

# FIRST PASSAGE PERCOLATION ON INHOMOGENEOUS RANDOM GRAPHS

ISTVÁN KOLOSSVÁRY,\* *Budapest University of Technology and Economics*

JÚLIA KOMJÁTHY,\*\* *Budapest University of Technology and Economics  
and Eindhoven University of Technology*

## Abstract

In this paper we investigate first passage percolation on an inhomogeneous random graph model introduced by Bollobás *et al.* (2007). Each vertex in the graph has a type from a type space, and edge probabilities are independent, but depend on the types of the end vertices. Each edge is given an independent exponential weight. We determine the distribution of the weight of the shortest path between uniformly chosen vertices in the giant component and show that the hopcount, i.e. the number of edges on this minimal-weight path, properly normalized, follows a central limit theorem. We handle the cases where the average number of neighbors  $\tilde{\lambda}_n$  of a vertex tends to a finite  $\tilde{\lambda}$  in full generality and consider  $\tilde{\lambda} = \infty$  under mild assumptions. This paper is a generalization of the paper of Bhamidi *et al.* (2011), where first passage percolation is explored on the Erdős–Rényi graphs.

*Keywords:* Inhomogeneous random graph; shortest-weight path; hopcount; first passage percolation; continuous-time multitype branching process

2010 Mathematics Subject Classification: Primary 05C80

Secondary 90B15; 60J85

## 1. Introduction and the main results

First passage percolation (FPP), generally speaking, deals with the asymptotic behavior of first passage times of a percolating fluid in some random environment. This topic has gained much attention due to its application in various fields such as interacting particle systems, statistical physics, epidemic models, and real-world networks, for example.

Particularly, when the random environment is a finite weighted random graph then FPP corresponds to the distance, i.e. the minimal-weight path between two vertices. Also of interest is the number of edges, often referred to as the hopcount, on this path. Without edge weights these quantities coincide. Other natural questions can arise such as how to determine the flooding time of the graph from a fixed vertex  $x$  or its diameter, i.e. the maximum of the shortest paths between  $x$  and all other vertices and the maximum of the flooding times, respectively. This paper investigates FPP on the inhomogeneous random graph (IHRG) model introduced in [12] with independent and identically distributed (i.i.d.) exponential edge weights with rate 1.

The addition of edge weights on the network can be interpreted as the cost of carrying the flow from one node to the other along the edge. Furthermore, edge weights can dramatically

---

Received 10 January 2013; revision received 2 May 2014.

\* Postal address: Budapest University of Technology and Economics, Inter-University Centre for Telecommunications and Informatics, 4028 Debrecen, Kassai út 26, Hungary. Email address: istvanko@math.bme.hu

\*\* Postal address: Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. Email address: j.komjathy@tue.nl

alter the geometry of the graph. For example, consider the complete graph on  $n$  vertices first without edge weights. The hopcount between any two vertices is of course 1. However, by adding i.i.d. exponential  $\text{Exp}(1)$  or uniform  $U(0, 1)$  edge weights, the weight of the shortest-weight path is of order  $\log n/n \ll 1$  and the hopcount is about  $\log n$  [19]. A similar phenomena can be observed for the IHRG, see Section 1.2.

The proofs usually rely on results from branching processes. The use of exponential weights imply that the exploration processes of the graph are Markovian. Only recently was FPP studied on random graphs with general continuous edge weights [9]. We suspect that similar results also hold in the IHRG for general weights. Other related results are discussed in Section 1.3. We begin by introducing the IHRG model in Section 1.1. The main results are stated in Section 1.2 and then the main ideas of the proofs are sketched in Section 1.4. In Section 2 we examine the multitype branching process and in Section 3 we relate the exploration process of the neighborhood of a vertex in the IHRG model to the branching process of Section 2. In section 4 we make the intuitive picture of colliding flows mathematically precise. Finally, in Section 5 we present the proofs of the main results.

**1.1. The model**

We consider a general IHRG model introduced by Bollobás *et al.* [12]. We briefly describe the model  $G(n, \kappa)$  on  $n$  vertices and kernel  $\kappa$  in the general setting and then turn to an important special case.

Each vertex of the graph will be assigned a type from a separable metric space  $\mathcal{X}$  which is equipped with a Borel probability measure  $\mu$ . For each  $n$  we have a deterministic or random sample of  $n$  points  $\mathbf{x}_n = (x_1, \dots, x_n)$  from  $\mathcal{X}$ . We assume that the empirical distribution converges in probability to  $\mu$  as  $n \rightarrow \infty$ , i.e.

$$v_n(S) := \frac{\#\{i : x_i \in S\}}{n} \xrightarrow{\mathbb{P}} \mu(S) \tag{1.1}$$

for every  $\mu$ -continuity set  $S \subset \mathcal{X}$  ( $S$  is measurable and there is no mass on the boundary of  $S$ ). The pair  $(\mathcal{X}, \mu)$  is called a ground space, and for a sequence  $(\mathbf{x}_n)_{n \geq 1}$  satisfying (1.1) we say that the triplet  $(\mathcal{X}, \mu, (\mathbf{x}_n)_{n \geq 1})$  defines a vertex space  $\nu$ . Furthermore, a kernel  $\kappa$  on a ground space is a symmetric nonnegative measurable function on  $\mathcal{X} \times \mathcal{X}$ . The natural interpretation of  $\kappa$  is that it measures the density of edges.

**Definition 1.1.** Given a vertex space  $\nu$  and kernel  $\kappa$ , for each pair  $(i, j)$ ,  $i \neq j$ ,  $i, j \in \{1, \dots, n\} := [n]$ , the edge  $\{i, j\}$  is present in the IHRG  $G(n, \kappa)$  with probability

$$p_{ij} := \min \left\{ \frac{\kappa(x_i, x_j)}{n}, 1 \right\}$$

independently of other edges. Conditioned on the edges of  $G(n, \kappa)$ , we assign each edge an independent  $\text{Exp}(1)$  edge weight.

**Definition 1.2.** (*Quasi-irreducibility.*) A kernel  $\kappa$  on a ground space  $(\mathcal{X}, \mu)$  is *irreducible* if  $A \subseteq \mathcal{X}$  and  $\kappa = 0$  almost everywhere (a.e.) on  $A \times (\mathcal{X} \setminus A)$  imply that  $\mu(A) = 0$  or  $\mu(\mathcal{X} \setminus A) = 0$ . A kernel  $\kappa$  is *quasi-irreducible* if there is a  $\mu$ -continuity set  $\mathcal{X}' \subseteq \mathcal{X}$  with  $\mu(\mathcal{X}') > 0$  such that the restriction of  $\kappa$  to  $\mathcal{X}' \times \mathcal{X}'$  is irreducible and  $\kappa(x, y) = 0$  if  $x \notin \mathcal{X}'$  or  $y \notin \mathcal{X}'$ .

Irreducibility of kernels ensures the emergence of a single giant component in the supercritical regime. The next definition ensures that the graph has the ‘right’ number of edges  $e(G)$  in expectation  $\mathbb{E}$ .

**Definition 1.3.** (*Graphical kernels.*) A sequence  $(\kappa_n)_{n \geq 1}$  of kernels on a vertex space  $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$  is *graphical* with limit  $\kappa$  if, for almost every  $(y, z) \in \mathcal{S}^2$ ,  $y_n \rightarrow y$  and  $z_n \rightarrow z$  imply that  $\kappa_n(y_n, z_n) \rightarrow \kappa(y, z)$ ,  $\kappa \in L^1(\mathcal{S} \times \mathcal{S}, \mu \times \mu)$  and is continuous a.e. Furthermore,

$$\frac{1}{n} \mathbb{E}\{e(G(n, \kappa_n))\} \rightarrow \frac{1}{2} \iint_{\mathcal{S}^2} \kappa(x, y) \, d\mu(x) \, d\mu(y). \tag{1.2}$$

When  $\kappa_n \equiv \kappa$  we simply say that  $\kappa$  is graphical.

For example, condition (1.2) holds whenever  $\kappa$  is bounded and  $\nu$  is a vertex space. For more general conditions, see [12, Lemma 8.1].

*Finite-type spaces.* In the proof of the main theorem below, we approximate general kernels by an appropriate sequence of step functions, referred to as *regular finitary kernels*. We call a kernel regular finitary if the type space  $\mathcal{S}$  has a finite partition into ( $\mu$ -continuity) sets  $S_1, \dots, S_r$  such that  $\kappa$  is constant on each  $S_i \times S_j$  for all  $1 \leq i, j \leq r$ . By identifying each  $S_i$  with a single point  $i$  with weight  $\mu_i = \mu(S_i)$ , a random graph  $G(n, \kappa)$  generated by a regular finitary kernel has the same distribution as a *finite-type kernel*. That is,  $\mathcal{S} = [r]$ ,  $\kappa = (\kappa(s, t))_{s, t=1}^r$  is a symmetric  $r \times r$  matrix with nonnegative entries. Denoting the number of type  $t$  vertices by  $n_t$ , condition (1.1) becomes

$$\frac{n_t}{n} \xrightarrow{\mathbb{P}} \mu_t \quad \text{holds for every } t \in \mathcal{S}. \tag{1.3}$$

All finite-type kernels are automatically graphical. The Erdős–Rényi random graph (ERRG) is a special case of a finite-type graph when  $r = 1$  and  $\kappa = c$ . In the finite-type case we also use the notation for  $s, t \in \mathcal{S}$ ,

$$\lambda_{st} := \kappa(s, t)\mu_t,$$

where  $\lambda_{st}$  gives the limiting average number of type  $t$  neighbors of a type  $s$  vertex, since this quantity is binomially distributed with parameters  $n_t - \mathbf{1}_{\{s=t\}}$  and  $\kappa(s, t)/n$ , where  $\mathbf{1}_{\{\cdot\}}$  is the indicator function.

*Stationary distributions and main assumption.* Let us define the integral operator  $T_\kappa$  on  $(\mathcal{S}, \mu)$  with kernel  $\kappa$  by

$$(T_\kappa f)(s) := \int_{\mathcal{S}} \kappa(s, t) f(t) \mu(dt).$$

Since  $T_\kappa f \geq 0$  if  $f \geq 0$  (for further properties, see [12]), by the Perron–Frobenius theory there is a single real main eigenvalue of  $T_\kappa$ , that we denote by  $\tilde{\lambda} + 1$ , and, assuming quasi-irreducibility of  $\kappa$ , let us denote the unique, nontrivial left eigenfunction by  $\pi(s)$  satisfying  $\int_{\mathcal{S}} \pi(dt) = 1$ , so

$$\int_{\mathcal{S}} \pi(ds) (\kappa(s, t) \mu(t)) = (\tilde{\lambda} + 1) \pi(t).$$

We call  $\pi$  the stationary type-distribution.

**Assumption 1.1.** *Throughout this paper we assume that  $\kappa$  is quasi-irreducible, the main eigenvalue of  $T_\kappa$ ,  $\tilde{\lambda} + 1 > 1$ , and the stationary type-distribution  $\pi$  satisfies*

$$\int_{\mathcal{S}} \frac{\pi(t)}{\mu(t)} \pi(dt) < \infty. \tag{1.4}$$

We also require that the second largest eigenvalue  $\lambda_2$  of  $T_\kappa - I$  in modulus satisfies  $2 \operatorname{Re}(\lambda_2) < \tilde{\lambda}$ . Equivalently, in the finite-type case we assume that the matrix  $A$  is irreducible, has a positive eigenvalue  $\tilde{\lambda} > 0$ , and  $2 \operatorname{Re}(\lambda_2(A)) < \tilde{\lambda}$  where

$$A_{st} = \lambda_{st} - \mathbf{1}_{\{s=t\}} \quad \text{for } 1 \leq s, t \leq r. \tag{1.5}$$

Note that the stationary distribution in the finite-type case satisfies  $\pi A = \tilde{\lambda}\pi, \pi(1, \dots, 1)^T = 1$ . The  $\tilde{\lambda} > 0$  condition is necessary and sufficient for a giant component to emerge in  $G(n, \kappa)$ , which is in direct accordance with whether  $\|T_\kappa\|_{2,2} > 1$  (see [12, Theorem 3.1]).

