

# CELEBRATING 50 YEARS OF THE APPLIED PROBABILITY TRUST

Edited by  
S. ASMUSSEN, P. JAGERS, I. MOLCHANOV and L. C. G. ROGERS

Part 4. Random graphs and particle systems

## THE FRONT OF THE EPIDEMIC SPREAD AND FIRST PASSAGE PERCOLATION

SHANKAR BHAMIDI, *University of North Carolina*

Department of Statistics, University of North Carolina, Chapel Hill, USA.  
Email address: bhamidi@email.unc.edu

REMCO VAN DER HOFSTAD, *Eindhoven University of Technology*

Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513,  
5600 MB Eindhoven, The Netherlands. Email address: rhofstad@win.tue.nl

JÚLIA KOMJÁTHY, *Eindhoven University of Technology*

Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513,  
5600 MB Eindhoven, The Netherlands. Email address: j.komjathy@tue.nl



APPLIED PROBABILITY TRUST  
DECEMBER 2014

# THE FRONT OF THE EPIDEMIC SPREAD AND FIRST PASSAGE PERCOLATION

BY SHANKAR BHAMIDI, REMCO VAN DER HOFSTAD AND JÚLIA KOMJÁTHY

---

## Abstract

We establish a connection between epidemic models on random networks with general infection times considered in Barbour and Reinert (2013) and first passage percolation. Using techniques developed in Bhamidi, van der Hofstad and Hooghiemstra (2012), when each vertex has infinite contagious periods, we extend results on the epidemic curve in Barbour and Reinert (2013) from bounded degree graphs to general sparse random graphs with degrees having finite second moments as  $n \rightarrow \infty$ , with an appropriate  $X^2 \log^+ X$  condition. We also study the epidemic trail between the source and typical vertices in the graph.

*Keywords:* Flow; random graph; random network; epidemics on random graphs; first passage percolation; hop count; interacting particle system

2010 Mathematics Subject Classification: Primary 60C05; 05C80; 90B15

---

## 1. Introduction and models

We consider the spread of an epidemic on the configuration model (of vertices and edges) with independent and identically distributed (i.i.d.) infection times having a general continuous distribution, and an infinite contagious period for each vertex. We describe the link between first passage percolation on sparse random graph models [6, 8], and general epidemics on the configuration model of Barbour and Reinert [3]. The work in [6, 8] is more general in terms of the graph models allowed, but more restrictive in terms of the epidemic process, requiring the assumption of infinite contagious periods and i.i.d. infection times. On the other hand, in [3] Barbour and Reinert allowed for more general epidemic processes, but assumed the graphs have bounded degrees. The main result, Theorem 2.1 below, extends our earlier work [6, 7, 8] to the study of the epidemic curve in the spirit of [3] by describing how the infection sweeps through the system. We also investigate the *epidemic trail*, namely the number of individuals that spread the infection from source to destination. Branching process approximations for the epidemic process and stable-age distribution theory for the corresponding branching processes developed by Jagers and Nerman [16, 15, 23] play a critical role in the proof of the main result.

### 1.1. Configuration model

We first describe the model for the underlying network (i.e. graph) on which the epidemic process takes place. The configuration model  $\text{CM}_n(\mathbf{d})$  (see [10] or [25, Chapters 7 and 10]) on  $n$  vertices with degree sequence  $\mathbf{d} = \mathbf{d}_n = (d_1, \dots, d_n)$  is constructed as follows. Let  $[n] := \{1, 2, \dots, n\}$  denote the vertex set. For each vertex  $i \in [n]$ , attach  $d_i$  half-edges to  $i$ , notating them suitably as the set  $\mathcal{X}_i$  (so  $|\mathcal{X}_i| = d_i$ ) and their union  $\mathcal{L}_n := \bigcup_{i \in [n]} \mathcal{X}_i$ . Assume the total degree  $|\mathcal{L}_n| = \sum_{i \in [n]} d_i$  to be even (if not, as, for example, may occur when the degrees  $d_i$  are drawn independently from some common degree distribution  $D$ , select  $i$  uniformly at

random from  $[n]$  and increase  $d_i$  by 1). For any given half-edge  $x \in \mathcal{L}_n$ , let  $V(x)$  denote the vertex incident to  $x$ .

Start by pairing the half-edges uniformly at random, i.e. pick an arbitrary unpaired half-edge,  $x$  say, and pair  $x$  to another unpaired half-edge,  $p_x$  say, chosen uniformly at random, to form the undirected edge  $\{V(x), V(p_x)\}$ . Once paired, remove both  $x$  and  $p_x$  from the set of unpaired half-edges; note that  $p_{p_x} = x$ . Now repeat until all half-edges are paired. Denote the resulting random multigraph by  $\text{CM}_n(\mathbf{d})$ . While self-loops (when  $V(x) = V(p_x)$ ) and multiple edges (when  $\{V(x'), V(p_{x'})\} = \{V(x), V(p_x)\}$  for some distinct pairs of half-edges  $\{x, p_x\}, \{x', p_{x'}\}$ ) can occur, under weak assumptions on the degree sequence as in, e.g. Condition 1.1 below, their number is a tight sequence as  $n \rightarrow \infty$  (see [17, Theorem 7.1] or [10] for more precise results in this direction). Write  $\mathcal{X}_i^\circ = \{x \in \mathcal{X}_i : p_x \in \mathcal{X}_i\}$  for the set of all half-edges in  $\mathcal{X}_i$  which pair to form self-loops;  $\mathcal{X}_i^\circ$  is empty for most  $i$ . For half-edges  $x'$  and  $x''$  for which  $(V(x'), V(x''))$  is an edge (and not a self-loop), we abuse notation and write the edge as  $(x', x'')$ .

We consider the configuration model for general degree sequences  $\mathbf{d}_n$ ; the model can be either deterministic or random, subject to mild regularity conditions as  $n \rightarrow \infty$ . To formulate these conditions, think of  $\mathbf{d}_n = (d_v)_{v \in [n]}$  as fixed and choose a vertex  $V_n$  uniformly from  $[n]$ . Then the distribution of  $D_n := d_{V_n} = |\mathcal{X}_{V_n}|$  is the degree of a uniformly chosen vertex  $V_n \in [n]$ , conditional on the degree sequence  $\mathbf{d}_n$ . To ensure that the majority of vertices are connected in the resulting graph, assume throughout that  $d_v \geq 2$  for each  $v \in [n]$  (see, e.g. [18] or [25, Chapter 10]). We make the following key assumption on the degree sequence.

**Condition 1.1.** (Degree regularity.) *The degrees  $D_n$  satisfy  $D_n \geq 2$  almost surely (a.s.) and, for some integer-valued random variable  $D$  with  $\mathbb{P}\{D > 2\} > 0$  and  $\mathbb{E}[D^2 \log^+(D)] < \infty$ ,*

$$\limsup_{n \rightarrow \infty} \mathbb{E}[D_n^2 \log^+(D_n)] = \mathbb{E}[D^2 \log^+(D)]. \tag{1.1}$$

The uniform integrability property (1.1) also implies that  $\mathbb{E}[D_n^p] \rightarrow \mathbb{E}[D^p]$  for  $p = 1, 2$ . When  $\mathbf{d}_n$  is itself random, we require the convergence in Condition 1.1 to hold in probability.

Let  $D_n^*$  denote a random variable with the *size-biased* distribution of  $D_n$  defined by

$$\mathbb{P}\{D_n^* = k\} = \frac{k \mathbb{P}\{D_n = k\}}{\mathbb{E}[D_n]}. \tag{1.2}$$

It is easily checked that the uniform integrability as in Condition 1.1 implies that  $\mathbb{E}[D_n^* - 1] \rightarrow \mathbb{E}[D^* - 1] = \mathbb{E}[D(D - 1)]/\mathbb{E}[D] < \infty$ , where  $D^*$  is the corresponding size-biased version of  $D$ . The assumption that  $D_n \geq 2$  and the nonvanishing variance  $\text{var}(D) > 0$  of the degrees implies that  $\mathbb{E}[D^* - 1] > 1$ .

### 1.2. Epidemic model

We now describe the infection model on  $\text{CM}_n(\mathbf{d})$ . Since multiple edges and self-loops play no role in the dynamics, replace multiple edges by a single edge and ignore self-loops. View each edge  $e = \{u, v\}$  in  $\text{CM}_n(\mathbf{d})$  as two directed edges  $(u, v)$  and  $(v, u)$ .

We consider an SIR (susceptible–infected–removed) process on  $\text{CM}_n(\mathbf{d})$ . Fix a continuous distribution  $G$  on  $\mathbb{R}_+$ . At time  $t = 0$ , start the infection at a uniformly chosen root vertex  $V_n$ . Each infected vertex infects its neighbours at times that start from the instant of its own infection and are i.i.d. with distribution  $G$ . This can be modelled by adding i.i.d. edge lengths  $X_e \sim G$  for every directed edge  $e = (v, u)$  from a vertex  $v \in \text{CM}_n(\mathbf{d})$  to a neighbour  $u$  of  $v$ . Suppose that each vertex  $v$  has an i.i.d. contagious period  $C_v \leq \infty$  after which it recovers, and that

once  $v$  gets infected, any neighbour  $u$  of  $v$  can become infected from  $v$  only if the infection time  $X_{(v,u)} < C_v$ . Denote the (possibly nonproper) tail distribution function of  $C$  by  $\bar{H}$ , i.e.  $\bar{H}(x) = \mathbb{P}\{C > x\}$ . Finally, assume that once a vertex is infected, it cannot be reinfected, so it transmits infection to its neighbours at most once.

Let  $\{\mathcal{F}_n(t)\}_{t \geq 0}$  denote this *epidemic process*, where, for any fixed  $t \geq 0$ ,  $\mathcal{F}_n(t)$  contains the entire  $\sigma$ -field of the process till time  $t$ ; thus,  $\mathcal{F}_n(t)$  contains information not only on the set and number of infected individuals by time  $t$ , but also on the entire sequence of transmissions on  $(0, t)$ .

Let  $|\mathcal{F}_n(t)|$  denote the total number of infected individuals by time  $t$ , and let  $|\mathcal{A}_n(t)|$  be the total size of the *coming generation*, by which we mean those vertices in the graph that at time  $t$  are not yet infected but have at time  $t$  an infectious neighbour that can infect them some time after  $t$ . (In fact, the set  $\mathcal{A}_n(t)$  will later be a set of half-edges, and when there are no self-loops or multiple edges, there is a bijection between vertices and half-edges in the coming generation; see Section 4.3.) Later we also define a related process  $\{\tilde{\mathcal{F}}_n(t), \tilde{\mathcal{A}}_n(t)\}_{t \geq 0}$  representing the collection of individuals that *would infect* a fixed target individual  $w$  by time  $t$  were the epidemic to start from them, and the corresponding *coming generation* in this process. We call this the *backward infection process* (see Section 4.3 for a precise definition).

