General epidemiological models: Law of large numbers and contact tracing

Jean-Jil Duchamps, Félix Foutel-Rodier, Emmanuel Schertzer

General Epidemiological Models: Law of Large Numbers and Contact Tracing

Jean-Jil Duchamps, Félix Foutel-Rodier, Emmanuel Schertzer

June 25, 2021

Abstract

We study a class of individual-based, fixed-population size epidemic models under general assumptions, e.g., heterogeneous contact rates encapsulating changes in behavior and/or enforcement of control measures. We show that the large-population dynamics are deterministic and relate to the Kermack–McKendrick PDE. Our assumptions are minimalistic in the sense that the only important requirement is that the basic reproduction number of the epidemic R_0 be finite, and allow us to tackle both Markovian and non-Markovian dynamics. The novelty of our approach is to study the "infection graph" of the population. We show local convergence of this random graph to a Poisson (Galton–Watson) marked tree, recovering Markovian backward-in-time dynamics in the limit as we trace back the transmission chain leading to a focal infection. This effectively models the process of contact tracing in a large population. It is expressed in terms of the Doob h-transform of a certain renewal process encoding the time of infection along the chain. Our results provide a mathematical formulation relating a fundamental epidemiological quantity, the generation time distribution, to the successive time of infections along this transmission chain.

Contents

1 Introduction

1.1 General individual-based epidemic model

In the present article, we study an extension of the general epidemiological framework introduced in [14] to model the COVID-19 epidemic. Let us briefly recall the main features of this model.

At time $t = 0$, we consider a population made of susceptible individuals, that have never encountered the disease, and infected individuals. Each infected individual is supposed to belong to one *compartment*, that models the stage of the disease of this individual. Classical examples of compartments are the exposed compartment (E), where the individual is infected but not yet infectious, the infectious compartment (I), and the recovered compartment (R), once the individual has become immunized. In the case of the COVID-19 epidemic, it might be relevant to add a hospitalized (H) and an intensive care unit (U) compartment, as monitoring the number of individuals in these states is typically important for policy-making. See Figure 1 for an example of compartmental model used for the COVID-19 epidemic. We denote by $\mathcal{S}$ the set of all compartments. For the sake of simplicity, we will also assume that $\mathcal{S}$ is finite.

We encode the compartment to which individual x belongs as a stochastic process $(X_x(a);\, a \geq 0)$ valued in $\mathcal{S}$, that we call the *life-cycle process*. The random variable $X_x(a)$ gives the compartment to which x belongs at *age of infection* a, that is, a unit of time after its infection. Moreover, individual x is endowed with a point measure $\mathcal{P}_x$ on $\mathbb{R}_+$ that we call the *infection point process*. The atoms of $\mathcal{P}_x$ encode the age at which x makes infectious contacts with the rest of the population. We think of the pair $(\mathcal{P}_x, X_x)$ as describing the course of the infection of individual x. We make the fundamental assumption that the pairs $(\mathcal{P}_x, X_x)$ are i.i.d. for distinct individuals in the population.

In [14] we assumed that susceptibles are in excess, and that any infectious contact leads to a new infection. The resulting population is then distributed as a Crump-Mode-Jagers (CMJ) process. In the current work, we consider an extension of this model that takes into account the saturation due to the finite pool of susceptibles in the population. More precisely, we consider a population of finite fixed size N. Each infectious contact is made with an individual uniformly chosen in this population, and it results in a new infection only if the targeted individual is susceptible. Finally, we model the impact of control measures, such as school closure, or national lockdown, with a *contact rate function* $(c(t);\, t \geq 0)$. This contact rate is such that an infection occuring at time t is only effective with

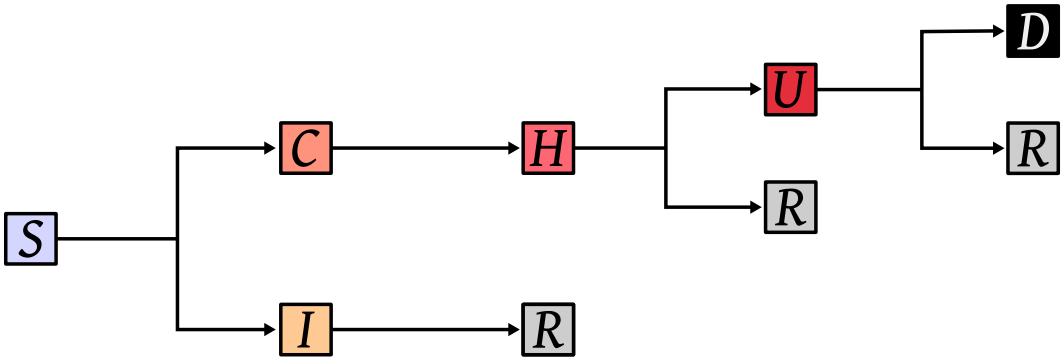

Figure 1: An example of compartmental model for the COVID-19 epidemic. The compartments are: S, susceptible; I, mildly infected; C, severely infected; H, hospitalized; U, admitted to intensive care unit; R, recovered; and D, dead.

probability $c(t) \in [0,1]$. With probability $1 - c(t)$, the infection is removed. A formal description of this model is provided in Section 2.

1.2 Assumptions and main result

A standard way to study compartmental models is to consider the dynamics of the number of individuals in each compartment. If the underlying probabilistic model is Markovian, this typically gives rise to systems of ODEs of the SIR type in the large population size limit, see [6] for a recent account. Here, we will not keep track of the count of individuals in the various compartments, but we will rather be interested in the age structure of the population. Our main result is a law of large number for the age structure of population, which is the equivalent of Theorem 7 of [14] for our non-linear extension of the model.

We anticipate the notation of Section 2 and denote the empirical measure of ages and compartments in the population at time t as

$$\forall i \in \mathcal{S}, \quad \mu_t^N(\mathrm{d}a, \{i\}) = \sum_{\sigma_x^N \leq t} \mathbb{1}_{\{X_x(t-\sigma_x^N)=i\}} \delta_{t-\sigma_x^N}(\mathrm{d}a),$$

where σ_x^N is the birth time of individual x, and the sum runs over all infected individuals at time t. (Note that $t - \sigma_x^N$ is just the age of x at time t.) The measure μ_t^N encodes the ages and compartments of infected individuals at time t. The limiting distribution of μ_t^N will depend on the following two quantities:

- The intensity measure of the infection point process defined as

  $$\tau(\mathrm{d}a) := \mathbb{E}\Big[\mathcal{P}(\mathrm{d}a)\Big].$$

  We assume that τ has a density w.r.t. the Lebesgue measure that we denote by $\tau(a)$ with a slight abuse of notation, and that $R_0 := \tau([0,\infty)) < \infty$.

- The one-dimensional marginals of the life-cycle process, denoted by

  $$\forall i \in \mathcal{S}, \forall a \geq 0, \quad p(a,i) := \mathbb{P}\Big(X(a) = i\Big).$$

Let us also denote by

$$\forall t \geq 0,\, \forall i \in \mathcal{S}, \quad Y_t^N(i) = \#\{\text{individuals in } i \text{ at time } t\}.$$

Let $I_0 \in (0,1)$. At time $t = 0$, we assume that every individual is independently infected with probability I_0 and that its age of infection is chosen independently according to a probability density g on $\mathbb{R}_+$. In the following, we define $\mathcal{I}_0^N \subseteq [N]$ as the set of infected individuals at $t = 0$.

We make the simplifying assumption that the underlying compartmental model is acyclic: we assume that for any two compartments $i, j \in \mathcal{S}$, if j can be accessed from i with positive probability, that is, if the event that we can find $s \leq t$ such that $X(s) = i$ and $X(t) = j$ has positive probability, then i cannot be accessed from j. In other words, the directed graph on $\mathcal{S}$ composed of all edges $i \to j$ such that j is accessible from i is a directed acyclic graph. This assumption is not very restrictive, most natural compartmental models enjoy this acyclic property. See Figure 1.

We can now state our main convergence result.

Theorem 1. *Let $t > 0$. Then, as $N \to \infty$, the following convergence holds in probability for the weak topology*

$$\frac{1}{N}\mu_t^N(\mathrm{d}a, \{i\}) \longrightarrow n(t,a)p(a,i)\,\mathrm{d}a$$

where $(n(t,a);\, t, a \geq 0)$ is the solution to

$$\begin{aligned} \partial_t n(t,a) + \partial_a n(t,a) &= 0 \\ \forall t \geq 0,\ n(t,0) &= c(t)S(t)\int_0^\infty n(t,a)\,\tau(\mathrm{d}a) \\ \forall a \geq 0,\ n(0,a) &= I_0 g(a) \\ \forall t \geq 0,\ S(t) &= 1 - \int_0^\infty n(t,a)\,\mathrm{d}a. \end{aligned} \tag{1}$$

Moreover, for any $i \in \mathcal{S}$, we have

$$\left(\frac{1}{N}Y_t^N(i);\, t \geq 0\right) \longrightarrow \left(\int_0^\infty n(t,a)p(a,i)\,\mathrm{d}a;\, t \geq 0\right)$$

in probability in the Skorohod topology.

Remark 2. *Theorem 1 can easily be extended to less restrictive assumptions on the initial condition. For instance, it is not hard to see from our proof that Theorem 1 holds true if we simply assume that*

$$\frac{1}{N}\sum_{x=1}^{N} \mathbb{1}_{\{\sigma_x^N < 0\}}\delta_{-\sigma_x^N}(\mathrm{d}u) \longrightarrow I_0 g(u)\mathrm{d}u$$

where the convergence holds in probability for the weak topology.

This result will follow from the more general Theorem 24, and is proved in Section 7. The definition of a solution to Equation (1) is provided in Section 3. Theorem 1 states that the age structure of the population converges to a limiting non-linear PDE of the Kermack–McKendrick type [23]. Moreover, it also

entails that the number of individuals in each compartment can be recovered by integrating the one-dimensional marginals $p(a, i)$ against the age structure.

There are two consequences of our result that we would like to emphasize. First, it shows that the macroscopic dynamics of the infected population is given by a universal equation, the Kermack–McKendrick PDE, which does not depend on the distribution of the life-cycle process. In order to recover the number of individuals in each compartment, one needs to decorate this PDE with a life-cycle process. The expression that links the age structure to the individual counts in each compartment is elementary.

Second, our approach allows to identify the characteristics of the microscopic model that impact the large population size dynamics. Recall that X and $\mathcal{P}$ are *a priori* correlated in a complex fashion: in time, because X is not a Markov process, and $\mathcal{P}$ is not a homogeneous Poisson process; and between them, as X and $\mathcal{P}$ are not independent. However, in the limit, the only two parameters that impact the dynamics of the epidemic are the intensity measure τ and the one-dimensional marginals p. These parameters are "first order quantities" of X and $\mathcal{P}$ in the sense that they only involve the distribution of the respective processes at one point in time, and are not influenced by the aforementioned correlations. Moreover, τ is the intensity measure of the infection point process, "averaged" over all life-cycles. Therefore, there is no need to assess which compartments are the most infectious in order to compute τ. Finally, let us argue that τ and p are two quantities that can be accessed in real epidemics. Write the intensity τ as

$$\tau = R_0 \nu,$$

where

$$R_0 = \int_0^\infty \tau(\mathrm{d}a), \quad \nu(\mathrm{d}a) = \frac{\tau(\mathrm{d}a)}{R_0}.$$

These two quantities have clear epidemiological interpretation:

- R_0 is the basic reproduction number, that is, the mean number of secondary infections induced by a single individual in an entirely susceptible population;
- ν is the distribution of the generation time, that is, the time between the infection of the source individual and that of the recipient individual in a typical infection pair [15].

The generation time distribution can be inferred from the time interval between the symptom onset of two individuals in an identified infectious contact, as in [11, 16] — see also the next section. The basic reproduction number R_0 is typically a quantity that needs to be estimated from the course of the epidemic, and plays an important role in assessing the efficiency of and planning control measures. The one-dimensional marginals can be inferred using a compartmental model as in [14].

1.3 Contact tracing: the historical process

We already argued that our approach allows to identify the individual characteristics that impact the large population size dynamics. We identified those parameters as R_0, the distribution of generation time ν together with the one-dimensional marginals of the life-cycle process. The estimation of those parameters is obviously of paramount importance. One possible approach to estimate the generation time

distribution consists in observing the generation times backwards in time using contact tracing, i.e., the time between the infection time of an individual (the infectee) and that of his/her infector (rather than the infection time of the individuals he/she infects). In [7], the authors addressed this specific question in a simplified setting. More specifically, they assumed that $c \equiv 1$ and that the susceptibles are in excess so that our microscopic model can be approximated by a Crump–Mode–Jagers process as in our earlier work [14]. They showed that the observation of backward generation times raises two serious issues:

(i) First, observations of past infections induce a strong observational bias: the backward generation time distribution differs from the actual generation time distribution. In the supercritical case (i.e., when $R_0 > 1$), the backward generation time has density

$$\exp(-\alpha u) R_0 \nu(u) \tag{2}$$

where $\alpha > 0$ is the Malthusian parameter of the model. As a consequence, observations of backward infection times tend to be biased towards lower values.

(ii) Infection times are difficult to observe. Instead, the onset of symptoms is generally observed. For this reason, the serial interval, which is defined as the time between symptom onsets in the two individuals mentioned above, is often used as a surrogate for the generation time. As discussed in [7], this can induce a second source of significant observational bias.

As already mentioned above, the authors in [7] address the previous bias in the case where $c \equiv 1$ and when susceptibles are in excess. In the present article, we will show if we (1) take into account saturation effect (i.e, when the population is out of the branching process regime), and (2) assume some heterogeneity in the contact rate, then those two components of the dynamics can induce a third source of bias.

In order to provide some intuition of the upcoming results, consider a newly infected individual at time t. Trace backward in time the chain of infection up to time $t = 0$. (The first individual along the chain is the infector of the focal individual, the second is the infector of the primary infector and so on.) Finally, along the chain, report the successive times of infection, see Figure 2. When susceptibles are in excess (branching approximation) Jagers and Nerman [24] showed under mild assumptions that as $t \to \infty$, the successive time of infections are well approximated by the values of a renewal process

$$\mathcal{R}^{(t)}(0) = t, \quad \mathcal{R}^{(t)}(k) = t - \sum_{i=1}^{k} \xi_i \text{ for } k \geq 1,$$

where the ξ_i's are i.i.d. and distributed according to (2).

