlàg $(M \times K)$ -matrix process

$$\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right) = \left(\bar{Y}_{ik}^{\mathbf{N}}(t)\right)_{ik}$$

collecting the fraction $\left(\bar{Y}_{ik}^{\mathbf{N}}(t)\right)$ of nodes infected with virus k (a.k.a. k -infected nodes)

$$k \in \{1, \dots, K\}$$

at each island i ,

$$i \in \{1, \dots, M\},$$

²Technically, malaria is not caused by a virus, but by a protozoan. For all our purposes, the relevant feature lies in the dynamics of spread and not in the nature of the viral agent.

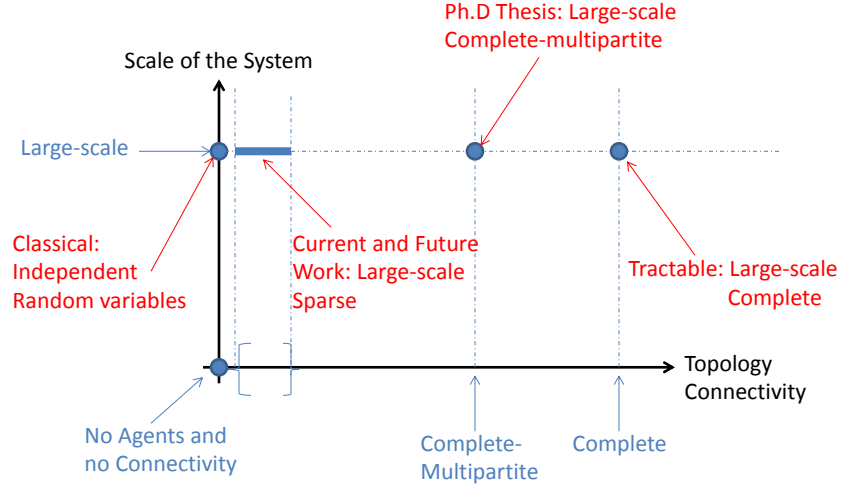


Figure 1.2: Informal summary of the goal of this Ph.D Thesis and future goal. This Ph.D concentrates on Large-scale complete-multipartite epidemics systems with preliminary results in systems with sparse topology, namely, regular.

converges weakly, with respect to the Skorokhod topology on the space of càdlàg sample paths, to the solution of an ODE. **Outline of Part I:** In Chapter 2, we introduce the main background on Skorokhod topology and weak convergence. In Chapter 3, we set the problem formulation, defining the peer-to-peer stochastic network model underlying the microscopic dynamics of the diffusion of the strains of virus, and we provide a pathwise characterization of the fraction of infected nodes process $(\bar{\mathbf{Y}}^N(t))$. In Chapter 4, we establish the fluid limit dynamics for a single virus spread over a complete-bipartite network, namely, we prove that the vector process

$$(\bar{\mathbf{Y}}^N(t)) = (\bar{Y}_1^N(t), \bar{Y}_2^N(t))$$

collecting the fraction of infected nodes at each island converges weakly to the solution of an ordinary differential equation. In Chapter 5, we extend the proof presented in Chapter 4 to the general case of a complete-multipartite network under multi-virus spread. In Chapter 6, we present preliminary results on the emergent behavior of epidemics over regular networks.

1.2 Part II – Qualitative Analysis of the Emergent Dynamics

The goal of Part II is to study the qualitative behavior of the limiting dynamics of bi-virus epidemics in networks obtained in the Part I, namely, the dynamics of the nonlinear coupled ordinary differential equations

$$\dot{x}_i(t) = \left(\sum_{j \in V} \gamma^x(j, i) x_j(t) \right) (1 - x_i(t) - y_i(t)) - x_i(t) \quad (1.1)$$

$$\dot{y}_i(t) = \left(\sum_{j \in V} \gamma^y(j, i) y_j(t) \right) (1 - x_i(t) - y_i(t)) - y_i(t). \quad (1.2)$$

As referred in the previous Section, the analytic study of this dynamical system can inform the regulatory policies of vaccination to preventing the persistence of an epidemics over large-scale complete-multipartite networks.

While the ODE (1.1)-(1.2) provides with the weak limit evolution of the fraction of infected nodes per strain in the limit of a large complete-multipartite network, the single-virus nonlinear ODE,

$$\dot{y}_i(t) = \left(\sum_{j \in V} \gamma(j, i) y_j(t) \right) (1 - y_i(t)) - y_i(t) \quad (1.3)$$

may conform to a model for the likelihood $y_i(t)$ of infection of a node i (instead of island i) at time t in a generic (not necessarily complete-multipartite) network G according to the **approximate** meanfield N-intertwined model motivated in [5] for single-virus. Similar ODE equations for single virus are also presented in [12]; this reference motivates them through full-mixing arguments. Therefore, these differential equations arise not only as the exact fluid limit dynamics over complete-multipartite networks as we show in this Thesis, see also [11], but also as describing the epidemics dynamics over general networks or as modeling the dynamics of epidemics in heterogenous interacting populations of individuals. In this sense, the analysis of (1.1)-(1.2) or (1.3) can be either interpreted as the **actual** evolution of the fraction of infected nodes per island in a large-scale complete-multipartite network G or as **modeling** the evolution of the likelihood of infection of nodes in a network \hat{G} whose topology abstracts the super-topology of G , in that the nodes and edges in \hat{G} play the role of super-nodes and inter-island super-edges in G , respectively (refer to Figure 3.2 in Chapter 3). In either case, the bi-virus dynamics (1.1)-(1.2) are a natural model to study the evolution of infected *versus* vaccinated individuals.

There is a vast body of literature dedicated to the study of single-virus epidemics over networks. References [13], [12] provide a comprehensive review of the epidemics literature. All these references focus mostly on studying the equilibria of imposed idealized differential equation models with a few providing heuristics on the *transitory* behavior of such dynamical systems. Reference [5] provides several bounds on the equilibria of a single virus epidemics. These references focus on finding the equilibria and do not study rigorously their qualitative behavior, i.e., do not characterize their attractors and basins of attraction of the dynamical system. Reference [12] also attempts to briefly study the qualitative behavior of some epidemics models resorting to heuristic arguments, e.g., linearizing the ODE's when the degree of infection is *small*. We do not evoke such approximation in this Thesis. Reference [14] does provide a rigorous qualitative analysis for the single virus epidemics dynamics – equation (1.3) – based on Lyapunov stability theory. As we explain in Chapter 7, the results in [14] for the single virus dynamics do not apply to the bi-virus case, and we need to develop a new methodology. Reference [15] studies via numerical simulations the epidemics

over big cities resorting to dynamical systems similar to the one in equation (1.3), to model the mixing of different interacting groups of individuals.

We stress that Part II studies the qualitative analysis of the bi-virus fluid limit dynamics. We do not focus on characterizing purely the equilibria (or properties of the equilibria), but we devise tools that guarantee to where the orbits of the dynamical system accumulate to. We are concerned with finding the attractors and corresponding basins of attraction of the dynamics. Namely, we show that under appropriate conditions on the parameters of the ODE's governing the bi-virus fluid limit epidemics the orbits accumulate about an attractor \mathcal{A} that we characterize. The attractor underlies a natural selection phenomenon taking course in the long run: the weaker strains die out whereas the stronger strain survives with a degree of infection that lies above a threshold that we determine.

The qualitative analysis of dynamical systems is usually studied through Lyapunov stability or empirically by numerical simulations. The bi-virus epidemics dynamical system defies the use of Lyapunov methods. Our approach lies rather in exploring special monotone properties of this dynamical system, namely, an appropriate bound on the initial conditions is preserved by the flow of the epidemics dynamical systems for all time t . For example, in the case of single virus dynamics, equation (1.3), if we let

$$(\mathbf{y}(t, \mathbf{y}(0)))$$

be the solution with initial condition $\mathbf{y}(0)$, and if

$$\mathbf{y}_0 \leq \mathbf{y}(0) \in [0, 1]^N$$

is another initial condition then,

$$\mathbf{y}(t, \mathbf{y}_0) \leq \mathbf{y}(t, \mathbf{y}(0))$$

for all $t, t \geq 0$. In the two virus case, if we have a *skewed* bound, i.e., if virus x is started with

$$\mathbf{x}_0 \leq \mathbf{x}(0)$$

and virus y is started with

$$\mathbf{y}_0 \geq \mathbf{y}(0)$$

then,

$$\mathbf{x}(t, (\mathbf{x}_0, \mathbf{y}_0)) \leq \mathbf{x}(t, (\mathbf{x}(0), \mathbf{y}(0)))$$

and

$$\mathbf{y}(t, (\mathbf{x}_0, \mathbf{y}_0)) \geq \mathbf{y}(t, (\mathbf{x}(0), \mathbf{y}(0)))$$

for all time $t \geq 0$. We will see that similar bounds are preserved when we refer to networks that are partially ordered as we study in Chapter 7. Namely, we will show that these dynamics over a network G are lower

(respectively, upper) bounded by the dynamics over a sub-network (respectively, supper-network) of G as will be rigorously stated in the corresponding Chapters within Part II.

Contribution of Part II: we provide techniques for the rigorous qualitative analysis of epidemics dynamical systems that overcome the lack of natural Lyapunov tools. Reference [16] describes some of the results of Part II. See also [17, 18].

Outline of Part II: Chapter 7 sets up the main notation, important inequalities between networks and the ODE's to be studied throughout Part II. Chapter 8 establishes the qualitative behavior of the ODE's (1.1)-(1.2) for the bi-viral epidemics over two-node networks alongside with the main results. Chapter 9 extends the analysis in Chapter 8 to the general network case. Proofs to Theorems 55, 56, 61 and 62 can be found in the Appendix.

Part I

Emergent Dynamics

Chapter 2

Background

The goal of this Chapter is to structure the background notation and briefly introduce some of the central technical objects needed in Part I, namely, Skorokhod topology and the weak convergence and tightness with respect to the Skorokhod topology. We emphasize that most of the results described in this Chapter are standard in [19, 20, 21] and do not entail our own contribution. Other than that, the presentation is our own and, as we explain in Section 2.2, we could not find in the literature an explicit proof for the measurability of some important operators – that often appear in Part I of this Thesis – over the Skorokhod space, and therefore, we provide with our own proof. Even though the reader can skip this Chapter, if already acquainted with these subjects, some of the main notation to be used in the rest of Part I is introduced in this Chapter. This is not an exhaustive account on the fundamentals of stochastic processes. For instance, concepts as martingale, Markov process and others in the foundations of probability theory do appear frequently in the text, but we assume the reader is familiar with them and we do not refer to them in this Chapter.

In this Thesis, we study sample mean processes such as the evolution of the fraction of infected nodes

$$\bar{Y}^N(t) = \frac{\sum_{i=1}^N X_i(t)}{N}$$

over time $t, t \geq 0$, in a network with N nodes with the state $X_i(t) = 1$, if node i is infected or $X_i(t) = 0$, if node i is healthy. Such process changes its state once an event is triggered: a new node is either infected or healed. For instance, if at time t , a new node is infected, then

$$\bar{Y}^N(t) = \bar{Y}^N(t-) + \frac{1}{N},$$

where we defined

$$\bar{Y}^N(t-) = \lim_{s \uparrow t} \bar{Y}^N(s).$$

Therefore, the process $(\bar{Y}^N(t))$ evolves through jumps. We can assume that $(\bar{Y}^N(t))$ is a càdlàg – acronym for the French *continue à droite limite à gauche* – process, that we define in Section 2.2. As

our goal is to establish emergent dynamics as the network grows large (or N goes to infinite), we must come to terms with the formalism of convergence, i.e., we shall set an underlying topology on the set of càdlàg paths. A usual topology – and the one we will assume in this Thesis – in the set of càdlàg sample paths is the Skorokhod topology that we describe in Section 2.2. Under an appropriate metric, the Skorokhod topological space of càdlàg functions conforms to a complete, separable, metric space – Polish space – thus enjoying several results from functional analysis, that are the building blocks of our results. In Section 2.1, we set up basic definitions of probability theory and notations that are relevant for this Part I. In Section 2.2, we define càdlàg functions and the Skorokhod topology. We also present examples of measurable operators on Skorokhod topological spaces. In Section 2.3, we characterize weak convergence and make sense of the statement: “*converges weakly with respect to the Skorokhod topology*”. In Section 2.4, we introduce further modes of convergence that inexorably pave the way to ultimately establishing weak convergence of càdlàg processes. In Section 2.5, we summarize the Chapter.

2.1 Basic Definitions in Probability Theory

In this Section, we recall some of the basic definitions from probability theory and set part of the notation to be used in the following Sections and the latter Chapters of Part I. We call a class Σ of subsets of a set M a σ -algebra whenever the following conditions hold

- $M \in \Sigma$;
- If $A_i \in \Sigma$ for all $i \in I$, then $\bigcup_{i \in I} A_i \in \Sigma$, for any countable index set I ;
- If $A \in \Sigma$, then $A^c \in \Sigma$,

where, we defined $A^c := M \setminus A$. In words, Σ is a σ -algebra if it is closed under countably many set operations. The tuple (M, Σ) is called a measurable space. For notational simplicity we may refer to the measurable space (M, Σ) simply as M . Let \mathcal{C} be a class of subsets of M and let Γ index the family of all σ -algebras Σ_α containing \mathcal{C} . We define the σ -algebra generated by \mathcal{C} as

$$\sigma\{\mathcal{C}\} = \bigcap_{\alpha \in \Gamma} \Sigma_\alpha,$$

the smallest σ -algebra containing \mathcal{C} . Provided that M is a topological space, we refer to the Borel σ -algebra $\mathcal{B}(M)$ associated with M , as the smallest σ -algebra containing the class of open sets of M . The pair $(M, \mathcal{B}(M))$ is called a Borel measurable space. In this Section, we assume that M is a topological space. For the sake of notational simplicity, we may refer to M as a set, or as a topological space, or as the Borel measurable space $(M, \mathcal{B}(M))$.

Definition 1 (Measurable Function). *Let (S, \mathcal{S}) be a measurable space. A function or operator*

$$\mathcal{G} : S \longrightarrow M$$

is measurable, if

$$\mathcal{G}^{-1}(B) := \{s \in S : \mathcal{G}(s) \in B\} \in \mathcal{S},$$

for all $B \in \mathcal{B}(M)$.

Let (Ω, \mathcal{F}) be a measurable space, and assume that $M \subset \mathbb{R}^{[0,T]}$, where $\mathbb{R}^{[0,T]}$ is the set of real valued functions defined over the compact interval $[0, T]$. The specific topology and hence Borel σ -algebra $\mathcal{B}(M)$ adopted in M is not important now. We define an M -valued stochastic process $(Z(t))$ as a function from the sample space Ω to the sample path space M

$$\begin{aligned} (Z(t)) &: \Omega \longrightarrow M \\ \omega \in \Omega &\longmapsto (Z(\omega, t)), \end{aligned}$$

that is measurable, where $(Z(\omega, t)) \in M$ is a sample path associated with the realization ω . In this Thesis, we will also refer to $Z(t)$ (without brackets), to mean the process $(Z(t))$ at the specific time t , i.e., to the real valued function

$$\begin{aligned} Z(t) &: \Omega \longrightarrow \mathbb{R} \\ \omega \in \Omega &\longmapsto Z(\omega, t). \end{aligned}$$

If the measurable space M is such that the canonical projection map over M

$$\begin{aligned} \pi_t &: M \longrightarrow \mathbb{R} \\ x &\longmapsto x(t) \end{aligned} \tag{2.1}$$

is measurable, then $Z(t)$ is measurable, since

$$Z(t) = \pi_t \circ (Z(t)), \tag{2.2}$$

and we refer to $Z(t)$ as a random variable, with ‘ \circ ’ representing the composition of functions.

• **Note:** It is important to keep in mind the differences among $(Z(t))$, $Z(t)$, $(Z(\omega, t))$, $Z(\omega, t)$. Remark that $(Z(t))$ is a stochastic process, i.e., a measurable function from the realization space Ω to the sample path space M ; $Z(t)$ is a real valued function over Ω and, under mild conditions on M , is a random variable; $(Z(\omega, t))$ is a sample path associated with the realization $\omega \in \Omega$; and $Z(\omega, t)$ is the real value assumed by the process $(Z(t))$ at the realization $\omega \in \Omega$ at time t , $t \geq 0$. Refer to Fig. 2.1 for a summary on these notations.

Let Σ_1 and Σ_2 be two σ -algebras, and define

$$\Sigma_1 \vee \Sigma_2 := \sigma(\Sigma_1 \cup \Sigma_2)$$

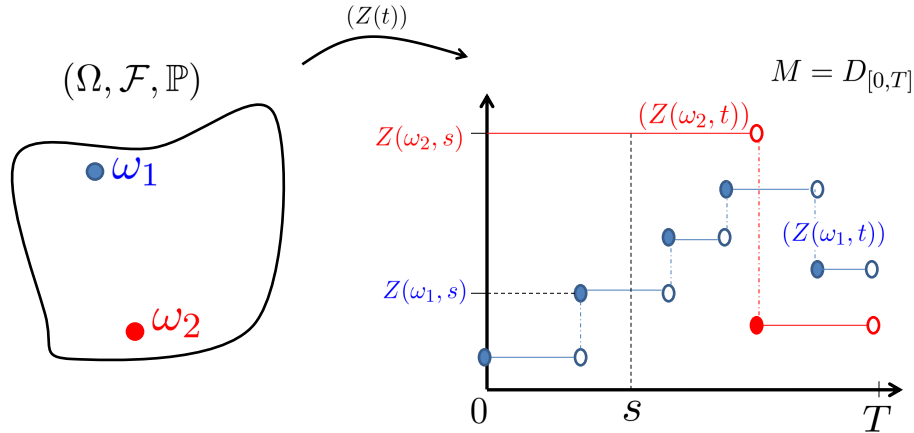


Figure 2.1: Illustration of an M -valued stochastic process $(Z(t))$ as a function from the realization space Ω to $M \subset \mathbb{R}^{[0,T]}$. The real-valued function $(Z(\omega_1, t))$, respectively $(Z(\omega_2, t))$, is represented by the blue (darker color), respectively red (lighter color) sample path. The values assumed by the random variable $Z(s)$ at ω_1 and ω_2 are given by $Z(\omega_1, s)$ and $Z(\omega_2, s)$.

as the smallest σ -algebra that contains Σ_1 and Σ_2 . Also, let (S, \mathcal{S}) be a measurable space and consider the function $f : C \rightarrow S$. We define

$$f^{-1}(\Sigma) := \{f^{-1}(B) : B \in \Sigma\},$$

as the σ -algebra generated by f on C . Now, we define a concept that often appears in this Thesis.

Definition 2 (Generated σ -algebra). Let (S, \mathcal{S}) be a measurable space and $\{f_t\}_{t \in \mathbb{T}}$ be a family of functions

$$f_t : \Omega \longrightarrow S$$

indexed by an arbitrary set \mathbb{T} . We define

$$\sigma\{f_t : t \in \mathbb{T}\} := \bigvee_{t \in \mathbb{T}} f_t^{-1}(\mathcal{S})$$

as the σ -algebra generated by the family $\{f_t\}_{t \in \mathbb{T}}$.

According to the previous definition, if $(Z(t))$ is a stochastic process, we define:

$$\mathcal{F}_{t'}^Z := \sigma\{Z(s) : 0 \leq s \leq t'\}$$

as the natural filtration associated with the process $(Z(t))$ up to time t' . Also, if $M = \mathbb{R}^{[0,T]}$, then the σ -algebra generated by the canonical maps π_t defined in equation (2.1)

$$\sigma \{ \pi_t : t \in [0, T] \} =: \bigotimes_{t \in [0, T]} \mathcal{B}(\mathbb{R}),$$

is often called the product σ -algebra on $\mathbb{R}^{[0,T]}$. If $M \subset \mathbb{R}^{[0,T]}$, then

$$\sigma \{ \pi_t : t \in [0, T] \}$$

is called the product σ -algebra on $\mathbb{R}^{[0,T]}$ relativized to $M \subset \mathbb{R}^{[0,T]}$. Note that, $Z(t)$ is a random variable for all t if

$$\sigma \{ \pi_t : t \in [0, T] \} \subset \mathcal{B}(M).$$

We call the function

$$\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$$

a probability measure on the measurable space (Ω, \mathcal{F}) , if it fulfills the following conditions

- $\mathbb{P}(\emptyset) = 0$;
- $\mathbb{P}(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mathbb{P}(A_i)$; (σ -additivity)
- $\mathbb{P}(\Omega) = 1$,

for any pair-wise disjoint family of measurable sets $A_i \in \mathcal{F}$ indexed by a countable set I . We refer to the triple $(\Omega, \mathcal{F}, \mathbb{P})$ as a probability space. Any measurable function f from a probability space Ω to a measurable space M induces a natural probability on M , that we may also refer to as the distribution of f on M , and we define it next.

Definition 3 (Probability Push-forward). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and*

$$\mathcal{G} : \Omega \longrightarrow M$$

be a measurable function. We define

$$\begin{aligned} \mathbb{P}_{*\mathcal{G}} : \mathcal{B}(M) &\longrightarrow [0, 1] \\ B \in \mathcal{B}(M) &\longmapsto \mathbb{P}(\mathcal{G}^{-1}(B)) \end{aligned},$$

as the probability push-forwarded by \mathcal{G} on M . The triple $(M, \mathcal{B}(M), \mathbb{P}_{\mathcal{G}})$ conforms to a probability space, and $\mathbb{P}_{*\mathcal{G}}$ may be also referred as the probability distribution of \mathcal{G} .*

If $\mathcal{G} = (Z(t))$ is a stochastic process, we write its push-forward probability simply as \mathbb{P}_Z . In the next Section, we define the Skorokhod topology.

2.2 Skorokhod Topology

In this Section, we describe the Skorokhod topology in the set of càdlàg functions; we briefly compare it with the uniform topology in the set of continuous functions; and we present some examples of important operators on the Skorokhod space that are measurable. We restrict the discussion throughout to real valued functions over the compact interval $[0, T]$. All the definitions and Theorems that follow are naturally extended to \mathbb{R}^n -valued functions over the interval $[0, T]$. Next, we define càdlàg functions.

Definition 4 (Càdlàg Function; [19]). *We say that $f : [0, T] \rightarrow S \subset \mathbb{R}$ is càdlàg if both of the following properties hold*

- $\lim_{s \downarrow t} f(s) = f(t)$
- $\limsup_{s \uparrow t} f(s) = \liminf_{s \uparrow t} f(s),$

for all $t \in (0, T)$, i.e., f is right-continuous with left limit (RCLL). Further,

- $\lim_{s \downarrow 0} f(s) = f(0)$
- $\limsup_{s \uparrow T} f(s) = \liminf_{s \uparrow T} f(s).$

As we would like to obtain results of convergence for càdlàg processes, we need to specify a topology on the set of càdlàg functions. The next example shows that the uniform topology, i.e., the topology induced by the norm

$$\|f\| = \sup_{s \in [0, T]} |f(s)|,$$

is not technically suitable in general.

Example: Let $f : [0, 2] \rightarrow [0, 1]$ with $f(t) = 1$, if $1 \leq t \leq 2$ and $f(t) = 0$, if $0 \leq t < 1$. Define the sequence of càdlàg functions $f_n : [0, 2] \rightarrow [0, 1]$ with

$$f_n(t) := \begin{cases} 1, & \text{if } 1 + \frac{1}{n} \leq t \leq 2 \\ 0, & \text{if } t < 1 + \frac{1}{n} \end{cases}.$$

Fig. 2.2 depicts the functions f_n and f . Even though one may be tempted to claim convergence in this case, f_n does not converge to f , under the uniform topology.

Indeed,

$$1 \geq \|f - f_n\| = \sup_{s \in [0, T]} |f(s) - f_n(s)| \geq \left| f_n\left(1 + \frac{1}{2n}\right) - f\left(1 + \frac{1}{2n}\right) \right| = |0 - 1| = 1,$$

that is,

$$\|f - f_n\| = 1,$$

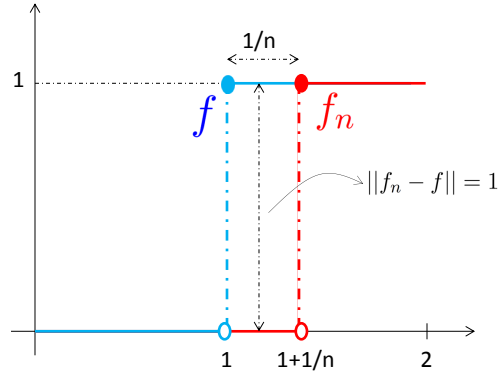


Figure 2.2: Illustration of the sequence f_n and the limiting function f . The sequence f_n does not converge to f , with respect to the uniform topology, as $\|f_n - f\| = 1$ for all n .

for all $n \in \mathbb{N}$. Moreover, f_n is not even a Cauchy sequence as

$$\|f_m - f_n\| = \sup_{s \in [0, T]} |f_m(s) - f_n(s)| = 1,$$

for any $m, n \in \mathbb{N}$.

One can slightly correct the uniform norm to overcome this technicality, by allowing a càdlàg function to be *approximated*, in the abscissa (or time) as well, instead of exclusively via the ordinate. In other words, two càdlàg functions f and g would be close together, $d(f, g) < \epsilon$, in light of the new metric d , if they are close after some *wiggling* on the abscissa of g , i.e., $|f(t) - g(\lambda(t))| < \epsilon$ (instead of $|f(t) - g(t)| < \epsilon$) for all $t \in [0, T]$, where $\lambda \in \Lambda$, and Λ is the set of strictly increasing functions $\lambda : [0, T] \rightarrow [0, T]$ with $\lambda(0) = 0$ and $\lambda(T) = T$. More formally, $d(f, g) < \epsilon$, if and only if there exists an ϵ -small change of coordinates $\lambda \in \Lambda$

$$\|\lambda - I\| = \sup_{t \in [0, T]} |\lambda(t) - t| < \epsilon$$

for g , such that

$$\|f - g \circ \lambda\| = \sup_{t \in [0, T]} |f(t) - g(\lambda(t))| < \epsilon,$$

where $I : [0, T] \rightarrow [0, T]$ is the identity map, that is, $I(t) = t$ for all $t \in [0, T]$. Equivalently, the metric d is given in closed form by

$$d(f, g) = \inf_{\lambda \in \Lambda} \{ \|\lambda - I\| \vee \|f - g \circ \lambda\| \},$$

where $a \vee b$ stands for the maximum value between the reals $a, b \in \mathbb{R}$. This metric defines the Skorokhod topology that we denote by $D_{[0, T]}$ – Skorokhod topological space of càdlàg real valued or \mathbb{R}^n -valued functions defined over the compact interval $[0, T]$. The Skorokhod topology is separable (refer to [19]). Now, a sequence f_n converges to f under the Skorokhod topology if and only if there exists a sequence $\lambda_n \in \Lambda$

with

$$\begin{aligned} \|f - f_n \circ \lambda_n\| &\longrightarrow 0 \\ \|\lambda_n - I\| &\longrightarrow 0. \end{aligned}$$

One can check that under the metric d , the sequence f_n in the previous example converges to f . In other words, f_n converges to f with respect to the Skorokhod topology. Indeed, take

$$\lambda_n(t) := \begin{cases} (1 + \frac{1}{n})t - \frac{2}{n}, & \text{if } t \geq 1 \\ (1 - \frac{1}{n})t, & \text{if } t < 1 \end{cases}$$

and observe that $f_n(\lambda_n(t)) = f(t)$ for all $t \in [0, 2]$; and

$$\|\lambda_n - I\| \leq \frac{1}{n} \longrightarrow 0.$$

An important result that bridges the metric d on $D_{[0,T]}$ with the uniform norm $\|\cdot\|$ on the set of continuous functions $C_{[0,T]}$ is that, if the limiting càdlàg function $f \in C_{[0,T]}$ is continuous, then

$$d(f, f_n) \longrightarrow 0 \iff \|f - f_n\| \longrightarrow 0, \quad (2.3)$$

i.e., if the limit f is continuous, then convergence with respect to the Skorokhod topology coincides with convergence with respect to the uniform topology (refer to [19]). This further implies that the Skorokhod topology induced in the subset of continuous functions coincides with the uniform topology on $C_{[0,T]}$. In fact, the right-to-left implication in (2.3) follows from the fact that

$$d(f, g) = \inf_{\lambda \in \Lambda} \{\|\lambda - I\| \vee \|f - g \circ \lambda\|\} \leq \|I - I\| \vee \|f - g \circ I\| = \|f - g\|.$$

This connection between the uniform norm $\|\cdot\|$ and the metric d , whenever the limit function is continuous, is of our interest. Indeed, and as will be clearer in Chapter 4, our goal will revolve about establishing weak convergence of càdlàg processes to the solution of a differential equation. That is, the limit process is continuous.

We call $(Z(t))$ a càdlàg stochastic process, or simply càdlàg process, a $D_{[0,T]}$ -valued function over the sample space Ω

$$\begin{aligned} (Z(t)) &: \Omega \longrightarrow D_{[0,T]} \\ \omega \in \Omega &\longmapsto (Z(\omega, t)), \end{aligned}$$

that is measurable with respect to the measurable space (Ω, \mathcal{F}) . Note that, for all $\omega \in \Omega$, the sample path $(Z(\omega, t))$ is a càdlàg function over the compact interval $[0, T]$, and thus, it is bounded and measurable

with respect to $\mathcal{B}([0, T])$ (refer to [19]). Therefore,

$$\int_0^T Z(\omega, s) ds < \infty \quad \text{or} \quad \int_0^T Z^2(\omega, s) ds < \infty, \quad \forall \omega \in \Omega$$

and this fact will be important for us later in the Thesis when exploring the stochastic dynamics of interest (peep ahead to equation (4.22) in Chapter 4). Note that in the definition of a stochastic process one often requires the measurability of the function

$$\begin{aligned} Z &: \Omega \times [0, T] \longrightarrow \mathbb{R} \\ (\omega, t) &\longmapsto Z(\omega, t) \end{aligned} \tag{2.4}$$

with respect to the product σ -algebra $\mathcal{F} \otimes \mathcal{B}([0, T])$. For all our purposes, we do not need to assure this further requirement for our stochastic processes constructs as there will be no need to evoke Fubini's Theorem for the product $\mathcal{F} \otimes \mathcal{B}([0, T])$, in this Thesis.

Since we will work with probability theory on the space $D_{[0, T]}$, for various technical reasons, it is important to understand what are the measurable subsets in the Borel σ -algebra

$$\mathcal{D}_{[0, T]} := \mathcal{B}(D_{[0, T]}).$$

The next Theorem states that the Borel σ -algebra $\mathcal{D}_{[0, T]}$ coincides with the smallest σ -algebra that makes the canonical projection maps

$$\begin{aligned} \pi_t &: D_{[0, T]} \longrightarrow \mathbb{R} \\ x &\longmapsto x(t) \end{aligned} \tag{2.5}$$

measurable, for all $t \in [0, T]$, i.e., the σ -algebra generated by the family $\{\pi_t\}_{t \in [0, T]}$. That is, $\mathcal{D}_{[0, T]}$ is equal to the product σ -algebra on the set of real valued functions $\mathbb{R}^{[0, T]}$ relativized to the set of càdlàg functions. In other words, $B \in \mathcal{D}_{[0, T]}$ if and only if B is measurable with respect to the product σ -algebra induced in the set of càdlàg functions.

Theorem 5 (Skorokhod and Product Topology; [20]). *The Borel σ -algebra $\mathcal{D}_{[0, T]}$ is equal to the product σ -algebra on $\mathbb{R}^{[0, T]}$ relativized to the set of càdlàg sample paths, i.e.,*

$$\mathcal{D}_{[0, T]} = \sigma(\pi_s : 0 \leq s \leq T),$$

where π_s is the canonical projection map defined in equation (2.5).

In what follows, we show that some of the main operators over $D_{[0, T]}$ that arise in this Thesis are measurable. First, it is rather surprising that a simple operator such as the sum

$$\begin{aligned} + &: D_{[0, T]} \times D_{[0, T]} \longrightarrow D_{[0, T]} \\ (x, y) &\longmapsto (x(t) + y(t)) \end{aligned}$$

is not continuous with respect to the product topology in $D_{[0,T]} \times D_{[0,T]}$. In other words, the Skorokhod space $D_{[0,T]}$ is not a topological group. For technical reasons, it is important to guarantee that the sum, as an operator, is at least measurable. We present in the following Corollary, the functional version of Theorem 5.

Corollary 6. *Let*

$$\mathcal{G} : S \longrightarrow D_{[0,T]}.$$

for some measurable space S . Then,

$$\mathcal{G} \text{ is measurable} \iff \mathcal{G}_t := \pi_t \circ \mathcal{G} : S \rightarrow \mathbb{R} \text{ is measurable}$$

for all $t \in [0, T]$.

In words, Corollary 6 states that an operator \mathcal{G} is measurable if and only if its projection on the t th coordinate is measurable. If S supports a probability space, then Corollary 6 reads as: \mathcal{G} is a stochastic process if and only if for each t , \mathcal{G}_t is a random variable. The next Corollary now clarifies that operators on the Skorokhod space constructed via pointwise (in time) composition with measurable functions (e.g. the sum operator) are measurable.

Corollary 7. *Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a measurable function with respect to the standard Borel σ -algebra in \mathbb{R}^2 . Then,*

$$\begin{aligned} O_f : D_{[0,T]} \times D_{[0,T]} &\longrightarrow D_{[0,T]} \\ (x, y) &\longmapsto (f(x(t), y(t))) \end{aligned}$$

is measurable with respect to the product σ -algebra $\mathcal{D}_{[0,T]} \otimes \mathcal{D}_{[0,T]}$, where

$$\mathcal{D}_{[0,T]} \otimes \mathcal{D}_{[0,T]} := \sigma \{ \hat{\pi}_1, \hat{\pi}_2 \} \tag{2.6}$$

with

$$\begin{aligned} \hat{\pi}_i : D_{[0,T]} \times D_{[0,T]} &\longrightarrow D_{[0,T]} \\ (x_1, x_2) &\longmapsto x_i \end{aligned}.$$

Aside Note: *Since $D_{[0,T]}$ is separable, the finite product topology commutes with the finite product σ -algebra (refer to [22]), i.e.,*

$$\mathcal{B}(D_{[0,T]} \times D_{[0,T]}) = \mathcal{B}(D_{[0,T]}) \otimes \mathcal{B}(D_{[0,T]}) = \mathcal{D}_{[0,T]} \otimes \mathcal{D}_{[0,T]},$$

and in the Corollary above one can adopt either σ -algebra. If we consider uncountable product though, then the product topology does not commute with the product σ -algebra, for instance,

$$\bigotimes_{t \in [0,T]} \sigma(\mathbb{R}) \neq \sigma \left\{ \mathbb{R}^{[0,T]} \right\},$$

where $\mathbb{R}^{[0,T]}$ bears the product topology $\Pi_{[0,T]}\mathbb{R}$. In the stochastic processes setting one often adopts the product σ -algebra rather than the σ -algebra of the product, as it leads to the assertion that $(Z(t))$ is a stochastic process if and only if $Z(t)$ is a random variable for all t .

It turns out that the sum and product, $(x(t) + y(t))$ and $(x(t)y(t))$, are measurable operators. Indeed, take $f(v, w) = v + w$ or $f(v, w) = vw$. Also, the modulus of continuity ω :

$$\begin{aligned} \omega_\delta : D_{[0,T]} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sup \{|x(u) - x(v)| : 0 \leq u, v \leq T, |u - v| \leq \delta\} \end{aligned}$$

is measurable, since $\mathcal{D}_{[0,T]}$ is generated by finite canonical projections on a dense subinterval of $[0, T]$. In other words, let $\epsilon > 0$,

$$\begin{aligned} \{x \in D_{[0,T]} : \omega_\delta(x) \leq \epsilon\} &= \{x \in D_{[0,T]} : \sup \{|x(u) - x(v)| : u, v \in \mathbb{Q} \cap [0, T] \mid |u - v| \leq \delta\} \leq \epsilon\} \\ &= \bigcap_{u, v \in \mathbb{Q} \cap [0, T]; |u - v| \leq \delta} \underbrace{\{x \in D_{[0,T]} : |x(u) - x(v)| \leq \epsilon\}}_{A_{u,v}} \end{aligned}$$

where the first equality holds from Lemma 1 of Chapter 3 in the reference [19]. From Corollary 7 and the previous discussion, the sets $A_{u,v}$ are measurable and thus, ω_δ is measurable. The measurability of the modulus of continuity is crucial to our purposes of establishing convergence of processes, please refer, for instance, to the Arzelà-Ascoli Theorem 20 in Chapter 4. Otherwise, the event

$$\{x \in D_{[0,T]} : \omega_\delta(x) \geq \epsilon\}$$

would not be necessarily measurable, and

$$\mathbb{P} \{x \in D_{[0,T]} : \omega_\delta(x) \geq \epsilon\}$$

would not make sense (and could lead to contradictions if taken forward).

Aside Note: We could not find in the literature the proof of the measurability of the modulus of continuity ω_δ with respect to the Skorokhod space, and that is why we displayed the proof above.

The fact that the integral construct $(H(t))$, defined as

$$\begin{aligned} (H(t)) : \Omega &\longrightarrow D_{[0,T]} \\ \omega \in \Omega &\longmapsto \left(\int_0^t Z(\omega, s) ds \right), \end{aligned}$$

is a càdlàg process, will be clearer later in the Thesis (see Lemma 23 in Chapter 4). This is an important requirement, for instance, to make sense of the expectation

$$E(H(t)) = E \left(\int_0^t Z(s) ds \right) = \int_\Omega \left(\int_0^t Z(\omega, s) ds \right) d\mathbb{P}, \quad (2.7)$$

for each t , $t \in [0, T]$, as in this case, and from Corollary 6, $H(t)$ is a random variable. We remark that there will be no need to interchange the integrals in equation (2.7) in this Thesis, and therefore, there is no need to guarantee that Z , as defined in equation (2.4), is measurable with respect to the product σ -algebra $\mathcal{F} \otimes \mathcal{B}(0, T)$.

2.3 Weak Convergence and Tightness

Let (Ω, \mathcal{F}) be a measurable space, where Ω is a topological space and \mathcal{F} is the corresponding Borel σ -algebra. Let $\mathbb{P}_n : \mathcal{F} \rightarrow [0, 1]$ be a sequence of probability measures in (Ω, \mathcal{F}) . Let \mathbb{P}_∞ be another probability measure in (Ω, \mathcal{F}) . We say that \mathbb{P}_n converges weakly to \mathbb{P}_∞ , with respect to the topology in Ω , and denote it by

$$\mathbb{P}_n \Rightarrow \mathbb{P}_\infty,$$

whenever,

$$\int f d\mathbb{P}_n \longrightarrow \int f d\mathbb{P} \quad (2.8)$$

for all bounded, continuous functions $f : \Omega \rightarrow \mathbb{R}$. This sense of convergence defines a topology on the space of probability measures $\mathcal{P}(\Omega)$ in (Ω, \mathcal{F}) , that we refer to as weak topology. Roughly speaking, two probability measures \mathbb{P}_1 and \mathbb{P}_2 are close, with respect to the weak topology, if their integrals return similar values for most bounded continuous functions.

There exist other notions of convergence. For instance, we say that \mathbb{P}_n converges pointwisely to \mathbb{P} if

$$\mathbb{P}_n(B) \longrightarrow \mathbb{P}(B) \quad (2.9)$$

for all $B \in \mathcal{F}$. Pointwise convergence is stronger than weak convergence.

In this Thesis we will adopt the weak convergence sense. The first reason lies in the structure that is preserved from the underlying metric space, namely, if Ω is a separable metric space, then the space of probability measures $\mathcal{P}(\Omega)$ under the weak topology is metrizable and separable (refer to [19]). This is an important structure when it comes to convergence and tightness of sequences of probability measures. The next example further motivates our choice for weak topology.

Example: Let μ_n be absolutely continuous – with respect to the Lebesgue measure Leb on \mathbb{R} – probability measures over $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with corresponding densities f_n characterized by the following properties:

- $\int f_n(t) d\text{Leb}(t) = \mu_n(\mathbb{R}) = 1$;
- $f_n(x) > 0$ only if $x \in [x^* - 1/n; x^* + 1/n]$, otherwise, $f_n(x) = 0$.

We note that the sequence of probability measures μ_n concentrates about the value x^* as n goes to infinity as depicted in Fig. 2.3. In other words, we would like to claim that μ_n converges to the Dirac measure δ_{x^*} about x^* . This is true if we adopt the weak convergence sense. It is not true that μ_n converges pointwise

to δ_{x^*} . Indeed,

$$\mu_n(x^*) = 0 \neq \delta_{x^*}(x^*) = 1,$$

where the first equality is due to the absolute continuity of μ_n . This example can be extended to the framework of a sequence of non-atomic probability measures – i.e., $\mathbb{P}_n(\{\omega\}) = 0$ for any singleton $\{\omega\} \subset \Omega$, – converging to a dirac measure.

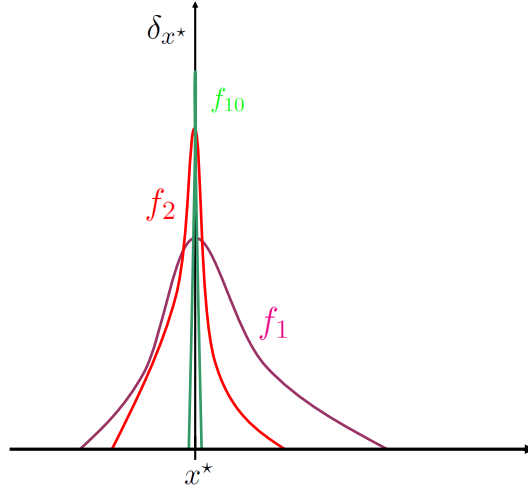


Figure 2.3: Informal illustration of a concentration of a sequence of absolutely continuous probability measures μ_n about a dirac measure δ_{x^*} . Each function f_i represents the density associated with each probability measure μ_i .