### 2. Results

In this section we state our main results. Let  $P_n(s)$  denote the proportion of vertices infected by time  $s$ , that is, with  $\mathbf{1}_{\{\cdot\}}$  an indicator function,

$$P_n(s) = \frac{1}{n} \sum_{w \in [n]} \mathbf{1}_{\{\text{vertex } w \text{ infected by time } s\}}.$$

We also investigate the *number of infected individuals* on the path from the initial source of the infection to other vertices in  $\text{CM}_n(\mathbf{d})$ . Let the vertex  $w$  be infected at some time  $s \in (0, \infty)$ . Then, since infection times are continuous random variables, there is a.s. a *unique* path that realizes the infection between the root  $V_n$  and such another fixed vertex  $w \in [n]$ , which we call the *infection trail* to vertex  $w$ . Let  $H_n(w)$  denote the number of infectives along the trail to  $w$  (including  $V_n$  and  $w$ ), and define

$$P_n(s, h) = \frac{1}{n} \sum_{w \in [n]} \mathbf{1}_{\{\text{vertex } w \text{ infected by time } s, \text{ and } H_n(w) \leq h\}}.$$

Now fix  $n \geq 1$ . In Section 4 we describe how to couple the *epidemic process* and the *backward infection process*  $\{\mathcal{F}_n(t), \tilde{\mathcal{F}}_n(t)\}_{t \geq 0}$  to two independent Crump–Mode–Jagers processes  $\{\text{BP}_n(t), \tilde{\text{BP}}_n(t)\}_{t \geq 0}$ , where each individual from the first generation onwards produces a random number of children with distribution  $D_n^*$  and birth times that are i.i.d. random variables with cumulative distribution function  $G$ , and with a possibly finite contagious period  $C_v$  whose tail distribution we write as  $\bar{H}$ . The root has a slightly different offspring distribution from the rest of the population. Recall that  $\mathcal{A}_n(t)$  and  $\tilde{\mathcal{A}}_n(t)$  denote the *coming generation* in the infection processes. Condition 1.1 and standard results [15, 16] that we describe in Section 4 imply that there exists a constant  $\lambda_n > 0$  and limit random variables  $W_n, \tilde{W}_n > 0$  a.s. such that  $\exp\{-\lambda_n t\}(|\mathcal{A}_n(t)|, |\tilde{\mathcal{A}}_n(t)|) \xrightarrow{\text{a.s.}} (W_n, \tilde{W}_n)$  as  $t \rightarrow \infty$  (see (4.14) below), where  $\lambda_n$  satisfies the equation

$$\mathbb{E}[D_n^* - 1] \int_{\mathbb{R}_+} e^{-\lambda_n x} \bar{H}(x) \, dG(x) = 1. \tag{2.1}$$

Furthermore,

$$(W_n, \tilde{W}_n) \xrightarrow{D} (W, \tilde{W}) \quad \text{and} \quad \lambda_n \rightarrow \lambda \quad \text{as } n \rightarrow \infty,$$

where  $W$  and  $\tilde{W}$  are the corresponding limit random variables for the branching processes  $\{\mathbf{BP}(t), \tilde{\mathbf{BP}}(t)\}_{t \geq 0}$  described below in Sections 4.1 and 4.4, and  $\lambda$  satisfies (2.1) with  $D_n^*$  replaced by  $D^*$ . Let  $\Lambda$  be a standard Gumbel random variable independent of  $(S, \tilde{S}) := (-\log W)/\lambda, -(\log \tilde{W})/\lambda$ . Define the function

$$P(t) = \mathbb{P}\left\{\tilde{S} - \frac{\Lambda}{\lambda} + c \leq t\right\}, \quad t \in \mathbb{R},$$

where (cf. Section 4.7)  $c = \lambda^{-1} \log\{\mathbb{E}[D]\{\mathbb{E}[D^* - 2]\}^2 / \{\lambda m^* \mathbb{E}[D^* - 1]\}$  and the constant  $m^*$  is defined in (4.8) below. Finally, let  $\Phi(\cdot)$  denote the standard normal cumulative distribution function.

Our main theorem describes the asymptotic behaviour of the functions  $P_n(t)$  and  $P_n(t, h)$ , showing that these functions follow a deterministic curve with a random time shift corresponding to the initial phase of the infection.

**Theorem 2.1.** (Epidemic curve.) *Consider the epidemic spread with i.i.d. continuous infection times on the configuration model  $\text{CM}_n(\mathbf{d})$  and infinite contagious periods. Assuming that Condition 1.1 holds, for each fixed  $t \in \mathbb{R}$ , the proportion of infected individuals satisfies*

$$P_n\left(t + \frac{\log n}{\lambda_n}\right) \xrightarrow{D} P(t - S) = \mathbb{P}\{\log \tilde{W} + \log W + \Lambda > \lambda(c - t)\}. \tag{2.2}$$

Furthermore,

$$P_n\left(t + \frac{\log n}{\lambda_n}, \alpha_n \log n + x\sqrt{\beta \log n}\right) \xrightarrow{D} P(t - S)\Phi(x), \tag{2.3}$$

where  $\alpha_n$  and  $\beta$  are constants arising from the branching processes  $\text{BP}_n(\cdot)$  and  $\text{BP}(\cdot)$ , and are defined in (4.25) below.

**Remark 2.1.** Theorem 2.1 implies that the epidemic sweeps through the graph in an almost deterministic fashion, where the dependence on the initial phase of the epidemic appears only in the random shift  $S$  in (2.2). Furthermore, (2.3) implies that the number of infectives needed to reach a typical vertex in the graph is asymptotically *independent* of the time at which the vertex is infected. Much information can be read off from the shape of the curve  $t \mapsto P(t)$ . For example, the fact that the infection grows *exponentially* at the start is related to the fact that  $P(t)$  decays exponentially at  $t = -\infty$ , which in turn follows from the fact that  $\mathbb{P}\{-\Lambda/\lambda + c \leq t\}$  decays exponentially for large and negative  $t$ .

**Remark 2.2.** We believe that the connection between first passage percolation and epidemic models used to prove Theorem 2.1 can be easily generalized to the case with *finite* contagious times. In this regime, the forward and backward branching processes have *identical* Malthusian rates of growth but different limit random variables (see Section 4.2). This would extend results in [3], where one assumes that the degree of all vertices is bounded by some constant  $K$ , to the general configuration model satisfying Condition 1.1. In this context, the missing ingredient of the proof is the coupling error of the joint construction of the exploration process on the graph and the epidemic with finite contagious period.

### 3. Discussion

Here we briefly describe some connections with work related to ours.

- (a) *Epidemic models on networks.* There is an enormous literature on general epidemic models, their behaviour on various network models, and their connections to other dynamic processes; see [4, 24] and the references therein for a description of the motivations from statistical physics, and see [1, 13, 14] and the references therein for pointers to more rigorous results. First passage percolation or shortest path problems play an integral role in our study and we use results in [8] for the analysis of such processes on general sparse graph models with general edge distributions.
- (b) *Connection to results of Barbour and Reinert [3].* Barbour and Reinert determined the epidemic curve for a mean-field model with a Poisson number of infections. This case is equivalent to the infection spread on the *Erdős–Rényi random graph*. They generalized this to multitype epidemics, and concluded that a similar result holds for the configuration model in which every vertex has degree bounded by  $K$  for some fixed constant  $K \geq 1$ . This restriction allowed them to consider infection rates with arbitrary dependence on the number of possible infections created by a vertex. They used an associated multitype branching process in their analysis. Using the connection to first passage percolation, we show that similar results can be derived for any degree distribution satisfying Condition 1.1. The price for allowing unbounded degree distributions is that the infection rates between two individuals now cannot depend on the degree of the infectious vertex.
- (c) *Links to the mathematical epidemics literature.* The seminal paper by Kermack and McKendrick from 1927 [22] was the first to model epidemic spread mathematically: the spread here is *deterministic* using differential equations. The history of stochastic epidemic models is somewhat less clear. Stochastic versions were first formulated by Bartlett in the 1940s (see in particular Bartlett's book [5] for some earlier references and historical discussion, and also Kendall's Berkeley Symposium paper [21]). The first rigorous treatment using Crump–Mode–Jagers processes was presented in the 1995 paper [2], where Ball and Donnelly derived a strong approximation of arbitrary infection dynamics using Crump–Mode–Jagers processes for mean-field models. More recently, Volz [26] used differential equations to describe the evolution of the (Markovian) epidemic spread on random graphs. Subsequently, Decreusefond *et al.* [12] proved these results rigorously under a fifth moment condition on the degrees, while Bohman and Picolleli [9] studied epidemics on the configuration model with bounded degrees using a combined method of differential equations and branching process approximations. Recently, Janson *et al.* [19] analysed (Markovian) SIR epidemics on the configuration model with an arbitrary initial set of infective vertices, under a second moment assumption on the degrees.

#### 3.1. Scientific relevance and organization of the paper

In Section 4 we give the idea underlying the proof of Theorem 2.1 via the connection to first passage percolation. The intuitive idea is as follows. The expected proportion of vertices infected by time  $t$  equals the probability that a *random* individual is infected by time  $t$ , that is, it is of distance less than  $t$  from the infection source in the weighted graph. Hence, we first state the crucial Proposition 4.2 (see [8, Theorem 1.3]) about the typical distance between two uniformly picked individuals in the graph, and then exploit first- and second-moment methods on the empirical proportion of infected individuals to obtain the epidemic curve.

This first- and second-moment approach first appeared in [3, Theorem 2.8], where the authors performed it directly on the number of infected vertices within a given distance. The result is the epidemic curve, which then implicitly gives the distribution of typical distances in the graph.

The novelty of our method is that we completely reverse this procedure. Instead of directly analysing the number of infected individuals at a given time, we first use results in [8] about the distribution of distances in a weighted graph: in graphs that are locally tree-like as here, these distances can usually be understood via a branching process approximation of local neighbourhoods. The infection process then equals the size of metric balls around the infection source as a function of the radius  $r$  of the ball, in the random metric space given by the weighted graph. This size is encoded by the distribution of distances, since a random vertex falls into a metric ball of radius  $r$  with the probability that its distance is less than  $r$  from the source. Moreover, results about the distribution of the hop count (i.e. the number of vertices on the path of shortest weight) can be turned into results on the number of vertices on the epidemic trail.

Our paper therefore provides a bridge between two different areas of literature: distances in random metric spaces (e.g. weighted random graphs) and the spread of epidemics on random networks. We believe that our paper helps communication between these two fields: if there is a random metric space where typical distances are well understood then this immediately implies results on the behaviour of epidemic processes on the metric space, and vice versa.

Finally, at the end of the paper, in Sections 4.6 and 4.7, we explain the idea of how to determine the distribution of typical distances—we describe the core ideas of the proof of [8, Theorem 1.3] and the similar result given implicitly in [3]. Both couple the initial phases of the infection to two branching processes, and describe how these clusters connect up. We explain how the connection is formed based on the Bhamidi–van der Hofstad–Hooghiemstra [8] connection process that describes the limiting Poisson process of possible connection edges, of which the first point corresponds to the infection time. Essentially, the same Poisson process appears in the connection process of [3]; hence, we just highlight the differences and similarities between these two approaches.

## 4. Proofs

In this section we prove our main result, Theorem 2.1, starting in Section 4.1 with the connection between the exploration process on the configuration model and branching processes. In Section 4.2 we describe the relevant forward and backward continuous-time branching processes (CTBPs). In Section 4.3 we provide the coupling between the infection process on the configuration model and the CTBPs. In Section 4.4 we investigate asymptotics for the CTBPs and give the main proposition about the distribution of typical distances. Then in Section 4.5 these results are used to prove Theorem 2.1. In the penultimate Section 4.6 we give the intuitive idea of the proof of typical distances, i.e. we describe how the forward and backward CTBPs from two uniform vertices meet. Finally, in Section 4.7 we intuitively describe how Barbour and Reinert [3] derived asymptotics for the connection time.

