Stochastic epidemics in a heterogeneous community (Part III of the book Stochastic Epidemic Models and Inference)

Viet-Chi Tran

Stochastic epidemics in a heterogeneous community

Viet Chi Tran

April 15, 2020

This document is the Part III of the book *Stochastic Epidemic Models with Inference* edited by Tom Britton and Etienne Pardoux [29].

Contents

Acknowledgements: This research has been supported by the “Chaire Modélisation Mathématique et Biodiversité" of Veolia Environnement-Ecole Polytechnique-Museum National d’Histoire Naturelle-Fondation X. V.C.T. also acknowledges support from Labex CEMPI (ANR-11-LABX-0007-01), GdR GeoSto 3477, ANR Project Cadence (ANR-16-CE32-0007) and ANR Project Econet (ANR-18-CE02-0010).

Introduction

Recently, network concepts have received much attention in infectious disease modelling, essentially for modeling purposes, and the reader is also referred to earlier references of Durrett [50], Newman [89], House [62] or Kiss et al. [71]. In the compartmental models presented in Part I of this volume, any infected individual can contaminate any susceptible individuals. In many public health problems, heterogeneity issues have to be taken into account, in particular some diseases such as AIDS or HCV (Hepatitis C Virus) may spread only along a social network: the network of people having sexual intercourse or of injecting drug partners. The need to take into account the network along which an epidemic spreads has been underlined by numerous papers, starting for example from [44, 51], and more recently [18, 62].

After introducing random networks and describing how the spread of disease can be modelled on such structures, we explain how to approximate the dynamics by deterministic differential equations when the graphs are large. Mathematical models for epidemics on large networks are obtained by mean-field approximation (e.g. [50, 72, 92]) or through large population approximations (e.g. [13, 45, 59, 19, 66]). They generally stipulate simple structures for the network: small worlds (e.g. [72, 82]), configuration models (e.g. [70, 76, 94, 111, 112]), random intersection graphs and graphs with overlapping communities (e.g. [27, 16, 41])...

In the last section, real data from the AIDS epidemic in Cuba is studied (data from [36] and that can be found in supplementary materials of this book). We show how to conduct descriptive statistical procedures. By performing clustering and simplification of the graph, we decompose it into smaller clusters where the probabilistic models of the previous sections can be used.

Notation 0.0.1. *In this part, we denote by $\mathbb{N}$ the set of strictly positive integers and by $\mathbb{Z}_+$ the set $\mathbb{N}\cup\{0\}$.*
For any real bounded function f on $\mathbb{Z}_+$, let $\|f\|_\infty$ denote the supremum of f on $\mathbb{Z}_+$. For all such f and $y\in\mathbb{Z}_+$, we denote by $\tau_y f$ the function $x\mapsto f(x-y)$. For all $n\in\mathbb{Z}_+$, χ^n is the function $x\mapsto x^n$, and in particular, $\chi\equiv\chi^1$ is the identity function, and $\mathbf{1}\equiv\chi^0$ is the function constant equal to 1.
We denote by $\mathscr{M}_F(\mathbb{Z}_+)$ the set of finite measures on $\mathbb{Z}_+$, equipped with the topology of weak convergence. For all $\mu\in\mathscr{M}_F(\mathbb{Z}_+)$ and real bounded function f on $\mathbb{Z}_+$, we write

$$\langle\mu,f\rangle=\sum_{k\in\mathbb{Z}_+}f(k)\,\mu(k), \tag{0.0.1}$$

where we use the notation $\mu(k)=\mu(\{k\})$.
For $k\in\mathbb{Z}_+$, we write δ_k for the Dirac measure at k. In particular, for any test function f from $\mathbb{Z}_+$ to $\mathbb{R}$, $\langle\delta_k,f\rangle=f(k)$.
For a sequence $D_1,\dots D_n\in\mathbb{Z}_+$, if $\mu=\sum_{k=1}^n\delta_{D_k}$, then

$$\langle\mu,f\rangle=\sum_{k=1}^n f(D_k),$$

implying in particular that $\langle\mu,1\rangle=n$ and $\langle\mu,\chi\rangle=\sum_{k=1}^n D_k$.

Chapter 1

Random Graphs

1.1 Definitions

Usually, social networks on which disease spread are very complex. It is thus convenient to model them by *random* networks. We start with some definitions, and then present some common families of random networks. There is a growing literature on random networks to which we refer the reader for further developments (e.g. [25, 110]).

Definition 1.1.1. *A random graph $\mathscr{G} = (V,E)$ is a set of vertices V and a set of edges $E \subset V \times V$. If $u, v \in V$ are connected in the random graph, then $(u,v) \in E$.*

The set of vertices of $\mathscr{G}$ is V, but when we will need to make precise that it is the set of vertices of $\mathscr{G}$, we will use the notation $V(\mathscr{G})$. The population size is $|V| = N$. In the sequel, we will label the vertices with integers, so that $V = \{1, \dots N\}$.

Definition 1.1.2. *The adjacency matrix of the graph $\mathscr{G}$ is a matrix $G \in \mathscr{M}_{V\times V}(\mathbb{R})$ such that $\forall u, v \in V$,*

$$G_{uv} = 1 \text{ if } (u,v) \in E,$$
$$G_{uv} = 0 \text{ if } (u,v) \notin E.$$

If the matrix is symmetric, the graph in undirected: to any edge from u to v corresponds an edge from v to u. Else, if $(u,v) \in E$ and $(v,u) \notin E$, the graph is oriented with only the directed edge from u to v belonging to E. We say that u is the *ego* and v the *alter* of the edge.
If we consider weighted graphs, we can generalize the entries of G to real non-negative numbers.

In this chapter, we will focus on undirected non-weighted graphs.

Definition 1.1.3. *The degree of a vertex $u \in V$ in the graph $\mathscr{G}$ is*

$$D_u = \sum_{v\in V} G_{uv}.$$

D_u hence corresponds to the number of neighbours of the vertex u, i.e. the number of the vertices of $\mathscr{G}$ that can be reached in one step starting from u.
If the graph is oriented, the above notion corresponds to the out-degree, and similarly we can define as in-degree the number of vertices of $\mathscr{G}$ that lead to u in one step:

$$D_u^{in} = \sum_{v\in V} G_{vu}.$$

For undirected graphs, the out and in-degrees coincide.

Definition 1.1.4. *The degree distribution of a finite graph $\mathscr{G}$ is:*

$$\frac{1}{N}\sum_{u\in V}\delta_{D_u}=\sum_{d\in\mathbb{Z}_+}\frac{\mathrm{Card}\{u\in V : D_u=d\}}{N}\delta_d.$$

For $d\in\mathbb{Z}_+$, $\mathrm{Card}\{u\in V : D_u=d\}/N$ is the proportion of vertices with degree d.

We see that the notion of degree distribution can be generalized to graphs with infinitely many vertices: the degree distribution is a probability measure on $\mathbb{Z}_+$, $\sum_{d\in\mathbb{Z}_+} p_d\delta_d$, where the weight p_d of the atom $d\in\mathbb{Z}_+$ is the proportion of vertices with degree d.

Let us consider the product of the matrix G with itself: $G^2=G\times G$. Notice that

$$G^2_{uv}=\sum_{w\in V}G_{uw}G_{wv},$$

and thus, $G^2_{uv}>0$ if there is a path consisting of two edges of G that links u and v. More precisely, G^2_{uv} counts the number of paths of length exactly 2 that link u and v. Generalizing this definition, and with the convention that $G^0=\mathrm{Id}$ the identity matrix of $\mathbb{R}^N$, we obtain that:

Definition 1.1.5. *Two vertices u and v of the graph $\mathscr{G}$ are connected if there is a path in $\mathscr{G}$ going from u to v, i.e. if there exists some integer $n\geq 1$ such that $G^n_{uv}>0$. We can then define the graph distance between u and v by:*

$$d_G(u,v)=\inf\{n\geq 0,\ G^n_{uv}>0\}. \tag{1.1.1}$$

By convention, $\inf\emptyset=+\infty$.
For $r\geq 0$, we define by $B_G(u,r)$ the ball of $\mathscr{G}$ with center u and radius r for the graph distance:

$$B_G(u,r)=\big\{v\in V : d_G(u,v)\leq r\big\}.$$

Several important descriptors of the graph depend on this graph distance. We remark for instance that $D_u=\mathrm{Card}(B_G(u,1))-1$. Also, we can define a shortest path (for the graph distance) between two vertices u and v. The diameter of the graph is:

$$\mathrm{diam}(\mathscr{G})=\sup\{d_G(u,v) : u,v\in V\}.$$

Definition 1.1.6. *For a vertex u in a graph $\mathscr{G}$, we denote by $\mathscr{C}(u)$ the connected component of u, i.e. the set of vertices $v\in V$ that are connected to u:*

$$\mathscr{C}(u)=\big\{v\in V : d_G(u,v)<+\infty\big\}.$$

1.2 Classical examples of random graphs

Random graphs, especially those arising from applications, can have very complex distributions and topologies. There are some simple families of random graphs. We now present the complete graph, the Erdös–Rényi graphs, the stochastic block model, the configuration model and the household model.

Definition 1.2.1 (Complete graph)**.** *The complete graph K_N is the graph where all the pairs of vertices are linked by an edge, i.e. $E=V\times V$.*

The complete graph is in fact a deterministic graph, and $\forall u,v\in V(K_N)$, $d_G(u,v)=1$ if $u\neq v$.

Definition 1.2.2 (Erdös–Rényi random graph (ER))**.** *Erdös–Rényi random graphs are undirected graphs where each pair of vertices $(u,v)\in V^2$ is linked by an edge with probability $p\in[0,1]$ independently from the other pairs. The distribution $\mathrm{ER}(N,p)$ of Erdös–Rényi random graphs is completely defined by the family $(G_{uv};\ u,v\in V,\ u<v)$ of i.i.d. random variables with Bernoulli distribution $Ber(p)$, $p\in[0,1]$.*

Notice that for $p = 1$, the Erdös–Rényi graph corresponds to the complete graph K_N.

These graphs can be generalized if we introduce a partition of the population according to a discrete type, taking K values, say $\{1,\dots,K\}$: to each vertex $u \in V$ is associated a type $k_u \in \{1,\dots K\}$. This corresponds to cases where a community contains different types of individuals that display specific roles in contact behaviour. Types might be related to age-groups, social behaviour or occupation.

Definition 1.2.3 (Stochastic block model graph (SBM))**.** *A stochastic block model graph is a undirected graph, where each vertex is given a type independently from the others, all with the same probability, and where each pair of vertices is linked independently of the other pairs with a probability depending on the types of the vertices. If there are K types, say* $\{1,\dots K\}$*, we will denote by* $(\rho_i)_{i\in\{1,\dots K\}}$ *the probability distribution of the types, and by* π_{ij} *the probability of linking a vertex of type* i *with a vertex of type* j*.*

If there is just one type of vertices ($K = 1$), the SBM resumes to ER graphs. For $K = 2$ where vertices of the same type cannot be connected ($\pi_{11} = \pi_{22} = 0$), we obtain *bipartite* graphs. For instance, sexual networks in heterosexual populations are bipartite networks. The interested reader is referred to the review of Abbe [1].

Proposition 1.2.4. *The degree distribution of a vertex* u *in an* $\mathrm{ER}(N,p)$ *random graph with* N *vertices and connection probability* p *is a binomial distribution* $\mathrm{Bin}(N,p)$*. When the connection probability is* λ/N*, with* $\lambda > 0$*, then for any integer* $d \geq 0$*,*

$$\lim_{N\to+\infty} \mathbb{P}_N(D_u = d) = \frac{\lambda^d}{d!} e^{-\lambda},$$

showing that the probability distribution converges to a Poisson distribution with expectation λ*.*

The proof of this result is easy and let to the reader.

A detailed presentation and study of Erdös–Rényi graphs and their limits when $N \to +\infty$ can be found in [110] for example. In particular, the case where the connection probability is λ/N, is carefully discussed. The case $\lambda > 1$ is termed the supercritical case, while the case $\lambda < 1$ is the subcritical case.

Proposition 1.2.4 emphasizes the importance of graphs defined from their degree distributions. The next class of graphs has been introduced by Bollobas [25] and Molloy and Reed [80]. The reader is referred to Durrett [50] and van der Hofstad [110] for more details.

Definition 1.2.5 (Configuration model graph (CM))**.** *Let* $\mathbf{p} = (p_k,\, k \in \mathbb{Z}_+)$ *be a probability distribution on* $\mathbb{Z}_+$*. The Bollobás–Molloy–Reed or Configuration model random graph with vertices* V *is constructed as follows. We associate with each vertex* $u \in V$ *an independent random variable* X_u *drawn from the distribution* $\mathbf{p}$*, that corresponds to the number of half edges attached to* u*. Conditionally on* $\{\sum_{u\in V} X_u \text{ even}\}$*, the Configuration model random graph is a multigraph (a graph with possibly self-loops and multiple edges) obtained by pairing the half-edges uniformly at random.*

A possible algorithm for pairing the half edges (also called stubs) is the following:

- Associate with each half edge an independent uniform random variable on $[0,1]$ and sort the half-edges by decreasing values.
- Pair each odd stub with the following even stub. Note that if the number of stubs $\sum_{u\in V} X_u$ is odd, it is possible to add or remove one stub arbitrarily.

Note that this linkage procedure does not exclude self-loops or multiple edges. When the size of the graph $N \to +\infty$ with a fixed degree distribution, self-loops and multiple edges become less and less apparent in the global picture (see e.g. [50, Theorem 3.1.2]).

In [110], it is carefully studied how one can turn a multigraph into a simple graph (without self-loop nor multi-edge), either by erasing self-loops and merging multi-edges, or by conditioning on obtaining a simple graph. Note that in this respect, a Configuration model with a Binomial distribution $\mathscr{B}(N, p/N)$ looks like an Erdös–Rényi graph with multiple-edges and self-loops.

Because of this construction, we see that in such a network, given an edge of ego u, the alter v is chosen proportionally to his/her number of half-edges (i.e. his/her degree). Thus, the following degree distribution $\mathbf{q} = (q_k, k \in \mathbb{Z}_+)$ defined as the size-biased degree distribution of $\mathbf{p}$ will play a major role in the understanding of disease dynamics on CM graphs:

$$q_k = \frac{kp_k}{\sum_{\ell \in \mathbb{Z}_+} \ell p_\ell}. \tag{1.2.1}$$

Example 1.2.6. *Particular graphs of this family include the regular graphs, where all the vertices have the same degree d (that is $p_d = 1$ and $\forall k \neq d, p_k = 0$) and the graphs whose degree distribution is a power law: for some $\alpha > 1$,*

$$p_k \overset{k\to+\infty}{\sim} k^{-\alpha}.$$

A key quantity when dealing with configuration models is the generating function of its degree distribution, defined as:

$$g(z) = \sum_{k\geq 0} z^k p_k = \mathbb{E}_{\mathbf{p}}\big(z^D\big), \tag{1.2.2}$$

where the notation in the right-hand side recalls that the random variable D has distribution $\mathbf{p}$.
In case it exists, the moment of order q of the degree distribution can be written by means of the generating function:

$$\forall q \geq 0,\ \mathbb{E}_{\mathbf{p}}\big(D^q\big) = g^{(q)}(1).$$

Example 1.2.7. *Let us recall the probability generating function of some usual parametric distributions:*

(i) For a Poisson distribution with parameter α: $g(z) = e^{\alpha(z-1)}$.

(ii) For a Geometric distribution with parameter ρ: $g(z) = \frac{\rho z}{1-z(1-\rho)}$.

(iii) For a Binomial with parameters (n, ρ): $g(z) = (z\rho + 1 - \rho)^n$.

Assumption 1.2.8. *Let us assume that $\mathbf{p} = (p_k, k \in \mathbb{Z}_+)$ admits a second order moment:*

$$m = g'(1) = \sum_{k\in\mathbb{Z}_+} kp_k, \qquad \sigma^2 = g''(1) + g'(1) - (g'(1))^2 = \sum_{k\in\mathbb{Z}_+} (k-m)^2 p_k.$$

Notice that under Assumptions 1.2.8, the size-biased degree distribution $\mathbf{q}$ defined in (1.2.1) admits a moment of order 1, which is referred to as the mean excess degree:

$$\kappa = \sum_{k\geq 0} \frac{k(k-1)p_k}{m} = \frac{\sigma^2}{m} + m - 1 = \frac{g''(1)}{g'(1)}. \tag{1.2.3}$$

The household models (see Part II of the present volume) can be built on the previous graph models. They were first analysed in detail in [12] and we also refer to Chapter **??** in Part II of this volume. They account for several levels of mixing, for instance local and global in case of 2 levels. In the latter case, the population is partitioned into clusters or households. A first possible approach is to consider a graph model on the entire population (for example a CM in [12, 13, 16]) on which the household structure is superposed independently. The links are considered stronger between individuals of the same household (for example they can transmit diseases at higher rates). Another possibility is to define the graph between individuals by taking into account the household structure, which results into clustering effects.

Definition 1.2.9 (Household models)**.** *A graph belong to the family of Household model if it is an SBM where the types are the households.*

Each household can be viewed as a vertex in a graph describing the global connections, while the intra-group connections between individuals of the same group are described by a local graph model.
How clustering affects epidemics using household models has for example been studied by [9, 40].

Let us also mention other families of random graphs: for example, the exponential random graphs, which are defined by their Radon–Nikodym densities. We refer to [32] for developments.

Definition 1.2.10 (Exponential random graph model (ERGM))**.** *A random graph belongs to the family of exponential random graphs if its distribution is of the following form. For a positive integer K, for a vector of parameters $\theta = (\theta_1, \dots \theta_K) \in \mathbb{R}^K$ and for a vector of statistics $(T_1, \dots T_K)$ of the graph, we have for any deterministic graph g:*

$$\mathbb{P}_\theta\big(G = g\big) = \exp\Big(\sum_{k=1}^{K} \theta_k T_k(g) - c(\theta)\Big).$$

The renormalizing constant $c(\theta)$ is also called partition function in statistical mechanics.

Examples of statistics T_k are the number of edges, the degrees of vertices, the number of triangles or other patterns. In Rolls et al. [101], ERGMs are for example used to estimate parameters describing the social networks of people who inject drugs in Australia. This has inspired a similar study for the French case, see [42].

1.3 Sequences of graphs

Let us consider a sequence of graphs $(\mathscr{G}_N)_{N\geq 1}$, such that for all $N \geq 1$, $\mathrm{Card}(V(\mathscr{G}_N)) = N$.

For a given graph $\mathscr{G}$ and for an integer $j \geq 1$, let us denote by $\mathscr{C}_{(j)}(\mathscr{G})$ the jth largest connected component of $\mathscr{G}$.

Definition 1.3.1 (Giant component)**.** *Consider a sequence of graphs $(\mathscr{G}_N)_{N\geq 1}$ such that for all $N \geq 1$, $\mathrm{Card}(V(\mathscr{G}_N)) = N$. If*

$$\liminf_{N\to+\infty} \frac{\mathrm{Card}\big(V(\mathscr{C}_{(1)}(\mathscr{G}_N))\big)}{N} > 0,$$

then we say that the sequence $(\mathscr{G}_N)_{N\geq 1}$ is highly connected and that the graph $\mathscr{G}_N$ admits a giant component, $\mathscr{C}_{(1)}(\mathscr{G}_N)$.

For $\mathrm{ER}(N,p)$ in the supercritical regime (with $Np > 1$), there exists a giant component [110, Theorem 4.8]. So does it for the CM, as shown by Molloy and Reed [80, 81]. The condition for the existence with positive probability of a giant component in CM graphs is that the expectation of the size biased distribution minus 1, κ, is larger than 1:

$$\kappa := \sum_{k\in\mathbb{Z}_+} (k-1)\frac{kp_k}{\sum_{\ell\in\mathbb{Z}_+}\ell p_\ell} = \mathbb{E}_{\mathbf{q}}(D-1) > 1.$$

This is connected with results on the super-criticality of Galton–Watson trees (see [50, Section 3.2 p. 75] for example). Heuristically, a CM graph looks like a tree locally, and a vertex of degree k of the graph corresponds in the tree to a node with 1 parent and $k-1$ offspring. From the construction of the CM graphs given after Definition 1.2.5, the degrees of the vertices encountered along the CM graph are given by the size-biased distribution.

If $\mathrm{Card}\big(V(\mathscr{C}_{(2)}(\mathscr{G}_N))\big) = o(N)$, then the giant component $\mathscr{C}_{(1)}(\mathscr{G}_N)$ is said to be unique. In many models such as ER, it is shown that the second largest component is of order $\log N$ (see [110, Corollary 4.13]).

The notion of being 'highly connected', as introduced in Definition 1.3.1, can also be extended.

Definition 1.3.2 (Sequence of dense graphs)**.** *We say that the graph sequence $(\mathscr{G}_N)_{N\geq 1}$ is a sequence of dense graphs if:*

$$\liminf_{N\to+\infty} \frac{\mathrm{Card}\big(E(\mathscr{G}_N)\big)}{N^2} > 0.$$

Of course, the next important notion is the notion of convergence of a sequence of graphs $(\mathscr{G}_N)_{N\geq 1}$. The topologies and notions of convergence depend on the order of the edge numbers. For graphs that are not dense, such as tree-like graphs, a large literature around the Hausdorff-Gromov topology has developed and we refer for instance to Addario-Berry et al. [2, 3]. When the graph is dense, the topology is inspired by ideas coming from the topologies of measure spaces (see Borgs et al. [26] or Lovasz and Szegedy [74]).

1.4 Definition of the SIR model on a random graph

We now describe the spread of infectious diseases on graphs. We consider a population of size N whose individuals are the vertices of a random graph $\mathscr{G}_N$. As in compartmental models, the population is partitioned into three classes that can change in time: susceptible individuals who can contract the disease (individuals of type S), infectious individuals who transmit the disease (type I) and removed individuals who were previously infectious and can not transmit the disease any more (type R). The corresponding sets of vertices, at time t, are respectively denoted by S_t, I_t and R_t, and the corresponding sizes by S_t, I_t and R_t.

On the graph $\mathscr{G}_N$, the dynamics is as follows. To each I individual is associated an exponential random clock with rate γ to determine its removal. To each edge with an infectious ego and a susceptible alter, we associate a random exponential clock with rate λ. When it rings, the edge transmits the disease and the susceptible alter becomes infectious.

Example 1.4.1 (Compartmental models)**.** *When the graph $\mathscr{G}_N = K_N$ is the complete graph, we recover the compartmental model of Part I of this volume.*

Example 1.4.2 (Household models)**.** *The above mechanisms can of course be generalized. For household models [13, 16], for example, the infection probability λ depends on whether ego and alter belong or not to the same household. See Part II of this volume.*

Notice also that for modelling real data, several studies require to take into account the dynamics of the social network itself (e.g. [52, 112]). For sexual network, for instance, accounting for the changes of sexual partners (contacts) is important (e.g. [73, 83, 104]). Also, the epidemics itself can act on the structure of the network (see [69]), such as the changes of sexual behaviour due to the spread of the AIDS epidemic (e.g. [75]). These aspects are however not treated here.

Chapter 2

The Reproduction Number R_0

We consider the early stage of the epidemics. Let us consider a single first infective of degree d_1 in a population of large size N.

For this, we proceed as in Section 1.2 of Part I of this volume and couple the process $(I_t)_{t\geq 0}$ with a branching process. As for the mixing case, it is more precisely a stochastic domination. The coupling remains exact as long as no infected or removed individual is contaminated for the second time, in which case the branching process creates an extra individual, who is named 'ghost'.

Definition 2.0.1 (R_0)**.** *The basic reproduction number of the epidemic, denoted by R_0, is the mean offspring number of the branching process approximating the infectious population in early stages.*
If we denote by $\beta(t)$ the birth rate at time $t>0$ in this branching process, then:

$$R_0 = \int_0^\infty \beta(t)dt. \tag{2.0.1}$$

Notice that in the above definition, the measure $\beta(t)dt$ represents the intensity measure of the point process describing the occurrence of new infections due to a chosen infective (e.g. [65]).

A large literature is devoted to this indicator R_0 and extensions. Recall indeed that the nature and importance of the disease is usually classified according to whether $R_0 > 1$ or $R_0 \leq 1$.
When $R_0 > 1$, the branching process is super-critical and with positive probability its size is infinite, in which case we say that there is a major outbreak of the disease. The probability for this to happen can be computed [48, Eq. 3.10] and is less than 1. When the branching process does not get extinct, its size grows roughly proportional to $e^{\alpha t}$, where α is termed the (initial) epidemic growth rate (see [65]). In this case, the positive constant α depends on the parameters of the model through the equation

$$1 = \int_0^\infty e^{-\alpha t}\beta(t)dt. \tag{2.0.2}$$

When $R_0 \leq 1$, the branching process is critical or subcritical and its size is almost surely finite. Then, the total number of individuals who have been infected when the epidemic stops (at the time t when $I_t = 0$) is upper bounded by an almost surely finite random variable with distribution independent of the total population size N, and we talk of a small epidemic. We refer to [7, 108] for reviews.

2.1 Homogeneous mixing

In the case where $\mathscr{G}_N = K_N$ is the complete graph, as stated in Part I of this volume, many results for epidemics in large homogeneous mixing populations can be obtained since the initial phase of the epidemic is well approximated by a branching process (see e.g. [11]).

Proposition 2.1.1 (R_0 for homogeneous mixing)**.** *The reproduction number is given by:*

$$R_0 = \frac{\lambda}{\gamma}.$$

In the case where $\lambda > \gamma$, then $\alpha = \lambda - \gamma$ and

$$R_0 = \frac{\lambda}{\gamma} = 1 + \frac{\alpha}{\gamma}.$$

Notice that the second expression of R_0 does not depend on λ, which is sometimes complicated to estimate, especially at the beginning of an epidemic, but only on the removal rate γ, that is usually documented, and on the Malthusian parameter α, that can be estimated from the dynamics of the emerging epidemics.

Proof. The reproduction number R_0 for the homogeneous mixing case has already been studied in Part I of this volume, but let us give here another proof of the proposition using (2.0.1). In this case, $\beta(t) = \lambda e^{-\gamma t}$. This can be understood by observing that λ is the rate at which an infected individual makes contacts if he or she is still infectious, while $e^{-\gamma t}$ is the probability that the individual is still infectious t time units after he or she became infected. Then, (2.0.2) and (2.0.1) translate to

$$1 = \frac{\lambda}{\gamma+\alpha} \qquad \text{and} \qquad R_0 = \frac{\lambda}{\gamma} = 1 + \frac{\alpha}{\gamma}. \tag{2.1.1}$$

This completes the proof. □

2.2 Configuration model

Assume that $\mathscr{G}_N$ is a configuration model graph whose degree distribution $\mathbf{p}$ admits a mean μ and a variance σ^2. Recall also the definition of the size-biased distribution $\mathbf{q}$ in (1.2.1), and of the mean excess degree κ in (1.2.3). The mean excess degree κ, is in the context of SIR epidemics spreading on graphs, the mean number of susceptibles that are contaminated by a typical infective (other than his or her own infector).

Let us consider the following continuous time birth-death process $(X_t)_{t\geq 0}$. Individuals live during exponential independent times with expectation $1/\gamma$. To each individual is associated a maximal number of offspring $k-1$, where k (the 'degree' of the individual) is drawn in the size-biased distribution $\mathbf{q}$. We associate to such an individual $k-1$ independent exponential random variables with expectations $1/\lambda$. The ages at which the individual gives birth are the exponential random variables that are smaller than the lifetime of the individual. There is an intuitive coupling between $(X_t)_{t\geq 0}$ and $(I_t)_{t\geq 0}$ such as $X_t \geq I_t$ for every t, with the equality as long as no 'ghost' has appeared. We can associate with the process $(X_t)_{t\geq 0}$ its discrete-time skeleton (time counting the generations) that is a Galton–Watson process $(Z_n)_{n\geq 0}$ ($Z_0 = 1$). Conditionally on the degree k and the fact that the chosen individual remains infectious for a duration y, the number of contacts contaminated by this individual follows a binomial distribution with parameters $k-1$ and $1-e^{-\lambda y}$. Summing over k and integrating with respect to y, we can write the probability that in this Galton–Watson process an individual of generation $n \geq 1$ has $\nu = \ell$ offspring:

$$\mathbb{P}(\nu = \ell) = \sum_{k=\ell+1}^{+\infty} \frac{kp_k}{m} \binom{k-1}{\ell} \Big(\frac{\lambda}{\lambda+\gamma}\Big)^{\ell} \Big(\frac{\gamma}{\lambda+\gamma}\Big)^{k-1-\ell}.$$

Proposition 2.2.1 (R_0 for CM)**.** *Recall the definition of the mean excess degree κ in* (1.2.3)*. We have:*

$$R_0 = \frac{\kappa\lambda}{\lambda+\gamma}. \tag{2.2.1}$$

In the super-critical case, R_0 can also be rewritten as

$$R_0 = \frac{\gamma+\alpha}{\gamma+\alpha/\kappa} = 1 + \frac{\alpha}{\lambda+\gamma}.$$

Proof. With the description of the process $(Z_n)_{n\geq 1}$:

$$\begin{aligned} R_0 &= \sum_{k\geq 0} \frac{kp_k}{m} \int_0^{+\infty} (k-1)(1-e^{-\lambda y})\ \gamma e^{-\gamma y} dy \\ &= \sum_{k\geq 0} (k-1)\frac{kp_k}{\mu}\frac{\lambda}{\lambda+\gamma} \\ &= \Big(\frac{g''(1)}{g'(1)} - 1\Big)\frac{\lambda}{\lambda+\gamma} \\ &= \frac{\kappa\lambda}{\lambda+\gamma}. \end{aligned}$$

We obtain

$$\beta(t) = \kappa\lambda e^{-(\lambda+\gamma)t}.$$

This can be seen by noting that κ is the expected number of susceptible acquaintances a typical newly infected individual has in the early stages of the epidemic, while $e^{-\lambda t}$ is the probability that a given susceptible individual is not contacted by the infective over a period of t time units, and $e^{-\gamma t}$ is the probability that the infectious individual is still infectious t time units after he or she became infected. From (2.0.2), we obtain that

$$\alpha = \kappa\lambda - \lambda - \gamma,$$

from which we conclude the proof. □

Example 2.2.2. *Let us compute R_0 for particular choices of degree distribution* **p***:*
(i) For a Poisson distribution with parameter $a > 0$,

$$R_0 = \frac{a\lambda}{\lambda+\gamma}.$$

Thus, $R_0 > 1$ if and only if $a > 1 + \gamma/\lambda$.
(ii) For a Geometric distribution with parameter $a \in (0,1)$, $R_0 = \frac{\lambda}{\lambda+\gamma}\frac{2(1-a)}{a}$. Thus, $R_0 > 1$ if and only if $a < 2\lambda/(3\lambda+\gamma)$. □

We can now connect the considerations on the skeleton with the epidemic in continuous time.

Proposition 2.2.3. *Let us consider the continuous time birth-death process $(X_t)_{t\geq 0}$.*

(i) If $R_0 \leq 1$, the process $(X_t)_{t\geq 0}$ dies out almost surely.

(ii) If $R_0 > 1$, the process $(X_t)_{t\geq 0}$ dies with a probability $z \in (0,1)$ that is the smallest solution of

$$z = \frac{\gamma}{g'(1)} \int_{\mathbb{R}_+} g'\big(z + e^{-\lambda y}(1-z)\big) e^{-\gamma y} dy. \tag{2.2.2}$$

(iii) Let us define the times $\tau_0 = \inf\{t \geq 0 \mid X_t = 0\}$ and $\tau_{\varepsilon n} = \inf\{t \geq 0 \mid X_t \geq \varepsilon n\}$. If $R_0 > 1$, then for all sequence $(t_n)_{n\in\mathbb{Z}_+}$ such that $\lim_{n\to+\infty} t_n/\log(n) = +\infty$,

$$\lim_{n\to+\infty} \mathbb{P}(\tau_0 \leq t_n \wedge \tau_{\varepsilon n}) = z \tag{2.2.3}$$

$$\lim_{n\to+\infty} \mathbb{P}(\tau_{\varepsilon n} \leq t_n \wedge \tau_0) = 1 - z. \tag{2.2.4}$$

Proof. Points (i) and (ii) are consequences of Proposition 2.2.1 and the connections between the discrete time Galton–Watson tree and the continuous time birth-death process $(X_t)_{t\geq 0}$ that is coupled with $(I_t)_{t\geq 0}$ as long as no ghost has appeared.

