

# An epidemic model in inhomogeneous environment

Daniela Bertacchi, Jürgen Kampf, Ecaterina Sava-Huss\*, and Fabio Zucca

November 24, 2021

## Abstract

The current work deals with an epidemic model on the complete graph  $\mathbb{K}_n$  on  $n$  vertices in a non-homogeneous setting, where the vertices may have distinct types. Different types differ in the probability of getting infected, and/or in the capacity of infecting other vertices. This generalizes the model in [4]. We prove in Theorem 5.1 and Theorem 6.3 laws of large numbers and central limit theorems for the total duration of the process and for the number of infected vertices, respectively, when  $n \rightarrow \infty$ . Moreover, we also prove that when all individuals have the same spread capacity, then a population with inhomogeneous susceptibility is less affected by the epidemics than a homogeneous population.

**Keywords:** multitype Galton-Watson process, coupon collector, branching process, limit theorems, Markov chain, Poisson process, random trees.

**AMS subject classification:** 60J10, 60J80, 60F17.

## 1 Introduction

We consider an epidemic model in an inhomogeneous environment, where a virus arrives from outside the system and infects one individual of a population. Then, the infected individual attempts to infect some other individuals of the population, and these individuals, in turn, try to infect other individuals, and so on. We start with a population of  $n$  individuals, that will be encoded in the vertices of the complete graph  $\mathbb{K}_n$  on  $n$  vertices. The virus can circulate along the edges of  $\mathbb{K}_n$ . When we refer to a vertex of a graph, we think of it as an individual of the population.

Once a target vertex  $u$  is hit by an infection attempt, two things may happen. Either  $u$  is visited for the first time, then it becomes infected and makes itself  $\mathcal{L}^{t(u)}(u)$  attempts to infect other vertices in the graph. The number  $\mathcal{L}^{t(u)}(u)$  is called the spread capacity of  $u$  and will be introduced in full detail later. Or the vertex  $u$  has been infected before, then it has antibodies, and it has already revealed its spread capacity, so nothing happens, but

---

\*The research of Ecaterina Sava-Huss is supported by the Austrian Science Fund (FWF): P 34129.

---

just one piece of the spread capacity of the source vertex is lost. The process stops when either all vertices have been infected, or there are non-infected vertices, but the total spread capacity is exhausted, which happens when too much capacity was wasted on vertices with antibodies.

A mathematically equivalent problem is that of information transmission in a network, as used in [4, 5]: the vertices of the graph are computers and when one of these computers receives a message, it spreads the message to a certain number of further computers, but it does so only if it has received the message for the first time.

During this work we shall refer to our model as the *virus spread model* or *infection model*. We recall that processes where individuals after infection become immune are often called SIR models (susceptible-infected-recovered) – see [3] for an introduction to this model and to other epidemic models. There is a vast literature on SIR models. On infinite graphs the aim is usually to study the shape of the set of the sites reached by infection (see for instance [6], [13], [1]) or to identify the critical parameter below which the epidemics dies out in a finite time (in these cases there is a parameter which tunes the intensity of the epidemics, see [10]). In the present paper the graph is finite, hence the epidemics ends almost surely in finite time. Natural questions in this context are how long the whole epidemic lasts, whether all vertices are infected during the epidemic and, if not, how large the number of infected/non-infected vertices is.

A slightly different, well-studied epidemic model is the so-called frog model, where infected individuals randomly walk on the graph and infect all the susceptible individuals they meet on their way. When the underlying graph is the complete graph, the frog model is a particular case of our model. In [9] this model was considered on complete graphs where all vertices are equally likely to be infected and they all have the same deterministic spread capacity. In that paper individuals are assumed to have a deterministic number  $L$  of “lives” and one life is lost when the infection attempt is unsuccessful. This translates into our language by assuming that each individual has a spread capacity equal to  $L + 1$ . The authors derive laws of large numbers and central limit theorems for the number of individuals eventually infected and law of large numbers for the number of infection attempts. The frog model on complete graphs has recently been studied in [11] (with geometric lifespans and simultaneous infection attempts), and [14] (where simultaneous infection attempts are allowed).

In [4] the model was generalized to random spread capacities and laws of large numbers, central limit theorems and large deviation results are derived. Later on, in [5] these results have been generalized from complete graphs to Erdős-Rényi graphs; in particular in [5] it is shown that the model behaves in the limit roughly like the one on complete graphs from [4] by using a coupling argument. However, in [4] and the other works mentioned so far the environment was homogeneous in the sense that all vertices had the same (distribution of the) spread capacity and the same probability of getting infected. As we all know, such assumptions are not very realistic, e.g. the probability to infect a member of the same family is much higher than infecting the member of another family, people with many social contacts have a higher probability of being infected and will infect more people once they are infected. Finally, also factors like age or other diseases can increase the probability of being

---

infected. So, a straightforward generalization of the model introduced in [4] is to consider  $J$  different types of vertices with distinct spread capacities and infection probabilities, for some fixed number  $J \in \mathbb{N}$ . The type of a vertex will be represented by a *type function*  $t : \mathbb{K}_n \rightarrow \{1, 2, \dots, J\}$  such that, for each vertex  $v$ ,  $t(v) = i \in \{1, 2, \dots, J\}$  means that  $v$  is of type  $i$ .

Somehow similarly to [4], our proof approach will be to consider a multitype Galton-Watson tree and a coupon collector with  $J$  different types of coupons on the same probability space as the infection model. In contrast to [4], in our model the coupons will be unequally likely, and thus the success epochs (time needed to collect a new coupon) will not be independent random variables. Independence was heavily used in [4]. To overcome the difficulties arising in the non-independent case, we use thinned Poisson processes as time intervals of the infection process.

**Structure of the paper.** In Section 2 we rigorously introduce our infection model and multitype Galton-Watson processes. We also show how to represent such a process as a Markov chain for which we write down the transition probabilities, and how to interpret it as a coupon collector's problem. Then in Section 3 we couple the infection model with a Poisson process, and then by using the thinning property of such processes, we get rid of the dependencies produced by the coupons of different types and probabilities of being chosen. Subsequently, in Section 3.1 we prove limit results for the number of infected individuals and for the acquisition times in the continuous-time setting. In Section 4, we look at our model as a random subtree of a multitype Galton-Watson tree with  $J$  different types, and we show that as in the homogeneous case, the infection takes place on a macroscopic time level. In Section 5 we prove law of large numbers for the total duration of the process  $\tau_n$  and for the total number  $N_n(\tau_n)$  of individuals infected by time  $\tau_n$  respectively, while in Section 6 we prove central limit theorems for the same quantities. Finally, in Section 7, we compare the homogeneous case with the inhomogeneous one. In particular we prove that when the spread capacities of the individuals are identically distributed, the epidemics in the inhomogeneous environment (that is, with non-uniform susceptibilities) lasts less and involves a smaller fraction of the population, when compared to the case with uniform susceptibility. On the other hand, if also the spread capacities are allowed to have different laws, then it may happen that a population with non-uniform susceptibility is less affected by the epidemics (for instance when more susceptible individuals have smaller expected spread capacity) or it may be more affected (for instance when more susceptible individuals have larger expected spread capacity).

## 2 Preliminaries

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a generic probability space on which all our random variables and processes are defined.

**Conventions and notations.** The number  $J \in \mathbb{N}$  will be reserved for the number of types, and we will sometimes write  $[J] := \{1, 2, \dots, J\}$ . We use the same notation  $[n] := \{1, 2, \dots, n\}$  for any natural number  $n \geq 1$ . All our vectors will be regarded as column

---

vectors, and for a vector  $\mathbf{u} \in \mathbb{R}^J$  (or infinite dimensional) the notation  $\mathbf{u} = (u_1, \dots, u_J)$  is purely for convenience. When we deal with row vectors, it will be either mentioned explicitly, or it will be clear from the context. For  $i \in [J]$ , denote by  $\mathbf{e}_i = (0, \dots, 1, \dots, 0)$  the  $i$ -th standard basis vector in  $\mathbb{R}^J$  that has 1 in the  $i$ -th coordinate, and all the other coordinates are 0.

**Stochastic processes.** For any discrete-time processes  $(X_t)_{t \in \mathbb{N}}$ ,  $(X_t)_{t \geq 0}$  stands for  $(X_{\lfloor t \rfloor})_{t \geq 0}$ , where  $\lfloor t \rfloor$  stands for the floor function. We may also use the ceiling function  $\lceil \cdot \rceil$ . For any interval  $I$  in  $\mathbb{R}$ , define the Skorohod space  $\mathcal{D}(I, \mathbb{R})$  as the space of càdlàg functions (right continuous with left limits) from  $I$  to  $\mathbb{R}$ . As usual, we endow  $\mathcal{D}(I, \mathbb{R})$  with the standard  $\mathbf{J}_1$  topology introduced by Skorohod (see [2] or [8]). We remind that convergence in the  $\mathbf{J}_1$  topology implies uniform convergence on compact sets when the limit function is continuous.

## 2.1 Virus spread model

We take  $\mathbb{K}_n$ , the complete graph on  $n$  vertices, as the base state space where the infection takes place and recall that, for  $J \in \mathbb{N}$ , the type of a vertex is represented by the *type function*  $t : \mathbb{K}_n \rightarrow \{1, 2, \dots, J\}$  that is, for each vertex  $v$ ,  $t(v) = i \in \{1, 2, \dots, J\}$  means that  $v$  is of type  $i$ . The type function is deterministic and fixed at the beginning. For  $i \in [J]$ , denote by  $n_i = |\{v \in \mathbb{K}_n : t(v) = i\}|$  the number of vertices in  $\mathbb{K}_n$  with type  $i$ , so that  $n = \sum_{i=1}^J n_i$ . We assume that each vertex of type  $i \in [J]$  has been assigned the weight  $p_i \in (0, 1)$  (the probability) so that the total weight is  $1 = n_1 p_1 + n_2 p_2 + \dots + n_J p_J$ . The goal of this paper is to extend part of the results in [4], where all vertices have the same type and equal weights  $1/n$ , to the non-homogeneous setting of vertices with different types and weights. If  $p_1 = \dots = p_J = \frac{1}{n}$ , then the proportion  $n_i/n$  of vertices of type  $i$  in the graph represents the weight of this group. We do not restrict to the case of equal probabilities, and the weight of the group of vertices of type  $i$  will be denoted by  $\alpha_i := p_i n_i$ , for  $i \in [J]$ .

**Description of the virus spread model.** We consider the following stochastic process in discrete time. At time  $t = 0$ , a fixed vertex  $v_0 \in \mathbb{K}_n$  receives a virus from outside, and it spreads the virus among the vertices of  $\mathbb{K}_n$  at random, considering that the spread capacity of each vertex is limited and encoded in the following random variables. Let  $\mathcal{L}^1, \dots, \mathcal{L}^J$  be  $J$  independent  $\mathbb{N}$ -valued random variables representing the number of infection attempts that vertices of different types make once they are infected. Furthermore, denote by  $L^{(i,j)}$  the  $\mathbb{N}$ -valued random variable that gives the number of vertices of type  $j$  that one infected vertex of type  $i$  can in turn infect, so  $\mathcal{L}^i = \sum_{j=1}^J L^{(i,j)}$ , and for  $i \in [J]$ . The random variables  $L^{(i,j)}$  will also be used as entries of an offspring distribution matrix  $L$  in a multitype Galton-Watson process in what follows.

For every  $i \in [J]$ , let  $(\mathcal{L}^i(u))_{\{u \in \mathbb{K}_n : t(u)=i\}}$  be a family (of length  $n_i$ ) of i.i.d. copies of  $\mathcal{L}^i$ . For a vertex  $u$  of type  $i$ , the random variable  $\mathcal{L}^i(u)$  represents *the spread capacity* of  $u$ , and has the same distribution as  $\mathcal{L}^i$ . For technical reasons we also need  $\mathcal{L}^i(u)$  to be defined if  $t(u) \neq i$ , but these random variables are void of meaning. We write  $\mathcal{L} := (\mathcal{L}^1, \mathcal{L}^2, \dots, \mathcal{L}^J)$  for the vector of spread capacities.

Notice that the offspring distribution  $(L^{(i,j)})_{i,j \in [J]}$  we can realize is far from arbitrary. In fact, given the event  $\{\mathcal{L}^i(u) = l\}$  that vertex  $u$  makes  $l \in \mathbb{N}$  infection attempts, the types of the vertices hit by these  $l$  attempts are i.i.d. with probabilities  $\alpha_1, \dots, \alpha_J$ , and thus the numbers  $(L^{(i,1)}(u), \dots, L^{(i,J)}(u))$  are multinomial distributed with parameter  $l$  and  $\alpha_1, \dots, \alpha_J$ , i.e.

$$\mathbb{P}[L^{(i,1)}(u) = k_1, \dots, L^{(i,J)}(u) = k_J | \mathcal{L}^i(u) = l] = \binom{l}{k_1, \dots, k_J} \alpha_1^{k_1} \dots \alpha_J^{k_J}.$$

for all  $k_1, \dots, k_J \in \mathbb{N}$  with  $\sum_{i=1}^J k_i = l$ , where  $\binom{0}{0, \dots, 0} := 1$ . Hence the distribution of the vector  $(L^{(i,1)}(u), \dots, L^{(i,J)}(u))$  is given by

$$\mathbb{P}[L^{(i,1)}(u) = k_1, \dots, L^{(i,J)}(u) = k_J] = \mathbb{P}[\mathcal{L}^i(u) = l] \cdot \binom{l}{k_1, \dots, k_J} \alpha_1^{k_1} \dots \alpha_J^{k_J}. \quad (1)$$

where  $l := k_1 + \dots + k_J$ . With the random variables introduced above, we are now ready to explain how the model works. Suppose that the vertex  $v_0 \in \mathbb{K}_n$  that receives the virus is of type  $i_0 \in [J]$ , i.e.  $t(v_0) = i_0$ . This first vertex reveals its spread capacity  $\mathcal{L}^{i_0}(v_0)$ , with  $\mathcal{L}^{i_0}(v_0) = \sum_{i=1}^J L^{(i_0,i)}(v_0)$  and distributes it randomly among the neighbors, not uniformly, but according to the probabilities  $p_i$ ,  $i = 1, \dots, J$ . One may think of having  $\mathcal{L}^{(i_0)}$  independent random walkers at  $v_0$  that perform independently one random step according to the transition probability  $p_j$ , so that  $L^{(i_0,j)}$  walkers reach a vertex of type  $j$  for each  $j \in [J]$ . That is, the spread capacity  $L^{(i_0,j)}(v_0)$  is spread only among vertices of type  $j$ . Recall that  $\sum_{u \in \mathbb{K}_n} p_{t(u)} = \sum_{i=1}^J n_i p_i = 1$ . Once another vertex is reached for the first time by such a random walker, this vertex reveals its capacity, and does the same thing as the first vertex  $v_0$ . If a random walker reaches a vertex infected previously, nothing happens. We continue until there is nothing to be spread around. We can describe the above model as a Markov chain in discrete time. Since the number  $n$  of vertices is finite, this process will terminate in finite time, because the spread capacity of each individual (vertex) is assumed to be a.s. finite.

Our main goal is to study the asymptotics, in  $n$ , for the probability of infecting every site, for the proportion of infected vertices of each type before exhaustion and for the total duration of the process.

### Virus spread model as a Markov chain

We describe now formally the model introduced heuristically above, as a Markov chain  $(M_t)_{t \in \mathbb{N}}$ , where  $M_t$  is given by the following  $(J+1)$ -uple in  $\mathbb{N}^{J+1}$ :

$$M_t = \underbrace{(N_t^1, N_t^2, \dots, N_t^J)}_{:=N_t}, S_t = (N_t, S_t)$$

where  $N_t^i$ , for  $i \in [J]$  represents the number of vertices of type  $i$  infected by time  $t$  and  $S_t$  represents the total spread capacity (revealed and available for distribution) by time  $t$ . We suppress the index  $n$  in the definition of the Markov chain  $M_t$  and in the random variables in

the  $(J+1)$ -uple, but it is clear that all those random variables depend on the total number  $n$  of available vertices. Suppose at time  $t = 0$ , the virus reaches a vertex  $v_0$  with type  $i_0$ , that is

$$\mathbf{N}_0 = \mathbf{e}_{i_0}, \quad \mathbf{S}_0 = \mathcal{L}^{(i_0)}(v_0) \sim \mathcal{L}^{(i_0)}, \quad \text{and} \quad \mathbf{M}_0 = (\mathbf{e}_{i_0}, \mathcal{L}^{(i_0)}).$$

Given the state  $(\mathbf{n}^1, \dots, \mathbf{n}^J, \mathbf{s})$  of the Markov chain  $(\mathbf{M}_t)_{t \in \mathbb{N}}$  at time  $t$ , the state at time  $t+1$  is given by

$$\begin{cases} (\mathbf{n}^1, \dots, \mathbf{n}^J, \mathbf{s} - 1), & \text{with probability } \sum_{i=1}^J \mathbf{n}^i p_i \\ (\mathbf{n}^1 + 1, \mathbf{n}^2, \dots, \mathbf{n}^J, \mathbf{s} + \mathcal{L}_t^1 - 1), & \text{with probability } p_1(\mathbf{n}_1 - \mathbf{n}^1) \\ \vdots \\ (\mathbf{n}^1, \dots, \mathbf{n}^i + 1, \dots, \mathbf{n}^J, \mathbf{s} + \mathcal{L}_t^i - 1), & \text{with probability } p_i(\mathbf{n}_i - \mathbf{n}^i) \\ \vdots \\ (\mathbf{n}^1, \dots, \dots, \mathbf{n}^J + 1, \mathbf{s} + \mathcal{L}_t^J - 1), & \text{with probability } p_J(\mathbf{n}_J - \mathbf{n}^J) \end{cases} \quad (2)$$

as long as  $\mathbf{s} \neq 0$ . In reality, the process stops if  $\mathbf{s} = 0$ . For technical reasons, we have, however, to assume that it continues, but its values are void of meaning then. Remark that  $\mathbf{N}_t = (\mathbf{N}_t^1, \mathbf{N}_t^2, \dots, \mathbf{N}_t^J)$  is a Markov process itself, the so-called *coupon collector's process*, where the coupons are chosen with unequal probabilities. We denote by  $\mathcal{N}_t = \sum_{i=1}^J \mathbf{N}_t^i$  the total number of vertices infected - or collected coupons - by time  $t$ .

The sequence of the revealed spread capacities  $(\mathcal{L}_t)_{t \in \mathbb{N}} = ((\mathcal{L}_t^1, \dots, \mathcal{L}_t^J))_{t \in \mathbb{N}}$  is an i.i.d. copy of the random variable  $\mathcal{L} = (\mathcal{L}^1, \dots, \mathcal{L}^J)$  as introduced at the beginning of this section. Since  $n$  is finite, and each vertex has a finite spread capacity, this process will terminate in finite time  $\tau := \tau_n$ ,

$$\tau_n := \min\{t \in \mathbb{N} : \mathbf{S}_t = 0\} \quad (3)$$

either when all vertices were reached and exhausted their own capacity, or when there is no more capacity to be spread even though some of the vertices were not visited at all. We will be interested in the asymptotic behavior of  $\tau_n$ , and in the total number of infected vertices at exhaustion  $\sum_{i=1}^J \mathbf{N}_{\tau_n}^i$ . It suffices to understand the law of the random vector  $(\mathbf{N}_{\tau_n}^1, \mathbf{N}_{\tau_n}^2, \dots, \mathbf{N}_{\tau_n}^J)$  in order to obtain the limit behavior for  $\mathcal{N}_{\tau_n} = \sum_{i=1}^J \mathbf{N}_{\tau_n}^i$ .