### 4.1. Exploration on the configuration model and branching processes

Consider the epidemic process  $\mathcal{F}_n(\cdot)$  of Section 1.2 with i.i.d. infection times and possibly infinite i.i.d. contagious periods  $\{C_v \in (0, \infty], v \in [n]\}$  with tail distribution  $\bar{H}$ . We show how this is connected to a shortest path problem on  $\text{CM}_n(\mathbf{d})$ . To all directed edges  $\{(v, u) \in \text{CM}_n(\mathbf{d})\}$ , assign the respective i.i.d. random edge lengths  $X_{(v,u)} \sim G$ . The epidemic process can be thought of as a flow starting at vertex  $V_n$  at  $t = 0$  and spreading at rate 1 through the graph using the corresponding edge lengths. When the infection hits a nonsource vertex  $v$  at

time  $\sigma_v$ , thus infecting vertex  $v$ , each neighbour  $u$  of  $v$  (other than the neighbour that spread the infection to  $v$ ) becomes infected at time  $\sigma_v + X_{(v,u)}$  if  $X_{(v,u)}$  is less than  $C_v$ . Thus, the offspring distribution of new infections created by vertex  $v$ —these describe the number of infections and infection times created by  $v$  after  $\sigma_v$ —has the same distribution as the counting measure

$$\xi_v = \sum_{i=1}^{d_v-1} \delta_{X_i} \mathbf{1}_{\{X_i \leq C_v\}}, \tag{4.1}$$

where  $d_v$  denotes the degree of  $v$ ,  $X_i \sim G$  are i.i.d.,  $C_v \sim H$  is the contagious period of  $v$ , and, for Borel subsets  $A$  of  $\mathbb{R}_+$  and  $a \geq 0$ ,  $\delta_a(A)$  is the Dirac measure (see, e.g. [11, p. 382]).

4.1.1. *Local neighbourhoods in  $\text{CM}_n(\mathbf{d})$ .* The initial source  $V_n$  of the epidemic is chosen uniformly at random from  $[n]$  and, thus, has degree distribution  $D_n$  in Condition 1.1. We now describe the neighbourhood of this vertex. By the definition of  $\text{CM}_n(\mathbf{d})$ , we can construct  $\text{CM}_n(\mathbf{d})$  from  $V_n$  by sequentially connecting the half-edges of  $V_n$  to uniformly chosen unpaired half-edges. For any  $j \geq 1$ , let  $N_j^*(n) \approx np_j$  (by Condition 1.1) be the number of vertices with degree  $j$ , where we exclude  $V_n$ . Then, for fixed  $k \geq 1$ , the probability that the first half-edge of  $V_n$  connects to a vertex  $v \in [n] \setminus \{V_n\}$  with degree  $d_v = k + 1$  equals

$$\frac{(k + 1)N_{k+1}^*(n)}{\sum_{v \in [n]} d_v - 1} \approx \frac{(k + 1) \mathbb{P}\{D_n = k + 1\}}{\mathbb{E}[D_n]}. \tag{4.2}$$

If  $V_n$  connects to such a vertex then this neighbour has  $k$  remaining half-edges that can be used to connect to vertices in  $\text{CM}_n(\mathbf{d})$ . Thus, the forward degree of each neighbour  $V_n$  has a distribution that is approximately equal to  $D_n^*$ . The same is true for the remaining half-edges of  $V_n$  and, in fact, the above approximation continues to hold as long as the neighbourhood is not too large. Equations (4.1) and (4.2) suggest that the epidemic process can be approximated by the following branching process  $\{\text{BP}_n(t)\}_{t \geq 0}$  with label set  $\text{BP}_n(t) \subset \mathcal{V} := \{0\} \cup \bigcup_{n=1}^\infty \mathbb{N}^n$ .

- At time  $t = 0$ , start with a single individual  $\rho = 0$  whose offspring distribution is constructed as follows. First generate  $D_n$  possible children, and let  $\{X_i^0\}_{1 \leq i \leq D_n}$  be i.i.d. with distribution  $G$  and independent of  $C_0 \sim H$ . Then the children of  $\rho$  comprise the set  $(0, i)$  such that  $X_i < C_0$ , labelled in an arbitrary order. The interpretation is that each of these vertices is born at time  $X_i^0$ . Thus, the offspring distribution of the root can be represented as

$$\xi_0 := \sum_{i=1}^{D_n} \delta_{X_i} \mathbf{1}_{\{X_i \leq C_0\}}. \tag{4.3}$$

- Every other individual  $v \in \mathcal{V}$  born into the process  $\text{BP}_n(\cdot)$  has i.i.d. offspring distribution  $\xi_v$  with

$$\xi_v := \sum_{i=1}^{D_n^*(v)-1} \delta_{X_i^v} \mathbf{1}_{\{X_i^v \leq C_v\}}, \tag{4.4}$$

where  $C_v \sim H$  is the contagious period,  $D_n^*(v)$  has the size-biased distribution (1.2), and the  $X_i^v$  are i.i.d. with distribution function  $G$ . Thus, conditionally on  $D_n^*(v)$ , a vertex  $(v, i) \in \mathcal{V}$  is born at time  $X_i^v$  after vertex  $v$  is born if and only if  $X_i^v \leq C_v$ .

When  $C = \infty$  a.s., this coupling between  $\mathcal{F}_n(\cdot)$  and the corresponding branching process  $\text{BP}_n(\cdot)$  is carried out in [8, Section 4]. The details and corresponding error bounds are rather

technical; we give an intuitive idea in Section 4.3 and a rigorous error bound for their difference in Theorem 4.1 below.

**4.2. Forward and backward processes**

In the previous section we described the branching process approximation to the epidemic *forward* in time. Another key aspect of [3] is the study of the *backward* branching process. For a uniformly chosen vertex  $w \in CM_n(\mathbf{d})$  and fixed time  $t > 0$ , the vertex  $w$  is infected by time  $t$  precisely when there is a chain of infections leading to  $w$ . Hence, for large time  $t$ , we can ask, when  $w$  is in the infection process of one of its neighbours, whether that neighbour is in the infection process of one of his/her neighbours, etc., i.e. we can *trace the infection path back*. In [3], this leads to a new approximating branching process, the *backward* branching process with offspring process  $\tilde{\xi}[0, \infty]$ .

To see the difference between the offspring processes  $\xi$  going forward and  $\tilde{\xi}$  going backward, consider the case where all contagious periods are a.s. finite and i.i.d. with cumulative distribution function  $H$ . Then, as before,  $\xi = \sum_{i=1}^{D_n^*-1} \delta_{X_i} \mathbf{1}_{\{X_i < C_v\}}$  denotes the offspring of the forward process. On the other hand, in the backward process each individual has to be *in the contagious period of its children*, thus resulting in the offspring distribution

$$\tilde{\xi} = \sum_{i=1}^{D_n^*-1} \delta_{X_i} \mathbf{1}_{\{X_i < C_i\}}, \tag{4.5}$$

where  $C_i \sim H$  are i.i.d. Thus, the distribution of the offspring distribution (4.5) is different from that of the forward process  $\xi$ , even though they have the same expectations. In more complicated infection models, the backward process is substantially more complicated to describe.

The crucial observation is that in the case (4.5),  $\mathbb{E}_n[\tilde{\xi}(a, b)] = \mathbb{E}_n[\xi(a, b)]$  for all  $0 \leq a < b \leq \infty$ , so the corresponding expected reproduction measures  $\tilde{\mu}_n(dt)$  and  $\mu_n(dt)$  are the same for all  $n$ . This implies that, when  $C < \infty$  a.s., the distribution of the limiting martingale variables defined in (4.14) below are not the same in the forward and backward processes, but the growth rate  $\lambda_n$  and the multiplying constants for every characteristic under consideration (see (4.9) below) are the same.

Note that, when  $C = \infty$  a.s., which is what we assume for the rest of the paper, then the branching processes corresponding to the backward and forward processes are the same with offspring distribution

$$\xi = \sum_{i=1}^{D_n^*} \delta_{X_i}, \quad X_i \sim G \text{ are i.i.d. random variables.}$$

Therefore, from now on, we write  $\tilde{Q}$  for the quantity in the backward process that corresponds to the quantity  $Q$  in the forward process.

**4.3. Labelling the BP with half-edges on the configuration model**

We now construct  $CM_n(\mathbf{d})$  together with the epidemic process  $\mathcal{F}_n(\cdot)$  on it. First we construct the forward process by describing the sequence of new vertices that are infected and the times when these vertices become infected. Informally, at each step  $k = 0, 1, \dots$ , one of two things can happen.

**Event I.** A new vertex is infected, either as the root ( $k = 0$ ) or via an active half-edge that is incident to a currently infected vertex connecting to a half-edge incident to an uninfected vertex, whereupon all other half-edges incident to the newly infected vertex become active half-edges, and the two newly paired half-edges that merge to create the connection are removed.

**Event II.** Occasionally, two active half-edges (i.e. incident to some infected vertex) merge. The number of times this happens before time  $\tau_k$  is a tight random variable. This leaves the cluster of infected vertices unaltered because no new vertex results.

Here is a precise description of the construction. Recall from Section 1.1 that each half-edge  $x \in \mathcal{L}_n$  is incident to a vertex  $V(x) \in [n]$ , so  $x \in \mathcal{X}_{V(x)}$ , and in  $\text{CM}_n(\mathbf{d})$   $x$  pairs with  $p_x \in \mathcal{L}_n$ , and  $p_x$  is incident to  $V(p_x)$ .

The successive steps  $k = 0, 1, \dots$  concern the pairings, one at each step  $k \geq 1$ , that occur between an active half-edge and all other edges not yet removed, and of the effects of those pairings on the vertices and sets of half-edges of  $\text{CM}_n(\mathbf{d})$ . Let  $\mathcal{F}'(\tau_k)$  denote the vertices infected by time  $\tau_k$ .

For  $k = 0$ , choose the infection root  $V_n \in [n]$  uniformly at random in  $[n]$ , set  $\tau_0 = 0$  and  $\mathcal{F}'(\tau_0) = \{V_n\}$ . Vertex  $V_n$  has the  $D_n$  offspring comprising  $x \in \mathcal{X}_{V_n}$  that are born immediately. Check whether any  $x \in \mathcal{X}_{V_n}$  belong to the self-loop set  $\mathcal{X}_{V_n}^\circ$ . The coming generation is  $\mathcal{A}_{\tau_0} = \mathcal{X}_{V_n} \setminus \mathcal{X}_{V_n}^\circ$ . Members of this initial set of active half-edges have residual times to birth given by  $\mathcal{B}_{\tau_0} := \{B_x(\tau_0)\}_{x \in \mathcal{A}_{\tau_0}}$ , where  $B_x(\tau_0) \sim G$  are i.i.d. For each  $x \in \mathcal{A}_{\tau_0}$ , the endpoint  $V(p_x)$  is revealed and infected at time  $X_x$  (if not infected earlier via some other infection chain). Write  $\mathcal{H}_{\tau_0} = \mathcal{L}_n \setminus \{x : V(x) = V_n\} = \mathcal{L}_n \setminus \mathcal{X}_{V_n}$  for the initial set of susceptible half-edges (i.e. half-edges incident to uninfected vertices).

For each  $k \geq 1$ , proceed recursively as below, starting from the three sets  $\mathcal{A}_{\tau_{k-1}}$  of active half-edges,  $\mathcal{H}_{\tau_{k-1}}$  of susceptible half-edges, and  $\mathcal{B}_{\tau_{k-1}}$  the residual times to birth of the active half-edges. A half-edge is ‘free’ if it is susceptible or active.