We call a kernel *homogeneous* (see [12, Example 4.6]), if it satisfies

$$\int_{\mathcal{S}} \kappa(s, t) \mu(dt) = \tilde{\lambda} + 1 \quad \text{for a.e. } s \in \mathcal{S}. \tag{1.6}$$

Roughly speaking, asymptotically the average degree of a vertex in  $G(n, \kappa)$  generated by a homogeneous kernel is independent of its type. In this case the symmetry of  $\kappa$  implies that  $\pi = \mu$ , thus, the quantity in (1.4) equals 1.

**1.2. Main results**

We investigate the weight and the number of edges on the shortest-weight path between two uniformly selected vertices  $U$  and  $V$ . Let  $\Gamma_{UV}$  denote the set of all  $\pi$  paths in  $G(n, \kappa)$  between  $U$  and  $V$ . Denote the weight of the shortest-weight path by

$$\mathcal{P}_n(U, V) = \min_{\pi \in \Gamma_{UV}} \sum_{e \in \pi} X_e,$$

where  $X_e$  is the exponential edge weight attached to edge  $e$  in the construction of  $G(n, \kappa)$ . Let  $\mathcal{H}_n(U, V)$  denote the number of edges or *hopcount* of this path. If the two vertices are in different components of the graph then  $\mathcal{P}_n, \mathcal{H}_n := \infty$ . In the following two theorems we describe the asymptotic behavior of these two quantities when the average degree remains finite (sparse regime) or goes to  $\infty$  (dense regime), respectively.

**Theorem 1.1.** (Sparse setting.) *Let  $(\mathcal{S}, \mu)$  be an arbitrary ground space and  $(\kappa_n)_{n \geq 1}$  be a sequence of uniformly continuous, graphical kernels on  $(\mathcal{S}, \mu)$  that satisfy Assumption 1.1 and  $\sup_{n,x,y} \kappa_n(x, y) < \infty$  with  $\lim_{n \rightarrow \infty} \tilde{\lambda}_n = \tilde{\lambda} < \infty$ . Conditioned on  $(U, V)$  being connected, it follows that*

$$\left( \frac{\mathcal{H}_n(U, V) - ((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}}, \mathcal{P}_n(U, V) - \frac{1}{\tilde{\lambda}} \log n \right) \xrightarrow{D} (Z, L),$$

where  $Z$  is a standard normal variable. Furthermore,  $L$  is a nondegenerate real-valued random variable whose distribution can be determined from the multitype branching process that arises when exploring a component of  $G(n, \kappa)$ , see Remark 2.2.

**Theorem 1.2.** (Dense setting.) (i) *Under the assumptions of Theorem 1.1 and (1.6) with  $\lim_{n \rightarrow \infty} \tilde{\lambda}_n = \infty$ , or*

(ii) *for a  $\kappa$  satisfying the assumptions of Theorem 1.1 and  $(\kappa_n)_{n \geq 1} := (\tilde{\lambda}_n + 1)\kappa$  with  $\lim_{n \rightarrow \infty} \tilde{\lambda}_n = \infty$ ,*

$$\left( \frac{\mathcal{H}_n(U, V) - ((\tilde{\lambda}_n + 1)/\tilde{\lambda}_n) \log n}{\sqrt{\log n}}, \tilde{\lambda}_n \mathcal{P}_n(U, V) - \log n \right) \xrightarrow{D} (Z, \tilde{L}),$$

where  $Z$  is a standard normal variable,  $\tilde{L}$  is equal in distribution to the sum of independent random variables  $\Lambda_1 + \Lambda_2 - \Lambda_3 - c$  with  $\Lambda_i$  i.i.d. standard Gumbel random variables,  $c = 0$  for case (i), and  $c = \log \int_{\mathcal{S}} \pi(s)/\mu(s)\pi(ds)$  for case (ii). Furthermore, we can substitute  $(\tilde{\lambda}_n + 1)/\tilde{\lambda}_n$  in the centering of the hopcount by 1 if and only if  $\tilde{\lambda}_n/\sqrt{\log n} \rightarrow \infty$ .

**Remark 1.1.** Comparing these results with the graph distance  $\mathcal{D}_n(U, V)$  in the unweighted IHRG, we have

- *Sparse setting.* In this case  $\mathcal{D}_n \simeq \log n/\log(\tilde{\lambda} + 1)$  (see [12, Theorem 3.14]), thus,  $\mathcal{P}_n < \mathcal{D}_n < \mathcal{H}_n$ . This shows that the structure of the graph changes with the addition of edge weights. The shortest-weight path is no longer the path with the least number of edges: among the paths with slightly more edges some paths have less weight than the path with the least number of edges.
- *Dense setting.* In this case with high probability (w.h.p.) the giant component contains  $n(1 - o(1))$  vertices (see [12]). The change in the graph structure is even more significant. Without edge weights the graph is ultra-small, i.e.  $\mathcal{D}_n(U, V) = o(\log n)$ . However, adding the edge weights,  $\mathcal{H}_n$  and  $\mathcal{P}_n$  remain  $O(\log n)$ .

### 1.3. Related work

First passage percolation has been investigated on various models such as the integer lattice, the mean-field model, configuration model, and graphs with i.i.d. vertex degrees both without edge weights (see, e.g. [6], [11], [12], [13], [16], [17], [22], [24], and [25]) and with exponential weights (see, e.g. [1], [5], [7], [8], [18], and [19]). The list of results are far from complete, we only attempt to discuss the results directly related to this paper.

The IHRG model was extensively investigated by Bollobás *et al.* [12] from many different aspects, including typical distances without edge weights. We have already shown what effect the addition of edge weights has on the structure of the graph in Section 1.2. Many other models are closely related to the IHRG model, for full details, see [12, Sections 4 and 16].

The classical supercritical ERRG  $G(n, c/n)$  ( $c > 1$ ) is the special case when  $|\mathcal{S}| = 1$ . Our results generalize the FPP results of Bhamidi *et al.* [8] on ERRGs with i.i.d. exponential edge weights. Finite-type graphs were previously studied by Soderberg [23].

A class of IHRGs is the *rank-1* class. The kernel  $\kappa$  has the special form  $\kappa(x, y) = \phi(x)\phi(y)$ , where the positive function  $\phi$  on  $\mathcal{S}$  can be interpreted as the ‘activity’ of a type  $x$  vertex. In the Chung–Lu model [15], [16], each vertex  $i$  is given a positive weight  $w_i$  and the edge probabilities  $p_{ij}$  are given by  $p_{ij} := w_i w_j / \ell_n$ , where  $\ell_n = \sum_{i=1}^n w_i$ . Norros and Reittu [22] presented results on the existence and size of a giant component with random  $w_i$ . For deterministic  $w_i$ , Chung and Lu [15], [16] showed that, under certain conditions, (without edge weights) the typical distance between two vertices is  $\log n / \log \bar{d}$ , where  $\bar{d} = \sum w_i^2 / \ell_n$ .

An asymptotically equivalent model is the generalized random graph introduced by Britton *et al.* [13] with edge probabilities  $p_{ij} = w_i w_j / (\ell_n + w_i w_j)$ ,  $\ell_n = \sum_k w_k$ . Conditioned on the vertex degrees, the resulting graph is uniformly distributed over all graphs with the given degree sequence. Hence, the result of Bhamidi *et al.* [9] concerning FPP for the configuration model with *general* continuous edge weights, implies these results by noting that the configuration model conditioned on it being simple is also uniformly distributed over all graphs with the given degree sequence.

First passage percolation results have a deep connection with epidemics with exponential spreading times on supercritical random graphs. Let  $I(t) := \{v \in [n], \mathcal{P}_n(x, v) \leq t\}$  denote the number of points that have distance less than  $t$  from vertex  $x$ ; equivalently, the number of

infected vertices up to time  $t$  in the epidemic started from  $x$ . Then the first and second moment method for  $I(t)$  combined with Theorem 1.1 yields, for  $t = \tilde{\lambda}^{-1} \log n + u$ , where  $W^1$  and  $W^2$  are i.i.d. copies of the limit variable below in Proposition 2.2,

$$I\left(u + \frac{\log n}{\tilde{\lambda}}\right) = \mathbb{P}\left\{\mathcal{P}_n(U, V) - \frac{\log n}{\tilde{\lambda}} \leq u\right\} \stackrel{D}{\rightarrow} \mathbb{P}\{-\tilde{\lambda}^{-1}(\log(W^1 W^2) + \Lambda) + c \leq u\},$$

i.e. a deterministic function of  $u$ , often called the *epidemic curve*. We omit the details and refer the reader to [4] and [10].

**1.4. Sketch of the proofs**

To find the shortest-weight path between two vertices  $x$  and  $y$ , we consider the edge weights as lengths and let fluid percolate through the edges of the graph simultaneously at a constant rate, starting from both vertices. We call this procedure the *exploration processes* of the neighborhoods of vertices  $x$  and  $y$ . Intuitively, when the flows collide, the shortest-weight path has been found: up to any time  $\tau$  each exploration process contains the vertices with distance at most  $\tau$  from  $x$  or  $y$ . It is standard to relate an exploration process to a branching process (BP). In our context this BP is a continuous-time multitype BP, which we introduce in Section 2.

We can apply BP theory once we embed the BP into the IHRG: in Section 3 we assign labels to the vertices of the BP according to their type in a way that we can rule out loop and multiple edges in the embedding in the IHRG. We thin the BP to deal with cycles.

The problem with the scenario of colliding flows in continuous time is that we can keep track of the flows only in discrete steps as new vertices are explored. So we first start the exploration process from  $x$  until a given number of vertices are explored and stop. Then we start the flow from  $y$  and wait until a vertex  $v_y$  which shares an edge with a wet vertex  $v_x$  in the flow of  $x$  becomes wet. We have then just found a path between  $x$  and  $y$  through  $v_x$  and  $v_y$ , which, however, might not be the shortest one. Thus, we have to minimize over all such possible choices of  $v_x$  and  $v_y$ . The rigorous treatment of the connection time is contained in Section 4. Finally, in Section 5 we prove our main results, Theorems 1.1 and 1.2.

**2. Multitype branching processes**

In this section we collect the required properties of the branching process that arises when exploring a component of  $G(n, \kappa)$ .

Let us define a multitype continuous-time branching process (CTBP) with type space  $\mathcal{S}$ , where a particle of type  $x \in \mathcal{S}$ , when it splits (dies), gives birth to a set of offsprings distributed as a Poisson process on  $\mathcal{S}$  with intensity measure  $\kappa(x, y) d\mu(y)$ : the number of children with types in a subset  $S \subset \mathcal{S}$  has a Poisson distribution with mean  $\int_S \kappa(x, y) d\mu(y)$ . Each offspring lives for an  $\text{Exp}(1)$  amount of time independently of everything else. We denote this branching process with root of type  $s$  up to time  $t$  by  $\Psi_\kappa^s(t)$ . For a set  $E \subseteq \mathcal{S}$ ,  $\mathcal{D}_m^E$  and  $\mathcal{A}_m^E$  stand for the set of dead and alive particles after the  $m$ th split whose type belongs to the set  $E$ . Let  $N_m^E := |\mathcal{D}_m^E|$  and  $S_m^E := |\mathcal{A}_m^E|$ , and we simply write  $\mathcal{D}_m := \mathcal{D}_m^\mathcal{S}$ ,  $\mathcal{A}_m := \mathcal{A}_m^\mathcal{S}$ , and  $S_m := |\mathcal{A}_m^\mathcal{S}|$ ,  $N_m := |\mathcal{D}_m^\mathcal{S}|$ . Let us also write  $\tau_m$  for the time of the  $m$ th split. With a slight abuse of notation we write  $A_t^E$  for the size of alive individuals in the set  $E$  at real time  $t \in \mathbb{R}$ .

