Combinatorics, Probability and Computing (2011) 20, 683–707. © Cambridge University Press 2011
doi:10.1017/S096354831100023X

First Passage Percolation on the Erdős–Rényi Random Graph

SHANKAR BHAMIDI[1], REMCO VAN DER HOFSTAD[2]
and GERARD HOOGHIEMSTRA[3]

[1]Department of Statistics and Operations Research, 304 Hanes Hall,
University of North Carolina, Chapel Hill, NC 27599, USA
(e-mail: bhamidi@email.unc.edu)

[2]Department of Mathematics and Computer Science, Eindhoven University of Technology,
PO Box 513, 5600 MB Eindhoven, The Netherlands
(e-mail: rhofstad@win.tue.nl)

[3]DIAM, Delft University of Technology, Mekelweg 4, 2628CD Delft, The Netherlands
(e-mail: g.hooghiemstra@tudelft.nl)

Received 21 May 2010; revised 13 May 2011; first published online 20 June 2011

In this paper we explore first passage percolation (FPP) on the Erdős–Rényi random graph $G_n(p_n)$, where we assign independent random weights, having an exponential distribution with rate 1, to the edges. In the sparse regime, *i.e.*, when $np_n \to \lambda > 1$, we find refined asymptotics both for the minimal weight of the path between uniformly chosen vertices in the giant component, as well as for the hopcount (*i.e.*, the number of edges) on this minimal weight path. More precisely, we prove a central limit theorem for the hopcount, with asymptotic mean and variance both equal to $(\lambda \log n)/(\lambda - 1)$. Furthermore, we prove that the minimal weight centred by $(\log n)/(\lambda - 1)$ converges in distribution.

We also investigate the dense regime, where $np_n \to \infty$. We find that although the base graph is *ultra-small* (meaning that graph distances between uniformly chosen vertices are $o(\log n)$), attaching random edge weights changes the geometry of the network completely. Indeed, the hopcount H_n satisfies the universality property that whatever the value of p_n, $H_n/\log n \to 1$ in probability and, more precisely, $(H_n - \beta_n \log n)/\sqrt{\log n}$, where $\beta_n = \lambda_n/(\lambda_n - 1)$, has a limiting standard normal distribution. The constant β_n can be replaced by 1 precisely when $\lambda_n \gg \sqrt{\log n}$, a case that has appeared in the literature (under stronger conditions on λ_n) in [4, 13]. We also find lower bounds for the maximum, over all pairs of vertices, of the optimal weight and hopcount.

This paper continues the investigation of FPP initiated in [4] and [5]. Compared to the setting on the configuration model studied in [5], the proofs presented here are much simpler due to a direct relation between FPP on the Erdős–Rényi random graph and thinned continuous-time branching processes.

1. Introduction

First passage percolation is one of the most fundamental problems in probability theory. The basic motivation to study this problem is the following. The goal is to model the flow of fluid

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

through some random medium. Suppose that we have a base graph on n vertices which represents the available pathways for the fluid. We attach to each edge in the graph some random edge weight, typically assumed to be independent and identically distributed (i.i.d.) positive random variables with some probability density function f. We then think of fluid percolating through the network at rate 1 from some source, so that the travel time of the fluid along the edges is equal to the corresponding weights. Letting $n \to \infty$, one is then interested in asymptotics of various statistics of the flow through this medium. See, *e.g.*, [11, 14] for results and a survey on the integer lattice. We refer to Janson [15] for a study of related problems in the setting of the complete graph.

Our aim in this paper is to rigorously analyse first passage percolation (FPP) on the Erdős–Rényi random graph (ERRG) denoted by $G_n(p_n)$. We shall see that for two randomly chosen vertices in the giant component the hopcount, *i.e.*, the number of edges on the *shortest*-weight path between these vertices, scales as $\log n$ and we shall find a central limit theorem (CLT) for this quantity. We shall also find that the weight of the shortest-weight path re-centred by a constant multiple of $\log n$ converges in distribution to some limiting random variable. We shall describe the explicit distribution of the limit. We shall also find lower bounds for the maximal optimal weight and hopcount between vertices in the giant component.

In [5], we have investigated FPP on the configuration model (CM) with degrees given by an i.i.d. sequence with distribution function F, satisfying $F(x) = 0$, $x < 2$. Consequently, all degrees are at least 2 and the giant component contains $n - o(n)$ vertices, so that with high probability (w.h.p.), two uniformly chosen vertices are connected. Furthermore, it was assumed that for all $x \geqslant 0$, there exist constants c_1, c_2, such that

$$c_1 x^{1-\tau} \leqslant 1 - F(x) \leqslant c_2 x^{1-\tau}, \quad \text{when} \quad \tau \in (2, 3),$$

whereas, for $\tau > 3$, it was assumed that F satisfies, for all $x \geqslant 0$,

$$1 - F(x) \leqslant c x^{1-\tau},$$

for some constant c. Apart from self-loops and multiple edges, the CM, with a binomial degree sequence, is not very different from the ERRG. The main challenge in studying the ERRG compared to the study of FPP on the CM in [5] is threefold.

(a) The degree sequence in [5] was assumed to satisfy $F(x) = 0$, $x < 2$, so that all degrees are at least 2, with probability 1. As said before, this implies that uniformly chosen vertices are w.h.p. connected. In the present paper a uniformly chosen pair of vertices does have a positive probability of being connected, but this probability does not equal 1, so that we have to condition on the uniformly chosen vertices to be in the giant component. Technically, this is a big step forward.

(b) We deal with the case where the average degree $\lambda_n = np_n \to \infty$ (see Corollary 2.4 below), a scenario that is not contained in the sparse setting in [5]. This is the first time a result of this generality has been proved in the regime $np_n \to \infty$.

(c) The ERRG admits an elegant embedding in a marked branching process in continuous time. Consequently, the proofs are short and non-technical compared to those in [5]. This technique should prove to be useful in a number of other random graph models. Moreover, we include a lower bound on the weight and length of the largest shortest-weight path for FPP on the ERRG.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

2. Results

In this section we formulate our main results. Throughout this paper, we work on the Erdős–Rényi random graph (ERRG) $G_n(p_n)$, with vertex set $[n] = \{1, \ldots, n\}$, and where there is an edge between every pair of vertices $i \neq j$ independently with probability p_n; the resulting edge set is denoted by $E_n = \{\{i, j\} : i, j \in [n], i \neq j\}$. Furthermore, each edge $e \in E_n$ is equipped with an independent edge weight E_e, having an exponential distribution with rate 1. We denote by $\xrightarrow{a.s.}$, $\xrightarrow{d}$, and $\xrightarrow{\mathbb{P}}$, convergence almost surely, in distribution, and in probability, respectively. The symbols o, O are the ordinary *Landau* symbols. We say that a sequence of random variables X_n satisfies $X_n = o_{\mathbb{P}}(b_n)$, $X_n = O_{\mathbb{P}}(b_n)$, respectively, if $X_n/b_n \xrightarrow{\mathbb{P}} 0$, X_n/b_n is tight, respectively. We write $X \stackrel{d}{=} Y$ to denote that the two random vectors X and Y are equal in distribution. We write that a sequence of events $(\mathcal{E}_n)_{n\geqslant 1}$ occurs *with high probability* (w.h.p.) when $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 1$. Binomial random variables are denoted by $\mathrm{Bin}(n, p)$, where n denotes the number of trials and p the success probability. We further denote an exponentially distributed random variable with rate μ and a Poisson random variable with mean λ by $\mathrm{Exp}(\mu)$ and $\mathrm{Poi}(\lambda)$, respectively.

In Theorems 2.1 and 2.2 below, we shall investigate the hopcount and weight of FPP on the ERRG.

Theorem 2.1 (CLT for hopcount). *Let* $\lim_{n\to\infty} np_n = \lambda > 1$ *and define* $\beta = \frac{\lambda}{\lambda-1} > 1$. *Then, the hopcount* H_n *between two uniformly chosen vertices, conditioned on being connected, satisfies a central limit theorem with the asymptotic mean and variance both equal to* $\beta \log n$, *i.e.,*

$$\frac{H_n - \beta \log n}{\sqrt{\beta \log n}} \xrightarrow{d} Z,$$

where Z *is a standard normal random variable.*

Theorem 2.2 (Limit distribution for minimal weight). *Let* $\lim_{n\to\infty} np_n = \lambda > 1$, *and define* $\gamma = \frac{1}{\lambda-1}$. *Then, there exists a non-degenerate real-valued random variable* X *such that the minimal weight* W_n *between two uniformly chosen vertices, conditioned on being connected, satisfies*

$$W_n - \gamma \log n \xrightarrow{d} X.$$

Remark 2.3 (Joint convergence). The proof reveals that the convergence in Theorems 2.1 and 2.2 holds *jointly*, with the limits being independent.

Let us identify the limiting random variable X in Theorem 2.2. Consider a continuous-time Galton–Watson branching-process with $\mathrm{Poi}(\lambda)$ offspring, where the individuals have an exponential lifetime with rate 1. Denote the number of alive individuals at time t by $N(t)$. It is well known [3] that there exists $\alpha > 0$ such that $\mathrm{e}^{-\alpha t} N(t)$ has an almost sure limit W. For the case under consideration it is readily verified that the Malthusian parameter α satisfies $\alpha = \lambda - 1 > 0$. Let D denote a $\mathrm{Poi}(\lambda)$ random variable representing the number of offspring from one individual, and, conditioned on D, denote by $E_1, E_2, \ldots, E_D$, the exponential lifetimes of the offspring. Then,

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

the limit W satisfies the stochastic equation

$$W \stackrel{d}{=} \sum_{i=1}^{D} \mathrm{e}^{-\alpha E_i} W_i, \tag{2.1}$$

where on the right-hand side all the involved random variables are mutually independent with W_i being copies of W. Hence, taking conditional expectations with respect to D, we see that $\phi(t) = \mathbb{E}[\mathrm{e}^{-tW}]$ satisfies the functional relation

$$\phi(t) = \mathbb{E}\big[\big(\mathbb{E}[\mathrm{e}^{-t\mathrm{e}^{-\alpha E_1} W}]\big)^D\big] = \exp\bigg(-\lambda \int_0^\infty \big[1 - \phi(t\mathrm{e}^{-(\lambda-1)x})\big]\mathrm{e}^{-x} dx\bigg). \tag{2.2}$$

From this, it is quite easy to prove that the random variable W has an atom at zero of size p_λ, where p_λ is the smallest non-negative solution of the equation

$$p_\lambda = \exp(-\lambda(1 - p_\lambda)),$$

i.e., p_λ is the extinction probability of a branching process with Poi(λ) offspring. Furthermore (see, *e.g.*, [3, Chapter 7, Theorem 2(iii)]), the random variable W, conditioned to be positive, admits a continuous density on $\mathbb{R}^+$, which we denote by

$$W_\lambda \stackrel{d}{=} (W | W > 0). \tag{2.3}$$

To construct the limiting random variable X in Theorem 2.2, we shall need the following random variables.

(a) Let $W_\lambda^{(1)}, W_\lambda^{(2)}$ be independent and identically distributed as W_λ.