- Find the active half-edge  $x_k^*$  with shortest residual time to birth,  $B_k^* := \min \mathcal{B}_{\tau_{k-1}}$ , and pair it to a uniformly chosen free half-edge  $p_{x_k^*} \in \mathcal{H}_{\tau_{k-1}} \cup \mathcal{A}_{\tau_{k-1}}$ . Update the time to  $\tau_k := \tau_{k-1} + B_k^*$ . For the rest of this algorithm, the vertex  $v_k := V(p_{x_k^*})$  is either (I) newly infected, so  $\mathcal{F}'(\tau_k) = \mathcal{F}'(\tau_{k-1}) \cup \{v_k\}$ , or (II) already infected, so  $\mathcal{F}'(\tau_k) = \mathcal{F}'(\tau_{k-1})$ .
- In case (II),  $\mathcal{A}_{\tau_k} = \mathcal{A}_{\tau_{k-1}} \setminus \{x_k^*, p_{x_k^*}\}$ . In case (I), form a putative new set of active half-edges  $\mathcal{A}'_{\tau_k} = (\mathcal{A}_{\tau_{k-1}} \setminus \{x_k^*\}) \cup \mathcal{X}'_{v_k}$ , where  $\mathcal{X}'_{v_k} = \mathcal{X}_{v_k} \setminus ((p_{x_k^*}) \cup \mathcal{X}_{v_k}^\circ)$ . For each  $x \in \mathcal{X}'_{v_k}$ , if  $p_x \in \mathcal{A}'_{\tau_k}$  then delete both  $x$  and  $p_x$  from  $\mathcal{A}'_{\tau_k}$ . The set  $\mathcal{A}'_{\tau_k}$  left after checking all such  $x$  is  $\mathcal{A}_{\tau_k}$ .
- Refresh residual times to birth, i.e. reduce all  $B_x \in \mathcal{B}_{\tau_{k-1}}$  by  $B_k^*$ , and in case (I), also add the i.i.d. edge weights  $X_y$  for newly active half-edges  $y \in (\mathcal{X}'_{v_k} \cap \mathcal{A}_{\tau_k})$  so that

$$\mathcal{B}_{\tau_k} := \{B_x(\tau_{k-1}) - B_k^* : x \in (\mathcal{B}_{\tau_{k-1}} \setminus \{x_k^*\})\} \cup \begin{cases} \{X_y : y \in \mathcal{X}'_{v_k}\} & \text{in case (I),} \\ \emptyset & \text{in case (II).} \end{cases}$$

- Refresh the set of free half-edges:

$$\mathcal{H}_{\tau_k} = \mathcal{H}_{\tau_{k-1}} \setminus \begin{cases} \mathcal{X}_{v_k} & \text{in case (I),} \\ \emptyset & \text{in case (II).} \end{cases}$$

Let  $\{\mathcal{F}'_n(k)\}_{k \geq 0}$  denote the above discrete-time process. We assert the truth of the following lemma as being obvious by construction.

**Lemma 4.1.** (Epidemic exploration process.) *For any  $t > 0$ , set  $k(t) = \sup\{k : \tau_k \leq t\}$ . Let  $\mathcal{F}_n^*(t) := \mathcal{F}'_n(k(t))$ . Then, for the epidemic process on  $\text{CM}_n(\mathbf{d})$ , the distributional equality  $\{\mathcal{F}_n(t)\}_{t \geq 0} \stackrel{D}{=} \{\mathcal{F}_n^*(t)\}_{t \geq 0}$  holds.*

4.3.1. *Coupling to a branching process.* It was shown in [8, Section 2] that the epidemic process constructed above can be coupled to a branching process  $\mathbf{BP}_n(\cdot)$  for which the root has offspring distribution (4.3) and all other individuals have distribution (4.4) (both with  $C_v = \infty$ ). The intuitive idea is as follows. In occurrences of Events I and II, the former correspond to the creation of new vertices in both  $\mathcal{F}_n$  and  $\mathbf{BP}_n$ , and the latter to the creation of *artificial vertices* in  $\mathbf{BP}_n$ . Now let  $\mathbf{BP}$  denote the ( $n$ -independent) branching process where the offspring distributions  $D_n$  in (4.3) and  $D_n^*$  in (4.4) are replaced by their distributional limits  $D$  and  $D^*$ . Let  $d_{\text{TV}}(\cdot, \cdot)$  denote the total variation distance between these mass functions on  $\mathbb{N}$ . Define sequences  $\{t_n\}_{n \geq 1}$  and  $\{s_n\}_{n \geq 1}$  satisfying both  $t_n, s_n \rightarrow \infty$  ( $n \rightarrow \infty$ ) and

$$t_n = \frac{\log n}{\lambda_n}, \quad e^{\lambda s_n} d_{\text{TV}}(D_n^*, D^*) \rightarrow 0. \tag{4.6}$$

One of the main coupling results in [8] is the following proposition.

**Proposition 4.1.** ([8, Proposition 2.4].) *There exists a coupling of the processes  $\{\mathcal{F}_n(t)\}_{0 \leq t \leq s_n}$  and  $\{\mathbf{BP}(t)\}_{0 \leq t \leq s_n}$  such that*

$$\mathbb{P}\{\{\mathcal{F}_n(t)\}_{0 \leq t \leq s_n} \neq \{\mathbf{BP}(t)\}_{0 \leq t \leq s_n}\} \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{4.7}$$

Replacing  $\mathbf{BP}$  here by  $\mathbf{BP}_n$  yields a coupling between  $\mathcal{F}_n$  and  $\mathbf{BP}_n$  satisfying (4.7).

4.3.2. *Exploration of the backward infection process.* After time  $t_n^* = \frac{1}{2}(\log n)/\lambda_n$ , freeze the forward cluster. The half-edges ‘sticking out’ of this cluster, namely the set of active edges, are exactly those in the coming generation  $\mathcal{A}_{t_n^*}$ . Start labelling the backward process *conditional on the presence of the forward process*. This labelling is slightly different from the labelling of the forward cluster, since we also want to keep track of when we connect to a half-edge in the coming generation  $\mathcal{A}_{t_n^*}$ .

At each step  $k = 0, 1, \dots$ , three things can happen in the backward process: exactly one of Events I and II in the forward process and the following event.

**Event III.** Occasionally a half-edge in the backward cluster is paired to a half-edge in the coming generation of the forward cluster  $\mathcal{A}_{t_n^*}$ . Call this a *collision* of the two processes. These collisions are of the utmost importance, as they let the infection spread between source and destination.

Here now is a precise description of the construction; note that our concern is now with sets of edges as well as of half-edges.

For  $k = 0$ , pick the source of the backward infection  $\tilde{v}_0 := \tilde{V}_n \in [n] \setminus \{\mathcal{F}_n(t_n^*)\}$  uniformly, so  $\tilde{\mathcal{F}}(0) := \{\tilde{V}_n\}$  is the initial (backward process) set of infected vertices. This vertex has  $\tilde{D}_n = d_{\tilde{V}_n}$  offspring and is born immediately. Set  $\tilde{\tau}_0 = 0$ . Pair the  $\tilde{D}_n$  outgoing half-edges immediately, uniformly at random without replacement from  $\mathcal{A}_{t_n^*} \cup \mathcal{H}_{t_n^*}$ . Check whether any of these half-edges are merged amongst themselves creating self-loops (Event II) or collision edges (Event III). Set the collision edges and residual collision times, and the coming generation or *active edges* for Event I by

$$\begin{aligned} \mathcal{C}_0 &:= \{(y, p_y), B_{p_y}(t_n^*)\}: V(y) = \tilde{V}_n, p_y \in \mathcal{A}_{t_n^*}\}, \\ \tilde{\mathcal{A}}_0 &:= \{(y, p_y): V(y) = \tilde{V}_n, p_y \notin \mathcal{A}_{t_n^*}, V(p_y) \neq \tilde{V}_n\}. \end{aligned}$$

For Event III, any  $(y, p_y)$  with  $p_y \in \mathcal{A}_{t_n^*}$  forms an edge between  $\tilde{V}_n$  and the forward cluster. The forward cluster has already ‘eaten up’ some time from this edge: the remaining time on this edge is  $B_{p_y}(t_n^*)$ . Remove Event II pairs  $(y, p_y)$  from the set of active edges; they form self-

loops which are irrelevant for the epidemic spread. For Event I, the initial remaining times to birth  $\tilde{\mathcal{B}}_0 := \{B_x(\tilde{\tau}_0) : x \in \tilde{\mathcal{A}}_0\}$  with  $B_x(\tilde{\tau}_0) \sim G$  are i.i.d. random variables. For each  $y \in \tilde{\mathcal{A}}_0$ , the endpoint  $V(p_y)$  is revealed immediately, but infected only at time  $X_y$ . (Here and below, the ‘half-edge  $y \in \tilde{\mathcal{A}}$ ’ is short for the ‘half-edge  $y$  is paired with  $p_y$  and  $(y, p_y) \in \tilde{\mathcal{A}}$ .’)

The initial set of free half-edges consists of the free half-edges at  $t_n^*$  (i.e.  $\mathcal{A}_{t_n^*} \cup \mathcal{H}_{t_n^*}$ ) adjusted by removing the half-edges of  $V_n$  and their half-edge pair partners:

$$\tilde{\mathcal{H}}_0 = (\mathcal{A}_{t_n^*} \cup \mathcal{H}_{t_n^*}) \setminus (\{y : V(y) = \tilde{V}_n\} \cup \{p_y : V(y) = \tilde{V}_n\}).$$

For  $k \geq 1$ , the construction proceeds as follows, starting from the three sets,  $\tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}}$  of active edges,  $\tilde{\mathcal{H}}_{\tilde{\tau}_{k-1}}$  of free half-edges, and  $\tilde{\mathcal{B}}_{\tilde{\tau}_{k-1}}$  of residual times of birth of the active edges (this is a marginal set of  $\mathcal{C}_{\tilde{\tau}_{k-1}}$ ). This is described in the following algorithm.

- Find the active edge  $(\tilde{x}_k^*, p_{\tilde{x}_k^*}) \in \tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}}$  with shortest residual time to birth  $\tilde{B}_k^* = \min_{x \in \tilde{\mathcal{B}}_{\tilde{\tau}_{k-1}}} \{B_x(\tilde{\tau}_{k-1})\}$ . Update the time:  $\tilde{\tau}_k := \tilde{\tau}_{k-1} + \tilde{B}_k^*$ . Noting that, by definition,  $\tilde{v}_k := V(p_{\tilde{x}_k^*})$  is newly infected at  $\tilde{\tau}_k$ , update the set of infected vertices to  $\tilde{\mathcal{F}}(\tilde{\tau}_k) := \tilde{\mathcal{F}}(\tilde{\tau}_{k-1}) \cup \{\tilde{v}_k\}$ .
- Refresh the coming generation and the collision edges: sequentially pair all half-edges  $\{y : V(y) = \tilde{v}_k\}$ , excluding  $p_{\tilde{x}_k^*}$ , to a uniformly chosen half-edge  $p_y \in \tilde{\mathcal{H}}_{\tilde{\tau}_{k-1}} \cup \tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}}$ . The new sets of collision and active edges are defined by

$$\begin{aligned} \mathcal{C}_{\tilde{\tau}_k} &:= \mathcal{C}_{\tilde{\tau}_{k-1}} \cup \{(y, p_y), B_{p_y}(t_n^*) : V(y) = \tilde{v}_k, p_y \in \mathcal{A}_{t_n^*}\}, \\ \tilde{\mathcal{A}}_{\tilde{\tau}_k} &:= \tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}} \cup \{(y, p_y) : V(y) = \tilde{v}_k, p_y \notin \mathcal{A}_{t_n^*} \cup \tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}}\} \setminus \{x_k^*, p_{x_k^*}\}, \end{aligned}$$

namely, the new collision edges are those among the  $d_{\tilde{v}_k} - 1$  newly found half-edges each of whose partner half-edge is an active half-edge in the forward process and for which the remaining time on the edge is  $B_{p_y}(t_n^*)$ . If  $p_y \in \tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}}$  then Event II happens: we have found a cycle. If none of these is the case, then the edge  $(y, p_y)$  becomes an active edge with residual time to birth  $B_y = X_y \sim G$  that is independent of all previous randomness.