In the presence of saturation, we show that the chain of infection is given by an h-transform of the renewal process $\mathcal{R}^{(t)}$. Intuitively, the h-transform tends to favor infection at times where there is a large fraction of susceptibles and a high contact rate. When the initial age structure of the population coincides with the "equilibrium" measure of the branching approximation, i.e.,

$$g(u) = \alpha \exp(-\alpha u),$$

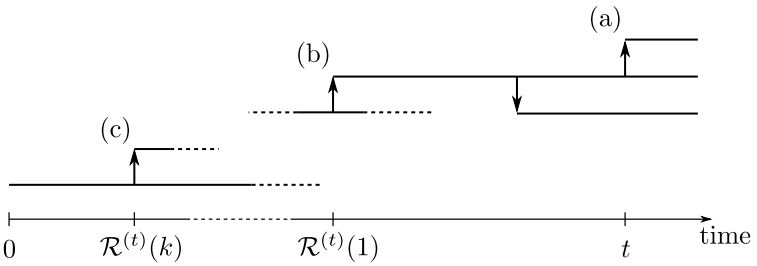

Figure 2: Chain of infection of a focal individual. Horizontal lines represent the lifetime of an infected individual, and vertical arrows represent new infections. Going backwards in time, the following events are recorded: (a) the focal individual i_0 is infected at time t; (b) its infector, individual i_1, is infected at time $\mathcal{R}^{(t)}(1)$; (c) going back a chain of k successive infections, individual i_k is infected at time $\mathcal{R}^{(t)}(k)$ by an individual that was already infected at time 0.

the h-transformed process can be reformulated in a simple manner. In Proposition 20 we show that it is identical in law to the original renewal process conditioned on survival assuming that at each step k the process is killed with probability $1 - c(\mathcal{R}^{(t)}(k))S(\mathcal{R}^{(t)}(k))$.

In Section 7, we introduce the historical process. Loosely speaking, the historical process is the empirical measure reporting the chain of infections for every individual infected at time $t = 0$ or who eventually gets infected in the future. It is constructed by reporting the successive age of infections along the chain, but also the stages of the life-cycle process for every "ancestor" along the chain, e.g. onset of the symptoms, latency period, etc. In Theorem 24, we show that the historical process converges to a deterministic probability law. Loosely speaking, our result shows that the chain of infections for a finite sample of infected individuals are asymptotically independent. Further, for each sampled individual, its chain of infection is distributed in such a way that

- the successive times of infection are expressed in terms of the h-transformed renewal process mentioned above.
- the life cycle of individuals along the chain is biased, and the bias can be expressed as a Palm modification of the original life cycle. This will be made more formal in Section 5.

Going back to the two epidemiological questions (i) and (ii), our results decipher how the epidemiological parameters R_0 and the generation time distribution ν relate (in a non-trivial way) to observables which can be directly collected from contact-tracing data.

1.4 A genealogical dual to the delay equation

The Kermack–McKendrick equation (1) can be reformulated in terms of a non-linear delay equation. To ease the exposition, let us consider the case $c \equiv 1$. In Section 5, the general case $c \not\equiv 1$ will be exposed.

If $(n(t, a); t, a \geq 0)$ denotes the solution to Equation (1) with $c \equiv 1$, let us define the number of infections between time 0 and t as

$$B(t) = \int_0^t n(s, 0)\, \mathrm{d}s = \int_0^t S(s) \int_0^\infty n(s, a)\, \tau(\mathrm{d}a)\, \mathrm{d}s.$$

Then we will derive in Section 3 that B solves the following non-linear delay equation:

$$B(t) = S_0\bigg(1 - \exp\bigg(-\int_0^t \tau(a)B(t-a)\,\mathrm{d}a - \int_0^\infty\int_0^t \tau(a+s)g(a)\,\mathrm{d}s\,\mathrm{d}a\bigg)\bigg), \quad (3)$$

where $S_0 = 1 - I_0$ is the initial number of susceptibles.

Our proof of Theorem 1 uses a genealogical approach, where we look backwards in time at the set of potential infectors of a focal individual. This approach leads to a genealogical dual to the delay equation that we think to be of independent interest. The dual is built out of the following branching process.

Recall that R_0 stands for the total mass of τ and $\nu = \tau/R_0$. We define the intensity

$$\bar\tau(u) = \int_0^\infty g(a)\tau(a+u)\,\mathrm{d}a, \qquad u \geq 0,$$

so that the measure $\bar\tau(u)\,\mathrm{d}u$ is the intensity measure of the infection point process of an individual with initial age distributed as g. Let us further set

$$\bar R_0 = \int_0^\infty \bar\tau(u)\,\mathrm{d}u, \quad \bar\nu(\mathrm{d}u) = \frac{\bar\tau(u)\,\mathrm{d}u}{\bar R_0}.$$

The branching process is constructed as follows. Let us assume that individuals in the branching process are either infected (I) or susceptibles (S). Suppose that the population starts from a single (S) individual. Then, at each generation, an (S) individual produces:

- a Poisson$\big(S_0R_0\big)$ distributed number of (S) individuals;
- a Poisson$\big((1-S_0)\bar R_0\big)$ distributed number of (I) individuals.

Individuals of type (I) have no offspring. Draw an oriented edge from each individual towards its parent. Assign a weight independently to each edge, such that the weight of an edge originating from an (S) individual is distributed as ν, and that of an edge coming from a (I) individual is distributed as $\bar\nu$.

The previous branching process corresponds to the large population size limit of the set of potential infectors of a fixed individual. Type (I) individuals correspond to individuals that were initially infected. Each edge corresponds to an infectious contact in the population, and the weight of that edge is the age of the infector when this contact occurs.

The corresponding object is a rooted *geometric* tree, where edges are endowed with a weight (or *length*). We define the infection path at the root as the (a.s. unique) geodesic connecting the set of infectious individuals (I) to the root. Finally, the time of infection σ^∞ is defined as the length of the geodesic. The following result connects the distribution of σ^∞ to the delay equation.

Proposition 3. *For any $t \geq 0$, define*

$$B(t) = S_0\mathbb{P}(\sigma^\infty \leq t).$$

Then $(B(t);\, t \geq 0)$ solves the delay equation (3).

In Section 5, we will derive a similar dual for the delay equation with $c \not\equiv 1$ — see Proposition 15.

1.5 Link with literature

The idea of considering an infection through its age structure dates back to at least the work of Kermack and McKendrick [23], who introduced the SIR model as a special case of a more general age-structured model. More generally, delay equations, which are an equivalent formulation of the Kermack–McKendrick PDE, have been widely proposed as models for the spread of epidemics, see for instance [4, 5, 8, 33]. Surprisingly, the convergence of probabilistic epidemic models towards the solution of a delay equation or a Kermack–McKendrick PDE has received only little attention. Let us briefly review the works in this direction that we are aware of.

In [2], the authors studied an epidemic model extremely similar to the one under consideration here. The only difference with our model is that they do not explicitly model the compartments, and do not allow for a contact rate function $(c(t); t \geq 0)$. They obtain a slightly different kind of law of large number. Instead of starting from a macroscopic fraction of infected individuals, they look at the epidemic started from one infected individual. They show that, after an appropriate time-shift so as to skip the long initial branching phase when there are few individuals, the number of susceptibles converges to the solution of an extension of the delay equation (3) to the whole real line, see their Theorem 2.10. It is remarkable that this delay equation is unique and does not require the knowledge of an initial age profile. It also corresponds to the limit of the Kermack–McKendrick model (1) when the initial fraction of infected individuals converges to zero [9]. (The existence and uniqueness of similar limits have been derived quite generically, and were coined "renewal theorems in epidemiology" [31, 32].) From an application point of view it is important to note that the age profile of the population cannot be directly assessed. The results in [2] prove that the limiting age profile is shaped by natural population growth, providing a natural candidate for this age profile when the epidemic is started from a few individuals. However, let us highlight some important differences with our work. First, [2] work under quite restrictive assumptions on the intensity measure of the infection point process τ, see Assumption 2 on the top of page 7. For instance, the case of the Markovian SIR model is not covered by these assumptions. Here, our only assumption is that τ has a finite mean, which is the minimal assumption to derive a law of large numbers. Second, we have accounted explicitly for time heterogeneity through the contact rate $(c(t); t \geq 0)$, and compartments through the life-cycle process $(X(a); a \geq 0)$, which are important features in view of applications. Third, our approach relies on tracing backward-in-time the chain of transmission leading to a focal infection. As far as we know it is new, it provides the limit of the historical process of the population, and it could be adapted to a more general setting where the contact rate can depend on the state of the whole infected population, in a similar way to [17].

In a recent series of work, [25] derived a functional law of large numbers and a functional central limit theorem for classical SIR-like models with general sojourn time distribution in each compartment. They also obtained similar results for extensions of these models incorporating spatial heterogeneity [26] and varying or random infectiosity [12, 27, 28], and applied these models to the COVID-19 epidemic in France [13]. The limiting equations that describe the dynamics of the density of individuals in each compartments are systems of so-called Volterra

integral equations. These equations are tightly linked to our PDE representation using the Kermack–McKendrick equation (1), as is acknowledged explicitly in [28], Proposition 2.1. All the models with non-vanishing immunity that they consider (SIR, SEIR) can be formulated within our framework. The infection point process $\mathcal{P}$ is obtained by starting either a homogeneous Poisson point process [25], an inhomogeneous Poisson point process [12, 27], or an inhomogeneous Poisson point process with random intensity [28], at the infection of an individual or after an exposed phase, depending on the model under consideration. Moreover, the proof techniques they use rely heavily on a representation of their model as the solution to a stochastic equation driven by a Poisson measure, which does not hold in our more general setting. Nevertheless, let us acknowledge that their techniques allow to derive central limit theorems for their models.

Finally, there exists a rich literature on general age-dependent population processes, not necessarily related to epidemic models. Let us first mention the Crump–Mode–Jagers (CMJ) processes, where the birth times of the children are allowed to depend in a very general way on the age of the parent, see [19] or [30] for a more recent account. The only restriction is that individuals should reproduce independently from each other. By assuming that the age structure of the population is a Markov process, it is possible to release this assumption and consider a much wider class of age-dependent models. Using the framework introduced in [20, 21], [17] and [10] derive respectively a law of large numbers and a central limit theorem for the age structure of a very general class of population models. The deterministic limit they obtain for the age structure corresponds to a more general form of the Kermack–McKendrick type of PDE that we have derived here. Our results are not trivially implied by that of [10, 17] as we do not require any Markov assumption, but we believe that our genealogical approach is interesting for its own sake, since it provides an expression for the observed transmission tree leading to a typical infection. We also want to stress again that our work provides an interesting epidemiological framework that can incorporate many modeling details and can be readily used for applications [14].

1.6 Outline

The rest of this paper is organized as follows. A formal description of the model is provided in Section 2, and the Kermack–McKendrick PDE is studied in Section 3.

Section 4 contains the graph construction of our model, as well as a rigorous definition of the ancestral process mentioned in Section 1.3. Our proofs rely on showing the convergence of the set of potential infectors of a focal infected individual to a limiting Poisson tree. Section 5 describes this limiting tree, and provides a characterization of the transmission chain leading to the infection of individual in terms of the h-transform of a renewal process. Finally, the convergence to the Poisson tree is carried out in Section 6 and our law of large numbers is proved in Section 7.

2 Description of the model

In the following, we will consider an epidemic model in which individuals' life trajectories are represented by independent stochastic processes. We distinguish between two types of individuals:

- Susceptible individuals that have never been infected before.
- Infected individuals that have been infected in the past. We emphasize that the meaning of infected is a bit broader than usual. For instance, a recovered or dead individual is considered as infected. To each infected individual, we associate an age. The age is the time elapsed since the beginning of the infection.

There are N individuals in the population. To each individual $x \in [N]$, we associate a pair of processes $(\mathcal{P}_x, X_x)$ describing respectively the process of secondary infections and the successive stages of the disease experienced by the focal individual x. More precisely:

- The *life-cycle process*, denoted by $(X_x(a);\, a \geq 0)$, is a random process valued in $\mathcal{S}$ where $X_x(a)$ is the stage of the disease (e.g., exposed, death, etc.) of x at age a.
- The *infection point process* $\mathcal{P}_x$ is a point measure describing the ages of potential infections.

Let us denote by $\mathcal{X}_x := (\mathcal{P}_x, X_x)$. We will always assume that $(\mathcal{X}_x;\, x \in [N])$ are i.i.d. and denote by $\mathcal{X} = (\mathcal{P}, X)$ their common distribution. The state space of $\mathcal{X}$ is denoted by $\mathscr{X}$.

Remark 4. *Note that we allow for non-trivial correlation between the life-cycle and the infection process. Examples of such correlations can be that a deceased individual is not infectious anymore, a hospitalized individual may have a reduced potential of infection due to quarantine, etc.*

We suppose that at $t = 0$, every individual is independently infected with probability I_0. Let $\mathcal{I}_0^N$ be the set of initially infected individuals. For each $x \in \mathcal{I}_0^N$ we need to prescribe an age, or equivalently, an infection time. We assume that, conditional on $\mathcal{I}_0^N$, the ages of the initial individuals $(Z_x;\, x \in \mathcal{I}_0^N)$ are i.i.d. with common distribution g. Let us denote by $(\sigma_x^N;\, x \in \mathcal{I}_0^N)$ the birth time of the initial infected, that is, $\sigma_x^N = -Z_x$.

The epidemic now spreads as follows. Suppose that, at some time t_0, we have defined a set $\mathcal{I}_{t_0}^N \subseteq [N]$ of infected individuals at time t_0, and a vector $(\sigma_x^N;\, x \in \mathcal{I}_{t_0}^N)$ of infection times. Let t_1 be the first atom after t_0 of the point measure

$$\sum_{x \in \mathcal{I}_{t_0}^N} \sum_{a \in \mathcal{P}_x} \delta(\sigma_x^N + a).$$

If there is no such atom, the infection stops. Otherwise, let U be uniformly chosen in $[N]$, independent of the rest, it is the first individual that comes in contact with any of the infected individuals after time t_0. If $U \in \mathcal{I}_{t_0}^N$, then nothing happens, and we carry out the same procedure for the next atom t_2. If $U \notin \mathcal{I}_{t_0}^N$, then, with probability $1 - c(t_1)$, the infection is ineffective in which case nothing happens and we consider the next infection time t_2. Otherwise, set $\mathcal{I}_{t_1}^N = \mathcal{I}_{t_0}^N \cup \{U\}$ and $\sigma_U^N = t_1$, and continue the procedure as if starting from time t_1 with the initial infected set $\mathcal{I}_{t_1}^N$. This inductive procedure will be reformulated in terms of an infection graph in Section 4.1.