(b) Let $E \stackrel{d}{=} \mathrm{Exp}(1)$ be independent of $W_\lambda^{(i)}$, $i = 1, 2$.

In terms of these random variables, the random variable X satisfies

$$X \stackrel{d}{=} -\gamma \log(\gamma W_\lambda^{(1)}) - \gamma \log(\gamma W_\lambda^{(2)}) + \gamma \log(E).$$

We next study the *dense graph setting*, where $np_n = \lambda_n \to \infty$. The proof in this setting follows the same lines as that of Theorems 2.1–2.2, and therefore we will only give a sketch of the proof. Observe that for $np_n \to \infty$, any pair of vertices is w.h.p. connected. It is not hard to see that, in this case, the giant component consists of $n(1 - o(1))$ vertices, and that the graph distance between two uniformly chosen vertices is $\log n / \log \lambda_n = o(\log n)$, so that the random graph is *ultra-small*. In the statement of the result, we denote by H_n and W_n, respectively, the hopcount and minimal weight of the shortest-weight path between two uniformly chosen vertices.

Corollary 2.4 (Limit of hopcount and weight for $np_n \to \infty$). *Set $\lambda_n = np_n$. For $\lambda_n \to \infty$, as $n \to \infty$,*

(a)

$$\frac{H_n - \beta_n \log n}{\sqrt{\log n}} \xrightarrow{d} Z, \tag{2.4}$$

where $\beta_n = \lambda_n/(\lambda_n - 1)$ and Z is a standard normal random variable;

(b)

$$(\lambda_n - 1) W_n - \log n \xrightarrow{d} \tilde{X}, \tag{2.5}$$

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

where the random variable $\tilde{X} \stackrel{d}{=} M_1 + M_2 - M_3$, *with* M_1, M_2, M_3 *independent Gumbel random variables, i.e.,* $\mathbb{P}(M_i \leqslant x) = \Lambda(x) = \exp(-\mathrm{e}^{-x})$, $1 \leqslant i \leqslant 3$, *for all* $x \in \mathbb{R}$.

Remark 2.5 (Dense setting). **(a)** In part (a) of Corollary 2.4 the centering $\beta_n \log n$ can be replaced by $\log n$ if and only if $\lambda_n/\sqrt{\log n} \to \infty$. In this case, the hopcount has the same limiting distribution as on the complete graph (see, *e.g.*, [13]).

(b) The distribution of H_n for the case where $np_n/(\log n)^3 \to \infty$ was obtained in [13]. In [4, Theorem 5] limit results for $H_n/\log n$ and $np_nW_n - \log n$ were obtained in the case where $\liminf_{n\to\infty}(np_n)/(\log n) = a$, with $1 < a \leqslant \infty$. Thus, the present paper essentially completes the study of FPP on ERRGs in all regimes.

Theorems 2.1–2.2 state that for uniformly chosen connected vertices, the weight and length of the optimal path between them scales as $\gamma \log n$ and $\beta \log n$, respectively. The following shows the existence of pairs of vertices with a much larger optimal path weight and hopcount.

Theorem 2.6 (Existence of long paths). *Let* $np_n \to \lambda > 1$ *and define by* $c(\lambda), d(\lambda)$ *the constants*

$$c(\lambda) = \frac{1}{\lambda - 1} + \frac{2}{|\log \mu_\lambda|}, \quad d(\lambda) = \frac{\lambda}{\lambda - 1} + \frac{2}{|\log \mu_\lambda|}, \tag{2.6}$$

where μ_λ *is the* dual *of* λ, *i.e., the unique* $\mu \in (0, 1)$ *such that*

$$\mu \mathrm{e}^{-\mu} = \lambda \mathrm{e}^{-\lambda}. \tag{2.7}$$

Then, for any given $\varepsilon > 0$, *with high probability, there exists a pair of vertices, say* i^*, j^*, *in the giant component such that the weight of the optimal path* $W_n(i^*, j^*)$ *and its number of edges* $H_n(i^*, j^*)$ *satisfy the inequalities*

$$W_n(i^*, j^*) \geqslant (1 - \varepsilon)c(\lambda) \log n, \quad H_n(i^*, j^*) \geqslant (1 - \varepsilon)d(\lambda) \log n. \tag{2.8}$$

In [2], the random variable $\max_{i,j\in[n]} W_n(i, j)$ was investigated on the CM with exponential edge weights. There, it was shown that $\max_{i,j\in[n]} W_n(i, j)/\log n \xrightarrow{\mathbb{P}} 1/(\lambda - 1) + 2/(1 - \mu_\lambda)$ in the context of Poi(λ) degrees. It is not hard to see that the same proof also applies to the ERRG (see also [1]). Thus, our lower bound on $\max_{i,j\in[n]} W_n(i, j)$ is not optimal. We conjecture that the above result on $\max_{i,j\in[n]} H_n(i, j)$ *is* optimal, *i.e.*, you can get between any pair of vertices within at most $(d(\lambda) + \varepsilon) \log n$ hops.

In [5], we have proved results parallel to those in Theorem 2.1 and 2.2 for the CM with degrees given by an i.i.d. sequence with distribution function F, where $[1 - F](x) \leqslant cx^{1-\tau}$, $\tau > 3$, for some constant c and all $x \geqslant 0$. We found that the asymptotic hopcount between two uniformly chosen vertices converges to a normal distribution with mean and variance equal to $\frac{\nu}{\nu-1} \log n$, where $\nu = \mathbb{E}[D(D - 1)]/\mathbb{E}[D]$ and where the random variable D is distributed as the degree distribution F. Note that in Theorem 2.1 the role of ν is taken over by the parameter λ. This is not surprising, since for a Poi(λ) variable D, we have $\nu = \mathbb{E}[D(D - 1)]/\mathbb{E}[D] = \lambda$.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

3. Setting the stage for the proofs

A rough idea of the proof is as follows. Fix two vertices, say 1 and 2, in the giant component. Think of fluid emanating from these two sources *simultaneously* at rate 1, so that at time t, $\mathcal{F}^{(i)}(t)$ is the flow cluster from vertex i and includes the minimal weight paths to all vertices wetted at or before time t from vertex i, $i = 1, 2$. When these two flows collide (namely, when the flow from one of the sources reaches the other flow) via the formation of an edge (v_1, v_2) between two vertices $v_1 \in \mathcal{F}^{(1)}(\cdot)$ and $v_2 \in \mathcal{F}^{(2)}(\cdot)$, then the shortest-weight path between the two vertices has been found. This collision time, which we denote by S_{12}, tells us that the weight between the two vertices, $W_n(1, 2)$, equals

$$W_n(1, 2) = 2S_{12}. \tag{3.1}$$

Furthermore, if $G_n(v_i), i = 1, 2$, denotes the number of edges between the source i and the vertex v_i along the tree $\mathcal{F}^{(i)}(S_{12})$, then the hopcount $H_n(1, 2)$ is given by

$$H_n(1, 2) = G_n(v_1) + G_n(v_2) + 1. \tag{3.2}$$

The above idea is indeed a very rough sketch of our proof. In the paper, we embed the flow on the ERRG in a continuous-time marked branching process (CTMBP), where the offspring distribution is binomial with parameters $n - 1$ and p (see Section 4.2). With high probability, the marks in the CTMBP correspond to the vertices in the ERRG. We denote by $\{\mathrm{SWT}_m^{(i)}\}_{m \geqslant 0}$ the marks of the individuals wetted by the flow after m splits and the arrival times of these splits, where the superscript i, $i = 1, 2$, denotes the root i of the flow. It is not mandatory to let the flows grow simultaneously, and for technical reasons, the proof is simpler if we first grow $\mathrm{SWT}_m^{(1)}$ to a size $a_n = \lceil \sqrt{n} \rceil$. After this, we grow $\mathrm{SWT}_m^{(2)}$, and we stop as soon as a mark of $\mathrm{SWT}_{a_n}^{(1)}$ appears in $\{\mathrm{SWT}_m^{(2)}\}_{m=0}^{\infty}$. The size $a_n = \lceil \sqrt{n} \rceil$ is the correct one, since if both flows are of size approximately $\sqrt{n}$, the probability that the second flow finds a mark of the first flow in some time interval of positive length is of order 1.

The CLT for the hopcount G_m between the mth-wetted mark and the root, as well as the limit distribution for the weight of the path between this mark and the root, are given in Section 4.3. This theory is based on the asymptotics for Bellman–Harris processes, which will be developed in Section 3.1.

In Section 4.4 we investigate the connection time C_n, *i.e.*, the random time until for the first time the second flow starting from mark 2 wets one of the individuals in $\mathrm{SWT}_{a_n}^{(1)}$. We prove that C_n/a_n converges in distribution to an Exp(1) random variable. This can informally be understood as follows. The number of distinct marks in $\mathrm{SWT}_{a_n}^{(1)}$ is $a_n(1 + o(1))$. Each of these marks is chosen with probability $1/n$ in $\mathrm{SWT}_m^{(2)}$, so that the first time that any of these marks is chosen is close to a geometric random variable with success parameter a_n/n. This random variable is close to $n/a_n\mathrm{Exp}(1)$, so that indeed $\mathrm{C}_n/a_n \xrightarrow{d} \mathrm{Exp}(1)$, since $a_n^2 = n(1 + o(1))$.

We furthermore show that, conditioned on C_n, the hopcount H_n is w.h.p. the independent sum of $G^{(1)}_{U_n, a_n}$ and $G^{(2)}_{\mathrm{C}_n}$. Here $G^{(1)}_{U_n, a_n}$ denotes the distance between root 1 and a mark in $\mathrm{SWT}_{a_n}^{(1)}$ that is chosen uniformly at random.

With the above sketch of proof in mind, the remainder of the paper is organized as follows.

(1) In Section 3.1, we analyse various properties of a *Bellman–Harris process* conditioned on non-extinction, including times to grow to a particular size and the generation of individuals

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

at this time. In this continuous-time branching process, the offspring will have a Poisson distribution.

(2) In Section 4.1, we introduce marked branching process trees with binomially distributed offspring and make the connection between these trees and the ERRG, by *thinning* the marked branching process tree.

(3) In Section 4.2, we replace the general weights on the edges by exponential ones so that we end up with a *continuous-time* marked branching processes (CTMBP) with binomial offspring. We further focus on some of the nice properties of the involved random variables, that are a consequence of the *memoryless property* of the exponential distribution.

(4) By *coupling* the CTMBP with binomial offspring to a Bellman–Harris process with Poisson offspring, we deduce in Section 4.3 the limit theorem for the generation G_m and the weight A_m from the results in Section 3.1.

(5) In Section 4.4, we present a refined analysis of the *connection time* of the two flow clusters.

(6) Finally, in Section 5, we *complete the proofs* of our main results.

The idea of the argument is quite simple, but making these ideas rigorous takes some technical work, particularly because we need to condition on our vertices being in the giant component.

3.1. Asymptotics for Bellman–Harris processes

Here we shall construct random trees where each vertex lives for an exponential amount of time, dies and gives birth to some number of children. Fix a sequence of non-negative integers $d_1, d_2, \ldots$. For future reference, let $s_i = \sum_{j=1}^{i} d_j - (i-1)$. We shall make the following blanket assumption:

$$s_i > 0, \qquad \text{for all } i \geqslant 1. \tag{3.3}$$

The condition in (3.3) is equivalent to the fact that the tree where the ith vertex has degree d_i is infinitely large. Consider the following *continuous-time* construction of a random tree.