- Refresh the residual times to birth, i.e. subtract  $\tilde{B}_k^*$  from all residual times to birth and add the i.i.d. edge weights  $X_y$  for newly active edges (but, do not add the remaining time of collision edges, and also remove cycle edges):

$$\begin{aligned} \tilde{\mathcal{B}}_{\tilde{\tau}_k} &:= \{B_x(\tilde{\tau}_{k-1}) - \tilde{B}_k^* : x \in \tilde{\mathcal{B}}_{\tilde{\tau}_{k-1}} \setminus \{\tilde{x}_k^*\}\} \\ &\cup \{X_y : V(y) = \tilde{v}_k, y \neq p_{\tilde{x}_k^*}, p_y \notin \mathcal{A}_{t_n^*} \cup \tilde{\mathcal{A}}_{\tilde{\tau}_{k-1}}\}. \end{aligned}$$

- Refresh the set of free half-edges, i.e. remove the half-edges of  $\tilde{v}_k$  and their partner half-edges:  $\tilde{\mathcal{H}}_{\tilde{\tau}_k} = \tilde{\mathcal{H}}_{\tilde{\tau}_{k-1}} \setminus (\{y : V(y) = \tilde{v}_k\} \cup \{p_y : V(y) = \tilde{v}_k\})$ .

The main difference between the forward process and this process is that here we pair the new outgoing half-edges  $y \in \{1, \dots, d_{\tilde{v}_k} - 1\}$  immediately at the birth of  $\tilde{v}_k$ , and check whether this edge collides with the forward cluster or becomes active. (Hence, in the backward process, the pairs  $(x, p_x)$  form the coming generation.) The statement of Proposition 4.1 remains valid for this process as well, i.e. the coupling between the backward cluster and BP can be established.

**4.3.3. The total length of collisions.** A collision happens at time  $\tilde{\tau}_k$  for some  $k$  when the vertex  $\tilde{v}_k$  has a half-edge  $y$  with a partner half-edge  $p_y \in \mathcal{A}_{t_n^*}$  of the forward process. Since this is checked exactly at the time when  $\tilde{v}_k$  becomes infected, and there is still a residual time  $B_{p_y}(t_n^*)$  on this edge, the length of this connection is exactly  $t_n^* + B_{p_y}(t_n^*) + \tilde{\tau}_k$ . This procedure yields several possible connection paths occurring at different times; the path with minimal time is the

one achieving the infection, its weight is the time of infection of  $\tilde{V}_n$ , and the number of edges in this path is the epidemic trail.

Note that  $p_y$  is a uniformly picked half-edge from the coming generation  $\mathcal{A}_{t_n^*}$ ; hence, its residual time to birth  $B_{p_y^*}(t_n^*)$  converges to the empirical residual time to birth distribution in (4.16) below. Also, note that this is independent of the backward process infection time  $\tilde{\tau}_k$ .

**4.4. Branching processes**

In this section we set up the branching process theory on which we rely. This includes stable-age distribution theory as in [23]. Fix a point process  $\xi$  on  $\mathbb{R}_+$ , and consider a branching process  $\text{BP}(\cdot)$  whose vertex set is a subset of  $\mathcal{N} := \{0\} \cup \bigcup_{n=1}^\infty \mathbb{N}^n$ , started with one individual 0 at  $t = 0$  with each vertex having an i.i.d. copy of  $\xi$ . Here an individual is labelled  $x = (i_1, i_2, \dots, i_n)$  if  $x$  is the  $i_n$ th child of the  $i_{n-1}$ th child of  $\dots$  of the  $i_1$ th child of the root.

Let  $\xi(A)$  denote the number of points of  $\xi$  in the Borel set  $A \subset \mathbb{R}_+$ . Write  $\mu(A) = \mathbb{E}[\xi(A)]$  for the corresponding first moment measure. Assume that  $\mu(\cdot)$  is nonlattice, that there exists a Malthusian parameter  $\lambda \in (0, \infty)$  satisfying

$$\int_0^\infty e^{-\lambda t} \mu(dt) = 1,$$

and that  $\xi$  satisfies the following integrability conditions with this parameter  $\lambda$ :

$$m^* := \int_0^\infty t e^{-\lambda t} \mu(dt) < \infty, \quad \mathbb{E} \left[ \left( \int_0^\infty e^{-\lambda t} \xi(dt) \right) \log^+ \left( \int_0^\infty e^{-\lambda t} \xi(dt) \right) \right] < \infty. \tag{4.8}$$

For  $v \in \text{BP}$ , write  $\sigma_v$  for its birth time and  $\xi_v$  for its offspring process. Let  $\{\{\phi_v(\cdot)\}: v \in \text{BP}\}$  be a family of i.i.d. stochastic processes with  $\{\phi_v(t)\}_{t \geq 0}$  measurable with respect to the offspring distribution  $\xi_v$ ,  $\phi_v(t) = 0$  for  $t < 0$  and  $\phi_v(t) \geq 0$  for  $t \geq 0$ . The interpretation of such a functional, often called a *characteristic* [16, 15, 23], is that it assigns a score  $\phi_v(t)$  when vertex  $v$  has age  $t$ . Write  $\phi := \phi_0$  to denote this process for the root. The branching process counted according to this characteristic is defined by

$$Z_t^\phi := \sum_{x \in \text{BP}(t)} \phi_x(t - \sigma_x).$$

Theorem 5.4 and Corollary 5.6 of [23] show that there exists a random variable  $W \geq 0$  with  $\mathbb{E}[W] = 1$  such that, for any characteristic  $\phi$  satisfying mild integrability conditions,

$$e^{-\lambda t} Z_t^\phi \rightarrow W \frac{\int_0^\infty e^{-\lambda t} \mathbb{E}[\phi(t)] dt}{m^*} \quad \text{a.s.} \tag{4.9}$$

Moreover, for two characteristics  $\phi_1$  and  $\phi_2$ ,

$$\frac{Z_t^{\phi_2}}{Z_t^{\phi_1}} \rightarrow \frac{\int_0^\infty e^{-\lambda t} \mathbb{E}[\phi_2(t)] dt}{\int_0^\infty e^{-\lambda t} \mathbb{E}[\phi_1(t)] dt} \quad \text{a.s. on } \{W > 0\}. \tag{4.10}$$

Now we apply this general theory to our epidemic-exploration process on  $\text{CM}_n(\mathbf{d})$ .

Since the BP results above hold both for finite and infinite infectious periods, and, for finite infectious period, only the joint construction of the exploration process and the epidemic spread is missing, we state results here (for possible future reference) for arbitrary i.i.d. infectious periods distributed as  $\mathbb{P}\{C \geq x\} = \bar{H}(x)$ . For  $C = \infty$ , we can set  $\bar{H}(x) \equiv 1$  everywhere below. (Referring to Remark 2.2, our hope is that, eventually, the connection between first passage percolation and epidemics used in this paper can be extended to epidemic models with finite contagious periods.)

First we fix  $n$ . Recall (Section 4.1) that the epidemic process  $\mathcal{F}_n(\cdot)$  on  $\text{CM}_n(\mathbf{d})$  is approximated by a branching process  $\text{BP}_n$  with offspring process  $\xi = \sum_{i=1}^{D_n^*} \delta_{X_i}$ . There is a slight modification for the distribution of the root; however, this does not effect the limit theorems above (other than the limit random variable having  $\mathbb{E}[W_n] \neq 1$ ). Recall the Malthusian rate of growth parameter  $\lambda_n$  from (2.1). The other parameters (with fixed  $n$ ) are calculated as

$$\begin{aligned} \mu_n(0, t] &:= \mathbb{E}(D_n^* - 1) \int_0^t \bar{H}(x) G(dx), \\ m_n^* &= \mathbb{E}[D_n^* - 1] \int_0^\infty t e^{-\lambda_n t} \bar{H}(t) G(dt), \quad \mu_n(dt) := \mathbb{E}[D_n^* - 1] \bar{H}(t) G(dt). \end{aligned} \tag{4.11}$$

Here  $m_n^*$  is called the *mean of the stable-age distribution* or mean age at child bearing.

In order to establish the connection between two infected clusters in the graph, we need the size of the so-called *coming generation* (i.e. those individuals that are born after time  $t$  to a mother born before time  $t$ ), and the empirical distribution of the *residual time to birth* of a uniformly picked individual in the coming generation. Asymptotics for these objects are derived by choosing appropriate characteristics  $\phi$ . Fix  $s > 0$ . Letting  $\phi^s(t) := \xi[t + s, \infty)$  and using  $\mathcal{F}$  to denote the particles born into the branching process,  $Z_t^{\phi^s} = \sum_{x \in \mathcal{F}} \xi_x[t - \sigma_x + s, \infty)$  counts the number of children of already born individuals whose birth date is at least  $s$  time units from now. In particular, we write  $A_t^d := Z_t^{\phi^0} = \sum_{x \in \mathcal{F}} \xi_x[t - \sigma_x, \infty)$ , so that  $A_t^d$  counts the size of the coming generation (usually referred to as the *alive* individuals in the CTBP literature) in a BP with expected intensity measure  $\mu_n$  in (4.11). (We add the superscript ‘d’ to denote *delaying* the process by one generation, i.e. the root here also has a  $\mu_n$  offspring law.) Use (4.11) to compute that in our case,  $\mathbb{E}[\phi_0] = \mathbb{E}[D_n^* - 1] \int_t^\infty \bar{H}(x) G(dx)$ ; hence,

$$\begin{aligned} e^{-\lambda_n t} A_t^d &= e^{-\lambda_n t} Z_t^{\phi^0} \\ &\xrightarrow{\text{a.s.}} W_n^d \frac{\int_0^\infty e^{-\lambda_n t} dt \mathbb{E}[D_n^* - 1] \int_t^\infty \bar{H}(x) G(dx)}{m_n^*} \\ &= W_n^d \frac{\mathbb{E}[D_n^* - 1] \int_0^\infty \bar{H}(x) G(dx) - 1}{m_n^* \lambda_n} \\ &= W_n^d \frac{\mu_n(\infty) - 1}{m_n^* \lambda_n}. \end{aligned} \tag{4.12}$$

Now, in order to match the BP to the exploration process  $\mathcal{F}_n(t)$  on  $\text{CM}_n(\mathbf{d})$  so as to have the same reproduction function at the root, introduce the following BP via the size of the coming generation:

$$A_t := \sum_{i=1}^{D_n} (\mathbf{1}_{\{t < X_i < C_v\}} + A_{t-X_i}^{d,(i)} \mathbf{1}_{\{X_i < t \wedge C_v\}}).$$

Here the  $A_t^{d,(i)}$  are i.i.d. copies of  $A_t^d$  in (4.12), and  $A_t$  corresponds to  $|\mathcal{A}_n(t)|$ , i.e. the number of active half-edges in  $\mathcal{F}_n(t)$ . Multiplying by  $e^{-\lambda_n t}$  and using (4.12) gives the convergence

$$\begin{aligned} e^{-\lambda_n t} A_t &= e^{-\lambda_n t} \sum_{i=1}^{D_n} \mathbf{1}_{\{t < X_i < C_v\}} + \sum_{i=1}^{D_n} e^{-\lambda_n X_i} \mathbf{1}_{\{X_i < t \wedge C_v\}} (e^{-\lambda_n(t-X_i)} A_{t-X_i}^{d,(i)}) \\ &\xrightarrow{\text{a.s.}} \sum_{i=1}^{D_n} e^{-\lambda_n X_i} \mathbf{1}_{\{X_i < C_v\}} W_n^{d,(i)} \frac{\mu_n(\infty) - 1}{\lambda_n m_n^*}, \end{aligned} \tag{4.13}$$

where the  $W_n^{d,(i)}$  are i.i.d. copies of  $W_n^d$ . Since  $\mathbb{E}[e^{-\lambda_n X_i} \mathbf{1}_{\{X_i < C_v\}}] = 1/\mathbb{E}[D_n^* - 1]$  by (2.1), and  $X_i$  is independent of  $W_n^{d,(i)}$ , we can introduce the limit random variable  $W_n$  in (2.1); it is the unique solution of the stochastic identity

$$W_n := \sum_{i=1}^{D_n} e^{-\lambda_n X_i} \mathbf{1}_{\{X_i < C_v\}} W_n^{d,(i)} \frac{\mu_n(\infty) - 1}{\lambda_n m_n^*}. \tag{4.14}$$

In terms of this, (4.13) implies that

$$e^{-\lambda_n t} A_t \xrightarrow{\text{a.s.}} W_n \quad \text{with} \quad \mathbb{E}[W_n] = \frac{\mathbb{E}[D_n](\mu_n(\infty) - 1)}{\mathbb{E}[D_n^* - 1] \lambda_n m_n^*}. \tag{4.15}$$

For an infinite contagious period (i.e.  $C = \infty$  a.s.),  $\mu_n(\infty) - 1 = \mathbb{E}[D_n^* - 2]$ .

