

# FIRST PASSAGE PERCOLATION ON INHOMOGENEOUS RANDOM GRAPHS

BY ISTVÁN KOLOSSVÁRY <sup>\*</sup>  
JÚLIA KOMJÁTHY <sup>†</sup>

*Budapest University of Technology and Economics  
Inter-University Centre for Telecommunications and Informatics 4028  
Debrecen, Kassai út 26.*

We investigate first passage percolation on inhomogeneous random graphs. The random graph model  $G(n, \kappa)$  we study is the model introduced by Bollobás, Janson and Riordan in [9], where each vertex has a type from a type space  $\mathcal{S}$  and edge probabilities are independent, but depending 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  $\tilde{\lambda}_n \rightarrow \tilde{\lambda}$  is finite or infinite, under the assumption that the average number of neighbors  $\tilde{\lambda}_n$  of a vertex is independent of the type. The paper is a generalization of [5] written by Bhamidi, van der Hofstad and Hooghiemstra, where FPP is explored on the Erdős-Rényi graphs.

**1. Introduction and the main results.** First passage percolation (FPP), generally speaking, deals with the asymptotic behavior of first-passage times of 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, just to name a few.

Particularly, two quantities of interest for FPP on finite weighted random graph models are the minimal weight of a path between two vertices  $x$  and  $y$  and the number of edges, often referred to the hopcount, on this path. Without edge weights these quantities coincide. Other natural questions

---

<sup>\*</sup>This research project was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001\*\*project, supported by the European Union, co-financed by the European Social Fund

<sup>†</sup>This research project was supported by the grant KTIA-OTKA # CNK 77778, funded by the Hungarian National Development Agency (NFÜ) from a source provided by KTIA.

*AMS 2000 subject classifications:* Primary 05C80, 90B15, 60J85

*Keywords and phrases:* inhomogeneous random graphs, shortest weight path, hopcount, first passage percolation, continuous-time multi-type branching process

can be 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 [9] with independent identically distributed (i.i.d.) exponential edge weights with rate one.

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 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 one. 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  $\log n$  [15]. A similar phenomena can be observed for the IHRG, see Subsection 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 [6]. Other related results are discussed in Subsection 1.4. We begin by introducing the IHRG model in Subsection 1.1. This section is continued with the statements of our main results in Subsection 1.2 and then the main ideas of the proofs are sketched in Subsection 1.3.

1.1. *The model.* The random graph model we consider is a general inhomogeneous random graph model introduced by Bollobás, Janson and Riordan in [9]. 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  $\mathcal{S}$  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  $\mathcal{S}$ . We assume that the empirical distribution

$$\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$$

converges in probability to  $\mu$  as  $n \rightarrow \infty$ , where  $\delta_x$  is the measure consisting of a point mass of weight 1 at  $x$ . This convergence condition is equivalent to the condition

$$(1.1) \quad \nu_n(S) := \frac{\#\{i : x_i \in S\}}{n} \xrightarrow{\mathbf{P}} \mu(S),$$

for every  $\mu$ -continuity set  $S \subset \mathcal{S}$  (i.e.  $S$  is measurable and there is no mass on the boundary of  $S$ ). The pair  $(\mathcal{S}, \mu)$  is called a ground space and for

a sequence  $(\mathbf{x}_n)_{n \geq 1}$  satisfying (1.1) we say that the triplet  $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$  defines a vertex space  $\nu$ . Further, a kernel  $\kappa$  on a ground space is a symmetric non-negative measurable function on  $\mathcal{S} \times \mathcal{S}$ . The natural interpretation of  $\kappa$  is that it measures the density of edges.

The simple random graph  $G(n, \kappa)$  for a given kernel  $\kappa$  and vertex space  $\nu$  is defined as follows. An edge  $\{ij\}$  (with  $i \neq j$ ) exists with probability

$$(1.2) \quad p_{ij} := \min \left\{ \frac{\kappa(x_i, x_j)}{n}, 1 \right\}.$$

Independently of all randomness each edge is given an  $\text{Exp}(1)$  edge weight.

We want to exclude cases where the vertex set of  $G(n, \kappa)$  can be split into two parts so that the probability of an edge from one part to the other is zero, i.e. we want to ensure the emergence of a single giant component later. To do so, further restrictions are needed for the kernel  $\kappa$ .

DEFINITION 1.1. *A kernel  $\kappa$  on a ground space  $(\mathcal{S}, \mu)$  is irreducible if*

$$A \subseteq \mathcal{S} \text{ and } \kappa = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus A) \text{ implies } \mu(A) = 0 \text{ or } \mu(\mathcal{S} \setminus A) = 0.$$

*As a slight modification we say that  $\kappa$  is quasi-irreducible if there is a  $\mu$ -continuity set  $\mathcal{S}' \subseteq \mathcal{S}$  with  $\mu(\mathcal{S}') > 0$  such that the restriction of  $\kappa$  to  $\mathcal{S}' \times \mathcal{S}'$  is irreducible and  $\kappa(x, y) = 0$  if  $x \notin \mathcal{S}'$  or  $y \notin \mathcal{S}'$ .*

$G(n, \kappa)$  is a sparse graph, i.e. the number of edges  $e(G(n, \kappa))$  is linear in  $n$ , since

$$\mathbf{E}[e(G(n, \kappa))] = \mathbf{E} \left[ \sum_{i < j} \min \{ \kappa(x_i, x_j)/n, 1 \} \right],$$

tends to  $n(\frac{1}{2} \iint \kappa)$  under certain conditions ensuring this "well-behavior" (see [9, Lemma 8.1]). This is formulated in the notion of graphical kernels.

DEFINITION 1.2. *A kernel  $\kappa$  is graphical on a vertex space  $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$  if it is continuous almost everywhere (a.e.) and in  $L^1(\mathcal{S} \times \mathcal{S}, \mu \times \mu)$ , furthermore*

$$(1.3) \quad \frac{1}{n} \mathbf{E}[e(G(n, \kappa))] \rightarrow \frac{1}{2} \iint_{\mathcal{S}^2} \kappa(x, y) d\mu(x) d\mu(y).$$

For example, condition (1.3) holds whenever  $\kappa$  is bounded and  $\nu$  is a vertex space. An important ingredient in the proof will be the use of approximating

kernels, where  $\kappa$  depends on  $n$ . We say that a sequence  $\kappa_n$  of kernels on  $(\mathcal{S}, \mu)$  is graphical on  $\nu$  with limit  $\kappa$  if, for a.e.  $(y, z) \in \mathcal{S}^2$ ,

$$y_n \rightarrow y \text{ and } z_n \rightarrow z \text{ imply that } \kappa_n(y_n, z_n) \rightarrow \kappa(y, z),$$

$\kappa \in L_1$  and continuous a.e., and

$$(1.4) \quad \frac{1}{n} \mathbf{E} [e(G(n, \kappa_n))] \rightarrow \frac{1}{2} \iint_{\mathcal{S}^2} \kappa(x, y) d\mu(x) d\mu(y).$$

The approximation of a general kernel will be done with an appropriate sequence of step functions. This motivates the special case of regular finitary kernels: 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 graph. If the type-space  $\mathcal{S} = \{1, 2, \dots, r\}$ , and  $n_t$  stands for the number of type  $t$  vertices (so  $\sum_{t \in \mathcal{S}} n_t = n$ ), condition 1.1 becomes

$$(1.5) \quad \frac{n_t}{n} \xrightarrow{\mathbf{P}} \mu_t \quad \text{holds for every } t \in \mathcal{S},$$

and  $\kappa = (\kappa(s, t))_{s, t=1}^r$  is a symmetric non-negative  $r \times r$  matrix. Note that finite-type kernels are automatically graphical. Further note that the Erdős-Rényi (ER) random graph is a special case of a finite-type graph when  $r = 1$  and  $\kappa = c$ . Then the probabilities  $p_{ij}$  are all simply  $c/n$  (for  $n > c$ ).

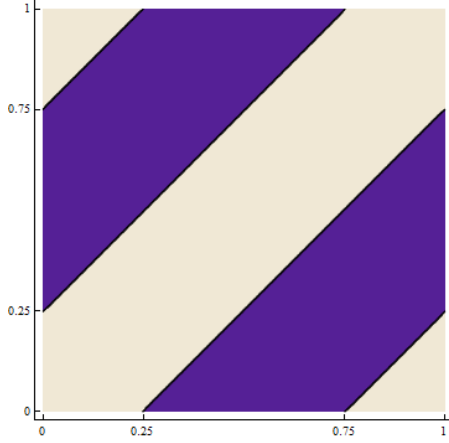
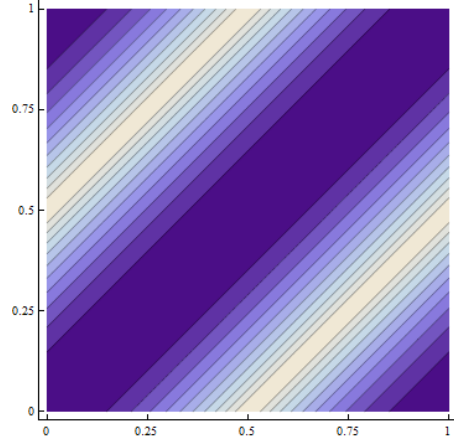
We make an important assumption on the random graphs  $G(n, \kappa)$ . We will assume that asymptotically the average degree of a vertex is independent of its type. This is referred to as the homogeneous case in [9, Example 4.6]. In this case the global behavior of  $G(n, \kappa)$  in the limit is the same as of the ER random graph, but the local behavior can be quite different. In the general setting this assumption can be formulated as

$$(AG) \quad \int_{\mathcal{S}} \kappa(x, y) d\mu(y) = \tilde{\lambda} + 1 \quad \text{for a.e. } x.$$

An example for a kernel  $\kappa$  satisfying (AG) can be given by taking  $\mathcal{S} = (0, 1]$  (interpreted as the 1 dimensional torus  $\mathbb{T}^1$ ),  $\mu$  as the Lebesgue measure and  $\kappa(x, y) = h(d(x, y))$  for an even function  $h \geq 0$  of period 1, where  $d(x, y)$  is the metric given on  $\mathcal{S}$ . Figures 1 and 2 show the contours of two such examples, where the dark purple strips indicate where the contour is zero.