3 Kermack–McKendrick PDE

In this section we provide our definition of the solution to the Kermack–McKendrick equation (1). We start with a formal resolution of the PDE using the method of characteristics.

Let I_0 be the initial density of infected individuals and g the initial age profile of the population. First, note that if n is solution of the PDE, then for every pair (t, a) of non-negative numbers, $s \mapsto n(t-s, a-s)$ is constant on $(0, t \wedge a)$. This yields

$$\forall t, a \geq 0, \quad n(t,a) = \begin{cases} I_0 g(a-t) & \text{when } a > t \\ b(t-a) & \text{when } a \leq t, \end{cases} \tag{4}$$

with

$$\forall t \geq 0, \quad b(t) \coloneqq n(t,0)$$

is the number of new infections at time t. Moreover,

$$\begin{aligned} \dot{S}(t) &= -\int_0^\infty \partial_t n(t,a)\,\mathrm{d}a = \int_0^\infty \partial_a n(t,a)\,\mathrm{d}a \\ &= -b(t) = -c(t)S(t)\int_0^\infty \tau(a)n(t,a)\,\mathrm{d}a. \end{aligned}$$

As a result, we have

$$\begin{aligned} S(t) &= S(0)\exp\Big(-\int_0^t c(s)\int_0^\infty \tau(a)n(s,a)\,\mathrm{d}a\,\mathrm{d}s\Big) \\ &= S(0)\exp\Big(-\int_0^t c(s)\Big(\int_0^s \tau(a)b(s-a)\,\mathrm{d}a + I_0\int_s^\infty \tau(a)g(a-s)\,\mathrm{d}a\Big)\,\mathrm{d}s\Big) \\ &= S(0)\exp\Big(-\int_0^t c(s)\Big(\int_0^s \tau(s-a)b(a)\,\mathrm{d}a + I_0\bar{\tau}(s)\Big)\,\mathrm{d}s,\Big) \end{aligned}$$

with $\bar{\tau}(s) = \int_0^\infty \tau(a+s)g(a)\,\mathrm{d}a$, so necessarily

$$B(t) \coloneqq \int_0^t b(s)\,\mathrm{d}s = S_0 - S(t)$$

solves the nonlinear delay equation

$$B(t) = S_0\Big[1 - \exp\Big(-\int_0^t c(s)\Big(\int_0^s \tau(s-a)\,B(\mathrm{d}a) + I_0\bar{\tau}(s)\Big)\,\mathrm{d}s\Big)\Big] \tag{5}$$

where $B(\mathrm{d}a) = b(a)\,\mathrm{d}a$ is the Stieltjes measure associated to the nondecreasing map B. This motivates the following definition of a solution to the Kermack–McKendrick equation.

Definition 5. *We say that $(n(t,a);\, t,a \geq 0)$ is a weak solution to (1) if there exists a nonnegative function $(b(t);\, t \geq 0)$ such that:*

1. *the functions n and b are related through (4);*
2. *the function $B(t) \coloneqq \int_0^t b(s)\,\mathrm{d}s$ solves the delay equation (5).*

If a nondecreasing function B satisfies (5), then we have the following inequality:

$$B(t+u) - B(t) \leq S(0) \int_t^{t+u} c(s)\Big(\int_0^s \tau(s-a) B(\mathrm{d}a) + I_0 \int_0^\infty \tau(a+s) g(a)\,\mathrm{d}a\Big)\,\mathrm{d}s.$$

The previous inequality readily entails that B is absolutely continuous, and thus that we can find b such that $B(t) = \int_0^t b(s)\,\mathrm{d}s$. Therefore, existence and uniqueness of solutions to (1) reduce to existence and uniqueness of nondecreasing solutions to (5), which is provided by the following result.

Lemma 6. *There is a unique nondecreasing, nonnegative solution to* (5).

Proof. Let us denote by E the set of all nondecreasing, nonnegative, càdlàg functions on $[0,\infty)$. For $\gamma > \alpha \vee 0$, define

$$E_\gamma = \{f \in E : \int_0^\infty e^{-\gamma t} f(t)\,\mathrm{d}t < \infty\}.$$

We endow E_γ with the metric

$$d_\gamma(f,g) = \int_0^\infty e^{-\gamma t} |f(t) - g(t)|\,\mathrm{d}t$$

which makes (E_γ, d_γ) a complete metric space. As any solution to (5) is bounded and continuous, it is sufficient to show existence and uniqueness of the solution in E_γ.

We introduce the operator $\Phi\colon E_\gamma \to E_\gamma$ such that

$$\Phi f(t) = S_0\Big(1 - \exp\Big(- \int_0^t c(s)\Big(\int_0^s \tau(s-a) f(\mathrm{d}a)\Big)\,\mathrm{d}s \\ - I_0 \int_0^\infty \Big(\int_0^t c(s)\tau(a+s)\,\mathrm{d}s\Big) g(a)\,\mathrm{d}a\Big)\Big),$$

where $f(\mathrm{d}a)$ denotes the Stieltjes measure associated to f. Note that $\Phi f \in E_\gamma$, since it is clear that Φf is bounded, continuous, nonnegative and nondecreasing. Let us show that Φ is a contraction. We have, for $f_1, f_2 \in E_\gamma$,

$$\begin{aligned} d_\gamma(\Phi f_1, \Phi f_2) &\leq S_0 \int_0^\infty e^{-\gamma t} \Big| \int_0^t c(s)\Big(\int_0^s \tau(s-a) f_1(\mathrm{d}a) - \int_0^s \tau(s-a) f_2(\mathrm{d}a)\Big)\,\mathrm{d}s\Big|\,\mathrm{d}t \\ &\leq \int_0^\infty e^{-\gamma t}\Big(\int_0^t \tau(s) |f_1(t-s) - f_2(t-s)|\,\mathrm{d}s\Big)\,\mathrm{d}t \\ &= d_\gamma(f_1, f_2) \int_0^\infty e^{-\gamma t}\tau(t)\,\mathrm{d}t. \end{aligned}$$

As $\gamma > \alpha$, we know that $\int_0^\infty e^{-\gamma t}\tau(t)\,\mathrm{d}t < 1$, showing that Φ is a contraction. The Banach fixed point theorem therefore shows that there exists a unique $B \in E_\gamma$ such that $\Phi B = B$, ending the proof. □

4 Graph of infection

4.1 Infection graph

Recall the infection model defined in Section 2, and the notation $(\mathcal{P}_x;\, x \in [N])$ for the infection point processes, $\mathcal{I}_0^N$ for the set of initially infected individuals, and $(\sigma_x^N = -Z_x;\, x \in \mathcal{I}_0^N)$ for their birth (or infection) time.

Let $(Z_x;\, x \in [N])$ be i.i.d. random variables with law

$$\delta_0(\mathrm{d}u)(1 - I_0) \;+\; I_0 g(u)\mathrm{d}u$$

Intuitively, Z_x encodes the age of infection of individual x at time $t = 0$. Susceptible individuals have age 0, whereas the age of an infected individual is chosen according to the density g. Define the shifted infection measure

$$\widehat{\mathcal{P}}_x \;= \sum_{u \in \mathcal{P}_x} \mathbb{1}_{\{u - Z_x \geq 0\}} \delta(u - Z_x)$$

Note that if x is susceptible (i.e., $Z_x = 0$), we have $\widehat{\mathcal{P}}_x = \mathcal{P}_x$. Vertices with $Z_x = 0$ will be said of type susceptibles (S). Vertices with $Z_x > 0$ will be said of type infected (I).

Recall that each atom of a point process $\widehat{\mathcal{P}}_x$ encodes a potential infectious contact, which is targeted to a uniformly chosen individual in the population. We enrich the infection point processes by adding the information about the label of the target individual. Formally, we consider a sequence of i.i.d. random variables $(U_{x,i};\, x \in [N], i \in \mathbb{N})$ uniformly distributed on $[N]$. Define

$$\forall x \in [N], \quad \widehat{\mathcal{N}}_x \;\coloneqq \sum_{u_i \in \widehat{\mathcal{P}}_x} \delta(u_i, U_{x,i})$$

where the sums $u_1 < u_2 < \cdots$ in the sum are the atoms of $\widehat{\mathcal{P}}_x$ listed in increasing order. We now build a graph out of the family $(\widehat{\mathcal{N}}_x;\, x \in [N])$ that records the potential infections in the population. For every $j \notin \mathcal{I}_0^N$, define the set of its potential infectors

$$A_j^N \coloneqq \{i \in [N] : \exists (a, j) \text{ s.t. } (a, j) \in \widehat{\mathcal{N}}_i\}.$$

(Note that this set is possibly empty in which case j is never infected.) We give the following definition of the infection graph.

Definition 7. *The* infection graph *built from the i.i.d. collection of triplets* $(\widehat{\mathcal{N}}_x, X_x, Z_x;\, x \in [N])$ *is the random oriented geometric marked graph* $\mathcal{G}^N = (V^N, E^N)$ *with* $V^N = [N]$ *and*

$$E^N = \bigcup_{i \in [N]} \bigsqcup_{(a,j) \in \widehat{\mathcal{N}}_i} \{(i, j)\},$$

where the second union is a disjoint union, meaning that for each pair (i, j) *we allow for multiple edges from* i *to* j *in the infection graph. The marks and weights are defined as follows.*

1. *Each edge* $e = (i_e, j_e)$ *corresponds to an atom* (a_e, j_e) *of the point process* $\widehat{\mathcal{N}}_{i_e}$*. The number* a_e *will be referred to as the weight of edge* e*.*

2. *For each vertex* $x \in V^N$*, we define the mark at* x *as*

$$m_x \coloneqq (Z_x, \mathcal{X}_x).$$

Remark 8. *As stated in the theorem,* $\mathcal{G}^N$ *is an oriented geometric marked graph. By geometric, we mean that edges are weighted. The orientation is dictated by the direction of potential infections, and the meaning of an edge* (i, j) *is that individual* j *is potentially infected by* i*. Finally, the first coordinate of the marking allows to distinguish between infected and susceptible individuals at time* $t = 0$*.*

A path in $\mathcal{G}^N$ is a set of edges $\pi = (e_1, \dots, e_n)$ such that, $j_{e_k} = i_{e_{k+1}}$, with the notation (i_e, j_e) for the origin and target vertices of the edge e. The length of a path $|\pi|$ is defined as

$$|\pi| = \sum_{e \in \pi} a_e.$$

The genealogical (or topological) distance is defined as the number of edges composing the path (n in our specific example). For $i \in \mathbb{N}$, we define the i-truncation of π as

$$\forall i \in \mathbb{N}, \quad \tau_i \pi \ := \ (e_1, \dots, e_{i \wedge n}).$$

We say that π is a path from i to j if $i_{e_1} = i$ and $j_{e_n} = j$. A path in $\mathcal{G}^N$ from i to j corresponds to a potential infection chain between i and j. The length of the path is the time elapsed between the infection of i and j. The genealogical length is the number of infectors along the chain.

4.2 Infection process

Conditional on a realization of the infection graph (V^N, E^N), we attach an additional independent random variable s_e uniform on $[0, 1]$ to every edge $e \in E^N$ of the graph. This random variable will encode what we will call the contact intensity of edge e. Roughly speaking, if the contact occurs at time t, this contact can only translates into an infection iff two conditions are satisfied. First, the contact intensity should be strong enough in the sense that $s_e \leq c(t)$ (see (6) below). Secondly, the target individual should not have been infected before (see (7) below). We make this more precise in the next definition; see also Figure 3.

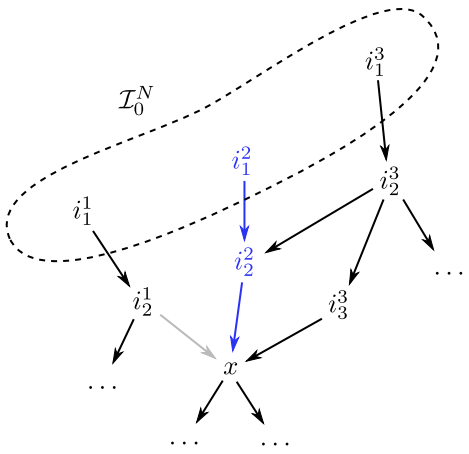

Figure 3: Ancestors of individual x in $\mathcal{G}^N$. In this picture, we have four potential infection paths from $\mathcal{I}_0^N$ to x: $\pi^1 = (e_1^1, e_2^1)$, $\pi^2 = (e_1^2, e_2^2)$, $\pi^3 = (e_1^3, e_2^3, e_3^3)$ and $\pi^4 = (e_1^3, e_*, e_2^2)$, where we write $e_k^\ell = (i_k^\ell, i_{k+1}^\ell)$ and $e_* = (i_2^3, i_2^2)$. Assume first that π^1 is the shortest path, but that $s_{e_2^1} > c(|\pi^1|)$ — the edge is grayed out in the figure. Then π^1 is not an active path. Now let us assume that $|\tau_1 \pi^2| = a_{e_1^2} < a_{e_1^3} + a_{e_*} = |\tau_2 \pi^4|$. This means that π^4 cannot be an active path. Finally, if π^2 and π^3 are the two active paths and $|\pi^2| < |\pi^3|$, then π^2 (in blue) is the active geodesic from $\mathcal{I}_0^N$ and $\sigma_x^N = |\pi^2|$.

Definition 9. *Let $\pi = (e_1, \cdots, e_n)$ be a path with $i_{e_1} \in \mathcal{I}_0^N$. The path is said to be active iff*

$$\forall k \in [n], \quad s_{e_k} \leq c\Big(|\tau_k \pi|\Big). \tag{6}$$

For every $x \notin \mathcal{I}_0^N$, let $\Xi^N(x)$ be the set of active paths from $\mathcal{I}_0^N$ to x. The path is said to be the active geodesic from $\mathcal{I}_0^N$ to x iff

$$\forall k \in [n], \quad \tau_k \pi \;=\; \underset{\pi' \in \Xi^N(j_k)}{\arg\min} |\pi'|. \tag{7}$$

Finally, we define the infection time of x — denoted by σ_x^N — as the length of the active geodesic from $\mathcal{I}_0^N$ to x, with the convention that $\sigma_x^N = \infty$ if the geodesic does not exist.

Remark 10.