In this Thesis, we will observe a concentration of non-atomic probability measures phenomenon: the stochastic process $(\bar{Y}^N(t))$ associated with certain fractions of infected nodes converges to the solution of a differential equation, as the network grows large, i.e., the number of agents N goes to infinite. In other words, the push-forward probability measures

$$\mathbb{P}_{\bar{Y}^N} \in \mathcal{P}(D_{[0,T]})$$

induced by the processes $(\bar{Y}^N(t))$ concentrate about a Dirac measure $\delta_{(y(t))}$ about a path $(y(t)) \in D_{[0,T]}$. In this case, and with some abuse of notation, we say that

$$(\bar{Y}^N(t)) \Rightarrow (y(t))$$

i.e., $(\bar{Y}^N(t))$ converges weakly to $(y(t))$, with respect to the Skorokhod topology $D_{[0,T]}$.

The criterion in equation (2.8) is not in general efficient to establish weak convergence. In the stochastic processes setting, it is often the case that we start by establishing the **tightness** of a sequence \mathbb{P}_n that we now define.

Definition 8 (Tightness). *A probability measure \mathbb{P} on (Ω, \mathcal{F}) is tight if for every $\epsilon > 0$, there exists a compact subset $K \subset \Omega$ such that*

$$\mathbb{P}(K) > 1 - \epsilon.$$

Further, a sequence of probability measures \mathbb{P}_n is tight if for any $\epsilon > 0$, there exists a compact subset $K \subset \Omega$, such that

$$\mathbb{P}_n(K) > 1 - \epsilon \tag{2.10}$$

for all $n \in \mathbb{N}$.

In words, a sequence \mathbb{P}_n is tight if all of its probabilities bear the same “modulus of tightness”. Now, Prohorov’s theorem bridges tightness of a sequence \mathbb{P}_n with convergence of probability measures. Namely, it states that a sequence \mathbb{P}_n is tight if and only if it is *relatively compact*, i.e., it admits weakly convergent subsequence.

Theorem 9 (Prohorov; [19]). *Let (M, d) be a Polish metric space. Then, the following holds*

- $\{\mathbb{P}_n\}_n$ is tight if and only if there exists a subsequence \mathbb{P}_{n_k} and a probability \mathbb{P}_∞ with $\mathbb{P}_{n_k} \Rightarrow \mathbb{P}_\infty$;
- If for any subsequence $(n_k)_k$ we have $\mathbb{P}_{n_k} \Rightarrow \mathbb{P}_\infty$, then $\mathbb{P}_n \Rightarrow \mathbb{P}_\infty$.

Some of the Theorems in Part I, provide sufficient conditions for tightness of a sequence of stochastic processes. If such conditions are fulfilled, then Prohorov’s Theorem implies relative compactness. Also, the uniqueness of the weak accumulation point \mathbb{P}_∞ implies the convergence of the sequence (as referred in the second bullet of Theorem 9). For example, the reader may be familiar with the Arzelà-Ascoli theorem associated with the uniform-tightness of a sequence of (deterministic) continuous functions.

Theorem 10 (Deterministic-Arzelà-Ascoli;[23]). *A sequence $f_n : [0, T] \rightarrow \mathbb{R}$ of continuous functions is uniformly-tight if*

- $\sup_{t \in [0, T]} |f_n(t)| < \infty$; (Uniform Boundedness)
- $\lim_{\delta \rightarrow 0} \limsup_n \omega_\delta(f_n) = 0$, (Equicontinuity)

i.e., there exists a subsequence f_{n_k} of continuous functions and a continuous function f , where

$$\|f_{n_k} - f\| \longrightarrow 0.$$

The modulus of continuity $\omega_\delta(\cdot)$ is an important function to detect tightness of a sequence of continuous functions. In fact, $\omega_\delta(\cdot)$ also allows to establish C -tightness of a sequence of càdlàg functions, i.e., a sequence that is tight and whose accumulation points are continuous functions.

Theorem 11 (Deterministic-Arzelà-Ascoli, BIS). *A sequence $f_n : [0, T] \rightarrow \mathbb{R}$ of càdlàg functions is C -tight if*

- $\sup_{t \in [0, T]} |f_n(t)| < \infty;$ (Uniform Boundedness)
- $\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \omega_\delta(f_n) = 0,$ (Equicontinuity)

i.e., there exists a subsequence f_{n_k} of càdlàg functions and a continuous function f , where

$$\|f_{n_k} - f\| \rightarrow 0.$$

The Arzelà-Ascoli Theorem 20 in Chapter 4, conforms to a stochastic generalization to Theorem 11 above, in that it provides similar conditions of equicontinuity and uniform boundedness of a sequence of càdlàg stochastic processes to guarantee C-tightness. Theorem 11 admits further extension to the case where the limit function f is càdlàg (not necessarily continuous) and the sequence is tight with respect to the Skorokhod topology. To achieve this extension, we just have to replace the modulus of continuity ω_δ in Theorem 11 by another modulus ω'_δ that appropriately extends ω_δ and can be found in [19].

Aside note: Prohorov's Theorem as stated relies on the Polishness of the underlying metric space. As referred to in Section 2.2, the topological space $D_{[0, T]}$ is separable. But, the metric space $(D_{[0, T]}, d)$ is not complete (refer to [19]). The metric d^o defined as

$$d^o(x, y) = \inf_{\lambda \in \Lambda} \{ \|\lambda\|^o \vee \|f - g \circ \lambda\| \},$$

with

$$\|\lambda\|^o := \sup_{s < t} \left| \log \left(\frac{\lambda(t) - \lambda(s)}{t - s} \right) \right|,$$

is topologically equivalent to d , i.e., induces the Skorokhod topology and leads to a complete (hence Polish) metric space $(D_{[0, T]}, d^o)$ (refer to [19]). In other words, d and d^o induce the same topology (Skorokhod), but they lead to different metric spaces. For other technical reasons, Polishness is a convenient structure. For instance, if the underlying metric space (M, d) is Polish then, the dual space $\mathcal{P}(M)$, with the weak topology, is Polishable – i.e., complete, separable and metrizable. Also, the probability space $(M, \mathcal{B}(M), \mathbb{P})$, where \mathbb{P} is any non-atomic Borel probability measure, is isomorphic to the Lebesgue space $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$, that is, there exists a measurable bijection $G : M \rightarrow [0, 1]$ with measurable inverse $G^{-1} : [0, 1] \rightarrow M$ that preserves the probability measures, i.e., $\mathbb{P}_{*G} = \text{Leb}$ and $\mathbb{P} = \text{Leb}_{*G^{-1}}$. Therefore, many results from probability theory on the Lebesgue space can be applied to Polish spaces. This isomorphism further offers an upper-bound on the cardinality of the set M which prevents from measurability issues.

2.4 Further Modes of Convergence

In this Section, we introduce the relevant modes of convergence that appear in Part I of this Thesis, besides the weak convergence with respect to the Skorokhod topology. We also set the notation for these modes. In what follows, we say that the sequence of stochastic processes $\left(\bar{Y}^N(t)\right)_N$ converges on the line in the

sense x , whenever the corresponding sequence of random variables $\bar{Y}^N(t)$ converges, in the sense x , for each time t , $t \geq 0$. This expression is coined from [19, 24]. For the convergence on the line modes, we reserve the symbol ‘ \longrightarrow ’ alongside with the sense of convergence x , as will be clearer next.

• **Convergence in \mathcal{L}_2 on the line:**

$(\bar{Y}^N(t))$ converges in \mathcal{L}_2 to $(\bar{Y}_\infty(t))$ on the line whenever,

$$\left\| \bar{Y}^N(t) - \bar{Y}_\infty(t) \right\|_{\mathcal{L}_2} := E \left(\bar{Y}^N(t) - \bar{Y}_\infty(t) \right)^2 \xrightarrow{N \rightarrow \infty} 0,$$

for all t , $t \geq 0$.

Notation:

$$(\bar{Y}^N(t)) \xrightarrow{\mathcal{L}_2} (\bar{Y}_\infty(t)).$$

• **Convergence in distribution on the line:**

$(\bar{Y}^N(t))$ converges in distribution to $(\bar{Y}_\infty(t))$ on the line whenever,

$$\bar{Y}^N(t) \xrightarrow{d} \bar{Y}_\infty(t)$$

for all t , $t \geq 0$, i.e., the sequence of random variables $\bar{Y}^N(t)$ converges weakly (or in distribution) to the random variable $\bar{Y}_\infty(t)$, for all t .

Notation:

$$(\bar{Y}^N(t)) \xrightarrow{d} (\bar{Y}_\infty(t)).$$

• **Convergence in finite-dimension distribution (f.d.d.):**

$(\bar{Y}^N(t))$ converges in finite-dimension distribution to $(\bar{Y}_\infty(t))$ whenever,

$$(\bar{Y}^N(t_1), \dots, \bar{Y}^N(t_n)) \xrightarrow{d} (\bar{Y}_\infty(t_1), \dots, \bar{Y}_\infty(t_n))$$

for any finite set

$$\{t_1, \dots, t_n\} \subset [0, T].$$

Notation:

$$(\bar{Y}^N(t)) \xrightarrow{f.d.d.} (\bar{Y}_\infty(t)).$$

• **Convergence in Probability on the line:**

$(\bar{Y}^N(t))$ converges in probability to $(\bar{Y}_\infty(t))$ on the line, whenever, for all $\epsilon > 0$

$$\mathbb{P} \left(\left| \bar{Y}^N(t) - \bar{Y}_\infty(t) \right| > \epsilon \right) \longrightarrow 0,$$

for all $t, t \geq 0$.

Notation:

$$(\bar{Y}^N(t)) \xrightarrow{\mathbb{P}} (\bar{Y}_\infty(t)).$$

• **Convergence in Probability:**

$(\bar{Y}^N(t))$ converges in probability to $(\bar{Y}_\infty(t))$, whenever, for all $\epsilon > 0$

$$\mathbb{P} \left(\left\| \bar{Y}^N(t) - \bar{Y}_\infty(t) \right\| > \epsilon \right) = \mathbb{P} \left(\sup_{t \in [0, T]} \left| \bar{Y}^N(t) - \bar{Y}_\infty(t) \right| > \epsilon \right) \longrightarrow 0,$$

i.e., $(\bar{Y}^N(t))$ converges in probability to $(\bar{Y}_\infty(t))$ with respect to the uniform norm. Note that this implies convergence in probability with respect to the Skorokhod metric d^o , since

$$d^o(f, g) = \inf_{\lambda \in \Lambda} \{ \|\lambda\|^o \vee \|f - g \circ \lambda\| \} \leq \|I\| \vee \|f - g \circ I\| = \|f - g\|.$$

for all $f, g \in D_{[0, T]}$. Convergence in probability with respect to the Skorokhod metric does not necessarily imply convergence in probability with respect to the uniform norm.

Notation:

$$(\bar{Y}^N(t)) \xRightarrow{\mathbb{P}} (\bar{Y}_\infty(t)).$$

• **Realization-wise convergence w.r.t. uniform topology:**

$(\bar{Y}^N(t))$ converges realization-wise to $(\bar{Y}_\infty(t))$, with respect to the uniform topology on the set of càdlàg functions, whenever,

$$\left\| \bar{Y}^N(\omega, t) - \bar{Y}_\infty(\omega, t) \right\| = \sup_{t \in [0, T]} \left| \bar{Y}^N(\omega, t) - \bar{Y}_\infty(\omega, t) \right| \longrightarrow 0,$$

for almost all $\omega \in \Omega$.

Notation:

$$(\bar{Y}^N(t)) \longrightarrow (\bar{Y}_\infty(t)).$$

• **Realization-wise convergence w.r.t. Skorokhod topology:**

$(\bar{Y}^N(t))$ converges realization-wise to $(\bar{Y}_\infty(t))$, with respect to the Skorokhod topology on $D_{[0,T]}$, whenever,

$$(\bar{Y}^N(\omega, t)) \xrightarrow{\text{SK}} (\bar{Y}_\infty(\omega, t)),$$

or equivalently,

$$d\left((\bar{Y}^N(\omega, t)), (\bar{Y}_\infty(\omega, t))\right) \rightarrow 0,$$

for almost all $\omega \in \Omega$.

Notation:

$$(\bar{Y}^N(t)) \xrightarrow{\text{SK}} (\bar{Y}_\infty(t)).$$

2.5 Summary of the Chapter

In this Chapter, we reviewed some of the main definitions and results from probability theory that are relevant for Part I of this Thesis. In Section 2.1, we set the notation and introduce some basic definitions from probability theory. In Section 2.2, we described the Skorokhod topology; provided a criterion for measurability in $D_{[0,T]}$; and stressed the measurability of some of the important operators on $D_{[0,T]}$ – e.g., the modulus of continuity, – that appear in the Thesis. Section 2.3 recalls the concept of weak convergence and motivates our choice for this sense of convergence in this Thesis. Also, we presented tightness as a means to establish weak convergence via Prohorov's theorem. In Section 2.4, we listed some of the main modes of convergence that are crucial as intermediate steps to establish weak convergence. Further details can be found in [19, 20, 21].

Chapter 3

Problem Formulation

This Chapter presents the peer-to-peer model of diffusion of the epidemics and constructs the integral stochastic equations governing the macroscopic dynamics, i.e., the dynamics of the fraction of infected nodes across islands over time. The model presented is in line with the Harris contact process (refer to [25]). In Section 3.1, we formally introduce the stochastic network process that models the peer-to-peer diffusion of multiple strains of virus over complete-multipartite networks. In Section 3.2, we build upon the microscopic model presented in Section 3.1 to provide a pathwise characterization of the number of infected nodes process $(\bar{Y}^N(t))$. In Section 3.3, we conclude the Chapter.

3.1 Stochastic Network Model

First, we set up the environment where the epidemics take place. Let $G = (V, E)$ be an undirected network, where

$$V = \{1, 2, \dots, N\} \text{ and } E = \{\{i, j\} : i, j \in V\}$$

represent the set of nodes and edges of the graph G , respectively. If $\{i, j\} \in E$, then we say that the nodes $i, j \in V$ are connected and represent it as $i \sim j$. In this Thesis, we establish the fluid limit dynamics of a multivirus epidemics over the class of complete-multipartite networks that is defined next.

Definition 12 (Complete-multipartite Network). *A network $G = (V, E)$ is complete-multipartite if there exists a partition $\bar{V} = \{V_1, \dots, V_M\}$ of V such that $\{a, b\} \notin E$ for any $a, b \in V_i$ and any $i \in \{1, \dots, M\}$. Moreover,*

$$u \in V_i, v \in V_j, u \sim v \Rightarrow w \sim r, \quad \forall w \in V_i, r \in V_j$$

with $i \neq j$. When $M = 2$, the complete-multipartite network is called complete-bipartite. ■

The elements V_i of the partition \bar{V} are referred to as *islands*. In words, if two nodes of different islands V_i and V_j are connected then any node from V_i is connected to any node from V_j . In this case, we say that

islands V_i and V_j are connected and refer to it as $V_i \sim V_j$, which allows us to abstract the supertopology structure of islands. Figure 3.1 depicts a complete-multipartite network with 5 islands. Also, we refer to

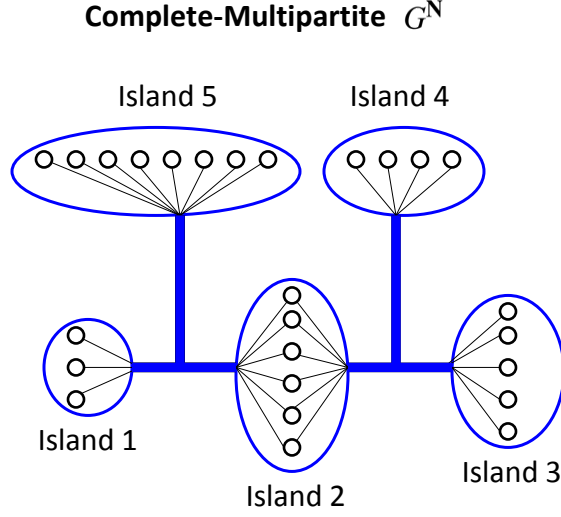


Figure 3.1: Complete-multipartite network representation. Nodes from the same island cannot transmit the infection amongst themselves. Nodes from an island can transmit the virus to nodes in neighboring islands. For instance, any node from island 5 can infect any node from islands 1 and 2. Nodes from island 1 cannot infect nodes from island 4 and vice-versa, as the islands 1 and 4 are not neighbors.

$$\mathcal{N}(V_i) = \{V_j : V_j \sim V_i\}$$

as the superneighborhood of island V_i and

$$d_i = |\mathcal{N}(V_i)|$$

refers to the superdegree or number of neighboring islands of island V_i . As an example, for the superneighbors of island 1, in Figure 3.1, we have

$$\mathcal{N}(1) = \{2, 5\}$$

and thus, $d_1 = 2$. Given a graph $G = (E, V)$, we define the sequence of induced complete-multipartite networks

$$G^{\mathbf{N}} = (E^{\mathbf{N}}, V^{\mathbf{N}}),$$

indexed by

$$\mathbf{N} = (N_1, \dots, N_M) \in \mathbb{R}^M,$$

where M is the fixed number of islands, N_i is the number of nodes at the i th island of $G^{\mathbf{N}}$, and

$$V_i^{\mathbf{N}} \sim V_j^{\mathbf{N}} \Leftrightarrow \ell_i \sim \ell_j$$

for all $\ell_i, \ell_j \in V$, where $V_i^{\mathbf{N}}$ is the i th island of $G^{\mathbf{N}}$, as depicted in Fig. 3.2. Note that the family $G^{\mathbf{N}}$ is uniquely determined, up to a relabeling of the nodes, by the graph G . We refer to the underlying graph G as the generator of the family of complete-multipartite networks $G^{\mathbf{N}}$. Note in Fig. 3.2 that, for instance, islands 4 and 5 are not super-neighbors, as in the generator network G the nodes 4 and 5 are not neighbors. In words, all the complete-multipartite graphs in the sequence $G^{\mathbf{N}}$ share the same supertopology imposed by the topology of its generator G , differing only on the number of nodes per island that is given by the upper-index $\mathbf{N} \in \mathbb{N}$. Given a generator graph G , we are interested in obtaining the limiting dynamics of the fraction of infected nodes per island and per strain over $G^{\mathbf{N}}$ as \mathbf{N} grows to infinity.

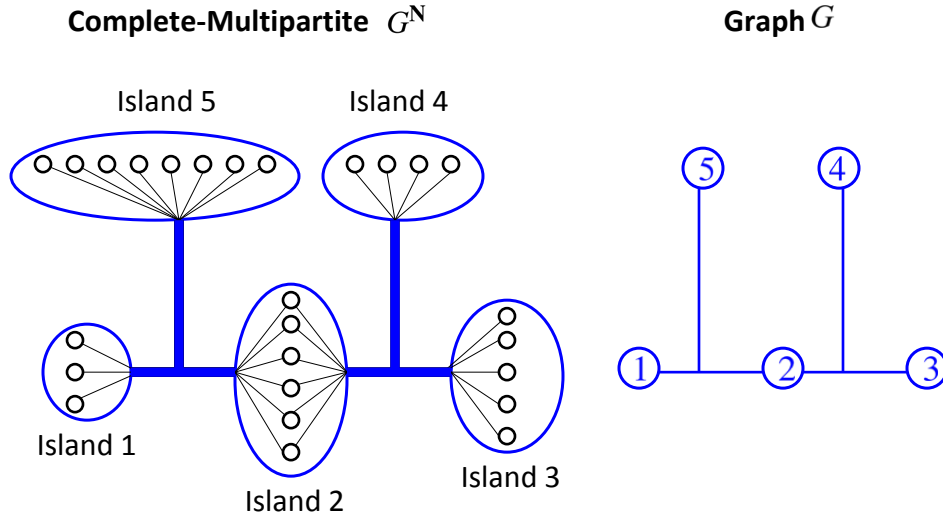


Figure 3.2: Illustration of the complete-multipartite network $G^{\mathbf{N}}$, with $\mathbf{N} = (3, 6, 5, 4, 8)$ and underlying graph G .

Remark that a complete-bipartite network models the underlying structure of contacts between humans and *Aedes Aegypti* mosquitoes in a dengue-like pandemics, as mentioned in Chapter 1. A complete-multipartite topology naturally extends to account for more than two species in the epidemics. Also, the motivation to characterize the limiting dynamics as \mathbf{N} goes to infinity lies in the large-scale nature of such epidemics.

Now, we introduce the dynamical model of diffusion. A node in a $G^{\mathbf{N}}$ complete-multipartite network may assume one of two possible states: 1) infected or 2) healthy and susceptible to infection. We define the binary *tensor* microprocess $(\mathbf{X}^{\mathbf{N}}(t))$ as describing the state of each node over time in $G^{\mathbf{N}}$. By

$$X_{nik}^{\mathbf{N}}(t) = 1,$$

we refer to node

$$n \in \{1, \dots, N_i\}$$

at island

$$i \in \{1, \dots, M\}$$

being infected at time $t, t \geq 0$, with virus strain

$$k \in \{1, 2, \dots, K\},$$

and by

$$X_{nik}^{\mathbf{N}}(t) = 0$$

if the corresponding node n is healthy or infected with other strain $l \neq k$; in this latter case,

$$X_{nil}^{\mathbf{N}}(t) = 1.$$

The upper-index \mathbf{N} will show in all relevant stochastic processes constructs to emphasize the underlying complete-multipartite network $G^{\mathbf{N}}$ induced by G . If only one strain of virus is present in the network then, for notational simplicity, we suppress the extra-index k and write simply

$$X_{ni}^{\mathbf{N}}(t) = 1$$

to represent that node n from island i is infected at time $t, t \geq 0$, and

$$X_{ni}^{\mathbf{N}}(t) = 0,$$

if otherwise.

Our microscopical model of diffusion of the virus is set at node level and goes as follows. If a node n from island i is k -infected – i.e., infected with the virus strain k – at time t , then it heals after an exponentially distributed random time

$$T_{ik}^h(n) \sim \text{Exp}(\mu_i^k)$$

whose distribution depends on the island i and on the type of infection k , but it does not depend on the specific node n . Once a node n in island i is k -infected, it transmits the infection to a randomly chosen node at the neighbor island $j \in \mathcal{N}(i)$ after an exponentially distributed random time

$$T_{ijk}^c(n) \sim \text{Exp}(\gamma_{ij}^k)$$

whose distribution only depends on the ordered pair

$$(i, j) \in \{1, 2, \dots, M\}^2 \text{ with } i \neq j,$$

of communicating islands i and j , and the type of infection k . The clocks for infection keep triggering according to the referred exponential law while the node is infected. Whenever considering only one strain of virus, we drop the strain subindex k . Also, if there is no room for ambiguity, we drop the node identity n , writing T_{ik}^h for the healing time and T_{ijk}^c for the infection time of some node at island i , whose identity is not relevant. We assume an *exclusion principle*: if the chosen node n_j at island j is already k -infected at the time of infection t , $t \geq 0$, then nothing happens. Therefore, $\sum_k X_{nik}(t) \leq 1$ for all nodes n in island i , for all

$$i \in \{1, \dots, M\},$$

and $t, t \geq 0$, or in words, a node can only be infected by one strain of virus at a time. The exclusion principle captures the real life inhibition of infection when an agent (or node) is vaccinated and the impossibility of vaccination while the agent is infected, as referred to in Chapter 1.

To summarize, an infected node $n \in i$ activates $d_i + 1$ independent exponentially distributed random *clocks*, where $d_i = |\mathcal{N}(i)|$ of them are associated with the times for infection and one clock is associated with the time for healing. Each of the d_i random clocks for infection is dedicated to one superneighbor $j \in \mathcal{N}(i)$ of island i as illustrated in Figure 3.3. As an example, if a node from island $i = 1$ in Figure 3.3

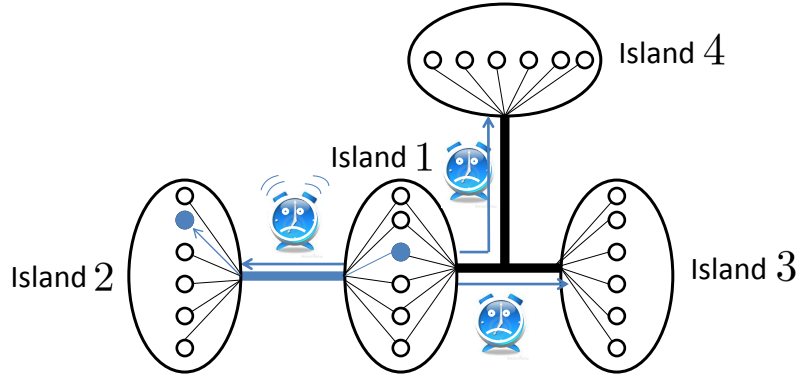


Figure 3.3: Illustration of an infection. The infected (dark colored) node at island 1 activates three exponentially distributed random *clocks*, each dedicated to one neighbor island. Assuming that the clock dedicated to island 2 rings first, after a time $T_{12k}^c \sim \text{Exp}(\gamma_{12}^k)$, a node from island 2 is randomly picked to be infected. The clocks for infection keep triggering while the node is infected. Also, the infected node in island 1 heals after a time $T_{1k}^h \sim \text{Exp}(\mu_1^k)$.

is k -infected, then, after a time interval of length $T_{12k}^c \sim \text{Exp}(\gamma_{12}^k)$, it picks randomly a node from island $j = 2$ and infects it, as long as the node from 1 is still infected. Also, after a time $T_{13k}^c \sim \text{Exp}(\gamma_{13}^k)$, that is independent of T_{12k}^c and of the healing time T_{1k}^h , it chooses randomly a node from island 3 and infects it.

The microscopic process $(\mathbf{X}^{\mathbf{N}}(t))$ thus, evolves through jumps according to the triggering of a sequence of independent exponentially distributed random variables. All time event-triggering random variables are assumed to be independent and have support in a single probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $(\mathcal{F}_t^{\mathbf{N}})_{t \geq 0}$ the natural filtration induced by the sequence of time independent random variables associated with the clocks. That is, $\mathcal{F}_t^{\mathbf{N}}$ gives us all the information on the values of all random variable times (healing or infection) involved in the evolution of $(\mathbf{X}^{\mathbf{N}}(t))$ up to time $t, t \geq 0$. Note that by construction $(\mathbf{X}^{\mathbf{N}}(t))$ is adapted to $(\mathcal{F}_t^{\mathbf{N}})_{t \geq 0}$, i.e.,

$$\sigma \{ \mathbf{X}^{\mathbf{N}}(s) : 0 \leq s \leq t \} \subset \mathcal{F}_t^{\mathbf{N}},$$

for all $t, t \geq 0$, where

$$\sigma \{ \mathbf{X}^{\mathbf{N}}(s) : 0 \leq s \leq t \}$$

represents the natural filtration of the process $(\mathbf{X}^{\mathbf{N}}(t))$.

Analyzing the full-microstate $(\mathbf{X}^{\mathbf{N}}(t))$ of the network $G^{\mathbf{N}}$ over time according to the local infection model just presented becomes quickly unfeasible with the number of nodes in the network as the microscopic process $(\mathbf{X}^{\mathbf{N}}(t))$ is high-dimensional. Instead, we are interested in characterizing macroscopically the virus evolution in the complete-multipartite network, namely, studying the dynamics of the number or fraction of infected nodes at each island.

To fix ideas, we assume single virus epidemics for the rest of this Section, unless otherwise stated. We refer to

$$(\mathbf{Y}^{\mathbf{N}}(t)) = (Y_1^{\mathbf{N}}(t), \dots, Y_M^{\mathbf{N}}(t))$$

as the macroprocess that stacks the number of infected nodes

$$Y_i^{\mathbf{N}}(t) = \sum_{n=1}^{N_i} \mathbf{X}_{ni}(t)$$

at each island i , where $\mathbf{X}_{ni}(t)$ is the state of node n at island i , and $N_i = |V_i|$ is the number of nodes at island i . The normalized vector process

$$(\bar{\mathbf{Y}}^{\mathbf{N}}(t)) = (\bar{Y}_1^{\mathbf{N}}(t), \dots, \bar{Y}_M^{\mathbf{N}}(t))$$

collects the corresponding fractions of infected nodes per island, where

$$\bar{Y}_i^{\mathbf{N}}(t) = \frac{Y_i^{\mathbf{N}}(t)}{N_i}$$

is the fraction of infected nodes at island i at time $t, t \geq 0$. The sequence of macroprocesses $(\mathbf{Y}^{\mathbf{N}}(t))_{\mathbf{N}}$ is indexed by

$$\mathbf{N} = (N_1, \dots, N_M),$$

the vector collecting the cardinality N_i of each island i in the underlying complete-multipartite network G^N . It turns out that, from the microscopic model of peer-to-peer infection previously described, $(\mathbf{Y}^N(t))$ is a jump Markov process, and, therefore, it is uniquely determined by its transition rate matrix Q and initial condition (refer to [20]), where

$$Q(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{v}) := \lim_{h \downarrow 0} \frac{\mathbb{P}(\mathbf{Y}^N(t+h) + \mathbf{v} | \mathbf{Y}^N(t))}{h}.$$

Note that two or more events – infection or healing of a node – happen at the same time with probability zero, that is, the evolution of the vector macroprocess $(\mathbf{Y}^N(t))$ is driven almost surely through unit jumps of e_i at each time, where $e_i \in \mathbb{R}^M$ is the canonical vector with the i th entry equal to 1 and the remaining entries equal to 0. Namely, and as we justify next,

$$Q(\mathbf{Y}^N(t), \mathbf{Y}^N(t) - e_i) = \mu_i Y_i^N(t) \quad (3.1)$$

$$Q(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + e_i) = \left(\sum_{j \sim i} \gamma_{ji} Y_j^N(t) \right) \left(\frac{N_i - Y_i^N(t)}{N_i} \right) \quad (3.2)$$

$$Q(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{v}) = 0, \quad \forall \mathbf{v} \in \mathbb{N}^M : \|\mathbf{v}\|_1 > 1, \quad (3.3)$$

where the lowercase subindexes in equations (3.1)-(3.2) refer to islands.

In equation (3.1), we represent the rate to decrease the infected population at island i by one. This happens once any infected node from island i heals,

$$T_i^h = \min \left\{ T_i^h(n) : X_{ni}^N(t) = 1, n \in i \right\} \sim \text{Exp}(\mu_i Y_i^N(t)).$$

In equation (3.2), we represent the rate to increase by one the infected population at island i . In this case, each neighboring island $j \in \mathcal{N}(i)$ of i will have $Y_j^N(t)$ infected nodes and, thus, after a time

$$T_i^c = \min \left\{ T_{ji}^c(n) : X_{nj}^N(t) = 1, n \in j, j \sim i \right\} \sim \text{Exp} \left(\sum_{j \sim i} \gamma_{ji} Y_j^N(t) \right), \quad (3.4)$$

an attempt of infection will be made by a neighboring node at some neighboring island, where $T_{ji}^c(n) \sim \text{Exp}(\gamma_{ji})$ is the time that the infected node $n \in j$ takes to make an attempt of infection towards a node at island i . The *minimum* in equation (3.4) runs over all infected nodes in all the neighboring islands of i . The rate at which an infection from the neighboring islands takes to strike island i is, thus,

$$\left(\sum_{j \sim i} \gamma_{ji} Y_j^N(t) \right).$$

From the exclusion principle, referred in the microscopic model description, if an infection is transmitted to

an already infected node, then the state of the latter node remains unchanged, that is, the infected population does not increase. Therefore, the effective rate of infection will be the rate at which infections arriving at island i hit a healthy node, that is, it is given by

$$\left(\sum_{j \sim i} \gamma_{ji} Y_j^{\mathbf{N}}(t) \right) \times \left(\frac{N_i - Y_i^{\mathbf{N}}(t)}{N_i} \right),$$

the arrival rate of infection at island i times the probability of hitting a healthy node, which is equal to the fraction of healthy nodes at island i at time t , since the chosen target node is drawn uniformly randomly. Note that the topology of the underlying network impacts the increasing rate $Q(\mathbf{Y}^{\mathbf{N}}(t), \mathbf{Y}^{\mathbf{N}}(t) + e_i)$, whereas the decreasing rate $Q(\mathbf{Y}^{\mathbf{N}}(t), \mathbf{Y}^{\mathbf{N}}(t) - e_i)$ only relies on the number of infected nodes at the network at time t , regardless of the peer-to-peer connections.

The goal of the next Chapter is to explore the Markov structure of the vector macroprocess $(\mathbf{Y}^{\mathbf{N}}(t))$ to establish weak convergence as the number of nodes per island \mathbf{N} goes to infinite keeping the underlying generator graph G (and thus, the number of islands M) fixed. Namely, the normalized process $\bar{\mathbf{Y}}^{\mathbf{N}}(t)$ admits a decomposition in terms of a martingale and a drift term that is a functional built upon the transition rates previously presented,

$$\bar{\mathbf{Y}}^{\mathbf{N}}(t) = \bar{\mathbf{Y}}^{\mathbf{N}}(0) + \underbrace{\bar{\mathbf{M}}^{\mathbf{N}}(t)}_{\text{martingale}} + \underbrace{\int_0^t F(\bar{\mathbf{Y}}^{\mathbf{N}}(s-)) ds}_{\text{drift}}.$$

This decomposition is also known as Dynkin's formula (refer to [21]) and holds for any Markov process. In the next Section, we provide an explicit pathwise characterization for the process $\bar{\mathbf{Y}}^{\mathbf{N}}(t)$, and in Chapters 4-5, we will explore this characterization, in particular, to establish that the underlying martingale converges weakly to zero as the underlying $G^{\mathbf{N}}$ complete-multipartite network grows large. As a result (to be also proved), the vector process will converge weakly to the solution of the deterministic ordinary differential equation

$$\frac{d}{dt} \bar{\mathbf{Y}}(t) = F(\bar{\mathbf{Y}}(t)),$$

where the vector field F will be characterized.

We briefly summarize the notation used throughout this Section and introduce useful notation for the next Section and further Chapters.

- $Y_i^{\mathbf{N}}(t)$: number of infected nodes at island i at time t , $t \geq 0$. The boldface upperscript

$$\mathbf{N} = (N_1, N_2, \dots, N_M)$$

stands for a vector stacking the number of nodes N_i at each island i .

- $\bar{Y}_i^{\mathbf{N}}(t)$: fraction of infected nodes at island i at time $t, t \geq 0$,

$$\bar{Y}_i^{\mathbf{N}}(t) = \frac{Y_i^{\mathbf{N}}(t)}{N_i}.$$

- $\mathbf{Y}^{\mathbf{N}}(t)$: vector stacking the number of infected nodes at each island i at time $t, t \geq 0$,

$$\mathbf{Y}^{\mathbf{N}}(t) = (Y_1^{\mathbf{N}}, \dots, Y_M^{\mathbf{N}}(t)).$$

- $\bar{\mathbf{Y}}^{\mathbf{N}}(t)$: vector stacking the fraction of infected nodes at each island i at time $t, t \geq 0$,

$$\bar{\mathbf{Y}}^{\mathbf{N}}(t) = (\bar{Y}_1^{\mathbf{N}}, \dots, \bar{Y}_M^{\mathbf{N}}(t)).$$

- $\mathcal{B}(\mathbb{R}^n)$: Borel σ -algebra over \mathbb{R}^n with the standard topology.
- $\text{Leb}(\cdot) : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$: Lebesgue measure over the real line \mathbb{R} .
- $\mathcal{N}_\alpha : \mathcal{B}_{[0, \infty)} \rightarrow \mathbb{N}$: Poisson point process with rate $\alpha, \alpha > 0$. The random variable $\mathcal{N}_\alpha(A)$ counts the number of *events* in the Borel-set $A \in \mathcal{B}_{[0, \infty)}$. We index the elements of a family of independent Poisson point processes by an upperscript $\mathcal{N}_\alpha^{(i)}$.

3.2 Pathwise Representation

In this Section, we provide a pathwise characterization for the macroprocess $(\mathbf{Y}^{\mathbf{N}}(t))$ built from the microscopic diffusion model. First, we present the relevant definitions regarding point processes over the real line \mathbb{R} that will be the building blocks for the pathwise description of $(\mathbf{Y}^{\mathbf{N}}(t))$. For more details, refer to [26].

Definition 13 (Point measure; [26]). $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ is a point measure on \mathbb{R} if there exists a sequence $a_n \in \mathbb{R}, n \in \mathbb{N}$, so that

$$\mu(A) = \#(A \cap \{a_n\}_{n=1}^\infty) = \sum_{i=1}^\infty \mathbf{1}_{\{a_i \in A\}}, \quad \forall A \in \mathcal{B}(\mathbb{R}),$$

that is, $\mu(A)$ counts the number of points of the sequence a_n in A for any Borelian set $A \in \mathcal{B}(\mathbb{R})$.

We represent the set of point measures on \mathbb{R} as $M_p(\mathbb{R})$. Therefore, for each point measure $\mu \in M_p(\mathbb{R})$ there exists an underlying real sequence $(a_n)_{n \in \mathbb{N}}$. A point measure μ on \mathbb{R} is called Radon, if each compact interval entails only a finite number of elements of the associated sequence, or in other words the set of accumulation points of (a_n) is empty. We assume that $M_p(\mathbb{R})$ is endowed with the topology of weak convergence of measures described in Chapter 2.

We now define a point process on the real line \mathbb{R} .

Definition 14 (Point process; [26]). \mathcal{N} is a point process if it is a Radon point measure valued random variable,

$$\begin{aligned}\mathcal{N} : \Omega &\rightarrow M_p(\mathbb{R}) \\ \omega &\rightarrow \mathcal{N}(\omega, \cdot),\end{aligned}$$

where $M_p(\mathbb{R})$ is endowed with the Borel σ -algebra $\mathcal{B}(M_p(\mathbb{R}))$.

In words, each realization $\omega \in \Omega$ leads to a sequence of points in the real line (void of accumulation points) that underlies the point measure $\mathcal{N}(\omega, \cdot)$. For simplicity, we refer to the random measure that counts the number of events in a Borel set as $\mathcal{N}(\cdot)$. Thus, $\mathcal{N}(A)$ is the random measure of $A \in \mathcal{B}(\mathbb{R})$.

We now introduce the definition of Poisson point process that will be central to building the macroscopic process $(\mathbf{Y}^N(t))$ from the local (at node level) rules of infection.

Definition 15 (Poisson point process; [26]). \mathcal{N}_γ is a Poisson point process on \mathbb{R} with rate $\gamma > 0$, if it satisfies the following two conditions

1. **[independent increments]** Given

$$I_1, I_2, \dots, I_n \subset \mathbb{R},$$

n disjoint intervals in the real line, then

$$\mathcal{N}_\gamma(I_1), \mathcal{N}_\gamma(I_2), \dots, \mathcal{N}_\gamma(I_n)$$

are independent random variables.

2. **[increment stationarity]** Let

$$I_1, I_2 \subset \mathbb{R}$$

be two intervals. Then,

$$L := \text{Leb}(I_1) = \text{Leb}(I_2) \Rightarrow \mathcal{N}_\gamma(I_1) \stackrel{d}{\sim} \mathcal{N}_\gamma(I_2),$$

and $\mathcal{N}_\gamma(I_1)$ and $\mathcal{N}_\gamma(I_2)$ are Poisson distributed random variables with rate (or intensity) parameter γL .

Definition 15 implies that the interarrival *time* interval between the elements of the underlying random sequence of the Poisson point process is exponentially distributed with mean $1/\gamma$ (refer to [27]). That is, the random sequence of points underlying the point process is constructed so that the time between events

$$T = a_{n+1} - a_n \sim \text{Exp}(\gamma)$$

is exponentially distributed. We may refer to a Poisson point process \mathcal{N}_γ with rate $\gamma > 0$ as γ -Poisson process.

If a process $Z(t)$ counts the number of events up to time t , $t \geq 0$, from a Poisson source with rate γ , then,

$$Z(t) = \mathcal{N}_\gamma((0, t]).$$

As a concrete example, consider a permanently infected node at island 2 from Figure 3.3. Then, according to the stochastic infection model described in Section 3.1, and if we consider $\gamma_{21} = \gamma$, the process $(Z(t))$ counts the number of infections that arrive at island 1 due to this permanently infected source up to time t , $t \geq 0$. Remark that the infections from island 2 flow to island 1 according to the triggering of exponentially distributed, i.e., $\text{Exp}(\gamma_{21})$, random clocks, that are attached to the infected nodes. Therefore, in this case the process $(Z(t))$ is γ_{21} -Poisson. Now, If instead of a single permanently infected node we had two, then the rate would double and the process counting the number of infections arriving at island 1 would still be Poisson given by

$$Z(t) = \mathcal{N}_{2\gamma}((0, t]).$$

If the process $(Z(t))$ counts the number of events from a *piecewise varying* Poisson source, that is, if the real line \mathbb{R} can be partitioned into intervals $\cup_{i=1}^{\infty} I_i = \mathbb{R}$ so that at each time interval I_i there is a γ_i -Poisson source $\mathcal{N}^{(i)}$ acting upon it, then

$$Z(t) = \sum_{i=1}^{\infty} \mathcal{N}_{\gamma_i}^{(i)}(I_i \cap (0, t]). \quad (3.5)$$

Recalling that $\mu(A) = \int \mathbf{1}_A d\mu$ for any measure μ (e.g., Lebesgue measure) then, from equation (3.5) and the fact that $\mathcal{N}_{\gamma_i}^{(i)}$ are random measures, we have the following integral characterization for the process $(Z(t))$

$$Z(t) = \sum_{i=1}^{\infty} \int_0^t \mathbf{1}_{\{s \in I_i \cap (0, t]\}} \mathcal{N}_{\gamma_i}^{(i)}(ds),$$

where the integrals are taken with respect to the Poisson random measures $\mathcal{N}_{\gamma_i}^{(i)}$. That is, for each realization $\omega \in \Omega$, $\mathcal{N}_{\gamma_i}^{(i)}(\omega, \cdot)$ is a measure over \mathbb{R} , and that is the measure under which the integral is defined

$$Z(\omega, t) = \sum_{i=1}^{\infty} \int_0^t \mathbf{1}_{\{s \in I_i \cap (0, t]\}} \mathcal{N}_{\gamma_i}^{(i)}(ds, \omega). \quad (3.6)$$

For notational simplicity, we will omit the realization component ω from such equations. In fact, in this Thesis, all equalities between processes under the symbol '=' are to be taken as pointwise in $\omega \in \Omega$, as in equation (3.6). For instance, they cannot be interpreted as equalities in distribution. For this latter sense, we reserve the symbol ' $\stackrel{d}{=}$ '. That is,

$$Y(t) \stackrel{d}{=} Z(t) \iff \mathbb{P}(Y(t) \in B) = \mathbb{P}(Z(t) \in B), \forall B \text{ Borel.}$$

For the virus spread case, we have to consider that the number of infected nodes at island 2 changes over time (instead of the static example with permanently infected sources), either due to new infections or healing. In this case, the process $(Z(t))$ counting the number of infections arriving at island 1 – but not necessarily hitting healthy nodes, i.e., increasing the infected population at 1 – due to infected nodes in island 2 is given by

$$Z(t) = \sum_{k=1}^{N_1} \int_0^t \mathbf{1}_{\{Y_2^N(s-) = k\}} \mathcal{N}_{k\gamma}^{(k)}(ds), \quad (3.7)$$

where, in this case, $\mathcal{N}_\gamma^{(1)}, \mathcal{N}_{2\gamma}^{(2)}, \dots, \mathcal{N}_{N_1\gamma}^{(N_1)}$ are **independent** Poisson processes. That is, we partitioned the real line up to time t ,

$$(0, t] = \bigcup_{k=1}^{N_1} \{0 < s \leq t : Y_2^N(s-) = k\}$$

according to the number of infected nodes $Y_2^N(s-) = k$ at island 2 during the time interval $(0, t]$. During the time intervals where island 2 has k infected nodes, $Y_2^N(s-) = k$, the source of infection that strikes island 1 due to island 2 is $(k\gamma)$ -Poisson. Equation (3.7) represents a sample path characterization for the process $(Z(t))$ that counts the number of infections that flow from island 1 towards island 2 up to time t , $t \geq 0$. It refers to the total number of infections from 2 to 1, and not only to the subset of infections that hit healthy nodes in island 1 and increases the infected population in this island.