The ratio convergence in (4.10) and  $\mathbb{E}[\phi^s(t)] = \mathbb{E}[D_n^* - 1] \int_{t+s}^\infty \bar{H}(x) G(dx)$  implies that the empirical ‘residual time to birth’ distribution converges:

$$\begin{aligned} \frac{Z_t^{\phi^s}}{Z_t^{\phi^0}} &\xrightarrow{\text{a.s.}} \frac{\mathbb{E}[D_n^* - 1] \int_0^\infty e^{-\lambda_n t} dt \int_{t+s}^\infty \bar{H}(x) G(dx)}{\mathbb{E}[D_n^* - 1] \int_0^\infty e^{-\lambda_n t} dt \int_t^\infty \bar{H}(x) G(dx)} \\ &= \frac{\mathbb{E}[D_n^* - 1]}{\mu(\infty) - 1} \int_s^\infty (1 - e^{\lambda_n(s-x)}) \bar{H}(x) G(dx) \\ &:= 1 - F_R^{(n)}(s). \end{aligned} \tag{4.16}$$

This is the limiting probability that a uniformly picked individual from the ‘coming generation’ will be born after an extra  $s$  time units.

We have now set the stage for the branching processes that approximate the initial phase of the infection and the backward infection process. First we state the main proposition on which our proof of the epidemic curve Theorem 2.1 is based, giving that proof in Section 4.5, and then indicating how to prove the proposition in Sections 4.6 and 4.7. To this end, denote the infection time from vertices  $v$  to  $w$  by  $L_n(v, w)$ . (Note that part (a) of the proposition is part of [8, Theorem 1.2], and part (b) is a two-vertex analogue which can be proved in a similar way.) In its statement,  $\mathcal{G}_n(s)$  for  $s > 0$  denotes the  $\sigma$ -algebra of all vertices that are infected before time  $s$ , as well as all edge weights of the half-edges that are incident to such vertices. Thus, as opposed to  $\mathcal{F}_n(s)$  which has information only about the sequence of transmissions that have transpired before time  $s$ ,  $\mathcal{G}_n(s)$  also contains information about the ‘coming generation’ of infections.

**Proposition 4.2.** (a) *Let  $s_n$  be as in (4.6), and let  $\Lambda$  be a standard Gumbel random variable. For  $n \rightarrow \infty$ , the shortest infection path between two uniformly picked vertices  $V_n$  and  $\tilde{V}_n$  satisfies*

$$\mathbb{P}\left\{L_n(V_n, \tilde{V}_n) - \frac{\log n}{\lambda_n} + \frac{\log W_{s_n}}{\lambda_n} + \frac{\log \tilde{W}_{s_n}}{\lambda_n} < t \mid \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n(s_n)\right\} \xrightarrow{D} \mathbb{P}\left\{-\frac{\Lambda}{\lambda} + c < t\right\}.$$

(b) *Let  $V_n, \tilde{V}_n^{(1)}$ , and  $\tilde{V}_n^{(2)}$  be three independent uniform vertices in  $[n]$ , with respective forward and backward infection processes  $\mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n)$ , and  $\tilde{\mathcal{G}}_n^{(2)}(s_n)$ . Then, for  $n \rightarrow \infty$ ,*

$$\begin{aligned} &\mathbb{P}\left\{L_n(V_n, \tilde{V}_n^{(i)}) - \frac{\log n}{\lambda_n} + \frac{\log W_{s_n}}{\lambda_n} + \frac{\log \tilde{W}_{s_n}^{(i)}}{\lambda_n} < t, i = 1, 2 \mid \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n), \tilde{\mathcal{G}}_n^{(2)}(s_n)\right\} \\ &\xrightarrow{D} \left[\mathbb{P}\left\{-\frac{\Lambda}{\lambda} + c < t\right\}\right]^2. \end{aligned} \tag{4.17}$$

**4.5. Proof of Theorem 2.1**

In this section we use Proposition 4.2 to explain how to get the epidemic curve in Theorem 2.1 and complete its proof which is based on the key proposition proved below. Let  $s_n \rightarrow \infty$  as in Proposition 4.1, and set  $W_{s_n} = e^{-s_n \lambda_n} |\mathcal{A}_{s_n}|$ , where, as before,  $|\mathcal{A}_t|$  is the size of the coming generation of infected individuals at time  $t$ .

**Proposition 4.3.** (Epidemic curve with an offset.) *Under Condition 1.1, consider the epidemic spread with i.i.d. continuous infection times on the configuration model  $CM_n(\mathbf{d})$  and infinite contagious periods. For every  $t > 0$ ,*

$$P_n \left( t + \frac{\log n}{\lambda_n} - \frac{\log W_{s_n}}{\lambda_n}, \alpha_n \log n + x \sqrt{\beta \log n} \right) \xrightarrow{\mathbb{P}} P(t) \Phi(x). \tag{4.18}$$

*Proof of Theorem 2.1 subject to Proposition 4.3.* Fix  $x \in \mathbb{R}$ . Because  $t \mapsto P_n(t, v)$  is nondecreasing, and the limit  $t \mapsto P(t)$  in (4.18) is nondecreasing, continuous, and bounded, Proposition 4.3 implies that the convergence in (4.18) is *uniform* in  $t$ , i.e.

$$\sup_{s \in \mathbb{R}} \left| P_n \left( s + \frac{\log n}{\lambda_n} - \frac{\log W_{s_n}}{\lambda_n}, \alpha_n \log n + x \sqrt{\beta \log n} \right) - P(s) \Phi(x) \right| \xrightarrow{\mathbb{P}} 0.$$

Applying this to  $s = t + (\log W_{s_n})/\lambda_n$ , we thus obtain

$$P_n \left( t + \frac{\log n}{\lambda_n}, \alpha_n \log n + x \sqrt{\beta \log n} \right) = P \left( t + \frac{\log W_{s_n}}{\lambda_n} \right) \Phi(x) + o_{\mathbb{P}}(1).$$

Since  $(\log W_{s_n})/\lambda_n \xrightarrow{D} (\log W)/\lambda = -S$  and  $t \mapsto P(t)$  is continuous, Theorem 2.1 is proved.

*Proof of Proposition 4.3.* Here we use Proposition 4.2, performing a second moment method on  $P_n(t + (\log n)/\lambda_n - (\log W_{s_n})/\lambda_n, \alpha_n \log n + x \sqrt{\beta \log n})$ , conditionally on  $\mathcal{G}_n(s_n)$ . To simplify the notation, set  $x = \infty$ ,  $S_n = -(\log W_{s_n})/\lambda_n$ ,  $\tilde{S}_n = -(\log \tilde{W}_{s_n})/\lambda_n$ . We show that

$$\begin{aligned} \mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \middle| \mathcal{G}_n(s_n) \right] &\xrightarrow{\mathbb{P}} P(t) \\ \text{and } \mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right)^2 \middle| \mathcal{G}_n(s_n) \right] &\xrightarrow{\mathbb{P}} P(t)^2. \end{aligned} \tag{4.19}$$

Equation (4.19) implies that, conditionally on  $\mathcal{G}_n(s_n)$ ,  $P_n(t + (\log n)/\lambda_n + S_n) \xrightarrow{\mathbb{P}} P(t)$ , as required. We start by identifying the first conditional moment. For this, note that

$$\begin{aligned} \mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \middle| \mathcal{G}_n(s_n) \right] &= \frac{1}{n} \sum_{w \in [n]} \mathbb{P} \left\{ L_n(V_n, w) \leq t + \frac{\log n}{\lambda_n} + S_n \middle| \mathcal{G}_n(s_n) \right\} \\ &= \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n \leq t \middle| \mathcal{G}_n(s_n) \right\}, \end{aligned}$$

where  $\tilde{V}_n^{(1)}$  is a uniform vertex independent of  $V_n$  and  $L_n(v, w)$  is the time for the infection starting from  $v$  to reach  $w$ . Thus, in the infinite-contagious period case,  $L_n(v, w)$  is nothing but the first passage time from  $v$  to  $w$ . For  $s > 0$ , let  $\tilde{\mathcal{G}}_n^{(1)}(s)$  denote the  $\sigma$ -algebra of all vertices that would infect  $\tilde{V}_n^{(1)}$  within time  $s$  if the infection started from them at time 0, as well as all edge weights of the edges that are incident to such vertices. Thus, by the argument about

the backward process in Section 4.2, these vertices are the same as the vertices that would be infected before time  $s$  from an infection started from  $\tilde{V}_n^{(1)}$  in the backward process.

Write  $\tilde{W}_{s_n} = e^{-\lambda_n s_n} |\tilde{\mathcal{A}}_{s_n}|$ , where  $\tilde{\mathcal{A}}_t$  consists of those half-edges that are in the coming generation of the backward infection process of  $\tilde{V}_n^{(1)}$  at time  $t$ . We now further condition on  $\mathcal{G}_n^{(1)}(s)$ , and obtain

$$\begin{aligned} & \mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \middle| \mathcal{G}_n(s_n) \right] \\ &= \mathbb{E} \left[ \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n \leq t \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n) \right\} \middle| \mathcal{G}_n(s_n) \right]. \end{aligned}$$

By Proposition 4.2, there exists a constant  $c > 0$  such that

$$\mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n - \tilde{S}_n \leq t \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n) \right\} \xrightarrow{\mathbb{P}} \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t \right\}.$$

Again, since  $t \mapsto \mathbb{P}\{-\Lambda/\lambda + c \leq t\}$  is increasing and continuous, the above convergence even holds uniformly in  $t$ , i.e.

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n - \tilde{S}_n \leq t \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n) \right\} - \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t \right\} \right| \xrightarrow{\mathbb{P}} 0.$$

Thus,

$$\begin{aligned} & \mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n) \right] \\ &= \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n \leq t \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n) \right\} \\ &= \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t - \tilde{S}_n \middle| \tilde{\mathcal{G}}_n^{(1)}(s_n) \right\} + o_{\mathbb{P}}(1), \end{aligned}$$

and since  $\tilde{W}_{s_n} \xrightarrow{\mathbb{P}} \tilde{W}$  and  $t \mapsto \mathbb{P}\{-\Lambda/\lambda + c \leq t\}$  is continuous and bounded,

$$\mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \middle| \mathcal{G}_n(s_n) \right] \xrightarrow{\mathbb{P}} \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t - \tilde{S} \right\} = P(t).$$

By bounded convergence, this also implies that

$$\mathbb{E} \left[ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \middle| \mathcal{G}_n(s_n) \right] \xrightarrow{\mathbb{P}} P(t),$$

which completes the proof of the convergence of the first moment.