The proof of (iii) is an adaptation of Lemma A.1 in Méléard and Tran [78] (see also [30, 107]). Heuristically, (iii) says that at the beginning of the epidemics, the population either gets extinct with probability z or, with probability $1-z$, reaches the size εn before time t_n and before extinction. The time t_n should be thought of as of order $\log(n)$, since the supercritical process has an exponential growth when it does not go to extinction.
For the birth-death process $(X_t)_{t\geq 0}$ there is no accumulation of birth and death events and almost surely,

$$\lim_{n\to+\infty} t_n \wedge \tau_{\varepsilon n} = +\infty.$$

So, we have by dominated convergence that $\lim_{n\to+\infty}\mathbb{P}(\tau_0 \leq t_n \wedge \tau_{\varepsilon n}) = \mathbb{P}(\tau_0 < +\infty)$. This last probability is the extinction probability of the process $(X_t)_{t\geq 0}$ which solves (2.2.2). For the second limit, we have:

$$\mathbb{P}(\tau_{\varepsilon n} \leq t_n \leq \tau_0) = \mathbb{P}(\tau_{\varepsilon n} \leq t_n \text{ and } \tau_0 = +\infty) + \mathbb{P}(\tau_{\varepsilon n} \leq t_n \leq \tau_0 < +\infty). \tag{2.2.5}$$

The second term of (2.2.5) is upper bounded by $\mathbb{P}(t_n \leq \tau_0 < +\infty)$ which converges to 0 by dominated convergence when $n \to +\infty$. For the second term, we can prove that with martingale techniques (e.g. [65]) that:

$$\lim_{t\to+\infty} \frac{\log X_t}{t} = \alpha, \tag{2.2.6}$$

where α is the initial epidemic growth rate defined in (2.0.2) and that is positive when $R_0 > 1$.

Let us consider $n > 1/\varepsilon$, so that $\log(\varepsilon n) > 0$. Since $\lim_{n\to+\infty}\tau_{\varepsilon n} = +\infty$ almost surely, we have on $\{\tau_0 = +\infty\}$ that:

$$\lim_{n\to+\infty} \frac{\log(\varepsilon n)}{\tau_{\varepsilon n}} \geq \lim_{n\to+\infty} \frac{\log(X_{\tau_{\varepsilon n}-})}{\tau_{\varepsilon n}} = \alpha > 0.$$

We deduce that:

$$\begin{aligned}\lim_{n\to+\infty} \mathbb{P}(\tau_{\varepsilon n} \leq t_n,\ \tau_0 = +\infty) &= \lim_{n\to+\infty} \mathbb{P}\big(\frac{\tau_{\varepsilon n}}{\log(\varepsilon n)} \leq \frac{t_n}{\log(\varepsilon n)},\ \tau_0 = +\infty\big)\\ &= \mathbb{P}(\tau_0 = +\infty) = 1 - z,\end{aligned}$$

since by our choice of t_n, $\lim_{n\to+\infty} t_n/\log(\varepsilon n) = +\infty$. □

Using similar results and fine couplings with branching properties, Barbour and Reinert [19] approximate the epidemic curve from the initial stages to the extinction of the disease.

2.3 Stochastic block models

We assume that there are K types of individuals, labeled $\{1,2,\cdots,K\}$ and that for $k = 1,\cdots,K$ a fraction η_k of the N individuals in the population is of type k. We assume that the infection rate from an ego of type i to an alter of type j is λ_{ij}/N.

Proposition 2.3.1 (R_0 for SBM)**.** *Consider a SBM as in Definition 1.2.3. Denote by ρ be the largest eigenvalue of the matrix with elements $\lambda_{ij}\rho_j$. Then:*

$$R_0 = \frac{\rho}{\gamma} = 1 + \frac{\alpha}{\gamma}.$$

Proof. We can hence couple here the infection process with a multi-type branching process. The rate at which a given i individual gives birth to a j individual corresponds to the rate, in the epidemic process, at which an i individual infects j individuals at time t since infection: it is $a_{ij}(t) = \lambda_{ij}\rho_j e^{-\gamma t}$. Here, λ_{ij}/N is the rate at which the i individual contacts a given j individual, $N\rho_j$ is the number of j individuals and $e^{-\gamma t}$ is the probability that

the i individual is still infectious t time units after being infected. For multi-type branching processes, it is well known (e.g. [10, 47, 48]) that the basic reproduction number $R_0 = \rho_M$ is the largest eigenvalue of the matrix M with elements $m_{ij} = \int_0^{+\infty} a_{ij}(t)dt$, and the epidemic growth rate α is such that $1 = \int_0^{\infty} e^{-\alpha t}\rho_{A(t)}dt$, where $\rho_{A(t)}$ is the largest eigenvalue of the matrix $A(t)$ with elements $a_{ij}(t)$. Note that $\rho_{A(t)} = \rho e^{-\gamma t}$. Therefore,

$$R_0 = \rho \int_0^{\infty} e^{-\gamma t}dt = \frac{\rho}{\gamma}$$

and

$$1 = \rho \int_0^{+\infty} e^{-(\alpha+\gamma)t}dt \quad \text{leading to} \quad \rho = \alpha + \gamma.$$

These equalities imply that

$$R_0 = 1 + \frac{\alpha}{\gamma},$$

which shows that the relation between R_0 and α for a multi-type Markov SIR epidemic is the same as for such an epidemic in a homogeneous mixing population (cf. equation (2.1.1)). □

2.4 Household structure

It is possible to define several different measures for the reproduction numbers for household models [14, 15, 23, 58]. For this model it is hard to find explicit expressions for R_0. We refer to Part II of this volume, for discussion on the early stages of the an epidemic spreading on a household graph or on a two-level mixing graph.

2.5 Statistical estimation of R_0 for SIR on graphs

Since we often have observations on symptom onset dates of cases for a new, emerging epidemic, as was the case for the Ebola epidemic in West Africa, it is often possible to estimate α from observations. In addition, we often have observations on the typical duration between time of infection of a case and infection of its infector, which allow us to estimate, assuming a Markov SIR model, the average duration of the infectious period, $1/\gamma$ [113].

In [108], it is shown that estimates of R_0 obtained by assuming homogeneous mixing are always larger than the corresponding estimates if the contact structure follows the configuration network model. For virtually all standard models studied in the literature, assuming homogeneous mixing leads to conservative estimates.

2.6 Control effort

Definition 2.6.1. *The control effort v_c is defined as the proportion of infected individuals that we should prevent from spreading the disease and immunize to stop the outbreak (have $R_0 < 1$), the immunized people being chosen uniformly at random.*

For the homogeneous mixing contact structure, the required control effort for epidemics on the network structures under consideration, is known to depend solely on R_0 through equation [28, p. 69]

Proposition 2.6.2. *On the complete graph K_N, we have that:*

$$v_c = 1 - \frac{1}{R_0} = \frac{\alpha}{\alpha + \gamma}. \tag{2.6.1}$$

Proof. Consider a given infectious non-immunized individual whose infectious period is of length $y > 0$. In case we immunize a fraction v_c of the infected individuals, the number of new infectious and non-immunized individuals contaminated by this individual is not a Poisson random variable with parameter λy, but a thinned Poisson random variable of parameter $\lambda(1-v_c)y$. The condition that the new $R_0 = \lambda(1-v_c)/\gamma$ is less than 1 provides the expression of v_c announced in the proposition. □

Notice that if we estimate the initial epidemic growth rate α and the mean duration of the infectious period $1/\gamma$ from the data, (2.6.1) allows us to propose a natural estimator of v_c.

For CM graphs, we can establish a similar formula for v_c that depends also on the mean excess degree κ:

Proposition 2.6.3 (v_c for CM graphs)**.** *For a CM graph with degree distribution* **p** *and mean excess degree* κ*:*

$$v_c = \frac{\kappa-1}{\kappa}\frac{\alpha}{\alpha+\gamma}.$$

The results obtained for Markov SIR epidemics in the complete graph model, CM and SBM are summarized in Table 2.6.1. The results from household models are not in the table, since the expressions are hardly insightful. These results are taken from [108].

Model	Quantity of interest	Quantity of interest as function of λ, γ and κ	α, γ and κ	Ratio with complete graph
Complete graph	α	$\lambda-\gamma$	-	-
	R_0	$\frac{\lambda}{\gamma}$	$1+\frac{\alpha}{\gamma}$	-
	v_c	$\frac{\lambda-\gamma}{\lambda}$	$\frac{\alpha}{\alpha+\gamma}$	-
CM	α	$(\kappa-1)\lambda-\gamma$	-	-
	R_0	$\frac{\kappa\lambda}{\lambda+\gamma}$	$\frac{\gamma+\alpha}{\gamma+\alpha/\kappa}$	$1+\frac{\alpha}{\gamma\kappa}$
	v_c	$1-\frac{\lambda+\gamma}{\kappa\lambda}$	$\frac{\kappa-1}{\kappa}\frac{\alpha}{\alpha+\gamma}$	$1+\frac{1}{\kappa-1}$
SBM	α	$\gamma(\rho_M-1)$	-	-
	R_0	ρ_M	$1+\frac{\alpha}{\gamma}$	1
	v_c	$1-\frac{1}{\rho_M}$	$\frac{\alpha}{\alpha+\gamma}$	1

Table 2.6.1: *The epidemic growth rate* α*, the basic reproduction number* R_0 *and required control effort* v_c *for a Markov SIR epidemic model as function of model parameters in the complete graph* K_N*, in the CM and in the SBM. In the fourth column, the ratio has been made between the* R_0 *in the CM and SBM cases (numerators) and the* R_0 *obtained in mixing populations (complete graphs) given the estimations of* α*,* γ *and* κ*.*

Let us comment on these results. First, we find that the estimator of R_0 obtained assuming homogeneous mixing (complete graph) overestimates by a factor $1+\frac{1}{\kappa-1}$ the R_0 in configuration models. This factor is always strictly greater than 1, since the mean excess degree κ is strictly greater than 1. Thus, v_c obtained by assuming homogeneous mixing is always larger than that of the configuration model. Consequently, if the actual infectious contact structure is made up of a CM and a perfect vaccine is available, we need to vaccinate a smaller proportion of the population than predicted assuming homogeneous mixing.
The overestimation of R_0 is small whenever R_0 is not much larger than 1 or when κ is large. The same conclusion applies to the required control effort v_c. The observation that the R_0 and v_c for the homogeneous mixing model exceed the corresponding values for the network model extends to the full epidemic model allowing for an arbitrarily distributed latent period followed by an arbitrarily distributed independent infectious period, during which the infectivity profile (the rate of close contacts) may vary over time but depends only on the time since the start of the infectious period. Figure 2.6.1(a) shows that for SIR epidemics with Gamma distributed infectious periods, the factor by which the homogeneous mixing estimator overestimates the actual R_0 increases with increasing epidemic growth rate α, and suggests that this factor increases with increasing standard deviation of the infectious period. Figure 2.6.1(b) shows that the factors by which the homogeneous mixing estimator overestimates the actual v_c, decreases with increasing α and increases with increasing standard deviation of the infectious period. When the standard deviation of the infectious period is low, which is a realistic assumption for most emerging infectious diseases (see e.g. [39]), and R_0 is not much larger than 1, then ignoring the contact structure in the network model and using the simpler estimators for the homogeneous mixing results in a slight overestimation of R_0 and v_c.

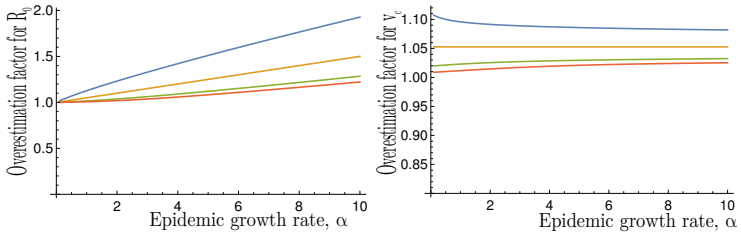


Figure 2.6.1: *The factor by which estimators based on homogeneous mixing will overestimate (a) the basic reproduction number R_0 and (b) the required control effort v_c for the network case. Here the epidemic growth rate α is measured in multiples of the mean infectious period $1/\gamma$. The mean excess degree $\kappa = 20$. The infectious periods are assumed to follow a gamma distribution with mean 1 and standard deviation $\sigma = 1.5$, $\sigma = 1$, $\sigma = 1/2$ and $\sigma = 0$, as displayed from top to bottom. Note that the estimate of R_0 based on homogeneous mixing is $1+\alpha$. Furthermore, note that $\sigma = 1$, corresponds to the special case of an exponentially distributed infectious period, while if $\sigma = 0$, the duration of the infectious period is not random.*

When considering epidemics spreading on SBM graphs (see [108, Supplementary materials]), we can derive that estimators for R_0 and (if control measures are independent of the types of individuals) v_c are exactly the same as for homogeneous mixing in a broad class of SEIR epidemic models. This class includes the full epidemic model allowing for arbitrarily distributed latent and infectious periods and models in which the rates of contacts between different types keep the same proportion all of the time, although the rates themselves may vary over time (cf. [49]).
We illustrate our findings on multitype structures through simulations of SEIR epidemics in an age stratified population with known contact structure as described in [114]. We use values of the average infectious period $1/\gamma$ and the average latent period $1/\delta$ close to the estimates for the 2014 Ebola epidemic in West Africa [116].
Two estimators for R_0 are computed. The first of these estimators is based on the average number of infections among the people who were infected early in the epidemic. This procedure leads to a very good estimate of R_0 if the spread of the disease is observed completely. The second estimator for R_0 is based on $\hat{\alpha}$, an estimate of the epidemic growth rate α, and known expected infectious period $1/\gamma$ and expected latent period $1/\delta$. This estimator of R_0 is $(1+\hat{\alpha}/\delta)(1+\hat{\alpha}/\gamma)$. We calculate estimates of R_0 using these two estimators for 250 simulation runs. As predicted by the theory, the simulation results show that for each run the estimates are close to the actual value (Figure 2.6.2(a)), without a systematic bias (Figure 2.6.2(b)).

Let us now consider an epidemic spreading on a household structure. It is also argued that the required control effort satisfies $v_c \geq 1 - 1/R_0$ for this model, which implies that if we know R_0 and we base our control effort on this knowledge, we might fail to stop an outbreak. However, we usually do not have direct estimates for R_0 and even though it is not true in general that using R_0 leads to conservative estimates for v_c [17], numerical computations suggest that the approximation of v_c using α and the homogeneous mixing assumption is often conservative.

To illustrate this last point, we consider in Figure 2.6.3 a household structure with within and global infectivities. The within household infection rate is λ_H. In the simulations, we show estimates for R_0 and v_c over a range of values for the relative contribution of the within-household spread. For each epidemic growth rate α, the estimated values remain below the value obtained for homogeneous mixing (neglecting the partition into households).
We use two types of epidemics: in (a) and (b) the Markov SIR epidemic is used, while in (c) the so-called Reed–Frost model is used, which can be interpreted as an epidemic in which infectious individuals have a long latent period of non-random length, after which they are infectious for a very short period of time. We note that for the Reed–Frost model the relationship between α and R_0 does not depend on the household structure (cf. [17]) and therefore, for this model, only the dependence of v_c on the relative contribution of the within household spread is shown in Figure 2.6.3.
The household size distribution is taken from a 2003 health survey in Nigeria [46]. For Markov SIR epidemics, as the within-household infection rate λ_H is varied, the global infection rate is varied in such a way that the computed epidemic growth rate α is kept fixed. For this model, α is calculated using the matrix method described in Section

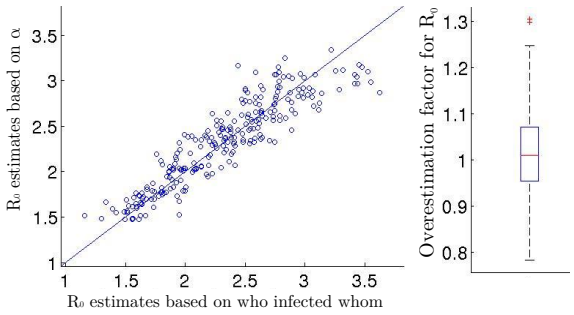


Figure 2.6.2: *The estimated basic reproduction number, R_0, for a Markov SEIR model in a multi-type population as described in [114], based on the real infection process (who infected whom) plotted against the computed R_0, assuming homogeneous mixing, based on the estimated epidemic growth rate, α, and given expected infectious period (5 days) and expected latent period (10 days). The infectivity is chosen at random, such that the theoretical R_0 is uniform between 1.5 and 3. The estimate of α is based on the times when individuals become infectious. In the right plot, a boxplot of the ratios is given.*

4.1 of [95].
For the Reed–Frost epidemic model, the probability that an infectious individual infects a given susceptible household member during its infectious period, p_H is varied, while the corresponding probability for individuals in the general population varies with p_H so that α is kept constant. For this model, R_0 coincides with the initial geometric rate of growth of infection, so $\alpha = \log(R_0)$. From Figure 2.6.3, we see that estimates of v_c assuming homogeneous mixing are reliable for Reed–Frost type epidemics, although as opposed to all other analysed models and structures, the estimates are not conservative. We see also that for the Markov SIR epidemic, estimating R_0 and v_c based on the homogeneously mixing assumption might lead to conservative estimates which are up to 40% higher than the real R_0 and v_c.

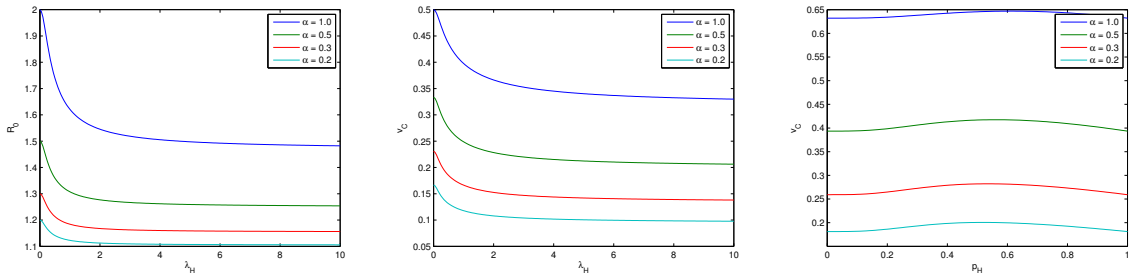


Figure 2.6.3: *Estimation of key epidemiological variables in a population structured by households (see Part II of this volume). The basic reproduction number R_0 for Markov SIR epidemics (a), critical vaccination coverage v_c for Markov SIR epidemics (b) and v_c for Reed–Frost epidemics (c), as a function of the relative influence of within household transmission, in a population partitioned into households. The household size distribution is given by $m_1 = 0.117, m_2 = 0.120, m_3 = 0.141, m_4 = 0.132, m_5 = 0.121, m_6 = 0.108, m_7 = 0.084, m_8 = 0.051, m_9 = 0.126$, for $i = 1, 2, \cdots, 9$, m_i is the fraction of the households with size i. The global infectivity is chosen so that the epidemic growth rate α is kept constant while the within household transmission varies. Homogeneous mixing corresponds to $\lambda_H = p_H = 0$.*

Chapter 3

SIR Epidemics on Configuration Model Graphs

We now turn to establishing limit theorems for approximating the dynamics of the disease in large populations, when $N \to +\infty$, similarly to Chapter **??** in Part I of this book. We focus here on the case where $\mathscr{G}_N$ is a Configuration model graph, and we will let $N \to +\infty$. Several strategies have been developed for epidemics spreading on such random graphs (see e.g. Newman [87, 89], Durrett [50], Barthélemy et al. [20], Kiss et al. [71]).

Contrarily to the classical mixing compartmental SIR epidemic models (e.g. [68, 21] see also Part I of this book for a presentation), heterogeneity in the number of contacts makes it difficult to describe the dynamical behaviour of the epidemic. An important literature, starting from Andersson [7], deals with moment closure, mean field approximations (e.g. [92, 20, 50, 71]) or large population approximations (e.g. [13], see also Eq. (3) of [6] in discrete time). In 2008, Ball and Neal [13] proposed to describe the dynamics with an infinite system of ordinary differential equations, by obtaining an equation for each subpopulation of individuals with same degree k, $k \in \mathbb{Z}_+$. The same year, Volz [111] proposed a large population approximation with only 5 ordinary differential equations and without moment closure, which was a major advance for prediction and tractability. The key concept behind his work was to focus not only on node-based quantities, but rather of edge-based ones (see also [79]). Rigorous proofs have then been proposed by [45, 19, 66]).

Recall that we have denoted the sets of S, I and R vertices at time t by S_t, I_t and R_t (see Section 1.4). The sizes of these sub-populations are S_t, I_t and R_t. We will say that an edge linking an infectious ego and susceptible alter is of type $\mathrm{I}-\mathrm{S}$ (accordingly $\mathrm{R}-\mathrm{S}$, $\mathrm{I}-\mathrm{I}$ or $\mathrm{I}-\mathrm{R}$).

3.1 Moment closure in large populations

For the presentation in this section, we follow the work of [7]. Let us introduce some notation. For $u \in V$, denote

$$S_u(t) = \mathbf{1}_{u \in \mathrm{S}_t} \qquad \text{and} \qquad I_u(t) = \mathbf{1}_{u \in \mathrm{I}_t}.$$

Then, $S_t = \sum_{u \in V} S_u(t)$ and $I_t = \sum_{u \in V} I_u(t)$. Because the size N of the graph $\mathscr{G}_N$ converges to infinity, we will be lead to study the proportions of susceptible, infectious and removed individuals, that are denoted by:

$$S_t^N = \frac{S_t}{N}, \qquad I_t^N = \frac{I_t}{N}, \qquad R_t^N = \frac{R_t}{N}. \tag{3.1.1}$$

Notice that $S_t^N + I_t^N + R_t^N = 1$ since our population is closed. Hence, knowing the evolution of $S_\cdot^N$ and $I_\cdot^N$ is sufficient for describing the size and evolution of the outbreak.
For A, B, C being S or I, we denote by

$$[a] = \lim_{N \to +\infty} \frac{1}{N} \sum_{u \in V} A_u = a, \qquad [ab] = \lim_{N \to +\infty} \frac{1}{N} \sum_{u,v \in V} A_u G_{uv} B_v,$$

$$[abc] = \lim_{N\to+\infty} \frac{1}{N} \sum_{u,v,w\in V} A_u G_{uv} B_v G_{vw} C_w,$$

where we recall that G is the adjacency matrix of the graph (see Definition 1.1.2).

In the sequel, we will work under the following assumption.

Assumption 3.1.1. *We assume that* $\lim_{N\to+\infty}(S_0^N, I_0^N) = (s_0, i_0) \in (\mathbb{R}_+ \setminus \{0\})^2$ *and that for all N,* $R_0^N = 0$.

The idea is that in the large population limit, the initial fraction of infectious individuals should be positive to allow the observation of an outbreak. That is why we assume that it is of order $i_0 N$ with $i_0 > 0$ but possibly small with respect to 1.

Let us present a system of limiting deterministic equations. The limit theorems allowing to obtain the following equations from the finite stochastic system are not shown here. In fact, we will later detail how Volz' equations are obtained.

Andersson [7] proposes the following ODEs for the sizes of the S and I classes.

$$\frac{ds_t}{dt} = -\lambda [s_t i_t], \qquad \frac{di_t}{dt} = \lambda [s_t i_t] - \gamma i_t. \tag{3.1.2}$$

Let us comment on these equations. In a closed population, susceptible individuals disappear when they are contaminated, i.e. when an edge with susceptible ego and infectious alter transmits the disease. Thus, the rate at which the number of susceptible individuals decreases due to infection (which equals to the rate at which the number of infectious individuals increases) should be proportional to the proportion of edges with susceptible ego and infectious alter, $[s_t i_t]$. The rate at which infectious individuals disappear is $-\gamma i_t$ as in the compartmental case, since removals are node-related events and not edge-related events like infections.

Equations 3.1.2 are not closed, and this leads Andersson to propose the following assumption.

Assumption 3.1.2. *Let A, B, C be S or I. If* $\{u,w\} \notin E$, *we assume that*

$$\mathbb{P}(A_u = 1 \mid B_v C_w = 1) = \mathbb{P}(A_u = 1 \mid B_v = 1) = \frac{\mathbb{P}(A_u = 1,\, B_v = 1)}{\mathbb{P}(B_v = 1)}.$$

Let us comment on this assumption. As the Bayes formula says that:

$$\mathbb{P}(A_u B_v C_w = 1) = \mathbb{P}(A_u = 1 \mid B_v C_w = 1)\mathbb{P}(B_v C_w = 1),$$

Assumption (3.1.2) implies that

$$\mathbb{P}(A_u C_w = 1 \mid B_v = 1) = \mathbb{P}(A_u = 1 \mid B_v = 1)\mathbb{P}(C_w = 1 \mid B_v = 1).$$

Thus, Assumption 3.1.2 amounts to assuming that conditionally on having a B friend, having an A and a C friends are independent events, and is heuristically true when

$$[abc] \approx \frac{[ab][bc]}{[b]}.$$

This assumption fails when we are in graphs with strong correlations between edges so that 'the friend of my friend is also my friend'.

Let us define the selection pressure by

$$\widetilde{i}_t = \frac{[s_t i_t]}{s_t}. \tag{3.1.3}$$

It is the mean number of edges toward I_t for individuals in S_t. This quantity allows Andersson [7] to close the system of ODEs (3.1.2) under Assumption 3.1.2.

Theorem 3.1.3. *Under Assumption 3.1.2, the epidemic on the network can be described by the following equations:*

$$\frac{ds_t}{dt} = -\lambda s_t \widetilde{i}_t, \tag{3.1.4}$$

$$\frac{di_t}{dt} = \lambda s_t \widetilde{i}_t - \gamma i_t \tag{3.1.5}$$

$$\frac{d\widetilde{i}_t}{dt} = \big(C\lambda s_t - \lambda - \gamma\big)\widetilde{i}_t. \tag{3.1.6}$$

Proof. The equations proposed in Theorem 3.1.3 are derived in several steps. Recall Equations (3.1.2). To close them, it is needed to describe how the quantities of edges $[s_t s_t]$ and $[s_t i_t]$ evolve. An edge $\mathrm{S}-\mathrm{S}$ disappears when one of its vertices is infected. For each motif $\mathrm{S}-\mathrm{S}-\mathrm{I}$, the edge $\mathrm{S}-\mathrm{I}$ transmits the disease independently with rate λ. Thus, the rate of disappearance of $\mathrm{S}-\mathrm{S}$ edges is proportional to the $\lambda[s_t s_t i_t]$.
Similarly, $\mathrm{S}-\mathrm{I}$ edges appear when edges $\mathrm{S}-\mathrm{S}$ become $\mathrm{S}-\mathrm{I}$, and disappear when becoming $\mathrm{I}-\mathrm{I}$ (which happens when the susceptible vertex is infected by its infectious alter, or by another infectious contact) or when becoming $\mathrm{S}-\mathrm{R}$ (when the infectious individual is removed). Then:

$$\begin{aligned}
\frac{d[s_t s_t]}{dt} &= -2\lambda[s_t s_t i_t], \\
\frac{d[s_t i_t]}{dt} &= \lambda\big([s_t s_t i_t] - [i_t s_t i_t] - [s_t i_t]\big) - \gamma[s_t i_t].
\end{aligned} \tag{3.1.7}$$

These equations are still not closed, as they depend on the numbers of motifs $\mathrm{S}-\mathrm{S}-\mathrm{I}$ and $\mathrm{I}-\mathrm{S}-\mathrm{I}$ renormalized by N. The equations that we might write for these quantities depend on motifs with four vertices etc. To close the equations, we use Assumption 3.1.2. Then, the equations (3.1.7) become:

$$\begin{aligned}
\frac{d[s_t s_t]}{dt} &= -2\lambda\frac{[s_t s_t][s_t i_t]}{s_t}, \\
\frac{d[s_t i_t]}{dt} &= \lambda\Big(\frac{[s_t s_t][s_t i_t]}{s_t} - \frac{[s_t i_t]^2}{s_t} - [s_t i_t]\Big) - \gamma[s_t i_t].
\end{aligned}$$

Notice that

$$\frac{d(s_t^2)}{dt} = 2s_t\frac{ds_t}{dt} = -2\lambda s_t[s_t i_t] = -2\lambda\frac{[s_t i_t]}{s_t}s_t^2.$$

Thus, (s_t^2) and $[s_t s_t]$ satisfy the same ODE and we deduce that there exists a $C>0$ such that $[s_t s_t] = Cs_t^2$.

Using the definition of the selection pressure $\widetilde{i}_t$,

$$\begin{aligned}
\frac{d\widetilde{i}_t}{dt} =& \frac{d[s_t i_t]}{dt}\frac{1}{s_t} - \frac{[s_t i_t]}{s_t^2}\frac{ds_t}{dt} \\
=& \frac{1}{s_t}\Big(\lambda\big(Cs_t^2 \times \widetilde{i}_t s_t \times \frac{1}{s_t} - \widetilde{i}_t^2 s_t^2 \times \frac{1}{s_t} - \widetilde{i}_t s_t\big) - \gamma\widetilde{i}_t s_t\Big) + \frac{\widetilde{i}_t s_t}{s_t^2} \times \lambda\widetilde{i}_t s_t \\
=& \big(C\lambda s_t - \lambda - \gamma\big)\widetilde{i}_t.
\end{aligned}$$

The system can then be reformulated as the announced system with three ODEs in s_t, i_t and $\widetilde{i}_t$. □

When the infection rate is low and the number of $\mathrm{S}-\mathrm{S}$ edges is very high, we recover the Kermack–McKendrick ODEs describing the dynamics of an epidemic in a homogeneous case:

Proposition 3.1.4. *If $C \to +\infty$ and $\lambda \to 0$ with $\lambda' = C\lambda$ constant, we recover in the limit the Kermack–McKendrick system of ODE:*

$$\begin{aligned}
\frac{ds_t}{dt} &= -\lambda' s_t i_t \\
\frac{di_t}{dt} &= \lambda' s_t i_t - \gamma i_t.
\end{aligned}$$

Proof. If $C \to +\infty$ and $\lambda \to 0$ with $\lambda' = C\lambda$ constant, then 'in the limit':

$$\frac{d\widetilde{i}_t}{dt} = \lambda' \, s_t \widetilde{i}_t - \gamma \widetilde{i}_t.$$

Consider $f(t) = \widetilde{i}_t - C i_t$. This quantity satisfies

$$\frac{df}{dt}(t) = -\gamma f_t.$$

Applying Gronwall's inequality, this yields that $\widetilde{i}_t = C i_t$. We recover as announced, the Kermack–McKendrick ODEs with infection rate λ'. □

From Equation (3.1.6), we can for example predict the total size of the epidemics, i.e. the number of removed individuals when the infective population vanishes and the epidemics stops.

Proposition 3.1.5. *Based on the equations* (3.1.6)*, we can compute the final size of the epidemics:*

$$z := s_0 - s_\infty = s_0 \, \big(1 - \exp\big(-\frac{\lambda}{\lambda+\gamma}(Cz + \widetilde{i}_0)\big)\big).$$

Proof. Because $t \mapsto s_t$ is a continuous non-negative decreasing function, it converges to a limit s_∞ when $t \to +\infty$. From (3.1.6):

$$\frac{d\widetilde{i}_t}{dt} = -\lambda s_t \widetilde{i}_t \big(-C + \frac{1}{s_t} + \frac{\gamma}{\lambda s_t}\big) = \frac{ds_t}{dt}\big(-C + \frac{1+\frac{\gamma}{\lambda}}{s_t}\big)$$

from which we obtain by integration:

$$\widetilde{i}_t - \widetilde{i}_0 = -C(s_t - s_0) + \big(1 + \frac{\gamma}{\lambda}\big)\log\frac{s_t}{s_0}.$$

Since $\widetilde{i}_\infty = 0$:

$$-\widetilde{i}_0 + C(s_\infty - s_0) = \big(1 + \frac{\gamma}{\lambda}\big)\log\frac{s_\infty}{s_0}.$$

Computing $z := s_0 - s_\infty$, we recover the announced result. □

For further and recent developments on moment closures, we refer the reader to e.g. [93] or [71].

3.2 Volz and Miller approach

In 2008, Volz [111] proposed a system of only 5 ODEs to describe the spread of an epidemic on a random CM graph. Volz approximation is based on an edge-centered point of view, in an 'infinite' CM graph setting, without any assumption of moment closure. We present Volz equations and then explain how to recover them with Miller's approach [79]. The derivation of these equations as limit of epidemics spreading on finite graphs is detailed following the approach of Decreusefond et al. [45].