Now, if each event a_n from a γ -Poisson source is only counted with probability p then,

$$Z(t) = \mathcal{N}_{\gamma p}((0, t]). \quad (3.8)$$

Namely, in this case $(Z(t))$ may count the number of infections from a permanently infected node from island 2 to 1 that strike healthy nodes, assuming that a fraction of $(1 - p)$ of the nodes at the sink island 1 are always infected.

We now consider all these effects together to build the sample path characterization of our macroprocess $(Y^N(t))$. For the sake of simplicity consider the complete-bipartite network single virus case, i.e., there are only islands 1 and 2 connected to each other. Let $(I_1(t))$ and $(H_1(t))$ be the processes counting the number of nodes that are or were infected (at least once) and number of healings, respectively, up to time t , $t \geq 0$, at island 1. We have

$$Y_1^N(t) - Y_1^N(0) = I_1(t) - H_1(t) \quad (3.9)$$

$$\begin{aligned} &= \sum_{\ell=1}^{N_1} \sum_{q=1}^{N_2} \mathcal{N}_{\gamma_{21}q}^{(\ell, q)} \left(\frac{N_1 - \ell}{N_1} \right) (\{0 \leq s \leq t : Y_1^N(s) = \ell, Y_2^N(s) = q\}) \\ &\quad - \sum_{\ell=1}^{N_1} \mathcal{N}_{\mu\ell}^{(\ell)} (\{0 \leq s \leq t : Y_1^N(s) = \ell\}) \end{aligned} \quad (3.10)$$

$$\begin{aligned}
&= \sum_{\ell=1}^{N_1} \sum_{q=1}^{N_2} \int_0^t \mathbf{1}_{\{Y_1^{\mathbf{N}}(s-)=\ell, Y_2^{\mathbf{N}}(s-)=q\}} \left(\mathcal{N}_{\gamma_{21}q\left(\frac{N_1-\ell}{N_1}\right)}^{(\ell,q)}(ds) \right) \\
&\quad - \sum_{\ell=1}^{N_1} \int_0^t \mathbf{1}_{\{Y_1^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\mu\ell}^{(\ell)}(ds) \right). \tag{3.11}
\end{aligned}$$

Note that the difference between the process $(I_1(t))$ and the process $(Z(t))$ in equation (3.7) is that $(Z(t))$ counts the number of arrival infections that may also hit already infected nodes, whereas $(I_1(t))$ counts the number of effective infections that hit healthy nodes and, thus, increases the infected population. In the latter case, one has to account for the effect described in equation (3.8), where only a fraction

$$\frac{N_1 - k_1}{N_1}$$

of nodes at island 1 is healthy at time t , if $Y_1^{\mathbf{N}}(t) = k_1$.

For a general complete-multipartite network with M islands, the pathwise dynamics of the single-virus epidemics is given by

$$\begin{aligned}
Y_i^{\mathbf{N}}(t) &= Y_i^{\mathbf{N}}(0) + \underbrace{\sum_{j=1}^M \sum_{q=1}^{N_j} \sum_{\ell=1}^{N_i} \int_0^t \sum_{j \sim i} \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\gamma_{ji}q\left(\frac{N_i-\ell}{N_i}\right)}^{(\ell,q)}(ds) \right)}_{\text{Inter-transmission}} \\
&\quad - \underbrace{\sum_{\ell=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_i^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\ell\mu_i}^{(\ell)}(ds) \right)}_{\text{Healing}}, \tag{3.12}
\end{aligned}$$

where all Poisson point processes

$$\mathcal{N}_{\gamma_{ji}q\left(\frac{N_i-\ell}{N_i}\right)}^{(\ell,q)} \text{ and } \mathcal{N}_{\ell\mu_i}^{(\ell)}$$

indexed by ℓ and q are **independent**. This is an important fact from the peer-to-peer model that will be evoked later in Chapters 4 and 5. Also, for notational simplicity, we drop from now on the upper-indexes of the Poisson processes. The *inter-transmission* term in equation (3.12) accounts for the number of infections transmitted from island j to healthy nodes in island i up to time t , $t \geq 0$. The *healing* term accounts for the number of healing's that occur in island i during the time interval $[0, t]$. One can check that almost surely the normalized process

$$\bar{Y}_i^{\mathbf{N}}(t) = \frac{Y_i^{\mathbf{N}}(t)}{N_i} \in [0, 1], \quad \forall t \geq 0$$

if the initial condition $\bar{Y}_i^{\mathbf{N}}(0) \in [0, 1]$ almost surely, that is, the set $[0, 1]^M$ is invariant for the stochastic dynamics of $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ (refer to [17]). In words, the macroscopic stochastic dynamics are consistent with our intuition about the underlying meaning of $(\bar{Y}_i^{\mathbf{N}}(t))$ that is the fraction of infected nodes at island i , and so it is clearly a quantity between 0 and 1, for all t , $t \geq 0$. Observe that the inter-transmission term is the

one that relies on the supertopology of the complete-multipartite network.

The pathwise characterization for the general case of multi-virus epidemics over complete-multipartite networks is now given by the $M \times K$ stochastic integral equations

$$\begin{aligned}
 Y_{ik}^{\mathbf{N}}(t) = & Y_{ik}^{\mathbf{N}}(0) + \underbrace{\sum_{j=1}^M \sum_{q=1}^{N_j} \sum_{\ell=1}^{N_i} \int_0^t \sum_{j \sim i} \mathbf{1}_{\{Y_{jk}^{\mathbf{N}}(s-)=q, \sum_{l=k}^K Y_{il}^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\gamma_{ji}^k q \left(\frac{N_i - \ell}{N_i} \right)}^{(\ell, q)}(ds) \right)}_{\text{Inter-transmission}} \\
 & - \underbrace{\sum_{\ell=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_{ik}^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\ell \mu_i^k}^{(\ell)}(ds) \right)}_{\text{Healing}},
 \end{aligned} \tag{3.13}$$

where $Y_{ik}^{\mathbf{N}}(t)$ stands for the number of k -infected nodes at island i at time t , $t \geq 0$, with

$$k \in \{1, \dots, K\} \text{ and } i \in \{1, \dots, M\}.$$

3.3 Concluding Remarks

In this Chapter, we presented the peer-to-peer stochastic network model assumed in this Thesis, under which the epidemics spreads. We described in Section 3.1 that nodes heal or infect their neighbors according to time-triggering clocks attached to them. These clocks are exponentially distributed with rates that depend on the islands and the virus strain, but not on the particular identity of the node. The topology of contacts that we assume is a complete-multipartite network. In Section 3.2, we provided a pathwise characterization for the process $(\mathbf{Y}^{\mathbf{N}}(t))$ that will be explored in Chapters 4 and 5 to establish convergence. In particular, in Chapters 4 and 5, we frame the normalized martingale term $(\overline{\mathbf{M}}^{\mathbf{N}}(t))$ hidden within the pathwise characterization of the normalized process $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$, and we prove that it converges weakly to zero as the network $G^{\mathbf{N}}$ grows large. This loosely implies that the randomness of the process $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$ dies out as the number of agents grows.

Chapter 4

Single-virus over Complete-bipartite Networks

In this Chapter, we focus on the single virus spread over complete-bipartite networks $G^{(N_1, N_2)}$ to establish that the empirical distribution sequence

$$\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right) = \left(\bar{\mathbf{Y}}_1^{\mathbf{N}}(t), \bar{\mathbf{Y}}_2^{\mathbf{N}}(t)\right),$$

defined over the compact interval $[0, T]$, converges weakly, as the network $G^{(N_1, N_2)}$ grows large, to the solution of a deterministic vector differential equation restricted to the time interval $[0, T]$. By the *network grows large*, we mean that the sizes N_1 and N_2 of the two islands of the complete-bipartite network go to infinity,

$$N_1, N_2 \rightarrow \infty,$$

with a finite asymptotic ratio

$$\frac{N_i}{N_j} \longrightarrow \alpha_{ji} < \infty.$$

We remark that the Markov jump process $\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right)$ admits a decomposition into a martingale term plus a drift term – this latter is obtained from the transition rates as characterized in Chapter 3 – as represented in equation (4.2) in Section 4.1.

In Section 4.1, we single out the pathwise representation of the martingale term from the equation (3.11) and establish the weak convergence to zero of its normalized counterpart

$$\left(\bar{\mathbf{M}}^{\mathbf{N}}(t)\right) = \left(\frac{M_1^{\mathbf{N}}(t)}{N_1}, \frac{M_2^{\mathbf{N}}(t)}{N_2}\right)$$

as the complete-bipartite network grows to infinite. In Section 4.2, we show that as a consequence $\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right)$ is a tight family indexed by \mathbf{N} whose accumulation points (w.r.t. the weak convergence) are necessarily

given by the solutions of a differential equation. By uniqueness of the resulting differential equation (the vector field is globally-Libpschitz), any convergent subsequence converges to the solution of this differential equation. Therefore, since the limiting process is unique, it follows that the whole sequence converges to the solution of the differential equation and the weak convergence of the normalized process $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is proved. Finally, Section 4.3 summarizes the Chapter.

To sum up, the main convergence result of this Chapter is stated in Theorem 26 whose proof is comprised of four main steps:

- (i) The martingale $(\bar{M}_1^{\mathbf{N}}(t), \bar{M}_2^{\mathbf{N}}(t))$ converges weakly to zero, see Section 4.1;
- (ii) The family $(\bar{Y}_1^{\mathbf{N}}(t), \bar{Y}_2^{\mathbf{N}}(t))$, indexed by $\mathbf{N} \in \mathbb{N}^2$, is tight, see Section 4.2;
- (iii) Any accumulation point of the tight family is solution to the ODE (4.26)-(4.27), see Section 4.2;
- (iv) Uniqueness of the ODE (4.26)-(4.27) implies convergence, see Section 4.2.

We extend the analysis to the multivirus epidemics over general complete-multipartite networks in the next Chapter, Chapter 5.

Preliminary Notation. To prevent the use of heavy notation, we adopt the following assumptions throughout, unless otherwise stated:

- $\mathbf{N} = (N_1, N_2) \longrightarrow \infty$ refers to $N_1, N_2 \rightarrow \infty$ with a finite asymptotic ratio $N_i/N_j \rightarrow \alpha_{ji} < \infty$.
- Let $\mathbf{N} \longrightarrow \infty$ in the sense defined in the previous bullet. We define

$$\lim_{\mathbf{N} \longrightarrow \infty} \sup F^{\mathbf{N}} := \inf_{\mathbf{K} \in \mathbb{N}^2} \sup_{\mathbf{N} \geq \mathbf{K}} F^{\mathbf{N}}. \quad (4.1)$$

Note that it is well defined as the right hand-side in (4.1) does not depend on the sequence $\mathbf{N} \longrightarrow \infty$.

- We refer to “weak convergence with respect to the Skorokhod topology $D_{\mathbb{R}}[0, T]$ over the compact interval $[0, T]$ ” simply by “weak convergence”.

- $(\bar{\mathbf{Y}}^{\mathbf{N}}(t)) \Rightarrow (\bar{\mathbf{Y}}(t))$ reads as “ $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ converges weakly to $(\bar{\mathbf{Y}}(t))$ ” as $\mathbf{N} \longrightarrow \infty$.
- We denote

$$(\bar{\mathbf{Y}}^{\mathbf{N}}(\omega, t))$$

as the càdlàg sample path associated with the realization $\omega \in \Omega$.

- All processes are studied over the compact interval $[0, T]$. For instance, we refer to

$$\begin{aligned} (\bar{\mathbf{Y}}^{\mathbf{N}}(t)) : \quad \Omega &\longrightarrow D_{[0, T]} \\ \omega \in \Omega &\longmapsto (\bar{\mathbf{Y}}^{\mathbf{N}}(\omega, t)) \end{aligned},$$

as the process associated with the evolution of the fraction of infected nodes in $G^{(N_1, N_2)}$ **up to time T** .

- $(Z(t)) \perp (W(t))$ refers to the independence of the corresponding random variables $Z(t)$ and $W(t)$ for every $t, t \in [0, T]$, and reads as “the stochastic processes $(Z(t))$ and $(W(t))$ are independent on the line.

4.1 Martingale Vanishes in Probability and Weakly

We start by characterizing the martingale term $(\mathbf{M}^{\mathbf{N}}(t))$ of our macroprocess $(\mathbf{Y}^{\mathbf{N}}(t))$

$$\mathbf{Y}^{\mathbf{N}}(t) = \mathbf{Y}^{\mathbf{N}}(0) + \mathbf{M}^{\mathbf{N}}(t) + \int_0^t F(\mathbf{Y}^{\mathbf{N}}(s-)) ds, \quad (4.2)$$

and afterwards we explore its structure to prove that its normalized counterpart

$$(\overline{\mathbf{M}}^{\mathbf{N}}(t)) := \left(\frac{M_1^{\mathbf{N}}(t)}{N_1}, \frac{M_2^{\mathbf{N}}(t)}{N_2} \right)$$

converges weakly to 0 as the number of nodes at each island goes to infinite. In words, this means that the randomness of the normalized macroprocess $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$ dies out as the network grows large. We start by showing that $\overline{\mathbf{M}}^{\mathbf{N}}(t)$ converges to zero in \mathcal{L}_2 for each $t, t \geq 0$. As mentioned in Chapter 2, we refer to it as *the martingale converges to zero in \mathcal{L}_2 on the line*. Then, by Doob's inequality, this will imply that it converges to zero in probability, with respect to the **uniform** topology, and thus, in the Skorokhod topological space $D_{[0,T]}$, as will be clearer momentarily. Finally, from Theorem 19, this implies that the martingale converges weakly to zero.

From Chapter 3, equation (3.11), the stochastic vector process $(\mathbf{Y}^{\mathbf{N}}(t))$ over the complete-bipartite network admits the following pathwise characterization:

$$\begin{aligned} Y_i^{\mathbf{N}}(t) = & Y_i^{\mathbf{N}}(0) + \underbrace{\sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \int_0^t \mathbf{1}_{\{Y_i^{\mathbf{N}}(s-)=\ell, Y_j^{\mathbf{N}}(s-)=q\}} \left(\mathcal{N}_{\gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right)}(ds) \right)}_{\text{Inter-transmission}} \\ & - \underbrace{\sum_{\ell=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_i^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\ell\mu_i}^{(\ell)}(ds) \right)}_{\text{Healing}}, \end{aligned} \quad (4.3)$$

for $i, j \in \{1, 2\}$ and $i \neq j$.

Now, we explicitly characterize the martingale

$$(\mathbf{M}^{\mathbf{N}}(t)) = (M_1^{\mathbf{N}}(t), M_2^{\mathbf{N}}(t))$$

in the Dynkin's decomposition of $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$:

$$Y_i^{\mathbf{N}}(t) = Y_i^{\mathbf{N}}(0) + M_i^{\mathbf{N}}(t) + \int_0^t \gamma_{ji} Y_j^{\mathbf{N}}(s-) \left(\frac{N_i - Y_i^{\mathbf{N}}(s-)}{N_i} \right) ds - \int_0^t \mu_i Y_i^{\mathbf{N}}(s-) ds, \quad (4.4)$$

for $i, j = 1, 2$, with $i \neq j$. It is well-known that the so-called compensated Poisson process

$$M(t) = \mathcal{N}_\gamma((0, t]) - \gamma t$$

is a martingale adapted to the natural filtration (refer to Proposition 1.21 from [26])

$$\Sigma_t := \sigma \{ \mathcal{N}_\gamma(s) : 0 \leq s \leq t \}$$

of $(\mathcal{N}_\gamma(t))$. The martingale property of a compensated Poisson process is stable under the integration of a predictable process, [26]; in particular, if $(Y(t))$ is a càdlàg process predictable w.r.t. Σ_t , then

$$M(t) = \int_0^t Y(s-) (\mathcal{N}_\lambda(ds) - \lambda ds)$$

is a martingale adapted to the natural filtration $\sigma \{ \mathcal{N}_\lambda(s) : 0 \leq s \leq t \}$ (refer to [28]).

We now apply these results to the underlying Poisson point processes in equation (4.3) to obtain

$$\begin{aligned} Y_i^{\mathbf{N}}(t) &= Y_i^{\mathbf{N}}(0) + \underbrace{\sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right)}(ds) - \gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right) ds \right)}_{M_{ai}^{\mathbf{N}}(t)} \\ &\quad - \underbrace{\sum_{k=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_i^{\mathbf{N}}(s-)=k\}} (\mathcal{N}_{k\mu_i}(ds) - k\mu_i ds)}_{M_{bi}^{\mathbf{N}}(t)} \\ &\quad + \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right) ds - \sum_{k=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_i^{\mathbf{N}}(s-)=k\}} k\mu_i ds. \end{aligned} \quad (4.5)$$

where, $M_{ai}^{\mathbf{N}}(t)$ and $M_{bi}^{\mathbf{N}}(t)$ are martingales since the space \mathcal{M} of martingales adapted to the same filtration conforms to a vector space and

$$M_i^{\mathbf{N}}(t) := M_{ai}^{\mathbf{N}}(t) - M_{bi}^{\mathbf{N}}(t) \quad (4.6)$$

is a martingale. Therefore, the normalized martingale

$$\overline{\mathbf{M}}^{\mathbf{N}}(t) = \left(\overline{M}_1^{\mathbf{N}}(t), \overline{M}_2^{\mathbf{N}}(t) \right)$$

with

$$\overline{M}_1^{\mathbf{N}}(t) = \frac{M_1^{\mathbf{N}}(t)}{N_1} \text{ and } \overline{M}_2^{\mathbf{N}}(t) = \frac{M_2^{\mathbf{N}}(t)}{N_2}$$

is now characterized. The terms $M_{ai}^{\mathbf{N}}(t)$ and $M_{bi}^{\mathbf{N}}(t)$ in equation (4.5) are martingales with respect to the natural filtration

$$\mathcal{F}_t = \sigma \left\{ \mathcal{N}_{\gamma_{ji}q\left(\frac{N_i-\ell}{N_i}\right)}^{(\ell,q)}((0,s)), \mathcal{N}_{\ell\mu_i}^{(\ell)}((0,s)) : i, j \in \{1, \dots, M\}; \ell \leq N_i, q \leq N_j; s \leq t \right\} \quad (4.7)$$

of the underlying family of Poisson point processes

$$\mathcal{N}_{\gamma_{ji}q\left(\frac{N_i-\ell}{N_i}\right)}^{(\ell,q)} \text{ and } \mathcal{N}_{\ell\mu_i}^{(\ell)},$$

respectively, indexed by $\ell \leq N_i$ and $q \leq N_j$. All processes that follow are càdlàg and (\mathcal{F}_t) -adapted.

Before proceeding, we evoke the next Theorem, Theorem 16, that will be important for the proof that the sequence of normalized martingales $\left(\overline{\mathbf{M}}^{\mathbf{N}}(t)\right)$ converges to zero in \mathcal{L}_2 on the line as stated in Theorem 17. Theorem 16 essentially states that orthogonality of compensated Poisson martingales is preserved under integration of some predictable processes. Theorem 16 can be found in [26] as Proposition A.10. We restate it here within our context and for completeness we provide with a proof in Appendix.

Theorem 16 (Orthogonality; Proposition A.10 in [26]). *Let $(\mathbf{Y}(t))$ be an (\mathcal{F}_t) -adapted càdlàg process with discrete range and piecewise constant (i.e., constant when it does not jump). Let $\mathcal{N}_\lambda(t)$ and $\mathcal{N}_\mu(t)$ be two independent (\mathcal{F}_t) -adapted Poisson processes (hence their compensated versions are (\mathcal{F}_t) -martingales, as it is trivial to establish). Assume the rates λ, μ are nonnegative. Let f, g be two bounded functions defined over the discrete range of $\mathbf{Y}(t)$. Then,*

$$\left(\int_0^t f(\mathbf{Y}(s-)) (\mathcal{N}_\lambda(ds) - \lambda ds) \int_0^t g(\mathbf{Y}(s-)) (\mathcal{N}_\mu(ds) - \mu ds) \right) \quad (4.8)$$

is an (\mathcal{F}_t) -martingale.

Next, we prove that the variances of the normalized zero-mean martingales

$$\left(\overline{M}_i^{\mathbf{N}}(t)\right) = \left(\frac{M_i^{\mathbf{N}}(t)}{N_i}\right)$$

converge to zero (as \mathbf{N} grows large) for all $t, t \geq 0$, and $i = 1, 2$, that is, $\overline{M}_i^{\mathbf{N}}(t)$ converges to zero in \mathcal{L}_2 on the line.

Theorem 17. *Let*

$$\left(\overline{M}_i^{\mathbf{N}}(t)\right) = \left(\overline{M}_{ai}^{\mathbf{N}}(t) - \overline{M}_{bi}^{\mathbf{N}}(t)\right)$$

be the normalized version of the martingale in equation (4.5). Then, we have that

$$\left(\overline{\mathbf{M}}^{\mathbf{N}}(t)\right) \xrightarrow{\mathcal{L}_2} 0,$$

or equivalently,

$$\text{Var} \left(\overline{M}_i^{\mathbf{N}}(t) \right) = \mathbb{E} \left(\overline{M}_i^{\mathbf{N}}(t) \right)^2 \longrightarrow 0, \quad \forall t \geq 0 \text{ and } i = 1, 2$$

as $N_i \rightarrow \infty$ for $i = 1, 2$ with $N_i/N_j \rightarrow \alpha_{ji} < \infty$.

Proof. Note first that for $s < t$

$$\mathbb{E} \left(M_i^{\mathbf{N}}(t) \right) = \mathbb{E} \left(\mathbb{E} \left(M_i^{\mathbf{N}}(t) | \mathcal{F}_s \right) \right) = \mathbb{E} \left(M_i^{\mathbf{N}}(s) \right).$$

Thus, $(M_i^{\mathbf{N}}(t))$ is zero-mean,

$$\mathbb{E} \left(M_i^{\mathbf{N}}(t) \right) = \mathbb{E} \left(M_i^{\mathbf{N}}(0) \right) = 0,$$

for all $t \geq 0$; therefore,

$$\text{Var} \left(\overline{M}_i^{\mathbf{N}}(t) \right) = \mathbb{E} \left(\overline{M}_i^{\mathbf{N}}(t) \right)^2.$$

Now,

$$\begin{aligned} \mathbb{E} \left(M_{ai}^{\mathbf{N}}(t) \right)^2 &= \mathbb{E} \left(\sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right)}(ds) - \gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right) ds \right) \right)^2 \\ &= \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbb{E} \left(\int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right)}(ds) - \gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right) ds \right) \right)^2 \\ &= \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbb{E} \left(\int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right) ds \right) \\ &\leq \mathbb{E} \left(\int_0^t \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \gamma_{ji}N_j ds \right) \\ &\leq \gamma_{ji}N_j t, \end{aligned}$$

where the second equality follows from Theorem 16 and the independence of all the underlying Poisson processes involved. The third equality is due to the Itô isometry Theorem (refer to [28] or [29]) and the fact that the quadratic variation of a compensated Poisson martingale is given by

$$\langle \mathcal{N}_\gamma(t) - \gamma t \rangle = \gamma t.$$

The first inequality is due to

$$q \leq N_j \text{ and } \frac{N_i - \ell}{N_i} \leq 1.$$

The last inequality holds since the family of subsets of \mathbb{R}

$$I_{\ell,q}(\omega) := \{s \in \mathbb{R} : Y_j^{\mathbf{N}}(\omega, s-) = q, Y_i^{\mathbf{N}}(\omega, s-) = \ell\},$$

indexed by ℓ, q , are realization-wise disjoint and thus

$$\sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbf{1}_{\{Y_j^{\mathbf{N}}(\omega, s-) = q, Y_i^{\mathbf{N}}(\omega, s-) = \ell\}} = \mathbf{1}_{\bigcup_{\ell, q} \{Y_j^{\mathbf{N}}(\omega, s-) = q, Y_i^{\mathbf{N}}(\omega, s-) = \ell\}} \leq \mathbf{1}_{[0, T]}(s).$$

for all $\omega \in \Omega$. Therefore,

$$\mathbb{E} \left(\overline{M}_{ai}^{\mathbf{N}}(t) \right)^2 = \frac{1}{N_i^2} \mathbb{E} \left(M_{ai}^{\mathbf{N}}(t) \right)^2 \leq \frac{\gamma_{ji}}{N_i} \left(\frac{N_j}{N_i} \right) t.$$

Thus, for all fixed $t, t \geq 0$,

$$\mathbb{E} \left(\overline{M}_{ai}^{\mathbf{N}}(t) \right)^2 \longrightarrow 0$$

as $N_i \rightarrow \infty$ and $N_j \rightarrow \infty$, and

$$\frac{N_j}{N_i} \rightarrow \alpha_{ij} < \infty.$$

Similarly, the variance of the martingale $\left(\overline{M}_{bi}^{\mathbf{N}}(t) \right)$ converges to zero on the line. Thus, the martingale vanishes in \mathcal{L}_2 with at least $O(1/N_i)$. \square

As a corollary, it follows that the $\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right)$ converges to zero in probability, with respect to the uniform (and hence Skorokhod) topology, as stated in the next Corollary.

Corollary 18. *The sequence of martingales $\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right)$ converges to zero in probability w.r.t. the Skorokhod topology*

$$\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right) \xrightarrow{\mathbb{P}} 0.$$

Proof. From Doob's inequality:

$$P \left(\sup_{0 \leq t \leq T} \left| \overline{M}_i^{\mathbf{N}}(t) \right| > \epsilon \right) \leq \frac{\mathbb{E} \left(\overline{M}_i^{\mathbf{N}}(T) \right)^2}{\epsilon^2} \longrightarrow 0, \quad \forall \epsilon > 0, \quad \forall T \geq 0. \quad (4.9)$$

That is, $\left(\overline{M}_i^{\mathbf{N}}(t) \right)$ converges to zero in probability. \square

The next Theorem is an extension for stochastic processes of the well-known result that convergence in probability implies convergence in distribution for real valued random variables. This Theorem will imply that $\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right)$ converges weakly to 0. In other words, the probability measure, $\mathbb{P}_{\overline{\mathbf{M}}^{\mathbf{N}}}$, pushforwarded onto $D_{\mathbb{R}}[0, T]$ by $\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right)$, converges weakly to δ_0 (Dirac measure about the zero path $m(t) = 0$ over $[0, T]$).

Theorem 19 ([26], Proposition C.5). *Let $\left(Z^{\mathbf{N}}(t) \right)$ be a sequence of càdlàg stochastic processes on the interval $[0, T]$ such that,*

$$\left(Z^{\mathbf{N}}(t) \right) \xrightarrow{\mathbb{P}} (z(t)),$$

where $(z(t))$ is a deterministic càdlàg function over $[0, T]$. Then,

$$(Z^N(t)) \Rightarrow (z(t)),$$

i.e., $(Z^N(t))$ converges weakly to $(z(t))$. In other words, the sequence of probability measures \mathbb{P}_{Z^N} over $D_{\mathbb{R}}[0, T]$ induced by $(Z^N(t))$ converges weakly to $\delta_{(z(t))}$ (Dirac measure about $(z(t))$).

Thus, we conclude that the martingale $(\bar{\mathbf{M}}^N(t))$ converges weakly to zero with respect to the Skorohod space $D_{[0, T]}$.

4.2 Weak convergence

In this Section, we prove that the sequence of processes $(\bar{\mathbf{Y}}^N(t))$, associated with the epidemics over complete-bipartite networks $G^{(N_1, N_2)}$, converges weakly to the solution of the ODE (4.26)-(4.27), below. We start by showing that the family $(\bar{\mathbf{Y}}^N(t))$ indexed by

$$\mathbf{N} = (N_1, N_2)$$

is tight. The next Theorem is a stochastic version of the well-known Arzelà-Ascoli Theorem (refer to [23]) and will provide sufficient conditions to guarantee the tightness (or more specifically, the C -tightness as defined in Chapter 2) of the sequence $(\bar{\mathbf{Y}}^N(t))$ – just as Arzelà-Ascoli provides sufficient conditions to guarantee the tightness of a family of *deterministic* functions.

Theorem 20 (Arzelà-Ascoli; [30]). *Let $(\bar{\mathbf{Z}}^N(t))$ be a sequence of càdlàg processes. Then, the sequence of probability measures $\mathbb{P}_{\bar{\mathbf{Z}}^N}$ induced on $D_{[0, T]}$ by $(\bar{\mathbf{Z}}^N(t))$ is tight and any weak limit point of this sequence is concentrated on the subset of continuous functions $C_{[0, T]} \subset D_{[0, T]}$ if and only if the following two conditions hold for each $\epsilon > 0$:*

$$\lim_{k \rightarrow \infty} \limsup_{\mathbf{N} \rightarrow \infty} P \left(\sup_{0 \leq t \leq T} |\bar{\mathbf{Z}}^N(t)| \geq k \right) = 0 \text{ (Uniform Boundness)} \quad (4.10)$$

$$\lim_{\delta \rightarrow 0} \limsup_{\mathbf{N} \rightarrow \infty} P \left(\omega(\bar{\mathbf{Z}}^N, \delta, T) \geq \epsilon \right) = 0 \text{ (Equicontinuity)} \quad (4.11)$$

where we defined the modulus of continuity

$$\omega(x, \delta, T) = \sup \{ |x(u) - x(v)| : 0 \leq u, v \leq T, |u - v| \leq \delta \}.$$

We are just left to show that our sequence $(\bar{\mathbf{Y}}^N(t))$ meets the requirements in equations (4.10) and (4.11), and, therefore, it is tight and hence relatively compact by Prohorov's Theorem 9 in Chapter 2. We prove this in the next Theorem, Theorem 21. In other words, the family $(\bar{\mathbf{Y}}^N(t))$ admits a weakly convergent

subsequence

$$\left(\bar{\mathbf{Y}}^{\mathbf{N}_k}(t)\right) \Rightarrow (\bar{\mathbf{Y}}(t))$$

where $(\bar{\mathbf{Y}}(t))$ is almost surely continuous.

Theorem 21. *The family of vector processes $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ indexed by \mathbf{N} is C -tight, i.e., all of its weak-accumulation points lie almost surely in $C_{[0,T]}$,*

$$\left\{ \left(\bar{\mathbf{Y}}^{\mathbf{N}_k}(t)\right) \Rightarrow (\bar{\mathbf{Y}}(t)) \right\} \Rightarrow \mathbb{P}((\bar{\mathbf{Y}}(t)) \in C_{[0,T]}) = 1.$$

Proof. We show that $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ obeys equations (4.10)-(4.11) in Theorem 20. First, from Theorems 17 and 19, we have that

$$(\bar{\mathbf{M}}^{\mathbf{N}}(t)) \Rightarrow 0.$$

Also,

$$P \left(\sup_{0 \leq t \leq T} \bar{Y}_i^{\mathbf{N}}(t) \geq k \right) = 0, \forall k > 1,$$

and the first condition holds trivially. The second condition is a stochastic version of the equicontinuity condition in the Arzelà-Ascoli Theorem.

$$\omega \left(\bar{Y}_i^{\mathbf{N}}, \delta, T \right) = \sup \left\{ \sup_{u,v \in [t, t+\delta]} \left| \bar{Y}_i^{\mathbf{N}}(u) - \bar{Y}_i^{\mathbf{N}}(v) \right| : 0 \leq t \leq t + \delta \leq T \right\} \quad (4.12)$$

$$\begin{aligned} &= \sup_{0 \leq t \leq t+\delta \leq T} \left\{ \sup_{u,v \in [t, t+\delta]} \left| \bar{M}_i^{\mathbf{N}}(u) - \bar{M}_i^{\mathbf{N}}(v) \right| \right. \\ &\quad \left. + \int_u^v \gamma_{ji} \frac{N_j}{N_i} \bar{Y}_j^{\mathbf{N}}(s-) \left(1 - \bar{Y}_i^{\mathbf{N}}(s-) \right) ds - \int_u^v \mu_i \bar{Y}_i^{\mathbf{N}}(s-) ds \right\} \end{aligned} \quad (4.13)$$

$$\leq \sup_{0 \leq t \leq t+\delta \leq T} \left\{ \sup_{u,v \in [t, t+\delta]} \left| \bar{M}_i^{\mathbf{N}}(u) - \bar{M}_i^{\mathbf{N}}(v) \right| \right\} + \gamma_{ji} \frac{N_j}{N_i} \delta + \mu_i \delta \quad (4.14)$$

$$\leq \sup_{0 \leq t \leq T} \left| \bar{M}_i^{\mathbf{N}}(t) \right| + \gamma_{ji} \frac{N_j}{N_i} \delta + \mu_i \delta \quad (4.15)$$

$$:= \omega_2 \left(\bar{Y}_i^{\mathbf{N}}, \delta, T \right). \quad (4.16)$$

Now, for any $\epsilon > 0$, we have

$$P \left(\omega \left(\bar{Y}_i^{\mathbf{N}}, \delta, T \right) \geq \epsilon \right) \leq P \left(\omega_2 \left(\bar{Y}_i^{\mathbf{N}}, \delta, T \right) \geq \epsilon \right).$$

From equation (4.9), given any $\alpha > 0$, we can choose N_1 and N_2 large enough so that

$$P \left(\sup_{0 \leq t \leq T} \left| \bar{M}_i^{\mathbf{N}}(t) \right| > \frac{\epsilon}{2} \right) < \alpha.$$

Now,

$$P\left(\omega_2\left(\bar{Y}_i^{\mathbf{N}}, \delta, T\right) \geq \epsilon\right) \leq P\left(\sup_{0 \leq t \leq T} \left|\bar{M}_i^{\mathbf{N}}(t)\right| > \frac{\epsilon}{2}\right) + \mathbb{P}\left(\gamma_{ji} \frac{N_j}{N_i} \delta + \mu_i \delta > \frac{\epsilon}{2}\right). \quad (4.17)$$

By applying the $\limsup_{\mathbf{N}}$ on both sides of equation (4.17), we obtain

$$\limsup_{\mathbf{N} \rightarrow \infty} \mathbb{P}\left(\omega_2\left(\bar{Y}_i^{\mathbf{N}}, \delta, T\right) \geq \epsilon\right) \leq \limsup_{\mathbf{N} \rightarrow \infty} \mathbb{P}\left(\gamma_{ji} \frac{N_j}{N_i} \delta + \mu_i \delta > \frac{\epsilon}{2}\right) = \mathbb{P}\left(\gamma_{ji} \alpha_{ji} \delta + \mu_i \delta > \frac{\epsilon}{2}\right), \quad (4.18)$$

with $\alpha_{ji} < \infty$, since we assumed

$$\frac{N_j}{N_i} \longrightarrow \alpha_{ji} < \infty.$$

Therefore, we can apply $\lim_{\delta \rightarrow 0}$ to equation (4.18)

$$\lim_{\delta \rightarrow 0} \limsup_{\mathbf{N} \rightarrow \infty} \mathbb{P}\left(\omega_2\left(\bar{Y}_i^{\mathbf{N}}, \delta, T\right) \geq \epsilon\right) \leq \lim_{\delta \rightarrow 0} \mathbb{P}\left(\gamma_{ji} \alpha_{ji} \delta + \mu_i \delta > \frac{\epsilon}{2}\right) = 0,$$

and thus,

$$\lim_{\delta \rightarrow 0} \limsup_{\mathbf{N} \rightarrow \infty} \mathbb{P}\left(\omega_2\left(\bar{Y}_i^{\mathbf{N}}, \delta, T\right) \geq \epsilon\right) \leq \lim_{\delta \rightarrow 0} \limsup_{\mathbf{N} \rightarrow \infty} \mathbb{P}\left(\omega_2\left(\bar{Y}_i^{\mathbf{N}}, \delta, T\right) \geq \epsilon\right) = 0.$$

We conclude that $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is a tight family with almost surely continuous weak-accumulation points, i.e.,

$$(\bar{\mathbf{Y}}^{\mathbf{N}_k}(t)) \Rightarrow (\bar{\mathbf{Y}}(t))$$

with

$$\mathbb{P}((\bar{\mathbf{Y}}(t)) \in C_{[0,T]}) = 1.$$

□

In fact, we will show in Theorem 25 that the sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is not only tight, but it converges to the solution of a deterministic differential equation. The main argument is that any weak accumulation point in the tight sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ should obey the equation

$$\frac{d}{dt} \bar{\mathbf{Y}}(t) = \mathbf{F}(\bar{\mathbf{Y}}(t)) \quad (4.19)$$

and from the uniqueness of equation (4.19), the whole sequence converges. To prove the uniqueness of the accumulation point, we will need to resort to the Skorokhod's representation Theorem along with the next two Lemmas. Lemma 22 states that any measurable operator on $D_{[0,T]}$ preserves equality in distribution.

Lemma 22. *Let $H : D_{[0,T]} \rightarrow D_{[0,T]}$ be a measurable operator. Then,*

$$(Z(t)) \stackrel{d}{=} (W(t)) \Rightarrow H(Z(s)) \stackrel{d}{=} H(W(s))$$

Proof. Let B be a Borel subset of $D_{[0,T]}$ then,

$$\mathbb{P}(H(Z(t)) \in B) = \mathbb{P}((Z(t)) \in H^{-1}(B)) = \mathbb{P}((W(t)) \in H^{-1}(B)) = \mathbb{P}(H(W(t)) \in B).$$

□

Now, with the next lemma, any linear bounded operator on $D_{[0,T]}$ preserves equality in distribution.

Lemma 23 ([31]). *Let $H : D_{[0,T]} \rightarrow D_{[0,T]}$ be a linear bounded operator. Then, H is measurable.*

Namely, if the càdlàg processes $(Z(t))$, $(W(t))$ are equal in distribution

$$(Z(t)) \stackrel{d}{=} (W(t)) \tag{4.20}$$

then, their integral counter-parts

$$\left(\int_0^t Z(s) ds \right) \stackrel{d}{=} \left(\int_0^t W(s) ds \right) \tag{4.21}$$

are equal in distribution, since the integral operator

$$\begin{aligned} \left(\int_0^t (\cdot) ds \right) : D_{[0,T]} &\longrightarrow D_{[0,T]} \\ x \in D_{[0,T]} &\longmapsto \int_0^t x(s) ds \end{aligned}$$

is linear and bounded in $D_{[0,T]}$. Next, we state the Skorokhod's representation theorem that will be crucial to establish convergence. Namely, the Theorem states that a weakly-convergent sequence of càdlàg processes admits an identically distributed representation sequence that converges realization-wise with respect to the Skorokhod topology.

Theorem 24 (Skorokhod's Representation Theorem; [26, 32]). *Let the sequence of càdlàg processes $(Z^N(t))$ converge weakly to the càdlàg process $(Z(t))$. Then, there exist processes $(\tilde{Z}^N(t))$ and $(\tilde{Z}(t))$ on $D_{[0,T]}$, with support in the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equal in distribution to $(Z^N(t))$ and $(Z(t))$, respectively, and for almost all $\omega \in \Omega$,*

$$(\tilde{Z}^N(\omega, t)) \xrightarrow{SK} (\tilde{Z}(\omega, t)),$$

with respect to the Skorokhod topology.

The realization-wise sense of convergence associated with the representation sequence $(\tilde{Z}^N(t))$, in Theorem 24, will allow for interchanging limit and integration, which is crucial to prove the next Theorem. Theorem 25 establishes that any weak accumulation point of $(\bar{\mathbf{Y}}^N(t))$ is necessarily solution to the realization-wise integral equation (4.22).

Theorem 25. *Let $(\bar{\mathbf{Y}}^{N_k}(t))$ be a subsequence converging weakly to $(\bar{\mathbf{Y}}(t))$ (an almost surely continuous process) with $\bar{\mathbf{Y}}^{N_k}(0) \xrightarrow{d} \bar{\mathbf{Y}}(0)$ and let $\frac{N_i}{N_j} \rightarrow \alpha_{ij} \in \mathbb{R}^+$. Then,*

$$\bar{Y}_i(\omega, t) = \bar{Y}_i(\omega, 0) + \int_0^t \bar{\gamma}_{ji} \bar{Y}_j(\omega, s) (1 - \bar{Y}_i(\omega, s)) ds - \int_0^t \mu_i \bar{Y}_i(\omega, s) ds, \quad (4.22)$$

for almost all $\omega \in \Omega$; with $i, j \in \{1, 2\}$ and $i \neq j$; and where we defined $\bar{\gamma}_{ji} := \gamma_{ji} \alpha_{ji}$.