**Acquisition and inter-arrival times.** In order to have a better understanding of the joint limit behavior of the process  $\mathbf{N}_t = (\mathbf{N}_t^1, \mathbf{N}_t^2, \dots, \mathbf{N}_t^J)$ , where for  $i \in [J]$ ,  $\mathbf{N}_t^i$  represents the number of vertices in  $\mathbb{K}_n$  of type  $i$  that have been infected by time  $t$ , we will also introduce *acquisition times*  $T_k^i$  and *inter-arrival times*  $\Delta_k^i$ : for  $i \in [J]$  and  $k = 1, \dots, n_i$  denote by  $T_k^i$  the time needed to reach the  $k$ -th vertex of type  $i$ , and by  $\Delta_k^i$  the time needed, after the  $(k-1)$ -th vertex of type  $i$  has been collected, to collect the  $k$ -th one, that is

$$T_k^i = \inf\{t \in \mathbb{N} : \mathbf{N}_t^i = k\}$$

$$\Delta_k^i = T_k^i - T_{k-1}^i \text{ or equivalently } T_k^i = \sum_{j=1}^k \Delta_j^i$$

---

and  $N_i^i = \sum_{k=1}^{n_i} \mathbb{1}_{\{T_k^i \leq t\}}$ . While in the case when all vertices have the same type and the same probability of being chosen, the inter-arrival times are independent geometrically distributed random variables, when one has different types of vertices with unequal probabilities of reaching them, we do not have independence anymore.

## 2.2 Multitype Galton-Watson processes

It may be useful to view Multitype Galton-Watson processes as random subtrees of the Ulam-Harris tree that we define now.

**Ulam-Harris tree.** The infinite Ulam-Harris tree  $\mathcal{U}_\infty$  is the infinite rooted tree with vertex set  $V_\infty := \bigcup_{n \in \mathbb{N}_0} \mathbb{N}^n$ , the set of all finite strings or words  $v_1 \cdots v_n$  of positive integers over  $n$  letters, including the empty word  $\emptyset$  which we take to be the root, and with an edge joining  $v_1 \cdots v_n$  and  $v_1 \cdots v_{n+1}$  for any  $n \in \mathbb{N}_0$  and any  $v_1, \dots, v_{n+1} \in \mathbb{N}$ . Thus every vertex  $v = v_1 \cdots v_n$  has outdegree  $\infty$ , and the children of  $v$  are the words  $v1, v2, \dots$  and we let them have this order so that  $\mathcal{U}_\infty$  becomes an infinite ordered rooted tree. We will identify  $\mathcal{U}_\infty$  with its vertex set  $V_\infty$ , where no confusion arises. For vertices  $v = v_1 \cdots v_n$  we also write  $v = (v_1, \dots, v_n)$ , and if  $u = (u_1, \dots, u_m)$  we write  $uv$  for the concatenation of the words  $u$  and  $v$ , that is  $uv = (u_1, \dots, u_m, v_1, \dots, v_n)$ . The parent of  $v_1 \cdots v_n$  is  $v_1 \cdots v_{n-1}$ . Further, if  $k \leq m$ , we set  $u|_k := u_1 \dots u_k$  for the vertex  $u$  truncated at height  $k$ . Finally, for  $u \in \mathcal{U}_\infty$ , we use the notation  $|u| = n$  for  $u \in \mathbb{N}^n$ , that is  $u$  is a word of length (or height)  $n$ , that is, at distance  $n$  from the root  $\emptyset$ . The family  $\mathcal{T}$  of ordered rooted trees can be identified with the set of all subtrees  $\mathbb{T}$  of  $\mathcal{U}_\infty$  that have the property that for all  $v \in V(\mathbb{T})$ :

$$vi \in V(\mathbb{T}) \Rightarrow vj \in V(\mathbb{T}), \quad \text{for all } j \leq i,$$

where  $V(\mathbb{T})$  denotes the set of vertices of  $\mathbb{T}$ . For a tree in  $\mathcal{T}$  which is not rooted at  $\emptyset$ , but at some other vertex  $u \in \mathcal{U}_\infty$ , we write  $\mathbb{T}_u$ . For two vertices  $u, v \in \mathcal{U}_\infty$  we denote by  $d(u, v)$  their graph distance, that is, the length of the shortest path between  $u$  and  $v$ . For trees rooted at  $\emptyset$ , we omit the root and we write only  $\mathbb{T}$ .

For  $\mathcal{T} = \{\text{all ordered rooted subtrees } \mathbb{T} \text{ of } \mathcal{U}_\infty\}$  and  $J \in \mathbb{N}$ , a  $J$ -type tree is a pair  $(\mathbb{T}, t^*)$  where  $\mathbb{T} \in \mathcal{T}$  and  $t^* : \mathbb{T} \rightarrow \{1, \dots, J\}$  is a function defined on the vertices of  $\mathbb{T}$  which returns for each vertex  $v$  its type  $t_{\mathbb{T}}(v)$ . We denote by  $\mathcal{T}^{[J]}$  the set of all  $J$ -type trees, and elements of  $\mathcal{T}^{[J]}$  will be referred to as  $\mathbb{T}$  without explicitly mentioning the function  $t^*$ .

A multitype Galton-Watson processes (MGW), called also multitype branching process, is a natural generalization of a Galton-Watson process, where a finite number of distinguishable types of particles with different probabilistic behavior are allowed. We use again  $J \in \mathbb{N}$  for the number of particle types, and the particle types will be the same as the  $J$  different types of vertices in the complete graph  $\mathbb{K}_n$  and will be denoted by  $\{1, \dots, J\}$ . In order to relate the infection model with multitype Galton-Watson processes, we proceed as follows. We consider the sequence  $(L^{(i)})_{i \in [J]}$  of  $J$  independent random (row) vectors in  $\mathbb{R}^J$ , with entries  $L^{(i)} = (L^{(i,1)}, \dots, L^{(i,J)})$ , whose distribution is given by (1), and with these  $J$  vectors we build our Galton-Watson tree. The vector  $L^{(i)}$  represents the offspring distribution vector of a vertex of type  $j$ , meaning that for  $j \in [J]$ , the entry  $L^{(i,j)}$  represents the number of

offspring of type  $j$  produced by a vertex of type  $i$ . Let now  $L$  be the  $J \times J$  random matrix whose rows are the vectors  $L^{(i)}$ :

$$L = \left( L^{(1)}, L^{(2)}, \dots, L^{(J)} \right) = \begin{pmatrix} L^{(1,1)} & L^{(1,2)} & \dots & \dots & L^{(1,J)} \\ L^{(2,1)} & L^{(2,2)} & \dots & \dots & L^{(2,J)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ L^{(J,1)} & L^{(J,2)} & \dots & \dots & L^{(J,J)} \end{pmatrix}$$

Starting with the random matrix  $L$  as introduced above, we can now define multitype Galton-Watson trees as  $\mathcal{T}^{[J]}$ -valued random variables, where the type function  $t^*$  is random and defined in terms of the  $J \times J$ -random matrix  $L$ . Let  $(L(u))_{u \in \mathcal{U}_\infty}$  be a family of i.i.d. copies of  $L$ , so  $(\mathcal{L}(u))_{u \in \mathcal{U}_\infty}$  is a family of i.i.d. copies of the random vector  $\mathcal{L} = (\mathcal{L}^1, \dots, \mathcal{L}^J)$  that has independent entries. For any  $i_0 \in [J]$ , we define the random labeled tree  $\mathbb{T}^{\text{MGW}} \in \mathcal{T}^{[J]}$  rooted at  $\emptyset$ , with the associated type function  $t^* : \mathbb{T}^{\text{MGW}} \rightarrow \{1, \dots, J\}$  defined recursively as follows:

$$\emptyset \in \mathbb{T}^{\text{MGW}} \quad \text{and} \quad t^*(\emptyset) = i_0.$$

Now suppose that  $u = u_1 \dots u_m \in \mathbb{T}^{\text{MGW}}$  with  $t^*(u) = i$ , for some  $i \in [J]$ . Then

$$u_1 \dots u_m k \in \mathbb{T}^{\text{MGW}} \quad \text{iff} \quad k \leq L^{(i,1)}(u) + \dots + L^{(i,J)}(u) = \mathcal{L}^i(u)$$

and we assume

$$|\{k \in \mathbb{N} \mid t^*(u_1 \dots u_m k) = j\}| = L^{(i,j)}(u) \quad \text{for all } j \in [J].$$

The multitype branching process  $Z_t = (Z_t^1, \dots, Z_t^J)$  associated with  $(\mathbb{T}^{\text{MGW}}, t^*)$ , and starting from a single particle of type  $i_0 \in [J]$  at the root  $\emptyset$ , that is  $t^*(\emptyset) = i_0$ , is defined as:  $Z_0 = e_{i_0}$  and for  $t \geq 1$

$$Z_t^i := \#\{u \in \mathbb{T}^{\text{MGW}} : |u| = t \text{ and } t^*(u) = i\} = \sum_{v \in \mathbb{T}^{\text{MGW}} : |v| = t-1} L^{(t^*(v), i)}(v), \quad \text{for } i \in [J],$$

that is,  $Z_t^i$  represents the number of particles of type  $i$  in the  $t$ -th generation, or more precisely the number of vertices  $u \in \mathbb{T}^{i_0}$  with  $|u| = t$  and  $t^*(u) = i$ . When referring to multitype Galton-Watson processes we shall always have in mind both  $(Z_t)_{t \in \mathbb{N}}$  and its genealogical Galton-Watson tree  $\mathbb{T}^{\text{MGW}}$ , with the corresponding type function  $t^*$ , i.e. the pair  $(\mathbb{T}^{\text{MGW}}, t^*)$ .

Denote by  $m_{ij} = \mathbb{E}[L^{(i,j)}]$  the expectation of the random variable  $L^{(i,j)}$ , for all  $i, j \in [J]$  and by  $M = (m_{ij})_{i,j \in [J]}$  the first moment or the mean offspring matrix. We have  $\mathbb{E}L = M$ . Also, for  $t \in \mathbb{N}$  we denote by  $Z_t$  the total number of offspring in the  $t$ -th generation, that is  $Z_t = \sum_{i=1}^J Z_t^i$ .

If there exists  $n \in \mathbb{N}$  such that all entries of  $M^n$  are strictly positive, then  $(Z_t)_{t \in \mathbb{N}}$  is called *positive regular*. If each particle has exactly one child, then  $(Z_t)_{t \in \mathbb{N}}$  is called *singular*. It is well known that if  $(Z_t)_{t \in \mathbb{N}}$  is positive regular and not singular, then if the spectral radius



---

$\rho(M) \leq 1$ , the MGW process  $(Z_t)_{t \in \mathbb{N}}$  dies out almost surely. If  $\rho(M) > 1$ , then it survives with positive probability  $\mathbb{P}[\text{Surv}^{\text{MGW}}]$  and we will denote by  $\sigma^{\text{MGW}} = 1 - \mathbb{P}[\text{Surv}^{\text{MGW}}]$  its complement, where

$$\text{Surv}^{\text{MGW}} = \left\{ Z_t = \sum_{i=1}^J Z_t^i > 0, \forall t \in \mathbb{N} \right\},$$

so

$$\sigma^{\text{MGW}} = 1 - \mathbb{P}[\text{Surv}^{\text{MGW}}] \begin{cases} = 1 & , \text{ if } \rho(M) \leq 1 \\ < 1 & , \text{ if } \rho(M) > 1. \end{cases}$$

Note that in the case of multitype Galton-Watson trees  $(\mathbb{T}^{\text{MGW}}, t^*)$ , the type function  $t^*$  is random and constructed with the help of the offspring distribution matrix  $L$ , while in the infection model the type function  $t$  is deterministic, and each vertex in  $\mathbb{K}_n$  comes at the beginning with its own type.

**Assumptions.** Our results will be proven under the following assumptions concerning the proportions and the weights of vertices of type  $i$ , and on the spectral radius  $\rho(M)$ .

**Assumption 1:** For every  $i \in [J]$ , both the proportions  $\gamma_i := \frac{n_i}{n} \in (0, 1)$  of vertices of type  $i$  and the weights  $\alpha_i := n_i p_i \in (0, 1)$  of type  $i$  group do not depend on  $n$ , thus are constant. We have  $\sum_{i=1}^J \gamma_i = 1$ .

**Assumption 2:**  $\rho(M) > 1$ .

### 3 Coupling coupons collector's with Poisson processes

Our proof strategy involves approximating the discrete-time process of collecting coupons (vertices in our case) by a continuous-time process. Instead of the coupon collector drawing a random coupon at integer time points, he draws a random coupon (from the same distribution) at times given by a Poisson process with parameter 1. The times at which any given vertex is drawn is then a thinned Poisson process, and the Poisson processes associated with different vertices are independent. The times at which any particular vertex  $v \in \mathbb{K}_n$  is reached will also be a thinned Poisson process. Working in the continuous rather than in the discrete setting simplifies calculations.

Suppose we are collecting coupons at times chosen according to a Poisson process with rate 1: let  $(\mathcal{P}_t)_{t \geq 0}$  be a Poisson process with rate  $1 = \sum_{i=1}^J n_i p_i$ , so that each Poisson event associated with this process is a sampled coupon (infected individual). Thus  $\mathcal{P}_t$  represents the total number of attempts to infect vertices by time  $t$  (out of these vertices, some of them may have been infected more than once, and we are only interested in the number of different ones).

We denote the number of vertices of type  $i$  infected by time  $t$  by  $\tilde{N}_t^i$ , when we refer to the continuous time process, while the notation  $N_t^i$  is kept for the number of vertices of type

$i$  infected up to time  $t$  in the discrete time process. The relation between the continuous versus discrete processes can be formalized in the following way:

$$\tilde{\mathbf{N}}_t^i := \mathbf{N}_{\mathcal{P}_t}^i \quad \text{and} \quad \tilde{\mathcal{N}}_t := \mathcal{N}_{\mathcal{P}_t} \quad \text{and} \quad \tilde{\mathbf{S}}_t := \mathbf{S}_{\mathcal{P}_t}.$$

We also use the notation  $\tilde{\mathbf{N}}_t = (\tilde{\mathbf{N}}_t^1, \dots, \tilde{\mathbf{N}}_t^J)$ . At times, it will be useful to enumerate the vertices of  $\mathbb{K}_n$  as  $\{v_1, v_2, \dots, v_n\}$ . With each vertex  $v_k \in \mathbb{K}_n$ ,  $k = 1, \dots, n$  we associate a Poisson process  $\mathcal{P}_t(k)$  which represents the number of times  $v_k$  has been collected (or contacted by an infected individual) by time  $t$ . Then  $(\mathcal{P}_t(k))_{t \geq 0}$  are independent Poisson processes with rates  $p_{t(v_k)}$ . More precisely  $\{(\mathcal{P}_t(k))_{t \geq 0} : v_k \in \mathbb{K}_n, t(v_k) = i\}$  is a sequence of  $n_i$  independent Poisson processes with rate  $p_i$ , for any  $i \in [J]$ . Therefore, we can write  $\mathcal{P}_t = \sum_{k=1}^n \mathcal{P}_t(k)$ . For each  $k = 1, \dots, n$  denote by  $E_k$  the time of the first event in the Poisson process  $(\mathcal{P}_t(k))_{t \geq 0}$  associated with  $v_k \in \mathbb{K}_n$ , that is, the first time when vertex  $v_k$  was reached and infected. The random variables  $(E_k)_{k=1, \dots, n}$  are independent (because they are associated to independent Poisson processes) and exponentially distributed with rate  $p_{t(v_k)}$ . We are interested only in the number of different coupons sampled, that is, only in the realization time of the first event in the Poisson processes mentioned above. With this notation and in view of the independence, we have now an alternative way of writing  $\tilde{\mathbf{N}}_t^i$  as sum of i.i.d. random variables: for  $i \in [J]$  we have

$$\tilde{\mathbf{N}}_t^i = \sum_{v_k \in \mathbb{K}_n; t(v_k)=i} \mathbb{1}_{\{E_k \leq t\}}, \quad \text{and} \quad E_k \sim \text{Exp}(p_i)$$

and hence

$$\mathbb{E}[\tilde{\mathbf{N}}_t^i] = n_i \mathbb{P}[E_k \leq t] = n_i(1 - e^{-p_i t}),$$

and

$$\text{Var}[\tilde{\mathbf{N}}_t^i] = n_i [(1 - e^{-p_i t}) - (1 - e^{-p_i t})^2] = n_i e^{-p_i t} (1 - e^{-p_i t}).$$

With this coupling, we are ready to proceed with limit theorems for the processes  $(\tilde{\mathbf{N}}_{n_1 s}^1, \dots, \tilde{\mathbf{N}}_{n_J s}^J)_{s \geq 0}$  and the corresponding acquisition times  $(\tilde{T}_{[n_1 q]}^1, \dots, \tilde{T}_{[n_J q]}^J)_{q \in [0,1]}$  as  $n \rightarrow \infty$ . In both random vectors, each random variable depends on  $n$ , but we will not write the dependence explicitly in order to keep the notation simpler.

**Continuous time-approximation of the coupon collector.** The Poisson process (the continuous coupon collector)  $\mathcal{P}_t$  receives a coupon each time an event occurs in  $(\mathcal{P}_t)_{t \geq 0}$ . List the coupons in the order they are received by the Poisson process as  $C_1, C_2, \dots$ . The distribution of the sequence  $C = (C_n)_{n \in \mathbb{N}}$  is identical to that of the sequence of coupons received by the (discrete-time) coupon collector  $(\mathbf{N}_t)_{t \in \mathbb{N}}$ . Furthermore, the time  $t_m$  at which the continuous coupon collector (Poisson process) receives his  $m$ -th coupon is sharply concentrated around  $m$ . Indeed, by a standard bound on the Poisson distribution, for any  $\epsilon > 0$

$$\mathbb{P}[|t_m - m| \geq m\epsilon] = O\left(\frac{1}{\sqrt{m\epsilon^2}} e^{-m\epsilon^2/2}\right).$$

In particular, since  $\tau_n$  grows with  $n$ , we have with high probability  $t_{\tau_n} = (1 + o(1))\tau_n$ . If we denote by

$$\tilde{\tau}_n := \inf\{t \in \mathbb{R} \mid \mathcal{P}_t \geq \tau_n\} = t_{\tau_n}$$

the total duration of the Poisson process, then  $\tau_n = \mathcal{P}_{\tilde{\tau}_n}$ .

### 3.1 Limit theorems for infected individuals and acquisition times

For simplicity, we introduce some additional notation.

- We will write  $\mathbf{p} = (p_1, \dots, p_J)$  and  $\gamma := (\gamma_1, \dots, \gamma_J)$ .
- We denote by  $\beta_i = \frac{\alpha_i}{\gamma_i} = np_i$ .
- We write  $w_i(s) := \gamma_i(1 - e^{-\beta_i s})$ , and  $\mathbf{w}(s) := (w_1(s), \dots, w_J(s))$ , for  $s \in \mathbb{R}$ , so we have

$$\langle \mathbf{w}(s), \mathbf{1} \rangle = \sum_{i=1}^J \gamma_i (1 - e^{-\beta_i s}) = 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i s}, \quad (4)$$

where  $\mathbf{1}$  is the all ones vector.

**Proposition 3.1.** *Suppose that Assumption 1 holds. For every  $i \in [J]$ , for the coupon's collector process  $(\tilde{\mathbf{N}}_t^i)_{t \in \mathbb{N}}$  we have the following weak convergence in the Skorohod space  $\mathcal{D}([0, \infty), \mathbb{R})$  endowed with the standard  $\mathbf{J}_1$  topology:*

$$\left\{ \frac{1}{\sqrt{n_i}} \left( \tilde{\mathbf{N}}_{ns}^i - n_i (1 - e^{-\beta_i s}) \right); s \geq 0 \right\} \xrightarrow{law} \left\{ \mathbf{X}_s^{(i)}; s \geq 0 \right\}, \quad \text{as } n \rightarrow \infty.$$

where, for every  $i \in [J]$ ,  $(\mathbf{X}_s^{(i)})_{s \geq 0}$  is a centered Gaussian process with

$$\text{Cov} \left[ \mathbf{X}_s^{(i)}, \mathbf{X}_t^{(i)} \right] = e^{-\beta_i t} (1 - e^{-\beta_i s}) > 0,$$

for all  $s, t \in [0, \infty)$  with  $s \leq t$ .

*Proof.* The claim follows from [2, Theorem 14.3], since  $\tilde{\mathbf{N}}_t^i = \sum_{k=1}^{n_i} \mathbf{1}_{\{E_k \leq t\}}$  for i.i.d. random variables  $E_k \sim \text{Exp}(p_i)$ ,  $k = 1, \dots, n_i$ .  $\square$

Since with the Poissonization technique, the processes  $(\tilde{\mathbf{N}}_t^i)_{t \in \mathbb{N}}$  are independent (because they are measurable functions of independent Poisson processes) we also have their joint convergence as noticed in the following proposition.