The spread of diseases on random graphs involves two sources of randomness: one for the random graph, the other for describing the way the epidemic propagates on this random environment. An idea coming from statistical mechanics is to build the random graph progressively as the epidemic spreads over it, instead of first constructing the random graph, conditioning on it and studying the epidemic on the frozen environment. We detail the process that we will consider in the rest of the section.
Assume that only the edges joining the I and R individuals are observed. This means that the cluster of infectious and removed individuals is built, while the network of susceptible individuals is still not defined. We further assume that the degree of each individual is known. To each I individual is associated an exponential random clock with rate γ to determine its removal. To each open edge (directed to S), we associate a random exponential

clock with rate λ. When it rings, an infection occurs. The infectious ego chooses the edge of a susceptible alter at random. Hence the latter individual is chosen proportionally to her/his degree, in the size biased distribution, as explained in (1.2.1). When this susceptible individual becomes infected, she/he is connected and uncovers the edges to neighbours that were already in the subgraph: we determine whether her/his remaining edges are linked with I, R-type individuals (already in the observed cluster) or to S, in which case the edges remains 'open' (the alter is not chosen yet).

Let us consider the limit when the size of the graph converges to infinity, and let us denote as before by s_t and i_t the proportion of susceptible and infectious individuals in the population at time t. A key quantity in the approach of Volz [111] and Miller [79] is the probability $\theta(t)$ that an directed edge picked uniformly at random at t has not transmitted the disease. Let $u \in V$ be a vertex of degree k. The vertex u is still susceptible at time t if none of its k edges has transmitted the disease. By the construction of the stochastic process, where the random graph is built simultaneously to the spread of the disease on it, any infectious individual that transmits the disease pairs one of her/his half-edge with a half-edge of a susceptible individual chosen uniformly at random. Thus, the probability that none of the k edges of a susceptible has transmitted the disease up to time t is $\theta^k(t)$. Hence,

$$s_t = \sum_{k=0}^{+\infty} \theta(t)^k p_k = g(\theta(t)), \tag{3.2.1}$$

where g is the generating function of the probability distribution $(p_k)_{k\geq 0}$ (see (1.2.2)). Notice that in Equation (3.2.1), the proportion s_t of susceptibles is assumed to coincide with the expectation of the proportion of the number of susceptible individuals at t. We recall that a rigorous derivation of Volz' equations is given in Section 3.3.7 below.

3.2.1 Dynamics of $\theta(t)$

To deduce an equation for s_t from (3.2.1), an equation for $\theta(t)$ is needed.

Proposition 3.2.1. *We have that:*

$$\frac{d\theta}{dt} = -\lambda\theta(t) + \gamma(1-\theta(t)) + \lambda\frac{g'(\theta(t))}{g'(1)}.$$

Proof. Denote by $h(t)$ the probability that the alter is still susceptible at time t. Define $\phi(t)$ as the probability that a random edge has not transmitted the disease and that its alter is infectious. Notice that

$$\frac{d\theta}{dt} = -\lambda\phi(t). \tag{3.2.2}$$

Given an edge satisfying the definition of $\phi(t)$ (an edge that has not transmitted the disease yet and whose alter is infectious), the probability that the alter is of degree k is given by (1.2.1) and given its degree, the probability that it is still susceptible at time t is $\theta^{k-1}(t)$, because the considered edge did not transmit the disease before t. Then:

$$h(t) = \sum_{k=0}^{+\infty} \frac{kp_k}{m}\theta^{k-1}(t) = \frac{g'(\theta(t))}{g'(1)},$$

from which we deduce that

$$\frac{dh}{dt} = \frac{g''(\theta(t))}{g'(1)}\frac{d\theta}{dt} = -\lambda\phi(t)\frac{g''(\theta(t))}{g'(1)}.$$

An equation for the evolution of $\phi(t)$ can be written by noticing that:

- An edge stops satisfying the definition of ϕ if it transmits the disease or if the alter is removed.
- An edge starts satisfying the definition of ϕ if its alter becomes infectious.

Thus

$$\begin{aligned}\frac{d\phi}{dt} &= -(\lambda+\gamma)\phi(t) - \frac{dh}{dt} \\ &= -(\lambda+\gamma)\phi(t) + \lambda\phi(t)\frac{g''(\theta(t))}{g'(1)} \\ &= \frac{\lambda+\gamma}{\lambda}\frac{d\theta}{dt} - \frac{g''(\theta(t))}{g'(1)}\frac{d\theta}{dt},\end{aligned} \tag{3.2.3}$$

which gives for a constant C:

$$\phi(t) = \frac{\lambda+\gamma}{\lambda}\theta(t) - \frac{g'(\theta(t))}{g'(1)} + C.$$

Using that $\phi(0) = 0$ and $\theta(0) = 1$, we deduce that $C = -\gamma/\lambda$ and hence

$$\phi(t) = \theta(t) - \frac{\gamma}{\lambda}(1-\theta(t)) - \frac{g'(\theta(t))}{g'(1)}. \tag{3.2.4}$$

We deduce the announced result from (3.2.2) and (3.2.4). □

3.2.2 Miller's equations

We can now deduce the equations for the proportions s_t, i_t and r_t of susceptible, infectious and recovered individuals proposed by Miller [79].

Proposition 3.2.2 (Miller's equations [79])**.** *We have:*

$$\begin{aligned}&s_t = g(\theta(t)) \\ &\frac{dr_t}{dt} = \gamma i_t \\ &\frac{di_t}{dt} = -g'(\theta(t))\big(-\lambda\theta(t) + \gamma(1-\theta(t)) + \lambda\frac{g'(\theta(t))}{g'(1)}\big) - \gamma i_t. \\ &\frac{d\theta}{dt} = -\lambda\theta(t) + \gamma(1-\theta(t)) + \lambda\frac{g'(\theta(t))}{g'(1)}.\end{aligned}$$

Proof. By (3.2.1), we have that $s_t = g(\theta(t))$. From the node-centered removal dynamics of infectious nodes, we have that $\frac{dr_t}{dt} = \gamma i_t$. Using $i_t = 1 - s_t - r_t$ and Proposition 3.2.1, we obtain the two last equations. □

We can now recover the equations proposed by Volz [111] by introducing the proportion of edges $\mathrm{I} - \mathrm{S}$ that have not transmitted the disease yet

$$p_I(t) = \frac{\phi(t)}{\theta(t)} \tag{3.2.5}$$

and the proportion of edges $\mathrm{S} - \mathrm{S}$ that have not transmitted the disease

$$p_S(t) = \frac{g'(\theta(t))}{\theta(t)g'(1)}. \tag{3.2.6}$$

From Miller's equations, we obtain by straightforward computation:

Proposition 3.2.3 (Volz' equations [111])**.** *We have:*

$$\begin{aligned}&\theta(t) = \exp\Big(-\lambda\int_0^t p_I(s)\,ds\Big), \qquad s_t = g(\theta(t)), \\ &\frac{di_t}{dt} = \lambda p_I(t)\theta(t)g'(\theta(t)) - \gamma i_t\end{aligned}$$

$$\frac{dp_I}{dt} = \lambda\, p_I(t)p_S(t)\theta(t)\frac{g''(\theta(t))}{g'(\theta(t))} - \lambda\, p_I(t)(1-p_I(t)) - \gamma p_T(t).$$
$$\frac{dp_S}{dt} = \lambda\, p_I(t)p_S(t)\big(1-\theta(t)\frac{g''(\theta(t))}{g'(\theta(t))}\big).$$

Let us compare Volz' equations with the Kermack–McKendrick equations:

$$\frac{ds}{dt} = -\lambda\, s_t i_t, \qquad \frac{di}{dt} = \lambda\, s_t i_t - \gamma i_t.$$

In Volz' equations, denoting by $\bar{N}_t^{\mathrm{S}} = p_I(t)\theta(t)g'(\theta(t))$ the 'quantity' of edges from I to S:

$$\frac{ds_t}{dt} = g'(\theta(t))\frac{d\theta}{dt} = -\lambda g'(\theta(t))\theta(t)p_I(t) = -\lambda \bar{N}_t^{\mathrm{S}} p_I(t) = -\lambda \bar{N}_t^{\mathrm{IS}}$$
$$\frac{di_t}{dt} = \lambda \times \bar{N}_t^{\mathrm{IS}} - \gamma i_t.$$

These equations account for the fact that not all the I and S vertices are connected, which modifies the infection pressure compared with the mixing models (Part I of this volume).

3.3 Measure-valued processes

Decreusefond et al. [45] proved the convergence that was left open by Volz [111]. The proof that we now present underlines the key objects that lie at the core of the phenomenon: because degree distributions are central in CMs, these objets are not surprisingly measures representing some particular degree distributions. Three degree distributions are sufficient to describe the epidemic dynamics which evolve in the space of measures on the set of nonnegative integers, and of which Volz' equations are a by-product.
A rigorous individual-based description of the epidemic on a random graph is provided. Starting with a node-centered description, we show that the individual dimension is lost in the large graph limit. Our construction heavily relies on the choice of a CM for the graph underlying the epidemic, which was also made in [111].

3.3.1 Stochastic model for a finite graph with N vertices

Recall the notation of Section 1.4. The idea of Volz is to use network-centric quantities (such as the number of edges from I to S) rather than node-centric quantities. For a vertex $u \in$ S, D_u corresponds to the degree of u. For $u \in$ I (respectively R), $D_u(\mathrm{S})$ represents the number of edges with u as infectious (resp. removed) ego and susceptible alter. The numbers of edges with susceptible ego (resp. of edges of types I − S and R − S) are denoted by N_t^{S} (resp. N_t^{IS} and N_t^{RS}). All these quantities are in fact encoded into three degree distributions, that we now introduce and on which we will work to establish Volz' equations. Notice that with the notations of Section 3.1, $\frac{1}{N}N_t^{\mathrm{IS}} = [SI]_t$ and $\frac{1}{N}N_t^{\mathrm{RS}} = [SR]_t$. However, we drop this notation with brackets for simplification of later formula and because we will not need motifs other than edges.

Definition 3.3.1. *We consider here the following three degree distributions of $\mathscr{M}_F(\mathbb{Z}_+)$, given for $t \geq 0$ as:*

$$\mu_t^{\mathrm{S}} = \sum_{u\in \mathrm{S}_t} \delta_{D_u}, \quad \mu_t^{\mathrm{IS}} = \sum_{u\in \mathrm{I}_t} \delta_{D_u(\mathrm{S}_t)}, \quad \mu_t^{\mathrm{RS}} = \sum_{u\in \mathrm{R}_t} \delta_{D_u(\mathrm{S}_t)}, \tag{3.3.1}$$

where we recall that δ_D is the Dirac mass at $D \in \mathbb{Z}_+$ (see Notation 0.0.1).

Notice that the measures μ_t^{S}/S_t, μ_t^{IS}/I_t and μ_t^{RS}/R_t are probability measures that correspond to usual (probability) degree distributions. The degree distribution μ_t^{S} of susceptible individuals is needed to describe the degrees of the new infected individuals. The measure μ_t^{IS} provides information on the number of edges from I_t to S_t, through which the disease can propagate. Similarly, the measure μ_t^{RS} is used to describe the evolution of the set of edges

linking S_t to R_t.
Using Notation 0.0.1, we can see that

$$I_t = \langle \mu_t^{\mathrm{IS}}, 1 \rangle, \qquad N_t^{\mathrm{IS}} = \langle \mu_t^{\mathrm{IS}}, \chi \rangle = \sum_{u \in \mathrm{I}_t} D_u(\mathrm{S}_t),$$

and accordingly for N_t^{S}, N^{RS}, S_t and R_t.

Definition 3.3.2 (Labelling the nodes)**.** *For an integer-valued measure $\mu \in \mathscr{M}_F(\mathbb{Z}_+)$, we can rank its atoms by increasing degrees and label them with this order. A way of deducing this labelling from μ by using its cumulative distribution function is proposed in [45]. We omit it here for the sake of simplicity.*

Example 3.3.3. *Consider for instance the measure $\mu = 2\delta_1 + 3\delta_5 + \delta_7$. If μ is a degree distribution, this means that 2 individuals have degree 1, 3 individuals have degree 5 and 1 individual has degree 7. Ranking the atoms by increasing degrees, we can label them from 1 to 6 such that $D_1 = D_2 = 1$, $D_3 = D_4 = D_5 = 5$, $D_6 = 7$.* □

3.3.2 Dynamics and measure-valued SDEs

Suppose that at initial time, we are given a set of S and I nodes together with their degrees. The graph of relationships between the I individuals is in fact irrelevant for studying the propagation of the disease. The minimal information consists in the sizes of the classes S, I, R and the number of edges to the class S for every infectious or removed node. Each node of class S comes with a given number of half-edges of undetermined types ; each node of class I (resp. R) comes with a number of $\mathrm{I}-\mathrm{S}$ (resp. $\mathrm{R}-\mathrm{S}$) edges. The numbers of $\mathrm{I}-\mathrm{R}$, $\mathrm{I}-\mathrm{I}$ and $\mathrm{R}-\mathrm{R}$ edges need not to be retained. The three descriptors in (3.3.1) are hence sufficient to describe the evolution of the SIR epidemic.

Recall the graph construction of Section 3.2 explaining how to handle simultaneously the two sources of randomness of the problem. The random network of social relationships is explored while the disease spreads on it: only the clusters of I and R individuals are observed and constructed, with $\mathrm{I}-\mathrm{S}$ and $\mathrm{R}-\mathrm{S}$ edges having their S alter still unaffected. Susceptible individuals remain unattached until they become infected, in which case their connections to the cluster of I's and R's are revealed. We assume that the degree distribution of S_0 and the size N of the total population are known.

We now explain the dynamics, that is summarized in Figure 3.3.1. Recall that to each half-edge of type $\mathrm{I}-\mathrm{S}$, an independent exponential clock with parameter λ is associated, and to each I vertex, an independent exponential clock with parameter γ is associated. The first of all these clocks that rings determines the next event.

Case 1 If the clock that rings is associated to an I individual, the latter is removed. Change her status from I to R and the type of her emanating half-edges accordingly: $\mathrm{I}-\mathrm{S}$ half-edges become $\mathrm{R}-\mathrm{S}$ half-edges for example.

Case 2 If the clock that rings is associated with a half $\mathrm{I}-\mathrm{S}$-edge (with unaffected susceptible alter), an infection occurs.

- **Step 1** Match randomly the $\mathrm{I}-\mathrm{S}$-half-edge whose clock has rung to a half-edge of a susceptible: this determines the susceptible becoming infected.
- **Step 2** Let k be the degree of the newly infected individual. Choose uniformly $k-1$ half edges among the open half-edges of the cluster of I and R individuals ($\mathrm{I}-\mathrm{S}$ or $\mathrm{R}-\mathrm{S}$ edges of this cluster, with susceptible alter still unaffected) and among the half edges of susceptible individuals. Let j, ℓ and m be the respective number of $\mathrm{I}-\mathrm{S}$, $\mathrm{R}-\mathrm{S}$ and $\mathrm{S}-\mathrm{S}$ edges chosen among the $k-1$ picks.
- **Step 3** The chosen half-edges of type $\mathrm{I}-\mathrm{S}$ and $\mathrm{R}-\mathrm{S}$ determine the infectious or removed neighbours of the newly infected individual who become the new (infectious) alter of these edges. The remaining m edges of type $\mathrm{S}-\mathrm{S}$ remain open in the sense that the susceptible neighbour is not fixed.
  Change the status of the newly infected from S to I. Change the status of the m (resp. j, ℓ) $\mathrm{S}-\mathrm{S}$-type (resp. $\mathrm{I}-\mathrm{S}$-type, $\mathrm{R}-\mathrm{S}$-type) edges considered to $\mathrm{I}-\mathrm{S}$-type (resp. $\mathrm{I}-\mathrm{I}$-type, $\mathrm{R}-\mathrm{I}$-type). □

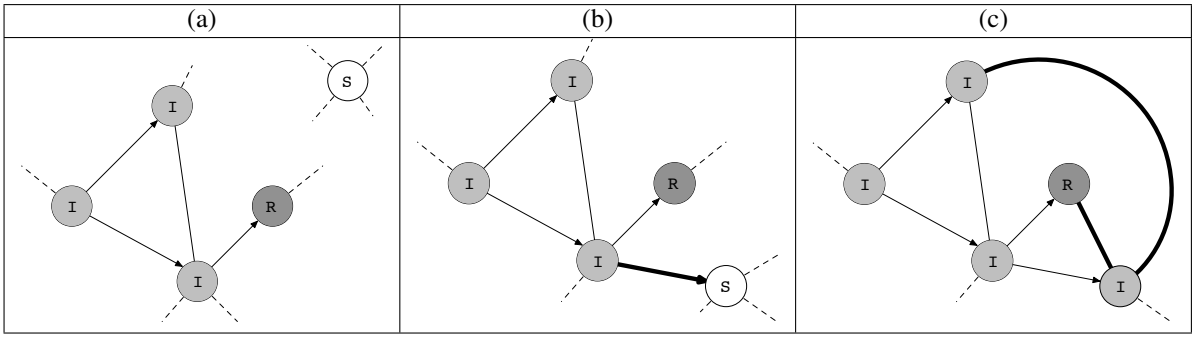


Figure 3.3.1: *Infection process. Arrows provide the infection tree. Susceptible, infectious and removed individuals are colored in white, grey and dark grey respectively. (a) The degree of each individual is known, and for each infectious (resp. removed) individual, we know his/her number of edges of type* IS *(resp.* RS*). (b) A contaminating half edge is chosen and a susceptible of degree k is infected at time t with the rate $\Lambda_t(k)$ defined in (3.3.13). The contaminating edge is drawn in bold line. The number $N^{\mathrm{IS}}_{t_-}$ of edges from* I *to* S *momentarily becomes $N^{\mathrm{IS}}_{t_-}-1+(k-1)$. (c) Once the susceptible individual has been infected, we determine how many of its remaining arrows are linked to the classes* I *and* R*. If we denote by j and ℓ these numbers, then $N^{\mathrm{IS}}_t = N^{\mathrm{IS}}_{t_-}-1+(k-1)-j-\ell$ and $N^{\mathrm{RS}}_t = N^{\mathrm{RS}}_{t_-}-\ell$.*

We then wait for another clock to ring and repeat the procedure.

From the dynamics described above, we can read that the global force of infection at time t is

$$\lambda N^{\mathrm{IS}}_{t_-}.$$

When an infection occurs, a half-edge of a susceptible individual is chosen and determines who is the contaminated person. Therefore, a given susceptible of degree k has a probability $k/N^{\mathrm{S}}_{t_-}$ to be the next infected individual. So that the rate of infection of a given susceptible of degree k at time t is:

$$\Lambda_{t_-}(k) = \lambda k \frac{N^{\mathrm{IS}}_{t_-}}{N^{\mathrm{S}}_{t_-}} = \lambda k p_I(t_-), \tag{3.3.2}$$

where $p_I(t)$ is defined by

$$p_I(t) = \frac{N^{\mathrm{IS}}_t}{N^{\mathrm{S}}_t},$$

is the proportion of edges linked to susceptible individuals that can transmit the disease. It is the discrete stochastic quantity that we expect will converge to (3.2.5).

Starting from t, and because of the properties of exponential distributions, the next event will take place after an exponentially distributed time with parameter $\lambda N^{\mathrm{IS}}_t + \gamma I_t$. Let T denote the time of this event after t.

Case 1 The next event corresponds to a removal, i.e., a node goes from status I to status R. Choose uniformly $u \in I_{T^-}$ (with probability $1/I_{T^-}$, then update the measures $\mu^{\mathrm{IS}}_{T_-}$ and $\mu^{\mathrm{RS}}_{T_-}$:

$$\mu^{\mathrm{IS}}_T = \mu^{\mathrm{IS}}_{T^-} - \delta_{D_u(\mathrm{S}_{T_-})} \text{ and } \mu^{\mathrm{RS}}_T = \mu^{\mathrm{RS}}_{T^-} + \delta_{D_u(\mathrm{S}_{T_-})}.$$

Case 2 The next event corresponds to a new infection. We choose uniformly a half-edge with susceptible alter, and this alter becomes infectious. The new infective has degree k with probability $k\mu^{\mathrm{S}}_{T_-}(k)/N^{\mathrm{S}}_{T_-}$. When the new individual is 'discovered' by the disease, she/he reveals her/his links with other infectious or removed individuals. The probability, given that the degree of the individual is k and that j (resp. ℓ) out of her $k-1$

other half-edges (all but the contaminating IS edge) are chosen to be of type II (resp. IR), according to Step 2', is given by the following multivariate hypergeometric distribution:

$$p_{T_-}(j,\ell\,|\,k-1) = \frac{\binom{N^{\mathrm{IS}}_{T_-}-1}{j}\binom{N^{\mathrm{RS}}_{T_-}}{\ell}\binom{N^{\mathrm{S}}_{T_-}-N^{\mathrm{IS}}_{T_-}-N^{\mathrm{RS}}_{T_-}}{k-1-j-\ell}}{\binom{N^{\mathrm{S}}_{T_-}-1}{k-1}}. \tag{3.3.3}$$

Finally, to update the values of μ^{IS}_T and μ^{RS}_T given k, j and ℓ, we have to choose the infectious and removed individuals to which the newly infectious is linked: some of their edges, which were IS or RS, now become II or RI. We draw two sequences of integers $\underline{n} = (n_1,\dots,n_{I_{T_-}})$ and $\underline{m} = (m_1,\dots,m_{R_{T_-}})$ that will indicate how many links each infectious or removed individual has to the newly contaminated individual. There exist constraints on these sequences: the number of edges recorded for each individual by the vectors $\underline{n}$ and $\underline{m}$ can not exceed the number of existing edges. Let us define the set

$$\mathscr{L} = \bigcup_{m=1}^{+\infty} \mathbb{Z}_+^m, \tag{3.3.4}$$

and for all finite integer-valued measure μ on $\mathbb{Z}_+$, corresponding to a degree distribution as in Section 3.3.1, and whose atoms are labelled say, according to Definition (3.3.2) and for all integer $\ell \in \mathbb{Z}_+$, we define the subset

$$\mathscr{L}(\ell,\mu) = \Big\{\underline{n} = (n_1,\dots,n_{\langle\mu,\mathbf{1}\rangle}) \in \mathbb{Z}_+^{\langle\mu,\mathbf{1}\rangle} \quad \text{such that} \\ \forall u \in \{1,\dots,\langle\mu,\mathbf{1}\rangle\},\, n_u \le D_u(\mu) \text{ and } \sum_{u=1}^{\langle\mu,\mathbf{1}\rangle} n_u = \ell\Big\}, \tag{3.3.5}$$

where $D_u(\mu)$ stands for the degree of the vertex u, read from the measure μ (see Example 3.3.3). Each sequence $\underline{n} \in \mathscr{L}(\ell,\mu)$ provides a possible configuration of how the ℓ connections of a given individual can be shared between neighbours whose degrees are summed up by μ. The component n_u, for $1 \le u \le \langle\mu,1\rangle$, provides the number of edges that this individual shares with the individual u. This number is necessarily smaller than the degree $D_u(\mu)$ of individual u. Moreover, the components of the vector $\underline{n}$ sum to ℓ. The probabilities of the draws of $\underline{n}$ and $\underline{m}$ that provide respectively the number of edges $\mathrm{I}-\mathrm{S}$ which become $\mathrm{I}-\mathrm{I}$ per infectious individual and the number of edges $\mathrm{R}-\mathrm{S}$ which become $\mathrm{R}-\mathrm{I}$ per removed individual are given by:

$$\rho(\underline{n}|j+1,\mu^{\mathrm{IS}}_{T_-}) = \frac{\prod_{u\in \mathrm{I}_{T_-}}\binom{D_u(\mathrm{S}_{T_-})}{n_u}}{\binom{N^{\mathrm{IS}}_{T_-}}{j+1}}\mathbf{1}_{\underline{n}\in\mathscr{L}(j+1,\mu^{\mathrm{IS}}_{T_-})} \\ \rho(\underline{m}|\ell,\mu^{\mathrm{RS}}_{T_-}) = \frac{\prod_{v\in \mathrm{R}_{T_-}}\binom{D_v(\mathrm{S}_{T_-})}{m_v}}{\binom{N^{\mathrm{RS}}_{T_-}}{\ell}}\mathbf{1}_{\underline{m}\in\mathscr{L}(\ell,\mu^{\mathrm{RS}}_{T_-})}. \tag{3.3.6}$$

Note that $I_{T_-} = \langle\mu^{\mathrm{IS}}_{T_-},1\rangle$ is the total mass of the measure $\mu^{\mathrm{IS}}_{T_-}$ and that $D_u(\mathrm{S}_{T_-})$ corresponds to the degree of the individual u encoded by $\mu^{\mathrm{IS}}_{T_-}$ with the labelling of Definition 3.3.2, i.e. to the number of edges from u to S before time T.

Then, we update the measures as follows:

$$\mu^{\mathrm{S}}_T = \mu^{\mathrm{S}}_{T^-} - \delta_k \\ \mu^{\mathrm{IS}}_T = \mu^{\mathrm{IS}}_{T^-} + \delta_{k-1-j-\ell} + \sum_{u\in \mathrm{I}_{T_-}}\big(\delta_{D_u(\mathrm{S}_{T_-})-n_u} - \delta_{D_u(\mathrm{S}_{T_-})}\big) \\ \mu^{\mathrm{RS}}_T = \mu^{\mathrm{RS}}_{T^-} + \sum_{v\in \mathrm{R}_{T_-}}\big(\delta_{D_v(\mathrm{S}_{T_-})-m_v} - \delta_{D_v(\mathrm{S}_{T_-})}\big). \tag{3.3.7}$$

Here, we propose stochastic differential equations (SDEs) driven by Poisson point measures (PPMs) to describe the evolution of the degree distributions (3.3.1) as in [45].

We consider two Poisson point measures Q^1 and Q^2 on $E_1 := \mathbb{Z}_+ \times \mathbb{R}_+ \times \mathbb{Z}_+ \times \mathbb{Z}_+ \times \mathbb{R}_+ \times \mathscr{L} \times \mathbb{R}_+ \times \mathscr{L} \times \mathbb{R}_+$ and $\mathbb{R}_+ \times \mathbb{Z}_+$ with intensity measures the product of Lebesgue measures on $\mathbb{R}_+$ and the of counting measures on each discrete set. The atoms of the point measure Q^1 are of the form $(s,k,\theta_1,j,\ell,\theta_2,\underline{n},\theta_3,\underline{m},\theta_4)$. They provide possible times s at which an infection may occur, and gives an integer k corresponding to the degree of the susceptible being possibly infected, the numbers $j+1$ and ℓ of edges that this individual has to the sets I_{s_-} and R_{s_-}. The marks $\underline{n}$ and $\underline{m} \in \mathscr{L}$ are as in the previous section. The marks θ_1, θ_2 and θ_3 are auxiliary variables used for the construction (see (3.3.9)–(3.3.10)) below.
The atoms of the point measure Q^2 are of the form (s,u) and give possible removal times s associated with the label u of the individual that may be removed.

The following SDEs describe the evolution of the epidemic: for all $t \geq 0$,

$$\mu_t^{\mathrm{S}} = \mu_0^{\mathrm{S}} - \int_0^t \int_{E_1} \delta_k \mathbf{1}_{\theta_1 \leq \Lambda_{s_-}(k)\mu_{s_-}^{\mathrm{S}}(k)} \tag{3.3.8}$$
$$\mathbf{1}_{\theta_2 \leq p_{s_-}(j,\ell|k-1)} \mathbf{1}_{\theta_3 \leq \rho(\underline{n}|j+1,\mu_{s_-}^{\mathrm{IS}})} \mathbf{1}_{\theta_4 \leq \rho(\underline{m}|\ell,\mu_{s_-}^{\mathrm{RS}})} \, \mathrm{d}Q^1$$

$$\mu_t^{\mathrm{IS}} = \mu_0^{\mathrm{IS}} + \int_0^t \int_{E_1} \Big(\delta_{k-(j+1+\ell)} + \sum_{u \in \mathrm{I}_{s_-}} \big(\delta_{D_u(\mu_{s_-}^{\mathrm{IS}})-n_u} - \delta_{D_u(\mu_{s_-}^{\mathrm{IS}})} \big) \Big) \tag{3.3.9}$$
$$\times \mathbf{1}_{\theta_1 \leq \Lambda_{s_-}(k)\mu_{s_-}^{\mathrm{S}}(k)} \mathbf{1}_{\theta_2 \leq p_{s_-}(j,\ell|k-1)} \mathbf{1}_{\theta_3 \leq \rho(\underline{n}|j+1,\mu_{s_-}^{\mathrm{IS}})} \mathbf{1}_{\theta_4 \leq \rho(\underline{m}|\ell,\mu_{s_-}^{\mathrm{RS}})} \, \mathrm{d}Q^1$$
$$- \int_0^t \int_{\mathbb{Z}_+} \delta_{D_u(\mu_{s_-}^{\mathrm{IS}})} \mathbf{1}_{u \in \mathrm{I}_{s_-}} \, \mathrm{d}Q^2$$

$$\mu_t^{\mathrm{RS}} = \mu_0^{\mathrm{RS}} + \int_0^t \int_{E_1} \Big(\sum_{v \in \mathrm{R}_{s_-}} \big(\delta_{D_v(\mu_{s_-}^{\mathrm{RS}})-m_v} - \delta_{D_v(\mu_{s_-}^{\mathrm{RS}})} \big) \Big) \tag{3.3.10}$$
$$\times \mathbf{1}_{\theta_1 \leq \Lambda_{s_-}(k)\mu_{s_-}^{\mathrm{S}}(k)} \mathbf{1}_{\theta_2 \leq p_{s_-}(j,\ell|k-1)} \mathbf{1}_{\theta_3 \leq \rho(\underline{n}|j+1,\mu_{s_-}^{\mathrm{IS}})} \mathbf{1}_{\theta_4 \leq \rho(\underline{m}|\ell,\mu_{s_-}^{\mathrm{RS}})} \, \mathrm{d}Q^1$$
$$+ \int_0^t \int_{\mathbb{Z}_+} \delta_{D_u(\mu_{s_-}^{\mathrm{IS}})} \mathbf{1}_{u \in \mathrm{I}_{s_-}} \, \mathrm{d}Q^2,$$

where we write $\mathrm{d}Q^1$ and $\mathrm{d}Q^2$ instead of $\mathrm{d}Q^1(s,k,\theta_1,j,\ell,\theta_2,\underline{n},\theta_3,\underline{m},\theta_4)$ and $\mathrm{d}Q^2(s,u)$ to simplify the notation.