**Remark 2.1.** The BP  $\Psi_\kappa^s$  survives with positive probability if and only if  $\|T_\kappa\|_{2,2} > 1$  (see [12, Theorem 6.1]). The survival probability  $\rho(s)$  is the maximal fixed point of the nonlinear operator  $\Phi_\kappa$ , defined by  $\Phi_\kappa f := 1 - \exp\{-T_\kappa f\}$  for  $f \geq 0$ . For homogeneous kernels  $\rho(s) \equiv \rho$  is the maximal solution of the equation  $\rho = 1 - \exp\{-\tilde{\lambda} + 1\}\rho$ .

Throughout the proofs, we mainly work with the finite-type  $\mathcal{J} = \{1, \dots, r\}$  case given with the mean-offspring matrix  $A$  in (1.5). Recall that  $\pi A = \lambda\pi$ . We need the following limit theorems for  $S_m^s$  and  $N_m^s$ , the number of alive and dead individuals of a given type  $s \in \mathcal{J}$ .

**Proposition 2.1.** (BP-asymptotics, see Athreya and Ney [3].) *For a finite-type continuous-time branching process as defined above, on the set of nonextinction as  $m \rightarrow \infty$ ,*

$$\frac{(S_m^1, \dots, S_m^r)}{\tilde{\lambda}m} \rightarrow \pi \quad \text{and} \quad \frac{(N_m^1, \dots, N_m^r)}{m} \rightarrow \pi \quad \text{almost surely (a.s.).} \tag{2.1}$$

This result implies that

$$\frac{1}{m} S_m \rightarrow \tilde{\lambda} \quad \text{and} \quad \frac{S_m^s}{S_m} \rightarrow \pi_s \quad \text{a.s.} \tag{2.2}$$

We provide the continuous-time version of these results in the following proposition.

**Proposition 2.2.** (BP-asymptotics [3].) *For a finite-type CTBP as defined above, on the set of nonextinction as  $t \rightarrow \infty$ ,*

$$e^{-\tilde{\lambda}t} (A_t^1, \dots, A_t^r) \rightarrow \pi W \quad \text{a.s.,}$$

where  $W$  is an absolutely continuous random variable with  $W > 0$  on nonextinction, and  $\mathbb{P}_x\{W > 0\} = \rho(x)$ .

In the following result we obtain a nontrivial consequence of Proposition 2.2 when setting  $t := \tau_m$ .

**Proposition 2.3.** (Split-time asymptotics [3].) *On the set of nonextinction*

$$\tau_m - \frac{1}{\tilde{\lambda}} \log m \rightarrow -\frac{1}{\tilde{\lambda}} \log\left(\frac{1}{\tilde{\lambda}} W\right) \quad \text{a.s.}$$

**Remark 2.2.** The distribution of  $W$  depends on the type of the root, say  $s$ . Since in our process the root dies immediately,  $A^s(t)$  (the number of alive particles at time  $t$ ) satisfies the stochastic equation  $A^s(t) = \int_{\mathcal{J}} A^u(t - E_i) \mathbf{1}_{\{E_i < t\}} + \mathbf{1}_{\{E_i > t\}} \xi_s(du)$ , with  $E_i$  being i.i.d. exponentials, and  $\xi_s(du) \sim \text{Poi}(\kappa(s, u)\mu(du))$ . Multiplying both sides by  $e^{-\tilde{\lambda}t}$  we obtain the distributional identity for  $W^s$ :  $W^s \stackrel{D}{=} \int_{\mathcal{J}} e^{-\tilde{\lambda}E_i} W^u \xi_s(du)$ . From here, by the Poissonian property we have  $\mathbb{E}\{\exp\{\int_{\mathcal{J}} f(u)\xi_s(du)\}\} = \exp\{\int_{\mathcal{J}} (e^{f(u)} - 1)\kappa(s, u)\mu(du)\}$ . Hence, first conditioning on all the Poisson points and then using this identity, it follows that the moment generating function  $M_{W^s}(\theta)$  of  $W^s$  satisfies the functional equation

$$M_{W^s}(\theta) = \exp\left\{ \int_{\mathcal{J}} \int_0^\infty (M_{W^u}(\theta e^{-\tilde{\lambda}x}) - 1)e^{-x} dx \kappa(s, u)\mu(du) \right\}.$$

The generation of a particle in the branching process corresponds to the hopcount of the vertex in the IHRG. We use the following theorem which is an immediate consequence of Kharlamov [21, Theorem 2].

**Theorem 2.1.** (Generation of a uniformly picked particle in a given type set [21].) *Let  $G_m^E$  denote the generation of a uniformly picked individual from  $\mathcal{A}^E(m)$ ,  $E \subset \mathcal{J}$ . Then, conditioned on survival, for  $m \rightarrow \infty$ , with  $Z$  a standard normal variable,*

$$\frac{G_m^E - ((\tilde{\lambda} + 1)/\tilde{\lambda}) \log m}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log m}} \xrightarrow{D} Z.$$

Note that the parameters are independent of the type. Kharlamov [21] proved the result in the continuous-time case (i.e. for  $G_t^E$ ), for an arbitrary set of types  $E \in \mathcal{S}$  under the conditions that the type distribution is tending to  $\pi$  exponentially fast, and that the expectation and variance of the generation of type  $s$  individuals after unit time for any type  $s \in \mathcal{S}$  is uniformly bounded in  $s$ . These conditions clearly hold if the lifetime is exponential and the total number of children is  $\text{Poi}(\int_{\mathcal{S}} \kappa(s, t)\mu(dt))$ . The result for discrete time (at  $\tau_m$ ), follows by aperiodicity of the types along generations.

To bound the error terms of coupling of the BP to the IHRG, we need the following lemma. Note that this is slightly stronger than what follows from (2.1).

**Lemma 2.1.** (Law of iterated logarithm for  $S_j$ .) *Suppose that Assumption 1.1 holds. Then, conditioned on survival, for every  $\varepsilon > 0$ , there exists a constant  $C > 0$  such that*

$$|S_j - \tilde{\lambda}j| \leq C(j)^{1/2+\varepsilon} \quad \text{for all } j \geq \log \log m$$

holds a.s.

*Proof (sketch).* The proof follows from Asmussen [2, Theorem 2]. There, Asmussen derives a law of iterated logarithm for (finite-type) multitype Markov branching processes: he shows the law of the iterated algorithm for scalar products of the vector of alive individuals with vectors that are perpendicular to the stationary distribution  $\pi$  for the process at time  $t$ . The conditions in [2, Theorem 2] correspond to the assumption that the real part of the second largest eigenvalue of  $A$  (corresponding to  $T_\kappa - I$ ) is at most  $\tilde{\lambda}/2$ . The normalization is different for equality and strict inequality, but in both case less than  $t^{1/2+\varepsilon}$  for  $\varepsilon > 0$ . Applying this theorem at the split times  $\tau_m$  and using the fact that  $\tau_m - \tilde{\lambda}^{-1} \log m$  converges, the statement of the lemma follows by using the projections of the vector  $S_j$  to the subspace formed by the eigenvectors that are perpendicular to  $\pi$ . This can be achieved in the exact same manner as in Corollaries 3.16 and 3.17 of Janson [20, Theorem 3.15]. Continuous-type spaces can then be handled by approximation methods.

### 3. Embedding the BP into the IHRG

In this section we relate the exploration process of the neighborhood of a vertex in the IHRG model  $G(n, \kappa)$  of Section 1 to the branching process  $\Psi_\kappa$  of Section 2.

*Coupling of BPs.* Denote the following continuous-time multitype BP by  $\Psi_\kappa^{\text{bin}}$ : the children of the root are born immediately, and each particle lives for an i.i.d.  $\text{Exp}(1)$  random time and upon death (split) gives birth to its children. The offspring distribution  $D^{\text{bin}}$  conditioned on the event that the particle to split is of type  $s$  is

$$D^{\text{bin},(s)} = \eta_{s1} + \eta_{s2} + \dots + \eta_{sr},$$

where  $\eta_{st} \stackrel{D}{=} \text{bin}(n_t - \mathbf{1}_{\{s=t\}}, \kappa(s, t)/n)$  is the distribution of the number of type  $t$  neighbors of a type  $s$  vertex. By the standard coupling of binomial and Poisson random variables we obtain the following lemma.

**Lemma 3.1.** (Coupling error to Poisson offsprings.) *The multitype branching processes  $\Psi_\kappa^{\text{bin}}$  and  $\Psi_\kappa$  (defined in Section 2) can be coupled up to the  $m$ th split with error*

$$\mathbb{P}\{\text{there exists } j \leq m, D_j^{\text{bin}} \neq D_j^{\text{Poi}}\} \leq \frac{m}{n}(\tilde{\lambda} + 1) \max \kappa(1 + o(1)). \quad (3.1)$$

*The coupling carries through for  $\tau_m$  and  $G_m^E$  for any set  $E \subseteq \mathcal{S}$  with the same error bound as in (3.1).*

*Proof.* By the standard coupling of binomial and Poisson random variables, we couple  $\eta_{st}$  to  $\xi_{st} \sim \text{Poi}(\lambda_{st})$  with error probability  $n_t(\kappa(s, t)^2/n^2) = \lambda_{st}\kappa(s, t)/n(1 + o(1))$ . Under the assumption that  $\max \kappa < \infty$ , the coupling error up to split  $m$  becomes

$$\sum_s N_m^s \sum_t \frac{\lambda_{st}\kappa(s, t)}{n} \leq \frac{m}{n} \max \kappa \sum_{s,t} \pi_s \lambda_{st} = \frac{m}{n} \max \kappa \sum_t (\tilde{\lambda} + 1)\pi_t(1 + o(1)),$$

where we used (2.1) in the inequality and  $\pi A = \tilde{\lambda}\pi$  in the equality. Clearly the error probabilities  $\mathbb{P}\{G_m^{\text{bin},E} \neq G_m^{\text{Poi},E}\}$  and  $\mathbb{P}\{\tau_m^{\text{bin}} \neq \tau_m^{\text{Poi}}\}$  can be bounded from above by the probability that the coupling fails between the two BPs up to the  $m$ th step.

Now, by using a two-step coupling, we can apply the following labeling procedure to  $\Psi_\kappa$  in order to couple it to the neighborhood of a vertex  $v$  in the IHRG.

*Labeling procedure.* Fix  $n \geq 1$  and denote the set of labels (vertices of the graph) by  $[n] = \{1, 2, \dots, n\}$ . The labels are distinguished according to the types, so  $[n]$  is the disjoint union of the sets of labels  $[n]^{(1)}, \dots, [n]^{(r)}$ , where there are  $n_t$  different labels in  $[n]^{(t)}$ . An individual of type  $s$  in  $\Psi_\kappa$  will be assigned a label from  $[n]^{(s)}$ . Given that the root is of type  $s$ , we choose its label  $i_0$  uniformly from  $[n]^{(s)}$ .

At any split time  $\tau_m$ , assuming that the vertex to split has label  $i_m$  and type  $s$ , we assign to its  $\eta_{st}, t \in [r]$  children different labels drawn without replacement uniformly at random for  $t \neq s$  from  $[n]^{(t)}$ ; for  $t = s$  from  $[n]^{(s)} \setminus \{i_m\}$ . As a result, siblings have different labels from each other and from their parents: this corresponds to excluding multiple and loop edges in  $G(n, \kappa)$ .

It can still happen that vertices of the same type that are not siblings are assigned the same label. Such multiple labels correspond to cycles in  $G(n, \kappa)$ . We keep only the shortest path by thinning the BP as follows.