Construction 3.1 (FPP on a tree).

(1) *Start with the root which dies immediately, giving rise to d_1 alive offspring.*

(2) *Each alive offspring lives for an* Exp(1) *amount of time, independent of all other randomness involved.*

(3) *When the ith vertex dies it leaves behind d_i alive offspring.*

In terms of the above construction, we can identify s_i as the number of alive vertices after the ith death. Let $T_1, T_2, \ldots$ be the time spacings between the ith and $(i+1)$st death $i \geqslant 1$. We shall often refer to $A_i = T_1 + T_2 + \cdots + T_i, i \geqslant 1$, as the time of the ith split. Note that at the graph topology level, due to the memoryless property of the exponential distribution, the above procedure is the same as the following construction, where exponential lifetimes do not appear.

Construction 3.2 (Discrete-time reformulation of FPP on a tree). *The shortest-weight graph on a tree with degrees $\{d_i\}_{i=1}^{\infty}$ is obtained as follows.*

(1) *At time* 0*, start with one alive vertex* (*the initial ancestor*).

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

(2) *At each time step i, $i \geqslant 1$, pick one of the alive vertices at random; this vertex dies giving birth to d_i children.*

Let G_m denote the graph distance between the root and a uniformly chosen vertex among all alive vertices at step m. We quote the following fundamental result from [7].[1]

Proposition 3.3 (Shortest-weight paths on a tree). *Pick a alive vertex at time $m \geqslant 1$ uniformly at random from all vertices alive at this time. Then the following hold.*

(a) *The generation of the mth chosen vertex is equal in distribution to*

$$G_m = \sum_{i=1}^{m} I_i, \tag{3.4}$$

where $\{I_i\}_{i=1}^{\infty}$ are independent Bernoulli random variables with

$$\mathbb{P}(I_i = 1) = d_i/s_i, \quad 1 \leqslant i \leqslant m. \tag{3.5}$$

(b) *The weight of the path between the root of the tree and the mth chosen vertex is equal in distribution to*

$$A_m = T_1 + T_2 + \cdots + T_m \stackrel{d}{=} \sum_{i=1}^{m} E_i/s_i, \tag{3.6}$$

where $\{T_i\}_{i=1}^{\infty}$ are i.i.d. exponential random variables with rate s_i, so that T_i is equal in distribution to E_i/s_i.

The above proposition states that the random variable G_m is the sum of m independent Bernoulli random variables. It is known that in a wide variety of settings, such quantities essentially follow a CLT. We shall be interested in showing that the standardization $\tilde{G}_m$ of G_m converges in distribution to a normal random variable Z, and the way we shall prove this is by showing that the Wasserstein distance between the distribution $\tilde{G}_m$ and that of Z goes to 0. Denote by Wass the metric on the space of probability measures defined as

$$\mathrm{Wass}(\mu_1, \mu_2) = \sup\left\{ \left| \int_{\mathbb{R}} g d\mu_1 - \int_{\mathbb{R}} g d\mu_2 \right| : g \text{ is 1-Lipschitz and bounded} \right\}.$$

It is well known [9, Theorem 11.3.3] that this metric on the space of probability measures induces the same topology as the topology of weak convergence. Thus, showing that $\tilde{G}_m \xrightarrow{d} Z$ is equivalent to showing that $\mathrm{Wass}(\mu_m, \mu) \to 0$ where $\tilde{G}_m \sim \mu_m$ and $Z \sim \mu$; here $X \sim \nu$ means that the law of the random variable X is equal to ν.

The following result is simple to prove using any of the methods for showing CLTs for sums of independent random variables, *e.g.*, Stein's method [8].

[1] A new *probabilistic* proof is given in [5], since there is some confusion in comparing the definition of s_i in this paper and the definition of s_i given in [7, below (3.1)].

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

Lemma 3.4 (CLT of hopcount in tree). *As before, let G_m denote the generation of a uniformly chosen vertex alive at time m. Define the sequence of random variables*

$$\tilde{G}_m = \frac{G_m - \sum_{i=1}^m \rho_i}{\sqrt{\sum_{i=1}^m \rho_i(1-\rho_i)}}, \tag{3.7}$$

where $\rho_i = d_i/s_i$. Let μ_m denote the distribution of $\tilde{G}_m$ and μ the distribution of a standard normal random variable. Then,

$$\mathrm{Wass}(\mu_m, \mu) \leqslant \frac{3}{\sqrt{\sum_{i=1}^m \rho_i(1-\rho_i)}}. \tag{3.8}$$

Proof. Combining [8, Theorems 3.1 and 3.2] yields

$$\mathrm{Wass}(\mu_m, \mu) \leqslant 3 \sum_{i=1}^m \mathbb{E}[|\eta_i|^3], \tag{3.9}$$

where

$$\eta_i = \frac{I_i - \rho_i}{\sqrt{\sum_{i=1}^m \rho_i(1-\rho_i)}},$$

and where I_i is a $\mathrm{Bin}(1, \rho_i)$ random variable. From this we obtain

$$\sum_{i=1}^m \mathbb{E}[|\eta_i|^3] = \frac{1}{[\sum_{i=1}^m \rho_i(1-\rho_i)]^{3/2}} \sum_{i=1}^m \rho_i(1-\rho_i)\{\rho_i^2 + (1-\rho_i)^2\} \leqslant \frac{1}{\sqrt{\sum_{i=1}^m \rho_i(1-\rho_i)}},$$

using $\rho_i^2 + (1-\rho_i)^2 \leqslant 1$. □

From now on, we take $D_1, D_2, \ldots$ i.i.d., where D_j has a $\mathrm{Poi}(\lambda)$ distribution with $\lambda > 1$. Furthermore, we consider the continuous-time construction of a random tree as in Construction 3.1, with $d_i = D_i$ and $S_i = \sum_{j=1}^i D_j - (i-1)$. Since the D_j are now *random*, for any fixed time $m \geqslant 1$, there is a non-zero probability that $S_j = 0$ for some $1 \leqslant j \leqslant m$. However, note that for any $m \geqslant 1$,

$$\mathbb{P}\big(S_i > 0 \quad \forall i = 1, 2, \ldots, m\big) \geqslant 1 - p_\lambda,$$

where p_λ is the extinction probability of a Galton–Watson branching process with $\mathrm{Poi}(\lambda)$ offspring. Let G_m be the generation of a randomly chosen alive individual among all alive vertices at time m, conditional on $S_i > 0$ for all $1 \leqslant i \leqslant m$. Let A_m be the time for the mth split to occur in the continuous-time construction (see Construction 3.1). Then we have the following asymptotics.

Theorem 3.5 (CLT for hopcount on trees conditioned to survive). *Conditioned on $S_i > 0$ for all $1 \leqslant i \leqslant m$, the following asymptotics for $m \to \infty$ hold.*

(a) *The generation G_m satisfies a CLT, i.e.,*

$$\frac{G_m - \beta \log m}{\sqrt{\beta \log m}} \xrightarrow{d} Z, \tag{3.10}$$

where Z is standard normal and $\beta = \frac{\lambda}{\lambda - 1}$.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

(b) *The mth split A_m satisfies the asymptotics*

$$(A_m|S_1 > 0, \dots, S_m > 0) - \gamma \log m \xrightarrow{d} -\gamma \log \gamma W_\lambda, \tag{3.11}$$

where W_λ has a distribution given by (2.3)*. The limits in* (a) *and* (b) *also hold jointly, where the limits are independent.*

The issue of conditioning is a technical annoyance as it removes the independence of increments of the random walk $\{S_i\}_{i=1}^m$. However, it should be intuitively clear that conditioning a random walk with positive drift to be positive is w.h.p. the same as conditioning the path to be positive in the first w_m steps, where we assume $w_m = o(\log \log m)$. Indeed, by the time w_m, the walk has reached height approximately $(\lambda - 1)w_m$ and if we just add independent Poisson random variables from this stage onwards, then the random walk will w.h.p. remain positive, since the walk has reached a high level by this time, and the probability that it would reach 0 after time m is exponentially small in w_m. This idea is made precise in the following construction. First we shall need some notation. Fix $w_m = o(\log \log m) \to \infty$ and let $Y_i = D_i - 1$ where the D_i are i.i.d. Poi(λ) random variables. Observe that

$$S_i = \sum_{j=1}^{i} D_j - (i-1) = 1 + \sum_{j=1}^{i} Y_j.$$

Now consider the following construction.

Construction 3.6. *Take the random vector $(S_1^*, S_2^*, \dots, S_{w_m}^*)$ equal in distribution to the random vector $(S_1, S_2, \dots, S_{w_m})$ conditioned on being positive, i.e.,*

$$(S_1^*, S_2^*, \dots, S_{w_m}^*) \stackrel{d}{=} (S_1, S_2, \dots, S_{w_m} | S_i > 0,\ 1 \leqslant i \leqslant w_m).$$

Let $Y_{w_m+1}, Y_{w_m+2}, \dots, Y_m$ be i.i.d. distributed as Poi(λ) $- 1$*, independent of $S_1^*, S_1^*, \dots, S_{w_m}^*$, and define*

$$S_i^* = S_{w_m}^* + \sum_{j=w_m+1}^{i} Y_j, \quad w_m + 1 \leqslant i \leqslant m. \tag{3.12}$$

Then the following proposition yields good bounds on the behaviour of the conditioned random walk $\{S_i^*\}_{1 \leqslant i \leqslant m}$.

Proposition 3.7 (Good bounds for the conditioned random walk). *The sequence in Construction 3.6 satisfies the following regularity properties.*

(a) $\mathbb{P}(S_1^* > 0, \dots, S_m^* > 0) = 1 - o(1)$ *as $m \to \infty$, and there exists a coupling such that $\mathbb{P}(S_j^* = \tilde{S}_j,\ 1 \leqslant j \leqslant m) = 1 - o(1)$, as $m \to \infty$, where $\tilde{S}_j, 1 \leqslant j \leqslant m$, is equal in distribution to the random walk $S_j = 1 + \sum_{i=1}^{j}(D_i - 1), 1 \leqslant j \leqslant m$, conditioned on $S_1 > 0, \dots, S_m > 0$.*

(b) *There exists a constant $C > 0$ such that w.h.p., for all $j > \log \log m$,*

$$(\lambda - 1)j - C\sqrt{j \log j} \leqslant S_j^* \leqslant (\lambda - 1)j + C\sqrt{j \log j}. \tag{3.13}$$

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

(c) *Let* $\rho_i^* = D_i^*/S_i^*$, *where* $D_i^* = S_i^* - S_{i-1}^* + 1, 1 \leqslant i \leqslant m$. *Then*

$$\sum_{i=1}^{m} \rho_i^* = \beta \log m + O_{\mathbb{P}}(\log\log m), \quad \sum_{i=1}^{m} (\rho_i^*)^2 = O_{\mathbb{P}}(\log\log m). \tag{3.14}$$

Assuming Proposition 3.7, let us show how to prove Theorem 3.5.