Proposition 3.3.4. *For any given initial conditions μ_0^{S}, μ_0^{SI} and μ_0^{RS} that are integer-valued measures on $\mathbb{Z}_+$ and for PPMs Q^1 and Q^2, there exists a unique strong solution to the SDEs (3.3.8)–(3.3.10) in the space $\mathbb{D}\big(\mathbb{R}_+, (\mathscr{M}_F(\mathbb{Z}_+))^3\big)$, the Skorokhod space of càdlàg functions with values in $(\mathscr{M}_F(\mathbb{Z}_+))^3$.*

Proof. For the proof, we notice that for every $t \in \mathbb{R}_+$, the measure μ_t^{S} is dominated by μ_0^{S} and the measures μ_t^{IS} and μ_t^{RS} have a mass bounded by $\langle \mu_0^{\mathrm{S}} + \mu_0^{\mathrm{IS}} + \mu_0^{\mathrm{RS}}, 1 \rangle$ and a support included in $[\![0, \max\{\max(\mathrm{supp}(\mu_0^{\mathrm{S}})), \max(\mathrm{supp}(\mu_0^{\mathrm{IS}})), \max(\mathrm{supp}(\mu_0^{\mathrm{RS}}))\}]\!]$ The result then follows the steps of [56] and [106] (Proposition 2.2.6) where a pathwise construction of the solution on the positive real line is given using the Poisson point processes Q^1 and Q^2. □

The course of the epidemic can be deduced from (3.3.8), (3.3.9) and (3.3.10). For the sizes $(S_t, I_t, R_t)_{t \in \mathbb{R}_+}$ of the different classes, for instance, we have with the choice of $f \equiv 1$ that for all $t \geq 0$, $S_t = \langle \mu_t^{\mathrm{S}}, \mathbf{1} \rangle$, $I_t = \langle \mu_t^{\mathrm{IS}}, \mathbf{1} \rangle$ and $R_t = \langle \mu_t^{\mathrm{RS}}, \mathbf{1} \rangle$ (see Notation 0.0.1). Writing the semi-martingale decomposition that results from standard stochastic calculus for jump processes and SDE driven by PPMs (e.g. [56, 63, 64]), we obtain for example:

$$I_t = \langle \mu_t^{\mathrm{IS}}, \mathbf{1} \rangle = I_0 + \int_0^t \Big(\sum_{k \in \mathbb{Z}_+} \mu_s^{\mathrm{S}}(k) \Lambda_s(k) - \gamma I_s \Big) \, \mathrm{d}s + M_t^{\mathrm{I}}, \tag{3.3.11}$$

where M^{I} is a square-integrable martingale that can be written explicitly as a stochastic integral with respect to the compensated PPMs of Q^1 and Q^2, and with predictable quadratic variation given for all $t \geq 0$ by

$$\langle M^{\mathrm{I}} \rangle_t = \int_0^t \sum_{k \in \mathbb{Z}_+} \left(\mu_s^{\mathrm{S}}(k) \Lambda_s(k) + \gamma I_s \right) \mathrm{d}s.$$

Other quantities of interest are the numbers of edges of the different types NS_t, N_t^{IS}, N_t^{RS}. The latter appear as the first moments of the measures μ_t^{S}, μ_t^{IS} and μ_t^{RS}:

$$NS_t = \langle \mu_t^{\mathrm{S}}, \chi \rangle, \quad N_t^{\mathrm{IS}} = \langle \mu_t^{\mathrm{IS}}, \chi \rangle \quad \text{and} \quad N_t^{\mathrm{RS}} = \langle \mu_t^{\mathrm{RS}}, \chi \rangle.$$

3.3.3 Rescaling

We consider a sequence of larger and larger graphs $(\mathscr{G}_N)_{N \geq 1}$ with $N \to +\infty$. The degree distribution $\mathbf{p}$ underlying these CM graphs remains unchanged with N.
The sequences of measures $(\mu^{\mathrm{N,S}})_{N \in \mathbb{N}}$, $(\mu^{\mathrm{N,IS}})_{N \in \mathbb{N}}$ and $(\mu^{\mathrm{N,RS}})_{n \in \mathbb{N}}$ are defined as

$$\mu_t^{\mathrm{N,S}} = \frac{1}{N} \mu_t^{\mathrm{S}}, \qquad \mu_t^{\mathrm{N,IS}} = \frac{1}{N} \mu_t^{\mathrm{IS}}, \qquad \mu_t^{\mathrm{N,RS}} = \frac{1}{N} \mu_t^{\mathrm{RS}} \tag{3.3.12}$$

where the measures non-rescaled μ^{S}, μ^{IS} and μ^{RS} are defined as in (3.3.1) and implicitly depend on N:

$$\langle \mu_t^{\mathrm{N,S}}, 1 \rangle + \langle \mu^{\mathrm{N,IS}}, 1 \rangle + \langle \mu^{\mathrm{N,RS}}, 1 \rangle = \frac{N}{N} = 1.$$

The proportions S_t^{N}, I_t^{N} and R_t^{N} defined in (3.1.1) can then be rewritten as $S_t^{\mathrm{N}} = \langle \mu^{\mathrm{N,S}}, 1 \rangle$, $I_t^{\mathrm{N}} = \langle \mu^{\mathrm{N,IS}}, 1 \rangle$ and $R_t^{\mathrm{N}} = \langle \mu^{\mathrm{N,RS}}, 1 \rangle$. Also, we have $N_t^{\mathrm{N,S}} = \langle \mu^{\mathrm{N,S}}, \chi \rangle$, $N_t^{\mathrm{N,IS}} = \langle \mu^{\mathrm{N,IS}}, \chi \rangle$ and $N_t^{\mathrm{N,RS}} = \langle \mu^{\mathrm{N,RS}}, \chi \rangle$, the numbers, renormalized by N, of edges with susceptible ego, infectious ego and susceptible alter, removed ego and susceptible alter.

We assume that the initial conditions satisfy:

Assumption 3.3.5. *The sequences $(\mu_0^{\mathrm{N,S}})_{n \in \mathbb{N}}$, $(\mu_0^{\mathrm{N,IS}})_{n \in \mathbb{N}}$ and $(\mu_0^{\mathrm{N,RS}})_{n \in \mathbb{N}}$ converge to measures $\bar{\mu}_0^{\mathrm{S}}$, $\bar{\mu}_0^{\mathrm{IS}}$ and $\bar{\mu}_0^{\mathrm{RS}}$ in $\mathscr{M}_F(\mathbb{Z}_+)$ equipped with the topology of weak convergence.*

Remark 3.3.6. *1. Assumption 3.3.5 entails that the initial (susceptible and infectious) population size is of order N if $\bar{\mu}_0^{\mathrm{S}}$ and $\bar{\mu}_0^{\mathrm{IS}}$ are nontrivial.*
2. If the distributions underlying the measures $\mu_0^{\mathrm{N,S}}$, $\mu_0^{\mathrm{N,IS}}$ and $\mu_0^{\mathrm{N,RS}}$ do not depend on the total number of vertices (e.g. Poisson, power-laws or geometric distributions), Assumption 3.3.5 can be viewed as a law of large numbers. When the distributions depend on the total number of vertices N (as in Erdös-Renyi graphs), there may be scalings under which Assumption 3.3.5 holds. For Erdös-Renyi graphs for instance, if the probability p_N of connecting two vertices satisfies $\lim_{N \to +\infty} N p_N = \lambda$, then we obtain in the limit a Poisson distribution with parameter λ.
3. Notice the appearance in Equation (3.3.2) of the size biased degree distribution. The latter reflects the fact that, in the CM, individuals having large degrees have higher probability to connect than individuals having small degrees. Thus, there is no reason why the degree distributions of the susceptible individuals $\bar{\mu}_0^{\mathrm{S}}/\bar{S}_0$ and the distribution $\sum_{k \in \mathbb{Z}_+} p_k \delta_k$ underlying the CM should coincide. This is developed in Section 3.3.6. □

It is possible to write rescaled SDEs which are the same as the SDEs (3.3.8)–(3.3.10) parameterized by N (see [45] for details). Several semi-martingale decompositions will be useful in the sequel. We focus on $\mu^{\mathrm{N,IS}}$ but similar decompositions hold for $\mu^{\mathrm{N,S}}$ and $\mu^{\mathrm{N,RS}}$, which we do not detail since they can be deduced by direct adaptation of the computation which follows.

Proposition 3.3.7. *Define:*

$$\Lambda_s^N(k) = \lambda k \frac{N_s^{\mathrm{N,IS}}}{N_s^{\mathrm{N,S}}}, \text{ and } p_s^N(j, \ell \mid k-1) = \frac{\binom{N_s^{\mathrm{N,IS}}-1}{j} \binom{N_s^{\mathrm{N,RS}}}{\ell} \binom{N_s^{\mathrm{N,S}} - N_s^{\mathrm{N,IS}} - N_s^{\mathrm{N,RS}}}{k-1-j-\ell}}{\binom{N_s^{\mathrm{N,S}}-1}{k-1}}. \tag{3.3.13}$$

For all $f \in \mathscr{B}_b(\mathbb{Z}_+)$, for all $t \geq 0$,

$$\langle \mu_t^{\mathrm{N,IS}}, f\rangle = \sum_{k\in\mathbb{Z}_+} f(k)\mu_0^{\mathrm{N,IS}}(k) + A_t^{\mathrm{N,IS},f} + M_t^{\mathrm{N,IS},f}, \tag{3.3.14}$$

where the finite variation part $A_t^{\mathrm{N,IS},f}$ of $\langle \mu_t^{\mathrm{N,IS}}, f\rangle$ reads

$$\begin{aligned} A_t^{\mathrm{N,IS},f} = \int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell+1\leq k} p_s^N(j,\ell|k-1) \sum_{\underline{n}\in\mathscr{L}} \rho(\underline{n}|j+1,\mu_s^{\mathrm{N,IS}}) \\ \times \Big(f(k-(j+1+\ell)) + \sum_{u\in I_s^{\mathrm{N}}} \big(f(D_u(\mathrm{S}_s)-n_u) - f(D_u(\mathrm{S}_s))\big)\Big)\, ds \\ - \int_0^t \gamma\langle \mu_s^{\mathrm{N,IS}}, f\rangle\, ds, \end{aligned} \tag{3.3.15}$$

and where the martingale part $M_t^{\mathrm{N,IS},f}$ of $\langle \mu_t^{\mathrm{N,IS}}, f\rangle$ is a square integrable martingale starting from 0 with quadratic variation

$$\begin{aligned} \langle M^{N,\mathrm{IS},f}\rangle_t = \frac{1}{N}\int_0^t \gamma\langle \mu_s^{\mathrm{N,IS}}, f^2\rangle\, ds \\ + \frac{1}{N}\int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell+1\leq k} p_s^N(j,\ell|k-1) \sum_{\underline{n}\in\mathscr{L}} \rho(\underline{n}|j+1,\mu_s^{\mathrm{N,IS}}) \\ \times \Big(f\left(k-(j+1+\ell)\right) + \sum_{u\in I_s^{\mathrm{N}}} \left(f\left(D_u(\mu_s^{\mathrm{N,IS}})-n_u\right) - f\left(D_u(\mu_s^{\mathrm{N,IS}})\right)\right)\Big)^2\, ds. \end{aligned}$$

Proof. The proof proceeds from standard stochastic calculus for jump processes, using the SDEs driven by Poisson point processes (see the appendices of Part I of this volume or [45, 63]). □

3.3.4 Large graph limit

We prove that the rescaled degree distributions mentioned above can then be approximated for large N, by the solution $(\bar\mu_t^{\mathrm{S}}, \bar\mu_t^{\mathrm{IS}}, \bar\mu_t^{\mathrm{RS}})_{t\geq 0}$ of a system of deterministic measure-valued equations, with initial conditions $\bar\mu_0^{\mathrm{S}}$, $\bar\mu_0^{\mathrm{IS}}$ and $\bar\mu_0^{\mathrm{RS}}$.

We denote by $\bar S_t$ (resp. $\bar I_t$ and $\bar R_t$) the mass of the measure $\bar\mu_t^{\mathrm{S}}$ (resp. $\bar\mu_t^{\mathrm{IS}}$ and $\bar\mu_t^{\mathrm{RS}}$). As for the finite graph, $\bar\mu_t^{\mathrm{S}}/\bar S_t$ (resp. $\bar\mu_t^{\mathrm{IS}}/\bar I_t$ and $\bar\mu_t^{\mathrm{RS}}/\bar R_t$) is the probability degree distribution of the susceptible individuals (resp. the probability distribution of the degrees of the infectious and removed individuals towards the susceptible ones). For all $t \geq 0$, we denote by $\bar N_t^{\mathrm{S}} = \langle \bar\mu_t^{\mathrm{S}}, \chi\rangle$ (resp. $\bar N_t^{\mathrm{IS}} = \langle \bar\mu_t^{\mathrm{IS}}, \chi\rangle$ and $\bar N_t^{\mathrm{RS}} = \langle \bar\mu_t^{\mathrm{RS}}, \chi\rangle$) the continuous number of edges with ego in S (resp. I − S edges, R − S edges). Following Volz [111], pertinent quantities are the proportions $\bar p_t^{\mathrm{I}} = \bar N_t^{\mathrm{IS}}/\bar N_t^{\mathrm{S}}$ (resp. $\bar p_t^{\mathrm{R}} = \bar N_t^{\mathrm{RS}}/\bar N_t^{\mathrm{S}}$ and $\bar p_t^{\mathrm{S}} = (\bar N_t^{\mathrm{S}} - \bar N_t^{\mathrm{IS}} - \bar N_t^{\mathrm{RS}})/\bar N_t^{\mathrm{S}}$) of edges with infectious (respectively removed, susceptible) alter among those having susceptible ego. We also introduce

$$\theta_t = \exp\Big(-\lambda\int_0^t \bar p_s^{\mathrm{I}}\, \mathrm{d}s\Big) \tag{3.3.16}$$

the probability that a degree one node remains susceptible until time t. The limiting measure-valued equation expresses for any bounded real function f on $\mathbb{Z}_+$ as:

$$\langle \bar\mu_t^{\mathrm{S}}, f\rangle = \sum_{k\in\mathbb{Z}_+} \bar\mu_0^{\mathrm{S}}(k)\,\theta_t^k f(k), \tag{3.3.17}$$

$$\langle \bar\mu_t^{\mathrm{IS}}, f\rangle = \langle \bar\mu_0^{\mathrm{IS}}, f\rangle - \int_0^t \gamma\langle \bar\mu_s^{\mathrm{IS}}, f\rangle\, \mathrm{d}s \tag{3.3.18}$$

$$
\begin{aligned}
&+\int_0^t \sum_{k\in\mathbb{Z}_+} \lambda k \bar{p}_s^{\mathrm{I}} \sum_{\substack{j,\ell,m\in\mathbb{Z}_+\\ j+\ell+m=k-1}} \binom{k-1}{j,\ell,m} (\bar{p}_s^{\mathrm{I}})^j (\bar{p}_s^{\mathrm{R}})^\ell (\bar{p}_s^{\mathrm{S}})^m f(m) \bar{\mu}_s^{\mathrm{S}}(k)\,\mathrm{d}s \\
&+\int_0^t \sum_{k\in\mathbb{Z}_+} \lambda k \bar{p}_s^{\mathrm{I}} (1+(k-1)\bar{p}_s^{\mathrm{I}}) \sum_{k'\in\mathbb{N}} \big(f(k'-1)-f(k')\big) \frac{k'\bar{\mu}_s^{\mathrm{IS}}(k')}{\bar{N}_s^{\mathrm{IS}}} \bar{\mu}_s^{\mathrm{S}}(k)\,\mathrm{d}s, \\
\langle \bar{\mu}_t^{\mathrm{RS}}, f\rangle =& \langle \bar{\mu}_0^{\mathrm{RS}}, f\rangle + \int_0^t \gamma \langle \bar{\mu}_s^{\mathrm{IS}}, f\rangle\,\mathrm{d}s \\
&+\int_0^t \sum_{k\in\mathbb{Z}_+} \lambda k \bar{p}_s^{\mathrm{I}} (k-1) \bar{p}_s^{\mathrm{R}} \sum_{k'\in\mathbb{N}} \big(f(k'-1)-f(k')\big) \frac{k'\bar{\mu}_s^{\mathrm{RS}}(k')}{\bar{N}_s^{\mathrm{RS}}} \bar{\mu}_s^{\mathrm{S}}(k)\,\mathrm{d}s.
\end{aligned}
\tag{3.3.19}
$$

Let us give a heuristic explanation of Equations (3.3.17)–(3.3.19). Notice that the limiting graph is infinite. The probability that an individual of degree k has been infected by none of her k edges is θ_t^k and Equation (3.3.17) follows. In Equation (3.3.18), the first integral corresponds to infectious individuals being removed. In the second integral, $\lambda k \bar{p}_s^{\mathrm{I}}$ is the rate of infection of a given susceptible individual of degree k. Once she gets infected, the multinomial term determines the number of edges connected to susceptible, infectious and removed neighbours. Multi-edges are not encountered in the limiting graph. Each infectious neighbour has a degree chosen according to the size-biased distribution $k'\bar{\mu}^{\mathrm{IS}}(k')/\bar{N}^{\mathrm{IS}}$ and the number of edges to S is reduced by 1. This explains the third integral. Similar arguments explain Equation (3.3.19).

Before stating the theorem, let us introduce the following state space. For any $\varepsilon \geq 0$ and $A > 0$, we define the following closed set of $\mathscr{M}_F(\mathbb{Z}_+)$ as

$$
\mathscr{M}_{\varepsilon,A} = \{\nu \in \mathscr{M}_F(\mathbb{Z}_+)\,;\, \langle \nu, \mathbf{1}+\chi^5\rangle \leq A \text{ and } \langle \nu, \chi\rangle \geq \varepsilon\}
\tag{3.3.20}
$$

and $\mathscr{M}_{0+,A} = \cup_{\varepsilon>0}\mathscr{M}_{\varepsilon,A}$.

Theorem 3.3.8. *Suppose that Assumption 3.3.5 holds and that there exists an $A > 0$ such that*

$$
\big(\mu_0^{\mathrm{N,S}}, \mu_0^{\mathrm{N,IS}}, \mu_0^{\mathrm{N,RS}}\big) \text{ in } (\mathscr{M}_{0,A})^3 \text{ for any } N, \text{ with } \langle \bar{\mu}_0^{\mathrm{IS}}, \chi\rangle > 0.
\tag{3.3.21}
$$

Then, as N converges to infinity, the sequence $(\mu^{\mathrm{N,S}}, \mu^{\mathrm{N,IS}}, \mu^{\mathrm{N,RS}})_{N\in\mathbb{N}}$ converges in distribution in $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A}^3)$ to $(\bar{\mu}^{\mathrm{S}}, \bar{\mu}^{\mathrm{IS}}, \bar{\mu}^{\mathrm{RS}})$ which is the unique solution of the deterministic system equations (3.3.17)–(3.3.19) in $\mathscr{C}(\mathbb{R}_+, \mathscr{M}_{0,A} \times \mathscr{M}_{0+,A} \times \mathscr{M}_{0,A})$.

The proof is detailed in Section 3.3.7 and follows standard arguments. First, tightness of the process is proved using the Roelly and Aldous–Rebolledo criteria [100, 67]. Then, the convergence of the generators is studied, which allows us to identify the limit, provided the number of edges S − I remains of order at least εN. For proving uniqueness of the limiting value, we show using Gronwall's lemma that any two solutions of the limiting equation have the same mass and the same moments of order 1 and 2. This allows us to show the uniqueness of the generating function of $\bar{\mu}^{\mathrm{IS}}$ which solves a transport equation.

The assumption of moments of order 5 are needed for the convergence of the generators and discussed in Section 3.3.6.

3.3.5 Ball–Neal and Volz' equations

Choosing $f(k) = \mathbf{1}_i(k)$, we obtain the following countable system of ordinary differential equations (ODEs).

$$
\begin{aligned}
\bar{\mu}_t^{\mathrm{S}}(i) =& \bar{\mu}_0^{\mathrm{S}}(i)\theta_t^i, \\
\bar{\mu}_t^{\mathrm{IS}}(i) =& \bar{\mu}_0^{\mathrm{IS}}(i) - \int_0^t \gamma \bar{\mu}_s^{\mathrm{IS}}(i)\,\mathrm{d}s \\
&+\int_0^t \lambda \bar{p}_s^{\mathrm{I}} \sum_{j,\ell\geq 0} (i+j+\ell+1)\bar{\mu}_s^{\mathrm{S}}(i+j+\ell+1) \binom{i+j+\ell}{i,j,\ell} (\bar{p}_s^{\mathrm{S}})^i (\bar{p}_s^{\mathrm{I}})^j (\bar{p}_s^{\mathrm{R}})^\ell\,\mathrm{d}s
\end{aligned}
$$

$$+\int_0^t \left(\lambda(\bar{p}_s^{\mathrm{I}})^2\langle\bar{\mu}_s^{\mathrm{S}},\chi^2-\chi\rangle+\lambda\bar{p}_s^{\mathrm{I}}\langle\bar{\mu}_s^{\mathrm{S}},\chi\rangle\right)\frac{(i+1)\bar{\mu}_s^{\mathrm{IS}}(i+1)-i\bar{\mu}_s^{\mathrm{IS}}(i)}{\langle\bar{\mu}_s^{\mathrm{IS}},\chi\rangle}\,\mathrm{d}s,$$

$$\begin{aligned}\bar{\mu}_t^{\mathrm{RS}}(i)=&\bar{\mu}_0^{\mathrm{RS}}(i)\\&+\int_0^t\left\{\beta\bar{\mu}_s^{\mathrm{IS}}(i)+\lambda\bar{p}_s^{\mathrm{I}}\langle\bar{\mu}_s^{\mathrm{S}},\chi^2-\chi\rangle\bar{p}_s^{\mathrm{R}}\frac{(i+1)\bar{\mu}_s^{\mathrm{RS}}(i+1)-i\bar{\mu}_s^{\mathrm{RS}}(i)}{\langle\bar{\mu}_s^{\mathrm{RS}},\chi\rangle}\right\}\mathrm{d}s,\end{aligned}\tag{3.3.22}$$

It is noteworthy to say that this system corresponds to that in Ball and Neal [13].

The system (3.3.17)–(3.3.19) allows us to recover the equations proposed by Volz [111, Table 3, p. 297] (see also Proposition 3.2.3). The latter are obtained directly from (3.3.17)–(3.3.19) and the definitions of $\bar{S}_t$, $\bar{I}_t$, $\bar{p}_t^{\mathrm{I}}$ and $\bar{p}_t^{\mathrm{S}}$ which relate these quantities to the measures $\bar{\mu}_t^{\mathrm{S}}$ and $\bar{\mu}_t^{\mathrm{IS}}$. Let

$$h(z)=\sum_{k\in\mathbb{Z}_+}\bar{\mu}_0^{\mathrm{S}}(k)z^k\tag{3.3.23}$$

be the generating function for the initial degree distribution of the susceptible individuals $\bar{\mu}_0^{\mathrm{S}}$. This generating function is *a priori* different from the generating function of the degree distribution of the total CM graph: $g(z)=\sum_{k\in\mathbb{Z}_+}p_kz^k$. Let also $\theta_t=\exp(-\lambda\int_0^t\bar{p}_s^{\mathrm{I}}\,\mathrm{d}s)$. Then:

$$\bar{S}_t=\langle\bar{\mu}_t^{\mathrm{S}},\mathbf{1}\rangle=h(\theta_t),\tag{3.3.24}$$

$$\bar{I}_t=\langle\bar{\mu}_t^{\mathrm{IS}},\mathbf{1}\rangle=\bar{I}_0+\int_0^t\left(\lambda\bar{p}_s^{\mathrm{I}}\theta_sh'(\theta_s)-\gamma\bar{I}_s\right)\mathrm{d}s,\tag{3.3.25}$$

$$\bar{p}_t^{\mathrm{I}}=\bar{p}_0^{\mathrm{I}}+\int_0^t\left(\lambda\,\bar{p}_s^{\mathrm{I}}\bar{p}_s^{\mathrm{S}}\theta_s\frac{h''(\theta_s)}{h'(\theta_s)}-\lambda\,\bar{p}_s^{\mathrm{I}}(1-\bar{p}_s^{\mathrm{I}})-\gamma\bar{p}_s^{\mathrm{I}}\right)\mathrm{d}s,\tag{3.3.26}$$

$$\bar{p}_t^{\mathrm{S}}=\bar{p}_0^{\mathrm{S}}+\int_0^t\lambda\,\bar{p}_s^{\mathrm{I}}\bar{p}_s^{\mathrm{S}}\left(1-\theta_s\frac{h''(\theta_s)}{h'(\theta_s)}\right)\mathrm{d}s.\tag{3.3.27}$$

Here, the graph structure appears through the generating function g. In (3.3.25), we see that the classical contamination terms $\lambda\bar{S}_t\bar{I}_t$ (mass action) or $\lambda\bar{S}_t\bar{I}_t/(\bar{S}_t+\bar{I}_t)$ (frequency dependence) of mixing SIR models (e.g. Part I of this volume or [5, 38]) are replaced by $\lambda\bar{p}_t^{\mathrm{I}}\theta_th'(\theta_t)=\lambda\bar{N}_t^{\mathrm{IS}}$. The fact that new infectious individuals are chosen in the size-biased distribution is hidden in the term $h''(\theta_t)/h'(\theta_t)$.

Proposition 3.3.9. *The system* (3.3.17)–(3.3.19) *implies Volz' equations* (3.3.24)–(3.3.27).

Before proving Proposition 3.3.9, we begin with a corollary of Theorem 3.3.8.

Corollary 3.3.10. *For all* $t\in\mathbb{R}_+$

$$\begin{aligned}\bar{N}_t^{\mathrm{S}}=&\theta_th'(\theta_t)\\\bar{N}_t^{\mathrm{IS}}=&\bar{N}_0^{\mathrm{IS}}+\int_0^t\lambda\,\bar{p}_s^{\mathrm{I}}\theta_sh'(\theta_s)\left((\bar{p}_s^{\mathrm{S}}-\bar{p}_s^{\mathrm{I}})\theta_s\frac{h''(\theta_s)}{h'(\theta_s)}-1\right)-\gamma\bar{N}_s^{\mathrm{IS}}\;\;ds\\\bar{N}_t^{\mathrm{RS}}=&\int_0^t\left(\gamma\bar{N}_s^{\mathrm{IS}}-\lambda\,\bar{p}_s^{\mathrm{R}}\bar{p}_s^{\mathrm{I}}\theta_s^2h''(\theta_s)\right)ds.\end{aligned}\tag{3.3.28}$$

Proof. In the proof of Proposition 3.3.12, we will show below that when $N\to+\infty$, $(N_t^{\mathrm{N,IS}})_{N\in\mathbb{N}}$ converges uniformly, as $N\to+\infty$ and on compact intervals $[0,T]$, and in probability to the deterministic and continuous solution $\bar{N}^{\mathrm{IS}}$ such that for all t, $\bar{N}_t^{\mathrm{IS}}=\langle\bar{\mu}_t^{\mathrm{IS}},\chi\rangle$. (3.3.17) with $f=\chi$ reads

$$\bar{N}_t^{\mathrm{S}}=\sum_{k\in\mathbb{Z}_+}\bar{\mu}_0^{\mathrm{S}}(k)k\theta_t^k=\theta_t\sum_{k=1}^{+\infty}\bar{\mu}_0^{\mathrm{S}}(k)k\theta_t^{k-1}=\theta_th'(\theta_t),\tag{3.3.29}$$

i.e. the first assertion of (3.3.28).

Choosing $f = \chi$ in (3.3.18), we obtain

$$\bar{N}_t^{\mathrm{IS}} = \bar{N}_0^{\mathrm{IS}} - \int_0^t \gamma \bar{N}_s^{\mathrm{IS}}\,\mathrm{d}s + \int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s(k) \sum_{j+\ell\le k-1} \big(k-2j-2-\ell\big) \times \Big[\frac{(k-1)!}{j!(k-1-j-\ell)!\ell!}(\bar{p}_s^{\mathrm{I}})^j(\bar{p}_s^{\mathrm{R}})^\ell(\bar{p}_s^{\mathrm{S}})^{k-1-j-\ell}\Big]\bar{\mu}_s^{\mathrm{S}}(k)\,\mathrm{d}s.$$

Notice that the term in the square brackets is the probability of obtaining $(j,\ell,k-1-j-\ell)$ from a draw in the multinomial distribution of parameters $(k-1,(\bar{p}_s^{\mathrm{I}},\bar{p}_s^{\mathrm{R}},\bar{p}_s^{\mathrm{S}}))$. Hence,

$$\sum_{j+\ell\le k-1} j \times \Big(\frac{(k-1)!}{j!(k-1-j-\ell)!\ell!}(\bar{p}_s^{\mathrm{I}})^j(\bar{p}_s^{\mathrm{R}})^\ell(\bar{p}_s^{\mathrm{S}})^{k-1-j-\ell}\Big) = (k-1)\bar{p}_s^{\mathrm{I}}$$

as we recognize the mean number of edges to I_s of an individual of degree k. Other terms are treated similarly. Hence, with the definition of $\Lambda_s(k)$, (3.3.2),

$$\bar{N}_t^{\mathrm{IS}} = \bar{N}_0^{\mathrm{IS}} + \int_0^t \lambda\,\bar{p}_s^{\mathrm{I}}\Big(\langle\bar{\mu}_s^{\mathrm{S}},\chi^2-2\chi\rangle - (2\bar{p}_s^{\mathrm{I}}+\bar{p}_s^{\mathrm{R}})\langle\bar{\mu}_s^{\mathrm{S}},\chi(\chi-1)\rangle\Big)\,\mathrm{d}s - \int_0^t \gamma\bar{N}_s^{\mathrm{IS}}\,\mathrm{d}s.$$

But since

$$\langle\bar{\mu}_t^{\mathrm{S}},\chi(\chi-1)\rangle = \sum_{k\in\mathbb{Z}_+}\bar{\mu}_0^{\mathrm{S}}(k)k(k-1)\theta_t^k = \theta_t^2 h''(\theta_t)$$
$$\langle\bar{\mu}_t^{\mathrm{S}},\chi^2-2\chi\rangle = \langle\bar{\mu}_t^{\mathrm{S}},\chi(\chi-1)\rangle - \langle\bar{\mu}_t^{\mathrm{S}},\chi\rangle = \theta_t^2 h''(\theta_t) - \theta_t h'(\theta_t),$$

we obtain by noticing that $1-2\bar{p}_s^{\mathrm{I}}-\bar{p}_s^{\mathrm{R}} = \bar{p}_s^{\mathrm{S}}-\bar{p}_s^{\mathrm{I}}$,

$$\bar{N}_t^{\mathrm{IS}} = \bar{N}_0^{\mathrm{IS}} + \int_0^t \lambda\,\bar{p}_s^{\mathrm{I}}\Big((\bar{p}_s^{\mathrm{S}}-\bar{p}_s^{\mathrm{I}})\theta_s^2 h''(\theta_s) - \theta_s h'(\theta_s)\Big)\,\mathrm{d}s - \int_0^t \gamma\bar{N}_s^{\mathrm{IS}}\,\mathrm{d}s, \tag{3.3.30}$$

which is the second assertion of (3.3.28). The third equation is obtained similarly. □

We are now ready to prove Volz' equations:

Proof of Proposition 3.3.9. We begin with the proof of (3.3.24) and (3.3.25). Fix again $t \ge 0$. For the size of the susceptible population, taking $f = \mathbf{1}$ in (3.3.17) gives (3.3.24). For the size of the infective population, setting $f = \mathbf{1}$ in (3.3.18) entails

$$\begin{aligned}\bar{I}_t =& \bar{I}_0 + \int_0^t \Big(\sum_{k\in\mathbb{Z}_+}\lambda k\bar{p}_s^{\mathrm{I}}\bar{\mu}_s^{\mathrm{S}}(k) - \gamma\bar{I}_s\Big)\,\mathrm{d}s\\ =& \bar{I}_0 + \int_0^t \Big(\lambda\bar{p}_s^{\mathrm{I}}\sum_{k\in\mathbb{Z}_+}\bar{\mu}_0^{\mathrm{S}}(k)k\theta_s^k - \gamma\bar{I}_s\Big)\,\mathrm{d}s\\ =& \bar{I}_0 + \int_0^t \Big(\lambda\bar{p}_s^{\mathrm{I}}\theta_s h'(\theta_s) - \gamma\bar{I}_s\Big)\,\mathrm{d}s\end{aligned}$$

by using (3.3.17) with $f = \chi$ for the second equality.

Let us now consider the probability that an edge with a susceptible ego has an infectious alter. Both equations (3.3.24) and (3.3.25) depend on $\bar{p}_t^{\mathrm{I}} = \bar{N}_t^{\mathrm{IS}}/\bar{N}_t^{\mathrm{S}}$. It is thus important to obtain an equation for this quantity. In Volz [111], this equation also leads to introduce the quantity $\bar{p}_t^{\mathrm{S}}$.
From Corollary 3.3.10, we see that $\bar{N}^{\mathrm{S}}$ and $\bar{N}^{\mathrm{IS}}$ are differentiable and:

$$\frac{\mathrm{d}\bar{p}_t^{\mathrm{I}}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\bar{N}_t^{\mathrm{IS}}}{\bar{N}_t^{\mathrm{S}}}\Big) = \frac{1}{\bar{N}_t^{\mathrm{S}}}\frac{\mathrm{d}}{\mathrm{d}t}(\bar{N}_t^{\mathrm{IS}}) - \frac{\bar{N}_t^{\mathrm{IS}}}{(\bar{N}_t^{\mathrm{S}})^2}\frac{\mathrm{d}}{\mathrm{d}t}(\bar{N}_t^{\mathrm{S}})$$

$$
\begin{aligned}
=&\Big(\lambda\bar{p}_t^{\mathrm{I}}(\bar{p}_t^{\mathrm{S}}-\bar{p}_t^{\mathrm{I}})\theta_t\frac{h''(\theta_t)}{h'(\theta_t)}-\lambda\bar{p}_t^{\mathrm{I}}-\gamma\bar{p}_t^{\mathrm{I}}\Big)\\
&\quad-\Big(\frac{\bar{p}_t^{\mathrm{I}}}{\theta_t h'(\theta_t)}\big(-\lambda\bar{p}_t^{\mathrm{I}}\theta_t h'(\theta_t)+\theta_t h''(\theta_t)(-\lambda\bar{p}_t^{\mathrm{I}}\theta_t)\big)\Big)\\
=&\lambda\bar{p}_t^{\mathrm{I}}\bar{p}_t^{\mathrm{S}}\theta_t\frac{h''(\theta_t)}{h'(\theta_t)}-\lambda\bar{p}_t^{\mathrm{I}}(1-\bar{p}_t^{\mathrm{I}})-\gamma\bar{p}_t^{\mathrm{I}},
\end{aligned}
$$

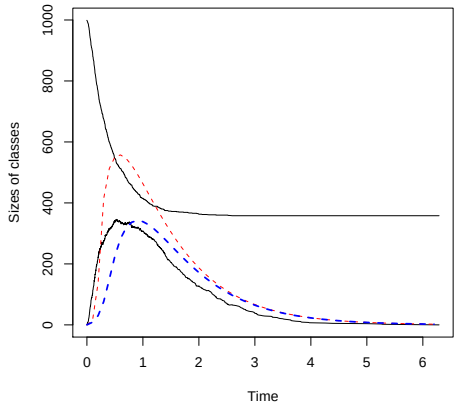
by using (3.3.28) for the derivatives of $\bar{N}^{\mathrm{S}}$ and $\bar{N}^{\mathrm{IS}}$ in the second line. This achieves the proof of (3.3.26).