*Thinning the BP.* We obtained a cycle in the exploration of the IHRG when a label in the labeled CTBP reappears among the labels of the previously split vertices, i.e.

$$i_k \in \mathcal{D}(k - 1) := \{i_0, i_1, \dots, i_{k-1}\}$$

for some  $k$ . By deleting  $i_k$  and the whole subtree starting from it in  $\Psi_\kappa$  we keep only the shortest-weight paths between pairs of vertices. We call the label  $i_k$  and its subtree thinned. We refer to the remaining process as  $\text{th}\Psi_\kappa$ .

The total number of labels is finite ( $= n$ ), so a.s. at some random time  $\text{th}\Psi_\kappa$  dies out. At this time the minimal-weight spanning tree of a component of  $G(n, \kappa)$  is found. It is clear that for each  $t \geq 0$ , the set of labels reached in  $\text{th}\Psi_\kappa(t)$  and the set of vertices reached by time  $t$  in  $G(n, \kappa)$  are equal in distribution. So we arrive at the following lemma.

**Lemma 3.2.** (FPP on  $G(n, \kappa)$  is thinned labeled CTBP.) *For any fixed  $n \geq 1$ , consider  $\text{th}\Psi_\kappa$  and  $G(n, \kappa)$  as defined above. Then, for any  $i_0 \in [n]$ , the weight  $\mathcal{P}_n(i_0, j)$  and the hopcount  $\mathcal{H}_n(i_0, j)$  in  $G(n, \kappa)$  is equal in distribution to the split time and generation of the individual with label  $j$  in  $\text{th}\Psi_\kappa$  with root  $i_0$ .*

**Remark 3.1.** Lemma 3.2 holds for i.i.d. edge weights with arbitrary continuous distribution supported on  $(0, \infty)$ , not only for  $\text{Exp}(1)$  weights.

### 3.1. Analysis of thinning

Here we bound the difference between  $\text{th}\Psi_\kappa$  and  $\Psi_\kappa$ . More precisely, the connection will happen at some alive vertex, so we need to give a bound of the proportion of thinned alive vertices among each type in  $\Psi_\kappa$ .

To keep track of the exploration processes at each split time  $\tau_k$ , we introduce shortest-weight trees  $\text{SWT}_k$ . With a slight misuse of notation, let  $\mathcal{D}(k)$  and  $\mathcal{A}(k)$  stand for the collection of dead and alive labels of the vertices that the flow reaches up to and including time  $\tau_k$  (as a list, with multiple occurrences). Clearly  $|\mathcal{A}(k)| = S_k$ , the number of alive vertices and  $|\mathcal{D}(k)| = k + 1$ . Let  $\text{SWT}_0 = (\{i_0\}, \tau_0 = 0)$ , and define

$$\text{SWT}_k = (\mathcal{D}(k), \mathcal{A}(k), \{\tau_0, \tau_1, \dots, \tau_k\}), \quad k \geq 1. \tag{3.2}$$

The labeled CTBP  $\Psi_\kappa(t)$  can be uniquely reconstructed from the sequence  $(\text{SWT}_k)_{k=1}^\infty$ . Note that  $\text{SWT}_k$  contains all the labels in  $\Psi_\kappa$ , also the thinned labels and possibly some multiple labels among alive vertices. Next, we bound the number of such labels in order to guarantee that the shortest-weight path w.h.p. does not fall in a thinned subtree.

**Lemma 3.3.** (Expected number of thinned alive labels.) *Let  $\mathcal{A}^t(k)$  and  $\text{th}^C \mathcal{A}^t(k)$  denote the number of alive and thinned alive labels of type  $t \in \mathcal{S}$  at split time  $\tau_k$  in the labeled CTBP. Then for  $k = n^\beta$ , under Assumption 1.1,*

$$\mathbb{E} \left\{ \frac{\text{th}^C \mathcal{A}^t(n^\beta)}{\mathcal{A}^t(n^\beta)} \right\} = o(1) \quad \text{whenever } \beta < \frac{3}{4}. \tag{3.3}$$

*Proof.* We calculate  $\text{th}^C \mathcal{A}^t(k)$  by checking to see if the particle that splits at time  $\tau_j$  is thinned, then see how many type  $t$  alive descendants it has in its subtree at the  $k$ th split. Denoting these descendants by  $\mathcal{A}_j^{\{i_j\} \rightarrow t}(k)$  and the type of  $i_j$  by  $t(i_j)$ , we obtain

$$\text{th}^C \mathcal{A}^t(k) = \sum_{j=1}^k \sum_{s \in \mathcal{S}} |\mathcal{A}_j^{\{i_j\} \rightarrow t}(k)| \mathbf{1}_{\{i_j \text{ is thinned} \mid t(i_j)=s\}} \mathbf{1}_{\{t(i_j)=s\}}. \tag{3.4}$$

Let  $\mathcal{A}^u(j_+)$  denote the type  $u$  alive labels immediately after  $\tau_j$  and

$$\mathcal{A}_j^{u \rightarrow t}(k) := \{v \in \mathcal{A}^t(k) : v \text{ is a descendant of } w \text{ with } w \in \mathcal{A}^u(j_+)\}.$$

In Figure 1 we give an example to illustrate the notation.

Note that by symmetry, immediately following  $\tau_j$ , the subtrees of the active individuals of the same type are i.i.d., so  $\mathcal{A}_j^{\{x\} \rightarrow t}(k)$  are i.i.d. over all alive particles  $x \in \mathcal{A}_j^u$ . Hence,

$$\mathbb{E} \left\{ \frac{|\mathcal{A}_{j_+}^{\{x\} \rightarrow t}(k)|}{|\mathcal{A}_{j_+}^{u \rightarrow t}(k)|} \mid S_j^u \right\} = \frac{1}{S_j^u}.$$

If the  $j$ th splitting particle was of type  $s$ , it had  $\eta_{su}$  type  $u$  children, so we obtain

$$\mathbb{E}\{\mathcal{A}_j^{\{i_j\} \rightarrow t}(k) \mid \eta_{su}, S_j^u, \mathcal{A}_j^{u \rightarrow t}(k)\} = \sum_{u \in \mathcal{S}} \frac{\eta_{su}}{S_j^u} |\mathcal{A}_j^{u \rightarrow t}(k)|.$$

Combining this and (3.4) with the fact that the event that  $i_j$  is thinned and  $\mathcal{A}_j^{\{i_j\} \rightarrow t}(k)$  are conditionally independent yields that the expectation in (3.3) is at most

$$\sum_{j=1}^k \sum_{s, u \in \mathcal{S}} \mathbb{E} \left\{ \frac{\eta_{su}}{|S_j^u|} \frac{|\mathcal{A}_j^{u \rightarrow t}(k)|}{|\mathcal{A}^t(k)|} \right\} \underbrace{\mathbb{P}\{i_j \text{ is thinned} \mid t(i_j) = s\}}_{(*)} \underbrace{\mathbb{P}\{t(i_j) = s\}}_{(\diamond)}.$$

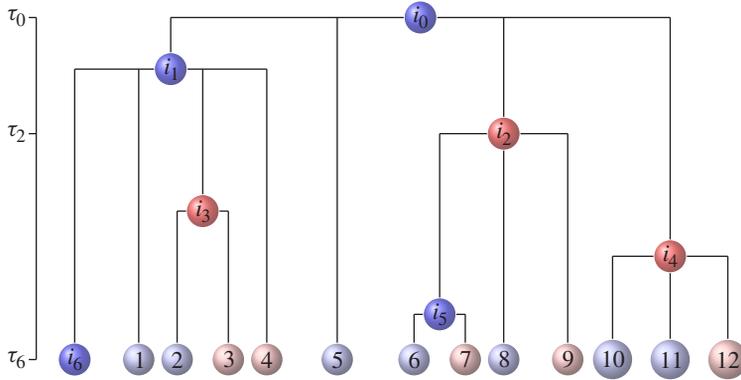


FIGURE 1: Illustration for the notation  $\mathcal{A}_j^{(i_j) \rightarrow t}(k)$  and  $\mathcal{A}_j^{u \rightarrow t}(k)$  with  $j = 2, k = 6$ , and two types  $s$  and  $t$ . Here,  $\mathcal{A}_2^{\{i_2\} \rightarrow s}(6) = \{6, 8\}$  and  $\mathcal{A}_2^{\{i_2\} \rightarrow t}(6) = \{7, 9\}$ . Furthermore,  $\mathcal{A}_2^{s \rightarrow s}(6) = \{1, 5, 6, 8\}$ ,  $\mathcal{A}_2^{s \rightarrow t}(6) = \{7\}$ ,  $\mathcal{A}_2^{t \rightarrow s}(6) = \{2, 10, 11\}$ , and  $\mathcal{A}_2^{t \rightarrow t}(6) = \{3, 4, 9, 12\}$ .

Recall that  $N_k^{(t)}$  denotes the number of type  $t$  vertices among the first  $k$  splits and  $n_t$  is the total number of type  $t$  labels. So  $(*)$  is at most  $N_k^s/n_s$ . Since the lifetimes are i.i.d. exponential, the  $j$ th split is uniform among the  $S_{j-1}$  alive vertices, thus,  $(\diamond)$  equals  $S_{j-1}^t/S_{j-1}$ . Hence,

$$\mathbb{E} \left\{ \frac{\text{th } \mathcal{A}^t(k)}{|\mathcal{A}^t(k)|} \right\} \leq \sum_{j=1}^k \sum_{s,u \in \mathcal{S}} \mathbb{E} \left\{ \frac{\eta_{su} |\mathcal{A}_j^{u \rightarrow t}(k)|}{S_j^u |\mathcal{A}^t(k)|} \right\} \frac{N_j^s S_{j-1}^s}{n_s S_{j-1}}. \tag{3.5}$$

The random variables in the expectation are positively correlated, thus, we cannot take expectations separately. Instead, we split the sum in  $j$  into two parts: from 1 to  $n^\alpha$  ( $\alpha < \beta$ ) and from  $n^\alpha$  to  $n^\beta$ .

In the first sum we use  $\eta_{su}/S_j^u \leq 1$ , while in the second  $\eta_{su} \leq \pi_u n^{\beta-\alpha}$ , the number of type  $u$  vertices in the whole subtree of  $i_j$  (from Propositions 2.2 and 2.3). Also observe that  $\sum_u |\mathcal{A}_j^{u \rightarrow t}(k)| = |\mathcal{A}^t(k)|$  for every  $j$ . Thus,

$$(3.5) \leq \sum_{j=1}^{n^\alpha} \sum_{s \in \mathcal{S}} \frac{\pi_s^2}{\mu_s} \frac{j}{n} (1 + o(1)) + \sum_{j=n^\alpha}^{n^\beta} \sum_{s \in \mathcal{S}} \frac{\pi_u n^{\beta-\alpha}}{\pi_u \tilde{\lambda} j} \frac{\pi_s^2}{\mu_s} \frac{j}{n} (1 + o(1)), \tag{3.6}$$

where we used (1.3), (2.1), and (2.2). Under Assumption 1.1 the sum in  $s$  is  $O(1)$  even as  $|\mathcal{S}| \rightarrow \infty$  later (for homogeneous kernels  $\pi = \mu$ , thus,  $\sum_{s \in \mathcal{S}} \pi_s^2/\mu_s = 1$ ).

As a result, the first sum in (3.6) is of order  $n^{2\alpha-1}$  which is  $o(1)$  if and only if  $\alpha < \frac{1}{2}$ . The second sum is of order  $n^{2\beta-\alpha-1} = o(1)$  if and only if  $\beta < (1 + \alpha)/2 < \frac{3}{4}$ .