Proof of Theorem 3.5 subject to Proposition 3.7. Denote by G_m^* a random variable with distribution equal to $G_m | S_i > 0$, $1 \leqslant i \leqslant m$. In order to prove that

$$\frac{G_m^* - \sum_{i=1}^m \rho^*}{\sqrt{\sum \rho_i^*(1-\rho_i^*)}} \xrightarrow{d} Z,$$

we note that according to Lemma 3.4 and part (a) of Proposition 3.7, it is sufficient to show that w.h.p. $\sum_{i=1}^m \rho_i^*(1-\rho_i^*) \to \infty$, as $m \to \infty$. This is immediate from part (c) of Proposition 3.7. Proposition 3.7(c) also implies that it is allowed in the above statement to replace both $\sum_{i=1}^m \rho^*$ and $\sum \rho_i^*(1-\rho_i^*)$ by $\beta \log m$. Together this proves Theorem 3.5(a).

The distributional convergence of $A_m = T_1 + \cdots + T_m$ conditioned on $S_1 > 0, S_2 > 0, \ldots, S_m > 0$ in (3.11) is a little more difficult. We refer to [3, Theorem 2, p. 120], where it has been proved that the time of the mth split, A_m, when the branching process has Poi(λ) offspring, satisfies the asymptotics

$$N(A_m)\mathrm{e}^{-(\lambda-1)A_m} \xrightarrow{a.s.} W, \tag{3.15}$$

where $N(t)$ denotes the number of alive individuals at time t. Observe that (3.15) can easily be deduced from the convergence $\mathrm{e}^{-(\lambda-1)t}N(t) \xrightarrow{a.s.} W$, by substituting A_m for t, and using that $A_m \xrightarrow{a.s.} \infty$. We now switch over to distributional convergence, since it is awkward to work with almost sure convergence for conditioned random variables. The *distribution* of a random variable conditioned on an event, however, *is* defined, so that we can define *distributional* convergence for a sequence of random variables conditioned on a sequence of events. Combining the fact that almost sure convergence implies distributional convergence with the fact that $\mathbb{P}(S_1 > 0, \ldots, S_m > 0) \to 1 - p_\lambda > 0$, the survival probability of the branching process, yields

$$\big(N(A_m)\mathrm{e}^{-(\lambda-1)A_m} | S_1 > 0, \ldots, S_m > 0\big) \xrightarrow{d} (W | W > 0) \stackrel{d}{=} W_\lambda. \tag{3.16}$$

Now $N(t)$ is the number of *alive* individuals at time t, so after m splits $N(A_m)/m = S_m/m \xrightarrow{a.s.} \lambda - 1$, by the strong law and from this it is not hard to see that $N(A_m)/m | S_i > 0$, $1 \leqslant i \leqslant m \xrightarrow{\mathbb{P}} \lambda - 1$. Hence, with $\gamma = 1/(\lambda - 1)$,

$$\big(A_m | S_1 > 0, \ldots, S_m > 0\big) - \gamma \log m \xrightarrow{d} -\gamma \log(\gamma W_\lambda). \tag{3.17}$$

This completes the proof of Theorem 3.5. □

Proof of Proposition 3.7. We start with the proofs of parts (b) and (c). Note that $w_m = o(\log\log m)$, and that the random variables D_j with $j > w_m$ are independent. The proof of (b) is then straightforward from the large deviation properties of the Poisson distribution and a simple union bound. Alternatively, one can prove (b) by proving (3.13) unconditionally, *i.e.*, with S_j^*

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

replaced by S_j and noting that the conditioning $\mathbb{P}(S_1 > 0, \dots, S_m > 0)$ is bounded from below by $1 - p_\lambda$.

For (c) we write

$$\sum_{i=1}^{m} \rho_i^* = \sum_{i=1}^{m} \frac{D_i^*}{S_i^*} = \sum_{i \leqslant \log\log m} \frac{D_i^*}{S_i^*} + \sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{S_i^*}. \tag{3.18}$$

From part (b), we can w.h.p. sandwich the second sum on the right-hand side, as follows:

$$\sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{(\lambda-1)i + C\sqrt{i \log i}} \leqslant \sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{S_i^*} \leqslant \sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{(\lambda-1)i - C\sqrt{i \log i}}.$$

Furthermore,

$$\sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{(\lambda-1)i - C\sqrt{i \log i}} - \sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{(\lambda-1)i} \leqslant \sum_{i=\lceil \log\log m \rceil}^{m} C' D_i^* \frac{\sqrt{\log i}}{i\sqrt{i}} = O_{\mathbb{P}}(1),$$

and

$$\sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{(\lambda-1)i} = \sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^* - \lambda}{(\lambda-1)i} + \sum_{i=\lceil \log\log m \rceil}^{m} \frac{\beta}{i}.$$

The above estimates can be repeated verbatim to deal with the error term on the other side of the sandwich. Using Chebyshev's inequality on the i.i.d. Poisson random variables D_i^*, $i \geqslant \lceil \log\log m \rceil$, we obtain that for $A \to \infty$,

$$\mathbb{P}\Bigg(\sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^* - \lambda}{(\lambda-1)i} > A \Bigg) \leqslant \frac{1}{A^2} \sum_{i=\lceil \log\log m \rceil}^{m} \frac{\lambda}{i^2(\lambda-1)^2} \to 0.$$

Together this shows that

$$\sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{S_i^*} = \sum_{i=\lceil \log\log m \rceil}^{m} \frac{D_i^*}{(\lambda-1)i} + O_{\mathbb{P}}(1) = \sum_{i=\lceil \log\log m \rceil}^{m} \frac{\beta}{i} + O_{\mathbb{P}}(1). \tag{3.19}$$

By definition, $D_i/S_i = D_i/(D_i + (S_{i-1} - 1))$. Hence, if we condition on $S_{i-1} > 0$ or $S_{i-1} \geqslant 1$, then we immediately get $D_i^*/S_i^* \leqslant 1$. Hence,

$$\sum_{j=1}^{i} \frac{D_i^*}{S_i^*} \leqslant i.$$

This implies that the first sum on the right-hand side of (3.18) is obviously of order $O_{\mathbb{P}}(\log\log m)$. Together with (3.19), this proves the first assertion in part (c). The proof that

$$\sum_{i=1}^{m} (\rho_i^*)^2 = O_{\mathbb{P}}(\log\log m),$$

goes in a similar fashion, since obviously

$$\sum_{i \leqslant \log\log m} (\rho_i^*)^2 \leqslant \log\log m,$$

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

whereas by a similar construction as before, we can bound $\sum_{i=\lceil \log\log m\rceil}^{m}(\rho_i^*)^2$ by

$$\sum_{i=\lceil \log\log m\rceil}^{m} C''(D_i^*)^2/i^2 = O_{\mathbb{P}}(1).$$

This completes the proof of part (c).

We now move to part (a). As before, we denote by $S_i = \sum_{j=1}^{i} D_j - (i-1)$, $i \geqslant 1$, the unconstrained random walk. We define for $k \geqslant 1$, the decreasing events

$$\mathcal{E}_k = \{S_1 > 0, \ldots, S_k > 0\}. \tag{3.20}$$

Then

$$\lim_{m\to\infty} \mathbb{P}(S_1^* > 0, \ldots, S_m^* > 0) = \lim_{m\to\infty} \mathbb{P}(\mathcal{E}_m | \mathcal{E}_{w_m}) = \lim_{m\to\infty} \frac{\mathbb{P}(\mathcal{E}_m)}{\mathbb{P}(\mathcal{E}_{w_m})} = 1, \tag{3.21}$$

since $\mathbb{P}(\mathcal{E}_k)$ tends to the positive limit $1 - p_\lambda$.

For the coupling we introduce two independent copies of the (unconditioned) random walk S_i, $i \geqslant 1$. With the help of Kolmogorov's extension theorem we introduce the random walks $S_i^{(1)}(\omega), i \geqslant 1$ and $S_i^{(2)}(\omega), i \geqslant 1$, where $S_i^{(1)}(\omega), i \geqslant 1$ is equal in distribution to the first copy conditioned on $\mathcal{E}_{w_m}^{(1)}$ and where $S_i^{(2)}(\omega), i \geqslant 1$ is equal in distribution to the second copy conditioned on $\mathcal{E}_m^{(2)}$. Finally we set $S_i^*(\omega) = S_i^{(1)}(\omega)$ and

$$\tilde{S}_i(\omega) = \begin{cases} S_i^{(1)}(\omega) & \omega \in \mathcal{E}_m^{(1)}, \\ S_i^{(2)}(\omega) & \text{otherwise.} \end{cases} \tag{3.22}$$

Then, clearly $\tilde{S}_i$ and S_i^* have the correct marginal distribution and

$$\mathbb{P}(S_j^* = \tilde{S}_j,\ 1 \leqslant j \leqslant m) \geqslant \mathbb{P}(\mathcal{E}_m^{(1)} | \mathcal{E}_{w_m}^{(1)}) \to 1, \tag{3.23}$$

by (3.21). □

4. Branching process trees and embedding of ERRG

We introduce marked branching process trees in Section 4.1. In Section 4.2, we attach the exponentially distributed weights to the edges and obtain in this way a continuous-time marked branching process (CTMBP) with binomially distributed offspring. Section 4.3 describes the distribution of the height and weight of minimal weight paths in these CTMBP. In Section 4.4, we treat the connection time of two independent CTMBPs with the same mark set $[n]$, one starting from a root with mark 1, whereas the second process starts from a root with mark 2.

4.1. Marked branching process trees

In this section we describe a marked branching process tree (MBPT) with binomially distributed offspring, and we show how this MBPT should be thinned to obtain the connection with the ERRG. We then attach i.i.d. weights having probability density f on $(0, \infty)$ to both the branching process tree and the ERRG. We interpret these weights as distances between marks, or vertices, respectively. We close the section with a proof that for each time $t \geqslant 0$, the set of marks that can be reached in the thinned process within time t from the root is identical in distribution to the set of vertices that can be reached within time t from the initial vertex i_0 in the ERRG.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

The set of marks is denoted by $[n] = \{1, 2, \ldots, n\}$. In the branching process tree, we start with a single individual with mark $i_0 \in [n]$, and we put $I_0 = \{i_0\}$. This individual reproduces a binomially distributed number of offspring X_1, with parameters $n-1$ and p. We attach each of these X_1 children to their father by a single edge, so that we obtain a tree structure. Each edge is assigned a weight taken from an i.i.d. sample having probability density f, with support $(0, \infty)$. To obtain the marks of the offspring we take at random a sample of size X_1 from the set $[n] \setminus \{i_0\}$. The corresponding set of marks is denoted by I_1. Hence, the marks are chosen uniformly at random from a set of size $n-1$ and are all different. Moreover, the mark of the father (i_0) is not present among the set of marks I_1 of the offspring. Observe that the marks can be seen as vertex numbers of the ERRG. From this viewpoint, the children of individual i_0, given by the vertex set I_1, are the direct neighbours of the vertex i_0 in the ERRG.