1. *Since τ has a density, there is at most one path satisfying the minimization problem (7).*
2. *If $c \equiv 1$, then any path in the infection graph is active, so that our definition coincides with the usual definition of a geodesic on a weighted graph. In particular, (7) just states that if $\pi = (e_1, \ldots, e_n)$ is the geodesic from $\mathcal{I}_0^N$ to x, then the truncated path $\tau_k \pi$ is the geodesic from $\mathcal{I}_0^N$ to j_{e_k}. Thus, when $c \equiv 1$, all the information about the infection process is contained in the infection graph and the extra variables s_e do not play any role.*

4.3 The ancestral path

Definition 11 (Infection and ancestral paths)**.**

- *Let us consider x of type (S) such that $\sigma_x^N < \infty$ and write $\pi = (e_1, \cdots, e_n)$ with $e_k = (i_k, j_k)$ for the active geodesic from $\mathcal{I}_0^N$ to x. We define the infection path $\mathcal{R}_x^N$ from x to $\mathcal{I}_0^N$ as*

$$\mathcal{R}_x^N(0) = \sigma_x^N, \quad \forall \ell \in [n], \quad \mathcal{R}_x^N(\ell) \;=\; \sigma_x^N - \sum_{k=0}^{\ell-1} (a_{e_{n-k}} + Z_{i_{n-k}}). \tag{8}$$

  Finally, we define the ancestral process as

$$\mathcal{A}_x^N \;\coloneqq\; \Big(\mathcal{R}_x^N(\ell), \mathcal{X}_{v_\ell}\Big)_{\ell=0}^{n}, \quad \text{where } v_k = \begin{cases} i_{n+1-\ell} & \text{if } \ell \neq 0 \\ x = j_n & \text{if } \ell = 0 \end{cases}$$

  to be the vector recording the information along the chain of infection (age of infection, infection measure, life-cycle).

- *If x is of type (S) but $\sigma_x^N = \infty$, then $\mathcal{A}_x^N$ is defined as the empty sequence.*
- *If x is of type (I), then $\mathcal{A}_x^N \coloneqq (\sigma_x^N, \mathcal{X}_x)$.*

In words, the random path $\mathcal{R}_x^N$ is obtained by tracing backward in time the chain of infection from the focal individual x to the set of infected individuals. The increments of the path are given by the successive age of infection. $\mathcal{R}_x^N(\ell)$ is the time of infection of the ℓ-th ancestor along the chain; the variable $\mathcal{X}_{v_\ell}$ encodes its infection and life-cycle processes. (Assuming that individuals are ranked from the focal individual x to the initially (I) individual.) Note that in the sum (8), all the $Z_{i_{n-k}}$ terms are equal to 0, except when $k = n-1$. In words, the oldest ancestor

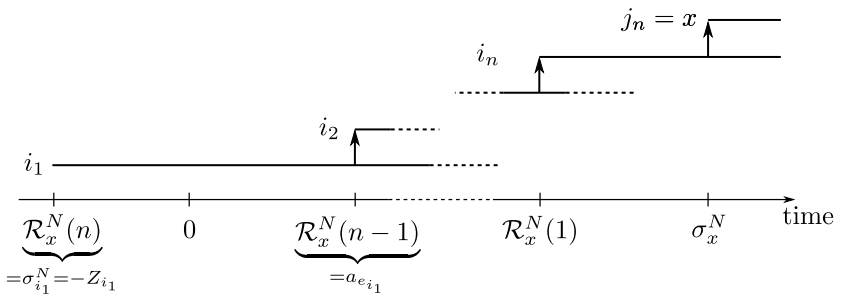

Figure 4: Infection path of individual x.

is the only type (I) individual along the chain of infection. For $k = n-1$, the corresponding increment is decomposed into two parts: (1) the undershoot a_{e_1} and (2) the overshoot Z_{i_1} corresponding to the age of infection of the (I) individual at time $t = 0$. See Figure 4.

We close this section by a brief description of the topology underlying the set of ancestral paths. Let $\mathcal{M}$ refer to the set of locally finite positive measures on $[0, \infty)$, and equip $\mathcal{M}$ with a metric $d_{\mathcal{M}}$ that induces the vague topology [22, Section 4.1]. We denote by $d_{\mathcal{S}}$ the Skorohod metric on the set of càdlàg processes valued in $\mathcal{S}$ denoted by $\mathbb{D}(\mathcal{S})$. The space $\mathbb{R}_+ \times \mathcal{M} \times \mathbb{D}(\mathcal{S})$ is equipped with the sup metric ρ defined as

$$\forall (x,y,z),(x',y',z'), \quad \rho\left((x,y,z),(x',y',z')\right) \;=\; \max\Big(|x-x'|, d_{\mathcal{M}}(y,y'), d_{\mathcal{S}}(z,z')\Big).$$

Each ancestral path is valued in the space

$$\bigcup_{n=0}^{\infty} \Big(\mathbb{R}_+ \times \mathcal{M} \times \mathbb{D}(\mathcal{S})\Big)^n$$

equipped with the metric $\mathcal{D}$ defined as follows

$$\mathcal{D}\Big((m_1,\cdots,m_n),(m'_1,\cdots,m'_{n'})\Big) = \begin{cases} 1 & \text{if } n \neq n' \\ \max(\rho(m_1,m'_1),\cdots,\rho(m_n,m'_{n'})) & \text{if } n = n'. \end{cases}$$

4.4 Local topology on graphs

For $x \in [N]$, define $\mathcal{G}^N(x)$ as the subgraph of $\mathcal{G}^N$ induced by all the vertices y with an oriented path from y to x (including x itself). We treat the graph $\mathcal{G}^N(x)$ as a pointed graph, with x as the reference vertex. Recall from Definition 7 that the edges of $\mathcal{G}^N(x)$ are endowed with weights and the vertices are equipped with marks $m_x = (Z_x, \mathcal{X}_x)$ where the coordinates are respectively (1) the age of infection at time $t = 0$, (2) the infection process, and the life cycle process of individual x. The set of marks is equipped with the sup metric ρ introduced in the previous section

$$\rho\left(m_x, m_y\right) = \max\Big(|Z_x - Z_y|, d_{\mathcal{M}}(\mathcal{P}_x,\mathcal{P}_y), d_{\mathcal{S}}(X_x,X_y)\Big).$$

We will say that $\mathcal{G}^N(x)$ is a random element of $\mathscr{H}$, the set of pointed oriented geometric marked (pogm) graphs. An element of $\mathscr{H}$ is characterized by five coordinates $(V, E, w, m, \emptyset)$, respectively the set of vertices, the set of edges,

$w = (w_e)_{e\in E}$ the weights on edges, $m = (m_x)_{x\in V}$ the set of marks, and $\emptyset$ the pointed vertex.

We equip $\mathscr{H}$ with a metric $d_{\mathscr{H}}$ so that $(\mathscr{H}, d_{\mathscr{H}})$ is a Polish space. A graph isomorphism ϕ between two *finite* pogm graphs $\mathcal{G} = (V, E, w, m, \emptyset)$ and $\mathcal{G}' = (V', E', w', m', \emptyset')$ is a bijection from V to V' such that

1. $(u, v) \in E$ iff $(\phi(u), \phi(v)) \in E'$.
2. ϕ maps the reference vertex of $\mathcal{G}$ to the reference vertex in $\mathcal{G}'$.

By convention, we set $\min(\emptyset) = \infty$ in the following. Let $\mathcal{G} = (V, E, w, m, \emptyset)$, $\mathcal{G}' = (V', E', w', m', \emptyset')$ be two elements of $\mathscr{H}$. Define

$$d(\mathcal{G}, \mathcal{G}') = \min\{1, \min_{\phi} \left(\max_{e:\, e\in E} |w_e - w'_{\phi(e)}| \vee \max_{x\in V} |\rho(m_x, m'_{\phi(x)})| \right)\}$$

where the minimum is taken over all possible graph isomorphisms between the two graphs (in the sense prescribed above, that is, we only consider the isomorphisms preserving the pointed vertex). If there is no such isomorphism between $\mathcal{G}$ and $\mathcal{G}'$, we set $d(\mathcal{G}, \mathcal{G}') = 1$.

For $\mathcal{G} \in \mathscr{H}$ and $y \in \mathcal{G}$, the topological (or *genealogical*) distance to the reference vertex x is defined as

$$\inf\{n : \text{there exists a path } (y = x_1, \ldots, x_n = x) \text{ in } \mathcal{G}\}.$$

For every $r \in \mathbb{N}^*$, we denote by $[\mathcal{G}]_r$, the subgraph induced by the vertices at a topological distance to the origin, that is, to the pointed vertex, less than r. For two elements $\mathcal{G}, \mathcal{G}' \in \mathscr{H}$, we define the (pseudo-)distance $d_{\mathscr{H}}$ as follows

$$d_{\mathscr{H}}(\mathcal{G}, \mathcal{G}') \;=\; \sum_r 2^{-r} d([\mathcal{G}]_r, [\mathcal{G}']_r).$$

The metric $d_{\mathscr{H}}$ naturally induces a notion of local convergence on (equivalence classes of) $\mathscr{H}$. Using standard arguments, we can see that $(\mathscr{H}, d_{\mathscr{H}})$ is a Polish space.

5 A limiting Poisson random tree

5.1 Palm infection measures

Recall that $\mathcal{X}_x = (\mathcal{P}_x, X_x)$ is the pair encoding the infection and the life-cycle process and $\mathcal{P}$ is a point process where each atom represents a potential infection event. Define $|\mathcal{P}| \coloneqq \int \mathrm{d}\mathcal{P}(a)$ which is interpreted as the total number of potential infections (or *contacts*) along the course of infection. We define a triplet of random variables $(W, \tilde{\mathcal{P}}, \tilde{X}_x) \equiv (W, \tilde{\mathcal{X}})$ valued in $\mathbb{R}_+ \times \mathcal{M} \times \mathcal{S} \equiv \mathbb{R}_+ \times \mathscr{X}$ such that for every bounded continuous function f

$$\begin{aligned}
\mathbb{E}\Big(f(W, \tilde{\mathcal{P}}, \tilde{X}) \Big) \;&=\; \frac{1}{R_0} \mathbb{E}\Big(\int f(a, \mathcal{P}, X) \mathrm{d}\mathcal{P}(a) \Big) \\
&=\; \frac{1}{R_0} \mathbb{E}\Big(|\mathcal{P}| \;\times \int \frac{1}{|\mathcal{P}|} f(a, \mathcal{P}, X) \mathrm{d}\mathcal{P}(a) \Big)
\end{aligned}$$

In words, we first bias the pair $\mathcal{X} = (\mathcal{P}, X)$ by $|\mathcal{P}|$. Conditional on the resulting biased pair $\tilde{\mathcal{X}} = (\tilde{\mathcal{P}}, \tilde{X})$, the r.v. W is obtained by picking an atom of the infection measure $\tilde{\mathcal{P}}$ uniformly at random.

Definition 12 (Campbell and Palm measures)**.** *The law of* $(W, \tilde{\mathcal{X}})$ *is the Campbell's measure associated to* $\mathcal{X}$ *[1]. The Palm measure at* $a \in \mathbb{R}_+^*$ *is defined as the distribution of the random pair* $\tilde{\mathcal{X}}$ *conditioned on the event* $\{W = a\}$*. See again [1] for a precise definition of this conditioning.*

Recall that τ is the intensity measure of $\mathcal{P}$, i.e.,

$$\tau(a)\,\mathrm{d}a = \mathbb{E}\big[\mathcal{P}(\mathrm{d}a)\big], \quad R_0 = \int_0^\infty \tau(a)\,\mathrm{d}a.$$

and recall that we denote $\nu = \tau/R_0$. The next result is standard from Palm theory.

Lemma 13. *W is distributed according to ν.*

5.2 Definition of the Poisson tree

Recall that we have defined $\bar{\tau}$ by

$$\bar{\tau}(u) = \int_0^\infty g(a)\tau(a+u)\,\mathrm{d}a, \qquad u \geq 0,$$

that $\bar{R}_0$ denotes the total mass $\int_0^\infty \bar{\tau}(u)\,\mathrm{d}u$, and that $\bar{\nu}$ denotes the normalization $\bar{\tau}(u)\,\mathrm{d}u/\bar{R}_0$. Let us now consider a pair of random variables $(\bar{W}, Z) \in \mathbb{R}_+^2$ with joint density

$$\forall w, z > 0, \quad G(w,z) \;=\; \frac{1}{\bar{R}_0} g(z)\tau(w+z). \tag{9}$$

In particular, the first coordinate is distributed according to $\bar{\nu}$.

We now construct a (pointed) Poisson marked random tree $\mathcal{H}$ in two consecutive steps. (This extends the construction of Section 1.4 to the case $c \not\equiv 1$.) First, the graph structure of $\mathcal{H}$ depends on the two positive real parameters S_0R_0, $I_0\bar{R}_0$, and second the random weights and the marks are assigned through two probability distribution ν, $\bar{\nu}$ and the Palm measures described in the previous section.

Step 1. Graph structure. The graph structure is given by a pointed-marked tree with Poisson offspring. We distinguish between susceptible (S) nodes and infected (I) nodes.

- Start from a root $\emptyset$ of type (S).
- At each vertex, susceptible (S) nodes have independent Poisson(S_0R_0) susceptible offspring, and Poisson$(I_0\bar{R}_0)$ infected (I) offspring.
- (I) nodes have no offspring.

Edges of the tree are *oriented towards the root.*

Step 2. Decoration. Given the tree structure with distinguished (S) and (I) vertices, we now assign a marking $m_x = (Z_x, \mathcal{X}_x)$ to every vertex x, and a weight a_e for every edge e as follows. If $i = \emptyset$, then $m_\emptyset = (0, \mathcal{X}_\emptyset)$ where $\mathcal{X}_\emptyset$ is distributed as $\mathcal{X}$. For every $i \neq \emptyset$, there exists a unique oriented edge $e = (i, j)$ originated from i and

- If $i \in (I)$, then (a_e, Z_i) is chosen according to the density G. If $i \in (S)$, then $Z_i = 0$ and a_e is chosen according to ν.

- Conditional on (a_e, Z_i), the variable $\mathcal{X}_i$ has the Palm measure evaluated at $a_e + Z_i$.

Remark 14.