In the finite-type case we use the following notation:

$$(1.6) \quad \lambda_{st} := \kappa(s, t) \mu_t,$$

FIG 1.  $h(z) = \mathbb{1}[|z| < 0.25]$ FIG 2.  $h(z) = z^2$ 

for all  $s, t \in \mathcal{S}$ . The number of type  $t$  neighbors of a type  $s$  vertex are binomially distributed with parameters  $n_t - \delta_{st}$  and  $\kappa(s, t)/n$  ( $\delta_{st} = 1$  if and only if  $s = t$ ). Thus from (1.5) we get that in the limit  $\lambda_{st}$  gives us the average number of type  $t$  neighbors of a type  $s$  vertex. From here we construct the matrix

$$(1.7) \quad A = \begin{pmatrix} \lambda_{11} - 1 & \lambda_{12} & \cdots & \lambda_{1r} \\ \lambda_{21} & \lambda_{22} - 1 & \cdots & \lambda_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{r1} & \lambda_{r2} & \cdots & \lambda_{rr} - 1 \end{pmatrix}.$$

ASSUMPTION 1.3. *We will assume that the row sums of the matrix  $A$  are the same and equal to  $\tilde{\lambda} > 0$ . Also assume that  $A$  is irreducible, i.e. there exists a  $k_0$  s.t.  $A^{k_0}$  has strictly positive entries.*

In the finite-type case Assumption 1.3 is the equivalent of (AG). The  $\tilde{\lambda} > 0$  condition is necessary and sufficient for a giant component to emerge in  $G(n, \kappa)$  [9, Theorem 3.1]. Let  $\tilde{\pi}$  denote the normalized left eigenvector corresponding to the eigenvalue  $\tilde{\lambda}$  of  $A$ . Under Assumption 1.3 we find that

$$(1.8) \quad \tilde{\pi} = \mu.$$

Indeed, using the symmetry of  $\kappa$ ,  $\mu A = \tilde{\lambda} \mu$  follows immediately.

Let us introduce some standard notation. Let  $\text{Bin}(n, p)$ ,  $\text{Poi}(\lambda)$ ,  $\text{Exp}(\mu)$  respectively denote a binomial, a Poisson and an exponential random variable with the parameters having their usual meaning. Convergence almost

surely, in distribution and in probability are denoted by  $\xrightarrow{a.s.}$ ,  $\xrightarrow{d}$ ,  $\xrightarrow{\mathbf{P}}$  respectively. A sequence of events holds with high probability (whp), if it holds with probability tending to 1 as  $n \rightarrow \infty$ . We use the Landau symbols  $O$  and  $o$  with their usual meaning. We say that a sequence of random variables  $X_n$  satisfies  $X_n = O_{\mathbf{P}}(b_n)$  if  $X_n/b_n$  is tight (i.e.  $\forall \varepsilon > 0 \exists M : \mathbf{P}[X_n > Mb_n] < \varepsilon$ ) or  $X_n = o_{\mathbf{P}}(b_n)$  if  $X_n/b_n \xrightarrow{\mathbf{P}} 0$ . Now we turn to the main results.

1.2. *Main results.* We investigate the weight and the number of edges on the shortest-weight path between two uniformly picked vertices  $x$  and  $y$ . Let  $\Gamma_{xy}$  denote the set of all  $\pi$  paths in  $G(n, \kappa)$  between  $x$  and  $y$ . Denote the weight of the shortest-weight path by

$$(1.9) \quad \mathcal{P}_n = \min_{\pi \in \Gamma_{xy}} \sum_{e \in \pi} X_e,$$

where  $X_e$  is the exponential edge weight attached to edge  $e$  in the construction of  $G(n, \kappa)$ . Let  $\mathcal{H}_n$  denote the number of edges or hopcount of this path. If the two vertices are in different components of the graph, then let  $\mathcal{P}_n, \mathcal{H}_n = \infty$ . The following theorem describes the asymptotic behavior of these two quantities.

**THEOREM 1.4** (Asymptotics of Hopcount & Shortest weight). *Let  $(\mathcal{S}, \mu)$  be an arbitrary ground space and  $\kappa$  be a uniformly continuous, quasi-irreducible, graphical kernel on  $(\mathcal{S}, \mu)$  that satisfies  $\sup \kappa(x, y) < \infty$  and*

$$\int_{\mathcal{S}} \kappa(x, y) d\mu(y) = \tilde{\lambda} + 1 < \infty \quad \text{for a.e. } x \in \mathcal{S}.$$

*Then the hopcount  $\mathcal{H}_n$  and the minimal weight  $\mathcal{P}_n$  between two uniformly chosen vertices, conditioned on being connected, satisfy*

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

*where  $Z$  is a standard normal variable. Furthermore,  $L$  is a non-degenerate real valued random variable whose distribution can be precisely determined from the behavior of the multi-type branching process that arises when exploring a component of  $G(n, \kappa)$ .*

A possible extension is to have  $\kappa$  depend on  $n$ . In this case we have a sequence of matrices  $A_n$ , each satisfying Assumption 1.3 with  $\tilde{\lambda}_n$  as the sum of the rows. Then we can trivially extend our theorems in the following form:

COROLLARY 1.5. *If  $\lim_{n \rightarrow \infty} \tilde{\lambda}_n = \tilde{\lambda} < \infty$  then for the hopcount we have*

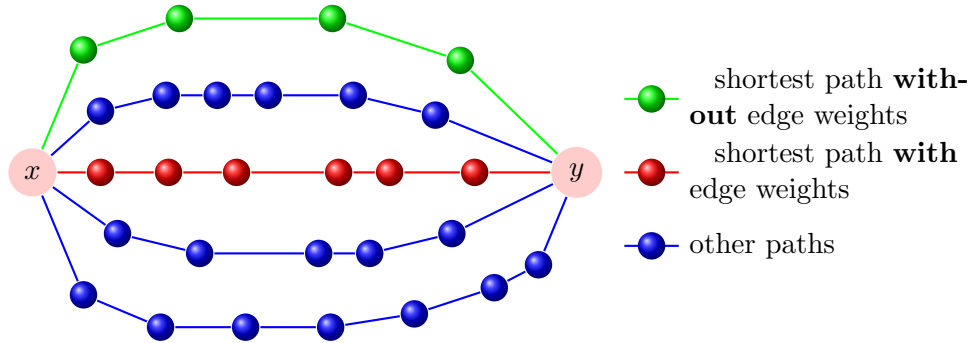
$$\left( \frac{\mathcal{H}_n - \frac{\tilde{\lambda}+1}{\tilde{\lambda}} \log n}{\sqrt{\frac{\tilde{\lambda}+1}{\tilde{\lambda}} \log n}}, \mathcal{P}_n - \frac{1}{\tilde{\lambda}} \log n \right) \xrightarrow{d} (Z, L),$$

where  $Z$  and  $L$  are as in Theorem 1.4.

It is interesting to compare these results with its counterpart where there are no edge weights. Then of course  $\mathcal{H}_n = \mathcal{P}_n$ . Theorem 3.14 of [9] in our context says that in this case the typical distance between two vertices in the giant component still scales as  $\log n$ , but with a different constant:  $1/\log(\tilde{\lambda}+1)$ . Comparing the constants, we see that

$$\frac{1}{\tilde{\lambda}} < \frac{1}{\log(1 + \tilde{\lambda})} < \frac{\tilde{\lambda} + 1}{\tilde{\lambda}}.$$

This shows that by adding edge weights the structure of the graph changes. Along the shortest-weight path more vertices are visited ( $\mathcal{H}_n$ ) than on the path with the least number of vertices. At the same time the weight of the path ( $\mathcal{P}_n$ ) becomes smaller than the number of edges on the path with the least number of vertices. The figure below illustrates this, where the red path is the shortest-weight path while the green is the one with the least number of vertices.



Similarly as in [5] we also investigate the dense regime, where  $\lim \tilde{\lambda}_n = \infty$ , i.e. the average degree tends to infinity. In this case whp any two vertices are connected, so the giant component contains  $n(1 - o(1))$  vertices (see [9]). Again comparing with the counterpart without edge weights, the change in the graph structure is even more significant. Without edge weights the graph is ultra small, meaning that graph distances between uniformly chosen

vertices are  $o(\log n)$ . However with the addition of edge weights the following theorem states that even the magnitudes do not coincide. We get the same type of behavior as observed in the sparse setting, which means that on the shortest-weight path many more vertices are visited.

**THEOREM 1.6** (Dense setting). *Under Assumption 1.3 and  $\lim_{n \rightarrow \infty} \tilde{\lambda}_n = \infty$ , then we have*

$$\left( \frac{H_n - \frac{\tilde{\lambda}_n + 1}{\lambda_n} \log n}{\sqrt{\log n}}, \tilde{\lambda}_n \mathcal{P}_n - \log n \right) \xrightarrow{d} (Z, \tilde{L}),$$

where  $Z$  is a standard normal variable,  $\tilde{L}$  is equal in distribution to the sum of independent random variables  $Y_1 + Y_2 - Y_3$ , with  $Y_i$  i.i.d. standard Gumbel random variables. Further, we can substitute  $\frac{\tilde{\lambda}_n + 1}{\lambda_n}$  in the centralizing of the hopcount by 1 if and only if  $\tilde{\lambda}_n / \sqrt{\log n} \rightarrow \infty$ .

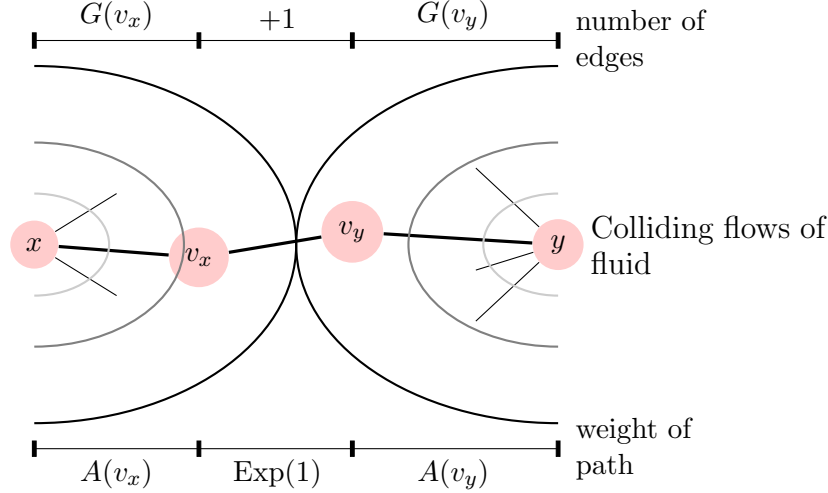
**1.3. Sketch of proofs.** When searching for the shortest-weight path between two vertices, the intuitive picture to keep in mind is to let fluid percolate at a constant rate simultaneously from each vertex. After time  $\tau$  the flow from  $x$  contains the vertices whose shortest-weight path is at most  $\tau$  from  $x$ . It is standard to relate this exploration process of the neighborhood of a vertex to a branching process (BP). The generation and lifetime of a particle in a BP corresponds to its hopcount and weight of optimal path in the exploration process. Section 2 introduces a continuous-time multi-type branching process that arises naturally in an exploration process of the giant component of  $G(n, \kappa)$ . An analog of the main results is proved in Theorem 2.2 in the BP setting.