We proceed by taking the individual for which the corresponding edge weight is minimal. In our language this individual is reached after T_1 time units, where T_1 is equal to the length of the minimal edge weight, since the fluid percolates at rate 1. The mark reached by time T_1 is denoted by i_1. The reproduction process for i_1 is identical to that of i_0: we take a $\mathrm{Bin}(n-1, p)$ number of children (denoted by X_2) and connect each of them by a single edge to i_1. The X_2 'new' edges are supplied by weights taken i.i.d. with probability density f. We update the 'old' edges by subtracting T_1 from their respective weights, because T_1 time units have been used. We then complete this step by assigning marks to the new individuals, taken uniformly at random from the set $[n] \setminus \{i_1\}$; this mark set is denoted by I_2.

We proceed by induction: suppose that we have reached the individuals $i_0, i_1, \ldots, i_k$, in this order, so that the (updated) weight attached to the edge with endpoint i_k was the minimal edge weight at time $A_k = T_1 + T_2 + \cdots + T_k$. We now form $X_{k+1} \stackrel{d}{=} \mathrm{Bin}(n-1, p)$ new edges and attach them to i_k. These edges are supplied by weights, taken i.i.d. with probability density f, whereas at the same time, the weights to all other edges emanating from $i_0, i_1, \ldots, i_{k-1}$ are updated by subtracting T_k. We assign marks analogously as before by drawing them uniformly at random from the set $[n] \setminus \{i_k\}$. The resulting mark set is denoted by I_{k+1}.

We thin the MBPT defined above as follows. At each time point $A_k = T_1 + T_2 + \cdots + T_k$ we delete the newly found individual with mark i_k and the entire tree emanating from i_k, when

$$i_k \in \{i_0, i_1, \ldots, i_{k-1}\}, \tag{4.1}$$

i.e., we delete i_k and all its offspring when the mark i_k has appeared previously. Observe that the probability of the event (4.1) is at most k/n. The sets $\{\hat{I}_k\}_{k \geqslant 0}$ are obtained from $\{I_k\}_{k \geqslant 0}$, by deleting the thinned vertices.

Obviously, for each fixed n, the thinned MBPT will become empty after a finite time, even in the super critical case ($(n-1)p_n > 1$), whereas the *un-thinned* MBPT will survive with positive probability. This is because we sample the marks from the finite set $[n]$, and hence with probability one at a certain random time all marks will have appeared. We consider the MBPT and its thinned version up to the first generation that the thinned process becomes empty. Furthermore, consider the ERRG with n vertices and attach weights to the edges independently with the same marginal density f as used in the branching process. Then, by the independence of binomial random variables $X_1, X_2, \ldots$ and the thinning described above, for each $t \geqslant 0$, the set of marks that are reached by time t in the thinned MBPT and the set of vertices that are reached by time t

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

in the ERRG, are equal in distribution. Since this holds for each $t \geqslant 0$, the proof of the following lemma is obvious.

Lemma 4.1 (FPP on ERRG is thinned CTMBP). *Fix $n \geqslant 1$ and $p \in (0,1)$. Consider the thinned MBPT until extinction with i.i.d. weights having density f with support $(0,\infty)$. If we apply independently an identical weight construction to the ERRG, then for any $i_0 \in [n]$, the weight $W_n(i_0, j)$ of the shortest path between initial vertex i_0 and any other vertex $j \in [n]$ in the ERRG is equal in distribution to the weight of the shortest path between i_0 and j in the thinned MBPT.*

Remark 4.2. **(a)** In both random environments (thinned MBPT and ERRG) it can occur that i_0 and j are not connected. We then put the weight equal to $+\infty$.

(b) Observe that since the weight distribution admits a density and is defined on finite objects, the minimal weight, if finite, uniquely identifies the minimal path and hence the number of edges on this minimal path, *i.e.*, the hopcount. Hence, Lemma 4.1 also proves that the hopcount between i_0 and j in the ERRG is equal in distribution to the hopcount between i_0 and j in the thinned MBPT.

4.2. The continuous-time branching process

We study shortest-weight distance between uniformly chosen *connected* vertices in $G_n(p_n)$ with exponentially distributed edge weights. Since the vertices of the ERRG are exchangeable we can as well study the distance between vertex 1 and vertex 2 conditioned on the event that 1 and 2 are in the giant component. Note that the condition that two vertices are in the giant component is w.h.p. equivalent to the condition that these vertices are connected, since other components are of size o(n). For the remainder of the paper we put

$$f(x) = \mathrm{e}^{-x}, \quad x > 0, \tag{4.2}$$

the density of the exponential distribution with rate 1. For the initial vertex we take either $i_0 = 1$ or $i_0 = 2$, and, if necessary, we will use a superscript to indicate from which vertex we start. Since we have a (*distributional*) embedding of the vertices of $G_n(p)$, reached by time t, in the marked branching process tree (see Lemma 4.1 and Remark 4.2), we can restrict our attention exclusively to the MBPT and its thinned version.

We start from the individual with mark 1 and let fluid percolate at unit speed from this individual until it reaches one of the children. Because the weights are exponentially distributed (see (4.2)), the number of individuals that has been reached by time $t \geqslant 0$ form a *continuous-time* branching process (CTMBP) [3, Chapter 2]. The *alive* individuals are defined as those individuals which are directly connected to the individuals reached, appropriately called *wetted* individuals. At time $t = 0$ this number is $S_1 = X_1$. Then, at the first splitting time T_1 at which time the second individual is wetted, the number of alive vertices equals $S_2 = X_1 + X_2 - 1$, since X_2 new individuals are born and one individual gave birth and died (became wet). So, starting from the i.i.d. sequence $X_1, X_2, \ldots$, we define the random walk

$$S_k = X_1 + \cdots + X_k - (k-1), \quad k \geqslant 1 \tag{4.3}$$

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

and the inter-splitting times T_k, $k \geqslant 1$, where conditioned on $X_1, X_2, \ldots, X_k$, the random variables $T_1, T_2, \ldots, T_k$ are independent and where T_j has an exponential distribution with parameter S_j. Indeed, by the memoryless property of the exponential distribution, when $t = A_j = T_1 + \cdots + T_j$, we have S_{j+1} *alive* individuals and all their S_{j+1} edges, with which they are attached to the dead or wet individuals, have an independent exponentially distributed random variable with rate 1 as weight. Therefore the inter-splitting time T_{j+1} until the next individual is wetted, conditioned on S_{j+1}, is equal in distribution to $\min_{1 \leqslant s \leqslant S_{j+1}} E_s$, where $E_1, E_2, \ldots$ are independent rate 1 exponentially distributed random variables. This shows that the conditional distribution of T_{j+1} is the correct one, since the minimum of q independent Exp(1) has an exponential distribution with rate q.

At each splitting time A_k, $k \geqslant 1$, the number of wetted individuals equals $k+1$, whereas the number of edges and endpoints competing to become wet equals S_{k+1}. We now introduce *shortest-weight trees* (SWTs). Since the ERRG is embedded in the CTMBP, we introduce the SWT in terms of marks and splitting times. The shortest-weight tree SWT_k is equal to the *ordered* collection of marks that are wetted at the kth splitting time A_k and we include in the definition the splitting times $A_1, A_2, \ldots, A_k$. So, $\mathrm{SWT}_0 = (i_0, A_0 = 0)$, the mark of the root, and

$$\mathrm{SWT}_k = \big((i_0, i_1, \ldots, i_k), (A_0, A_1, \ldots, A_k)\big), \quad k \geqslant 1, \tag{4.4}$$

where $i_0, i_1, \ldots, i_k$, denote the marks reached at the splitting times $A_0, A_1, \ldots, A_k$, respectively. To connect i_j with the splitting time A_j, we introduce the mapping $t(i_j) = j$, so that mark $i \in \mathrm{SWT}_k$ was reached at time $A_{t(i)}$. Also, with some misuse of notation we will write $i \in \mathrm{SWT}_k$ meaning that the first $k+1$-tuple of SWT_k contains the index i. For a fixed mark $i \in \mathrm{SWT}_k$ and all marks $j \neq i$, the random variables $\mathbf{1}_{ij}|\mathrm{SWT}_k$ that indicate whether edge ij is in the branching process tree (or, in other words, j is one of the alive individuals born out of vertex i) are *independent* and satisfy

$$\mathbb{P}\big(\mathbf{1}_{ij} = 1|\mathrm{SWT}_k\big) = p_n. \tag{4.5}$$

Furthermore, let us fix $k \geqslant 0$ and consider the in total S_{k+1} alive individuals at the splitting time A_k. The edges between the wet mark i and the mark j, corresponding to a alive individual with mark j, has weight E_{ij}, $i \in \mathrm{SWT}_k$. Now observe from the memoryless property of the exponential distribution, that, conditional on SWT_k,

$$E_{ij}|\mathrm{SWT}_k \stackrel{d}{=} A_k - A_{t(i)} + E'_{ij}, \tag{4.6}$$

where E'_{ij} are independent Exp(1) random variables. This completes the description of our CTMBP and its relation to FPP on the ERRG.

4.3. Hopcount and weight from root to newly added individuals

Regarding the height and the weight of the path with minimal weight from root 1 in $\{\mathrm{SWT}_k\}_{k \geqslant 0}$ to the mth individual that has been wetted we refer to (3.4) and (3.6), respectively, where the random variables $I_1, I_2, \ldots, I_m$ are conditionally independent given $X_1, X_2, \ldots, X_m$, the binomially distributed offspring in the tree with root 1, and where $d_j = X_j$, $s_j = S_j$. Indeed, the height of the mth individual that has been wetted in the CTMBP is equal to the generation G_m of that individual and its weight is equal to the splitting time $A_m = T_1 + \cdots + T_m$, where, conditional

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

on $X_1, X_2, \ldots, X_m$, the distribution of T_j is given by

$$T_j \stackrel{d}{=} E_j/S_j, \tag{4.7}$$

and where $\{E_i\}_{i=1}^{\infty}$ are i.i.d. exponential random variables with mean 1, independent of all random variables introduced earlier. The CLT for the hopcount G_m and the limit law for the weight A_m in the CTMBP can be deduced from Theorem 3.5 in the following way.