**Proposition 3.2.** *Suppose that Assumption 1 holds. For the (continuous) coupon's collector process  $(\tilde{\mathbf{N}}_t)_{t \in \mathbb{N}}$  with  $J$  different types of coupons and  $\tilde{\mathbf{N}}_t = (\tilde{\mathbf{N}}_t^1, \tilde{\mathbf{N}}_t^2, \dots, \tilde{\mathbf{N}}_t^J)$ , we have the following (joint) weak convergence in the Skorohod space  $\mathcal{D}([0, \infty), \mathbb{R}^J)$  endowed with the standard  $\mathbf{J}_1$  topology:*

$$\left\{ \frac{1}{\sqrt{n}} \left( \vec{\mathbf{N}}_{ns} - n \mathbf{E}_s \right); s \geq 0 \right\} \xrightarrow{law} \left\{ \mathbf{X}_s; s \geq 0 \right\}$$

where the  $J$ -dimensional processes above are defined as

$$\vec{\mathbf{N}}_{ns} := \left( \frac{\tilde{\mathbf{N}}_{ns}^1}{\sqrt{\gamma_1}}, \dots, \frac{\tilde{\mathbf{N}}_{ns}^J}{\sqrt{\gamma_J}} \right).$$

The expectation vector  $\mathbf{E}_s$  is given by  $\mathbf{E}_s := (\sqrt{\gamma_1}(1 - e^{-\beta_1 s}), \dots, \sqrt{\gamma_J}(1 - e^{-\beta_J s}))$ , and the  $J$ -dimensional stochastic process  $(\mathbf{X}_s)_{s \geq 0}$  with  $\mathbf{X}_s = (\mathbf{X}_s^{(1)}, \dots, \mathbf{X}_s^{(J)})$  has independent components  $\mathbf{X}_s^{(1)}, \dots, \mathbf{X}_s^{(J)}$  which are centered Gaussian processes with covariances

$$\text{Cov}[\mathbf{X}_s^{(i)}, \mathbf{X}_t^{(i)}] = e^{-\beta_i t} (1 - e^{-\beta_i s}) > 0, \quad \text{for all } s, t \in [0, \infty) \text{ with } s \leq t.$$

Since  $\tilde{\mathcal{N}}_t = \sum_{i=1}^J \tilde{\mathcal{N}}_t^i$ , an immediate corollary of Proposition 3.2 is the following.

**Corollary 3.3.** *We have the following weak convergence in the Skorohod space  $\mathcal{D}([0, \infty), \mathbb{R})$  endowed with the standard  $\mathbf{J}_1$  topology:*

$$\left\{ \frac{1}{\sqrt{n}} \left( \tilde{\mathcal{N}}_{ns} - n \langle \mathbf{w}(s), \mathbf{1} \rangle \right); s \geq 0 \right\} \xrightarrow{\text{law}} \{ \mathbf{X}_s^{\mathcal{N}}; s \geq 0 \}$$

where  $\mathbf{X}_s^{\mathcal{N}}$  is a centered Gaussian processes with

$$\text{Cov}[\mathbf{X}_s^{\mathcal{N}}, \mathbf{X}_t^{\mathcal{N}}] = \sum_{i=1}^J \gamma_i \cdot (1 - e^{-\beta_i s}) \cdot e^{-\beta_i t}, \quad \text{for all } s, t \in [0, \infty) \text{ with } s \leq t.$$

For the vector of acquisition times processes  $(\tilde{T}_{[n_1 q]}^1, \dots, \tilde{T}_{[n_J q]}^J)_{q \in [0, 1]}$  of the continuous time process we obtain a similar behavior. Recall that by  $\tilde{T}_{n_i q}^i := \tilde{T}_{[n_i q]}^i$  for  $i \in [J]$ , we denote the acquisition time of the  $[n_i q]$ -th coupon of type  $i$ .

**Proposition 3.4.** *Suppose that Assumption 1 holds. For every  $i \in [J]$ , for the acquisition times  $\tilde{T}_k^i$  associated with the process  $\tilde{\mathcal{N}}_t^i$  of collecting vertices of type  $i$ , we have the following weak convergence in the Skorohod space  $\mathcal{D}([0, 1], \mathbb{R})$  endowed with the standard  $\mathbf{J}_1$  topology:*

$$\left\{ \sqrt{n_i} \left( \frac{1}{n} \tilde{T}_{n_i q}^i - \frac{1}{\beta_i} \log \frac{1}{1 - q} \right); q \in [0, 1) \right\} \longrightarrow \{ \mathbf{B}^i(\sigma^i(q)); q \in [0, 1) \}$$

where, for  $i \in [J]$ ,  $(\mathbf{B}^i(\sigma^i(q)))_{q \in [0, 1)}$  is a standard Brownian motion with  $\sigma^i(q) = \frac{q}{\beta_i^2(1-q)}$ .

*Proof.* We have the following reciprocal equations that relate the distribution of the number of collected coupons  $\tilde{\mathcal{N}}_t^i$  to the distribution of the arrival times  $\tilde{T}_s^i$  of the respective coupons: for every  $i = 1, 2, \dots, J$

$$\{\tilde{\mathcal{N}}_t^i < s\} = \{\tilde{T}_s^i > t\},$$

that is  $\tilde{T}_s^i$  is a generalized inverse of  $\tilde{\mathcal{N}}_t^i$ . For  $i \in [J]$ , we write

$$f_i(q) := \frac{1}{\beta_i} \cdot \log \frac{1}{1 - q}, \quad q \in [0, 1), \quad \text{and} \quad f_i^{-1}(s) := 1 - e^{-\beta_i s}, \quad s \in [0, \infty),$$

so that  $f_i^{-1}(s)$  is the inverse of  $f_i(s)$ . We apply a Taylor series expansion to  $f_i$  in  $\frac{\tilde{\mathcal{N}}_{ns}^i}{n_i}$ . Indeed, there is  $q_s \in [\min\{\tilde{\mathcal{N}}_{ns}^i/n_i, f_i^{-1}(s)\}, \max\{\tilde{\mathcal{N}}_{ns}^i/n_i, f_i^{-1}(s)\}]$  such that

$$f_i \left( \frac{\tilde{\mathcal{N}}_{ns}^i}{n_i} \right) - s = \underbrace{f_i'(f_i^{-1}(s))}_{= \frac{-1}{\beta_i e^{-\beta_i s}}} \cdot \left( \frac{\tilde{\mathcal{N}}_{ns}^i}{n_i} - f_i^{-1}(s) \right) + \frac{1}{2} f_i''(q_s) \cdot \left( \frac{\tilde{\mathcal{N}}_{ns}^i}{n_i} - f_i^{-1}(s) \right)^2. \quad (5)$$

If we show that for every fixed  $i \in [J]$

$$\sup_{s \in C} f_i''(q_s) \cdot \left( \frac{\tilde{\mathbf{N}}_{ns}^i}{n_i} - f_i^{-1}(s) \right)^2 \longrightarrow 0, \quad \text{in probability, as } n \rightarrow \infty \quad (6)$$

for every compact set  $C \subseteq [0, \infty)$ , then according to [8, Proposition VI.1.17 b)]

$$f_i''(q_s) \cdot \left( \frac{\tilde{\mathbf{N}}_{ns}^i}{n_i} - f_i^{-1}(s) \right)^2 \longrightarrow 0$$

in law in the Skorohod topology as  $n \rightarrow \infty$ . So let  $C \subseteq [0, \infty)$  be compact, and denote by  $s_0$  the supremum of  $C$ . Choose  $q_1 \in (f_i^{-1}(s_0), 1)$ . From Proposition 3.1 we also have that  $\frac{\tilde{\mathbf{N}}_{ns}^i}{n_i} \rightarrow f_i^{-1}(s)$  in probability as  $n \rightarrow \infty$ , uniformly on compact sets of  $\mathbb{R}_+$ , so

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{\tilde{\mathbf{N}}_{ns_0}^i}{n_i} > q_1 \right] = 0.$$

Put  $M := f_i''(q_{s_0}) = \sup_{s \in C} f_i''(q_s)$ . Observe that  $f_i''(q) = \frac{1}{\beta_i(1-q)^2} > 0$  is monotonically increasing on  $[0, 1)$ . Then

$$\mathbb{P} [M > f_i''(q_1)] \leq \mathbb{P} \left[ \frac{\tilde{\mathbf{N}}_{ns_0}^i}{n_i} > q_1 \right] \longrightarrow 0,$$

i.e.  $M$  is bounded in probability. Since the operator which maps a function  $x : [0, \infty) \rightarrow \mathbb{R}$  to  $\sup_{t \in C} x(t)$  is continuous at continuous functions  $x : [0, \infty) \rightarrow \mathbb{R}$ , the continuous mapping theorem in its sharp version (see e.g. [2, Theorem 2.7]) implies (6). From (5) it follows that  $\left( f_i \left( \frac{\tilde{\mathbf{N}}_{ns}^i}{n_i} \right) - s \right)$  and  $f_i'(f_i^{-1}(s)) \cdot \left( \frac{\tilde{\mathbf{N}}_{ns}^i}{n_i} - f_i^{-1}(s) \right)$  have the same limit law, since the remainder in the expansion converges to 0 in probability, and so Proposition 3.1 yields

$$\sqrt{n_i} \left( f_i \left( \frac{\tilde{\mathbf{N}}_{ns}^i}{n_i} \right) - s \right) \stackrel{\text{law}}{\sim} f_i'(f_i^{-1}(s)) \underbrace{\frac{1}{\sqrt{n_i}} \left( \tilde{\mathbf{N}}_{ns}^i - n_i f_i^{-1}(s) \right)}_{\xrightarrow{\text{law}} \mathbf{X}_s^{(i)}} \stackrel{\text{law}}{\rightarrow} f_i'(f_i^{-1}(s)) \mathbf{X}_s^{(i)}.$$

By [12, Theorem 1] this is equivalent to

$$\left\{ \sqrt{n_i} \left( \frac{1}{n} \tilde{T}_{n_i f_i^{-1}(t)}^i - t \right); t \geq 0 \right\} \longrightarrow - \left\{ f_i'(f_i^{-1}(t)) \cdot \mathbf{X}_t^{(i)}; t \geq 0 \right\},$$

which, in turn is equivalent to

$$\left\{ \sqrt{n_i} \left( \frac{1}{n} \tilde{T}_{n_i q}^i - f_i(q) \right); q \in [0, 1) \right\} \longrightarrow - \left\{ f_i'(q) \cdot \mathbf{X}_{f_i(q)}^{(i)}; q \in [0, 1) \right\}.$$

Observe now that, for  $p, q \in [0, 1)$  with  $p \leq q$

$$\begin{aligned} \text{Cov} \left[ -f_i'(p) \cdot \mathbf{X}_{f_i(p)}^{(i)}, -f_i'(q) \cdot \mathbf{X}_{f_i(q)}^{(i)} \right] &= f_i'(p) \cdot f_i'(q) \cdot f_i^{-1}(f_i(p)) \cdot [1 - f_i^{-1}(f_i(q))] \\ &= \frac{p}{\beta_i^2(1-p)} \end{aligned}$$

and hence

$$\begin{aligned} - \left\{ f'_i(q) \cdot \mathbf{X}_{f_i(q)}^{(i)}; q \in [0, 1] \right\} &= - \left\{ \frac{1}{\beta_i(1-q)} \cdot \mathbf{X}_{-\log(1-q)/\beta_i}^{(i)}; q \in [0, 1] \right\} \\ &= \left\{ \mathbf{B}^i(\sigma^i(q)); q \in [0, 1] \right\}, \end{aligned}$$

in law in the Skorohod space, with  $\sigma^i(q) = \frac{q}{\beta_i^2(1-q)}$ . Therefore

$$\left\{ \sqrt{n_i} \left( \frac{1}{n} \tilde{T}_{n_i q}^i - \frac{1}{\beta_i} \log \frac{1}{1-q} \right); q \in [0, 1] \right\} \longrightarrow \left\{ \mathbf{B}^i(\sigma^i(q)); q \in [0, 1] \right\}$$

and this proves the claim.  $\square$

Since for each  $i \in [J]$ , the random variables  $(\tilde{T}_k^i)$  are measurable functions of  $\tilde{\mathbf{N}}_k^i$ , and  $\tilde{\mathbf{N}}_t^1, \dots, \tilde{\mathbf{N}}_t^J$  are independent, the random variables  $\tilde{T}_{n_1}^1, \dots, \tilde{T}_{n_J}^J$  are independent as well, so we also have their joint convergence.

**Proposition 3.5.** *Suppose that Assumption 1 hold. For the continuous time acquisition process  $(\tilde{T}_k)_{k \in \mathbb{N}}$  with  $\tilde{T}_k = (\tilde{T}_k^1, \dots, \tilde{T}_k^J)$ , we have the following (joint) weak convergence in the Skorohod space  $\mathcal{D}([0, 1], \mathbb{R})$  endowed with the standard  $\mathbf{J}_1$  topology:*

$$\left\{ \left( \frac{1}{\sqrt{n}} \vec{T}_{nq} - \sqrt{n} \log \frac{1}{1-q} \mathbf{E}[\tilde{T}_{nq}] \right); q \in [0, 1] \right\} \xrightarrow{law} \left\{ \mathbf{B}(\sigma(q)); q \in [0, 1] \right\}$$

where the  $J$ -dimensional processes above are defined as following:

$$\vec{T}_{nq} := \left( \sqrt{\gamma_1} \tilde{T}_{n_i q}^1, \dots, \sqrt{\gamma_J} \tilde{T}_{n_j q}^J \right),$$

the expectation vector  $\mathbf{E}$  is given by

$$\mathbf{E}[\tilde{T}_{nq}] := \left( \frac{\sqrt{\gamma_1}}{\beta_1}, \dots, \frac{\sqrt{\gamma_J}}{\beta_J} \right),$$

and  $\mathbf{B}(\sigma(q)) := (\mathbf{B}^1(\sigma^1(q)), \dots, \mathbf{B}^J(\sigma^J(q)))$  is a  $J$ -dimensional standard Brownian motion, and  $\mathbf{B}^i(\sigma^i(q))$  are independent for  $i \in [J]$  with  $\sigma^i(q) = \frac{q}{\beta_i^2(1-q)}$ .

## 4 Virus spread model as a random subtree of a MGW

We will now show how to obtain a virus spread model by coupling a coupon collector with a multitype Galton-Watson process (MGW) with  $J$  types. We will cut from the multitype Galton-Watson tree  $\mathbb{T}^{\text{MGW}}$  nodes that correspond to vertices in  $\mathbb{K}_n$  that are reached by more than one random walker. We will construct, on a joint probability space  $(\Omega, \mathcal{F}, \mathbb{P})$

- a multitype Galton-Watson process with offspring distribution matrix  $L$ , where the row sums are  $(\mathcal{L}^1, \dots, \mathcal{L}^J)$ , and  $J \in \mathbb{N}$  is the number of types,
- for each integer  $n$ , a coupon collector process with  $J$  different types of coupons,

- for each integer  $n$ , sequences  $(\mathbf{S}_t)_{t \in \mathbb{N}}$  and  $(\mathbf{N}_t)_{t \in \mathbb{N}} = ((\mathbf{N}_t^1, \dots, \mathbf{N}_t^J))_{t \in \mathbb{N}}$ , with  $\mathbf{N}_0 = \mathbf{e}_{i_0}$  for a fixed  $i_0 \in [J]$ ,  $\mathbf{S}_0 \sim \mathcal{L}^{i_0}$ , and such that  $(\mathbf{N}_t, \mathbf{S}_t)_{t \in \mathbb{N}}$  is a Markov chain with transition probabilities as in (2).

Notice that the situation here is slightly different from the situation in [4]. In [4] the Galton-Watson process and the coupon collector were assumed to be independent and then the sequence  $(\mathbf{N}_t, \mathbf{S}_t)_{t \in \mathbb{N}}$  was constructed out of them (by pruning the Galton-Watson tree). This is not possible anymore here: the types of the Galton-Watson process and the coupon collector have to match and thus we cannot assume that these processes are independent of each other.

In order to construct the processes mentioned above, we start with a multitype Galton-Watson tree  $(\mathbb{T}^{\text{MGW}}, \mathbf{t}^*)$  where  $\mathbf{t}^*$  is the type function in the tree rooted at  $\emptyset$  with  $\mathbf{t}^*(\emptyset) = i_0$ , with offspring distribution matrix  $L$ , which has independent rows  $L^{(i)} = (L^{(i,1)}, \dots, L^{(i,J)})$  and distribution given as in (1). Also  $\mathcal{L}^i = \sum_{j=1}^J L^{(i,j)}$  represents the total offspring number of a vertex of type  $i$ , and the random variables  $\mathcal{L}^1, \dots, \mathcal{L}^J$  are independent but not identically distributed. For the following construction it will be crucial that the types of the children  $v_1, \dots, v_l$  of each vertex  $v$  of  $\mathbb{T}^{\text{MGW}}$  have an exchangeable distribution, i.e. the distribution of  $(\mathbf{t}^*(v_{\pi(1)}), \dots, \mathbf{t}^*(v_{\pi(l)}))$  does not depend on the permutation  $\pi : \{1, \dots, l\} \rightarrow \{1, \dots, l\}$ .

Recursively, for  $t = 0, 1, \dots$  we construct  $X(t) \in \mathbb{T}^{\text{MGW}}$  representing the  $t$ -th tentative infection, and disjoint subtrees  $\mathcal{T}^\circ(t), \mathcal{T}^\boxplus(t), \mathcal{T}^\dagger(t) \subset \mathbb{T}^{\text{MGW}}$ . Note that if  $X(t)$  has type  $i$ , then the corresponding infection attempt in  $\mathbb{K}_n$  is at a vertex  $u$  with type  $i$ . Moreover, if this infection attempt is successful (that is,  $u$  was not infected before), then the further infection attempts made by  $u$  are represented by the children of  $X(t)$ . We define

$$\mathcal{T}^\circ(t) = \bigcup_{i=1}^J \underbrace{\{v \in \mathcal{T}^\circ(t) : \mathbf{t}^*(v) = i\}}_{:= \mathcal{T}_i^\circ(t)}$$

representing the set of vertices (written as a partition in vertices of types  $i \in [J]$ , contained in the set  $\mathcal{T}_i^\circ(t)$ ) already infected by time  $t$ . The set  $\mathcal{T}^\boxplus(t)$  is the set of tentative infections scheduled but not performed by time  $t$ , that we write again as a disjoint union of vertices of type  $i$

$$\mathcal{T}^\boxplus(t) = \bigcup_{i=1}^J \underbrace{\{v \in \mathcal{T}^\boxplus(t) : \mathbf{t}^*(v) = i\}}_{:= \mathcal{T}_i^\boxplus(t)}.$$

Finally  $\mathcal{T}^\dagger(t)$  represents the set of failed spreads (the set of failures in the coupon collecting process), i.e. the part that we have to cut out from the MGW tree; we do not partition this part into subsets with different types. At time  $t = 0$ , we start with the following data: for some fixed  $i_0 \in [J]$  and root  $\emptyset$  of type  $i_0$

$$\begin{cases} X(0) &= \emptyset \text{ (the root of the MGW tree) with } t^*(\emptyset) = i_0 \\ \mathcal{T}^\circ(0) &= \{\emptyset\} \text{ where } \mathcal{T}_{i_0}^\circ(0) = \{\emptyset\} \text{ and } \mathcal{T}_i^\circ(0) = \emptyset, \forall i \neq i_0 \\ \mathcal{T}^\boxplus(0) &= \underbrace{\bigcup_{i=1}^J \{v \in \mathbb{T}^{\text{MGW}} : t^*(v) = i; |v| = 1\}}_{:= \mathcal{T}^\boxplus(0)} \\ \mathcal{T}^\dagger(0) &= \emptyset \end{cases}$$

Given the process  $(X(t), \mathcal{T}^\circ(t), \mathcal{T}^\boxplus(t), \mathcal{T}^\dagger(t))$  at time  $t$ , its value at time  $t+1$  can be defined as follows. If  $\mathcal{T}^\boxplus(t)$  is non-empty, let

$$v = X(t+1) = \inf\{w \in \mathcal{T}^\boxplus(t)\},$$

where the vertices of  $\mathbb{T}^{\text{MGW}}$  are sorted according to their generation, within the generation lexicographically. The type  $j_0 = t^*(v)$  determines the type of the vertex  $u \in \mathbb{K}_n$  which is hit by the infection attempt at time  $t+1$ . Each of the  $n_{j_0}$  candidates, i.e. each vertex  $w \in \mathbb{K}_n$  with  $t(w) = j_0$ , has the same probability  $\frac{1}{n_{j_0}}$  of being hit. In particular, the decision, whether this infection attempt is successful will be made at random independently of  $(\mathbb{T}^{\text{MGW}}, t^*)$ . Let  $\mathbf{n}^{j_0} = |\mathcal{T}_{j_0}^\circ(t)|$ .