We use similar ideas to identify the second conditional moment. Start by writing

$$\begin{aligned} & \mathbb{E} \left[ \left\{ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \right\}^2 \middle| \mathcal{G}_n(s_n) \right] \\ &= \frac{1}{n^2} \sum_{i, j \in [n]} \mathbb{P} \left\{ L_n(V_n, i) + \frac{\log n}{\lambda_n} + S_n \leq t, L_n(V_n, j) + \frac{\log n}{\lambda_n} + S_n \leq t \middle| \mathcal{G}_n(s_n) \right\} \\ &= \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) + \frac{\log n}{\lambda_n} + S_n \leq t, L_n(V_n, \tilde{V}_n^{(2)}) + \frac{\log n}{\lambda_n} + S_n \leq t \middle| \mathcal{G}_n(s_n) \right\}, \end{aligned}$$

where  $V_n, \tilde{V}_n^{(1)},$  and  $\tilde{V}_n^{(2)}$  are three i.i.d. uniform vertices in  $[n]$ . For  $s > 0$  and  $j \in \{1, 2\}$ , let  $\tilde{\mathcal{G}}_n^{(j)}(s)$  denote the  $\sigma$ -algebra of all vertices that would infect  $\tilde{V}_n^{(j)}$  within time  $s$  if the infection started from them at time 0, as well as all edge weights of the edges that are incident to such vertices. Thus, these vertices are the same as the vertices that would be infected before time  $s$  in the backward infection process started from  $\tilde{V}_n^{(j)}$ .

Write  $\tilde{W}_{s_n}^{(j)} = e^{-\lambda_n s_n} | \tilde{\mathcal{A}}_{s_n}^{(j)} |$  and  $\tilde{S}_n^{(i)} = -\log W^{(i)} / \lambda_n$ . Now condition also on  $\tilde{\mathcal{G}}_n^{(1)}(s_n)$  and  $\tilde{\mathcal{G}}_n^{(2)}(s_n)$ :

$$\begin{aligned} & \mathbb{E} \left[ \left( P_n \left\{ t + \frac{\log n}{\lambda_n} - S_n \right\} \right)^2 \middle| \mathcal{G}_n(s_n) \right] \\ &= \mathbb{E} \left[ \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n \leq t, \right. \right. \\ & \quad \left. \left. L_n(V_n, \tilde{V}_n^{(2)}) - \frac{\log n}{\lambda_n} - S_n \leq t \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n), \tilde{\mathcal{G}}_n^{(2)}(s_n) \right\} \middle| \mathcal{G}_n(s_n) \right]. \end{aligned}$$

By (4.17), there exists a constant  $c > 0$  such that

$$\begin{aligned} & \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(i)}) - \frac{\log n}{\lambda_n} - S_n - \tilde{S}_n^{(i)} \leq t, i = 1, 2 \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n), \tilde{\mathcal{G}}_n^{(2)}(s_n) \right\} \\ & \xrightarrow{\mathbb{P}} \left( \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t, -\frac{\Lambda'}{\lambda} + c \leq t \right\} \right)^2 \\ &= \left( \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t \right\} \right)^2, \end{aligned}$$

since  $\Lambda$  and  $\Lambda'$  are two independent Gumbel variables. Now the argument proving convergence of the first moment can be repeated to yield

$$\mathbb{E} \left[ \left\{ P_n \left( t + \frac{\log n}{\lambda_n} + S_n \right) \right\}^2 \middle| \mathcal{G}_n(s_n) \right] \xrightarrow{\mathbb{P}} [P(t)]^2,$$

and this completes the proof of the convergence of the second moment for  $x = \infty$ .

The extension to  $x < \infty$  follows in an identical fashion, now using [8, Theorem 2.2], i.e.

$$\begin{aligned} & \mathbb{P} \left\{ L_n(V_n, \tilde{V}_n^{(1)}) - \frac{\log n}{\lambda_n} - S_n - \tilde{S}_n \leq t, \right. \\ & \quad \left. H_n(V_n, \tilde{V}_n^{(1)}) \leq \alpha_n \log n + x \sqrt{\beta \log n} \middle| \mathcal{G}_n(s_n), \tilde{\mathcal{G}}_n^{(1)}(s_n) \right\} \\ & \xrightarrow{\mathbb{P}} \mathbb{P} \left\{ -\frac{\Lambda}{\lambda} + c \leq t \right\} \Phi(x), \end{aligned}$$

as well as a three-vertex extension involving  $V_n, \tilde{V}_n^{(1)},$  and  $\tilde{V}_n^{(2)}$ . We omit further details.

#### 4.6. The Bhamidi–van der Hofstad–Hooghiemstra connection process

In this section we give the idea of the proof of Proposition 4.2. We have seen in Section 4.1 that the early stages of the exploration processes can be coupled to two branching processes. Now we describe how these branching processes connect up and explain the results on the connection process in [8]. We start by setting the stage. Fix the deterministic sequence  $s_n \rightarrow \infty$

as in (4.6). Then, define

$$t_n = \frac{1}{2\lambda_n} \log n, \quad \bar{t}_n = \frac{1}{2\lambda_n} \log n - \frac{1}{2\lambda_n} \log(W_{s_n} \tilde{W}_{s_n}). \tag{4.20}$$

Note that  $e^{\lambda_n t_n} = \sqrt{n}$ , so that at time  $t_n$ , both  $|\mathcal{F}_n(t_n)|$  and  $|\tilde{\mathcal{F}}_n(t_n)|$  have size of order  $\sqrt{n}$ . Consequently, the variable  $t_n$  denotes the typical time when collision edges start appearing. The time  $\bar{t}_n$  allows for stochastic fluctuations in the size of these infected (and backward-infected) clusters.

By Proposition 4.1,  $s_n \rightarrow \infty$  in such a way that each of  $\{\mathcal{F}_n(t)\}_{t \leq s_n}$  and  $\{\tilde{\mathcal{F}}_n(t)\}_{t \leq s_n}$  can be coupled to independent CTBPs. For the present part, it is crucial that the forward CTBP from  $V_n$  and the backward CTBP from  $\tilde{V}_n$  should run *simultaneously*, i.e. we run the two exploration processes described in Section 4.3 at the same time (meaning the same actual continuous time).

A *collision edge* is formed when a half-edge, on pairing, connects to a half-edge in the other CTBP, i.e. either a half-edge in the coming generation of the forward cluster of  $V_n$  pairs to a half-edge in the coming generation of the backward cluster of  $\tilde{V}_n$  or *vice versa*. The main result in this section describes the limiting stochastic process of the appearance of collision edges and their properties. For this, we need some more notation.

Denote the  $i$ th collision edge by  $(x_i, p_{x_i})$ , where  $p_{x_i}$  is an active half-edge (in either the forward or backward cluster) and  $x_i$  is the half-edge which pairs to  $p_{x_i}$ . Furthermore, let  $T_i^{(\text{col})}$  denote the time at which the  $i$ th collision edge is formed, which is the same as the birth time of the vertex incident to  $x_i$ . We let  $R_{T_i^{(\text{col})}}(p_{x_i})$  be the remaining lifetime of the half-edge  $p_{x_i}$ , which, by construction, is equal to the time after time  $2T_i^{(\text{col})}$  that the edge will be found completely by the flow. Thus, the path that the edge  $(x_i, p_{x_i})$  completes has length equal to  $2T_i^{(\text{col})} + R_{T_i^{(\text{col})}}(p_{x_i})$  and it has  $N_e(x_i) + \tilde{N}_e(p_{x_i}) + 1$  edges, where  $N_e(x_i)$  and  $\tilde{N}_e(p_{x_i})$  denote the number of edges between the respective roots and the vertices  $V(x_i)$  and  $V(p_{x_i})$  incident to  $x_i$  and  $p_{x_i}$ , respectively (our notation assumes that the chain with  $p_{x_i}$  at one end has  $\tilde{V}_n$  at the other; if not then simply swap  $x_i$  and  $p_{x_i}$ ). We conclude that the shortest weight path has weight equal to  $L_n(V_n, \tilde{V}_n) = \min_{i \geq 1} \{2T_i^{(\text{col})} + R_{T_i^{(\text{col})}}(p_{x_i})\}$ . Let  $J$  be the minimizer of this minimization problem. Then the number of edges equals  $N_e := N_e(x_J) + \tilde{N}_e(p_{x_J}) + 1$ . Finally, for a collision edge  $(x_i, p_{x_i})$ , let

$$I(x_i) = \begin{cases} 1 & \text{when } x_i \text{ is incident to a vertex in } \mathcal{F}_n(T_i^{(\text{col})}), \\ 2 & \text{when } x_i \text{ is incident to a vertex in } \tilde{\mathcal{A}}_n(T_i^{(\text{col})}). \end{cases}$$

In order to describe the properties of the shortest weight path, define

$$\bar{T}_i^{\text{col}} = T_i^{(\text{col})} - \bar{t}_n, \quad \bar{N}_i = \frac{m_n^* N_e(x_i) - t_n}{\sqrt{(\sigma_n^*)^2 t_n / m_n^*}}, \quad \tilde{\bar{N}}_i = \frac{m_n^* \tilde{N}_e(p_{x_i}) - t_n}{\sqrt{(\sigma_n^*)^2 t_n / m_n^*}}, \tag{4.21}$$

where  $m_n^*$  and  $\sigma_n^*$  are the mean and standard deviation of the stable-age distribution in (4.11). Introduce the space  $\mathcal{A} := \mathbb{R} \times \{1, 2\} \times \mathbb{R} \times \mathbb{R} \times [0, \infty)$ , and define the  $\mathcal{A}$ -valued random variables  $\{\Xi_i\}_{i \geq 1}$  by

$$\Xi_i = (\bar{T}_i^{\text{col}}, I(x_i), \bar{N}_i, \tilde{\bar{N}}_i, R_{T_i^{(\text{col})}}(p_{x_i})).$$

Then, for sets  $A$  in the Borel  $\sigma$ -algebra of the space  $\mathcal{A}$ , define the point process

$$\Pi_n(A) = \sum_{i \geq 1} \delta_{\Xi_i}(A), \tag{4.22}$$

where  $\delta$  is the Dirac measure as in (4.1). Let  $\mathcal{M}(\mathcal{S})$  denote the space of all simple locally finite point processes on  $\mathcal{S}$  equipped with the vague topology (see, e.g. [20]). Use the natural definition of weak convergence of a sequence of random point processes  $\Pi_n \in \mathcal{M}(\mathcal{S})$  on this space. This is the notion of convergence referred to in the following theorem in which  $\Phi$  denotes the distribution function of a standard normal random variable. Finally, define the density  $f_R$  of the limiting residual time to birth distribution  $F_R$  in (4.16) by

$$f_R(x) = \frac{\int_0^\infty e^{-\lambda y} g(x + y) dy}{\int_0^\infty e^{-\lambda y} [1 - G(y)] dy}.$$

Then the main result about the appearance of collision edges in [8, Theorem 3.1] is the following theorem.