For the results to carry through we need an embedding of the BP into the IHRG. This is dealt with in Section 3. The vertices of the BP and the vertices of the IHRG need to be identified with one another with high probability. To achieve this labels are assigned to the vertices according to their type in such a way that we can rule out loop and multiple edges in the IHRG. To deal with the problem of cycles in the graph, the notion of thinning is introduced on branching processes.

Intuitively it is clear that when the two flows of fluid collide, then the shortest-weight path between  $x$  and  $y$  has been found. Thus it is crucial to determine when the connection actually happens. The random time of the collision will be referred to as the connection time  $C_{xy}$ . In the figure below, the fluids flow continuously and the vertices  $v_x$  and  $v_y$  denote the



last wetted vertices (along the shortest path) by the flows from  $x$  and  $y$  respectively before time  $C_{xy}$ .



Let  $G(v_x)$  denote the number of edges between the source  $x$  and vertex  $v_x$  and  $A(v_x)$  the weight of this path. Then for the hopcount  $\mathcal{H}_n(x, y)$  and for the weight  $\mathcal{P}_n(x, y)$  of the shortest-weight path we get that

$$\begin{aligned} \mathcal{H}_n(x, y) &= G(v_x) + G(v_y) + 1, \\ \mathcal{P}_n(x, y) &= A(v_x) + A(v_y) + \text{Exp}(1). \end{aligned}$$

The problem is that we can't let the fluids flow continuously, but only in discrete steps as new vertices are wetted by the flows. Thus, the sum of the path length plus the remaining edge-weight between  $v_x$  and  $v_y$  must be minimized over all possible choices of last vertices  $v_x$  and  $v_y$ . The rigorous treatment of the connection time is done in Section 4

After these preparations we begin Section 5 with the proof of Theorem 1.6 for finite-type graphs. To prove the main result in the general setting a sequence of finite partitions of  $\mathcal{S}$  is used. The approximating step function on  $\mathcal{S} \times \mathcal{S}$  will be given by the average of the kernel  $\kappa$  determined by the partition. This ensures that if  $\kappa$  is homogeneous then so are all of its approximations.

1.4. *Related work.* First passage percolation has been investigated on various models, such as the integer lattice, the mean-field model, configuration model, graphs with i.i.d. vertex degrees both without edge weights (see e.g. [3, 8, 9, 11, 12, 18, 19, 21, 22]) and with exponential weights (see e.g. [2, 4, 5, 13–15]). The list of results are far from complete, we only attempt to discuss the ones directly related to the present paper.

The IHRG model was extensively investigated by Bollobás, Janson and Riordan in [9] from many different aspects, including typical distances without edge weights. We already showed what effect the addition of edge weights has on the structure of the graph in Subsection 1.2. Many other models are closely related to the IHRG model, for full details see [9, Sections 4 and 16].

The classical supercritical Erdős-Rényi (ER) random graph  $G(n, c/n)$  ( $c > 1$ ) is the special case when  $|\mathcal{S}| = 1$ . Our results generalize the FPP results of Bhamidi, van der Hofstad and Hooghiemstra on ER random graphs with i.i.d. exponential edge weights in [5]. Finite-type graphs were previously studied by Soderberg [20].

A restrictive, yet natural class of inhomogeneous random graphs 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 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 [18] give results on the existence and size of a giant component with random  $w_i$ . For deterministic  $w_i$ , Chung and Lu [10, 11] show that, under certain conditions, (without edge weights) the typical distance between two vertices is  $\log n / \log \bar{d}$ , where  $\bar{d} = \sum w_i^2 / \ell_n$ .

A closely related model is the generalized random graph introduced by Britton, Deijfen and Martin-Löf [21] with edge probabilities  $p_{ij} = w_i w_j / (n + w_i w_j)$ . They show that conditioned on the vertex degrees, the resulting graph is uniformly distributed over all graphs with the given degree sequence. As a result Bhamidi, van der Hofstad and Hooghiemstra in [6] prove FPP results for the latter two models with *general* continuous edge weights. This is a corollary of their result for the configuration model.

**2. Multi-type branching processes.** In this section we collect the needed properties of the branching process that arises when exploring a component of  $G(n, \kappa)$ . Let us define a multi-type continuous time branching process with type space  $\mathcal{S}$ , where a particle of type  $x \in \mathcal{S}$ , when it splits, gives birth to a set of offsprings distributed as a Poisson process on  $\mathcal{S}$  with intensity measure  $\kappa(x, y)d\mu(y)$ . That is, 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  $E \in \mathcal{S}$ ,  $\mathcal{D}_m^E$  and  $\mathcal{A}_m^E$  stands for the set of dead and alive particles after the  $m$ -th split whose type belongs to the set  $E \in \mathcal{S}$ . For  $E = \mathcal{S}$  we simply write  $\mathcal{D}_m$  and  $\mathcal{A}_m$ , respectively. Similarly,  $S_m^E$  and  $N_m^E$  stand for the number of alive and dead individuals with type in the set  $E \in \mathcal{S}$ , and

we simply write  $S_m$  for  $E = \mathcal{S}$ . Let us also write  $\tau_m$  for the time of the  $m$ -th split.

The branching process  $\Psi_\kappa^x$  arises naturally when exploring a component of  $G(n, \kappa)$ , analogously to that of the exploration of FPP on the Erdős-Rényi graph. In this exploration process no size-biasing of the degrees happens, due to the independence of edges.

Analogous to the lines of [9], we define an integral operator, whose norm establishes a direct connection with the emergence of a giant component in the random graph  $G(n, \kappa)$  and the survival of  $\Psi_\kappa$ . Let

$$(2.1) \quad (T_\kappa f)(x) := \int_{\mathcal{S}} \kappa(x, y) f(y) d\mu(y),$$

for any measurable function  $f$  such that this integral is defined. The norm of  $T_\kappa$  is

$$\|T_\kappa\| := \sup \{ \|T_\kappa f\|_2 : f \geq 0, \|f\|_2 \leq 1 \} \leq \infty,$$

where  $\|\cdot\|_2$  is the norm of  $L^2(\mathcal{S}, \mu)$ .

In the finite-type case each type- $t$  particle gives birth to a Poisson number of type  $s$  children with parameter  $\lambda_{st} = \kappa(s, t)\mu(t)$ , so in this case  $T_\kappa f = (A + I)f$  with  $A$  defined in (1.7). Easy calculations yields the norm:

$$\|T_\kappa\| = \|A + I\| = \tilde{\lambda} + 1,$$

where the last equality holds because of Assumption 1.3.

Let us recall Theorem 3.1 of [9]: a giant component emerges in  $G(n, \kappa_n)$  with  $\kappa_n \rightarrow \kappa$  and its corresponding branching process survives with positive probability if and only if  $\|T_\kappa\| > 1$ . This is why we assumed throughout that  $\tilde{\lambda} > 0$ . More precisely, the survival probability of the branching process is the maximal solution to the functional equation  $\rho(x) = 1 - \exp\{-T_\kappa \rho(x)\}$ . Under assumption (AG), the maximal solution is independent of the type  $\rho(x) = \rho = 1 - \exp\{-(\tilde{\lambda} + 1)\rho\}$ .

Throughout the proofs, we will mainly work with finite-type case  $\mathcal{S} = \{1, \dots, r\}$  given with the mean-offspring matrix  $A$  in (1.7). Recall that  $\pi$  stands for the normalized left main eigenvector of  $A$  (i.e.  $\pi$  is a probability measure on  $\mathcal{S}$ ). We will need the following limit theorems for  $S_m^s$  and  $N_m^s$ , the number of alive and dead individuals of a given type  $s \in \mathcal{S}$ : (these results do not require Assumption (1.3)), and all can be found in Athreya-Ney [1].

**THEOREM 2.1 (BP-asymptotics).** *For a finite-type continuous time Branching Process as defined above,*

(i) On the set of non-extinction as  $m \rightarrow \infty$

$$(2.2) \quad \lim_{m \rightarrow \infty} \frac{(S_m^1, \dots, S_m^r)}{\tilde{\lambda}m} = \pi \text{ a.s.},$$

(ii) and similarly

$$(2.3) \quad \frac{(N_m^1, \dots, N_m^r)}{m} \xrightarrow{\text{a.s.}} \pi.$$

As a consequence we also get that

$$(2.4) \quad \frac{1}{m} S_m \rightarrow \tilde{\lambda}, \quad \text{and} \quad \frac{S_m^t}{S_m} \xrightarrow{(2.2)} \pi_s \text{ a.s.}$$

Estimates can also be given for the magnitude of the error terms in (2.4). More precisely, the exponent of the error terms is of order  $0.5 \wedge (\mathcal{R}e(\lambda_2)/\tilde{\lambda})$ , where  $\lambda_2$  is the eigenvalue of  $A$  with second largest real part. If this quotient is less than  $1/2$ , then  $(S_m - \tilde{\lambda}m)/\sqrt{m}$  and  $(S_m^t - \tilde{\lambda}\pi_t m)/\sqrt{m}$  tends to a multidimensional normal random variable. For details we refer to [16, Theorems 3.22-24.]. We will later make use of these in some of our arguments.

The generation of a particle in the Branching Process corresponds to the hopcount of the vertex in the IHRG. Thus, when we establish the connection between the exploration processes of  $x$  and  $y$ , the hopcount of the vertices at the connection are needed. We will later see that the generation of the particles in the two BP-s will be only independent when conditioned on their type. The following lemma for the multi-type BP will handle these issues:

**THEOREM 2.2** (Generation of a uniformly picked particle in a given type-set). *Let  $G_m^E$  denote the generation of a uniformly picked individual from  $\mathcal{A}^E(m)$ ,  $E \subset \mathcal{S}$ . Then, conditioned on survival, the following holds for  $m \rightarrow \infty$ :*

$$(2.5) \quad \frac{G_m^E - \frac{\tilde{\lambda}+1}{\lambda} \log m}{\sqrt{\frac{\tilde{\lambda}+1}{\lambda} \log m}} \xrightarrow{d} Z,$$

where  $Z$  is a standard normal variable.

Thus, the generation of a uniformly picked alive of a given type follows a central limit theorem with parameters independent of the type. The lemma is a consequence of the result by Kharlamov [17]. He proved the result in continuous time (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 the stationary  $\pi$  exponentially fast in time and that the expectation and variance of the generation of the density of type  $y$  individuals after unit time for any type  $y \in \mathcal{S}$  is uniformly bounded in  $y$ . These conditions clearly hold if the life-time is exponential and the total number of children is  $\text{Poi}(\tilde{\lambda} + 1)$ .

To get the result from continuous time to discrete time, one needs to replace  $t$  by the  $m$ -th split time  $\tau_m$  and use aperiodicity of the types along generations.