Proposition 4.3 (Asymptotics for shortest-weight paths in the CTMBP). *Consider an i.i.d. sequence $X_1, X_2, \ldots$, with a binomial distribution where the parameters $n-1$ and p_n satisfy $\lim_{n\to\infty} np_n = \lambda > 1$, and put $\beta = \lambda/(\lambda-1)$. Furthermore, let G_m and A_m be defined as in* (3.4) *and* (3.6)*, respectively, with $d_i = X_i, i \geqslant 1$. Then, conditional on $S_i = X_1 + \cdots + X_i - (i-1) > 0$ for all $1 \leqslant i \leqslant m$, the asymptotics in Theorem 3.5*(a, b) *continue to hold, where the limits are independent.*

Proof. According to [12, Theorem 2.9], we can couple $X \sim \mathrm{Bin}(n-1, p_n)$ and $D \sim \mathrm{Poi}((n-1)p_n)$, by means of a pair of variables $(\tilde{X}, \tilde{D})$ with $\tilde{X} \sim \mathrm{Bin}(n-1, p_n)$ and $\tilde{D} \sim \mathrm{Poi}((n-1)p_n)$, so that for any $\varepsilon > 0$, and n sufficiently large,

$$\mathbb{P}(\tilde{X} \neq \tilde{D}) \leqslant (n-1)p_n^2 \leqslant \frac{(\lambda+\varepsilon)^2}{n}. \tag{4.8}$$

Hence, for $m_n = o(n)$, we can couple the i.i.d. sequence $X_1, X_2, \ldots, X_{m_n}$ to an i.i.d. sequence of Poisson variables $D_1, D_2, \ldots, D_{m_n}$, such that

$$\mathbb{P}(\tilde{X}_i \neq \tilde{D}_i, \text{ for some } i \leqslant m_n) \leqslant \sum_{i=1}^{m_n} \frac{(\lambda+\varepsilon)^2}{n} = m_n \frac{(\lambda+\varepsilon)^2}{n} \to 0. \tag{4.9}$$

Part (a) of Theorem 3.5 yields that

$$\frac{G_{m_n} - \beta \log m_n}{\sqrt{\beta \log m_n}}$$

is asymptotically normal when

$$G_{m_n} = I_1 + I_2 + \cdots + I_{m_n},$$

where

$$\mathbb{P}(I_j = 1 | D_1, D_2, \ldots, D_{m_n}) = D_j/S_j, \quad 1 \leqslant j \leqslant m_n,$$

are based on *Poisson* random variables $D_1, D_2, \ldots, D_{m_n}$. We now introduce (conditional) Bernoulli random variables $J_1, J_2, \ldots, J_{m_n}$ and $K_1, K_2, \ldots, K_{m_n}$ belonging to the coupled sequences $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{m_n}$ and $\tilde{D}_1, \tilde{D}_2, \ldots, \tilde{D}_{m_n}$ so that we can conclude the asymptotic normality of

$$\tilde{G}^X_{m_n} = J_1 + J_2 + \cdots + J_{m_n},$$

based on *binomially* distributed random variables $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_{m_n}$, from that of

$$\tilde{G}^D_{m_n} = K_1 + K_2 + \cdots + K_{m_n},$$

which is based on *Poissonian* distributed random variables $\tilde{D}_1, \tilde{D}_2, \ldots, \tilde{D}_{m_n}$, and hence follows from Theorem 3.5(a). In order to do so, we define $(J_i, K_i), 1 \leqslant i \leqslant m_n$, conditioned on

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

$(\tilde{X}_1, \tilde{D}_1), \ldots, (\tilde{X}_{m_n}, \tilde{D}_{m_n})$ as follows. Let $U_1, U_2, \ldots, U_{m_n}$ be an i.i.d. sequence of uniform $(0,1)$ random variables. For each i, the sample space of (J_i, K_i) is $\{0,1\}^2$, and the conditional probabilities are defined by

$$\mathbb{P}(J_i = 1 \mid U_i) = \mathbf{1}_{\{U_i \leqslant \frac{\tilde{X}_i}{\tilde{S}_i^X}\}}, \quad \mathbb{P}(K_i = 1 \mid U_i) = \mathbf{1}_{\{U_i \leqslant \frac{\tilde{D}_i}{\tilde{S}_i^D}\}} \tag{4.10}$$

where $\tilde{S}_i^X = \tilde{X}_1 + \cdots + \tilde{X}_i - (i-1)$, $\tilde{S}_i^D = \tilde{D}_1 + \cdots + \tilde{D}_i - (i-1)$, and where $\mathbf{1}_{\mathcal{A}}$ denotes the indicator of the set $\mathcal{A}$. Note that the joint distribution of (J_i, K_i), conditioned on $(\tilde{X}_1, \tilde{D}_1), \ldots, (\tilde{X}_m, \tilde{D}_m)$, is completely specified by the probabilities in (4.10). Finally, set

$$\mathcal{A}_m = \{(\tilde{X}_1 = \tilde{D}_1), \ldots, (\tilde{X}_m = \tilde{D}_m)\}. \tag{4.11}$$

Then, by (4.9), we have that $\mathbb{P}(\mathcal{A}_{m_n}) \to 1$, and

$$\mathbb{P}\left(\sum_{i=1}^{m_n} J_i = \sum_{i=1}^{m_n} K_i\right) \geqslant \mathbb{P}(\mathcal{A}_{m_n}) \to 1, \tag{4.12}$$

because, on $\mathcal{A}_{m_n}$, we have $J_i = K_i$, $1 \leqslant i \leqslant m_n$. In order to finish the proof of part (a) of the present proposition, we have to show that $\sum_{i=1}^{m_n} J_i$ conditioned on $\tilde{S}_i^X > 0$, $1 \leqslant i \leqslant m_n$ has w.h.p. the same distribution as $\sum_{i=1}^{m_n} K_i$ conditioned on $\tilde{S}_i^D > 0$, $1 \leqslant i \leqslant m_n$. Observe that this statement follows from (4.11) and (4.12) since, by the coupling introduced above,

$$\mathbb{P}(\tilde{S}_i^D > 0,\ 1 \leqslant i \leqslant m_n) \to 1 - p_\lambda > 0. \tag{4.13}$$

Hence, part (a) follows from Theorem 3.5(a).

Referring to Theorem 3.5(b) and the definition of A_m in (3.6), we conclude that the proof of part (b) also follows in a straightforward manner from Theorem 3.5(b) using the coupling between

$$\tilde{A}_{m_n}^X = \sum_{i=1}^{m_n} E_i / \tilde{S}_i^X \quad \text{and} \quad \tilde{A}_{m_n}^D = \sum_{i=1}^{m_n} E_i / \tilde{S}_i^D. \qquad \square$$

4.4. The connection time

In definition (4.4) we introduced $\{\mathrm{SWT}_k\}_{k \geqslant 0}$, which includes the set of marks of the individuals that are wet after the kth split. More precisely, we grow a CTMBP from a root with mark 1 and include the marks that are successively reached and their splitting times in $\{\mathrm{SWT}_k^{(1)}\}_{k \geqslant 0}$. Our plan is to grow $\mathrm{SWT}_k^{(1)}$ until the set of marks has reached size $a_n = \lceil \sqrt{n} \rceil$; the branching process then contains a_n wet individuals and the same marks can appear more than once. Then, we independently grow a second marked branching process starting from a root with mark 2. We denote the shortest-weight tree, *i.e.*, the wetted marks and their splitting times of this second process, by $\{\mathrm{SWT}_k^{(2)}\}_{k \geqslant 0}$. We stop the first time that this second process contains a mark from $\mathrm{SWT}_{a_n}^{(1)}$, *i.e.*, we stop growing $\{\mathrm{SWT}_k^{(2)}\}_{k \geqslant 0}$ at the *random* time C_n defined by

$$\mathrm{C}_n = \min\{m \geqslant 0 : \mathrm{SWT}_m^{(2)} \cap \mathrm{SWT}_{a_n}^{(1)} \neq \varnothing\}, \tag{4.14}$$

where the set $\mathrm{SWT}_m^{(2)} \cap \mathrm{SWT}_{a_n}^{(1)}$ contains the common marks in $\mathrm{SWT}_m^{(2)}$ and $\mathrm{SWT}_{a_n}^{(1)}$. We now reach the *main* theorem of this section, where we establish the connection between the weight

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

W_n and the hopcount H_n in the ERRG and the weight and height in the shortest-weight trees $\{\mathrm{SWT}_k^{(i)}\}_{k\geqslant 0}$, $i = 1, 2$, introduced in this section.

Theorem 4.4 (Connecting hopcount to heights in CTMBP). *Let H_n, W_n denote the hopcount and the weight, respectively, of the minimal path between the vertices* 1 *and* 2 *in the ERRG $G_n(p_n)$, where we condition on the vertices* 1 *and* 2 *to be in the giant component.*

(a) *For $n \to \infty$,*

$$\mathbb{P}(H_n = G^{(1)}_{U_n,a_n} + G^{(2)}_{\mathrm{C}_n}) = 1 - o(1), \tag{4.15}$$

where G_{U_n,a_n} is the height (or generation) of a uniformly chosen mark in $\mathrm{SWT}^{(1)}_{a_n}$, and $G^{(2)}_{\mathrm{C}_n}$ is the height of the mark attached at the random time C_n defined in (4.14)*, and where, conditioned on C_n, the variables $G^{(1)}_{U_n,a_n}$ and $G^{(2)}_{\mathrm{C}_n}$ are independent.*

(b) *Similarly, for $n \to \infty$,*

$$\mathbb{P}(W_n = A^{(1)}_{a_n} + A^{(2)}_{\mathrm{C}_n}) = 1 - o(1), \tag{4.16}$$

where $A^{(i)}_m$, $i = 1, 2$, is the splitting time of the mth mark in $\{\mathrm{SWT}^{(i)}_k\}_{k\geqslant 0}$.

For the moment, we postpone the proof of Theorem 4.4. We first need to establish two intermediate results. Observe that if in $\mathrm{SWT}^{(i)}_k$, $i = 1, 2$, all marks on the minimal weight path between the root i and some other mark $U \in \mathrm{SWT}^{(i)}_k$ appeared for the first time, then the entire path is contained in the thinned branching process, and by Lemma 4.1 and Remark 4.2, both the weight and the hopcount in the ERRG between the vertices i and U are equal in distribution to $A^{(i)}_U$ and $G^{(i)}_U$, respectively, where $A^{(i)}_U$ denotes the splitting time of mark U and $G^{(i)}_U$ denotes the generation of mark U. Hence, in order to prove Theorem 4.4, we need an upper bound on the expected number of marks in $\mathrm{SWT}^{(i)}_k$ whose minimal weight path contains thinned marks. This will be the content of the first lemma below. The second lemma below shows that C_n/a_n converges to an Exp(1) random variable, so that we have a handle on the size of C_n.

Lemma 4.5 (A coupling bound). *Fix $k \geqslant 1$ and denote by $M^{(i)}_k$, $i = 1, 2$, the number of marks in $\mathrm{SWT}^{(i)}_k$ for which the minimal path from root i to that mark contains a thinned mark. Then*

$$\mathbb{E}[M^{(i)}_k] \leqslant \frac{k^2}{n - a_n - 1}, \quad i = 1, 2. \tag{4.17}$$

The above inequality holds in $\mathrm{SWT}^{(1)}_k$ for $1 \leqslant k \leqslant a_n$ and in $\mathrm{SWT}^{(2)}_k$ for $1 \leqslant k \leqslant \mathrm{C}_n$.