- *If $e = (i, j)$ with $i \in (S)$ then $(a_e, \mathcal{X}_i)$ has the Campbell measure introduced in Definition 12.*
- *If $i \in (I)$, then a_e is distributed according to $\bar{\nu}$.*

The random tree $\mathcal{H}$ will correspond to the local limit of the pogm graph $\mathcal{G}^N(x)$ conditioned on $\{Z_x = 0\}$. Let us now consider the infection process on $\mathcal{H}$ introduced in Section 4.2. Conditioned on $\mathcal{H}$, we endow each oriented edge e with a uniform random variable s_e (the intensity of the contact). As pointed out in Definition 9, those r.v's allow to determine whether a path is active or not and to determine the active geodesic at the root.

Define σ^∞ as the length of the active geodesic in $\mathcal{H}$ from the set of (I) leaves to the root $\emptyset$. The following key result connects the distribution of σ^∞ to the delay equation.

Proposition 15. *Define*

$$\forall t \geq 0, \quad B(t) \coloneqq S_0 \mathbb{P}(\sigma^\infty \leq t).$$

Then B solves the delay equation (5).

Proof. As we have assumed that τ has a density w.r.t. the Lebesgue measure, it is clear that this also holds for the distribution of σ^∞. We denote its density by f. Let K, resp. $\bar{K}$, be the number of unmarked, resp. marked, children of the root of $\mathcal{H}$. Let $(\mathcal{H}_1, \dots, \mathcal{H}_N)$ denote the subtrees attached to the root $\emptyset$, and let $(\sigma_1^\infty, \dots, \sigma_K^\infty)$ be the corresponding birth times obtained by determining the active geodesic from $\mathcal{I}_0^N$ to those vertices. Moreover, let $(W_1, \dots, W_K)$ and $(\bar{W}_1, \dots, \bar{W}_{\bar{K}})$ be the weights of the edges ending at $\emptyset$ and starting from unmarked and marked children respectively. (Recall that the edges of the Poisson tree are directed towards the root.) Finally, with a slight abuse of notation, let s_i be the contact intensity on the edge with weight W_i. $\bar{s}_i$ is defined analogously. Define

$$B_i \coloneqq \mathbb{1}_{\{c(W_i + \sigma_i^\infty) \leq s_i\}}, \quad \bar{B}_i = \mathbb{1}_{\{c(\bar{W}_i) \leq \bar{s}_i\}}.$$

Since $\mathcal{H}$ is a tree, one can readily check that

$$\sigma^\infty = \Big(\min_{1 \leq i \leq K} \big\{ B_i(W_i + \sigma_i^\infty) + (1 - B_i) \times \infty \big\} \Big) \wedge \Big(\min_{1 \leq i \leq \bar{K}} \big\{ \bar{B}_i \bar{W}_i + (1 - \bar{B}_i) \times \infty \big\} \Big),$$

with the convention $0 \times \infty = 0$. Define $G(t) = \mathbb{P}(\sigma^\infty > t)$. Let V and $\bar{V}$ be distributed according to τ and $\bar{\tau}$ respectively. As by the branching property, conditional on K and $\bar{K}$, all previously introduced variables are independent, we have

$$\begin{aligned}
G(t) &= \mathbb{E}\Big\{ \Big(1 - \mathbb{E}\big(c(\sigma^\infty + V) \mathbb{1}_{\{\sigma^\infty + V \leq t\}} \big) \Big)^K \Big(1 - \mathbb{E}\big(c(\bar{V}) \mathbb{1}_{\{\bar{V} \leq t\}} \big) \Big)^{\bar{K}} \Big\} \\
&= \mathbb{E}\Big\{ \Big(1 - \int_0^t \int_0^{t-a} c(a+s) f(s) \, \mathrm{d}s \, \nu(\mathrm{d}a) \Big)^K \Big(1 - \int_0^t c(s) \, \bar{\nu}(\mathrm{d}s) \Big)^{\bar{K}} \Big\} \\
&= \exp\Big(- S_0 \int_0^t \int_0^{t-a} c(a+s) f(s) \tau(a) \, \mathrm{d}s \, \mathrm{d}a \\
&\qquad\qquad - I_0 \int_0^t g(a) \int_a^\infty c(u-a) \tau(u) \, \mathrm{d}u \, \mathrm{d}a \Big),
\end{aligned}$$

where, in the last equality, we have used the generating function of a Poisson distribution. It now follows that $B(t) = S_0(1 - G(t))$ satisfies (5). □

5.3 The infection path conditioned on its length

Let us consider the infection process on $\mathcal{H}$ as described in the previous section. For every realization in $\{\sigma^\infty < \infty\}$, define $\mathcal{R}^\infty$ to be the infection path from $\emptyset$ to the (I) leaves in $\mathcal{H}$, and let $\mathcal{A}^\infty$ be the ancestral process defined analogously to Definition 11. In this section, we ask the following question: *conditional on the active geodesic to be of length t, what is the distribution of the vector of infection times along the geodesic?* In order to give an answer to this question, we start with some definition.

Let us consider $\mathcal{R}^\infty$ to be the infection path from $\emptyset$ to the (I) leaves in $\mathcal{H}$ — see Definition 11. Our aim is to provide a description of $\mathcal{R}^\infty$ conditional on $\{\sigma^\infty = t\}$. Define the process $\hat{R}^{(t)} \equiv \hat{R}$ as the $\mathbb{R}$-valued, nonincreasing Markov chain, started from t and stopped upon reaching $(-\infty, 0]$, with transition kernel $Q(x, y)$ defined for all $x > 0$ by

$$\begin{aligned} \forall y \geq x, \quad Q(x,y) &:= 0 \\ \forall y < x, \quad Q(x,y) &:= \frac{S(x)c(x)b(y)}{b(x)}\tau(x-y), \end{aligned}$$

where b is extended to the negative half-line with $b(-t) := I_0 g(t)$. The fact that Q defines a transition kernel follows from the renewal equation for b, which is obtained by differentiating (5) with respect to t:

$$\forall t \geq 0, \qquad b(t) \;=\; c(t)S(t)\int_{-\infty}^{t} b(a)\tau(t-a)da. \tag{10}$$

Define

$$\hat{L}^{(t)} \;:=\; \hat{L} \;=\; \inf\{k : \hat{R}_k^{(t)} \leq 0\}.$$

In the next proposition, we slightly abuse notation and identify $\hat{R}^{(t)}$ with its finite-length restriction to $[\hat{L}]$.

Proposition 16. *Let $\mathcal{R}^\infty$ be the infection path from $\emptyset$ to the (I) leaves. Conditional on $\{\mathcal{R}^\infty(0) = \sigma^\infty = t\}$,*

$$\mathcal{R}^\infty \;=\; \hat{R}^{(t)} \quad \textit{in law.}$$

Proof. Recall that $\sigma^\infty = \mathcal{R}^\infty(0)$ is a random variable valued in $\mathbb{R}_+ \cup \{\infty\}$. By Proposition 15, the density of the random variable σ^∞ on $\mathbb{R}_+$ is given by $S_0^{-1}b(t)$. Let F be the joint probability density of the pair of random variables $(\mathcal{R}^\infty(1), \mathcal{R}^\infty(0) - \mathcal{R}^\infty(1))$, and define

$$\forall t > 0 \text{ and } x < t, \quad F^{(t)}(t-x) \;:=\; \mathbb{1}_{\{x \leq t\}}\frac{F(x, t-x)}{S_0^{-1}b(t)},$$

so that $F^{(t)}$ corresponds to the density of the increment $\mathcal{R}^\infty(0) - \mathcal{R}^\infty(1)$ conditioned on $\{\mathcal{R}^\infty(0) = t\}$. Since $\mathcal{H}$ is a Poisson random tree it is sufficient to understand the first step of the infection path, i.e., we need to show that

$$F^{(t)}(t-x) \;=\; \frac{c(t)S(t)b(x)\tau(t-x)}{b(t)} \tag{11}$$

We distinguish between two cases.

Case 1: $x \in [0,t]$**.** We use the same notation as in the proof of Proposition 15. By construction of the active geodesic and the branching property in $\mathcal{H}$, an argument analogous to the one in Proposition 15 yields that

$$\begin{aligned}
F^{(t)}(t-x) &= \frac{c(t)S_0^{-1}b(x)\nu(t-x)}{S_0^{-1}b(t)} \\
&\quad \times \mathbb{E}\left(\mathbb{1}_{\{K\geq 1\}}K\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K-1}\right)\times \mathbb{E}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}} \\
&= \frac{c(t)b(x)\nu(t-x)}{b(t)}\ S_0R_0 \\
&\quad \times \mathbb{E}\left(\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K}\right)\times \mathbb{E}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}} \\
&= \frac{c(t)b(x)\tau(t-x)}{b(t)} \\
&\quad \times S_0\mathbb{E}\left(\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K}\right)\times \mathbb{E}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}}
\end{aligned}$$

where in the second line, we used the fact that K is Poisson(S_0R_0) (so that the size-biased version of K is identical in law to $K+1$). In the third line, we used the relation $\tau(u) = R_0\nu(u)$. In Proposition 15, we showed that

$$S_0\mathbb{E}\left(\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K}\right)\mathbb{E}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}} = S_0-B(t) = S(t).$$

This shows (11).

Case 2: $x \leq 0$**.** As above

$$\begin{aligned}
F^{(t)}(t-x) &= \frac{c(t)g(-x)\ \tau(t-x)/\bar{R}_0}{S_0^{-1}b(t)} \\
&\quad \times \mathbb{E}\left(\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K}\right)\mathbb{E}\left(\mathbb{1}_{\{\bar{K}\geq 1\}}\bar{K}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}-1}\right) \\
&= \frac{c(t)b(x)\ \tau(t-x)/\bar{R}_0}{I_0S_0^{-1}b(t)}\ I_0\bar{R}_0 \\
&\quad \times \mathbb{E}\left(\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K}\right)\mathbb{E}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}} \\
&= \frac{c(t)b(x)\tau(t-x)}{b(t)} \\
&\quad \times S_0\mathbb{E}\left(\left(1-\mathbb{E}\left(c(V+\sigma^\infty)\mathbb{1}_{\{V+\sigma^\infty\leq t\}}\right)\right)^{K}\right)\mathbb{E}\left(1-\mathbb{E}\left(c(\bar{V})\mathbb{1}_{\{\bar{V}\leq t\}}\right)\right)^{\bar{K}} \\
&= \frac{c(t)S(t)b(x)\tau(t-x)}{b(t)}.
\end{aligned}$$

□

5.4 Harmonic transform

In this section, we prove that the path $\hat{R}^{(t)}$ is the h-transform of a renewal process stopped upon reaching $(-\infty, 0]$. Throughout this section, we assume the existence of a unique Malthusian parameter $\alpha \in \mathbb{R}$ such that

$$\int \exp(-\alpha a)\tau(a)\,\mathrm{d}a \;=\; 1.$$

We define the probability density on $\mathbb{R}_+^*$

$$\forall a > 0, \quad r(a) \;\coloneqq\; \exp(-\alpha a)\tau(a).$$

Let (ξ_i) be a sequence of i.i.d. random variables with probability density r. Let $t > 0$ and define the renewal process $R^{(t)} \equiv R$ as follows

$$\forall k \geq 1, \quad R_k^{(t)} \;=\; t - \sum_{i=1}^{k} \xi_i, \quad R_0^{(t)} = t.$$

We couple the renewal process R with a random variable $K^{(t)} \equiv K$ valued in $\mathbb{N} \cup \{\infty\}$ such that conditional on R,

$$\forall j \geq 0,\ \mathbb{P}(K = j \mid R) = \ell(R_0) \cdots \ell(R_{j-1}) \Big(1 - \ell(R_j)\Big),$$

with $\ell(x) = \mathbb{1}_{\{x>0\}} S(x)c(x) + \mathbb{1}_{\{x\leq 0\}}$.

Remark 17. *Recall that c and S are valued in $[0, 1]$. Think of K as a killing time for the process R, i.e., at site $x > 0$, R dies with probability $1 - S(x)c(x)$, or makes a transition according to the distribution r with the remaining probability. Since by definition $\ell(x) = 1$ for all $x \leq 0$, if R reaches a negative state without being killed, it can no longer be killed.*

Consider the filtration $(\mathcal{F}_k; k \geq 0)$ where

$$\mathcal{F}_k \;=\; \sigma((R_0, \chi_0), \cdots, (R_k, \chi_k)), \quad \text{where } \chi_k \;=\; \mathbb{1}_{\{K\geq k\}},$$

and define the reaching time of $(-\infty, 0]$ as $L \coloneqq \inf\{k : R_k \leq 0\}$.

Lemma 18. *Define*

$$M_k \coloneqq b(R_{k\wedge L}) e^{-\alpha R_{k\wedge L}} \chi_k.$$

The process $(M_k; k \geq 0)$ is a martingale with respect to the filtration $(\mathcal{F}_k; k \geq 0)$.

Proof. Let us compute the conditional expectation $\mathbb{E}(M_{k+1} \mid \mathcal{F}_k)$ for a realization on the event $A_k \coloneqq \{R_k > 0, K \geq k\}$. The martingale property is obviously satisfied for any realization on the complementary event. Using the renewal equation (10) for b, we have

$$\begin{aligned}
\mathbb{1}_{A_k}\mathbb{E}(M_{k+1} \mid \mathcal{F}_k) &= \mathbb{1}_{A_k}\mathbb{E}\Big(b(R_{k+1})e^{-\alpha R_{k+1}}\mathbb{1}_{\{K\geq k+1\}} \mid \mathcal{F}_k\Big) \\
&= \mathbb{1}_{A_k} S(R_k)c(R_k) \int_0^\infty b(R_k - a)e^{-\alpha(R_k - a)}\tau(a)e^{-\alpha a}\,\mathrm{d}a \\
&= \mathbb{1}_{A_k} e^{-\alpha R_k} S(R_k)c(R_k) \int_0^\infty b(R_k - a)\tau(a)\,\mathrm{d}a \\
&= \mathbb{1}_{A_k} b(R_k)e^{-\alpha R_k}.
\end{aligned}$$

□

Proposition 19. *Let $h(s,u) := b(s)e^{-\alpha s}u$ and consider the h-transform of the two dimensional process (R,χ). Then the process $\hat{R}$ is the first coordinate of the h-transformed process.*

Proof. On the one hand, the previous lemma implies that h is an harmonic function for the bivariate process (R,χ). On the other hand, the transition kernel $\hat{Q}$ for the h-transformed process can be rewritten explicitly as

$$\begin{aligned}
\forall x,y;\ \hat{Q}\Big((x,1),(y,0)\Big) &:= 0\\
\forall x \le y; \forall \epsilon \in \{0,1\}\ ,\ \hat{Q}\Big((x,1),(y,\epsilon)\Big) &:= 0\\
\forall x > y;\ \hat{Q}\Big((x,1),(y,1)\Big) &:= \frac{b(y)e^{-\alpha y}}{b(x)e^{-\alpha x}}\ S(x)c(x)\ \tau(x-y)e^{-\alpha(x-y)}\\
&= \frac{b(y)S(x)c(x)}{b(x)}\ \tau(x-y)
\end{aligned}$$

It is now straightforward to check that $\hat{R}$ is identical in law with the first coordinate of the h-transformed process. □

Let P be the law of the bivariate path (R,χ) stopped at $L = \inf\{k : R_k \le 0\}$. Let $\hat{P}$ be the law of h-transform $(\hat{R},\hat{\chi})$ stopped at $\hat{L} = \inf\{k : \hat{R}_k \le 0\}$. Then $\hat{P} \ll P$ and the Radon–Nykodim derivative is given by

$$\frac{\mathrm{d}\hat{P}}{\mathrm{d}P} \;=\; \frac{b(R_L)\exp(-\alpha R_L)}{b(t)\exp(-\alpha t)}\chi_L.$$

This immediately entails the following result.