For the asymptotic behavior of split times  $\tau_m$  we cite again [1]:

THEOREM 2.3.

$$(2.6) \quad \tau_m - \frac{1}{\tilde{\lambda}} \log m \xrightarrow{a.s.} -\frac{1}{\tilde{\lambda}} \log \left( \frac{1}{\tilde{\lambda}} W \right)$$

on the set of non-extinction, and  $\mathbf{P}_x(W > 0) = \rho(x)$ . The distribution of  $W$  is given by distribution of the limit of the continuous time martingale  $\lim_{t \rightarrow \infty} e^{-\tilde{\lambda}t} |\underline{S}_t| = W$ , where  $\underline{S}_t$  is vector of alive particles in  $\Psi_\kappa(t)$ .

Under assumption (1.3), it is not hard to see that the distribution of  $W$  is independent of the initial type  $x$ , and is the same as for a single-type BP with  $\text{Poi}(\tilde{\lambda})$  offsprings. Further,  $\mathbf{E}(W) = 1$  and the moment generating function  $M_W(t)$  of  $W$  satisfies the functional equation

$$(2.7) \quad M_W(t) = \int_0^\infty \exp \left\{ (\tilde{\lambda} + 1) \left( M_W(te^{-\tilde{\lambda}y}) - 1 \right) \right\} e^{-y} dy.$$

In some calculations, we will need the following lemma for the total number of alive individuals. This lemma is slightly stronger than what follows from the almost sure convergence, and in fact this is the crucial element which is missing for the general  $A$  matrix case, i.e. without the Assumption (1.3).

LEMMA 2.4. [Summable error terms for  $S_j$ ] *Let us suppose Assumption 1.3 holds. Then, Conditioned on the survival, there exists a constant  $C > 0$  such that*

$$(2.8) \quad |S_j - \tilde{\lambda}j| \leq C(j \log j)^{1/2} \quad \forall j \geq \log \log m$$

*holds with high probability.*

The lemma is the direct analog of [5, Proposition 3.7]. The heuristics of the proof is the following: The total number of alives at the split times

is just a random walk on  $\mathbb{Z}$  with independent  $\text{Poi}(\tilde{\lambda} + 1) - 1$  increments, but conditioning on survival destroys this independence. However, with a coupling argument we can reduce the error probabilities to the independent case as follows. First condition on the first  $\log \log m$  steps of the walk to stay positive. On this event, the random walk is whp close to its expected value, and hitting zero from this point has exponentially small probability. Then, after the  $\log \log m$ -th step we can forget about the conditioning on survival and work with independent  $\text{Poi}(\tilde{\lambda} + 1) - 1$  increments. For these, a simple large deviation estimate already yields a summable error term in  $j$  for the events in (2.8).

**3. Embedding the BP into the IHRG.** This section relates the exploration process of the neighborhood of a vertex in the inhomogeneous random graph (IHRG) model  $G(n, \kappa)$  of Section 1 and the branching process  $\Psi_\kappa$  of Section 2 to one another. To obtain an embedding, we give each BP-particle a label (= the vertex it corresponds) to get a continuous-time labeled BP (CTLBP) on which we define thinning. This procedure basically deals with the problem of finding the shortest weight path amongst multiple possible paths between any two given vertices.

*3.1. Labeled branching processes.* The first step is to determine for a particle in  $\Psi_\kappa$  to which vertex of the graph  $G(n, \kappa)$  it corresponds. Thus, we describe continuous-time labeled branching processes (CTLBP). By assigning labels to the individuals in  $\Psi_\kappa$  according to their type, we will be able to couple the BP to the exploration of  $G(n, \kappa)$  from a given initial vertex. Fix  $n \geq 1$  and denote the set of labels (vertices of the graph) by  $[n] = \{1, 2, \dots, n\}$ . It is important to distinguish the labels 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  $[n]^{(s)}$ .

The construction goes as follows. Assume that the root of the BP is of type  $s$ . Assign the label  $i_0 \in [n]^{(s)}$  to it. The root immediately dies and gives birth to a random number  $\eta_{st}$  of type  $t$  offspring, for each  $t \in \mathcal{S}$ . By (1.2), we get that the distribution of the number of type  $t$  neighbors of a type  $s$  vertex in  $G(n, \kappa)$  is

$$\eta_{st} \stackrel{d}{=} \text{Bin} \left( n_t - \delta_{st}, \frac{\kappa(s, t)}{n} \right).$$

For  $t \neq s$ , the  $\eta_{st}$  "new" individuals of type  $t$  are assigned different labels from  $[n]^{(t)}$  drawn without replacement uniformly at random. For  $t = s$ ,

we choose the labels the same way from  $[n]^{(s)} \setminus i_0$ . Further, by denoting the number of children of the  $j$ -th dying particle in the BP by  $D_j^{Bin}$ , we see that conditioned on the event that the  $j$ -th split is of a type  $s$  vertex, the distribution of  $D_j^{Bin}$  is a sum of independent random variables

$$(3.1) \quad D_j^{Bin,(s)} = \eta_{s1} + \eta_{s2} + \dots + \eta_{sr}.$$

Since the edge weights are exponential, each offspring lives for an exponentially distributed random time with rate one, then dies and gives birth to its children, thus the natural embedding requires a continuous time multi-type BP with offspring distribution given in (3.1). We denote this BP up to time  $t$  by  $\Psi_\kappa^{Bin}(t)$ .

For the number of alive vertices after the  $m$ -th split we write  $S_m^{Bin} = D_1^{Bin} + \dots + D_m^{Bin} - (m-1)$ . Due to the memoryless property of the exponential distribution, the next vertex ( $m+1$ -st) to split is uniformly distributed among the  $S_m^{Bin}$  alive vertices.

Inductively, the individual that splits at the  $j$ -th split time can be uniformly chosen from the  $S_j^{Bin}$  alive vertices. Assume it is of type  $\tilde{s}$  with label  $i_j$ . It gives birth to  $\eta_{\tilde{s}t}$  offspring of type  $t$  and choose  $\eta_{\tilde{s}t}$  different labels from  $[n]^{(t)}$  for  $t \neq \tilde{s}$  and  $\eta_{\tilde{s}\tilde{s}}$  different labels from  $[n]^{(\tilde{s})} \setminus i_j$  for  $t = \tilde{s}$ .

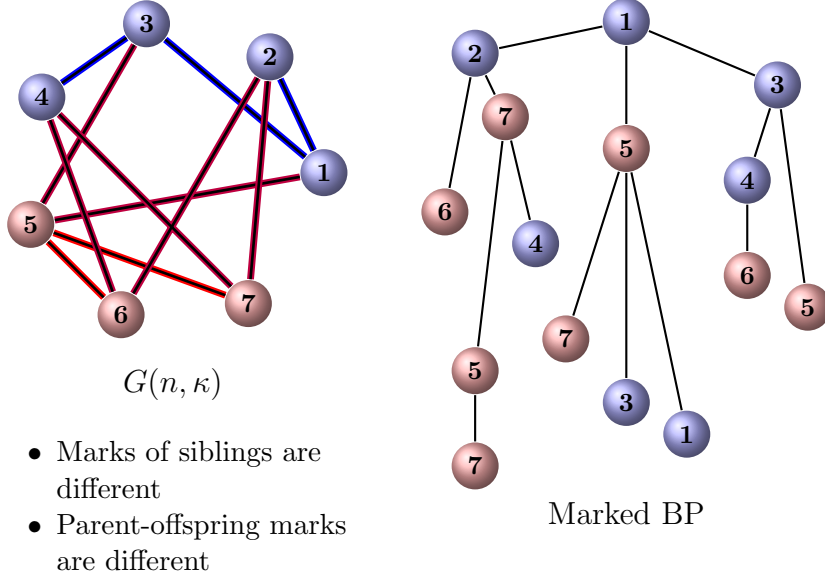
LEMMA 3.1 (Coupling error to Poi offsprings). *The multi-type branching processes  $\Psi_\kappa^{Bin}$  and  $\Psi_\kappa$  (described at the beginning of Section 2) can be coupled until the  $m$ -th split with error*

$$\mathbf{P} [\exists j \leq m, D_j^{Bin} \neq D_j^{Poi}] \leq \frac{m}{n} (\tilde{\lambda} + 1) \max \kappa (1 + o(1)).$$

*As a consequence we immediately get that the time of the  $m$ -th split and the generation of a uniformly picked alive particle in a given type-set  $E \in S$  can be coupled in the two processes with the same upper bound for the error probability.*

PROOF. By the usual coupling of Binomial and Poisson distribution, we can couple  $\eta_{st}$  in (3.1) to  $\xi_{st} \sim \text{Poi}(\lambda_{st})$  distribution with a coupling error of  $n_t(\kappa(s,t)^2/n^2) = \lambda_{st}\kappa(s,t)(1+o(1))/n$ . Under the assumption that  $\max \kappa < \infty$ , the coupling to Poisson offspring distribution of a single split has an error at most  $\max \kappa(\tilde{\lambda} + 1)/n$ . Summing up the error terms up to  $m$  yields the statement. Clearly the error probabilities  $\mathbf{P}[G_m^{Bin,E} \neq G_m^{Poi,E}]$  and  $\mathbf{P}[\tau_m^{Bin} \neq \tau_m^{Poi}]$  can be bounded from above by the error probability that the coupling fails between the two BP-s up to the  $m$ -th step.  $\square$

As a consequence of the lemma, we apply the above described labeling procedure to  $\Psi_\kappa$ , to embed it in  $G(n, \kappa)$ , and the error term stays small until we do  $o(n/(\tilde{\lambda} + 1))$  steps. However, there is another error arising from the embedding: By assigning the labels the above described way, we rule out the possibility of loop edges and multiple edges between vertices during an exploration of  $G(n, \kappa)$ . The error of the coupling may arise from cycles: we have to find the shortest weight path amongst multiple possible paths between any two given vertices. We handle this problem by thinning  $\Psi_\kappa$ .



3.2. *Thinning the branching process.* The notion of thinning is introduced on branching processes to identify the shortest weight path between any two given vertices. The percolating fluid first reaches  $y$  along the shortest path, then other paths are found later whenever the label of  $y$  reappears in the labeling procedure. Thus, we only have to keep track of the first occurrences of each label.

In terms of the CTLBP this means that when we reach a label

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

then we found a cycle in the exploration of the IHRG. The longer paths are irrelevant, thus we delete  $i_k$  and the whole subtree starting from it in  $\Psi_\kappa$ . We call the label  $i_k$  and its subtree thinned. As a result we only keep the shortest weight paths between pairs of vertices. We refer to the resulting process as  $\text{th}\Psi_\kappa$ .