Proof. We start with $\mathrm{SWT}^{(1)}_k$. Let $\mathrm{SWT}^{(1)}_k$ contain $M^{(1)}_k$ marks whose minimal path contains at least one thinned mark. The new mark which is drawn to be the mark of the newly wetted vertex in $\mathrm{SWT}^{(1)}_{k+1}$ attaches to one of these $M^{(1)}_k$ marks with probability $M^{(1)}_k/(k+1)$, since $|\mathrm{SWT}^{(1)}_k| = k + 1$. If the new mark is attached to a vertex with one of the $k + 1 - M^{(1)}_k$ other marks, then, with probability at most $(k+1)/n$, this mark has appeared previously and has to be thinned. Hence,

$$\mathbb{E}[M^{(1)}_{k+1} - M^{(1)}_k | M^{(1)}_k] \leqslant \frac{k+1}{n} + \frac{M^{(1)}_k}{k+1}. \tag{4.18}$$

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

For $\mathrm{SWT}_k^{(2)}$ the fraction $\frac{k+1}{n}$ in the recursion should be replaced by $\frac{k+1}{n-a_n-1}$, because for each $1 \leqslant k \leqslant \mathrm{C}_n$ we know that $\mathrm{SWT}_{a_n}^{(1)}$ contains at most a_n+1 different marks. Write $b_k = \mathbb{E}[M_k^{(1)}]$. Taking double expectations in (4.18) and solving the recursive inequality yields

$$\frac{b_{k+1}}{k+1} \leqslant \frac{1}{n} + \frac{b_k}{k}, \quad \text{so that} \quad \frac{b_{k+1}}{k+1} \leqslant \frac{k}{n}.$$

This yields $\mathbb{E}[M_k^{(1)}] \leqslant \frac{k^2}{n}$. The proof for $\mathrm{SWT}_k^{(2)}$ is similar; however, since at most a_n marks are occupied by $\mathrm{SWT}_{a_n}^{(1)}$, we need to replace n by $n-a_n-1$ and this yields (4.17). □

Lemma 4.6 (Weak convergence of connection time). *Conditioned on the event that both CTMBPs survive, the connection time* C_n *defined in* (4.14) *satisfies the asymptotics*

$$\mathrm{C}_n/a_n \xrightarrow{d} E, \tag{4.19}$$

where E *has an exponential distribution with rate* 1.

Proof. Define by $\mathbb{Q}_n^{(j)}$ the conditional probability given both $\mathrm{SWT}_{a_n}^{(1)}$ and $\mathrm{SWT}_j^{(2)}$. As for [5, Lemma B.1], the probability $\mathbb{P}(\mathrm{C}_n > m)$ has the product form

$$\mathbb{P}(\mathrm{C}_n > m) = \mathbb{E}\left[\prod_{j=1}^{m} \mathbb{Q}_n^{(j-1)}(\mathrm{C}_n > j | \mathrm{C}_n > j-1)\right], \tag{4.20}$$

where the expectation is taken with respect to the conditional events introduced by $\mathbb{Q}_n^{(j)}$. We omit the proof of (4.20), which follows by suitable conditioning arguments and is identical to [5, Proof of Lemma B.1]. Since both branching processes survive, $\mathrm{SWT}_{a_n}^{(1)}$ contains $a_n(1+o(1))$ different marks and by (4.17) the expected number of multiple marks is $O(1)$. From this we obtain that as long as the second branching process survives up to size at least j,

$$\mathbb{Q}_n^{(j-1)}(\mathrm{C}_n > j | \mathrm{C}_n > j-1) = 1 - \frac{a_n(1+o(1))}{n}.$$

Substitution of this into (4.20) with $m = a_n x$, and $x > 0$, yields

$$\mathbb{P}(\mathrm{C}_n > a_n x) = \left[1 - \frac{a_n(1+o(1))}{n}\right]^{a_n x} \to \mathrm{e}^{-x}, \tag{4.21}$$

and we arrive at (4.19). □

Proof of Theorem 4.4. We start with the proof of the hopcount. Since 1 and 2 are conditioned to be connected, with high probability, 1 and 2 are contained in the giant component of the ERRG. This implies w.h.p. that both CTMBPs do not die out. Consider the mark set $\mathrm{SWT}_{a_n}^{(1)}$ and choose one of the wet individuals at random. The mark number of this individual will be denoted by U_n. Then the probability that the shortest-weight path from root 1 to U_n contains a thinned mark is at most

$$\frac{\mathbb{E}[M_{a_n}^{(1)}]}{a_n+1} \leqslant \frac{a_n^2}{(a_n+1)(n-a_n-1)} \to 0. \tag{4.22}$$

This shows that w.h.p. the hopcount between vertex 1 and vertex U_n in $G_n(p_n)$ is given by $G_{U_n,a_n}^{(1)}$. By Lemma 4.6, $\mathrm{C}_n = o_{\mathbb{P}}(\sqrt{n}\log n)$, and a similar argument shows that w.h.p. the hopcount between 2 and the individual corresponding to the last mark added to $\mathrm{SWT}_{\mathrm{C}_n}^{(2)}$ is equal to $G_{\mathrm{C}_n}^{(2)}$.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

Now recall the discussion around (4.5) and (4.6). We conclude from this discussion that indeed at time C_n one of the marks in $\mathrm{SWT}_{a_n}^{(1)}$ is chosen uniformly at random and this establishes (4.15). Finally, we introduce the notation $p(i)$ for $i \in \mathrm{SWT}_k$, to denote the parent of mark i, *i.e.*, the mark to which i was attached in SWT_k. Because of the way the exponential weight in (4.6) can be re-allocated to edges in $\mathrm{SWT}_{a_n}^{(1)}$, the edge $\{p(i), i\}$ gets weight $A_{a_n} - A_{t(i)}$ and the weight E'_{ij} is allocated to ij, with $i \in \mathrm{SWT}_{a_n}^{(1)}$ and $j \notin \mathrm{SWT}_{a_n}^{(1)}$, so that we get (4.16). □

5. Proof of the main results

In this section, we complete the proof of our main results.

Proof of Theorem 2.1. Observe that the statistical dependence between the two CTMBPs, *i.e.*, the one that grows from the root with mark 1 and the one that grows from the root with mark 2, is introduced *only* through their marks. Moreover, in each tree, the hopcount of the mth added individual is independent of the marks. Hence, conditioned on the event $\{\mathrm{C}_n = m\}$, the random variables $G^{(1)}_{U_n,a_n}$ and $G^{(2)}_{\mathrm{C}_n} = G^{(2)}_m$ are statistically independent.

Let $Z_n^{(1)} = (G^{(1)}_{U_n,a_n} - \beta \log a_n)/\sqrt{\beta \log a_n}$, $Z_n^{(2)} = (G^{(2)}_{\mathrm{C}_n} - \beta \log \mathrm{C}_n)/\sqrt{\beta \log \mathrm{C}_n}$, and ξ_1, ξ_2 denote two independent standard normal random variables. Since the branching process is Markovian, new individuals are attached to the SWT uniformly at random. Hence the limit behaviour of $G^{(1)}_{U_n,a_n}$ is the same as that of $G^{(1)}_{a_n}$. Then, by Proposition 4.3, for any sequence $m_n = o(n) \to \infty$, and for any bounded continuous function g of two real variables,

$$\lim_{n\to\infty} \mathbb{E}[g(Z_n^{(1)}, Z_n^{(2)}) | \mathrm{C}_n = m_n] = \mathbb{E}[g(\xi_1, \xi_2)]. \tag{5.1}$$

Hence, by bounded convergence (the function g is bounded), and using (4.19),

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}[g(Z_n^{(1)}, Z_n^{(2)})] &= \int_0^\infty \lim_{n\to\infty} \mathbb{E}\big[g(Z_n^{(1)}, Z_n^{(2)}) | \mathrm{C}_n = \lceil a_n y \rceil\big]\, d\mathbb{P}(\mathrm{C}_n \leqslant \lceil a_n y \rceil) \\ &= \int_0^\infty \mathbb{E}[g(\xi_1, \xi_2)] \mathrm{e}^{-y}\, dy = \mathbb{E}[g(\xi_1, \xi_2)].\end{aligned}$$

This shows the joint weak convergence of the pair $(Z_n^{(1)}, Z_n^{(2)})$ to (ξ_1, ξ_2). By the continuous mapping theorem [6, Theorem 5.1] applied to $(x, y) \mapsto x + y$, we then find that

$$\begin{aligned}\frac{H_n - \beta \log n}{\sqrt{\log n}} &= \frac{G^{(1)}_{U_n,a_n} - (\beta/2)\log n}{\sqrt{\log n}} + \frac{G^{(2)}_{\mathrm{C}_n} - (\beta/2)\log n}{\sqrt{\log n}} \\ &= \frac{G^{(1)}_{U_n,a_n} - \beta \log a_n}{\sqrt{\log a_n}} \cdot \frac{\sqrt{\log a_n}}{\sqrt{\log n}} + \frac{G^{(2)}_{\mathrm{C}_n} - \beta \log \mathrm{C}_n}{\sqrt{\log \mathrm{C}_n}} \cdot \frac{\sqrt{\log \mathrm{C}_n}}{\sqrt{\log n}} + o_{\mathbb{P}}(1) \\ &\xrightarrow{d} \xi_1/\sqrt{2} + \xi_2/\sqrt{2},\end{aligned} \tag{5.2}$$

and we note that $(\xi_1 + \xi_2)/\sqrt{2}$ is again standard normal. □

Proof of Theorem 2.2. Similarly to the proof of Theorem 2.1, for any sequence $m_n \to \infty$,

$$\big(A^{(1)}_{a_n} - \gamma \log a_n, A^{(2)}_{\mathrm{C}_n} - \gamma \log \mathrm{C}_n | \mathrm{C}_n = m_n\big) \xrightarrow{d} (X_1, X_2), \tag{5.3}$$

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

with X_1, X_2 two independent copies of $X = -\gamma \log(\gamma W_\lambda)$. We rewrite the second coordinate as

$$A^{(2)}_{\mathrm{C}_n} - \gamma \log \mathrm{C}_n = A^{(2)}_{\mathrm{C}_n} - \gamma \log a_n - \gamma \log(\mathrm{C}_n/a_n).$$

Hence, from (5.3) and the limit law for C_n/a_n given in (4.19), we obtain

$$\begin{aligned} W_n - \gamma \log n &= A^{(1)}_{a_n} + A^{(2)}_{\mathrm{C}_n} - 2\gamma \log a_n + o_{\mathbb{P}}(1) \\ &= A^{(1)}_{a_n} - \gamma \log a_n + A^{(2)}_{\mathrm{C}_n} - \gamma \log \mathrm{C}_n + \gamma \log(\mathrm{C}_n/a_n) + o_{\mathbb{P}}(1) \\ &\xrightarrow{d} X_1 + X_2 + \gamma \log(E), \end{aligned}$$

where $E \stackrel{d}{=} \mathrm{Exp}(1)$ is independent of (X_1, X_2). We can reformulate this as

$$(\lambda - 1)W_n - \log n \xrightarrow{d} -\log(\gamma W^{(1)}_\lambda) - \log(\gamma W^{(2)}_\lambda) - M, \tag{5.4}$$

where $M = -\log(E)$ has a Gumbel distribution, *i.e.*, $\mathbb{P}(M \leqslant x) = \Lambda(x) = \exp(-\mathrm{e}^{-x})$. □