Proposition 20. *Assume that $g(t) = \alpha\exp(-\alpha t)$. Then $\hat{P}$ is obtained by conditioning the renewal process R on not being killed before time L, and $b(t)$ can be written:*

$$b(t) = \alpha e^{\alpha t} P(R^{(t)} \text{ is not killed before time } L).$$

Remark 21. *Consider the linearized version of the Kermack–McKendrick equation*

$$\begin{aligned}
\partial_t n(t,a) + \partial_a n(t,a) &= 0\\
\forall t \ge 0,\ n(t,0) &= c(t)\int_0^\infty n(t,a)\,\tau(\mathrm{d}a)\\
\forall a \ge 0,\ n(0,a) &= I_0 g(a)
\end{aligned}$$

obtained from (1) by assuming $S(t) = 1$. This can be thought as the age structure of a population where susceptibles are in excess. One can check that if $g(a) = \alpha\exp(-\alpha t)$, then $b_{lin}(t) := n(0,t) = \alpha e^{\alpha t}$. As a consequence, Proposition 20 can be rewritten as

$$b(t) \;=\; b_{lin}(t)\ P(R^{(t)} \text{ is not killed before time } L).$$

We close this section by a brief discussion on the previous result. In [14], we considered a "linearized" version of the present model by making the simplifying

assumption that susceptible individuals are always in excess (branching assumption), so that the epidemic is described by a Crump-Mode-Jagers process. When $c \equiv 1$ and $R_0 > 1$, the process is supercritical. Starting from a single infected individual, there is a positive probability of non-extinction and conditional on this event, the number of infected grows exponentially at rate $\alpha > 0$. Further, it is well known from the seminal work of Jagers and Nerman [24] that under mild assumptions,

1. the age structure of the population converges to the exponential profile $g(t) = \alpha \exp(-\alpha t)$ mentioned in Proposition 20.

2. the infection path — interpreted as the ancestral line in the work of Jagers and Nerman — is well described by the renewal process R. More precisely, if one sample an infected individual at a large time t, its infection path converges to the renewal process R.

We can draw two conclusions out of those observations. As a consequence of the first item, the age structure $g(t) = \alpha \exp(-\alpha t)$ could be interpreted as the age structure emerging from a single infected individual in the past (provided that the initial faction of infected individuals in our model is small). The second conclusion is that the effect of the conditioning in Proposition 20 encodes the effect of the saturation and the contact rate c on the genealogy. Recall that in the absence of saturation (branching approximation) and full contact rate ($c \equiv 1$), the infection path is distributed as the renewal process. When those effects are taken into account, Proposition 20 indicates that the law of the infection path is twisted in such a way that infection paths with infection occurring at low susceptible frequency (i.e low values of S) and high contact rates c are favored. This is consistent with the intuition that ancestral infections tend to be biased towards periods when many infections occurred.

6 Convergence of the infection graph

Recall the infection graph $\mathcal{G}^N$ defined in Section 4.1, and the notation $\mathcal{X}_x = (\mathcal{P}_x, X_x)$. The following result proves the convergence of the local structure of $\mathcal{G}^N$ to the tree described in Section 5.

Proposition 22. *Let x, y be two distinct elements of $[N]$. Conditional on $\{Z_x = 0, Z_y = 0\}$,*

$$\Big(\mathcal{G}^N(x), \mathcal{G}^N(y)\Big) \Longrightarrow \Big(\mathcal{H}, \mathcal{H}'\Big),$$

where $\mathcal{H}$ and $\mathcal{H}'$ are distributed as two independent Poisson marked trees as defined in Section 5.2.

Proof. The proof is not so difficult with the right approach, but somewhat heavy in notation.

Define the genealogical projection of the pogm graph $\mathcal{G}^N(x)$ as the graph obtained by erasing all the information about the weights and the marking, except for the status of each vertex at time $t = 0$: i.e., an edge x is either of type (I) if $Z_x > 0$ or of type (S) if $Z_x = 0$. The genealogical projection of $\mathcal{G}^N(x)$ will be denoted by $G^N(x)$. If $r \in \mathbb{N}$, define $[G^N(x)]_r$ as the subgraph of $G^N(x)$ induced by

the vertices at a genealogical distance at most r from the pointed vertex. $G^N(y)$ and $[G^N(y)]_r$ are defined analogously.

Let $r \in \mathbb{N}$. Let (T_1, T_2) be a pair of discrete, rooted (i.e., pointed), *typed planar* trees with genealogical depth $\leq r$. As for the graph $G^N(x)$, we assume that vertices are endowed with a type: either (S) or (I). By typed planar, we mean that at every vertex, children of the same type (i.e., (I) or (S)) are ordered. We will further assume that the roots of T_1 and T_2 are always of type (S). To ease the terminology, we will often shorten *typed planar* to *planar*.

We separate $E_1 = E_1' \cup E_1''$, where E_1 is the set of edges of T_1, E_1'' is the set of edges that originate from a (I) leaf, and $E_1' = E_1 \setminus E_1''$. Let V_1' be the set of (S) vertices of T_1, and $\mathring{V}_1'$ be the subset of V_1' at distance $< r$ from the root. For any vertex $u \in V_1'$, define n_u (resp. m_u) as the number of (S) (resp. (I)) offspring of u. We define the analogous quantities $E_2 = E_2' \cup E_2''$ for T_2. Also, for any $e \in E_1 \cup E_2$, we fix a continuous bounded map $f_e \colon [0,\infty)^2 \times \mathcal{M} \times \mathcal{S} \to [0,\infty)$.

Let $\mathcal{H}^{\mathrm{pl}}$ denote the *planarization* of a Poisson marked tree. By planar, we mean again that at every vertex, a uniform random ordering is prescribed on the set of (I) children, and on the set of (S) children. More precisely, conditional on a realization of the tree, at every vertex u, we pick a given ordering uniformly at random among the $n_u! m_u!$ possibilities. Let H^{pl} be the genealogical projection of $\mathcal{H}^{\mathrm{pl}}$. (As before, the genealogical projection is obtained by stripping $\mathcal{H}$ from its marks and weights, while preserving the status of the nodes at time $t = 0$.) The following functional F defined by

$$F(T_1, (f_e)_{e\in E_1}) \coloneqq \mathbb{E}\bigg(\mathbb{1}([H^{\mathrm{pl}}]_r = T_1) \prod_{e=(i,j)\in E_1} f_e(M_{(i,j)}) \bigg),$$

where

$$M_{(i,j)} \coloneqq (a_{(i,j)}, Z_i, \mathcal{P}_i, X_i) \;\equiv\; (a_{(i,j)}, Z_i, \mathcal{X}_i),$$

can be explicitly computed as follows

$$F(T_1, (f_e)_{e\in E_1}) = \prod_{u\in \mathring{V}_1'} \Big(e^{-S_0R_0 - (1-S_0)\bar{R}_0} \frac{(S_0R_0)^{n_u}((1-S_0)\bar{R}_0)^{m_u}}{n_u! m_u!} \Big)$$
$$\prod_{e\in E_1'} \mathbb{E}\bigg(\frac{1}{R_0} \int_0^\infty f_e(a, 0, \mathcal{X}) d\mathcal{P}(a) \bigg) \prod_{e\in E_1''} \mathbb{E}\bigg(\frac{1}{\bar{R}_0} \int_0^\infty g(u) \int_u^\infty f_e(a, u, \mathcal{X}) d\mathcal{P}(a) du \bigg).$$

Let us us consider a realization such that $[G^N(x)]_r, [G^N(y)]_r$ are trees and have an empty intersection. Let $[G^N(x), G^N(y)]_r^{\mathrm{pl}}$ denotes the *planarization* of the pair $([G^N(x)]_r, [G^N(y)]_r)$ induced by the labeling of the nodes. Namely, $[G^N(x), G^N(y)]_r^{\mathrm{pl}}$ is the *unlabeled* pointed planar marked pair of trees (or forest) where the ordering of the children of a given vertex in the tree is inherited from the original labelling of the nodes.

Let $\bar{\mathbb{P}}$ be the measure $\mathbb{P}$ conditioned on the event $\{Z_x, Z_y = 0\}$. We aim at showing that

$$\bar{\mathbb{E}}\bigg(\mathbb{1}\Big([G^N(x), G^N(y)]_r^{\mathrm{pl}} = (T_1, T_2) \Big) \Big[\prod_{e\in E_1\cup E_2} f_e(M_e) \Big] g_1(\mathcal{X}_x) g_2(\mathcal{X}_y) \bigg)$$
$$\underset{N\to\infty}{\longrightarrow} F(T_1, (f_e)_{e\in E_1}) F(T_2, (f_e)_{e\in E_2}) \mathbb{E}[g_1(\mathcal{X})] \mathbb{E}[g_2(\mathcal{X})], \tag{12}$$

where g_1 and g_2 are continuous bounded maps $\mathscr{X} \to \mathbb{R}$. To see why it is sufficient to show (12), note that the RHS is simply

$$\mathbb{E}\left(\mathbb{1}\left([H^{\mathrm{pl}}]_r, [H'^{\mathrm{pl}}]_r = (T_1, T_2) \right) \left[\prod_{e=(i,j)\in E_1 \cup E_2} f_e(M_e) \right] g_1(\mathcal{X}_\emptyset) g_2(\mathcal{X}'_\emptyset) \right),$$

where $\mathcal{H}, \mathcal{H}'$ are independent Poisson marked trees, H and H' their genealogical projection, for each edge $e = (i,j)$ in either trees, M_{ij} denotes the mark $(a_{(i,j)}, Z_i, \mathcal{X}_i)$, and $\mathcal{X}_\emptyset$ and $\mathcal{X}'_\emptyset$ denote the marks at the root of each tree. We see now that it is indeed enough to consider the event where $[G^N(x)]_r, [G^N(y)]_r$ are trees and have an empty intersection: define

$$B^N \coloneqq \{[G^N(x)]_r, [G^N(y)]_r \text{ not a pair of tree with disjoint union}\}.$$

Then by Fatou's lemma,

$$\begin{aligned} \limsup_{N\to\infty} \bar{\mathbb{P}}(B^N) &= 1 - \liminf \sum_{(T_1,T_2)} \bar{\mathbb{P}}([G^N(x), G^N(y)]_r^{\mathrm{pl}} = (T_1, T_2)) \\ &\leq 1 - \sum_{(T_1,T_2)} \mathbb{P}([H^{\mathrm{pl}}]_r = T_1)\mathbb{P}([H^{\mathrm{pl}}]_r = T_2) = 0, \end{aligned}$$

where the sum is taken over every pair of pointed discrete planar trees with height at most r. Therefore, Equation (12) shows local convergence of the planarized pair $(\mathcal{G}^N(x)^{\mathrm{pl}}, \mathcal{G}^N(y)^{\mathrm{pl}})$ to $(\mathcal{H}^{\mathrm{pl}}, \mathcal{H}'^{\mathrm{pl}})$. To get the (non-planarized) statement of the proposition, let T_1° and T_2° be two (unordered) rooted marked trees. Since (12) holds regardless of the orientation of T_1 and T_2, we can sum over every planar pair of trees (T_1, T_2) with unordered graph structure T_1° and T_2°, which shows that the convergence $(\mathcal{G}^N(x)^{\mathrm{pl}}, \mathcal{G}^N(y)^{\mathrm{pl}}) \Rightarrow (\mathcal{H}^{\mathrm{pl}}, \mathcal{H}'^{\mathrm{pl}})$ holds in the absence of planarization, completing the proof of Proposition 22.

We now turn to the proof of (12). Let us now consider any two *labeled* pointed marked rooted trees $(\tilde{T}_1, \tilde{T}_2)$, where (1) the vertices are labeled by $[N]$ and the roots are labeled by x and y respectively; and (2) $[\tilde{T}_1, \tilde{T}_2]^{\mathrm{pl}} = (T_1, T_2)$. Let $n' \coloneqq |T_1| + |T_2|$ and $(N-2)_{n'-2} \coloneqq (N-2)(N-3)\cdots(N-n'+1)$. Define the combinatorial factor

$$K^N(T_1, T_2) \coloneqq (N-2)_{n'-2} \prod_{u \in \mathring{V}'_1} \frac{1}{n_u! m_u!} \prod_{u \in \mathring{V}'_2} \frac{1}{n_u! m_u!}$$

which corresponds to be the number of *disjoint* labeled pointed marked trees $(\tilde{t}, \tilde{t}')$ (with labels in $[N]$) such that $([\tilde{t}, \tilde{t}']^{\mathrm{pl}}) = (T_1, T_2)$ and the pointed vertices are respectively x and y. By exchangeability, it is clear that the LHS of (12) is equal to

$$K^N(T_1, T_2)\bar{\mathbb{E}}\left(\mathbb{1}([G^N(x)]_r = \tilde{T}_1 \text{ and } [G^N(y)]_r = \tilde{T}_2) \left[\prod_{e=(i,j)\in E_1\cup E_2} f_e(M_{(i,j)}) \right] g_1(\mathcal{X}_x) g_2(\mathcal{X}_y) \right).$$

It remains to estimate the latter expected value. Let us identify the vertices $i \in \tilde{T}_1$ with their labels in $[N]$, and define:

- for $i \in \tilde{T}_1 \setminus \{x\}$, $p(i)$ is the unique element in $\tilde{T}_1$ such that $e(i) = (i, p(i))$ is an oriented edge in $\tilde{T}_1$;

- if $i = x$, then for convenience let $p(i) \coloneqq \dagger$ denote any element not in $[N]$.