For (3.3.27), we notice that $\bar{p}_t^{\mathrm{S}}=1-\bar{p}_t^{\mathrm{I}}-\bar{p}_t^{\mathrm{R}}$ and achieve the proof by showing that

$$
\bar{p}_t^{\mathrm{R}}=\int_0^t\Big(\gamma\bar{p}_s^{\mathrm{I}}-\lambda\bar{p}_s^{\mathrm{I}}\bar{p}_s^{\mathrm{R}}\Big)\,\mathrm{d}s \tag{3.3.31}
$$

by using arguments similar as for $\bar{p}_t^{\mathrm{I}}$. □

3.3.6 Degree distribution of the "initial condition"



The assumption of moments of order 5 in the Theorem 3.3.8 may seem restrictive. Janson et al. [66] showed that this assumption was not necessary if Volz' equations are established by considering the process $(S_t^{\mathrm{N}},I_t^{\mathrm{N}},R_t^{\mathrm{N}},N_t^{\mathrm{N,S}},N_t^{\mathrm{N,I}},N_t^{\mathrm{N,R}})_{t\in\mathbb{R}_+}$ where $N_t^{\mathrm{N,S}}=\langle\mu_t^{\mathrm{N,S}},k\rangle$, $N_t^{\mathrm{N,I}}$ and $N_t^{\mathrm{N,R}}$ are respectively the numbers of half-edges of the susceptible, infectious and removed individuals that are not attached to the cluster. This process contains less information than the process $(\mu_t^{\mathrm{N,S}},\mu_t^{\mathrm{N,IS}},\mu_t^{\mathrm{N,RS}})_{t\in\mathbb{R}_+}$, and an assumption on the existence of moments of order 2 uniformly bounded in N is sufficient. Janson and coauthors emphasize that if we allow the CM graph to have self-loops and multiple edges, then only the uniform integrability of the degree distribution of an individual chosen at random is needed, which seems to be the minimal assumption...

However, when considering the beginning of the epidemics, it appears that the assumption corresponding to Equation (3.3.21) is not so restrictive. Indeed, we emphasize that it should be distinguished between the degree distribution of the graph $\mathbf{p}$, associated with the generating function g, and the degree distribution of the S individuals when the proportion of infectious individuals has reached a non-negligible value, and which we associate with the generating function h. If we consider the degree distribution of the susceptible individuals, we see that the individuals with highest degrees will be infected first, since individuals are chosen in the size-biased distribution (1.2.1) when pairing the half-edges at random. After the $[\varepsilon N]$ first infections, with $\varepsilon>0$, when the Theorem 3.3.8 starts to apply, all the susceptible individuals of highest degree have disappeared from $\mu^{\mathrm{N,S}}$. Then, $\mu^{\mathrm{N,S}}$ will even admit exponential moments.

For a population of size N, whose individuals have degrees $D_1, \dots D_N$, let us define, for all $k \in \mathbb{Z}_+$, the number of vertices with degree k among them by

$$N_k^{\mathrm{N}} = \mathrm{Card}\{u \in \{1,\dots,N\},\, D_u = k\}.$$

To each of the D_u half-edges of individual u, we associate an independent uniform random variable on $[0,1]$. The vertex u is infected before the vertex v if the minimal value Z_u of the random variables attached to its half-edges is smaller than the minimal value Z_v of the random variables attached to the half-edges of v. This construction has been used by Riordan [98] and is related to size-biased orderings.

Proposition 3.3.11. *(i) The degree distribution $(\widehat{p}_k^{\varepsilon,\mathrm{N}})_{k\geq 1}$ of the remaining susceptible individuals after the $[\varepsilon N]$ first infections is:*

$$\widehat{p}_k^{\varepsilon,\mathrm{N}} = \frac{1}{N-[\varepsilon N]} \sum_{u=1}^{N} \mathbf{1}_{D_u=k} \mathbf{1}_{Z_u > Z_{([\varepsilon N])}} \tag{3.3.32}$$

where $(Z_{(1)},\dots,Z_{(N)})$ are the order statistics of $(Z_1,\dots,Z_N)$, and where

$$\mathbb{P}(Z_u \leq z \,|\, D_u) = 1-(1-z)^{D_u}.$$

(ii) For $z \in (0,1)$, let $M(z)$ be the survival function of the distribution of the Z_i and let $M_N(z)$ be the empirical survival function of $(Z_1,\dots,Z_N)$:

$$M_N(z) = \frac{1}{N}\sum_{u=1}^{N} \mathbf{1}_{Z_u>z}, \quad \text{and} \quad M(z) = \sum_{k\geq 0} p_k (1-z)^k = g(1-z),$$

where $g(z) = \sum_{k\geq 0} p_k z^k$ is the generating function of the degree distribution $\boldsymbol{p}$ of the CM graph. Let ε defined by $z^\varepsilon = \inf\{z \in (0,1),\, M(z) \geq \varepsilon\}$ be the quantile of order ε of the Z_u. Then, provided M is continuous and strictly increasing at z^ε,

$$\lim_{N\to+\infty} Z_{[\varepsilon N]} = z^\varepsilon \qquad \text{almost surely}.$$

(iii) For such an ε, the degree distribution of the remaining susceptible individuals after the $[\varepsilon N]$ first infections converges weakly to:

$$\lim_{N\to+\infty} \sum_{k\geq 0} \widehat{p}_k^{\varepsilon,\mathrm{N}} \delta_k = \frac{1}{1-\varepsilon} \sum_{k\geq 0} p_k (1-z^\varepsilon)^k \,\delta_k, \tag{3.3.33}$$

where z^ε is solution of $1-\varepsilon = g(1-z^\varepsilon)$. Moreover, we have convergence of the moments of order 5:

$$\lim_{N\to+\infty} \sum_{k\geq 0} k^5 \widehat{p}_k^{\varepsilon,\mathrm{N}} = \frac{1}{1-\varepsilon} \sum_{k\geq 0} k^5 p_k (1-z^\varepsilon)^k < +\infty. \tag{3.3.34}$$

In particular, the limiting distribution (3.3.33) admits moments of all orders.

Proof. Let us prove (ii). Let $z \in [0,1]$. The proportion of vertices of degree k whose minimal value of the Z_u is smaller than z is $M_k^N(z) = \frac{1}{N}\sum_{u=1}^N \mathbf{1}_{D_u=k}\mathbf{1}_{Z_u>z}$. By the law of large numbers, $\lim_{N\to+\infty} M_k^N(z) = p_k(1-z)^k$ a.s., which implies that

$$\lim_{N\to+\infty} \frac{1}{N}\sum_{u=1}^{N} \mathbf{1}_{Z_u>z}\delta_{D_u} = \sum_{k\geq 0} p_k (1-z)^k \delta_k$$

for the weak convergence.

Assume that $\varepsilon > 0$ is such that M is continuous and strictly increasing at z^ε. Then, $M(z^\varepsilon) = \varepsilon$. Let $\delta > 0$ and

$$\eta = \min(|M(z^\varepsilon - \delta) - M(z^\varepsilon)|, |M(z^\varepsilon + \delta) - M(z^\varepsilon)|).$$

By the Kolmogorov–Smirnov theorem: $\lim_{N\to+\infty}\|M_N - M\|_\infty = 0$. Then, there exists $\mathbb{P}(d\omega)$-a.s. an integer $N_0(\omega)$ sufficiently large such that for all $N \geq N_0$, $\|M_N - M\|_\infty < \eta/2$. Since M is non-decreasing and since $Z_{[\varepsilon N]}$ is such that $M_N(Z_{[\varepsilon N]}) = \frac{[\varepsilon N]}{N}$, then,

$$\begin{aligned}\big|M(Z_{([\varepsilon N])}) - \varepsilon\big| \leq& \big|M(Z_{([\varepsilon N])}) - M_N(Z_{([\varepsilon N])})\big| + \big|M_N(Z_{([\varepsilon N])}) - \varepsilon\big| \\ \leq& \frac{\eta}{2} + \Big|\frac{[\varepsilon N]}{N} - \varepsilon\Big|.\end{aligned}$$

Thus, for $N \geq \max(N_0, 2/\eta)$, $|M(Z_{[\varepsilon N]}) - \varepsilon| < \eta$ and hence $Z_{[\varepsilon N]} \in (z^\varepsilon - \delta, z^\varepsilon + \delta)$ a.s. This implies that $(Z_{[\varepsilon N]})_{N\geq 1}$ converges a.s. to z^ε.

If $(\widehat{p}_k^{\varepsilon,N})_{k\in\mathbb{Z}_+}$ is the degree distribution after the $[\varepsilon N]$ first infections, then

$$\lim_{N\to+\infty}\sum_{k\geq 0}\widehat{p}_k^{\varepsilon,N}\delta_k = \frac{1}{1-\varepsilon}\sum_{k\geq 0}p_k(1-z^\varepsilon)^k\delta_k. \tag{3.3.35}$$

The convergence, for every $k \in \mathbb{Z}_+$, of $\widehat{p}_k^{\varepsilon,N}$ to $p_k(1-z^\varepsilon)^k/(1-\varepsilon)$ implies the convergence of (3.3.35) for the vague topology. Because (3.3.35) deals with probability measures, the criterion of [77, Proposition 2] implies that the convergence also holds for the weak topology.

Since

$$\lim_{N\to+\infty}\mathbb{E}\Big(\frac{1}{N}\sum_{i=1}^{N}\big|\mathbf{1}_{Z_i>Z_{[\varepsilon N]}} - \mathbf{1}_{Z_i>z^\varepsilon}\big|\mathbf{1}_{d_i=k}\Big) = \lim_{N\to+\infty}\mathbb{P}\Big(Z_1 \in [Z_{[\varepsilon N]}\wedge z^\varepsilon, Z_{[\varepsilon N]}\vee z^\varepsilon],\ \ d_1 = k\Big) = 0, \tag{3.3.36}$$

and since $N/(N-[\varepsilon N])$ converges to $1/(1-\varepsilon)$, we obtain (3.3.33).

For the convergence of the moments of order 5, we notice that for large $K \in \mathbb{N}$,

$$\begin{aligned}&\mathbb{E}\Big(\Big|\sum_{k\geq 0}k^5 M_k^N(Z_{[\varepsilon N]}) - \sum_{k\geq 0}k^5 p_k(1-z^\varepsilon)^k\Big|\Big) \\ \leq&\mathbb{E}\Big(\Big|\sum_{k\leq K}k^5\big(M_k^N(Z_{[\varepsilon N]}) - p_k(1-z^\varepsilon)^k\big)\Big|\Big) + \mathbb{E}\Big(\sum_{k>K}k^5 M_k^N(Z_{[\varepsilon N]})\Big) + \sum_{k>K}k^5 p_k(1-z^\varepsilon)^k.\end{aligned}$$

The first term converges to 0 with the preceding arguments. The third term is controlled for K sufficiently large. For the second term, we use that for all $z \in (0,1)$,

$$\mathbb{E}\big(\sum_{k>K}k^5 M_k^N(z)\big) = \sum_{k>K}k^5 p_k(1-z)^k$$

and that $Z_{[\varepsilon N]}$ converges a.s. to z^ε. □

3.3.7 Proof of the limit theorem

We now prove Theorem 3.3.8.

In the proof, we will see that the epidemic remains large and described by a deterministic equation provided the number of edges from I to S remains of the order of N. Let us thus define, for all $\varepsilon > 0$, $\varepsilon' > 0$ and $n \in \mathbb{N}$,

$$t_{\varepsilon'} := \inf\{t \geq 0, \langle\bar{\mu}_t^{\mathrm{IS}}, \chi\rangle < \varepsilon'\} \tag{3.3.37}$$

and:

$$\tau_\varepsilon^N = \inf\{t \geq 0, \langle\mu_t^{\mathrm{N,IS}}, \chi\rangle < \varepsilon\}. \tag{3.3.38}$$

In the sequel, we choose $0 < \varepsilon < \varepsilon' < \langle\bar{\mu}_0^{\mathrm{IS}}, \chi\rangle$.

Step 1 Let us prove that $(\mu^{\mathrm{N,S}}, \mu^{\mathrm{N,IS}}, \mu^{\mathrm{N,RS}})_{N\in\mathbb{N}}$ is tight. Let $t \in \mathbb{R}_+$ and $N \in \mathbb{N}$. By hypothesis, we have that

$$\langle \mu_t^{\mathrm{N,S}}, \mathbf{1}+\chi^5\rangle + \langle \mu_t^{\mathrm{N,IS}}, \mathbf{1}+\chi^5\rangle + \langle \mu_t^{\mathrm{N,RS}}, \mathbf{1}+\chi^5\rangle \le \langle \mu_0^{\mathrm{N,S}}, \mathbf{1}+\chi^5\rangle + \langle \mu_0^{\mathrm{N,IS}}, \mathbf{1}+\chi^5\rangle \le 2A. \quad (3.3.39)$$

Thus the sequences of marginals $(\mu_t^{\mathrm{N,S}})_{N\in\mathbb{N}}$, $(\mu_t^{\mathrm{N,IS}})_{N\in\mathbb{N}}$ and $(\mu_t^{\mathrm{N,RS}})_{N\in\mathbb{N}}$ are tight for each $t\in\mathbb{R}_+$. Now by the criterion of Roelly [100], it remains to prove that for each bounded function f on $\mathbb{Z}_+$, the sequence $(\langle \mu_\cdot^{\mathrm{N,S}}, f\rangle, \langle \mu_\cdot^{\mathrm{N,IS}}, f\rangle, \langle \mu_\cdot^{\mathrm{N,RS}}, f\rangle)_{N\in\mathbb{N}}$ is tight in $\mathbb{D}(\mathbb{R}_+,\mathbb{R}^3)$. Since we have the semi-martingale decompositions of these processes, it is sufficient, by using the Rebolledo criterion, to prove that the finite variation part and the bracket of the martingale satisfy the Aldous criterion (see e.g. [67]). We only prove that $\langle \mu_\cdot^{\mathrm{N,IS}}, f\rangle$ is tight. The computations are similar for the other components.

The Rebolledo–Aldous criterion is satisfied if for all $\alpha>0$ and $\eta>0$ there exists $N_0\in\mathbb{Z}_+$ and $\delta>0$ such that for all $N>N_0$ and for all stopping times S_N and T_N such that $S_N<T_N<S_N+\delta$,

$$\begin{aligned} &\mathbb{P}\big(|A_{T_N}^{\mathrm{N,IS},f} - A_{S_N}^{\mathrm{N,IS},f}| > \eta\big) \le \alpha, \quad \text{and} \\ &\mathbb{P}\big(|\langle M^{\mathrm{N,IS},f}\rangle_{T_N} - \langle M^{\mathrm{N,IS},f}\rangle_{S_N}| > \eta\big) \le \alpha. \end{aligned} \quad (3.3.40)$$

For the finite variation part,

$$\begin{aligned} \mathbb{E}\Big[|A_{T_N}^{\mathrm{N,IS},f} - A_{S_N}^{\mathrm{N,IS},f}|\Big] \le{}& \mathbb{E}\left[\int_{S_N}^{T_N} \gamma\|f\|_\infty \langle \mu_s^{\mathrm{N,IS}}, 1\rangle \,\mathrm{d}s\right] \\ &+ \mathbb{E}\left[\int_{S_N}^{T_N} \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell\le k-1} p_s^N(j,\ell|k-1)(2j+3)\|f\|_\infty \,\mathrm{d}s\right]. \end{aligned}$$

The term $\sum_{j+\ell\le k-1} j p_s^N(j,\ell|k-1)$ is the mean number of links to I_{s-}^N that the newly infected individual has, given that this individual is of degree k. It is bounded by k. Then, with (3.3.13),

$$\mathbb{E}\Big[|A_{T_N}^{\mathrm{N,IS},f} - A_{S_N}^{\mathrm{N,IS},f}|\Big] \le \delta\mathbb{E}\Big[\beta\|f\|_\infty(S_0^{\mathrm{N}}+I_0^{\mathrm{N}}) + \lambda\|f\|_\infty\langle \mu_0^{\mathrm{N,S}}, 2\chi^2+3\chi\rangle\Big],$$

by using that the number of infectives is bounded by the size of the population and that $\mu_s^{\mathrm{N,S}}(k)\le\mu_0^{\mathrm{N,S}}(k)$ for all k and $s\ge 0$. From (3.3.21), the r.h.s. is finite. Using Markov's inequality,

$$\mathbb{P}\big(|A_{T_N}^{\mathrm{N,IS},f} - A_{S_N}^{\mathrm{N,IS},f}| > \eta\big) \le \frac{(5\lambda+2\gamma)A\delta\|f\|_\infty}{\eta},$$

which is smaller than α for δ small enough.

We use the same arguments for the bracket of the martingale:

$$\begin{aligned} \mathbb{E}\big[|\langle M^{\mathrm{N,IS},f}\rangle_{T_N} - \langle M^{\mathrm{N,IS},f}\rangle_{S_N}|\big] &\le \mathbb{E}\Big[\frac{\delta\gamma\|f\|_\infty^2(S_0^{\mathrm{N}}+I_0^{\mathrm{N}})}{N} + \frac{\delta\lambda\|f\|_\infty^2\langle \mu_0^{\mathrm{N,S}}, \chi(2\chi+3)^2\rangle}{N}\Big] \\ &\le \frac{(25\lambda+2\gamma)A\delta\|f\|_\infty^2}{N}, \end{aligned} \quad (3.3.41)$$

using Assumption 3.3.5 and (3.3.21). The r.h.s. can be made smaller than $\eta\alpha$ for a small enough δ, so the second inequality of (3.3.40) follows again from Markov's inequality. By [100], this provides the tightness in $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A}^3)$, with $\mathscr{M}_{0,A}$ defined in (3.3.20).

By Prohorov's theorem (e.g. [53], p. 104) and Step 1, we obtain that the distributions of $(\mu^{\mathrm{N,S}}, \mu^{\mathrm{N,IS}}, \mu^{\mathrm{N,RS}})$, for $N\in\mathbb{N}$, form a relatively compact family of bounded measures on $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A}^3)$, and so do the laws of the stopped processes $(\mu_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,S}}, \mu_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,IS}}, \mu_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,RS}})_{N\in\mathbb{N}}$ (recall (3.3.38)). Because of the moment assumptions for the degree distributions, the limiting process is continuous. Let $\bar\mu := (\bar\mu^{\mathrm{S}}, \bar\mu^{\mathrm{IS}}, \bar\mu^{\mathrm{RS}})$ be a limiting point in $\mathscr{C}(\mathbb{R}_+, \mathscr{M}_{0,A}^3)$ of the sequence of stopped processes and let us consider a subsequence again denoted by $\mu^N := (\mu^{\mathrm{N,S}}, \mu^{\mathrm{N,IS}}, \mu^{\mathrm{N,RS}})_{N\in\mathbb{N}}$, with an abuse of notation, and that converges to $\bar\mu$. Because the limiting values are continuous, the convergence of $(\mu^N)_{N\in\mathbb{N}}$ to $\bar\mu$ holds for the uniform convergence on every compact subset of $\mathbb{R}_+$ (e.g. [24] p. 112).

Now, let us define for all $t \in \mathbb{R}_+$ and for all bounded functions f on $\mathbb{Z}_+$, the mappings $\Psi_t^{\mathrm{S},f}$, $\Psi_t^{\mathrm{IS},f}$ and $\Psi_t^{\mathrm{RS},f}$ from $\mathbb{D}\big(\mathbb{R}_+, \mathscr{M}_{0,A}^3\big)$ into $\mathbb{D}\big(\mathbb{R}_+, \mathbb{R}\big)$ such that (3.3.17)–(3.3.19) read

$$\big(\langle \bar\mu_t^{\mathrm{S}}, f\rangle, \langle \bar\mu_t^{\mathrm{IS}}, f\rangle, \langle \bar\mu_t^{\mathrm{RS}}, f\rangle\big) \quad = \quad \Big(\Psi_t^{\mathrm{S},f}\left(\bar\mu^{\mathrm{S}}, \bar\mu^{\mathrm{IS}}, \bar\mu^{\mathrm{RS}}\right), \Psi_t^{\mathrm{IS},f}\left(\bar\mu^{\mathrm{S}}, \bar\mu^{\mathrm{IS}}, \bar\mu^{\mathrm{RS}}\right), \Psi_t^{\mathrm{RS},f}\left(\bar\mu^{\mathrm{S}}, \bar\mu^{\mathrm{IS}}, \bar\mu^{\mathrm{RS}}\right)\Big). \quad (3.3.42)$$

Our purpose is to prove that the limiting values are the unique solution of (3.3.17)–(3.3.19).
Before proceeding to the proof, a remark is in order. A natural way of reasoning would be to prove that $\Psi^{\mathrm{S},f}$, $\Psi^{\mathrm{IS},f}$ and $\Psi^{\mathrm{RS},f}$ are Lipschitz continuous in some spaces of measures. To avoid doing so by considering the set of measures with moments of any order, which is a set too small for applications, we circumvent this difficulty by first proving that the mass and the first two moments of any solutions of the system are the same. Then, we prove that the generating functions of these measures satisfy a partial differential equation known to have a unique solution.

Step 2 We now prove that the differential system (3.3.17)–(3.3.19) has at most one solution in $\mathscr{C}(\mathbb{R}_+, \mathscr{M}_{0,A} \times \mathscr{M}_{0+,A} \times \mathscr{M}_{0,A})$. Let $T > 0$. Let $\bar\mu^i = (\bar\mu^{\mathrm{S},i}, \bar\mu^{\mathrm{IS},i}, \bar\mu^{\mathrm{RS},i})$, $i \in \{1,2\}$ be two solutions of (3.3.17)–(3.3.19), started with the same initial conditions in $\mathscr{M}_{0,A} \times \mathscr{M}_{\varepsilon,A} \times \mathscr{M}_{0,A}$ for some small $\varepsilon > 0$. Set

$$\Upsilon_t \quad = \quad \sum_{j=0}^{3} |\langle \bar\mu_t^{\mathrm{S},1}, \chi^j\rangle \; - \; \langle \bar\mu_t^{\mathrm{S},2}, \chi^j\rangle| \; + \; \sum_{j=0}^{2} \Big(|\langle \bar\mu_t^{\mathrm{IS},1}, \chi^j\rangle \; - \; \langle \bar\mu_t^{\mathrm{IS},2}, \chi^j\rangle| \; + \; |\langle \bar\mu_t^{\mathrm{RS},1}, \chi^j\rangle \; - \; \langle \bar\mu_t^{\mathrm{RS},2}, \chi^j\rangle|\Big).$$

Let us first remark that for all $0 \le t < T$, $\bar N_t^{\mathrm{S}} \ge \bar N_t^{\mathrm{IS}} > \varepsilon$ and then

$$\begin{aligned} |\bar p_t^{\mathrm{I},1} - \bar p_t^{\mathrm{I},2}| &= \Big|\frac{\bar N_t^{\mathrm{IS},1}}{\bar N_t^{\mathrm{S},1}} - \frac{\bar N_t^{\mathrm{IS},2}}{\bar N_t^{\mathrm{S},2}}\Big| \le \frac{A}{\varepsilon^2}\Big|\bar N_t^{\mathrm{S},1} - \bar N_t^{\mathrm{S},2}\Big| + \frac{1}{\varepsilon}\Big|\bar N_t^{\mathrm{IS},1} - \bar N_t^{\mathrm{IS},2}\Big| \\ &= \frac{A}{\varepsilon^2}\Big|\langle \bar\mu_t^{\mathrm{S},1}, \chi\rangle - \langle \bar\mu_t^{\mathrm{S},2}, \chi\rangle\Big| + \frac{1}{\varepsilon}\Big|\langle \bar\mu_t^{\mathrm{IS},1}, \chi\rangle - \langle \bar\mu_t^{\mathrm{IS},2}, \chi\rangle\Big| \le \frac{A}{\varepsilon^2}\Upsilon_t. \quad (3.3.43) \end{aligned}$$

The same computations show a similar result for $|\bar p_t^{\mathrm{S},1} - \bar p_t^{\mathrm{S},2}|$.

Using that $\bar\mu^i$ are solutions to (3.3.17)–(3.3.18) let us show that Υ satisfies a Gronwall inequality which implies that it is equal to 0 for all $t \le T$. For the degree distributions of the susceptible individuals, we have for $p \in \{0,1,2,3\}$ and $f = \chi^p$ in (3.3.17):

$$\begin{aligned} |\langle \bar\mu_t^{\mathrm{S},1}, \chi^p\rangle - \langle \bar\mu_t^{\mathrm{S},2}, \chi^p\rangle| &= \Big|\sum_{k\in\mathbb{Z}_+} \bar\mu_0^{\mathrm{S}}(k) k^p \big(e^{-\lambda\int_0^t \bar p_s^{\mathrm{I},1} ds} - e^{-\lambda\int_0^t \bar p_s^{\mathrm{I},2} ds}\big)\Big| \\ &\le \lambda \sum_{k\in\mathbb{Z}_+} k^p \bar\mu_0^{\mathrm{S}}(k) \int_0^t \big|\bar p_s^{\mathrm{I},1} - \bar p_s^{\mathrm{I},2}\big| ds \le \lambda \frac{A^2}{\varepsilon^2}\int_0^t \Upsilon_s ds, \end{aligned}$$

by using (3.3.43) and the fact that $\bar\mu_0^{\mathrm{S}} \in \mathscr{M}_{0,A}$.
For $\bar\mu^{\mathrm{IS}}$ and $\bar\mu^{\mathrm{RS}}$, we use (3.3.18) and (3.3.19) with the functions $f = \chi^0 = \mathbf{1}$, $f = \chi$ and $f = \chi^2$. We proceed here with only one of the computations, others can be done similarly. From (3.3.18):

$$\langle \bar\mu_t^{\mathrm{IS},1}, \mathbf{1}\rangle \; - \; \langle \bar\mu_t^{\mathrm{IS},2}, \mathbf{1}\rangle \quad = \quad \gamma\int_0^t \langle \bar\mu_s^{\mathrm{IS},1} \; - \; \bar\mu_s^{\mathrm{IS},2}, \mathbf{1}\rangle \, \mathrm{d}s \; + \; \lambda\int_0^t (\bar p_s^{\mathrm{I},1}\langle \bar\mu_s^{\mathrm{S},1}, \chi\rangle \; - \; \bar p_s^{\mathrm{I},2}\langle \bar\mu_s^{\mathrm{S},2}, \chi\rangle) \, \mathrm{d}s.$$

Hence, with (3.3.43),

$$\Big|\langle \bar\mu_t^{\mathrm{IS},1} - \bar\mu_t^{\mathrm{IS},2}, \mathbf{1}\rangle\Big| \le C(\lambda,\gamma,A,\varepsilon)\int_0^t \Upsilon_s \, \mathrm{d}s.$$

By analogous computations for the other quantities, we show that

$$\Upsilon_t \le C'(\lambda,\gamma,A,\varepsilon)\int_0^t \Upsilon_s \, \mathrm{d}s,$$

hence $\Upsilon \equiv 0$. It follows that for all $t < T$, and for all $j \in \{0,1,2\}$,

$$\langle \bar{\mu}_t^{\mathrm{S},1}, \chi^j \rangle = \langle \bar{\mu}_t^{\mathrm{S},2}, \chi^j \rangle \quad \text{and} \quad \langle \bar{\mu}_t^{\mathrm{IS},1}, \chi^j \rangle = \langle \bar{\mu}_t^{\mathrm{IS},2}, \chi^j \rangle, \tag{3.3.44}$$

and in particular, $\bar{N}_t^{\mathrm{S},1} = \bar{N}_t^{\mathrm{S},2}$ and $\bar{N}_t^{\mathrm{IS},1} = \bar{N}_t^{\mathrm{IS},2}$. This implies that $\bar{p}_t^{\mathrm{S},1} = \bar{p}_t^{\mathrm{S},2}$, $\bar{p}_t^{\mathrm{I},1} = \bar{p}_t^{\mathrm{I},2}$ and $\bar{p}_t^{\mathrm{R},1} = \bar{p}_t^{\mathrm{R},2}$. From (3.3.17), we have that $\bar{\mu}^{\mathrm{S},1} = \bar{\mu}^{\mathrm{S},2}$.

Our purpose is now to prove that $\bar{\mu}^{\mathrm{IS},1} = \bar{\mu}^{\mathrm{IS},2}$. Let us introduce the following generating functions: for any $t \in \mathbb{R}_+$, $i \in \{1,2\}$ and $\eta \in [0,1)$,

$$\mathscr{G}_t^i(\eta) = \sum_{k \geq 0} \eta^k \bar{\mu}_t^{\mathrm{IS},i}(k).$$

Since we already know that these measures have the same total mass, it remains to prove that $\mathscr{G}^1 \equiv \mathscr{G}^2$. Let us define

$$\begin{aligned} H(t,\eta) &= \int_0^t \sum_{k \in \mathbb{Z}_+} \lambda k \bar{p}_s^{\mathrm{I}} \sum_{\substack{j,\ell,m \in \mathbb{Z}_+ \\ j+\ell+m=k-1}} \binom{k-1}{j,\ell,m} (\bar{p}_s^{\mathrm{I}})^j (\bar{p}_s^{\mathrm{R}})^\ell (\bar{p}_s^{\mathrm{S}})^m \eta^m \bar{\mu}_s^{\mathrm{S}}(k)\, ds, \\ K_t &= \sum_{k \in \mathbb{Z}_+} \lambda k \bar{p}_t^{\mathrm{I}}(k-1) \bar{p}_t^{\mathrm{R}} \frac{\bar{\mu}_t^{\mathrm{S}}(k)}{\bar{N}_t^{\mathrm{IS}}}. \end{aligned} \tag{3.3.45}$$

The latter quantities are respectively of class $\mathscr{C}^1$ and $\mathscr{C}^0$ with respect to time t and are well-defined and bounded on $[0,T]$. Moreover, H and K do not depend on the chosen solution because of (3.3.44). Applying (3.3.18) to $f(k) = \eta^k$ yields

$$\begin{aligned} \mathscr{G}_t^i(\eta) =& \mathscr{G}_0^i(\eta) + H(t,\eta) + \int_0^t \Big(K_s \sum_{k' \in \mathbb{N}} \big(\eta^{k'-1} - \eta^{k'}\big) k' \bar{\mu}_s^{\mathrm{IS},i}(k') - \gamma \mathscr{G}_s^i(\eta) \Big) ds \\ =& \mathscr{G}_0^i(\eta) + H(t,\eta) + \int_0^t \Big(K_s (1-\eta) \partial_\eta \mathscr{G}_s^i(\eta) - \gamma \mathscr{G}_s^i(\eta) \Big) ds. \end{aligned}$$

Then, the functions $t \mapsto \widetilde{\mathscr{G}}_t^i(\eta)$ defined by $\widetilde{\mathscr{G}}_t^i(\eta) = e^{\beta t} \mathscr{G}_t^i(\eta)$, $i \in \{1,2\}$, are solutions of the following transport equation (of unknown function g):

$$\partial_t g(t,\eta) - (1-\eta) K_t\, \partial_\eta g(t,\eta) = \partial_t H(t,\eta) e^{\beta t}. \tag{3.3.46}$$

In view of the regularity of H and K, it is known that this equation admits a unique solution (see e.g. [54]). Hence $\mathscr{G}_t^1(\eta) = \mathscr{G}_t^2(\eta)$ for all $t \in \mathbb{R}_+$ and $\eta \in [0,1)$. The same method applies to $\bar{\mu}^{\mathrm{RS}}$. Thus there is at most one solution to the differential system (3.3.17)–(3.3.19).

Step 3 We now show that μ^N nearly satisfies (3.3.17)–(3.3.19) as N gets large. Recall (3.3.14) for a bounded function f on $\mathbb{Z}_+$. To identify the limiting values, we establish that for all $N \in \mathbb{N}$ and all $t \geq 0$,

$$\langle \mu_{t \wedge \tau_\varepsilon^N}^{\mathrm{N,IS}}, f \rangle = \Psi_{t \wedge \tau_\varepsilon^N}^{\mathrm{IS},f}(\mu^N) + \Delta_{t \wedge \tau_\varepsilon^N}^{\mathrm{N},f} + M_{t \wedge \tau_\varepsilon^N}^{\mathrm{N,IS},f}, \tag{3.3.47}$$

where $M^{N,\mathrm{IS},f}$ is defined in (3.3.14) and where $\Delta_{\cdot \wedge \tau_\varepsilon^N}^{N,f}$ converges to 0 when $N \to +\infty$, in probability and uniformly in t on compact time intervals.