- With probability  $\frac{\mathbf{n}^{j_0}}{n_{j_0}}$ , the infection attempt fails (resulting from sending the virus to a vertex that was previously infected). Then we set

$$\begin{cases} \mathcal{T}^\circ(t+1) &= \mathcal{T}^\circ(t), \text{ i.e. } \mathcal{T}_i^\circ(t+1) = \mathcal{T}_i^\circ(t) \text{ for all } i \in [J] \\ \mathcal{T}^\boxplus(t+1) &= \mathcal{T}^\boxplus(t) \setminus \{v\} \\ \mathcal{T}^\dagger(t+1) &= \mathcal{T}^\dagger(t) \cup \{v\} \end{cases} \quad (7)$$

which means that at time  $t+1$  the vertex  $v$ , that was on standby, is now put in the tree  $\mathcal{T}^\dagger(t+1)$  with failures, and we have lost one piece of spread capacity without infecting a new vertex. In this case we define  $\mathbf{N}_{t+1} = \mathbf{N}_t, \mathbf{S}_{t+1} = \mathbf{S}_t - 1$ .

- With probability  $\frac{n_{j_0} - \mathbf{n}^{j_0}}{n_{j_0}}$  the infection is successful. Then

$$\begin{cases} \mathcal{T}_{j_0}^\circ(t+1) &= \mathcal{T}_{j_0}^\circ(t) \cup \{v\} \text{ and } \mathcal{T}_j^\circ(t+1) = \mathcal{T}_j^\circ(t), \text{ for all } j \neq j_0 \\ \mathcal{T}^\boxplus(t+1) &= \mathcal{T}^\boxplus(t) \setminus \{v\} \cup \{v_1, \dots, v_{\mathcal{L}^{j_0}(v)}\} \\ \mathcal{T}^\dagger(t+1) &= \mathcal{T}^\dagger(t) \end{cases} \quad (8)$$

where  $v_1, \dots, v_{\mathcal{L}^{j_0}(v)}$  are the children of  $v$  in  $\mathbb{T}^{\text{MGW}}$ . In this case, in the tree  $\mathcal{T}^\circ(t+1)$  of infected vertices, the vertex  $v$  of type  $j_0$  will be added and it will be deleted from the set  $\mathcal{T}^\boxplus$  of stand-by vertices. In addition, to the later tree  $\mathcal{L}^{j_0}(v) = L^{(j_0,1)}(v) + \dots + L^{(j_0,J)}(v)$  offsprings of  $v$  with the corresponding types will be added. We put  $\mathbf{N}_{t+1} = \mathbf{N}_t + \mathbf{e}_{j_0}, \mathbf{S}_{t+1} = \mathbf{S}_t + \mathcal{L}^{j_0} - 1$ .

We proceed with the algorithm as long as  $\mathcal{T}^\boxplus(t)$  is non-empty. If  $\mathcal{T}^\boxplus(t)$  is empty then we have no vertices on standby, and the process stops; so we set  $\tau_n = t$ . The set

$$\mathcal{T}^\circ(t) = \mathcal{T}^\circ(\tau_n) = \bigcup_{i=1}^J \mathcal{T}_i^\circ(\tau_n)$$



is the set of vertices infected during the entire process.

The sequence  $(\mathbf{N}_t^1, \dots, \mathbf{N}_t^J, \mathbf{S}_t)_{t \in \mathbb{N}}$  constructed above, up to time  $\tau_n$ , is a Markov chain with transition probabilities as in (2). This follows from the fact that for any  $n \in \mathbb{N}$ , the type of the vertex  $X(t+1)$  in the above construction is distributed as  $\mathbb{P}[t^*(X(t+1)) = j] = \alpha_j$  for  $j \in [J]$ .  $T_k^i := \inf\{t \in \mathbb{N} : \mathbf{N}_t^i = k\}$ . We also have that  $\tau_n$  is a.s. finite and bounded by

$$\tau_n \leq \max\{T_{n_1}^1, \dots, T_{n_J}^J\} \quad \text{and} \quad |\mathcal{T}^\circ(\tau_n)| = \sum_{i=1}^J |\mathcal{T}_i^\circ(\tau_n)|.$$

Again, for technical reasons and in order to obtain a complete coupon collector model, we extend the sequences  $\mathbf{N}_t$  and  $\mathbf{S}_t$  beyond the end of the epidemic  $\tau_n$ . Hence for  $t > \tau_n$  we choose a type  $\Lambda$  according to  $\mathbb{P}[\Lambda = j] = \alpha_j$  for  $j \in [J]$ . Then we let the collection of the next coupon be successful and put  $\mathbf{N}_{t+1} = \mathbf{N}_t + e_\Lambda$  with probability  $\frac{n_\Lambda - \mathbf{N}_t^\Lambda}{n_\Lambda}$  (independent of anything else) and we let it fail with probability  $\frac{\mathbf{N}_t^\Lambda}{n_\Lambda}$ , in which case we put  $\mathbf{N}_{t+1} = \mathbf{N}_t$ .

Consider now  $J$  i.i.d. (infinite) sequences of vectors, also independent of each other:  $(\tilde{\mathcal{L}}_l^i)_{l \geq 1}$  with law  $\mathcal{L}^i$ ,  $i = 1, 2, \dots, J$  and define for  $l = 1, \dots, n_i$

$$\mathcal{L}_l^i = \begin{cases} \mathcal{L}^i(X(T_l^i)) & \text{if } l \leq |\mathcal{T}_i^\circ(\tau_n)| \\ \tilde{\mathcal{L}}_l^i & \text{if } l > |\mathcal{T}_i^\circ(\tau_n)|. \end{cases} \quad (9)$$

This means that  $\mathcal{L}_l^i$  is the number of children of the vertex which represents the  $l$ -th infected vertex of type  $i$ , if in the infective process we reached at least  $l$  vertices of type  $i$ ; otherwise  $\mathcal{L}_l^i$  is the  $l$ -th element of the sequence  $\tilde{\mathcal{L}}_l^i$ . We also use the following notation: for  $i \in [J]$ , let  $\mathcal{R}_t^i$  be the spread capacity revealed by vertices of type  $i$  up to time  $t$ , that is:

$$\mathcal{R}_t^i = \sum_{l=1}^{\mathbf{N}_t^i} \mathcal{L}_l^i, \quad (10)$$

and as before, we also write  $\tilde{\mathcal{R}}_t^i = \mathcal{R}_{\mathcal{P}_t}^i$ . The total capacity available at time  $t \in \mathbb{N}$  is

$$\mathbf{S}_t = \underbrace{\sum_{l=1}^{\mathbf{N}_t^1} \mathcal{L}_l^1}_{=: \mathcal{R}_t^1} + \dots + \underbrace{\sum_{l=1}^{\mathbf{N}_t^J} \mathcal{L}_l^J}_{=: \mathcal{R}_t^J} - t = \sum_{i=1}^J \mathcal{R}_t^i - t. \quad (11)$$

Note that, by construction, for  $t \leq \tau_n$ , we have:

$$\mathcal{N}_t = \sum_{i=1}^J \mathbf{N}_t^i = |\mathcal{T}^\circ(t)| = \sum_{i=1}^J \underbrace{|\mathcal{T}_i^\circ(t)|}_{=: \mathbf{N}_t^i} \quad \text{and} \quad \mathbf{S}_t = |\mathcal{T}^\boxplus(t)|.$$

Remark that only  $\mathcal{T}^\circ(t)$  is a connected random tree, as subtree of  $\mathbb{T}^{\text{MGW}}$ , but  $\mathcal{T}_i^\circ(t)$ ,  $i \in [J]$  are not connected, they are only subsets of vertices of  $\mathcal{T}^\circ(t)$ .

**Proposition 4.1.** *For each  $i \in [J]$  and  $l \in [n_i]$  the random variable  $\mathcal{L}_l^i$  has law  $\mathcal{L}^i$ . The random variables  $(\mathcal{L}_l^i)_{1 \leq l \leq n_i, 1 \leq i \leq J}$  are independent (of each other) and also independent of the acquisition times  $(T_l^i)_{1 \leq l \leq n_i, 1 \leq i \leq J}$ . Also*

$$\tau_n = \inf\{t \in \mathbb{N} : \mathbf{S}_t = \mathbf{0}\}.$$

*Proof.* The formula for  $\tau_n$  follows from (11) and from the fact that  $\tau_n$  is finite.

For the other claim, consider the system  $\mathcal{H}$  of functions  $h : \{1, \dots, n\} \rightarrow \{1, \dots, J\}$  such that

$$|\{k \in [n] \mid h(k) = j\}| = n_j, \quad j \in [J].$$

Associate with each  $h \in \mathcal{H}$  a further function

$$\bar{h} : [J] \times [n] \rightarrow [n] \cup \{\infty\}, \quad (i, k) \mapsto \inf\{l \in [n] \mid |\{l' \in [l] \mid h(l') = i\}| \geq k\}$$

so that  $\bar{h}(i, k)$  represents the  $k$ -th time that  $h$  takes the value  $i$ . Furthermore put

$$H_h = \bigcap_{i_1, i_2, l_1, l_2 : \bar{h}(i_1, l_1) < \bar{h}(i_2, l_2)} \left\{ T_{l_1}^{i_1} < T_{l_2}^{i_2} \right\} \cap \bigcap_{i_1, i_2, l_1, l_2 : \bar{h}(i_1, l_1) > \bar{h}(i_2, l_2)} \left\{ T_{l_1}^{i_1} > T_{l_2}^{i_2} \right\},$$

where  $i_1, i_2$  run in  $[J]$  and  $l_1, l_2$  run in  $[n_{i_1}]$  and  $[n_{i_2}]$ , respectively. Write

$$\mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]} = ((T_1^1, \dots, T_{n_1}^1), \dots, (T_1^J, \dots, T_{n_J}^J)).$$

Our aim is to show that

$$\begin{aligned} & \mathbb{P}[\mathcal{L}_l^i = \lambda_l^{(i)} \text{ for all } i \in [J], l \in [n_i], \mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]} \in B_1 \mid H_h] \\ &= \prod_{i=1}^J \prod_{l=1}^{n_i} \mathbb{P}[\mathcal{L}_l^{(i)} = \lambda_l^{(i)}] \cdot \mathbb{P}[\mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]} \in B_1 \mid H_h] \end{aligned} \quad (12)$$

for every  $h \in \mathcal{H}$ ,  $\lambda_l^{(i)} \in \mathbb{N}$ ,  $i \in [J]$ ,  $l \in [n_i]$  and  $B_1 \subseteq \mathbb{N}^n$ . In order to see this put, for a fixed  $h \in \mathcal{H}$ ,

$$K_l := \mathcal{L}_{\bar{h}(l)}^{h(l)} \quad \text{where} \quad \tilde{h}(l) := |\{l' \in [l] \mid h(l) = h(l')\}|$$

counts how often the type of  $h(l)$  has appeared in  $h$  up to time  $l$ . With this notation we are then going to show

$$\mathbb{P} \left[ K_m = \lambda \mid K_1, \dots, K_{m-1}, \mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]} \right] = \mathbb{P} \left[ \mathcal{L}^{h(m)} = \lambda \right]. \quad (13)$$

For this set  $T_l := T_{\tilde{h}(l)}^{h(l)}$ , and for  $0 = k_1 < k_2 < \dots < k_m$ ,  $m \in \mathbb{N}$ ,  $w \in \mathcal{U}_\infty$ , put also

$$A = \{(T_1, \dots, T_m) = (k_1, \dots, k_m)\} \cap \{X(k_{m-1}) = w\} \cap \{T_m \leq \tau_n\}.$$

Then, on  $A$ ,  $X(T_m) = X(k_m)$  is  $\mathcal{F}$ -measurable, where  $\mathcal{F} := \sigma(K_1, \dots, K_{m-1})$ . In order to see this we notice that an easy induction argument yields that

$$(X(t), \mathcal{T}_1^\circ(t), \dots, \mathcal{T}_J^\circ(t), \mathcal{T}^\boxplus(t), \mathcal{T}^\dagger(t))$$

is  $\mathcal{F}$ -measurable for  $t = 1, \dots, k_{m-1}$  on  $A$  in the sense that the event

$$A \cap \{X(t) = w_1\} \cap \bigcap_{i=1}^J \{w_{\nu_{i-1}+1}, \dots, w_{\nu_i} \in \mathcal{T}_i^\circ(t)\} \\ \cap \{w_{\nu_{J+1}}, \dots, w_{\nu_{J+1}} \in \mathcal{T}^\boxplus(t)\} \cap \{w_{\nu_{J+1}+1}, \dots, w_{\nu_{J+2}} \in \mathcal{T}^\dagger(t)\}$$

is measurable w.r.t. the trace- $\sigma$ -algebra of  $\mathcal{F}$  on  $A$  for any numbers  $1 = \nu_0 \leq \nu_1 \leq \dots \leq \nu_{J+2}$  and  $w_1, \dots, w_{\nu_{J+2}} \in \mathcal{U}_\infty$ . In particular,

$$X(k_m) = \inf\{w \in \mathcal{T}^\circ(k_m - 1)\} = \inf\{w \in \bigcup_{i=1}^J \mathcal{T}_i^\circ(k_m - 1)\}$$

is  $\mathcal{F}$ -measurable on  $A$ . Notice that  $X(k_m) > X(k_m - 1) = w$ . Next we will show that the acquisition times  $T_l^i, i \in [J], l \in [n_i]$  are independent of  $K_1, \dots, K_n$ . For this we first observe that on  $\{t \leq \tau_n\}$  the sequence  $(t^*(X(1)), t^*(X(2)), \dots, t^*(X(t)))$  is independent of  $(K_1, \dots, K_n)$ . Indeed,

$$\mathbb{P}[t^*(X(t)) = j \mid t \leq \tau_n, t^*(X(1)), \dots, t^*(X(t-1)), K_1, \dots, K_n] \\ = \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\nu} \sum_{t'=0}^t \mathbb{P}[X(t) = X(t')\mu, \mathcal{L}(X(t')) = \nu, t^*(X(t')\mu) = j \mid E] \\ = \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\nu} \sum_{t'=0}^{t-1} \mathbb{P}[X(t) = X(t')\mu \mid E] \cdot \mathbb{P}[\mathcal{L}(X(t')) = \nu \mid X(t) = X(t')\mu, E] \\ \cdot \mathbb{P}[t^*(X(t')\mu) = j \mid X(t) = X(t')\mu, \mathcal{L}(X(t')) = \nu, E]$$

where by  $E$  we have denoted

$$E := \{t \leq \tau_n, t^*(X(1)), \dots, t^*(X(t-1)), K_1, \dots, K_n\}$$

and  $X(t')\mu$  denotes the  $\mu$ -th child of  $X(t')$ . Putting

$$E' := \{t \leq \tau_n, t^*(X(1)), \dots, t^*(X(t-\mu)), t^*(X(t')1), \dots, t^*(X(t')(\mu-1)), K_1, \dots, K_n\}$$

we get

$$\mathbb{P}[t^*(X(t')\mu) = j \mid X(t) = X(t')\mu, \mathcal{L}(X(t')) = \nu, E] \\ = \mathbb{P}[t^*(X(t')\mu) = j \mid X(t) = X(t')\mu, \mathcal{L}(X(t')) = \nu, E'] \\ = \mathbb{P}[t^*(X(t')\mu) = j \mid \mathcal{L}(X(t')) = \nu, t \leq \tau_n, t^*(X(t')1), \dots, t^*(X(t')(\mu-1))] = \alpha_j$$

due to the properties of the multinomial distribution. So we get

$$\mathbb{P}[t^*(X(t)) = j \mid t \leq \tau_n, t^*(X(1)), \dots, t^*(X(t-1)), K_1, \dots, K_n] = \alpha_j.$$

Furthermore, let  $U(t) \sim U(0, 1)$ ,  $t \in \mathbb{N}$ , be independent random variables (independent of  $(K_1, \dots, K_n)$ ) such that the transmission at time  $t$  fails iff  $U(t) \geq (n_j - \mathbf{n}^j)/n_j$ . Then an easy induction w.r.t.  $t$  shows that on  $\{t \leq \tau_n\}$  the event  $\{T_l^i = t_l^i \text{ for all } i \in [J], l \in$

$[n_i]$  is  $\sigma(t^*(X(\tilde{t})1), \dots, t^*(X(\tilde{t})\mathcal{L}(X(\tilde{t}))), U(\tilde{t}) | \tilde{t} < t)$  measurable as long as  $\max\{t_l^i \mid i \in [J], l \in [n_i]\} \leq t$ . Since this holds for arbitrary  $t \in \mathbb{N}$  and since the random variables  $t^*(X(\tilde{t})1), \dots, t^*(X(\tilde{t})\mathcal{L}(X(\tilde{t}))), U(\tilde{t}), \tilde{t} \in \mathbb{N}$  are independent of  $K_1, \dots, K_n$ , we get that  $T_l^i, i \in [J], l \in [n_i]$  are independent of  $K_1, \dots, K_n$ .

For every  $v \in \mathcal{U}_\infty$  with  $v > w$ , there is  $B_2 = B_2(v) \in \sigma(K_1, \dots, K_{m-1})$  with  $\{X(k_m) = v\} \cap A = B_2 \cap A$ . Notice that  $t^*(X(T_m)) = h(m)$ . So for every  $B_3 \in \sigma(K_1, \dots, K_{m-1})$  and  $B_4 \in \sigma(\mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]})$  with  $B_3 \cap B_4 \subseteq A$  and for every  $\lambda \in \mathbb{N}$  we get

$$\begin{aligned} \mathbb{P}[\{K_m = \lambda\} \cap B_3 \cap B_4] &= \mathbb{P}[\{\mathcal{L}(X(k_m)) = \lambda\} \cap B_3 \cap B_4] \\ &= \sum_{v \in \mathcal{U}_\infty, v > w} \mathbb{P}[\{\mathcal{L}(v) = \lambda\} \cap \{t^*(v) = h(m)\} \cap \{v = X(k_m)\} \cap B_3 \cap B_4] \\ &= \sum_{v \in \mathcal{U}_\infty, v > w} \mathbb{P}[\mathcal{L}(v) = \lambda, t^*(v) = h(m)] \cdot \mathbb{P}[B_2 \cap B_3] \cdot \mathbb{P}[B_4] \\ &= \mathbb{P}[\mathcal{L}^{h(m)} = \lambda] \sum_{v \in \mathcal{U}_\infty, v > w} \mathbb{P}[B_2 \cap B_3 \cap B_4] \\ &= \mathbb{P}[\mathcal{L}^{h(m)} = \lambda] \sum_{v \in \mathcal{U}_\infty, v > w} \mathbb{P}[\{v = X(k_m)\} \cap B_3 \cap B_4] \\ &= \mathbb{P}[\mathcal{L}^{h(m)} = \lambda] \cdot \mathbb{P}[B_3 \cap B_4]. \end{aligned}$$

Summing up over all possible values  $k_1, \dots, k_m$  and  $w \in \mathcal{U}_\infty$  we can relax the assumption  $B_3 \cap B_4 \subseteq A$  to  $B_3 \cap B_4 \subseteq \{T_m \leq \tau_n\}$ . Since the sets  $B_3 \cap B_4$  with  $B_3 \in \sigma(K_1, \dots, K_{m-1})$  and  $B_4 \in \sigma(\mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]})$  form an intersection-stable generating system of  $\sigma(K_1, \dots, K_{m-1}, \mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]})$ , we get

$$\mathbb{P}[\{K_m = \lambda\} \cap B_5] = \mathbb{P}[\mathcal{L}^{h(m)} = \lambda] \cdot \mathbb{P}[B_5] \quad (14)$$

for all  $B_5 \in \sigma(K_1, \dots, K_{m-1}, \mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]})$  with  $B_5 \subseteq \{T_m \leq \tau_n\}$ . Since (14) is trivial for  $B_5 \in \sigma(K_1, \dots, K_{m-1}, \mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]})$  with  $B_5 \subseteq \{T_m > \tau_n\}$ , we can drop the assumption  $B_5 \subseteq \{T_m \leq \tau_n\}$  in order to arrive at (13). Applying (13) inductively, we get

$$\mathbb{P}[K_1 = \lambda_1, \dots, K_n = \lambda_n, \mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]} \in B_1] = \prod_{m=1}^n \mathbb{P}[\mathcal{L}^{h(m)} = \lambda_m] \cdot \mathbb{P}[\mathbb{T}_{\{n_1, \dots, n_J\}}^{[J]} \in B_1]$$

and this shows (12). Summing up over all  $h \in \mathcal{H}$ , we get the claimed independence properties.  $\square$

Finally, we can also state in our case of different types of vertices, a result similar to [4, Proposition 2.1].