Proof.

Define

$$\mathcal{F}_i : D_{[0,T]} \times D_{[0,T]} \longrightarrow \mathbb{R}$$

for $i = 1, 2$, with

$$\mathcal{F}_i((\mathbf{y}(t))) := y_i(t) - y_i(0) + \int \gamma_{ji} y_j(s) (1 - y_i(s)) ds.$$

We first observe that \mathcal{F}_i is measurable: indeed, the sum ‘+’ operator is measurable (with respect to the product topology $D_{[0,T]} \times D_{[0,T]}$) as referred in Chapter 2; the integral operator ‘ $(\int_0^t(\cdot)ds)$ ’ is measurable; and composition of measurable operators is measurable.

Now, from Skorokhod’s representation Theorem 24,

$$\begin{aligned} \exists \left(\tilde{\mathbf{Y}}^{N_k}(t) \right), \left(\tilde{\mathbf{Y}}(t) \right) : \quad & \left(\tilde{\mathbf{Y}}^{N_k}(t) \right) \stackrel{d}{=} \left(\bar{\mathbf{Y}}^{N_k}(t) \right), \left(\tilde{\mathbf{Y}}(t) \right) \stackrel{d}{=} \left(\bar{\mathbf{Y}}(t) \right) \\ & \left(\tilde{\mathbf{Y}}^{N_k}(\omega, t) \right) \xrightarrow{SK} \left(\tilde{\mathbf{Y}}(\omega, t) \right). \end{aligned}$$

for almost all $\omega \in \Omega$, with respect to the Skorokhod topology.

Define $\mathcal{F} := (\mathcal{F}_1, \mathcal{F}_2)$. We now show that,

$$\mathcal{F} \left(\tilde{\mathbf{Y}}^{N_k}(\omega, t) \right) \longrightarrow \mathcal{F} \left(\tilde{\mathbf{Y}}(\omega, t) \right) \quad (4.23)$$

for almost all $\omega \in \Omega$, with respect to the Skorokhod topology (in fact, with respect to the uniform topology). Indeed, since $(\bar{\mathbf{Y}}(t))$ is almost surely continuous, its distribution $\mathbb{P}_{\bar{\mathbf{Y}}}$ concentrates about the subspace of continuous functions $C_{\mathbf{R}}[0, 1] \subset D_{\mathbf{R}}[0, 1]$, and therefore $(\tilde{\mathbf{Y}}(t))$ is a.s. continuous as well given that

$$\mathbb{P} \left((\tilde{\mathbf{Y}}(t)) \in C_{[0,T]} \right) = \mathbb{P} \left((\bar{\mathbf{Y}}(t)) \in C_{[0,T]} \right) = 1.$$

Therefore, the pathwise convergence w.r.t. the Skorokhod topology implies pathwise uniform convergence as referred in Chapter 2, and

$$\left(\tilde{Y}_i^k(\omega, t)\right) \longrightarrow \left(\tilde{Y}_i(\omega, t)\right)$$

a.s. uniformly over the compact interval $[0, T]$. Therefore, we can interchange the limit with the integral via the dominated convergence Theorem (refer to [33]),

$$\int_0^t \gamma_{ji} \frac{N_j}{N_i} \tilde{Y}_j^{N_k}(\omega, s) \left(1 - \tilde{Y}_i^{N_k}(\omega, s)\right) ds \longrightarrow \int_0^t \bar{\gamma}_{ji} \tilde{Y}_j(\omega, s) \left(1 - \tilde{Y}_i(\omega, s)\right) ds$$

with realization-wise convergence almost surely uniformly. Therefore,

$$\mathcal{F} \left(\bar{\mathbf{Y}}^{N_k}(t) \right) \stackrel{d}{=} \mathcal{F} \left(\tilde{\mathbf{Y}}^{N_k}(t) \right) \longrightarrow \mathcal{F} \left(\tilde{\mathbf{Y}}(t) \right) \stackrel{d}{=} \mathcal{F} \left(\bar{\mathbf{Y}}(t) \right) \quad (4.24)$$

where the first and last equality are due to the measurability of \mathcal{F} ; and the convergence ‘ \longrightarrow ’ is in a realization-wise sense with respect to the uniform (and thus, Skorokhod) topology. In particular, this implies convergence in probability, and thus, convergence (4.24) holds in a weak sense (refer to corollary 1.6 from [20]), i.e.,

$$\mathcal{F} \left(\bar{\mathbf{Y}}^{N_k}(t) \right) \Rightarrow \mathcal{F} \left(\bar{\mathbf{Y}}(t) \right).$$

Remark that $\left(\bar{\mathbf{Y}}^{N_k}(t)\right)$ obeys the following stochastic dynamics

$$\mathcal{F} \left(\bar{\mathbf{Y}}^{N_k}(\omega, t) \right) = \bar{\mathbf{M}}^{N_k}(\omega, t),$$

and since

$$\left(\bar{\mathbf{M}}^{N_k}(t)\right) \Rightarrow 0 \quad (4.25)$$

then,

$$\mathcal{F} \left(\bar{\mathbf{Y}}(\omega, t) \right) \stackrel{d}{=} 0,$$

or in other words,

$$\bar{Y}_i(\omega, t) = \bar{Y}_i(\omega, 0) + \int_0^t \bar{\gamma}_{ji} \bar{Y}_j(\omega, s) \left(1 - \bar{Y}_i(\omega, s)\right) ds - \int_0^t \mu_i \bar{Y}_i(\omega, s) ds,$$

for almost all $\omega \in \Omega$; with $i, j = 1, 2$ and $i \neq j$. Fig. 4.1 summarizes the proof. □

Finally, the next theorem rigorously states the emergent dynamics of the single virus spread over complete-bipartite networks as the number of agents grows large.

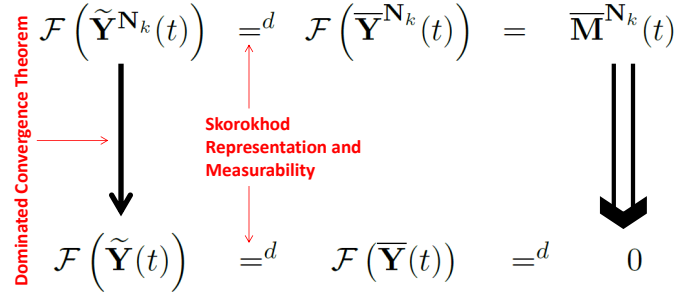


Figure 4.1: Summary of the proof to Theorem 25.

Theorem 26. Let $(\bar{\mathbf{Y}}_1^{\mathbf{N}}(0), \bar{\mathbf{Y}}_2^{\mathbf{N}}(0)) \xrightarrow{d} \mathbf{y}_0 \in \mathbb{R}^2$. Then, the normalized sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ converges weakly to the solution $(y_1(t, \mathbf{y}_0), y_2(t, \mathbf{y}_0))$ of the following ODE:

$$\frac{d}{dt}y_1(t) = (\bar{\gamma}_{21}y_2(t))(1 - y_1(t)) - \mu_1y_1(t) \quad (4.26)$$

$$\frac{d}{dt}y_2(t) = (\bar{\gamma}_{12}y_1(t))(1 - y_2(t)) - \mu_2y_2(t) \quad (4.27)$$

Proof. As the underlying vector field of (4.26)-(4.27),

$$\mathbf{F} = (F_1, F_2) : [0, 1]^2 \rightarrow \mathbb{R}^2,$$

where

$$F_i(y_1, y_2) := \bar{\gamma}_{ji}y_j(1 - y_i) - \mu_iy_i,$$

is Lipschitz, the continuous (and thus, differentiable) solution $(\bar{\mathbf{Y}}(t))$ of (4.22) is unique. Thus, any weak limit of $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ with initial condition given by $\bar{\mathbf{Y}}^{\mathbf{N}}(0)$ and converging in distribution to $\bar{\mathbf{Y}}(0)$ is equal to the unique solution $(\bar{\mathbf{Y}}(t))$ of (4.22) with initial condition $(\bar{\mathbf{Y}}(0))$. Therefore, the whole sequence converges

$$(\bar{\mathbf{Y}}^{\mathbf{N}}(t)) \Rightarrow (\bar{\mathbf{Y}}(t))$$

to the solution of (4.22). Equation (4.22) is the integral version of the ODE (4.26)-(4.27). Theorem (26) is concluded. \square

4.3 Concluding Remarks

We showed in this Chapter that the sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ associated with the fraction of infected nodes per island on complete-bipartite networks $G^{(N_1, N_2)}$ converges weakly to the solution of the vector differential equation (4.26)-(4.27). We explored the martingale structure on the dynamics of the process $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ to show that it converges weakly to zero. As a Corollary to this fact, the family $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is tight with a single

accumulation point given by the (unique) solution of a limiting differential equation. Since any convergent subsequence converges to the same accumulation point, then the whole sequence $(\bar{\mathbf{Y}}^N(t))$ converges to the unique accumulation point, namely, the solution of the ODE (4.26)-(4.27). In the next Chapter, we extend the convergence result to the multivirus complete-multipartite network case.

Chapter 5

Multivirus over Complete-Multipartite Networks

In Chapter 4, we established that the sequence of single virus macroprocesses $(\bar{Y}_1^{\mathbf{N}}(t), \bar{Y}_2^{\mathbf{N}}(t))$ over the underlying sequence of complete-bipartite networks $G^{\mathbf{N}}$ converges weakly to the solution $(y_1(t), y_2(t))$ of a deterministic ODE given by equations (4.26)-(4.27), as $\mathbf{N} \rightarrow \infty$ with $N_i/N_j \rightarrow \alpha_{ji} < \infty$ for $i, j \in \{1, 2\}$ and $i \neq j$. We divided the proof of Theorem 26 into four main steps:

- (i) The martingale $(\bar{M}_1^{\mathbf{N}}(t), \bar{M}_2^{\mathbf{N}}(t))$ converges weakly to zero;
- (ii) The family $(\bar{Y}_1^{\mathbf{N}}(t), \bar{Y}_2^{\mathbf{N}}(t))$, indexed by \mathbf{N} , is tight;
- (iii) Any accumulation point of the tight family is solution of (4.26)-(4.27);
- (iv) Uniqueness of the differential equation (4.26)-(4.27) implies convergence.

In this Chapter, we extend Theorem 26 to the more general case of multi-virus epidemics over complete-multipartite networks. We consider in Section 5.1 the single-virus over complete-multipartite networks case, and in Section 5.2 we study the general case of multivirus over complete-multipartite networks. In Section 5.3, we summarize the Chapter and present a numerical simulation that corroborates with the fluid limit dynamics established in this Chapter.

5.1 Single-virus over Complete-Multipartite Networks

For single virus spread,

$$(\bar{\mathbf{Y}}^{\mathbf{N}}(t)) = (\bar{Y}_1^{\mathbf{N}}(t), \dots, \bar{Y}_M^{\mathbf{N}}(t))$$

stands for the process associated with the fraction of infected nodes at each island

$$i \in \{1, \dots, M\}$$

over the complete-multipartite network $G^{\mathbf{N}}$ with M islands and supertopology induced by the topology of the fixed graph G . From the pathwise characterization in equation (3.12) in Chapter 3, we can rewrite the stochastic integral equation for $(\mathbf{Y}^{\mathbf{N}}(t))$ as

$$Y_i^{\mathbf{N}}(t) = Y_i^{\mathbf{N}}(0) + M_i^{\mathbf{N}}(t) + \sum_{j \sim i} \int_0^t \gamma_{ji} Y_j^{\mathbf{N}}(s-) \left(\frac{N_i - Y_i^{\mathbf{N}}(s-)}{N_i} \right) ds - \int_0^t \mu_i Y_i^{\mathbf{N}}(s-) ds,$$

for $i = 1, \dots, M$. The infections from all neighboring islands are now coupled by these M equations. The corresponding martingales are given by

$$\begin{aligned} M_i^{\mathbf{N}}(t) = & \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \underbrace{\int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-)=q, Y_i^{\mathbf{N}}(s-)=\ell\}} \times \left(\mathcal{N}_{\gamma_{ji}q\left(\frac{N_i-\ell}{N_i}\right)}(ds) - \gamma_{ji}q \left(\frac{N_i-\ell}{N_i} \right) ds \right)}_{:=M_{aij}^{\mathbf{N}}(t,\ell,q)} \\ & - \underbrace{\sum_{\ell=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_i^{\mathbf{N}}(s-)=\ell\}} (\mathcal{N}_{\ell\mu_i}(ds) - \ell\mu_i ds)}_{:=M_{bi}^{\mathbf{N}}(t)} . \end{aligned}$$

Next, we prove that the normalized sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ over the underlying sequence of multipartite networks $G^{\mathbf{N}}$ converges weakly to the solution of an M -dimensional ODE.

Theorem 27. *Let $\bar{\mathbf{Y}}^{\mathbf{N}}(0) \xrightarrow{d} \mathbf{y}_0 \in \mathbb{R}^M$ with $M < \infty$, $\mathbf{N} = (N_1, \dots, N_M) \rightarrow \infty$ and $\frac{N_i}{N_j} \rightarrow \alpha_{ij} < \infty$ for all $i, j \in \{1, \dots, M\}$ and $i \sim j$. Then, the normalized sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ converges weakly to the solution $(\mathbf{y}(t, \mathbf{y}_0)) = (y_1(t, \mathbf{y}_0), \dots, y_M(t, \mathbf{y}_0))$ of the following ODE:*

$$\frac{d}{dt} y_i(t) = \underbrace{\left(\sum_{j \sim i} \bar{\gamma}_{ji} y_j(t) \right)}_{:=F_i(y_1(t), \dots, y_M(t))} (1 - y_i(t)) - \mu_i y_i(t) \quad (5.1)$$

with $\bar{\gamma}_{ji} := \alpha_{ji} \gamma_{ji}$.

Proof. For the sake of clarity, we revisit each of the points (i)-(iv) applied to prove the Theorem 26 in Chapter 4 and referred to in the beginning of this Chapter.

(i) **Martingale vanishes.** We start by observing that, for fixed

$$i, j \in \{1, \dots, M\},$$

the underlying Poisson point processes

$$\mathcal{N}_{\gamma_{ji}q\left(\frac{N_i-\ell}{N_i}\right)}$$

indexed by

$$(\ell, q) \in \{1, \dots, N_i\} \times \{1, \dots, N_j\}$$

are independent as explained in Section 3.2. Now, fixing the island subindex i , the Poisson processes are still independent due to the independence of the exponential time services associated with different nodes.

Let

$$M_{ai}^{\mathbf{N}}(t) := \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} M_{aij}^{\mathbf{N}}(t, \ell, q).$$

Then

$$\begin{aligned} \mathbb{E} \left(M_{ai}^{\mathbf{N}}(t) \right)^2 &= \mathbb{E} \left(\sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} M_{aij}^{\mathbf{N}}(t, \ell, q) \right)^2 = \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbb{E} \left(M_{aij}^{\mathbf{N}}(t, \ell, q) \right)^2 \\ &= \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbb{E} \left(\int_0^t \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-) = q, Y_i^{\mathbf{N}}(s-) = \ell\}} \gamma_{ji} q \left(\frac{N_i - \ell}{N_i} \right) ds \right) \\ &\leq \sum_{j \sim i} \mathbb{E} \left(\int_0^t \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbf{1}_{\{Y_j^{\mathbf{N}}(s-) = q, Y_i^{\mathbf{N}}(s-) = \ell\}} \gamma_{ji} N_j ds \right) \\ &\leq \sum_{j \sim i} \gamma_{ji} N_j t \leq M \max_{j=1, \dots, M} \{ \gamma_{ji} N_j \} t, \end{aligned}$$

and thus,

$$\mathbb{E} \left(\overline{M}_{ai}^{\mathbf{N}}(t) \right)^2 = \frac{1}{N_i^2} \mathbb{E} \left(M_{ai}^{\mathbf{N}}(t) \right)^2 \leq \frac{M}{N_i} \left(\max_{j: j \sim i} \left\{ \frac{\gamma_{ji} N_j}{N_i} \right\} \right) t \rightarrow 0.$$

Similarly,

$$\mathbb{E} \left(\overline{M}_{bi}^{\mathbf{N}}(t) \right)^2 \rightarrow 0$$

for each $t, t \geq 0$. Therefore, the normalized martingale $\left(\overline{M}_i^{\mathbf{N}}(t) \right)$ converges to zero in \mathcal{L}_2 on the line, that is,

$$\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right) \xrightarrow{\mathcal{L}_2} 0.$$

Now, from Doob's inequality in Corollary 18 and from Theorem 19, we conclude that

$$\left(\overline{M}_i^{\mathbf{N}}(t) \right) \xrightarrow{P} 0 \text{ and } \left(\overline{M}_i^{\mathbf{N}}(t) \right) \Rightarrow 0.$$

(ii) The family $\left(\overline{\mathbf{Y}}^{\mathbf{N}}(t) \right)$ is tight. As in equations (4.12)-(4.16), we have

$$\omega \left(\overline{Y}_i^{\mathbf{N}}, \delta, T \right) \leq \sup_{0 \leq t \leq t+\delta \leq T} \left\{ \sup_{u, v \in [t, t+\delta]} \left| \overline{M}_i^{\mathbf{N}}(u) - \overline{M}_i^{\mathbf{N}}(v) \right| \right\} + \sum_{j \sim i} \gamma_{ji} M \frac{N_j}{N_i} \delta + \mu_i \delta \quad (5.2)$$

$$:= \omega_2 \left(\overline{Y}_i^{\mathbf{N}}, \delta, T \right). \quad (5.3)$$

Due to the convergence in probability, with respect to the uniform topology, of the martingale

$$\left(\overline{\mathbf{M}}^{\mathbf{N}}(t)\right) \xrightarrow{\mathbb{P}} 0$$

and the assumption that the number of islands $M < \infty$ is finite, it follows similarly to as done in Section 4.2 that $\left(\overline{\mathbf{Y}}^{\mathbf{N}}(t)\right)$ fulfills the conditions of the Arzelà-Ascoli Theorem 20, and we conclude that $\left(\overline{\mathbf{Y}}^{\mathbf{N}}(t)\right)$ is a tight family, that is, it admits a convergent subsequence

$$\left(\overline{\mathbf{Y}}^{\mathbf{N}_k}(t)\right) \Rightarrow \left(\overline{\mathbf{Y}}(t)\right).$$

where $\left(\overline{\mathbf{Y}}(t)\right)$ is necessarily almost surely continuous.

(iii) If $\left(\overline{\mathbf{Y}}^{\mathbf{N}_k}(t)\right) \Rightarrow \left(\overline{\mathbf{Y}}(t)\right)$ then, $\left(\overline{\mathbf{Y}}(t)\right)$ is solution to the ODE (5.1). It follows similarly to as done in the proof of Theorem 25, remarking that we assume a finite (fixed) number of islands M , which prevents us from handling technical issues regarding the interchanging between the limit and integral.

(iv) $\left(\overline{\mathbf{Y}}^{\mathbf{N}}(t)\right) \Rightarrow \left(\overline{\mathbf{Y}}(t)\right)$, where $\left(\overline{\mathbf{Y}}(t)\right)$ is solution to the ODE (5.1). Note that the underlying vector field

$$\mathbf{F}(y_1, \dots, y_M) = (F_1(y_1, \dots, y_M), \dots, F_M(y_1, \dots, y_M))$$

in equation (5.1) is differentiable and thus, \mathbf{F} is locally Lipschitz. Therefore, the solution of (5.1) exists locally and it is unique. Since the state space of interest $[0, 1]^M$ is compact and invariant, \mathbf{F} is globally Lipschitz over $[0, 1]^M$, and any solution of (5.1) is defined for all time $t \geq 0$ and is unique. In particular, any convergent subsequence $\left(\overline{\mathbf{Y}}^{\mathbf{N}_k}(t)\right)$ converges to the same weak limit given by the unique solution of (5.1) and, thus, the whole sequence converges. This concludes the proof of Theorem 27. \square

5.2 Multivirus Case

We denote as

$$\left(\overline{\mathbf{Y}}^{\mathbf{N}}(t)\right) = \left[\overline{Y}_{ik}^{\mathbf{N}}(t)\right]_{ik}$$

the matrix process collecting the fraction of k -infected nodes at island

$$i \in \{1, \dots, M\} \text{ with } k \in \{1, \dots, K\}$$

over time $t, t \geq 0$, where $K < \infty$ is the number of virus strains. In this Section, we refer to

$$\left(\overline{\mathbf{Y}}_i^{\mathbf{N}}(t)\right) = \left(\overline{Y}_{i1}^{\mathbf{N}}(t), \dots, \overline{Y}_{iK}^{\mathbf{N}}(t)\right)$$

as the distribution of infected nodes at island i across the K strains of virus. We define

$$\mathcal{Y}_i^{\mathbf{N}} := \{\mathbf{y} \in \mathbb{N}^K : 0 \leq \langle \mathbf{y}, \mathbf{1}_K \rangle \leq N_i\}$$

as the state-space of the process $(\mathbf{Y}_i^{\mathbf{N}}(t))$.

In the case of multivirus epidemics, we assume the exclusion principle, i.e., a node can only be infected with at most one strain of virus at a time, as referred in Chapter 3. Applying the same reasoning as in Section 3.2, we obtain the following pathwise description for $(\mathbf{Y}^{\mathbf{N}}(t))$

$$Y_{ik}^{\mathbf{N}}(t) = Y_{ik}^{\mathbf{N}}(0) + M_{ik}^{\mathbf{N}}(t) + \sum_{j \sim i} \int_0^t \gamma_{ji}^k Y_{jk}^{\mathbf{N}}(s-) \left(\frac{N_i - \langle \mathbf{Y}_i^{\mathbf{N}}(s-), \mathbf{1} \rangle}{N_i} \right) ds - \int_0^t \mu_i^k Y_{ik}^{\mathbf{N}}(s-) ds,$$

where the martingale is given by

$$\begin{aligned} M_{ik}^{\mathbf{N}}(t) &= \sum_{j \sim i} \sum_{y \in \mathcal{Y}_i^{\mathbf{N}}} \sum_{q=1}^{N_j} \underbrace{\int_0^t \mathbf{1}_{\{Y_{jk}^{\mathbf{N}}(s-)=q, \mathbf{Y}_i^{\mathbf{N}}(s-)=y\}} \left(\mathcal{N}_{\gamma_{ji}^k q \left(\frac{N_i - \langle y, \mathbf{1} \rangle}{N_i} \right)}(ds) - \gamma_{ji}^k q \left(\frac{N_i - \langle y, \mathbf{1} \rangle}{N_i} \right) ds \right)}_{M_{j,y,q}^{\mathbf{N}}(t)} \\ &\quad - \underbrace{\sum_{\ell=1}^{N_i} \int_0^t \mathbf{1}_{\{Y_{ik}^{\mathbf{N}}(s-)=\ell\}} \left(\mathcal{N}_{\ell \mu_i^k}(ds) - \ell \mu_i^k ds \right)}_{M_b^{\mathbf{N}}(t)} \end{aligned}$$

and, from the construction in Chapter 3,

$$\left\{ \mathcal{N}_{\gamma_{ji}^k q \left(\frac{N_i - \langle y, \mathbf{1} \rangle}{N_i} \right)} \right\}_{i,j,q,y}$$

is a family of independent Poisson processes. The next Theorem finally extends Theorems 27 and 25 to establish that the sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ associated with the fraction of infected nodes per island and per strain converges weakly to the solution of the vector ordinary differential equation given by (5.4).

Theorem 28. *Let $\bar{\mathbf{Y}}^{\mathbf{N}}(0) \xrightarrow{d} \mathbf{y}_0 \in \mathbb{R}^{M \times K}$, with $M, K < \infty$, $\mathbf{N} = (N_1, \dots, N_M) \rightarrow \infty$ and $\frac{N_i}{N_j} \rightarrow \alpha_{ij} < \infty$ for all $i \sim j$. Then, the normalized sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ converges weakly to the solution $(\mathbf{y}(t, \mathbf{y}_0))$ of the following ODE:*

$$\frac{d}{dt} y_{ik}(t) = \underbrace{\left(\sum_{j \sim i} \bar{\gamma}_{ji}^k y_{jk}(t) \right) \left(1 - \sum_{m=1}^K y_{im}(t) \right)}_{F_{ik}([y_{mn}(t)]_{mn})} - \mu_i^k y_{ik}(t) \quad (5.4)$$

with $\bar{\gamma}_{ji}^k := \alpha_{ji} \gamma_{ji}^k$.

Proof. **i) Martingale vanishes.** Since the underlying Poisson point processes

$$\mathcal{N}_{\gamma_{ji}^k q \left(\frac{N_i - \langle y, \mathbf{1} \rangle}{N_i} \right)}$$

are independent, then,

$$\begin{aligned}
\mathbb{E} \left(\sum_{j \sim i} \sum_{y \in \mathcal{Y}_i^{\mathbf{N}}} \sum_{q=1}^{N_j} M_{j,y,q}^{\mathbf{N}}(t) \right)^2 &= \sum_{j \sim i} \sum_{y \in \mathcal{Y}_i^{\mathbf{N}}} \sum_{q=1}^{N_j} \mathbb{E} \left(M_{j,y,q}^{\mathbf{N}}(t) \right)^2 \\
&= \sum_{j \sim i} \sum_{y \in \mathcal{Y}_i^{\mathbf{N}}} \sum_{q=1}^{N_j} \mathbb{E} \left(\int_0^t \mathbf{1}_{\{Y_{jk}^{\mathbf{N}}(s-) = q, \mathbf{Y}_i^{\mathbf{N}}(s-) = y\}} \left(\gamma_{ji}^k q \left(\frac{N_i - \langle y, 1 \rangle}{N_i} \right) ds \right) \right) \\
&\leq \sum_{j \sim i} \mathbb{E} \left(\int_0^t \sum_{y \in \mathcal{Y}_i^{\mathbf{N}}} \sum_{q=1}^{N_j} \mathbf{1}_{\{Y_{jk}^{\mathbf{N}}(s-) = q, \mathbf{Y}_i^{\mathbf{N}}(s-) = y\}} \gamma_{ji}^k N_j ds \right) \\
&\leq M \max_{j: j \sim i} \left\{ \gamma_{ji}^k N_j \right\} t,
\end{aligned}$$

and thus,

$$\mathbb{E} \left(\frac{1}{N_i} \sum_{j \sim i} \sum_{y \in \mathcal{Y}_i^{\mathbf{N}}} \sum_{q=1}^{N_j} M_{j,y,q}^{\mathbf{N}}(t) \right)^2 \leq \frac{M}{N_i} \left(\max_{j: j \sim i} \left\{ \gamma_{ji}^k N_j \right\} \right) t \rightarrow 0.$$

Similarly

$$\mathbb{E} \left(\overline{M}_b^{\mathbf{N}}(t) \right)^2 \rightarrow 0$$

and, therefore, the normalized martingale $\overline{M}^{\mathbf{N}}(t)$ converges to zero in \mathcal{L}_2 on the line. Now, from Doob's inequality and from Theorem 19, we conclude that

$$\left(\overline{M}_{ik}^{\mathbf{N}}(t) \right) \xrightarrow{\mathbb{P}} 0.$$

ii) The family $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$ is tight. As in equations (4.12)-(4.16), we have

$$\omega \left(\overline{Y}_{ik}^{\mathbf{N}}, \delta, T \right) \leq \sup_{0 \leq t \leq t+\delta \leq T} \left\{ \sup_{u,v \in [t, t+\delta]} \left| \overline{M}_{ik}^{\mathbf{N}}(u) - \overline{M}_{ik}^{\mathbf{N}}(v) \right| \right\} + M \frac{\max_{j: j \sim i} \left\{ \gamma_{ji}^k N_j \right\}}{N_i} \delta + \mu_i \delta.$$

Again, due to

$$\left(\overline{\mathbf{M}}^{\mathbf{N}}(t) \right) \xrightarrow{\mathbb{P}} 0$$

and similarly to as done in Section 4.2, the matrix-process $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$ fulfills the conditions of Theorem 20, and we conclude that $(\overline{\mathbf{Y}}^{\mathbf{N}}(t))$ is a tight family whose accumulation points are almost surely continuous functions.

iii) If $(\overline{\mathbf{Y}}^{\mathbf{N}_k}(t)) \Rightarrow (\overline{\mathbf{Y}}(t))$ then, $(\overline{\mathbf{Y}}(t))$ is solution of the ODE (5.4). Remark that we assume $K, M < \infty$ and thus, the result follows similarly to as done in the proof to Theorem 25.

iv) $(\overline{\mathbf{Y}}^{\mathbf{N}}(t)) \Rightarrow (\overline{\mathbf{Y}}(t))$, where $(\overline{\mathbf{Y}}(t))$ is solution of the ODE (5.4). Note that for similar reasons as exposed in Section 5.1, the solution to (5.4) exists and is unique. Thus, any convergent subsequence

$(\bar{\mathbf{Y}}^{\mathbf{N}_k}(t))$ converges to the same weak limit given by the unique solution to (5.4) and thus, the whole sequence converges. \square

5.3 Concluding Remarks

In this Chapter, we established that the sequence of multi-virus macroprocesses $\left[\bar{\mathbf{Y}}_{ik}^{\mathbf{N}}(t)\right]_{ik}$ over the underlying sequence of complete-multipartite networks $G^{\mathbf{N}}$ converges weakly to the solution $(\mathbf{y}(t, \mathbf{y}_0))$ of the deterministic ODE given by equations (5.4), as $\mathbf{N} \rightarrow \infty$ with $N_i/Nj \rightarrow \alpha_{ji} < \infty$ for $i, j \in \{1, \dots, M\}$ and $i \neq j$. To prove Theorem 28, we revisited the following four main steps:

- (i) The martingale $(\bar{\mathbf{M}}^{\mathbf{N}}(t))$ converges weakly to zero;
- (ii) The family $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$, indexed by \mathbf{N} , is tight;
- (iii) Any accumulation point of the tight family is solution to (5.4);
- (iv) Uniqueness of the differential equation (5.4) implies convergence.

Theorem 28 is our main contribution in Part I. We stress that the weak convergence takes place over the compact interval $[0, T]$. In words, while $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ weakly approximates the solution to the ODE (5.4) over the compact interval $[0, T]$, the process $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is eventually absorbed to zero. This follows simply by the fact that for any finite time interval with length l , the process $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ can reach any of its finite possible state-values with a probability of at least $\epsilon > 0$, where ϵ depends on l . When $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ reaches zero (no virus in the network), it remains in the zero state for all future times. This induces a dichotomy of attractors between the ODE and the approximating stochastic dynamical system:

- $\limsup_{t \rightarrow \infty} \bar{\mathbf{Y}}^{\mathbf{N}}(\omega, t) = 0$, for all \mathbf{N} and almost all $\omega \in \Omega$;
- $\liminf_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0) = \mathbf{y}_\infty \neq 0$.

The asymptotics associated with the approximating process $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ and the solution to the differential equation $(\mathbf{y}(t, \mathbf{y}_0))$ (being approximated) are different. In fact, when the time-scale of $(\mathbf{y}(t, \mathbf{y}_0))$ associated with the convergence to the attractors of (5.4) is much faster than the time-scale to absorption of $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$, then, for large \mathbf{N} , the attractors of (5.4) depict the quasi-stationary state of $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$, and the phenomenon is often cast as metastability (refer to [25]). We do not address this phenomenon here, but in future work.

To conclude the Chapter, we present in Fig. 5.1 a numerical experiment of two strains of virus x and y spreading across a complete-bipartite network via our SIS stochastic peer-to-peer law of infection. Figure 5.1 illustrates the Matlab results for the evolution of two strains in the complete-bipartite network (refer to the noisy curves), and we superimpose on it the corresponding ODE fluid-limit evolution (refer to the smooth curves). Figures 5.1a, 5.1b, and 5.1c illustrate the evolution of the fractions of x -infected (blue/solid curves) and y -infected (red/dashed curves) nodes at islands 1 and 2 for different number of nodes. We

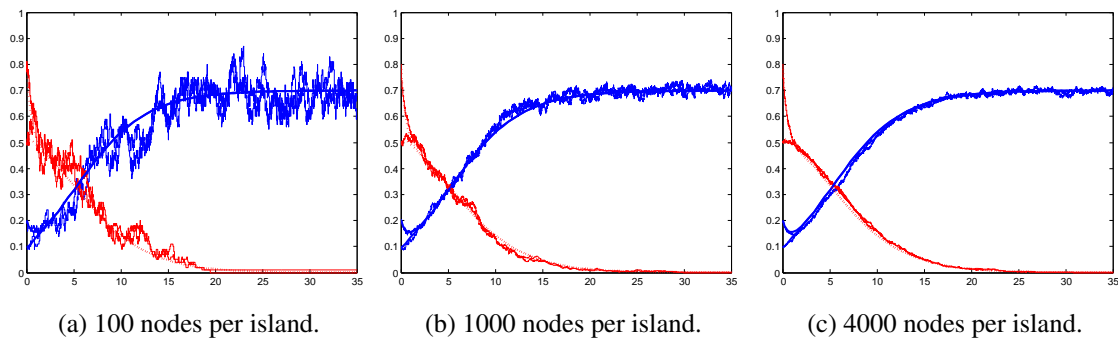


Figure 5.1: The plots represent the numerical evolution of the fractions of x -infected (in blue/solid) and y -infected (in red/dashed) nodes at each island 1 and 2. The boldfaced curves represent the solution of the limiting vector differential equation of a bi-viral epidemics in a bipartite network.

observe that, as the number of nodes grows large, the randomness decreases and the infected population dynamics fits the fluid-limit dynamics, corroborating with our result in Theorem 28.

Chapter 6

Epidemics over Regular Networks: Preliminary Results

In this Chapter, we present some preliminary results on our attempt to extend the weak convergence results of Chapters 4-5 – in particular, Theorems 28-27 – to general sparse networks. Namely, our future goal is to build on the results from this Chapter to determine the emergent dynamics associated with the fraction of infected nodes process $(\bar{Y}^N(t))$ under a peer-to-peer model of infection in regular networks. As discussed in Section 6.4, numerical simulations depicted in Fig. 6.5 and Fig. 6.6 suggest that certain ODE's, e.g., the ODE in equation (6.16), that are commonly assumed models to explain the dynamics of epidemics over regular (or sometimes coined as *homogeneous*) networks, are not appropriate models in the framework of peer-to-peer epidemics over **static** regular networks – i.e., when nodes and edges are fixed. On the other hand, such ODE's may be appropriate to model peer-to-peer epidemics over more *dynamical* regular networks as we will show in our intermediate results in Section 6.3 and in the numerical simulation depicted in Fig. 6.4, in Section 6.4. This confirms that deriving rigorously the macroscopic emergent dynamics in sparse static networks is fundamental to understand their correct behavior.

In this Chapter, we focus on single virus epidemics. It turns out that, under the sparse network setting, obtaining the weak limit dynamics of $(\bar{Y}^N(t))$ is quite a challenging problem that remains open in the literature. The major difficulty lies in the loss of Markovian property of the sequence $(\bar{Y}^N(t))$ whose stochastic dynamics is now coupled to the high-dimensional microstate $(X^N(t))$. Establishing the fluid limit dynamics of $(\bar{Y}^N(t))$ from the peer-to-peer rules of infection amounts to proving that certain functionals of the microstate $(X^N(t))$, representing the rates of $(\bar{Y}^N(t))$, concentrate about functionals of weak accumulation points $(\bar{Y}(t))$ of $(\bar{Y}^N(t))$. This roughly implies that the rate processes associated with any weak accumulation point $(\bar{Y}(t))$ are not coupled to the limit microstate $(X^\infty(t))$, and then, $(\bar{Y}(t))$ is Markov. In other words, the nonMarkov sequence $(\bar{Y}^N(t))$ is asymptotically Markov or at least bear Markov accumulation points. For the complete-multipartite networks addressed in Chapters 4-5, the vector process $(\bar{\mathbf{Y}}^N(t))$ collecting the fraction of infected nodes by virus and per island is Markov, and this was explored in our

weak convergence proof in these Chapters. For general sparse networks, e.g., regular networks, this is not the case. More precisely, and as we explain in the next Section, proving weak convergence of the rates of $(\bar{Y}^N(t))$ requires to prove that they are tight and converge in finite dimensional distribution to certain functionals of $(\bar{Y}(t))$, as will be clearer. When the sequence $(\bar{Y}^N(t))$ is non-Markov – as it is for sparse networks – obtaining both tightness and finite dimensional distribution convergence seems difficult. In this Chapter, we present two peer-to-peer epidemics models, to which we refer to as static network model and shuffling model. The shuffling model is a simplistic model for the peer-to-peer epidemics over *dynamical* networks. We prove tightness for the rates of $(\bar{Y}^N(t))$ under the static network model, but convergence in finite-dimensional distribution is not easy and we pursue as future work. On the other hand, for the shuffling model, we can prove finite-dimensional distribution convergence, but tightness is challenging.

In Section 6.1 we lay down our approach in obtaining the emergent dynamics. In Section 6.2, we present the static network model, and we prove tightness for the rates of $(\bar{Y}^N(t))$, but the finite-dimensional distribution convergence is challenging to estimate and will be addressed in future work. In Section 6.3, we introduce the shuffling model, and we show that, under this model, the rates of $(\bar{Y}^N(t))$ converge in finite-dimensional distribution, but tightness and therefore, convergence of $(\bar{Y}^N(t))$ is challenging to guarantee and will be addressed in future work. In Section 6.4, we present numerical simulations that suggest concentration of probability measures about the solution to the ODE (6.16) for the dynamical shuffling regular network model, while showing that the same ODE (6.16) is not an appropriate model for the static regular network case. In Section 6.5, we conclude the Chapter.

6.1 Weak Convergence Framework

To establish the weak convergence

$$(\bar{Y}^N(t)) \Rightarrow (\bar{Y}(t))$$

of a general sequence of càdlàg processes $(\bar{Y}^N(t))$ to an almost surely continuous process $(\bar{Y}(t))$, one needs roughly to guarantee two minimal conditions:

- (i) $(\bar{Y}^N(t))$ is tight;
- (ii) $(\bar{Y}^N(t))$ converges in finite-dimensional distribution (f.d.d.) to an almost surely continuous process $(\bar{Y}(t))$.

While in the complete-multipartite case, these two conditions were met for the matrix-process $(\bar{\mathbf{Y}}^N(t))$ via the Arzelà-Ascoli Theorem and the uniqueness of the limiting ODE (5.4), it turns out to be quite challenging to prove similar results for peer-to-peer epidemics models over large-scale *sparse* networks, e.g., regular networks. One reason is that, in a sparse network, the fraction of infected nodes process $(\bar{Y}^N(t))$ is not Markov. Indeed, the rate driving the evolution of $(\bar{Y}^N(t))$ is tied to the microstate $(X^N(t))$. In

words, the *tendency* to increase or decrease $\bar{Y}^N(t)$ at time t , does not depend solely on $\bar{Y}^N(t)$, but it is also coupled with the microscopic organization of the nodes, which is high-dimensional. Fig. 6.1 illustrates two different microstate configurations with the same number of infected nodes, namely $\bar{Y}^N(t) = 3$. Under any peer-to-peer model of infection, the rate to increase $\bar{Y}^N(t)$ is greater at the top network, in Fig. 6.1, as the corresponding microscopic configuration $X^N(t)$ offers more links of contact between infected and healthy nodes as opposed to the clustered configuration at the bottom network. As a more concrete example, under

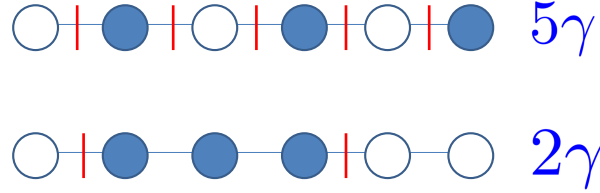


Figure 6.1: Illustration of two line networks with the same number of infected nodes (blue/colored nodes). The bottom line network shows a clustered configuration and therefore the only way to spread the infection is via the two links of contact between infected and healthy nodes. On the other hand, the top line network has more contact links between infected and healthy nodes. Therefore, and despite both networks having the same number of infected nodes, the rate to increase the fraction of infected nodes is greater in the top line network.

a peer-to-peer model of infection (to be described in the following Sections) the pathwise characterization of $\left(\bar{Y}^N(t)\right)$ is given by

$$\bar{Y}^N(t) = \bar{Y}^N(0) + \bar{M}^N(t) - \int_0^t \underbrace{F^N(X^N(s))}_{:=\bar{R}^N(s)} ds. \quad (6.1)$$

where $\left(\bar{M}^N(t)\right)$ is a martingale. Roughly, under mild conditions on the dynamical system (6.1), e.g., the martingale $\left(\bar{M}^N(t)\right)$ converges weakly to zero, one can establish weak convergence of $\left(\bar{Y}^N(t)\right)$ by assuring that the conditions (i) – (ii) (above) hold rather for the rate process $\left(\bar{R}^N(t)\right)$, defined in equation (6.1), i.e.,

- (i') $\left(\bar{R}^N(t)\right)$ is tight;
- (ii') $\left(\bar{R}^N(t)\right)$ converges in finite-dimensional distribution (f.d.d.) to an almost surely continuous process $(R(t))$ of the form

$$(\bar{R}(t)) = (F(\bar{Y}(t))),$$

where F is Lipschitz, and $(\bar{Y}(t))$ is a weak accumulation point of $\left(\bar{Y}^N(t)\right)$.

As will be clear momentarily, if these two conditions are met, along with the weak convergence to zero of the normalized martingale, then

$$\left(\bar{Y}^N(t)\right) \Rightarrow (\bar{Y}(t))$$

with

$$\bar{Y}(t) = \bar{Y}(0) + \int_0^t F(\bar{Y}(s)) ds.$$

Remark that, in the complete-multipartite case, the total fraction of infected nodes

$$\left(\bar{Y}^N(t)\right) = \left(\frac{1}{N} \sum_{i=1}^N X_i^N(t)\right)$$

is not a Markov process, but one could break this process into the vector

$$\left(\bar{\mathbf{Y}}^N(t)\right) = \left(\bar{Y}_1^N(t), \dots, \bar{Y}_M^N(t)\right)$$

collecting the fraction of infected nodes per island, among the M islands, and this latter vector is a Markov process. This moderate zoom in to recover the Markov property seems not to apply to regular networks.