**Lemma 3.4.** (Expected number of different labels.) *For all  $k \geq 1$  and  $t \in \mathcal{S}$ , the number of different labels of alive vertices after the  $k$ th split is*

$$|\mathcal{A}^t(k)| = S_k^t \left( 1 - \frac{\tilde{\lambda} \pi_t}{2\mu_t} \frac{k}{n} \right).$$

*Proof.* We assign the labels of new active vertices of type  $t$  by sampling with constraints on the set  $[n]^{(t)}$ , see Section 3. As a result, the number of different labels in  $\mathcal{A}^t(k)$  is dominated by uniform sampling from  $[n]^{(t)}$  with replacement. Since by (2.1),  $S_k^t = \tilde{\lambda} \pi_t k (1 + o(1))$ , we

sample  $\tilde{\lambda}\pi_t k(1 + o(1))$  labels from  $[n]^{(t)}$ . Let  $X_i, i \in [n]^{(t)}$ , be the number of times label  $i$  was chosen:  $X_i \sim \text{bin}(\tilde{\lambda}\pi_t k(1 + o(1)), 1/n_t)$ . Hence,

$$\mathbb{P}\{X_i \geq 2\} = \frac{\tilde{\lambda}^2 \pi_t^2 k^2}{2 n_t^2} + O\left(\frac{k^3}{n_t^3}\right).$$

Thus, the expected number of type  $t$  labels chosen multiple times satisfies

$$\mathbb{E}\left\{ \sum_{i \in [n]^{(t)}} \mathbf{1}_{\{X_i \geq 2\}} \right\} = \sum_{i \in [n]^{(t)}} \mathbb{P}\{X_i \geq 2\} = \frac{\tilde{\lambda}^2 \pi_t^2 k^2}{2\mu_t n} + o(1).$$

Similarly, the expected number of labels chosen more than three times is  $O(k^3/n^2)$ . The statement of the lemma immediately follows since the number of different labels equals the number of alive vertices  $S_k^t = \tilde{\lambda}\pi_t k(1 + o(1))$  minus the multiple labels.

### 4. Connection time

In this section we make the intuitive picture of colliding flows mathematically precise. Computationally, it is slightly easier to let the fluid flow from source  $U$  until it reaches some  $a_n = o(n)$  vertices, then ‘freeze’ it, and then start a flow from source  $V$  until the random time of connection, i.e. when the two flows collide.

The exploration process from  $U := x$  until the split time  $\tau_{a_n}^x$  is coded in  $\text{SWT}_{a_n}^x$  (see (3.2)). Conditioned on the frozen flow of  $x$ , the flow from  $V := y$  can only connect to the flow from  $x$  via an alive vertex in  $\text{SWT}_{a_n}^x$ . Hence, we must leave out the labels  $\mathcal{D}^x(a_n)$  from the possible labels  $[n]$  in the labeling procedure of the BP from  $y$ . A collision edge appears when a label from  $\mathcal{A}^x(a_n)$  appears among the labels in  $\mathcal{D}^y(k)$ . The  $i$ th ( $i \geq 1$ ) collision edge appears at split

$$C_n^{(i)} = \min\{k \geq C_n^{(i-1)} : |\mathcal{A}^x(a_n) \cap \mathcal{D}^y(k)| = i\} \text{ at time } \tau_{C_n^{(i)}}^y.$$

Note that the weight of a path between  $x$  and  $y$  is  $\tau_{a_n}^x + \tau_{C_n^{(i)}}^y + E_i$ , where  $E_i$  is the remaining lifetime of the collision edge after time  $\tau_{a_n}^x$  (This is the remaining time needed for the collision vertex to split in the flow of  $x$ ). The memoryless property of exponential weights implies that  $E_i \stackrel{D}{=} \text{Exp}(1)$ . So the shortest path is through the collision edge that minimizes the expression  $\tau_{C_n^{(i)}}^y + E_i$ . Hence, the shortest path has length

$$\mathcal{P}_n = \tau_{a_n}^x + \min_i \{\tau_{C_n^{(i)}}^y + E_i\}. \tag{4.1}$$

Let us denote the split which minimizes the above expression by  $C_n^{\text{con}}$ . In Figure 2 we illustrate the connection time.

Next, we determine the distribution of the minimum in (4.1): the following proposition states that collision edges appear around time  $O(n/a_n)$ , forming a Poisson point process (PPP).

**Proposition 4.1.** (PPP limit of collision edges.) *Denote a homogeneous PPP with intensity  $\lambda$  by  $\text{PPP}(\lambda)$  and let  $\hat{\lambda} = \tilde{\lambda} \sum_{s \in \mathcal{S}} \pi_s^2 / \mu_s$ . Conditioned on the event that both CTBPs survive, the point process*

$$\left\{ \frac{C_n^{(i)} a_n}{n} \right\}_i \stackrel{D}{\rightarrow} \text{PPP}(\hat{\lambda}) \text{ as } n \rightarrow \infty.$$

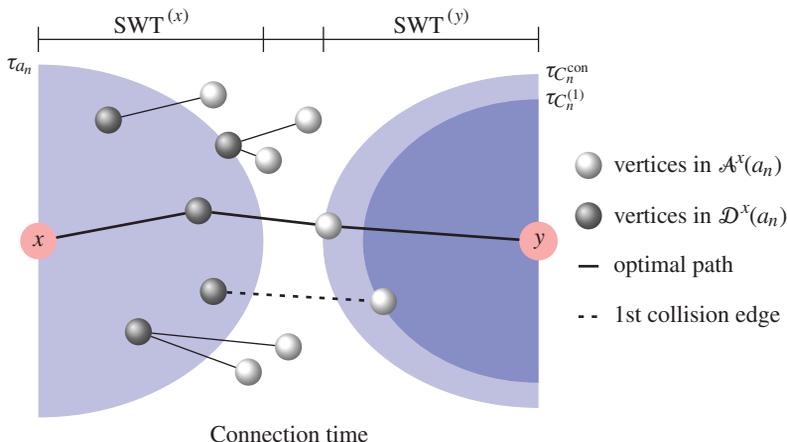


FIGURE 2: Illustration of the connection time. The first collision edge to appear is not necessarily the one which will be on the optimal path.

*Proof.* We first show by induction that for fixed  $n$ ,

$$\frac{C_n^{(i)} a_n}{n} \stackrel{D}{=} \text{gamma}(i, \hat{\lambda}_n) \quad \text{with } \hat{\lambda}_n := \frac{n}{a_n} \sum_{s \in \mathcal{S}} \frac{S_{a_n}^{x,s}}{n_s} \pi_s (1 + o(1)) + o(1). \quad (4.2)$$

For  $C_n^{(1)}$  we can write

$$\mathbb{P} \left\{ C_n^{(1)} > x \frac{n}{a_n} \right\} = \prod_{j=1}^{xn/a_n} \mathbb{E} \{ \mathbb{P} \{ C_n^{(1)} > j \mid C_n^{(1)} > j-1, \mathcal{F}_{j-1} \} \},$$

where  $\mathcal{F}_{j-1} := \sigma(\text{SWT}_{a_n}^x, \text{SWT}_{j-1}^y)$ . To calculate  $\mathbb{P} \{ C_n^{(1)} > j \mid C_n^{(1)} > j-1, \mathcal{F}_{j-1} \}$ , we condition on the type of the splitting vertex in  $\text{SWT}^y$  and find the probability that its label is not an alive label in  $\text{SWT}^x$ . Combining the results of Lemmas 3.3 and 3.4, we can substitute the number of nonthinned different marks in  $\mathcal{A}^{x,t}(a_n)$  by  $S_{a_n}^{x,t}$ , and neglect the error factor of order  $(1 - \hat{\lambda}_{a_n}/n)$  along the lines. This error can be included in the  $o(1)$  term of the last line of the following:

$$\begin{aligned} \mathbb{P} \left\{ C_n^{(1)} > \frac{xn}{a_n} \right\} &= \prod_{j=1}^{xn/a_n} \left[ \sum_{t \in \mathcal{S}} \frac{S_{j-1}^{y,t}}{S_{j-1}^y} \left( 1 - \frac{S_{a_n}^{x,t}}{n_t} \right) \right] \\ &= \prod_{j=1}^{xn/a_n} \left[ 1 - \sum_{t \in \mathcal{S}} \frac{S_{j-1}^{y,t}}{S_{j-1}^y} \frac{S_{a_n}^{x,t}}{n_t} \right] \\ &= \exp \left( -x \frac{n}{a_n} \sum_{t \in \mathcal{S}} \frac{S_{a_n}^{x,t}}{n_t} \frac{a_n}{xn} \underbrace{\sum_{j=1}^{xn/a_n} \frac{S_{j-1}^{y,t}}{S_{j-1}^y}}_{(*)} + o(1) \right), \end{aligned}$$

where  $(*)$  equals  $\pi_t(1 + o(1))$  by (2.2). So  $C_n^{(1)} a_n/n \stackrel{D}{=} \text{Exp}(\hat{\lambda}_n) = \text{gamma}(1, \hat{\lambda}_n)$ . From the

induction hypothesis,

$$\begin{aligned} & \mathbb{P}\left\{C_n^{(i+1)} > \frac{xn}{a_n}\right\} \\ &= \int_0^x \mathbb{P}\left\{C_n^{(i+1)} > \frac{xn}{a_n} \mid \frac{C_n^{(i)} a_n}{n} = s + o(1)\right\} \frac{\hat{\lambda}_n^i s^{i-1}}{(i-1)!} e^{-\hat{\lambda}_n s} ds + \mathbb{P}\left\{C_n^{(i)} > \frac{xn}{a_n}\right\} \\ &= \int_0^x \prod_{j=sn/a_n}^{xn/a_n} \left[1 - \sum_{t \in \delta} \frac{S_{j-1}^{y,t} S_{a_n}^{x,t}}{S_{j-1}^y n_t}\right] \frac{\hat{\lambda}_n^i s^{i-1}}{(i-1)!} e^{-\hat{\lambda}_n s} ds + \mathbb{P}\left\{C_n^{(i)} > \frac{xn}{a_n}\right\} \\ &= \frac{\hat{\lambda}_n^i x^i}{i!} e^{-\hat{\lambda}_n x} + \mathbb{P}\left\{C_n^{(i)} > \frac{xn}{a_n}\right\}. \end{aligned}$$

Differentiating the cumulative distribution function of  $C_n^{(i+1)} a_n/n$  with respect to  $x$  yields the probability density function of  $\text{gamma}(i + 1, \hat{\lambda}_n)$ , which proves (4.2). Thus, for fixed  $n$  the point process,

$$\left\{\frac{C_n^{(i)} a_n}{n}\right\}_i \text{ is a PPP}(\hat{\lambda}_n).$$

From (1.3) and (2.1) it follows that  $\hat{\lambda}_n \rightarrow \hat{\lambda}$  a.s. as  $n \rightarrow \infty$ . Hence, the assertion holds.

We will see below in Section 5.1 that since  $\tau_{C_n^{(i)}} - (1/\tilde{\lambda}) \log C_n^{(i)} \rightarrow -(1/\tilde{\lambda}) \log(W/\tilde{\lambda})$  a.s., to determine the distribution of  $\tau_{C_n^{\text{con}}} + E_{\text{con}}$  we need the following lemma.

**Lemma 4.1.** *Let  $(P_i)_i$  denote the points of a PPP(1) process and independent of the  $(P_i)_i$  let  $(E_i)_i$  be i.i.d. Exp(1) random variables. Then*

$$\min_i \left\{ \frac{1}{\tilde{\lambda}} \log P_i + E_i \right\} \stackrel{D}{=} -\frac{1}{\tilde{\lambda}} \Lambda + \frac{1}{\tilde{\lambda}} \log(\tilde{\lambda} + 1),$$

where  $\Lambda$  follows a standard Gumbel distribution, i.e.  $\mathbb{P}\{\Lambda \leq x\} = e^{-e^{-x}}$ .