For a fixed  $n$ , the total number of labels is finite ( $= n$ ), so a.s. at some random time all labels will have appeared. This means that  $\text{th}\Psi_\kappa$  dies out a.s., and at this time we found the minimal weight spanning tree of a component of  $G(n, \kappa)$ . 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:

LEMMA 3.2 (FPP on  $G(n, \kappa)$  is thinned CTLBP). *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  $\mathcal{H}_n(i_0, j)$  of the shortest weight path between vertices  $i_0, j \in [n]$  in  $G(n, \kappa)$  is equal in distribution to the weight and hopcount of the shortest weight path between the root  $i_0$  and  $j$  in  $\text{th}\Psi_\kappa$ .*

REMARK 3.3. *We did not use that the edge weights are exponentially distributed. So Lemma 3.2 holds for i.i.d. edge weights with arbitrary continuous distribution supported on  $(0, \infty)$ .*

To make the intuitive picture of colliding flows of fluid precise we formally introduce the notion of shortest-weight trees  $SWT_k$ , for  $k \geq 1$ . Since we cannot follow the progress continuously in time, we keep track of the flows at each split time  $\tau_k$ . 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  $|\mathcal{A}(k)| = S_k$ . Let  $SWT_0 = (\{i_0\}, \tau_0 = 0)$ , and define

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

The CTLBP  $\Psi_\kappa(t)$  can be uniquely reconstructed from the sequence  $(SWT_k)_{k=1}^\infty$ . Note that  $SWT_k$  contains all the labels in  $\Psi_\kappa$ , also the thinned labels and possibly some multiple labels among alive vertices. When we investigate the collision of the two flows from  $x$  and from  $y$  later, we want to avoid the case that this connection happens at thinned vertices in the BP-s. Thus, we will need an upper bound on the proportion of thinned alive labels of a given type.

LEMMA 3.4 (Expected number of thinned alive labels). *Fix  $k \geq 1$  and denote by  $\mathcal{A}^t(k)$  and  $\text{th}\mathcal{A}^t(k)$  the number of alive and thinned alive particles of type  $t \in \mathcal{S}$  after the  $k$ -th split in the CTLBP. Then under Assumption 1.3*

$$(3.3) \quad \mathbf{E} \left[ \frac{\text{th}\mathcal{A}^t(k)}{\mathcal{A}^t(k)} \right] \leq \frac{\tilde{\lambda} + 1}{\tilde{\lambda}} \frac{k}{n} (1 + o(1)).$$

PROOF. We calculate the number of alive thinned vertices of type- $t$  by checking whether 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 get that

$$(3.4) \quad \text{th}\mathcal{A}^t(k) = \sum_{j=1}^k \sum_{s \in \mathcal{S}} |\mathcal{A}_j^{i_j \rightarrow t}(k)| \mathbb{1}[i_j \text{ is thinned} | t(i_j) = s] \mathbb{1}[t(i_j) = s].$$

Let us further introduce

$$\mathcal{A}_j^{u \rightarrow t}(k) = \{v \in \mathcal{A}^t(k) : v \text{ is a descendant of } w \text{ with } w \in \mathcal{A}^u(j+1)\}.$$

For  $|\mathcal{A}_j^{i_j \rightarrow t}(k)|$  we can argue that if the  $j$ -th particle to split was of type- $s$ , it had  $\eta_{su}$  type- $u$  children, then by symmetry and by the memoryless property of the lifetimes we have that

$$\mathbf{E} \left[ \mathcal{A}_j^{i_j \rightarrow t}(k) \mid \eta_{su}, \mathcal{A}^u(j), \mathcal{A}_j^{u \rightarrow t}(k) \right] = \sum_{u \in \mathcal{S}} \frac{\eta_{su}}{|\mathcal{A}^u(j+1)|} |\mathcal{A}_j^{u \rightarrow t}(k)|.$$

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) can be bounded from above by

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

Recall that  $N_k^{(t)}$  denotes the number of splits of type- $t$  vertices among the first  $k$  splits and there are  $n_t$  different marks that correspond to vertices of type- $t$ . The conditional probability in  $(*)$  can be bounded simply by  $N_k^s/n_s$ , while  $(\diamond)$  equals  $S_j^t/S_j$ . Thus

$$\begin{aligned} \mathbf{E} \left[ \frac{\text{th}\mathcal{A}^t(k)}{|\mathcal{A}^t(k)|} \right] &\leq \sum_{j=1}^k \sum_{s, u \in \mathcal{S}} \mathbf{E} \left[ \frac{\eta_{su}}{S_j^u} \frac{|\mathcal{A}_j^{u \rightarrow t}(k)|}{|\mathcal{A}^t(k)|} \right] \frac{N_j^s}{n_s} \frac{S_j^t}{S_j} \\ &= \sum_{j=1}^k \sum_{s, u \in \mathcal{S}} \frac{\kappa(s, u) \mu_u}{\tilde{\lambda}(j+1) \pi_u} \frac{|\mathcal{A}_j^{u \rightarrow t}(k)|}{|\mathcal{A}^t(k)|} \frac{\pi_s^2}{\mu_s} \frac{j}{n} (1 + o(1)), \end{aligned}$$

where we used (1.5), (2.2), (2.3) and (2.4). Under Assumption 1.3  $\pi = \mu$  (see (1.8)) and using the symmetry of  $\kappa$  the above expression simplifies to

$$\frac{1}{\tilde{\lambda}} \sum_{j=1}^k \frac{1}{n} \sum_{u \in \mathcal{S}} \frac{|\mathcal{A}_j^{u \rightarrow t}(k)|}{|\mathcal{A}^t(k)|} \underbrace{\sum_{s \in \mathcal{S}} \kappa(u, s) \mu_s}_{\tilde{\lambda}+1} (1 + o(1)) = \frac{\tilde{\lambda} + 1}{\tilde{\lambda}} \frac{k}{n} (1 + o(1)).$$

□

We show a similar result to the assertion of Lemma 3.4 for the number of multiple labels among  $\mathcal{A}(k)$ . We need to guarantee that the number of different labels in the set  $\mathcal{A}_k$  is approximately the same as the size of the set,  $S_k$ . More precisely,

LEMMA 3.5 (Expected number of multiple labels). *For all  $k \geq 1$  and  $t \in \mathcal{S}$ , the number of different labels of alive vertices after the  $k$ -th split*

$$(3.5) \quad |\mathcal{A}^t(k)| = S_k^t \left( 1 - \frac{\tilde{\lambda}\pi_t k}{2\mu_t n} \right).$$

PROOF. When assigning labels to new vertices of type  $t$  we prescribed some constraints on the set  $[n]^{(t)}$  from which we choose its label. We can dominate this by repeatedly choosing from  $[n]^{(t)}$  without any constraints. From (2.2) we know that  $S_k^t = \tilde{\lambda}\pi_t k(1+o(1))$ . So let us sample  $\tilde{\lambda}\pi_t k(1+o(1))$  marks from  $[n]^{(t)}$  with replacement. For  $i \in [n]^{(t)}$  let  $X_i$  be the number of times  $i$  was chosen, thus  $X_i \sim \text{Bin}(\tilde{\lambda}\pi_t k(1+o(1)), 1/n_t)$ . The probability that  $i$  is chosen at least twice is

$$\mathbf{P}[X_i \geq 2] = 1 - \mathbf{P}[X_i = 0] - \mathbf{P}[X_i = 1] = \frac{\tilde{\lambda}^2 \pi_t^2 k^2}{2 n_t^2} + O\left(\frac{k^3}{n_t^3}\right).$$

Thus for the expected number of type- $t$  multiple labels

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

The claim 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 rigorously examine the intuitive picture of colliding flows. For technical reasons we do not let the fluids flow simultaneously from both vertices, rather we let the fluid flow from  $x$  until it reaches some  $a_n = o(n)$  vertices, then we "freeze" it, and start a flow from  $y$  until the random time of connection, i.e. when the two flows collide.

The exploration process from  $x$  until the split time  $\tau_{a_n}^x$  is coded in  $SWT_{a_n}^x$  (see (3.2)). Afterwards, the flow from  $y$  can only connect to the flow from  $x$  via an alive vertex in  $SWT_{a_n}^x$ . So when assigning the labels to the vertices in the BP from  $y$  we must leave out the labels  $\mathcal{D}^x(a_n)$  from the possible labels

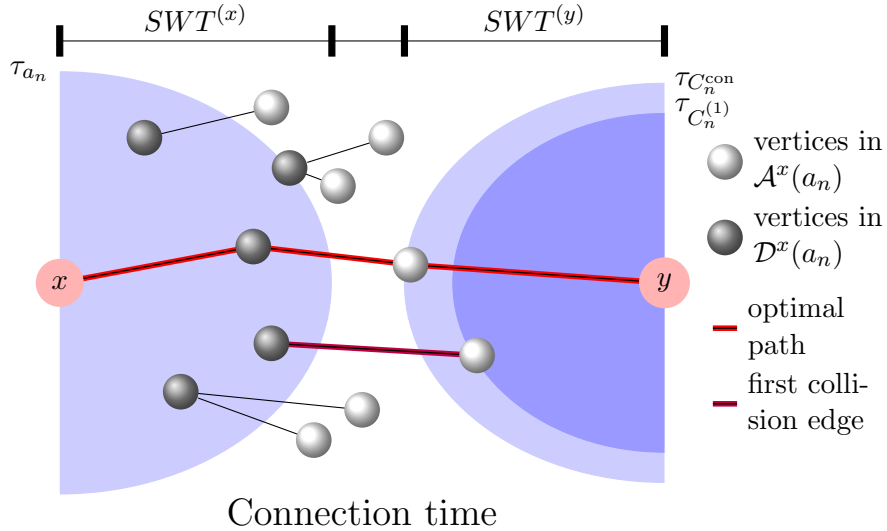
[ $n$ ]. A possible collision edge appears when a label from  $\mathcal{A}^x(a_n)$  appears among the labels in  $\mathcal{D}^y(k)$ . The first possible collision edge appears at split  $C_n^{(1)} = \min\{k \geq 0 : \mathcal{A}^x(a_n) \cap \mathcal{D}^y(k) \neq \emptyset\}$  at time  $\tau_{C_n^{(1)}}^y$ . The  $i$ -th appears at split

$$(4.1) \quad C_n^{(i)} = \min\{k \geq C_n^{(i-1)} : |\mathcal{A}^x(a_n) \cap \mathcal{D}^y(k)| = i\}, \text{ at time } \tau_{C_n^{(i)}}^y.$$

Thus 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 possible collision edge after time  $\tau_{a_n}^x$ . From the memoryless property of the weights it follows that  $E_i \stackrel{d}{=} \text{Exp}(1)$ . So the actual connection happens through the possible collision edge that minimizes the expression  $\tau_{C_n^{(i)}}^y + E_i$ . Thus the shortest weight path equals

$$(4.2) \quad \mathcal{P}_n = \tau_{a_n}^x + \min_i \left\{ \tau_{C_n^{(i)}}^y + E_i \right\}.$$

Let us denote the split which minimizes the above expression by  $C_n^{\text{con}}$ . The figure below illustrates the connection time.