In the next two Sections we will introduce two peer-to-peer stochastic models for epidemics. The first model (refer to Section 6.2) is similar to the complete-multipartite model assumed in Chapters 4-5, but now once a node is infected, it activates one exponential clock per neighbor node (instead of per super-neighbor island), as will be clearer. The second model (refer to Section 6.3) evolves similarly, but with the proviso that, once an infection or healing strikes the system, the nodes in the system are randomly shuffled, and the epidemics spreads again according to the peer-to-peer model. For the sake of clarity, all stochastic processes constructed under the shuffling model will be indexed by “sh”, e.g., $\left(\bar{Y}_{\text{sh}}^N(t)\right)$ represents the process associated with the fraction of infected nodes over time, under the shuffling model. The binary-vector microprocess is represented as $\left(X_{\text{sh}}^N(t)\right)$.

6.2 Epidemics over Regular Networks: Tightness

In this Section, we introduce a peer-to-peer model of diffusion over regular networks that resembles the one assumed for complete-multipartite networks, and we show that under this model, the fraction of infected nodes process $\left(\bar{Y}^N(t)\right)$ is tight (and thus, relatively compact), but finite-dimensional distribution convergence is difficult to attain, and we leave it as a future work. If one proves convergence in finite-dimensional distribution for some rate constructs within the dynamics of $\left(\bar{Y}^N(t)\right)$, then $\left(\bar{Y}^N(t)\right)$ converges weakly to the solution of a fluid limit ODE as stated in Theorem 32.

The epidemics dynamics is defined as follows. Once a node i is infected, it activates d clocks (one per neighbor) associated with infection and another clock for healing. The clocks for infection and healing are all independent and exponentially distributed with rates γ and μ , respectively. Whenever an infection clock of an infected node i ticks, the corresponding neighbor is infected. If this latter is already infected, then nothing happens (exclusion principle). If the healing clock ticks, node i heals and the ringing of its clocks induce no further infections until it gets infected again. This peer-to-peer model is equivalent to the

model in Chapter 5 for complete-multipartite networks, if now we assume only one node per island. Fig. 6.2 subsumes the infection model.

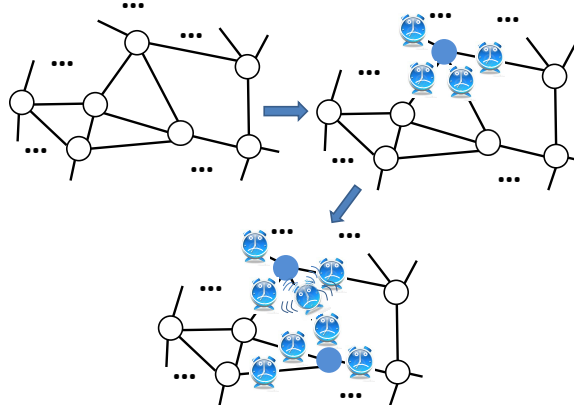


Figure 6.2: Illustration of the peer-to-peer diffusion model over a regular network with degree $d = 4$. Once a node is infected, it activates d clocks – one per neighbor – associated with infection, plus one clock for healing. For simplicity, we omitted the healing clock in the figure. All clocks are exponentially distributed and independent. The clocks for infection bear a rate of γ , while the healing clocks have rate μ .

Under this model, the process associated with the fraction of infected nodes over time $\left(\bar{Y}^N(t)\right)$ admits the following pathwise characterization:

$$\bar{M}^N(\omega, t) = \bar{Y}^N(\omega, t) - \bar{Y}^N(\omega, 0) + (\mu - d\gamma) \int_0^t \bar{Y}^N(\omega, s) ds \quad (6.2)$$

$$+ \gamma \int_0^t \underbrace{\frac{X^{N\top}(\omega, s) A^N X^N(\omega, s)}{N}}_{\bar{R}^N(\omega, s)} ds, \quad (6.3)$$

where $\left(\bar{M}^N(t)\right)$ is a martingale. Note that $\left(\bar{Y}^N(t)\right)$ is not Markov since $\left(\bar{R}^N(t)\right)$ is coupled with the microstate $\left(X^N(t)\right)$. Also note that the topology of the network plays a role in the dynamics of $\left(\bar{Y}^N(t)\right)$ via the adjacency matrix A^N in equation (6.2)-(6.3). In our case, the network is regular.

For technical reasons, it is convenient to re-write (6.2)-(6.3) as

$$\begin{aligned} \bar{M}^N(t) &= \bar{Y}^N(t) - \bar{Y}^N(0) + \gamma \int_0^t F^N\left(\bar{Y}^N(s)\right) ds + (\mu - d\gamma) \int_0^t \bar{Y}^N(s) ds \\ &\quad + \gamma \int_0^t \left(\bar{R}^N(s) - F^N\left(\bar{Y}^N(s)\right)\right) ds, \end{aligned}$$

where $F^N : [0, 1] \rightarrow \mathbb{R}$, is some sequence of functions with the following properties

- F^N is differentiable for every N ;
- $\limsup_{N \rightarrow \infty} \sup_{y \in [0, 1]} |F^N(y)| = M < \infty$ (Uniform Boundedness).

Our goal in this Section is to show that

$$\left(\bar{R}^N(t) - F^N(\bar{Y}^N(t)) \right)$$

is tight, and we further prove that, if

$$\left(\bar{R}^N(t) - F^N(\bar{Y}^N(t)) \right) \xrightarrow{f.d.d.} 0$$

for some uniformly-convergent sequence F^N with Lipschitz limit then, as will be clearer, we obtain weak convergence for $(\bar{Y}^N(t))$.

The next Theorem provides a criterion for tightness and weak convergence of càdlàg stochastic processes. Before proceeding, we say that T^N is a stopping time with respect to the natural filtration $\mathcal{F}_t^{\bar{Z}^N}$ of the process $(\bar{Z}^N(t))$, when

$$\{\omega : T^N(\omega) \leq t\} \in \mathcal{F}_t^{\bar{Z}^N},$$

for all time $t, t \geq 0$. In what follows, we assume that the stopping times T^N have finite alphabet, i.e.,

$$T^N(\omega) \in \mathcal{A}^N \subset [0, T]$$

for almost all $\omega \in \Omega$, for all N , with \mathcal{A}^N a finite subset of $[0, T]$.

Theorem 29 (Aldous' Tightness Criterion;[24]). *Let $\bar{Z}_\infty(t)$ be almost surely continuous and*

$$(\bar{Z}^N(t)) \xrightarrow{f.d.d.} (\bar{Z}_\infty(t)).$$

Then,

$$(\bar{Z}^N(t)) \Rightarrow (\bar{Z}_\infty(t))$$

if and only if

$$\bar{Z}^N(T^N - \Delta_N) - \bar{Z}^N(T^N) \xrightarrow{\mathbb{P}} 0, \tag{6.4}$$

for any sequence Δ_N with $\Delta_N \rightarrow 0$; and any sequence of finite-alphabet stopping times T^N , with respect to $\mathcal{F}_t^{\bar{Z}^N}$.

We observe that

$$R^N(t) = X^{N\top}(t)A^N X^N(t)$$

represents the number of unbalanced edges – i.e., edges that connect an infected node to a healthy node – at time t . At each clock triggered event – healing or infection, – there are at most $d - 1$ balanced edges that become unbalanced (in case of infection) or d unbalanced edges that become balanced (in case of healing),

i.e., $(R^N(t))$ has bounded jumps

$$j(R^N(t)) = \left| R(t) - \lim_{s \uparrow t} R(s) \right| \leq d$$

almost surely and for all t and N . Therefore, we have

$$\left| \bar{R}^N(\omega, t + \Delta_N) - \bar{R}^N(\omega, t) \right| \leq \frac{1}{N} \mathcal{N}_{\lambda(N)}(\omega, (0, \Delta_N))$$

for all $\omega \in \Omega$, where $\mathcal{N}_{\lambda(N)}(\omega, (0, s))$ is the Poisson process, with rate

$$\lambda(N) := (\gamma + \mu)dN,$$

that counts d times the number of clocks for infection or healing triggered at the time interval $(0, s)$ at the d -regular network with N nodes. The next Theorem shows that

$$\frac{1}{N} \mathcal{N}_{\lambda(N)}(\omega, (0, \Delta_N)) \xrightarrow{\mathbb{P}} 0, \quad (6.5)$$

whenever $\Delta_N \rightarrow 0$.

Theorem 30. *Let $\epsilon > 0$ be fixed and consider any sequence Δ_N with*

$$\Delta_N \rightarrow 0. \quad (6.6)$$

Then,

$$\mathbb{P}(\mathcal{N}_{\lambda(N)}(0, \Delta_N) \geq N\epsilon) \rightarrow 0$$

where we defined

$$\lambda(N) := (\gamma + \mu)dN \quad (6.7)$$

Proof. We have

$$\mathbb{P}(\mathcal{N}_{\lambda(N)}(0, \Delta_N) \geq N\epsilon) = e^{-\lambda(N)\Delta_N} \sum_{k \geq N\epsilon} \frac{(\lambda(N)\Delta_N)^k}{k!} \quad (6.8)$$

Define

$$g_k(N) := \frac{(\lambda(N)\Delta_N)^k}{k!}$$

First, let $r := \gamma + \mu$ and note that

$$g_{(\epsilon N)}(N) = \frac{(rd)^{\epsilon N} N^{\epsilon N} \Delta_N^{\epsilon N}}{(\epsilon N)!} = \frac{(rd)^{\epsilon N} (\epsilon N)^{\epsilon N} \Delta_N^{\epsilon N}}{\epsilon^{\epsilon N} (\epsilon N)!} = \left(\frac{(\epsilon N)^{\epsilon N}}{(\epsilon N)!} \right) \times \left(\frac{rd \Delta_N}{\epsilon} \right)^{\epsilon N}.$$

Therefore,

$$\begin{aligned}
 \lim_{N \rightarrow \infty} g_{(\epsilon N)}(N) &= \lim_{N \rightarrow \infty} \left(\frac{(\epsilon N)^{\epsilon N}}{(\epsilon N)!} \right) \times \lim_{N \rightarrow \infty} \left(\frac{rd\Delta_N}{\epsilon} \right)^{\epsilon N} \\
 &= \lim_{N \rightarrow \infty} \frac{e^{\epsilon N}}{\sqrt{2\pi\epsilon N}} \times \lim_{N \rightarrow \infty} \left(\frac{rd\Delta_N}{\epsilon} \right)^{\epsilon N} \\
 &= \lim_{N \rightarrow \infty} \left(\frac{erd\Delta_N}{\epsilon} \right)^{\epsilon N} \times \frac{1}{\sqrt{2\pi\epsilon N}} = 0,
 \end{aligned} \tag{6.9}$$

where the second equality is due to the Stirling's formula

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} = 1.$$

Therefore, $g_{(K)}(N)$ converges to zero faster than

$$O\left(\frac{\alpha^N}{\sqrt{N}}\right)$$

for any $\alpha \in (0, 1)$ and any $K \geq N\epsilon$. Thus, the sum in equation (6.8) converges to zero, since

$$\sum_{K \geq N} \frac{\alpha^K}{\sqrt{K}} \leq \sum_{K \geq N} \alpha^K \leq \int_{N-1}^{\infty} \alpha^x dx = \int_{N-1}^{\infty} e^{-x|\ln(\alpha)|} dx = \frac{e^{-N+1}}{|\ln(\alpha)|} \rightarrow 0.$$

□

The previous Theorem will imply that the process

$$\tilde{F}^N(X^N(t)) := \left(\bar{R}^N(t) - F^N(\bar{Y}^N(t)) \right)$$

is tight. First, let us define

$$M := \limsup_{N \rightarrow \infty} \sup_{\{y \in [0,1]\}} |F^N(y)| < \infty$$

and note that $\tilde{F}^N(X^N(t))$ has bounded jumps

$$j\left(\bar{R}^N(t) - F^N(\bar{Y}^N(t))\right) \leq d + M.$$

Theorem 31. *The process*

$$\tilde{F}^N(X^N(t)) := \left(\frac{X^{N\top}(t)A^N X^N(t)}{N} - F^N(\bar{Y}^N(t)) \right)$$

fulfills the condition entailed in equation (6.4).

Proof. Note that

$$\left| \tilde{F}^N (X^N (t + \Delta^N)) - \tilde{F}^N (X^N (t)) \right| \leq_{a.s.} \mathcal{N}_{\lambda(N)} (t, t + \Delta^N),$$

where now we defined,

$$\lambda(N) := (\gamma + \mu)(d + M)N,$$

and therefore,

$$\begin{aligned} & \mathbb{P} \left(\frac{1}{N} \left| \tilde{F}^N (X^N (T^N + \Delta^N)) - \tilde{F}^N (X^N (T^N)) \right| > \epsilon \mid T^N = s \right) = \\ &= \mathbb{P} \left(\left| \tilde{F}^N (X^N (T^N + \Delta^N)) - \tilde{F}^N (X^N (T^N)) \right| > N\epsilon \mid T^N = s \right) \\ &= \mathbb{P} \left(\left| \tilde{F}^N (X^N (s + \Delta^N)) - \tilde{F}^N (X^N (s)) \right| > N\epsilon \mid T^N = s \right) \\ &\leq \mathbb{P} (\mathcal{N}_{\lambda(N)} (s, s + \Delta^N) > N\epsilon \mid T^N = s) \\ &\leq \mathbb{P} (\mathcal{N}_{\lambda(N)} (0, \Delta^N) > N\epsilon) \rightarrow 0. \end{aligned}$$

for all $s \in \mathcal{A}^N$. Therefore,

$$\begin{aligned} & \mathbb{P} \left(\left| \tilde{F}^N (X^N (T^N + \Delta^N)) - \tilde{F}^N (X^N (T^N)) \right| > \epsilon \right) = \\ & E \left[\mathbb{P} \left(\left| \tilde{F}^N (X^N (T^N + \Delta^N)) - \tilde{F}^N (X^N (T^N)) \right| > \epsilon \mid T^N \right) \right] \rightarrow 0, \end{aligned}$$

and thus we have convergence in probability

$$\tilde{F}^N (X^N (T^N + \Delta^N)) - \tilde{F}^N (X^N (T^N)) \xrightarrow{\mathbb{P}} 0$$

and the process $(\tilde{F}^N (X^N (t)))$ is tight. □

Establishing convergence in finite-dimensional distribution has shown to be challenging, and we shall pursue it as future work. Next, we present the main Theorem of this Section that states that if $(\tilde{F}^N (X^N (t)))$ converges in finite-dimensional distribution, then we can characterize the emergent dynamics of the process $(\bar{Y}^N (t))$ as an ODE.

Theorem 32. *Let $F^N : [0, 1] \rightarrow \mathbb{R}$ be a sequence of differentiable functions converging uniformly to a differentiable map $F : [0, 1] \rightarrow \mathbb{R}$. If the sequence of initial conditions $\bar{Y}^N (0)$ is such that*

$$(\tilde{F}^N (X^N (t))) = \left(\frac{X^{N\top} (t) A^N X^N (t)}{N} - F^N (\bar{Y}^N (t)) \right) \xrightarrow{f.d.d.} 0$$

and

$$\bar{Y}^N (0) \xrightarrow{d} y_0,$$

then,

$$(\bar{Y}^N (t)) \Rightarrow (y(t, y_0))$$

where, $(y(t, y_0))$ is the solution to

$$\dot{y}(t) = (d\gamma - \mu)y(t) - \gamma F(y(t)). \quad (6.10)$$

with initial condition y_0 .

Proof. **i)** $(\bar{M}^N(t)) \Rightarrow 0$: The characterization of the martingale and the proof that it converges weakly and in probability to zero follows similarly to as done in Chapters 4-5.

(ii) The family $(\bar{Y}^N(t))$ is tight: This follows since $(\bar{M}^N(t)) \Rightarrow_{\mathbb{P}} 0$, and due to the fact that at each new event – healing or infection – the jumps

$$j\left(\tilde{F}^N(X^N(t))\right) \leq_{a.s.} d + M.$$

are bounded.

iii) If $(\bar{Y}^{N_k}(t)) \Rightarrow (\bar{Y}(t))$ then, $(\bar{Y}(t))$ is solution to the ODE (6.10). The stochastic integral dynamics associated with $(\bar{Y}^N(t))$ is given by

$$\begin{aligned} \bar{M}^N(\omega, t) &= \underbrace{\bar{Y}^N(\omega, t) - \bar{Y}^N(\omega, 0) + \gamma \int_0^t F^N(\bar{Y}^N(\omega, s)) ds + (\mu - d\gamma) \int_0^t \bar{Y}^N(\omega, s) ds}_{:= \mathcal{F}_a(\bar{Y}^N(t))} \\ &\quad + \underbrace{\gamma \int_0^t \left(\frac{X^{N\top}(\omega, s) A^N X^N(\omega, s)}{N} - F^N(\bar{Y}^N(\omega, s)) \right) ds}_{:= \mathcal{F}_b(\bar{R}^N(t), \bar{Y}^N(t))}. \end{aligned}$$

It follows that,

$$\begin{aligned} \bar{M}^N(t) &= \mathcal{F}_a(\bar{Y}^N(t)) + \mathcal{F}_b(\bar{R}^N(t), \bar{Y}^N(t)) \\ \Downarrow 1 \quad \quad \quad \Downarrow 2 \quad \quad \quad \Downarrow 3 \\ 0 &\stackrel{d}{=} \mathcal{F}_a(\bar{Y}(t)) + 0 \end{aligned} \quad (6.11)$$

where the convergence 1 is due to $(\bar{M}^N(t)) \Rightarrow 0$; convergence 2 is due to the measurability of \mathcal{F}_a , the Skorokhod's representation theorem, similarly to as done in Chapters 4-5, and also the uniform convergence of F^N ; convergence 3 is due to Aldous' tightness criterion along with the Theorem 30. The equality in distribution in the last line of (6.11) holds due to Slutsky's Theorem (applied to càdlàg processes, refer to [20]).

iv) $(\bar{Y}^N(t)) \Rightarrow (\bar{Y}(t))$, **where $(\bar{Y}(t))$ is solution to the ODE (6.10).** Indeed, the underlying vector field in the limit equation (6.11) is Lipschitz, therefore the solution to (6.11) exists and is unique for each initial condition. Thus, any convergent subsequence $(\bar{Y}^{N_k}(t))$ converges to the same weak limit given by the unique solution to (6.10) and thus, the whole sequence converges. \square

6.3 Epidemics over Regular Networks with Shuffling: F.D.D Convergence

In this Section, we introduce the peer-to-peer model of diffusion over a regular network with shuffling, and we show that the rate constructs associated with the fraction of infected nodes process $(\bar{Y}^N(t))$ converge in finite dimensional distribution to a continuous function. If further tightness conditions hold, then we have weak convergence for the process $(\bar{Y}^N(t))$ with emergent dynamics characterized by an ODE, as stated in the main Theorem of this Section, Theorem 37. Whether the tightness conditions are met is left as future work.

The model of infection under shuffling goes as follows. Once an infection or healing takes place, the nodes are randomly rearranged. Then, the infected nodes activate d infection clocks (one per neighbor) and one clock for healing, and the next infection spreads similarly to the model of the previous section. The vector microprocess $(X_{\text{sh}}^N(t))$ collects the state – infected or healthy – of each node **after the shuffling**. That is, the intermediate state between infection and shuffling (refer to Fig. 6.3) is not observable through $(X_{\text{sh}}^N(t))$. Fig. 6.3 summarizes the model, for a cycle network.

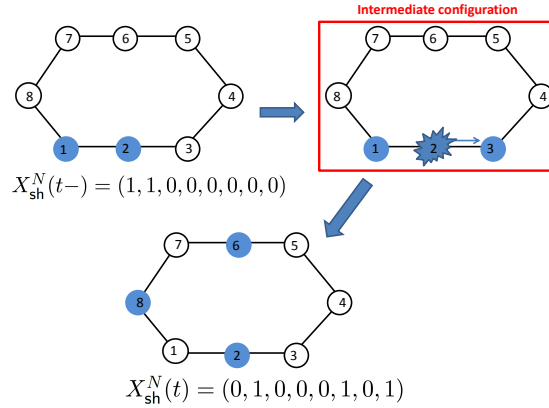


Figure 6.3: Illustration of the diffusion with shuffling. We omit the infection and healing clocks for simplicity. After a node is infected, the micro-state is permuted to display the new state $X_{\text{sh}}^N(t)$. Note that the intermediate state just after infection is not accounted by $X_{\text{sh}}^N(t)$.

The shuffling after event-triggering may be interpreted as a simplistic attempt to model a diffusion of a virus over a population that is wandering randomly around and preserving the degree d , as opposed to the static network framework. In other words, the diffusion under shuffling is a simple model for epidemics over a dynamical network of agents.

The process $(\bar{Y}_{\text{sh}}^N(t))$ admits a pathwise characterization that is similar to the process $(\bar{Y}^N(t))$ of the previous Section, but now in terms of the shuffling dynamics

$$\bar{Y}_{\text{sh}}^N(\omega, t) = \bar{Y}_{\text{sh}}^N(\omega, 0) + \bar{M}^N(\omega, t) + d\gamma \int_0^t (\bar{Y}_{\text{sh}}^N(\omega, s))^2 ds + (\mu - d\gamma) \int_0^t \bar{Y}_{\text{sh}}^N(\omega, s) ds$$

$$+ \gamma \int_0^t \left(\frac{X_{\text{sh}}^{N\top}(\omega, s) A^N X_{\text{sh}}^N(\omega, s)}{N} - d \left(\bar{Y}_{\text{sh}}^N(\omega, s) \right)^2 \right) ds.$$

Define the càdlàg process

$$\left(\bar{R}_{\text{sh}}^N(t) \right) := \left(\frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} \right),$$

and its conditional expectation with respect to $\left(\bar{Y}_{\text{sh}}^N(t) \right)$ as

$$\left(\hat{R}_{\text{sh}}^N(t) \right) := E \left(\frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} \middle| \bar{Y}_{\text{sh}}^N(t) \right).$$

Due to the random shuffle, we have that for each t ,

$$\mathbb{P} \left(X_{\text{sh}}^N(t) \in A \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha \right) = \frac{\#A}{\#S^N(\alpha)} = \frac{\#A}{\binom{N}{\alpha N}},$$

where

$$\alpha \in \mathcal{Y}^N := \left\{ 0, \frac{1}{N}, \frac{2}{N}, \dots, 1 \right\}$$

and

$$A \subset S^N(\alpha) := \left\{ x \in \{0, 1\}^N : \frac{\mathbf{1}^\top x}{N} = \alpha \right\}.$$

In particular,

$$\mathbb{P} \left(X_{\text{sh}}^N(t) \in A \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha \right) \tag{6.12}$$

only depends on α and N , and does not depend on the time t , $t \geq 0$. The next lemma characterizes some statistics of the process.

Lemma 33. *The expectation and variances of the rate process $\left(\bar{R}_{\text{sh}}^N(t) \right)$ over time conditioned to $\left(\bar{Y}_{\text{sh}}^N(t) \right)$ is given by*

$$E \left(\frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha \right) = d \frac{\alpha(\alpha N - 1)}{N - 1}$$

Proof. We start by computing for $i \neq j$

$$\begin{aligned} E \left(X_{i\text{sh}}^N(t) X_{j\text{sh}}^N(t) \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha \right) &= \mathbb{P} \left(X_{i\text{sh}}^N(t) X_{j\text{sh}}^N(t) = 1 \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha \right) \\ &= \mathbb{P} \left(X_{i\text{sh}}^N(t) = 1, X_{j\text{sh}}^N(t) = 1 \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha \right) \end{aligned}$$

$$= \frac{\binom{N-2}{\alpha N-2}}{\binom{N}{\alpha N}} = \frac{\alpha(\alpha N-1)}{N-1}.$$

Now, the quadratic term

$$X_{\text{sh}}^{N\top}(t)A^N X_{\text{sh}}^N(t)$$

is the sum of dN monomials $X_{i\text{sh}}^N(t)X_{j\text{sh}}^N(t)$. Therefore,

$$E\left(\frac{X_{\text{sh}}^{N\top}(t)A^N X_{\text{sh}}^N(t)}{N} \middle| \bar{Y}_{\text{sh}}^N(t) = \alpha\right) = d\frac{\alpha(\alpha N-1)}{N-1}.$$

□

The next Theorem states that the conditional variance associated with the normalized quadratic rate converges uniformly to zero.

Theorem 34. Define $f^N : [0, 1] \rightarrow \mathbb{R}_+$ with

$$f^N(\alpha) := \begin{cases} \text{Var}\left(\frac{X_{\text{sh}}^{N\top}A^N X_{\text{sh}}^N}{N} \middle| \bar{Y}_{\text{sh}}^N = \alpha\right) & , \text{ if } \alpha \in \mathcal{Y}^N \\ 0 & , \text{ if otherwise.} \end{cases}$$

Then, f^N converges uniformly (in α) to zero, i.e.,

$$\lim_{N \rightarrow \infty} \sup_{\alpha \in [0,1]} |f^N(\alpha)| = 0.$$

Proof.

$$\begin{aligned} f^N(\alpha) &= \frac{2\alpha d(\alpha N-1)}{N(N-1)} + \frac{4\alpha d(d-1)(\alpha N-1)(\alpha N-2)}{N(N-1)(N-2)} + \frac{d^2\alpha(\alpha N-1)(\alpha N-2)(\alpha N-3)}{(N-1)(N-2)(N-3)} \\ &\quad - \frac{2d(2d-1)\alpha(\alpha N-1)(\alpha N-2)(\alpha N-3)}{N(N-1)(N-2)(N-3)} - \left(\frac{2d\alpha(\alpha N-1)}{N-1}\right)^2 \\ &\leq \frac{2d}{N} + \frac{4d(d-1)}{N} + \left\{ \frac{d^2\alpha(\alpha N-1)(\alpha N-2)(\alpha N-3)}{(N-1)(N-2)(N-3)} - \frac{d^2\alpha^2(\alpha N-1)^2}{(N-1)^2} \right\} \\ &\leq \frac{2d(2d-1)}{N} + d^2 \left\{ \frac{\alpha(\alpha N-1)}{N-1} - \frac{(\alpha N-2)(\alpha N-3)}{(N-2)(N-3)} \right\} \\ &\leq \frac{2d(2d-1)}{N} + d^2 \left\{ \alpha^2 - \frac{(\alpha N-2)(\alpha N-3)}{(N-2)(N-3)} \right\} \\ &= \frac{2d(2d-1)}{N} + d^2 \left\{ \frac{5\alpha(1-\alpha)N + 6(\alpha^2-1)}{(N-2)(N-3)} \right\} \end{aligned}$$

$$\leq \frac{2d(2d-1)}{N} + d^2 \frac{5N}{(N-2)(N-3)}.$$

Therefore,

$$\begin{aligned} \lim_{N \rightarrow \infty} \sup_{\alpha \in [0,1]} |f^N(\alpha)| &\leq \lim_{N \rightarrow \infty} \sup_{\alpha \in [0,1]} \frac{2d(2d-1)}{N} + d^2 \frac{5N}{(N-2)(N-3)} \\ &= \lim_{N \rightarrow \infty} \frac{2d(2d-1)}{N} + d^2 \frac{5N}{(N-2)(N-3)} = 0. \end{aligned}$$

□

Corollary 35.

$$\left\| \frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} - d \frac{\bar{Y}_{\text{sh}}^N(t) (\bar{Y}_{\text{sh}}^N(t) N - 1)}{N-1} \right\|_{\mathcal{L}_2} \rightarrow 0$$

for all $t, t \geq 0$.

Proof. It follows from Theorem 34 that for all $\epsilon > 0$,

$$\exists N(\epsilon) \in \mathbb{N} : \text{Var} \left(\frac{X_{\text{sh}}^{N\top} A^N X_{\text{sh}}^N}{N} \middle| \bar{Y}_{\text{sh}}^N \right) <_{a.s.} \epsilon, \quad \forall N \geq N(\epsilon),$$

therefore,

$$E \left[E \left[\left(\frac{X_{\text{sh}}^{N\top} A^N X_{\text{sh}}^N}{N} - d \frac{\bar{Y}_{\text{sh}}^N (\bar{Y}_{\text{sh}}^N N - 1)}{N-1} \right)^2 \middle| \bar{Y}_{\text{sh}}^N \right] \right] \leq \epsilon.$$

or equivalently,

$$\left\| \frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} - d \frac{\bar{Y}_{\text{sh}}^N(t) (\bar{Y}_{\text{sh}}^N(t) N - 1)}{N-1} \right\|_{\mathcal{L}_2} \rightarrow 0.$$

for all $t, t \geq 0$.

□

Corollary 36.

$$\left(\frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} - d \frac{\bar{Y}_{\text{sh}}^N(t) (\bar{Y}_{\text{sh}}^N(t) N - 1)}{N-1} \right) \rightarrow_{f.d.d} 0.$$

Proof. Let $t_1, \dots, t_n \in [0, T]$. From the corollary 35,

$$\bar{Z}^N(t_i) \xrightarrow{\mathcal{L}_2} 0, \tag{6.13}$$

for all $i = 1, 2, \dots, n$. Thus, from Chebyshev's inequality

$$\bar{Z}^N(t_i) \xrightarrow{\mathbb{P}} 0, \tag{6.14}$$

for all i and therefore,

$$\begin{aligned} \mathbb{P}\left(\left\|\left(\bar{Z}^N(t_1), \dots, \bar{Z}^N(t_n)\right)\right\| > \epsilon\right) &\leq \mathbb{P}\left(\left|\bar{Z}^N(t_1)\right| + \dots + \left|\bar{Z}^N(t_n)\right| > \epsilon\right) \\ &\leq \mathbb{P}\left(\left|\bar{Z}^N(t_1)\right| > \frac{\epsilon}{n}\right) + \dots + \mathbb{P}\left(\left|\bar{Z}^N(t_n)\right| > \frac{\epsilon}{n}\right) \\ &\rightarrow 0. \end{aligned} \quad (6.15)$$

In particular,

$$\left(\bar{Z}^N(t_1), \dots, \bar{Z}^N(t_n)\right) \rightarrow_d 0.$$

Equivalently,

$$\left(\bar{Z}^N(t)\right) \rightarrow_{f.d.d} 0.$$

□

Establishing tightness has shown to be challenge and we leave as future work.

Theorem 37. *If the sequence*

$$\left(\frac{X_{\text{sh}}^{N\top}(t) A^N X_{\text{sh}}^N(t)}{N} - d \left(\bar{Y}_{\text{sh}}^N(t)\right)^2\right)$$

is tight then,

$$\left(\bar{Y}_{\text{sh}}^N(t)\right) \Rightarrow (y(t, y_0))$$

where, $(y(t, y_0))$ is the solution to

$$\dot{y}(t) = \gamma dy(t) (1 - y(t)) - \mu y(t). \quad (6.16)$$

Proof. We do not repeat all the steps here. We just refer that now it is not clear that tightness hold as the jump $j(R^N(t))$ is not almost surely bounded as in the non-shuffling case. □

6.4 Numerical Simulations

In this Section, we present two numerical simulations for epidemics over a cycle network under the peer-to-peer models explored in the previous Sections: static network and shuffling. All simulations were ran in Matlab.

In Fig. 6.4, we contrast the evolution of the epidemics over a cycle network under the shuffling model (refer to the noisy blue/darker color curves) with the solution to the ODE (6.16) (refer to the smooth red/lighter color curves), where both evolutions depart from the same initial condition. Fig. 6.4 suggests a concentration of mass phenomenon as the number of nodes grows large. The simulation does not *disprove* weak convergence of the sequence $\left(\bar{Y}^N(t)\right)$ to the solution to the ODE (6.16).

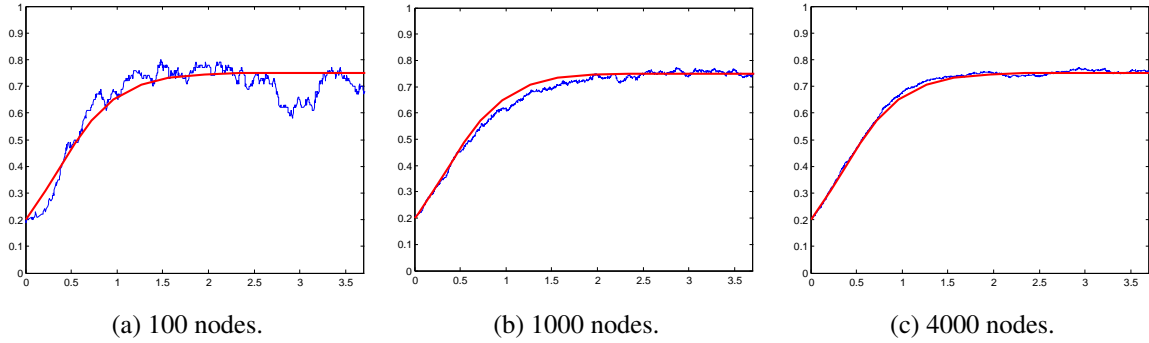


Figure 6.4: Evolution of the peer-to-peer epidemics with shuffling over cycle networks with different number of nodes: 100, 1000 and 4000, respectively. The noisy blue/darker color curves represent the evolution of the epidemics under the shuffling model, and the smooth red/lighter color curves represent the solution to the ODE (6.16) with the same initial condition. We assumed $\mu = 1$ and $\gamma = 2$.

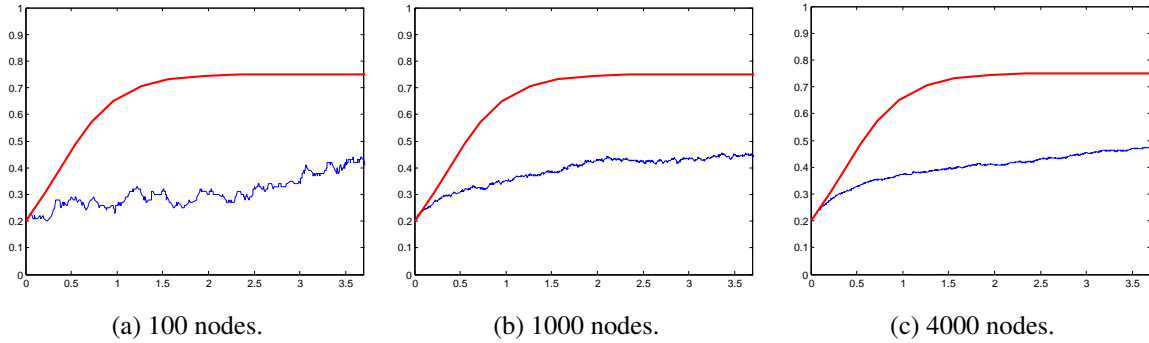


Figure 6.5: Evolution of the peer-to-peer epidemics without shuffling over cycle networks with different number of nodes: 100, 1000 and 4000, respectively. The noisy blue/darker color curves represent the evolution of the peer-to-peer epidemics driven by the random clocks, and the smooth red/lighter color curves represent the solution to the ODE (6.16) with the same initial condition. We assumed $\mu = 1$ and $\gamma = 2$.

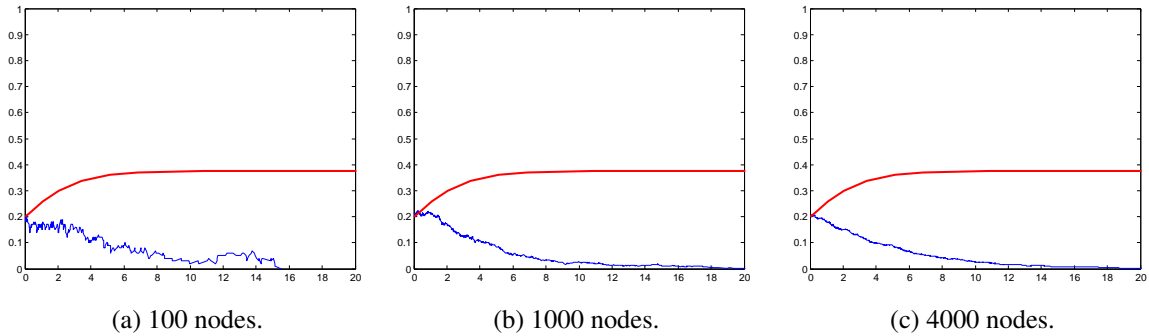


Figure 6.6: Evolution of the peer-to-peer epidemics without shuffling over cycle networks with different number of nodes: 100, 1000 and 4000, respectively. The noisy blue/darker color curves represent the evolution of the peer-to-peer epidemics driven by the random clocks, and the smooth red/lighter color curves represent the solution to the ODE (6.16) with the same initial condition. We assumed $\mu = 1$ and $\gamma = 0.8$.

In Fig. 6.5 and Fig 6.6, we set the evolution of the epidemics over a cycle network under the static network model (refer to the noisy blue/darker color curves) against the solution to the ODE (6.16) (refer to the smooth red/lighter color curves), where both evolutions depart from the same initial condition. Fig. 6.5 and Fig 6.6 suggest that the ODE (6.16) turns out to be a rough model for the peer-to-peer epidemics over static networks. In particular, in Fig. 6.6, and for particular rate parameters $\gamma = 0.8$ and $\mu = 1$, the ODE (6.16) predicts resilience of the virus (red/dark curves), whereas in the actual peer-to-peer epidemics under the static network model, the virus dies out for the same rate parameters $\gamma = 0.8$ and $\mu = 1$. Similar ODE's are often coined in the literature of epidemics (refer for instance to [12, 13]) with the purpose of studying this phenomenon over arbitrary networks.

6.5 Concluding Remarks

In this Chapter, we presented our current work on the framework of emergent dynamics associated with peer-to-peer epidemics over large-scale regular networks. As explained, to establish the functional weak law of large numbers dynamics for the fraction of infected nodes process, one must guarantee that such process and its built-in rate processes are tight and converge in finite-dimensional distribution to Lipschitz functions. The major difficulty in obtaining both conditions lies in the lack of Markov property associated with the macroscopic fraction of infected nodes state process. We presented two models of diffusion: static network and shuffling. In the static network model, we showed that the fraction of infected nodes $(\bar{Y}^N(t))$ process is tight as well as its corresponding rate processes, but convergence in finite dimensional distribution is hard to obtain and we leave it as future work. On the other hand, under the shuffling model, the rates associated with the sequence $(\bar{Y}_{sh}^N(t))$ converge in finite-dimensional distribution, but tightness is hard to determine. Establishing that $(\bar{Y}^N(t))$ or $(\bar{Y}_{sh}^N(t))$ converges weakly to the solution of an ODE would ultimately imply that such processes recover the Markov property at the limit of an infinite network. As illustrated in the numerical simulations in Fig. 6.5 and Fig. 6.6, in Section 6.4, deriving rigorously the macroscopic fluid limit dynamics in regular static networks is crucial to comprehend their correct behavior. Also, commonly assumed and used ODE's in the literature, as the ODE (6.16), can still model peer-to-peer epidemics, but under a rather dynamical environment as described in Section 6.4 via the shuffling of the nodes at each new infection or healing event.

Part II

Qualitative Analysis of the Emergent Dynamics

Chapter 7

Model Analysis

The goal of this Part II is to study the qualitative behavior of the weak limit (in the number of nodes) dynamics of bi-virus epidemics over large-scale complete-multipartite networks obtained in Chapter 5 of Part I, namely, we are to study the ODE's

$$\frac{d}{dt}y_i(t) = \left(\sum_{j \sim i} \gamma_{ji}^y y_j(t) \right) (1 - x_i(t) - y_i(t)) - y_i(t) \quad (7.1)$$

$$\frac{d}{dt}x_i(t) = \left(\sum_{j \sim i} \gamma_{ji}^x x_j(t) \right) (1 - x_i(t) - y_i(t)) - x_i(t), \quad (7.2)$$

where we assume the healing rate $\mu_i^{x,y} = 1$, and the component $y_i(t)$ (respectively, $x_i(t)$) stands for the weak limit fraction of nodes infected with the virus strain y (respectively, x) at island i at time t , $t \geq 0$. By studying the qualitative behavior of the ODE's (7.1)-(7.2), we mean to characterize the attractors and corresponding basins of attraction of the dynamical system (7.1)-(7.2), that will entail long-term information on the occurrence of natural selection between two strains of virus, or on the resilience of a single virus strain in a large-scale network.

There is a vast body of literature that studies the dynamics of epidemics via nonlinear ordinary differential equations models, similar to the ODE in equations (7.1)-(7.2), please refer to [12, 13, 34] and references therein. For instance, reference [5] characterizes properties of the equilibria of the nonlinear ordinary differential equations (7.1)-(7.2), restricted to the single virus case (i.e., $x_i(t) = 0$, for all t and all i) and with constant rates $\gamma_{ij} = \gamma$ for all i and j . Reference [5] considers such ODE's as an approximating model for the likelihood of infection of each node i in a general network under a single virus epidemics. In other words, in reference [5], $y_i(t)$ approximates the probability of infection of node i at time t . Therefore, the ODE's (7.1)-(7.2) may stand not only for the exact fluid limit dynamics of a bi-virus epidemics over large-scale complete-multipartite networks, but may also conform to a **postulated/approximating model** for the bi-virus dynamics over general networks.

Most of the literature in epidemics explore these ODE's by characterizing the equilibrium points of the equations. We are not concerned with characterizing the equilibria, but rather with studying the qualitative behavior of the bi-virus ODE's. To our knowledge, only reference [14] explores rigorously the qualitative behavior of such systems, for single virus dynamics via Lyapunov stability. For the bi-virus case, we consider that there is no natural Lyapunov function for the ODE's (7.1)-(7.2). Our results are based on a rigorous analysis using a novel methodology that we develop and present in this Part II, namely, in Chapters 8-9, and that allow us to overcome the lack of Lyapunov tools.

In this Chapter, we introduce the main notation and background definitions to be used in the Chapters 8-9. In Section 7.1, we define the main network constructs and introduce a partial order on the set of networks. This partial order is preserved by the virus dynamical flow over the networks. Section 7.2 introduces a special notation for the solutions of the ODE's (7.1)-(7.2) that is used in Chapters 8-9. Section 7.3 concludes the Chapter.