Let us fix $t \in \mathbb{R}_+$. Computation similar to (3.3.41) give:

$$\mathbb{E}\big((M_t^{\mathrm{N,IS},f})^2\big) = \mathbb{E}\big(\langle M^{\mathrm{N,IS},f} \rangle_t\big) \leq \frac{(25\lambda + 2\gamma) A t \|f\|_\infty^2}{N}. \tag{3.3.48}$$

Hence the sequence $(M_t^{\mathrm{N,IS},f})_{N \in \mathbb{Z}_+}$ converges in L^2 and in probability to zero.

We now consider the finite variation part of (3.3.14), given in (3.3.15). The sum in (3.3.15) corresponds to the links to I that the new infected individual has. We separate this sum into cases where the new infected individual only has simple edges to other individuals of I, and cases where multiple edges exist. The latter term is expected to vanish for large populations.

$$A_t^{\mathrm{N,IS},f} = B_t^{\mathrm{N,IS},f} + C_t^{\mathrm{N,IS},f}, \tag{3.3.49}$$

where

$$\begin{aligned} B_t^{\mathrm{N,IS},f} = &-\int_0^t \gamma\langle \mu_s^{\mathrm{N,IS}}, f\rangle \,\mathrm{d}s \\ &+\int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell+1\le k} p_s^N(j,\ell|k-1)\Bigg\{ f(k-(j+1+\ell)) \\ &+ \sum_{\substack{u\in\mathscr{L}(j+1,\mu_s^{\mathrm{N,IS}});\\ \forall u\le I_{s_-}^{\mathrm{N}},\, n_u\le 1}} \rho(\underline{n}|j+1,\mu_s^{\mathrm{N,IS}}) \sum_{u\in I_{s_-}^{\mathrm{N}}} \left(f\left(D_u(\mu_{s_-}^{\mathrm{N,IS}})-n_u\right) - f\left(D_u(\mu_{s_-}^{\mathrm{N,IS}})\right)\right)\Bigg\} \,\mathrm{d}s \end{aligned} \tag{3.3.50}$$

and

$$\begin{aligned} C_t^{\mathrm{N,IS},f} = &\int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell+1\le k} p_s^N(j,\ell|k-1) \\ &\times \sum_{\substack{\underline{n}\in\mathscr{L}(j+1,\mu_s^{\mathrm{N,IS}});\\ \exists u\le I_{s_-}^{\mathrm{N}},\, n_u>1}} \rho(\underline{n}|j+1,\mu_s^{\mathrm{N,IS}}) \sum_{u\in I_{s_-}^{\mathrm{N}}} \left(f\left(D_u(\mu_{s_-}^{\mathrm{N,IS}})-n_u\right) - f\left(D_u(\mu_{s_-}^{\mathrm{N,IS}})\right)\right) \,\mathrm{d}s. \end{aligned} \tag{3.3.51}$$

We first show that $C_t^{\mathrm{N,SI},f}$ is a negligible term. Let $q_{j,\ell,s}^N$ denote the probability that the newly infected individual at time s has a double (or of higher order) edge to some alter in $I_{s_-}^{\mathrm{N}}$, given j and ℓ. The probability to have a multiple edge to a given infectious i is less than the number of pairs of edges linking the newly infected to i, times the probability that these two particular edges linking i to a susceptible alter at time s_- actually lead to the newly infected. Hence,

$$\begin{aligned} q_{j,\ell,s}^N &= \sum_{\substack{\underline{n}\in\mathscr{L}(j+1,\mu_s^{\mathrm{N,IS}});\\ \exists u\in I_{s_-}^{\mathrm{N}},\, n_u>1}} \rho(\underline{n}|j+1,\mu_s^{\mathrm{N,IS}}) \\ &\le \binom{j}{2} \sum_{u\in I_{s_-}^{\mathrm{N}}} \frac{D_u(\mathrm{S}_{s_-}^{\mathrm{N}})(D_u(\mathrm{S}_{s_-}^{\mathrm{N}})-1)}{N_{s_-}^{\mathrm{N,IS}}(N_{s_-}^{\mathrm{N,IS}}-1)} = \binom{j}{2}\frac{1}{N}\frac{\langle \mu_{s_-}^{\mathrm{N,IS}}, \chi(\chi-1)\rangle}{\frac{N_{s_-}^{\mathrm{N,IS}}}{N}\left(\frac{N_{s_-}^{\mathrm{N,IS}}}{N}-\frac{1}{N}\right)} \\ &\le \binom{j}{2}\frac{1}{N}\frac{A}{\varepsilon(\varepsilon-1/N)} \quad \text{if } s<\tau_\varepsilon^N \text{ and } N>1/\varepsilon. \end{aligned} \tag{3.3.52}$$

Then, since for all $u\in\mathscr{L}(j+1,\mu_s^{\mathrm{N,IS}})$,

$$\left|\sum_{u\in I_{s_-}^{\mathrm{N}}} \left(f\left(D_u(\mu_{s_-}^{\mathrm{N,IS}})-n_u\right) - f\left(D_u(\mu_{s_-}^{\mathrm{N,IS}})\right)\right)\right| \le 2(j+1)\|f\|_\infty, \tag{3.3.53}$$

we have by (3.3.52) and (3.3.53), for $N>1/\varepsilon$,

$$\begin{aligned} &|C_{t\wedge\tau_\varepsilon^N}^{\mathrm{N,IS},f}| \\ &\le \int_0^{t\wedge\tau_\varepsilon^N} \sum_{k\in\mathbb{Z}_+} \lambda k \mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell+1\le k} p_s^N(j,\ell|k-1) 2(j+1)\|f\|_\infty \frac{j(j-1)A}{2N\varepsilon(\varepsilon-1/N)} \,\mathrm{d}s \end{aligned} \tag{3.3.54}$$

$$\leq \frac{A\lambda t\|f\|_\infty}{N\varepsilon(\varepsilon-1/N)}\langle \mu_0^{\mathrm{N,S}}, \chi^4\rangle,$$

which tends to zero in view of (3.3.21) and thanks to the fact that $\mu_s^{\mathrm{N,S}}$ is dominated by $\mu_0^{\mathrm{N,S}}$ for all $s \geq 0$ and $N \in \mathbb{N}$.

We now aim at proving that $B^{\mathrm{N,IS},f}_{\cdot\wedge\tau_\varepsilon^N}$ is close to $\Psi^{\mathrm{IS},f}_{\cdot\wedge\tau_\varepsilon^N}(\mu^N)$. First, notice that

$$\begin{aligned}
&\sum_{\substack{\underline{n}\in\mathscr{L}(j+1,\mu_s^{\mathrm{N,IS}});\\ \forall u\in I_{s_-}^{\mathrm{N}}, n_u\leq 1}} \rho(u|j+1,\mu_s^{\mathrm{N,IS}}) \sum_{i\in I_{s_-}^{\mathrm{N}}} \left(f\left(D_u\left(\mu_{s_-}^{\mathrm{N,IS}}\right)-n_u\right)-f\left(D_u\left(\mu_{s_-}^{\mathrm{N,IS}}\right)\right)\right)\\
&= \sum_{u_0\neq\cdots\neq u_j\in I_{s_-}^{\mathrm{N}}} \left(\frac{\prod_{k=0}^j D_{u_k}(\mathrm{s}_s^{\mathrm{N}})}{N_{s_-}^{\mathrm{N,IS}}\ldots(N_{s_-}^{\mathrm{N,IS}}-(j+1))}\right)\\
&\qquad\qquad \times \sum_{m=0}^j \left(f\left(D_{u_m}(\mathrm{s}_{s_-}^{\mathrm{N}})-1\right)-f\left(D_{u_m}(\mathrm{s}_{s_-}^{\mathrm{N}})\right)\right)\\
&= \sum_{m=0}^j \sum_{u_0\neq\cdots\neq u_j\in I_{s_-}^{\mathrm{N}}} \left(\frac{\prod_{k=0}^j D_{u_k}(\mathrm{s}_s^{\mathrm{N}})}{N_{s_-}^{\mathrm{N,IS}}\ldots(N_{s_-}^{\mathrm{N,IS}}-(j+1))}\right)\\
&\qquad\qquad \times \left(f\left(D_{u_m}(\mathrm{s}_{s_-}^{\mathrm{N}})-1\right)-f\left(D_{u_m}(\mathrm{s}_{s_-}^{\mathrm{N}})\right)\right) \qquad (3.3.55)\\
&= \sum_{m=0}^j \left(\sum_{x\in I_{s_-}^{\mathrm{N}}} \frac{D_x(\mathrm{s}_{s_-}^{\mathrm{N}})}{N_{s_-}^{\mathrm{N,IS}}}\left(f\left(D_x(\mathrm{s}_{s_-}^{\mathrm{N}})-1\right)-f\left(D_x(\mathrm{s}_{s_-}^{\mathrm{N}})\right)\right)\right)\\
&\qquad\qquad \times \left(\sum_{u_0\neq\cdots\neq u_{j-1}\in I_{s_-}^{\mathrm{N}}\setminus\{x\}} \frac{\prod_{k=0}^{j-1} D_{u_k}(\mathrm{s}_s^{\mathrm{N}})}{(N_{s_-}^{\mathrm{N,IS}}-1)\ldots(N_{s_-}^{\mathrm{N,IS}}-(j+1))}\right)\\
&= (j+1)\frac{\langle \mu_{s_-}^{\mathrm{N,IS}}, \chi\left(\tau_1 f - f\right)\rangle}{N_{s_-}^{\mathrm{N,IS}}}\left(1-q^N_{j-1,\ell,s}\right),
\end{aligned}$$

where we recall (see Notation 0.0.1) that $\tau_1 f(k) = f(k-1)$ for every function f on $\mathbb{Z}_+$ and $k \in \mathbb{Z}_+$. In the third equality, we split the term u_m from the other terms $(u_{m'})_{m'\neq m}$. The last sum in the r.h.s. of this equality is the probability of drawing j different infectious individuals that are not u_m and that are all different, hence $1-q^N_{j-1,\ell,s}$.

Define for $t > 0$ and $N \in \mathbb{Z}_+$,

$$\begin{aligned}
p_t^{\mathrm{N,I}} &= \frac{\langle \mu_t^{\mathrm{N,IS}}, \chi\rangle - 1}{\langle \mu_t^{\mathrm{N,S}}, \chi\rangle - 1},\\
p_t^{\mathrm{N,R}} &= \frac{\langle \mu_t^{\mathrm{N,RS}}, \chi\rangle}{\langle \mu_t^{\mathrm{N,S}}, \chi\rangle - 1},\\
p_t^{\mathrm{N,S}} &= \frac{\langle \mu_t^{\mathrm{N,S}}, \chi\rangle - \langle \mu_t^{\mathrm{N,IS}}, \chi\rangle - \langle \mu_t^{\mathrm{N,RS}}, \chi\rangle}{\langle \mu_t^{\mathrm{N,S}}, \chi\rangle - 1},
\end{aligned}$$

the proportion of edges with infectious (resp. removed and susceptible) alters and susceptible egos among all the edges with susceptible egos but the contaminating edge. For all integers j and ℓ such that $j+\ell \leq k-1$ and $N \in \mathbb{N}$, denote by

$$\tilde{p}_t^N(j,\ell \mid k-1) = \frac{(k-1)!}{j!(k-1-j-\ell)!\ell!}(p_t^{\mathrm{N,I}})^j (p_t^{\mathrm{N,R}})^\ell (p_t^{\mathrm{N,S}})^{k-1-j-\ell},$$

the probability that the multinomial variable counting the number of edges with infectious, removed and susceptible alters, among $k-1$ given edges, equals $(j,\ell,k-1-j-\ell)$. We have that

$$|\Psi^{\mathrm{IS},f}_{t\wedge\tau_\varepsilon^N}(\mu^N) - B^{\mathrm{N,IS},f}_{t\wedge\tau_\varepsilon^N}| \leq |D^{\mathrm{N,IS},f}_{t\wedge\tau_\varepsilon^N}| + |E^{\mathrm{N,IS},f}_{t\wedge\tau_\varepsilon^N}|, \qquad (3.3.56)$$

where

$$D_t^{\mathrm{N,IS},f} = \int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \sum_{j+\ell+1\le k} \left(p_s^N(j,\ell|k-1) - \tilde{p}_s^N(j,\ell|k-1)\right) \times \left(f(k-(j+\ell+1)) + (j+1)\frac{\langle \mu_{s-}^{\mathrm{N,IS}}, \chi\left(\tau_1 f - f\right)\rangle}{N_{s-}^{\mathrm{N,IS}}} \right) \mathrm{d}s,$$

$$E_t^{\mathrm{N,IS},f} = \int_0^t \sum_{k\in\mathbb{Z}_+} \Lambda_s^N(k)\mu_s^{\mathrm{N,S}}(k) \times \sum_{j+\ell+1\le k} p_s^N(j,\ell|k-1)(j+1)\frac{\langle \mu_{s-}^{\mathrm{N,IS}}, \chi\left(\tau_1 f - f\right)\rangle}{N_{s-}^{\mathrm{N,IS}}} q_{j-1,\ell,s}^N \ \mathrm{d}s.$$

First,

$$|D_{t\wedge\tau_\varepsilon^N}^{\mathrm{N,IS},f}| \le \int_0^{t\wedge\tau_\varepsilon^N} \sum_{k\in\mathbb{Z}_+} \lambda k \alpha_s^N(k)\|f\|_\infty \left(1+\frac{2kA}{\varepsilon}\right) \mu_s^{\mathrm{N,S}}(k) \ \mathrm{d}s, \tag{3.3.57}$$

where for all $k\in\mathbb{Z}_+$

$$\alpha_t^N(k) = \sum_{j+\ell+1\le k} \left| p_t^N(j,\ell|k-1) - \tilde{p}_t^N(j,\ell|k-1)\right|.$$

The multinomial probability $\tilde{p}_s^N(j,\ell|k-1)$ approximates the hypergeometric one, $p_s^N(j,\ell|k-1,s)$, as N increases to infinity, in view of the fact that the total population size, $\langle \mu_0^{\mathrm{N,S}}, \mathbf{1}\rangle + \langle \mu_0^{\mathrm{N,IS}}, \mathbf{1}\rangle$, is of order n. Hence, the r.h.s. of (3.3.57) vanishes by dominated convergence.

On the other hand, using (3.3.52),

$$\begin{aligned} |E_{t\wedge\tau_\varepsilon^N}^{\mathrm{N,IS},f}| &\le \int_0^{t\wedge\tau_\varepsilon^N} \sum_{k\in\mathbb{Z}_+} \lambda k^2 \mu_s^{\mathrm{N,S}}(k) \frac{2\|f\|_\infty A}{\varepsilon} \frac{k^2 A}{2N\varepsilon(\varepsilon - 1/N)} \ \mathrm{d}s \\ &\le \frac{A^3 \lambda t \|f\|_\infty}{N\varepsilon^2(\varepsilon-1/N)}, \end{aligned} \tag{3.3.58}$$

in view of (3.3.21). Gathering (3.3.48), (3.3.49), (3.3.54), (3.3.56), (3.3.57) and (3.3.58) concludes the proof that the rest of (3.3.47) vanishes in probability uniformly over compact intervals.

As a consequence, the sequence $(\Psi_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{IS},f}(\mu^N))_{N\in\mathbb{N}}$ is also tight in $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A}\times\mathscr{M}_{\varepsilon,A}\times\mathscr{M}_{0,A})$.

Step 4 Recall that in this proof, $\bar{\mu} = (\bar{\mu}^{\mathrm{S}}, \bar{\mu}^{\mathrm{IS}}, \bar{\mu}^{\mathrm{RS}})$ is the limit of the sequence $(\mu_{\cdot\wedge\tau_\varepsilon^N}^N)_{N\in\mathbb{N}} = (\mu_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,S}}, \mu_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,IS}}, \mu_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,RS}})_{N\in\mathbb{N}}$, and recall that these processes take values in the closed set $\mathscr{M}_{0,A}^3$. Our purpose is now to prove that $\bar{\mu}$ satisfies (3.3.17)–(3.3.19). Using Skorokhod's representation theorem, there exists, on the same probability space as $\bar{\mu}$, a sequence, again denoted by $(\mu_{\cdot\wedge\tau_\varepsilon^N}^N)_{N\in\mathbb{N}}$ with an abuse of notation, with the same marginal distributions as the original sequence, and that converges a.s. to $\bar{\mu}$.

The maps $\nu_\cdot := (\nu_\cdot^1, \nu_\cdot^2, \nu_\cdot^3) \mapsto \langle \nu_\cdot^1, \mathbf{1}\rangle / (\langle \nu_0^1, \mathbf{1}\rangle + \langle \nu_0^2, \mathbf{1}\rangle + \langle \nu_0^3, \mathbf{1}\rangle)$ (respectively $\langle \nu_\cdot^2, \mathbf{1}\rangle / (\langle \nu_0^1, \mathbf{1}\rangle + \langle \nu_0^2, \mathbf{1}\rangle + \langle \nu_0^3, \mathbf{1}\rangle)$ and $\langle \nu_\cdot^3, \mathbf{1}\rangle / (\langle \nu_0^1, \mathbf{1}\rangle + \langle \nu_0^2, \mathbf{1}\rangle + \langle \nu_0^3, \mathbf{1}\rangle)$) are continuous from $\mathscr{C}(\mathbb{R}_+, \mathscr{M}_{0,A}\times\mathscr{M}_{\varepsilon,A}\times\mathscr{M}_{0,A})$ into $\mathscr{C}(\mathbb{R}_+,\mathbb{R})$. Using the moment assumption (3.3.21), the following mappings are also continuous for the same spaces: $\langle \nu_\cdot^1, \chi\rangle / \langle \nu_\cdot^2, \chi\rangle$, $\nu_\cdot \mapsto \mathbf{1}_{\langle \nu_\cdot^1, \chi\rangle > \varepsilon} / \langle \nu_\cdot^2, \chi\rangle$ and $\nu_\cdot \mapsto \langle \nu_\cdot^2, \chi\left(\tau_1 f - f\right)\rangle$, for bounded function f on $\mathbb{Z}_+$ and where we recall that $\tau_1 f(k) = f(k-1)$ for every $k\in\mathbb{Z}_+$ (see Notation 0.0.1). Thus, using the continuity of the mapping $y\in\mathbb{D}([0,t],\mathbb{R}) \mapsto \int_0^t y_s \ \mathrm{d}s$, we obtain the continuity of the mapping Ψ_t^f defined in (3.3.42) on $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A}\times\mathscr{M}_{\varepsilon,A}\times\mathscr{M}_{0,A})$.

By (3.3.21), the process $(N_{\cdot\wedge\tau_\varepsilon^N}^{\mathrm{N,IS}})_{N\in\mathbb{N}}$ converges in distribution to $\bar{N}_\cdot^{\mathrm{IS}} = \langle \bar{\mu}_\cdot^{\mathrm{IS}}, \chi\rangle$. Since the latter process is continuous, the convergence holds in $(\mathbb{D}([0,T],\mathbb{R}_+), \|\cdot\|_\infty)$ for any $T>0$ (see [24, p. 112]). As $y\in\mathbb{D}(\mathbb{R}_+,\mathbb{R}) \mapsto$

$\inf_{t\in[0,T]} y(t) \in \mathbb{R}$ is continuous, we have a.s. that:

$$\inf_{t\in[0,T]} \bar{N}^{\mathrm{IS}}_t = \lim_{N\to+\infty} \inf_{t\in[0,T]} N^{\mathrm{N,IS}}_{t\wedge\tau^N_\varepsilon} \quad \big(\geq \varepsilon\big).$$

Analogously to (3.3.37), we consider $\bar{t}_{\varepsilon'} = \inf\{t \in \mathbb{R}_+, \bar{N}^{\mathrm{IS}}_t \leq \varepsilon'\}$ for $\varepsilon' > \varepsilon > 0$. A difficulty lies in the fact that we do not know yet whether this time is deterministic. We have a.s.:

$$\varepsilon' \leq \inf_{t\in[0,T]} \bar{N}^{\mathrm{IS}}_{t\wedge\bar{t}_{\varepsilon'}} = \lim_{N\to+\infty} \inf_{t\in[0,T]} N^{\mathrm{N,IS}}_{t\wedge\tau^N_\varepsilon\wedge\bar{t}_{\varepsilon'}}. \tag{3.3.59}$$

Hence, using Fatou's lemma:

$$\begin{aligned} 1 =& \mathbb{P}\Big(\inf_{t\in[0,\bar{t}_{\varepsilon'}]} \bar{N}^{\mathrm{IS}}_t > \varepsilon\Big) \\ \leq& \lim_{N\to+\infty} \mathbb{P}\Big(\inf_{t\in[0,T\wedge\bar{t}_{\varepsilon'}]} N^{\mathrm{N,IS}}_{t\wedge\tau^N_\varepsilon} > \varepsilon\Big) = \lim_{N\to+\infty} \mathbb{P}\Big(\tau^N_\varepsilon > T\wedge\bar{t}_{\varepsilon'}\Big). \end{aligned} \tag{3.3.60}$$

We have hence

$$\Psi^{\mathrm{IS},f}_{.\wedge\tau^N_\varepsilon\wedge\bar{t}_{\varepsilon'}\wedge T}(\mu^N) = \Psi^{\mathrm{IS},f}_{.\wedge\tau^N_\varepsilon\wedge T}(\mu^N)\mathbf{1}_{\tau^N_\varepsilon\leq\bar{t}_{\varepsilon'}\wedge T} + \Psi^{\mathrm{IS},f}_{.\wedge\bar{t}_{\varepsilon'}\wedge T}(\mu^N_{.\wedge\tau^N_\varepsilon})\mathbf{1}_{\tau^N_\varepsilon>\bar{t}_{\varepsilon'}\wedge T}. \tag{3.3.61}$$

From the estimates of the different terms in (3.3.47), $\Psi^{\mathrm{IS},f}_{.\wedge\tau^N_\varepsilon\wedge T}(\mu^N)$ is upper bounded by a moment of μ^N of order 4. In view of (3.3.21) and (3.3.60), the first term in the r.h.s. of (3.3.61) converges in L^1 and hence in probability to zero. Using the continuity of $\Psi^{\mathrm{IS},f}$ on $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A}\times\mathscr{M}_{\varepsilon,A}\times\mathscr{M}_{0,A})$, $\Psi^{\mathrm{IS},f}(\mu^N_{.\wedge\tau^N_\varepsilon})$ converges to $\Psi^{\mathrm{IS},f}(\bar{\mu})$ and therefore, $\Psi^{\mathrm{IS},f}_{.\wedge\bar{t}_{\varepsilon'}\wedge T}(\mu^N_{.\wedge\tau^N_\varepsilon})$ converges to $\Psi^{\mathrm{IS},f}_{.\wedge\bar{t}_{\varepsilon'}\wedge T}(\bar{\mu})$. Thanks to this and (3.3.60), the second term in the r.h.s. of (3.3.61) converges to $\Psi^{\mathrm{IS},f}_{.\wedge\bar{t}_{\varepsilon'}\wedge T}(\bar{\mu})$ in $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$.

Then, the sequence $(\langle\mu^{\mathrm{N,IS}}_{.\wedge\tau^N_\varepsilon\wedge\bar{t}_{\varepsilon'}\wedge T}, f\rangle - \Psi^{\mathrm{IS},f}_{.\wedge\tau^N_\varepsilon\wedge\bar{t}_{\varepsilon'}\wedge T}(\mu^N))_{N\in\mathbb{N}}$ converges in probability to $\langle\bar{\mu}^{\mathrm{IS}}_{.\wedge\bar{t}_{\varepsilon'}\wedge T}, f\rangle - \Psi^{\mathrm{IS},f}_{.\wedge\bar{t}_{\varepsilon'}\wedge T}(\bar{\mu})$. From (3.3.47), this sequence also converges in probability to zero.

By identification of these limits, $\bar{\mu}^{\mathrm{IS}}$ solves (3.3.18) on $[0,\bar{t}_{\varepsilon'}\wedge T]$. If $\langle\bar{\mu}^{\mathrm{RS}}_0,\chi\rangle > 0$ then similar techniques can be used. Else, the result is obvious since for all $t\in[0,t_{\varepsilon'}\wedge T]$, $\langle\mu^{\mathrm{N,IS}}_t,\chi\rangle > \varepsilon$ and the term $p^N_t(j,\ell|k-1)$ is negligible when $\ell > 0$. Thus $\bar{\mu}$ coincides a.s. with the only continuous deterministic solution of (3.3.17)–(3.3.19) on $[0,\bar{t}_{\varepsilon'}\wedge T]$. This implies that $\bar{t}_{\varepsilon'}\wedge T = t_{\varepsilon'}\wedge T$ and yields the convergence in probability of $(\mu^N_{.\wedge\tau^N_\varepsilon})_{N\in\mathbb{N}}$ to $\bar{\mu}$, uniformly on $[0,t_{\varepsilon'}\wedge T]$ since $\bar{\mu}$ is continuous.

We finally prove that the non-localized sequence $(\mu^N)_{N\in\mathbb{N}}$ also converges uniformly and in probability to $\bar{\mu}$ in $\mathbb{D}([0,t_{\varepsilon'}],\mathscr{M}_{0,A}\times\mathscr{M}_{\varepsilon,A}\times\mathscr{M}_{0,A})$. For a small positive η,

$$\begin{aligned} &\mathbb{P}\Big(\sup_{t\in[0,t_{\varepsilon'}]}\Big|\langle\mu^{\mathrm{N,IS}}_t,f\rangle - \Psi^{\mathrm{IS},f}_t(\bar{\mu})\Big| > \eta\Big) \\ &\qquad\leq \mathbb{P}\Big(\sup_{t\in[0,t_{\varepsilon'}]}\Big|\Psi^{\mathrm{IS},f}_{t\wedge\tau^N_\varepsilon}(\mu^N) - \Psi^{\mathrm{IS},f}_t(\bar{\mu})\Big| > \frac{\eta}{2}\,;\,\tau^n_\varepsilon \geq t_{\varepsilon'}\Big) \\ &\qquad\qquad + \mathbb{P}\Big(\sup_{t\in[0,t_{\varepsilon'}]}\Big|\Delta^{\mathrm{N},f}_{t\wedge\tau^N_\varepsilon} + M^{\mathrm{N,IS},f}_{t\wedge\tau^N_\varepsilon}\Big| > \frac{\eta}{2}\Big) + \mathbb{P}\Big(\tau^N_\varepsilon < t_{\varepsilon'}\Big). \end{aligned} \tag{3.3.62}$$

Using the continuity of Ψ^f and the uniform convergence in probability proved above, the first term in the r.h.s. of (3.3.62) converges to zero. We can show that the second term converges to zero by using Doob's inequality together with the estimates of the bracket of $M^{\mathrm{N,IS},f}$ (similar to (3.3.41)) and of $\Delta^{\mathrm{N},f}$ (Step 2). Finally, the third term vanishes in view of (3.3.60).

The convergence of the original sequence $(\mu^N)_{N\in\mathbb{N}}$ is then implied by the uniqueness of the solution to (3.3.17)–(3.3.19) proved in Step 2.

Step 5 When $N \to +\infty$, by taking the limit in (3.3.12), $(\mu^{\mathrm{N,S}})_{N\in\mathbb{N}}$ converges in $\mathbb{D}(\mathbb{R}_+, \mathscr{M}_{0,A})$ to the solution of the following transport equation: for every bounded function $f : (k,t) \mapsto f_t(k) \in \mathscr{C}_b^{0,1}(\mathbb{Z}_+ \times \mathbb{R}_+, \mathbb{R})$ of class $\mathscr{C}^1$ with bounded derivative with respect to t,

$$\langle \bar{\mu}_t^{\mathrm{S}}, f_t\rangle = \langle \bar{\mu}_0^{\mathrm{S}}, f_0\rangle - \int_0^t \langle \bar{\mu}_s^{\mathrm{S}}, \lambda\chi \bar{p}_s^{\mathrm{I}} f_s - \partial_s f_s\rangle \, \mathrm{d}s. \tag{3.3.63}$$

Choosing $f(k,s) = \varphi(k)\exp\big(-\lambda k \int_0^{t-s} \bar{p}^{\mathrm{I}}(u)du\big)$, we obtain that

$$\langle \bar{\mu}_t^{\mathrm{S}}, \varphi\rangle = \sum_{k\in\mathbb{Z}_+} \varphi(k)\theta_t^k \bar{\mu}_0^{\mathrm{S}}(k). \tag{3.3.64}$$

where $\theta_t = \exp\big(-\lambda \int_0^t \bar{p}^{\mathrm{I}}(u)du\big)$ is the probability that a given degree 1 node remains susceptible at time t. This is the announced Equation (3.3.17).

The proof of Theorem 3.3.8 is now completed. □

Recall that the time $t_{\varepsilon'}$ has been defined in (3.3.37). We end this section with a lower bound of the time $t_{\varepsilon'}$ until which we proved that the convergence to Volz' equations holds.

Proposition 3.3.12. *Under the assumptions of Theorem 3.3.8,*

$$t_{\varepsilon'} > \bar{\tau}_{\varepsilon'} := \frac{\log\big(\langle \bar{\mu}_0^{\mathrm{S}}, \chi^2\rangle + \bar{N}_0^{\mathrm{IS}}\big) - \log\big(\langle \bar{\mu}_0^{\mathrm{S}}, \chi^2\rangle + \varepsilon'\big)}{\max(\gamma,\lambda)}. \tag{3.3.65}$$

Proof. Because of the moment Assumption 3.3.5 and (3.3.21), we can prove that (3.3.47) also holds for $f = \chi$. This is obtained by replacing in (3.3.48), (3.3.54), (3.3.57) and (3.3.58) $\|f\|_\infty$ by k and using the Assumption of boundedness of the moments of order 5 in (3.3.54) and (3.3.58). This shows that $(N^{\mathrm{N,IS}})_{N\in\mathbb{N}}$ converges, uniformly on $[0,t_{\varepsilon'}]$ and in probability, to the deterministic and continuous solution $\bar{N}^{\mathrm{IS}} = \langle \bar{\mu}^{\mathrm{IS}}, \chi\rangle$. We introduce the event $\mathscr{A}_\xi^N = \{|\, N_0^{\mathrm{N,IS}} - N\bar{N}_0^{\mathrm{IS}} \,| \leq \xi\}$ where their differences are bounded by $\xi > 0$. Recall the definition (3.3.38) and let us introduce the number of edges Z_t^N that were IS at time 0 and that have been removed before t. For $t \geq \tau_{\varepsilon'}^N$, we have necessarily that $Z_t^N \geq N_0^{\mathrm{N,IS}} - N\varepsilon'$. Thus,

$$\begin{aligned}\mathbb{P}\big(\{\tau_{\varepsilon'}^N \leq t\} \cap \mathscr{A}_\xi^N\big) &\leq \mathbb{P}\big(\{Z_t^N > N_0^{\mathrm{N,IS}} - N\varepsilon'\} \cap \mathscr{A}_\xi^N\big) \\ &\leq \mathbb{P}\Big(\big\{Z_t^N > N(\bar{N}_0^{\mathrm{IS}} - \varepsilon') - \xi\big\} \cap \mathscr{A}_\xi^N\Big).\end{aligned} \tag{3.3.66}$$

When susceptible (resp. infectious) individuals of degree k are contaminated (resp. removed), at most k I$-$S-edges are lost. Let $X_t^{N,k}$ be the number of edges that, at time 0, are I$-$S with susceptible alter of degree k, and that have transmitted the disease before time t. Let $Y_t^{N,k}$ be the number of initially infectious individuals x with $d_x(\mathrm{S}_0) = k$ and who have been removed before time t. $X_t^{N,k}$ and $Y_t^{N,k}$ are bounded by $k\mu_0^{\mathrm{N,S}}(k)$ and $\mu_0^{\mathrm{N,IS}}(k)$. Thus:

$$Z_t^N \leq \sum_{k\in\mathbb{Z}_+} k\big(X_t^{N,k} + Y_t^{N,k}\big). \tag{3.3.67}$$

Let us stochastically bound Z_t^N from above. Since each I$-$S-edge transmits the disease independently at rate λ, $X_t^{N,k}$ is stochastically dominated by a binomial r.v. of parameters $k\mu_0^{\mathrm{N,S}}(k)$ and $1-e^{-\lambda t}$. We proceed similarly for $Y_t^{N,k}$. Conditional on the initial condition, $X_t^{N,k} + Y_t^{N,k}$ is thus stochastically dominated by a binomial r.v. $\tilde{Z}_t^{N,k}$ of parameters $(k\mu_0^{\mathrm{N,S}}(k) + \mu_0^{\mathrm{N,IS}}(k))$ and $1 - e^{-\max(\lambda,\gamma)t}$. Then (3.3.66) and (3.3.67) give:

$$\mathbb{P}\big(\{\tau_{\varepsilon'}^N \leq t\} \cap \mathscr{A}_\xi^N\big) \leq \mathbb{P}\Big(\sum_{k\in\mathbb{Z}_+} \frac{k\tilde{Z}_t^{N,k}}{N} > \bar{N}_0^{\mathrm{IS}} - \varepsilon' - \frac{\xi}{N}\Big). \tag{3.3.68}$$

Thanks to Assumption 3.3.5 and (3.3.21), the series $\sum_{k\in\mathbb{Z}_+} k\tilde{Z}_t^{N,k}/N$ converges in L^1 and hence in probability to $(\langle\bar{\mu}_0^{\mathrm{S}},\chi^2\rangle+\bar{N}_0^{\mathrm{IS}})(1-e^{-\max(\lambda,\gamma)t})$ when $N\to+\infty$. Thus, for sufficiently large N,

$$\mathbb{P}\big(\{\tau_{\varepsilon'}^N\le t\}\cap\mathscr{A}_\xi^N\big)=1 \text{ if } t>\bar{\tau}_{\varepsilon'} \text{ and } 0 \text{ if } t<\bar{\tau}_{\varepsilon'}.$$

For all $t<\bar{\tau}_{\varepsilon'}$, it follows from Assumption 3.3.5, (3.3.21) and Lemma A.0.4 that:

$$\lim_{N\to+\infty}\mathbb{P}\Big(\tau_{\varepsilon'}^N\le t\le\lim_{N\to+\infty}\mathbb{P}\big(\{\tau_{\varepsilon'}^N\le t\}\cap\mathscr{A}_\xi^N\big)\Big)+\mathbb{P}\big((\mathscr{A}_\xi^N)^c\big)=0,$$

so that by Theorem 3.3.8

$$1=\lim_{N\to+\infty}\mathbb{P}(\tau_{\varepsilon'}^N\ge\bar{\tau}_{\varepsilon'})=\lim_{n\to+\infty}\mathbb{P}\Big(\inf_{t\le\bar{\tau}_{\varepsilon'}}N_t^{\mathrm{N,IS}}\ge\varepsilon'\Big)=\mathbb{P}\Big(\inf_{t\le\bar{\tau}_{\varepsilon'}}\bar{N}_t^{\mathrm{IS}}\ge\varepsilon'\Big).$$

This shows that $t_{\varepsilon'}\ge\bar{\tau}_{\varepsilon'}$ a.s., which concludes the proof. □

Chapter 4

Statistical Description of Epidemics Spreading on Networks: The Case of Cuban HIV

In this section, we turn our attention to epidemics spreading on networks. Probability models have been described in Section 1.4. We now deal with the statistical treatment of data obtained from diseases propagating on networks. The statistical methods described here are illustrated on the sexual network obtained from the Cuban HIV contact-tracing system that we now describe. For a complete description of the Cuban network, we refer to [36]. The Cuban graph is available as supplementary material of this book.