**Lemma 4.2.** *If we denote by  $Z_{\text{tot}}^{\text{MGW}} = |V(\mathbb{T}^{\text{MGW}})|$  the total number of individuals in the Galton-Watson process, then the full transmission event is asymptotically included in the survival event in the sense that*

$$\text{Trans}_n \subset \{Z_{\text{tot}}^{\text{MGW}} \geq n\}, \quad \text{where} \quad \{Z_{\text{tot}}^{\text{MGW}} \geq n\} \downarrow \text{Surv}^{\text{MGW}}, \quad n \rightarrow \infty.$$

---

Conversely, the event of termination at time  $o(n)$  in the epidemic model converges to the event of extinction in the multitype Galton-Watson process.

*Proof.* The first claim follows directly from the construction. The converse claim is a consequence of Lemma 4.3.  $\square$

In the sequel we assume Assumption 2 holds, in order to ensure that the MGW is supercritical and hence  $\mathbb{P}[\text{Surv}^{\text{MGW}}] > 0$ . In fact, it is close to being necessary for  $\mathbb{P}[\text{Surv}^{\text{MGW}}] > 0$  – we have only excluded the case  $\mathcal{L}^1 + \dots + \mathcal{L}^J = 1$  a.s., in which we would get very pathological problems.

We want to show next, that also in the case of  $J$  different types of coupons, the transmission takes place on a macroscopic time level. The proof is more involved than the one in the homogeneous case and uses a different idea.

**Lemma 4.3.** *Suppose that Assumption 1 and Assumption 2 hold. There exists  $\epsilon_0 > 0$  such that, for all  $\epsilon \in (0, \epsilon_0)$ :*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \tau_n \geq n\epsilon, \text{Surv}^{\text{MGW}} \right] = \mathbb{P} \left[ \text{Surv}^{\text{MGW}} \right] = 1 - \sigma^{\text{MGW}}.$$

*Proof.* Put

$$\epsilon_1 := \left( 1 - \frac{1}{\rho(M)} \right) \quad \text{and} \quad \epsilon_0 := \min\{\gamma_1, \dots, \gamma_J\} \cdot \epsilon_1$$

and let  $\epsilon \in (0, \epsilon_0)$ . Then there exists an  $\epsilon' \in (0, \epsilon_1)$  with  $\epsilon = \min\{\gamma_1, \dots, \gamma_J\} \cdot \epsilon'$ . We have

$$\{\tau_n \geq n\epsilon\} \supseteq \{|\mathcal{T}^\circ(\tau_n)| \geq n\epsilon\}$$

since  $|\mathcal{T}^\circ(t)| \leq t$  for all  $t \in \mathbb{N}$ . In order to find a lower bound on the probability  $\mathbb{P} [ |\mathcal{T}^\circ(\tau_n)| \geq n\epsilon, \text{Surv}^{\text{MGW}} ]$  we construct a new process  $(\mathcal{U}^\circ(r), \mathcal{U}^\boxplus(r), \mathcal{U}^\dagger(r))$ ,  $r \in \mathbb{N}$ , similar to  $(\mathcal{T}^\circ(t), \mathcal{T}^\boxplus(t), \mathcal{T}^\dagger(t))$ . To this aim, given a fixed  $k$  we let  $g_k$  be the last generation in the Galton-Watson process such that the total number of individuals ever born does not exceed  $k$ , that is

$$g_k = \max \left\{ g \in \mathbb{N} \left| \sum_{t=1}^g \mathcal{Z}_t \leq k \right. \right\}.$$

Then

- at each epoch  $r$ , a new vertex of the GW tree is added to  $\mathcal{U}^\circ(r)$ , until the process reaches generation  $g_k + 1$ ;
- from generation  $g_k + 1$ , vertices of the GW-tree are added with probability  $1 - \epsilon'$ ;
- moreover, from generation  $g_k + 1$ , as long as only a small fraction - smaller than  $\epsilon'$  - of the available "coupons" of each type has been collected, the vertices that are added to  $\mathcal{U}^\circ(r)$  are chosen from the ones in  $\mathcal{T}^\circ(t_r)$  ( $t_r$  is a function of  $r$ ).

Now let  $k = \lceil (\log n)^2 \rceil$  and let

$$k_1 := \mathcal{Z}_1 + \dots + \mathcal{Z}_{g_k}$$

denote the total number of individuals up to the  $g_k$ -th generation. We set

$$\left( \mathcal{U}^\circ(0), \mathcal{U}^\boxplus(0), \mathcal{U}^\dagger(0) \right) := \left( \mathcal{T}^\circ(0), \mathcal{T}^\boxplus(0), \mathcal{T}^\dagger(0) \right)$$

and assume that  $(\mathcal{U}^\circ(r), \mathcal{U}^\boxplus(r), \mathcal{U}^\dagger(r))$  has been defined for some  $r$ . Conditionally on  $\mathcal{U}^\boxplus(r) \neq \emptyset$  we define  $(\mathcal{U}^\circ(r+1), \mathcal{U}^\boxplus(r+1), \mathcal{U}^\dagger(r+1))$  by picking the element  $v \in \mathcal{U}^\boxplus(r)$  of least generation, among elements of same generation the first one in the lexicographic order, and distinguishing the following cases:

- If  $|v| \leq g_k$ , i.e.  $v$  belongs to the first  $g_k$  generations, then assume that the transmission is successful, i.e.

$$\begin{aligned} \mathcal{U}^\circ(r+1) &:= \mathcal{U}^\circ(r) \cup \{v\} \\ \mathcal{U}^\boxplus(r+1) &:= \mathcal{U}^\boxplus(r) \setminus \{v\} \cup \{v_1, \dots, v_{\mathcal{L}^i(v)}\} \\ \mathcal{U}^\dagger(r+1) &:= \mathcal{U}^\dagger(r), \end{aligned}$$

where  $v_1, \dots, v_{\mathcal{L}^i(v)}$  are the children of  $v$  in the Galton-Watson tree  $\mathcal{T}$ .

- If  $|v| > g_k$  and  $v \in \mathcal{T}^\dagger(t+1) \setminus \mathcal{T}^\dagger(t)$  for some  $t \in \{1, \dots, \tau_n\}$  such that  $\mathbf{N}_t^j \leq \epsilon' n_j$  for all  $j \in \{1, \dots, J\}$ , i.e. the transmission in the old  $\mathcal{T}$ -model is attempted at some medium time and fails, then the transmission fails also in the new  $\mathcal{U}$ -model, i.e.

$$\begin{aligned} \mathcal{U}^\circ(r+1) &:= \mathcal{U}^\circ(r) \\ \mathcal{U}^\boxplus(r+1) &:= \mathcal{U}^\boxplus(r) \setminus \{v\} \\ \mathcal{U}^\dagger(r+1) &:= \mathcal{U}^\dagger(r) \cup \{v\}. \end{aligned}$$

- If  $|v| > g_k$ ,  $v \in \mathcal{T}^\circ(t+1) \setminus \mathcal{T}^\circ(t)$  for some  $t \in \{1, \dots, \tau_n\}$  such that  $\mathbf{N}_t^j \leq \epsilon' n_j$  for all  $j \in [J]$ , then let the transmission be successful with probability  $\frac{1-\epsilon'}{(n_i - \mathbf{N}_t^i)/n_i}$ , where  $i = t(v)$ , conditionally on everything defined so far.
- If either  $v \in (\mathcal{T}^\circ(t+1) \cup \mathcal{T}^\dagger(t+1)) \setminus (\mathcal{T}^\circ(t) \cup \mathcal{T}^\dagger(t))$  for some  $t \in \{1, \dots, \tau_n\}$  such that  $\mathbf{N}_t^j > \epsilon' n_j$  for some  $j \in \{1, \dots, J\}$  or  $v \notin \mathcal{T}^\circ(\tau_n) \cup \mathcal{T}^\dagger(\tau_n)$ , then the transmission fails with probability  $\epsilon'$  conditionally on everything defined so far and is successful otherwise.

This construction is summarized in Table 1. If in any of these cases  $t \in \{1, \dots, \tau_n\}$  with  $v \in (\mathcal{T}^\circ(t+1) \cup \mathcal{T}^\dagger(t+1)) \setminus (\mathcal{T}^\circ(t) \cup \mathcal{T}^\dagger(t))$  exists, then denote  $t+1$  by  $t_{r+1}$ .

If  $\mathcal{U}^\boxplus(r) = \emptyset$ , the process stops and we set  $\rho_n = r$ . Put

$$r_0 := \max \{r' \in \{1, \dots, \rho_n\} \mid t_r \text{ is defined for all } r \leq r'\}.$$

Then either  $r_0 = \rho_n$  or for the vertex  $v \in \mathbb{T}^{\text{MGW}}$  chosen in the  $r_0 + 1$ -st step there is no  $t \in \{1, \dots, \tau_n\}$  with  $v \in (\mathcal{T}^\circ(t+1) \cup \mathcal{T}^\dagger(t+1)) \setminus (\mathcal{T}^\circ(t) \cup \mathcal{T}^\dagger(t))$ . The latter however

		transmission in $\mathcal{T}$ -model successful	transmission in $\mathcal{T}$ -model fails
generation $\leq g_k$		transmission successful	
generation $> g_k$	$\mathbf{N}_t^j \leq \epsilon' n_j$ for all $j \in [J]$	transmission successful with probability $\frac{1-\epsilon'}{(n_i - \mathbf{N}_t^i)/n_i}$	transmission fails
generation $> g_k$	$\mathbf{N}_t^j > \epsilon' n_j$ for some $j \in [J]$ or $t$ not defined	transmission successful with probability $1 - \epsilon'$	

Table 1: Summary of the construction of the  $\mathcal{U}$ -model

means, that at the time step  $r_1 \in \{1, \dots, r_0\}$ , when  $v$  was added to  $\mathcal{U}^{\boxplus}(r)$ , it was not added to  $\mathcal{T}^{\boxplus}(t)$ , so the transmission in the  $\mathcal{T}$ -model at time  $t_{r_1}$  failed, while the transmission in the  $\mathcal{U}$ -model at time  $r_1$  was successful. This is possible in two cases: either the vertex  $v_1$  treated at time  $r_1$  belongs to the first  $g_k$  generations or  $\mathbf{N}_{t_{r_1}}^j > \epsilon' n_j$  for some  $j \in \{1, \dots, J\}$ . Now assume that the event  $\text{Succ}_{g_k}$ , that in the first  $g_k$  generations all transmissions are successful in the  $\mathcal{T}$ -model, occurs. Then  $v_1$  belonging to the first  $g_k$  generations cannot explain anymore that the transmission in the  $\mathcal{U}$ -model at time  $r_1$  is successful, while the transmission in the  $\mathcal{T}$ -model at time  $t_{r_1}$  fails, and we are left with the case that  $\mathbf{N}_{t_{r_1}}^j > \epsilon' n_j$  for some  $j \in \{1, \dots, J\}$ . Since  $r_0 \geq r_1$  we get  $\mathbf{N}_{t_{r_0}}^j > \epsilon' n_j$  for some  $j \in \{1, \dots, J\}$ . In conclusion we have

$$\text{Succ}_{g_k} \subseteq \{\rho_n = r_0\} \cup \{\mathbf{N}_{t_{r_0}}^j > \epsilon' n_j \text{ for some } j \in \{1, \dots, J\}\}.$$

Since  $\tau_n \geq t_{r_0}$  by the definition of  $r_0$  and since  $\mathbf{N}_{t_{r_0}}^j = |\mathcal{T}_j^\circ(t_{r_0})| \leq |\mathcal{T}^\circ(t_{r_0})|$ , we get

$$\begin{aligned} \mathbb{P} \left[ \tau_n \geq n\epsilon, \text{Surv}^{\text{MGW}} \right] &\geq \mathbb{P} \left[ |\mathcal{T}^\circ(\tau_n)| \geq n\epsilon, \text{Surv}^{\text{MGW}} \right] \\ &\geq \mathbb{P} \left[ |\mathcal{T}^\circ(t_{r_0})| \geq n\epsilon, \text{Surv}^{\text{MGW}} \right] \\ &\geq \mathbb{P} \left[ \mathbf{N}_{t_{r_0}}^j > \epsilon' n_j \text{ for some } j \in \{1, \dots, J\} \right] \\ &\geq \mathbb{P} [\rho_n = \infty, \text{Succ}_{g_k}] \geq \mathbb{P} [\rho_n = \infty] + \mathbb{P} [\text{Succ}_{g_k}] - 1. \end{aligned}$$

In order to estimate  $\mathbb{P}[\text{Succ}_{g_k}]$ , put  $\bar{\beta} := \max\{\beta_1, \dots, \beta_n\}$ . Notice that the probability that in one step a new vertex is infected is bounded from below by  $1 - \bar{\beta} \cdot j/n$  given that  $j$  vertices have already been infected. Then

$$1 - \mathbb{P}[\text{Succ}_{g_k}] \leq 1 - \left(1 - \bar{\beta} \frac{1}{n}\right) \cdots \left(1 - \bar{\beta} \frac{k}{n}\right) \leq 1 - \left(1 - \bar{\beta} \frac{k}{n}\right)^k \sim \bar{\beta} \frac{k^2}{n} \rightarrow 0,$$

since  $k^2/n \rightarrow 0$  as  $n \rightarrow \infty$ .

Let us now turn to the probability of the event  $\{\rho_n = \infty\}$  that the process in the  $\mathcal{U}$ -model survives forever. For any vertex  $v$  of the  $g_k$ -th generation, let  $\text{Surv}(v)$  denote the event that  $v$  has infinitely many descendants in the  $\mathcal{U}$ -model. Clearly  $\{\rho_n = \infty\} = \bigcup_{l=1}^{Z_{g_k}} \text{Surv}(v_l)$ , where  $v_1, \dots, v_{Z_{g_k}}$  are the members of the  $g_k$ -th generation. Notice that the systems of

descendants of these vertices form independent Galton-Watson trees that are independent of the process  $\mathcal{U}^\circ$  up to time  $k_1$ . This is a consequence of the fact that the remainder of an i.i.d. sequence after a stopping time is again an i.i.d. sequence.

Observe that the offspring distribution of these Galton-Watson trees is given by  $\bar{L}^{(i,j)} = \sum_{k=1}^{L^{(i,j)}} \delta_k^{(i,j)}$ , where  $\delta_1^{(i,j)}, \dots, \delta_{L^{(i,j)}}^{(i,j)}$  are independent random variables with distribution

$$\mathbb{P}[\delta_k^{(i,j)} = 1] = 1 - \epsilon' \quad \text{and} \quad \mathbb{P}[\delta_k^{(i,j)} = 0] = \epsilon' \quad \text{for} \quad i, j = 1, \dots, J, k = 1, \dots, L^{(i,j)}.$$

Hence the mean value matrix is  $(1 - \epsilon')M$ , and in particular it has spectral radius bigger than 1. So the probability of the event  $\text{Surv}(v_l)$ , that this Galton-Watson process survives, is positive and independent of  $l$  or  $n$ . Hence we get for any  $z \in \mathbb{N}$  that

$$\mathbb{P}[\rho_n < \infty] \leq \mathbb{P}\left[\{\mathcal{Z}_{g_k} \leq z\} \cup \left(\text{Surv}(v_1)^{\complement} \cap \dots \cap \text{Surv}(v_{z+1})^{\complement}\right)\right] \leq \mathbb{P}[\mathcal{Z}_{g_k} \leq z] + (1 - \delta)^{z+1},$$

where  $\delta := \mathbb{P}[\text{Surv}(v)] > 0$ . Therefore

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\tau_n \geq n\epsilon, \text{Surv}^{\text{MGW}}\right] \geq \lim_{n \rightarrow \infty} \mathbb{P}[\rho_n = \infty] \geq 1 - \lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{Z}_{g_k} \leq z] - (1 - \delta)^{z+1}.$$

Since  $g_k \rightarrow \infty$  as  $n \rightarrow \infty$  almost surely and  $\mathcal{Z}_g \rightarrow \infty$  as  $g \rightarrow \infty$  almost surely conditioned on  $\text{Surv}^{\text{MGW}}$ , where the latter is a consequence of [7, Section II, Theorem 6.1], we have that the total number  $\mathcal{Z}_{g_k}$  of individuals in the  $g_k$ -th generation tends to infinity as  $n \rightarrow \infty$  conditionally on  $\text{Surv}^{\text{MGW}}$  almost surely. So

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{Z}_{g_k} \leq z] &= \lim_{n \rightarrow \infty} \mathbb{P}\left[\mathcal{Z}_{g_k} \leq z, \text{Surv}^{\text{MGW}}\right] + \lim_{n \rightarrow \infty} \mathbb{P}\left[\mathcal{Z}_{g_k} \leq z, (\text{Surv}^{\text{MGW}})^{\complement}\right] \\ &= \mathbb{P}\left[(\text{Surv}^{\text{MGW}})^{\complement}\right] \end{aligned}$$

and thus

$$\lim_{n \rightarrow \infty} \mathbb{P}\left[\tau_n \geq n\epsilon, \text{Surv}^{\text{MGW}}\right] \geq \mathbb{P}\left[\text{Surv}^{\text{MGW}}\right] - (1 - \delta)^{z+1},$$

which shows the assertion, since  $z \in \mathbb{N}$  was arbitrary.  $\square$

## 5 Law of large numbers

Let  $\theta$  denote the unique solution in  $(0, \infty)$  of

$$f(\theta) := \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \underbrace{\gamma_i(1 - e^{-\beta_i \theta})}_{:= w_i(\theta)} - \theta = 0, \quad (15)$$

or written in another way  $\langle \mathbb{E}[\mathcal{L}], \mathbf{w}(\theta) \rangle = \theta$ , where  $\mathbb{E}[\mathcal{L}] = (\mathbb{E}[\mathcal{L}^1], \dots, \mathbb{E}[\mathcal{L}^J])$  and

$$\mathbf{w}(\theta) = (\gamma_1(1 - e^{-\beta_1 \theta}), \dots, \gamma_J(1 - e^{-\beta_J \theta}))$$

and  $\langle \mathbb{E}[\mathcal{L}], \mathbf{w}(\theta) \rangle$  is the usual scalar product of the two involved vectors. Observe that there is a solution, since  $f(0) = 0$ ,  $f'(0) = \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i \beta_i - 1 = \rho(M) - 1 > 0$  due to Assumption 2,  $\lim_{\theta \rightarrow \infty} f(\theta) = -\infty$  and  $f$  is continuous. Moreover, the solution is unique, since  $f''(\theta) = -\sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i \beta_i^2 e^{-\beta_i \theta} < 0$  for  $\theta \in [0, \infty)$  and therefore  $f(\theta) = 0$  can have at most two solutions in  $[0, \infty)$  and one of these solutions is given by  $\theta = 0$ .



---

**Theorem 5.1.** Let  $0 < \mathbb{E}[\mathcal{L}^i] < \infty$ , for all  $i = 1, \dots, J$ .

1. Then

$$\frac{\tau_n}{n} \longrightarrow \theta \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}} \quad \text{and} \quad \frac{\tilde{\tau}_n}{n} \longrightarrow \theta \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$$

in probability as  $n \rightarrow \infty$ .