Preliminary Notation. We summarize the main notation used throughout this Chapter. \mathbb{R}_+ and \mathbb{R}_{++} represent the set of nonnegative and positive, respectively, real numbers. The set

$$\mathbb{N} = \{1, 2, \dots\}$$

represents the set of natural numbers. We define $\mathbf{1}_N \in \mathbb{R}^N$ (and $\mathbf{0}_N \in \mathbb{R}^N$) as the vectors with all entries equal to one (respectively, zero) – the subindex may be omitted whenever the dimension is clear from the context. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ and define $\mathbf{x} \geq \mathbf{y}$ as the pointwise inequality, i.e., $x_i \geq y_i$ for all $i = 1, \dots, N$. We refer to $\mathbf{x} \not\geq \mathbf{0}$ as $\mathbf{x} \notin \mathbb{R}_{++}^N$. Define $|V|$ as the cardinality of the set V . Let $x, y \in \mathbb{R}$ and define $x \wedge y$ as:

$$x \wedge y = x, \text{ if } x \leq y \text{ and } x \wedge y = y, \text{ if otherwise.}$$

Let $x, y \in \mathbb{R}$ and define $x \vee y$ as:

$$x \vee y = x, \text{ if } x \geq y \text{ and } x \vee y = y, \text{ if otherwise.}$$

We define the Hadamard pointwise product $\mathbf{x} \odot \mathbf{y}$ as

$$\mathbf{x} \odot \mathbf{y} = (x_1 y_1, \dots, x_N y_N).$$

We define $\mathbf{A} \otimes \mathbf{y}$ as the Kronecker product

$$\mathbf{A} \otimes \mathbf{y} = (a_{ij} y_j)_{ij}.$$

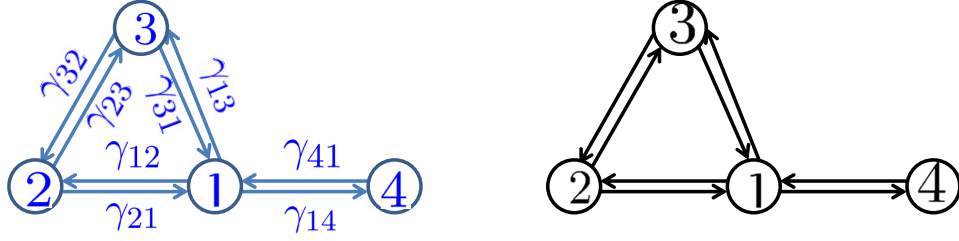


Figure 7.1: Graphical illustration on the left (blue/lighter color) of an e-network (all γ_{ij} sketched are positive). On the right (black/darker color) lies its support graph counterpart.

7.1 E-networks

In this Section, we introduce the main notation and definitions regarding the network constructs to be used throughout this part of the Thesis. We start by the basic concepts on graphs.

Definition 38 (Graph). *The ordered pair $G = (V, E)$, where $V \subset \mathbb{N}$ is finite and $E \subset V \times V$, is called a graph. The symbols V and E stand for the set of nodes and edges, respectively.*

Unless otherwise stated, for a node set V of cardinality $|V| = N$, we assume by default that

$$V = \{1, \dots, N\}.$$

Next, we define an e-network that we associate with a virus: it is a weighted directed graph whose edges determine the topology of infection and the weights represent the rates of infection between nodes. Refer to Fig. 7.1 for an illustration.

Definition 39 (E-Network). *We call the ordered pair $G = (V, \gamma)$, where $V \subset \mathbb{N}$ is finite and*

$$\gamma : V \times V \rightarrow \mathbb{R}_+$$

is a nonnegative function, an e-network. We define

$$\mathcal{E}(V) := \{G = (V, \gamma) : \gamma : V \times V \rightarrow \mathbb{R}_+\}$$

as the set of e-networks with fixed node set V . Whenever the particulars of V are not important, we write $\mathcal{E}(|V|)$ instead of $\mathcal{E}(V)$.

The $N \times N$ matrix/function γ is the infection rate of a virus between any two arbitrary nodes with $\gamma(i, j)$ standing for the rate of infection from node i towards node j . For convenience, we will write γ_{ij} – as in equations (7.1)-(7.2) – interchangeably with $\gamma(i, j)$. Different virus strains have different e-networks

associated. Whenever studying the bi-virus dynamics, we will refer to the pair

$$G^{x,y} = (G^x, G^y)$$

as the e-network associated to the bi-virus diffusion, where G^x and G^y represent the e-networks associated with the virus strains x and y , respectively. The path-flow graph of infection is specified by the support of γ that we now define.

Definition 40 (Support Graph of an E-Network). *Let $G = (V, \gamma)$ be an e-network. We call*

$$\text{supp}(G) = (V, \gamma^{-1}(\mathbb{R}_{++}))$$

the support graph of the e-network G , where $\gamma^{-1}(\mathbb{R}_{++})$ is the inverse image of \mathbb{R}_{++} by

$$\gamma : V \times V \rightarrow \mathbb{R}_+$$

or, in words, the set of edges (i, j) where $\gamma(i, j) > 0$.

The support graph represents the graph where an edge from node i to node j is placed whenever

$$\gamma(i, j) > 0$$

and not placed if otherwise

$$\gamma(i, j) = 0.$$

This provides us with the path-flow of infection of a virus associated with the e-network G . Thus, an e-network G not only specifies the topology of the network environment as seen by the virus, but also the quantitative rates $\gamma(i, j)$ of infection attached to each edge (i, j) . Figure 7.1 summarizes the definitions of an e-network and its support graph. Next, we define the neighborhood of a node.

Definition 41 (Neighborhood). *Let $G = (V, \gamma)$ be an e-network. Let $i \in V$ and define $\mathcal{N}(i, G)$ as the 1-hop neighborhood of node i in G . Write it as*

$$\mathcal{N}(i, G) := \{j \in V : j \rightarrow_G i\}$$

where $j \rightarrow_G i$ means that $\gamma(j, i) > 0$, i.e., node j can directly infect node i . We write $\mathcal{N}^2(i, G)$ as the 2nd order neighborhood of i , that is, $j \in \mathcal{N}^2(i, G)$ if and only if the shortest path connecting j to i (a.k.a. geodesic) has a length of 2 hops. Now, inductively, we say $j \in \mathcal{N}^n(i, G)$ if and only if there exists $k \in \mathcal{N}^{n-1}(i, G)$ with $j \rightarrow_G k$, i.e., the geodesic connecting j to i comprises n hops. Whenever there is no room for ambiguity, we suppress the e-network G in the notation, and we write instead $\mathcal{N}^n(i)$ and $i \rightarrow j$.

We will also refer to

$$d_i(G) := |\mathcal{N}(i, G)|$$

as the degree of node $i \in V$ for the e-network G . Note from the definition of neighborhood that

$$\mathcal{N}^l(i) \cap \mathcal{N}^m(i) = \emptyset$$

whenever $l \neq m$, i.e., the family $(\mathcal{N}^n(i))_n$ induces a partition on the node set V of the underlying e-network G . For certain symmetric conditions on the rate parameters γ of the strains of virus, the dynamical system (7.1)-(7.2) can be fully characterized from a qualitative point of view. We define next, the right object that turns the analysis easier.

Definition 42 (Regular E-Network). *We call $G \in \mathcal{E}(V)$ a regular e-network, whenever $\text{supp}(G) = (V, \overline{E})$ is a regular graph and γ is constant over \overline{E} , i.e., $d_i(G) = |\mathcal{N}(i, G)| = d$ for all $i = 1, \dots, N$ for some $d \in \mathbb{N}$, and the infection rates $\gamma(i, j)$ are the same across the edges $(i, j) \in \overline{E}$ of the support graph. We denote by $\mathcal{R}(V) \subset \mathcal{E}(V)$ the subset of regular e-networks with node set V . We refer to a d -regular e-network G , whenever all nodes $i \in V$ have degree d .*

The solutions to (7.1)-(7.2) over an e-network $G^{x,y}$ can be appropriately lower/upper bounded (in a sense that will be precise later) by flows under sub/super e-networks $G^{x,y}$. Next, we define an inequality that will provide with the appropriate notion of dominance of a virus (or an e-network) over another one.

Definition 43 (Inequalities). *Let $G_1 = (V_1, \gamma_1)$, $G_2 = (V_2, \gamma_2)$ be two e-networks. 1) We say $G_1 \leq G_2$ whenever $\text{supp}(G_1)$ is a subgraph of $\text{supp}(G_2)$; 2) We say $G_1 \preceq G_2$ whenever $V_1 = V_2$ and $\gamma_1 \leq \gamma_2$, where the latter inequality is pointwise; 3) We say $G_1 \prec G_2$ whenever $V_1 = V_2$ and $\gamma_1 < \gamma_2$. ■*

Note that the inequality \preceq is stronger than \leq . In words, $G_1 \preceq G_2$ if the rates of infection underlying G_1 lower-bound the rates of infection of G_2 at every edge. We will be able to show fairly general qualitative results for arbitrary e-networks $G^{x,y}$ on the infection dynamics by upper/lower bounding its dynamics by solutions over well understood regular e-networks that appropriately upper/lower bound the networks G^x and G^y – we will refer to them as outer/inner regular e-networks of G^x and G^y .

Definition 44 (Inner/Outer Regular E-Network). *Let $G = (V, \gamma)$ be an e-network. We define*

$$\underline{G} = \max \{S : S \preceq G, S \in \mathcal{R}(V)\}$$

as the greatest regular e-network dominated by G . Correspondingly,

$$\overline{G} := \min \{S : S \succeq G, S \in \mathcal{R}(V)\}$$

is defined as the smallest regular e-network that dominates G . We will refer to \underline{G} and \overline{G} as the inner and outer, respectively, regular e-networks of G .

In fact, not all graphs admit nontrivial regular subgraphs, and in such cases

$$\{S : S \preceq G, S \in \mathcal{R}(V)\} = \emptyset \text{ and } \underline{G} = (V, \mathbf{0})$$

is defined as the e-network of isolated nodes, i.e., with no connections ($\gamma(i, j) = 0$ for all $i, j \in V$). Reference [35] establishes that a typical large (Erdos-Rényi random) graph admits regular subgraphs with high probability. Since the complete graph is regular and it is a supper graph for any graph, any graph admits a non-degenerate regular outer-graph.

Aside Note: For generic e-networks G , the inner/outer regular e-networks \underline{G} and \overline{G} may not be uniquely determined. In fact, \underline{G} and \overline{G} are just representatives from the class of maximizers/minimizers in the definition. As will be clear, the results derived in our qualitative analysis are insensitive to the choice of the representatives \underline{G} and \overline{G} .

Definition 45 (Strong Inequality). Let $G^1 = (V, \gamma^1)$ and $G^2 = (V, \gamma^2)$ be two e-networks. We say $G^1 \preceq^* G^2$ whenever $\overline{G}^1 \preceq \underline{G}^2$. In words, we say that G^1 is **strongly-dominated** by G^2 whenever the outer-regular e-network \overline{G}^1 of G^1 is dominated by the inner-regular e-network \underline{G}^2 of G^2 .

One can readily check that ' \preceq^* ' defines a partial order on the set $\mathcal{E}(V)$. Note that the inequality ' \preceq^* ' is stronger than ' \preceq '.

7.2 Dynamical Flow in E-networks

The dynamical systems that we are interested in studying model the evolution over time of an epidemics over networks. Namely, as determined in Part I and for the single-virus case, $y_i(t, \mathbf{y}_0)$ represents the fluid limit of the fraction of infected nodes in island $i \in \{1, \dots, N\}$ at time t , $t \geq 0$, with initial distribution across islands

$$\mathbf{y}_0 \in [0, 1]^N,$$

in a large-scale complete-multipartite network whose (arbitrary) super-topology plus inter-island rates of infection are given by the e-network $G = (V, \gamma)$. Next, we define what we will refer to as the *flow* of a virus over an e-network G .

Definition 46 (Flow of Single-Virus ODE). Let $\mathcal{D}(V) = [0, 1]^{|V|}$. Define the function

$$\begin{aligned} \mathbf{y} : \mathbb{R}_+ \times \mathcal{D}(V) \times \mathcal{E}(V) &\longrightarrow \mathcal{D}(V), \\ (t, \mathbf{y}_0; G) &\longmapsto \mathbf{y}(t, \mathbf{y}_0; G) \end{aligned}$$

where $(\mathbf{y}(t, \mathbf{y}_0; G))_{t \geq 0}$ is the solution to the ODE

$$\dot{y}_i(t) = \underbrace{\left(\sum_{j \in V} \gamma(j, i) y_j(t) \right)}_{F_i(\mathbf{y}(t); G)} (1 - y_i(t)) - y_i(t) \quad (7.3)$$

for $i = 1, \dots, N$ with $N = |V|$, with initial condition $\mathbf{y}_0 \in \mathcal{D}(V)$ and underlying e-network $G = (V, \gamma) \in \mathcal{E}(V)$. For a fixed G , we refer to $(\mathbf{y}(t, \mathbf{y}_0; G))_{t \geq 0}$ as the flow of virus y over the e-network G . Whenever V is clear from the context, we will write \mathcal{D} and \mathcal{E} instead of $\mathcal{D}(V)$ or $\mathcal{E}(V)$.

We observe that, given an e-network $G \in \mathcal{E}(V)$, the flow $(\mathbf{y}(t, \mathbf{y}_0; G))$ to the ODE (7.3) is well defined – it is uniquely determined for all time $t, t \geq 0$. Indeed, the underlying vector field

$$\mathbf{F}(\cdot, G) = (F_1(\cdot, G), \dots, F_N(\cdot, G))$$

is (globally) Lipschitz over the domain

$$\mathcal{D} = [0, 1]^N$$

for all $G \in \mathcal{E}(V)$. Note that the set \mathcal{D} is invariant with respect to the dynamics, that is, $\mathbf{y}(t, \mathbf{y}_0; G) \in \mathcal{D}$ for all time t whenever $\mathbf{y}_0 \in \mathcal{D}$. The fact that \mathcal{D} is compact further implies that the solutions are defined for all $t, t \geq 0$. Whenever the underlying e-network G is clear from the context or previously fixed, we will write $y_i(t)$ or $y_i(t, \mathbf{y}_0)$ instead of $y_i(t, \mathbf{y}_0; G)$.

As an aside note, the notation $(\mathbf{y}(t, \mathbf{y}_0; G))_{t \geq 0}$ carrying the e-network G as an argument will be helpful as we will often compare the dynamics of evolution of the epidemics over different e-networks. For instance, we will assert that, if the e-network G^1 is dominated by G^2 , i.e., $G^1 \preceq G^2$, then

$$(\mathbf{y}(t, \mathbf{y}_0, G^1)) \leq (\mathbf{y}(t, \mathbf{y}_0, G^2)).$$

Our main focus in Part II of the Thesis is the study of the bi-virus epidemics dynamics. Let

$$\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{D}^2 \text{ and } G^{x,y} = (G^x, G^y) \in \mathcal{E}^2,$$

then, as proved in Part I, $x_i(t, \mathbf{z}_0, G^{x,y})$ (respectively, $y_i(t, \mathbf{z}_0, G^{x,y})$) can be interpreted as the fraction of nodes infected with virus x (respectively, y) at time $t, t \geq 0$, at island i , of a bi-virus epidemics with initial distribution $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$ and associated complete-multipartite networks whose super-topology plus inter-island rates of infection are given by the e-networks G^x and G^y , in the weak limit of a large network. We now define flow of a bi-virus in an e-network $G^{x,y}$.

Definition 47 (Flow of bi-Virus ODE). *Let $\mathcal{D}_2(V) := \{(\mathbf{x}, \mathbf{y}) \in \mathcal{D}(V) \times \mathcal{D}(V) : \mathbf{x} + \mathbf{y} \leq \mathbf{1}_N\}$. Define the function*

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) : \mathbb{R}_+ \times \mathcal{D}_2(V) \times \mathcal{E}^2(V) &\longrightarrow \mathcal{D}_2(V) \\ (t, \mathbf{z}_0; G^{x,y}) &\longmapsto (\mathbf{x}(t, \mathbf{z}_0; G^{x,y}), \mathbf{y}(t, \mathbf{z}_0; G^{x,y})). \end{aligned}$$

where $((\mathbf{x}, \mathbf{y})(t, \mathbf{z}_0; G^{x,y}))_{t \geq 0}$ is the solution to the ODE

$$\dot{x}_i(t) = \underbrace{\left(\sum_{j \in V} \gamma^x(j, i) x_j(t) \right) (1 - x_i(t) - y_i(t)) - x_i(t)}_{=: F_i(\mathbf{x}(t), \mathbf{y}(t); G^x)} \quad (7.4)$$

$$\dot{y}_i(t) = \underbrace{\left(\sum_{j \in V} \gamma^y(j, i) y_j(t) \right) (1 - x_i(t) - y_i(t)) - y_i(t)}_{=: F_i(\mathbf{x}(t), \mathbf{y}(t); G^y)}, \quad (7.5)$$

for $i = 1, \dots, N$ with $N = |V|$, with initial condition $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{D}_2(V)$ and underlying e -networks $G^{x,y} = (G^x, G^y)$ with $G^x = (V, \gamma^x)$ and $G^y = (V, \gamma^y)$. For a fixed $G^{x,y}$, we refer to

$$((\mathbf{x}, \mathbf{y})(t, (\mathbf{x}_0, \mathbf{y}_0)(G^x, G^y)))_{t \geq 0}$$

as the flow of viruses x and y over the corresponding e -networks G^x and G^y .

For the same reasons as in the single-virus dynamics, the flow

$$((\mathbf{x}, \mathbf{y})(t, (\mathbf{x}_0, \mathbf{y}_0); (G^x, G^y)))$$

is well defined – it is uniquely determined for all time $t, t \geq 0$, as the underlying vector field

$$\mathbf{F}(\cdot, G^{x,y}) = (\mathbf{F}(\cdot, G^x), \mathbf{F}(\cdot, G^y))$$

is Lipschitz over the invariant compact domain of interest \mathcal{D}_2 . Again, whenever clear from the context, we will write

$$y_i(t), x_i(t) \text{ or } y_i(t, \mathbf{z}_0), x_i(t, \mathbf{z}_0)$$

instead of

$$y_i(t, \mathbf{z}_0, G^{x,y}) \text{ or } x_i(t, \mathbf{z}_0, G^{x,y}).$$

We will also refer to $\overset{(n)}{y}_i(t, \mathbf{z}_0, G^{x,y})$ or $\overset{(n)}{x}_i(t, \mathbf{z}_0, G^{x,y})$ as the n th derivative of the flow to (7.3) or (7.4)-(7.5) at time $t, t \geq 0$.

7.3 Concluding Remarks

In this Chapter, we presented the central objects and notations to be adopted in the later Chapters 8-9. In Section 7.1, we defined an e-network $G = (V, \gamma)$ as a weighted directed graph, where the weights γ_{ij} represent the rates of infection of a virus between nodes i and j . If we have a bi-virus spread over a population V , then we have two e-networks G^x and G^y each associated with one virus strain. Also, we presented regular e-networks, whose symmetry leads to a simplification of the dynamical system (7.4)-(7.5), as will be shown in Chapters 8-9. Finally, we introduced two partial orders on the set of e-networks that represent the dominance of a strain of virus over another. In Section 7.2, we defined flows in e-networks as the solutions to (7.4)-(7.5) indexed by the underlying e-networks and initial conditions. We briefly referred to the framework underlying our approach in the next Chapters, namely, appropriate upper and lower bounds on the initial conditions and underlying e-networks are preserved by the flow of the dynamical system (7.4)-(7.5).

Chapter 8

Qualitative Analysis: Two-node E-networks

This Chapter considers the dynamics of epidemics in a two-node e-network that we refer to as

$$G = (V, \gamma) \in \mathcal{E}(V) \text{ with } |V| = 2.$$

In Section 8.1, we determine the qualitative behavior of the coupled system of two nonlinear ODE's (7.3) associated with the single virus epidemics. In Section 8.2, we study the dynamics (7.4)-(7.5) associated with the bi-virus case. In Section 8.3, we conclude the Chapter. For the sake of simplicity, in both Sections 8.1 and 8.2, we restrict the analysis to two-node e-networks. We extend the results for general e-networks in Chapter 9. Some of the results in this Chapter can be obtained via Lyapunov tools, but these tools are not naturally applicable to tackle bi-virus epidemics over general e-networks. The goal of this Chapter is to introduce our techniques in the simpler two-node e-network setting that admit the natural generalization presented in the next Chapter.

8.1 Single Virus Epidemics

In this Section, we determine the global attractors and corresponding basins of attraction of the dynamical system (7.3) restricted to single virus epidemics in a two-node e-network.

The next Theorem reveals a monotone aspect of the dynamical system (7.3) (refer to Chapter 7) that we now prove for two-node e-networks, but that will be further established in the setting of general e-networks – an upper-bound on the initial conditions is preserved by the flow of the dynamical system (7.3) through all time $t \geq 0$.

Theorem 48 (Monotonicity). *Let $G_1 \preceq G_2 \in \mathcal{E}(2)$ be two-node e-networks and $\mathbf{y}(0) \leq \mathbf{y}_0 \in \mathcal{D}(2)$. Then,*

$$\mathbf{y}(t, \mathbf{y}(0), G_1) \leq \mathbf{y}(t, \mathbf{y}_0, G_2), \quad \forall t \geq 0,$$

where $(\mathbf{y}(t, \mathbf{z}_0, G_i))$ is the flow of the dynamical system (7.3) with underlying e -network G_i and initial condition $\mathbf{z}_0 \in \mathcal{D}(2)$.

Proof. If $\mathbf{y}(0) = \mathbf{y}_0$ and $G_1 = G_2$ then, by uniqueness,

$$\mathbf{y}(t, \mathbf{y}(0), G_1) = \mathbf{y}(t, \mathbf{y}_0, G_2)$$

for all $t \geq 0$, and the result holds. Now, let $\mathbf{y}(0) \leq \mathbf{y}_0$ with $\mathbf{y}(0) \neq \mathbf{y}_0$ and $G_1 \preceq G_2$. Define

$$T = \inf \{t \geq 0 : \mathbf{y}(t, \mathbf{y}(0), G_1) \not\leq \mathbf{y}(t, \mathbf{y}_0, G_2)\}$$

and assume that $T < +\infty$. Since the flow is continuous and uniqueness is preserved for all $t \geq 0$, then,

$$y_1(T, \mathbf{y}(0), G_1) = y_1(T, \mathbf{y}_0, G_2) \text{ and } y_2(T, \mathbf{y}(0), G_1) < y_2(T, \mathbf{y}_0, G_2)$$

(up to a relabeling). Observe from equations (7.3) that this implies

$$\dot{y}_1(T, \mathbf{y}(0), G_1) < \dot{y}_1(T, \mathbf{y}_0, G_2).$$

Therefore, from Theorem 68 (refer to the Appendix),

$$\exists \epsilon_1 > 0 : y_1(t, \mathbf{y}(0), G_1) < y_1(t, \mathbf{y}_0, G_2), \forall T < t < T + \epsilon_1.$$

Moreover,

$$y_2(T, \mathbf{y}(0), G_1) < y_2(T, \mathbf{y}_0, G_2) \Rightarrow \exists \epsilon_2 > 0 : y_2(t, \mathbf{y}(0), G_1) < y_2(t, \mathbf{y}_0, G_2), \forall T < t < T + \epsilon_2.$$

Thus, we conclude that

$$\mathbf{y}(T + \epsilon, \mathbf{y}(0)) \leq \mathbf{y}(T + \epsilon, \mathbf{y}_0), \text{ for all } 0 < \epsilon < \epsilon_1 \wedge \epsilon_2.$$

This contradicts the assumption that T is finite. □

Before proceeding, we consider the simple case where the initial infection at each node is the same, i.e.,

$$y_1(0) = y_2(0) =: y_0$$

and γ is symmetric, i.e.,

$$\gamma(1, 2) = \gamma(2, 1) =: \gamma_0.$$

Then, we claim,

$$y_1(t, y_0 \mathbf{1}_2; G) = y_2(t, y_0 \mathbf{1}_2; G), \forall t \geq 0.$$

Indeed, if $(z(t, y_0))$ is the (scalar) solution to

$$\frac{d}{dt} z(t) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = (\gamma_0 z(t) (1 - z(t)) - z(t)) \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad (8.1)$$

then,

$$(\mathbf{z}(t, y_0 \mathbf{1}_2)) := (z(t, y_0) \mathbf{1}_2)$$

is the vector solution to (7.3) and by uniqueness

$$(\mathbf{z}(t, y_0 \mathbf{1}_2)) = (\mathbf{y}(t, y_0 \mathbf{1}_2)).$$

Now, it is easy to check that $z(t) \rightarrow 0$ if $\gamma_0 \leq 1$, and that $z(t) \rightarrow 1 - \frac{1}{\gamma_0}$ if $\gamma_0 > 1$, regardless of $y_0 \in [0, 1]$.

As an aside note, we refer that the scalar equation (8.1) on z – that represents the dynamics of the fraction of infected nodes in a complete network in the weak limit of large number of nodes – is the same as equation (9.6) in Section 9.1 in [12] – motivated through a full-mixing reasoning over *homogeneous* networks. We note that Chapter 9 of [12] studies only the **equilibrium** of similar SIS equations. In contrast, we provide a rigorous **qualitative analysis** for the general bi-virus over e-networks case. Such analysis relies on the generalization of the tools presented in this and the next Section.

The next Theorem relies on the previous observation for symmetric two-node e-networks and on Theorem 48 to establish the first asymptotic result on general two-node e-networks. Namely, it implies that, if $\gamma(1, 2) \wedge \gamma(2, 1) > 1$, then the virus survives, and if $\gamma(1, 2) \vee \gamma(2, 1) \leq 1$ then, it dies out, irrespective to the initial distribution $\mathbf{y}_0 \in \mathcal{D}$.

Theorem 49. *Let $(\mathbf{y}(t, \mathbf{y}_0, G))$ be the flow of (7.3) with an underlying two node e-network $G = (V, \gamma) \in \mathcal{E}(2)$ and $\mathbf{y}_0 \in \mathcal{D}(2)$. Then, if $\gamma_{12} \wedge \gamma_{21} > 1$ we have*

$$\left(1 - \frac{1}{\gamma_{12} \wedge \gamma_{21}}\right) \mathbf{1} \leq \liminf_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0, G) \leq \limsup_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0, G) \leq \left(1 - \frac{1}{\gamma_{12} \vee \gamma_{21}}\right) \mathbf{1}.$$

If $\gamma_{12} \vee \gamma_{21} \leq 1$ then,

$$\mathbf{y}(t, \mathbf{y}_0, G) \rightarrow 0.$$

Proof. First, let $\mathbf{y}_0 > \mathbf{0}$ and $\gamma_{12} \wedge \gamma_{21} > 1$ and define the symmetric two-node e-network $\underline{G} := (V, \underline{\gamma})$ such that

$$\underline{\gamma}_0 := \underline{\gamma}_{12} = \underline{\gamma}_{21} = \gamma_{12} \wedge \gamma_{21}.$$

Choose $\delta > 0$ so that $\mathbf{y}_0 > \delta \mathbf{1}_2 > \mathbf{0}$ and observe that $\underline{G} \preceq G$. From Theorem 48,

$$\mathbf{y}(t, \mathbf{y}_0; G) \geq \mathbf{y}(t, \delta \mathbf{1}_2; \underline{G}), \forall t \geq 0.$$

Thus,

$$\liminf_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0; G) \geq \lim_{t \rightarrow \infty} \mathbf{y}(t, \delta \mathbf{1}_2; \underline{G}) = \left(1 - \frac{1}{\underline{\gamma}_0}\right) \mathbf{1}_2.$$

Similarly, define $\overline{G} := (V, \overline{\gamma})$ such that

$$\overline{\gamma}_0 := \overline{\gamma}_{12} = \overline{\gamma}_{21} = \gamma_{12} \vee \gamma_{21}.$$

We have

$$\mathbf{y}(t, \mathbf{y}_0, G) \leq \mathbf{y}(t, \mathbf{1}_2, \overline{G}), \forall t \geq 0.$$

Thus,

$$\limsup_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0, G) \leq \lim_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{1}_2; \overline{G}) = \left(1 - \frac{1}{\overline{\gamma}_0}\right) \mathbf{1}_2.$$

Now, assume

$$\mathbf{y}(0) \not\geq 0,$$

i.e.,

$$y_1(0) = 0 \text{ and } y_2(0) > 0$$

(up to a relabeling of the nodes). Then,

$$\dot{y}_1(0) = \gamma_{21} y_2(0) > 0.$$

Therefore, by the same argument as in the proof to Theorem 48, there exists $\epsilon > 0$ so that

$$\mathbf{y}(t, \mathbf{y}_0) > 0, \forall 0 < t < \epsilon.$$

Choose, $t_0 \in (0, \epsilon)$. Then,

$$\mathbf{y}(t, \mathbf{y}_0) = \mathbf{y}(t - t_0, \mathbf{y}(t_0)), \forall t \geq t_0.$$

Since $\mathbf{y}(t_0) > 0$,

$$\begin{aligned} \liminf_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0) &= \liminf_{t \rightarrow \infty} \mathbf{y}(t - t_0, \mathbf{y}(t_0)) \geq \left(1 - \frac{1}{\underline{\gamma}_0}\right) \mathbf{1}_2 \\ \limsup_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0) &= \limsup_{t \rightarrow \infty} \mathbf{y}(t - t_0, \mathbf{y}(t_0)) \leq \left(1 - \frac{1}{\overline{\gamma}_0}\right) \mathbf{1}_2. \end{aligned}$$

The argument repeats for $\gamma_{12} \vee \gamma_{21} \leq 1$. □

Next, we present a Corollary to Theorem 49, and we complete the analysis for the two-node single virus case. The Corollary states that, when the two node e-network is symmetric, the bounds in Theorem 49 are

tight and the flow converges to the equilibrium regardless of the initial conditions (without being trapped in a limit cycle).

Corollary 50. *Let $G = (V, \gamma)$ be a two-node e-network with symmetric $\gamma : \{1, 2\}^2 \rightarrow \mathbb{R}_+$, i.e., $\gamma_0 := \gamma(1, 2) = \gamma(2, 1)$. Then,*

$$\begin{aligned} \gamma_0 > 1 &\Rightarrow \mathbf{y}(t, \mathbf{y}_0, G) \rightarrow \left(1 - \frac{1}{\gamma_0}, 1 - \frac{1}{\gamma_0}\right) \\ \gamma_0 \leq 1 &\Rightarrow \mathbf{y}(t, \mathbf{y}_0, G) \rightarrow 0, \end{aligned}$$

for all $\mathbf{y}_0 \in \mathcal{D}$ with $\mathbf{y}_0 \neq 0$.

We would like to stress that our focus is not on computing the equilibrium points neither for the simplest single virus two-node case nor for the general bi-virus over general e-networks case. Our main goal is devising tools to study the qualitative behavior of the general dynamical system. Also, devising a Lyapunov function to establish Corollary 50 is not difficult as one may consider the error function

$$w(\mathbf{y}(t)) := \frac{1}{2} (y_1(t) - y_2(t))^2 \geq 0,$$

where

$$\frac{d}{dt} w(\mathbf{y}(t)) = -(1 + \gamma) (y_1(t) - y_2(t)) \leq 0$$

for all time $t, t \geq 0$ and any solution $(\mathbf{y}(t))$ of (7.3). In words, w is a Lyapunov function for the attractor given by the straight line

$$r = \{(y_1, y_2) \in \mathbb{R}^2 : y_1 = y_2\} \cap \mathcal{D}(2),$$

i.e., the set of configurations where nodes are evenly infected. Since the set $\mathcal{D}(2)$ is compact and the singleton

$$\left\{ \left(1 - \frac{1}{\gamma_0}, 1 - \frac{1}{\gamma_0}\right) \right\}$$

is the maximally invariant subset of the straight line r for the dynamics (7.3), Lasalle's principle (refer to [36]) now implies Corollary 50.

It is not clear how to extend the above Lyapunov based proof to the Corollary 50 to general e-networks. In fact, our approach for the more general framework will depend upon the appropriate generalization of the previous theorems. In what follows, we explore the monotonicity property of the dynamical system to analyze these more general cases, starting by extending the analysis to bi-viral infection in two-node e-networks in the next Section. We would like to emphasize that we are not simply restricting ourselves to the analysis of the equilibrium points, i.e., algebraic solutions x^* to $\mathbf{F}(x^*; G) = 0$, but rather we are studying the global stability from the initial conditions to the attractors – where orbits accumulate to. Namely, what Corollary 50 states is that regardless of the initial conditions, the orbit of a solution will converge (in time) to the equilibrium without being attracted by a limit cycle.

8.2 Bi-virus Epidemics

This Section considers bi-virus epidemics in two-node e-networks that we refer to as

$$G^{x,y} = (G^x, G^y)$$

with

$$G^x = (V, \gamma^x), G^y = (V, \gamma^y) \in \mathcal{E}(V) \text{ and } |V| = 2.$$

We determine the qualitative behavior of the coupled system of nonlinear ODEs (7.4)-(7.5). We observed in Section 8.1 that, under symmetry on the initial conditions and the e-network – namely, if $y_1(0) = y_2(0) =: y_0$ and $\gamma_{12} = \gamma_{21}$, – the qualitative behavior of the corresponding solutions $(\mathbf{y}(t, y_0 \mathbf{1}_2; G))$ were easily characterized, and any other solution associated with an arbitrary underlying two-node e-network could be appropriately lower/upper bounded by such *easy* solutions, from which we could extract valuable information about the long term behavior of any solution. In this Section, we extend this to bi-virus epidemics over two-node e-networks, namely, as we will observe, now the flow of the dynamical system preserves *skewed* bounded initial conditions as will be clearer next. We start by extending Theorem 48.

Theorem 51 (Bi-virus Monotonicity). *Let $G^x, G^y \in \mathcal{E}(2)$ be two-node e-networks. Consider $\mathbf{x}_0 \leq \mathbf{x}(0)$, $\mathbf{y}_0 \geq \mathbf{y}(0)$; and $G_S^y \succeq G^y$, $G_s^x \preceq G^x$. Define $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$, $\mathbf{z}(0) = (\mathbf{x}(0), \mathbf{y}(0))$; and $G^{x,y} = (G^x, G^y)$, $G_{s,S}^{x,y} = (G_s^x, G_S^y)$. Then,*

$$\mathbf{y}\left(t, \mathbf{z}_0; G_{s,S}^{x,y}\right) \geq \mathbf{y}\left(t, \mathbf{z}(0); G^{x,y}\right) \quad (8.2)$$

$$\mathbf{x}\left(t, \mathbf{z}_0; G_{s,S}^{x,y}\right) \leq \mathbf{x}\left(t, \mathbf{z}(0); G^{x,y}\right) \quad (8.3)$$

for all time $t, t \geq 0$.

Proof. For the sake of notational simplicity, we write $\mathbf{y}(t, \mathbf{z}_0)$ and $\mathbf{x}(t, \mathbf{z}_0)$ instead of

$$\mathbf{y}\left(t, \mathbf{z}_0, G_{s,S}^{x,y}\right) \text{ and } \mathbf{x}\left(t, \mathbf{z}_0, G_{s,S}^{x,y}\right);$$

and $\mathbf{y}(t, \mathbf{z}(0))$ and $\mathbf{x}(t, \mathbf{z}(0))$ instead of

$$\mathbf{y}(t, \mathbf{z}(0), G^{x,y}) \text{ and } \mathbf{x}(t, \mathbf{z}(0), G^{x,y}).$$

Assume that $\mathbf{z}(0) \neq \mathbf{z}_0$. Similarly to the proof of Theorem 48, define

$$T = \inf \{t \geq 0 : \mathbf{y}(t, \mathbf{z}(0)) \not\leq \mathbf{y}(t, \mathbf{z}_0) \text{ or } \mathbf{x}(t, \mathbf{z}(0)) \not\geq \mathbf{x}(t, \mathbf{z}_0)\}.$$

Assume $T < \infty$. Then, with $i \neq j$, $i, j \in \{1, 2\}$, we have one of the two configurations below:

$$y_i(T, \mathbf{z}(0)) = y_i(T, \mathbf{z}_0) \text{ and } y_j(T, \mathbf{z}(0)) < y_j(T, \mathbf{z}_0) \quad (8.4)$$

$$x_i(T, \mathbf{z}(0)) = x_i(T, \mathbf{z}_0) \text{ and } x_j(T, \mathbf{z}(0)) > x_j(T, \mathbf{z}_0). \quad (8.5)$$

Without loss of generality, choose configuration (8.4) with $i = 1$ and $j = 2$.

Case 1: If $x_1(T, \mathbf{z}(0)) + y_1(T, \mathbf{z}(0)) < 1$, then, from (7.4)-(7.5), we have

$$\dot{y}_1(T, \mathbf{z}(0)) < \dot{y}_1(T, \mathbf{z}_0).$$

Therefore (refer to Theorem (68) in Appendix),

$$\exists \epsilon_1 > 0 : y_1(t, \mathbf{z}(0)) < y_1(t, \mathbf{z}_0), \quad \forall T < t < T + \epsilon_1.$$

Also,

$$y_2(T, \mathbf{z}(0)) < y_2(T, \mathbf{z}_0) \Rightarrow \exists \epsilon_2 > 0 : y_2(t, \mathbf{z}(0)) < y_2(t, \mathbf{z}_0) \quad \forall T < t < T + \epsilon_2.$$

Thus,

$$y(t, \mathbf{z}(0)) \leq y(t, \mathbf{z}_0), \quad \forall T < t < T + \epsilon$$

with $\epsilon = \epsilon_1 \wedge \epsilon_2$. In the same vein, we can conclude that for some $\alpha > 0$:

$$x(t, \mathbf{z}(0)) \geq x(t, \mathbf{z}_0), \quad \forall T < t < T + \alpha.$$

which contradicts the finite assumption on T .

Case 2: If $x_1(T, \mathbf{z}(0)) + y_1(T, \mathbf{z}(0)) = 1$, then,

$$\dot{x}_1(T, \mathbf{z}(0)) + \dot{y}_1(T, \mathbf{z}(0)) = -(x_1(T) + y_1(T)) < 0 \Rightarrow \exists \epsilon > 0 : x_1(t, \mathbf{z}(0)) + y_1(t, \mathbf{z}(0)) < 1,$$

for all $t \in (T, T + \epsilon)$, and we are back to case 1 to reach a contradiction on the assumption that $T < \infty$.

The Theorem is proved. \square

Figure 8.1 depicts geometrically Theorem 51 as the monotonous property in Theorem 51 is equivalent to the invariance of the set given by the Cartesian product of the dark (colored) triangles in Figure 8.1.

Theorems 48 and 51 will be extended to general e-networks in Chapter 9, but the corresponding proofs require additional features, namely, that the n th derivative

$$y_i^{(n)}(t, \mathbf{z}_0, G^{x,y})$$

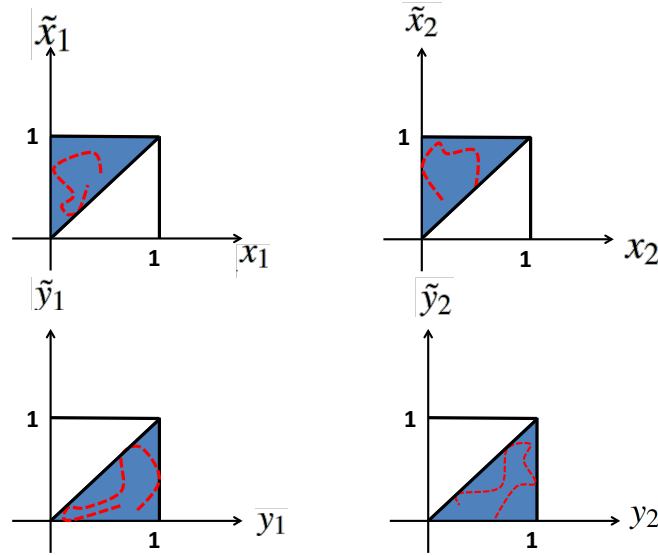


Figure 8.1: Phase space of the augmented dynamical system with $(x_i(t), y_i(t)) := (x_i, y_i)(t, \mathbf{z}_0)$, and $(\tilde{x}_i(t), \tilde{y}_i(t)) := (x_i, y_i)(t, \mathbf{z}(0))$. The red (dashed) curves represent the components of $(\tilde{x}_1(t), x_1(t), \tilde{y}_1(t), y_1(t), \tilde{x}_2(t), x_2(t), \tilde{y}_2(t), y_2(t))$. The fact that the red (dashed) curve cannot escape the blue (dark) region in finite time is equivalent to $x_i(t) \leq \tilde{x}_i(t)$ and $y_i(t) \geq \tilde{y}_i(t)$ for all time t .

is sensitive to $y_j(t, \mathbf{z}_0, G^{x,y})$ (in a sense that will be clear later) if j lies at most n hops away from i , which qualitatively interrelates the topology of the e-network with the geometry of the flow of the dynamical system.

Similarly to as pointed in Section 8.1, if symmetry on the underlying e-networks G^x and G^y and the initial conditions $x_1(0) = x_2(0)$, $y_1(0) = y_2(0)$ hold then, the dimension of (7.4)-(7.5) is reduced, as stated in the next Lemma, and the analysis is simplified.

Lemma 52. *Let $G^y \preceq G^x$ be symmetric two-node e-networks, and define $\gamma_0^x := \gamma_{12}^x = \gamma_{21}^x$ and $\gamma_0^y := \gamma_{12}^y = \gamma_{21}^y$; and $x_0 := x_1(0) = x_2(0)$, $y_0 := y_1(0) = y_2(0)$ then,*

$$x_1(t, (x_0, y_0) \otimes \mathbf{1}_2, G^{x,y}) = x_2(t, (x_0, y_0) \otimes \mathbf{1}_2, G^{x,y}) \quad (8.6)$$

$$y_1(t, (x_0, y_0) \otimes \mathbf{1}_2, G^{x,y}) = y_2(t, (x_0, y_0) \otimes \mathbf{1}_2, G^{x,y}), \quad (8.7)$$

for all $t, t \geq 0$. Moreover, if $\gamma_0^x > \gamma_0^y$ then,

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbf{x}(t) &= \left(1 - \frac{1}{\gamma_0^x}\right) \mathbf{1}_2 \\ \lim_{t \rightarrow \infty} \mathbf{y}(t) &= \mathbf{0} \end{aligned}$$

regardless of the symmetric initial conditions x_0 and y_0 .