Intuitively, $p(i)$ corresponds to an individual that i can potentially infect. We define the analogous $p(i), e(i)$ for $i \in \tilde{T}_2$. For $i \notin \tilde{T}_1 \cup \tilde{T}_2$, we define for convenience of notation $p(i) = \dagger$. Let us define the event

$$Q_i(\tilde{T}_1, \tilde{T}_2) \equiv Q_i = \bigcap_{k \in \mathring{V}_1' \cup \mathring{V}_2' \setminus \{p(i)\}} \{i \not\to k\}.$$

where $i \to k$ (resp., $i \not\to k$) means that (i, k) is an oriented edge of the infection graph (resp., (i, k) is not an oriented edge of the infection graph).

Recall that $\mathring{V}_1'$ (resp. $\mathring{V}_2'$), denotes the set of (S) vertices of $\tilde{T}_1$ (resp. $\tilde{T}_2$) at genealogical distance $< r$ from the root. We will also write V_1'' and V_2'' for the set of (I) vertices in $\tilde{T}_1$ and $\tilde{T}_2$. The key behind the definition of the (Q_i) is the following decomposition:

$$\begin{aligned}
&\Big\{[G^N(x)]_r = \tilde{T}_1 \text{ and } [G^N(y)]_r = \tilde{T}_2\Big\} \\
&\quad = \bigcap_{i \in (\tilde{T}_1 \cup \tilde{T}_2) \setminus \{x,y\}} \Big(\{i \to p(i)\} \cap Q_i\Big) \cap \bigcap_{i \notin (\tilde{T}_1 \cup \tilde{T}_2) \setminus \{x,y\}} Q_i \\
&\qquad \cap \bigcap_{i \in V_1'' \cup V_1''} \{i \in \mathcal{I}_0^N\} \bigcap_{i \in V_1' \cup V_2'} \{i \notin \mathcal{I}_0^N\}.
\end{aligned}$$

Using the independence of the marking, we get that

$$\begin{aligned}
&\bar{\mathbb{E}}\Bigg(\mathbb{1}([G^N(x)]_r = \tilde{T}_1 \text{ and } [G^N(y)]_r = \tilde{T}_2)\Big[\prod_{e=(i,j) \in E_1 \cup E_2} f_e(M_{ij})\Big] g_1(\mathcal{X}_x) g_2(\mathcal{X}_y)\Bigg) \\
&\quad = \prod_{i \in (V_1' \cup V_2') \setminus \{x,y\}} \mathbb{E}\Big[f_{e(i)}(M_{ip(i)}) \mathbb{1}(\{i \to p(i)\} \cap Q_i) \;\mid Z_i = 0\Big] \\
&\qquad \times \prod_{i \in (V_1'' \cup V_2'')} \mathbb{E}\Big[f_{e(i)}(M_{ip(i)}) \mathbb{1}(\{i \to p(i)\} \cap Q_i) \;\mid Z_i > 0\Big] \\
&\times (1 - S_0)^{|E_1''| + |E_2''|} S_0^{|E_1'| + |E_2'|} \;\times\; \mathbb{E}[g_1(\mathcal{X})] \mathbb{E}[g_2(\mathcal{X})] \times \prod_{i \notin (\tilde{T}_1 \cup \tilde{T}_2) \setminus \{x,y\}} \bar{\mathbb{P}}(Q_i),
\end{aligned} \tag{13}$$

where we used the fact that $|E_i'| = |V_i' \setminus \{x\}|$ and $|E_i''| = |V_i''|$ for $i = 1, 2$. We now estimate (13) in three consecutive steps.

Step 1. We start by showing that

$$\prod_{i \notin (\tilde{T}_1 \cup \tilde{T}_2) \setminus \{x,y\}} \bar{\mathbb{P}}(Q_i) = e^{-n(S_0 R_0 + (1 - S_0)\bar{R}_0)} + o(1).$$

First, note that $\bar{\mathbb{P}}(Q_i) = \mathbb{P}(Q_i)$ for each $i \notin \{x, y\}$. Since it will be shown in the following that $\bar{\mathbb{P}}(Q_x) = \mathbb{P}(Q_x \mid x \notin \mathcal{I}_0^N) \to 1$, we may ignore the factor $\bar{\mathbb{P}}(Q_x)\bar{\mathbb{P}}(Q_y)$ in the approximation of the product above. Recall the notation $\widehat{\mathcal{P}}_i$, which is the shifted infection point process for $i \in \mathcal{I}_0^N$, and the plain infection point process for $i \notin \mathcal{I}_0^N$. If $i \notin (\tilde{T}_1 \cup \tilde{T}_2)$, we have

$$\begin{aligned}
\mathbb{P}(Q_i) &= \mathbb{E}\Big[\Big(1 - \frac{n}{N}\Big)^{|\widehat{\mathcal{P}}_i|}\Big] \\
&= 1 + \mathbb{E}\Big[\Big(1 - \frac{n}{N}\Big)^{|\widehat{\mathcal{P}}_i|} - 1\Big],
\end{aligned}$$

with $n \coloneqq |\mathring{V}_1' \cup \mathring{V}_2'|$. Noticing that $|1-(1-n/N)^{|\widehat{\mathcal{P}}_i|}| \le n|\widehat{\mathcal{P}}_i|/N$, an application of dominated convergence shows that

$$\lim_{N\to\infty} N\mathbb{E}\Big[\Big(1-\frac{n}{N}\Big)^{|\widehat{\mathcal{P}}_i|}-1\Big] = n\mathbb{E}\Big[|\widehat{\mathcal{P}}_i|\Big]$$

so that

$$\mathbb{E}\Big[\Big(1-\frac{n}{N}\Big)^{|\widehat{\mathcal{P}}_i|}-1\Big] \sim \frac{n\mathbb{E}\Big[|\widehat{\mathcal{P}}_i|\Big]}{N}. \tag{14}$$

The same reasoning holds under a conditioning on $\{i \notin \mathcal{I}_0^N\}$ or on the complementary event, yielding $\mathbb{P}(Q_i \mid i \notin \mathcal{I}_0^N) = 1 - nR_0/N + o(N^{-1})$, and similarly, $\mathbb{P}(Q_i \mid i \in \mathcal{I}_0^N) = 1 - n\bar{R}_0/N + o(N^{-1})$. By the law of large numbers, this yields the following approximation for the product

$$\prod_{i\notin(\tilde{T}_1\cup\tilde{T}_2)\setminus\{x,y\}} \bar{\mathbb{P}}(Q_i) = e^{-n(S_0R_0+(1-S_0)\bar{R}_0)} + o(1).$$

Step 2. Let us now show that for every $i \in (\tilde{T}_1 \cup \tilde{T}_2) \setminus \{x,y\}$,

$$\begin{aligned}\mathbb{E}\Big[f_{e(i)}(M_{ip(i)})\mathbb{1}\Big(\{i\to p(i)\}\cap Q_i\Big) \;\Big|\; Z_i=0\Big]\\ = \frac{1}{N}\mathbb{E}\bigg(\int_0^\infty f_{e(i)}(a,0,\mathcal{X})\,\mathrm{d}\mathcal{P}_i(\mathrm{d}a)\bigg) + o(N^{-1}).\end{aligned}$$

Let us define $\tilde{Q}_i$ as the event

$$\tilde{Q}_i = \{i \text{ has no multiple edges to } p(i)\}.$$

We have

$$\mathbb{P}\Big(\{i\to p(i)\}\setminus\tilde{Q}_i\Big) = \mathbb{E}\Big[1-\Big(1-\frac{1}{N}\Big)^{|\widehat{\mathcal{P}}_i|} - |\widehat{\mathcal{P}}_i|\frac{1}{N}\Big(1-\frac{1}{N}\Big)^{|\widehat{\mathcal{P}}_i|-1}\Big].$$

Dominated convergence shows that

$$\lim_{N\to\infty} N\mathbb{E}\Big[|\widehat{\mathcal{P}}_i|\frac{1}{N}\Big(1-\frac{1}{N}\Big)^{|\widehat{\mathcal{P}}_i|-1}\Big] = \mathbb{E}\Big[|\widehat{\mathcal{P}}_i|\Big]$$

and combined with (14) with $n=1$, this proves that

$$\mathbb{P}\Big(\{i\to p(i)\}\setminus\tilde{Q}_i\Big) = o(N^{-1})$$

Therefore it remains to show that

$$\begin{aligned}\mathbb{E}\Big[f_{e(i)}(M_{ip(i)})\mathbb{1}\Big(\{i\to p(i)\}\cap\tilde{Q}_i\cap Q_i\Big) \;\Big|\; Z_i=0\Big]\\ = \frac{1}{N}\mathbb{E}\bigg(\int_0^\infty f_{e(i)}(a,0,\mathcal{X})\,\mathrm{d}\mathcal{P}_i(\mathrm{d}a)\bigg) + o(N^{-1}).\end{aligned}$$

To show this, check that with our construction, this expression can be computed explicitly:

$$\begin{aligned}\mathbb{E}\Big[f_{e(i)}(M_{ip(i)})\mathbb{1}\Big(\{i\to p(i)\}\cap\tilde{Q}_i\cap Q_i\Big) \;\Big|\; Z_i=0\Big]\\ = \mathbb{E}\Big[\frac{1}{N}\Big(1-\frac{n}{N}\Big)^{|\mathcal{P}_i|-1}\int_0^\infty f_{e(i)}(a,0,\mathcal{X})\,\mathrm{d}\mathcal{P}_i(\mathrm{d}a)\Big],\end{aligned}$$

which proves our claim. The same argument works if we condition on the event $\{Z_i > 0\}$ and one can prove that

$$\mathbb{E}\Big[f_{e(i)}(M_{ip(i)})\mathbb{1}\Big(\{i \to p(i)\} \cap Q_i\Big) \;\Big|\; Z_i > 0\Big] \\ = \frac{1}{N}\mathbb{E}\bigg(\int_0^\infty g(u)\int_0^\infty f_{e(i)}(a,u,\mathcal{X})\,\mathrm{d}\mathcal{P}_i(\mathrm{d}a)du\bigg) \;+\; o(N^{-1}).$$

Step 3. We can now leverage the previous estimates to get the desired approximation. As before, we set $n' \coloneqq |T_1| + |T_2|$. From (13), we deduce from the previous approximations

$$\bar{\mathbb{E}}\bigg(\mathbb{1}([G^N(x)]_r = \tilde{T}_1 \text{ and } [G^N(y)]_r = \tilde{T}_2)\Big[\prod_{e=(i,j)\in E_1\cup E_2} f_e(M_{ij})\Big]g_1(\mathcal{X}_x)g_2(\mathcal{X}_y)\bigg) \\ = \frac{\mathbb{E}[g_1(\mathcal{X})]\mathbb{E}[g_2(\mathcal{X})]}{N^{n'-2}}(1-S_0)^{|E_1''|+|E_2''|}S_0^{|E_1'|+|E_2'|}\Big(e^{-n(S_0R_0+(1-S_0)\bar{R}_0)}\Big)R_0^{|V_1'|+|V_2'|-2}\bar{R}_0^{|V_1''|+|V_2''|} \\ \prod_{i\in(V_1'\cup V_2')\setminus\{x,y\}}\mathbb{E}\bigg(\frac{1}{R_0}\int_0^\infty f_e(a,0,\mathcal{X})d\mathcal{P}(a)\bigg)\prod_{i\in V_1''\cup V_2''}\mathbb{E}\bigg(\frac{1}{\bar{R}_0}\int_0^\infty g(u)\int_u^\infty f_e(a,u,\mathcal{X})d\mathcal{P}(a)du\bigg) \\ + o(\frac{1}{N^{n'-2}}).$$

We get the desired limit on the RHS of (12) by multiplying by the combinatorial factor $K^N(T_1,T_2)$ and by noting that

$$\forall i = 1,2, \quad |E_i''| = |V_i''|, \qquad |E_i'| = |V_i'| - 1,$$

and the proof is complete. □

7 Convergence of the historical process

Corollary 23. *Let $x, y \in [N]$ and $(\mathcal{A}_x^N, \mathcal{A}_y^N)$ be the ancestral paths starting resp. from x and y. Conditional on $\{Z_x = 0, Z_y = 0\}$,*

$$(\mathcal{A}_x^N, \mathcal{A}_y^N) \Longrightarrow (\mathcal{A}_1^\infty, \mathcal{A}_2^\infty)\,,$$

where $\mathcal{A}_1^\infty$ (resp., $\mathcal{A}_2^\infty$) are the limiting random variables defined in Section 5.2.

Proof. It is sufficient to show that the distribution of $\mathcal{A}_x^N$, obtained from $\mathcal{G}^N(x)$ out of the infection process described in Section 4.2, converges to that of $\mathcal{A}^\infty$ derived from the Poisson limiting tree $\mathcal{H}$. Up to using Skorohod's representation theorem, see Theorem 6.7 in [3], we might assume that $\mathcal{G}^N(x)$ converges a.s. to $\mathcal{H}$ in the topology defined in Section 4.4. In what follows, we work conditional on $\mathcal{G}^N(x)$ and $\mathcal{H}$ and consider them as deterministic. It is now sufficient to prove that for any continuous bounded function F

$$\mathbb{E}\Big(F(\mathcal{A}_x^N)\mathbb{1}_{\{\sigma_x^N \le t\}} \mid \mathcal{G}^N(x)\Big) \longrightarrow \mathbb{E}\Big(F(\mathcal{A}^\infty)\mathbb{1}_{\{\sigma^\infty \le t\}} \mid \mathcal{H}\Big), \quad \text{a.s.} \tag{15}$$

where the expected value is taken with respect to the contact intensity variables s_e (as defined in Section 4.2). It is clear that the latter statement holds if $\mathcal{H}$ is a

finite tree. In the following, we show that we can actually restrict our attention to this specific case.

Step 1. For every $r \in \mathbb{N}$, define $[\mathcal{H}]_r$ the limiting marked Poisson tree restricted to the vertices at genealogical distance at most r from the root. We first claim that there exists r_0 such that,

$$\forall r \geq r_0, \quad \min\{|\pi_v| : v \in [\mathcal{H}]_r\} > t, \tag{16}$$

where π_v denotes the unique path from v to the root in $\mathcal{H}$ with the convention that $\min\{\emptyset\} = \infty$. If $\mathcal{H}$ is finite there is nothing to show. Let us now assume that $\mathcal{H}$ is supercritical, i.e., that $R_0 S_0 > 1$. Further, let us condition the process on non-extinction.