2. Also

$$\frac{\mathcal{N}_{\tau_n}}{n} \longrightarrow \langle \mathbf{w}(\theta), \mathbf{1} \rangle \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}} \quad \text{and} \quad \frac{\tilde{\mathcal{N}}_{\tilde{\tau}_n}}{n} \longrightarrow \langle \mathbf{w}(\theta), \mathbf{1} \rangle \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$$

in probability as  $n \rightarrow \infty$ , where  $\mathcal{N}_t = \sum_{i=1}^J \mathbf{N}_t^i$  and  $\tilde{\mathcal{N}}_t = \sum_{i=1}^J \tilde{\mathbf{N}}_t^i$ .

*Proof.* Proof of  $\frac{\tau_n}{n} \rightarrow \theta \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$  in probability as  $n \rightarrow \infty$ .

We use a similar approach as in [4, Theorem 2.2]. We have  $\mathcal{R}_t^i = \sum_{l=1}^{\mathbf{N}_t^i} \mathcal{L}_l^i$ , where for each  $i = 1, \dots, J$ ,  $\mathcal{L}_l^i$  are i.i.d. distributed like  $\mathcal{L}^i$  with  $\mathbb{E}[\mathcal{L}^i] < \infty$ . Also  $\tilde{\mathbf{N}}_t^1, \dots, \tilde{\mathbf{N}}_t^J$  are independent and for every  $i$ ,  $\frac{\tilde{\mathbf{N}}_{n_s}^i}{n} \rightarrow \gamma_i(1 - e^{-\beta_i s}) = w_i(s)$  in probability as  $n \rightarrow \infty$ , and

$$\tilde{\mathcal{S}}_t = \mathcal{S}_{\mathcal{P}_t} = \sum_{l=1}^{\mathbf{N}_{\mathcal{P}_t}^1} \mathcal{L}_l^1 + \dots + \sum_{l=1}^{\mathbf{N}_{\mathcal{P}_t}^J} \mathcal{L}_l^J - \mathcal{P}_t.$$

Using that  $\tilde{\mathbf{N}}_t^i = \mathbf{N}_{\mathcal{P}_t}^i$  for all  $i \in [J]$ , we can in turn write

$$\tilde{\mathcal{S}}_t = \mathcal{S}_{\mathcal{P}_t} = \sum_{l=1}^{\tilde{\mathbf{N}}_t^1} \mathcal{L}_l^1 + \dots + \sum_{l=1}^{\tilde{\mathbf{N}}_t^J} \mathcal{L}_l^J - \mathcal{P}_t.$$

Therefore

$$\begin{aligned} \frac{\tilde{\mathcal{S}}_{n_s}}{n} &= \frac{\mathcal{S}_{\mathcal{P}_{n_s}}}{n} = \frac{1}{n} \left( \sum_{l=1}^{\tilde{\mathbf{N}}_{n_s}^1} \mathcal{L}_l^1 + \dots + \sum_{l=1}^{\tilde{\mathbf{N}}_{n_s}^J} \mathcal{L}_l^J - \mathcal{P}_{n_s} \right) \\ &= \underbrace{\frac{\sum_{l=1}^{\tilde{\mathbf{N}}_{n_s}^1} \mathcal{L}_l^1}{\tilde{\mathbf{N}}_{n_s}^1}}_{\rightarrow \mathbb{E}[\mathcal{L}^1]} \cdot \underbrace{\frac{\tilde{\mathbf{N}}_{n_s}^1}{n}}_{\rightarrow \gamma_1(1 - e^{-\beta_1 s})} + \dots + \underbrace{\frac{\sum_{l=1}^{\tilde{\mathbf{N}}_{n_s}^J} \mathcal{L}_l^J}{\tilde{\mathbf{N}}_{n_s}^J}}_{\rightarrow \mathbb{E}[\mathcal{L}^J]} \cdot \underbrace{\frac{\tilde{\mathbf{N}}_{n_s}^J}{n}}_{\rightarrow \gamma_J(1 - e^{-\beta_J s})} - \underbrace{\frac{\mathcal{P}_{n_s}}{n}}_{\rightarrow s} \\ &\rightarrow \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i(1 - e^{-\beta_i s}) - s = \langle \mathbb{E}[\mathcal{L}], \mathbf{w}(s) \rangle - s \end{aligned}$$

in probability, uniformly on compacts of  $\mathbb{R}_+$ , where in the second line above we have applied the weak law of large numbers for the sequences of involved random variables, and the last line holds in view of the continuous mapping theorem. Now, for any  $\delta > 0$  and  $\epsilon > 0$ , we

have

$$\begin{aligned} \mathbb{P} \left[ \left| \tau_n - n\theta \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}} \right| > n\delta \right] &= \mathbb{P} \left[ |\tau_n - n\theta| > n\delta, \text{Surv}^{\text{MGW}} \right] + \mathbb{P} \left[ \tau_n > n\delta, \left( \text{Surv}^{\text{MGW}} \right)^{\mathbb{C}} \right] \\ &\leq \mathbb{P} \left[ |\tau_n - n\theta| > n\delta, \text{Surv}^{\text{MGW}}, \tau_n \geq n\epsilon \right] \\ &\quad + \mathbb{P} \left[ \tau_n < n\epsilon, \text{Surv}^{\text{MGW}} \right] + \mathbb{P} \left[ \tau_n > n\delta, \left( \text{Surv}^{\text{MGW}} \right)^{\mathbb{C}} \right]. \end{aligned}$$

On the right hand side of the above inequality, the second term tends to 0 as  $n \rightarrow \infty$  in view of Lemma 4.3. The third term vanishes also as  $n \rightarrow \infty$ , since  $\tau_n$  is smaller than the extinction time of the multitype Galton-Watson process, which is a.s. finite on the extinction event  $(\text{Surv}^{\text{MGW}})^{\mathbb{C}}$ . Finally, for the first term, by picking  $\epsilon < \theta - \delta$ , we have the following:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P} \left[ |\tau_n - n\theta| > n\delta, \text{Surv}^{\text{MGW}}, \tau_n \geq n\epsilon \right] \\ \leq \limsup_{n \rightarrow \infty} \left( \mathbb{P} \left[ n\epsilon \leq \tau_n \leq n\theta - n\delta, \text{Surv}^{\text{MGW}} \right] + \mathbb{P} \left[ \tau_n > n\theta + n\delta, \text{Surv}^{\text{MGW}} \right] \right) \\ \leq \limsup_{n \rightarrow \infty} \left( \mathbb{P} \left[ \min\{\mathcal{S}_{\lfloor ns \rfloor}; s \in [\epsilon, \theta - \delta]\} \leq 0 \right] + \mathbb{P} \left[ \mathcal{S}_{\lfloor n(\theta + \delta) \rfloor} > 0 \right] \right) \\ \leq \limsup_{n \rightarrow \infty} \left( \mathbb{P} \left[ \min\{\mathcal{S}_{\mathcal{P}_{ns}}; s \in [\epsilon/2, \theta - \delta/2]\} \leq 0 \right] + \mathbb{P} \left[ \mathcal{P}_{n\epsilon/2} > \lfloor n\epsilon \rfloor \right] \right) \\ + \mathbb{P} \left[ \mathcal{P}_{n \cdot (\theta - \delta/2)} < \lfloor n \cdot (\theta - \delta) \rfloor \right] + \mathbb{P} \left[ \mathcal{S}_{\mathcal{P}_{n(\theta + \delta/2)}} > 0 \right] \\ + \mathbb{P} \left[ \mathcal{P}_{n(\theta + \delta/2)} > \lfloor n \cdot (\theta + \delta) \rfloor \right] = 0. \end{aligned}$$

For the second inequality we have used that since  $\tau_n = \min\{t \in \mathbb{N} \mid \mathcal{S}_t = 0\}$  we have  $\tau_n \leq t_0$  if and only if there is  $t \leq t_0$  with  $\mathcal{S}_t = 0$ . To see the third inequality observe that  $\mathcal{P}_{ns}$ ,  $s \in [\epsilon/2, \theta - \delta/2]$  attains with probability one every integer from  $\mathcal{P}_{n\epsilon/2}$  through  $\mathcal{P}_{n \cdot (\theta - \delta/2)}$ . The final limit relation holds, since  $\frac{\tilde{\mathcal{S}}_{ns}}{n} = \frac{\mathcal{S}_{\mathcal{P}_{ns}}}{n}$  converges in probability to  $\langle \mathbb{E}[\mathcal{L}], \mathbf{w}(s) \rangle - s$  uniformly on compacts of  $\mathbb{R}_+$ , and, moreover,  $\langle \mathbb{E}[\mathcal{L}], \mathbf{w}(s) \rangle - s > 0$  for  $s \in (0, \theta)$  and  $\langle \mathbb{E}[\mathcal{L}], \mathbf{w}(s) \rangle - s < 0$  for any  $s \in (\theta, \infty)$ . This completes the proof of the first part.

Proof of  $\frac{\tilde{\tau}_n}{n} \rightarrow \theta \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$  in probability as  $n \rightarrow \infty$ .

We write  $\tilde{\tau}_n = \sum_{k=1}^{\tilde{\tau}_n} E_k$  for independent, exponentially distributed random variables  $E_k \sim \text{Exp}(1)$  with rate 1 and set  $\theta' := \theta \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$ . Pick any  $\epsilon > 0$ . We first split the event  $\left[ \left| \frac{\tilde{\tau}_n}{n} - \theta' \right| > \epsilon \right]$  intersecting it with  $A_n := \left[ \left| \frac{\tilde{\tau}_n}{n} - \theta' \right| \leq \frac{\epsilon}{9} \right]$  and its complement. Note that on  $A_n$ ,  $\tilde{\tau}_n$  differs from  $\sum_{k=1}^{\lfloor n\theta' \rfloor} E_k$  by a random variable which is the sum of at most  $\lfloor n\epsilon/9 \rfloor$  exponentials of rate 1. Let  $W_n$  be this variable (with its sign) so that  $\tilde{\tau}_n = \sum_{k=1}^{\lfloor n\theta' \rfloor} E_k + W_n$ .

$$\mathbb{P} \left[ \left| \frac{\tilde{\tau}_n}{n} - \theta' \right| > \epsilon \right] \leq \mathbb{P} \left[ A_n^{\mathbb{C}} \right] + \mathbb{P} \left[ \left| \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) + \frac{\lfloor n\theta' \rfloor}{n} - \theta' + \frac{W_n}{n} \right| > \epsilon \right].$$

We already know that the first term vanishes, so we concentrate on the second one, which we call  $\mathbb{P}(B_n)$ . We split the event by intersecting with  $C_n := \left[ |W_n/n| \leq \frac{\epsilon}{3} \right]$  and its complement.

By the weak law of large numbers,  $\mathbb{P}[C_n^c] \rightarrow 0$ , since  $|W_n|$  is a.s. positive and bounded by a sum of  $n\epsilon/9$  terms with mean 1. Now, use  $\theta' - \lfloor n\theta' \rfloor/n \in [0, 1/n]$  to get

$$\begin{aligned} \mathbb{P}[B_n \cap C_n] &= \mathbb{P} \left[ \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) > \epsilon - \frac{W_n}{n}, C_n \right] \\ &\quad + \mathbb{P} \left[ \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) < -\epsilon + \frac{1}{n} - \frac{W_n}{n}, C_n \right]. \end{aligned}$$

For the first term,

$$\mathbb{P} \left[ \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) > \epsilon - \frac{W_n}{n}, C_n \right] \leq \mathbb{P} \left[ \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) > \frac{2}{3}\epsilon \right] \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

since by the weak law of large numbers  $\frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) \rightarrow 0$ . For the second term

$$\mathbb{P} \left[ \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) < -\epsilon + \frac{1}{n} - \frac{W_n}{n}, C_n \right] \leq \mathbb{P} \left[ \frac{1}{n} \sum_{k=1}^{\lfloor n\theta' \rfloor} (E_k - 1) < -\frac{2}{3}\epsilon + \frac{1}{n} \right] \rightarrow 0,$$

by the same argument.

Proof of  $\frac{\tilde{\mathcal{N}}_{\tilde{\tau}_n}}{n} \rightarrow \langle \mathbf{w}(\theta), \mathbf{1} \rangle \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$  in probability as  $n \rightarrow \infty$ .

Since  $\frac{\tilde{\mathcal{N}}_{\tilde{\tau}_n}^i}{n} \rightarrow \gamma_i(1 - e^{-\beta_i s}) = w_i(s)$  in probability, the law of large numbers together with a random time change (see e.g. [2, p. 151]) and the first part of the proof yields

$$\begin{aligned} \frac{\tilde{\mathcal{N}}_{\tilde{\tau}_n}}{n} &= \sum_{i=1}^J \underbrace{\frac{\tilde{\mathcal{N}}_{\tilde{\tau}_n}^i}{n}}_{\rightarrow \gamma_i(1 - e^{-\beta_i \theta}) \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}} \\ &\rightarrow \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}} \sum_{i=1}^J \gamma_i(1 - e^{-\beta_i \theta}) = \langle \mathbf{w}(\theta), \mathbf{1} \rangle \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}, \end{aligned}$$

in probability, as  $n \rightarrow \infty$ . So this part of the proof is complete.

Proof of  $\frac{\mathcal{N}_{\tau_n}}{n} \rightarrow \langle \mathbf{w}(\theta), \mathbf{1} \rangle \mathbb{1}_{\{\text{Surv}^{\text{MGW}}\}}$  in probability as  $n \rightarrow \infty$ .

Since  $\tilde{\mathcal{N}}_t = \mathcal{N}_{\mathcal{P}_t}$  for  $t \in \mathbb{N}$  and  $\tau_n = \mathcal{P}_{\tilde{\tau}_n}$ , we have  $\tilde{\mathcal{N}}_{\tilde{\tau}_n} = \mathcal{N}_{\tau_n}$  and the claim follows from the previous one.  $\square$

## 6 Central Limit Theorems

In this section we prove central limit theorems for the duration  $\tau_n$  of the process and for the total infected vertices at this time. We begin with a preliminary result.

**Theorem 6.1.** For  $t \in [0, \infty)$ , and  $n \in \mathbb{N}$  denote by

$$\begin{aligned} Z_{nt}^{(i)} &= \frac{1}{\sqrt{n}} \left( \tilde{\mathbf{N}}_{nt}^i - n\gamma_i \cdot (1 - e^{-\beta_i t}) \right) \\ Z_{nt}^{\mathcal{P}} &= \frac{1}{\sqrt{n}} (\mathcal{P}_{nt} - nt). \end{aligned}$$

Then it holds

$$\left\{ \left( Z_{ns}^{(1)}, \dots, Z_{ns}^{(J)}, Z_{ns}^{\mathcal{P}} \right); s \in [0, \infty) \right\} \rightarrow \left\{ \left( \sqrt{\gamma_1} \mathbf{X}_s^{(1)}, \dots, \sqrt{\gamma_J} \mathbf{X}_s^{(J)}, \mathbf{B}_s^{\mathcal{P}} \right), s \in [0, \infty) \right\},$$

as  $n \rightarrow \infty$  in law in the Skorohod space  $\mathcal{D}([0, \infty), \mathbb{R}^{J+1})$  endowed with the standard  $\mathbf{J}_1$  topology, where  $(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(J)}, \mathbf{B}^{\mathcal{P}})$  is a centered Gaussian process with the first  $J$  components being independent and with covariances given by

$$\begin{aligned} \text{Cov}[\mathbf{X}_s^{(i)}, \mathbf{X}_t^{(i)}] &= (1 - e^{-\beta_i s}) e^{-\beta_i t} & s, t \in [0, \infty), \quad s \leq t, \\ \text{Cov}[\mathbf{X}_s^{(i)}, \mathbf{B}_t^{\mathcal{P}}] &= \min\{s, t\} \cdot \gamma_i e^{-s\beta_i} & s, t \in [0, \infty), \\ \text{Cov}[\mathbf{B}_s^{\mathcal{P}}, \mathbf{B}_t^{\mathcal{P}}] &= \min\{s, t\} & s, t \in [0, \infty). \end{aligned}$$

*Proof.* The limit of  $\{Z_{ns}^{(i)}; s \in [0, \infty)\}$  is given in Proposition 3.1. We next investigate the limit of  $\{Z_{ns}^{\mathcal{P}}; s \in [0, \infty)\}$ . Since  $\mathcal{P}_t$  is a Poisson process with rate 1, there are independent random variables  $E_j \sim \text{Exp}(1)$ ,  $j \in \mathbb{N}$  with

$$\{\mathcal{P}_t; t \in [0, \infty)\} = \left\{ \sum_{k=1}^{\infty} \mathbb{1}_{\{\sum_{j=1}^k E_j \leq t\}}; t \in [0, \infty) \right\}.$$

By using Donsker's invariance principle we first have that

$$\left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor ns \rfloor} (E_j - 1); s \in [0, \infty) \right\} \rightarrow \{\mathbf{B}_s; s \in [0, \infty)\}, \quad \text{as } n \rightarrow \infty$$

in law in the Skorohod topology. We then obtain

$$\{Z_{ns}^{\mathcal{P}}; s \in [0, \infty)\} \xrightarrow{\text{law}} \{\mathbf{B}_s; s \in [0, \infty)\}$$

in view of [2, Theorem 14.6], and we set  $\mathbf{B}_s^{\mathcal{P}} := \mathbf{B}_s$ . Now let us turn to  $\text{Cov}[\tilde{\mathbf{N}}_{ns}^i, \mathcal{P}_{nt}]$ . We assume w.l.o.g. that the vertices  $v_1, \dots, v_{n_i}$  are of type  $i$ , while the vertices  $v_{n_i+1}, \dots, v_n$  are not of type  $i$ . The time at which vertex  $v_k$ ,  $k = 1, \dots, n$  receives the virus for the  $l$ -th time can be modeled by  $t_{k,l} = \sum_{j=1}^l E_{k,j}$ , where again  $E_{k,j}$ ,  $k = 1, \dots, n$ ,  $j \in \mathbb{N}$ , are independent exponentially distributed random variables with rate  $p_{\mathbf{t}(v_k)}$ . Then

$$\{\tilde{\mathbf{N}}_t^i; t \in [0, \infty)\} \stackrel{\text{law}}{=} \left\{ \sum_{k=1}^{n_i} \mathbb{1}_{\{t_{k,1} \leq t\}}; t \in [0, \infty) \right\}$$

and

$$\{\mathcal{P}_t; t \in [0, \infty)\} \stackrel{\text{law}}{=} \left\{ \sum_{k=1}^n \sum_{l=1}^{\infty} \mathbb{1}_{\{t_{k,l} \leq t\}}; t \in [0, \infty) \right\}.$$

Observe that  $\mathcal{P}_t(k) := \sum_{l=1}^{\infty} \mathbb{1}_{\{t_{k,l} \leq t\}} \sim \text{Poi}(t p_t(v_k))$  and  $\mathbb{1}_{\{t_{k,1} \leq t\}} = \mathbb{1}_{\{\mathcal{P}_t(k) \geq 1\}}$ . For  $s \leq t$ :

$$\begin{aligned} \frac{1}{n} \text{Cov}[\tilde{\mathbf{N}}_{ns}^i, \mathcal{P}_{nt}] &= \frac{1}{n} \sum_{k=1}^{n_i} \left( \text{Cov} \left[ \mathbb{1}_{\{t_{k,1} \leq ns\}}, \mathcal{P}_{ns}(k) \right] + \text{Cov} \left[ \mathbb{1}_{\{t_{k,1} \leq ns\}}, \mathcal{P}_{nt}(k) - \mathcal{P}_{ns}(k) \right] \right) \\ &= \frac{1}{n} \sum_{k=1}^{n_i} \left( \mathbb{E}[\mathcal{P}_{ns}(k)] - \mathbb{E}[\mathcal{P}_{ns}(k)] \cdot \mathbb{E}[\mathbb{1}_{\{\mathcal{P}_{ns}(k) \geq 1\}}] \right) = \frac{1}{n} \sum_{k=1}^{n_i} s \beta_i \cdot e^{-s \beta_i} = s \cdot \gamma_i e^{-s \beta_i}. \end{aligned}$$