Proof. Indeed, in this case the flow

$$((\mathbf{x}, \mathbf{y})(t, (x_0, y_0) \otimes \mathbf{1}_2; G^{x,y}))$$

is solution to the following reduced 2-dimensional ODE

$$\dot{x}(t)\mathbf{1}_2 = (\gamma_0^x x(t)(1 - x(t) - y(t)) - x(t))\mathbf{1}_2 \quad (8.8)$$

$$\dot{y}(t)\mathbf{1}_2 = (\gamma_0^y y(t)(1 - x(t) - y(t)) - y(t))\mathbf{1}_2. \quad (8.9)$$

These equations (8.8)-(8.9) represent the fluid-limit dynamics of a bi-viral epidemics over complete networks and were studied in [17], from where the asymptotics follows. \square

The next Theorem provides a sufficient condition on the two-node e-networks G^x and G^y to assert a *survival of the fittest* phenomenon – the weaker virus strain dies out while the dominant one persists as long as it is strong enough. The Theorem follows from the previous symmetry observation and Theorem 51.

Theorem 53 (Natural Selection). *Let $G^x \succeq G^y$ be two-node e-networks. If $\gamma_{12}^x \wedge \gamma_{21}^x > \gamma_{12}^y \vee \gamma_{21}^y$ then,*

$$\mathbf{y}(t, \mathbf{z}_0; G^{x,y}) \longrightarrow 0.$$

Additionally, if $\gamma_{12}^x \wedge \gamma_{21}^x > 1$ then,

$$\left(1 - \frac{1}{\gamma_{12}^x \wedge \gamma_{21}^x}\right) \mathbf{1} \leq \liminf_{t \rightarrow \infty} \mathbf{x}(t, \mathbf{z}_0, G^{x,y}) \leq \limsup_{t \rightarrow \infty} \mathbf{x}(t, \mathbf{z}_0, G^{x,y}) \leq \left(1 - \frac{1}{\gamma_{12}^x \vee \gamma_{21}^x}\right) \mathbf{1}.$$

If $\gamma_{12}^x \vee \gamma_{21}^x \leq 1$ then,

$$\mathbf{x}(t, \mathbf{z}_0; G^{x,y}) \longrightarrow 0.$$

Proof. Define the symmetric e-networks,

$$\underline{G}^x := (V, \underline{\gamma}^x) \text{ and } \overline{G}^y := (V, \overline{\gamma}^y)$$

with

$$\underline{\gamma}^x := \gamma_{12}^x \wedge \gamma_{21}^x \text{ and } \overline{\gamma}^y := \gamma_{12}^y \vee \gamma_{21}^y.$$

Note from the assumptions that $\underline{G}^x \succeq \overline{G}^y$ – or, in other words, G^x strongly dominates G^y . Define

$$\overline{G}^{x,y} = (\underline{G}^x, \overline{G}^y).$$

Let $\mathbf{x}_0 > \epsilon \mathbf{1}_N > 0$ for some $\epsilon > 0$ then,

$$\liminf_{t \rightarrow \infty} \mathbf{x}(t, \mathbf{z}_0; G^{x,y}) \geq \lim_{t \rightarrow \infty} \mathbf{x}(t, (\epsilon \mathbf{1}_N, \mathbf{1}_N); \overline{G}^{x,y}) \longrightarrow \left(1 - \frac{1}{\underline{\gamma}^x}\right)$$

$$\lim_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{z}_0; G^{x,y}) \leq \lim_{t \rightarrow \infty} \mathbf{y}(t, (\epsilon \mathbf{1}_N, \mathbf{1}_N); \overline{G}^{x,y}) \rightarrow 0,$$

where the asymptotics in the right hand side follows from Lemma 52, due to the symmetry of $\overline{G}^{x,y}$ and the initial condition $(\epsilon \mathbf{1}_N, \mathbf{1}_N)$. The case when $\mathbf{x}_0 \not\geq 0$ is treated similarly to as done in the proof to Theorem 48, in that there exists $\delta > 0$ so that

$$\mathbf{x}(t, \mathbf{z}_0; G^{x,y}) > \epsilon > 0$$

for all $t \in (0, \delta)$ for some ϵ . Otherwise, if $\mathbf{x}_0 = 0$ or $\mathbf{y}_0 = 0$ (no virus of type x or y in the system) then, clearly

$$(\mathbf{x}(t, \mathbf{z}_0; G^{x,y})) = 0 \text{ or } (\mathbf{y}(t, \mathbf{z}_0; G^{x,y})) = 0.$$

□

Figure 8.2 illustrates the possibility of bounding any configuration by simpler symmetric well-characterized configurations. Such bounds are preserved for all $t, t \geq 0$ as established in Theorem 51.

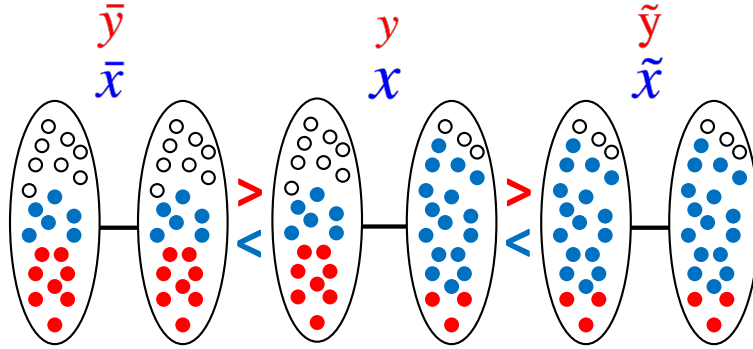


Figure 8.2: Population of blue (lighter color) in the center bipartite network is lower and upper bounded by the corresponding populations in the left and right bipartite networks, respectively. The same goes, in the other way around, for the red (darker) population. The symmetric configurations in the left and right bipartite networks induce well-known solutions that bound the qualitative behavior of the middle configuration for all time $t \geq 0$.

Theorem 53 states that if G^x strongly dominates G^y , i.e., $G^x \succeq^* G^y$, then strain y dies out in the long run while strain x persists if strong enough.

Corollary 54. *Let $G^x \succ^* G^y$ be two-node e-networks with G^x symmetric and G^y arbitrary. Let $\mathbf{x}_0 \neq 0$. Then,*

$$\mathbf{y}(t, \mathbf{z}_0; G^{x,y}) \rightarrow 0.$$

Additionally, if $\gamma^x > 1$ then,

$$\mathbf{x}(t, \mathbf{z}_0; G^{x,y}) \rightarrow 1 - \frac{1}{\gamma^x},$$

otherwise, if $\gamma_0^x \leq 1$,

$$\mathbf{x}(t, \mathbf{z}_0, G^{x,y}) \longrightarrow 0.$$

In words, Theorem 53 states that, if $\gamma_{12}^x \wedge \gamma_{21}^x > 1$, the set

$$\mathcal{A} = \left[1 - \frac{1}{\gamma_{12}^x \wedge \gamma_{21}^x}, 1 - \frac{1}{\gamma_{12}^x \vee \gamma_{21}^x} \right]^2 \times \{0\}^2$$

is an attractor of the dynamical system (7.4)-(7.5), with basin of attraction given by

$$\mathcal{B} = \{(\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{D}_2(V) : \mathbf{x}_0 \neq 0\},$$

otherwise, if $\gamma_{12}^x \vee \gamma_{21}^x \leq 1$ then, the origin $\{0\}^4$ is the global attractor of (7.4)-(7.5). When the two-node e-network G^x is symmetric, then the attractor \mathcal{A} collapses to be a singleton, namely, the non-trivial equilibrium

$$\left\{ 1 - \frac{1}{\gamma_0} \right\}^2 \times \{0\}^2$$

as stated in Corollary 54. In the next Chapter, we study the qualitative dynamics of (7.3) and (7.4)-(7.5) over general e-networks.

8.3 Concluding Remarks

The goal of this Chapter is to introduce our techniques of analysis of the dynamical system (7.4)-(7.5) in the simple framework of two-node e-networks. Namely, we observed that an upper/lower bound on the initial conditions and on the parameters of the dynamical system is preserved by the flow $(\mathbf{y}(t, \mathbf{y}_0))$ of the ODE's (7.4)-(7.5). Even though some of the results in this Chapter can be obtained via a Lyapunov approach, there seems to be no natural Lyapunov extension to the more general case of bi-virus over general e-networks. In Section 8.1, we determined the qualitative behavior of the single virus dynamics (7.3). In Section 8.2, we studied the bi-virus dynamics (7.4)-(7.5) over two node e-networks. By resorting to the monotonous structure of (7.4)-(7.5), we determined that if a virus is strongly-dominated, then it necessarily dies out while the stronger strain survives as long as it is strong enough. In the next Chapter, we extend the results for bi-virus epidemics over general e-networks.

Chapter 9

Qualitative Analysis: General E-networks

This Chapter considers the dynamics of epidemics in general e-networks. In Section 9.1, we determine the qualitative behavior of the coupled system of nonlinear ODE's (7.3) (refer to Chapter 7) associated with the single virus epidemics. In Section 9.2, we study the dynamics (7.4)-(7.5) associated with the bi-virus case. Our approach consists in exploring the monotonicity of the system to appropriately upper/lower-bound any solution by well-characterized *symmetric solutions* and then determine the attractors of the system. Section 9.3 concludes the Chapter.

9.1 Single Virus Epidemics

The goal of this Section is to extend the results and techniques on the qualitative behavior of the single-virus dynamical flow $(\mathbf{y}(t, \mathbf{y}_0, G))$ over two-node e-networks treated in Section 8.1 to the framework of arbitrary e-networks $G \in \mathcal{E}(V)$. Again, our main focus is on the qualitative dynamics of the vector flow $(\mathbf{y}(t, \mathbf{y}_0; G))$ rather than on determining its equilibrium points. We remark that some of the results in this Section follow from the Lyapunov-based qualitative analysis on (7.3) provided in reference [14] but do not extend to the bi-virus dynamics. Therefore, the goal of this Section is to provide the main techniques of analysis to be extended in the next Section to bi-virus epidemics – where one of our main contributions lies – and that do not rely on Lyapunov analysis.

The next two Theorems reveal the qualitative coupling between the topology of the e-network and the regularity of the solutions, namely, they state that the degree of infection $y_j(t, \mathbf{y}_0, G)$ at node j has an impact on the degree of infection $y_i(t, \mathbf{y}_0, G)$ at node i , n -hops away from j , via its n th-order (or higher than n) derivatives. They are crucial to establishing the main Theorems 57 and 59 of this Section. For notational simplicity, in the next two Theorems, we write $y_i(t)$ instead of $y_i(t, \mathbf{y}_0, G)$. The proofs to Theorems 55-56 are included in the Appendix (refer to Theorems 69 and 70 in Appendix, respectively).

Theorem 55. *Let $y_i(t) = 0$ and $y_j(t) = 0, \forall j \in \cup_{l=1}^n \mathcal{N}^l(i)$ for some time $t \geq 0$. Then,*

$$y_i^{(\ell)}(t) = 0, \quad \forall 0 < \ell \leq n, \quad (9.1)$$

that is, if there is no infection within a neighborhood up to order n of node i , then all derivatives of $(y_i(t))$ up to order n are zero. ■

The next Theorem states that higher order moments are sensitive to further away infected nodes – node i located n -hops away from node j , affects only the n th-order derivative of $(y_j(t))$.

Theorem 56. *Let $y_i(t) > 0$ and $y_j(t) = 0, \forall j \neq i$ for some time $t \geq 0$. Then, $\forall \ell < n$*

$$j \in \mathcal{N}^n(i) \Rightarrow y_j^{(n)}(t) > 0 \text{ and } y_j^{(\ell)}(t) = 0. \quad (9.2)$$

■

Theorems 55 and 56 reveal the impact of the topology of $\text{supp}(G)$ on the inter-dependence among the geometric aspects (e.g., derivative, curvature) of the infection across nodes, namely, a perturbation on the infection $y_i(t)$ at node i at time t will perturb: *i*) its immediate neighbors $j \in \mathcal{N}(i)$ by perturbing their first derivatives $\dot{y}_j(t)$; *ii*) its n -th order neighbors $j \in \mathcal{N}^n(i)$ by perturbing their n th order curvature $y_j^{(n)}(t)$.

Next, we extend Theorem 48 to general e-networks, namely, the flow $(\mathbf{y}(t, \mathbf{y}_0, G))$ to (7.3) preserves upper/lower bounds on the initial conditions and e-networks.

Theorem 57. *Let $G_2 = (V, \gamma^{(2)}) \preceq G_1 = (V, \gamma^{(1)})$ be two e-networks and $\mathbf{y}(0) \leq \mathbf{y}_0$. Then,*

$$\mathbf{y}(t, \mathbf{y}(0), G_2) \leq \mathbf{y}(t, \mathbf{y}_0, G_1), \quad \forall t \geq 0.$$

Proof. We show that the set

$$B = \{(\mathbf{y}, \tilde{\mathbf{y}}) \in [0, 1]^{2N} : \mathbf{y} \geq \tilde{\mathbf{y}}\}$$

is invariant with respect to the augmented single-virus dynamics

$$\frac{d}{dt}(\mathbf{y}(t), \tilde{\mathbf{y}}(t)) = (\mathbf{F}(\mathbf{y}(t), G_1), \mathbf{F}(\tilde{\mathbf{y}}(t), G_2)), \quad (9.3)$$

where the vector field $\mathbf{F}(\cdot, G)$ is defined in (7.3). Let

$$(\mathbf{y}(t, \mathbf{y}_0, G_1), \mathbf{y}(t, \mathbf{y}(0), G_2))$$

be solution to (9.3). Then, it is enough to investigate the decoupled augmented vector field

$$\bar{\mathbf{F}}_{G_1, G_2}(\mathbf{y}, \tilde{\mathbf{y}}) = (\mathbf{F}(\mathbf{y}, G_1), \mathbf{F}(\tilde{\mathbf{y}}, G_2))$$

over the boundary of B to establish that, once started in B , the solution

$$(\mathbf{y}(t, \mathbf{y}_0, G_1), \mathbf{y}(t, \mathbf{y}(0), G_2))$$

never escapes the set B , i.e.,

$$\mathbf{y}(t, \mathbf{y}_0, G_1) \geq \mathbf{y}(t, \mathbf{y}(0), G_2)$$

for all $t, t \geq 0$, if $\mathbf{y}_0 \geq \mathbf{y}(0)$. The set B is depicted in Figure 9.1 as the Cartesian product of triangles, and one has to assure that no solution components can leave the triangular regions. For notational simplicity, define

$$\mathbf{y}(t) := \mathbf{y}(t, \mathbf{y}_0, G_1) \text{ and } \tilde{\mathbf{y}}(t) := \mathbf{y}(t, \mathbf{y}(0), G_2).$$

Let $t > 0$ be such that:

Case 1: $y_i(t) = 1, 0 < \tilde{y}_i(t) < 1$ for some i : then,

$$F_i(\mathbf{y}(t), G_1) = \frac{d}{dt} y_i(t) = -y_i(t) = -1 < 0$$

and the vector field points inwards B .

Case 2: $\tilde{y}_i(t) = 0, 0 < y_i(t) < 1$ for some i : then,

$$F_i(\tilde{\mathbf{y}}(t)) = \frac{d}{dt} \tilde{y}_i(t) = \sum_j \gamma_{ji}^{(1)} y_j(t) > 0$$

Case 3: $0 \leq y_i(t) = \tilde{y}_i(t) \leq 1$: then,

$$\begin{aligned} F_i(\mathbf{y}(t)) &= \frac{d}{dt} y_i(t) = \sum_j \gamma_{ji}^{(1)} y_j(t) (1 - y_i(t)) - y_i(t) \\ F_i(\tilde{\mathbf{y}}(t)) &= \frac{d}{dt} \tilde{y}_i(t) \\ &= \sum_j \gamma_{ji}^{(2)} \tilde{y}_j(t) (1 - \tilde{y}_i(t)) - \tilde{y}_i(t) \\ &= \sum_j \gamma_{ji}^{(2)} \tilde{y}_j(t) (1 - y_i(t)) - y_i(t). \end{aligned} \tag{9.4}$$

If

$$y_j(t) \geq \tilde{y}_j(t), \forall j \in \mathcal{N}(i),$$

with strict inequality for at least some $j \in \mathcal{N}(i)$, then

$$F_i(\mathbf{y}(t)) = \dot{y}_i(t) > \dot{\tilde{y}}_i(t) = F_i(\tilde{\mathbf{y}}(t))$$

and, therefore, from the analyticity of the vector field \mathbf{F} (thus, the analyticity of the solutions), we have

$y_j(t) > \tilde{y}_j(t)$ for all $t \in (0, \epsilon)$ for some $\epsilon > 0$ small enough (refer to Theorem 68 in the Appendix). Therefore, in this case the solution $(\mathbf{y}(t), \tilde{\mathbf{y}}(t))$ cannot escape the set B . More generally, if

$$y_j(t) = \tilde{y}_j(t), \forall j \in \cup_{l=0}^{n-1} \mathcal{N}^l(i)$$

for some $n \geq 2$ with strict inequality $y_j(t) > \tilde{y}_j(t)$ for some $j \in \mathcal{N}^n(i)$ then, from Theorem 56 (and here is exactly where we need this Theorem) it follows that

$${}^{(n)}y_j(t) > {}^{(n)}\tilde{y}_j(t)$$

and, thus, Theorem 68 yields $y_j(t) > \tilde{y}_j(t)$ for all $t \in (0, \epsilon)$ for some $\epsilon > 0$ small enough and again, the solution $(\mathbf{y}(t), \tilde{\mathbf{y}}(t))$ cannot escape B . Otherwise, if $y_j(t) = \tilde{y}_j(t), \forall j$, then, both $(\mathbf{y}(t))$ and $(\tilde{\mathbf{y}}(t))$ obey the same differential equation with a Lipschitz continuous vector field over the compact domain B . The solution is thus unique and $y_j(t) = \tilde{y}_j(t), \forall t \geq 0$. Figure 9.1 depicts the main idea of the proof. \square

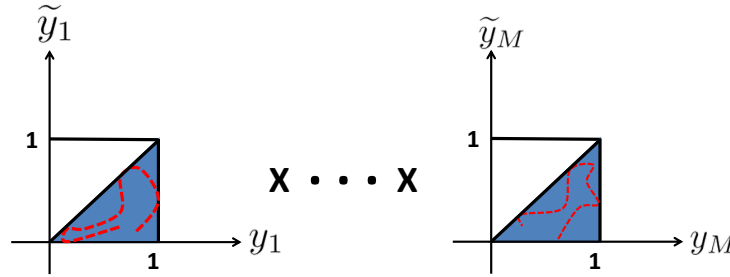


Figure 9.1: Illustration of an orbit of the augmented system. The set B is invariant, which implies that if $\mathbf{y}(0) \geq \tilde{\mathbf{y}}(0)$ then $\mathbf{y}(t, \mathbf{y}(0)) \geq \mathbf{y}(t, \tilde{\mathbf{y}}(0)), \forall t \geq 0$. Whenever a component hits one of the 3 sides of a triangle, it returns inwards the corresponding triangle.

The next Lemma is an extension of the observation made in Chapter 8 that the dimension of the dynamics (7.3) is reduced under symmetry on the underlying e-networks and initial conditions.

Lemma 58. *Let $G = (V, \gamma)$ be regular, i.e., $\text{supp}(G)$ is a regular graph and $\gamma \equiv \gamma_0$ is constant across the edges, then,*

$$y_1(t, y_0 \mathbf{1}; G) = \dots = y_N(t, y_0 \mathbf{1}; G),$$

for all $t \geq 0$. Moreover, if $\gamma_0 > 1$ and $y(0) \neq 0$

$$\mathbf{y}(t) \longrightarrow \left(1 - \frac{1}{d\gamma}\right) \mathbf{1},$$

otherwise,

$$\mathbf{y}(t) \longrightarrow \mathbf{0}.$$

Proof. We have that $(\mathbf{y}(t, y_0 \mathbf{1}; G))$ is solution to

$$\dot{z}(t) \mathbf{1}_N = [d\gamma_0 z(t)(1 - z(t)) - z(t)] \mathbf{1}_N \quad (9.5)$$

where $\mathbf{1}_N \in \mathbb{R}^M$ is the vector of ones and d is the degree of the regular e-network G . Now, note that equation (9.5) captures the dynamics of the complete network as discussed before (refer to equation (8.1) in Chapter 8) and thus, the asymptotics follows. \square

The next Theorem characterizes the global attractor of the dynamical system (7.3).

Theorem 59. *Let G be an e-network. Let $\underline{G} = (V, \underline{\gamma})$ and $\overline{G} = (V, \overline{\gamma})$ be the inner-regular and outer-regular e-networks of G . Then, if $\underline{\gamma}d(\underline{G}) > 1$*

$$1 - \frac{1}{\underline{\gamma}d(\underline{G})} \leq \liminf_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0, G) \leq \limsup_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0; G) \leq 1 - \frac{1}{\overline{\gamma}d(\overline{G})}. \quad (9.6)$$

If $\overline{\gamma}d(\overline{G}) \leq 1$ then,

$$\lim_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{y}_0, G) \longrightarrow \mathbf{0}.$$

Proof. First, assume that $\mathbf{y}(0) > \mathbf{0}$ and choose $\epsilon > 0$ such that $\mathbf{y}(0) > \epsilon \mathbf{1} > \mathbf{0}$. Therefore,

$$\begin{aligned} \left(1 - \frac{1}{\overline{\gamma}d(\overline{G})}\right) \mathbf{1} &= \lim_{t \rightarrow +\infty} \mathbf{y}(t, \mathbf{1}, \overline{G}) \\ &\geq \limsup_{t \rightarrow +\infty} \mathbf{y}(t, \mathbf{y}(0), G) \\ &\geq \liminf_{t \rightarrow +\infty} \mathbf{y}(t, \mathbf{y}(0), G) \\ &\geq \lim_{t \rightarrow +\infty} \mathbf{y}(t, \epsilon \mathbf{1}; \underline{G}) \\ &= \left(1 - \frac{1}{\underline{\gamma}d(\underline{G})}\right) \mathbf{1}. \end{aligned}$$

where, the first and third inequalities follow from Theorem 57; and the equalities follow from Lemma 58. It is left to prove the case where $\mathbf{y}(0) \not> \mathbf{0}$, i.e., $y_j(0) = 0$ for some j . We assume the minimal nonzero configuration, where $y_i(0) > 0$ and $y_j(0) = 0, \forall j \neq i$. Let $j \in \mathcal{N}^k(i)$. Then, from Theorem 56,

$$y_j^{(k)}(0) > 0, \quad y_j^{(m)}(0) = 0, \quad \forall m < k.$$

Theorem 68 (refer to the Appendix) yields $y_j(t) > 0$ for all $t \in (0, \epsilon)$ for some $\epsilon > 0$ small enough. Now, $y_j(T^*) > 0, \forall j$ for some $T^* \in (0, \epsilon)$. Then, from T^* on, we recover the case of strictly positive initial conditions – i.e.,

$$\mathbf{y}(t, \mathbf{y}(0)) = \mathbf{y}(t - T^*, \mathbf{y}(T^*)) =: \tilde{\mathbf{y}}(t, \tilde{\mathbf{y}}(0)),$$

for all $t \geq T^*$, where $\tilde{\mathbf{y}}(0) := \mathbf{y}(T^*)$. \square

Corollary 60. *Let $G = (V, \gamma) \in \mathcal{E}(V)$ be a d -regular e-network with $\gamma \equiv \gamma_0$. If $\gamma_0 d(G) > 1$ and $\mathbf{y}(0) \neq 0$ then,*

$$\mathbf{y}(t, \mathbf{y}(0), G) \longrightarrow \left(1 - \frac{1}{d(G) \gamma_0}\right) \mathbf{1}$$

otherwise, $\mathbf{y}(t, \mathbf{y}(0), G) \longrightarrow \mathbf{0}$. ■

Theorem 59 states that if $\underline{\gamma} d(\underline{G}) > 1$ then, the set

$$\mathcal{A} = \left[1 - \frac{1}{\underline{\gamma} d(\underline{G})}, 1 - \frac{1}{\bar{\gamma} d(\bar{G})}\right]^N \times \{0\}^N$$

is an attractor to (7.3) with basin of attraction

$$\mathcal{B} = \{\mathbf{y}_0 \in \mathcal{D}(V) : \mathbf{y}_0 \neq 0\}.$$

Corollary 60 states that when symmetry on the underlying e-network (but not necessarily on the initial conditions) is assumed, namely, G is regular then, the attractor \mathcal{A} reduces to the singleton

$$\left\{1 - \frac{1}{d\gamma_0}\right\}^N \times \{0\}^N.$$

9.2 Bi-virus Epidemics

This is the main Section of Part II of this Thesis. We extend the results of the previous Sections to study the qualitative behavior of (7.4)-(7.5) for the bi-virus over general e-networks dynamics and that is one of our main contributions.

For notational simplicity, in the next two Theorems 61 and 62 we define the flows

$$(\mathbf{z}(t)) = (\mathbf{x}(t), \mathbf{y}(t)) := ((\mathbf{x}, \mathbf{y})(t, \mathbf{z}(0); G^{x,y}))$$

and

$$(\tilde{\mathbf{z}}(t)) = (\tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t)) := ((\tilde{\mathbf{x}}, \tilde{\mathbf{y}})(t, \tilde{\mathbf{z}}(0); G^{x,y}))$$

under the same e-networks $G^{x,y}$ but, with possibly different initial conditions

$$\mathbf{z}(0) = (\mathbf{x}(0), \mathbf{y}(0)) \text{ and } \tilde{\mathbf{z}}(0) = (\tilde{\mathbf{x}}(0), \tilde{\mathbf{y}}(0)).$$

The next Theorem extends Theorem 55 to the bi-virus case. It states that if the initial conditions $\tilde{\mathbf{z}}(0)$ and $\mathbf{z}(0)$ are the same in a neighborhood $\mathcal{N}^k(i)$ about i up to order $k = n$ then, the derivatives up to order n of the associated flows $(\mathbf{z}(t))$ and $(\tilde{\mathbf{z}}(t))$ will coincide at time $t = 0$. The proofs to Theorems 61 and 62 are in the Appendix (refer to Theorems 71 and 72 in the Appendix).

Theorem 61. Let $y_i(0) = \tilde{y}_i(0)$ and $x_i(0) = \tilde{x}_i(0)$. Let

$$\overline{\mathcal{N}}^n(i) := \bigcup_{\ell=1}^n \mathcal{N}^\ell(i).$$

Then:

$$\begin{cases} y_k(0) = \tilde{y}_k(0) \\ x_k(0) = \tilde{x}_k(0) \end{cases} \quad \forall k \in \overline{\mathcal{N}}^n(i) \implies y_i^{(\ell)}(0) = \tilde{y}_i^{(\ell)}(0) \quad \text{and} \quad x_i^{(\ell)}(0) = \tilde{x}_i^{(\ell)}(0), \quad \forall \ell \leq n. \quad \blacksquare$$

The next Theorem extends Theorem 56 to the bi-virus case. It states that, if

$$\mathbf{z}(0) = (\mathbf{x}(0), \mathbf{y}(0)) \quad \text{and} \quad \tilde{\mathbf{z}}(0) = (\tilde{\mathbf{x}}(0), \tilde{\mathbf{y}}(0))$$

have the same degree of infection at each node about i , except at some node j , n -hops away from i , then there will be a mismatch between the n th-order derivatives $z_i^{(n)}(0)$ and $\tilde{z}_i^{(n)}(0)$ in the e-networks $G^{x,y}$.

Theorem 62. Let $y_i(0) = \tilde{y}_i(0)$, $x_i(0) = \tilde{x}_i(0)$ and

$$\begin{cases} y_k(0) = \tilde{y}_k(0) \\ x_k(0) = \tilde{x}_k(0) \end{cases} \quad \forall k \in \overline{\mathcal{N}}^{n-1}(i) \quad \text{and} \quad \begin{cases} y_k(0) \geq \tilde{y}_k(0) \\ x_k(0) \leq \tilde{x}_k(0) \end{cases} \quad \forall k \in \mathcal{N}^n(i) \setminus \{m\} \quad (9.7)$$

with strict inequality $y_m(0) > \tilde{y}_m(0)$ for some $m \in \mathcal{N}^n(i)$. Then, $y_i^{(n)}(0) > \tilde{y}_i^{(n)}(0)$. ■

The next Theorem is an extension of the monotonous property for a single virus spread established in Theorem 57 to the bi-viral epidemics case: *skewed* bounds on the initial conditions and e-networks are preserved by the flow of the dynamical system (7.4)-(7.5).

Theorem 63 (Bi-virus Monotonicity). Let $G^{x,y} = (G^x, G^y)$ and $G_{s,S}^{x,y} = (G_s^x, G_S^y)$ be two pairs of e-networks. Consider $\mathbf{x}_0 \leq \mathbf{x}(0)$ and $\mathbf{y}_0 \geq \mathbf{y}(0)$; $G_s^x \preceq G^x$ and $G_S^y \succeq G^y$. Define $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$, $\mathbf{z}(0) = (\mathbf{x}(0), \mathbf{y}(0))$. Then,

$$\mathbf{y} \left(t, \mathbf{z}_0, G_{s,S}^{x,y} \right) \geq \mathbf{y} \left(t, \mathbf{z}(0), G^{x,y} \right) \quad (9.8)$$

$$\mathbf{x} \left(t, \mathbf{z}_0, G_{s,S}^{x,y} \right) \leq \mathbf{x} \left(t, \mathbf{z}(0), G^{x,y} \right) \quad (9.9)$$

for all time $t, t \geq 0$.

Proof. Assume that

$$\mathbf{y}(0) \neq \mathbf{y}_0 \quad \text{or} \quad \mathbf{x}(0) \neq \mathbf{x}_0.$$

We write $\mathbf{y}(t, \mathbf{z}_0)$, $\mathbf{x}(t, \mathbf{z}_0)$ instead of

$$\mathbf{y}(t, \mathbf{z}_0, G_{s,S}^{x,y}) \text{ and } \mathbf{x}(t, \mathbf{z}_0, G_{s,S}^{x,y});$$

or $\mathbf{y}(t, \mathbf{z}(0))$, $\mathbf{x}(t, \mathbf{z}(0))$ instead of

$$\mathbf{y}(t, \mathbf{z}(0), G^{x,y}) \text{ and } \mathbf{x}(t, \mathbf{z}(0), G^{x,y}).$$

Define

$$T = \inf \{t : t \geq 0, \mathbf{y}(t, \mathbf{z}(0)) \not\leq \mathbf{y}(t, \mathbf{z}_0) \text{ or } \mathbf{x}(t, \mathbf{z}(0)) \not\geq \mathbf{x}(t, \mathbf{z}_0)\}.$$

Assume that $T < \infty$. Then, for some

$$i, j \in \{1, \dots, M\} \text{ with } i \neq j,$$

we have one of the following:

$$y_i(T, \mathbf{z}(0)) = y_i(T, \mathbf{z}_0) \text{ and } y_j(T, \mathbf{z}(0)) < y_j(T, \mathbf{z}_0) \quad (9.10)$$

$$x_i(T, \mathbf{z}(0)) = x_i(T, \mathbf{z}_0) \text{ and } x_j(T, \mathbf{z}(0)) > x_j(T, \mathbf{z}_0). \quad (9.11)$$

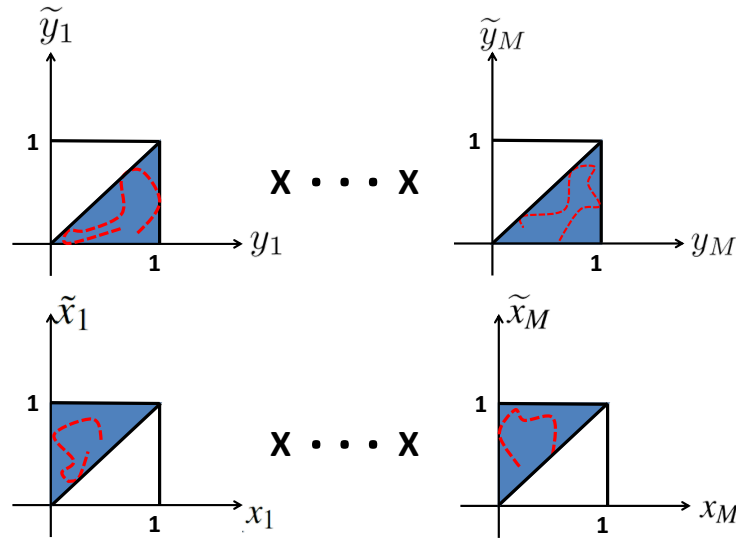


Figure 9.2: Phase space of the augmented dynamical system with $(x_i(t), y_i(t)) := (x_i, y_i)(t, \mathbf{z}_0)$, and $(\tilde{x}_i(t), \tilde{y}_i(t)) := (x_i, y_i)(t, \mathbf{z}(0))$ for $i = 1, \dots, M$. The red (dashed) curves represent the components of $(\mathbf{x}(t), \mathbf{y}(t), \tilde{\mathbf{x}}(t), \tilde{\mathbf{y}}(t))$. The fact that the red (dashed) curve cannot escape the blue (dark) region in finite time is equivalent to $x_i(t) \leq \tilde{x}_i(t)$ and $y_i(t) \geq \tilde{y}_i(t)$ for all time t and $i = 1, \dots, M$.

Without loss of generality, choose configuration (9.10) and assume $j \in \mathcal{N}^n(i)$ is the closest node to i where we have strict inequality $y_j(T, \mathbf{z}(0)) < y_j(T, \mathbf{z}_0)$.

Case 1: If $x_1(T, \mathbf{z}(0)) + y_1(T, \mathbf{z}(0)) < 1$, then, from Theorem 62 we have

$$y_i^{(n)}(T, \mathbf{z}(0)) < \tilde{y}_i^{(n)}(T, \mathbf{z}_0).$$

Therefore, from Theorem 68 in the Appendix, we have that

$$\exists \epsilon_1 > 0 : y_1(t, \mathbf{z}(0)) < y_1(t, \mathbf{z}_0), \quad \forall T < t < T + \epsilon_1.$$

Also,

$$y_j(T, \mathbf{z}(0)) < y_j(T, \mathbf{z}_0) \Rightarrow \exists \epsilon_2 > 0 : y_j(t, \mathbf{z}(0)) < y_j(t, \mathbf{z}_0), \quad \forall T < t < T + \epsilon_2.$$

Thus,

$$\mathbf{y}(t, \mathbf{z}(0)) \leq \mathbf{y}(t, \mathbf{z}_0), \quad \forall T < t < T + \epsilon$$

with $\epsilon = \epsilon_1 \wedge \epsilon_2$. Similarly, we have that

$$\mathbf{x}(t, \mathbf{z}(0)) \geq \mathbf{x}(t, \mathbf{z}_0), \quad \forall T < t < T + \alpha$$

for some $\alpha > 0$.

Case 2: If $x_1(T, \mathbf{z}(0)) + y_1(T, \mathbf{z}(0)) = 1$, then,

$$\dot{x}_1(T, \mathbf{z}(0)) + \dot{y}_1(T, \mathbf{z}(0)) = -(x_1(T) + y_1(T)) < 0 \Rightarrow \exists \epsilon > 0 : x_1(t, \mathbf{z}(0)) + y_1(t, \mathbf{z}(0)) < 1,$$

for all $t \in (T, T + \epsilon)$. In any case, we reach a contradiction on the definition of T , and the Theorem is proved. \square

The next Lemma observes that if all nodes are evenly infected with viruses x and y in regular e-networks G^x and G^y then, they will remain equally infected for all time t .

Lemma 64. *Let $G^x, G^y \in \mathcal{R}(V)$ be two regular e-networks. Let $x_0, y_0 \in \mathbb{R}$. Then,*

$$\begin{aligned} \mathbf{x}(t, (x_0, y_0) \otimes \mathbf{1}_N; G^{x,y}) &= x(t, (x_0, y_0); G^{x,y}) \mathbf{1}_N \\ \mathbf{y}(t, (x_0, y_0) \otimes \mathbf{1}_N; G^{x,y}) &= y(t, (x_0, y_0); G^{x,y}) \mathbf{1}_N \end{aligned}$$

for all $t, t \geq 0$. In this case, if $\gamma_0^x > \gamma_0^y$ with $\gamma_0^x > \frac{1}{d}$, we have

$$\mathbf{x}(t) \longrightarrow \left(1 - \frac{1}{\gamma_0^x d(G^x)}\right) \mathbf{1}_N \text{ and } \mathbf{y}(t) \longrightarrow \mathbf{0}_N$$

otherwise, if $\gamma^x \leq \frac{1}{d}$, $\mathbf{x}(t) \longrightarrow \mathbf{0}_N$ and $\mathbf{y}(t) \longrightarrow \mathbf{0}_N$.

Proof. We observe that, when the nodes are evenly infected, i.e., $\mathbf{y}(0) = y_0 \mathbf{1}$ and $\mathbf{x}(0) = x_0 \mathbf{1}$; and when the underlying e-networks G^x and G^y are regular then, the dynamics of the flow

$$((\mathbf{x}, \mathbf{y})(t, (x_0, y_0) \mathbf{1}; G^{x,y}))$$

reduce to the 2D ODE

$$\begin{aligned}\dot{y}(t) \mathbf{1}_N &= (\gamma_0^y d(G^y) y(t) (1 - x(t) - y(t)) - y(t)) \mathbf{1}_N \\ \dot{x}(t) \mathbf{1}_N &= (\gamma_0^x d(G^x) x(t) (1 - x(t) - y(t)) - x(t)) \mathbf{1}_N.\end{aligned}$$

The equations also describe the dynamics of diffusion of two virus in a complete network explored in Reference [17], and the asymptotics follows. \square

Theorem 65 (Natural Selection). *Let $G^y \preceq^* G^x$, i.e., $\overline{G}^y \preceq \underline{G}^x$, where \overline{G}^y (respectively, \underline{G}^x) is the outer (respectively, inner) regular e-network of G^y (respectively, G^x). Then,*

$$\mathbf{y}(t, \mathbf{z}_0; G^{x,y}) \rightarrow 0.$$

Additionally, if $\underline{\gamma}^x d(\underline{G}^x) > 1$ then,

$$1 - \frac{1}{\underline{\gamma}^x d(\underline{G}^x)} \leq \liminf_{t \rightarrow \infty} \mathbf{x}(t, \mathbf{z}_0; G^{x,y}) \leq \limsup_{t \rightarrow \infty} \mathbf{x}(t, \mathbf{z}_0; G^{x,y}) \leq 1 - \frac{1}{\overline{\gamma}^x d(\overline{G}^x)}. \quad (9.12)$$

Proof. Let $\mathbf{x}_0 > \epsilon \mathbf{1}_N > 0$ for some $\epsilon > 0$ then,

$$\begin{aligned}\liminf_{t \rightarrow \infty} \mathbf{x}(t, \mathbf{z}_0; G^{x,y}) &\geq \lim_{t \rightarrow \infty} \mathbf{x}(t, (\epsilon \mathbf{1}_N, \mathbf{1}_N); \overline{G}^{x,y}) \rightarrow \left(1 - \frac{1}{\underline{\gamma}^x d(\underline{G}^x)}\right) \\ \limsup_{t \rightarrow \infty} \mathbf{y}(t, \mathbf{z}_0; G^{x,y}) &\leq \lim_{t \rightarrow \infty} \mathbf{y}(t, (\epsilon \mathbf{1}_N, \mathbf{1}_N); \overline{G}^{x,y}) \rightarrow 0,\end{aligned}$$

where the convergence on the right-hand side is due to Lemma 64 and the inequalities follow from Theorem 63. The case when $\mathbf{x}_0 \not> 0$ is treated similarly as in the proof of Theorem 48, in that there exists $\delta > 0$ so that

$$\mathbf{x}(t, \mathbf{z}_0; G^{x,y}) > \epsilon > 0$$

for all $t \in (0, \delta)$ for some ϵ . Otherwise, if $\mathbf{x}_0 = 0$ or $\mathbf{y}_0 = 0$ (no virus of type x or y in the system) then, clearly

$$(\mathbf{x}(t, \mathbf{z}_0; G^{x,y})) = 0 \quad \text{or} \quad (\mathbf{y}(t, \mathbf{z}_0; G^{x,y})) = 0.$$

\square

Corollary 66. Let $G^{x,y} = (G^x, G^y) \in \mathcal{E}^2(V)$ be d -regular e -networks with $\gamma^x \equiv \gamma_0^x$ and $\gamma^y \equiv \gamma_0^y$. If $\gamma_0^x d^x > \gamma_0^y d^y$ and $\gamma_0^x d^x > 1$ then,

$$\begin{aligned} \mathbf{x}(t, \mathbf{z}(0); G^{x,y}) &\longrightarrow \left(1 - \frac{1}{d^x \gamma_0^x}\right) \mathbf{1}_N \\ \mathbf{y}(t, \mathbf{z}(0); G^{x,y}) &\longrightarrow \mathbf{0}_N \end{aligned}$$

otherwise, if $\gamma_0^x d^x \leq 1$

$$\mathbf{x}(t, \mathbf{z}(0); G^{x,y}) \longrightarrow \mathbf{0}_N.$$

■

The next Theorem extends Theorem 65 to the case of multivirus epidemics. It states that, if G^1 strongly dominates all the other strains in the network, then it will survive and the remaining strains will die out. The ODE governing the multivirus epidemics is given by,

$$\frac{d}{dt} y_{ik}(t) = \left(\sum_j \gamma_{ji}^k y_{jk}(t) \right) \left(1 - \sum_{\ell=1}^K y_{i\ell}(t) \right) - y_{ik}(t), \quad (9.13)$$

for $i = 1, \dots, N$ and $k = 1, \dots, M$, where $y_{ik}(t)$ stands for the degree of infection at node i by virus type k . This is the corresponding dynamics obtained from the peer-to-peer rules of infection in the limit of large networks for the multivirus case (refer to Part I of this Thesis). Also, for simplicity, we write

$$\mathbf{y}_k(t) = (\mathbf{y}_{1k}(t), \dots, \mathbf{y}_{Nk}(t))$$

(with only one sub-index) as the vector stacking the degree of infection across nodes due to the virus k .