To be able to determine the distribution of the minimum in (4.2), we need a handle on the size of  $C_n^{(i)}$ . The following proposition, roughly speaking, states that all the possible collision edges appear at  $O(n/a_n)$  time with some random constant.

**PROPOSITION 4.1** (PPP limit of collision edges). *Denote a homogeneous Poisson Point Process 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 CTLBPs survive, the point process

$$\left\{ \frac{C_n^{(i)} a_n}{n} \right\}_i \xrightarrow{d} \text{PPP}(\hat{\lambda}) \text{ as } n \rightarrow \infty,$$

where  $\hat{\lambda}$  simplifies to rate  $\tilde{\lambda}$  under Assumption 1.3.

PROOF. We first show by induction that for fix  $n$

$$(4.3) \quad \frac{C_n^{(i)} a_n}{n} \stackrel{d}{=} \text{Gamma}(i, \hat{\lambda}_n), \text{ where}$$

$$\hat{\lambda}_n = \frac{n}{a_n} \sum_{s \in \mathcal{S}} \frac{S_{a_n}^{x,s}}{n_s} \pi_s (1 + o(1)) + o(1).$$

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

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

where  $\mathcal{F}_{j-1}$  is the  $\sigma$ -algebra generated by  $SWT_{a_n}^x$  and  $SWT_{j-1}^y$ . To calculate  $\mathbf{P} [C_n > j \mid C_n > j-1, \mathcal{F}_{j-1}]$  we have to sum over the types in  $SWT^y$  and find the probability that it does not connect to an alive vertex in  $SWT^x$  of the same type. Below we use the result of Lemma 3.5, and we substitute the number of different marks in  $\mathcal{A}^{x,t}(a_n)$  by  $S_{a_n}^{x,t}$ , and neglect the error factor of order  $(1 - (\tilde{\lambda} a_n)/n)$  along the lines. This error can be included in the  $o(1)$  term of the last line of the display below.

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

where  $(*)$  equals  $\pi_t(1+o(1))$  by (2.4). So  $\frac{C_n^{(1)} a_n}{n} \stackrel{d}{=} \text{Exp}(\hat{\lambda}_n) = \text{Gamma}(1, \hat{\lambda}_n)$ .

From the induction hypothesis

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

Differentiating the cdf of  $C_n^{(i+1)} a_n/n$  with respect to  $x$  yields the pdf of Gamma( $i+1, \hat{\lambda}_n$ ), which proves (4.3). 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 (1.5) and (2.2) it follows that  $\hat{\lambda}_n \xrightarrow{a.s.} \hat{\lambda}$  as  $n \rightarrow \infty$ . The weak convergence now immediately follows.  $\square$

This result shows that when finding the shortest path it makes no difference to let the fluids flow simultaneously or to delay one of them. When they flow simultaneously, the number of vertices explored by both flows is of order  $\sqrt{n}$  which is optimal in the sense that in every other case the explored vertices are of larger magnitude.

We will see in Subsection 5.1 that to determine the distribution of  $\tau_{C_n^{\text{con}}} + E_{\text{con}}$  we need the following.

LEMMA 4.2. *Let  $(P_i)_i$  denote the points of a PPP(1) process and independent of that,  $E_i \stackrel{d}{=} \text{Exp}(1)$ , also independent. Then*

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

where  $X$  follows a standard Gumbel distribution, i.e.  $\mathbf{P}[X \leq x] = e^{-e^{-x}}$ .

PROOF. For convenience let  $X_i \sim \frac{1}{\tilde{\lambda}} \log P_i$ . We will calculate the tail

distribution of the minimum by conditioning on the Poisson points first:

$$\begin{aligned} \mathbf{P} \left[ \min_i X_i + E_i \geq z \right] &= \mathbf{E} [\mathbf{P} [\forall i X_i + E_i > z | X_1, X_2 \dots]] \\ &= \mathbf{E} \left[ \prod_i e^{-(z-X_i)_+} \right] \\ &= \mathbf{E} \left[ \prod_{i: P_i < e^{\tilde{\lambda}z}} e^{-z(P_i)^{\frac{1}{\tilde{\lambda}}}} \right] \end{aligned}$$

Now, 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 this number  $Z$ , the points  $P_i, i \leq Z$  are independent and uniform in the interval  $[0, e^{\tilde{\lambda}z}]$ . Thus, we can calculate the expected value of the product on the right hand side of the previous display as follows:

$$\begin{aligned} \mathbf{E} \left[ \prod_{i: P_i < e^{\tilde{\lambda}z}} e^{-z(P_i)^{\frac{1}{\tilde{\lambda}}}} \right] &= \mathbf{E} \left[ \mathbf{E} \left[ e^{-zU_i^{\frac{1}{\tilde{\lambda}}}} \right]^Z \right] \\ &= \mathbf{E} \left[ \left( \frac{\tilde{\lambda}}{\tilde{\lambda}+1} \right)^Z \right] = \exp \left\{ e^{\tilde{\lambda}z} \left( 1 - \frac{\tilde{\lambda}}{\tilde{\lambda}+1} \right) \right\}. \end{aligned}$$

It is easy to see that if  $X$  is a standard Gumbel random variable with  $\mathbf{P}(X \leq x) = e^{-e^{-x}}$ , then  $\mathbf{P}(-aX + b > x) = e^{-e^{x/a}e^{-b/a}}$ . Thus, here with  $a = \frac{1}{\tilde{\lambda}}$  and  $b = \frac{1}{\tilde{\lambda}} \log(\tilde{\lambda}+1)$  we get the claim.  $\square$

We finish the section with a result that shows that the index  $i$  where the minimum is taken is stochastically bounded by a Geometric distribution of parameter  $1/(\tilde{\lambda}+1)$ .

LEMMA 4.3. *The probability that the shortest weight path is not among the first  $k$  collision edges is decaying exponentially in  $k$ , i.e.*

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

As a consequence of the lemma we immediately get that the distribution of the rescaled connection time is stochastically dominated by a sum of independent exponentials with parameter  $i$  up to a geometric random variable

independent of them, namely

$$(4.4) \quad P_n^{\text{con}} := \frac{a_n C_n^{\text{con}}}{n} \leq \sum_{i=1}^N \tilde{E}_i,$$

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

**PROOF OF LEMMA 4.3.** We use again the notation  $X_i \sim \frac{1}{\tilde{\lambda}} \log(P_i)$ , where  $P_i$  is the  $i$ -th point in a  $\text{PPP}(1)$  process. To show that the minimum is taken at an index at least  $k+1$ , we will condition on the value of the minimum ( $=z$ ) and also on the value of the  $X_{k+1} = c$  with  $z \geq c$ . Thus we have

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

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  $U_i, i = 1, \dots, k$  independent uniform points on  $[0, e^{\tilde{\lambda}c}]$ . Thus, the expectation can be calculated as follows:

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

Then clearly we have

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

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



**5. Proof of main results.** We begin the section with the proof of Theorem 1.4 for finite-type graphs. We continue with the discussion of approximating kernels and then prove Theorem 1.4 in the general setting. The section is concluded with the proof of Theorem 1.6. The proofs are analogous to the ones in [5]. The idea of using approximating kernels comes from [9, Section 7].

5.1. *Proof of Theorem 1.4: finite-type setting.* Let  $(\mathcal{S}, \mu)$  be an arbitrary finite-type ground space and  $\kappa$  a kernel that satisfies Assumption 1.3. We first argue that the probability that the shortest weight path contains thinned vertices -i.e. it is not a real shortest path - is  $o(1)$ . 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  $x$ .

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

$$\sum_{t \in \mathcal{S}} (\mathbf{P} [v \in \text{th}\mathcal{A}^{x,t}(a_n)] + \mathbf{P} [v \in \text{th}\mathcal{A}^{y,t}(C_n^{\text{con}})]) \mathbf{P} [\text{con}, t^\dagger].$$

Using Lemma 3.4 with  $k = a_n$  and  $k = C_n^{\text{con}}$  respectively yields that this term equals

$$\begin{aligned} & \sum_{t \in \mathcal{S}} \frac{S_{C_n^{\text{con}}}^{y,t}}{S_{C_n^{\text{con}}}^{x,t}} \left( \frac{\tilde{\lambda} + 1}{\tilde{\lambda}} \left( \frac{a_n}{n} + \frac{C_n^{\text{con}}}{n - a_n} \right) (1 + o(1)) \right) \\ &= \frac{\tilde{\lambda} + 1}{\tilde{\lambda}} \left( \frac{a_n}{n} + \frac{n}{n - a_n} \frac{P_n^{\text{con}}}{a_n} \right) (1 + o(1)) \rightarrow 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

We divide in the flow of  $y$  by  $n - a_n$  since the labels are chosen from  $[n] \setminus \mathcal{D}^x(a_n)$ . Further note that by Lemma 3.5, in the proof of Lemma 3.4  $|\mathcal{A}^{x,t}(a_n)|$  can be replaced by  $S_{a_n}^{x,t} = \pi_t a_n (1 + o(1))$ .

Then, to get the second line we used that  $a_n C_n^{\text{con}}/n \sim P_n^{\text{con}}$ , is tight by (4.4). This shows that if  $a_n/n$  and  $1/a_n$  are both tending to zero, - i.e. for all  $a_n = o(n)$ , the shortest weight path whp. does not contain a thinned vertex.

Now we turn to determine the distribution of the shortest weight path. We know from (4.2) that

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

In  $\Psi_\kappa(t)$ , the rescaled total number of alive individuals,  $M_t = e^{\tilde{\lambda}t} |\underline{S}(t)|$  is a

martingale. Thus,  $\tau_k$  can be expressed as

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

Applying this formula to the minimum in (5.1) we have

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

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

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

For the last term we can apply Lemma 4.2 to get

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

with  $X$  denoting a standard Gumbel random variable. We can also use that under Assumption 1.3  $\hat{\lambda} = \tilde{\lambda}$ . Further, applying (5.2) and (5.3) to the expressions in (5.1), we arrive at

$$\mathcal{P}_n = \frac{1}{\tilde{\lambda}} \log n - \frac{1}{\tilde{\lambda}} \log \hat{W}^x \hat{W}^y - \frac{1}{\tilde{\lambda}} X + \frac{1}{\tilde{\lambda}} \log \left( \tilde{\lambda}(\tilde{\lambda} + 1) \right),$$

with  $X$  being a standard Gumbel random variable,  $\hat{W}^z$ ,  $z = x, y$  denoting the limiting random variable of the (independent) martingales  $e^{-\lambda t} |S(t)^z|$  in  $\Psi_\kappa^z$ ,  $z = x, y$  conditioned on non-extinction. We know that  $W^i > 0$  a.s. on non-extinction of the processes, so these quantities are well-defined.