Since 1986, a contact-tracing detection system has been set up in Cuba in order to bring the spread of the HIV epidemic under control. It has also enabled the gathering of a considerable amount of detailed epidemiological data at the individual level. In the resulting database, any individual tested as HIV positive is indexed and anonymized for confidentiality reasons. Information related to uninfected individuals is not recorded in the data, and of course infected individuals not diagnosed yet are also absent. The network only consists of detected HIV+ individuals. However, note that the network is age-structured and data related to the infectious population of the first six years of the epidemic seems to show (e.g. [38]) that this population has been discovered by now.
Individuals in the database are described through several attribute variables: gender and sexual orientation, way of detection, age at detection, date of detection, area of residence, etc. In the sequel, we will mainly focus on the gender/sexual orientation, for which three modalities are identified: 'woman', 'heterosexual man', 'MSM' (Men who have Sex with Men; men who reported at least one sexual contact with another man in the two years preceding HIV detection). Because Female-to-female transmission is neglected, no sexual orientation is distinguished for women (e.g. [31]). It is worth recalling that in Cuba HIV spreads essentially through sexual transmission. Infection by blood transfusion or related to drug use are neglected. We refer to [8] for a preliminary overview of the HIV/AIDS epidemics in Cuba, as well as a description and the context of the construction of the database used in the present study and the context in which it was constructed.

Importantly, for each HIV+ individual that is detected, the list of indices corresponding to the sexual partners appearing in the database she/he possibly named for contact-tracing is also available. In [34, 35, 36] the graph of sexual partners that have been diagnosed HIV positive on the Cuban data repository is reconstructed and an exploratory statistical analysis of the resulting sexual contact network is carried. The network is composed of 5,389 vertices, or nodes, that correspond to the individuals diagnosed as HIV positive between 1986 and 2006 in Cuba, i.e. 1,109 women (20.58%) and 4,280 men (79.42%); 566 (10.50%) of which are heterosexual and 3,714 (68.92%) are MSMs. Individuals declared as sexual contacts but who are not HIV positive are not listed in the database: the only observed vertices correspond to individuals who have been detected as HIV positive or AIDS. The vertices that depict the fact that two individuals have been sexual partners during the two years that preceded the detection of either one are linked by 4,073 edges. Only edges between observed HIV cases are hence observed, but the degree (total number of sexual partners) is known. Also, some information is documented on who infects whom, giving access to a partial infection tree. Our data exhibit a "giant component", counting 2,386 nodes. The second largest component has only 17 vertices and there are about 2000 isolated individuals or couples. It is remarkable

that in the existing literature on sexually transmitted diseases graph networks are generally smaller and/or do not exhibit such a large connected component and/or contain a very small number of infected persons (e.g. [103, 117]).

In Section 4.2, using graph-mining techniques, the connectivity/communication properties of the sexual contact network are described to understand the impact of heterogeneity (with respect to the attributes observed) in the graph structure. Particular attention is paid to the graphical representation of the data, as conventional methods cannot be used with databases of the size of the one used in this study. A clustering of the population is performed so as to represent structural information in an interpretable way. Beyond global graph visualization, the task of partitioning the network into groups, with dense internal links and low external connectivity, is known as clustering. In contrast to standard multivariate analysis, in which the network structure of the data is ignored, our method has shed light on how different mechanisms (e.g. social behaviour, detection system) have affected the epidemics of HIV in the past, and provide a way of predicting the future evolution of this disease. This study paves the way for building more realistic network models in the field of mathematical modelling of infectious diseases.

4.1 Modularity and assortative mixing

Assortative mixing coefficients can be computed to highlight the possible existence of selective linking in the network structure. Various measures have been proposed in the literature for quantifying the tendency for individuals to have connections with other individuals that are similar in regards to certain attributes, depending on the nature of the latter (quantitative vs. qualitative). For a partition of J classes, $\mathscr{P} = C_1, \ldots, C_J$, one may calculate the proportion $m_{i,j}$ of edges in the graph connecting a node lying in group i to another one in group j, $1 \leq i \leq j \leq J$ and build the $J \times J$ mixing matrix $\mathscr{M} = (m_{i,j})$ (notice it is symmetric since edges are not directed here). We can then define the modularity coefficient $Q_{\mathscr{P}}$ (e.g. [85]) by:

$$Q_{\mathscr{P}} = \mathrm{Tr}(\mathscr{M}) - ||\mathscr{M}^2|| = \sum_i \left\{ m_{i,i} - \left(\sum_{j=1}^{N} m_{i,j} \right)^2 \right\}, \tag{4.1.1}$$

where $||A|| = \sum_i \sum_j a_{i,j}$ denotes the sum of all the entries of a matrix $A = (a_{i,j})$ and $\mathrm{Tr}(A)$ its trace when the latter is square.

We can define the assortative coefficient as

$$r = Q_{\mathscr{P}} / (1 - ||\mathscr{M}^2||).$$

As pointed out in [88], large values of r indicate "selective linking": values around 0 correspond to randomly mixed network, whereas values close to 1 are associated with perfectly assortative network. The assortative coefficient can also be negative.

A first class of partitions are constituted by nodes taking the same modalities of qualitative variables: area of residence, sexual orientation, age, detection mode... Let us comment on the partition defined by the gender/sexual orientation variable (see Table 4.1.1). As edges correspond to sexual contacts in the present graph, the gender/sexual orientation of adjacent vertices cannot be arbitrary of course. More than a half of the edges (56.47%) link two MSM. Links between MSM and women make 1,208 edges (29.66%) and there are 439 edges (10.78%) between women and heterosexual men. Looking at the infection tree provided similar proportions: 1,202 edges (52.56%), 667 edges (29.16%) and 375 edges (16.40%) respectively. Figures reveal an asymmetry in HIV infection: among (oriented) infection edges involving women, the latter are more often alters than egos (66.13% of the edges shared with heterosexual men and 74.21% of the edges shared with MSM). The declarative degree shows a smaller mean degree for heterosexual men and comparable degree distributions between women and MSM. MSM are expected to contribute most to the connectivity of the graph, especially bisexual men who act as contact points between women and MSM who declare only contacts with men.

Ego is a	Alter is a woman	Alter is a heterosexual man	Alter is an MSM	Total
Woman	77 (*1.9%*)	157 (*3.9%*)	408 (*10.0%*)	642 (*15.8%*)
HT man	282 (*6.9%*)	4 (*0.1%*)	20 (*0.5%*)	306 (*7.5%*)
MSM	800 (*19.6%*)	25 (*0.6%*)	2300 (*56.5%*)	3125 (*76.7%*)
Total	1159 (*28.5%*)	186 (*4.6%*)	2728 (*67.0%*)	

Table 4.1.1: *Sexual orientation of egos and alters for the edges in the whole graph. The figures presented here account for the direction of the edges: egos are detected first and alters are the partners they refer to during the contact-tracing interviews. Frequencies are given together with row and column proportions between brackets. The diagonal of the contingency table represents 58,46% of the whole edges. The assortative mixing coefficient is* $r = 0.0512$*. The independence between the sexual orientation of egos and alters is rejected by a* χ^2*-test with a p-value smaller than* $2.2\,10^{-16}$*. In theory, there should be no sexual contact between two heterosexual men or between a heterosexual man and an MSM. The semantic of the database also exclude sexual contact between women. However, those events actually occur in the dataset.*

Of course, a natural question is to see whether we can define other partitions that are more closely related to the modularity defined in (4.1.1). This is the topic of the next section, which is related with visual-mining and modularity clustering.

4.2 Visual-mining

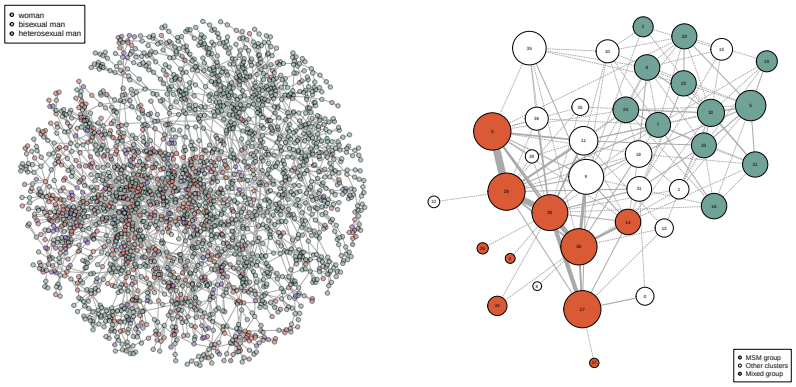
Graph visualization techniques are used routinely to gain insights about medium size graph structures, but their practical relevance is questionable when the number of vertices and the density of the graph are high both for computational issues (as many graph drawing algorithms have high complexities) and for readability issues [22, 60]. We illustrate the clustering and visualization on the Cuba HIV data where the situation is borderline as the giant component of the graph contains 2,386 vertices and 3,168 edges (respectively 44.28% and 77.78% of the global quantities). As the graph is of medium size from a computational point of view and has a low density, it is a reasonable candidate for state-of-the-art global and detailed visualization techniques. We use the optimised force directed placement algorithm proposed in [109]. It recasts the classical force directed paradigm [57] into a nonlinear optimization problem in which the following energy is minimised over the vertex positions in the euclidean plane, $(z_1, \ldots, z_n)$,

$$\mathscr{E}(z_1, \ldots, z_n) = \sum_{1 \le i \ne j \le n} \left(a_{i,j} \frac{1}{3\delta} \|z_i - z_j\|^3 - \delta^2 \ln \|z_i - z_j\| \right),$$

where, δ is a free parameter that is roughly proportional to the expected average distance between vertices in the plane at the end of the optimization process, $a_{i,j}$ are the terms of the adjacency matrix of the network and $\|\cdot\|$ denotes the Euclidean distance in the plane.

However, the structure of the graph under study, in particular its uneven density, has adverse effects on the readability of its global representation. We rely therefore on the classical simplification approach [60] that consists in building a clustering of the vertices of the graph and in representing the simpler graph of the clusters. More precisely, the general idea is to define a partition composed of groups with dense internal links but low inter-group connectivity. Each group can then be considered as a vertex of a new graph: two such vertices are connected if there is at least one pair of original vertices in each group that are connected in the original graph.

Following [35, 34, 102], we compute a maximal modularity clustering [85] as the obtained clusters are well adapted to subsequent visual representation, as shown in [90]. Maximizing $Q_{\mathscr{P}}$ over all the partitions $\mathscr{P}$ provides an optimal J classes partition. This is an NP-Hard and can only be solved via some heuristics. As in [102], we use a modified version of the multi-level greedy merging approach proposed in [91]: our modification guarantees that the final clusters are connected. The optimization process is carried out on the partitions for a given number of clusters J but also over the number of clusters J itself which is then automatically selected. This makes the method



(a) (b)

Figure 4.2.1: *(a): Raw view of the giant component for the Cuban HIV epidemics. (b) Modularity clustering of the giant component in 37 classes.*

essentially parameter free.

It should be noted however that one can find partitions with a rather high modularity even in completely random graphs (configuration model graphs where vertices have different degrees but are paired independently) where no modular structure actually exists (see [96] for an estimation of the expected value of this spurious modularity in the limit of large and dense graphs). To check that the modular structure found in a network cannot be explained by this phenomenon, we use the simulation approach proposed in [36, 102]. Using a Markov Chain Monte Carlo (MCMC) approach inspired by [99], we generate configuration model graphs with exactly the same size and degree distribution as the epidemics graph. Using the above algorithm, we compute a maximal modularity clustering on each of those graphs. The modularities of the clustering provide an estimate of the distribution of the maximal modularity in random graphs with our degree distribution. If a partition of this graph exhibits a higher modularity, we conclude that it must be the result of some actual modular structure rather than a random outcome.

The maximal modularity clustering is visualised using the force directed placement algorithm described above. In addition to giving a general idea of the global structure of the graph, the obtained visual representation can be used to display distributions of covariates at the cluster level. Homogeneity tests are performed in order to assess possible significant differences between these statistical subpopulations.

However, as demonstrated in [55], finding the maximal modularity clustering can lead to ignoring small modular structures that fall below the resolution limit of the modularity measure. It is then recommended in [55] to recursively apply maximal modularity clustering to the original clusters in order to investigate potential smaller scale modules. We follow this strategy coupled with the MCMC approach described above: each cluster is tested for substructure by applying the maximal modularity clustering technique from [102] and by assessing the actual significance of a potential sub-structure via comparison with similar random graphs.

To sum up, we recall the procedure that we recommend for clustering a large network:

- maximization of the modularity (4.1.1) (see [85]).
  - this favours dense clusters and produces interesting partitions for visualization (Fortunato 2010)
  - the optimisation is an NP-hard problem but high quality sub-optimal solutions can be obtained by annealing (Rossi Villa-Vialaneix 2010) or other methods (Noak Rotta, 2009)
- Clustering significance:
  - compute the modularity of the partition that is obtained,

  - test the significance of the obtained partition by simulating configuration models with same degree distribution and compute modularity.

- **Hierarchical clustering:** if the first clustering is relevant, and if the classes have large sizes, we can refine the partition.
  - Reiterate the clustering for each element of the partition, without taking inter-cluster connections.
  - Test the significance of the cluster's partition
  - Test the significance of the global clustering of the graph.
- **Coarsening:** merge clusters that induce the least reduction in modularity as long as we remain above the original graph.
- **Visualization:** use the Fruchterman–Reingold algorithm to display the network of clusters

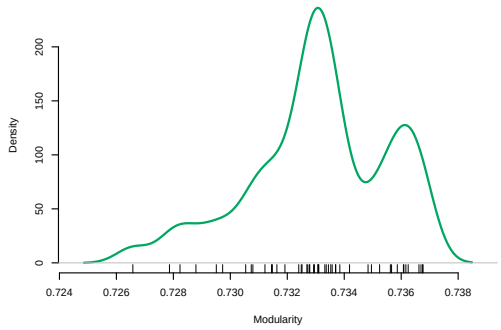
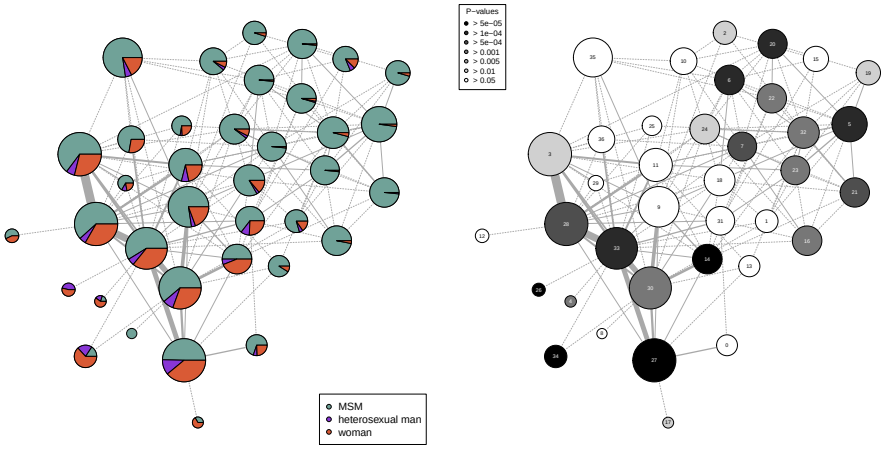


Figure 4.2.2: *In Figure 4.2.1, a modularity clustering is performed on the Cuban HIV data. The modularity of the partition obtained is* $\simeq 0.85$*. To test the significancy of this partition, 100 configuration model graphs with same size and same degree distribution as the observed one are simulated. The empirical distribution of the random modularity obtained by these simulations is depicted with small black bars on the abscissa axis and has a support bounded by* 0.74*. This shows that the partition obtained by maximizing the modularity is significant (at level 95% for instance).*

4.3 Analysis of the "giant component"

The network density is globally low and very heterogeneous. But although the connectivity of the network seems fragile at first glance, density may be locally very high. The harmonic average of the geodesic path lengths equals 10.24 and 12.2 for the directed graph (taking into account the information of who mentions whom). Most of the graph connectivity is concentrated in the largest component (3,168 edges out of 4,073). The largest component has a diameter of 26 (36 when taking into account the direction of the infections) and the harmonic average of the geodesic path lengths are the same inside the largest component. These values are slightly higher than those of other real networks mentioned in [89] but remain well below the number of vertices and compatible with the logarithmic scaling related to the so-termed small world effect.

Figure 4.2.1 (b) seems quite clear, with what appears to be two parts in the graph: the lower part of the graph (on the figure) seems to be dominated by MSM while the upper part gathers almost all persons from the giant component that have only heterosexual contacts. However, the upper part is quite difficult to read as it seems denser than the lower part. The layout shows what might be interpreted as cycles and also a lot of small trees connected to denser parts. The actual connection patterns between the upper part and the lower part are also very unclear. Because of these crowding effects, structural properties of the network from Figure 4.2.1 appears



(a) (b)

Figure 4.3.1: *The giant component divided into 37 clusters. (a) Each disk of representation corresponds to one cluster and has an area proportional to the number of persons (original vertices) gathered in the associated cluster. The pie chart of the disk displays the percentage of MSM (green), of heterosexual men (blue) and of women (red) in the cluster. Links between clusters summarise the connectivity pattern between members of the clusters. The thinnest edge width corresponds to only one connection between a member of one cluster and another person in the connected cluster (the corresponding edges are drawn using dashed segments). Thicker edges have a width proportional to the number of connected persons. (b) Disk areas and edges thicknesses are chosen as in (a). The grey level of a disk encode the p-value of a χ^2 test of homogeneity in which the distribution of the sexual orientations in the associated cluster is compared to the distribution in the giant component.*

quite difficult and probably misleading. We rely therefore on the simplification technique outlined in Section 4.2 leveraging a clustering of the giant component to get an insight into its general organization.

A graphical representation of the partition obtained by the method from [102] is displayed in Figure 4.3.1 (a). The clustering thus produced exhibits a modularity of 0.8522 and is made up of 37 clusters. This modularity is very high compared to the random level and strongly supports the hypothesis of a specific (“non-random") underlying community structure. For comparison purpose, the average maximal modularity attained by random graphs built from a configuration model with the same size and degree distribution as those of the giant component observed over a collection of 100 simulated replications (using the same partitioning method) is of the order 0.74, with a maximum of 0.7435.

Considering that the modules are meaningful, the visual representation provided by Figure 4.3.1 (a) is more faithful to the underlying graphical structure than the finer displays of Figure 4.2.1 (b). That said, the two graphs tend to agree as the pie charts of Figure 4.3.1 clearly show two parts in the network: the lower left part seems to gather most of the women and heterosexual men (as the upper part of Figure 4.2.1 (b)), while the upper right part contains clusters made almost entirely of MSM, as the lower part of Figure 4.2.1 (b). While the display of Figure 4.3.1 (a) might seem cluttered, it is in fact very readable if one considers that only 328 edges of the giant component connect persons from different clusters while 2,840 connections happen inside clusters. Then most of the edges on Figure 4.3.1 (a) could be disregarded as they corresponds to only one pair of connected persons (this is the case of 94 of such edges out of 142 and the former are represented as dashed segments). Taken this aspect into account, it appears that the MSM part of the giant component (upper right part) is made of loosely connected clusters while the bulk of the connectivity between clusters is gathered in the mixed part of the component, in which most women and heterosexual men are gathered. The fact that the mixed part is more dense was already visible in Figure 4.2.1 (b), but Figure 4.3.1 (a) provides a much stronger demonstration.

The pie chart based visualization of Figure 4.3.1 (a) shows the sexual orientation distribution in the clusters

and hence sheds light on its relationship with the graphical structure. In Figure 4.3.1 (b), a visual representation of the corresponding p-values is given. The darker the node, the more statistically significant the difference between the cluster distribution of sexual orientation and the distribution of the giant component.

Combining Figures 4.3.1 (a) and (b) is very useful: Figure (b) highlights atypical clusters while Figure (a) identifies why they are atypical. It appears that among the 37 clusters, 22 exhibit a χ^2 p-value below 5%. They will be abusively referred to as "atypical clusters" in the following. The set of those clusters can be split into two subsets, depending on the percentage of MSM in the cluster: above or below the global value of 76% (the percentage in the giant component), as illustrated by Figure 4.3.1 (b). Almost two thirds (67%) of the individuals of the largest connected component lie in the atypical clusters. Among the latter, 774 individuals belong to the 12 clusters which display a large domination of MSM (denoted the MSM group of clusters in the sequel) and 825 to the 10 clusters that contain an unexpectedly large number of heterosexual persons (denoted the mixed group of clusters in the sequel).

According to Figure 4.3.1, the two subsets of atypical clusters seem to be almost disconnected. This is confirmed by a detailed connectivity analysis. There are indeed 864 internal connections in the MSM group, 1,276 in the heterosexual group, and only 10 links between pairs of individuals belonging to the two different groups. This asymmetry was expected, given the quality of the clustering with only 328 inter-cluster connections. Nevertheless, the number of connections between the two groups of clusters is also small compared to connections between the clusters of the groups: 129 connections between persons of distinct clusters in the group of mixed clusters and 55 in the group of MSM clusters. Finally, there are 83 connections from persons in the group of mixed clusters to persons in non-atypical clusters, and 36 connections from persons in the group of MSM clusters to persons in non-atypical clusters. Mean geodesic distances inside the MSM group are larger than in the mixed group (respectively 9.95 and 7.28, computed without orientation). To conclude, the two groups are weakly connected to the outside, with a small number of direct connections, and rather internally more connected than expected.

4.4 Descriptive statistics for epidemic on networks

We now review some basic descriptive statistics for networks. Exhaustive statistical exploration of networks has been described by Newman [89] for example.

4.4.1 Estimating degree distributions

For the Cuban HIV data, we want to calculate for instance the degree distribution $(p_k : k \in \mathbb{N})$ using the number of declared sexual partners in the two years preceding detection, where p_k is the proportion of vertices having declared k sexual partners.

The degree distributions of most real-world networks, referred to as scale-free networks, often exhibit a power-law behaviour in their right tails (see [50]), i.e.

$$p_k \sim k^{-\alpha}, \text{ as } k \text{ becomes large,}$$

for some exponent $\alpha > 1$ (notice that $\sum_{k=1}^{\infty} 1/k^{\alpha} < \infty$ in this case). Roughly speaking, this describes the situations where the majority of vertices have few connections, but a small fraction of the vertices are highly connected (e.g. Chapter 4 in [84] for further details). We propose to fit a power-law exponent and consider two methods for this purpose, see also [33]. First, we minimize, over $\alpha > 1$, the following measure of dissimilarity between the observed degree distribution and the power-law distribution with exponent α based on degree values larger than k_0

$$\mathscr{K}_{k_0}(p,\alpha) = \sum_{k \geq k_0} \frac{p_k}{c_{p,k_0}} \log\left(\frac{C_\alpha \cdot p_k}{c_{p,k_0} \cdot k^{-\alpha}}\right), \tag{4.4.1}$$

where log denotes the natural logarithm, $c_{p,k_0} = \sum_{k \geq k_0} p_k$ and $C_\alpha = \sum_{k \geq k_0} 1/k^{\alpha}$. Notice that, when k_0 is larger than the maximum observed degree distribution $k_{\max}$, we have $\mathscr{K}_{k_0}(p,\alpha) = 0$ no matter the exponent α. Also, the computation of (4.4.1) involves summing a finite number of terms only, since the empirical frequency p_k is equal

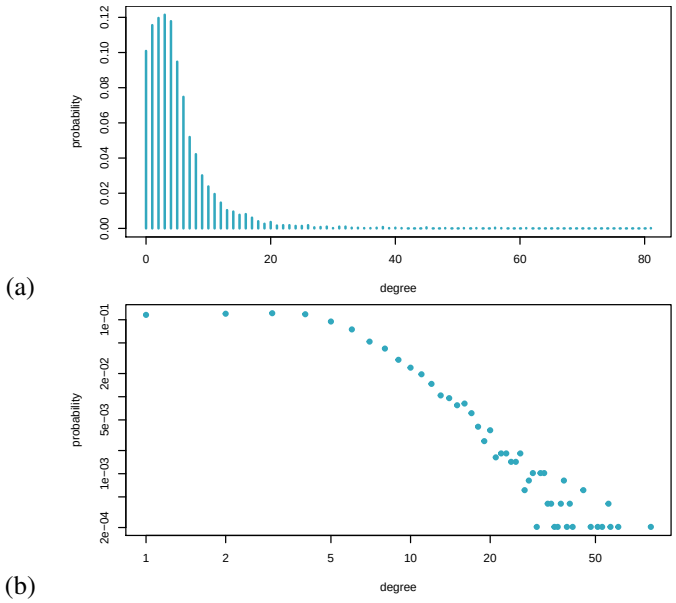


Figure 4.4.1: *(a) Distribution of the declared number of sexual partners for the HIV+ individuals detected and present in the Cuban database. (b) Preceding degree distribution plotted in a log-log scale: the graph exhibits a power-law behaviour.*

to zero for any degree k sufficiently large. The criterion $\mathscr{K}_{k_0}(p,\alpha)$ is known as the Kullback–Leibler divergence between the empirical and theoretical conditional distributions given that the degree is larger than k_0. Incidentally, we point out that other dissimilarity measures could be considered for the purpose of fitting a power-law, such as the χ^2-distance for instance. For a fixed threshold $k_0 \geq 1$, it is natural to select the value of the power-law exponent that provides the best fit, that is:

$$\widehat{\alpha}_{k_0} = \arg\min_{\alpha>1} \mathscr{K}_{k_0}(p,\alpha).$$

Choosing k_0 precisely being a challenging question to statisticians. Following in the footsteps of the heuristic selection procedures proposed in the context of heavy-tailed continuous distributions (see Chapter 4 in [97]), when possible, we suggest to choose $\widehat{\alpha}_{k_0}$ with k_0 in a region where the graph $\{(k,\ \widehat{\alpha}_k):\ k=1,\ \ldots,\ k_{\max}\}$ is becoming horizontal, or at least shows an inflexion point. For completeness, we also compute the Hill estimator:

$$\widetilde{\alpha}_m = \left(\frac{1}{m} \sum_{j=1}^{m} \frac{k_{(j)}}{k_{(m)}} \right)^{-1},$$

where n is the number of vertices of the graph under study, $1 \leq m \leq n$ and $k_{(1)} = k_{\max},\ k_{(2)},\ \ldots,\ k_{(m)}$ denote the m largest observed degrees sorted in decreasing order of their magnitude. The tuning parameter m is selected graphically, by plotting the graph $\{(m, \widetilde{\alpha}_m):\ m=1,\ \ldots,\ n\}$. In the case when the degrees of the vertices of the graph are independent, as for the configuration model [86], this statistic can be viewed as a conditional maximum likelihood estimator and arguments based on asymptotic theory supports its pertinence in this situation, see [61].

Let us consider the declared degree distribution in the Cuban database (see Fig. 4.4.1). Among the 5,389 individuals appearing in the database, 483 declared no sexual partners during this period. Degree distributions for the whole population exhibit a clear power-law behaviour. Power laws are fitted to the declared degree distributions, for the whole population and for the strata defined by the variable gender/sexual orientation respectively. Both methods present similar results. The resulting estimates (see Table 4.4.1) reveal the thickness of the upper tails: the smaller the tail exponent α, the heavier the distribution tail. Women correspond to the heaviest tail, followed by MSM and heterosexual men. However, an ANOVA reveals no statistically significant impact of the covariates

gender/sexual orientation. All the same, using the observed degree distribution, we obtain $(k_0,\ \alpha) = (3,\ 2.99)$ which is very close to the result when using the number of neighbours having been detected positive.
All the tail exponent estimates are below the critical value of $\alpha_c = 3.4788$, below which a giant component exists in scale-free networks generated by means of the configuration model, and above the value 2, below which the whole graph reduces to the giant component with probability one (see [80, 89]).

	$\widehat{k}_0$	$\widehat{\alpha}_{k_0}$	Mean	Std dev.	Min	Max
Whole population	7	3.06	6.17	5.54	1	82
Women	6	2.71	5.88	5.03	1	39
Heterosexual men	7	3.36	4.98	4.11	1	30
MSM	7	3.02	6.43	5.84	1	82

Table 4.4.1: Degree distribution for the Cuban HIV+ network, for the whole population and by sexual orientation.

For completeness, we can also compare with the Hill estimator (4.4.1) to the estimator (4.4.1) in each case, obtained by plotting the curves $(m, \widetilde{\alpha}_m)$ in Fig. 4.4.2: reassuringly, we found that both estimation methods yield similar results.

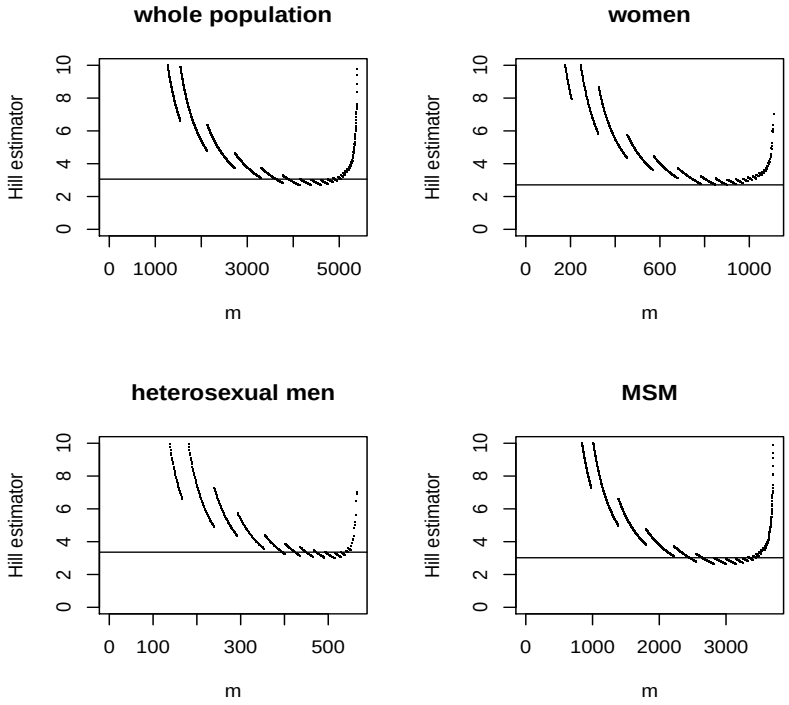


Figure 4.4.2: *Graph of $(m, \widetilde{\alpha}_m)$ for $m \in \{1, \ldots, n\}$. This graph allows us to choose the Hill estimator. The horizontal line $y = \widehat{\alpha}_{k_0}$ permits to visualize the estimator $\widehat{\alpha}_{k_0}$ and compare it with the Hill estimator.*

4.4.2 Joint degree distribution of sexual partners.

The independence assumption between the degrees of adjacent vertices does not hold here, see Fig. 4.4.3, in contrast to what is assumed for the vast majority of graph-based SIR models of epidemic disease, e.g. [50, 89].