Theorem 67. Let $G^1 \succeq^* G^2 \succeq \dots \succeq G^M$ be the e -networks associated with viruses $k = 1, \dots, M$ governed by the ODE (9.13). Then,

$$\mathbf{y}_m(t) \longrightarrow \mathbf{0}$$

for all $m \geq 2$. Additionally, if $\underline{\gamma}^1 d(\underline{G}_1) > 1$ then, $(\mathbf{y}_1(t))$ obeys

$$1 - \frac{1}{\underline{\gamma}^1 d(\underline{G}^1)} \leq \liminf_{t \rightarrow \infty} \mathbf{y}_1(t, \mathbf{z}_0; G) \leq \limsup_{t \rightarrow \infty} \mathbf{y}_1(t, \mathbf{z}_0; G) \leq 1 - \frac{1}{\bar{\gamma}^1 d(\bar{G}^1)}. \quad (9.14)$$

with $G := (G^1, \dots, G^M)$.

Proof. First, it is easy to check that if $(\mathbf{y}(t))$ is solution to the ODE (9.13) and if $\mathbf{y}_k(0) = \mathbf{0}_N$ for some

$$k \in \{1, \dots, K\}$$

then, $\mathbf{y}_k(t) = \mathbf{0}$ for all time $t \geq 0$. In words, if a virus strain is not present in the network at time $t_0 \geq 0$ then, it will remain extinct for all future times $t \geq t_0$. Now, let

$$\begin{cases} \mathbf{y}_1(0) & \geq \tilde{\mathbf{y}}_1(0) \\ \sum_{k \neq 1} \mathbf{y}_k(0) & \leq \tilde{\mathbf{y}}_2(0) \end{cases} . \quad (9.15)$$

Define

$$(\tilde{\mathbf{y}}(t)) = \left((\mathbf{y}_1, \mathbf{y}_2) \left(t, (\tilde{\mathbf{y}}_1(0), \tilde{\mathbf{y}}_2(0)); \overline{G}^{1,2} \right) \right)$$

as the flow over the regular e-networks

$$\overline{G}^{1,2} = (\overline{G}_1, \overline{G}^2) .$$

The inequalities (9.15) are preserved by the dynamics

$$\begin{cases} \mathbf{y}_1(t) & \geq \tilde{\mathbf{y}}_1(t) \\ \sum_{k \neq 1} \mathbf{y}_k(t) & \leq \tilde{\mathbf{y}}_2(t) \end{cases} , \quad (9.16)$$

for all $t \geq 0$, where $(\mathbf{y}(t))$ and $(\tilde{\mathbf{y}}(t))$ are solutions to (9.13) with initial conditions $\mathbf{y}(0)$ and $\tilde{\mathbf{y}}(0)$ obeying inequalities (9.15). We can establish this fact through similar invariance type of arguments as, for instance, in the proof of Theorem 63: let T be the hitting time to invalidate any of the inequalities in equation (9.16), assume that $T < \infty$ and reach a contradiction (we do not repeat the steps here). Now, from Corollary 66

$$\begin{cases} \mathbf{y}_1(t) & \geq \tilde{\mathbf{y}}_1(t) \rightarrow \left(1 - \frac{1}{\underline{\gamma}^1 d(\overline{G}^1)} \right) \mathbf{1}_N \\ \sum_{k \neq 1} \mathbf{y}_k(t) & \leq \tilde{\mathbf{y}}_2(t) \rightarrow \mathbf{0}_N \end{cases} \quad (9.17)$$

and the theorem is proved. \square

Theorems 65 and 67 state that if $\underline{\gamma}^x \underline{G}^x > 1$ and $G^x \preceq^* G^y$ then, the set

$$\mathcal{A} = \left[1 - \frac{1}{\underline{\gamma}^x d(\underline{G}^x)}, 1 - \frac{1}{\overline{\gamma}^x d(\overline{G}^x)} \right]^N \times \{0\}^N \times \dots \times \{0\}^N$$

is an attractor to (7.4)- (7.5) with basin of attraction

$$\mathcal{B} = \{(\mathbf{x}_0, \mathbf{y}_0) \in \mathcal{D}_2(V) : \mathbf{x}_0 \neq \mathbf{0}\} .$$

Corollary 66 states that, when symmetry is assumed, namely, G^x is regular then, the attractor \mathcal{A} reduces to the singleton

$$\left\{ 1 - \frac{1}{d(G^x) \gamma^x} \right\}^N \times \{0\}^N \times \dots \times \{0\}^N .$$

9.3 Concluding Remarks

In this Chapter we studied the dynamics (7.4)-(7.5) of bi-virus epidemics over general e-networks. We started by showing, in Section 9.1, that the flow $(\mathbf{y}(t, \mathbf{z}(0), G))$ associated with the single virus dynamics (7.3) preserves upper/lower bounds on the initial conditions and on the underlying e-networks, i.e.,

$$\mathbf{z}(0) \geq \tilde{\mathbf{z}}(0) \text{ and } G \succeq \tilde{G} \Rightarrow \mathbf{y}(t, \mathbf{z}(0), G) \geq \mathbf{y}(t, \tilde{\mathbf{z}}(0), \tilde{G})$$

for all $t, t \geq 0$. This was further extended to the bi-virus case in Section 9.2, where now the bi-virus flow $((\mathbf{y}, \mathbf{x})(t, \mathbf{z}(0), G^{x,y}))$ is shown to preserve *skew* bounds on the initial conditions and on the e-networks of the underlying strains. We showed that a certain subclass of dynamical systems, namely, symmetric and regular, were easy to characterize. This allowed us to establish the attractors and the corresponding basins of attraction of the system by bounding any solution to (7.4)-(7.5) by appropriately initialized solutions over regular networks. Namely, we showed that if a strain is strongly dominated, then it necessarily dies out whereas the stronger strain persists if it is strong enough. Our contribution consists in bypassing the lack of Lyapunov tools in the case of bivirus over general networks to, instead, exploring the monotonicity of the system to upper/lower-bounding any solution by well-characterized *symmetric solutions* and then determine the attractors of the system.

Chapter 10

Conclusion and Future Work

In this Thesis, we studied bi-virus epidemics over large-scale networks. We divided the Thesis into two parts. **In Part I**, we determined the macroscopic emergent dynamics of diffusion of the epidemics, associated with the evolution over time of the fraction of infected nodes $\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right)$, from the microscopic peer-to-peer simple rules of diffusion. The problem was addressed under the framework of weak convergence of stochastic processes. By emergent dynamics, we mean a set of nonlinear coupled ordinary differential equations whose solutions are the weak limit of $\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right)$, when the network grows large, in a precise sense. **In Part II**, we study the qualitative behavior of the emergent dynamics, i.e., the nonlinear coupled ordinary differential equations, obtained in Part I. We determine the attractors and corresponding basins of attraction of the fluid limit dynamical system. We showed that, under appropriate conditions on the rates of infection associated to each strain of virus, we observe a natural selection phenomenon, namely, if a strain of virus is *strongly* dominated by another strain, then it necessarily dies out.

In this Chapter, we provide a concluding summary of our work (refer to Sections 10.1-10.3), highlighting our contributions, and outlining possible lines of research for future work (refer to Sections 10.2-10.4).

10.1 Part I: Concluding Summary

Our contribution in Part I lies in establishing rigorously the exact *weak functional law of large numbers* of the macroscopic dynamics of a multivirus epidemics over a non-complete, namely, complete-multipartite, network **from the microscopic peer-to-peer rules of diffusion**. More precisely, we defined simple peer-to-peer rules of infection at the node level, by assuming a stochastic network model for the dynamics of the microstate process $\left(\mathbf{X}^{\mathbf{N}}(t)\right)$, stacking the state of each node n , i.e, $X_{nik}^{\mathbf{N}}(t) = 1$, if node n at island i is infected with virus k at time t , or $X_{nik}^{\mathbf{N}}(t) = 0$, if otherwise. Then, we proved that the normalized macrostate càdlàg process $\left(\bar{\mathbf{Y}}^{\mathbf{N}}(t)\right)$ defined over the compact interval $[0, T]$, and collecting the fraction of k -infected nodes

$$\left(\bar{Y}_{ik}^{\mathbf{N}}(t)\right) = \sum_{n=1}^{N_i} X_{nik}^{\mathbf{N}}(t)$$

per island $i \in \{1, \dots, M\}$ with $k \in \{1, \dots, K\}$ over the complete-multipartite network $G^{\mathbf{N}}$, converges weakly as \mathbf{N} grows large, with respect to the Skorokhod topology $D_{[0,T]}$ on the space of *càdlàg* sample paths, to the solution $(y(t))$ of an $(M \times K)$ -dimensional ordinary differential equation given by (5.4) and restricted to the interval $[0, T]$. For the formal statement, please refer to Theorem 28, in Chapter 5. We first proved that the underlying martingale perturbation $(\bar{\mathbf{M}}^{\mathbf{N}}(t))$ converges weakly and in probability to zero as \mathbf{N} grows large. We further proved that the macrostate family $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is tight (in fact, C -tight). Then, we showed that any weak accumulation point of the family $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ is necessarily solution to the vector ordinary differential equation (5.4) that has a Lipschitz vector field. From the uniqueness of the solutions of the resulting fluid limit differential equations (5.4), we concluded that the whole sequence $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ converges weakly to the solution of (5.4) (refer to the second bullet of Prohorov's Theorem 9, in Chapter 2).

In Chapter 6, we presented our preliminary results on the emergent dynamics of the fraction of infected nodes process $(\bar{Y}^{\mathbf{N}}(t))$ over regular networks. We showed that, under the peer-to-peer model without shuffling (explored in Section 6.2), the process $(\bar{Y}^{\mathbf{N}}(t))$ and its rates are tight. For the process $(\bar{Y}_{\text{sh}}^{\mathbf{N}}(t))$ under shuffling (refer to Section 6.3), we proved convergence in finite-dimensional distribution to almost surely continuous functions. We pursue to prove or disprove tightness of $(\bar{Y}_{\text{sh}}^{\mathbf{N}}(t))$ and convergence in f.d.d. of $(\bar{Y}^{\mathbf{N}}(t))$ in future work. This would in particular imply that $(\bar{Y}_{\text{sh}}^{\mathbf{N}}(t))$ and $(\bar{Y}^{\mathbf{N}}(t))$ converge weakly.

10.2 Part I: Future Work

We suggest two possible lines of research for future work:

1. **(Sparse Networks):** Complete the preliminary work laid down in Chapter 6 to determine the emergent dynamics over regular networks in two cases: i) infection with shuffling; ii) infection without shuffling. For the first case, establish whether the rate process associated with the process $(\bar{Y}_{\text{sh}}^{\mathbf{N}}(t))$ is tight. If tightness holds, then, from Theorem 37, the process $(\bar{Y}_{\text{sh}}^{\mathbf{N}}(t))$ converges weakly to the solution of (6.16). For the second case, we have already shown, in Section 6.2, that the rate process associated with $(\bar{Y}^{\mathbf{N}}(t))$ is tight. But, to determine whether there exists a fluid limit dynamics given by an ODE, one should guarantee that the rate process converges in finite dimensional distribution to a continuous function of $(\bar{Y}^{\mathbf{N}}(t))$. Such function will be the vector field of the limit ODE.
2. **(Pre-limit Analysis):** Extend the results on the qualitative analysis of the limit dynamics (5.4) obtained in Part II to the stochastic case when \mathbf{N} is *large*, but finite. As observed in Section 5.3, in Chapter 5, the process $(\bar{\mathbf{Y}}^{\mathbf{N}}(t))$ weakly approximates the solution $(y(t))$ of the ODE (5.4) in the compact interval $[0, T]$. In other words, we did not prove weak convergence with respect to the topology $D_{[0,\infty)}$, but rather w.r.t. $D_{[0,T]}$. This fact is not surprising as the time to extinction of the virus strains in the network is almost surely finite, regardless how virulent they are. Therefore, in a pre-limit

analysis, one should rather address natural selection under this metastability framework: what is the strain that is likely to die out first? It dies out exponentially fast? What is the strain that is likely to last the longer? These are important questions that should be addressed in future work.

10.3 Part II: Concluding Summary

We presented a qualitative analysis of the bi-virus epidemics dynamics (7.4)-(7.5) over general e-networks, i.e., we determined the attractors and corresponding basins of attraction of the system. We built our analysis in steps: in Section 8.1, we analyzed the simplest case – the single virus over two-node e-network; in Section 8.2, we tackled the bi-virus case over two-node e-networks; in Section 9.1, we treated the single virus epidemics over general e-networks; and in Section 9.2, we studied the bi-virus epidemics over general e-networks. Our contribution lies in establishing the qualitative behavior of the bi-virus epidemics over general e-networks and on the techniques to study such a system and overcome the lack of a natural Lyapunov function. We presented sufficient conditions on the e-networks G^x and G^y upon which a survival of the fittest phenomenon is observed. Namely, if $G^x \succeq^* G^y$ then, $\mathbf{y}(t) \rightarrow \mathbf{0}$, i.e., if virus x strongly-dominates virus y , then the strain y will die out in the long run. Also, if $\underline{\gamma}^x d(\underline{G}^x) > 1$, then the strain x persists, where $\underline{\gamma}^x$ is the (constant) rate across all edges of the inner-regular graph \underline{G}^x of G^x . We provided a qualitative analysis of the bi-virus epidemics dynamics given by (7.4)-(7.5), studying the behavior of the dynamical system from the initial conditions to the corresponding attractors. Indeed, we have shown that

$$\mathcal{A} = \left[1 - \frac{1}{\underline{\gamma}^x d(\underline{G}^x)}, 1 - \frac{1}{\overline{\gamma}^x d(\overline{G}^x)} \right]^N \times \{0\}^N \times \dots \times \{0\}^N$$

is an attractor of the dynamical system (7.4)-(7.5), with basin of attraction given by $\mathcal{B} = \{(\mathbf{x}_0, \mathbf{y}_0) : \mathbf{x}_0 \neq \mathbf{0}\}$. Our method to derive the qualitative analysis of such coupled nonlinear dynamical system is not Lyapunov theory nor numerical simulations based. Instead, we explored a monotonous structure of the system. Namely, the flow to (7.4)-(7.5) of bi-virus epidemics in a general pair of e-networks $G^{x,y} = (G^x, G^y)$ preserves skewed bounded initial conditions and e-networks: if $x_0 \geq x(0)$, $y_0 \leq y(0)$ and $S^x \preceq G^x$, $S^y \succeq G^y$ then,

$$\begin{aligned} \mathbf{x}(t, (\mathbf{x}_0, \mathbf{y}_0); G^{x,y}) &\geq \mathbf{x}(t, (\mathbf{x}(0), \mathbf{y}(0)); S^{x,y}) \\ \mathbf{y}(t, (\mathbf{x}_0, \mathbf{y}_0); G^{x,y}) &\leq \mathbf{y}(t, (\mathbf{x}(0), \mathbf{y}(0)); S^{x,y}). \end{aligned}$$

Also, when the underlying e-networks are regular, the qualitative dynamics of the system can be fully understood. Combining these two properties, we could bound the bi-virus dynamics over general e-networks by the dynamics over inner/outer-regular e-networks to characterize the attractors of the general system.

10.4 Part II: Future Work

We suggest three possible lines of research for future work:

1. **(Finer Attractors):** Determine whether the attractors characterized in Chapter 9 contain finer attractors. For instance, are the equilibrium points of (5.4) global attractors? In other words, do the attractors that we determined possibly hide limit cycles for some configuration of rate parameters? The simulations suggest that the equilibrium of the system is a global attractor.
2. **(Weaker Dominance):** We proved that if $G^x \succeq^* G^y$, i.e., the virus strain y is strongly dominated by x , then the strain y dies out. Can we extend this to claim that, if $G^x \succeq G^y$, then the strain y dies out? The simulations suggest that the answer is yes.
3. **(Co-resilience):** What are the conditions on the rate parameters to enable two strains of virus to co-exist in the network? Fig. 10.1 illustrates that this is possible. Fig. 10.1b depicts the evolution of a solution to the ODE (5.4), for the bi-virus case. Fig. 10.1a depicts the underlying 3 node e-network associated with the strains of virus x and y . Fig. 10.1 illustrates that both strains persist for the set of rate parameters γ_{ij}^x and γ_{ij}^y represented in Fig. 10.1a.

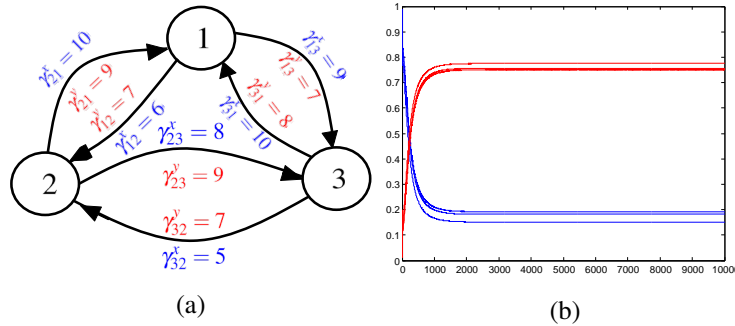


Figure 10.1: Fluid limit epidemics evolution over a 3 node e-network. The rates of virus x and y from node i to node j are given by γ_{ij}^x and γ_{ij}^y , respectively. Each curve in blue/lighter color and red/darker color represent the evolution of each $(x_i(t, \mathbf{x}_0, \mathbf{y}_0))$ and $(y_i(t, \mathbf{x}_0, \mathbf{y}_0))$, respectively (one curve per index i).

Appendix A

Appendix

In this Appendix we present the proofs to the Theorems 55, 56, 61 and 62 relabeled in this Appendix as Theorems 69, 70, 71 and 72, respectively. Theorem 68 conforms to an important intermediate theorem from analysis to establish the Theorems 69, 70, 71 and 72, and we include its proof for completeness.

Theorem 68. *Let $f : (0, +\infty) \rightarrow \mathbb{R}$ be an analytic function. If for some $T \in \mathbb{R}$ we have*

$${}^{(k)}f(T) > 0, \quad {}^{(m)}f(T) = 0, \quad \forall \quad m = 0, 1, \dots, k-1 \text{ and } k \geq 1$$

then, there exists $\epsilon > 0$ such that $f(t) > 0$ for all $t \in (T, T + \epsilon)$.

Proof. Without loss of generality, assume $T = 0$. Since $f \in C^\omega(\mathbb{R})$ then,

$$f(t) = f(0) + \dot{f}(0)t + \ddot{f}(0)t^2 + \dots + \frac{{}^{(k)}f(0)}{k!}t^k + r(t) = f(0) + \left(\frac{{}^{(k)}f(0)}{k!} + \frac{r(t)}{t^k} \right) t^k,$$

with

$$\frac{|r(t)|}{t^k} \rightarrow 0 \text{ as } t \rightarrow 0.$$

Choose δ such that

$$\frac{|r(t)|}{t^k} < \frac{{}^{(k)}f(0)}{2}, \quad \forall \quad t \in (0, \delta).$$

Then,

$$\frac{{}^{(k)}f(0)}{k!} + \frac{r(t)}{t^k} > 0, \quad \forall \quad t \in (0, \delta).$$

Then,

$$f(t) = f(0) + \left(\frac{{}^{(k)}f(0)}{k!} + \frac{r(t)}{t^k} \right) t^k > 0, \quad \forall \quad t \in (0, \delta).$$

□

Theorem 69 (Corresponds to Theorem 55 in Chapter 9). *Let $y_i(t) = 0$ and $y_j(t) = 0, \forall j \in \cup_{l=1}^n \mathcal{N}^l(i)$ for some time $t \geq 0$. Then,*

$$y_i^{(\ell)}(t) = 0, \quad \forall 0 < \ell \leq n, \quad (\text{A.1})$$

that is, if there is no infection within a neighborhood up to order n of node i , then all derivatives of $(y_i(t))$ up to order n are zero. ■

Proof. We apply induction on the order n .

Step 1: For $n = 1$:

$$\dot{y}_i(t) = \left(\sum_q \gamma_{qi} y_q(t) \right) (1 - y_i(t)) - y_i(t) = 0.$$

Step 2: Induction step. We assume that Theorem 55 holds for $n - 1$ and prove it holds for order n . By algebraic and reordering manipulations, we can show:

$$y_i^{(n)}(t) = \underbrace{\left(\sum_k \gamma_{ki} y_k^{(n-1)}(t) \right)}_{\mathbf{A}} - \underbrace{\sum_{\ell=0}^{n-1} \binom{n-1}{\ell} y_i^{(\ell)}(t) \left(\sum_q \gamma_{qi} y_q^{(n-1-\ell)}(t) \right)}_{\mathbf{B}} - \underbrace{y_i^{(n-1)}(t)}_{\mathbf{C}} \quad (\text{A.2})$$

holds for all $n \in \mathbb{N}$. We analyze now each term **A**, **B** and **C** in equation (A.2). First, note that from the induction hypothesis $y_i^{(\ell)}(t) = 0$ for all $\ell = 1, \dots, n - 1$.

A: Since by assumption $y_j(t) = 0, \forall j \in \cup_{l=1}^n \mathcal{N}^l(i)$, then, if $k \in \mathcal{N}(i)$, by induction, $y_k^{(\ell)}(t) = 0$ for all $\ell = 1, \dots, n - 1$. Therefore,

$$\sum_k \gamma_{kj} y_k^{(n-1)}(t) = 0,$$

i.e., term **A** is zero.

B: By assumption, $y_i(t) = 0$, and by induction, for all $\ell = 1, \dots, n - 1$, $y_i^{(\ell)}(t) = 0$, hence term **B** is zero.

C: Term **C** is zero, since by induction $y_i^{(n-1)}(t) = 0$. □

Theorem 70 (Corresponds to Theorem 56 in Chapter 9). *Let $y_i(t) > 0$ and $y_j(t) = 0, \forall j \neq i$ for some time $t \geq 0$. Then, $\forall \ell < n$*

$$j \in \mathcal{N}^n(i) \Rightarrow y_j^{(n)}(t) > 0 \text{ and } y_j^{(\ell)}(t) = 0. \quad (\text{A.3})$$

■

Proof. Now, we apply induction on the number of hops n .

Step 1: For $n = 1$, we have that $j \in \mathcal{N}(i)$ and

$$\dot{y}_j(t) = \left(\sum_k \gamma_{kj} y_k(t) \right) (1 - y_j(t)) - y_j(t) = \sum_k \gamma_{kj} y_k(t) = \gamma_{ij} y_i(t) + \underbrace{\sum_{k \neq i} \gamma_{kj} y_k(t)}_{=0} = \gamma_{ij} y_i(t) > 0.$$

That is, $y_j(t) = 0$ and $\dot{y}_j(t) > 0$.

Step 2: Induction step. Assume assertion (9.2) holds for $n - 1$. We consider successively the terms **A**, **B**, and **C**.

A: By definition,

$$j \in \mathcal{N}^n(i) \Rightarrow \exists k \in \mathcal{N}^{n-1}(i) : j \rightarrow k.$$

From the induction hypothesis, $\binom{n-1}{k} y_k(t) > 0$. Therefore,

$$\sum_k \gamma_{kj} \binom{n-1}{k} y_k(t) > 0$$

and term **A** is strictly positive.

B: From Theorem 55, for $j \in \mathcal{N}^n(i)$, $y_j^{(\ell)}(t) = 0$, $\forall \ell = 1, \dots, n-1$, and, thus, term **B** is zero.

C: From Theorem 55, for $j \in \mathcal{N}^n(i)$, $y_j^{(n-1)}(t) = 0$, and term **C** is zero.

Therefore, $y_j^{(n)}(t) > 0$ with $y_j^{(\ell)}(t) = 0$, $\forall \ell < n$, and the Theorem is proved. \square

Theorem 71 (Corresponds to Theorem 61 in Chapter 9). *Let $y_i(0) = \tilde{y}_i(0)$ and $x_i(0) = \tilde{x}_i(0)$. Let*

$$\overline{\mathcal{N}}^n(i) := \bigcup_{\ell=1}^n \mathcal{N}^\ell(i).$$

Then:

$$\begin{cases} y_k(0) = \tilde{y}_k(0) \\ x_k(0) = \tilde{x}_k(0) \end{cases} \quad \forall k \in \overline{\mathcal{N}}^n(i) \implies y_i^{(\ell)}(0) = \tilde{y}_i^{(\ell)}(0) \quad \text{and} \quad x_i^{(\ell)}(0) = \tilde{x}_i^{(\ell)}(0), \quad \forall \ell \leq n. \quad \blacksquare$$

Proof. We apply induction on n . For $n = 1$,

$$\begin{aligned} \frac{d}{dt} y_i(t) &= \left(\sum_j \gamma_{ji}^y y_j(t) \right) (1 - x_i(t) - y_i(t)) - y_i(t) \\ \frac{d}{dt} \tilde{y}_i(t) &= \left(\sum_j \gamma_{ji}^y \tilde{y}_j(t) \right) (1 - \tilde{x}_i(t) - \tilde{y}_i(t)) - \tilde{y}_i(t). \end{aligned}$$

Note that $\tilde{y}_j(0) = y_j(0)$, $\forall j \in \mathcal{N}(i)$ and $x_i(0) = \tilde{x}_i(0)$. By inspection, $\tilde{\dot{y}}_i(0) = \dot{y}_i(0)$ and (by assumption) $y_i(0) = \tilde{y}_i(0)$.

Now, assume Theorem 61 holds for $n - 1$. We establish that it holds for n . We have:

$$\binom{n}{i} y_i(0) = \underbrace{\left(\sum_j \gamma_{ji}^y \binom{n-1}{j} y_j(0) \right) (1 - y_i(0) - x_i(0))}_{\mathbf{A}} - \underbrace{\sum_{\ell=1}^{n-1} \binom{n-1}{\ell} y_i^{(\ell)}(0) \left(\sum_q \gamma_{qi}^y \binom{n-1-\ell}{q} y_q(0) \right)}_{\mathbf{B}} \quad (\text{A.4})$$

$$\begin{aligned}
& - \underbrace{\binom{n-1}{\ell} y_i(0)}_{\mathbf{C}} - \underbrace{\sum_{\ell=1}^{n-1} \binom{n-1}{\ell} \binom{\ell}{\ell} x_i(0) \left(\sum_j \gamma_{ji}^y \binom{n-1-\ell}{\ell} y_j(0) \right)}_{\mathbf{D}}, \\
\binom{n}{\ell} \dot{y}_i(0) &= \underbrace{\left(\sum_j \gamma_{ji}^y \binom{n-1}{\ell} \tilde{y}_j(0) \right) (1 - \tilde{y}_i(0) - \tilde{x}_i(0))}_{\mathbf{A}} - \underbrace{\sum_{\ell=1}^{n-1} \binom{n-1}{\ell} \binom{\ell}{\ell} \tilde{y}_i(0) \left(\sum_q \gamma_{qi}^y \binom{n-1-\ell}{\ell} \tilde{y}_q(0) \right)}_{\mathbf{B}} \\
& - \underbrace{\binom{n-1}{\ell} \tilde{y}_i(0)}_{\mathbf{C}} - \underbrace{\sum_{\ell=1}^{n-1} \binom{n-1}{\ell} \binom{\ell}{\ell} \tilde{x}_i(0) \left(\sum_j \gamma_{ji}^y \binom{n-1-\ell}{\ell} \tilde{y}_j(0) \right)}_{\mathbf{D}}.
\end{aligned} \tag{A.5}$$

Recall the assumption

$$\begin{cases} y_k(0) = \tilde{y}_k(0) \\ x_k(0) = \tilde{x}_k(0) \end{cases} \quad \forall k \in \overline{\mathcal{N}}^{(n)}(i). \tag{A.6}$$

From the induction hypothesis, $\forall j \in \mathcal{N}(i), \forall \ell = 1, \dots, n-1$,

$$\binom{\ell}{\ell} y_j(0) = \binom{\ell}{\ell} \tilde{y}_j(0), \quad \binom{\ell}{\ell} x_j(0) = \binom{\ell}{\ell} \tilde{x}_j(0)$$

and also

$$\binom{\ell}{\ell} y_i(0) = \binom{\ell}{\ell} \tilde{y}_i(0), \quad \forall \ell \leq n-1.$$

Therefore, by inspection, we conclude that the terms **A**, **B**, **C**, and **D** for both equations (A.4) and (A.5) match together, and, thus, $\binom{n}{\ell} y_j(0) = \binom{n}{\ell} \tilde{y}_j(0)$. The same reasoning goes with $\binom{n}{\ell} x_j(0) = \binom{n}{\ell} \tilde{x}_j(0)$, and we conclude the proof of the Theorem. \square

Theorem 72 (Corresponds to Theorem 62 in Chapter 9). *Let $y_i(0) = \tilde{y}_i(0)$, $x_i(0) = \tilde{x}_i(0)$ and*

$$\begin{cases} y_k(0) = \tilde{y}_k(0) \\ x_k(0) = \tilde{x}_k(0) \end{cases} \quad \forall k \in \overline{\mathcal{N}}^{n-1}(i) \text{ and } \begin{cases} y_k(0) \geq \tilde{y}_k(0) \\ x_k(0) \leq \tilde{x}_k(0) \end{cases} \quad \forall k \in \mathcal{N}^n(i) \setminus \{m\} \tag{A.7}$$

with strict inequality $y_m(0) > \tilde{y}_m(0)$ for some $m \in \mathcal{N}^n(i)$. Then, $\binom{n}{\ell} y_i(0) > \binom{n}{\ell} \tilde{y}_i(0)$. \blacksquare

Proof. We apply induction on the number of hops n .

Case 1: For $n = 1$, from the assumptions of the Theorem, namely, $y_m(0) > \tilde{y}_m(0)$ for some $m \in \mathcal{N}(i)$ we conclude that $\dot{y}_i(0) > \dot{\tilde{y}}_i(0)$ since

$$\dot{y}_i(0) = \left(\sum_j \gamma_{ji}^y y_j(0) \right) (1 - x_i(0) - y_i(0)) - y_i(0)$$

$$\begin{aligned}
&> \left(\sum_j \gamma_{ji}^y \tilde{y}_j(0) \right) (1 - x_i(0) - y_i(0)) - y_i(0) \\
&= \dot{\tilde{y}}_i(0).
\end{aligned}$$

Case 2: Induction step. Assume that Theorem 62 holds for $n - 1$ and let us prove that it holds for n . We consider successively the terms **A**, **B**, **C**, and **D** in equations (A.4) and (A.5).

A: Note that for some $j \in \mathcal{N}(i)$ we have that $m \in \mathcal{N}^{(n-1)}(j)$ where m is defined in the assumptions of the Theorem. Thus, by the induction hypothesis, we have $y_j(0) > \tilde{y}_j(0)$, and, hence, the term **A** in equation (A.4) is greater than its counterpart in equation (A.5).

B and **C:** From Theorem 61, it should be now clear that these terms match together between equations (A.4) and (A.5).

D: From Theorem 61, it follows that $x_j^{(\ell)}(0) = \tilde{x}_j^{(\ell)}(0)$ for all $\ell = 1, \dots, n - 1$ and thus, term **D** is the same for both equations.

Therefore, $y_j^{(n)}(0) > \tilde{y}_j^{(n)}(0)$ and the Theorem is proved. \square

Theorem 73 (Orthogonality; Proposition A.10 in [26]). *Let $(\mathbf{Y}(t))$ be an (\mathcal{F}_t) -adapted càdlàg process with discrete range and piecewise constant (i.e., constant when it does not jump). Let $\mathcal{N}_\lambda(t)$ and $\mathcal{N}_\mu(t)$ be two independent (\mathcal{F}_t) -adapted Poisson processes (hence their compensated versions are (\mathcal{F}_t) -martingales, as it is trivial to establish). Assume the rates λ, μ are nonnegative. Let f, g be two bounded functions defined over the discrete range of $\mathbf{Y}(t)$. Then,*

$$\left(\int_0^t f(\mathbf{Y}(s-)) (\mathcal{N}_\lambda(ds) - \lambda ds) \int_0^t g(\mathbf{Y}(s-)) (\mathcal{N}_\mu(ds) - \mu ds) \right) \quad (\text{A.8})$$

is an (\mathcal{F}_t) -martingale.

Proof. This proof goes in line with the one provided in Proposition A.10, page 356, in [26], but we provide more details.

First, define

$$\begin{aligned}
(A_1(t)) &:= \left(\int_0^t f(\mathbf{Y}(s-)) (\mathcal{N}_\lambda(ds) - \lambda ds) \right) \\
(A_2(t)) &:= \left(\int_0^t g(\mathbf{Y}(s-)) (\mathcal{N}_\mu(ds) - \mu ds) \right).
\end{aligned}$$

Note that $(A_i(t))$ are \mathcal{F}_t -martingale. Indeed, it follows from Theorem 1.5.1 in [37] that if $(H(t))$ is a bounded \mathcal{F}_t -predictable process then,

$$\left(\int_0^t H(s) (\mathcal{N}_\lambda(ds) - \lambda ds) \right) \quad (\text{A.9})$$

is an \mathcal{F}_t -martingale. Moreover, $(A_i(t))$ are finite variation martingales.

Now, from Ito integration by parts formula, remark that if $(X(t))$ and $(Y(t))$ are semimartingales, then,

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(s-)dY(s) + \int_0^t Y(s-)dX(s) + [X, Y](t). \quad (\text{A.10})$$

Now, if $(Y(t))$ is a finite variation process, then

$$[X, Y](t) = \sum_{0 \leq s \leq t} \Delta X(s) \Delta Y(s), \quad (\text{A.11})$$

refer to [38]. Therefore, and since $(A_i(t))$ are martingales (and thus, semimartingales) of finite variation, we have

$$A_1(t)A_2(t) = A_1(0)A_2(0) + \int_0^t A_1(s-)dA_2(s) + \int_0^t A_2(s-)dA_1(s) + \sum_{0 \leq s \leq t} \Delta A_1(s) \Delta A_2(s). \quad (\text{A.12})$$

As referred in [26], since the underlying Poisson processes are independent, almost surely, they will have no common discontinuity point between 0 and t and therefore,

$$\sum_{0 \leq s \leq t} \Delta A_1(s) \Delta A_2(s) = 0 \quad (\text{A.13})$$

and thus,

$$A_1(t)A_2(t) = \int_0^t A_1(s-)dA_2(s) + \int_0^t A_2(s-)dA_1(s). \quad (\text{A.14})$$

Now, since $(A_i(t))$ are of finite variation, the stochastic integrals are almost surely the Stieltjes integrals and the integration by part formula is now given (as done in [26])

$$A_1(t)A_2(t) = \int_0^t A_2(s-)f(\mathbf{Y}(s-))(\mathcal{N}_\lambda(ds) - \lambda ds) + \int_0^t A_1(s-)g(\mathbf{Y}(s-))(\mathcal{N}_\mu(ds) - \mu ds). \quad (\text{A.15})$$

Again, from Theorem 1.5.1 in [37], it follows that $(A_1(t)A_2(t))$ is an \mathcal{F}_t -martingale.

□

Bibliography

- [1] F. Cucker and S. Smale, “Emergent behavior in flocks,” *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 852–862, May 2007.
- [2] S. Kar and J. M. F. Moura, “Global emergent behaviors in clouds of agents,” in *2011 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011, pp. 5796–5799.
- [3] M. P. Richardson, “Large scale brain models of epilepsy: dynamics meets connectomics,” *Journal of Neurology, Neurosurgery, and Psychiatry*, no. 83, pp. 1238–1248, December 2012.
- [4] N. Antunes, C. Fricker, P. Robert, and D. Tibbi, “Stochastic networks with multiple stable points,” *The Annals of Applied Probability*, vol. 36, no. 1, pp. 255–278, 2008.
- [5] P. Van Mieghem, J. Omic, and R. Kooij, “Virus spread in networks,” *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, February 2009.
- [6] L. C. d. C. Medeiros, C. A. R. Castilho, C. Braga, W. V. de Souza, L. Regis, and A. M. V. Monteiro, “Modeling the Dynamic Transmission of Dengue Fever: Investigating Disease Persistence,” *PLoS Negl Trop Dis*, vol. 5, no. 1, 01 2011.
- [7] S. Flint, *Principles of Virology: Molecular Biology, Pathogenesis, and Control*. ASM Press, 2000. [Online]. Available: <http://books.google.pt/books?id=IBVrAAAAMAAJ>
- [8] E. K. Brunson, “The impact of social networks on parents’ vaccination decisions,” *Pediatrics*, vol. 131, no. 5, pp. e1397–e1404, 2013.
- [9] T. Oraby, V. Thampi, and C. T. Bauch, “The influence of social norms on the dynamics of vaccinating behaviour for pediatric infectious diseases,” *Proc. of the Royal Society B: Biological Sciences*, vol. 281, no. 1780, 2014.
- [10] O. C. Finney, G. J. Keitany, H. Smithers, A. Kaushansky, S. Kappe, and R. Wang, “Immunization with genetically attenuated p. falciparum parasites induces long-lived antibodies that efficiently block hepatocyte invasion by sporozoites,” *Vaccine*, vol. 32, no. 19, pp. 2135–2138, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0264410X14002400>
- [11] A. Santos, J. M. F. Moura, and J. M. F. Xavier, “Emergent behavior in multipartite large networks: Multi-virus epidemics,” 2013, submitted. <http://arxiv.org/abs/1306.6198>.
- [12] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks*. London, UK: Cambridge University Press, November 2012.
- [13] M. O. Jackson, *Social and Economic Networks*. Princeton University Press, 2008.

- [14] A. Lahmanovich and A. Y. James, "A deterministic model for gonorrhea in a nonhomogeneous population," *Mathematical Biosciences*, vol. 28, pp. 221–236, 1976.
- [15] N. Parikh, M. Youssef, S. Swarup, and S. Eubank, "The influence of social norms on the dynamics of vaccinating behaviour for paediatric infectious diseases," *Scientific Reports*, vol. 3, no. 3152, 2013.
- [16] A. Santos, J. M. F. Moura, and J. M. F. Xavier, "Bi-virus SIS epidemics over networks: Qualitative analysis," 2014, submitted.
- [17] A. Santos and J. M. F. Moura, "Emergent behavior in large scale networks," in *2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, December 2011, pp. 4485–4490.
- [18] A. Santos, J. M. F. Moura, and J. M. F. Xavier, "Diffusion and topology: Large densely connected bipartite networks," in *2012 51th IEEE Conference on Decision and Control and European Control Conf. (CDC-ECC)*, December 2012, pp. 4485–4490.
- [19] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., ser. Probability and Statistics. Wiley, August 1999.
- [20] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, ser. Probability and Statistics. New York, NY: John Wiley & Sons, Inc., 1986.
- [21] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales: Foundations*, 2nd ed., ser. Cambridge Mathematical Library. Cambridge, UK: Cambridge University Press, April 2000, vol. 1.
- [22] V. Bogachev, *Measure Theory*, ser. Measure Theory. Springer, 2007, no. v. 1-2. [Online]. Available: <http://books.google.pt/books?id=5vbYMQEACAAJ>
- [23] J. Hale, *Ordinary Differential Equations*, ser. Dover Books on Mathematics Series. Dover Publications, 2009. [Online]. Available: <http://books.google.pt/books?id=LdTZJ4HwCv4C>
- [24] D. Aldous, "Stopping times and tightness," *The Annals of Probability*, vol. 6, no. 2, pp. 335–340, 04 1978. [Online]. Available: <http://dx.doi.org/10.1214/aop/1176995579>
- [25] E. Olivieri and M. E. Vares, *Large Deviations and Metastability*, ser. Encyclopedia of Mathematics and its Applications. London, UK: Cambridge University Press, 2005, no. 100.
- [26] P. Robert, *Stochastic Networks and Queues*, ser. Stochastic Modelling and Applied Probability. Springer-Verlag, 2003, vol. 52.
- [27] P. Billingsley, *Probability and Measure*, 3rd ed., ser. Probability and Statistics. John Wiley & Sons, Inc., 1995.
- [28] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales: Foundations*, 2nd ed., ser. Cambridge Mathematical Library. Cambridge, UK: Cambridge University Press, September 2000, vol. 2.
- [29] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 1st ed., ser. Graduate Texts in Mathematics. New York: Springer-Verlag, September 1988, vol. 113.
- [30] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed., ser. Grundlehren der mathematischen Wissenschaften. New York, NY: Springer-Verlag, 2003, vol. 288.

-
- [31] W. R. Pestman, “Measurability of linear operators in the Skorokhod topology,” *Bulletin of the Belgian Mathematical Society - Simon Stevin*, vol. 2, no. 4, pp. 381–388, 1995. [Online]. Available: <http://projecteuclid.org/euclid.bbms/1103408695>
- [32] A. Skorokhod, “Limit theorems for stochastic processes,” *Theory of Probability and Its Applications*, vol. 1, no. 3, pp. 261–290, 1956.
- [33] D. Williams, *Probability with Martingales*, ser. Cambridge Mathematical Textbooks. Cambridge University Press, 1991. [Online]. Available: <http://books.google.pt/books?id=RnOJeRpk0SEC>
- [34] D. J. Daley and J. Gani, *Epidemic Modelling: An Introduction*. Cambridge, UK: Cambridge University Press, 2001.
- [35] B. Bollobas, J. Han, and K. J. Verstraete, “Regular subgraphs of random graphs,” in *Random Structures & Algorithms*, vol. 29, no. 1. Wiley, August 2006, pp. 1–13.
- [36] N. P. Bhatia and G. P. Szego, *Stability Theory of Dynamical Systems*. Springer, 2002.
- [37] T. Fleming and D. Harrington, *Counting Processes and Survival Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2013. [Online]. Available: <http://books.google.pt/books?id=i85dAAAACAAJ>
- [38] [Online]. Available: <http://almostsure.wordpress.com/2010/01/19/properties-of-quadratic-variations/>