Now we turn to the derivation of the limit theorem for the hopcount  $\mathcal{H}_n$ . We start first 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

$a_n$ .  $(\mathcal{A}^{x,t}(a_n), \mathcal{D}^{x,t}(a_n))$  denotes the labels of the alive and dead vertices of type  $t$ , respectively. Now the evolution of the flow of  $y$  conditioned on  $SWT^x(a_n)$  is as follows: leaving out the labels in  $\mathcal{D}^x(a_n)$ , 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 of the corresponding type.

LEMMA 5.1. *We have the following statements given the event of*

$\{ \text{Collision happens from } SWT^{(y)} \text{ to } SWT^{(x)} \text{ at the } C_n\text{-th split at a type } t \text{ vertex} \}$ ,

1. *the label  $v_C$  at which this happens is uniform among all labels in  $\mathcal{A}_{a_n}^{x,t}$ .*
2. *Further, the hopcounts  $G_{a_n}^{x,t}$ ,  $G_{C_n}^{y,t}$  are independent given that the collision happens at a type  $t$  vertex.*
3. *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 in the processes  $\mathcal{A}_{a_n}^x$  and  $\mathcal{A}_{C_n}^y$ .*

PROOF. The first two statement of the Lemma is 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 do the following procedure: in the  $k$ -th step we draw  $d_k$  balls without replacement, add the label of type  $A$  and type  $U$  balls to sets  $L_A$  and  $L_U$ , respectively, and then put them back top the urn. Thus,  $L_A(k)$  consists of all the labels of type- $A$  balls which has been drawn before or at 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

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

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 says 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 possible collision edges between the processes  $x$  and  $y$  are established such that in each step  $k$  with some probability  $p_k$  (which is the probability that the next dying particle in  $\Psi_\kappa^y$  is of type  $t$ ), we pick a uniform label among  $L_A(k) \cup L_U(k)$ , and check if it is of type  $A$ . Clearly, whenever this holds true, a possible collision edge is formed in the two shortest weight trees. Also, it is clear from the previous argument that conditioned on the picked label to be in  $L_A(k)$ , the label of it is a uniformly picked label among  $\mathcal{A}_{a_n}^{x,t}$ , (and clearly also uniform in  $L_A(k)$ ). Further, 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  $SWT^x$  and in  $SWT^y$  are independent and equal to the generation of a uniformly picked alive individual of type  $t$ .  $\square$

Now we are ready to 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  $C_n^{\text{con}}$ . Then we have

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

Thus

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

First use 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). Further, recall the the distributional bound on  $P_n^{\text{con}} = C_n^{\text{con}} a_n / n$  in (4.4) to see that the last term tends to zero. Lemma 5.1 ensures the independence of the two limiting normal variables, thus we get the following distributional limit of the right hand side of the last display:

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

Combining this with Lemma 4.3 and (4.4) we get that the last term vanishes as  $n \rightarrow \infty$  and the variance of the normal distribution is also tending to 1.

5.2. *Approximation of kernels.* For a sequence of partitions of  $\mathcal{S}$  a sequence of regular finitary approximating kernels will be defined, each satisfying (AG). In the regular finitary case we may assume that the type space  $\mathcal{S}$  is finite (the regular finitary case and the finite-type case differ only in

notation). Furthermore, we can assume that  $\mu_s > 0$  for every  $s \in \mathcal{S}$ , however an argument is needed. We cannot simply ignore such types, since measure zero sets can alter  $G(n, \kappa)$  significantly. We can argue as follows, [9].

Suppose that  $\mu_s = 0$  for some  $s \in \mathcal{S}$ . Start by redefining the kernel  $\kappa'(s, t) = \kappa'(t, s) := \max \kappa$  for every  $t \in \mathcal{S}$  and leaving it alone otherwise. Then define a new probability measure  $\mu'$  by shifting some small mass  $\eta$  over to  $s$  from the other types. Clearly  $\mu'_t > 0$  for every  $t \in \mathcal{S}$ . Possibly the types of some of the vertices changes, so change them correspondingly. This way we obtain a vertex space  $\nu' = (\mathcal{S}, \mu', (\mathbf{x}'_n)_{n \geq 1})$  with kernel  $\kappa'$ . It is not hard to see that we can couple  $G(n, \kappa)$  and  $G'(n, \kappa')$  so that  $G(n, \kappa) \subseteq G'(n, \kappa')$ . Finally, letting  $\eta \rightarrow 0$ , the norm of  $T_{\kappa'}$  with respect to  $\mu'$  tends to the norm of  $T_{\kappa'}$  with respect to  $\mu$ , which is equal to the norm of  $T_\kappa$ , since  $\kappa = \kappa'$  a.e. Iterating for other measure-0 types, we can see that it suffices to consider cases where  $\mu_t > 0 \forall t \in \mathcal{S}$ .

We continue with the definitions of the approximating kernels. 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  $i_m(x)$  as the element of  $\alpha_m$  in which  $x$  falls, formally  $x \in A_{m, i_m(x)}$ . 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 our metric space  $\mathcal{S}$ . Lemma 7.1 of [9] states that for any ground space  $(\mathcal{S}, \mu)$  there exists a sequence of finite partitions of  $\mathcal{S}$  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.  $x \in \mathcal{S}$ ,  $\text{diam}(A_{m, i_m(x)}) \rightarrow 0$ , as  $m \rightarrow \infty$ .

For such a sequence of partitions we can define a sequence of approximations of  $\kappa$  by taking its average on each  $A_{mi} \times A_{mj}$ :

$$(5.4) \quad \bar{\kappa}_m(x, y) := \frac{1}{\mu(A_{m, i_m(x)}) \cdot \mu(A_{m, i_m(y)})} \iint_{A_{m, i_m(x)} \times A_{m, i_m(y)} \kappa(s, t) d\mu(s) d\mu(t).$$

If  $\kappa$  is continuous a.e. then property 3) implies that  $\bar{\kappa}_m(x, y) \rightarrow \kappa(x, y)$  for a.e. every  $(x, y) \in \mathcal{S}^2$ . To be able to apply our theorems for finite-type kernels we need to guarantee that Assumption 1.3 holds for all  $\bar{\kappa}_m$ . Thus, considering  $\bar{\kappa}_m$  as a finite-type kernel with respect to the partition  $(A_{m1}, \dots, A_{mM_m})$ , the row sums of  $\bar{\kappa}_m$  weighted by  $\mu$  must be equal to some constant  $c_m$ . In fact easy calculations show that independently of the partition sequence or the ground space, only using assumption (AG) and the fact that  $\kappa$  is symmetric, this holds with  $c_m \equiv \tilde{\lambda} + 1$ .

5.3. *Proof of Theorem 1.4: general setting.* The extension of the proof for general  $(\mathcal{S}, \mu)$  goes with usual discretization techniques, however, one must be careful with the error terms to maintain the distributional convergence.

Let  $(\mathcal{S}, \mu)$  be an arbitrary ground space and kernel  $\kappa$  satisfies the conditions of the theorem. These define the sequence of random graphs  $(G(n, \kappa))_{n \geq 1}$ . Take any sequence of finite partitions  $\mathcal{P}_m = \{A_{m1}, \dots, A_{mM_m}\}$ ,  $m \geq 1$ , that satisfy properties 1), 2) and 3) described in Subsection 5.2. For each  $m$ , consider the finite type approximating kernel  $\bar{\kappa}_m$  defined in (5.4), with ground space  $(\mathcal{S}_m, \mu)$  (where  $|\mathcal{S}_m| = M_m$ ). As a result we obtain 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 the cardinality of  $\mathcal{S}_m$ , 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.

To prove the results we let  $n$  and  $m$  tend to  $\infty$  simultaneously in a carefully chosen way. For fixed  $m$ , from the proof of [7, Lemma 2.1 and Theorem 3.1] it is easy to see that

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

where  $C(\tilde{\lambda})$  is a  $\tilde{\lambda}$ -dependent constant. Thus, with the choice  $a_n = \sqrt{n}$  we get the error of order  $1/\sqrt{\log n}$ .

To be able to couple the graphs  $G(n, \kappa)$  and  $G(n, \bar{\kappa}_m)$ , we need a fine relation between  $\kappa(x, y)$  and  $\bar{\kappa}_m(x, y)$ . Since  $\kappa$  is uniformly continuous,  $\exists \varepsilon_m$  s.t. for all  $x, y$ , and all  $(u, v) \in A_{m, i_m(x)} \times A_{m, i_m(y)}$ :

$$\bar{\kappa}_m(u, v) \leq \kappa(x, y)(1 \pm \varepsilon_m) \quad \text{if } |u - x| < \delta_m \text{ and } |v - y| < \delta_m,$$

where  $\text{diam}(A_{m, i_m(x)}) < \delta_m$  and  $\text{diam}(A_{m, i_m(y)}) < \delta_m$ . Abbreviate  $A_{m, i_m(x)}$  by  $A_{mx}$ . In  $G(n, \kappa)$  the edge-probability between two vertices of types  $x$  and  $y$  is  $\kappa(x, y)/n$ , while in  $G(n, \bar{\kappa}_m)$ , between types  $A_{mx}$  and  $A_{my}$  this probability is  $\bar{\kappa}_m(x, y)/n \in \kappa(x, y)(1 \pm \varepsilon_m)$ . Thus

$$\mathbf{P} [\mathbb{1} [\{x, y\} \in e(G(n, \kappa))] \neq \mathbb{1} [\{A_{mx}, A_{my}\} \in e(G(n, \bar{\kappa}_m))] ] \leq \frac{2\varepsilon_m}{n}.$$

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

$$(5.6) \quad \mathbf{P} [e(G(n, \kappa)) \neq e(G(n, \bar{\kappa}_m))] \leq \frac{2n^2\varepsilon_m}{2n} = n\varepsilon_m.$$

For a fix  $m$ . The  $\delta_m$  and uniform continuity of  $\kappa$  defines  $\varepsilon_m$ . Let

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

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

$$\mathbf{P} \left[ \mathcal{H}_n \neq \mathcal{H}_n^{m(n)} \right] \leq 1/\sqrt{\log n} = o(1).$$

Combining this error bound with the one in (5.5) we obtain that

$$\begin{aligned} & \mathbf{P} \left[ \frac{\mathcal{H}_n - \frac{\tilde{\lambda}+1}{\lambda} \log n}{\sqrt{\frac{\tilde{\lambda}+1}{\lambda} \log n}} < x \right] \\ &= \mathbf{P} \left[ \frac{\mathcal{H}_n^{m(n)} - \frac{\tilde{\lambda}+1}{\lambda} \log n}{\sqrt{\frac{\tilde{\lambda}+1}{\lambda} \log n}} < x \mid \mathcal{H}_n = \mathcal{H}_n^{m(n)} \right] (1 - o(1)) + \mathbf{P} \left[ \mathcal{H}_n \neq \mathcal{H}_n^{m(n)} \right] \\ &= \Phi(x) \left( 1 - O \left( \sqrt{\log n}^{-1} \right) \right) + C(\tilde{\lambda})/\sqrt{\log n} + 1/\sqrt{\log n}. \end{aligned}$$

Finally letting  $n \rightarrow \infty$  (thus  $m(n) \rightarrow \infty$  also), we obtain the desired result for the hopcount.