*Proof.* For convenience let  $X_i \sim (1/\tilde{\lambda}) \log P_i$ . We will calculate the tail distribution of the minimum by conditioning on the Poisson points first. Thus,

$$\begin{aligned} \mathbb{P}\left\{\min_i (X_i + E_i) \geq z\right\} &= \mathbb{E}\{\mathbb{P}\{\text{for all } i, X_i + E_i > z \mid X_1, X_2, \dots\}\} \\ &= \mathbb{E}\left\{\prod_i e^{-(z-X_i)_+}\right\} \\ &= \mathbb{E}\left\{\prod_{i: P_i < e^{\tilde{\lambda}z}} e^{-z(P_i)^{1/\tilde{\lambda}}}\right\}. \end{aligned}$$

The number of Poisson points  $Z$  in the interval  $[0, e^{\tilde{\lambda}z}]$  follows a Poisson random variable with parameter  $e^{\tilde{\lambda}z}$ . Conditioning on  $Z$ , the points  $P_i, i \leq Z$ , are independent and uniform in the interval  $[0, e^{\tilde{\lambda}z}]$ . Thus,

$$\mathbb{E}\left\{\prod_{i: P_i < e^{\tilde{\lambda}z}} e^{-z(P_i)^{1/\tilde{\lambda}}}\right\} = \mathbb{E}\{\mathbb{E}\{e^{-zU_i^{1/\tilde{\lambda}}}\}^Z\} = \mathbb{E}\left\{\left(\frac{\tilde{\lambda}}{\tilde{\lambda} + 1}\right)^Z\right\} = \exp\left\{e^{\tilde{\lambda}z} \left(1 - \frac{\tilde{\lambda}}{\tilde{\lambda} + 1}\right)\right\}.$$

It is easy to see that if  $\Lambda$  is a standard Gumbel then  $\mathbb{P}\{-a\Lambda + b > x\} = e^{-e^{x/a} e^{-b/a}}$ . Thus, here  $a = 1/\tilde{\lambda}$  and  $b = (1/\tilde{\lambda}) \log(\tilde{\lambda} + 1)$ , and the lemma follows.

Note that the Poisson convergence is true only for finitely many collision edges, hence, we need a stochastic bound on the index  $i$  where the minimum in (4.1) is attained.

**Lemma 4.2.** *The probability that the shortest-weight path is not among the first  $k$  collision edges decays exponentially in  $k$ , i.e.*

$$\mathbb{P}\{\arg(C_n^{(i)} = C_n^{\text{con}}) > k\} \leq \left(\frac{\tilde{\lambda}}{\tilde{\lambda} + 1}\right)^k.$$

Consequently, the distribution of the rescaled connection time is stochastically dominated by the  $\text{geo}(1/(\tilde{\lambda} + 1))$ th point in the PPP in Proposition 4.1. Thus,

$$P_n^{\text{con}} := \frac{a_n C_n^{\text{con}}}{n} \leq \sum_{i=1}^N \tilde{E}_i, \tag{4.3}$$

where  $N \sim \text{geo}(1/(\tilde{\lambda} + 1))$  and independently  $\tilde{E}_i$ s are independent  $\text{Exp}(\tilde{\lambda})$ s.

*Proof of Lemma 4.2.* We use again the notation  $X_i \sim (1/\tilde{\lambda}) \log(P_i)$ , where  $P_i$  is the  $i$ th point in a PPP(1) process. To bound the event that the minimum is taken at an index at least  $k + 1$ , we condition on the value of the minimum ( $= z$ ) and also on the value of the  $X_{k+1} = c$  with  $z \geq c$  so that

$$\begin{aligned} \mathbb{P}\{\arg \min > k\} &= \mathbb{P}\left\{\min_{i \leq k} (X_i + E_i) > \min_{j \geq k+1} (X_j + E_j)\right\} \\ &= \mathbb{E}\left\{\mathbb{P}\left\{\text{for all } i \leq k, E_i > z - X_i \mid X_{k+1} = c, \min_{j \geq k+1} (X_j + E_j) = z\right\}\right\} \\ &= \mathbb{E}\left\{\mathbb{E}\left[\prod_{i \leq k} e^{-(z-X_i)} \mid X_{k+1} = c, \min = z\right]\right\}. \end{aligned}$$

Now  $X_{k+1} = c$  means that the  $(k + 1)$ th point in the Poisson process is  $P_{k+1} = e^{\tilde{\lambda}c}$ . Conditioning on this information means that the first  $k$  points have the same distribution as  $(U_i)_{1 \leq i \leq k}$  independent uniform points on  $[0, e^{\tilde{\lambda}c}]$ . Thus,

$$\mathbb{E}\left\{\prod_{i \leq k} e^{-(z-X_i)} \mid X_{k+1} = c\right\} = \prod_{i \leq k} \mathbb{E}\{e^{-z} U_i^{1/\tilde{\lambda}}\} = \left(e^{c-z} \frac{\tilde{\lambda}}{\tilde{\lambda} + 1}\right)^k.$$

Then clearly we have

$$\mathbb{P}\{\arg \min > k\} = \left(\frac{\tilde{\lambda}}{\tilde{\lambda} + 1}\right)^k \mathbb{E}\left\{\mathbb{E}\left[e^{k(c-z)} \mid X_{j+1} = c, \min_{j \geq k+1} = z\right]\right\} \leq \left(\frac{\tilde{\lambda}}{\tilde{\lambda} + 1}\right)^k,$$

where the last inequality comes from the fact that  $z - c \geq 0$  a.s. This is immediate since the sequence of  $X_j$ s are increasing  $X_j \geq X_{k+1}$  for all  $j \geq k + 1$ . Thus,  $z = \min_{j \geq k+1} (X_j + E_j) \geq X_{k+1} = c$ .

### 5. Proof of the main results

We begin this section with the proof of Theorem 1.1 for finite-type graphs. Then, for the general setting, we approximate general kernels by regular finitary kernels. We conclude with the proof of Theorem 1.2. The proofs are analogous to the proofs of [8].

**5.1. Proof of Theorem 1.1: finite-type space**

Let  $(\mathcal{S}, \mu)$  be an arbitrary finite-type ground space and  $\kappa$  a kernel that satisfies Assumption 1.1. Denote the split of a type  $t$  vertex by  $t^\dagger$ , the event that the connection happens by ‘con’ and recall that  $\mathcal{A}^{x,t}(k)$  is the collection of labels of alive type  $t$  vertices after the  $k$ th split in the flow of  $x$ . We first argue that the probability that the shortest-weight path contains thinned vertices, i.e. it is not a real shortest path, is  $o(1)$ .

From a union bound, the probability that the connection happens through a thinned alive vertex  $v$  can be bounded from above by

$$\sum_{t \in \mathcal{S}} \left( \underbrace{\mathbb{P}\{v \in \text{th}^C \mathcal{A}^{x,t}(a_n)\}}_{(*)} + \underbrace{\mathbb{P}\{v \in \text{th}^C \mathcal{A}^{y,t}(C_n^{\text{con}})\}}_{(\diamond)} \right) \mathbb{P}\{\text{con}, t^\dagger\}.$$

Using Lemma 3.3, the probability  $(*)$  is  $o(1)$  for every  $t \in \mathcal{S}$  if  $a_n = n^\alpha$  with  $\alpha < \frac{3}{4}$ . Similarly, by Lemma 3.3, Proposition 4.1, and (4.3) the probability  $(\diamond)$  is also  $o(1)$  for every  $t \in \mathcal{S}$  if  $\alpha > \frac{1}{4}$ . Hence, for all choices of  $\alpha \in (\frac{1}{4}, \frac{3}{4})$ , the shortest-weight path w.h.p. does not contain a thinned vertex.

Now we determine the distribution of the shortest-weight path. Recall (4.1),

$$\mathcal{P}_n = \tau_{a_n}^x + \min_i \{ \tau_{C_n^{(i)}}^y + E_i \}. \tag{5.1}$$

In  $\Psi_\kappa(t)$ ,  $M_t = e^{-\tilde{\lambda}t} |\underline{S}(t)|$  is a martingale. Thus,  $\tau_k$  can be expressed as

$$\tau_k = -\frac{1}{\tilde{\lambda}} \log \frac{M_{\tau_k}}{\tilde{\lambda}} + \frac{1}{\tilde{\lambda}} \log \frac{S_k}{\tilde{\lambda}k} + \frac{1}{\tilde{\lambda}} \log k. \tag{5.2}$$

Applying (5.2) to the minimum in (5.1), we have

$$\min_i \{ \tau_{C_n^{(i)}} + E_i \} = \min_i \left\{ -\frac{1}{\tilde{\lambda}} \log \left( \frac{\hat{\lambda} M_{\tau_{C_n^{(i)}}}^y}{\tilde{\lambda}} \right) + \frac{1}{\tilde{\lambda}} \log \left( \frac{S_{C_n^{(i)}}^y}{\tilde{\lambda} C_n^{(i)}} \right) + \frac{1}{\tilde{\lambda}} \log (\hat{\lambda} C_n^{(i)}) + E_i \right\}.$$

For  $i$  fixed,  $n \rightarrow \infty$ ,  $\hat{\lambda} C_n^{(i)} = (n/a_n) \hat{\lambda} P_i \rightarrow \infty$  and, thus, conditioned on the survival of the branching process,  $\tau_{C_n^{(i)}} \rightarrow \infty$  holds also, implying that  $M_{\tau_{C_n^{(i)}}}^y \rightarrow (W^y \mid W^y > 0) := \hat{W}^y$  a.s. and in  $L_2$ , and  $S_{C_n^{(i)}}^y / (\tilde{\lambda} C_n^{(i)}) \rightarrow 1$  also a.s. Furthermore, we also know from Proposition 4.1 that the law of  $(\hat{\lambda} C_n^{(i)})_i$  converges to a PPP(1) process. Thus, the minimum becomes asymptotically as  $n \rightarrow \infty$ :

$$\min_i \{ \tau_{C_n^{(i)}} + E_i \} = -\frac{1}{\tilde{\lambda}} \log \frac{\hat{W}^y \hat{\lambda}}{\tilde{\lambda}} + \frac{1}{\tilde{\lambda}} \log \frac{n}{a_n} + \min_i \left\{ \frac{1}{\tilde{\lambda}} \log (\hat{\lambda} P_i) + E_i \right\}.$$

For the last term we can apply Lemma 4.1 to obtain

$$\min_i \{ \tau_{C_n^{(i)}} + E_i \} = -\frac{1}{\tilde{\lambda}} \log \frac{\hat{W}^y \hat{\lambda}}{\tilde{\lambda}(\tilde{\lambda} + 1)} + \frac{1}{\tilde{\lambda}} \log \frac{n}{a_n} - \frac{1}{\tilde{\lambda}} \Lambda. \tag{5.3}$$

Now (5.2), (5.3), and  $\hat{\lambda} = \tilde{\lambda} \sum_{s \in \mathcal{S}} \pi(s)^2 / \mu(s)$  imply that

$$\mathcal{P}_n = \frac{1}{\tilde{\lambda}} \log n - \frac{1}{\tilde{\lambda}} \log \hat{W}^x \hat{W}^y - \frac{1}{\tilde{\lambda}} \Lambda + \frac{1}{\tilde{\lambda}} \log \frac{\tilde{\lambda}(\tilde{\lambda} + 1)}{\sum_{s \in \mathcal{S}} \pi(s)^2 / \mu(s)},$$

with  $\Lambda$  being a standard Gumbel,  $\hat{W}^z$ ,  $z = x, y$ , the limits of the (independent) martingales  $e^{-\lambda t} |S(t)^z|$  in  $\Psi_\kappa^z$ ,  $z = x, y$ , conditioned on nonextinction. Since  $W^i > 0$  a.s. on nonextinction and  $\sum_{s \in \mathcal{S}} \pi(s)^2 / \mu(s) < \infty$  under Assumption 1.1, these quantities are well defined.

Next, we derive the central limit theorem for the hopcount  $\mathcal{H}_n$ . We start by proving that the hopcount of the connecting vertices in the processes  $\Psi_\kappa^x$  and  $\Psi_\kappa^y$  are independent conditioned on their types. We remind the reader that first the flow of  $x$  is constructed, which is then frozen at time  $a_n$ . Then, each time a vertex splits in  $\Psi_\kappa^y$ , the label of its type  $t$  children are picked uniformly at random without replacement amongst the possible labels  $n_t \setminus \{\mathcal{D}^x(a_n)\}$ .