**Theorem 4.1.** (Poisson process limit of collision edges [8].) *Consider the distribution of the point process  $\tilde{\Pi}_n \in \mathcal{M}(\mathcal{S})$  defined in (4.22) conditional on  $\{(\mathcal{F}_n(t), \tilde{\mathcal{F}}_n(t))\}_{t \in [0, s_n]}$  such that  $W_{s_n} > 0$  and  $\tilde{W}_{s_n} > 0$ . Then  $\tilde{\Pi}_n$  converges in distribution as  $n \rightarrow \infty$  to a Poisson process  $\Pi$  on  $\mathcal{S}$  with intensity measure*

$$\lambda(dt \times i \times dx \times dy \times dr) = \frac{2\mathbb{E}[D^* - 1]f_R(0)}{\mathbb{E}[D]} e^{2\lambda t} dt \otimes \left\{ \frac{1}{2}, \frac{1}{2} \right\} \otimes \Phi(dx) \otimes \Phi(dy) \otimes F_R(dr). \tag{4.23}$$

Given a realization of the Poisson process  $\Pi$  (PPP) as in Theorem 4.1, denote the marginal process of the first component with points in  $(-\infty, \infty)$  by  $\{P_i\}_{i \geq 1}$ . It was shown in [8] that Theorem 4.1 implies that  $L_n(V_n, \tilde{V}_n) - 2\tilde{t}_n \xrightarrow{D} \min_{i \geq 1} \{2P_i + R_i\}$ . Furthermore, it follows that

$$\lambda \min_{i \geq 1} \{2P_i + R_i\} \stackrel{D}{=} -\Lambda - \log\left(\frac{\mathbb{E}[D^* - 1]f_R(0)B}{\mathbb{E}[D]}\right), \tag{4.24}$$

where  $B = \int_0^\infty F_R(z)e^{-\lambda z} dz = m^*/\mathbb{E}[D^* - 2]$  and  $m^*$  is the mean of the so-called stable-age distribution in (4.11). It was shown in [8, Lemma 2.3] that  $f_R(0) = \lambda/\mathbb{E}[D^* - 2]$ , so

$$\lambda c = -\log\left(\frac{\mathbb{E}[D^* - 1]f_R(0)B}{\mathbb{E}[D]}\right) = \log\left(\frac{\mathbb{E}[D] (\mathbb{E}[D^* - 2])^2}{\lambda m^* \mathbb{E}[D^* - 1]}\right).$$

We thus see here that the Gumbel distribution arises from the minimization of the points of the PPP  $\{2P_i + R_i\}_{i \geq 1}$ . Interestingly, the Gumbel distribution also arises in  $\min_{i \geq 1} \{P_i : i \geq 1\}$ , but with a different constant  $c$ . Thus, the addition of the residual lifetime only changes the constant. Since  $2(\tilde{t}_n - t_n) \xrightarrow{D} \lambda^{-1} \log(W\tilde{W})$ , this proves that, as  $n \rightarrow \infty$ ,

$$L_n(V_n, \tilde{V}_n) - 2t_n = \min_{i \geq 1} \{2T_i^{(col)} + R_{T_i^{(col)}}(p_{x_i})\} - 2t_n \xrightarrow{D} c + \frac{\log(W\tilde{W}) - \Lambda}{\lambda}. \tag{4.25}$$

Also, by (4.21), the trail of the epidemic, which is equal to  $N_e = N_e(x_J) + \tilde{N}_e(p_{x_J}) + 1$ , when normalized to  $(m_n^* N_e - 2t_n) / \sqrt{(\sigma_n^*)^2 t_n / m_n^*}$ , converges in distribution to the sum of two i.i.d. standard normal random variables, where  $m_n^*$  and  $\sigma_n^*$  are the mean and standard deviation of the stable-age distribution in (4.11). This explains (2.3), and identifies  $\alpha_n = 1/(\lambda_n m_n^*)$  and  $\beta = (\sigma^*)^2 / [\lambda(m^*)^3]$ , where  $(m^*, \sigma^*) = \lim_{n \rightarrow \infty} (m_n^*, \sigma_n^*)$ .

To prove Theorem 4.1, in [8] the expected number of collision edges that are created are investigated. The branching process theory in Section 4.4 suggests that when a collision edge occurs, the processes generating the two vertices of the collision edge satisfy central limit theorems. Furthermore, the residual time to birth of the active half-edge to which we have

paired the newly found half-edges converges in distribution to the residual lifetime distribution. Thus, we only need to argue that the stochastic process that describes the times of finding the collision edges and centered by  $\bar{t}_n$  as in (4.21) converges to a PPP whose intensity measure has density function on  $(-\infty, 0]$  equal to  $t \mapsto (2\mathbb{E}[D^* - 1]f_R(0)/\mathbb{E}[D])e^{2\lambda t}$ . For this, we note that the rate at which new half-edges are found at time  $t + \bar{t}_n$  is roughly equal to  $2f_R(0)|\mathcal{A}_n(t + \bar{t}_n)|/|\tilde{\mathcal{A}}_n(t + \bar{t}_n)|/|\mathcal{L}_n|$ , where the factor  $f_R(0)$  is due to the fact that half-edges with remaining lifetime equal to 0 are those that die, and the factor 2 is due to the fact that both  $\mathcal{F}_n$  and  $\tilde{\mathcal{F}}_n$  can give rise to the birth of the half-edge.

Here we also note that  $|\mathcal{A}_n(t + \bar{t}_n)|$  and  $|\tilde{\mathcal{A}}_n(t + \bar{t}_n)|$  are of order  $\sqrt{n}$ , and, thus, the total number of half-edges is equal to  $|\mathcal{L}_n|(1 + o_{\mathbb{P}}(1))$ . When a half-edge dies, it has a random number of children with distribution close to  $D_n^* - 1$ , and each of the corresponding half-edges can create a collision edge; hence, we add an extra  $\mathbb{E}[D_n^* - 1]$  factor. Furthermore, we can approximate  $|\mathcal{L}_n| \approx n\mathbb{E}[D_n]$ ,  $|\mathcal{A}_n(t)| \approx e^{\lambda nt} W_{s_n}$ , and  $|\tilde{\mathcal{A}}_n(t)| \approx e^{\lambda nt} \tilde{W}_{s_n}$ , so that, using (4.20),

$$\begin{aligned} \frac{\mathbb{E}[D_n^* - 1] f_R(0) |\mathcal{A}_n(t + \bar{t}_n)| |\tilde{\mathcal{A}}_n(t + \bar{t}_n)|}{|\mathcal{L}_n|} &\approx \frac{\mathbb{E}[D_n^* - 1] f_R(0)}{\mathbb{E}[D]n} e^{2\lambda_n(t + \bar{t}_n)} W_{s_n} \tilde{W}_{s_n} \\ &= \frac{\mathbb{E}[D_n^* - 1] f_R(0)}{\mathbb{E}[D_n]} e^{2\lambda_n t}. \end{aligned}$$

This explains the intuition behind Theorem 4.1.

**4.7. The Barbour–Reinert connection process: differences**

The main difference between the Barbour–Reinert proof of Proposition 4.2 given in [3] and the previous section is that in the proof in [3], the forward and backward clusters are run sequentially, one after the other, not simultaneously.

This is achieved by coupling the infection process together with the exploration on  $CM_n(\mathbf{d})$  to the forward branching process with small errors up to time  $t_n^* := \tau_{\sqrt{n}}$  in the forward process ( $\tau_{\sqrt{n}}$  denotes the time when the  $\sqrt{n}$ th vertex enters the infection), which we freeze after this time. We then couple the backward process conditionally on the frozen cluster of the forward process up to time  $(\log n)/2\lambda_n + K$  for some large  $K > 0$ . Then, by (4.15), for any  $u \in \mathbb{R}$ , at time  $t_n(u) := (\log n)/2\lambda_n + u$  the size of the coming generation in the forward process is  $|\mathcal{A}_{\tau_{\sqrt{n}}}| = c_n^A \sqrt{n}(1 + o(1))$  for a specific constant  $c_n^A$  and the size of the backward cluster is  $|\tilde{\mathcal{A}}_{t_n(u)}| = c_n^A \sqrt{n} e^{\lambda u} \tilde{W}_{t_n(u)}(1 + o(1))$ . From here the formation of collision edges leads to a similar two-dimensional Poisson process to that described by the first and last coordinates in (4.23), i.e. here the intensity measure, conditioned on  $\tilde{W}_{s_n}$ , is given by

$$\frac{\mathbb{E}[D^*]f_R(0)}{\mathbb{E}[D]} e^{\lambda x} \tilde{W}_{s_n} dx \otimes F_R(dy).$$

From here onwards, the two proofs are essentially the same: the factor  $W$  from the forward process appears in the formula  $\tau_{\sqrt{n}} \approx (\log n)/2\lambda_n - (\log W_{s_n})/\lambda$ . The minimization problem (4.24) is then solved by calculating the probability that there are no PPP points in the infinite triangle  $x + y \leq t$ , yielding the statement of Proposition 4.2.

**Acknowledgements**

The work of RvdH and JK was supported in part by The Netherlands Organisation for Scientific Research (NWO). SB was partially supported by the NSF-DMS grants 1105581 and 1310002. We thank the anonymous referee for numerous suggestions that significantly improved the organization of the paper.

## References

- [1] ALDOUS, D. (2013). Interacting particle systems as stochastic social dynamics. *Bernoulli* **19**, 1122–1149.
- [2] BALL, F. AND DONNELLY, P. (1995). Strong approximations for epidemic models. *Stoch. Process. Appl.* **55**, 1–21.
- [3] BARBOUR, A. D. AND REINERT, G. (2013). Approximating the epidemic curve. *Electron. J. Prob.* **18**, 30pp.
- [4] BARRAT, A., BARTHÉLEMY, M. AND VESPIGNANI, A. (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press.
- [5] BARTLETT, M. S. (1955). *An Introduction to Stochastic Processes, with Special Reference to Methods and Applications*. Cambridge University Press.
- [6] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2010). First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Prob.* **20**, 1907–1965.
- [7] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2011). First passage percolation on the Erdős–Rényi random graph. *Combinatorics Prob. Comput.* **20**, 683–707.
- [8] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2012). Universality for first passage percolation on sparse random graphs. Preprint. Available at <http://arxiv.org/abs/1210.6839v1>.
- [9] BOHMAN, T. AND PICOLLELLI, M. (2012). SIR epidemics on random graphs with a fixed degree sequence. *Random Structures Algorithms* **41**, 179–214.
- [10] BOLLOBÁS, B. (2001). *Random Graphs* (Camb. Stud. Adv. Math. **73**), 2nd edn. Cambridge University Press.
- [11] DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*, Vol. I, 2nd edn. Springer, New York.
- [12] DECREUSEFOND, L., DHERSIN, J.-S., MOYAL, P. AND TRAN, V. C. (2012). Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Prob.* **22**, 541–575.
- [13] DRAIEF, M. AND MASSOULIÉ, L. (2010). *Epidemics and Rumours in Complex Networks* (London Math. Soc. Lecture Note Series **369**). Cambridge University Press.
- [14] DURRETT, R. (2007). *Random Graph Dynamics*. Cambridge University Press.
- [15] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, London.
- [16] JAGERS, P. AND NERMAN, O. (1984). The growth and composition of branching populations. *Adv. Appl. Prob.* **16**, 221–259.
- [17] JANSON, S. (2009). The probability that a random multigraph is simple. *Combinatorics Prob. Comput.* **18**, 205–225.
- [18] JANSON, S. AND LUCZAK, M. J. (2009). A new approach to the giant component problem. *Random Structures Algorithms* **34**, 197–216.
- [19] JANSON, S., LUCZAK, M. AND WINDRIDGE, P. (2014). Law of large numbers for the SIR epidemic on a random graph with given degrees. Preprint. Available at <http://arxiv.org/abs/1308.5493v3>.
- [20] KALLENBERG, O. (1976). *Random Measures*. Akademie, Berlin.
- [21] KENDALL, D. G. (1956). Deterministic and stochastic epidemics in closed populations. In *Proc. 3rd Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*, Vol. IV, University of California Press, Berkeley, pp. 149–165.
- [22] KERMACK, W. O. AND MCKENDRICK, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London. A* **115**, 700–721.
- [23] NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitsth.* **57**, 365–395.
- [24] NEWMAN, M., BARABÁSI, A.-L. AND WATTS, D. J. (eds) (2006). *The Structure and Dynamics of Networks*. Princeton University Press, Princeton, NJ.
- [25] VAN DER HOFSTAD, R. (2014). *Random Graphs and Complex Networks*. In preparation.
- [26] VOLZ, E. (2008). SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.* **56**, 293–310.

SHANKAR BHAMIDI, *University of North Carolina*

Department of Statistics, University of North Carolina, Chapel Hill, USA. Email address: bhamidi@email.unc.edu

REMCO VAN DER HOFSTAD, *Eindhoven University of Technology*

Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email address: rhofstad@win.tue.nl

JÚLIA KOMJÁTHY, *Eindhoven University of Technology*

Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email address: j.komjathy@tue.nl