If  $s \geq t$  we get

$$\begin{aligned} \frac{1}{n} \text{Cov}[\tilde{\mathbf{N}}_{ns}^i, \mathcal{P}_{nt}] &= \frac{1}{n} \sum_{k=1}^{n_i} \text{Cov} \left[ \mathbb{1}_{\{t_{k,1} \leq ns\}}, \mathcal{P}_{nt}(k) \right] \\ &= \frac{1}{n} \sum_{k=1}^{n_i} \left( \mathbb{E}[\mathcal{P}_{nt}(k)] - \mathbb{E}[\mathcal{P}_{nt}(k)] \cdot \mathbb{E}[\mathbb{1}_{\{t_{k,1} \leq ns\}}] \right) = \frac{1}{n} \sum_{k=1}^{n_i} t \beta_i \cdot e^{-s \beta_i} = t \gamma_i e^{-s \beta_i}. \end{aligned}$$

The covariances  $\text{Cov}[\mathbf{X}_s^{(i)}, \mathbf{X}_t^{(i)}]$  have been computed in Proposition 3.1. Hence the finite-dimensional convergence of the processes of  $\{Z_{ns}^{(i)}; s \in [0, \infty)\}$  to  $\{\sqrt{\gamma_i} \mathbf{X}_s^{(i)}; s \in [0, \infty)\}$  for all  $i \in [J]$  and of  $\{Z_{ns}^{\mathcal{P}}; s \in [0, \infty)\}$  to  $\{\mathbf{B}_s^{\mathcal{P}}; s \in [0, \infty)\}$  imply the finite-dimensional convergence of  $\left\{ \left( Z_{ns}^{(1)}, \dots, Z_{ns}^{(J)}, Z_{ns}^{\mathcal{P}} \right); s \in [0, \infty) \right\}$  to  $\left\{ \left( \sqrt{\gamma_1} \mathbf{X}_s^{(1)}, \dots, \sqrt{\gamma_J} \mathbf{X}_s^{(J)}, \mathbf{B}_s^{\mathcal{P}} \right); s \in [0, \infty) \right\}$ . In order to prove tightness of the involved processes, let  $\epsilon > 0$ . Then there are sets  $\mathcal{K}_1, \dots, \mathcal{K}_{J+1} \subseteq \mathcal{D}([0, \infty), \mathbb{R})$  which are compact w.r.t. the Skorohod topology such that

$$\begin{aligned} \mathbb{P} \left[ (s \mapsto Z_{ns}^{(i)}) \in \mathcal{K}_i \text{ for all } n \in \mathbb{N} \right] &\geq 1 - \epsilon / (2J) \\ \mathbb{P} \left[ (s \mapsto Z_{ns}^{\mathcal{P}}) \in \mathcal{K}_{J+1} \text{ for all } n \in \mathbb{N} \right] &\geq 1 - \epsilon / 2. \end{aligned}$$

Hence

$$\mathbb{P} \left[ (s \mapsto (Z_{ns}^{(1)}, \dots, Z_{ns}^{(J)}, Z_{ns}^{\mathcal{P}})) \in \mathcal{K}_1 \times \dots \times \mathcal{K}_{J+1} \text{ for all } n \in \mathbb{N} \right] \geq 1 - \epsilon$$

and this implies tightness, so the claim is proved.  $\square$

Let  $r_i(n) := \sum_{l=1}^n \mathcal{L}_l^i$  for  $i \in [J]$  and  $n \in \mathbb{N}$ .

**Lemma 6.2.** *Denote for  $i \in [J]$*

$$Z_i^L(nq) = n^{-1/2} (r_i(nq) - nq \mathbb{E}[\mathcal{L}^i]) \quad \text{for } q \geq 0$$

and

$$Z_i^N(ns) = n^{-1/2} \left( \tilde{\mathbf{N}}_{ns}^i - n \gamma_i (1 - e^{-\beta_i s}) \right) \quad \text{for } s \geq 0.$$

If

$$\left( \left\{ Z_i^L(nq); q \geq 0 \right\} \right) \xrightarrow{\text{law}} \left( \left\{ \mathbf{Y}_i^L(q); q \geq 0 \right\} \right), \quad \text{as } n \rightarrow \infty \quad (16)$$

in the product topology of the Skorohod  $\mathbf{J}_1$  topology under  $\mathbb{P}$  for some stochastic process  $(\{Y_i^L(q); q \geq 0\}, \{Y_i^N(s); s \geq 0\})$ , then the same holds also under  $\mathbb{P}[\cdot | \text{Surv}^{\text{MGW}}]$  and under  $\mathbb{P}[\cdot | (\text{Surv}^{\text{MGW}})^c]$ .

*Proof.* Recall first the survival event  $\text{Surv}^{\text{MGW}} = \left\{ \sum_{i=1}^J Z_t^i > 0, \forall t \in \mathbb{N} \right\}$ , and  $Z_{\text{tot}}^{\text{MGW}} = \sum_{t=0}^{\infty} \sum_{i=1}^J Z_t^i$  represents the total number of individuals of the MGW  $(Z_t)_{t \in \mathbb{N}}$ . From (9) we know that, for the event  $\{Z_{\text{tot}}^{\text{MGW}} \leq A\}$  for  $A \geq 0$ , the variables  $\mathcal{L}_l^i$  in the definitions of  $r_i(t)$  coincide with  $\tilde{\mathcal{L}}_l^i$  for  $l > A$ , so that conditionally on  $\{Z_{\text{tot}}^{\text{MGW}} \leq A\}$  for  $i = 1, \dots, J$  the processes

$$\sum_{l=A}^{\lfloor nq \rfloor} \tilde{\mathcal{L}}_l^i = r_i(nq) - \sum_{l=1}^A \mathcal{L}_l^i$$

are independent of the MGW, and also independent among them. Moreover, for each vertex  $j = 1, \dots, n_i$  of type  $i$  let  $Y_l^{(j)}$ ,  $l \in \mathbb{N}$ , denote the arrival times of the infection attempts of vertex  $i$ . Then

$$\tilde{N}_t^i = \sum_{j=1}^{n_i} \max_{l \in \mathbb{N}} \mathbb{1}_{\{Y_l^{(j)} \leq t\}}.$$

Now the process

$$\tilde{N}_t^{A,i} = \sum_{j=1}^{n_i} \max_{l \in \mathbb{N}} \mathbb{1}_{\{A < Y_l^{(j)} \leq t\}}$$

is independent of  $\{Z_{\text{tot}}^{\text{MGW}} \leq A\}$  and

$$\sup_{t \in [0, \infty)} \left| \tilde{N}_t^{A,i} - \tilde{N}_t^i \right| \leq \sum_{j=1}^{n_i} \max_{l \in \mathbb{N}} \mathbb{1}_{\{Y_l^{(j)} \leq A\}} \leq \sum_{j=1}^{n_i} \sum_{l=1}^{\infty} \mathbb{1}_{\{Y_l^{(j)} \leq A\}}$$

and the right-hand side is a Poisson distributed random variable whose parameter does not depend on  $n$ . We have then that the event  $\{Z_{\text{tot}}^{\text{MGW}} \leq A\}$  and the process

$((Z_i^L(nq))_{q \geq 0}, (Z_i^N(ns))_{s \geq 0})$  are asymptotically independent, thus also  $(\text{Surv}^{\text{MGW}})^c$  and  $((Z_i^L(nq))_{q \geq 0}, (Z_i^N(ns))_{s \geq 0})$  are asymptotically independent. Therefore

$$\left( \begin{array}{l} \{n^{-1/2} (r_i(nq) - nq\mathbb{E}[\mathcal{L}^i]); q \geq 0\} \\ \{n^{-1/2} (\tilde{N}_{ns}^i - n\gamma_i(1 - e^{-\beta_i s})); s \geq 0\} \end{array} \right)$$

has the same limit in law under  $\mathbb{P}$ , under  $\mathbb{P}[\cdot | (\text{Surv}^{\text{MGW}})^c]$  and hence also under  $\mathbb{P}[\cdot | \text{Surv}^{\text{MGW}}]$ .  $\square$

If  $\sigma_i^2 := \text{Var}[\mathcal{L}^i] < \infty$  for all  $i = 1, \dots, J$ , we define the following constants:

$$\sigma_{\tilde{\tau}}^2 = \frac{\sum_{i=1}^J \sigma_i^2 \gamma_i (1 - e^{-\beta_i \theta}) + \sum_{i=1}^J (\mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i})^2 \sigma_{\mathcal{N},i}^2 + \theta - 2\theta \sum_{i=1}^J \gamma_i^{3/2} e^{-\beta_i \theta} \mathbb{E}[\mathcal{L}^i]}{(1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i \theta})^2} \quad (17)$$

with  $\sigma_{\mathcal{N},i}^2 = \gamma_i(1 - e^{-\beta_i\theta})e^{-\beta_i\theta}$ ,

$$\begin{aligned} \sigma_\tau^2 &= \frac{1}{(1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i\theta})^2} \left[ \sum_{i=1}^J \sigma_i^2 \gamma_i (1 - e^{-\beta_i\theta}) + \sum_{i=1}^J (\mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i})^2 \sigma_{\mathcal{N},i}^2 \right. \\ &\quad \left. + \theta \cdot \left( \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i\theta} \right)^2 \right] \\ &\quad - \frac{2\theta}{1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i\theta}} \left( \sum_{i=1}^J \gamma_i^{3/2} e^{-\beta_i\theta} \mathbb{E}[\mathcal{L}^i] \right) \left( \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i\theta} \right), \end{aligned} \quad (18)$$

and

$$\sigma_w^2 = c_w^2 \sum_{i=1}^J \sigma_i^2 \gamma_i (1 - e^{-\beta_i\theta}) + \sum_{i=1}^J (1 - c_w \mathbb{E}[\mathcal{L}^i])^2 \sigma_{\mathcal{N},i}^2 - 2 \left( \sum_{i=1}^J (1 - c_w \mathbb{E}[\mathcal{L}^i]) c_w \theta \gamma_i e^{-\beta_i\theta} \right) + c_w^2 \theta \quad (19)$$

with

$$c_w := \frac{1 - \sum_{i=1}^J \beta_i \gamma_i e^{-\beta_i\theta}}{1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i\theta}}.$$

**Theorem 6.3. Central Limit Theorems for  $\tau_n, \tilde{\tau}_n$ , and  $\tilde{\mathcal{N}}_{\tilde{\tau}_n}$ .**

Suppose Assumption 2 holds, and  $\sigma_i^2 := \text{Var}[\mathcal{L}^i] < \infty$  for all  $i = 1, \dots, J$ , i.e. all entries of the offspring distribution matrix  $L$  of the MGW have finite variances. Then we have

$$\begin{aligned} n^{-1/2}(\tilde{\tau}_n - n\theta) &\xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_{\tilde{\tau}}^2) && \text{as } n \rightarrow \infty \\ n^{-1/2}(\tau_n - n\theta) &\xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_\tau^2), && \text{as } n \rightarrow \infty \\ n^{-1/2}(\tilde{\mathcal{N}}_{\tilde{\tau}_n} - nw) &\xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_w^2), && \text{as } n \rightarrow \infty \end{aligned}$$

conditionally on  $\text{Surv}^{\text{MGW}}$ , i.e. under the probability measure  $\mathbb{P}[\cdot | \text{Surv}^{\text{MGW}}]$ , where  $w := \langle \mathbf{w}(\theta), \mathbf{1} \rangle = \sum_{i=1}^J \gamma_i (1 - e^{-\beta_i\theta})$  and  $\sigma_{\tilde{\tau}}^2, \sigma_\tau^2, \sigma_w^2$  are given by (17), (18), (19) respectively.

*Proof.* The random variables  $\mathcal{L}^1, \dots, \mathcal{L}^J$  are independent by construction. By Proposition 4.1  $(\mathcal{L}_l^i)_{l \in \mathbb{N}}$  is an i.i.d. copy of  $\mathcal{L}^i$ , and also independent of  $\mathbf{N}_t^i$ , so the random variables  $\mathcal{R}_t^i = \sum_{l=1}^{\mathbf{N}_t^i} \mathcal{L}_l^i$ , for  $i = 1, \dots, J$  are independent as well, and we can apply for each of the processes  $\{\sum_{l=1}^{nq} \mathcal{L}_l^i; q \geq 0\}$  the Donsker's invariance principle, and then a random time change.

For simplicity of notation, for  $i \in [J]$  and  $n \in \mathbb{N}$ , let us write  $r_i(n) := \sum_{l=1}^n \mathcal{L}_l^i$  so that  $\mathcal{R}_t^i = r_i(N_t^i)$ . By the Donsker's invariance principle we have

$$\left\{ \frac{1}{\sqrt{n}} (r_i(nq) - nq\mathbb{E}[\mathcal{L}^i]); q \geq 0 \right\} \xrightarrow{\text{law}} \left\{ \sigma_i \tilde{\mathbf{B}}_q^i; q \geq 0 \right\}, \quad \text{as } n \rightarrow \infty$$

in the Skorohod topology, where  $\tilde{\mathbf{B}}_q^i$  is a Brownian motion, and  $\sigma_i^2 = \text{Var}[\mathcal{L}^i]$ . Let  $\mathbf{X}_s^{(i)}$  be the centered Gaussian process from Proposition 3.1 which arises as scaling limit of  $\tilde{\mathbf{N}}_{ns}^i$ . Since

the random variables  $\mathcal{L}_l^i$  are independent of  $\mathbf{N}_{ns}^i$  we also have the following joint convergence:

$$\left( \begin{array}{l} \{n^{-1/2}(r_i(nq) - nq\mathbb{E}[\mathcal{L}^i]); q \geq 0\} \\ \{n^{-1/2}(\tilde{\mathbf{N}}_{ns}^i - n\gamma_i(1 - e^{-\beta_i s})); s \geq 0\} \end{array} \right) \xrightarrow{\text{law}} \left( \begin{array}{l} \{\sigma_i \tilde{\mathbf{B}}_q^i; q \geq 0\} \\ \{\sqrt{\gamma_i} \mathbf{X}_s^{(i)}; s \geq 0\} \end{array} \right)$$

in the product topology, as  $n \rightarrow \infty$ , where both components are equipped with the Skorohod topology. By Lemma 6.2 this convergence holds under  $\mathbb{P}[\cdot | \text{Surv}^{\text{MGW}}]$  as well. Using that, by definition we have the sequence of equalities

$$r_i(\tilde{\mathbf{N}}_{ns}^i) = r_i(\mathbf{N}_{\mathcal{P}_{ns}}^i) = \mathcal{R}_{\mathcal{P}_{ns}} = \sum_{l=1}^{\mathbf{N}_{\mathcal{P}_{ns}}^i} \mathcal{L}_l^i = \sum_{l=1}^{\tilde{\mathbf{N}}_{ns}^i} \mathcal{L}_l^i = \tilde{\mathcal{R}}_{ns}^i$$

and doing a random change of time ([2, Lemma p. 151]) in  $r_i(nq)$ , we obtain

$$\left\{ \frac{1}{\sqrt{n}} \left( \tilde{\mathcal{R}}_{ns}^i - \tilde{\mathbf{N}}_{ns}^i \mathbb{E}[\mathcal{L}^i] \right); s \geq 0 \right\} \xrightarrow{\text{law}} \left\{ \sigma_i \tilde{\mathbf{B}}_{w_i(s)}^i; s \geq 0 \right\},$$

where we recall that  $w_i(s) = \gamma_i(1 - e^{-\beta_i s})$ . Hence

$$\left\{ \frac{1}{\sqrt{n}} \left( \tilde{\mathcal{R}}_{ns}^i - n\gamma_i \mathbb{E}[\mathcal{L}^i](1 - e^{-\beta_i s}) \right); s \geq 0 \right\} \xrightarrow{\text{law}} \left\{ \sigma_i \tilde{\mathbf{B}}_{w_i(s)}^i + \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_s^{(i)}; s \geq 0 \right\}.$$

Since the processes  $r_i(t)$ ,  $i = 1, \dots, J$ , (and thus also the processes  $\tilde{\mathcal{R}}_t^i$ ) are independent, we also have convergence in law for the sum  $\tilde{\mathcal{R}}_{ns} := \sum_{i=1}^J \tilde{\mathcal{R}}_{ns}^i$ . Therefore the process

$$\left\{ \frac{1}{\sqrt{n}} \left( \tilde{\mathcal{R}}_{ns} - n \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i (1 - e^{-\beta_i s}) \right); s \geq 0 \right\}$$

converges in law to

$$\left\{ \sum_{i=1}^J \left( \sigma_i \tilde{\mathbf{B}}_{w_i(s)}^i + \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_s^{(i)} \right); s \geq 0 \right\}.$$

Now we know that at the end of the process (which occurs at time  $\tau_n$ ) we have

$$\tilde{\mathcal{R}}_{\tau_n} = \mathcal{R}_{\tau_n} = \tau_n = \mathcal{P}_{\tau_n}.$$

So  $\tilde{\tau}_n$  is the point at which the process  $\{\tilde{\mathcal{R}}_t - \mathcal{P}_t \mid t \in [0, \infty)\}$  becomes zero. Due to the independence mentioned above together with the scaling limit for  $Z_{ns}^{\mathcal{P}}$  from Theorem 6.1, we get the convergence in law

$$\left\{ \frac{1}{\sqrt{n}} \left( \tilde{\mathcal{R}}_{ns} - \mathcal{P}_{ns} + ns - n \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i (1 - e^{-\beta_i s}) \right); s \geq 0 \right\} \xrightarrow{\text{law}} \left\{ \sum_{i=1}^J \sigma_i \tilde{\mathbf{B}}_{w_i(s)}^i + \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_s^{(i)} - \mathbf{B}_s^{\mathcal{P}}; s \geq 0 \right\},$$



where  $\mathbf{B}^{\mathcal{P}}$  is a Brownian motion independent of the family  $\{\tilde{\mathbf{B}}^i\}_{1 \leq i \leq J}$  and correlated to  $\{\mathbf{X}^{(i)}\}_{1 \leq i \leq J}$  as in Theorem 6.1. By Theorem 5.1, both  $\tau_n/n$  and  $\tilde{\tau}_n/n$  converge in probability to  $\theta \mathbf{1}_{\{\text{Surv}_{\text{MGW}}\}}$ , with  $ns = \tilde{\tau}_n$  and we deduce that

$$n^{1/2} \cdot \underbrace{\left( \frac{\tilde{\tau}_n}{n} - \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i (1 - e^{-\beta_i \tilde{\tau}_n/n}) \right)}_{:= -f(\tilde{\tau}_n/n)} \xrightarrow{\mathcal{D}} \sum_{i=1}^J \sigma_i \tilde{\mathbf{B}}_{w_i(\theta)}^i + \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_\theta^{(i)} - \mathbf{B}_\theta^{\mathcal{P}}$$

as  $n \rightarrow \infty$ , with  $f(s)$  as defined in (15). Since  $f(\theta) = 0$ , a Taylor series expansion around  $\theta$  gives

$$\frac{\tilde{\tau}_n}{n} - \theta = \frac{-f(\frac{\tilde{\tau}_n}{n})}{-f'(\theta) - \frac{1}{2}(\frac{\tilde{\tau}_n}{n} - \theta)f''(A)}$$

for some  $A \in [\min\{\tilde{\tau}_n/n, \theta\}, \max\{\tilde{\tau}_n/n, \theta\}]$ . By the law of large numbers, Theorem 5.1, we have  $f'(\theta) + \frac{1}{2}(\frac{\tilde{\tau}_n}{n} - \theta)f''(A) \rightarrow f'(\theta)$  as  $n \rightarrow \infty$  in probability. Thus

$$n^{-1/2}(\tilde{\tau}_n - n\theta) \xrightarrow{\mathcal{D}} \frac{1}{-f'(\theta)} \cdot \left( \sum_{i=1}^J \sigma_i \tilde{\mathbf{B}}_{w_i(\theta)}^i + \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_\theta^{(i)} - \mathbf{B}_\theta^{\mathcal{P}} \right) \sim \mathcal{N}(0, \sigma_\tau^2),$$

with  $\sigma_\tau^2$  defined as in equation (17) and  $\sigma_{\mathcal{N},i}^2 = \gamma_i(1 - e^{-\beta_i \theta})e^{-\beta_i \theta}$  and this completes the first of the three claims.