Indeed, the linear correlation coefficient between the degree distributions of alters and egos is equal to 0.68. Testing the significance of this coefficient, that describes the correlation of these degree distributions, allows us to test the independence of the latter. Independence between the degree distributions of alters and egos is rejected by a χ^2-test with a p-value of $6.85\,10^{-6}$. In particular, highly connected vertices tend to be connected to vertices with a high number of connections too. From the perspective of mathematical modeling, this suggests to consider graph models with a dependence structure between the degrees of adjacent nodes, in opposition to most percolation processes on (configuration model) networks used to describe the spread of epidemics [80, 13, 111, 45, 59]. However, it is worth noticing that, if we restrict our analysis to some specific, more homogeneous, subgroups, the independence assumption may be grounded in evidence. So if assumptions such that the network is generated by a configuration model do not hold globally, they may be valid for smaller clusters, which is another motivation for clustering.

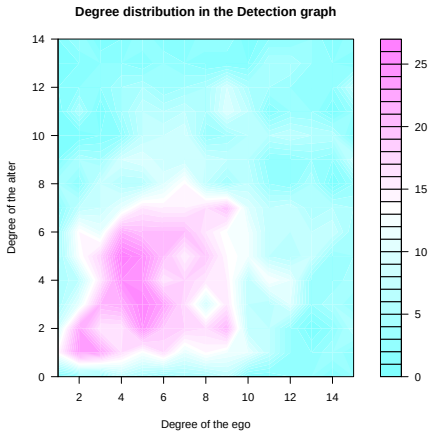


Figure 4.4.3: *Joint degree distribution of the number of contacts for connected vertices.*

4.4.3 Computation of geodesic distances and other connectivity properties

There is a large literature on describing the social networks on which epidemics might propagate (see Newman [89] for a more exhaustive list of descriptive statistics, and [36, 37] for an application to the Cuban HIV epidemics). Here, we mention some of them, related to community and connexity. All results presented here are obtained with the R-package `igraph` [43].

A set of connected vertices with the corresponding edges, constitutes a component of the graph. The collection of components forms a partition of the graph. We identify the components of the network and compute their respective sizes. When the size of the largest component is much larger than the size of the second largest component, see section IV A in [89] and the references therein, one then refers to the notion of giant component.

A geodesic path between two connected vertices x and y is a path with shortest length that connects them, its length $d(x,y)$ being the geodesic distance between x and y. One also defines the mean geodesic distance:

$$\mathscr{L} = \frac{1}{n(n+1)} \sum_{(x,y)\in\mathscr{V}^2} d(x,y),$$

where $\mathscr{V}$ denotes the set of all vertices of the connected graph and n its size. For non-connected graphs, one usually computes a harmonic average. Mean geodesic distances measure “how far" two randomly chosen vertices are, given the network structure. When $\mathscr{L}$ is much smaller than n, one says that a “small-world effect" is observed. In this regard, the diameter of a connected graph, that is to say the length of the longest geodesic path, is also a

quantity of major interest:

$$\delta = \max_{(x,y)\in\mathscr{V}^2} d(x,y).$$

Computations have been made for each component of the network of sexual contacts among individuals diagnosed as HIV positive before 2006 in Cuba, using the dedicated “burning algorithm" for the mean geodesic distances, see [4].

Along these lines, we also investigate how the connectivity properties of the network evolve when removing various fractions of specific strata of the population: we studied the resilience to various strata (robustness of certain statistics such as mean geodesic distance or size of the largest component to deletion of points in these strata), the clustering coefficients (defined as the number of triangles over the number of connected triples of vertices) and the articulation points (points that disconnect the component they belong to into two components when removed; see Section 6 of [37]). Indicators show an apparent weak resilience: 1,157 articulation points (out of 2,386 nodes), only 187 cliques (among them 177 triangles) and low assortative mixing coefficients. Global statistics thus indicate a low density of the graph (many articulation points, resilient structure, low clustering coefficients), the clustering emphasised the important heterogeneity in the network, with some dense regions that are internally more connected than average and with few links to the outside. We found subgroups with atypical covariate distributions, each reflecting a different stage of the evolution of the epidemic. Clustering the graph also allows us to unfold the complex structure of the Cuban HIV contact-tracing network. As a byproduct, the clustering indicates sub-structures that may be considered as random graphs resulting from configuration models, bridging the gap between the modelling papers whose assumptions on network structures do not often match reality.

Appendix: Finite Measures on $\mathbb{Z}_+$

First, some notation is needed in order to clarify the way the atoms of a given element of $\mathscr{M}_F(\mathbb{Z}_+)$ are ranked. For all $\mu \in \mathscr{M}_F(\mathbb{Z}_+)$, let F_μ be its cumulative distribution function and F_μ^{-1} be its right inverse defined as

$$\forall x \in \mathbb{R}_+, F_\mu^{-1}(x) = \inf\{i \in \mathbb{Z}_+, F_\mu(i) \geq x\}. \tag{A.0.2}$$

Let $\mu = \sum_{n\in\mathbb{Z}_+} a_n \delta_n$ be an integer-valued measure of $\mathscr{M}_F(\mathbb{Z}_+)$, i.e. such that the a_n's are themselves integers. Then, for each atom $n \in \mathbb{Z}_+$ of μ such that $a_n > 0$, we duplicate the atom n with multiplicity a_n, and we rank the atoms of μ by increasing values, sorting arbitrarily the atoms having the same value. Then, we denote for any $i \leq \langle \mu, \mathbf{1} \rangle$,

$$\gamma_i(\mu) = F_\mu^{-1}(i), \tag{A.0.3}$$

the level of the i^{th} atom of the measure, when ranked as described above. We refer to Example 3.3.3 for a simple illustration.

We now make precise a few topological properties of spaces of measures and measure-valued processes. For $T > 0$ and a Polish space (E, d_E), we denote by $\mathbb{D}([0,T], E)$ the Skorokhod space of càdlàg (right-continuous and left-limited) functions from $[0,T]$ into E (e.g. [24, 67]) equipped with the Skorokhod topology induced by the metric

$$d_T(f,g) := \inf_{\alpha\in\Delta([0,T])} \left\{ \sup_{\substack{(s,t)\in[0,T]^2,\\ s\neq t}} \left| \log \frac{\alpha(s)-\alpha(t)}{s-t} \right| + \sup_{t\leq T} d_E\big(f(t), g(\alpha(t))\big) \right\}, \tag{A.0.4}$$

where the infimum is taken over the set $\Delta([0,T])$ of continuous increasing functions $\alpha : [0,T] \to [0,T]$ such that $\alpha(0) = 0$ and $\alpha(T) = T$.

Limit theorems are heavily dependent on the topologies considered. We introduce here several technical lemmas on the space of measures related to these questions. For any fixed $0 \leq \varepsilon < A$, recall the definition of $\mathscr{M}_{\varepsilon,A}$ in (3.3.20). Note that for any $\nu \in \mathscr{M}_{\varepsilon,A}$, and $i \in \{0,\ldots,5\}$, $\langle \nu, \chi^i \rangle \leq A$ since the support of ν is included in $\mathbb{Z}_+$.

Lemma A.0.1. *Let $\mathfrak{I}$ be an arbitrary set and consider a family $(\nu_\tau, \tau \in \mathfrak{I})$ of elements of $\mathscr{M}_{\varepsilon,A}$. Then, for any real-valued function f on $\mathbb{Z}_+$ such that $f(k) = o(k^5)$, we have that*

$$\lim_{K\to\infty} \sup_{\tau\in\mathfrak{I}} |\langle \nu_\tau, f\mathbf{1}_{[K,\infty)} \rangle| = 0.$$

Proof. By Markov inequality, for any $\tau \in \mathfrak{I}$, for any K, we have

$$\sum_{k\geq K} |f(k)| \nu_\tau(k) \leq A \sup_{k\geq K} \frac{|f(k)|}{k^5},$$

hence

$$\lim_{K\to\infty} \sup_{\tau\in\mathfrak{I}} |\langle \nu_\tau, f \rangle| \leq A \limsup_{k\to\infty} \frac{|f(k)|}{k^5} = 0.$$

The proof is thus complete. □

Lemma A.0.2. *For any $A > 0$, the set $\mathscr{M}_{\varepsilon,A}$ is a closed subset of $\mathscr{M}_F(\mathbb{Z}_+)$ embedded with the topology of weak convergence.*

Proof. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of $\mathscr{M}_{\varepsilon,A}$ converging to $\mu \in \mathscr{M}_F(\mathbb{Z}_+)$ for the weak topology, Fatou's lemma for sequences of measures implies

$$\langle \mu, \chi^5 \rangle \leq \liminf_{n\to\infty} \langle \mu_n, \chi^5 \rangle.$$

Since $\langle \mu_n, \mathbf{1} \rangle$ tends to $\langle \mu, \mathbf{1} \rangle$, we have that $\langle \mu, \mathbf{1} + \chi^5 \rangle \leq A$.

Furthermore, by uniform integrability (Lemma A.0.1), it is also clear that

$$\varepsilon \leq \lim_{n\to\infty} \langle \mu_n, \chi \rangle = \langle \mu, \chi \rangle,$$

which shows that $\mu \in \mathscr{M}_{\varepsilon,A}$. □

Lemma A.0.3. *The traces on $\mathscr{M}_{\varepsilon,A}$ of the total variation topology and of the weak topology coincide.*

Proof. It is well known that the total variation topology is coarser than the weak topology. In the reverse direction, assume that $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of weakly converging measures all belonging to $\mathscr{M}_{\varepsilon,A}$. Since,

$$d_{TV}(\mu_n, \mu) \leq \sum_{k\in\mathbb{Z}_+} |\mu_n(k) - \mu(k)|.$$

according to Lemma A.0.1, it is then easily deduced that the right-hand side converges to 0 as n goes to infinity. □

Lemma A.0.4. *If the sequence $(\mu_n)_{n\in\mathbb{N}}$ of $\mathscr{M}_{\varepsilon,A}^{\mathbb{N}}$ converges weakly to the measure $\mu \in \mathscr{M}_{\varepsilon,A}$, then $(\langle \mu_n, f \rangle)_{n\in\mathbb{N}}$ converges to $\langle \mu, f \rangle$ for all function f such that $f(k) = o(k^5)$ for all large k.*

Proof. The triangle inequality implies that:

$$|\langle \mu_n, f \rangle - \langle \mu, f \rangle| \leq |\langle \mu_n, f\mathbf{1}_{[0,K]} \rangle - \langle \mu, f\mathbf{1}_{[0,K]} \rangle| + |\langle \mu, f\mathbf{1}_{(K,+\infty)} \rangle| + |\langle \mu_n, f\mathbf{1}_{(K,+\infty)} \rangle|.$$

We then conclude by uniform integrability and weak convergence. □

Recall that $\mathscr{M}_{\varepsilon,A}$ can be equipped with the total variation distance topology, hence the topology on $\mathbb{D}([0,T], \mathscr{M}_{\varepsilon,A})$ is induced by the distance

$$\rho_T(\mu_\cdot, \nu_\cdot) = \inf_{\alpha\in\Delta([0,T])} \left(\sup_{\substack{(s,t)\in[0,T]^2, \\ s\neq t}} \left| \log \frac{\alpha(s) - \alpha(t)}{s - t} \right| + \sup_{t\leq T} d_{TV}(\mu_t, \nu_{\alpha(t)}) \right).$$

Bibliography

[1] E. Abbe. Community detection and stochastic block models: recent development. *Journal of Machine Learning Research*, 18:1–86, 2018.

[2] L. Addario-Berry, N. Broutin and C. Goldschmidt, Critical random graphs: limiting constructions and distributional properties, *Electronic Journal of Probability*, 15(25):741–775, 2010.

[3] L. Addario-Berry, N. Broutin and C. Goldschmidt, The continuum limit of critical random graphs, *Probability Theory and Related Fields*, 152(3-4):367–406, 2012.

[4] R. Ahuja, T. Magnanti and J. Orlin, *Network flows: theory, algorithms and applications*, Prentice Hall, New Jersey, 1993.

[5] H. Anderson and T. Britton, *Stochastic Epidemic models and Their Statiatical Analysis*, volume 151 of *Lecture Notes in Statistics*, Springer, New York, 2000.

[6] H. Andersson, Limit theorems for a random graph epidemic model, *Annals of Applied Probability*, 8(4):1331–1349, 1998.

[7] H. Andersson, Epidemic models and social networks, *Mathematical Scientist*, 24(2):128–147, 1999.

[8] H. De Arazoza, J. Joanes, R. Lounes, C. Legeai, S. Clémencon, J. Perez and B. Auvert, The HIV/AIDS epidemic in Cuba: description and tentative explanation of its low prevalence, *BMC Infectious Disease*, 7:130, 2007.

[9] F. Ball, T. Britton and D. Sirl, A network with tunable clustering, degree correlation and degree distribution, and an epidemic thereon, *Journal of Mathematical Biology*, 66(4-5):979–1019, 2013.

[10] F. Ball and D. Clancy, The final size and severity of a generalised stochastic multitype epidemic model, *Advances in Applied Probability*, 25(4):721–736, 1993.

[11] F. Ball and P. Donnelly, Strong approximations for epidemic models, *Stochastic Processes and their Applications*, 55(1):1–21, 1995.

[12] F. Ball, D. Mollison and G. Scalia-Tomba, Epidemics with two levels of mixing, *The Annals of Applied Probability*, 7:46–89, 1997.

[13] F. Ball and P. Neal, Network epidemic models with two levels of mixing, *Mathematical Biosciences*, 212:69–87, 2008.

[14] F. Ball, L. Pellis and P. Trapman, Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R_0, *Mathematical Biosciences*, 235(1):85–97, 2012.

[15] F. Ball, L. Pellis and P. Trapman, Reproduction numbers for epidemic models with households and other social structures II: Comparisons and implications for vaccination, *Mathematical Biosciences*, 274:108–139, 2016.

[16] F. Ball, D. Sirl and P. Trapman, Epidemics on random intersection graphs, *Annals of Applied Probability*, 24(3):1081–1128, 2014.

[17] F. Ball, L. Pellis and P. Trapman, Reproduction numbers for epidemic models with households and other social structures II: comparisons and implications for vaccination, to appear in Math. Biosci.; arXiv preprint arXiv:1410.4469, 2016.

[18] S. Bansal, B.T. Grenfell and L.A. Meyers, When individual behaviour matters: homogeneous and network models in epidemiology, *Journal of the Royal Society Interface*, 4(16):879–891, 08 2007.

[19] A.D. Barbour and G. Reinert, Approximating the epidemic curve, *Electronic Journal of Probability*, 18(54):2557, 2013.

[20] M. Barthélemy, A. Barrat, R. Pastor-Satorras and A. Vespignani, Dynamical patterns of epidemic outbreaks in complex heterogeneous networks, *Journal of Theoretical Biology*, 235:275–288, 2005.

[21] M.S. Bartlett, *Stochastic Population Models in Ecology and Epidemiology*, London, methuen edition, 1960.

[22] G. Di Battista, P. Eades, R. Tamassia and I.G. Tollis, *Graph Drawing: Algorithms for the Visualization of Graphs*, Prentice Hall, 1999.

[23] N.G. Becker and K. Dietz, The effect of household distribution on transmission and control of highly infectious diseases, *Math. Biosci.*, 127(2):207–219, 1995.

[24] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[25] B. Bollobás, *Random graphs*, Cambridge University Press, 2 edition, 2001.

[26] C. Borgs, J. Chayes, L. Lovász, V. Sós and K. Vesztergombi, Limits of randomly grown graph sequences, *European Journal of Combinatorics*, 32(7):985–999, 2011.

[27] T. Britton, M. Deijfen, A.N. Lagerås and M. Lindholm, Epidemics on random graphs with tunable clustering, *Journal of Applied Probability*, 45:743–756, 2008.

[28] T. Britton, S. Janson, and A. Martin-Löf, Graphs with specified degree distributions, simple epidemics, and local vaccination strategies, *Adv. Appl. Probab.*, 39(4):922–948, 2007.

[29] F. Ball, T. Britton, C. Larédo, E. Pardoux, D. Sirl and V.C. Tran, *Stochastic Epidemic Models with Inference*, T. Britton and E. Pardoux eds., Lecture Notes in Mathematics, Mathematical Biosciences, Vol. 2255, 2019.

[30] N. Champagnat, S. Méléard and V.C. Tran, Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer, arXiv:1901.02385, 2019.

[31] S.K. Chan, L.R. Thornton, K.J. Chronister, J. Meyer, M. Wolverton, C.K. Johnson, R.R. Arafat, P. Joyce, W.M. Switzer, W. Heneine, A. Shankar, T. Granade, S. Michele Owen, P. Sprinkle and V. Sullivan, Likely Female-to-Female sexual transmission of HIV-Texas, *Morbidity and Mortality Weekly Report*, 63(10):209–212, 2014, Centers for Disease Control and Prevention.

[32] S. Chatterjee, *Large Deviations for Random Graphs*, volume 2197 of *Lecture Notes in Mathematics, Ecole d'Eté de Probabilités de Saint-Flour XLV - 2015*, Springer, Cham, 1 edition, 2017.

[33] A. Clauset, C. Shalizi and M. Newman, Power-law distributions in empirical data, *SIAM Review*, 51:661–703, 2009.

[34] S. Clémencon, H. De Arazoza, F. Rossi and V.C. Tran, Hierarchical clustering for graph vizualization, In *Proceedings of XVIIIth European Symposium on Artificial Neural Networks (ESANN 2011)*, pages 227–232, Bruges, Belgium, April 2011, `http://hal.archives-ouvertes.fr/hal-00603639/fr/`.

[35] S. Clémencon, H. De Arazoza, F. Rossi and V.C. Tran, Visual mining of epidemic networks, In *Proceedings of the International Work conference of Artificial Neural Networks (IWANN)*, volume 6692 of *Lecture Notes in Computer Sciences*, pages 276–283. Springer, June 2011.

[36] S. Clémencon, H. De Arazoza, F. Rossi and V.C. Tran, A statistical network analysis of the hiv/aids epidemics in cuba, *Social Network Analysis and Mining*, 5:Art.58, 2015.

[37] S. Clémencon, H. De Arazoza, F. Rossi and V.C. Tran, Supplementary materias for "a statistical network analysis of the HIV/AIDS epidemics in Cuba", *Social Network Analysis and Mining*, 5, 2015. Supplementary materials.

[38] S. Clémencon, V.C. Tran and H. De Arazoza, A stochastic SIR model with contact-tracing: large population limits and statistical inference, *Journal of Biological Dynamics*, 2(4):391–414, 2008.

[39] A. Cori, A.J. Valleron, F. Carrat, G. Scalia-Tomba, G. Thomas and P.Y. Boëlle, Estimating influenza latency and infectious period durations using viral excretion data, *Epidemics*, 4(3):132–138, 2012.

[40] E. Coupechoux and M. Lelarge, How clustering affects epidemics in random networks, *Advances in Applied Probability*, 46:985–1008, 2014.

[41] E. Coupechoux and M. Lelarge,
Contagions in random networks with overlapping communities, *Advances in Applied Probability*, 47(4):973–988, 2015.

[42] A. Cousien, V.C. Tran, S. Deuffic-Burban, M. Jauffret-Roustide, G. Mabileau, J.S. Dhersin and Y. Yazdanpanah, Effectiveness and cost-effectiveness of interventions targeting harm reduction and chronic hepatitis c cascade of care in people who inject drugs: the case of France, *Journal of Viral Hepatitis*, 25(10):1197–1207, 2018.

[43] G. Csardi and T. Nepusz, The igraph software package for complex network research, *InterJournal*, Complex Systems:1695, 2006.

[44] L. Danon, A.P. Ford, T. House, C.P. Jewell, M.J. Keeling, G.O. Roberts, J.V. Ross and M.C. Vernon, Networks and the epidemiology of infectious disease, In *Interdisciplinary Perspectives on Infectious Diseases*, volume 2011, pages 1–28, 2011.

[45] L. Decreusefond, J.-S. Dhersin, P. Moyal and V.C. Tran, Large graph limit for a sir process in random network with heterogeneous connectivity, *Annals of Applied Probability*, 22(2):541–575, 2012.

[46] Demographic, Nigeria, Health survey (NDHS), *Problems in accessing health care. NDHS/National Population Commission*, page 140, 2003.

[47] O. Diekmann, M. Gyllenberg, J.A.J. Metz and H.R. Thieme, On the formulation and analysis of general deterministic structured population models. I. Linear theory, *Journal of Mathematical Biology*, 36(4):349–388, 1998.

[48] O. Diekmann, H. Heesterbeek and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*, Princeton Series in Theoretical and Computational Biology. Princeton University Press, New Jersey, 2012.

[49] O. Diekmann, M. Gyllenberg, J.A.J. Metz and H.R. Thieme, On the formulation and analysis of general deterministic structured population models. I. Linear theory, *J. Math. Biol.*, 36(4):349–388, 1998.

[50] R. Durrett, *Random graph dynamics*, Cambridge University Press, New York, 2007.

[51] K.T.D. Eames and M.J. Keeling, Modelling dynamic and network heterogeneities in the spread of sexually transmitted diseases, *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):13330–13335, 2002.

[52] J. Enright and R.R. Kao, Epidemics on dynamic networks, *Epidemics*, 24:88–97, 2018.

[53] S.N. Ethier and T.G. Kurtz, *Markov Processus, Characterization and Convergence*. John Wiley & Sons, New York, 1986.

[54] L.C. Evans, *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*, American Mathematical Society, 1998.

[55] S. Fortunato and M. Barthélemy, Resolution limit in community detection, *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007.

[56] N. Fournier and S. Méléard, A microscopic probabilistic description of a locally regulated population and macroscopic approximations, *Ann. Appl. Probab.*, 14(4):1880–1919, 2004.

[57] T. Fruchterman and B. Reingold, Graph drawing by force-directed placement, *Software-Practice and Experience*, 21:1129–1164, 1991.

[58] E. Goldstein, K. Paur, C. Fraser, E. Kenah, J. Wallinga and M. Lipsitch, Reproductive numbers, epidemic spread and control in a community of households, *Mathematical Biosciences*, 221(1):11–25, 2009.

[59] M. Graham and T. House, Dynamics of stochastic epidemics on heterogeneous networks, *Journal of Mathematical Biology*, 68(7):1583–1605, 2014.

[60] I. Herman, G. Melancon and M. Scott Marshall, Graph visualization and navigation in information visualisation, *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, 2000.

[61] B. Hill, A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, 3(5):1163–1174, 01 1975.

[62] T. House, Modelling epidemics on networks, *Contemporary Physics*, 53(3):213–225, 2012.

[63] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, volume 24, North-Holland Publishing Company, 1989, Second Edition.

[64] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin, 1987.

[65] P. Jagers, *Branching Processes with Biological Applications*, Wiley Series in Probability and Mathematical Statistics, Wiley-Interscience, London-New York-Sydney, 1975.

[66] S. Janson, M. Luczak and P. Windridge, Law of large numbers for the SIR epidemic on a random graph with given degrees, *Annals of Applied Probability*, 2014, accepted.

[67] A. Joffe and M. Métivier, Weak convergence of sequences of semimartingales with applications to multitype branching processes, *Advances in Applied Probability*, 18:20–65, 1986.

[68] W.O. Kermack and A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. Roy. Soc. Lond. A*, 115:700–721, 1927.

[69] I.Z. Kiss, L. Berthouze, J.C. Miller and P.L. Simon, Mapping out emerging network structures in dynamic network models coupled with epidemics, In *Temporal Network Epidemiology*, Theoretical Biology, pages 267–289. Springer, 2017.

[70] I.Z. Kiss, D.M. Green and R.R. Kao, Infectious disease control using contact tracing in random and scale-free networks, *J. R. Soc. Interface*, 3(6):55–62, 2013.

[71] I.Z. Kiss, J.C. Miller and P. Simon, *Mathematics of Epidemics on Networks*, volume 46 of *Interdisciplinary Applied Mathematics*, Springer, 1 edition, 2017.

[72] A. Kleczkowski and B.T. Grenfell, Mean-field-type equations for spread of epidemics: The small world model, *Physica A*, 274:355–360, 1999.

[73] M. Kretzschmar and M. Morris, Measures of concurrency in networks and the spread of infectious disease, *Math. Biosci.*, 133:165–195, 1996.

[74] L. Lovàsz and B. Szegedy, Limits of dense graph sequences, *Journal of Combinatorial Theory, Series B*, 96:933–957, 2006.

[75] T.L. Mah and J.D. Shelton, Concurrency revisited: increasing and compelling epidemiological evidence, *Journal of the International AIDS Society*, 14(33), 2011.

[76] R.M. May and A.L. Lloyd, Infection dynamics on scale-free networks, *Phys. Rev. E*, 64:066112, 2001.

[77] S. Méléard and S. Roelly, Sur les convergences étroite ou vague de processus à valeurs mesures, *CRAS de l'Acad. des Sci. Paris*, t. 317, Série I, 785–788, 1993.

[78] S. Méléard and V.C. Tran, Trait substitution sequence process and canonical equation for age-structured populations, *Journal of Mathematical Biology*, 58(6):881–921, 2009.

[79] J.C. Miller, A note on a paper by Erik Volz: SIR dynamics in random networks, *Journal of Mathematical Biology*, 62(3):349–358, 2011, `http://arxiv.org/abs/0909.4485`.

[80] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random structures and algorithms*, 6:161–180, 1995.

[81] M. Molloy and B. Reed, The size of the giant component of a random graph with a given degree sequence, *Combinatorics probability and computing*, 7(3):295–305, 1998.

[82] C. Moore and M.E.J. Newman, Epidemics and percolation in small-world networks, *Phys. Rev. E*, 61:5678–5682, 2000.

[83] M. Morris and M. Kretzschmar, Concurrent partnerships and transmission dynamics in networks, *Social Networks*, 17:299–318, 1995.

[84] M. Newman, A. Barabási and D. Watts, *The structure and dynamics of networks*, Princeton University Press, 2006.

[85] M. Newman and M. Girvan, Finding and evaluating community structure in network., *Physical Review E*, 69:026113, 2004.

[86] M. Newman, S. Strogatz and D. Watts, Random graphs with arbitrary degree distributions and their applications, *Physical Review E*, 64(2):026118, 2001.

[87] M.E.J. Newman, The spread of epidemic disease on networks, *Physical Reviews E*, 66, 2002.

[88] M.E.J. Newman, Mixing patterns in networks, *Phys. Rev. E*, 67:026126, 2003.

[89] M.E.J. Newman, The structure and function of complex networks, *SIAM Review*, 45:167–256, 2003.

[90] A. Noack, Modularity clustering is force-directed layout, *Physical Review E*, 79:026102, 2009.

[91] A. Noack and R. Rotta, Multi-level algorithms for modularity clustering, In *SEA '09: Proceedings of the 8th International Symposium on Experimental Algorithms*, pages 257–268,Springer-Verlag, Berlin, Heidelberg, 2009.

[92] R. Pastor-Satorras and A. Vespignani, Epidemics and immunization in scale-free networks, In *Handbook of Graphs and Networks: From the Genome to the Internet*, pages 113–132, Berlin, 2002. Wiley-VCH.

[93] L. Pellis, T. House and M.J. Keeling, Exact and approximate moment closures for non-Markovian network epidemics, *Journal of Theoretical Biology*, 382:160–177, 2015.

[94] L. Pellis, S.E.F. Spencer and T. House, Real-time growth rate for general stochastic sir epidemics on unclustered networks, *Mathematical biosciences*, 265:65–81, 2015.

[95] L. Pellis, N.M. Ferguson and C. Fraser, Epidemic growth rate and household reproduction number in communities of households, schools and workplaces, *J. Math. Biol.*, 63(4):691–734, 2011.

[96] J. Reichardt and S. Bornholdt, Partitioning and modularity of graphs with arbitrary degree distribution, *Physical Review E*, 76(1):015102, 2007.

[97] S. Resnick, *Heavy-tail phenomena*, Springer, 2007.

[98] O. Riordan, The phase transition in the configuration model, *Combinatorics, Probability and Computing*, 21(1-2):265–299, 2012.

[99] J.M. Roberts Jr., Simple methods for simulating sociomatrices with given marginal totals, *Social Networks*, 22(3):273–283, 2000.

[100] S. Roelly, A criterion of convergence of measure-valued processes: Application to measure branching processes, *Stochastics*, 17:43–65, 1986.

[101] D.A. Rolls, P. Wang, R. Jenkinson, P.E. Pattison, G.L. Robins, R. Sacks-Davis, G. Daraganova, M. Hellard and E. McBryde, Modelling a disease-relevant contact network of people who inject drugs, *Social Networks*, 35(4):699–710, 2013.

[102] F. Rossi and N. Villa-Vialaneix, Représentation d'un grand réseau à partir d'une classification hiérarchique de ses sommets, *Journal de la Société Francaise de Statistique*, 152(3):34–65, 2011.

[103] R.B. Rothenberg, D.E. Woodhouse, J.J. Potterat, S.Q. Muth, W.W. Darrow and A.S. Klovdahl, Social networks in disease transmission: The Colorado Springs study, In R.H. Needle, S.L. Coyle, S.G. Genser and R.T. Trotter II, editors, *Social networks, drug abuse and HIV transmission*, volume 151 of *Research Monographs*, pages 3–18. National Instit, 1995.

[104] B.V. Schmid and M. Kretzschmar, Determinants of sexual network structure and their impact on cumulative network measures, *PLoS Computational Biology*, 8(4):e1002470, 2012.

[105] Statistics Sweden, *Statistical Yearbook of Sweden 2014*, Statistics Sweden, 2014.

[106] V.C. Tran, *Modèles particulaires stochastiques pour des problèmes d'évolution adaptative et pour l'approximation de solutions statistiques*, Phd thesis, Université Paris X - Nanterre, 12 2006, `http://tel.archives-ouvertes.fr/tel-00125100`.

[107] V.C. Tran, *Une ballade en forêts aléatoires. Théorèmes limites pour des populations structurées et leurs généalogies, étude probabiliste et statistique de modèles SIR en épidémiologie, contributions à la géométrie aléatoire*, Habilitation à diriger des recherches, Université de Lille 1, 11 2014, `http://tel.archives-ouvertes.fr/tel-01087229`.

[108] P. Trapman, F. Ball, J.-S. Dhersin, V.C. Tran, J. Wallinga and T. Britton. Inferring r0 in emerging epidemics-the effect of common population structure is small, *Journal of the Royal Society Interface*, 13:20160288, 2016.

[109] D. Tunkelang, *A Numerical Optimization Approach to General Graph Drawing*. PhD thesis, School of Computer Science, Carnegie Mellon University, 01 1999.

[110] R. Van der Hofstad, *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge, 2017.

[111] E. Volz, SIR dynamics in random networks with heterogeneous connectivity, *Mathematical Biology*, 56:293–310, 2008.

[112] E. Volz and L. Ancel Meyers, Susceptible-infected-recovered epidemics in dynamic contact networks, *Proceeding of the Royal Society B*, 274:2925–2933, 2007.

[113] J. Wallinga and M. Lipsitch, How generation intervals shape the relationship between growth rates and reproductive numbers, *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604, 2007.

[114] J. Wallinga, P. Teunis and M. Kretzschmar, Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents, *Am. J. Epidemiol.*, 164(10):936–944, 2006.

[115] W. Whitt, Blocking when service is required from several facilities simultaneously, *AT&T Tech. J.*, 64:1807–1856, 1985.

[116] WHO Ebola Response Team, Ebola virus disease in West Africa − the first 9 months of the epidemic and forward projections, *N Engl J Med*, 371:1481–1495, 2014.

[117] J.L. Wylie and A. Jolly, Patterns of Chlamydia and Gonorrhea infection in sexual networks in Manitoba, Canada, *Sexually transmitted diseases*, 28(1):14–24, January 2001.