**Lemma 5.1.** *Conditioned on the event {collision happens from  $\text{SWT}^{(y)}$  to  $\text{SWT}^{(x)}$  at the  $C_n$ th split at a type  $t$  vertex},*

- (a) *the label  $v_C$  where this happens is uniform among all labels in  $\mathcal{A}_{a_n}^{x,t}$ ,*
- (b) *the hopcounts  $G_{a_n}^{x,t}$ ,  $G_{C_n}^{y,t}$  are independent given that the collision happens at a type  $t$  vertex,*
- (c) *the distribution of  $G_{a_n}^{x,t}$  and  $G_{C_n}^{y,t}$  is the same as of a uniformly picked type  $t$  alive individual from  $\mathcal{A}_{a_n}^x$  and  $\mathcal{A}_{C_n}^y$ , respectively.*

*Proof.* The first two statements of the lemma are a straightforward consequence of the following urn problem: in an urn there are  $M$  balls of type  $A$  (alive) and  $N$  balls of type  $U$  (untouched), each of them labeled. We carry out the following procedure: in the  $k$ th step we draw  $d_k$  balls without replacement, record the label of type  $A$  and type  $U$  balls in sets  $L_A$  and  $L_U$ , respectively, and then put the balls back in to the urn. Thus,  $L_i(k)$  consists of all the labels of type  $i$  balls  $i = A, U$  that have been drawn up to step  $k$ . It is easy to show that at any time the content of the set  $L_A$  and  $L_U$  is a uniformly picked set of size  $|L_A|$  and  $|L_U|$  among all the labels in  $A$  and  $B$ , respectively. In particular, for every label  $v \in A$  we have

$$\mathbb{P}\{v \notin L_A(k)\} = \prod_{j \leq k} \left(1 - \frac{d_j}{M + N}\right).$$

Now let  $A = \mathcal{A}_{a_n}^{x,t}$  and  $U = [n_t] \setminus (\mathcal{D}_{a_n}^{x,t} \cup \mathcal{A}_{a_n}^{x,t})$ . Then,  $L_A(k) \cup L_U(k) = \mathcal{A}^{y,t}(k)$ . The previous argument states that at any time the labels in  $L_A = \mathcal{A}^{y,t} \cap \mathcal{A}^{x,t}$  are uniformly picked from the labels of  $\mathcal{A}^{x,t}$ .

The collision edges between the processes  $x$  and  $y$  are established such that, in each step  $k$ , if a type  $t$  particle dies in  $\Psi_\kappa^y$ , we pick a uniform label among  $L_A(k) \cup L_U(k)$ , and collision happens if it is of type  $A$ . Clearly, by the previous argument, whenever this is the case, conditioned on the picked label being in  $L_A(k)$ , the its label is a uniformly picked label among  $\mathcal{A}_{a_n}^{x,t}$  (and clearly also uniform in  $L_A(k)$ ). Furthermore, the step  $k$  when a label of type  $A$  enters  $L_A$  is independent of the label itself, thus, the generation of the label at the connection in  $\text{SWT}^x$  and in  $\text{SWT}^y$  are independent and equal to the generation of a uniformly picked alive individuals of type  $t$ .

Next, we determine the limit distribution of the hopcount. Let  $G_k^{z,t}$  denote the generation of a uniformly picked alive individual of type  $t$  in  $\mathcal{A}^z(k)$ ,  $z = x, y$ , and recall the definition of  $C_n^{\text{con}}$  given just after (4.1). Then

$$\mathcal{H}_n = \sum_{t \in \mathcal{S}} \mathbf{1}_{\{C_n^{\text{con}} \cap t\}} (G_{a_n}^{x,t} + G_{C_n^{\text{con}}}^{y,t}).$$

Thus,

$$\begin{aligned} & \frac{\mathcal{H}_n - ((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}} \\ &= \sum_{t \in \mathcal{S}} \mathbf{1}_{\{C_n^{\text{con}} \cap t\}} \frac{G_{a_n}^{x,t} - ((\tilde{\lambda} + 1)/\tilde{\lambda}) \log a_n}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log a_n}} \sqrt{\frac{\log a_n}{\log n}} \\ &+ \sum_{t \in \mathcal{S}} \mathbf{1}_{\{C_n^{\text{con}} \cap t\}} \frac{G_{C_n^{\text{con}}}^{y,t} - ((\tilde{\lambda} + 1)/\tilde{\lambda}) \log C_n^{\text{con}}}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log C_n^{\text{con}}}} \sqrt{\frac{\log C_n^{\text{con}}}{\log n}} \\ &+ \frac{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log(C_n^{\text{con}} a_n/n)}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}}. \end{aligned}$$

First, using the fact that conditioned on  $C_n^{\text{con}}$  and the type, the two terms containing  $G_*^{z,t}$  converge to independent standard normal variables (independently of the type). Furthermore, by Lemma 4.2, the last term tends to 0. From Lemma 5.1 we ensure the independence of the two limiting normal variables, thus, it follows that the right-hand side is tending to

$$N\left(0, \frac{\log a_n}{\log n}\right) + N\left(0, \frac{\log(n C_n^{\text{con}}/a_n)}{\log n}\right) \rightarrow N(0, 1).$$

We used the fact that Lemma 4.2 and (4.3) imply that the total variance is tending to 1.

**5.2. Proof of Theorem 1.1: general kernels**

We approximate a general kernel  $\kappa$  by a sequence of regular finitary kernels, where we assume that  $|\mathcal{S}| < \infty$  (the regular finitary case and the finite-type case differ only in notation). Types with zero measure cannot be simply ignored, since they can alter  $G(n, \kappa)$  significantly. However, with a simple argument from [12], we can assume that  $\mu_s > 0$  for every  $s \in \mathcal{S}$ .

Given a sequence of finite partitions  $\alpha_m = \{A_{m1}, \dots, A_{mM_m}\}$ ,  $m \geq 1$ , of  $\mathcal{S}$  and an  $x \in \mathcal{S}$ , we define  $A_m^x$  as the element of  $\alpha_m$  for which  $x \in A_m^x$ . As usual,  $\text{diam}(A)$  denotes  $\sup\{d(x, y) : x, y \in A\}$  for  $A \subset \mathcal{S}$ , where  $d$  is the metric on  $\mathcal{S}$ . For any ground space  $(\mathcal{S}, \mu)$  there exists a sequence of finite partitions  $\alpha_m$  (see [12, Lemma 7.1]) such that

- (1) each  $A_{mi}$  is a  $\mu$ -continuity set,
- (2) for each  $m$ ,  $\alpha_{m+1}$  refines  $\alpha_m$ ,
- (3) for a.e.  $x \in \mathcal{S}$ ,  $\text{diam}(A_m^x) \rightarrow 0$ , as  $m \rightarrow \infty$ .

For such a sequence  $\alpha_m$ ,  $m \geq 1$ , we define the regular finitary kernel

$$\bar{\kappa}_m(x, y) := \frac{1}{\mu(A_m^x)\mu(A_m^y)} \iint_{A_m^x \times A_m^y} \kappa(s, t) \, d\mu(s) \, d\mu(t). \tag{5.4}$$

If  $\kappa$  is continuous a.e. then property (3) implies that  $\bar{\kappa}_m(x, y) \rightarrow \kappa(x, y)$  for a.e.  $(x, y) \in \mathcal{S}^2$ . If the original  $\kappa$  is homogeneous, then so is  $\bar{\kappa}_m$ , with  $\int_{\mathcal{S}} \bar{\kappa}_m(x, y) \, d\mu(y) = \tilde{\lambda} + 1$ .

Let  $(\mathcal{S}, \mu)$  be an arbitrary ground space and assume that kernel  $\kappa$  satisfies the conditions of Theorem 1.1. These define the sequence of random graphs  $(G(n, \kappa))_{n \geq 1}$ . Take any sequence of finite partitions  $\alpha_m = \{A_{m1}, \dots, A_{mM_m}\}$ ,  $m \geq 1$ , that satisfy properties (1), (2), and (3)

described above. For each  $m$ ,  $\bar{\kappa}_m$  (defined in (5.4), with ground space  $(\mathcal{G}_m, \mu)$ ) defines the sequence  $(G(n, \bar{\kappa}_m))_{n,m \geq 1}$ . Note that in the proofs for finite-type kernels, none of the estimates depend on  $\mu_t$  or  $|\mathcal{G}_m|$ , so all the error terms are uniform. The condition  $\sup \kappa(x, y) < \infty$  is necessary because it is used in the proof of Lemma 3.1.

In the proof we let  $n$  and  $m$  tend to  $\infty$  simultaneously in a carefully chosen way and couple  $(G(n, \bar{\kappa}_m))_{n,m \geq 1}$  and  $(G(n, \kappa))_{n \geq 1}$ . For fixed  $m$ , from the proof of [14, Lemma 2.1 and Theorem 3.1] it is easy to see that

$$\left| \mathbb{P} \left\{ \frac{H_n^{m(n)} - ((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}{\sqrt{((\tilde{\lambda} + 1)/\tilde{\lambda}) \log n}} < x \right\} - \Phi(x) \right| \leq C(\tilde{\lambda}) \left( \frac{1}{\sqrt{\log a_n}} + \frac{1}{\sqrt{\log(n/a_n)}} \right), \tag{5.5}$$

where  $C(\tilde{\lambda})$  is a  $\tilde{\lambda}$ -dependent constant. Setting  $a_n := \sqrt{n}$ , the error is  $O(1/\sqrt{\log n})$ .

Let  $\delta_m = \max_i \text{diam}(A_{mi})$ . Since  $\kappa$  is uniformly continuous, there exists  $\varepsilon_m = \varepsilon_m(\kappa, \delta_m)$  such that for all  $x, y$ , and all  $(u, v) \in A_m^x \times A_m^y$ :  $|\bar{\kappa}_m(u, v) - \kappa(x, y)| \leq \varepsilon_m$ . Recall that an edge is present with probability  $\kappa(\cdot, \cdot)/n$ , thus,

$$\mathbb{P} \{ \mathbf{1}_{\{\{x,y\} \in e(G(n,\kappa))\}} \neq \mathbf{1}_{\{\{A_m^x, A_m^y\} \in e(G(n,\bar{\kappa}_m))\}} \} \leq \frac{2\varepsilon_m}{n}.$$

Summing over all possible edges, we find for the edge sets that

$$\mathbb{P} \{ e(G(n, \kappa)) \neq e(G(n, \bar{\kappa}_m)) \} \leq \frac{2n^2\varepsilon_m}{2n} = n\varepsilon_m.$$

For a fixed  $n$ , define

$$m(n) := \inf \{ m : \varepsilon_m n \sqrt{\log n} \leq 1 \}.$$

Then for all  $m > m(n)$ , the coupling between  $G(n, \bar{\kappa}_m)$  and  $G(n, \kappa)$  fails with probability less than  $1/\sqrt{\log n}$ . Under this coupling and also for the hopcount, we have

$$\mathbb{P} \{ \mathcal{H}_n \neq \mathcal{H}_n^{m(n)} \} \leq \frac{1}{\sqrt{\log n}} = o(1).$$

Combining this error bound with the error bound in (5.5), we obtain the desired central limit theorem (CLT) for the hopcount  $\mathcal{H}_n$  in the sequence  $(G(n, \kappa))_{n \geq 1}$  as  $n \rightarrow \infty$ .