Now we turn to the proof of the convergence of the shortest weight path. To avoid conflicting notation we will denote  $P_n(\kappa)$  the shortest weight path belonging to  $G(n, \kappa)$ . We can use the same coupling argument as for the hopcount to get the estimate

$$\mathbf{P} \left[ P_n(\kappa) \neq P_n(\bar{\kappa}_{m(n)}) \right] \leq \frac{1}{\sqrt{\log n}}$$

We know from the finite type case that

$$P_n(\bar{\kappa}_{m(n)}) - \frac{1}{\lambda} \log n \xrightarrow{d} -\frac{1}{\lambda} \hat{W}_{(m(n))}^x \hat{W}_{(m(n))}^y - \frac{1}{\lambda} X,$$

where  $X$  is a standard Gumbel variable,  $\hat{W}_{(m(n))}^i, i = x, y$  is i.i.d. random variables, distributed as the limit of the martingales arising from the branching processes with kernel  $\bar{\kappa}_{m(n)}$ , conditioned on being positive. Since for all  $m$ , the row sums of  $\bar{\kappa}_m$  equals  $\tilde{\lambda}$  i.e. each particle has a  $\text{Poi}(\tilde{\lambda})$  total number of children, it is not hard to see that the limit of the martingales

$$W_{(m(n))} \stackrel{d}{\equiv} W.$$

This finishes the proof of the distributional convergence of the shortest weight path.

REMARK 5.2. Besides Lemma 2.4, this is the other spot where the generalization for  $\kappa$  not satisfying Assumption (AG) would fail: to get the distributional convergence without Assumption (AG), we should show that  $W_{(m(n))} \rightarrow W$ . But the relation of the limits of the approximating Branching Process martingales is not clear at this point to us.

5.4. *Proof of Theorem 1.6: dense setting.* In the dense setting, where  $\tilde{\lambda}_n \rightarrow \infty$ , we have a sequence of kernels  $\kappa_n$ ,  $n = 1, 2, \dots$ . The type  $t$  neighbors of a type  $s$  vertex have distribution  $\eta_{st}^{(n)} \stackrel{d}{=} \text{Bin}(n_t - \delta_{st}, \kappa_n(s, t)/n)$ . We avoid the coupling with Poisson random variables done in Lemma 3.1 by immediately applying the CLT result of [17] to  $\Psi_{\kappa_n}$  where the offspring distribution  $(D_i^{(n)} | \text{type } s \text{ splits})$  is the sum of independent binomial random variables  $\eta_{st}^{(n)}$ ,  $t \in \mathcal{S}$ . We will apply a similar argument than the one in the general case. Namely, we get that in  $\Psi_{\kappa_n}$ , for a uniformly picked type- $t$  individual at step  $k$

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

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 1.6.

The centering constant can be replaced by  $\log n$  if and only if  $(\frac{\tilde{\lambda}_n + 1}{\lambda_n} - 1)\sqrt{\log n} \rightarrow 0$ , or equivalently  $\sqrt{\log n} = o(\tilde{\lambda}_n)$ .

To prove the part concerning the length shortest-weight path, with the limit taken diagonally, we need to be a bit more careful to determine 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 just the minimum of that many independent exponential random variables, we have for every  $m$

$$(5.7) \quad \tau_m \stackrel{d}{=} \sum_{i=1}^m E_i / S_i^{(n)},$$

with  $E_i$  i.i.d.  $\text{Exp}(1)$ . Now let  $\xi^{(1)}, \dots, \xi^{(r)}, \xi_1, \xi_2, \dots$  denote independent standard normal random variables. Recall that  $D_i^{(n)}$  denotes the number of



children of the  $i$ -th dying particle. Then, by the CLT we have

$$(D_i^{(n)} | \text{type } s \text{ splits}) \xrightarrow{d} \sum_{t \in \mathcal{S}} \left( \lambda_{st}^{(n)} + \sqrt{\lambda_{st}^{(n)}} \cdot \xi^{(t)} \right) \stackrel{d}{=} \tilde{\lambda}_n + 1 + \sqrt{\tilde{\lambda}_n + 1} \cdot \xi_i.$$

From here, with the usual notation  $S_i^{(n)} = \sum_{j=1}^i D_j^{(n)} - (i-1)$ :

$$S_i^{(n)} \xrightarrow{d} i\tilde{\lambda}_n + 1 + \sqrt{i(\tilde{\lambda}_n + 1)} \cdot \tilde{\xi}_i.$$

Applying this result to (5.7) yields  $\tilde{\lambda}_n \tau_m \approx \sum_{i=1}^m E_i/i$ .

Notice that the sequence  $E_m/m, E_{m-1}/m-1, \dots, E_1/1$  gives in distribution the spacings of the exponential random variables  $E_1, \dots, E_m$ . So the sum  $\sum_{i=1}^m E_i/i$  is equal in distribution to a random variable  $B_m$  that is the maximum of  $m$  independent exponentially distributed random variables with rate 1. The distribution function of  $B_m$  is

$$\mathbf{P}[B_m \leq x] = (1 - e^{-x})^m.$$

From here we find that

$$\begin{aligned} \mathbf{P}\left[\tilde{\lambda}_n \tau_{a_n}^{(x)} - \log a_n \leq x\right] &= \mathbf{P}[B_{a_n} \leq x + \log a_n] = \left(1 - e^{-(x + \log a_n)}\right)^{a_n} \\ &= \exp\left(-e^{-x} + O(1/a_n)\right) \rightarrow \exp(-e^{-x}) = \Lambda(x), \end{aligned}$$

where  $\Lambda$  denotes the distribution function of a standard Gumbel random variable. Similarly to the proof in Section 5.1, we conclude that

$$(\tilde{\lambda}_n \tau_{a_n}^{(x)} - \log a_n, \tilde{\lambda}_n \tau_{C_n^{\text{con}}}^{(y)} - \log C_n^{\text{con}}) \xrightarrow{d} (Y_1, Y_2),$$

where  $Y_1$  and  $Y_2$  are two independent copies of a standard Gumbel random variable. From Theorem (5.1) we get that by the continuous mapping theorem

$$\begin{aligned} \tilde{\lambda}_n \mathcal{P}_n - \log n &\stackrel{d}{=} \underbrace{\tilde{\lambda}_n \tau_{a_n}^{(x)} - \log a_n}_{\xrightarrow{d} Y_1} - \log \frac{n}{a_n} \\ (5.8) \quad &+ \min_i \left\{ \underbrace{\tilde{\lambda}_n \tau_{C_n^{(i)}}^{(y)} - \log C_n^{(i)}}_{\xrightarrow{d} Y_2} + \tilde{\lambda}_n \left( \frac{1}{\tilde{\lambda}_n} \log C_n^{(i)} + E_i \right) \right\}. \end{aligned}$$

Now using Proposition 4.1 to see that  $C_n^{(i)} = n/a_n P_n^{(i)}$ , with  $P_n^{(i)}$  PPP( $\hat{\lambda}_n$ ) points, and then Lemma 4.2 yields that the last term in the minimum equals

$$\tilde{\lambda}_n \min_i \left\{ \frac{1}{\tilde{\lambda}_n} \log C_n^{(i)} + E_i \right\} = \log \frac{n}{a_n} - X_3 + \log \frac{(\tilde{\lambda}_n + 1)}{\hat{\lambda}_n},$$

with  $X_3$  a standard Gumbel variable. Then, under Assumption 1.3  $\hat{\lambda}_n = \tilde{\lambda}_n$ , thus the last term vanishes in the limit. Combining this with (5.8) finishes the proof.

## References.

- [1] K.B. Athreya and P.E. Ney. *Branching Processes*. Dover Publications, 2004.
- [2] S. Bhamidi. First passage percolation on locally tree-like networks. i. dense random graphs. *Journal of Mathematical Physics*, 49(12):125218, 2008.
- [3] S. Bhamidi and R. van der Hofstad. Weak disorder asymptotics in the stochastic mean-field model of distance. *Ann. Appl. Probab.*, 22(1):29–69, 2012.
- [4] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.*, 20(5):1907–1965, 2010.
- [5] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on the Erdős-Rényi random graph. *Combinatorics, Probability and Computing*, 20:683–707, 2011.
- [6] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. Universality for first passage percolation on sparse random graphs. arXiv:1210.6839 [math.PR], 2012.
- [7] W.J. Bühler. Generations and degree of relationship in supercritical markov branching processes. *Probability Theory and Related Fields*, 18(2):141–152, 1971.
- [8] B. Bollobás and W. F. de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.
- [9] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct Algor*, 31:3–122, 2007.
- [10] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, 6:125–145, 2002.
- [11] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Math.*, 1:91–113, 2003.
- [12] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures and Algorithms*, 31:482–516, 2007.
- [13] Marc Lelarge Hamed Amini. The diameter of weighted random graphs. arXiv:1112.6330 [math.PR], 2011.
- [14] C. Howard. Models of first-passage percolation in probability on discrete structures. pages 125–173, 2004.
- [15] S. Janson. One, two and three times  $\log n/n$  for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.
- [16] Svante Janson. Functional limit theorems for multitype branching processes and generalized pólya urns. *Appl*, 110:177–245, 2004.
- [17] B. P. Kharlamov. The numbers of generations in a branching process with an arbitrary set of particle types. *Theory of Probability and Its Applications*, 14(3):432–449, 1969.
- [18] I. Norros and H. Reittu. On a conditionally poissonian graph process. *Adv. Appl. Probab.*, 38:59–75, 2006.
- [19] G. Hooghiemstra R. van der Hofstad and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electron J. Probab.*, 12(25):703–766 (electronic), 2007.
- [20] B. Söderberg. General formalism for inhomogeneous random graphs. *Phys. Rev. E*, 66:066121, 2002.
- [21] M. Deijfen T. Britton and A. Martin-Löf. Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, 124(6):1377–1397, 2006.

- [22] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures Algorithms*, 27:76–123, 2005.

BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS [istvanko@math.bme.hu](mailto:istvanko@math.bme.hu) [konyju@math.bme.hu](mailto:konyju@math.bme.hu)