Let $(V_r;\, r \geq 0)$ be the process that records the ages of the (S) vertices of the Poisson tree $\mathcal{H}$, defined as

$$V_r := \sum_{\substack{u \in \mathcal{H},\, d(u,\emptyset)=r \\ u \text{ of type } (S)}} \delta(|\pi_u|),$$

where π_u is the unique path connecting u to the root $\emptyset$, and $d(\emptyset, u)$ is the genealogical distance between u and $\emptyset$. $(V_r)_{r\geq 0}$ is a branching random walk with Poisson$(R_0 S_0)$ offspring distribution, and it follows from general results that, conditional on non-extinction, its minimum drifts to ∞, see for instance Theorem 5.12 in [29]. As $\mathcal{H}$ is obtained by attaching independently to any unmarked vertex a Poisson$(I_0 \bar{R}_0)$ distributed number of (I) leaves, this also shows that

$$\lim_{r\to\infty} \min_{\substack{u\in\mathcal{H} \\ d(u,\emptyset)>r}} |\pi_u| = \infty.$$

This completes Step 1.

Step 2. The first step entails that, for a.e. realization of $\mathcal{H}$, we can find a large enough r such that for all $v \notin [\mathcal{H}]_r$, the path from v to the root has length larger than t. For N large enough, $[\mathcal{G}^N(x)]_r$ is isomorphic to $[\mathcal{H}]_r$, and the weights of the edges of $[\mathcal{G}^N(x)]_r$ converge to those of $[\mathcal{H}]_r$. The first consequence is that (16) holds if we replace $\mathcal{H}$ by $\mathcal{G}^N(x)$ for N large enough. It follows that (15) is equivalent to proving the same convergence result if we replace the infinite graph $\mathcal{G}^N(x)$ (resp., $\mathcal{H}$) by $[\mathcal{G}^N(x)]_{r_0}$ (resp., $[\mathcal{H}]_{r_0}$). The latter holds from the definition of the convergence of $\mathcal{G}^N(x)$ to $\mathcal{H}$ in the local topology. □

Let X be an arbitrary random variable. In the following, $\mathcal{L}(X)$ will denote the law of the random variable X.

Theorem 24 (Convergence of the historical process)**.** *Define the historical process as the following empirical measure*

$$H^N := \frac{1}{N} \sum_{x\in[N]} \mathbb{1}_{\{\sigma_x^N < \infty\}}\, \delta_{\mathcal{A}_x^N}. \tag{17}$$

Let $\mathcal{A}^\infty$ be the limiting ancestral process in the Poisson tree $\mathcal{H}$. Finally, let $(\sigma_0, \mathcal{X})$ denote a pair of independent random variables where $-\sigma_0$ is distributed according to the density g. Then

$$H^N \longrightarrow S_0 \mathbb{P}(\sigma^\infty < \infty) \mathcal{L}\left(\mathcal{A}^\infty | \sigma^\infty < \infty\right) + I_0 \mathcal{L}(\sigma_0, \mathcal{X}).$$

where $\mathcal{L}\left(\mathcal{A}^\infty | \sigma^\infty < \infty\right)$ is the law of the random variable $\mathcal{A}^\infty$ conditioned on the event $\{\sigma^\infty < \infty\}$, and the convergence is in distribution for the weak topology.

Proof. Let π be an ancestral path. We will denote by $\sigma(\pi)$ the infection time associated with π.

Let $x \in \mathcal{I}_0^N$. The ancestral path at x coincides with $(-Z_x, \mathcal{X}_x)$. Conditional on the set of infected individuals $\mathcal{I}_0^N$, the ancestral paths of infected individuals are i.i.d. with law $g(-\sigma)\,\mathrm{d}\sigma \otimes \mathcal{L}(\mathcal{X})$. By the law of large numbers,

$$\mathbb{1}_{\sigma(\pi)<0} H^N(d\pi) \longrightarrow I_0 g(-\sigma)\,\mathrm{d}\sigma \otimes \mathcal{L}(\mathcal{X})$$

It remains only to show that

$$\mathbb{1}_{\sigma(\pi)\geq 0} H^N(\mathrm{d}\pi) \longrightarrow S_0 \mathbb{P}(\sigma^\infty < \infty)\mathcal{L}(\mathcal{A}^\infty \mid \sigma^\infty < \infty).$$

It is sufficient to check convergence in distribution of

$$\int \mathbb{1}_{\sigma(\pi)\geq 0} f(\pi) H^N(\mathrm{d}\pi) = \frac{1}{N} \sum_{x\in[N]} \mathbb{1}_{\{Z_x=0, \sigma_x^N<\infty\}} f(\mathcal{A}_x^N),$$

for any continuous bounded maps f (see for instance [22, Theorem 4.11]). First, using Corollary 23, we obtain

$$\mathbb{E}[f(\mathcal{A}_1^N)\mathbb{1}_{\{\sigma_1^N<\infty\}} \mid Z_1 = 0] \underset{N\to\infty}{\longrightarrow} \mathbb{E}[f(\mathcal{A}^\infty)\mathbb{1}_{\{\sigma_1^\infty<\infty\}}]$$

so that

$$\mathbb{E}\left(\int \mathbb{1}_{\sigma(\pi)\geq 0} f(\pi) H^N(\mathrm{d}\pi)\right) = S_0 \mathbb{E}[f(\mathcal{A}_1^N)\mathbb{1}_{\{\sigma_1^N<\infty\}} | Z_1 = 0] \underset{N\to\infty}{\longrightarrow} S_0 \mathbb{E}[f(\mathcal{A}^\infty)\mathbb{1}_{\{\sigma^\infty<\infty\}}]$$

Furthermore we have

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{N^2}\Big(\sum_{x\in[N]} \mathbb{1}_{\{Z_x=0,\sigma_x^N<\infty\}} f(\mathcal{A}_x^N)\Big)^2\right) \\
&\quad = \frac{1}{N^2}\sum_{x\in[N]} \mathbb{E}\Big[\mathbb{1}_{\{Z_x=0,\sigma_x^N<\infty\}} f(\mathcal{A}_x^N)^2\Big] + \frac{1}{N^2}\sum_{x\neq y} \mathbb{E}\Big[\mathbb{1}_{\{Z_x=0,Z_y=0\}}\mathbb{1}_{\{\sigma_y^N,\sigma_x^N<\infty\}} f(\mathcal{A}_x^N) f(\mathcal{A}_y^N)\Big] \\
&\quad \longrightarrow \Big(S_0 \mathbb{E}[f(\mathcal{A}^\infty)\mathbb{1}_{\{\sigma^\infty<\infty\}}]\Big)^2
\end{aligned}$$

again by Corollary 23. This shows that $\int \mathbb{1}_{\sigma(\pi)\geq 0} f(\pi) H^N(\mathrm{d}\pi)$ converges in distribution to the constant $S_0\mathbb{E}[f(\mathcal{A}^\infty)\mathbb{1}_{\{\sigma^\infty<\infty\}}]$, concluding the proof of the theorem. □

We can now prove Theorem 1 using Theorem 24. Recall the notation

$$\mu_t^N = \sum_{x\in[N]} \mathbb{1}_{\{\sigma_x^N\leq t\}} \delta_{(t-\sigma_x^N, X_x(t-\sigma_x^N))}$$

for the empirical distribution of ages and compartments at time t, and the notation

$$Y_t^N(i) = \sum_{x\in[N]} \mathbb{1}_{\{\sigma_x^N\leq t, X_x(t-\sigma_x^N)=i\}} = \mu_t^N\Big([0,\infty), \{i\}\Big) \tag{18}$$

for the number of individuals in compartment i at time t. Note that $\mu_t^N = \mu_t^N(\mathrm{d}a, \mathrm{d}i)$ can be written in terms of H^N as follows

$$\mu_t^N = N \int \mathbb{1}_{\{0\leq\sigma_0\leq t\}} \delta_{(t-\sigma_0, X(t-\sigma_0))}\, H^N(\mathrm{d}\pi), \tag{19}$$

where $\pi = (\sigma_\ell, \mathcal{X}_\ell)_{\ell=0}^k = (\sigma_\ell, (\mathcal{P}_\ell, X_\ell))_{\ell=0}^k$ denotes a generic ancestral path.

Proof of Theorem 1. By Theorem 24 and (19), we get for fixed t and $i \in \mathcal{S}$,

$$\mu_t^N(\mathrm{d}a, \{i\}) \longrightarrow S_0 \mathbb{P}(t - \sigma^\infty \in \mathrm{d}a)\mathbb{P}(X(a) = i),$$

where σ^∞ is the length of the active geodesic in $\mathcal{H}$, and X is a life-cycle process. Using Proposition 15, we can further identify $S_0\mathbb{P}(t - \sigma^\infty \in \mathrm{d}a) = n(t, a)\,\mathrm{d}a$, concluding the proof of convergence of μ_t^N.

Because of the expressions of $Y^N(i)$ in terms of μ_t^N in (18), identification of their limit is trivial. All there is to check is tightness of the processes. Recall that the compartments of the life-cycle process enjoy an "acyclic orientation" property. See statement before Theorem 1. Writing $i \preceq j$ if j can be accessed from i, the process

$$\sum_{j:\, i \preceq j} \frac{1}{N} Y_t^N(j),$$

is nondecreasing in time. Since the finite-dimensional marginals of this nondecreasing process converge towards a bounded limit, tightness follows, see for instance Theorem 3.37, Chapter VI of [18]. The tightness of $Y_t^N(i)/N$ follows by subtracting the previous processes in an appropriate way. □

References

[1] François Baccelli, Bartłomiej Błaszczyszyn, and Mohamed Karray. Random measures, point processes, and stochastic geometry, 2020.

[2] Andrew Barbour and Gesine Reinert. Approximating the epidemic curve. *Electronic Journal of Probability*, 18:30 pp., 2013.

[3] Patrick Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., second edition, 1999.

[4] Fred Brauer. The Kermack–McKendrick epidemic model revisited. *Mathematical Biosciences*, 198(2):119–131, 2005.

[5] Fred Brauer and Carlos Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. Texts in Applied Mathematics. Springer, New Yord, 2012.

[6] Tom Britton and Etienne Pardoux. Stochastic epidemics in a homogeneous community. *arXiv preprint arXiv:1808.05350*, 2018.

[7] Tom Britton and Gianpaolo Scalia Tomba. Estimation in emerging epidemics: Biases and remedies. *Journal of the Royal Society Interface*, 16(150):20180670, 2019.

[8] Anne Cori, Neil M. Ferguson, Christophe Fraser, and Simon Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9):1505–1512, 2013.

[9] Odo Diekmann. Limiting behaviour in an epidemic model. *Nonlinear Analysis: Theory, Methods & Applications*, 1(5):459–470, 1977.

[10] Jie Yen Fan, Kais Hamza, Peter Jagers, and Fima C. Klebaner. Convergence of the age structure of general schemes of population processes. *Bernoulli*, 26:893–926, 2020.

[11] Luca Ferretti, Chris Wymant, Michelle Kendall, Lele Zhao, Anel Nurtay, Lucie Abeler-Dörner, Michael Parker, David Bonsall, and Christophe Fraser. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, 368(6491), 2020.

[12] Raphaël Forien, Guodong Pang, and Etienne Pardoux. Epidemic models with varying infectivity, 2020.

[13] Raphaël Forien, Guodong Pang, and Etienne Pardoux. Estimating the state of the covid-19 epidemic in france using a non-markovian model. *medRxiv*, 2020.

[14] Félix Foutel-Rodier, François Blanquart, Philibert Courau, Peter Czuppon, Jean-Jil Duchamps, Jasmine Gamblin, Élise Kerdoncuff, Rob Kulathinal, Léo Régnier, Laura Vuduc, Amaury Lambert, and Emmanuel Schertzer. From individual-based epidemic models to McKendrick-von Foerster PDEs: A guide to modeling and inferring COVID-19 dynamics, 2020.

[15] Christophe Fraser. Estimating individual and household reproduction numbers in an emerging epidemic. *PLOS ONE*, 2(8):1–12, 2007.

[16] Tapiwa Ganyani, Cécile Kremer, Dongxuan Chen, Andrea Torneri, Christel Faes, Jacco Wallinga, and Niel Hens. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, march 2020. *Eurosurveillance*, 25, 2020.

[17] Kais Hamza, Peter Jagers, and Fima C. Klebaner. The age structure of population-dependent general branching processes in environments with a high carrying capacity. *Proceedings of the Steklov Institute of Mathematics*, 282:90–105, 2013.

[18] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. Grundlehren Der Mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg, second edition, 2003.

[19] Peter Jagers. *Branching Process with Biological Applications*. Wiley, London, 1975.

[20] Peter Jagers and Fima C. Klebaner. Population-size-dependent and age-dependent branching processes. *Stochastic Processes and their Applications*, 87:235–254, 2000.

[21] Peter Jagers and Fima C. Klebaner. Population-size-dependent, age-structured branching processes linger around. *Journal of Applied Probability*, 48A:249–260, 2011.

[22] Olav Kallenberg. *Random Measures, Theory and Applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, Cham, 2017.

[23] William O. Kermack and Anderson G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721, 1927.

[24] Olle Nerman and Peter Jagers. The stable doubly infinite pedigree process of supercritical branching populations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65(3):445–460, 1984.

[25] Guodong Pang and Etienne Pardoux. Functional limit theorems for non-markovian epidemic models, 2020.

[26] Guodong Pang and Etienne Pardoux. Multi-patch epidemic models with general infectious periods, 2020.

[27] Guodong Pang and Étienne Pardoux. Functional central limit theorems for epidemic models with varying infectivity, 2021.

[28] Guodong Pang and Étienne Pardoux. Functional law of large numbers and PDEs for epidemic models with infection-age dependent infectivity, 2021.

[29] Zhan Shi. *Branching Random walks*, volume 2151 of *École d'Été de Probabilités de Saint-Flour*. Springer, Cham, 2015.

[30] Ziad Taib. *Branching Processes and Neutral Evolution*. Lecture Notes in Biomathematics. Springer-Verlag Berlin Heidelberg, 1992.

[31] Horst R. Thieme. Renewal theorems for linear periodic volterra integral equations. *Journal of Integral Equations*, 7(3):253–277, 1984.

[32] Horst R. Thieme. Renewal theorems for some mathematical models in epidemiology. *Journal of Integral Equations*, 8(3):185–216, 1985.

[33] Jacco Wallinga and Marc 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.