Now we turn to the proof of the convergence of the shortest-weight path. To avoid conflicting notation we will denote  $\mathcal{P}_n(\kappa)$  to be the shortest-weight path belonging to  $G(n, \kappa)$ . The same coupling argument as for the hopcount yields

$$\mathbb{P} \{ \mathcal{P}_n(\kappa) \neq \mathcal{P}_n(\bar{\kappa}_{m(n)}) \} \leq \frac{1}{\sqrt{\log n}}. \tag{5.6}$$

The result for the finite-type case is

$$\mathcal{P}_n(\bar{\kappa}_m) - \frac{1}{\tilde{\lambda}} \log n \xrightarrow{D} -\frac{1}{\tilde{\lambda}} \hat{W}_{(m)}^x \hat{W}_{(m)}^y - \frac{1}{\tilde{\lambda}} \Lambda + \frac{1}{\tilde{\lambda}} \log \frac{\tilde{\lambda}(\tilde{\lambda} + 1)}{\sum_{s \in \mathcal{S}} \pi_m(s)^2 / \mu_m(s)},$$

where  $\Lambda$  is a standard Gumbel,  $\hat{W}_{(m)}^i$ ,  $i = x, y$ , are i.i.d. limits of the martingales from the branching processes with kernel  $\bar{\kappa}_{m(n)}$ , conditioned on being positive. Note that we can take  $m = m(n)$  so that  $\kappa_{m(n)} \rightarrow \kappa$  as  $n \rightarrow \infty$ . The coupling (5.6) implies that  $\mathcal{P}_n(\kappa)$  and  $\mathcal{P}_n(\bar{\kappa}_{m(n)})$

must have the same limit, hence,  $\widehat{W}_{(m(n))}^x \widehat{W}_{(m(n))}^y \rightarrow \widehat{W}^x \widehat{W}^y$ . Also, by the properties of the partition  $\mathcal{A}_{m(n)}$ , we have

$$\sum_{s \in \mathcal{S}_{m(n)}} \frac{\pi_{m(n)}(s)^2}{\mu_{m(n)}(s)} \rightarrow \int_{\mathcal{S}} \frac{\pi(s)}{\mu(s)} \pi(ds).$$

This completes the proof.

**5.3. Proof of Theorem 1.2: dense setting**

In the dense setting, where  $\tilde{\lambda}_n \rightarrow \infty$ , we have a sequence of kernels  $\kappa_n$ ,  $n \geq 1$ . The type  $t$  neighbors of a type  $s$  vertex have distribution  $\eta_{st}^{(n)} \stackrel{D}{=} \text{bin}(n_t - \delta_{st}, \kappa_n(s, t)/n)$ . Note that under both assumptions  $\mathbb{E}\{\eta_{s,t}^{(n)}\} \rightarrow \infty$ , hence, we avoid the coupling to Poisson variables as in Lemma 3.1, and immediately apply the CLT result of [21] to  $\Psi_{\kappa_n}$ , where the offspring distribution  $(D_i^{(n)} \mid \text{type } s \text{ splits})$  is the sum of independent binomial variables  $\eta_{st}^{(n)}$ ,  $t \in \mathcal{S}$ . We apply a similar argument as for the general case. Namely, in  $\Psi_{\kappa_n}$ , for a uniformly picked type  $t$  individual at step  $k$ ,

$$\left| \mathbb{P} \left\{ \frac{G_k^{(n),t} - ((\tilde{\lambda}_n + 1)/\tilde{\lambda}_n) \log k}{\sqrt{((\tilde{\lambda}_n + 1)/\tilde{\lambda}_n) \log k}} < x \right\} - \Phi(x) \right| \leq C(\tilde{\lambda}) \frac{1}{\sqrt{\log k}},$$

which, when considering the connection of the flows at  $k = a_n$  and  $C_n^{\text{con}} = \Theta(n/a_n)$ , will yield an error term of  $1/\sqrt{\log n}$  for  $\mathcal{H}_n$ . Considering  $\tilde{\lambda}_n \rightarrow \infty$ , the term  $(\tilde{\lambda}_n + 1)/\tilde{\lambda}_n \rightarrow 1$  in the denominator and this immediately yields the desired result for the hopcount in Theorem 1.2. The centering constant can be replaced by  $\log n$  if and only if  $\sqrt{\log n} = o(\tilde{\lambda}_n)$ .

For the shortest-weight path, we have to reinvestigate the distribution of the split times  $\tau_{a_n}$  and  $\tau_{C_n^{\text{con}}}$ . Since the time between two consecutive splits, given the number of alive individuals in the BP, is the minimum of that many independent exponentials, for every  $m$ ,

$$\tau_m \stackrel{D}{=} \sum_{i=1}^m \frac{E_i}{S_i^{(n)}}, \tag{5.7}$$

with  $E_i$  i.i.d.  $\text{Exp}(1)$ . Recall that  $D_j^{(n)}$  denotes the number of children of the  $j$ th dying particle. Then, since both in case (1)  $D_j^{(n)} \sim \text{Poi}(\tilde{\lambda}_n + 1)$  and in case (2)  $D_j^{(n)} \sim \text{Poi}((\tilde{\lambda}_n + 1) \int_{\mathcal{S}} \kappa(s, t) \mu(dt))$  if  $j$  is of type  $s$ , by the usual CLT for Poisson variables, we have, for  $S_i^{(n)} = \sum_{j=1}^i D_j^{(n)} - (i - 1)$ ,

$$\frac{S_i^{(n)} - i\tilde{\lambda}_n}{\sqrt{i(\tilde{\lambda}_n + 1)}} \xrightarrow{D} N(0, 1).$$

From this we only need  $S_i^{(n)} = \tilde{\lambda}_n i(1 + o(1))$  for all  $i \geq 1$ , combined with (5.7), yielding  $\tilde{\lambda}_n \tau_m \approx \sum_{i=1}^m E_i/i$ . Note that the sequence  $E_m/m, E_{m-1}/m-1, \dots, E_1/1$  has the same distribution as the spacings between  $m$  i.i.d. exponentials  $E'_i$ , so  $\sum_{i=1}^m E_i/i \stackrel{D}{=} \max_{1 \leq i \leq m} E'_i := B_m$ . From here it is straightforward that

$$\mathbb{P}\{\tilde{\lambda}_n \tau_{a_n}^{(x)} - \log a_n \leq x\} = \mathbb{P}\{B_{a_n} \leq x + \log a_n\} = (1 - e^{-(x+\log a_n)})^{a_n} \rightarrow \exp(-e^{-x})$$

is the distribution function of a standard Gumbel. Similarly to the proof in Section 5.1, we conclude that

$$(\tilde{\lambda}_n \tau_{a_n}^{(x)} - \log a_n, \tilde{\lambda}_n \tau_{C_n^{\text{con}}}^{(y)} - \log C_n^{\text{con}}) \xrightarrow{D} (\Lambda_1, \Lambda_2),$$

where  $\Lambda_1, \Lambda_2$  are i.i.d. standard Gumbels. From (5.1) we obtain

$$\begin{aligned} \tilde{\lambda}_n \mathcal{P}_n - \log n &\stackrel{D}{=} \underbrace{\tilde{\lambda}_n \tau_{a_n}^{(x)} - \log a_n - \log \frac{n}{a_n}}_{\xrightarrow{D} \Lambda_1} \\ &+ \min_i \left\{ \underbrace{\tilde{\lambda}_n \tau_{C_n^{(i)}}^{(y)} - \log C_n^{(i)}}_{\xrightarrow{D} \Lambda_2} + \tilde{\lambda}_n \left( \frac{1}{\tilde{\lambda}_n} \log C_n^{(i)} + E_i \right) \right\}. \end{aligned} \tag{5.8}$$

Proposition 4.1 yields  $C_n^{(i)} = n/a_n P_n^{(i)}$ , with  $P_n^{(i)}$  PPP( $\hat{\lambda}_n$ ) points, and then Lemma 4.1 yields that the last term in the minimum equals

$$\tilde{\lambda}_n \min_i \left\{ \frac{1}{\tilde{\lambda}_n} \log C_n^{(i)} + E_i \right\} = \log \frac{n}{a_n} - \Lambda_3 + \log \frac{(\tilde{\lambda}_n + 1)}{\tilde{\lambda}_n \int_{\mathcal{S}} (\pi(t)/\mu(t)) \pi(dt)},$$

with  $\Lambda_3$  again a standard Gumbel. Since  $\tilde{\lambda}_n \rightarrow \infty$ , the  $(\tilde{\lambda}_n + 1)/\tilde{\lambda}_n$  term vanishes in the limit, and, under (1.6), it follows that  $\int_{\mathcal{S}} (\pi(t)/\mu(t)) \pi(dt) = 1$ . Combining these with (5.8) completes the proof.

### Acknowledgements

This research was supported by a KTIA-OTKA grant (no. CNK 77778), funded by the Hungarian National Development Agency (NFÜ) from a source provided by the Research and Technological Innovation Fund (KTIA). This publication was supported by the Társadalmi Megújulás Operatív Program TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund. The work of JK is supported in part by The Netherlands Organization for Scientific Research (NWO). We thank Remco van der Hofstad for useful discussions.

### References

- [1] AMINI, H. AND LELARGE, M. (2015). The diameter of weighted random graphs. *Ann. Appl. Prob.* **3**, 1686–1727.
- [2] ASMUSSEN, S. (1977). Almost sure behavior of linear functionals of supercritical branching processes. *Trans. Amer. Math. Soc.* **231**, 233–248.
- [3] ATHREYA, K. B. AND NEY, P. E. (2004). *Branching Processes*. Dover, Mineola, NY.
- [4] BARBOUR, A. D. AND REINERT, G. (2013). Approximating the epidemic curve. *Electron. J. Prob.* **18**, 30pp.
- [5] BHAMIDI, S. (2008). First passage percolation on locally treelike networks. I. Dense random graphs. *J. Math. Phys.* **49**, 125–218.
- [6] BHAMIDI, S. AND VAN DER HOFSTAD, R. (2012). Weak disorder asymptotics in the stochastic mean-field model of distance. *Ann. Appl. Prob.* **22**, 29–69.
- [7] 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.
- [8] BHAMIDI, S., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2011). First passage percolation on the Erdős–Rényi random graph. *Combin. Prob. Comput.* **20**, 683–707.
- [9] 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/1210.6839>.
- [10] BHAMIDI, S., VAN DER HOFSTAD, R. AND KOMJÁTHY, J. (2014). The front of the epidemic spread and first passage percolation. *J. Appl. Prob.* **51**, 101–121.
- [11] BOLLOBÁS, B. AND FERNANDEZ DE LA VEGA, W. (1982). The diameter of random regular graphs. *Combinatorica* **2**, 125–134.
- [12] BOLLOBÁS, B., JANSON, S. AND RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31**, 3–122.
- [13] BRITTON, T., DEIJFEN, M. AND MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Statist. Phys.* **124**, 1377–1397.

- [14] BÜHLER, W. J. (1971). Generations and degree of relationship in supercritical Markov branching processes. *Prob. Theory Relat. Fields* **18**, 141–152.
- [15] CHUNG, F. AND LU, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6**, 125–145.
- [16] CHUNG, F. AND LU, L. (2003). The average distance in a random graph with given expected degrees. *Internet Math.* **1**, 91–113.
- [17] FERNHOLZ, D. AND RAMACHANDRAN, V. (2007). The diameter of sparse random graphs. *Random Structures Algorithms* **31**, 482–516.
- [18] HOWARD, C. D. (2004). Models of first-passage percolation. In *Probability on Discrete Structures*, Springer, Berlin, pp. 125–173.
- [19] JANSON, S. (1999). One, two and three times  $\log n/n$  for paths in a complete graph with random weights. *Combin. Prob. Comput.* **8**, 347–361.
- [20] JANSON, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110**, 177–245.
- [21] KHARLAMOV, B. P. (1969). The numbers of generations in a branching process with an arbitrary set of particle types. *Theory Prob. Appl.* **14**, 432–449.
- [22] NORROS, I. AND REITTU, H. (2006). On a conditionally Poissonian graph process. *Adv. Appl. Prob.* **38**, 59–75.
- [23] SÖDERBERG, B. (2002). General formalism for inhomogeneous random graphs. *Phys. Rev. E* **66**, 066121.
- [24] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND VAN MIEGHEM, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures Algorithms* **27**, 76–123.
- [25] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. AND ZNAMENSKI, D. (2007). Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Prob.* **12**, 703–766.