Using that  $\tau_n = \mathcal{P}_{\tilde{\tau}_n}$  together with Theorem 6.1 we also get

$$\begin{aligned} n^{-1/2}(\tau_n - n\theta) &= n^{-1/2}(\mathcal{P}_{\tilde{\tau}_n} - \tilde{\tau}_n + \tilde{\tau}_n - n\theta) \\ &\xrightarrow{\mathcal{D}} \mathbf{B}_\theta^{\mathcal{P}} + \frac{1}{1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i \theta}} \cdot \left( \sum_{i=1}^J \sigma_i \tilde{\mathbf{B}}_{w_i(\theta)}^i + \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_\theta^{(i)} - \mathbf{B}_\theta^{\mathcal{P}} \right) \\ &\sim \mathcal{N}(0, \sigma_\tau^2) \end{aligned}$$

with  $\sigma_\tau^2$  given by equation (18) and this gives the second claim as well. Finally, for the third claim we proceed as follows:

$$n^{-1/2} \left( \tilde{\mathcal{N}}_{\tilde{\tau}_n} - n \langle \mathbf{w}(\theta), \mathbf{1} \rangle \right) = n^{-1/2} \left( \tilde{\mathcal{N}}_{\tilde{\tau}_n} - n \langle \mathbf{w}(\tilde{\tau}_n/n), \mathbf{1} \rangle + n \langle \mathbf{w}(\tilde{\tau}_n/n), \mathbf{1} \rangle - n \langle \mathbf{w}(\theta), \mathbf{1} \rangle \right).$$

Now Theorem 6.1 together with

$$n \langle \mathbf{w}(\tilde{\tau}_n/n) - \mathbf{w}(\theta), \mathbf{1} \rangle = \sum_{i=1}^J \beta_i \gamma_i e^{-\beta_i \theta} (\tilde{\tau}_n - n\theta) + O \left( n \left( \frac{\tilde{\tau}_n}{n} - \theta \right)^2 \right)$$

and

$$n^{1/2} \left( \frac{\tilde{\tau}_n}{n} - \theta \right)^2 \rightarrow 0$$

in probability as  $n \rightarrow \infty$  yields

$$\begin{aligned} &n^{-1/2} \left( \tilde{\mathcal{N}}_{\tilde{\tau}_n} - n \langle \mathbf{w}(\theta), \mathbf{1} \rangle \right) \\ &\xrightarrow{\mathcal{D}} \sum_{i=1}^J \sqrt{\gamma_i} \mathbf{X}_\theta^{(i)} - \frac{\sum_{i=1}^J \beta_i \gamma_i e^{-\beta_i \theta}}{1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i \theta}} \cdot \left( \sum_{i=1}^J \sigma_i \tilde{\mathbf{B}}_{w_i(\theta)}^i + \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \sqrt{\gamma_i} \mathbf{X}_\theta^{(i)} - \mathbf{B}_\theta^{\mathcal{P}} \right) \\ &\sim \mathcal{N}(0, \sigma_w^2), \end{aligned}$$

where we put

$$c_w := \frac{\sum_{i=1}^J \beta_i \gamma_i e^{-\beta_i \theta}}{1 - \sum_{i=1}^J \beta_i \gamma_i \mathbb{E}[\mathcal{L}^i] e^{-\beta_i \theta}}$$

and  $\sigma_w^2$  is defined as in (19) and this completes the whole proof.  $\square$

**Corollary 6.4.** *Since  $\tilde{\mathcal{N}}_t = \mathcal{N}_{\mathcal{P}_t}$  for  $t \in \mathbb{N}$  and  $\tau_n = \mathcal{P}_{\tilde{\tau}_n}$ , we have  $\tilde{\mathcal{N}}_{\tilde{\tau}_n} = \mathcal{N}_{\tau_n}$  and thus*

$$n^{-1/2}(\mathcal{N}_{\tau_n} - nw) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_w^2), \quad \text{as } n \rightarrow \infty$$

with  $w$  and  $\sigma_w^2$  as in Theorem 6.3.

## 7 Inhomogeneous versus homogeneous population

In this section we want to compare what happens in the uniform and non-uniform case. As we have seen in Theorem 5.1, the results are apparently undistinguishable. On the event  $\text{Surv}^{\text{MGW}}$ , we have

$$\frac{\tau_n}{n} \rightarrow \theta \quad \text{and} \quad \frac{N_{\tau_n}}{n} \rightarrow p, \quad \text{in probability, as } n \rightarrow \infty.$$

The difference lies in the computation of  $\theta$  and  $p$ .

**Homogeneous spread capacity.** Let us first deal with the case where all individuals have identically distributed spread capacities, let this law be  $\mathcal{L}$ . In a uniform population, all individuals have the same susceptibility and  $\theta$  is the solution of the equation

$$\theta = \mathbb{E}[\mathcal{L}] \cdot (1 - e^{-\theta}).$$

Let  $\theta_{un}$  be this solution: the corresponding fraction of infected individuals is then

$$p = p_{un} := 1 - e^{-\theta_{un}} = \theta_{un}/\mathbb{E}[\mathcal{L}].$$

In the non-uniform case,  $\theta$  is the solution of the equation

$$\theta = \mathbb{E}[\mathcal{L}] \left( 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i \theta} \right).$$

We call this solution  $\theta_\beta$  and we have  $p = p_\beta = 1 - \sum \gamma_i e^{-\beta_i \theta_\beta} = \theta_\beta/\mathbb{E}[\mathcal{L}]$ .

The following result says that in the non-uniform case, the epidemics dies out earlier and spreads in a smaller portion of the population.

**Proposition 7.1.** *For any fixed value of  $\mathbb{E}[\mathcal{L}] > 1$ , and any choice of the parameters  $\beta_i$ ,*

$$\theta_\beta \leq \theta_{un} \quad \text{and} \quad p_\beta \leq p_{un},$$

and equality holds if and only if  $\beta_i = 1$  for all  $i \in [J]$ .

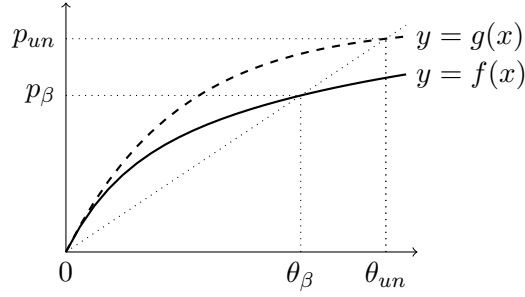


Figure 1: Inhomogeneous (solid) vs homogeneous (dashed)

*Proof.* If  $\beta_i = 1$  for all  $i \in [J]$  the statement is trivial. Let us assume that there exists  $\beta_i \neq 1$ . Define

$$f(x) = 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i x} \quad \text{and} \quad g(x) = 1 - e^{-x}.$$

Note that  $(\theta_{un}, p_{un})$  and  $(\theta_\beta, p_\beta)$  are the intersections of  $y = x/\mathbb{E}[\mathcal{L}]$  with  $y = g(x)$  and  $y = f(x)$ , respectively. It is enough to prove that  $f(x) \leq g(x)$  for all  $x \geq 0$  and equality holds only at 0 (see Figure 1). The two functions are clearly equal at 0. Since the exponential function is a strictly convex function and  $\sum_{i=1}^J \gamma_i = 1$ , for  $x > 0$  we get

$$f(x) = 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i x} < 1 - e^{-\sum_{i=1}^J \gamma_i \beta_i x} = 1 - e^{-x} = g(x),$$

where we have used  $\sum_{i=1}^J \gamma_i \beta_i = \sum_{i=1}^J \alpha_i = 1$ . This proves the statement.  $\square$

**Inhomogeneous spread capacity.** When types differ in the expected spread capacity  $\mathbb{E}[\mathcal{L}^i]$ , as well as in the susceptibility, then  $\theta$  is the solution of

$$\theta = f(\theta) := \sum_{i=1}^J \mathbb{E}[\mathcal{L}^i] \gamma_i (1 - e^{-\beta_i \theta}),$$

and  $p = 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i \theta}$ . We want to compare this case with the case where all  $\beta_i = 1$  (homogeneous susceptibility, inhomogeneous spread –  $\mathbb{E}[\mathcal{L}^i]$  are unaltered). If we define

$$g(x) = \sum_i \mathbb{E}[\mathcal{L}^i] \gamma_i (1 - e^{-x}),$$

the parameters  $\theta_\beta$  and  $\theta_{un}$  are the  $x$ -coordinates of the intersection of  $y = x$  with  $y = f(x)$  and  $y = g(x)$ , respectively. Depending on the choice of the parameters, the "homogeneous" case ( $\beta_i = 1$ ) may last longer or not, that is, both  $\theta_{un} > \theta_\beta$  and  $\theta_{un} < \theta_\beta$  are possible scenarios. Indeed there are examples where the inhomogeneous setting, compared to the homogeneous one, has:

- a) a shorter epidemics involving a smaller percentage of cases;

- b) a longer epidemics involving a smaller percentage of cases;
- c) a longer epidemics involving a larger percentage of cases.

Here are examples of the three possible situations described before.

- a) Take  $J = 3$ ,  $\gamma_i = 1/3$  for all  $i$ ,  $\mathbb{E}[\mathcal{L}^1] = 1$ ,  $\mathbb{E}[\mathcal{L}^2] = 2$ ,  $\mathbb{E}[\mathcal{L}^3] = 4$ ,  $\beta_1 = 2$ ,  $\beta_2 = 0.5$ ,  $\beta_3 = 0.5$ , then  $\theta_{un} \approx 2.0255$ ,  $\theta_\beta \approx 1.2139$ ,  $p_{un} \approx 0.8681$ ,  $p_\beta \approx 0.6073$ .
- b) Take  $J = 3$ ,  $\gamma_i = 1/3$  for all  $i$ ,  $\mathbb{E}[\mathcal{L}^1] = 1$ ,  $\mathbb{E}[\mathcal{L}^2] = 2$ ,  $\mathbb{E}[\mathcal{L}^3] = 4$ ,  $\beta_1 = 0.6$ ,  $\beta_2 = 1.2$ ,  $\beta_3 = 1.2$ :  $\theta_{un} \approx 2.0255$ ,  $\theta_\beta \approx 2.0703$ ,  $p_{un} \approx 0.8681$ ,  $p_\beta \approx 0.8482$ .
- c) Let  $J = 2$ ,  $\gamma_1 = \gamma_2 = 1/2$ ,  $\mathbb{E}[\mathcal{L}^1] = 0.2$ ,  $\mathbb{E}[\mathcal{L}^2] = 3.6$ ,  $\beta_1 = 0.6$ ,  $\beta_2 = 1.4$ ,  $\theta_{un} \approx 1.4578$ ,  $\theta_\beta \approx 1.6964$ ,  $p_{un} \approx 0.7672$ ,  $p_\beta \approx 0.7728$ .

Notice that in the inhomogeneous setting we cannot have a shorter epidemics involving a larger percentage of cases. Indeed, if  $\theta_\beta < \theta_{un}$ , then the convexity of the exp-function implies

$$p_\beta = 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i \theta_\beta} < 1 - e^{-\sum_{i=1}^J \gamma_i \beta_i \theta_\beta} = 1 - e^{-\theta_\beta} < 1 - e^{-\theta_{un}} = 1 - \sum_{i=1}^J \gamma_i e^{-\beta_i \theta_{un}} = p_{un}.$$

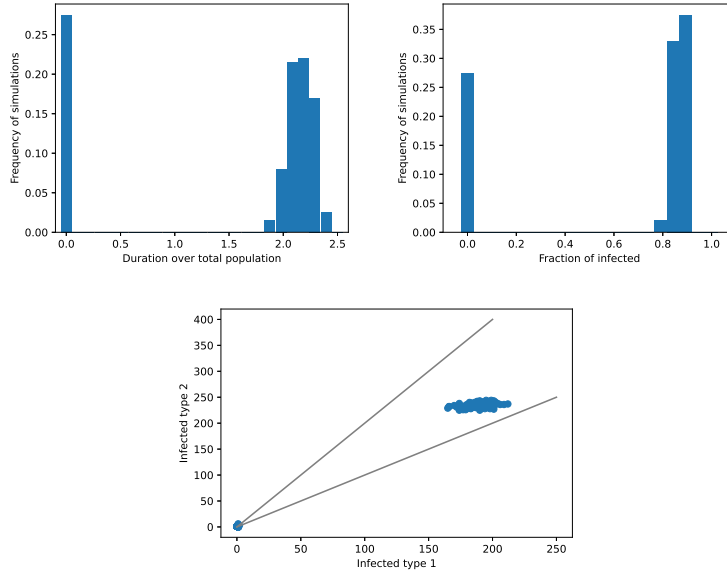


Figure 2: Equal groups, second more susceptible

One may also wonder what happens if we compare an epidemics which is inhomogeneous both in the spread and susceptibility, with the case where the  $\beta_i$  are unaltered and the spread capacities are homogeneous in the sense that all types have the same spread capacity  $\mathcal{L}$ , with  $\mathbb{E}[\mathcal{L}] = \sum_i \gamma_i \mathbb{E}[\mathcal{L}^i]$ . It is not difficult to find examples where the homogeneous case lasts longer (or less) than in the inhomogeneous case.

**Simulations.** Below we give the plots of some simulations, which agree with our theoretical results. In all simulations we have  $n = 500$  and we run 200 simulations. Moreover, the spread capacity is homogeneous with expectation 2.5 (1/2 being the probability of 4 attempts and 1/4 each the probability of 0 or 2 attempts). The probability of extinction of the associated GW process is 0.2711. There are two subpopulations which differ in susceptibility. We have the histograms of the frequencies of  $N_n(\tau_n)/n$  and of  $\tau_n/n$ . We also plot a scatterplot of the numbers of infected of the two types, with the two lines  $y = x \cdot \gamma_2/\gamma_1$  and  $y = x \cdot \beta_2/\beta_1$ .

In the first example depicted in Figure 2,  $\gamma_1 = 0.5$ ,  $\gamma_2 = 0.5$ ,  $\alpha_1 = 1/3$  and  $\alpha_2 = 2/3$ . The limiting values are  $\theta = 2.12261$  and  $p = 0.849044$ .

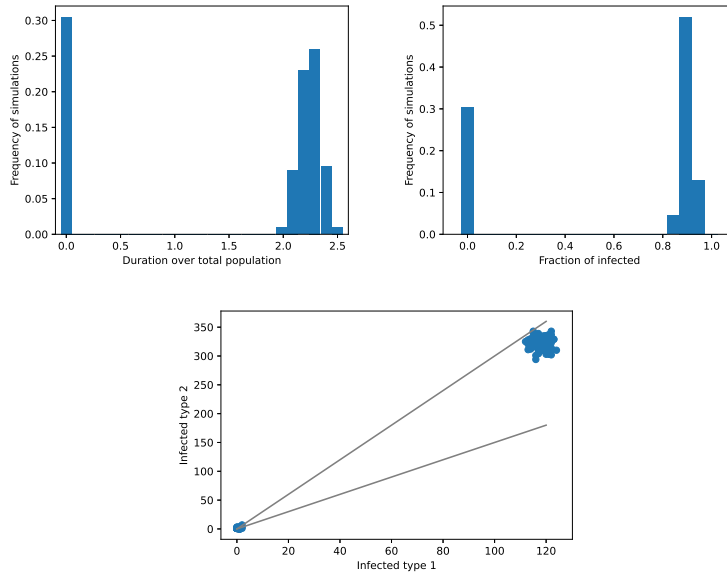


Figure 3: Larger group is more likely

In the second example depicted in Figure 3,  $\gamma_1 = 0.25$ ,  $\gamma_2 = 0.75$ ,  $\alpha_1 = 1/3$  and  $\alpha_2 = 2/3$ . The larger group is more likely to be chosen as a target, but single individuals in this group are less likely to be hit than individuals in the other group. The limiting values are  $\theta = 2.20202$  and  $p = 0.880807$ .

In the third example depicted in Figure 4,  $\gamma_1 = 0.75$ ,  $\gamma_2 = 0.25$ ,  $\alpha_1 = 1/3$  and  $\alpha_2 = 2/3$ . The larger group is less likely to be chosen as a target, and single individuals are less likely to be hit than individuals in the smaller group. The limiting values are  $\theta = 1.54723$  and  $p = 0.618904$ .

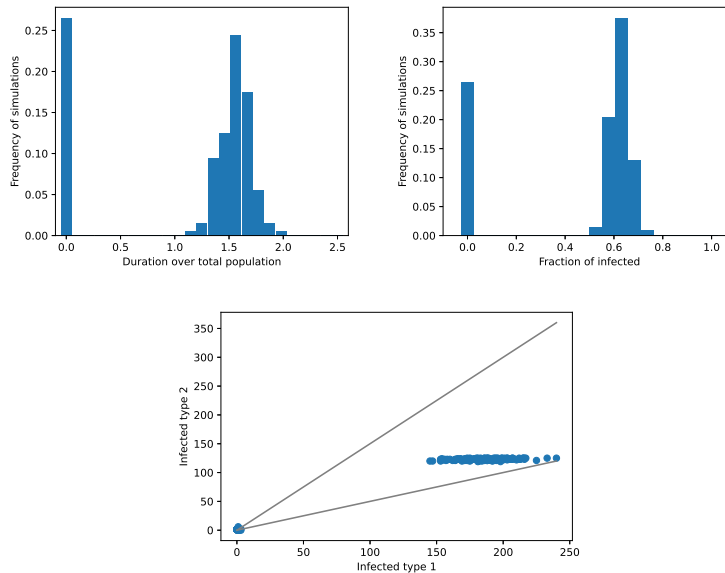


Figure 4: Larger group is less likely

## References

- [1] E. D. Andjel, N. Chabot, and E. Saada. A shape theorem for an epidemic model in dimension  $d \geq 3$ . *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(2):917–953, 2015.
- [2] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [3] T. Britton and E. Pardoux. *Stochastic Epidemic Models with Inference*, volume 2255 of *Lecture Notes in Math*. Springer, 2019.
- [4] F. Comets, F. Delarue, and R. Schott. Information transmission under random emission constraints. *Combin. Probab. Comput.*, 23(6):973–1009, 2014.
- [5] F. Comets, C. Gallesco, S. Popov, and M. Vachkovskaia. Constrained information transmission on Erdős-Rényi graphs. *Markov Process. Related Fields*, 22(1):111–138, 2016.
- [6] J. T. Cox and R. Durrett. Limit theorems for the spread of epidemics and forest fires. *Stochastic Process. Appl.*, 30(2):171–191, 1988.
- [7] T. E. Harris. *The Theory of Branching Processes*. Die Grundlehren der Mathematischen Wissenschaften, Band 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

- 
- [8] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.
- [9] T. G. Kurtz, E. Lebensztayn, A. R. Leichsenring, and F. P. Machado. Limit theorems for an epidemic model on the complete graph. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:45–55, 2008.
- [10] S. P. Lalley, E. A. Perkins, and X. Zheng. A phase transition for measure-valued SIR epidemic processes. *Ann. Probab.*, 42(1):237–310, 2014.
- [11] E. Lebensztayn and M. A. Estrada. Laws of large numbers for the frog model on the complete graph. *J. Math. Phys.*, 60(12):123302, 15, 2019.
- [12] W. Vervaat. Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 23:245–253, 1972.
- [13] Y. Zhang. A shape theorem for epidemics and forest fires with finite range interactions. *Ann. Probab.*, 21(4):1755–1781, 1993.
- [14] M. E. Zhukovskii. The law of large numbers for the number of active particles in an epidemic model. *Theory Probab. Appl.*, 58(2):297–313, 2014.

DANIELA BERTACCHI, Department of Mathematics, University of Milano-Bicocca, Italy.  
daniela.bertacchi@unimib.it

JÜRGEN KAMPF, Department of Mathematics, University of Innsbruck, Austria.  
Juergen.kampf@uibk.ac.at

ECATERINA SAVA-HUSS, Department of Mathematics, University of Innsbruck, Austria.  
Ecatarina.Sava-Huss@uibk.ac.at

FABIO ZUCCA, Department of Mathematics, Polytechnic University of Milan, Italy.  
fabio.zucca@polimi.it