Proof of Corollary 2.4. We start with a proof of (3.10), with β replaced by $\beta_n = \lambda_n/(\lambda_n - 1)$. We give the proof for the CTMBP with $\mathrm{Bin}(n, \lambda_n/n)$ offspring distribution (thus avoiding the coupling with Poisson random variables in Proposition 4.3). Since, for $\lambda_n \to \infty$, any two vertices are w.h.p. connected, it is not necessary to condition on the event $\{S_i > 0,\ 1 \leqslant i \leqslant m\}$, and therefore it is straightforward from (3.4) and (3.5) that

$$G_m \stackrel{d}{=} \sum_{i=1}^{m} I_i, \quad \text{where} \quad \mathbb{P}(I_i = 1 | \{D^{(n)}_i\}_{1 \leqslant i \leqslant m}) = D^{(n)}_i / S^{(n)}_i, \tag{5.5}$$

with $D^{(n)}_1, D^{(n)}_2, \ldots, D^{(n)}_m$, independent $\mathrm{Bin}(n, \lambda_n/n)$ random variables and with

$$S^{(n)}_i = \sum_{j=1}^{i} D^{(n)}_j - (i-1).$$

By [16, Chapter 3], the random variables $D^{(n)}_j$ concentrate around λ_n, whereas $S^{(n)}_i$ concentrates around $i(\lambda_n - 1) + 1$. Therefore G_m in (5.5) has the same distributional limit as

$$\sum_{i=1}^{m} \hat{I}^{(n)}_i, \quad \text{where} \quad \mathbb{P}(\hat{I}^{(n)}_i = 1) = \frac{\lambda_n}{i(\lambda_n - 1)} = \frac{\beta_n}{i},$$

with $\hat{I}^{(n)}_1, \hat{I}^{(n)}_2, \ldots, \hat{I}^{(n)}_m$ independent. Applying Lemma 3.4, we conclude that for $m = m_n \to \infty$,

$$\frac{G_m - \beta_n \sum_{i=1}^{m} 1/i}{\sqrt{\sum_{i=1}^{m} (\beta_n/i)(1 - \beta_n/i)}} \xrightarrow{d} Z, \tag{5.6}$$

where Z is standard normal. Note that the denominator of (5.6) can be replaced by $\sqrt{\log m}$, since $\beta_n \to 1$ and hence $\sum_{i=1}^{m} (\beta_n/i)(1 - \beta_n/i)/\log m \to 1$. This proves (3.10), with β replaced by $\beta_n = \lambda_n/(\lambda_n - 1)$. The centering constant $\beta_n \sum_{i=1}^{m} 1/i$ can be replaced by $\beta_n \log m$, but in general not by $\log m$, except when $(\beta_n - 1)\sqrt{\log m_n} \to 0$, or equivalently $\lambda_n = o(\sqrt{\log m_n})$. Since $m_n = \Theta(\log n)$, this condition is equivalent to $\lambda_n = o(\sqrt{\log m_n})$.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

We now turn to the proof of part (b) of Corollary 2.4. Again, we will use that $S_i^{(n)}$ concentrates around $i(\lambda_n - 1) + 1$. Hence, from (3.6),

$$A_m \stackrel{d}{=} \sum_{i=1}^{m} E_i / S_i^{(n)} \approx \frac{1}{\lambda_n - 1} \sum_{i=1}^{m} E_i / i.$$

Now define B_n as a random variable with distribution equal to the maximum of n independent exponentially distributed random variables with rate 1, *i.e.*,

$$\mathbb{P}(B_n \leqslant x) = (1 - \mathrm{e}^{-x})^n.$$

Then E_i/i, $1 \leqslant i \leqslant n$, are equal in distribution to the spacings of these exponentially distributed variables, *i.e.*, $B_n \stackrel{d}{=} \sum_{i=1}^{n} E_i/i$. Consequently, with $\gamma_n = 1/(\lambda_n - 1)$,

$$\begin{aligned} \lim_{n\to\infty} \mathbb{P}(\gamma_n^{-1} A_{a_n}^{(1)} - \log a_n \leqslant x) &= \lim_{n\to\infty} \mathbb{P}(B_{a_n} - \log a_n \leqslant x) \\ &= \lim_{n\to\infty} (1 - \mathrm{e}^{-x - \log a_n})^{a_n} = \Lambda(x), \end{aligned} \tag{5.7}$$

where $\Lambda(x) = \exp(-\mathrm{e}^{-x})$ denotes the distribution function of a Gumbel random variable. Parallel to the proof of Theorem 2.2, we conclude that

$$\left(\gamma_n^{-1} A_{a_n}^{(1)} - \log a_n, \gamma_n^{-1} A_{\mathrm{C}_n}^{(2)} - \log \mathrm{C}_n\right) \stackrel{d}{\to} (M_1, M_2),$$

where M_1 and M_2 are two independent copies of a random variable with distribution function Λ. Finally, since C_n/a_n converges in distribution to an Exp(1) random variable, we conclude from the continuous mapping theorem that (2.5) holds. □

Proof of Theorem 2.6. This proof is a consequence of Theorems 2.1–2.2 and the results on the diameter of $G_n(\lambda/n)$ proved in [10]. We shall sketch the proof. We start by recalling some notation and properties of the ERRG from [10]. Let the 2-*core* of $G_n(\lambda/n)$ be the maximal subgraph of which each vertex has degree at least 2. The study of the diameter in [10] is based on the crucial observation that any longest shortest path in a random graph will consist of three segments: a path from its initial vertex to the 2-core, a path within the 2-core, and a path from the 2-core to its end vertex. While this is used in [10] only for graph distances, this also gives valid lower bounds for FPP paths on the giant component.

Now, when we pick two uniform vertices i and j (as in Theorems 2.1–2.2), then the paths from i and j to the 2-core are (a) unique, (b) of length $O_{\mathbb{P}}(1)$, and (c) the vertices i^* and j^* in the 2-core to which they are connected are asymptotically *uniform elements* of the 2-core. We shall first prove these three statements. By construction, the giant component with the 2-core removed is a collection of trees rooted at the vertices of the 2-core, so that any path linking a vertex that is not in the 2-core to the 2-core is unique, which proves (a).

To see (b), we note that the size of most trees is $O_{\mathbb{P}}(1)$, since the tree rooted at a given vertex converges to a Poisson branching process with parameter λ, conditioned to die out (see, *e.g.*, [17] and [10, Section 6.1]). It is well known that this branching process is the same as a Poisson branching process with *dual parameter* $\mu = \mu_\lambda < 1$ satisfying $\mu \mathrm{e}^{-\mu} = \lambda \mathrm{e}^{-\lambda}$. Since $\mu = \mu_\lambda < 1$, its total population size or total progeny is bounded, which shows that its size is $O_{\mathbb{P}}(1)$. This proves (b).

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

For (c), we note that the asymptotic law of the tree size rooted at a given vertex is the *same* for every vertex, since each of these trees is close to a Poisson branching process with parameter λ, conditioned to die out. Moreover, the sizes of these trees are asymptotically independent, so that they are in fact asymptotically i.i.d. When choosing a vertex uniformly at random from the giant component, the probability of choosing a vertex rooted at a given vertex of the 2-core is proportional to the size of the tree rooted at that vertex. Since these trees asymptotically i.i.d., the probability of choosing a particular vertex in the 2-core is asymptotically equal for all vertices in the 2-core.

As a result, Theorem 2.1 and the convergence $W_n/\log n \xrightarrow{\mathbb{P}} \gamma$ in Theorem 2.2 also hold when picking two *uniform vertices in the* 2*-core*. The weak convergence in Theorem 2.2 does not follow from this argument, due to the fact that the weight of the path from a uniform vertex to the 2-core probably contributes a bounded non-zero random variable to the weight of the shortest-weight path. However, in what follows, we shall only rely on the *first-order asymptotics* in Theorems 2.1–2.2.

Next we investigate the maximal weight and length of shortest-weight paths in the ERRG. In [10, Section 6.1], it is shown that w.h.p. there are two paths of length at least $(1/|\log \mu_\lambda| - \varepsilon)\log n$ in the giant component with the 2-core removed, connecting two vertices i^* and j^* of degree 1 to two vertices U, V in the 2-core. Since the paths connecting i^* and j^* to U and V, respectively, are *unique*, they must also be the shortest-weight paths for FPP on the ERRG between the vertices i^* and U, and j^* and V, respectively. The weights along these paths are thus a sum of $(1/|\log \mu_\lambda| - \varepsilon)\log n$ i.i.d. Exp(1) random variables. As explained above, U, V are asymptotically uniform vertices in the core. Thus, the shortest-weight path from i^* to j^* has at least $\lambda/(\lambda-1)\log n + o_{\mathbb{P}}(\log(n)) + (2/|\log \mu_\lambda| - 2\varepsilon)\log n$ hops, as required, while its weight is at least $1/(\lambda-1)\log n + o_{\mathbb{P}}(\log(n)) + (2/|\log \mu_\lambda| - 2\varepsilon)\log n$. □

Acknowledgements

The work of RvdH was supported in part by the Netherlands Organisation for Scientific Research (NWO). We thank the referee for carefully reading the paper and for many helpful suggestions.

References

[1] Amini, H. (2011) Epidemics and percolation in random networks. PhD thesis, ENS Inria Rocquencourt.
[2] Amini, H., Draief, M. and Lelarge, M. (2010) Flooding in weighted random graphs. Preprint. arxiv.org/abs/1011.5994
[3] Athreya, K. B. and Ney, P. E. (2004) *Branching Processes*, Dover.
[4] Bhamidi, S. (2008) First passage percolation on locally tree like networks I: Dense random graphs. *J. Math. Phys.* **49** 125218.
[5] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2010) First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.* **20** 1907–1965.
[6] Billingsley, P. (1999) *Convergence of Probability Measures*, second edition, Wiley-Interscience.
[7] Bühler, W. J. (1971) Generations and degree of relationship in supercritical Markov branching processes. *Probab. Theory Rel. Fields* **18** 141–152.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press

[8] Chen, L. H. Y. and Shao, Q. M. (2005) Stein's method for normal approximation. In *An Introduction to Stein's Method* (A. D. Barbour and L. H. Y. Chen, eds), IMS Lecture Note Series, National University of Singapore, pp. 1–59.
[9] Dudley, R. M. (2002) *Real Analysis and Probability*, Cambridge University Press.
[10] Fernholz, D. and Ramachandran, V. (2007) The diameter of sparse random graphs. *Random Struct. Alg.* **31** 482–516.
[11] Hammersley, J. M. and Welsh, D. J. A. (1965) First-passage percolation, sub-additive process, stochastic network and generalized renewal theory. In *Bernoulli, 1713; Bayes, 1763; Laplace, 1813: Anniversary Volume*, Springer.
[12] van der Hofstad, R. (2011) Random graphs and complex networks. In preparation. `www.win.tue.nl/~rhofstad/NotesRGCN.pdf`.
[13] van der Hofstad, R., Hooghiemstra, G. and Van Mieghem, P. (2001) First-passage percolation on the random graph. *Probab. Engng Inform. Sci.* **15** 225–237.
[14] Howard, C. D. (2004) Models of first-passage percolation. In *Probability on Discrete Structures*, Springer, pp. 125-173.
[15] Janson, S. (1999) One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combin. Probab. Comput.* **8** 347–361.
[16] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience.
[17] Pittel, B. (1990) On tree census and the giant component in sparse random graphs. *Random Struct. Alg.* **1** 311–342.

https://doi.org/10.1017/S096354831100023X Published online by Cambridge University Press